

Prédire la circulation à l'aide de l'apprentissage machine

Olivier Audet ing., Chef de projet, Ville de Montréal
Patrice O'Carroll, Scientifique des données-IoT, Mnubo

Communication préparée pour présenter à la séance:
Intelligence artificielle et apprentissage machine à l'appui de la mobilité intelligente

Congrès-Exposition conjointe ATC-STI Canada 2019
Halifax (Nouvelle-Écosse)

Résumé

La Ville de Montréal (la Ville) possède et gère plus de 2 500 feux de circulation sur son territoire. Plusieurs systèmes simples et complexes sont actuellement en place actuellement afin de gérer dynamiquement les feux de circulation, par exemple des systèmes adaptatifs. Évidemment, les systèmes complexes sont performants, mais également plus dispendieux à l'achat, au déploiement et en frais de maintenance. Ils nécessitent habituellement de déployer des équipements de détection à toutes les approches d'une intersection à contrôler.

Notre vision est de créer un système simplifié, moins coûteux et flexible. Ce système sera peut-être légèrement moins performant au départ, mais offrira plus de potentiel d'amélioration. L'approche de la Ville est d'utiliser une technique de gestion des feux de circulation existante, en l'occurrence le TRP (Traffic Responsive Plan), bien maîtrisée et d'y ajouter un volet de prédiction. En utilisant la prédiction des débits de circulation, on prévoit que les feux de circulation soient proactifs sur le trafic plutôt qu'en réaction à ce dernier. La technique de prédiction présentement utilisée est l'apprentissage machine et s'alimente à partir d'un flux de données temps-réel Internet des Objets (IdO) issu de plusieurs détecteurs véhiculaires. La technique d'apprentissage machine utilisée est eXtreme Gradient Boosting, XGBoost, (Chen & Guestrin, 2016) et sa performance est évaluée à l'aide de l'erreur moyenne absolue (mean absolute error, MAE).

Nous présenterons les différentes phases du projet, l'étape à laquelle nous sommes rendus, ainsi que les différentes équipes impliquées à chaque étape. Sera aussi démontré le potentiel d'amélioration qu'apporte l'apprentissage machine en comparaison des autres techniques de gestion des feux de circulation.

Gestion des feux de circulation à Montréal

De ses débuts par la gestion via des polices du trafic au premier feu de circulation au gaz en 1868 et ensuite par l'invention du feu de circulation électrique en 1912 pour finalement évoluer vers la gestion automatisée dès les années 1960, la gestion du trafic a continuellement évolué pour assurer la sécurité aux intersections des villes (« Traffic light », 2019). Évidemment, la Ville a fait évoluer au fil des années ses systèmes de gestion de feux de circulation. Ci-bas est faite une description sommaire des différents modes de gestion de feux de circulation utilisés à la Ville.

Mode libre, isolé ou non synchronisé

Ce mode de gestion de feux de circulation, qu'il fonctionne à temps fixe ou en mode libre consiste en une séquence de phases où des mouvements spécifiques sont permis tantôt dans un sens tantôt dans l'autre. Cette séquence de phases est appelée un cycle puisque la même séquence se répète. À temps fixe, la longueur des intervalles est prédéterminée, c'est-à-dire que la durée de chacun des affichages demeure constante de cycle en cycle. Un ingénieur en circulation détermine la séquence des phases, appelée phasage, en fonction des mouvements de circulation à un carrefour et des volumes de véhicules comptés par mouvement. L'ingénieur

calcule la durée de chacun des intervalles la plupart du temps à l'aide de logiciels d'optimisation et de simulation.

