



**DÉVELOPPEMENT D'UNE MÉTHODE D'OPTIMISATION POUR LA
GÉNÉRATION DES HORAIRES DES INFIRMIÈRES AFIN DE MINIMISER
L'ABSENTÉISME.**

PAR YOAN VILLENEUVE

**MÉMOIRE PRÉSENTÉ À L'UNIVERSITÉ DU QUÉBEC À CHICOUTIMI
COMME EXIGENCE PARTIELLE EN VUE DE L'OBTENTION DU GRADE DE
MAÎTRE ÈS SCIENCES EN INFORMATIQUE PROFIL RECHERCHE**

QUÉBEC, CANADA

© YOAN VILLENEUVE, 2022

RÉSUMÉ

Depuis quelques années, le réseau public de la santé du Québec fait face à une pénurie grandissante de main-d'oeuvre. Cette pénurie est reliée à une réforme du système de santé au Québec ayant introduit de nombreux changements administratifs et plusieurs coupes budgétaires. L'analyse des infirmières montre que celles-ci travaillent dans des conditions difficiles, ce qui cause une détérioration de leur santé mentale et physique. Entre autres, le taux d'absentéisme des infirmières a atteint un sommet historique et continue d'augmenter chaque année. De plus, un grand nombre de cas d'épuisements professionnels ont été recensés dans les dernières années, causant des arrêts de travail pour des périodes prolongées. Les conséquences sont ressenties par une augmentation de la charge de travail des infirmières, qui doivent effectuer des quarts de travail en temps supplémentaire afin de compenser le manque de personnel soignant. Pour toutes ces raisons, les infirmières ont besoin d'horaires de travail convenables et adaptés à leur situation.

Ce mémoire propose une solution à la problématique de gestion des horaires des infirmières afin de réduire les risques d'absentéisme à l'aide d'une valeur permettant de quantifier le bien-être des infirmières. Un modèle d'optimisation en nombres entiers est utilisé afin de construire des horaires pour les infirmières d'une unité de soins. Un algorithme *Branch-and-bound* est utilisé afin de résoudre le modèle. Ce modèle d'optimisation est supporté par un autre modèle issu de l'apprentissage machine, où l'un des paramètres d'optimisation est calculé par un algorithme Random Forest Regression (RFR). Ce dernier permet de prédire le paramètre de bien-être des infirmières en utilisant l'historique des horaires antérieurs. À partir de ces données, la sommation du nombre de quarts travaillés (ex. à temps régulier, à temps supplémentaire, etc.), des congés et des absences est utilisée pour la construction d'une variable de prédiction de l'état de chaque infirmière. L'objectif du modèle d'optimisation est de diminuer la charge de travail des infirmières à risque en assignant le nombre minimal de quarts de travail requis, en leur offrant de meilleures assignations à des quarts de préférence et en évitant les quarts de travail problématique (ex. temps supplémentaire).

Les résultats obtenus à l'aide du modèle de prédiction et d'optimisation sont tous deux concluants. Le modèle de prédiction est en mesure de suivre l'évolution de la tendance du bien-être des infirmières. En analysant les données d'horaire des infirmières, les semaines menant à un arrêt de travail se traduisent généralement par une dégradation de la valeur de bien-être précédent l'absence d'une infirmière. Les horaires réels provenant des données fournies sont comparés avec ceux générés en répliquant leurs paramètres dans le modèle d'optimisation. Les horaires générés satisfont toutes les contraintes du modèle, diminuent le nombre d'affectations de temps supplémentaires et répartissent mieux les quarts de travail aux infirmières. En général, les infirmières à risque sont rarement affectées à des quarts en temps supplémentaires, obtiennent leurs quarts de préférence et sont affectées moins fréquemment à l'horaire.

TABLE DES MATIÈRES

RÉSUMÉ	ii
LISTE DES TABLEAUX	vi
LISTE DES FIGURES	x
LISTE DES ABRÉVIATIONS	xii
DÉDICACE	xiii
REMERCIEMENTS	xiv
CHAPITRE I – INTRODUCTION	1
1.1 PROBLÉMATIQUE	1
1.2 CONFECTION D'HORAIRES DANS LE SYSTÈME DE LA SANTÉ	2
1.3 OBJECTIFS DU PROJET	4
1.4 STRUCTURE DU MÉMOIRE	4
CHAPITRE II – REVUE DE LITTÉRATURE	6
2.1 L'ABSENTÉISME AU TRAVAIL ET L'ÉTAT DES INFIRMIÈRES AU QUÉBEC	6
2.2 GÉNÉRATION D'HORAIRE	12
2.2.1 EXEMPLE D'OPTIMISATION LINÉAIRE	12
2.2.2 MODÈLES DE CONFECTION D'HORAIRE DANS LA LITTÉRATURE	18
2.2.3 AUTRES APPROCHES À LA CONFECTION D'HORAIRE	22
2.3 PRÉDICTION DE DONNÉES ET APPRENTISSAGE MACHINE	25
2.3.1 ARBRES DÉCISIONNELS	27
2.3.2 FORêt D'ARBRES DÉCISIONNELS ALÉATOIRES	30
2.3.3 PRÉDICTION DE L'ABSENTÉISME	33
CHAPITRE III – MÉTHODOLOGIE	35
3.1 MANIPULATION ET ANALYSE DES DONNÉES HISTORIQUE	36
3.1.1 RÉCEPTION ET NETTOYAGE DES DONNÉES	36

3.1.2 ANALYSE DES DONNÉES	41
3.2 MODÈLE DE PRÉDICTION DU BIEN-ÊTRE	47
3.2.1 CRÉATION D'UN NOUVEL ENSEMBLE DE DONNÉES POUR LA PRÉDICTION DU BIEN-ÊTRE	47
3.2.2 CRÉATION D'UN SCORE DE PRÉDICTION DU BIEN-ÊTRE	50
3.2.3 AJUSTEMENT DE L'ENSEMBLE DE DONNÉES POUR LA PRÉDICTION	56
3.2.4 NETTOYAGE DE L'ENSEMBLE DE DONNÉES DE PRÉDICTION	57
3.2.5 PROGRAMMATION DU MODÈLE DE PRÉDICTION DU BIEN-ÊTRE	58
3.3 MODÈLE D'OPTIMISATION	62
3.3.1 ENSEMBLE	62
3.3.2 VARIABLES	63
3.3.3 PARAMÈTRES	64
3.3.4 FONCTION OBJECTIF	65
3.3.5 CONTRAINTES	66
CHAPITRE IV – RÉSULTATS	72
4.1 PERFORMANCE DU MODÈLE DE PRÉDICTION	72
4.2 ANALYSE DE LA QUALITÉ DU SCORE DE BIEN-ÊTRE PRÉDIT	84
4.2.1 INFIRMIÈRE EN ARRÊT DE TRAVAIL	90
4.3 RÉSULTATS DU MODÈLE D'OPTIMISATION	95
4.3.1 DONNÉES UTILISÉES	96
4.3.2 RÉSULTATS DU MODÈLE	98
4.3.3 COMPARAISON DES RÉSULTATS	110
4.3.4 PERSPECTIVE SANS TEMPS SUPPLÉMENTAIRE	132
CHAPITRE V – CONCLUSION	134
5.1 DISCUSSION DES RÉSULTATS	134
5.2 LIMITATION DU MODÈLE ET AMÉLIORATIONS FUTURES	136
BIBLIOGRAPHIE	138

ANNEXE A – CERTIFICAT ÉTHIQUE 143

LISTE DES TABLEAUX

TABLEAU 2.1 : PARAMÈTRE DES QUOTAS FICTIFS.	14
TABLEAU 2.2 : RÉSULTAT APRÈS OPTIMISATION DE LA FONCTION OBJECTIF À L'ÉQ. (2.2).	15
TABLEAU 3.1 : TABLE DES VARIABLES DE L'ENSEMBLE DE DONNÉES HORAIRE.XLSX	37
TABLEAU 3.2 : ÉCHANTILLON DE L'ENSEMBLE À LA SUITE DU NETTOYAGE DES DONNÉES.	38
TABLEAU 3.3 : QUOTAS DE TRAVAIL DES INFIRMIÈRES (7471) ET INFIRMIÈRES (3455).	40
TABLEAU 3.4 : MÉDIANE DU NOMBRE D'INFIRMIÈRE AYANT EFFECTUÉ AU MINIMUM UN QUART DE TRAVAIL SELON LE JOUR DE LA SEMAINE.	43
TABLEAU 3.5 : EXEMPLE D'UNE INSTANCE DE DONNÉES FOURNIES AU MODÈLE D'APPRENTISSAGE MACHINE.	48
TABLEAU 3.6 : EXEMPLE D'UNE INSTANCE DE DONNÉES DE L'INFIRMIÈRE 12 POUR L'HORAIRE DU 4 AOÛT AU 17 AOÛT 2019.	54
TABLEAU 3.7 : ENSEMBLES DES INFIRMIÈRES.	62
TABLEAU 3.8 : ENSEMBLES DES QUARTS ET DES JOURS.	63
TABLEAU 3.9 : MAXIMUM DE TEMPS SUPPLÉMENTAIRE ATTRIBUABLE SELON L'ÉQ. (3.29).	70
TABLEAU 4.1 : RÉSULTAT DE LA PERFORMANCE DU MODÈLE DE PRÉDICTION DU BIEN-ÊTRE LORSQUE DU RETRAIT DES INSTANCES DE L'ENSEMBLE DE DONNÉES COMPORTANT UN VALEUR DE ASSURANCE À 5 ET PLUS.	74
TABLEAU 4.2 : RÉSULTAT DE LA PERFORMANCE DU MODÈLE DE PRÉDICTION DU BIEN-ÊTRE UTILISANT LA STRATÉGIE <i>LEAVE-ONE-SUBJECT-OUT</i>	76

TABLEAU 4.3 : RÉSULTAT DE LA PERFORMANCE DU MODÈLE DE PRÉDICTION DU SCORE DE BIEN-ÊTRE LORSQUE LES VARIABLES LES MOINS SIGNIFICATIVES SONT RETIRÉES DE L'ENSEMBLE DE DONNÉES.	80
TABLEAU 4.4 : INTERVALLE DES PARAMÈTRES POSSIBLES POUR LA RECHERCHE D'HYPERRPARAMÈTRE DANS LE MODÈLE DE PRÉDICTION DU BIEN-ÊTRE	81
TABLEAU 4.5 : PARAMÈTRES POUR LE MODÈLE DE PRÉDICTION DU SCORE DE BIEN-ÊTRE	82
TABLEAU 4.6 : RÉSULTAT DE LA PERFORMANCE DU MODÈLE DE PRÉDICTION DU BIEN-ÊTRE À LA SUITE DE LA MODIFICATION DES HYPERPARAMÈTRES DU RFR.	82
TABLEAU 4.7 : RÉSULTAT FINAL DE LA PERFORMANCE DU MODÈLE DE PRÉDICTION DU SCORE DE BIEN-ÊTRE POUR LES INFIRMIÈRES.	83
TABLEAU 4.8 : RÉPARTITION DES TYPES DE QUARTS TRAVAILLÉS PAR INFIRMIÈRE SELON LES DONNÉES HISTORIQUES.	97
TABLEAU 4.9 : RÉPARTITION DES TYPES DE QUARTS TRAVAILLÉS PAR INFIRMIÈRE AUXILIAIRE SELON LES DONNÉES HISTORIQUES.	97
TABLEAU 4.10 : PARAMÈTRES UTILISÉS POUR LES RÉSULTATS DU TABLEAU 4.11	99
TABLEAU 4.11 : RÉSULTAT DE LA GÉNÉRATION D'HORAIRE DE L'UNITÉ D'INFIRMIÈRES AVEC LE PARAMÉTRAGE PRÉSENTÉ AU TABLEAU 4.10	99
TABLEAU 4.12 : RÉSULTAT DE LA GÉNÉRATION D'HORAIRE DE L'UNITÉ D'INFIRMIÈRES AUXILIAIRES AVEC PARAMÉTRAGE DE BASE.	102
TABLEAU 4.13 : PARAMÈTRES UTILISÉS POUR LES RÉSULTATS DU TABLEAU 4.14	104
TABLEAU 4.14 : RÉSULTAT DE LA GÉNÉRATION D'HORAIRE DE L'UNITÉ D'INFIRMIÈRES AVEC SCORE DE BIEN-ÊTRE FICTIF	105

TABLEAU 4.15 : PARAMÈTRES UTILISÉS POUR LES RÉSULTATS DU TABLEAU 4.16	106
TABLEAU 4.16 : RÉSULTAT DE LA GÉNÉRATION D'HORAIRE DE L'UNITÉ D'INFIRMIÈRES AVEC SCORE DE BIEN-ÊTRE PRÉDIT AVEC LA SEMAINE DU 8 DÉCEMBRE 2019.	107
TABLEAU 4.17 : PARAMÈTRES UTILISÉS POUR LES RÉSULTATS DU TABLEAU 4.18	108
TABLEAU 4.18 : RÉSULTAT DE LA GÉNÉRATION D'HORAIRE DE L'UNITÉ D'INFIRMIÈRES AVEC SCORE DE BIEN-ÊTRE PRÉDIT AVEC LA SEMAINE DU 8 DÉCEMBRE 2019.	109
TABLEAU 4.19 : DONNÉES HISTORIQUES TIRÉES DE L'HORAIRE DE L'UNITÉ D'INFIRMIÈRES DU 8 DÉCEMBRE 2019 AU 21 DÉCEMBRE 2019	112
TABLEAU 4.20 : PARAMÈTRES UTILISÉS POUR LES RÉSULTATS DU TABLEAU 4.21	113
TABLEAU 4.21 : RÉSULTAT DE LA GÉNÉRATION DE L'HORAIRE DE L'UNITÉ D'INFIRMIÈRES À DES FINS DE COMPARAISON AVEC L'HORAIRE DU 8 DÉCEMBRE 2019.	114
TABLEAU 4.22 : DONNÉES NUMÉRIQUES DES HORAIRES DE TRAVAIL ET DE CONGÉ PARTANT DE LA SEMAINE DU 12 MAI 2019 AU 22 JUIN 2019.	116
TABLEAU 4.23 : HORAIRE DÉTAILLÉ DE CHAQUE JOUR POUR L'INFIRMIÈRE 11 POUR LES HORAIRES DE LA SEMAINE 15 À 19.	118
TABLEAU 4.24 : PARAMÈTRES UTILISÉS POUR LES RÉSULTATS DU TABLEAU 4.25	119
TABLEAU 4.25 : RÉSULTAT DE LA GÉNÉRATION D'UN HORAIRE SE BASANT SUR LES DONNÉES HISTORIQUES DE LA SEMAINE 17 (26 MAI 2019).	120
TABLEAU 4.26 : DONNÉES TIRÉES DE L'HORAIRE DE L'UNITÉ D'INFIRMIÈRES DU 21 JUILLET 2019 AU 3 AOÛT 2019.	122

TABLEAU 4.27 : PARAMÈTRES UTILISÉS POUR LES RÉSULTATS DE LA FIGURE 4.28	124
TABLEAU 4.28 : RÉSULTAT DE LA GÉNÉRATION D'UN HORAIRE SE BASANT SUR LES DONNÉES HISTORIQUES DE LA SEMAINE 25 (21 JUILLET 2019).	124
TABLEAU 4.29 : DONNÉES TIRÉES DE L'HORAIRE DE L'UNITÉ D'INFIRMIÈRES DU 21 JUILLET 2019 AU 3 AOÛT 2019.	127
TABLEAU 4.30 : PARAMÈTRES UTILISÉS POUR LES RÉSULTATS DU TABLEAU 4.31	129
TABLEAU 4.31 : RÉSULTAT DE LA GÉNÉRATION D'UN HORAIRE SE BASANT SUR LES DONNÉES HISTORIQUES DE LA SEMAINE 25 (21 JUILLET 2019).	130
TABLEAU 4.32 : RÉSULTAT DE LA GÉNÉRATION D'HORAIRE DE L'UNITÉ D'INFIRMIÈRES AVEC SCORE DE BIEN-ÊTRE PRÉDIT AVEC LA SEMAINE DU 8 DÉCEMBRE 2019.	133

LISTE DES FIGURES

FIGURE 2.1 – EXEMPLE DE L’ÉTAPE DE SÉPARATION ET ÉVALUATION D’UN PROBLÈME FICTIF D’OPTIMISATION EN NOMBRES ENTIERS À L’AIDE DE L’ALGORITHME <i>BRANCH-AND-BOUND</i>	17
FIGURE 2.2 – VISUALISATION D’UN ENSEMBLE DE DONNÉES.	28
FIGURE 2.3 – EXEMPLE D’ARBRE DÉCISIONNEL CLASSIFIANT LE SURMENAGE.	29
FIGURE 2.4 – VISUALISATION D’UNE FORêt D’ARBRES DÉCISIONNELS À TROIS VARIABLES DE CLASSIFICATION.	32
FIGURE 3.1 – PROCESSUS DE DÉVELOPPEMENT DU MODÈLE DE GÉNÉRATION D’HORAIRE OPTIMALE AVEC VALEUR DE PRÉDICTION DE L’ABSENTÉISME.	35
FIGURE 3.2 – NOMBRE D’INFIRMIÈRES AYANT EFFECTUÉ UN QUART DE TRAVAIL POUR CHAQUE JOURNÉE ENTRE LE 3 FÉVRIER 2019 ET LE 28 DÉCEMBRE 2019.	42
FIGURE 3.3 – PROPORTION DES QUARTS AFFECTÉS POUR CHAQUE INFIRMIÈRE PAR RAPPORT AU NOMBRE D’HEURES PAR AFFECTATION DE QUART.	44
FIGURE 3.4 – PROPORTION DU NOMBRE D’HEURES TRAVAILLÉES POUR CHAQUE INFIRMIÈRE RELATIF À LA PÉRIODE D’AFFECTATION DU QUART.	46
FIGURE 3.5 – MISE AU POINT DE LA POSITION DE L’ÉTIQUETTE ET DES INSTANCES DE L’ENSEMBLE DE DONNÉES AFIN DE PRÉDIRE LE SCORE DE BIEN-ÊTRE.	56
FIGURE 3.6 – PROCESSUS DE CRÉATION D’UN ENSEMBLE D’ENTRAÎNEMENT ET DE TEST POUR LA CRÉATION D’UNE FORêt D’ARBRES DE RÉGRESSION.	60
FIGURE 4.1 – MATRICE DES COEFFICIENTS DE CORRÉLATION ENTRE LES DIFFÉRENTES VARIABLES DU MODÈLE DE PRÉDICTION DE L’ABSENTÉISME.	77

FIGURE 4.2 – MATRICE DES VALEURS DE <i>P</i> -VALUE ENTRE CHAQUE VARIABLE DE L'ENSEMBLE DE DONNÉES DE PRÉDICTION.	78
FIGURE 4.3 – ANALYSE DE L'ÉVOLUTION DU SCORE DE BIEN-ÊTRE PRÉDIT ET RÉEL DE L'INFIRMIÈRE 43.	85
FIGURE 4.4 – ANALYSE DE L'ÉVOLUTION DU SCORE DE BIEN-ÊTRE PRÉDIT ET RÉEL DE L'INFIRMIÈRE 12.	86
FIGURE 4.5 – ANALYSE DE L'ÉVOLUTION DU SCORE DE BIEN-ÊTRE PRÉDIT ET RÉEL DE L'INFIRMIÈRE 90.	88
FIGURE 4.6 – ANALYSE DE L'ÉVOLUTION DU SCORE DE BIEN-ÊTRE PRÉDIT ET RÉEL DE L'INFIRMIÈRE 96.	89
FIGURE 4.7 – ANALYSE DE L'ÉVOLUTION DU SCORE DE BIEN-ÊTRE PRÉDIT ET RÉEL DE L'INFIRMIÈRE 11.	91
FIGURE 4.8 – ANALYSE DE L'ÉVOLUTION DU SCORE DE BIEN-ÊTRE PRÉDIT ET RÉEL DE L'INFIRMIÈRE 16.	92
FIGURE 4.9 – ANALYSE DE L'ÉVOLUTION DU SCORE DE BIEN-ÊTRE PRÉDIT ET RÉEL DE L'INFIRMIÈRE 63.	94

LISTE DES ABRÉVIATIONS

API	Application Programming Interface
APA	American Psychological Association
BP	Backpropagation
CBC	COIN Branch-and-cut
COIN	Computational Infrastructure for Operations Research
CH	Centre Hospitalier
CHSLD	Centre d'Hébergement et de Soins de Longue Durée
CLSC	Centre Local de Services Communautaires
CPEJ	Centre de Protection de l'Enfance et de la Jeunesse
CR	Centre de réadaptation
CCHST	Centre Canadien d'Hygiène et de Sécurité au Travail
CIUSSS	Centre Intégré Universitaire de Santé et de Services Sociaux
CISSS	Centre Intégré de Santé et de Services Sociaux
CSSS	Centre Santé et de Services Sociaux
EV	Variance Expliquée
FIQ	Fédération Interprofessionnelle de la santé du Québec
IA	Intelligence Artificielle
LSTM	Long Short-Term Memory
MSE	Mean Square Error
MAE	Mean Absolute Error
RF	Random Forest
R^2	Coefficient de Détermination
RBF	Radial Basis Function
RFR	Random Forest Regression
RNA	Réseaux de Neurones Artificiels
RNC	Réseau de Neurones Classique
SVR	Support Vector Regression
TSO	Temps Supplémentaire Obligatoire
TS	Temps Supplémentaire

DÉDICACE

À ma grand-maman Diane Boivin,

Tu es l'infirmière la plus importante de ma vie. Merci d'avoir passé 34 années de carrière à travailler pour le bien-être et la santé de la population. Je suis fier de pouvoir contribuer au bien-être de personne comme toi qui consacre leur vie à un métier aussi gratifiant.

REMERCIEMENTS

Je remercie toutes les personnes qui m'ont aidé pour la rédaction de ce mémoire. Ce projet n'aurait pas pu avoir lieu sans l'implication de plusieurs personnes ayant offert leur support dans la réalisation de celui-ci.

Je remercie grandement ma directrice de recherche Sara Séguin pour m'avoir offert la possibilité d'effectuer ce projet de recherche sous sa tutelle. Son savoir en matière d'optimisation et sa supervision tout au long de cette recherche ont rendu possible l'accomplissement de ce mémoire. Ce projet a demandé l'investissement d'un temps précieux de sa part et j'en suis considérablement reconnaissant.

Je remercie également mon co-directeur de recherche Julien Maitre pour avoir été tout aussi présent dans le cadre de cette recherche. Son expertise en apprentissage machine et en analyse de données mon été indispensable tout au long de ce projet. Son approche analytique m'a grandement aidé dans chaque étape de développement de ce projet de recherche.

Je remercie l'entreprise Timesphere et particulièrement Renaud Grimard pour m'avoir partagé son expertise en matière de confection d'horaires pour la réalisation de ce projet. Je remercie aussi tous les employés pour m'avoir accueilli dans leur équipe au cours de ce projet. Travailler dans vos bureaux fut une très belle expérience.

Merci au Centre Intégré Universitaire de Santé et de Services Sociaux (CIUSSS) du Saguenay-Lac-Saint-Jean de m'avoir fourni les données nécessaires au développement de ce projet. Je remercie également toutes les infirmières ayant participé de façon anonyme aux données collectées.

Un grand merci à mes parents pour m'avoir supporté et encouragé durant toutes mes études. C'est votre soutien constant qui m'a permis d'arriver où j'en suis à ce jour.

Enfin, je remercie tous mes amis et surtout mon amour, Cynthia Tremblay, pour son support moral et ces encouragements. Sa présence à mes côtés et son sourire m'ont donné le courage et la motivation nécessaire pour compléter ce mémoire.

CHAPITRE I

INTRODUCTION

Depuis plusieurs années, le système de santé du Québec a été sujet à de nombreux changements propres à la réorganisation du système administratif et à la centralisation des instances de santé à leurs secteurs d'opération. Depuis l'adoption de la loi 10 en 2015, le Québec a vu l'apparition des Centre Intégré de Santé et de Services Sociaux (CISSS), soit des organismes publics chargés par le ministère de la Santé et des Services sociaux de prodiguer les soins et les services sociaux dans chacune des régions sociosanitaires. L'objectif de la loi 10 est de diminuer les frais gouvernementaux et de simplifier l'accès aux soins et aux services de santé du Québec. Pour ce faire, les CISSS permettent la diminution du nombre de postes administratifs en centralisant les soins de santé pour les différentes régions du Québec. Ceux-ci assurent un lien facile entre la population et les soins de santé offerts selon les directives gouvernementales. Il est dit Centre Intégré Universitaire de Santé et de Services Sociaux (CIUSSS) lorsqu'un secteur englobe une université offrant un programme pré-doctoral en médecine ou exploite un institut universitaire du domaine social. Les CISSS font principalement la gestion de cinq types de Centre Santé et de Services Sociaux (CSSS), soit les Centre Hospitalier (CH), Centre d'Hébergement et de Soins de Longue Durée (CHSLD), Centre Local de Services Communautaires (CLSC), Centre de Protection de l'Enfance et de la Jeunesse (CPEJ) et Centre de réadaptation (CR).

1.1 PROBLÉMATIQUE

Ces établissements emploient tous à différents degrés des infirmières afin de subvenir aux besoins médicaux de leurs patients. Dans plusieurs cas, le besoin de la présence de personnel soignant est nécessaire 24 heures sur 24. Or, les établissements des CIUSSS du Québec font

face à une pénurie grandissante d'infirmières dans le domaine public de la santé. La cause de cette pénurie est attribuable à de nombreux facteurs, comme des coupes budgétaires, une transition difficile de la centralisation des CSSS et une mauvaise gestion des ressources offertes pour le personnel médical. Les difficultés grandissantes des infirmières ont été observées à la suite de la réforme causée par l'adoption de la loi 10. En 2017, Gagné (2017) observe que 25% d'un échantillon de 891 infirmières du Québec se retrouve en situation de détresse. Une étude plus récente de Stelnicki & Carleton (2021) indique que 47.9% d'un échantillon de 4267 infirmières canadiennes ont été dépistées positives pour un trouble mental, notant un très haut taux de trouble anxieux, dépressif et d'épuisement professionnel. Par conséquent, une baisse de la qualité des soins aux patients est ressentie, ce qui implique des risques accrus d'erreurs pouvant être faits par une infirmière (Rogers *et al.*, 2004).

1.2 CONFECTION D'HORAIRES DANS LE SYSTÈME DE LA SANTÉ

Un horaire pour infirmières de la Fédération Interprofessionnelle de la santé du Québec (FIQ) englobe 14 jours de travail, où chaque journée se divise en période de jour, soir et nuit. Un gestionnaire commence le processus de confection de l'horaire huit semaines avant le début des jours de travail de celui-ci. Les infirmières doivent envoyer leurs préférences, leurs demandes de remplacement et demandes de congé, au plus tard six semaines avant le début de l'horaire concerné. L'horaire est construit durant les deux semaines qui suivent par le gestionnaire et est rendu disponible aux infirmières quatre semaines avant le début de celui-ci.

La création d'horaires pour les infirmières est un processus long demandant beaucoup de temps aux gestionnaires. Les personnes travaillant sur la confection d'horaire rapportent que cette tâche représente la majeure partie de leur temps au travail. Il est souvent question d'utiliser une liste d'appel afin de trouver des infirmières pouvant combler les quotas de travail de chaque période. Si chaque infirmière a été appelée sans succès, le gestionnaire

doit recommencer la liste avec l'ajout d'une prime. Si personne ne répond positivement, les infirmières sont soumises à du temps supplémentaire obligatoire, une pratique fortement controversée et lourde de conséquences sur la santé mentale et physique des infirmières.

À ce jour, encore beaucoup de gestionnaires construisent leur horaire de travail à la main. Or, il existe des méthodes permettant l'automatisation de ce processus. Depuis plusieurs années, la problématique de confection d'horaire est un sujet que plusieurs chercheurs ont abordé dans le domaine de l'optimisation et de la recherche opérationnelle. Néanmoins, encore trop peu de chercheurs ont abordé le bien-être des infirmières dans leur modèle de confection d'horaires à un niveau significatif. La confection d'horaire optimisé passe par différentes étapes. Après l'analyse et la compréhension du problème, celui-ci doit être transcrit sous forme de variables, paramètres et équations mathématiques afin de composer un modèle d'optimisation. Un modèle d'optimisation a pour objectif de minimiser ou maximiser le résultat de sa fonction objectif en manipulant la valeur des variables du problème. Les variables d'un modèle d'optimisation sont sujettes à un ensemble de contraintes limitant l'ensemble de solutions possibles du problème à celles désirées. La résolution d'un modèle est possible grâce à un grand nombre d'algorithmes dédiés à ce type de problème, tel le *Branch-and-bound* qui utilise le principe des arbres de recherche afin d'obtenir une solution optimale. Un modèle de gestion d'horaire pour les hôpitaux demande beaucoup de flexibilité, puisque les infirmières sont soumises à un grand nombre de contraintes propres à la logistique de leur horaire. La conception d'un tel modèle demande l'obtention de données par rapport aux horaires des infirmières, de l'information qui n'est pas toujours facile à récolter. Dans le cadre de ce mémoire, des données provenant du CIUSSS de la région du Saguenay Lac-Saint-Jean ont été obtenues sur une durée d'un an. Ces données contiennent de l'information sur chaque quart travail, congés et absence des employés (majoritairement des infirmières) d'une unité de soins d'un CHSLD.

1.3 OBJECTIFS DU PROJET

Dans le cadre de ce projet, deux modèles sont développés et interviennent ensemble afin de concevoir des horaires pour les infirmières qui visent à minimiser les risques d'absence au travail. La prévention de l'absentéisme est possible avec le développement d'un modèle d'apprentissage machine. Ce modèle a pour but de prédire le bien-être d'une infirmière après chaque horaire travaillé en utilisant les données des horaires antérieurs. Le modèle de prédiction du bien-être des infirmières utilise l'algorithme de forêts aléatoires régressives afin d'entraîner un modèle. Le résultat de ce modèle interagit ensuite avec le second modèle, un modèle d'optimisation en nombres entiers permettant la confection d'horaire optimisé pour maximiser le bien-être des infirmières tout en respectant les restrictions propres à la logistique des horaires pour les infirmières du Québec. Le paramètre prédit par le modèle d'apprentissage machine est fourni en paramètre d'entrée au modèle d'optimisation en nombres entiers, permettant ainsi d'attribuer plus de quarts de préférence, diminuer les chances d'affectation de quart à temps supplémentaire et diminuer le nombre d'affectations de quarts réguliers à un horaire. Cet objectif relie le domaine mathématique et informatique afin de répondre à la problématique reliée à la confection d'horaire respectant la santé physique et mentale des infirmières du Québec.

1.4 STRUCTURE DU MÉMOIRE

Le chapitre 1 de ce mémoire donne une mise en contexte de la réalité vécue par les infirmières travaillant dans le domaine public du Québec. Le chapitre 2 est une analyse de la littérature scientifique portant sur les conditions de travail des infirmières, les travaux antérieurs portant la génération d'horaires optimisés et sur la prédiction de données à l'aide de l'apprentissage machine. Entre autres, la publication de Gagné (2017) a permis de dresser un portait des problématiques dans le système de santé québécois en se rapportant aux

conditions de travail des infirmières. Concernant l'optimisation des horaires, un bon nombre de publications abordant le problème de gestion d'horaires sont présentées. Malgré un grand nombre de publications sur le sujet, trop peu considère réellement l'état des infirmières, optant plus souvent pour une approche avantageant la productivité des centres hospitaliers.

Le chapitre 3 aborde la méthodologie du travail. À l'aide d'un ensemble de données d'horaire d'infirmière, un ensemble de données pour la prédiction d'un score de bien-être a été créé. Le score de bien-être est une variable créée afin de quantifier le bien-être pour chaque infirmière. Enfin, le modèle d'optimisation pour la génération d'horaire est détaillé. Il s'y trouve un total de 19 contraintes reliées à la confection d'horaire et à la réduction de la charge des infirmières avec haut score prédit.

Le chapitre 4 montre les résultats obtenus par rapport aux deux modèles développés dans ce mémoire. D'abord, l'évolution du résultat de performance du modèle est faite en expliquant chaque étape de modification de l'ensemble de données et du modèle de prédiction qui a mené à une amélioration. Ensuite, le score de bien-être réel qui a été calculé à partir des données historiques est comparé avec celui obtenu à l'aide du modèle de prédiction. Enfin, les résultats du modèle d'optimisation pour générer des horaires pour les infirmières sont analysés. Puisque le modèle n'a pas pu être testé dans le milieu hospitalier, il est comparé avec les horaires de l'ensemble de données historiques à l'aide d'un paramétrage similaire.

CHAPITRE II

REVUE DE LITTÉRATURE

L'objectif principal de ce projet est la création d'un modèle d'optimisation générant des horaires pour les infirmières afin d'éviter l'absentéisme. Ce projet de recherche fait intervenir plusieurs domaines et explore différentes perspectives. Toutefois, celui-ci se concentre sur l'apprentissage machine afin de prédire un score décrivant l'état de bien-être pour chaque infirmière et la recherche opérationnelle afin de concevoir un modèle de génération d'horaires soumis à plusieurs contraintes. Ce travail aborde aussi les enjeux de l'absentéisme au travail et la situation actuellement vécue par les infirmières du milieu professionnel de la santé au Québec. Ainsi, la première partie de ce chapitre est une mise en situation afin de justifier ce projet à l'aide de la documentation sur l'absentéisme et l'expérience de travail des infirmières. La suite de la revue de littérature aborde des sujets plus techniques concernant la génération des horaires, principalement dans le domaine hospitalier, à l'aide de l'optimisation. La dernière partie se concentre sur la littérature traitant de la prédiction de données à l'aide de l'apprentissage machine, abordant le domaine des arbres décisionnels, des arbres décisionnels aléatoires et la prédiction de l'absentéisme au travail.

2.1 L'ABSENTÉISME AU TRAVAIL ET L'ÉTAT DES INFIRMIÈRES AU QUÉBEC

Selon l'*American Psychological Association (2021)* (APA), le terme *épuisement professionnel*, inspiré de son équivalent anglais *burnout*, décrit un sentiment d'épuisement émotionnel, de détachement ou de cynisme, un faible niveau de réalisation personnelle ou encore une dépersonnalisation de l'individu en question à l'intention de sa condition professionnelle. Cette définition est acceptée par le Centre canadien d'hygiène et de sécurité au travail (2021) (CCHST). Ce phénomène est observé chez les personnes subissant un haut niveau de stress

dû à un surmenage physique ou mental sur une période prolongée. Le CCHST cite différents facteurs provenant d'un milieu de travail pouvant mener à l'épuisement professionnel :

- exigences déraisonnables,
- manque de contrôle ou incapacité d'influer sur les décisions qui ont une incidence sur son travail (ex. horaire, charge de travail et ressources),
- attentes relatives au travail mal définies,
- milieu de travail médiocre (intimidation, harcèlement, etc.),
- non concordance des valeurs ou incompatibilité personne-emploi,
- soutien insuffisant, que ce soit au travail ou à la maison,
- déséquilibre entre le travail et la vie personnelle.

Parmi ces facteurs, le manque de contrôle, le soutien insuffisant ainsi que le déséquilibre entre le travail et la vie personnelle sont trois points pouvant être influencés de façons positives à l'aide d'une méthode de création d'horaires. La prise en considération des préférences de quart de travail couplé avec un modèle d'optimisation mathématique peut influencer positivement le moral des infirmières en leur offrant une meilleure gestion de leur horaire.

Kane *et al.* (2007) proposent une étude sur l'état du personnel infirmier et la qualité des soins prodigués. Les auteurs ont fait une analyse de 94 travaux sur le sujet. Dans ces 94 travaux, sept se sont focalisés sur la gestion d'horaire de travail. Ils ont conclu que les infirmières travaillent en moyenne 55 minutes de plus que leur horaire planifié. Une analyse de 5317 quarts de travail effectué au cours d'une période de 28 jours montre que 38,7% d'entre eux ont été d'une durée de 12,5 heures ou plus. Durant cette période, le quart des infirmières ont travaillé 50 heures ou plus par semaine sur deux semaines. Plus de la moitié des infirmières travaillent 12 heures ou plus par jour et la moitié sont susceptibles de travailler six à sept jours par semaine. Cette situation représente un risque selon Rogers *et al.* (2004), qui ont observé

une augmentation des chances qu'une infirmière effectue des erreurs après 8.5 heures de travail consécutif. Après avoir dépassé un quart de travail de 12.5 heures, les chances d'effectuer une erreur peuvent augmenter jusqu'au triple du risque initial pour chaque infirmière. Enfin, les infirmières de 50 ans et plus sont moins enclines à effectuer de longues heures de travail.

Une analyse similaire sur le sujet, réalisée par Bae & Fabry (2014), se base sur des publications datant de 2000 à 2013. Un total de 24 articles ont été pris en compte dans cette revue de littérature dont le but est d'évaluer l'effet des longues périodes de travail et des heures supplémentaires sur le corps infirmier et ses patients. Les auteurs constatent que la durée des quarts de travail et le nombre d'heures de travail hebdomadaire ont un impact sur la santé physique et mentale des infirmières, quel que soit le type d'heures supplémentaires (obligatoires ou volontaires). Dans l'ensemble, les résultats de cette analyse suggèrent que le travail par quarts de 12 heures affecte négativement la santé des infirmières et le fait de travailler plus de 40 heures par semaine influe indéniablement sur la santé des infirmières et les soins prodigués aux patients. Enfin, l'étude suggère que la durée des quarts de travail et le nombre d'heures de travail hebdomadaires devraient être contrôlés afin de prévenir d'éventuels résultats indésirables pour les infirmières et les patients. Les auteurs suggèrent ainsi que les personnes responsables de la confection d'horaire de travail envisagent d'inclure des limitations par rapport à la durée des quarts de travail et le nombre d'heures de travail hebdomadaires, le tout afin de promouvoir un environnement de travail sain et sûr.

Gagné (2017) a soumis un mémoire dressant un bilan de l'expérience au travail des infirmières au Québec. L'auteure y souligne, entre autres, une augmentation de la charge de travail des infirmières au courant des dernières années avec un manque de ressources disponibles dans le milieu, menant à la détérioration de la santé mentale et physique des infirmières et à l'augmentation de l'absentéisme chez celles-ci. Ce mémoire est une analyse transversale de 891 questionnaires remplis auprès du personnel infirmier suivi d'une étude avancée de la

littérature afin de cerner les différents enjeux du métier d'infirmière dans le système de santé québécois. Avec cette étude, l'auteure cherche à dresser les profils d'expérience des infirmières oeuvrant dans le réseau de la Santé et des Services sociaux du Québec. À la suite de l'analyse des questionnaires distribués, les 891 infirmières ayant répondu ont été classifiées en situation favorable, modérément favorable ou de détresse. Les infirmières en situation moyennement favorable constituent 50,1% (446 infirmières) des cas à l'étude. Celles-ci se caractérisent par une perception modérée de l'expérience de travail, autant au point de vue structurel, des ressources disponibles liées au processus de travail ou encore au soutien psychologique offert. Mis à part l'accès aux possibilités de développement professionnel, ce groupe d'infirmières a des réponses généralement positives et ne montre pas de signes d'épuisement professionnel ou d'intentions de quitter leurs emplois. Les infirmières en situation de détresse constituent 25,9% (231 infirmières) des cas. Ce profil d'expérience se caractérise par une évaluation modérée à basse des ressources disponibles dans le milieu de travail ainsi qu'une forte pression psychologique. Des trois profils étudiés, ce dernier contient le plus grand nombre d'infirmières décrivant l'effort physique comme étant élevé. Malgré une évaluation généralement basse de l'effort physique, ce profil d'infirmière contient le plus grand nombre d'évaluations négatives. Ce groupe d'infirmières démontre tout de même un engagement modéré envers leur emploi. L'auteure remarque que les infirmières les moins satisfaites de leur condition de travail vivent un fort déséquilibre entre la charge de travail et les ressources à leur disposition. L'étude met en valeur la situation des infirmières ayant beaucoup de difficultés d'accès aux ressources de travail structurelles (environnement de travail, relation avec le personnel supérieur, soutien physique et psychologique, dotation, etc.). Pour ces raisons, les infirmières en détresse ressentent beaucoup de mécontentement envers leurs supérieurs et le personnel administratif. En parallèle avec les objectifs de cette recherche, la prise en considération des préférences de travail des infirmières ainsi qu'une métrique considérant le bien-être peut alléger la tension en offrant la possibilité aux gestionnaires d'offrir des horaires adaptés et la possibilité de mieux

cerner la charge de travail du corps infirmier. Enfin, la catégorisation des infirmières ayant participé à ce questionnaire démontre que plus du quart des infirmières vivent des difficultés au travail.

Par ailleurs, ce même mémoire par Gagné (2017) inclut une analyse approfondie des publications sur l'expérience du métier d'infirmière. Les études empiriques analysées dans ce mémoire sont divisées selon trois perspectives de l'expérience au travail des infirmières, soit : les perspectives structurelles, les processus de travail et les résultats. Les perspectives structurelles se concentrent sur l'aspect pratique et environnemental des infirmières, ce qui n'est pas en lien direct avec l'objectif de ce projet. Les perspectives des processus de travail et des résultats abordent la problématique de gestion d'horaire et du bien-être au travail.

Les travaux empiriques des processus de travail incluent la division et distribution des tâches, les conditions d'emploi et les relations entretenues au travail. Par rapport aux conditions de l'emploi, Gagné (2017) note les restructurations fréquentes ayant lieu dans le système de santé. Ces restructurations engendrent une augmentation de la charge de travail des infirmières, ce qui a un fort impact sur le stress ressenti par ces dernières et dégrade la qualité de vie au travail. Selon Vagharseyedin *et al.* (2011), l'enjeu de la surcharge de travail est l'un des problèmes les plus alarmants. Pour améliorer la qualité de vie des infirmières, les auteurs proposent le développement et la mise en oeuvre de stratégies telles que l'auto-planification, le travail à temps partiel, la flexibilité des horaires ainsi que des systèmes de quarts alternatifs. Il note également la pénurie d'infirmières dans le domaine de la santé, encourageant ainsi la création de nouvelles méthodes et stratégies adaptatives à ce phénomène de surcharge. Les quarts de soir et de nuit, même volontaires, ont aussi un impact négatif sur la santé psychologique des infirmières. En effet, les auteurs Berger & Hobbs (2006) démontrent que ces quarts ont des effets négatifs, autant du point de vue physique que psychologique, en plus d'encourager indirectement des comportements à risque, telle une augmentation de

la consommation de caféine, d'alcool et de médicament pour le sommeil. Il est néanmoins possible de s'adapter aux horaires de soir et de nuit pour les travailleurs permanents, même si cela demande plusieurs jours ou semaines et, de préférence, un suivi adapté. Pour remédier aux problématiques de conditions d'emploi des infirmières, Gagné (2017) propose d'offrir plus d'autonomie décisionnelle aux infirmières, car celle-ci est hautement valorisée par les infirmières et représente un facteur important dans la satisfaction vis-à-vis leur travail. Il est toutefois mentionné d'éviter que cette autonomie devienne elle-même une source de stress pour l'employée.

La perspective des résultats de l'expérience au travail des infirmières met l'accent sur la satisfaction, la santé au travail, l'intention de quitter son emploi et l'engagement au travail. En ce sens, Nolan *et al.* (1995) citent le désir des infirmières à prodiguer des soins de qualité comme étant l'une de leurs plus grandes préoccupations. Malheureusement, les infirmières rapportent un manque de personnel dans leur milieu de travail, ce qui diminue leurs habiletés à assurer des soins de haute qualité. De plus, celles-ci notent un manque de support, de reconnaissance, d'accessibilité, de communication et de flexibilité venant de leurs supérieurs. Pour ces raisons, il serait bénéfique d'améliorer le processus de création des horaires, ce qui laisserait plus de temps aux gestionnaires pour subvenir aux besoins des infirmières. Le stress vécu par les infirmières est aussi un facteur dans la diminution de la qualité des soins qu'elles prodiguent. Bien que le stress puisse provenir du milieu professionnel que de la personne, la surcharge de travail, la sous-dotation (attribution de quart inadéquat pour la charge de travail demandée) et le manque de soutien contribuent également à ce sentiment. Selon Jennings (2008), lorsque le stress devient chronique, celui-ci évolue en épuisement professionnel qui vient à son tour augmenter les risques d'absentéisme. Enfin, Gagné (2017) indique que la responsabilité parentale des infirmières est significativement associée à l'intention de quitter leur emploi.

Les problématiques recensées ci-dessus montrent l'intérêt de développer des moyens pour améliorer la qualité de vie professionnelle des infirmières du Québec. En ce qui a trait à l'absentéisme, la surcharge de travail est probablement l'aspect le plus prédominant et le plus abordé lorsqu'il est fait mention de la qualité de vie au travail des infirmières. Il faut donc considérer le surmenage des infirmières comme un facteur important dans la prédiction de l'absentéisme. De plus, le temps gagné par l'automatisation des horaires de travail permet aux gestionnaires de s'attarder aux problématiques structurelles de l'expérience au travail des infirmières qui s'apparentent à l'environnement de travail. La prise en considération des préférences de quarts des infirmières risque foncièrement d'augmenter la satisfaction au travail de ces dernières tout en leur permettant de mieux concilier leur travail et leurs vies personnelles.

2.2 GÉNÉRATION D'HORAIRE

Le problème de création d'horaires de travail a été largement abordé dans le domaine de la recherche opérationnelle. Plusieurs modélisations et algorithmes d'optimisation ont été proposés pour sa résolution. On y trouve l'optimisation linéaire, non-linéaire, déterministe, stochastique, en nombres entiers et d'autres encore. Parmi ces travaux, ceux portant sur l'optimisation linéaire sont analysés, en plus des méthodes utilisant les métaheuristiques, qui permettent l'obtention relativement rapide de solutions approchées pour des instances où la résolution à l'aide de méthodes exactes prendrait un temps excessivement long.

2.2.1 EXEMPLE D'OPTIMISATION LINÉAIRE

L'optimisation linéaire, ou programmation linéaire, est une méthode mathématique permettant de maximiser ou de minimiser la valeur d'une fonction objectif linéaire, tout en respectant un ensemble de contraintes étant elles aussi linéaires, Schrijver (1998). En program-

mation linéaire, la solution d'un modèle est atteinte lorsque la valeur optimale est trouvée, mais il est possible qu'un problème soit irréalisable selon les restrictions imposées. Malgré un format généralement linéaire de résolution en temps polynomial, certains algorithmes (ex. l'algorithme *branch-and-bound*) peuvent atteindre une complexité de calcul allant de quadratique à exponentielle (Zhang, 1996). Il est question d'optimisation en nombres entiers lorsque toutes les variables d'un modèle sont des entiers. Les problèmes d'optimisation linéaire en nombres entiers font partie de la classe des problèmes **NP-Complets**. **NP** signifie que le problème décisionnel en question est vérifiable en temps polynomial, tandis que **Complets** fait référence à la propriété que la complexité du problème ne peut être réduite davantage, Cook (1971). Dans le cadre de ce projet de recherche, l'optimisation linéaire en nombres entiers est utilisée pour la création d'un modèle de génération d'horaire.

Les problèmes linéaires sont composés de trois parties, soit les variables de décision, la fonction objectif et les contraintes opérationnelles. Les variables de décision sont les variables du problème à déterminer. Les valeurs des variables de décision permettent de maximiser ou de minimiser la fonction selon la nature du problème. Les variables du problème sont aussi soumises à certaines restrictions ou contraintes qui permettent de délimiter l'ensemble des solutions, soit les bornes comprenant les solutions réalisables du problème.

Supposons un horaire de trois jours avec cinq employés. Notons l'ensemble d'employées I et l'ensemble de jours J . Le but est de générer un horaire en assignant des employées aux journées. La variable décisionnelle x_{ij} représentera l'affectation d'un employé $i \in I$ à un jour $j \in J$:

$$x_{ij} = \begin{cases} 1, & \text{si l'employé } i \in I \text{ a été affecté au quart } j \in J \\ 0, & \text{sinon.} \end{cases} \quad (2.1)$$

Comme l'objectif de ce modèle est d'affecter des employés à différents jours de l'horaire, la fonction objectif vise à maximiser le nombre d'affectations d'employés à l'horaire :

$$\max \sum_{i \in I} \sum_{j \in J} x_{ij}. \quad (2.2)$$

Le modèle est aussi constitué de contraintes limitant les solutions possibles. Par exemple, l'Éq. (2.3) permet de limiter l'affectation d'un employé i à une seule affectation par journée j :

$$x_{ij} \leq 1, \quad i \in I, j \in J. \quad (2.3)$$

L'Éq. (2.4) permet de limiter le nombre d'affectations d'employé à l'aide d'un paramètre de quota Q pour chaque jour j . Le quota représente le nombre d'employés requis selon la journée. Le Tableau 2.1 montre la valeur de chacun des paramètres par rapport à la journée.

Tableau 2.1 : Paramètre des quotas fictifs.

	Jour 1	Jour 2	Jour 3
Quota	2	3	2

Ainsi, une contrainte de quota peut être appliquée au modèle en faisant la sommation des employés durant chaque journée de l'horaire :

$$\sum_{i \in I} x_{ij} = Q_j, \quad \forall j \in J. \quad (2.4)$$

Enfin, par souci d'équité, chaque employé i devra travailler au moins une fois dans l'horaire :

$$\sum_{j \in J} x_{ij} \geq 1 \quad \forall i \in I. \quad (2.5)$$

Un modèle d’optimisation est généralement écrit sous forme d’expressions algébriques dans un langage donné avant d’être compilé. AMPL (AMPL Optimization Inc. (2021)), JuMP (Dunning *et al.* (2017)) et PuLP (Mitchell *et al.* (2011)) sont des interfaces permettant l’écriture d’un modèle d’optimisation sous forme de formules mathématiques. Après compilation, le modèle est envoyé à un solveur afin de résoudre le problème. Il existe plusieurs solveurs commerciaux permettant la résolution de problème d’optimisation linéaire en nombres entiers, tel Xpress (FICO (2021)) et Gurobi (Gurobi Optimization, LLC (2022)), mais plusieurs autres solveur sont aussi disponible en source libre, comme GLPK (Makhorin (2012)) et LPSolve (Berkelaar *et al.* (2004)). Pour les problèmes d’optimisation linéaire, le résultat peut être : une solution optimale, plusieurs solutions optimales, ou encore aucune solution si le modèle n’est pas capable de satisfaire les contraintes imposées. Une solution au problème ci-dessus peut prendre la forme du Tableau 2.2. Cet exemple considère un ensemble de cinq infirmières (I) travaillant sur une période de trois jours (J).

Tableau 2.2 : Résultat après optimisation de la fonction objectif à l’Éq. (2.2).

	Jour 1	Jour 2	Jour 3
Affectation	$i_1 \ i_2$	$i_3 \ i_4 \ i_5$	$i_1 \ i_2$

ALGORITHME DE SÉPARATION ET ÉVALUATION

Dans le cadre de ce projet, l'algorithme *Branch-and-bound*, ou séparation et évaluation, est utilisé comme méthode de résolution du modèle d'optimisation linéaire pour la confection d'horaire. Cette méthode a été proposée par Land & Doig (1960). L'algorithme fait partie de la catégorie de problèmes NP-Complets, signifiant un temps de résolution exponentiel dans les pires cas. Il existe plusieurs variations de l'algorithme (ex. *Branch-and-cut*, *Branch-and-price*), mais le principe de base reste le même.

La recherche de solution s'effectue à l'aide d'un arbre de recherche. La recherche est effectuée par un processus d'élagage, où chaque noeud d'un niveau est résolu avant d'explorer le niveau suivant. Pour obtenir les noeuds de l'arbre de recherche, le problème d'optimisation est subdivisé en sous-problème. Ces noeuds sont des sous-problèmes où une opération de relaxation est appliquée sur les variables entières. La relaxation des variables signifie que les variables entières ne sont plus imposées par ce domaine, mais qu'elles peuvent prendre des valeurs réelles. Cette opération permet de faciliter la résolution du modèle puisqu'un algorithme d'optimisation linéaire, tel le simplexe, peut être utilisé pour résoudre le problème.

La séparation du noeud s'effectue en créant de nouveaux noeuds selon les variables fractionnaires à la suite de la résolution de la relaxation. Ce faisant, la résolution du nouveau noeud aboutit nécessairement à une valeur entière pour le branchement effectué sur ce noeud. Les résultats des noeuds sont ensuite comparés avec la borne inférieure et supérieure. Si la solution d'un noeud se trouve à l'extérieur des bornes, l'exploration de ce noeud prend fin. Les bornes du problème doivent alors être remises à jour. La Figure 2.1 montre un exemple du processus de séparation et l'évaluation d'un problème fictif.

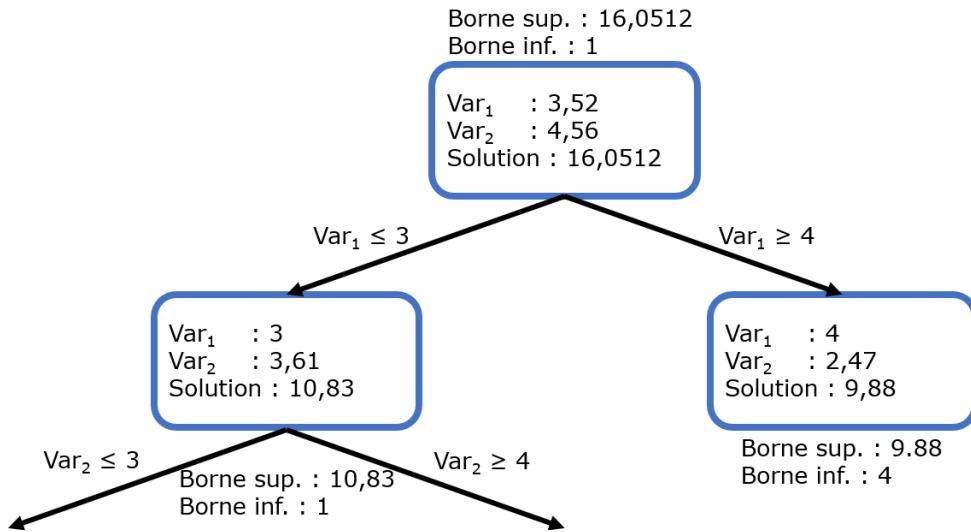


Figure 2.1 : Exemple de l'étape de séparation et évaluation d'un problème fictif d'optimisation en nombres entiers à l'aide de l'algorithme *Branch-and-bound*.

© 2021 Yoan Villeneuve.

Dans cet exemple, une opération de relaxation a été faite sur les variables Var_1 et Var_2 afin d'obtenir la solution 16,0512, soit la borne supérieure pour ce noeud. La solution relaxée est trouvée à l'aide de deux variables fractionnaires devant être transformées en entier. Deux sous-problèmes sont créés en récupérant la variable inférieure et supérieure entière de Var_1 . Pour chaque sous-problème, la valeur de Var_2 est calculée avec la solution de celui-ci. Comme l'objectif est de maximiser la solution, le noeud lorsque la variable Var_1 est égale ou plus grande que 4 est abandonné en faveur de celui de gauche. La valeur de la borne supérieure devient celle de la nouvelle solution maximale et l'opération de séparation recommence pour la valeur de Var_2 . En calculant les noeuds résultant de la séparation, il se peut que la valeur de Var_1 redevienne fractionnaire, devant ainsi continuer l'algorithme jusqu'à atteindre deux valeurs entières avec la valeur des bornes inférieures et supérieures égalisées.

La répétition de ce processus permet de trouver la solution optimale en nombres entiers lorsque les valeurs de la borne inférieure et supérieure sont égales entre elles. Il est aussi possible que la solution soit impossible ou encore qu'il y ait un ensemble de solutions possibles.

2.2.2 MODÈLES DE CONFECTION D'HORAIRE DANS LA LITTÉRATURE

Trilling *et al.* (2006) ont proposé un modèle linéaire de création d'horaires pour un groupe d'infirmières en anesthésiologie, formant un groupe de neuf employées qualifiées travaillant à temps plein et se partageant les horaires de jour et de nuit. En temps normal, les horaires sont faits par l'infirmière en chef. Celle-ci tente de construire un horaire à la main en tenant compte des préférences de chacune. Ces préférences sont considérées comme des contraintes dures, c'est-à-dire, des contraintes que le modèle d'optimisation doit absolument respecter. De plus, par souci d'égalité entre les membres du personnel, les horaires de nuit sont distribués de façons égalitaires entre eux. L'algorithme *branch-and-bound* est utilisé pour résoudre le problème. La fonction objectif du modèle doit permettre de distribuer les heures de travail d'une façon équitable entre chaque employée. Chaque infirmière obtient donc un nombre d'heures de travail similaire à celui de ses consoeurs avec une prise en considération de la difficulté des quarts de travail. Le but est donc de minimiser l'inégalité de distribution d'horaire entre le personnel soignant. Cependant, le modèle conçu ne permet pas d'atteindre la solution optimale dans un intervalle de temps raisonnable. Par conséquent, l'exécution du programme est arrêtée après une heure et la meilleure solution trouvée est retenue. Même après une exécution sur une période de 12 heures, la solution optimale n'a toujours pas été atteinte, mais la qualité de la solution retenue est dite semblable ou identique à celle retenue après la première heure de calcul. Toutefois, la solution obtenue après résolution reste intéressante puisqu'elle offre une amélioration par rapport aux horaires conçus par l'infirmière en chef.

Un autre article, par Widystiti *et al.* (2016), présente également un modèle d'optimisation linéaire de création d'horaires en milieu hospitalier. Ce dernier se base sur la qualification du personnel infirmier et l'objectif est de maximiser le nombre de jours de congé. L'étude se base sur un échantillon de 18 infirmières, où chacune possède deux qualifications différentes et peut être assignée à deux unités de soins demandant des qualifications particulières, soit

l’unité de soins intensifs et la salle d’urgence. Neuf infirmières détiennent la qualification 1 et peuvent être assignées autant à l’unité de soins intensifs que la salle d’urgence, tandis que les neuf autres infirmières ne peuvent seulement qu’être affectées à la salle d’urgence. Par rapport au modèle, le problème d’optimisation est constitué de quatre variables influencées par un ensemble de 14 contraintes. Les horaires sont conçus sur une période de 30 jours avec deux à quatre infirmières par plage horaire pour chacun des départements. Les résultats du modèle respectent toutes les normes de travail en place et satisfont les quotas de travail pour chaque quart à l’horaire. Les horaires générés sont comparés à ceux faits à la main, dans lesquels se trouvent plusieurs inégalités et incohérences dans le nombre d’heures travaillées. Chaque infirmière ferait six ou huit quarts de nuit sur 30 jours, mais l’horaire fait à la main contient une infirmière avec seulement cinq quarts de nuit. Cependant, aucune mention n’est faite par rapport au temps de calcul de l’algorithme ou au solveur utilisé. De plus, malgré la complexité du modèle, ce dernier ne considère pas les temps supplémentaires et ne semble pas prendre en considération la possibilité de travailler avec une unité d’infirmière en sous-effectif.

Un article de Mohan (2008) présent un modèle d’optimisation en nombres entiers dans un contexte où les employés à temps partiel représentent la majorité de la force de travail d’une entreprise. Le but est de pouvoir gérer les horaires des temps partiels tout en gardant un niveau de satisfaction élevé et un taux d’absentéisme bas. L’objectif principal est la conception d’horaires de travail qui maximise le niveau de satisfaction des employés. Une méthode *branch-and-cut* est utilisée pour maximiser la satisfaction des employés et répondre aux exigences de chaque quart de travail. En se basant sur un niveau de disponibilité allant de 25% à 75%, le logiciel d’optimisation CPLEX est capable de produire des horaires pour 50 et 100 employés à temps partiel. Malgré tout, le modèle prend seulement en compte les préférences d’heure de travail des employés. La prise en considération de l’état des employés peut ainsi être plus étendue pour les travaux futurs.

Brazeau (2013) a publié un mémoire portant sur la résolution du problème d'affectation des infirmières au sein de l'unité d'endoscopie digestive du centre hospitalier universitaire de Sherbrooke. Dans ce mémoire, un modèle d'optimisation linéaire en nombres entiers utilisant le solveur *COIN Branch-and-cut* (CBC) est développé en utilisant l'algorithme de séparation et d'évaluation progressive (*Branch-and-bound*). Le modèle permet d'obtenir des horaires mensuels à l'aide du solveur CPLEX avec l'objectif d'offrir des horaires plus stables et rapides à générer. Il contient plusieurs contraintes douces avec différents seuils de tolérance quant à la violation de ces dernières. Ces contraintes visent à favoriser la constance dans les quarts des infirmières en imposant, par exemple, l'affectation à une seule unité pour la durée complète du quart et d'un autre côté à éviter la surcharge de travail. Les résultats ont montré plusieurs avantages comparativement à la méthode de confection manuelle d'horaires puisque le modèle développé prend en compte plusieurs contraintes normalement difficiles à prendre en considération simultanément. Les résultats obtenus par le modèle sont équivalents ou dominent tous les objectifs étant normalement considérés par les ressources humaines lors de la confection d'un horaire de travail. L'auteur note toutefois quelques limitations du programme linéaire. Celui-ci ne permet pas l'ajout ou le retrait de contraintes, ni la modification du nombre d'infirmières affectées à une salle. De plus, les horaires générés sont ensuite modifiés à la main pour mieux plaire aux besoins des infirmières (congé maladie, échange de salle, etc.). En général, l'auteure déplore l'absence de possibilité permettant le paramétrage ou la modification de la part des gestionnaires et propose la création d'une interface plus interactive dans les travaux futurs.

Huang *et al.* (2016) proposent une approche en deux étapes à la construction d'horaires pour des infirmières en considérant plusieurs quarts par jour. Ceux-ci utilisent un modèle basé sur les horaires antécédents et, après la création d'un horaire, offrent aux infirmières la possibilité de modifier leurs affectations de quarts. L'objectif de ce modèle en deux phases

est la maximisation de la satisfaction des préférences d'horaire des infirmières. L'étude se base sur un modèle publié par Hsia (2009). Ce modèle cherche à concevoir un horaire respectant les besoins hospitaliers et les normes de travail des infirmières, tout en se basant sur les horaires précédemment utilisés. À partir de l'horaire généré, un second modèle est développé pour satisfaire les préférences d'horaire des infirmières. Ce modèle considère quatre niveaux d'absence ordonnés par importance. Dans l'ordre d'importance, il y a les congés reliés aux funérailles de famille proche et les vacances de mariage, suivi des absences reliées aux funérailles et aux mariages, suivi des vacances touristiques et enfin les absences ordinaires (rendez-vous à la banque, affaire gouvernementale, etc.). D'autres paramètres sont aussi considérés, comme les saisons estivales et les congés maladie. Le modèle a été testé sur une unité de pédiatrie fictive. Celle-ci est composée de 20 infirmières avec un besoin de 10 infirmières le jour et cinq infirmières le matin et la nuit. Les résultats obtenus satisfont la quasi-totalité des préférences d'horaire des infirmières et se calculent en moins d'une seconde pour chaque modèle. Cet article est très intéressant puisqu'il introduit un concept de niveau d'importance aux jours de congé. Cependant, le projet a aussi des limites. Par exemple, l'horaire doit être calculé à l'aide de deux modèles d'optimisation, comportant chacune différentes phases entre la résolution de chaque horaire, ce qui rend l'exécution de ce modèle complexe. Dans de futurs travaux, le tout pourrait être centralisé en un seul modèle, diminuant ainsi la complexité de résolution du problème.

Shuib & Kamarudin (2019) ont publié un modèle utilisant la programmation linéaire en nombres entiers pour résoudre le problème de planification d'horaires pour les employés d'une centrale électrique basée en Malaisie. L'étude a été effectuée sur 43 employés de l'un des départements de l'usine sur une durée de 28 jours comportant des quarts de travail de huit heures chacun, soit le matin, le jour et la nuit. Le modèle est multi-objectif avec des contraintes dures et douces. Les contraintes dures portent sur les besoins de la centrale et

les normes de travail en place, tandis que les contraintes douces sont les préférences de jour de congé des employés. Le problème est résolu avec le logiciel MATLAB et une solution optimale est atteinte en moins d'une minute. L'horaire obtenu montre que parmi les 172 préférences des employés, 138 ont démontré leurs satisfactions comparativement à 74 pour les horaires effectués à la main par un gestionnaire. Ce modèle se démarque par son utilisation de MATLAB pour résoudre le problème d'optimisation. Néanmoins, les préférences des employés ne sont pas considérées comme prioritaires par rapport au reste des contraintes. De plus, ces contraintes de préférence expriment seulement celle-ci relativement aux jours de congé. Une amélioration du modèle pourrait être faite en ajoutant les préférences de jour de travail comme contrainte souple, ou encore en considérant les préférences comme variables dures. D'autre part, la considération des préférences pourrait être étendue au-delà des jours de congé, tel que les quarts de travail.

2.2.3 AUTRES APPROCHES À LA CONFECTON D'HORAIRE

Certains problèmes d'optimisation présentent une difficulté inhérente faisant qu'ils ne peuvent être résolus de façon optimale dans des temps de calcul raisonnables par des méthodes exactes. À noter que ce travail n'utilise pas la métaheuristique comme approche de résolution pour la confection d'horaire optimisé. Or, il est pertinent d'observer les travaux accomplis à l'aide de cette approche, car l'utilisation de métaheuristique implique généralement des modèles hautement complexes. Les méthodes approchées, en particulier les métaheuristiques, permettent d'obtenir plus rapidement des solutions à ce genre de problèmes en sacrifiant la garantie d'optimalité, Gandomi *et al.* (2013). Il est à noter que ces méthodes permettent empiriquement de trouver des solutions de bonne qualité, voire les solutions optimales des problèmes considérés. En pratique, on se contente en général de ces solutions en raison de contraintes de temps de résolution. Une métaheuristique est une sorte d'ossature générale

à adapter au problème à résoudre. Ainsi, une même métaheuristique peut être appliquée de façons très différentes lorsqu'adaptée à deux problèmes différents. Parmi les métaheuristiques les plus populaires se trouvent les algorithmes évolutionnaires, le recuit simulé, les algorithmes d'optimisation par essaims particulaires et la recherche avec tabous, Silberholz & Golden (2010).

Legrain *et al.* (2015) ont publié un article sur le problème de génération d'horaires des hôpitaux du Québec. Plus précisément, l'étude porte sur les équipes régulières et flottantes des hôpitaux Notre-Dame et Sainte-Justine situés dans la métropole de Montréal. L'équipe flottante est le groupe d'infirmières qui couvre le manque de personnel dans les différents secteurs d'un hôpital ou des hôpitaux voisinant. Une infirmière flottante n'appartient donc pas à un secteur particulier et peut changer de secteur à chaque nouvel horaire. Il est noté que les préférences des infirmières sont généralement sous forme de jour de travail ou de congé. L'article propose que le quota d'infirmières ne soit pas une contrainte dure, tout en essayant de rester le plus près de la valeur du quota (contrainte douce). Les auteurs proposent une heuristique simple d'utilisation, pouvant être implémentées sur des feuilles de calcul Excel, demandant ainsi peu d'investissement de la part du planificateur d'horaire. Le planificateur se base sur trois contraintes liées aux normes de travail, aux quotas de travail et aux préférences des infirmières. Deux méthodes de résolution ont été utilisées : soit avec un modèle sur le logiciel d'optimisation commerciale CPLEX et avec une méthode heuristique. Les résultats montrent que le solveur CPLEX atteint une solution de manière quasi instantanée comparativement à une demi-journée pour l'heuristique. Malgré cette économie de temps, la méthode heuristique obtient des résultats aussi satisfaisants. Les objectifs de préférence et d'équité sont atteints. En raison d'une collecte de données difficile et dite inefficace, les auteurs suggèrent que le planificateur établisse des règles claires lorsque la liste des préférences est prévue et une date limite après laquelle plus aucun changement ne sera accepté.

Aux États-Unis, Bard & Purnomo (2005) proposent une solution au manque grandissant de personnel soignant dans les hôpitaux américains en prenant en compte leurs préférences pour ainsi augmenter le taux de rétention. Ce modèle multi-objectif est résolu avec une approche par génération de colonnes qui combine la programmation linéaire en nombres entiers avec une méthode heuristique. Les colonnes correspondent à des horaires alternatifs d'une infirmière sur l'horizon de planification. Une heuristique de permutation double est utilisée pour générer chaque colonne. Le but principal de ce modèle est de réussir à concevoir des horaires respectant au maximum les préférences de travail du personnel. Le modèle prend en considération cinq types de quarts. Ainsi, une journée de 24h sera divisée en trois quarts de 8h et deux quarts de 12h. Le manque de personnel est aussi considéré dans ce modèle en incluant plusieurs variables correspondant au nombre d'infirmières externes et à leurs périodes de séjour. Dans le but d'avoir un modèle équitable pour chaque infirmière, une variable du modèle assure que le nombre de violations de leurs préférences reste équitable. Le problème commence avec N colonnes correspondant à l'horaire préféré de chaque infirmière. Puis, si la demande ne satisfait pas les besoins en horaire, le nombre d'horaires possible (colonnes) par infirmière augmente exponentiellement avec différentes violations de préférence. La méthode de génération de colonnes a été codée en C++ en incluant un appel au solveur CPLEX. Le modèle développé est capable de générer un horaire pour une centaine d'infirmières sur quatre semaines en quelques minutes et avec un faible taux de violation des préférences. La création de plusieurs horaires pour chaque infirmière suivis d'une optimisation heuristique de la solution s'avère être robuste. Même si cet article n'est relativement pas récent, il semble apporter des éléments permettant de résoudre le problème de manque de personnel infirmier. Cela dit, le modèle qui y est proposé nécessite une mise à jour ainsi qu'une adaptation au contexte des hôpitaux du Québec.

Pour tenter de remédier au problème de construction d'horaires pour le personnel infirmier dans les salles d'opération, Xiang (2017) a conçu un algorithme métaheuristique multi-objectif. Les objectifs de ce modèle considèrent plusieurs facteurs par rapport aux patients, aux personnels et à la gestion. Pour les patients, le *makespan*, soit la durée entre l'entrée du premier patient et la sortie du dernier, et le *Throughput*, qui représente le nombre total d'opérations effectuées, sont les variables considérées. Le modèle cherche donc à minimiser le *makespan* et maximiser le *Throughput*. Pour les infirmières, les heures supplémentaires ainsi que les préférences de quart de travail sont les problèmes principaux. Enfin, la gestion des infirmières et des salles d'opération demande d'éviter les situations de surutilisation ou sous-utilisation des ressources. L'optimisation par colonies de fourmis est utilisée comme méthode de résolution, Dorigo & Blum (2005). Elle permet de trouver des ensembles de solutions non-dominées dits ensembles de Pareto. Le modèle conçu est comparé avec un projet antérieur par Xiang *et al.* (2015) ainsi qu'un autre article par Mobasher *et al.* (2011) sur le sujet. La première comparaison confirme que le modèle conçu atteint une valeur *makespan* optimale en plus d'avoir une meilleure balance des ressources allouées. La deuxième comparaison montre une diminution du temps de travail pour les infirmières des salles d'opération. Quatre valeurs de rendement sont utilisées pour cette comparaison : l'heure de fin, le coefficient de variation du temps de travail, les heures supplémentaires maximales des infirmières et les heures supplémentaires totales des infirmières.

2.3 PRÉDICTION DE DONNÉES ET APPRENTISSAGE MACHINE

L'intelligence artificielle, plus particulièrement l'apprentissage machine, est un outil privilégié par les chercheurs afin de réaliser la tâche de prédiction de données numériques. La prédiction de données est l'analyse de plusieurs variables constituant un ensemble de données afin de prédire la ou les valeurs recherchées (catégorielles ou numériques) selon un contexte

fixe. Pour effectuer une prédiction de l'absentéisme, il nous faut donc analyser le profil temporel de chaque infirmière afin d'en extraire les variables significatives. Cette information servira à fournir une métrique de risque. Pour cela, l'étude repose sur l'exploitation de données historiques.

Le langage de programmation Python est reconnu pour avoir une grande diversité de libraires et d'interfaces de programmation. Ces derniers sont semblables à des blocs de programmation externes au langage de base que l'on peut rajouter à notre environnement de programmation pour effectuer différentes tâches. Parmi celles-ci, *scikit-learn* (Buitinck *et al.*, 2013) est une librairie permettant de faire de l'apprentissage machine avec différents algorithmes et modèles mis à disposition. Parmi ces modèles, des algorithmes de classification, de régression et de *clustering* sont disponibles. Pour ce projet, nous avons choisi d'exploiter les algorithmes des forêts d'arbres décisionnels (Random Forest). Ce choix a été fait en raison du type de données à disposition ainsi que de la taille de l'ensemble de données.

Les forêts d'arbres décisionnels (*Random Forest*), Ho (1995), se séparent en deux catégories : les méthodes de classification et celles de régression. L'objectif d'un algorithme de classification en apprentissage machine est de classifier une instance de données par rapport aux caractéristiques la définissant. Par exemple, on cherche à classifier une infirmière "absente" ou "présente" en analysant les caractéristiques de travail des jours antérieurs, comme le cumul des quarts travaillés ou le temps supplémentaire effectué. Les forêts d'arbres décisionnels sont reconnues comme étant parmi les algorithmes les plus efficaces lorsqu'on a peu de données pour effectuer l'apprentissage d'un modèle de prédiction.

2.3.1 ARBRES DÉCISIONNELS

Une forêt d’arbre décisionnel est composée de plusieurs arbres de décision. Un arbre décisionnel consiste en une suite d’opérations de décision sur une instance de données menant à sa classification selon un ensemble de classes prédéfinies. Il existe plusieurs algorithmes d’arbres de décision telle que ID3 (Quinlan, 1986), C4.5 (Quinlan, 1993) et CART (Breiman *et al.*, 1984). Pour effectuer une classification, il faut avoir un ensemble de données X constitué de n instances et de m variables. Une instance x_i ($i = [1, 2, 3, \dots, n]$) représente une entrée de l’ensemble de données (une ligne) et les variables servent à définir l’information de chaque donnée contenue dans une entrée. À partir des variables X , il nous faut déterminer une classe $y_i \in Y$, soit la variable à prédire. Comparable aux autres variables, la classe permet de définir une instance. Celle-ci est une variable prédite à partir des données d’une instance. Une donnée peut être sous forme numérique ou catégorielle . La Figure 2.2 présente une visualisation d’un ensemble de données d’horaire de travail de trois employés.

Diagram illustrating the structure of a data set:

- Instances, lignes.** (Instances, lines) points to the rows of the table.
- Ensemble de données catégoriques.** (Categorical data set) and **Ensemble de données numériques.** (Numerical data set) point to the first four columns of the table.
- Variables prédictives.** (Predictive variables) points to the last column of the table.
- Ensemble de données.** (Data set) points to the entire table.