Le mode libre utilise des valeurs minimales et maximales de temps de vert déterminées par l'ingénieur pour chacun des intervalles. Les détecteurs de véhicules permettent de faire apparaître une phase ou non et d'en prolonger la durée à l'intérieur de la fourchette de seuils minimum et maximum prédéfinis. Aucun plan de synchronisation n'est défini puisque la longueur des cycles varie en fonction de l'arrivée aléatoire des véhicules. Ce type de mode de gestion est habituellement utilisé pour des intersections isolées, c'est-à-dire à plus de 200m d'une autre intersection ou même au coeur d'un réseau maillé, par exemple entre deux rues locales pour maintenir des cycles courts et efficaces à un endroit donné et une desserte plus adéquate

Synchronisation en mode Time of Day (TOD)

La synchronisation des feux de circulation consiste à faire fonctionner deux feux de circulation ou plus de manière synchronisée l'un par rapport à l'autre. Le mode de synchronisation "TOD" est un mode de gestion rudimentaire basé sur l'heure du jour et sur le jour de la semaine. Il est utilisé pour synchroniser les différents feux de circulation d'un même réseau (20 à 30 intersections) entre eux. Cette synchronisation des feux de circulation d'un même réseau permet de créer des "bandes vertes" directionnelles sur les artères, c'est-à-dire de permettre à un peloton de véhicules de se déplacer à la vitesse prescrite dans une direction à travers plusieurs feux de circulation sans s'arrêter.

Ce mode de synchronisation des feux de circulation utilise une horloge, via un GPS ou un serveur central, pour gérer l'affichage des feux. Des plans de synchronisation sont définis au préalable pour gérer le cycle d'affichage des feux. Ceux-ci comprennent chacun:

- une longueur de cycle prédéfinie, c'est-à-dire le temps pour compléter une séquence de phases complète
- un décalage spécifique, soit le temps de référence du début d'un cycle par rapport au début du cycle d'une autre intersection ou par rapport à un autre référentiel
- une répartition de temps entre les différentes phases de la séquence

Les trois plans de synchronisation définis par la Ville le sont pour les périodes "heure de pointe A.M.", "heure de pointe P.M." et "hors pointe". Les références pour créer les différents plans de synchronisation sont les comptages véhiculaires (volumes) effectués sur le réseau. Puisque ces volumes sont mesurés pour la mise en place du mode TOD, les feux de circulation sont performants au début de leur utilisation, mais leur performance se dégrade avec le temps au fil de l'évolution des débits de circulation. Il est de pratique courante dans le domaine de mettre à jour les plans de feux aux cinq ans environ pour maintenir une performance acceptable. L'utilisation de détecteurs de présence est optionnelle; elle permet d'escamoter des phases du cycle lorsqu'on ne détecte aucun véhicule en présence et puisque la longueur du cycle est fixe, de donner ce temps à une autre phase plus sollicitée, ce qui améliore l'expérience des utilisateurs (moins d'attente).

Synchronisation en mode Traffic Responsive Plan (TRP)

À l'instar du mode TOD, ce mode de gestion des feux de circulation utilise aussi des plans de synchronisation prédéfinis, mais utilise cette fois le volume et le taux d'occupation des voies (auquel on fera référence comme étant l'occupation) en temps réel à certains endroits spécifiques d'un réseau de feux de circulation pour choisir le plan de synchronisation à implanter plutôt que l'heure de la journée et le jour de la semaine. Le volume et l'occupation sont mesurés à l'aide de détecteurs véhiculaires. Les plans de synchronisation sont définis pour des seuils de volume et d'occupation et non plus pour des périodes temporelles telle "l'heure de pointe A.M.", "l'heure de pointe P.M.", la période "hors pointe". Par exemple, lors d'un jour férié le système pourra utiliser un plan de synchronisation "hors pointe" alors que nous sommes en heure de pointe habituellement à ce moment de la journée.

Les références pour créer les différents plans de synchronisation demeurent les comptages véhiculaires (volumes) effectués sur le réseau ou des prévisions de débits. Le système demeurera performant tant et aussi longtemps que le système peut choisir dans sa banque de plans un plan adapté aux seuils de volume et d'occupation mesurés en temps réel à certains endroits spécifiques ; un sérieux avantage par rapport au mode TOD. De surcroît, l'utilisation de détecteurs de présence pour ajuster le temps de phase au niveau de l'intersection même est optionnel, mais améliore davantage l'expérience des utilisateurs (moins d'attente). Le coût de ce système est d'environ 100 000\$ plus élevé que les modes libre et TOD puisqu'il nécessite au minimum deux détecteurs en tout temps pour fonctionner au coût d'environ 50 000\$ l'unité par 25 intersections. À long terme (après cinq ans de mise en place), le mode TRP est environ 5% plus efficace que le mode TOD puisqu'il ne se dégrade pas avec le temps. L'efficacité est mesurée à l'aide des retards aux intersections, à toutes les approches, et des retards totaux sur l'ensemble du réseau de feux de circulation.