Table data:

Employé (no.)	Jour	Quart	Heures Travailées	Temps Supplémentaires
1	Lundi	Jour	8.5	Oui
2	Lundi	Jour	8	Non
3	Lundi	Soir	7.5	Non
2	Mardi	Jour	8	Non
3	Mardi	Jour	9	Oui
1	Mardi	Soir	7.8	Non
3	Mercredi	Jour	8.5	Oui
...

Figure 2.2 : Visualisation d'un ensemble de données.

© 2021 Yoan Villeneuve.

Pour effectuer une classification, il faut concevoir un modèle en effectuant un apprentissage. Un arbre de classification est un modèle permettant d'effectuer la classification d'un ensemble de données. Il est représenté graphiquement comme une arborescence orientée selon la théorie des graphes soit acyclique (impossible de relier deux points par deux chemins distincts) et connexe (tous les sommets du graphe sont accessibles entre eux). Chaque sommet de l'arbre représente un noeud. En partant de la racine de l'arbre (noeud de départ), on fournit une instance à l'arbre de classification. Chaque noeud représente une variable de l'instance, où un test est effectué sur la valeur de cette variable. Ce test mène à d'autres noeuds jusqu'à atteindre une feuille de l'arbre, représentant la classification de l'instance fournie. L'exemple à la Figure 2.3 montre un arbre décisionnel effectuant la classification d'une instance de la Figure 2.2. À chaque niveau de l'arbre, une opération sur une variable de l'instance est effectuée jusqu'à atteindre une classification au bas de l'arbre.

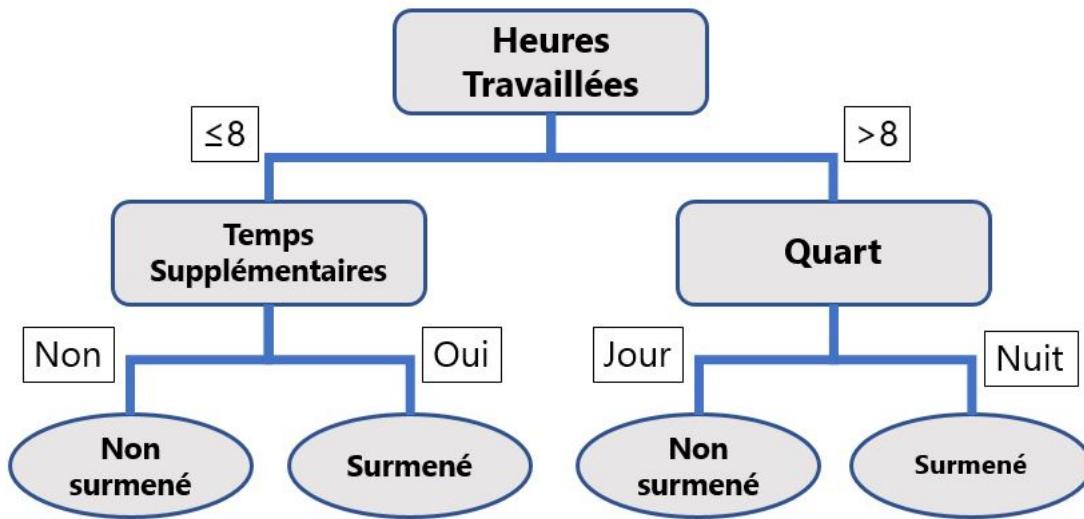


Figure 2.3 : Exemple d'arbre décisionnel classifiant le surmenage.

© 2021 Yoan Villeneuve.

Lorsque l'on veut construire un arbre de décision ID3, Quinlan (1986), il nous faut un ensemble de données comportant plusieurs variables définissant chaque instance de l'ensemble. D'abord, une étape de fractionnement (*splitting*) est effectuée sur les données. Cette étape consiste à effectuer des tests sur les variables de l'ensemble afin de générer des branches pour chaque résultat possible. Pour les valeurs numériques, des opérateurs relationnels ($<$, $=$ ou $>$) sont utilisés. Lorsque chaque branche a été reliée à une feuille, une opération d'élagage est réalisée, ce qui permet d'éviter les biais ayant pu se former sur une variable particulière. Cette méthode permet aussi de repérer les variables les plus significatives pour la prédiction de la variable cible. Pour les valeurs catégoriques, le branchement est effectué par rapport aux valeurs de classifications possibles pour la variable catégorique. Le choix des variables pour construire l'arbre est fait en calculant l'entropie pour chaque variable. L'entropie mesure l'incertitude dans un ensemble de données, représente généralement une valeur de zéro à un et est calculée à l'aide de l'Éq. (2.6) :

$$E(Var) = \sum_{i=1} -s_i \log_2 s_i. \quad (2.6)$$

La valeur de s_i représente la division du nombre d'instances présentes dans les sous-ensembles générés par le branchement de l'arbre sur la variable Var et le nombre total d'instances. Pour connaître la prochaine variable pour le branchement de l'arbre, le gain d'information GI de l'Éq. (2.7) est calculé en faisant la différence entre l'entropie du noeud parent et le poids moyen de l'entropie avec une autre variable :

$$GI = E(Var_1) - E(Var_1|Var_2). \quad (2.7)$$

La variable offrant le plus grand gain d'information constitue le prochain noeud de l'arbre. L'entropie est le gain d'information est alors recalculer pour les prochains noeuds de l'arbre.

2.3.2 FORÊT D'ARBRES DÉCISIONNELS ALÉATOIRES

Les arbres décisionnels ont tout de même leurs défauts. Ceux-ci sont reconnus pour sur-apprendre les données d'apprentissage. Ce phénomène est plus connu sous le terme de sur-apprentissage (*overfitting*), où un algorithme devient trop performant à prédire son propre ensemble de données pour l'entraînement et le test du modèle. Par conséquent, ce phénomène cause une baisse de performance lorsque de la prédiction est effectuée sur de nouvelles données de la même nature. Pour remédier à ce problème, l'algorithme des forêts d'arbres décisionnels aléatoires a été développé (*Random Forest*). Ce dernier est plus robuste face au phénomène de sur-apprentissage. Le principe des forêts d'arbres décisionnels aléatoires a d'abord été proposé par Ho (1995) avant d'être proprement abordé dans un article par Breiman (2001). Expliqué

simplement, il consiste à construire plusieurs arbres de classification avec l'algorithme *CART* (*Classification And Regression Trees*), de Breiman *et al.* (1984). L'algorithme CART est une méthode de confection d'arbres de décision avec la particularité d'être binaire (deux fils par noeud) et d'utiliser l'indice de Gini pour l'évaluation des noeuds, au lieu du gain d'information vu dans la sous-section précédent.

Dans une forêt d'arbres décisionnels aléatoires, on cherche à générer plusieurs arbres aléatoires. Les arbres aléatoires qui composent la forêt sont tous similaires sans être identiques. La force de l'algorithme réside sur l'aspect aléatoire appliqué sur ses arbres ainsi que dans son intelligence de groupe. Pour remédier au défaut de l'*over-fitting* des arbres, chaque arbre est entraîné avec un sous-ensemble de données choisi au hasard. De plus, le choix des variables utilisées pour chacun des arbres est aussi fait de façon aléatoire. Ce concept permet d'obtenir des résultats souvent très fiables, puisque la faiblesse d'un arbre est compensée par ses voisins. Dans une forêt aléatoire, les arbres composant la forêt donnent leur avis sur la valeur de la variable à prédire et le choix de la prédiction en revient à un vote de la majorité. Un modèle régressif (qui cherche à prédire une valeur numérique) effectue une moyenne des scores obtenus dans la forêt.

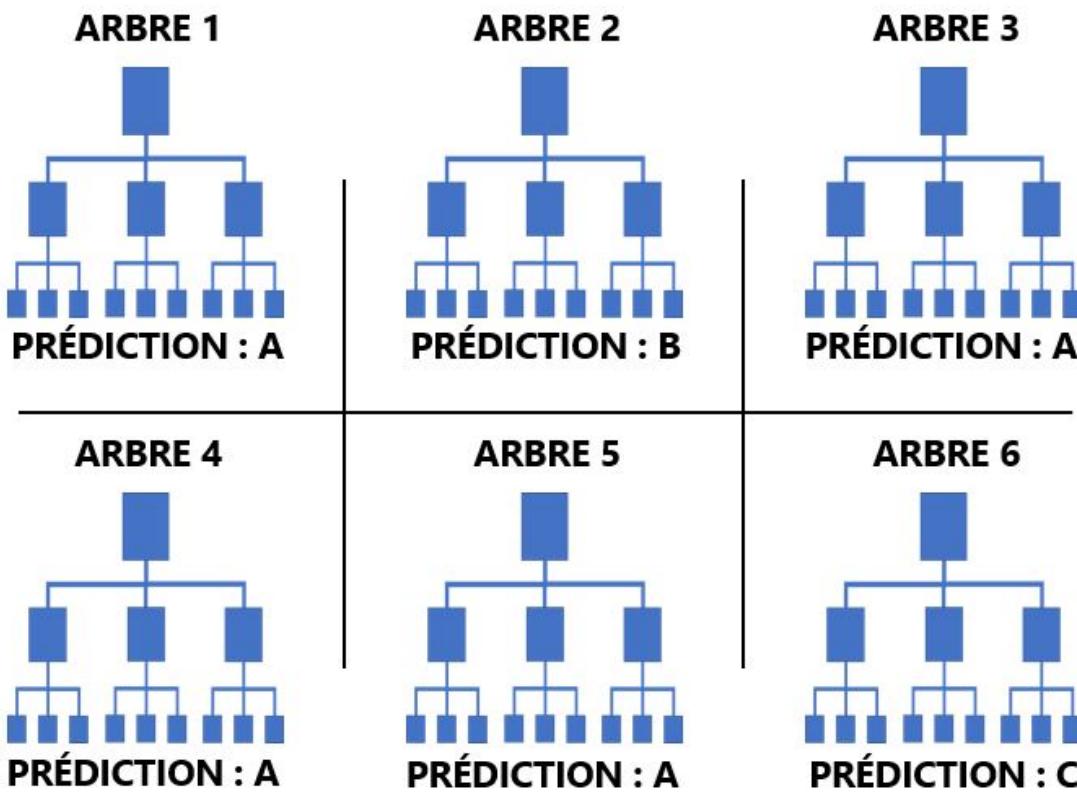


Figure 2.4 : Visualisation d'une forêt d'arbres décisionnels à trois variables de classification.

© 2021 Yoan Villeneuve.

La Figure 2.4 montre une forêt d'arbres de classification à trois classes. Pour donner suite à la prédition de chaque arbre, un vote est fait et la prédition la plus fréquente devient la classification finale du modèle. Dans cet exemple, le modèle classifie une donnée à la valeur A pour avoir été prédit quatre fois sur les six arbres. L'intelligence de groupe, où les arbres se protègent entre eux de leurs erreurs, permet une prédition plus fiable, mais au risque d'une baisse de la précision. Ce type d'algorithme est très prisé lorsqu'il faut travailler avec de petits ensembles de données tests ou ayant de faibles corrélations entre les variables.

2.3.3 PRÉDICTION DE L'ABSENTÉISME

Un article de Tewari *et al.* (2020) aborde le sujet de la prédiction de l'absentéisme à l'aide de l'apprentissage machine. Dans cet article, il est soutenu que l'absentéisme peut être le résultat de différents facteurs d'un point de vue personnel ou professionnel. Des facteurs tels que la dépression, le stress et l'anxiété, qui mènent au travail à des interactions sociales difficiles, un sentiment de surcharge, de l'intimidation ou un sentiment de sous-appréciation. La régression linéaire et les machines à vecteurs de support pour la régression (Support Vector Regression (SVR)) sont les deux algorithmes d'apprentissage utilisés dans ce travail. Les données exploitées pour effectuer leur prédiction sont en libre accès et disponible sur *Kaggle* (www.kaggle.com). Les données rapportent 741 observations de 21 paramètres concernant de l'information personnelle et professionnelle. L'article se divise en quatre parties, soit l'exploration des données, le prétraitement des données, l'agencement des données et l'apprentissage machine. Les paramètres ayant été considérés après l'exploration des données sont l'absence (oui ;non), la catégorisation de l'employé (1,2 ou 3), le nom, l'âge, le sexe, la distance entre la maison et le travail, les transports saisonniers, le temps de travail, les quarts de jour, la charge de travail et le niveau d'éducation. Après avoir appliqué la régression linéaire sur chaque paramètre, une relation entre l'âge d'un individu et l'absentéisme a été observée. En effectuant la même analyse avec la méthode SVR, il a été découvert que la tranche d'âge de 35 à 40 ans a le plus haut taux d'absentéisme, contrairement au groupe de 50 à 60 ans. En ajoutant le paramètre des jours de la semaine, le plus haut taux d'absentéisme se trouve au 6^e jour de la semaine et un minimum observé au second jour de la semaine. Une analyse du taux d'absentéisme par rapport aux saisons a aussi été réalisée, mais les résultats montrent peu de fluctuation entre les saisons.

Un autre article par Dogruyol & Sekeroglu (2020) propose l'application des Réseaux de Neurones Artificiels (RNA) pour tenter de prédire l'absentéisme sans préavis. Trois méthodes d'entraînement de RNA sont appliquées à la problématique : un Réseau de Neurones Classique (RNC), Backpropagation (BP), Radial Basis Function (RBF) et Long Short-Term Memory (LSTM). L'ensemble de données se nomme *absenteeism at work* provenant de *UC Irvine Machine Learning Repository* (<https://archive.ics.uci.edu/>), soit un ensemble de 740 instances de 20 variables amassées entre 2007 et 2010. Deux échantillons de 60% et 70% de l'ensemble de données est pris comme test d'ensemble d'entraînement, utilisant le reste de l'ensemble comme échantillon test. Ce type d'échantillonnage est propre à la méthode *hold-out* pour l'entraînement et le test de modèle en apprentissage machine. Les critères d'évaluation sont basés sur le score pour la régression (R^2), l'erreur quadratique moyenne (Mean Square Error (MSE)) et la Variance Expliquée (EV). Les résultats obtenus montrent une claire supériorité du réseau de neurones LSTM, suivi par le RBF et le RNC. Les auteurs ont observé que le modèle LSTM a l'avantage d'avoir une mémoire à court terme permettant d'apprendre les relations temporelles dans les données chronologiques. En considérant que plusieurs articles ont déjà démontré que l'absentéisme au travail est un phénomène progressif résultant d'une accumulation de facteurs négatifs (professionnel ou non), il semble très probable que le modèle puisse identifier des *pattern* menant à une augmentation des chances d'absentéisme d'une infirmière. Malheureusement, en raison de la nature complexe des réseaux neuronaux, il est difficile d'interpréter le mécanisme interne menant aux solutions obtenues. Dans les travaux futurs, les auteurs proposent d'effectuer d'autres tests avec les Machines à vecteurs de support (SVR) et les arbres décisionnels, puis de comparer les résultats avec ceux obtenus.

CHAPITRE III

MÉTHODOLOGIE

Dans le cadre de ce projet, une méthode de génération d'horaire est développée pour les infirmières en milieu hospitalier afin de réduire les arrêts de travail et l'absentéisme. Pour ce faire, le projet est constitué de trois étapes. Premièrement, un nettoyage et une analyse des données historiques obtenues grâce à la collaboration du CIUSSS ont été faits. Deuxièmement, un modèle de prédiction du bien-être des infirmières est créé. Enfin, un modèle d'optimisation permettant la création d'horaires pour les unités infirmières par rapport à un ensemble de paramètres et de contraintes a été développé. Ainsi, ce chapitre fera mention des différents moyens et algorithmes utilisés afin d'obtenir un modèle d'optimisation d'horaire pour les infirmières avec la prédiction du bien-être. La Figure 3.1 montre les étapes de développement pour la création d'un horaire optimisé pour le bien-être des infirmières.

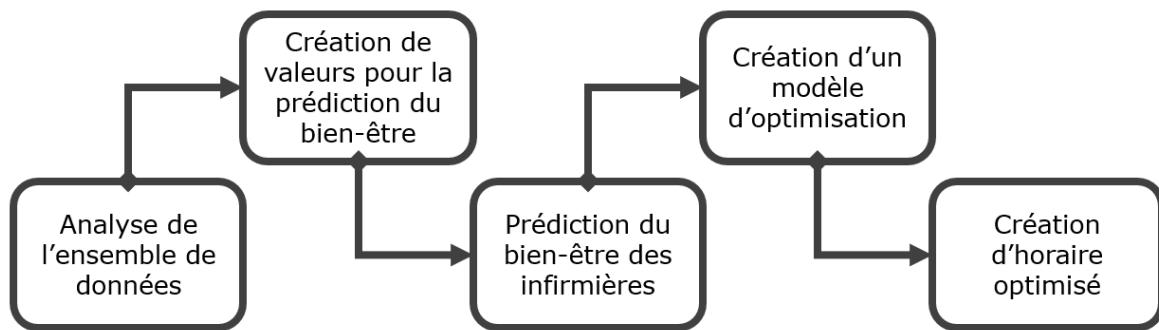


Figure 3.1 : Processus de développement du modèle de génération d'horaire optimale avec valeur de prédiction de l'absentéisme.

© 2021 Yoan Villeneuve.

3.1 MANIPULATION ET ANALYSE DES DONNÉES HISTORIQUE

Pour concevoir un modèle de prédiction du bien-être, il faut avant tout des données historiques d'horaires dans le milieu infirmier. Ces données seront utiles à l'analyse de la problématique de confection d'horaire ainsi qu'à la conception de variables utile à la prédiction du bien-être. Il a été possible de recevoir ces données grâce à la collaboration du CIUSSS du Saguenay—Lac-Saint-Jean, un organisme public chargé par le ministère de la Santé et des Services sociaux visant à fournir divers services de santé dans leur région attitrée. Ceux-ci ont fourni plusieurs données provenant d'une unité de soins de l'hôpital de Roberval.

3.1.1 RÉCEPTION ET NETTOYAGE DES DONNÉES

Parmi les fichiers de données, le fichier HORAIRE.xlsx présente une année de données historiques sur les horaires de l'unité. Plus précisément, le fichier contient les données historiques de chaque employée ayant travaillé dans le département de soins sur la période du 28 décembre 2018 au 28 décembre 2019. Le Tableau 3.1 présente les attributs et variables présents dans l'ensemble de données, ainsi que le type de données de chacune d'elle. Le fichier contient un total de 11314 instances, où chaque instance représente un quart de travail d'un employé.

Tableau 3.1 : Table des variables de l'ensemble de données HORAIRE.xlsx .

Variable	Type de données
DATE_JOUR	Date
MATRICULE	Chaîne de caractères
UNITEADMCODE	Nombre entier
TITRE EMPLOI	Nombre entier
QUART_CODE	Chaîne de caractères
J_S_N	Caractère
HR_DEBUT	Nombre entier
HR_FIN	Nombre entier
HR_EN_MOINS	Nombre entier
HR_TOTAL	Nombre entier
HR_MIN	Nombre à virgule flottante
CODE_PAIE	Nombre entier
INDICE_PAIE	Nombre entier
NOTE	Chaîne de caractères
ASSIGNATION_CODE	Chaîne de caractères

Puisque les données reçues sont à l'état brut, une phase de nettoyage est nécessaire avant de pouvoir interpréter et analyser celles-ci. Les colonnes UNITEADMCODE, NOTE et ASSIGNATION CODE ont été retirées de l'ensemble par manque de pertinence pour notre analyse. Ces informations ne sont pas pertinentes pour ce projet. Par exemple, la variable UNITEADMCODE égale toujours 605151. Les variables NOTE et ASSIGNATION CODE sont de l'information dédiée au personnel responsable de la gestion d'horaire. Ces variables sont rarement utilisées et l'information qu'ils fournissent ne peut être interprétée (ex : *ES*, *med* ou *BON+*). De plus, les données comprises du 28 décembre 2018 au 3 février 2019 exclusivement ont été supprimées de l'ensemble de données en raison de modifications apportées à la convention collective en vigueur depuis le 3 février 2019. Cette suppression de données a été effectuée par souci de cohérence entre les données. Ainsi, le nombre d'instances exploitables pour l'analyse de l'ensemble de données est de 10225. Le Tableau 3.2 montre un échantillon de ces données après nettoyage.

Tableau 3.2 : Échantillon de l'ensemble à la suite du nettoyage des données.

DATE_JOUR	2019-02-03	2019-02-03	2019-02-03	2019-02-03
MATRICULE	EMPLOYÉ 93	EMPLOYÉ 96	EMPLOYÉ 41	EMPLOYÉ 50
TITRE EMPLOI	3480	7471	3455	3480
QUART_CODE	S14	N23 ³ / ₄	N00	S14
J_S_N	s	n	n	s
HR_DEBUT	14 :00 :00	23 :45 :00	00 :00 :00	14 :00 :00
HR_FIN	22 :00 :00	08 :15 :00	08 :00 :00	22 :00 :00
HR_EN_MOINS	00 :45 :00	00 :30 :00	00 :45 :00	00 :45 :00
HR_TOTAL	07 :15 :00	07 :30 :00	07 :15 :00	07 :15 :00
HR_MIN	7.25	7.5	7.25	7.25
CODE_PAIE	961	256	961	235
INDICE_PAIE	3	0	3	8

Après le nettoyage de l'ensemble de données, il faut extraire l'information pertinente par rapport à chaque instance de données. L'énumération suivante décrit chaque variable et leur signification :

- DATE_JOUR : La date en années, mois et jour où l'instance du quart s'est déroulée.
- MATRICULE : Le code pour différencier les infirmières entre elles. Le matricule remplace le nom des employées, une démarche qui a été faite par le CIUSSS afin d'assurer la confidentialité des infirmières utilisées dans ce projet.
- TITRE EMPLOI : Le poste occupé par un employé pour chaque instance enregistrée.
- QUART_CODE : La période et l'heure à laquelle un quart commence. La période est indiquée par *J* pour jour, *S* pour soir et *N* pour nuit. Les chiffres et fractions à la suite représentent l'heure à laquelle le quart a débuté (ex. 23 23 heures et 45 minutes).
- J_S_N : La période à laquelle un quart commence. La période est indiquée par *J* pour jour, *S* pour soir et *N* pour nuit.
- HR_DEBUT : L'heure de début du quart.
- HR_FIN : L'heure de fin du quart.

- HR_EN_MOINS : Le nombre d'heures retiré de la paie. Généralement, ce temps est la pause repas, car celle-ci n' est pas rémunérée.
- HR_TOTAL : Le nombre d'heures total du quart travaillé.
- HR_MIN HR_TOTAL : Le nombre d'heures total du quart travaillé en nombre.
- CODE_PAIE : Permet de connaître le type de quart assigné à l'instance de données.
- INDICE_PAIE : Spécification du quart travaillé, relié à la variable CODE_PAIE.

À partir de l'information fournie par la variable TITRE EMPLOI, le poste de chacune des employées de l'unité peut être identifié. L'unité est constituée de plusieurs employées par poste, mais comme ce projet porte sur les horaires des infirmières, les données sont isolées en deux sous-ensembles d'infirmière, soit l'unité d'infirmière et l'unité d'infirmière auxiliaire. Après avoir observé les données, il s'y retrouve 16 infirmières et 14 infirmières auxiliaires. L'information par rapport aux quotas de travail des unités est fournie dans les données historiques du CIUSSS. Le quota représente le nombre d'employés nécessaire à chaque quart de travail. Les données de l'étude contiennent l'information sur les besoins en employés pour chaque titre d'employé par rapport à la date, le jour de la semaine, le moment de la journée (jour, soir ou nuit) et l'heure de commencement du quart de travail. Le Tableau 3.3 résume l'information pertinente de ce fichier par rapport aux infirmières (7471) et aux infirmières auxiliaires (3455). La variable DÉBUT représente le moment auquel un quart commence. La variable JOUR permet de déterminer les jours de la semaine affectés par le quota. Les variables J_S_N et DÉBUT informent sur la période de la journée et l'heure de début du quart. La variable QUOTA indique le nombre d'infirmières nécessaire pour la période du quart, selon le type d'infirmière (TEMP_CODE).

Tableau 3.3 : Quotas de travail des infirmières (7471) et infirmières (3455).

TEMP_CODE	J_S_N	DÉBUT	JOUR	QUOTA
3455	Nuit	00H00	[0...6]	1
3455	Jour	08H00	[0...6]	3
3455	Soir	16H00	[0...6]	2
7471	Nuit	23H45	[0...6]	2
7471	Jour	07H45	[0...6]	4
7471	Soir	15H45	[0...6]	3

Une fois les infirmières séparées en deux sous-ensembles, l'étape suivante consiste à sélectionner chaque quart de l'ensemble de données afin de mieux l'étudier. L'ensemble de données contient non seulement de l'information par rapport au quart travaillé par les infirmières, mais aussi de l'information sur les temps supplémentaires, les congés, les absences et d'autres classifications de quart se rapportant à un horaire de travail (ex. l'assurance salaire). Par exemple, la première colonne du Tableau 3.2 fait référence à l'employé 93 ayant effectué un quart de soir de 14 à 22 heures. Ce quart est étiqueté avec le code de paie 961 ainsi qu'un indice de paie définie à 3. Ces codes font référence à un quart de remplacement régulier à long terme, signifiant un quart de travail classifié à temps régulier dans ce projet. Il y a plus de 326 combinaisons de codes et indices de paie fournies avec les données. Pour gérer cette quantité importante de codes et indices de paie, ceux-ci ont été catégorisés selon les types de quarts les plus prédominants ayant été observés. Chaque instance de travail est classifiée comme un quart régulier, de l'équipe volante, du temps supplémentaire, du Temps Supplémentaire Obligatoire (TSO), un jour de congé ou de vacances planifié, du chômage, une absence, un congé maladie ou de l'assurance travail. Cette classification permet une meilleure analyse des horaires des infirmières de l'unité et permet de concevoir le modèle de prédiction du bien-être dans la suite de ce projet.

3.1.2 ANALYSE DES DONNÉES

Avec un ensemble de données propre et étiqueté selon le quart effectué, une analyse de cet ensemble permet de mieux visualiser le comportement des infirmières et les enjeux vécus par celle-ci.

FRÉQUENCE DE TRAVAIL DES INFIRMIÈRES

Comme nous avons les données sur les horaires des infirmières sur presque toute l'année 2019, il est pertinent d'observer la fréquence à laquelle les infirmières des unités travaillent. Ainsi, la Figure 3.2 montre un graphe décrivant le nombre de quarts travaillés chaque jour. Chaque barre indique le nombre d'infirmières s'étant présentées au travail durant cette journée. Un quart travaillé peut être étiqueté comme du temps régulier, temps supplémentaire, TSO ou de l'équipe volante. Si une infirmière a travaillé deux quarts ou plus dans une même journée (par exemple, un temps régulier de jour et supplémentaire de soir), celle-ci est comptée comme une seule infirmière dans le cumulatif du nombre d'infirmières ayant travaillé.

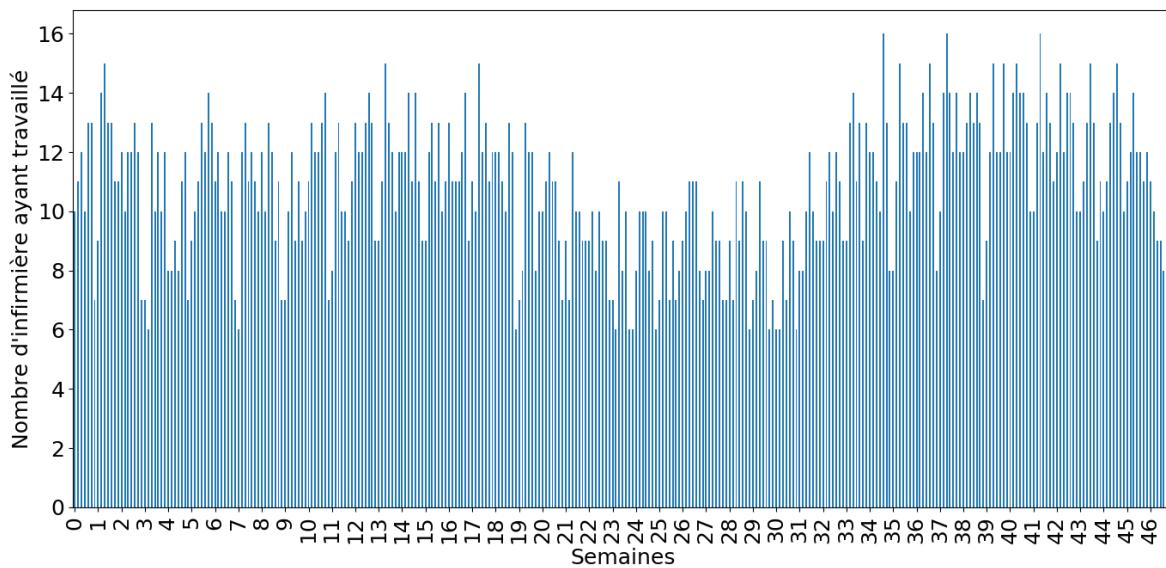


Figure 3.2 : Nombre d'infirmières ayant effectué un quart de travail pour chaque journée entre le 3 février 2019 et le 28 décembre 2019.

© 2021 Yoan Villeneuve.

Comme le quota de travail est fixe au courant de l'année, il ne devrait pas y avoir de fluctuation majeure dans le nombre d'infirmières au travail. Or, certaines périodes de l'année ont une baisse de personnel distinct, avec des fluctuations allant de cinq à 16 infirmières selon la journée observée. En général, 8 à 14 infirmières semblent toujours venir au travail, avec une médiane se trouvant à 11. Les semaines 18 à 33 ont un taux particulièrement bas d'infirmières présentes par jour. Cette période est comprise entre les mois de juin à septembre. Cet intervalle de temps compte de 6 à 13 employés présents au travail par jour. Cela signifie que cette période ne remplit jamais son quota de travail, avec certaines journées comportant moins de la moitié du quota nécessaire au fonctionnement des unités. Certaines journées dans la figure semblent fluctuer de manière consistante, ce qui laisse porter à croire que certains jours de la semaine voient un nombre différent d'infirmière active. Le Tableau 3.4 montre la médiane du nombre d'infirmières ayant travaillé pour chaque jour de la semaine.

Tableau 3.4 : Médiane du nombre d'infirmière ayant effectué au minimum un quart de travail selon le jour de la semaine.

Jour	Médiane
Dimanche	9
Lundi	11
Mardi	12
Mercredi	11
Jeudi	12
Vendredi	11
Samedi	9

Selon les quotas de travail vu dans le Tableau 3.3, un total de 15 quarts de travail sont à remplir pendant une journée (la nuit, le jour et le soir). En se fiant à la médiane, il y a toujours moins d'infirmières par jour que le quota de travail. Cela implique que certaines infirmières doivent compenser pour le manque de personnel, ou alors, que le département fonctionne en sous-effectif. L'un ou l'autre, ces situations restent problématiques en augmentant la charge de travail. Les samedis et dimanches se démarquent des jours de semaine avec deux à trois infirmières en moins que ces derniers.

RÉPARTITION DES QUARTS PAR INFIRMIÈRE

Puisque les infirmières sont toutes étiquetées avec leurs quarts de travail et d'absence, une analyse de chaque infirmière permet de cerner leur comportement au courant de l'année 2019. La Figure 3.3 permet d'observer le nombre d'heures affectées à une infirmière selon le type de quart de celle-ci. Chaque heure est groupée dans les barres de la figure selon son type. Le type *Autre* représente les quarts n'appartenant à aucune autre classification de quart existant, mais ceux-ci représentent une promotion minime des quarts totaux.

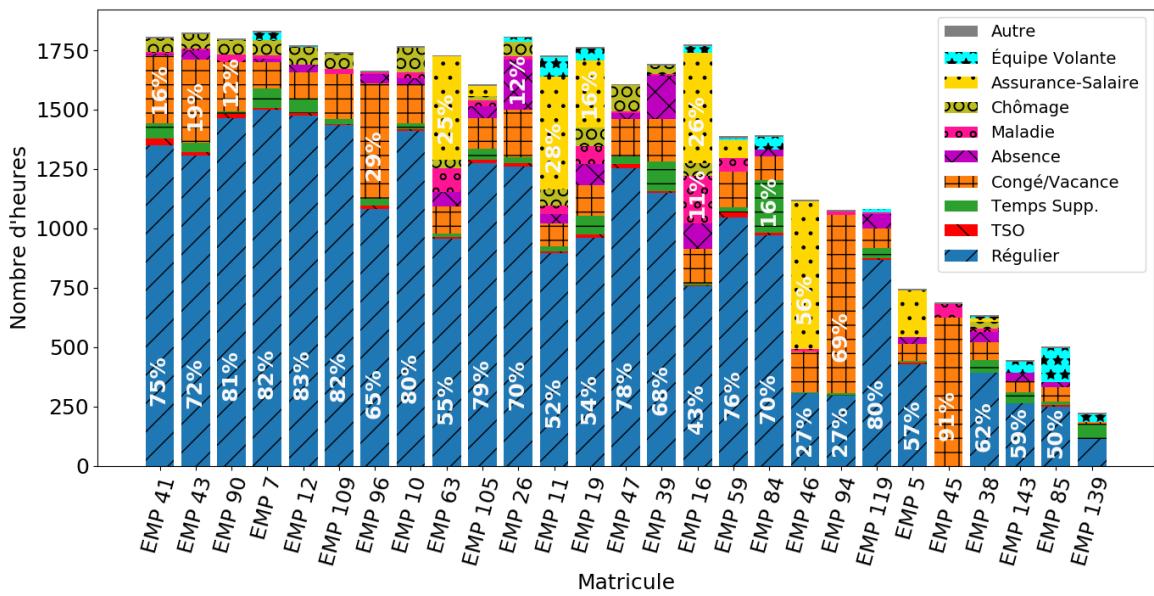


Figure 3.3 : Proportion des quart affectés pour chaque infirmière par rapport au nombre d'heures par affectation de quart.

© 2021 Yoan Villeneuve.

Les infirmières étant catégorisées selon le nombre de temps supplémentaire travaillé, il ne semble pas y avoir de différence entre les infirmières lorsque l'on observe la différence entre l'unité infirmière et l'unité d'infirmière auxiliaire. Les infirmières semblent presque toutes avoir certains points en commun. D'abord, le nombre d'heures de temps supplémentaire et de TSO semble assez bas, atteignant rarement plus de 4% des heures cumulées. Or, ce n'est pas autant le nombre d'heures que la fréquence des temps supplémentaires qui affectent négativement une infirmière, Rogers *et al.* (2004).

En termes de congés et vacances, ils représentent 6% à 11% du nombre d'heures total à l'horaire par infirmière. Les infirmières 10, 41 et 96 se démarquent par un nombre plus élevé de congés et vacances, mais ne semblent pas avoir d'autres différences par rapport au reste des infirmières. En revanche, les infirmières 46, 94 et 45 se démarquent. L'infirmière 46 a effectué beaucoup d'assurance salaire, signifiant un arrêt de travail. En inspectant celle-ci, elle était en arrêt de travail dès le début de l'enregistrement des données étudié, avec un retour au

travail progressif vers la fin de la période de données disponible. Tandis que les infirmières 94 et 45 ont leurs congés étiquetés comme étant des congés parentaux. En raison d'un manque de pertinence par rapport au besoin de ce projet, les infirmières 45 et 46 ont été retirées du problème de prédiction du bien-être.

Les absences non planifiées au travail sont présentes chez la majorité des infirmières de la figure. Les infirmières avec le plus grand nombre d'heures à temps régulier semblent n'avoir que peu de cas d'absences non planifiés, mais semblent de plus en plus présentes lorsque le temps régulier est à la baisse. Les infirmières 26 et 39 se démarquent le plus en rapport à leur absence, mais ne semblent pas avoir davantage de différence avec le reste des infirmières. Autrement, il ne semble pas y avoir de corrélation visible entre les heures d'absence et le reste des heures de travail. La corrélation entre les variables de l'ensemble est détaillée à la Sous-Section 4.1 du chapitre des résultats.

Les infirmières ayant fait un arrêt de travail sont facilement identifiables dans la figure 3.3. L'assurance salaire est le titre des quarts attribué lorsqu'une infirmière ne peut se présenter au travail sur une période d'invalidité de courte ou longue durée. Ici, les infirmières 19, 63, 11, 16, 5 et 46 présentent toutes un taux d'assurance salaire élevé par rapport au reste des heures attribuées. Ces infirmières représentent les cas devant être évités grâce au modèle développé au cours de ce projet. L'infirmière 46 est déjà considérée comme non pertinente pour le projet. L'infirmière 5 ne pourra pas être un cas d'arrêt de travail analysable, puisque celle-ci était déjà en assurance salaire lorsque les données historiques débutent. Ce qui laisse les infirmières 11, 16, 19 et 63 à analyser. L'analyse de celle-ci est effectuée à la Sous-Section 4.3.2 en parallèle avec le modèle de prédiction du score de bien-être.

Une autre classe de types de quarts peut être analysée pour les infirmières. Comme il est supposé que le nombre d'heures attribué par type de quart ait un impact sur le personnel

infirmier, il est supposé que la période de travail puisse aussi jouer un rôle dans les performances des infirmières. Semblable à la figure précédente, la Figure 3.4 montre la proportion du nombre d'heures ayant été travaillées selon la période de la journée. Ce nombre d'heures travaillées inclut le temps régulier, le temps supplémentaire et le TSO.

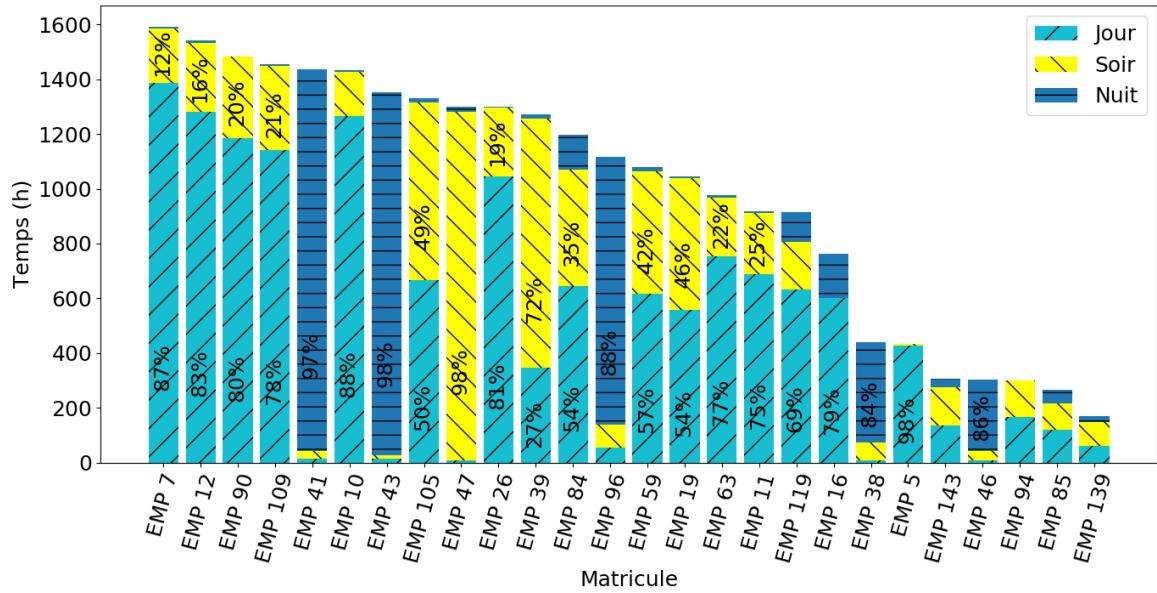


Figure 3.4 : Proportion du nombre d'heures travaillées pour chaque infirmière relativement à la période d'affectation du quart.

© 2021 Yoan Villeneuve.

À première vue, certaines infirmières sont affectées à des périodes fixes avec très peu de quart travaillés à l'extérieur de ces périodes. Ceci est relié au contrat de travail des infirmières, où celles-ci sont employées pour travailler sur des périodes précises selon le contrat. En revanche, certaines infirmières semblent vivre le contraire, pouvant atteindre une différence d'environ 50% des quartes hors de leur contrat. Par rapport à la répartition des quartes, les infirmières de jour et de soir semblent partager beaucoup de leur période de travail entre elles, tandis que les infirmières de nuit ne font pratiquement aucun quart à l'extérieur de la période de leur contrat. En se fiant aux proportions des données de la figure 3.4, il n'y a que deux

infirmières de soir, où la majorité des quarts semble être divisée avec le reste des infirmières de jour. En regardant les infirmières ayant fait des arrêts de travail, soit l'infirmière 11, 16, 19 et 63, rien ne sort de l'ordinaire par rapport aux autres infirmières, sauf un nombre d'heures de travail similaire.

Finalement, cette analyse a permis de cerner les infirmières ayant fait un arrêt de travail et d'observer le comportement de ces dernières afin de mieux orienter les étapes suivantes.

3.2 MODÈLE DE PRÉDICTION DU BIEN-ÊTRE

Cette section porte sur le développement du modèle de prédiction du score de bien-être du personnel infirmier. L'objectif de celui-ci est d'offrir un score reflétant l'état dans lequel une infirmière se trouve. Cet état est déterminé sous forme d'un score nommé *score de bien-être*. Différentes étapes sont accomplies afin d'obtenir un modèle de prédiction. D'abord, un nouvel ensemble de données est construit avec l'information nécessaire à la prédiction du bien-être des infirmières. Par la suite, la formule permettant de calculer le bien-être d'une infirmière est expliquée. Enfin, deux étapes de manipulation sont appliquées sur l'ensemble de données, une pour permettre au modèle de prédire le score de bien-être et une phase de nettoyage de l'ensemble.

3.2.1 CRÉATION D'UN NOUVEL ENSEMBLE DE DONNÉES POUR LA PRÉDICTION DU BIEN-ÊTRE

L'objectif de ce projet est de concevoir un modèle générant des horaires qui minimiseront l'absentéisme des infirmières. À cet effet, l'algorithme d'apprentissage machine impose que les données qui lui sont fournies soient numériques. Cet ensemble de données quantifie le nombre de quarts assignés à chaque infirmière en catégorisant ceux-ci. Une instance de l'ensemble

de données reflète un horaire de deux semaines pour chaque infirmière. Une variable appelée SCORE sera notre étiquette, soit la variable cible à prédire par le modèle. Le calcul du score est présenté à la section 3.2.2. Cette valeur est comprise dans l'intervalle de zéro à 10. Le Tableau 3.5 montre une instance de données de l'ensemble pour la prédiction.

Tableau 3.5 : Exemple d'une instance de données fournies au modèle d'apprentissage machine.

MATRICULE	WEEK	Q_JOUR	Q_SOIR	Q NUIT	RÉGULIER
EMPLOYÉ 7	30	7	2	0	6
TSO	TEMP_SUPP	ABSENCE	CONGÉ_VACANCE	CHOMAGE	MALADIE
0	3	0	0	0	0
ASSURANCE	VOLANTE	SURMENAGE	CONSÉCUTIF	HORS CONTRAT	SCORE
0	0	8	3	2	4,5

Dans cette instance, seule la variable MATRICULE n'est pas fournie au modèle de prédiction. L'ensemble ne contient pas assez de données pour espérer construire un modèle pouvant focaliser sur le comportement de chaque infirmière individuellement. Néanmoins, cette information est importante dans les étapes précédant l'entraînement d'un modèle de prédiction. La variable WEEK représente la semaine à laquelle une instance fait référence. Le reste des données, excluant CONSÉCUTIF, SURMENAGE et SCORE, sont un cumulatif de tous les quarts attribués à l'infirmière (MATRICULE) au courant de la semaine (WEEK). Voici un descriptif de chaque variable de l'ensemble de données :

- Q_JOUR, Q_SOIR, Q NUIT : Cumul à l'horaire des quarts réguliers travaillé à l'horaire pour le jour, le soir et la nuit.
- RÉGULIER : Cumul des quarts travaillés à temps régulier.
- TSO : Cumul à l'horaire des quarts travaillés à temps supplémentaire obligatoire à l'horaire.

- TEMP_SUPP : Cumul à l'horaire des quarts travaillés à temps supplémentaires à l'horaire.
- ABSENCE : Cumul à l'horaire des absences non planifiées à l'horaire.
- CONGÉ_VACANCE : Cumul à l'horaire des absences planifiées étiquetées comme congés ou vacances à l'horaire.
- CHOMAGE : Cumul à l'horaire des absences liées au chômage à l'horaire.
- MALADIE : Cumul à l'horaire des congés maladie à l'horaire
- ASSURANCE : Cumul à l'horaire des congés en assurance salaire à l'horaire.
- VOLANTE : Cumul à l'horaire des quarts de travail liés à l'équipe volante.

Les variables CONSÉCUTIF, SURMENAGE et HORS CONTRAT sont calculées en utilisant les informations par horaire (instance).

- CONSÉCUTIF : Total des jours travaillés consécutivement (jour après jour) dans l'instance de l'horaire concernée. La valeur de la variable est égale au nombre de jours excédant cinq jours travaillés consécutivement. Ce choix a été fait suite aux recommandations et à l'expertise des membres de l'entreprise Timesphere, étant spécialisée dans le domaine de la confection d'horaire de travail. De plus, Bae & Fabry (2014) suggère une plus grande limitation face au nombre de quarts de travail hebdomadaire par infirmière.
- SURMENAGE : Cumul des quarts considérés problématiques pour une infirmière au courant des six semaines (trois instances). Le calcul de cette variable est expliqué à la Sous-Section 3.2.2.
- HORS CONTRAT : Le nombre de quarts travaillés à temps régulier ou en équipe volante hors de la période fixée par leur contrat de travail. Chaque infirmière travaille selon son contrat de travail pour travailler durant un ou parfois deux quarts de travail à une période précise de la journée.

Enfin, la variable SCORE représente l'étiquette de chaque instance de données ainsi que la valeur à prédire par le modèle de prédiction. Le calcul du score de bien-être est détaillé à la Sous-Section 3.2.2. Finalement, cet ensemble de données est constitué de 647 instances de données chacune elle-même constituée de 18 variables.

3.2.2 CRÉATION D'UN SCORE DE PRÉDICTION DU BIEN-ÊTRE

Cette section aborde le score de bien-être. Celui-ci est utilisé comme paramètre dans le modèle d'optimisation afin de reconnaître et ainsi accommoder les infirmières à risque. Pour ce faire, trois facteurs de risque influencent la valeur du score : le nombre de quarts de travail en dehors du contrat de travail (F_1), le cumul du surmenage des six dernières semaines (F_2) et le nombre d'absences non planifié (F_3). Les particularités de chaque sous-équation des facteurs de risque sont discutées dans les sous-sections suivantes. L'Éq. (3.1) montre la formule utilisée pour calculer le score :

$$Score = 10 - (10 \times F_1 \times F_2 \times F_3), \quad (3.1)$$

Où chacune des variables est liée à un facteur de risque, soit une sous-équation calculée avec les données de l'instance concernées. F_1 , F_2 et F_3 représentent toute une valeur entre 0 et 1 inclusivement, de sorte à toujours obtenir une valeur de score allant de 0 à 10. Les facteurs n'ont pas tous la même pondération, de sorte à prioriser certaines situations pouvant affecter le bien-être par rapport à d'autres.

F1 : RISQUE LIÉ AUX QUARTS DE TRAVAIL HORS CONTRAT

Les quarts travaillés hors du contrat de travail représentent un risque pour les infirmières. Chaque infirmière est engagée avec un contrat de travail représentant le ou les quarts principaux auxquelles elle doit être assignée. Ainsi, une infirmière engagée pour travailler les quarts de soir doit en théorie toujours travailler durant les quarts de soir. En réalité, il est commun dans les données historiques que les infirmières travaillent au moins un quart hors de leur contrat de travail. Cela risque d'augmenter l'insatisfaction de l'infirmière face à ses horaires de travail et augmenter le risque à l'intention de quitter son emploi (Berger & Hobbs, 2006). Ainsi, les quarts de travail hors contrat devraient être évités et chaque quart de ce genre travaillé représente une diminution du bien-être de l'infirmière. Pour chaque horaire, la somme des quarts réguliers travaillés hors contrat est calculée, puis l'Éq. (3.2) est calculée :

$$F_1 = \frac{RÉGULIER - \frac{HORS CONTRAT}{3}}{RÉGULIER}, \quad (3.2)$$

Où RÉGULIER et HORS CONTRAT représentent les variables de l'ensemble de données. La valeur de F_1 varie entre 0.67 et 1, où 0.67 représente un horaire de travail où une infirmière n'a travaillé durant aucun de ses quarts principaux selon son contrat. L'influence de cette variable est réduite en divisant HORS CONTRAT par 3. Ce choix a été fait afin de diminuer l'impact de F_1 par rapport aux deux autres facteurs, à la suite des résultats obtenus après avoir expérimenté sur l'ensemble de données historiques.

F2 : RISQUE LIÉ AU SURMENAGE

La surcharge du personnel infirmier est une situation bien réelle affectant un grand nombre d'infirmières dans les hôpitaux du Québec (Gagné, 2017). Comme vu dans la revue

de littérature à la Section 2.1, la surcharge est une cause majeure de la dégradation de la santé mentale des infirmières (Vagharseyyedin *et al.*, 2011).

La variable SURMENAGE n'est pas issue de l'ensemble de données historiques. Elle représente la somme de tous les quarts de travail considérés négatifs pour la santé d'une infirmière. Ces quarts négatifs incluent le temps supplémentaire, le TSO et l'équipe volante. La variable CONSÉCUTIF est aussi ajoutée dans le calcul du SURMENAGE, puisqu'elle représente un risque de surcharge lorsqu'une infirmière travaille plus de cinq jours de suite. Enfin, le contrat de travail est aussi considéré. Si une infirmière travaille plus de quarts que le maximum selon son contrat de travail, la variable *reg_surplus* est égale au nombre de quarts travaillés en surplus (ex : une infirmière avec un contrat de 10 quarts obtient une valeur de *reg_surplus* de deux pour avoir travaillé 12 quarts). Enfin, une infirmière ne devient pas surmenée sur une courte période, mais plutôt sur une accumulation de plusieurs semaines et horaires de travail. L'équation suivante exprime le surmenage de l'horaire h :

$$SURMENAGE_h = TEMP_SUPP + TSO + VOLANTE + CONSÉCUTIF + reg_surplus, \quad (3.3)$$

Où h représente l'horaire sur 14 jours de l'instance analysée. La variable SURMENAGE cumule les valeurs des deux $SURMENAGE_h$ des instances d'horaires précédents, soit $SURMENAGE_{h-1}$ et $SURMENAGE_{h-2}$. Ainsi, SURMENAGE est calculé sur une période de six semaines (incluant l'horaire de l'instance courante) à l'aide de l'équation :

$$SURMENAGE = SURMENAGE_h + SURMENAGE_{h-1} + SURMENAGE_{h-2}. \quad (3.4)$$

Dans les données historiques, la valeur la plus haute atteinte par une infirmière est 16. La variable SURMENAGE est utilisée afin de calculer la variable F_2 en effectuant une

division de la variable par 20. Le choix de la valeur de division a été fait par rapport au maximum obtenu par une infirmière. Dans l'ensemble de données, la valeur la plus haute atteinte par une infirmière est 16, mais il serait possible qu'une infirmière atteigne une valeur de SURMENAGE plus élevée, d'où le choix de placer la variable de division à 20. La variable F_2 est calculée avec l'Éq. (3.5) :

$$F_2 = 1 - \frac{\text{SURMENAGE}}{20}, \quad (3.5)$$

Où F_2 est une valeur entre 0 et 1 en considérant SURMENAGE plus petit que 20. Une valeur de SURMENAGE de 20 et plus est extrêmement improbable puisque la valeur maximale de temps supplémentaire est de 7 et que la médiane se trouve à 0.

F3 : RISQUE LIÉ AUX ABSENCES NON PLANIFIÉES

Le dernier facteur F_3 prend en considération les absences non planifiées à l'horaire. Ce projet vise à éviter les absences non planifiées liées à la surcharge. Cette situation est à éviter et représente un indicatif très révélateur de la détresse d'une infirmière. L'impact d'une absence non planifiée doit grandement pénaliser le score de bien-être d'une infirmière, car selon Jennings (2008), les absences au travail peuvent être causé par un grand niveau de stress. De plus, plusieurs absences non planifiées sur une courte durée sont une indicatrice qu'une infirmière est en surcharge. Ainsi, si l'absence d'une infirmière est due à un grand stress vécu au travail, les facteurs F_1 et F_2 refléteront aussi le stress de l'infirmière analysée. L'équation de F_3 se calcule ainsi :

$$F_3 = \left(1 - \frac{\text{ABSENCE} + \text{CHOMAGE} + \text{MALADIE} + \text{ASSURANCE}}{10} \right)^{1,5}, \quad (3.6)$$

Où la valeur possible de F_3 se trouve dans l'intervalle de 0 à 1. Selon les données de l'ensemble, la somme des absences non planifiées n'est jamais plus grande que 10. Les seules instances atteignant un F_3 de 0 sont celles où une infirmière est en arrêt de travail, soit une variable ASSURANCE égale à 10. Les congés malades (MALADIE) sont inclus dans le calcul puisque cette variable indique que l'état de santé d'une infirmière est précaire. Le facteur 1,5 est utilisé afin d'augmenter l'impact des absences non planifiées sur le score de bien-être final.

DÉMONSTRATION DU CALCUL

Afin de bien comprendre le calcul du score de bien-être lié à une instance de données, l'instance pour la semaine 26 de l'infirmière 12 est utilisée dans cette sous-section afin de montrer le fonctionnement de l'Éq. (3.1). Le tableau 3.6 contient la valeur de chaque variable décrivant la semaine 26 de l'infirmière 12 dans l'ensemble de données de prédiction.

Tableau 3.6 : Exemple d'une instance de données de l'infirmière 12 pour l'horaire du 4 août au 17 août 2019.

MATRICULE	WEEK	Q_JOUR	Q_SOIR	Q NUIT	RÉGULIER
EMPLOYÉ 12	26	6	3	0	9
TSO	TEMP_SUPP	ABSENCE	CONGÉ_VACANCE	CHOMAGE	MALADIE
0	2	1	0	0	0
ASSURANCE	VOLANTE	SURMENAGE	CONSÉCUTIF	HORS CONTRAT	SCORE
0	0	12	3	3	3,04

Le facteur F_1 demande de connaître le nombre de quarts à temps régulier travaillés ainsi que le nombre de quarts à temps régulier travaillés hors contrat. L'infirmière 12 étant une infirmière de jour, elle a effectué un total de trois quarts hors contrat sur neuf quarts à temps régulier. En appliquant ces données à l'Éq. (3.2), le facteur F_1 se calcule :

$$F_1 = \frac{RÉGULIER - \frac{HORS CONTRAT}{3}}{RÉGULIER} = \frac{9 - \frac{3}{3}}{9} = 0,88. \quad (3.7)$$

Le facteur F_2 calcul la surcharge de l'infirmière à l'aide de la variable SURMENAGE. Puisque cette variable se base sur les horaires antérieurs, celle-ci a déjà été calculée lors de la création de l'ensemble de données de prédition, pour une valeur de 12. À l'aide de l'Eq. (3.5), le facteur F_2 se calcule :

$$F_2 = 1 - \frac{SURMENAGE}{20} = 1 - \frac{12}{20} = 0,4. \quad (3.8)$$

Le facteur F_3 calcul les quarts d'absence à être évité dans le quart de ce projet. Durant cet horaire, l'infirmière 12 a été absente une seule fois sans avoir été planifiée au préalable. Ainsi, l'Eq. (3.6) pour la variable F_3 se calcule :

$$\begin{aligned} F_3 &= \left(1 - \frac{ABSENCE + CHOMAGE + MALADIE + ASSURANCE}{10} \right)^{1,5} \\ &= \left(1 - \frac{1+0+0+0}{10} \right)^{1,5} = 0,85. \end{aligned} \quad (3.9)$$

Par suite du calcul de chaque facteur, ceux-ci ont appliqué à l'Eq. 3.1 afin d'obtenir le score de bien-être de cette infirmière pour l'horaire de la semaine 26. Le score de bien-être est égal à :

$$\begin{aligned} Score &= 10 - (10 \times F_1 \times F_2 \times F_3) \\ &= 10 - (10 \times 0,89 \times 0,4 \times 0,85) = 3,04. \end{aligned} \quad (3.10)$$

Cette valeur est emmagasinée dans l'ensemble de données de prédition comme étiquette pour l'instance de données sous la variable SCORE.

3.2.3 AJUSTEMENT DE L'ENSEMBLE DE DONNÉES POUR LA PRÉDICTION

Dans la Sous-Section 3.2.2, il est discuté que chaque instance de données est reliée à une étiquette décrivant le score de bien-être d'une infirmière par rapport à l'information disponible dans l'instance. Or, le modèle de prédiction du score de bien-être doit être utilisé afin de prédire le score d'une infirmière à la semaine suivante par rapport aux horaires précédents. Ainsi, l'étiquette d'une instance ne doit pas être le score de l'instance avec lequel il a été calculé, mais plutôt le score de la semaine suivante. La Figure 3.5 montre la manipulation faite sur l'ensemble de données.

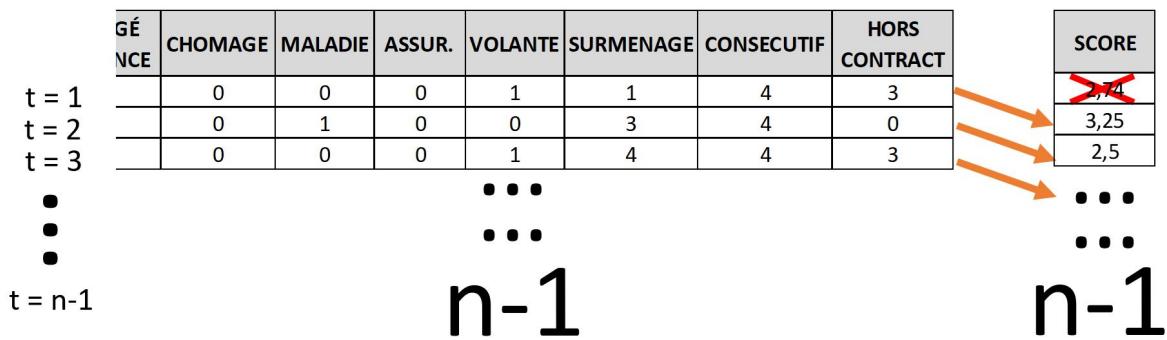


Figure 3.5 : Mise au point de la position de l'étiquette et des instances de l'ensemble de données afin de prédire le score de bien-être.

© 2021 Yoan Villeneuve.

Pour ce faire, l'ensemble est séparé en sous-ensembles pour chaque infirmière présente. Pour chaque sous-ensemble, la variable cible SCORE a été retirée de l'ensemble de données afin de pouvoir manipuler ceux-ci séparément. Par la suite, la première instance de l'ensemble de variables SCORE ainsi que la dernière instance de l'ensemble de données sont retirées. Enfin, le score est remplacé dans le sous-ensemble, avec maintenant l'instance de la semaine t liée au score de la semaine $t + 1$. Lorsque cette opération est terminée pour toutes les infirmières, l'ensemble de données est reconstruit. Les données sont maintenant prêtes à être

fournies en entraînement au modèle de prédiction du score de bien-être, qui prédira le score de chaque infirmière pour la semaine suivante, se basant sur l'horaire courant.

3.2.4 NETTOYAGE DE L'ENSEMBLE DE DONNÉES DE PRÉDICTION

Avant l'entraînement du modèle de prédiction, une étape de nettoyage des données est nécessaire. Certaines instances ne sont pas pertinentes à la prédiction du score de bien-être, représentant une forme de biais pour le modèle. Cette sous-section explique les démarches appliquées afin d'épurer l'ensemble de données pour la prédiction d'un score de bien-être.

Plusieurs instances comportant de l'assurance salaire ont été retirées de l'ensemble de données pour la prédiction. Lorsqu'une infirmière est en arrêt de travail, cela se traduit à l'horaire comme étant 10 quarts d'assurance salaire. Cela implique que plusieurs instances de l'ensemble de données représentent un bruit pour le modèle de prédiction, soit de l'information pouvant biaiser le modèle. De plus, il n'est pas pertinent de faire la prédiction d'une infirmière étant en arrêt de travail. La variable reste tout de même pertinente, puisqu'elle indique au modèle qu'une infirmière est en arrêt de travail. Comme un horaire est construit sur deux semaines, il faut aussi considérer qu'une infirmière peut partir en arrêt de travail après une semaine. Ainsi, toutes les instances ayant un cumul de cinq assurances salaires ou plus sont retirées de l'ensemble de données. Ce choix a été fait après avoir effectué plusieurs tests de performance expliquée à la Section 4.1.

La création de l'ensemble de données de prédiction a été automatisée en langage *Python*. Chaque instance de l'ensemble représente l'information d'un horaire de travail par infirmière. Or, certaines instances ne sont pas pertinentes pour la création d'un modèle de prédiction du bien-être des infirmières. Ces instances ont été retirées sous deux conditions, soit :

1. l'infirmière n'effectue aucun quart de travail dans l'instance,
2. si le point 1 est vrai, retirer l'instance si la valeur de SCORE est égale à 0.

En général, ces instances représentent des horaires d'infirmière vides ou sans quart de travail. Au courant des tests effectués sur le modèle, ces instances se sont avérées nuisibles à la prédiction du score des infirmières, car l'ensemble de prédiction est trop petit, ce qui représente un bruit pouvant biaiser la prédiction.

Pour faire suite aux modifications apportées à l'ensemble de données dans cette section et la Sous-Section 3.2.1, l'ensemble de données final pour le modèle de prédiction du bien-être est constitué de 533 instances. Les sections suivantes montrent comment l'information contenue dans cet ensemble de données est appliquée à l'algorithme d'apprentissage machine Random Forest (RF) pour la régression.

3.2.5 PROGRAMMATION DU MODÈLE DE PRÉDICTION DU BIEN-ÊTRE

Afin de prédire le score de bien-être pour chaque infirmière, un algorithme d'apprentissage machine est utilisé. L'apprentissage machine est un domaine appartenant à l'intelligence artificielle. À l'aide d'une phase d'entraînement, le modèle analyse en profondeur les données qui lui ont été fournies relatives à la solution qui s'y rattache, soit l'étiquette. La performance d'un modèle est calculée par rapport à un autre ensemble d'instances de données en comparant les réponses issues du modèle d'entraînement et la réalité. Dans le cadre de ce projet, l'algorithme de Random Forest est utilisé comme modèle d'apprentissage machine, afin de prédire le score de bien-être. De plus amples détails sur le fonctionnement de cet algorithme se trouvent à la Sous-Section 2.3.2 de la revue de littérature. La suite de cette section concerne la manipulation de l'ensemble de données et la méthodologie appliquée afin d'obtenir la prédiction du score.

APPLICATION DE L'ALGORITHME DE FORêt D'ARBRES DÉCISIONNELS ALÉATOIRES

Dans le domaine de l'Intelligence Artificielle (IA), les algorithmes d'arbre décisionnel sont reconnus pour leur performance avec de petits ensembles de données. L'ensemble de données généré pour la prédiction du score de bien-être comporte 535 instances de données. Dans le cadre de ce projet, la méthode d'ensemble est utilisée afin de consolider les performances du modèle de prédiction.

La méthode d'ensemble consiste à construire un ensemble d'arbres décisionnels aléatoires formant ainsi une forêt. Chaque arbre fournit une prédiction ou une décision, selon des variables sélectionnées aléatoirement. Cela permet d'obtenir un nombre défini de prédictions pour le vote de la valeur prédite par l'ensemble d'arbres de la forêt. Les arbres s'appuient entre eux afin d'éliminer le bruit causé par les arbres ayant une valeur prédite erronée. Pour la méthode d'ensemble appliquée à la régression, ce principe est réutilisé. Un ensemble de RFR est construit, où la valeur obtenue est la moyenne de la prédiction pour chaque forêt régressive. Afin d'obtenir une forêt différente à chaque nouvelle itération, un ensemble d'entraînement et de test comportant un ensemble d'infirmières unique est utilisé. La Figure 3.6 montre le principe utilisé afin de créer un ensemble d'entraînement et de test pour la création d'une forêt d'arbres de régression.

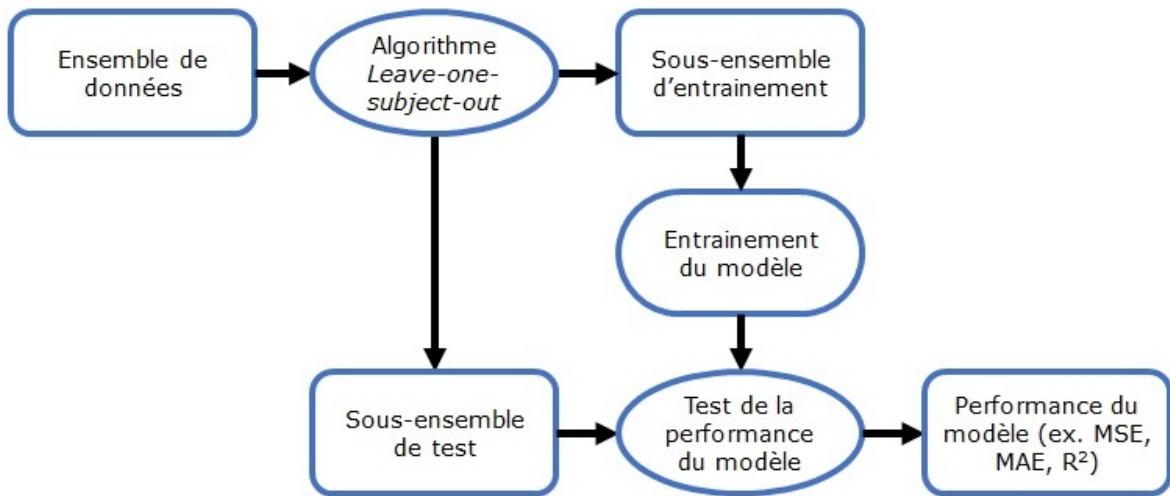


Figure 3.6 : Processus de création d'un ensemble d'entraînement et de test pour la création d'une forêt d'arbres de régression.

© 2021 Yoan Villeneuve.

Afin d'obtenir les ensembles nécessaires pour l'entraînement et le test, l'ensemble de données construit pour le modèle de prédiction est divisé en deux sous-ensembles. Une variation de l'algorithme *leave-one-subject-out* est utilisée afin de construire les sous-ensembles d'entraînement et de test. Cet algorithme vise à retirer un seul individu de l'ensemble d'entraînement pour le placer dans l'ensemble test. Ainsi, le procédé du choix des infirmières pour la taille des ensembles d'entraînement et de test pour la prédiction est discuté dans la Section 4.1.

HYPERPARAMÈTRAGE DU MODÈLE

La plupart des algorithmes d'apprentissage machine ont des hyperparamètres permettant de modifier l'entraînement d'un modèle afin d'influencer la performance pour la classification ou la régression. Les hyperparamètres sont des paramètres qui influence la phase d'apprentissage d'un modèle d'apprentissage machine, Claesen & Moor (2015). En d'autres termes, le réglage des hyperparamètres permet de raffiner le modèle d'apprentissage machine. Dans

ce projet, la librairie *scikit-learn* (Buitinck *et al.*, 2013) est utilisée avec la fonction *RandomForestRegressor* afin de générer un modèle. La recherche d'hyperparamètre est faite de façons aléatoires avec la fonction *RandomizedSearchCV*. Plusieurs ensembles de paramètres possibles sont fournis à cette dernière fonction afin de contrôler les combinaisons possibles d'hyperparamètre recherché. La partie Hyperparamétrage de la Section 4.1 discute du choix des hyperparamètres ainsi que leur conséquence sur la performance du modèle de prédiction du score de bien-être.

3.3 MODÈLE D'OPTIMISATION

Cette section présente le modèle d'optimisation développé pour l'affectation de quarts aux employés. Le modèle est détaillé en différente partie, soit les ensembles, les variables, la fonction objectif, les paramètres et les contraintes. Le modèle contient 64708 variables de décision en nombres entiers et est soumis à un total de 109295 contraintes. Ces valeurs peuvent varier dépendant du nombre d'infirmières disponible à la confection d'un horaire. La meilleure solution est atteinte en moins d'une seconde pour le scénario d'horaire de l'unité d'infirmières de ce travail, ce qui sous-entend qu'un plus grand ensemble d'infirmières en provenance d'une unité différente pourrait être utilisé sans impacter drastiquement le temps de calcul.

3.3.1 ENSEMBLE

Les tableaux 3.7 et 3.8 montrent les ensembles utilisés pour le modèle d'optimisation. Il existe plusieurs sous-ensembles d'infirmières dérivés de l'ensemble d'infirmières. Les infirmières internes I^{int} englobent les sous-ensembles d'infirmières I^{inf} et d'infirmières auxiliaires I^{aux} . Le sous-ensemble d'infirmières externes I^{ext} représentent la main d'oeuvre temporaire ne faisant pas partie de l'unité (ex. infirmière volante).

Tableau 3.7 : Ensembles des infirmières.

I	Ensemble des infirmières
I^{int}	Sous-ensemble des infirmières de l'unité
I^{ext}	Sous-ensemble des infirmières externes
I^{inf}	Sous-ensemble des infirmières
I^{aux}	Sous-ensemble des infirmières auxiliaires

Tableau 3.8 : Ensembles des quarts et des jours.

J_i	Ensemble des Jours dans l'horaire $i \in [1, 14]$
Q_j	Ensemble des quarts de travail dans une journée $j = \{0, 1, 2\}$

3.3.2 VARIABLES

Les variables sont les valeurs que le modèle d'optimisation doit déterminer. À l'exception de la variable $penalty_i$, toutes les variables sont binaires. Chaque variable constitue l'information sur l'affectation d'une infirmière à un quart, soit un oui ou un non :

$$X_{ijq} = \begin{cases} 1, & \text{si l'infirmière } i \in I \text{ est affecté au quart } q \in Q \text{ du jour } j \in J \\ 0, & \text{sinon.} \end{cases}, \quad (3.11)$$

$$TS_{ijq} = \begin{cases} 1, & \text{si l'infirmière } i \in I \text{ travaille à temps supplémentaire au quart } q \in Q \\ & \text{du jour } j \in J \\ 0, & \text{sinon.} \end{cases}, \quad (3.12)$$

$$Fds_{i[0;1]} = \begin{cases} 1, & \text{si l'infirmière } i \in I \text{ a été affecté à travailler durant la première ou} \\ & \text{la deuxième fin de semaine à l'horaire} \\ 0, & \text{sinon.} \end{cases}, \quad (3.13)$$

3.3.3 PARAMÈTRES

Les paramètres sont les suivants :

- Aff_{ijq} : Indique le quart $q \in Q$ du jour $j \in J$ où l'infirmière $i \in I$ peut être affectée à travailler.
- Aff_{ijq}^{sec} : Indique si le quart $q \in Q$ du jour $j \in J$ fait partie des quarts pouvant être affectés à l'infirmière $i \in I$, mais ne faisant pas partie du contrat de travail.
- B^{inf}, B^{aux} : Nombre minimum de quarts affectés par chaque infirmière de l'unité.
- Con_{ij} : Indique si le jour $j \in J$ fait partie d'un jour de congé demandé par l'infirmière $i \in I$.
- Exp_i : Valeur binaire désignant l'expertise ou non d'une infirmière $i \in I$.
- Max_i : Maximum de quart pouvant être attribué à une infirmière $i \in I$ selon son contrat de travail.
- QQ_{jq} : Nombre d'infirmières nécessaires durant un quart $q \in Q$ au jour $j \in J$.
- $Pref_{ijq}$: Indique si le quart $q \in Q$ du jour $j \in J$ fait partie des préférences de l'infirmière $i \in I$.
- SC_i : Valeur du score de bien-être d'absentéisme pour une infirmière $i \in I$ fournie par l'algorithme de régression.

3.3.4 FONCTION OBJECTIF

Comme mentionné précédemment, l'objectif de ce projet est de construire un modèle d'optimisation capable de générer des horaires qui minimisent le risque d'absentéisme. La fonction objectif vise à maximiser le nombre d'infirmières affectées à l'horaire :

$$\begin{aligned}
 \max \sum_{i \in I^{int}} \sum_{j \in J} \sum_{q \in Q} & ((x_{ijq} \times Pref_{ijq} \times SC_i) - (x_{ijq} \times Aff_{ijq}^{sec} \times 100) \\
 & - ts_{ijq} \times SC_i \times 400) - \sum_{i \in I^{ext}} \sum_{j \in J} \sum_{q \in Q} (x_{ijq} \times 10000) \\
 & - \sum_{i \in I} (Penalty_i \times 35).
 \end{aligned} \tag{3.14}$$

Dans cette fonction, trois méthodes permettent l'affectation de quarts de travail, soit régulier et temps supplémentaire, aux infirmières internes (I^{int}) :

1. Le seul moyen d'augmenter le résultat de la fonction objectif est par l'affectation de quarts de travail (x_{ijq}), défini comme préférence ($Pref_{ijq}$) pour une infirmière. Cette valeur est multipliée par le score de bien-être (SC_i), ce qui avantage l'affectation des quarts en préférence des infirmières avec un score plus élevé que leurs consœurs.
2. Le résultat est diminué pour chaque affectation de quart secondaire (Aff_{ijq}^{sec}). Une multiplication par 100 permet d'appliquer un poids à ce type d'affectation sur l'optimisation de la fonction. Le poids permet de garder la priorité sur l'affectation de quarts de préférence, car les valeurs de x_{ijq} et Aff_{ijq}^{sec} sont binaires.
3. Pour l'affectation de quarts à temps supplémentaires (ts_{ijq}), l'impact sur le résultat dépend du score (SC_i) de l'infirmière (I^{int}) concernée. Une multiplication par 400 assure que l'affectation de quarts de temps supplémentaire soit toujours plus pénalisée que l'affectation d'un temps quart à temps régulier (x_{ijq}).

La plus grande pénalité possible est l'affectation d'infirmières externes (I^{ext}). Dans un cas idéal, il n'y a pas de besoin en infirmières supplémentaires afin de combler les périodes ne remplissant pas le quota d'employés (contrainte de l'Éq. (3.30)). Or, l'analyse des données historiques démontre que les infirmières travaillent souvent en sous-effectif. Si le modèle d'optimisation n'arrive pas à remplir tous les quarts avec l'ensemble d'infirmières internes, celui-ci affecte des infirmières externes. Comme cette situation est à éviter, chaque affectation d'infirmière externe est multipliée par 1000.

La variable $Penalty_i$ est liée à la contrainte Éq. (3.23). Puisque cette contrainte peut être violée par le modèle, $Penalty_i$ pénalise la solution de la fonction objectif pour chaque non-respect de cette contrainte.

3.3.5 CONTRAINTES

Les contraintes de ce modèle portent sur l'affectation de quart régulier et à temps supplémentaire, le respect des quotas de travail, les contraintes d'affectation de fin de semaine et la limite de travail consécutif.