Synchronisation en mode Adaptive Traffic Control System (ATCS) ou mode adaptatif

Ce mode de gestion des feux de circulation utilise le volume, préférablement par mouvement, et l'occupation en temps réel à toutes les approches d'une intersection pour calculer de cycle en cycle la longueur des intervalles de feux verts et donc la longueur des cycles. Le volume et l'occupation sont mesurés à l'aide de détecteurs véhiculaires. Aucun plan de synchronisation n'est défini au préalable contrairement aux modes TOD et TRP pour gérer les feux de circulation. C'est plutôt un système expert, à chaque intersection ou centralisé, qui coordonne les différents feux de circulation d'un réseau de feux de circulation adaptatif selon des algorithmes spécifiques. Le mode adaptatif gère les véhicules en peloton et peut ainsi optimiser leur passage à toutes les approches des intersections d'un réseau. Ce système de gestion est performant, mais nécessite que les détecteurs de toutes les approches des intersections du réseau soient fonctionnels pour un fonctionnement optimal. Ceci implique beaucoup de maintenance et une supervision rigoureuse du système. Ce type de mode de gestion est utilisé pour coordonner chacun des différents feux de circulation d'un même réseau. Un système a été implanté à la Ville avec trois intersections. Comme mentionné précédemment ce n'est pas le

concept de “bandes vertes” qui est appliqué, mais plutôt la gestion de peloton de véhicules pour optimiser leur temps de passage. Le coût de ce système est d’environ 750 000\$ plus élevé que les modes libre et TOD puisqu’il nécessite des détecteurs (25 000\$ par intersections) à toutes les approches des intersections du réseau, pour un réseau de 25 intersections. Il nécessite également 25 systèmes adaptatifs, un à chaque intersection (5 000\$ par système). À long terme (après cinq ans de mise en place), le mode adaptatif est environ 8% plus efficace que le mode TOD puisqu’il ne se dégrade pas avec le temps. Il est également environ 3% plus efficace que le mode TRP. L’efficacité est mesurée à l’aide des retards aux intersections, de toutes les approches, et du retard total sur le réseau de feux de circulation.

Comparaison des méthodes

Avant de débiter, voici quelques clarifications sur les prémisses de base:

- Pour l’évaluation des coûts, nous utilisons comme références de base le “mode libre” et le mode TOD comme dépense minimale pour contrôler une intersection. Bien que ce prix de base puisse varier si un ou plusieurs détecteurs de présence sont installés à l’intersection, nous postulons pour nos analyses que le coût de ces détecteurs de présence est négligeable puisque la majorité des intersections de la Ville n’ont pas de détecteur de présence.
- Le nombre d’intersections considérées pour les comparaisons est de 25.
- L’efficacité est mesurée en fonction des retards aux intersections, de toutes les approches, et du retard total du réseau de feux de circulation.
- Pour l’évaluation de l’efficacité nous utilisons comme référence de base le mode TOD puisque ça ne s’applique pas au mode libre, car ce sont des intersections isolées.

Le tableau 1 résume les comparaisons des différents modes de gestion des feux de circulation.