AFFECTATION DE QUART RÉGULIER

Une infirmière ne peut travailler qu'un seul quart à temps régulier par jour :

$$\sum_{q \in Q} x_{ijq} \leq 1, \quad \forall i \in I, j \in J. \quad (3.15)$$

Une infirmière peut seulement être affectée à un quart si celui-ci respecte son contrat de travail ou fait partie des quarts secondaires de celle-ci :

$$x_{ijq} \leq Aff, \quad \forall i \in I^{int}, j \in J, q \in Q. \quad (3.16)$$

Cette contrainte limite le nombre d'attributions de quart selon le contrat de travail de l'infirmière. Les congés de l'infirmière visée sont considérés dans cette contrainte. Lorsque l'on accorde un congé à une infirmière, celui-ci est soustrait du nombre de quarts à travailler selon le contrat de travail de celle-ci :

$$\sum_{j \in J} \sum_{q \in Q} x_{ijq} \leq Max_i - \sum_{j \in J} Con_{ij}, \quad \forall i \in I^{int}. \quad (3.17)$$

Il y a également une contrainte d'affectation d'un quart de travail durant un jour de congé pour une infirmière $i \in I$. La même contrainte est appliquée sur les affectations de temps supplémentaires :

$$x_{ijq} < Con_{ij}, \quad \forall i \in I^{int}, j \in J, q \in Q. \quad (3.18)$$

$$ts_{ijq} < Con_{ij}, \quad \forall i \in I^{int}, j \in J, q \in Q. \quad (3.19)$$

Afin d'assurer que le modèle affecte des quarts aux infirmières de façons équitables, une borne minimale d'affectation de quart est ajoutée. La borne est différente pour chaque catégorie d'infirmières, en raison des différences entre celles-ci. Les unités d'infirmières et d'infirmières auxiliaires ne partagent pas le même nombre d'infirmières et ont des quotas de travail différent. La borne minimale diffère donc pour ces deux raisons, mais dépend aussi du personnel infirmier disponible durant la période courante.

$$\sum_{j \in J} \sum_{q \in Q} x_{ijq} + \sum_{j \in J} con_{ij} \geq B^{inf}, \quad \forall i \in I^{inf}. \quad (3.20)$$

$$\sum_{j \in J} \sum_{q \in Q} x_{ijq} + \sum_{j \in J} con_{ij} \geq B^{aux}, \quad \forall i \in I^{aux}. \quad (3.21)$$

Cette contrainte limite le nombre d'affectations d'infirmière selon le contrat de travail. Le nombre de quarts pouvant être affectés à une infirmière est fixe selon son contrat (Max_i). La contrainte diminue ce maximum lorsque la partie entière du score d'une infirmière $\lfloor SC_i \rfloor$ est plus grande ou égale 4. Si ce n'est pas le cas, la contrainte Éq. (3.17) a priorité sur le maximum de quarts par horaire d'une infirmière. Ceci est une contrainte douce, soit une contrainte qui peut ne pas être respectée par le modèle d'optimisation. Le non-respect d'une contrainte a généralement comme conséquence la dégradation du résultat de la fonction objectif. Dans ce cas, la variable *penalty* (Éq. (3.22)) est égale au nombre d'affectations supérieures à la limite imposée par les autres variables de la contrainte :

$$Penalty_i = \text{variable en nombres entiers visant à pénaliser le modèle lorsque} \quad (3.22)$$

la contrainte 3.23 n'est pas respectée.

$$\sum_{j \in J} \sum_{q \in Q} x_{ijq} \leq Max_i + 3 - \lfloor SC_i \rfloor + penalty_i, \quad \forall i \in I^{int}. \quad (3.23)$$

Chaque affectation de quart pour une infirmière doit être distancée de deux quarts de repos ou plus avant la prochaine affectation. Ainsi, une infirmière ne peut pas faire plus d'un quart par jour (Éq. (3.15)) et bénéficie d'un temps de repos d'environ 16 heures (2 quarts) pour les quarts de soir (Éq. (3.24 et 3.25)) et de nuit (contrainte 3.26) :

$$x_{ij1} + x_{i(j+1)2} \leq 1, \quad \forall i \in I, j \in J. \quad (3.24)$$

$$x_{ij1} + x_{i(j+1)0} \leq 1, \quad \forall i \in I, j \in J. \quad (3.25)$$

$$x_{ij0} + x_{i(j+1)2} \leq 1, \quad \forall i \in I, j \in J. \quad (3.26)$$

Cette contrainte limite l'affectation d'infirmière par rapport à leur expertise. Pour chacun des quarts de travail, au moins une des infirmières affectées doit avoir de l'expertise, qui est un paramètre décrivant si une infirmière a l'expérience nécessaire pour faire un quart seul ou accompagné d'infirmières ayant moins d'expérience :

$$\sum_{i \in I} ((x_{ijq} + ts_{ijq}) \times Exp_i) \geq 1, \quad \forall j \in J, q \in Q. \quad (3.27)$$

AFFECTATION DE TEMPS SUPPLÉMENTAIRE

Cette contrainte évite l'affectation d'un temps supplémentaire et un quart régulier durant la même période pour une infirmière. Une infirmière ne peut pas effectuer un quart régulier et à temps supplémentaire au même moment :

$$x_{ijq} + ts_{ijq} \leq 1, \quad \forall i \in I, j \in J, q \in Q. \quad (3.28)$$

Cette contrainte permet de limiter le nombre maximum de temps supplémentaires pouvant être attribués à une infirmière selon son score de bien-être :

$$\sum_{j \in J} \sum_{q \in Q} ts_{ijq} \leq \left\lfloor \frac{8}{SC_i + 1} \right\rfloor, \quad \forall i \in I^{int}. \quad (3.29)$$

Le maximum de temps supplémentaire est déterminé à partir du score de bien-être. Une infirmière avec un score plus élevé obtient un nombre maximum de temps supplémentaires attribuables plus petits. Le Tableau 3.9 détaille le nombre maximum selon le score :

Tableau 3.9 : Maximum de temps supplémentaire attribuable selon l’Éq. (3.29).

score	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5
TS max	8	5	4	3	2	2	2	1	1	1	1	1	1	1	1	0

QUOTAS DE TRAVAIL

Le modèle doit respecter les quotas de travail de l’unité. Les quotas de travail sont disponibles au Tableau 3.3. Ceux-ci diffèrent selon le moment de la journée et la catégorie auquel est reliée chaque infirmière (infirmière ou auxiliaire) :

$$\sum_{i \in I} (x_{ijq} + ts_{ijq}) = QQ_{jq}, \quad \forall j \in J, q \in Q. \quad (3.30)$$

FIN DE SEMAINE

Les contraintes qui suivent sont liées aux restrictions de travail durant les fins de semaine provenant de la convention collective des infirmières de la Fédération Interprofessionnelle de la santé du Québec (FIQ). Une infirmière ne doit travailler qu’une seule fin de semaine maximum sur une période de deux semaines.

Les contraintes des Éq. (3.31 et 3.32) s’assurent qu’une infirmière travaille durant une même fin de semaine, étiqueté comme étant les jours 0, 6, 7 et 13. L’Éq. (3.33) limite une infirmière à travailler une seule fin de semaine.

$$\sum_{q \in Q} x_{6jq} + x_{7jq} = 2_{i0}, \quad \forall i \in I. \quad (3.31)$$

$$\sum_{q \in Q} x_{0jq} + x_{13jq} = 2_{i1}, \quad \forall i \in I. \quad (3.32)$$

$$\sum_{q \in Q} x_{0jq} + x_{6jq} + x_{7jq} + x_{13jq} \leq 2, \quad \forall i \in I. \quad (3.33)$$

En plus de travailler une seule fin de semaine sur deux, une infirmière doit travailler durant la même période pendant les deux jours de travail attribués (ex. une infirmière qui travaille samedi de jour doit aussi travailler dimanche de jour).

$$(x_{6,0,i} \times x_{7,0,i}) + (x_{6,1,i} \times x_{7,1,i}) + (x_{6,2,i} \times x_{7,2,i}) \leq 2 \quad \forall i \in I. \quad (3.34)$$

TRAVAIL CONSÉCUTIF

Afin de limiter la surcharge de travail des infirmières, celle-ci ne devrait pas travailler plus de cinq quarts consécutifs, à temps régulier ou supplémentaire, sur une période de cinq jours :

$$\begin{aligned} & \sum_{q \in Q} x_{ijq} + x_{i(j+1)q} + x_{i(j+2)q} + x_{i(j+3)q} + x_{i(j+4)q} + x_{i(j+5)q} \\ & + ts_{ijq} + ts_{i(j+1)q} + ts_{i(j+2)q} + ts_{i(j+3)q} + ts_{i(j+4)q} + ts_{i(j+5)q} \leq 5 \\ & \forall i \in I, j \in J. \end{aligned} \quad (3.35)$$

CHAPITRE IV

RÉSULTATS

Ce chapitre aborde les résultats obtenus à la suite des tests effectués sur les modèles développés dans ce mémoire. Dans ce projet, une méthode est développée afin de prédire le score de bien-être à l'aide de l'apprentissage machine. De plus, un modèle d'optimisation linéaire en nombres entiers a été développé afin de générer des horaires, selon le score de bien-être prédit. Les résultats obtenus sont analysés dans ce chapitre par rapport au modèle analysé. Ce chapitre se divise en trois principales sections, soit la performance du modèle de prédiction, l'analyse des résultats de la prédiction du score des infirmières et l'analyse des horaires construits à l'aide des résultats du modèle d'optimisation.

4.1 PERFORMANCE DU MODÈLE DE PRÉDICTION

Cette section présente les résultats relatifs à la performance du modèle d'apprentissage machine de la prédiction du score de bien-être. Plusieurs manipulations sont faites afin d'améliorer la précision de la prédiction du modèle. Ces tests impliquent la modification de l'ensemble de données afin de réduire la marge d'erreur du modèle et le paramétrage de l'algorithme de RFR.

En apprentissage machine, plusieurs méthodes de calculs existent afin de juger la performance d'un modèle, tel qu'une matrice de confusion, le *F-score* ou encore l'erreur quadratique moyenne. Le choix de la méthode dépend entre autres de la variable cible que l'on cherche à prédire. Dans le cadre de ce projet, la prédiction est faite à partir d'une régression. Il est donc possible de calculer la performance du modèle en comparant les résultats numériques prédits avec les valeurs réelles. Ainsi, deux métriques de performance sont régulièrement

utilisées pour déterminer la performance du modèle, soit le Mean Square Error (MSE) et Mean Absolute Error (MAE). Le MSE représente la moyenne du carré de la marge d'erreur observée pour toutes les instances prédictes par le modèle. Elle est calculée en faisant la différence entre chaque valeur calculée et prédictée. Chaque différence est mise au carré afin d'éviter les valeurs négatives. Un MSE se rapprochant de 0 signifie un excellent modèle ayant prédicté à la quasi-perfection les valeurs calculées. Autrement dit, plus le nombre de prédictions ayant une grande marge d'erreur est important, plus la valeur du MSE est élevée. Similairement, le MAE représente la différence absolue entre la valeur calculée et la valeur prédictée. Le MSE et le MAE sont des scores de performance similaires, car ceux-ci représentent un score de prédictions entre les valeurs prédictes et réelles. Dans le cadre de ce projet, le MSE (Éq. (4.1)) et le MAE (Éq. (4.2)) sont calculés à l'aide des équations :

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2, \quad (4.1)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \quad (4.2)$$

Où N représente le nombre d'instances ayant été testées, y_i représente la valeur réelle et \hat{y}_i la valeur prédictée. Comme le MAE est la différence absolue entre le réel et la prédition, le score obtenu est un indicateur des performances générales du modèle. En élevant la différence au carré, le MSE permet d'obtenir un score de performance influencé grandement par les valeurs aberrantes obtenues.

Les sous-sections suivantes montrent l'évolution de la performance du modèle par rapport aux différentes manipulations effectuées sur ce dernier. Plusieurs modèles sont générés afin de s'assurer de la constance des résultats obtenus.

BIAIS DES INFIRMIÈRES EN ASSURANCE SALAIRE

Comme discuté dans la Sous-Section 3.2.4, les instances comportant des infirmières en arrêt de travail doivent être retirées de l'ensemble de données. Ces instances ne sont pas pertinentes à la prédiction d'un score de bien-être et risquent de biaiser l'entraînement du modèle. Un arrêt de travail est synonyme d'une instance comportant une variable ASSURANCE égale à 10. Néanmoins, cette variable ne doit pas être retirée en entier de l'ensemble de données. En effet, celle-ci est un indicatif de l'arrêt de travail d'une infirmière. Après différents tests de performance, la première instance comportant une valeur de cinq ou plus pour la variable ASSURANCE est conservée, tandis que les instances suivantes (où l'infirmière est en arrêt de travail) sont retirées jusqu'à un retour au travail, pourvu que celle-ci retourne au travail. Ainsi, les périodes d'arrêt de travail des infirmières ne font pas partie du modèle d'entraînement. Le Tableau 4.1 montre les résultats de la performance de quatre différents modèles de prédiction du score de bien-être généré. Quatre modèles ont été générés afin d'analyser les variations de performance entre chaque modèle.

Tableau 4.1 : Résultat de la performance du modèle de prédiction du bien-être lorsque du retrait des instances de l'ensemble de données comportant un valeur de ASSURANCE à 5 et plus.

Modèle	MSE	MAE
1	2,792	1,250
2	2,795	1,250
3	2,796	1,248
4	2,816	1,253
Moyenne	2,800	1,250

Les quatre résultats obtenus sont tous très similaires entre chaque test, ce qui montre que l'algorithme reste constant entre chaque itération de modèle. Le MSE et MAE sont assez élevés dans ce premier test de performance. Cela montre qu'il y a une grande variance entre les valeurs de la variable cible SCORE et les valeurs prédictives par le modèle. Le MAE indique un manque de précision différent d'environ 1.25 de la valeur calculée. Un MSE élevé montre qu'il y a une grande disparité entre les valeurs calculées et prédictives. Ces deux valeurs doivent être réduites afin de démontrer de meilleurs résultats provenant du modèle.

TAILLE DE L'ENSEMBLE D'ENTRAÎNEMENT ET DE L'ENSEMBLE TEST

La performance d'un modèle d'apprentissage machine est directement reliée à la qualité des données fournies. C'est pourquoi il est très important de séparer l'ensemble de données en sous-ensembles d'entraînements et de tests afin d'optimiser le gain d'information pour l'apprentissage du modèle. De plus, la méthode d'ensembles génère plusieurs forêts d'arbres régressifs, ce qui demande la création de plusieurs sous-ensembles d'entraînements et de tests. Pour que chaque forêt générée soit différente des autres, les sous-ensembles d'entraînements et de tests doivent différer d'un modèle à l'autre. Après plusieurs tests, la stratégie *leave-one-subject-out* a été appliquée à la création des sous-ensembles d'entraînements et de tests. Comme le nom l'indique, toutes les infirmières, à l'exception d'une, font partie de l'ensemble d'entraînement, où les données de l'infirmière retirée sont utilisées comme test de performance. Cette méthode fait référence à la Figure 3.6, soit l'étape *Choix des infirmières pour la prédiction*. Afin d'obtenir des résultats de performance plus constants, l'infirmière 11 a aussi été retirée de l'ensemble d'entraînement, faisant partie de l'ensemble test pour chaque itération de forêt générée. L'infirmière 11 est une infirmière ayant fait un arrêt de travail à la suite d'une augmentation graduelle de son score de bien-être. Donc, le choix de cette infirmière pour tester la performance du modèle est très pertinent. Une analyse plus

détaillée de cette infirmière se trouve au début de la Sous-Section 4.2.1. Ainsi, les données provenant de deux infirmières constituent l'ensemble test pour tous les modèles de RFR. Le Tableau 4.2 montre les résultats de performance du modèle lorsque le choix de la taille de l'ensemble test est influencé par la méthode *leave-one-subject-out* et l'infirmière 11.

Tableau 4.2 : Résultat de la performance du modèle de prédiction du bien-être utilisant la stratégie *leave-one-subject-out*.

Modèle	MSE	MAE
1	2,697	1,196
2	2,729	1,205
3	2,714	1,198
4	2,702	1,198
Moyenne	2,711	1,199

Les résultats entre chaque modèle restent constants, le MSE et le MAE moyens ont été réduits de .89 et 0.06 respectivement. L'amélioration du MSE indique une légère diminution de la variance pour les valeurs aberrantes. La configuration précédente du tableau 4.1 retire deux infirmières de leur ensemble d'entraînement. Plus le nombre d'infirmières retirées est élevé, plus les résultats de performance se détériorent. Ainsi, la méthode *leave-one-subject-out* est la meilleure pour construire l'ensemble d'entraînement et de test.

CHOIX DES VARIABLES SIGNIFICATIVES DU MODÈLE

Le modèle de prédiction est construit à l'aide d'un ensemble de données de 533 instances de données, dont chacune a 18 variables. Toutes ces variables sont nécessaires à la création de la variable SCORE, l'étiquette représentant le score de bien-être d'une infirmière. Or, certaines variables sont nuisibles à la création de modèles pour la régression. Celles-ci risquent d'agir

comme bruit, mitigeant l'impact des variables plus importantes à la prédiction sur la qualité des résultats obtenus. Une étude des variables est faite afin de retirer les moins importantes de l'ensemble de données. La Figure 4.1 montre la matrice de corrélation entre toutes les variables de l'ensemble de données. Une matrice de corrélation permet d'analyser si certaines variables ont une corrélation entre elles, ou si celles-ci sont complètement indépendantes. Une valeur absolue de 1 représente une corrélation parfaite entre deux variables, tandis que 0 signifie qu'aucun lien n'existe entre celles-ci.

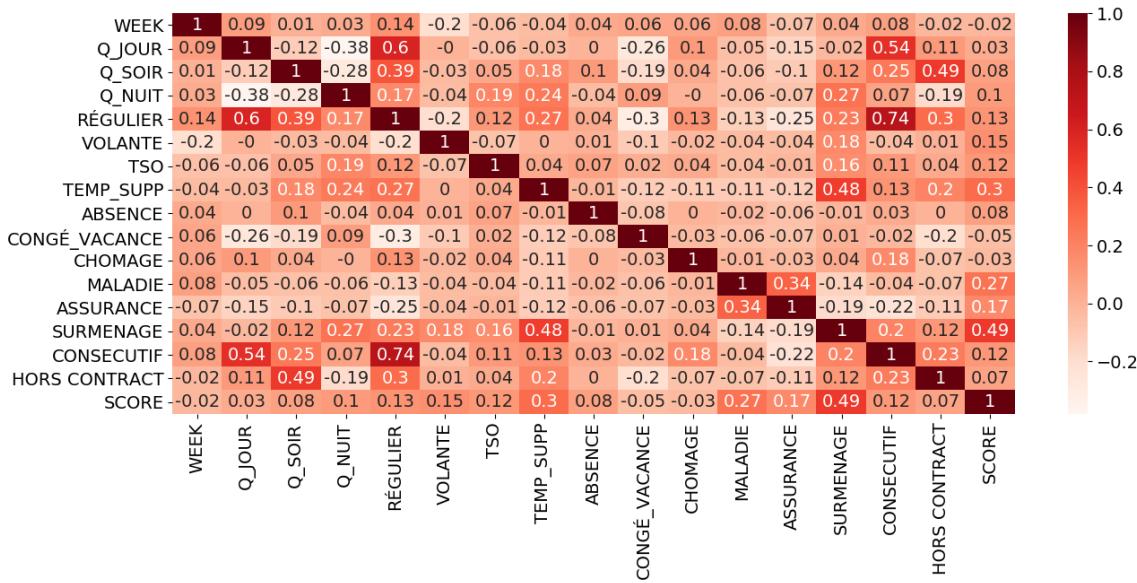


Figure 4.1 : Matrice des coefficients de corrélation entre les différentes variables du modèle de prédiction de l'absentéisme.

© 2021 Yoan Villeneuve.

Or, pour savoir si une corrélation est réellement significative avec les données de la matrice de corrélation, il est nécessaire de connaître les *p-values* (valeur p) reliées à la variable cible. La valeur de *p-value* permet d'affirmer ou infirmer l'*hypothèse nulle*, soit de savoir si les données d'un ensemble sont équivalentes entre elles. Si l'*hypothèse nulle* est fausse, il y a donc une inégalité entre les données de l'ensemble. Si tel est le cas, cela implique donc

l'*hypothèse alternative*. Nous cherchons ainsi à démontrer que les coefficients de corrélation de la Figure 4.1 sont significatifs. La Figure 4.2 montre les *p-values* calculées entre chaque variable de l'ensemble de données de prédiction.

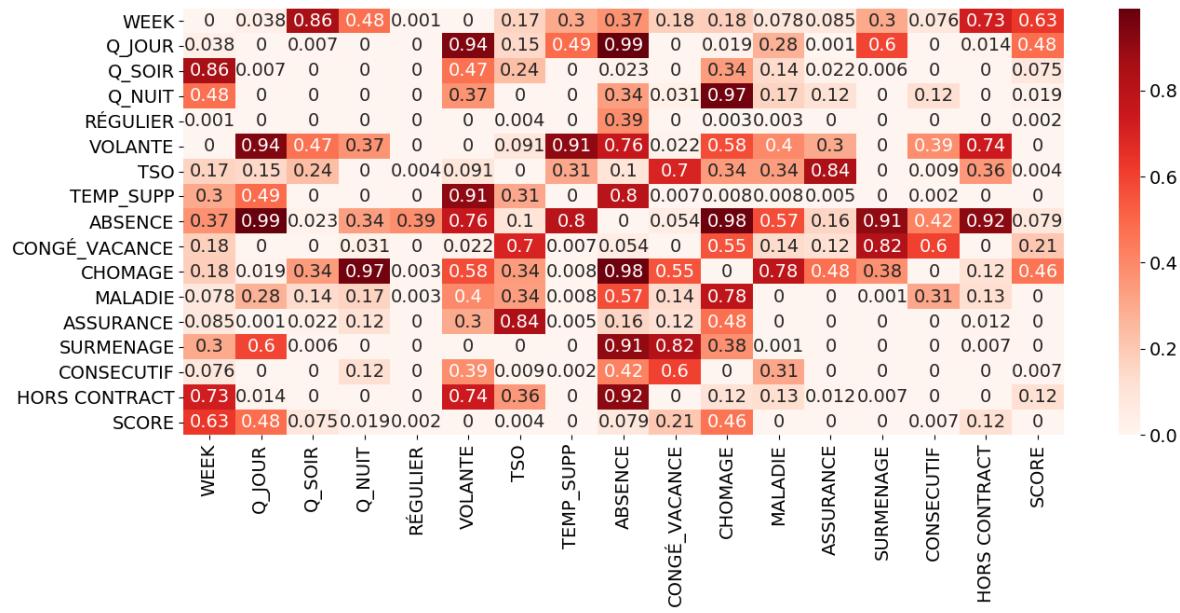


Figure 4.2 : Matrice des valeurs de *p-value* entre chaque variable de l'ensemble de données de prédiction.

© 2021 Yoan Villeneuve.

Les valeurs de *p-value* obtenues dans cette matrice ont été arrondies à trois décimales. Les valeurs possibles se trouvent entre 0 et 1. Mis à part le calcul de la *p-value* sur une même variable (ex. une comparaison entre la variable WEEK avec WEEK), les zéros obtenus sont en fait de très petites valeurs. Il est maintenant possible d'analyser la corrélation des variables à l'aide de la signification statistique obtenue par l'analyse de leur *p-value*. Une valeur de *p-value* plus petite que 0.05 indique une corrélation significative entre les données, confirmant l'*hypothèse alternative*. Une valeur de *p-value* de plus de 0.05 confirme plutôt l'*hypothèse nulle*, signifiant que la relation entre les deux variables n'est probablement pas statistiquement

significative. Les variables WEEK, VOLANTE, TSO, ABSENCE, CONGÉ_VACANCE et CHOMAGE semblent avoir le moins de corrélation avec le reste des variables du problème, appuyé par une *p-value* assez élevée pour la plupart des variables du problème. Ces variables ont ainsi de bonnes chances de biaiser l'apprentissage du modèle de prédiction. D'un autre côté, les variables SURMENAGE et CONSECUTIF ont une haute corrélation avec le reste des variables et une *p-value* généralement basse. Cela est normal, car ceux-ci sont construits à partir de beaucoup d'information provenant des autres variables de l'ensemble de données. Il est intéressant de voir la variable HORS CONTRAT légèrement corrélée avec la variable Q_SOIR. Concernant la variable SCORE, celle-ci semble être majoritairement corrélée avec la variable SURMENAGE, une variable construite, entre autres, à l'aide de la variable TEMP_-SUPP. Les variables ASSURANCE et MALADIE sont les quarts d'absences non planifiées qui sont corrélées avec le plus avec le score de bien-être. Enfin, les variables WEEK, Q_JOUR et CHOMAGE ont les plus petites valeurs de corrélation et la plus grande *p-value*.

Soutenus par la matrice des coefficients de corrélation, visible à la Figure 4.1 et des *p-values* à la Figure 4.2, plusieurs tests ont été faits afin de trouver la meilleure combinaison de variables. Il a été supposé que la variable ASSURANCE était défavorable pour le modèle en relation aux faibles coefficients de corrélation dans la Figure 4.1. Après avoir effectué plusieurs tests, la présence de la variable ASSURANCE s'avère avoir un impact bénéfique sur les performances du modèle. Les variables Q NUIT, VOLANTE, TSO et CHOMAGE ont été retirés de l'ensemble en raison d'une augmentation de la précision causée par leur absence. Le Tableau 4.3 montre les résultats de la performance du modèle.

Tableau 4.3 : Résultat de la performance du modèle de prédiction du score de bien-être lorsque les variables les moins significatives sont retiré de l'ensemble de données.

Modèle	MSE	MAE
1	2,654	1,211
2	2,693	1,218
3	2,658	1,215
4	2,717	1,225
Moyenne	2,680	1,217

Après le retrait des variables non significatives à la prédiction du score de bien-être, on observe une légère amélioration du MSE et du MAE, précédemment à 1.199. Cela signifie une diminution de la différence entre la valeur prédite et calculée pour les valeurs aberrantes, mais pas nécessairement pour le reste des prédictions.

HYPERPARAMÈTRAGE

Les hyperparamètres sont des paramètres qui influencent l'entraînement d'un modèle d'apprentissage machine, Claesen & Moor (2015). La librairie *scikit-learn* (Buitinck *et al.*, 2013) permet une certaine optimisation des hyperparamètres, testant plusieurs combinaisons de paramètres. L'utilisateur doit fournir l'intervalle des valeurs pour chaque hyperparamètre à tester. Ensuite, une recherche aléatoire de la valeur des paramètres est faite afin de trouver la meilleure combinaison possible à l'aide de la fonction *RandomizedSearchCV*. Le choix des intervalles de valeurs d'hyperparamètre a été fait en raison du choix de la fonction de recherche d'hyperparamètre. Cette fonction permet une recherche non exhaustive des hyperparamètres, améliorant considérablement le temps de calcul. Au total, 100 itérations ont été lancées pour chaque modèle. Le Tableau 4.4 montre le choix de l'intervalle des paramètres à explorer.

Tableau 4.4 : Intervalle des paramètres possibles pour la recherche d'hyperparamètre dans le modèle de prédiction du bien-être

Paramètre	Intervalle
n_estimators	[10, 20, 30, ..., 200]
max_features	[auto, sqrt]
max_depth	[None, 10, 20, ..., 110]
min_samples_split	[2, 5, 7, 10]
min_samples_leaf	[1, 2, 4]
bootstrap	[True, False]

Voici la description des hyperparamètres sélectionnés pour le RFR :

- n_estimators : Nombre d'arbres constituant la forêt du modèle.
- max_features : Méthode pour déterminer le nombre de variables (*features*) considérées à chaque embranchement de noeud pour un arbre. Considérant n variables, *auto* signifie que n variables sont considérées et *sqrt* signifie que seulement la racine du nombre de variables (\sqrt{n}) est considérée.
- max_depth : Profondeur maximale des arbres.
- min_samples_split : Spécifie le nombre minimum d'instances requises pour diviser un noeud.
- min_samples_leaf : Spécifie le nombre minimum d'instances requises pour qu'un noeud soit une feuille, soit un noeud terminal.
- bootstrap : Si vrai (*True*), alors le choix des données pour la création d'un arbre est aléatoire, ce qui crée des variations entre chaque arbre. Si fausses (*False*), alors toutes les données sont utilisées.

La recherche d'hyperparamètre pour tous les modèles a duré 26.39 minutes. Le choix des hyperparamètres a été fait en fonction des meilleures performances obtenues et de la récurrence des paramètres obtenus pour chaque modèle. Le programme a été exécuté sur

un ordinateur portable avec l'édition Windows 8.1 pro 64-bits, 8 Gigabits de mémoire vive (RAM) et un processeur Intel(R) Core(TM) i5-4300U @ 2.50 GHz. Le Tableau 4.5 montre le résultat final des hyperparamètres retenus.

Tableau 4.5 : Paramètres pour le modèle de prédiction du score de bien-être

Paramètre	Valeur
n_estimators	160
max_features	auto
max_depth	80
min_samples_split	10
min_samples_leaf	1
bootstrap	True

Ainsi, les paramètres obtenus ont été appliqués au modèle de prédiction du score de bien-être des infirmières. Le Tableau 4.6 montre les résultats de performance du modèle après la modification des hyperparamètres de l'algorithme RFR.

Tableau 4.6 : Résultat de la performance du modèle de prédiction du bien-être à la suite de la modification des hyperparamètres du RFR.

Exécution	MSE	MAE
1	2,519	1,184
2	2,499	1,183
3	2,504	1,178
4	2,514	1,186
5	2,504	1,185
Moyenne	2,508	1,183

L'influence du paramétrage des hyperparamètres du modèle de prédiction est très positive pour le modèle. Une forte amélioration est observable pour le MSE et ainsi qu'une légère

amélioration pour le MAE. Ceci implique que l'hyperparamétrage du modèle a permis l'amélioration de l'étape d'apprentissage du modèle sans avoir eu le besoin de modifier l'ensemble de données.

PERFORMANCE FINALE DU MODÈLE

Le résultat final de la performance du modèle se trouve dans le Tableau 4.7.

Tableau 4.7 : Résultat final de la performance du modèle de prédiction du score de bien-être pour les infirmières.

Modèle	MSE	MAE
1	2,508	1,175
2	2,496	1,170
3	2,514	1,172
4	2,507	1,168
5	2,475	1,163
Moyenne	2,500	1,169

Le MAE égal à 1.169 signifie que le modèle n'est généralement pas si loin de la valeur réelle du score. Considérant que le modèle effectue une prédiction du score en se basant sur l'information de l'horaire (l'instance) précédent pour la semaine suivante, cette valeur est très satisfaisante.

Finalement, la performance du modèle a été améliorée, mais un modèle global d'apprentissage machine ne peut pas être jugé en se basant seulement sur la performance de prédiction. La Section 4.2 fait l'observation des résultats de la prédiction sur l'évolution du score des infirmières en comparant le score réel calculé par rapport aux prédictions faites par le modèle.

4.2 ANALYSE DE LA QUALITÉ DU SCORE DE BIEN-ÊTRE PRÉDIT

Dans cette section, l'évolution du score de bien-être est analysée individuellement pour chaque infirmière. Chaque infirmière est retirée de l'ensemble d'entraînement pour son analyse, ce qui permet d'éviter que les données d'horaire de celle-ci entrent en conflit avec la prédiction. Pour les analyses de cette sous-section et la Sous-Section 4.2.1, les infirmières analysées sont placées dans l'ensemble test du modèle de prédiction. L'analyse d'une infirmière débute par la justification du choix de celle-ci pour être analysée. Ensuite, un modèle est construit et les résultats sont représentés dans un graphe. Les horaires des infirmières sont construits sur une durée de deux semaines. Ainsi, les graphes analysent le score de chaque infirmière aux deux semaines, soit la durée d'une instance dans l'ensemble de données. Le graphe fait la distinction entre les données réelles et prédites. Il s'en suit une analyse des résultats obtenus.

ANALYSE DE L'INFIRMIÈRE 43

Cette infirmière a travaillé avec constance et sans absences fréquentes, travaillant de huit à neuf quarts de travail régulier par semaine. Celle-ci travaille parfois un ou deux quarts à temps supplémentaires et a cumulé un total de cinq absences sur les 45 semaines observées. À deux occasions, elle a effectué quatre et cinq quarts à temps supplémentaires dans ses horaires de travail. La Figure 4.3 montre l'évolution du score de bien-être de l'infirmière 43 par rapport aux valeurs prédites (Prédit) et calculées (Réel). Le modèle a pris un total de 13.58 secondes pour être généré, avec un MSE de 2.2632 et un MAE de 1.1891.

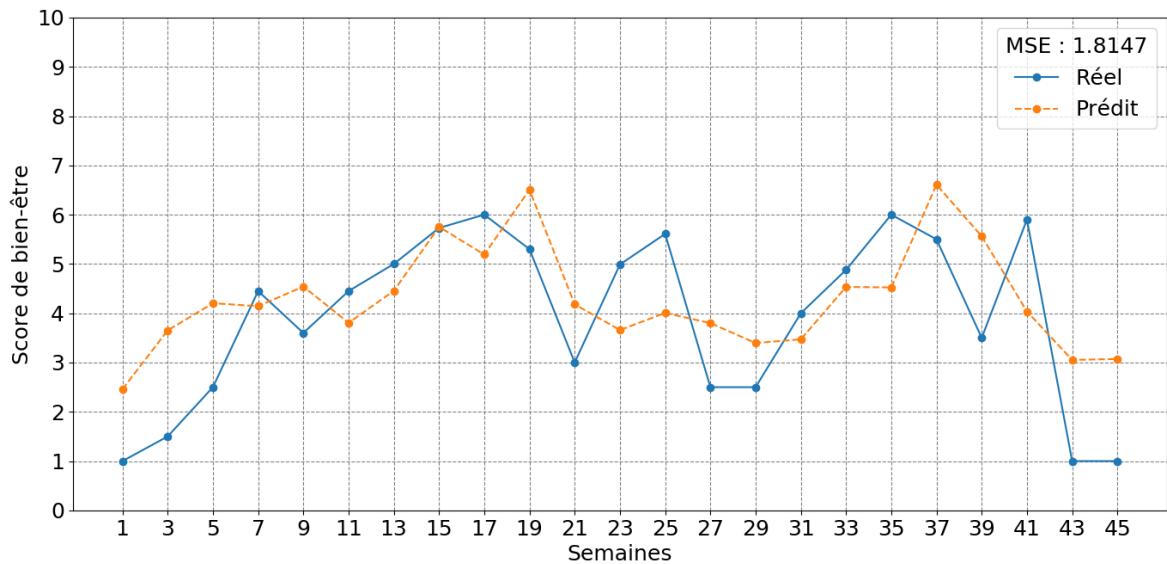


Figure 4.3 : Analyse de l'évolution du score de bien-être prédit et réel de l'infirmière 43.
 © 2021 Yoan Villeneuve.

La ligne pleine représente l'étiquette, soit la valeur réelle. L'étiquette ne semble pas suivre de tendance particulière, mais les valeurs prédites (la ligne en pointillée) suivent de façons relativement exactes les valeurs réelles. En étudiant la semaine 1 à 17, le modèle réussit à suivre l'évolution à la hausse du score de l'infirmière. De la semaine 19 à 29, le modèle reconnaît une baisse du niveau de bien-être malgré les semaines 23 et 25, où le modèle prédit une augmentation du score. Dans l'intervalle des semaines 31 à 45, le score augmente, suivi d'une baisse à la semaine 37, toutes deux reconnues par le modèle. Le modèle donne une prédiction erronée à la semaine 41, mais continue à prédire la baisse du score. Il est pertinent de noter que l'infirmière 43, à la semaine 45, a travaillé six quarts réguliers, pris un congé, effectué deux jours de chômage et un congé maladie. Comme cette période de l'année (mi-décembre) se rapproche de la fête de Noël, une corrélation entre les deux est probable.

ANALYSE DE L'INFIRMIÈRE 12

L'infirmière auxiliaire 12 effectue en général de 8 à 11 quarts réguliers aux deux semaines (par horaire) avec un à cinq quarts de temps supplémentaire. Elle a aussi eu quatre absences au total dans les 45 semaines disponibles à l'analyse. Cette infirmière effectue beaucoup de temps supplémentaire, ce qui lui vaut en général un haut score de bien-être. Or, celle-ci semble choisir certaines périodes où elle effectue plusieurs temps supplémentaires par horaire, puis diminue le nombre de temps supplémentaires sur une autre période. La Figure 4.4 montre l'évolution du score de bien-être de l'infirmière 12 par rapport aux valeurs prédites (Prédit) et calculées (Réel). Le modèle a pris un total de 11.62 secondes pour être généré avec un MSE de 2.5172 et un MAE de 1.2198.

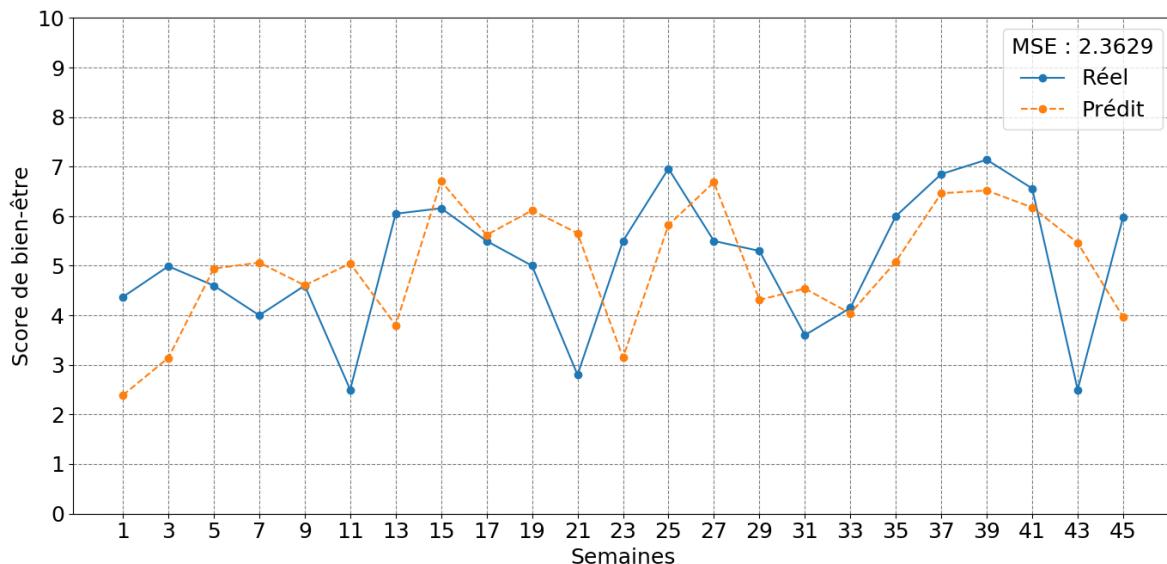


Figure 4.4 : Analyse de l'évolution du score de bien-être prédit et réel de l'infirmière 12.

© 2021 Yoan Villeneuve.

Cette infirmière à un score fréquemment au-dessus de 4, mais celui-ci semble descendre à 3 pendant les semaines 11, 21 et 43. Ces variations semblent soudaines et le score retourne toujours à 4 et plus pour la semaine suivante. Cette infirmière travaille souvent plus de 10 quarts par deux semaines (par horaire) et il arrive souvent qu'elle effectue trois quarts et plus de travail à temps supplémentaires, ce qui explique ce haut score. En général, le modèle fait un très bon travail pour prédire l'état de celle-ci, mais a beaucoup de difficultés à prédire les soudaines baisses de score observées à la semaine 11, 21 et 43. Cette difficulté de prédiction est normale, puisque l'infirmière effectue des changements soudains difficilement reconnaissables à la prédiction. De la semaine 29 à 32, celle-ci part en vacances, ce qui explique le suivi de la tendance de l'étiquette calculée par les prédictions du modèle. Ainsi le modèle semble assez performant pour prédire le bien-être des infirmières à partir des informations de la semaine précédente, mais la Figure 4.4 montre que les changements soudains du score sont difficiles à prédire correctement. Un plus grand ensemble de données contenant des infirmières similaires peut avoir un impact positif sur la prédiction de ces changements brusques.

ANALYSE DE L'INFIRMIÈRE 90

L'infirmière 90 de l'unité est un cas d'intérêt, car celle-ci travaille généralement de 9 à 11 quarts réguliers par deux semaines (par horaire), avec occasionnellement un à trois temps supplémentaires travaillés sur un horaire. De plus, celle-ci n'a eu aucune absence non planifiée. La Figure 4.5 montre l'évolution du bien-être de l'infirmière 90 par rapport aux valeurs prédites (Prédit) et calculées (Réel). Le modèle a pris un total de 11.4 secondes pour être généré, avec un MSE de 2.3378 et un MAE de 1.1375.

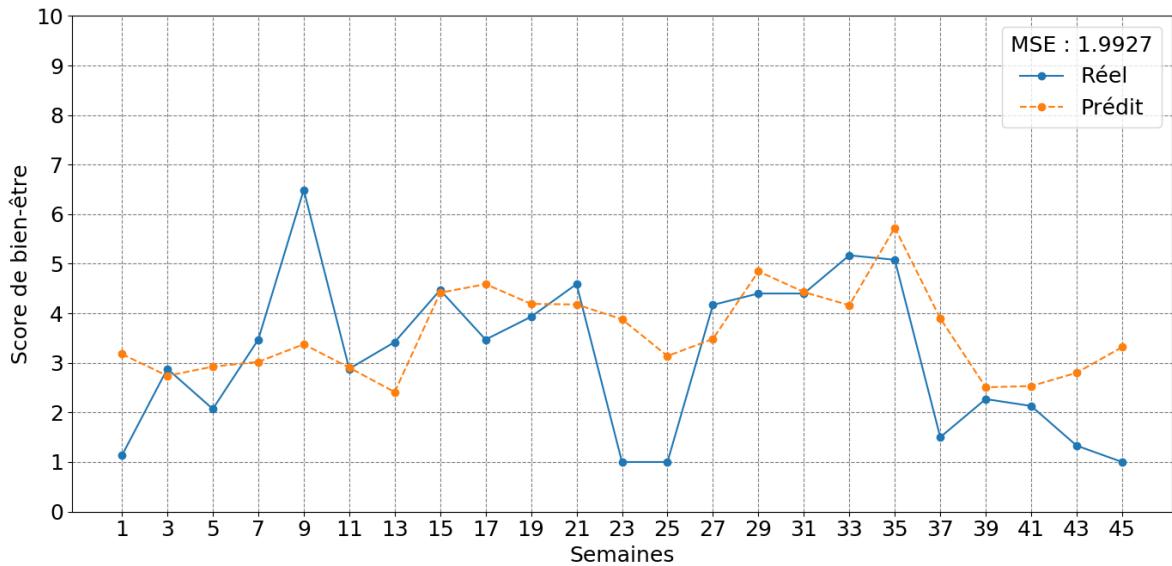


Figure 4.5 : Analyse de l'évolution du score de bien-être prédit et réel de l'infirmière 90.

© 2021 Yoan Villeneuve.

Encore une fois, les prédictions semblent suivre assez précisément l'étiquette calculée. De la semaine 1 à 21, seule la semaine 9 a une variance élevée, mais le modèle prédit tout de même une hausse du score pour cette instance. Durant l'horaire de la semaine 11, l'infirmière a pris quatre jours de chômage, ce qui explique cette hausse soudaine du score. Les semaines 23 et 25 sont des vacances, où le modèle ne semble pas avoir réussi à prédire adéquatement le score du bien-être, mais celui-ci se rattrape en suivant la tendance de l'étiquette de la semaine 27 à 41. Les prédictions du modèle réussissent à suivre le score en baisse de l'infirmière de la semaine 35 à 41. Durant les dernières semaines, celle-ci a travaillé huit à neuf quarts réguliers avec peu de temps supplémentaires et part en vacances pour la semaine 47. Le modèle de prédiction ne réussit pas à prédire les deux derniers scores, probablement en raison du temps des fêtes durant la semaine 47.

ANALYSE DE L'INFIRMIÈRE 96

L'employée 96 fait partie de l'unité d'infirmières. Celle-ci est une infirmière qui travaille peu de quarts et qui effectue beaucoup d'absences planifiées. Celle-ci travaille entre cinq et huit quarts réguliers par deux semaines (par horaire) et effectue entre un et trois quarts de temps supplémentaires. Cette infirmière a toujours environ deux congés sans solde autorisés par deux semaines. L'analyse de cette infirmière permet de voir les performances du modèle chez une infirmière moins active. La Figure 4.6 montre l'évolution du score de bien-être de l'infirmière 96 par rapport aux valeurs prédites (Prédit) et calculées (Réel). Le modèle a pris un total de 12.3 secondes pour être généré, avec un MSE de 2.0486 et un MAE de 1.073.

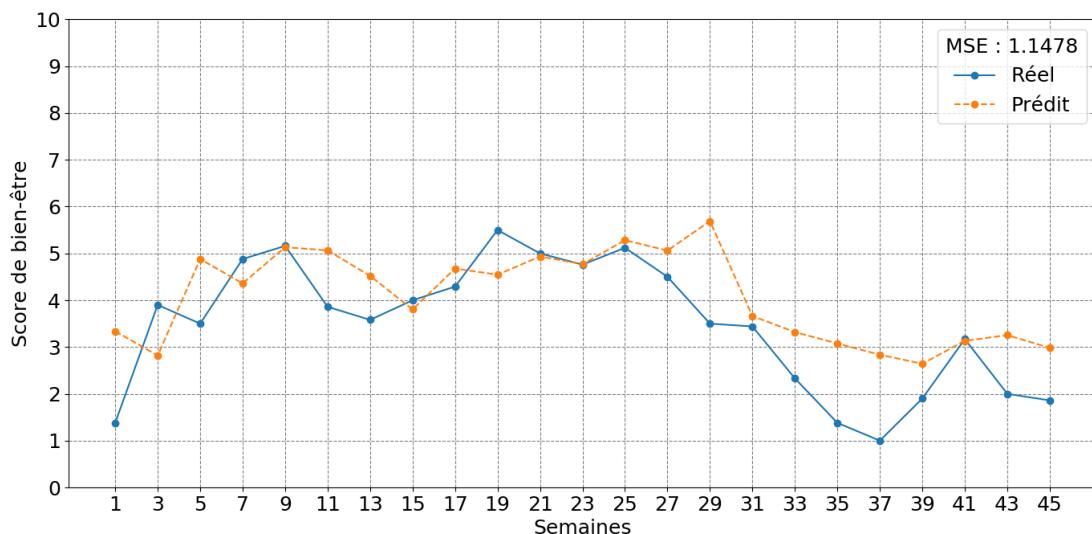


Figure 4.6 : Analyse de l'évolution du score de bien-être prédit et réel de l'infirmière 96.

© 2021 Yoan Villeneuve.

Cette infirmière a un score réel très constant dans l'évolution de son bien-être. La différence des résultats entre le score réel et prédit ne contient aucune donnée aberrante, mais les prédictions du modèle sont généralement un peu plus élevées que le score réel. Ainsi, la prédition du score de bien-être de cette infirmière est concluante.

4.2.1 INFIRMIÈRE EN ARRÊT DE TRAVAIL

Dans cette section, l'évolution du score de bien-être des infirmières ayant fait un arrêt de travail sera analysée. Comme dans les analyses précédentes, l'infirmière est introduite, les résultats du modèle généré pour l'infirmière sont présentés et une analyse des résultats est faite. Par souci de confidentialité pour le personnel soignant, cette section se concentre seulement sur l'évolution du bien-être des infirmières. Une étude plus approfondie des horaires menant à l'arrêt de travail de celles-ci est effectuée à la Sous-Section 4.3.3.

ANALYSE DE L'INFIRMIÈRE 11

Cette infirmière auxiliaire était en arrêt de travail à la semaine 19, allant jusqu'à la semaine 33. Considérant que les raisons de cet arrêt de travail ne sont pas connues, ce cas est celui ayant le plus de chance d'avoir été causé par une surcharge. La Figure 4.6 montre l'évolution du bien-être de l'infirmière 11 par rapport aux valeurs prédictes (Prédit) et calculées (Réel). Le modèle a pris un total de 11.31 secondes pour être généré, avec un MSE de 2.5184 et un MAE de 1.1776.

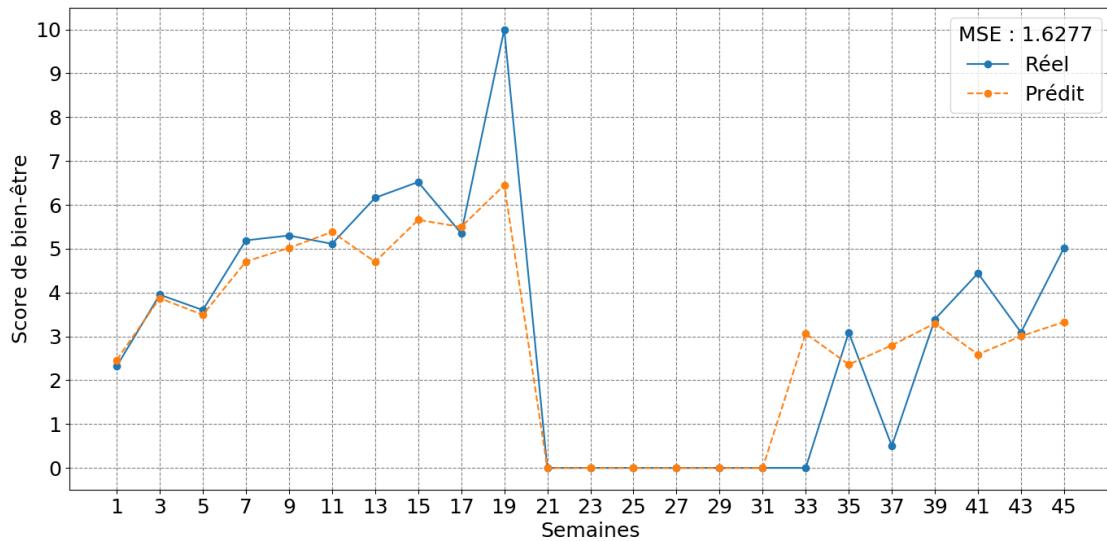


Figure 4.7 : Analyse de l'évolution du score de bien-être prédit et réel de l'infirmière 11.

© 2021 Yoan Villeneuve.

En analysant les semaines de 1 à 19, il est visible que le score de bien-être de l'infirmière se détériore à un rythme constant. À partir de la semaine 7, le score augmente au-dessus de 5 et reste au-dessus de cette valeur jusqu'à son arrêt de travail. Si l'on compare les prédictions du modèle avec le score calculé, le modèle prédit très bien l'augmentation du score, en plus de prédire avec une précision relativement satisfaisante le score réel. La dernière prédiction, soit le score de la semaine 19, n'est pas nécessairement précis, mais reste tout de même acceptable, puisqu'il est prédit une augmentation du score accru de l'infirmière. De la semaine 33 à 39, l'infirmière 11 retourne graduellement au travail, effectuant d'abord des heures de travail définies comme *Retour progressif en assurance salaire en régulier*. De la semaine 33 à 45, le score de l'infirmière atteint à nouveau une valeur de 5 à la semaine 45. Les valeurs prédites ne semblent pas en mesure de suivre les valeurs réelles. Supposant que l'infirmière 11 a été victime d'un arrêt de travail causé par une surcharge, le modèle de prédiction semble en mesure d'offrir des prédictions pertinentes sur l'état de l'infirmière à risque. Ainsi, les prédictions de ce modèle pourraient être utilisées afin de réduire les chances d'une augmentation de la surcharge.

ANALYSE DE L'INFIRMIÈRE 16

Cette infirmière a aussi effectué un arrêt de travail, commençant à la semaine 29. Ce cas est pertinent, puisqu'en comparaison avec le cas de l'infirmière 11 (Figure 4.7), celle-ci subit une dégradation graduelle de son score. En observant les données historiques, cette infirmière est partie en vacances planifiées, mais n'est jamais revenue au travail à la suite de cette période de vacances. La Figure 4.8 montre l'évolution du bien-être de l'infirmière 16 par rapport aux valeurs prédites (Prédit) et calculées (Réel). Le modèle a pris un total de 10.07 secondes pour être généré, avec un MSE de 3.3264 et un MAE de 1.3754.

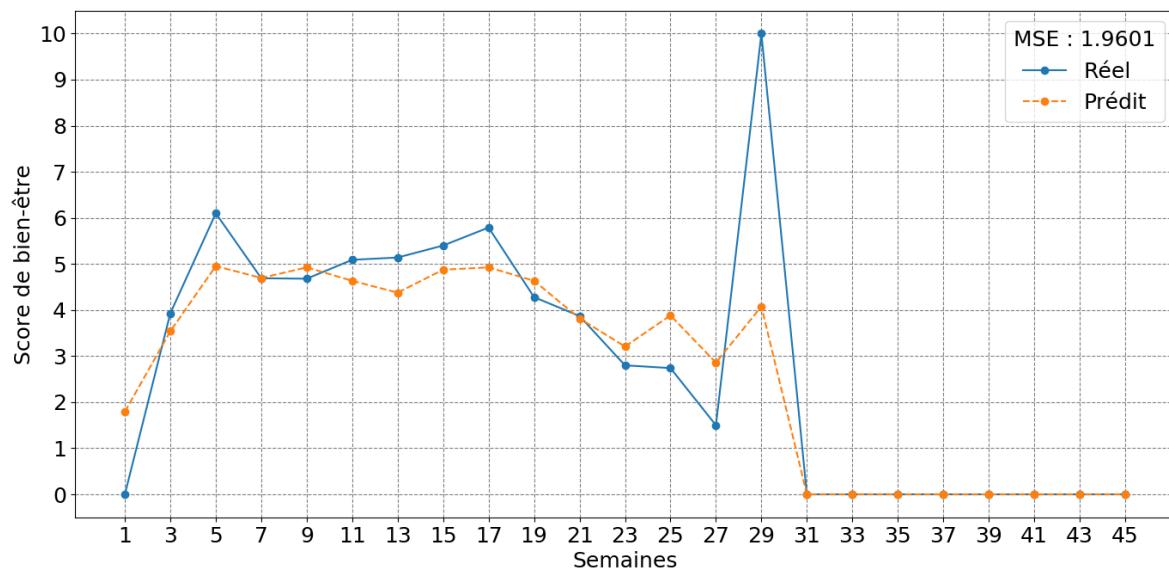


Figure 4.8 : Analyse de l'évolution du score de bien-être prédit et réel de l'infirmière 16.

© 2021 Yoan Villeneuve.

D'abord, le modèle fait un très bon travail de prédiction par rapport aux valeurs réelles. De la semaine 1 à la semaine 29, les prédictions sont représentatives des scores réels et la tendance de l'évolution du score est maintenue. De la semaine 17 à 27, une diminution

graduelle du score est observable, passant d'un score d'une valeur de 6 à environ 1.5. Cette baisse est justifiable du fait qu'elle n'a pas fait immédiatement de l'assurance salaire. L'arrêt de travail de cette infirmière a débuté à la semaine 27. À partir de cette date, celle-ci part deux semaines en vacances, mais elle n'est pas retournée au travail par la suite, optant pour de l'assurance salaire. Ainsi, le modèle peut difficilement prévoir un arrêt de travail, car les infirmières ont généralement deux semaines de vacances planifiées durant cette période de l'année, soit la période des mois de juillet et août. Ainsi, la prédiction par le modèle n'est pas entièrement satisfaisante, mais celui-ci prédit tout de même un score plus haut que le score réel calculé à partir de la semaine 23. Cette observation sous-entend que le modèle considère que le score de l'infirmière devrait être plus haut que prévu. Pendant le retour de l'infirmière à la semaine 30, le modèle de prédiction prédit un score de 4, semblable à la semaine 25 qui précède son départ en vacances.

ANALYSE DE L'INFIRMIÈRE 63

Cette infirmière auxiliaire a commencé un arrêt de travail à la semaine 35. Ce cas est particulier puisque celle-ci est partie en congé maladie avant d'être placée en assurance salaire. Cet arrêt de travail a eu lieu quelques semaines suivant le congé de vacances d'été. L'évolution du score de bien-être est observée à la Figure 4.9, montrant l'évolution du bien-être de l'infirmière 16 par rapport aux valeurs prédites (Prédit) et calculées (Réel). Le modèle a pris un total de 11.13 secondes pour être généré, avec un MSE de 3.2992 et un MAE de 1.3505.

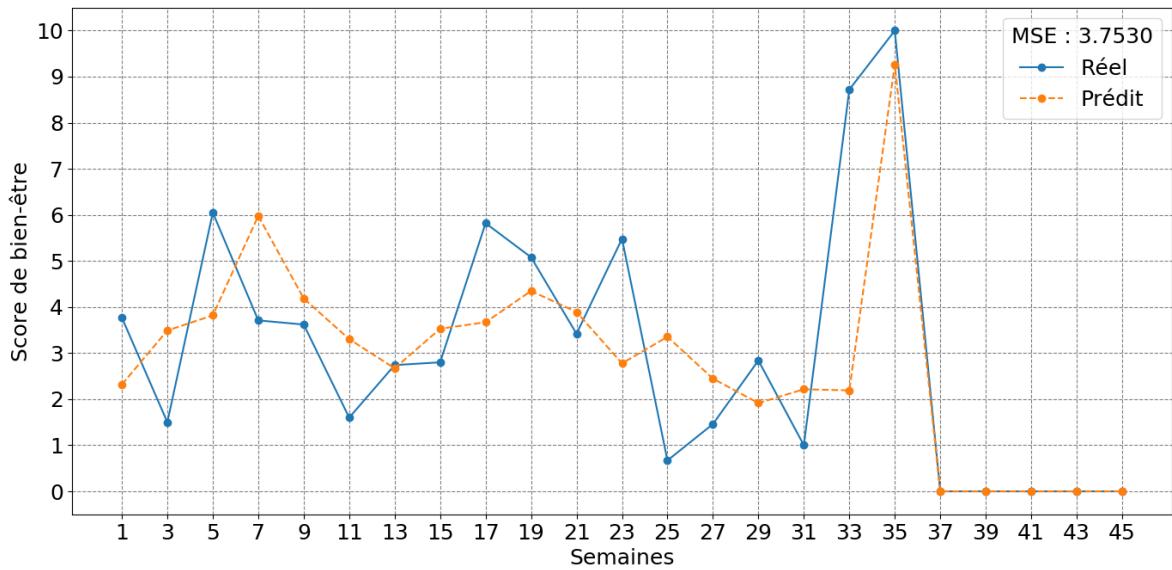


Figure 4.9 : Analyse de l'évolution du score de bien-être prédict et réel de l'infirmière 63.
 © 2021 Yoan Villeneuve.

De toutes les infirmières analysées dans le cadre des arrêts de travail, celle-ci a le score le moins constant. Le score calculé varie d'un extrême à un autre en l'espace d'une semaine. En analysant l'ensemble de données historiques, cette infirmière travaille des semaines avec un grand nombre de temps supplémentaires, suivis de semaines seulement des quarts de travail réguliers. Celle-ci a quelques quarts en congé maladie, soit durant les semaines 5, 19 et 21. Concernant le modèle de prédiction, les variations de score élevées de l'infirmière semblent rendre la prédiction difficile. Les prédictions semblent tout de même suivre une tendance relativement en accord avec l'évolution du score réel. De la semaine 1 à 15, les prédictions sont relativement justes, mis à part les horaires des semaines 5 et 7. Allant de la semaine 17 à 31, les prédictions suivent une tendance vers le bas, ce qui concorde avec la tendance du score réel, mais le score calculé varie trop de manière spontanée pour obtenir une prédiction globale de bonne qualité. Cela est supporté par une haute valeur du MSE. En effet, la valeur du MSE est pratiquement le double des valeurs de MSE obtenues pour les autres infirmières analysées dans ce chapitre, ce qui témoigne d'une très grande disparité entre les valeurs prédites et

réelles. L'infirmière 63 est partie en congé maladie vers la fin de la semaine 33, pour ensuite effectuer de l'assurance salaire pour le reste des données accessibles. Notons que celle-ci effectue des quarts étiquetés *Ret. Prog. Ass.Sal* (dans les données historiques) au courant de la semaine 44 à 47, ce qui signifie un possible retour au travail, tel que vu dans l'arrêt de travail de l'infirmière 11 à la Figure 4.7.

Le modèle de prédiction de l'absentéisme est en mesure de prédire le bien-être d'une infirmière avec une précision satisfaisante. Même si les prédictions ne sont pas toujours identiques aux valeurs réelles calculées, les prédictions sont en mesure de suivre la tendance de l'évolution du score pour l'ensemble des infirmières. Le modèle est ainsi capable de prévoir le score de bien-être des infirmières et de suivre la tendance de l'évolution du score de façon satisfaisante. Néanmoins, certains cas où une infirmière montre un manque de constance s'avèrent difficiles pour le modèle. Le modèle a été en mesure de prédire une augmentation du score de bien-être menant à l'arrêt de travail de deux infirmières (11 et 16). Il doit être mentionné que la nature de ces arrêts de travail est inconnue, ce qui rend les conclusions purement spéculatives. Un plus grand ensemble de données avec un plus grand nombre d'arrêts de travail pourrait améliorer le modèle.

4.3 RÉSULTATS DU MODÈLE D'OPTIMISATION

Cette section se concentre sur les résultats obtenus par le modèle d'optimisation pour la génération d'horaire de travail. D'abord, les résultats du modèle sont présentés sous un format simple. Pour ce faire, plusieurs paramètres sont simplifiés pour obtenir les premiers résultats d'optimisation. Graduellement, les paramètres du modèle se complexifient, jusqu'à la prédiction de réels horaires à partir des données historiques. Cette approche permet de visualiser l'influence des paramètres sur la confection d'horaire, mais aussi des contraintes liées à ces paramètres. La deuxième partie expose la comparaison du modèle avec des horaires

en provenance des données historiques. Ces comparaisons décrivent des horaires concernant les intervalles de temps où certaines infirmières ont fait un arrêt de travail. Enfin, une dernière section est dédiée à un modèle alternatif où aucun temps supplémentaire n'est assigné.

4.3.1 DONNÉES UTILISÉES

Tout d'abord, les infirmières se séparent en deux groupes, soit l'unité d'infirmières et l'unité d'infirmières auxiliaires. Ces dernières se différencient par leur spécialisation et par les besoins en quotas pour chaque quart. L'unité d'infirmières à un besoin de quatre infirmières de jour, trois de soir et deux de nuit. L'unité d'infirmières auxiliaires a un besoin de trois auxiliaires de jour, deux de soir et une de nuit. L'unité d'infirmières comprend 14 infirmières actives à différents niveaux pendant la durée des horaires en provenance des données historiques et de 11 infirmières pour l'unité auxiliaire.

Chaque infirmière est reliée à un quart prioritaire propre à son contrat de travail. Un quart prioritaire représente la période de la journée où une infirmière est censée travailler selon son contrat de travail. Les quarts de travail secondaire sont des quarts en temps régulier ou volant et travaillés à l'extérieur de la période décrite par le contrat de travail d'une infirmière. Par exemple, une infirmière de jour travaillant un quart de soir est considérée comme quart secondaire. La période où une infirmière a été la plus présente est son quart principal (**P**). Pour les autres périodes, elles sont étiquetées comme secondaires (**S**) si une infirmière a travaillé plus de 20% du total de ses quarts de travail. Le Tableau 4.8 et 4.9 montrent les quarts de travail principaux et secondaires ayant été attribués pour chaque infirmière, selon le nombre de quarts travaillés pour chaque période.

Tableau 4.8 : Répartition des types de quarts travaillés par infirmière selon les données historiques.

(a) Pourcentage des types de quarts travaillés.

MATRICULE	JOUR	SOIR	NUIT
Employé 105	59.21	39.66	1.11
Employé 143	39.65	50.00	10.34
Employé 26	73.20	26.79	0.00
Employé 59	67.85	32.14	0.00
Employé 16	75.54	0.00	24.45
Employé 90	80.34	19.65	0.00
Employé 96	15.89	5.64	78.46
Employé 84	54.81	37.34	7.83
Employé 47	1.36	97.27	1.36
Employé 10	80.73	18.34	0.91
Employé 94	51.78	48.21	0.00
Employé 139	30.76	50.00	19.23
Employé 85	51.56	26.56	21.87
Employé 119	72.05	15.44	12.50

(b) Périodes de travail principal et secondaire pour chaque infirmière.

MATRICULE	JOUR	SOIR	NUIT
Employé 105	P	S	
Employé 143	S	P	
Employé 26	P	S	
Employé 59	P	S	
Employé 16	P		S
Employé 90	P		
Employé 96			P
Employé 84	P	S	
Employé 47		P	
Employé 10	P		
Employé 94	P	P	
Employé 139	S	P	
Employé 85	P	S	S
Employé 119	P		

Tableau 4.9 : Répartition des types de quarts travaillés par infirmière auxiliaire selon les données historiques.

(a) Pourcentage des types de quarts travaillés.

MATRICULE	JOUR	SOIR	NUIT
Employé 7	84.80	15.20	0.00
Employé 41	0.00	1.67	98.32
Employé 11	73.50	26.49	0.00
Employé 19	50.41	49.16	0.41
Employé 12	84.45	15.12	0.42
Employé 109	76.37	23.62	0.00
Employé 5	99.01	0.99	0.00
Employé 39	29.13	70.00	0.87
Employé 63	75.63	24.37	0.00
Employé 38	7.06	10.59	82.35
Employé 46	9.74	4.54	85.71

(b) Périodes de travail principal et secondaire pour chaque infirmière.

MATRICULE	JOUR	SOIR	NUIT
Employé 7	P		
Employé 41			P
Employé 11	P	S	
Employé 19	P	P	
Employé 12	P		
Employé 109	P	S	
Employé 5	P		
Employé 39	S	P	
Employé 63	P	S	
Employé 38			P
Employé 46			P

4.3.2 RÉSULTATS DU MODÈLE

Dans cette section, les résultats obtenus servent à montrer l'influence des paramètres dans le modèle et la capacité du modèle à générer des horaires. Le tout est fait graduellement en augmentant la complexité de l'utilisation des paramètres à chaque nouvelle sortie de résultats. Les horaires sont construits sur une durée de 14 jours, commençant la nuit de dimanche. Fidèle à la convention collective des infirmières, chaque infirmière doit travailler seulement une fin de semaine par horaire, soit aux deux semaines (voir les contraintes dans la Sous-Section 3.3.5). La première fin de semaine est considérée comme le premier dimanche et le dernier samedi à l'horaire.

Le calcul des résultats a été effectué à partir d'un ordinateur portable avec l'édition Windows 8.1 pro 64-bits, 8 Gigabits de mémoire vive (RAM) et un processeur Intel(R) Core(TM) i5-4300U @ 2.50 GHz. Le modèle d'optimisation a été écrit et modélisé à l'aide de l'outil AMPL, développé par AMPL Optimization Inc. (2021). Le solveur Xpress par FICO (2021) est utilisé pour résoudre le modèle d'optimisation en nombres entiers.

HORAIRE DE BASE

La première version des résultats est obtenue en utilisant des paramètres diminuant la complexité du problème afin d'obtenir des résultats d'horaire simplifiés. Le Tableau 4.10 résume les paramètres utilisés afin d'obtenir l'horaire de base du Tableau 4.11.

En utilisant ces paramètres avec le modèle d'optimisation linéaire appliqué sur l'unité d'infirmières, l'horaire du Tableau 4.11 est obtenu. La durée de résolution du problème a été de 0.14 secondes.

Tableau 4.10 : Paramètres utilisés pour les résultats du Tableau 4.11 .

Paramètre	Description	Valeur
SC_i	Score de bien-être	1
I^{inf}	Nombre d'infirmière	14
Con_i	Absence planifiée	0
Exp_i	Expertise	0
$Pref_{ijq}$	Préférence	0
B^{inf}	Nombre d'affectation min.	8

Tableau 4.11 : Résultat de la génération d'horaire de l'unité d'infirmières avec le paramétrage présenté au Tableau 4.10 .

Inf.	JOURS														
	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	
105	J	J		J	J			*N	J	J	J			J	
143				S	S			S	S	S	S		S	S	
26		J	J			J	J	J		J	J	J			
59		S		J	J			J	J	J	J		J	J	
16	N	N		N	N			*S		N		N	N	N	
90	J		J	J		J			J			J	J	J	
96	N	N	N		N	N	S		N	N	N		N	N	
84		J			J	J	J	J		J	S		J	*S	
47	S	S	S		S	S			S		S		S	S	
10	J		J	J		J			J		J	J		J	
94	*S	S	S	S		S	S	S		S	S	S		S	
139	S		S	S	S	S	*N		S	S		S		S	
85	J		N	N		N		*N	N			N	N	J	
119		J	J		J		J*N			J		J	J		
Matricule	105	143	26	59	16	90	96	84	47	10	94	139	85	119	
Bien-être	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Contrat	J	S	J	J	J	J	N	J	S	J	J-S	S	J	J	
Régulier	8	8	8	8	8	8	10	8	9	8	10	9	8	8	
Temps Supp.	1	0	0	0	1	0	1	1	0	0	1	1	1	1	
Préférence	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Le Tableau 4.11 montre le résultat de l'horaire généré. Le premier étant l'horaire de quart de travail assigné selon le matricule d'une infirmière et le jour de l'affectation. Lorsqu'une

infirmière est affectée à un quart, la lettre **J** (jour) ; **S** (soir) ou **N** (nuit) est affichée pour montrer le type de quart effectué durant cette journée. Lorsqu'un astérisque (*) se trouve devant une lettre, cela signifie une affectation de temps supplémentaire. Par exemple, **J*S** représente une affectation de temps régulier à un quart de jour (**J**) et une affectation de temps supplémentaire de soir (***S**). Le bas du tableau montre les données statistiques de l'horaire généré. Les trois premières lignes du tableau montrent le score de bien-être calculé par le modèle de prédiction et le contrat de travail primaire relié à chaque infirmière. Les trois dernières lignes affichent de l'information sur l'affectation de quarts pour chaque infirmière. Les lignes **Régulier** et **Temps Supp.** donnent de l'information sur le nombre de quarts affectés par infirmière. **Préférence** représente le nombre de quarts de préférence assignés pour chaque infirmière. Les quarts de préférence sont des quarts de travail choisis par les infirmières. Une infirmière peut choisir jusqu'à 10 quarts primaires à l'horaire avant chaque génération d'horaire. Comme ce projet n'a pas pu être testé en temps réel, les préférences des infirmières sont toujours générées aléatoirement, selon leur contrat de travail.

En ce qui a trait à l'analyse du Tableau 4.11, chaque infirmière est affectée de huit à 10 quarts à l'horaire. Sauf quelques exceptions, celles-ci sont pratiquement toutes affectées au bon quart par rapport à leur contrat de travail. Tous les quarts de l'employé 94 sont assignés le soir, bien qu'elle soit en mesure de travailler le jour et le soir. Cela est dû au fait qu'il y a beaucoup plus de besoins pour une infirmière de soir qu'une infirmière de jour. L'infirmière 16 est de jour, mais a pourtant reçu seulement des quarts de nuit. Un phénomène similaire est observable chez l'infirmière 85. Le problème provient d'un manque d'infirmière de nuit dans l'unité. Seule l'infirmière 96 est reconnue, selon son contrat de travail, comme une infirmière de nuit. Vu le quota de deux infirmières par quart de nuit (dont 28 plages horaires à remplir), il est impossible de satisfaire cette exigence. En ce qui concerne les temps supplémentaires, six infirmières sur 14 sont affectées à un quart de temps supplémentaires. En portant attention

aux jours de la semaine, les quarts de temps supplémentaire affectés sont tous durant les fins de semaine, soit samedi et dimanche. De plus, seuls les quarts de soir et de nuit sont donnés en temps supplémentaire. Ce manque de personnel durant les fins de semaine est dû aux contraintes de la Sous-Section 3.3.5. Cette contrainte provient de la convention collective des infirmières. Entre autres, cette dernière limite le nombre de fins de semaine de travail à une sur deux. Cela montre qu'il y a un manque d'infirmière de soir et de nuit pour pouvoir satisfaire cette contrainte.

En appliquant les paramètres du Tableau 4.10 sur l'unité d'infirmières auxiliaires, le résultat de l'horaire dans le Tableau 4.12 est obtenu. Seule la borne de l'Éq. (3.21) a été modifiée à 4 en raison du surplus d'infirmières par rapport aux quotas. Cette unité est composée de 11 infirmières auxiliaires, avec un quota de travail de trois infirmières le jour, deux la nuit et une le soir. La durée de résolution du problème est la même que pour l'horaire précédent (Tableau 4.11).

Tableau 4.12 : Résultat de la génération d'horaire de l'unité d'infirmières auxiliaires avec paramétrage de base.

Aux.	Dim.	JOURS													
		Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	
7		J			J	J	J+S	J			J	J	J		
41	N					N		S				N		N	
11		J	J	J			J	J	J		J	J	J		
19	S	S	S		S	S			S	S	S	S		S	
12	J				J	J			J		J			J	
109	J	J	J	S					J	J	J	J		J	
5	J		J	J		J			J				J	J	
39	S		S	S	S	S			S	S		S	S	S	
63		S		J	J		J	J		J	S		S		
38					N		N+S	N+S			N				
46		N	N	N					N	N				N	
Matricule	7	41	11	19	12	109	5	39	63	38	46				
Bien-être	1	1	1	1	1	1	1	1	1	1	1				
Contrat	J	N	J	J-S	J	J	J	S	J	N	N				
Occurrence	8	4	9	10	5	9	7	10	8	4	6				
Temps Supp	1	1	0	0	0	0	0	0	0	2	0				
Préférence	0	0	0	0	0	0	0	0	0	0	0				

Une grande disparité dans le nombre de quarts attribués est observable. Le nombre d'affectations varie entre 4 et 10. Les infirmières de nuit sont celles recevant le moins de quarts. Puisqu'il y a un quota d'une infirmière auxiliaire par nuit, le modèle n'a simplement pas assez de quarts à offrir pour ces dernières. Ceci est causé par la contrainte de l'Éq. (3.21) puisque la borne B^{aux} est paramétrée à 4. Or, dans la Sous-Section 4.3.3, les horaires disponibles dans les données historiques montrent qu'il n'est pas rare que deux infirmières de nuit travaillent sur un même quart, malgré le quota de nuit. Il arrive aussi que les infirmières de nuit ne soient pas toutes disponibles à l'horaire. Ainsi, le modèle tentera toujours de respecter les quotas, mais il en revient au gestionnaire de décider s'il y a un besoin pour plus ou moins d'infirmières durant cette période. D'autre part, l'auxiliaire 19 porte des similarités avec l'infirmière 94. Celle-ci a

un contrat de travail englobant deux quarts de travail principal, soit jour et soir. Néanmoins, elle n'est assignée qu'à des quarts de soir. Cette unité souffre aussi d'un manque de personnel, mais seulement pour les infirmières de soir. Cette remarque est accentuée par les quatre quarts de temps supplémentaires assignés à l'horaire, tous des quarts de soirée au courant de la fin de semaine. Comme l'unité d'infirmières, l'unité d'auxiliaires fait face à un déséquilibre de la main-d'oeuvre similaire, ayant un surplus d'infirmières de jour et un manque d'infirmières de nuit.