Mode de Gestion	Coût	Utilisation	Efficacité	Complexité
Libre	Réf. de base	Intersection isolée	N/A	Simple
TOD	Réf. de base	Réseau de 25 intersections	Réf. de base	Horloge requise
TRP	+100 000\$	Réseau de 25 intersections	+5%	- 2 détecteurs - 1 système TRP
Adaptatif	+750 000\$	Réseau de 25 intersections	+8%	- 25 postes de détection - 25 systèmes adaptatifs

Tableau 1: Comparaison des différents modes de gestion de feux de circulation à la Ville de Montréal

Autres constats:

- Le système adaptatif est très coûteux, 650 000\$ de plus, pour seulement 3% de gain supplémentaire par rapport au mode TRP.
- Puisque le système adaptatif nécessite d'équiper de détecteurs et de systèmes adaptatifs les 25 intersections du réseau, il est plus difficile à maintenir fonctionnel et nécessite plus de personnel pour l'entretien.
- Le transit routier le plus important est majoritairement linéaire (nord/sud, est/ouest) à la Ville et est très bien géré par le mode TRP.

Pour toutes ces raisons, la Ville a décidé d'investiguer la possibilité d'améliorer les performances du mode TRP que l'équipe d'ingénierie de circulation maîtrise très bien.

TRP + Prédiction = Traffic Predictive Plan (TPP)

Après avoir fait ces constats sur les différents modes de gestion des feux de circulation, nous avons regardé plus en détail le mode TRP dans le but de trouver des points à améliorer. Voici les faits qui ressortent du mode TRP:

1. Le réseau de feux de circulation prend environ 15 minutes à se stabiliser après le changement du plan de feux.
2. Les changements sont engendrés par un changement de la combinaison volume/occupation sur le réseau de feux de circulation.
3. Il est très bien maîtrisé par notre équipe d'ingénierie de circulation.

Par la suite, nous avons émis des hypothèses à utiliser la prédiction de la circulation pour améliorer le TRP. Voici nos pistes de solutions en utilisant la prédiction:

1. La prédiction pourra prédire aux moins 15 minutes à l'avance le volume et l'occupation.
2. Ces volumes et occupations pourront être validés par notre équipe d'ingénierie de circulation en comparant avec les données historiques.
3. La prédiction de circulation permettra au mode TRP d'être proactif sur les flux de trafic à venir et non de réagir à ces derniers, tel que fait normalement le mode TRP. Son nom le dit bien "Responsive", donc réactif au flux de trafic.

Fort de ces hypothèses positives nous avons débuté un projet pilote pour valider le concept de Traffic Predictive Plan (TPP) qui serait le mode TRP avec comme intrant un volume et une occupation prédits de 15 minutes toutes les minutes.

Revue de littérature des méthodes de prédictions en circulation

Différentes villes travaillent à prévoir la circulation urbaine pour adresser les problèmes de congestion. Différentes études utilisent les méthodes d'analyse de données statistiques et d'apprentissage machine pour prévoir les séries temporelles de la circulation. L'objectif des prédictions est d'aider les utilisateurs à considérer les chemins les plus rapides vers leurs destinations, d'aider les systèmes de gestion des feux de circulation à prendre de meilleures décisions ou d'aider les villes à faire une meilleure planification urbaine.

En 2012, à Los Angeles, les incidents ont été utilisés pour améliorer les performances de prédictions à long terme et à court terme de la circulation sur le réseau routier des grandes artères de la ville (Pan, Demiryurek, & Shahabi, 2012).

Une autre approche utilisée à Hangzhou, en Chine, est de modéliser la congestion dans le but de faire de meilleurs choix de planification urbaine. Une approche basée sur les données GPS de taxi est proposée pour mesurer la capacité de chacun des tronçons de route (Castro, Zhang, & Li, 2012).

À Lyon, les prédictions sont utilisées pour améliorer la suggestion de trajet ainsi que d'améliorer la prise de décisions de leur système de control adaptatif (Wynter et al., 2015). Les auteurs utilisent la méthode de prédiction spatio-temporelle qui utilise de l'information à plusieurs endroits sur le réseau (Min & Wynter, 2011). Ils font face aux difficultés des données urbaines qui sont plus variables que les données de réseaux d'autoroutes. Dans ce contexte, ils proposent une méthode pour déterminer un seuil d'occupation et de volume qui correspond à la congestion. Cette méthode permet de nettoyer le jeu de données, d'entraîner et évaluer des modèles de prédictions sur un jeu de données plus prévisible.