L'horaire de base permet d'avoir un premier point de vue sur les résultats à venir. Il montre les défauts intrinsèques de la méthodologie de gestion du personnel de cette unité de soins. Avant même d'effectuer l'optimisation des horaires, ce premier jet de résultat montre qu'il y a un manque de personnel de soir et de nuit. En ce qui concerne le quota de l'unité d'infirmières, un horaire offre 56 plages horaires de jour à remplir pour 10 infirmières ; 42 plages horaires de soir pour quatre infirmières et 28 plages horaires de nuit pour une seule infirmière. Pour l'unité d'infirmières auxiliaires, l'horaire offre 42 plages horaires de jour pour sept infirmières ; 28 plages horaires de soir pour deux infirmières et 14 plages horaires de nuit pour trois infirmières. Considérant qu'une infirmière doit être affectée de 8 à 10 quarts par horaire, un réajustement de l'effectif est à considérer. Malgré tout, le modèle génère des horaires satisfaisant toutes les contraintes lui étant imposées et atteignant le quota de travail pour chaque jour de la semaine.

SCORE DE BIEN-ÊTRE

Dans cette section, le score de bien-être est ajouté afin d'observer son impact sur les résultats obtenus. Certaines infirmières reçoivent un score de bien-être initialisé à 6 afin d'observer l'influence de celui-ci. Ces infirmières sont les matricules 143, 59, 90, 96, 139 et

85. Au total, six infirmières sont assignées un score de bien-être de 6, soit trois infirmières de jour, deux de soir et une de nuit. Afin d'obtenir une représentation juste, les préférences des infirmières sont générées aléatoirement. Chaque infirmière se voit attribuer 10 quarts de préférence pouvant seulement être attribués à leur quart primaire, selon leur contrat de travail.

Le Tableau 4.14 montre l'horaire généré selon les paramètres au Tableau 4.13.

Tableau 4.13 : Paramètres utilisés pour les résultats du Tableau 4.14 .

Paramètre	Description	Valeur
SC_i	Score de bien-être	6 infirmières avec un score de 6 8 infirmières avec un score de 1
I^{inf}	Nombre d'infirmière	14
Con_i	Absence planifiée	0
Exp_i	Expertise	0
$Pref_{ijq}$	Préférence	Aléatoire selon le contrat de travail
B^{inf}	Nombre d'affectation min.	8

Tableau 4.14 : Résultat de la génération d'horaire de l'unité d'infirmières avec score de bien-être fictif.

Inf.	JOURS														
	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	
105	*S	J	J	J			J	J	J	J			J		
143	S		S	S		S					S	S	S	S	
26	*N		J		J	J	J	J	S		J	J		*S	
59		J		J	J		J	J		J		J	J		
16		N	N		N	N	N	N	N		N	N	N		
90	J	J			J	J			J	J	J			J	
96	N			N	N	N			N	N			N	N	
84	J	J	J		J	S	*S		J				J	J	
47		S	S	S	S		S	S		S	S	S	S		
10	J			J		J	*N	*N	J		J	J	J	J	
94		S	S	S	S		S	S	S	S			S		
139	S	S			S	S			S	S		S		S	
85		N	N	N			J	J		N	N	N			
119	J		J	J		J	*S	J		J	J			J*N	
Matricule	105	143	26	59	16	90	96	84	47	10	94	139	85	119	
Bien-être	1	6	1	6	1	6	6	1	1	1	1	6	6	1	
Contrat	J	S	J	J	J	J	N	J	S	J	J-S	S	J	J	
Occurrence	8	8	8	8	10	8	8	8	10	8	10	8	8	8	
Temps Supp.	1	0	2	0	0	0	0	1	0	2	0	0	0	2	
Préférence	7	8	6	8	0	8	7	7	9	8	9	7	1	7	

Avec le score de bien-être, toutes les infirmières ayant un score de 6 ont vu leur nombre de quarts diminués à 8, dû à la contrainte souple de l'Éq. (3.23). Cette contrainte risque de ne pas toujours être respectée en situation de manque d'effectif. À l'exception des infirmières 16 et 85, toutes les infirmières ont reçu la majorité de leurs quarts mis en préférence. Les infirmières 16 et 85 sont les infirmières ayant le quart de nuit étiqueté comme secondaire, tel que montré dans le Tableau 4.9. Comme l'infirmière 85 a un score de bien-être de 6, le modèle lui a attribué un seul quart de préférence et a diminué son nombre de quarts à 8, en raison de la borne B^{inf} de l'Éq. (3.20) ajustée à 8. Aucune infirmière avec un score plus grand que 1 n'a reçu de temps supplémentaire.

Ainsi, le modèle se comporte comme désiré avec les infirmières à haut score de bien-être, c'est-à-dire en limitant le nombre de temps supplémentaire, en augmentant les chances d'obtention de quart de préférence et en diminuant le nombre d'affectations totales de quarts de travail.

PARAMÈTRE RÉALISTE

Pour faire suite à l'introduction du score de bien-être comme paramètre et à la démonstration des résultats après l'ajout de ce dernier, un score différent pour chaque infirmière est appliqué. Le score est choisi par rapport à la dernière semaine en provenance des données historiques, en date du 22 décembre 2019. Cet horaire est incomplet, avec seulement une semaine sur deux enregistrée. De plus, afin d'avoir une représentation plus juste, la borne B^{inf} est réduite à 6. À l'aide des données de l'horaire du 8 décembre 2019 et des paramètres résumés dans le Tableau 4.15, les résultats du Tableau 4.16 sont obtenus.

Tableau 4.15 : Paramètres utilisés pour les résultats du Tableau 4.16 .

Paramètre	Description	Valeur
SC_i	Score de bien-être	Prédiction du 22 décembre 2019
I^{inf}	Nombre d'infirmière	14
Con_i	Absence planifiée	0
Exp_i	Expertise	0
$Pref_{ijq}$	Préférence	Aléatoire selon le contrat de travail
B^{inf}	Nombre d'affectation min.	6

Tableau 4.16 : Résultat de la génération d'horaire de l'unité d'infirmières avec score de bien-être prédit avec la semaine du 8 décembre 2019.

Inf.	JOURS														
	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	
105		J	J		J		J	J		J					
143	S		S		S	S			S		S	S	S	S	
26	S			J		J			J	J	J				S
59					J		J	J	J	J		J	J		
16		N			N		J	J			N		N		
90	*N	J		J	J	J	J	J	J	J		J			*N
96	*N	N	N	N		N	N	N	N	N		N	N		
84	J	J	J		J							J	J		J
47		S		S	S	S	S	S		S	S	S	S		*S
10	J		J	J		J		*N	*N	J	J	J	J	J	
94		S	S	S	S			S	S	S	S			S	
139	*S	S	S	S		S	S	S	S						*N
85	J		N	N	N	N			N	N	N	N			J
119	J	J	J	J	J				J	J	J	J			J
Matricule	105	143	26	59	16	90	96	84	47	10	94	139	85	119	
Bien-être	4,35	3,34	3,3	9,48	9,54	1,93	2,25	4,31	2,81	1,8	2,96	2,41	3,15	3,09	
Contrat	J	S	J	J	J	N	J	S	J	J-S	S	J	J		
Occurrence	6	9	8	6	6	9	10	6	10	9	10	9	10	10	10
Temps Supp.	0	0	0	0	0	2	1	0	1	2	0	2	0	0	0
Préférence	6	8	6	6	2	7	8	6	8	9	9	8	2	9	

Le score de bien-être des infirmières 59 et 16 est élevé, en raison du fait que celles-ci n'ont pas travaillé dans l'horaire réel de ces semaines. Ces dernières se caractérisent dans l'horaire comme ayant obtenu plusieurs congés maladie et assurances salaire. Tous de même, elles n'ont reçu aucun temps supplémentaire et ont obtenu le plus petit nombre de quarts possible. Dans le même ordre d'idée, les matricules 105 et 84 représentent les seules infirmières avec un score de bien-être de 4 et plus. Celles-ci n'ont reçu que le minimum d'affectation de quarts de travail possible, aucun temps supplémentaire et toutes leurs préférences de quarts. Les infirmières ayant été affectées au temps supplémentaire sont celles ayant un score de bien-être plus bas, allant d'une à deux affectations de temps supplémentaire par infirmière. Aucune infirmière ayant un score plus grand que 3 n'a reçu de temps supplémentaire.

RÉSULTAT D'HORAIRE AVEC SCORE DE BIEN-ÊTRE PRÉDIT

Contrairement au résultat de la Sous-Section 4.3.2, l'horaire suivant résulte de l'utilisation de chaque contrainte et paramètre du modèle. Cet horaire utilise comme paramètre les données de l'horaire du 8 décembre 2019 de l'unité d'infirmières, soit le dernier horaire travaillé disponible dans l'ensemble de données historiques. Les infirmières 56, 16 et 139 ont été retirées de l'horaire, car celles-ci n'ont reçu aucune affectation de quart. Les congés pris au courant de cet horaire sont aussi utilisés comme paramètre du modèle. Afin d'expérimenter avec le paramètre **Expertise** (Exp_i), celui-ci est retiré pour les infirmières 105, 26, 47, 94 et 119. Ces dernières ne peuvent pas être assignées seules pour un quart de travail et doivent être accompagnées d'une autre infirmière ayant l'étiquette d'expertise. Une infirmière obtient son expertise lorsque les gestionnaires considèrent que celle-ci a accumulé assez d'expérience dans les soins du département. L'expertise réelle des infirmières n'est pas disponible dans les données de ce projet, mais reste assez importante dans la confection d'horaire pour être incluse dans le modèle. Le Tableau 4.18 montre le résultat de l'horaire obtenu à l'aide des paramètres résumés dans le Tableau 4.17.

Tableau 4.17 : Paramètres utilisés pour les résultats du Tableau 4.18 .

Paramètre	Description	Valeur
SC_i	Score de Bien-être	Prédiction du 8 décembre 2019
I^{inf}	Nombre d'infirmière	11
Con_i	Absence planifiée	7
Exp_i	Expertise	Retiré pour les infirmières 05, 26, 47, 94 et 119
$Pref_{ijq}$	Préférence	Aléatoire selon le contrat de travail
B^{inf}	Nombre d'affectation min.	7

Tableau 4.18 : Résultat de la génération d'horaire de l'unité d'infirmières avec score de bien-être prédit avec la semaine du 8 décembre 2019.

Inf.	Dim.	JOURS												Sam.
		Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	
105	*N	J	S	S		J	J	J	S	S		J	J	
143		S	S		S	S	S	S		S	S	S	S	*N
26		J	J	J	J		J	J		J	S		S	*N
90	J	J		J	J	J	*S		J	J	*N		J	J
96	*S	*N		N	N	N	N	N		N		N	N	
84	J	S	S	S		J	*N		J	J	J	J		J
47	S	S	*N		S	S		*J	S	S	S	S		S
10	J		J	J	J			*N	J	J	J		J	J
94	S		J	S	S	S	*J		S		J	S	S	S
85	J*N			N	N	N			N		N	J		J
119		J	J	J	J		J	J	J		J	J	J	*S

Matricule	105	143	26	90	96	84	47	10	94	85	119
Bien-être	4,35	3,34	3,3	1,93	2,25	4,31	2,81	1,8	2,96	3,15	3,09
Contrat	J	S	J	J	N	J	S	J	J-S	J	J
Occurrence	10	10	9	9	8	10	9	10	10	8	10
Temps Supp.	1	1	1	2	2	1	2	1	1	1	1
Préférence	6	9	6	7	7	7	7	8	8	2	8

En raison du nombre d'infirmières réduit, comparativement aux cas précédents, le nombre d'affectations et de temps supplémentaires ont grandement augmentés afin de pouvoir remplir les quotas de travail. Huit jours sur 14 n'atteignent pas le quota d'infirmières pour au moins un de leurs quarts. Ce faisant, il faut alors faire appel aux infirmières externes pour remplir les manquements de personnel infirmier. Les infirmières ne travaillant pas 10 quarts à l'horaire se sont vue accorder un à deux jours de congés, tel que l'horaire original le demande. Les infirmières 105, 26 et 85 ont reçu le moins de quarts de préférence comparativement au reste de l'unité, particulièrement pour l'infirmière 85. L'employée 105 fait partie du groupe sans expertise et a reçu plusieurs quarts secondaires de soir, ce qui explique un nombre plus bas de quarts préférentiels. L'employée 85 a les quarts de nuit en priorité secondaire. Certains soirs dans la semaine n'ont aucune affectation régulière d'infirmière, ce qui explique le grand

nombre de quarts secondaires pour cette dernière. Il est difficile de créer un horaire, avec ou sans l'utilisation d'un modèle d'optimisation, lorsque l'unité en question est en sous-effectif.

Le modèle n'est pas conçu pour prendre en compte un si grand manque de personnel, mais génère tout de même le meilleur horaire théoriquement possible. La Sous-Section 4.3.3 montre la comparaison de cet horaire avec l'horaire réellement travaillé et bien d'autres comparaisons de scénarios en provenance des données historiques. La Sous-Section 4.3.4 reprend ces scénarios, mais avec une modification du modèle afin de favoriser davantage l'utilisation d'infirmières externes.

4.3.3 COMPARAISON DES RÉSULTATS

Dans cette sous-section, les horaires obtenus avec le modèle de génération d'horaires développé dans ce projet sont comparés avec des situations réelles récupérées des données historiques fournies par le CIUSSS. Parmi les cas analysés, certains cas présentent des infirmières ayant fait un arrêt de travail. Ainsi, les arrêts de travail étudiés dans cette section se basent sur l'observation des données historiques et les conclusions tirées de l'analyse de ces cas sont spéculatives de cette même analyse. Cette partie commence par la comparaison de différents horaires de travail et discute des différences avec ceux générés par le modèle d'optimisation, puis les arrêts de travail des matricules 11, 16 et 19 sont abordés.

Afin de mieux refléter les cas réels avec ceux générés à partir du modèle de programmation linéaire en nombres entiers, une contrainte ne faisant pas partie du modèle de base a été ajoutée. Celle-ci permet une représentation plus fidèle de la semaine de travail en faisant l'addition du nombre d'affectations de quarts réguliers, de TS et de TSO effectués dans l'horaire. Ensuite, le modèle est contraint à affecter le même nombre de quarts de travail que totalisé précédemment. La contrainte suivante est ainsi ajoutée au modèle de la Section 4.3 :

$$\sum_{i \in I} \sum_{j \in J} \sum_{q \in Q} x_{ijq} + ts_{ijq} \leq Total_{reg.} + Total_{vol.} + Total_{ts} + Total_{tso} \quad (4.3)$$

Où les variables *Total* représentent le nombre de quarts travaillés par les infirmières et où l'indice de la variable représente le type de quart concerné (quart régulier, volant, à temps supplémentaire ou TSO). Les variables x_{ijq} et ts_{ijq} représentent chaque affectation de quart régulier et de temps supplémentaire d'une infirmière à la période de travail d'une journée à l'horaire.

COMPARAISON DE L'HORAIRE DU 8 DÉCEMBRE 2019

L'horaire construit dans le Tableau 4.18 est comparé avec l'horaire réel travaillé à la date du 8 décembre 2019. Comme l'expertise des infirmières n'est pas connue, celle-ci est dorénavant toujours paramétrée à 1, c'est-à-dire, toutes les infirmières sont considérées comme étant aptes à effectuer un quart de travail sans avoir besoin de supervision provenant d'une infirmière plus expérimentée. Sinon, le reste du paramétrage reste le même. Le Tableau 4.19 est une représentation de l'horaire travaillé construit à partir des données historiques. Elle permet ainsi de comprendre la réalité des horaires des infirmières et permet d'effectuer de meilleures comparaisons avec les résultats du modèle d'optimisation.

Tableau 4.19 : Données historiques tirées de l'horaire de l'unité d'infirmières du 8 décembre 2019 au 21 décembre 2019

Matricule	105	143	26	59	16	90	96	84	47	10	94	139	85	119
Bien-être	3,88	3,05	3,51	9,03	9,24	2,72	2,9	3,72	3,32	2,54	2,69	2,37	3,75	3,26
Contrat	J	S	J	J	J	J	N	J	S	J	J-S	S	J	J
Régulier	10	5	7	0	0	9	6	7	9	9	5	0	6	7
Temps Supp.	1	0	0	0	0	0	0	1	0	0	1	0	0	0
TSO	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Volante	0	0	1	0	0	0	0	0	0	0	0	0	0	0
ABSENCE	0	0	1	0	0	0	0	1	0	1	0	0	0	0
MALADIE	0	0	0	1	2	0	0	0	0	0	0	0	0	0
ASSURANCE	0	0	0	5	8	0	0	0	0	0	0	0	0	0
CONGÉ	0	0	1	0	0	1	2	0	0	0	0	0	2	0
CHOMAGE	0	0	0	0	0	0	0	0	1	0	0	0	0	0

En retirant les employées n'ayant pas travaillé au courant de la semaine, la moyenne du score de bien-être est 3,21. En ce qui concerne les absences, trois employées ne sont pas présentes pour leurs quarts au courant de cet horaire, deux d'entre elles ayant un score plus haut que la moyenne de l'horaire. Six congés ont été distribués entre quatre infirmières et une autre infirmière a effectué un jour de chômage. Le chômage est affecté comme paramètre de congé pour les infirmières concernées, car ce type de quart reste tout de même une absence planifiée. Par rapport au quart de travail, le Tableau 4.20 montre les paramètres du modèle utilisé et la valeur pour le type de quart effectué, utilisé comme paramètres pour l'Éq. (4.3).

Tableau 4.20 : Paramètres utilisés pour les résultats du Tableau 4.21 .

Paramètre	Description	Valeur
SC_i	Score de bien-être	Prédiction du 8 décembre 2019
I^{inf}	Nombre d'infirmière	11
Con_i	Absence planifiée	7
Exp_i	Expertise	0
$Pref_{ijq}$	Préférence	Aléatoire selon le contrat de travail
B^{inf}	Nombre d'affectation min.	7
$Total_{reg}$	paramètre pour l'Éq. (4.3)	80
$Total_{vol}$	paramètre pour l'Éq. (4.3)	1
$Total_{ts}$	paramètre pour l'Éq. (4.3)	3
$Total_{iso}$	paramètre pour l'Éq. (4.3)	0

En utilisant l'information précédente afin de générer un horaire, les résultats obtenus sont interprétés dans le Tableau 4.21. Le score de bien-être a été calculé à partir de l'horaire précédemment travaillé, soit celui du 24 novembre 2019. La prédiction du score de bien-être a pris 41,15 secondes, avec un MSE de 2,8278 et un MAE de 1,3305. Une baisse de la précision du modèle est normale puisque la semaine prédite a été retirée du processus de création du modèle. Le temps de calcul de cet horaire a été de 0,79 secondes. Il y a un total de 84 affectations de quart de travail à l'horaire, ce qui correspond à l'historique.

Tableau 4.21 : Résultat de la génération de l'horaire de l'unité d'infirmières à des fins de comparaison avec l'horaire du 8 décembre 2019.

Inf.	JOURS														
	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	
105				J		J	J	J	J			J	J	J	
143		S	S		S	S	S	S		S	S		S		
26		J					J	J		J	J		J		
59	J	J		J	J	J			J		J		J	J	
96	N			N	N	N				N		N	N	N	
84	J		J	J		J			J		J		J	J	
47	S		S		S				S			S		S	
10	J	J	*N		J	J			J	J	J		J	J	
94	S	S		S		S			S				S	S	
85		N		J	J		N	N	N		N	J			
119		J		J		J	I	J	J		J	J			
Matricule	105	143	26	90	96	84	47	10	94	85	119				
Bien-être	6,09	3,08	3,58	2,1	2,58	4,64	3,63	1,82	3,33	2,83	3,54				
Contrat	J	S	J	J	N	J	S	J	J-S	J	J				
Régulier	7	9	6	9	8	7	6	9	7	8	7				
Temps Supp.	0	0	0	0	0	0	0	1	0	0	0				
Préférence	7	9	6	8	7	7	6	8	7	3	6				

L'horaire généré montre une amélioration dans le nombre d'affectations de quarts à temps régulier. Chaque infirmière obtient de six à neuf quarts de travail réguliers et seule l'infirmière 10 a reçu un temps supplémentaire avec un score de bien-être de 1,82. En raison de la limite de quarts attribuables (Éq. (4.3)) ainsi que du quota de travail de l'unité d'infirmières (Tableau 3.3), il est impossible de remplir les quotas de travail pour chacun des quarts. Malgré tout, huit quarts de jour ont atteint leurs quotas et aucun quart ne se retrouve sans infirmière. En général, les quarts ne respectant pas le quota de travail sont en sous-effectif d'une seule infirmière. De toutes les journées, le premier mardi de l'horaire diffère de la norme. Ce dernier n'a qu'une seule infirmière affectée de jour, comparativement à un minimum de trois pour le reste des jours de l'horaire, en plus de détenir le seul temps supplémentaire pour le quart de

nuit. Cette situation s'explique par les préférences des infirmières choisies aléatoirement. En étudiant les paramètres de préférence générés aléatoirement, le premier mardi est la journée avec le moins de préférence d'infirmière, avec cinq au total. Le deuxième mardi de la semaine a aussi un total de cinq préférences choisies. Ce quart a aussi peu d'affectations comparativement au reste de l'horaire. Pratiquement toutes les préférences de travail ont été accordées aux infirmières, sauf le matricule 85. Cette dernière travaille cinq quarts de nuit, bien qu'elle soit une infirmière de jour (Tableau 4.9). Dans un contexte réel, il y aurait probablement plus d'affectation de quarts de nuit, ou encore l'utilisation d'infirmières externes pour remplir les quotas manquants. Concernant les infirmières à haut score, soit l'infirmière 105 et 84, celles-ci ont reçu tous leurs quarts de préférence, un nombre d'affectations réduit à l'horaire et aucun temps supplémentaire. D'autres infirmières, matricules 26 et 47, ont aussi reçu un nombre d'affectations réduit. Ce bas nombre d'affectations est dû à leur score de bien-être relativement élevés ainsi qu'à un nombre d'affectations limité à 84 pour l'horaire.

ARRÊT DE TRAVAIL DE L'INFIRMIÈRE 11

L'arrêt de travail de l'infirmière auxiliaire 11 est analysé dans cette partie. Il est établi que celle-ci a effectué un arrêt de travail de la semaine 19 à la semaine 35. Au cours de ces semaines, l'infirmière était en assurance salaire, comme indiqué dans les données historiques. Afin de comprendre l'arrêt de travail de l'infirmière 11, observons les données des horaires à la semaine 15, 17 et 19. Le Tableau 4.22 montre les horaires de travail de la semaine 15, 17 et 19 dans cet ordre.

Tableau 4.22 : Données numériques des horaires de travail et de congé partant de la semaine du 12 mai 2019 au 22 juin 2019.

(a) Horaire de la semaine 15-16 de l'unité d'infirmières.

MATRICULE	7	41	11	19	12	109	5	39	63	38	46
Bien-être	2.5	7	6.52	3.56	6.16	3.95	0	4.44	2.8	1.5	7.47
Contrat	j	n	j	js	j	j	j	s	j	n	n
Régulier	9	9	6	9	9	9	4	5	10	0	0
Temps Supp	2	4	0	0	4	1	0	0	1	2	0
TSO	0	0	0	0	0	0	0	0	1	0	0
Volante	0	0	2	1	0	0	0	0	0	0	0
ABSENCE	0	0	0	0	0	0	0	5	1	0	0
MALADIE	0	0	0	0	0	0	0	0	0	0	0
ASSURANCE	0	0	0	1	0	0	3	0	0	0	5
CONGÉ	0	1	0	0	0	0	0	0	1	0	0
CHOMAGE	0	0	0	0	0	0	0	0	0	0	0

(b) Horaire de la semaine 17-18 de l'unité d'infirmières.

MATRICULE	7	41	11	19	12	109	5	39	63	38	46
Bien-être	2.5	7	6.52	3.56	6.16	3.95	0	4.44	2.8	1.5	7.47
Contrat	j	n	j	js	j	j	j	s	j	n	n
Régulier	9	9	6	9	9	9	4	5	10	0	0
Temps Supp	2	4	0	0	4	1	0	0	1	2	0
TSO	0	0	0	0	0	0	0	0	1	0	0
Volante	0	0	2	1	0	0	0	0	0	0	0
ABSENCE	0	0	0	0	0	0	0	5	1	0	0
MALADIE	0	0	0	0	0	0	0	0	0	0	0
ASSURANCE	0	0	0	1	0	0	3	0	0	0	5
CONGÉ	0	1	0	0	0	0	0	0	1	0	0
CHOMAGE	0	0	0	0	0	0	0	0	0	0	0

(c) Horaire de la semaine 19-20 de l'unité d'infirmières.

MATRICULE	7	41	11	19	12	109	5	39	63	38	46
Bien-être	2.5	7	6.52	3.56	6.16	3.95	0	4.44	2.8	1.5	7.47
Contrat	j	n	j	js	j	j	j	s	j	n	n
Régulier	9	9	6	9	9	9	4	5	10	0	0
Temps Supp	2	4	0	0	4	1	0	0	1	2	0
TSO	0	0	0	0	0	0	0	0	1	0	0
Volante	0	0	2	1	0	0	0	0	0	0	0
ABSENCE	0	0	0	0	0	0	0	5	1	0	0
MALADIE	0	0	0	0	0	0	0	0	0	0	0
ASSURANCE	0	0	0	1	0	0	3	0	0	0	5
CONGÉ	0	1	0	0	0	0	0	0	1	0	0
CHOMAGE	0	0	0	0	0	0	0	0	0	0	0

Avant d'effectuer les premières observations, prenons note que les infirmières 38 et 46 ne sont pas actives à l'horaire. L'infirmière 38 n'effectue que deux temps supplémentaires et l'infirmière 46 est absente, comme l'indique son nombre d'assurances salaire obtenues par horaire. De plus, l'infirmière 5 a un score de bien-être à 0, car celle-ci était en arrêt de travail dans les semaines précédentes. Malheureusement, cet arrêt de travail date d'un horaire antérieur à ceux disponibles dans le cadre de cette recherche. L'arrêt de travail de celle-ci ne peut donc pas être documenté. Ces infirmières sont omises dans l'analyse qui suit, par souci de simplicité.

Par rapport au score de bien-être des infirmières au cours de ces six semaines, les infirmières 41, 11 et 12 sont celles ayant le plus haut score calculé. Le score de l'infirmière 41 est justifiable par le manque d'infirmière de nuit, étant la seule infirmière de ce genre active à l'horaire. De plus, l'état de cette infirmière s'améliore graduellement à chaque horaire. L'infirmière 12 suit une progression similaire à celle de l'infirmière 11. Celle-ci effectue des semaines complètes de travail (9 à 10 quarts par horaire) et a travaillé un nombre anormalement élevé de temps supplémentaires pour la première semaine (semaine 15) observée. Or, une analyse des habitudes de travail de cette infirmière montre qu'il est rare que le nombre de quarts à temps supplémentaire effectués excède un ou deux par horaire. Celle-ci montre un niveau de constance très élevé dans ses horaires de travail, puisqu'elle respecte toujours son contrat de travail et prend rarement congé, planifié ou non.

Concernant l'infirmière 11, celle-ci est l'infirmière ayant reçu les plus hauts scores de bien-être cumulatifs des trois horaires observés. Elle a travaillé huit quarts (six réguliers et deux volantes) durant la semaine 15, 10 quarts réguliers à la semaine 17 avec une absence et deux quarts à la semaine 19 suivi de congés maladie et d'assurances salaire. Le score d'une infirmière est influencé par trois facteurs, tels que vus à la Sous-Section 3.2.2, soit le surmenage, les absences non planifiées et une mauvaise affectation de quart. Pour le surmenage,

celle-ci ne semble pas trop avoir travaillé au courant des dernières semaines, soit huit et 10 quarts de travail, mais elle a tout de même participé deux fois à l'équipe volante, ce qui constitue un stress. L'infirmière a aussi été absente une fois dans la semaine 17. Pour aider à mieux comprendre la situation de l'infirmière, regardons son horaire de travail pour les trois horaires dans le Tableau 4.23.

Tableau 4.23 : Horaire détaillé de chaque jour pour l'infirmière 11 pour les horaires de la semaine 15 à 19.

sem.	JOURS														
	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	
15		J	J	J			J		S	S			J	J	
17	J		S	S			J	J	J	J			S	S	S
19	J	J													

D'abord, l'infirmière 11 semble effectuer beaucoup de quarts secondaires à son contrat de travail. Celle-ci est une infirmière de jour, mais durant la semaine 17-18, celle-ci a effectué la moitié des quarts de travail en secondaire. De plus, celle-ci a travaillé durant toutes les fins de semaine des trois horaires. Les cinq derniers quarts travaillés ont été majoritairement de soir et travaillés consécutivement. Avant ceux-ci, quatre autres quarts consécutifs de jour ont été travaillés, précédant une absence faisant suite à deux quarts secondaires. Voyant l'horaire travaillé dans ces trois derniers horaires, avec une absence qui semble témoigner de la surcharge de cette infirmière, il est plausible de relier cet arrêt de travail à une surcharge. Avec l'aide du modèle de prédiction du bien-être et du modèle d'optimisation linéaire en nombres entiers pour concevoir des horaires, voyons s'il est possible de générer un meilleur horaire pour cette infirmière. Prenons l'horaire de la semaine 17-18 puisqu'elle semble avoir été critique dans le processus menant à l'arrêt de travail de cette infirmière. Les paramètres ainsi que le nombre de quarts travaillés sont affichés dans le Tableau 4.24.

Tableau 4.24 : Paramètres utilisés pour les résultats du Tableau 4.25 .

Paramètre	Description	Valeur
SC_i	Score de bien-être	Prédiction du 26 mai 2019
I^{aux}	Nombre d'infirmière	9
Con_i	Absence planifiée	8
Exp_i	Expertise	0
$Pref_{ijq}$	Préférence	Aléatoire selon le contrat de travail
B^{aux}	Nombre d'affectation min.	6
$Total_{reg}$	paramètre pour l'Éq. (4.3)	69
$Total_{vol}$	paramètre pour l'Éq. (4.3)	1
$Total_{ts}$	paramètre pour l'Éq. (4.3)	7
$Total_{iso}$	paramètre pour l'Éq. (4.3)	0

Le score de bien-être a été calculé à partir de l'horaire de la semaine 15, soit du 12 au 25 mai 2019. La prédiction du score de bien-être a pris 13.6122 secondes avec un MSE de 2.7735 et un MAE de 1.3105, des performances légèrement meilleures que pour le modèle de prédiction précédent. Le temps de calcul de cet horaire a été de 0.031 secondes. Il y a eu un total de 77 affectations de quart de travail à l'horaire.

Tableau 4.25 : Résultat de la génération d'un horaire se basant sur les données historiques de la semaine 17 (26 mai 2019).

Aux.	Dim.	JOURS														Sam.
		Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.		
7		J	J		J	J	J	J		J	J	J	J	J	*N	
41		N		N	N	N			N	N	N	N	N	N		
11	J	J		J	J							J		J		
19		S	S	S		S	S	S		S	S	S	S	S		
12				J			J	J	J	J	J	J				
109			J	J	S				S							
5	J*N		*N		J	J	*N	*N	J		J		J	J		
39	S	S	S		S	S			S	S	S		S	S		
63		J	J	S		J	J	J	J	J	J	J	J	J		

Matricule	7	41	11	19	12	109	5	39	63
Bien-être	2,93	6,28	5,95	3,25	6,48	3,07	1	3,38	3,36
Contrat	J	N	J	J-S	J	J	J	S	J
Régulier	10	9	6	10	6	4	7	10	10
Temps Supp.	1	0	0	0	0	0	4	0	0
Préférence	9	7	6	8	6	2	5	8	8

Le nombre de quarts attribué à l'infirmière 11 a été réduit à 6, contrairement à 10 dans l'horaire provenant de l'historique. Sur ces six quarts, ceux-ci sont tous des quarts de préférence (générés aléatoirement) faisant partie du contrat de travail de celle-ci. Cet horaire réduit considérablement la pression sur cette infirmière lorsque comparé avec l'horaire réel. D'autre part, des 77 quarts travaillés à l'horaire, 59 font partie des quarts de préférence. L'employée 5 a reçu quatre des cinq quarts à temps supplémentaires assignés à l'horaire. Cela est dû au score anormalement bas de celle-ci. Comme discuté ci-dessus, ce score est bas en raison d'un retour au travail effectué à la semaine 15. Un score de bien-être de 1 signifie un maximum de quatre quarts de temps supplémentaires. Avec un ensemble de données plus grand, le score de celle-ci aurait probablement été plus élevé, résultant à une répartition des temps supplémentaires entre les infirmières 7, 19, 109 et 5. Concernant les quotas de travail, malgré un nombre réduit d'affectations possibles, pratiquement tous les quotas sont remplis.

Les fins de semaine et le dernier jeudi à l'horaire sont les seuls jours ayant des quotas qui n'ont pas été atteints, ce qui est normal en raison de la complexité ajoutée aux fins de semaine par la contrainte de l'Éq. (3.3.5). L'infirmière 19 se fait attribuer que des quarts de soir malgré les quarts de jour et de soir comme étant primaire à son contrat, ce qui pourrait être retravaillé par la gestionnaire d'horaire de cette unité. L'horaire est malléable, c'est-à-dire que certaines infirmières peuvent échanger leur affectation de quart après discussion avec le gestionnaire d'horaire de l'unité. Néanmoins, le modèle a généré un horaire qui semble diminuer la charge de travail des infirmières à haut score de bien-être tout en respectant les besoins de l'unité, dans la mesure du possible.

HORAIRE COMPORTANT PLUSIEURS CONGÉS

L'horaire de la semaine 25 et 26, soit du 21 juillet 2019 au 3 août 2019, est analysé. Cet horaire comporte plusieurs particularités. D'abord, à ce moment de l'année, plusieurs infirmières prennent deux semaines de vacances d'été. Ce faisant, les unités se retrouvent avec un nombre anormalement bas d'infirmières disponibles. De plus, l'unité d'infirmières comporte deux infirmières notables. D'abord, l'infirmière 26 effectue plusieurs jours de congé et est absente trois jours dans la semaine. Ensuite, l'infirmière 16 effectue son dernier horaire de travail avant de partir en vacances, mais celle-ci tombe en arrêt de travail immédiatement après. L'évolution du score de bien-être est observée dans le Tableau 4.8 de la Section 4.2. Le Tableau 4.26 montre la grille de travail de l'unité d'infirmières au courant de la semaine 25 et 26. Les données numériques des quarts travaillés et des absences sont aussi disponibles dans cette même figure.

Tableau 4.26 : Données tirées de l'horaire de l'unité d'infirmières du 21 juillet 2019 au 3 août 2019.

Inf.	JOURS													
	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.
105	J*J	J		J					S	S		S+S	S+S	S+N
143									*S			*J		
26	J		J		J									S
59	S		S	S		J			J	J		J		J
16	J	J		N*J	N	N*N			J	J	J			
90														
96	N	N		J		N*N				N	J		N	N
84		J			J	S	J	J*J*N	J			J	J	
47		S	S	S		S	S+S	S+S+N	S		S	S+S		
10														
94														
139												*J		
85									J	J*J				
119		J	J		J	J	J	J+S	J		J*J	J*J		
Matricule	105	143	26	59	16	90	96	84	47	10	94	139	85	119
Bien-être	6.87	2	4.93	2.5	2.74	1	5.12	5.88	4	0	0	0.5	1.5	2.5
Contrat	J	S	J	J	J	J	N	J	S	J	J-S	S	J	J
Régulier	8	0	4	8	8	0	8	8	9	0	0	0	2	8
Temps Supp.	4	2	0	0	2	0	1	2	0	0	0	1	0	0
TSO	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Volante	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ABSENCE	0	0	3	0	1	0	0	0	0	0	0	0	0	0
MALADIE	1	0	0	0	0	0	0	0	0	0	0	0	0	0
ASSURANCE	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CONGÉ	1	0	2	0	1	10	2	0	1	10	5	0	0	0
CHOMAGE	0	0	1	0	0	0	0	0	0	0	0	0	0	0

Comme décrit précédemment, les infirmières 143, 90, 10, 94, 139 et 85 ne sont pas ou très peu présentes dans l'horaire de travail. Bien que certaines effectuent quelques quarts réguliers ou à temps supplémentaires, ceux-ci ont un impact trop minime pour être considérés dans la génération d'horaires. Un total de trois infirmières sont officiellement en congé, mais aucune information sur les autres infirmières absentes n'est disponible. En matière de quarts travaillés, toutes les infirmières ont travaillé de huit à neuf quarts, mise à part l'infirmière 26

qui a travaillé quatre quarts et a été absente pendant trois autres quarts. L'infirmière 16 a aussi été absente une fois dans l'horaire, soit la journée du dernier vendredi pour un quart de jour. Les infirmières actives à l'horaire ont aussi des congés prévus à l'horaire. Un total de sept congés divisés entre chaque infirmière a été pris. Dans cet horaire, il arrive à plusieurs reprises de retrouver des quarts de travail régulier avec des quarts de temps supplémentaires pour une même infirmière sur un même quart, comme le premier dimanche de l'infirmière 105. Ces doubles quarts sont seulement comptabilisés comme un quart de régulier, puisque le modèle de génération d'horaire est limité à l'attribution d'un seul quart de travail par infirmière. En matière de temps supplémentaire, un total de 12 temps supplémentaires ont été considérés. En considérant les doubles quarts, un total de 18 temps supplémentaires sont à l'horaire, ce qui est beaucoup plus élevé que la normale. Considérant les dernières observations de l'horaire, il est compréhensible que le score de bien-être des infirmières actives soit élevé. Deux sous-groupes d'infirmières se forment à l'observation du bien-être, soit les infirmières 105, 26, 96, 84 et 47 qui ont un score de 4 et plus, et les infirmières 56, 16 et 119 qui ont un score de moins de 3. Notons encore que l'infirmière 16 a effectué un arrêt de travail après un congé de deux semaines qui a suivi cet horaire. Le Tableau 4.27 montre les paramètres appliqués au modèle ainsi que le nombre de quarts travaillé dans l'horaire historique.

Avec les paramètres obtenus à la suite de l'analyse des données historiques, l'horaire dans le Tableau 4.28 a été généré à l'aide du modèle. Le score de bien-être a été calculé à partir de l'horaire de la semaine 23, soit du 7 au 20 juillet 2019. La prédiction du score de bien-être a pris 12.150 secondes avec un MSE de 2.7776 et un MAE de 1.3132. Ces performances sont similaires à la prédiction précédente. Le temps de calcul de cet horaire a été de 0.015 secondes. Il y a eu un total de 75 affectations de quart de travail à l'horaire.

Tableau 4.27 : Paramètres utilisés pour les résultats de la figure 4.28 .

Paramètre	Description	Valeur
SC_i	Score de Bien-être	Prédiction du 21 juillet 2019
I^{inf}	Nombre d'infirmière	8
Con_i	Absence planifiée	8
Exp_i	Expertise	0
$Pref_{ijq}$	Préférence	Aléatoire selon le contrat de travail
B^{inf}	Nombre d'affectation min.	6
$Total_{reg}$	paramètre pour l'Éq. (4.3)	63
$Total_{vol}$	paramètre pour l'Éq. (4.3)	0
$Total_{ts}$	paramètre pour l'Éq. (4.3)	12
$Total_{iso}$	paramètre pour l'Éq. (4.3)	0

Tableau 4.28 : Résultat de la génération d'un horaire se basant sur les données historiques de la semaine 25 (21 juillet 2019).

Inf.	JOURS														
	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	
105	J		J	J		J	*N	*N	J	J		J	J	J	
26		*S			J	J	S	S			J		S		
59	J	J		J	J	J			J	J	J		J	J	
16		N	N	J	J				N	N	J	J	J		
96	N			N	N	N					N	N	N	N	
84		J	J	J	S		J	J	J	J		J	J		
47	S		S	S		S			S	S	S	S		S	
105	J	J	J		J	J			J	J	J			J	

Matricule	105	26	59	16	96	84	47	119
Bien-être	2,75	2,8	4,4	3,19	4,73	6,88	3,98	3,77
Contrat	J	J	J	J	N	J	S	J
Régulier	9	7	10	9	8	10	9	10
Temps Supp.	2	1	0	0	0	0	0	0
Préférence	8	4	8	5	8	8	8	8

D'abord, les quotas de travail ne sont pratiquement jamais remplis dans cet horaire. Puisqu'il y a un besoin de neuf infirmières par jour (quatre de jour, trois de soir et deux

de nuit) et que l'unité d'infirmières de cet horaire est composée de huit infirmières, il est normal que l'unité ne puisse pas subvenir à leur besoin en personnel soignant. Seulement quelques quarts de jour remplissent les quotas demandés, ce qui est normal vu le nombre d'infirmières de jour restant dans l'unité. Il n'y a qu'une seule infirmière pour les quarts de soir et une autre pour les quarts de nuit, le reste étant compensé avec l'affectation de quarts secondaire pour les autres infirmières. Malgré le manque d'infirmières, tous les quarts ont au moins une infirmière assignée. Par rapport au score de bien-être, pratiquement toutes les infirmières ont obtenu un score de 3 et plus, ce qui témoigne de la surcharge de travail vécu par l'unité. Comme les infirmières sont en sous-nombre, le score n'influence pas sur le nombre de quarts réguliers affectés (Éq. (3.23)), puisqu'il faut toujours remplir au mieux les quotas de travail de l'unité. Les seules infirmières n'ayant pas travaillé à l'horaire sont celles ayant pris des congés, ce qui réduit proportionnellement le nombre de quarts pouvant être affectés à celle-ci. Sur les 75 quarts devant être à l'horaire, seulement deux temps supplémentaires ont été attribués, soit à l'infirmière 105 et 26, les deux infirmières ayant le plus bas score de bien-être. Les préférences des infirmières sont bien respectées, mais les infirmières 26 et 16 ont reçu plus de quarts secondaires pour remplir les manquements de personnel que le reste des infirmières de l'unité. L'infirmière 26 a été trois fois absente dans l'horaire réel et a travaillé quatre quarts. L'horaire généré offre sept quarts de travail avec un quart en temps supplémentaire. Le score de bien-être calculé et plus haut que celui prédit, ce qui est justifiable du fait que cette infirmière était de retour au travail à la suite de ses deux semaines de vacances d'été. Le modèle peut alors difficilement donner un haut score. Semblablement, l'infirmière 105 revenait aussi de son congé de vacances d'été dans cet horaire. Puisqu'elle n'a aucun surmenage dans les semaines antérieures de vacances, cela explique son bas score de bien-être prédict, en comparaison avec celui calculé. Enfin, l'infirmière 16 ne semble pas être reconnue comme une infirmière à risque. Celle-ci a fait un arrêt de travail après être partie en vacances d'été et n'est pas revenue au travail par la suite. Dans l'horaire généré, elle se voit

attribuer quatre quarts de nuit, faisant partie de ces quarts de travail secondaires. Dans l'horaire historique, celle-ci travaille trois quarts de nuit sur huit quarts et deux temps supplémentaires. En raison de la similarité des deux horaires, il n'est pas concluant que cet horaire puisse avoir un effet positif sur celle-ci, mis à part la considération des quarts de préférence. À la suite de l'analyse de l'unité au courant de cet horaire, il est possible que l'environnement de travail stressant puisse avoir découragé celle-ci à ne pas retourner au travail. Ainsi, l'introduction d'une méthode de génération d'horaire avec prédition de l'absentéisme pourrait avoir un impact à long terme sur la qualité de vie au travail des infirmières en diminuant le stress lié à la surcharge. Un environnement de travail moins stressant peut ainsi inciter le personnel médical, telle l'infirmière 16, à revenir au travail après la période de vacances.

HORAIRE D'INFIRMIÈRE AUXILIAIRE COMPORTANT PLUSIEURS CONGÉS

Comme les circonstances de cet horaire sont pertinentes pour les résultats de cette recherche, l'unité d'infirmières auxiliaires de la semaine 25 et 26 (21 juillet 2019 au 3 août 2019) a été analysée. Une fois de plus, cet horaire comporte plusieurs infirmières en vacances d'été, ce qui a un impact sur le nombre d'infirmières disponibles. Les données sur l'horaire travaillé sont affichées dans le Tableau 4.29.

Tableau 4.29 : Données tirées de l'horaire de l'unité d'infirmières du 21 juillet 2019 au 3 août 2019.

aux.	JOURS														
	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	
7															
41	N*N	N	N		N	N*N				N*J	N	N			N*J
11															
19			I	I*PS	S	S			S	S	S				J
12	J		J	J	I*S				J	J*J				S	S*J
109		J	J	J	J										
5							J*J	JPS	J						
39		J			J	J	S*N	S*S*N	*J						
63		S	S	S			J	J	S	S		J	J		
38								*N							
46															
Matricule	7	41	11	19	12	109	5	39	63	38	46				
Bien-être	4.31	5.00	9.68	6.78	6.96	1.5	1.5	5.23	0.67	1.00	4.14				
Contrat	j	n	j	js	j	j	j	s	j	n	n				
Régulier	0	9	0	8	8	4	3	5	9	0	0				
Temps Supp.	0	2	0	2	3	0	2	4	0	1	0				
TSO	0	2	0	0	0	0	0	0	0	0	0				
Volante	0	0	0	0	0	0	0	0	0	0	0				
ABSENCE	0	0	0	1	1	0	1	2	1	0	0				
MALADIE	0	0	0	0	0	0	0	0	0	0	0				
ASSURANCE	0	0	10	0	0	0	0	0	0	0	0				5
CONGÉ	10	1	0	0	0	5	0	1	0	2	0				
CHOMAGE	0	0	0	1	0	1	0	0	0	0	0				

Sur les 11 infirmières auxiliaires de cette unité, quatre d'entre elles ne sont pas à l'horaire, incluant l'infirmière 38 qui ne travaille qu'une seule fois à l'horaire. Des infirmières absentes à l'horaire, l'infirmière 11 et 46 est en arrêt de travail. Le cas de l'employée 11 a été discuté à la Sous-Section 4.3.3 de ce chapitre et l'infirmière 46 est en assurance maladie depuis le début des données historiques analysées. En analysant l'employée 38, celle-ci travaille par intermittence de la semaine 1 à 35, effectuant de temps à autre du temps supplémentaire.

En matière de congés planifiés, les infirmières 7 et 109 sont en vacances, mais l'employé 109 commence ses vacances après la première semaine d'horaire. Pour les absences, cinq infirmières sur les sept actives ont une absence non planifiée à l'horaire. Cela témoigne de la surcharge vécue par les infirmières de cette unité, appuyé par les scores de bien-être de celle-ci. Quatre infirmières sur sept sont grandement à risque, ayant un score calculé de 5 et plus. Contrairement à ces dernières, les trois autres infirmières affichent un score de bien-être de 1,5 et moins, ce qui est étonnamment bas. Pour justifier ces scores, l'analyse des horaires antérieurs des infirmières 105 et 63 montre qu'elles travaillent habituellement 9 à 10 quarts par horaire avec de rares temps supplémentaires. L'infirmière 5 est revenue de deux semaines de vacances à l'horaire précédent et n'effectue que cinq quarts de travail à cet horaire. Notons que ces scores ne sont pas les scores prédicts par le modèle de prédiction, mais plutôt de la formule du calcul du score de bien-être présenté dans la Sous-Section 3.2.2. Un total de 46 quarts en temps régulier, 14 temps supplémentaires et deux temps supplémentaires obligatoires ont été travaillés à l'horaire. Entre autres, les employées 41 et 39 ont chacune travaillé quatre quarts à temps supplémentaires. Comme dans l'horaire de l'unité d'infirmières, certaines infirmières ont effectué un quart régulier et de temps supplémentaires sur la même plage horaire. Considérant les limitations du modèle de génération d'horaire, ces quarts sont considérés comme des temps réguliers dans les paramètres de comparaison du modèle (Éq. (4.3)). Le Tableau 4.30 montre les paramètres appliqués au modèle ainsi que le nombre de quarts travaillés dans l'horaire historique.

Tableau 4.30 : Paramètres utilisés pour les résultats du Tableau 4.31 .

Paramètre	Description	Valeur
SC_i	Score de bien-être	Prédiction du 21 juillet 2019
I^{aux}	Nombre d'infirmière	7
Con_i	Absence planifiée	9
Exp_i	Expertise	Aucune
$Pref_{ijq}$	Préférence	Aléatoire selon le contrat de travail
B^{aux}	Borne	5
$Total_{reg}$	paramètre pour l'Éq. (4.3)	46
$Total_{vol}$	paramètre pour l'Éq. (4.3)	0
$Total_{ts}$	paramètre pour l'Éq. (4.3)	14
$Total_{tso}$	paramètre pour l'Éq. (4.3)	2

Le modèle de génération d'horaire affecte donc un total de 56 quarts à l'horaire. Avec les paramètres précédemment discutés, l'horaire dans le Tableau 4.31 a été généré. Le score de bien-être a été calculé à partir de l'horaire de la semaine 23, soit du 7 au 20 juillet 2019. La prédiction du score de bien-être et la même qu'à l'horaire précédent, soit 12.150 secondes avec un MSE de 2.7776 et un MAE de 1.3132. Ces performances sont similaires à celles du modèle utilisé précédemment. Le temps de calcul de cet horaire a été de 0.015 secondes. La borne de l'Éq. (3.21) a été paramétrée à 5.

Tableau 4.31 : Résultat de la génération d'un horaire se basant sur les données historiques de la semaine 25 (21 juillet 2019).

aux.	Dim.	JOURS													
		Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	
41		N	N		N	N	N	N	N	N	N		N	N	
19				S	S		S	S						S	
12	J	J							J	J		J		J	
109	*N		J	J*N	J	J									
5		J	J	J	J		J	J	J	J			*N	J	*N
39	S	S	S			S			S*N	S	S	S		S	
63	J		J	J	J	J			J	J		J	J	J	

Matricule	41	19	12	109	5	39	63
Bien-être	4,18	3,99	4,11	2,71	2,62	2,88	3,74
Contrat	N	J-S	J	J	J	S	J
Régulier	9	5	5	4	9	9	10
Temps Supp.	0	0	0	2	2	1	0
Préférence	8	5	5	3	8	8	8

Comme les quotas en infirmières auxiliaires sont moins grands que ceux de l'unité d'infirmières, plus de plages horaires ont atteint leur quota de personnel auxiliaire. Cumulativement, seulement 56 affectations étaient disponibles pour remplir les 84 quarts à cet horaire. Néanmoins, les quarts de nuit remplissent tout leur quota d'une infirmière, même si quatre des quarts de nuit à l'horaire sont en temps supplémentaire. De plus, cinq quarts de jour ont été remplis par rapport au quota, mais le reste des quarts de jour et de soir sont en sous-effectif. Néanmoins, tous les quarts ont au moins une infirmière auxiliaire d'affectée. Par rapport au résultat de la prédition du bien-être pour les infirmières, les résultats obtenus sont plus bas pour les infirmières ayant eu un score calculé élevé et plus haut pour celles ayant eu un score calculé plus bas. Les infirmières 19, 12 et 109 ont reçu le moins d'affectations régulières. L'employé 109 avait pris la deuxième semaine de l'horaire en vacances, mais son bas score en comparaison avec le reste de l'unité lui a valu deux affectations de temps supplémentaire. Les employées 19 et 12 ont deux des plus hauts scores de l'unité, il est donc normal que le modèle

affecte un nombre de quart de travail n'allant pas au-delà des bornes de la contrainte de l'Éq. (4.3). Dans l'horaire réellement travaillé, ces infirmières faisaient partie des plus surmenées. L'infirmière 19, pouvant être affectée au quart de jour et de soir, se fait affecter seulement de soir par manque de personnel à ce quart. L'infirmière 41, qui a le plus haut score et qui est la seule infirmière de nuit, travaille neuf quart réguliers, en raison de son rôle indispensable dans le contexte de cet horaire. Même si celles-ci sont générées aléatoirement, pratiquement toutes les préférences des infirmières ont été remplies.

L'employé 39 est un cas intéressant en raison de score de bien-être prédit de 2.88, alors que son score réel calculé est de 5.23. La cause de cette disparité des résultats provient du retour des vacances de celle-ci. Le manque de données historiques précédent l'horaire analysé semble avoir un impact sur le résultat de la prédiction obtenu. Ce scénario semble avoir un impact sur toutes les infirmières revenant d'une longue période de congés, laissant croire que l'infirmière est reposée après ce moment de répit. Or, il semble difficile pour le modèle de prédire adéquatement le score des infirmières dans ce cas. Ce scénario semble ne pas être reconnaissable par le modèle de prédiction. Un ensemble de données étendu sur une plus grande période avec plusieurs autres cas similaires pourrait permettre au modèle d'apprentissages machine à reconnaître ce scénario. Ces scénarios pourraient aussi être balisés manuellement, permettant à un autre modèle de reconnaître le cas.

Dans cette partie, les horaires réels ont été comparés avec ceux générés à partir du modèle de programmation linéaire en nombres entiers pour générer des horaires optimaux. La comparaison des deux horaires avec les mêmes paramètres et contraintes en se basant sur les données historiques a su démontrer que le modèle est en mesure de générer des horaires diminuant la surcharge des infirmières ayant un mauvais score de bien-être. En utilisant les données d'horaires antérieurs pour générer un score de bien-être pour chaque infirmière, le modèle peut considérer avec une précision satisfaisant les infirmières qui ne doivent pas

être surchargées. Une analyse du modèle déployé en temps réel n'a pas pu être faite, mais les résultats montrent qu'une implémentation à long terme de ce type d'algorithme pourrait diminuer de façons graduelles le score moyen des infirmières.

4.3.4 PERSPECTIVE SANS TEMPS SUPPLÉMENTAIRE

Dans cette Sous-Section, une utilisation alternative du modèle de programmation en nombres entiers pour générer des horaires optimaux pour les infirmières en milieu médical est explorée. Bien qu'il soit préférable que les infirmières puissent subvenir en permanence au besoin en quotas de travail, la réalité des horaires travaillés par celles-ci montre qu'en général, les infirmières font souvent appel aux infirmières externes pour remplir les manquements de personnel à l'horaire. Considérant que les infirmières font beaucoup de temps supplémentaire, cette Sous-Section présente une solution sans affectation de temps supplémentaire, utilisant ainsi de façon active les infirmières externes dans le modèle de création d'horaires.

Une reprise du scénario de la Sous-Section 4.3.2 est utilisée pour tester ce modèle altéré. Les valeurs appliquées au paramètre sont les mêmes que vu dans le Tableau 4.17. L'horaire construit à partir des résultats obtenus dans le Tableau 4.18 montre qu'il y a huit jours sur 14 ayant une infirmière manquant durant l'une des périodes de la journée. De plus, 14 temps supplémentaires ont été attribués aux infirmières de l'unité. Au total, un surplus de 22 quarts dans le quota du travail de la semaine est imposé sur l'unité d'infirmières. Cette dernière affirmation implique que l'unité a un besoin d'environ deux à trois infirmières de plus. En retirant l'affectation de quart en temps supplémentaire du modèle d'optimisation de génération d'horaire de travail, l'horaire dans le Tableau 4.32 a été construit à partir des résultats obtenus. Le modèle d'optimisation a pris un total de 0.14 secondes pour être résolu.

Tableau 4.32 : Résultat de la génération d'horaire de l'unité d'infirmières avec score de bien-être prédit avec la semaine du 8 décembre 2019.

Inf.	JOURS														
	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	Dim.	Lundi	Mardi	Merc.	Jeudi	Ven.	Sam.	
105		J		J	J		J	J	J			J			
143		S	S		S		S	S	S	S		S		S	
26		J	J				J	J		J		J			
90	J	J		J	J				J	J	J		J	J	
96	N			N	N	N				N		N	N	N	
84	J		J		J			J	J		J		J	J	
47		S	S	S		S		S			S		S		
10	J		J	J	J	J		J	J	J		J	J	J	
94	S			S		S					S	S	S	S	
85				J	J	J	J	J			J	J	J		
119	J	J	J		J						J		J	J	

Matricule	105	143	26	90	96	84	47	10	94	85	119
Bien-être	6,09	3,08	3,58	2,1	2,58	4,64	3,63	1,82	3,33	2,83	3,54
Contrat	J	S	J	J	N	J	S	J	J-S	J	J
Régulier	7	9	6	9	8	7	6	10	7	8	7
Temps Supp.	0	0	0	0	0	0	0	0	0	0	0
Préférence	7	9	5	7	7	7	6	8	7	7	5

L'horaire construit avec les résultats obtenus montre un total de 84 affectations de quarts réguliers. Avec un quota total de 126 quarts à remplir dans un horaire et seulement 84 affectations, cela implique un manque de quatre à cinq infirmières à l'horaire. Comme discuté précédemment dans ce chapitre, les manquements de personnel infirmier se font ressentir majoritairement durant les quarts de soir et de nuit. Particulièrement le soir, plusieurs quarts n'ont aucune affectation. Il semble malgré tout difficile de trouver des infirmières prêtes à travailler durant les soirs et nuits. Une responsabilité est ainsi imposée aux infirmières externes. L'ensemble des infirmières externes est généralement composé d'infirmières de la relève et d'équipe volante. Il y a ainsi un risque de surcharge pour ce type d'infirmières également.

CHAPITRE V

CONCLUSION

Dans le cadre de ce projet, une méthode permettant de générer des horaires optimisés pour les infirmières afin de prévenir l'absentéisme a été développée. Cette méthode utilise deux types de modèles, un algorithme de forêt d'arbres aléatoires régressifs (RFR) et un modèle d'optimisation en nombres entiers résolus à l'aide de l'algorithme *Branch-and-bound*. L'algorithme RFR permet la prédiction d'un score de bien-être des infirmières en se basant sur les données antérieures. Dans le premier chapitre, une introduction à la problématique des infirmières a été faite. Il a été conclu que la méthode utilisée pour générer des horaires n'est actuellement pas adaptée à la surcharge vécue par les infirmières. Un modèle d'optimisation *Branch-and-bound* est utilisé afin de générer des horaires optimisés respectant un ensemble de contraintes relatives aux normes de la FIQ et au bien-être des infirmières. Les deux modèles interagissent au moyen d'un paramètre dans le modèle d'optimisation contenant le score prédit pour chaque infirmière. Le résultat de l'optimisation est un horaire de 14 jours contenant les quarts de travail régulier et à temps supplémentaire pour toutes les infirmières. L'horaire résultant du modèle permet, entre autres, de donner plus de quarts de préférence aux infirmières avec un biais par rapport aux infirmières à risque selon la prédiction.

5.1 DISCUSSION DES RÉSULTATS

Un grand nombre de modifications ont été appliquées sur le modèle de prédiction du bien-être et sur l'ensemble de données pour la prédiction afin d'obtenir un bon score de performance, évalué avec le MSE et le MAE. Ces modifications incluent le retrait d'instances non significatives de l'ensemble de prédiction, où beaucoup d'instances avec un haut taux d'assurance salaire étaient problématiques. L'utilisation de la méthode *leave-one-subject-out*

permet de déterminer l'ensemble d'entraînement et de test du modèle, où un seul individu d'un ensemble est retiré de l'entraînement et utilisé pour le test. Cette méthode est couplée avec la rétention d'une infirmière ayant fait un arrêt de travail afin d'obtenir des résultats de performance pertinents. Le choix des variables retirées de l'ensemble a été fait en analysant la matrice des coefficients de corrélation et des *p-values*. Enfin, le paramétrage du modèle a été fait en effectuant une recherche des hyperparamètres, soit les paramètres du modèle de prédiction exerçant une influence sur l'algorithme d'apprentissage. Les résultats finaux du modèle sont un MSE de 2.5 et un MAE de 1.1696, signifiant que la différence absolue entre les valeurs réelles et prédites se trouve autour de 1. Considérant que le score de bien-être a un intervalle de valeur possible entre 0 et 10, les résultats des prédictions sont concluants. Les prédictions effectuées en regardant l'évolution des scores prédits et réels d'une infirmière montrent que le modèle est en mesure de prédire la tendance du score de l'infirmière à un degré satisfaisant. Le modèle est aussi en mesure de prédire les absences prolongées des infirmières, mais un plus grand nombre de tests sur des infirmières ayant fait un arrêt de travail est nécessaire.

Les résultats du modèle d'optimisation pour la génération d'horaire ont été aussi concluants. Entre autres, l'analyse des horaires réellement travaillés a montré que les infirmières et les infirmières auxiliaires sont toujours en manque de personnel pour avoir un horaire respectant les quotas. Dans les pires cas observés, les ressources disponibles afin de générer un horaire permettent tout juste à remplir la moitié des quotas de travail pour plusieurs journées. Dans les cas avec un plus grand nombre d'infirmières disponibles, les samedis et dimanches de l'horaire sont toujours des journées problématiques. Les unités analysées ne contiennent pas assez d'infirmières pour satisfaire les restrictions de la FIQ concernant les fins de semaine. De plus, le nombre d'infirmières de jour disponibles est souvent assez élevé pour satisfaire les quotas de travail durant cette période. Le manque de personnel infirmier est

le plus souvent observé durant les périodes de soir et de nuit. Il est fortement probable que beaucoup de quarts ne remplissant pas leurs quotas font appel à des infirmières externes, mais cette pratique n'est qu'une solution à court terme. La surutilisation des infirmières externes risque de surcharger celle-ci, étant déjà des infirmières occupant un poste précaire.

L'objectif du modèle de génération d'horaires est de réduire les risques d'absentéisme des infirmières. Comme ce modèle n'a pas été mis en pratique, les impacts du modèle sont observés en comparant les horaires résultant du modèle aux horaires réellement travaillés. Les paramètres des horaires réels ont été répliqués pour le modèle d'optimisation afin d'obtenir un maximum de similarités. Les résultats obtenus sont relativement toujours meilleurs concernant la répartition des affectations et la réduction de la charge de travail pour les infirmières. Entre autres, les quarts de temps supplémentaire sont moins nombreux dans les horaires générés, mais généralement durs à éviter en raison du manque de personnel. Malgré une génération aléatoire des préférences de travail des infirmières, celles-ci reçoivent la majorité de leurs préférences pour tous les résultats d'horaires générés. Les seuls cas où une infirmière obtient peu de préférences sont lorsqu'un de ses quarts en priorité secondaire est fortement en sous-effectif. Les horaires pour les infirmières démontrant un haut risque sont concluants. Afin de solidifier cette analyse, il serait préférable d'effectuer de plus amples tests sur d'autres types d'unité d'infirmière. De plus, une mise en pratique serait idéale afin de confirmer les observations des résultats d'horaire générés obtenus.

5.2 LIMITATION DU MODÈLE ET AMÉLIORATIONS FUTURES

Le modèle de prédiction du score de bien-être a permis d'obtenir des résultats concluants. Or, ce modèle a été construit avec un ensemble de données considéré petit pour le domaine de l'intelligence artificielle. Bien que la forêt d'arbres aléatoire soit un algorithme reconnu pour sa performance avec de petits ensembles de données, il serait pertinent dans de futurs

travaux d'expérimenter avec de plus grands ensembles de données, permettant par le fait même l'utilisation d'autres algorithmes d'intelligence artificielle. L'ensemble de données pour la prédiction du score de bien-être pourrait aussi être bonifié de plusieurs autres paramètres. Par exemple, les temps de travail des quarts pourraient être utilisés ou de nouvelles catégories pourraient être ajoutées. Si les données le permettent, il serait intéressant de créer un profil pour chaque infirmière. Cette pratique permettrait au modèle de prédire avec une plus grande certitude les variations de score. En rapport à la variable de bien-être prédite, celle-ci pourrait être peaufinée davantage afin d'obtenir des valeurs plus représentatives de l'état des infirmières. Plus de tests sont nécessaires pour le modèle d'optimisation. Ces tests pourraient être faits sur une unité comportant un plus grand ensemble d'infirmières ou être mis en pratique afin d'observer les conséquences à long terme de son utilisation sur l'évolution du score des infirmières.

BIBLIOGRAPHIE

American Psychological Association (2021, May). *APA dictionary of psychology*. Repéré à <https://dictionary.apa.org/burnout>

AMPL Optimization Inc. (2021). *Home - AMPL*. Repéré le 2021-10-26, à <https://ampl.com/>

Bae, S.-H. & Fabry, D. (2014). Assessing the relationships between nurse work hours/overtime and nurse and patient outcomes : Systematic literature review. *Nursing Outlook*, 62(2), 138 – 156. doi: 10.1016/j.outlook.2013.10.009

Bard, J. F. & Purnomo, H. W. (2005). Preference scheduling for nurses using column generation. *European Journal of Operational Research*, 164(2), 510 – 534. doi: 10.1016/j.ejor.2003.06.046

Berger, A. & Hobbs, B. (2006). Impact of shift work on the health and safety of nurses and patients. *Clinical Journal of Oncology Nursing*, 10(4), 465 – 499.

Berkelaar, M., Eikland, K. & Notebaert, P. (2004, May 1). *lp_solve 5.5, Open source (Mixed-Integer) Linear Programming system*. Software. Available at <<http://lp-solve.sourceforge.net/5.5/>>. Last accessed Dec, 18 2009, Repéré à http://lp_solve.sourceforge.net/5.5/

Brazeau, M. (2013, Avril). *Affectation des infirmières aux salles de l'unité d'endoscopie digestive du Centre hospitalier universitaire de Sherbrooke*. Repéré à <https://publications.polymtl.ca/1086/>

Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5–32. doi: 10.1023/A:1010933404324

Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. (1984). *Classification and regression trees*. Monterey, CA: Wadsworth Brooks/Cole Advanced Books and Software.

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B.

& Varoquaux, G. (2013). API design for machine learning software : Experiences from the scikit-learn project. Dans *ECML PKDD workshop : Languages for data mining and machine learning*, pp. 108–122.

Centre canadien d'hygiène et de sécurité au travail (2021). *Santé mentale – Épuisement professionnel : Réponses SST*. Repéré à https://www.cchst.ca/oshanswers/psychosocial/mentalhealth_jobburnout.html

Claesen, M. & Moor, B. D. (2015). Hyperparameter Search in Machine Learning. *CoRR*, *abs/1502.02127*. Repéré à <http://arxiv.org/abs/1502.02127>

Cook, S. A. (1971). The Complexity of Theorem-Proving Procedures. Dans *Proceedings of the Third Annual ACM Symposium on Theory of Computing*, STOC '71, p. 151–158., New York, NY, USA. Association for Computing Machinery. doi: 10.1145/800157.805047

Dogruyol, K. & Sekeroglu, B. (2020). Absenteeism prediction : A comparative study using machine learning models. Dans R. A. Aliev, J. Kacprzyk, W. Pedrycz, M. Jamshidi, M. B. Babanli, & F. M. Sadikoglu (Éds.). *10th International conference on theory and application of soft computing, computing with words and perceptions - ICSCCW-2019*, pp. 728–734., Cham. Springer International Publishing.

Dorigo, M. & Blum, C. (2005). Ant colony optimization theory : A survey. *Theoretical computer science*, 344(2-3), 243–278.

Dunning, I., Huchette, J. & Lubin, M. (2017). JuMP : A Modeling Language for Mathematical Optimization. *SIAM Review*, 59(2), 295–320. doi: 10.1137/15M1020575

FICO (2021). *FICO® Xpress Solver*. Repéré le 2021-10-26, à <https://www.fico.com/en/products/fico-xpress-solver>

Gagné, M.-A. (2017, Décembre). *Étude transversale descriptive de l'expérience au travail des infirmières québécoises*. Repéré à <http://hdl.handle.net/1866/20083>

Gandomi, A. H., Yang, X.-S., Talatahari, S. & Alavi, A. H. (2013). Metaheuristic algorithms in modeling and optimization. *Metaheuristic applications in structures and infrastructures*, pp. 1–24.

Gurobi Optimization, LLC (2022). *Gurobi Optimizer Reference Manual*. Repéré à <https://www.gurobi.com>

Ho, T. K. (1995). Random decision forests. Dans *Proceedings of 3rd international conference on document analysis and recognition*, Vol. 1, pp. 278–282 vol.1. doi: 10.1109/ICDAR.1995.598994

Hsia, F. Y. (2009). Optimal weight design for nurse scheduling problem. *The Journal of Health Science*, (11), 148 – 158.

Huang, Y.-C., Hsieh, Y.-H. & Hsia, F.-Y. (2016). A study on nurse day-off scheduling under the consideration of binary preference. *Journal of Industrial Production Engineering*, 33(6), 363 – 372.

Jennings, B. M. (2008). Work stress and burnout among nurses : Role of the work environment and working conditions. *Patient safety and quality : An evidence-based handbook for nurses*.

Kane, R., Shamliyan, T., Mueller, C., Duval, S. & Wilt, T. (2007). Nurse staffing and quality of patient care. *Evidence report/technology assessment*, (151), 1–115.

Land, A. H. & Doig, A. G. (1960). An automatic method of solving discrete programming problems. *Econometrica*, 28(3), 497–520. Repéré à <http://www.jstor.org/stable/1910129>

Legrain, A., Bouarab, H. & Lahrichi, N. (2015). The nurse scheduling problem in real-life. *Journal of medical systems*, 39(1), 160. doi: 10.1007/s10916-014-0160-8

Makhorin, A. (2012). *GLPK (GNU Linear Programming Kit)*. Repéré à <https://www.gnu.org/software/glpk/glpk.html>

Mitchell, S., Consulting, S. M. & Dunning, I. (2011). *PuLP : A Linear Programming Toolkit for Python*.

Mobasher, A., Lim, G., Bard, J. F. & Jordan, V. (2011). Daily scheduling of nurses in

operating suites. *IIE Transactions on Healthcare Systems Engineering*, 1(4), 232–246. doi: 10.1080/19488300.2011.631097

Mohan, S. (2008). Scheduling part-time personnel with availability restrictions and preferences to maximize employee satisfaction. *Mathematical and Computer Modelling*, 48(11), 1806 – 1813. doi: 10.1016/j.mcm.2007.12.027

Nolan, M., Nolan, J. & Grant, G. (1995). Education and debate. Maintaining nurses' job satisfaction and morale. *British Journal of Nursing*, 4(19), 1149 – 1154.

Quinlan, J. R. (1986). Induction of decision trees. *MACH. LEARN*, 1, 81–106.

Quinlan, J. R. (1993). *C4.5 : Programs for machine learning*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Rogers, A. E., Hwang, W.-T., Scott, L. D., Aiken, L. H. & Dinges, D. F. (2004). The working hours of hospital staff nurses and patient safety. *Health affairs*, 23(4), 202–212.

Schrijver, A. (1998). *Theory of linear and integer programming*. John Wiley & Sons.

Shuib, A. & Kamarudin, F. I. (2019). Solving shift scheduling problem with days-off preference for power station workers using binary integer goal programming model. *Annals of Operations Research*, 272(1/2), 355 – 372.

Silberholz, J. & Golden, B. (2010). Comparison of metaheuristics. Dans *Handbook of metaheuristics* pp. 625–640. Springer.

Stelnicki, A. M. & Carleton, R. N. (2021). Mental disorder symptoms among nurses in Canada. *Canadian Journal of Nursing Research*, 53(3), 264–276.

Tewari, K., Vandita, S. & Jain, S. (2020). Predictive analysis of absenteeism in mnics using machine learning algorithm. Dans P. K. Singh, A. K. Kar, Y. Singh, M. H. Kolekar, & S. Tanwar (Éds.). *Proceedings of ICRIC 2019*, pp. 3–14., Cham, Suisse. Springer International Publishing.

Trilling, L., Guinet, A. & Magny, D. L. (2006). Nurse scheduling using integer linear programming and constraint programming. *IFAC Proceedings Volumes*, 39(3), 671 – 676. 12th IFAC Symposium on Information Control Problems in Manufacturing, doi: 10.3182/20060517-3-FR-2903.00340

Vagharseyyedin, S. A., Vanaki, Z. & Mohammadi, E. (2011). The nature nursing quality of work life : An integrative review of literature. *Western Journal of Nursing Research*, 33(6), 786–804. PMID : 20719995, doi: 10.1177/0193945910378855

Widyastiti, M., Aman, A. & Bakhtiar, T. (2016). Nurses scheduling by considering the qualification using integer linear programming. *Telkomnika*, 14(3), 933 – 940.

Xiang, W. (2017). A multi-objective ACO for operating room scheduling optimization. *Natural Computing*, 16(4), 607 – 617.

Xiang, W., Yin, J. & Lim, G. (2015). An ant colony optimization approach for solving an operating room surgery scheduling problem. *Computers Industrial Engineering*, 85, 335 – 345. doi: 10.1016/j.cie.2015.04.010

Zhang, W. (1996). *Branch-and-Bound search algorithms and their computational complexity*. University of Southern California, Information Sciences Institute, Marina Del Rey, CA.

ANNEXE A
CERTIFICAT ÉTHIQUE

Le 01 octobre 2021

RENOUVELLEMENT DE L'APPROBATION ÉTHIQUE

La présente atteste que le projet de recherche décrit ci-dessous a fait l'objet d'un renouvellement de l'approbation éthique émise par le CER-UQAC et qu'il satisfait aux exigences de la politique de l'Université du Québec à Chicoutimi en matière d'éthique de la recherche avec des êtres humains.

Projet # : 2021-562

Titre du projet de recherche: Développement d'une méthode d'optimisation pour la génération des horaires du personnel de soins afin de minimiser l'absentéisme et les arrêts de travail

Chercheur principal à l'UQAC

Maud-Christine Chouinard
département des sciences de la santé, UQAC

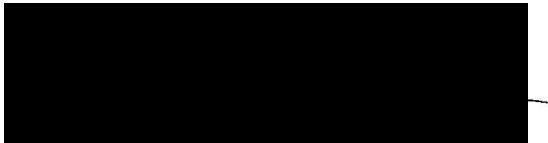
Financement : MITACS, FUQAC

Date de l'approbation éthique initiale du projet : 08 juillet 2020

Date du prochain renouvellement : 08 juillet 2022.

N.B. Un rappel automatique vous sera envoyé par courriel quelques semaines avant l'échéance de votre certificat afin de remplir le formulaire F7 - Renouvellement annuel.

- Si votre projet se termine avant la date du prochain renouvellement, vous devrez remplir le formulaire **F9 - Fin de projet**.
 - Si des modifications sont apportées à votre projet avant l'échéance du certificat, vous devrez remplir le formulaire **F8 - Modification de projet**.
 - Tout nouveau membre de votre équipe de recherche devra être déclaré au CER-UQAC lors de votre prochaine demande de renouvellement ou lors de la fin de votre projet si le renouvellement n'est pas requis. ATTENTION: Vous devez faire signer une déclaration d'honneur aux personnes ayant accès aux participants (ou à des données nominatives sur les participants) et la conserver dans vos dossiers de recherche.
 - Si vous avez des cochercheurs dans d'autres universités, veuillez leur transmettre ce certificat.
 - Si votre projet est financé, le Décanat de la recherche et de la création sera mis en copie conforme afin de l'informer du renouvellement de votre certification éthique.
-



Stéphane Allaire