À Hong Kong, les approches statistiques *AutoRegressive Integrated Moving Average* (ARIMA) et *Periodical Moving Average* (PMA) ont été combinées à un réseau de neurones pour éviter les limitations des deux approches statistiques. Les méthodes statistiques permettent de bien représenter les données historiques et le réseau de neurones permet de pondérer les prédictions par ces méthodes avec les valeurs de la circulation en temps réel. Les auteurs concluent que leur méthode d'apprentissage machine permet d'améliorer la performance des méthodes statistiques aux frontières des heures de pointe (Xie & Choi, 2017).

Dans cet article, nous montrons les résultats de prédiction d'un modèle d'apprentissage automatique multivarié. Notre approche par voie de circulation donne une information granulaire aux systèmes de prise de décision en plus de pouvoir être indépendante des autres informations des autres tronçons.

Méthodologie

La tâche des algorithmes est de faire des prédictions de la somme des véhicules et du taux d'occupation moyen de chaque zone(voie) d'un tronçon pendant les 15 prochaines minutes et ce, à une fréquence de prédiction à la minute. Pour ce faire, un algorithme a été entraîné pour chaque détecteur sélectionné dans le cadre de l'étude.

La figure 1 montre une vue détaillée de l'étape de transformation de la donnée.

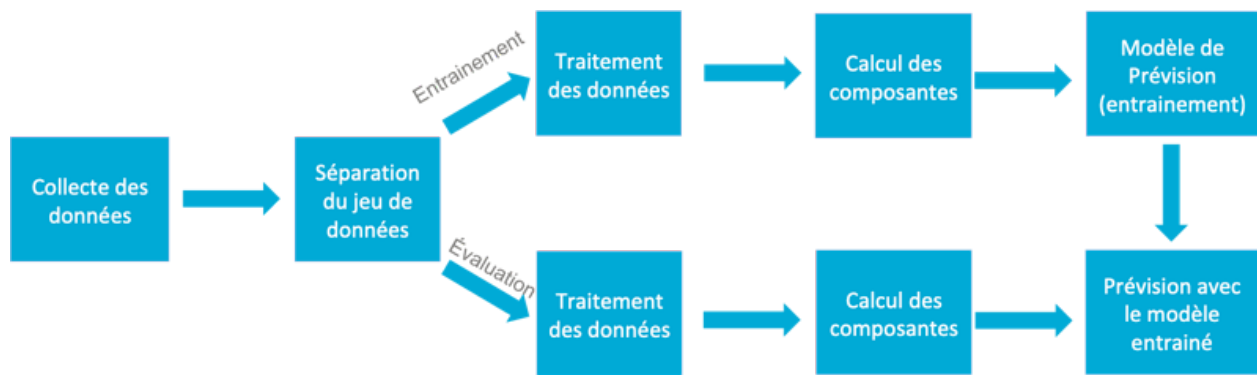


Figure 1: Schéma du flux de travail entre les données brutes et la prédiction du volume et du taux d'occupation de la circulation.

Description du jeu de données

Les données brutes sont collectées à partir du flux MQTT de la Ville (Ville de Montréal, 2019). Voici la liste de mesures utilisées:

- volume (*vehicle-count*) : Le nombre de véhicules de cette classe et zone (qui ont quitté la zone) dans cet intervalle.
- occupation (*occupancy*) : Pourcentage de l'intervalle d'intégration occupé par un véhicule.

Les données sont séparées par zone et par tronçon. Un détecteur a une vision d'un tronçon, qui est défini par un ensemble de zones dans lesquelles les véhicules circulent dans la même direction.

Les événements contenant ces valeurs sont envoyés en moyenne aux 20 secondes pour chaque zone de chaque tronçon. La figure 2 montre une représentation d'un tronçon où l'on observe deux détecteurs thermiques (det-00757-02 et le det-00757-01) ainsi que leurs zones correspondantes. Pour chaque zone, une mesure de volume et d'occupation est utilisée par l'algorithme.

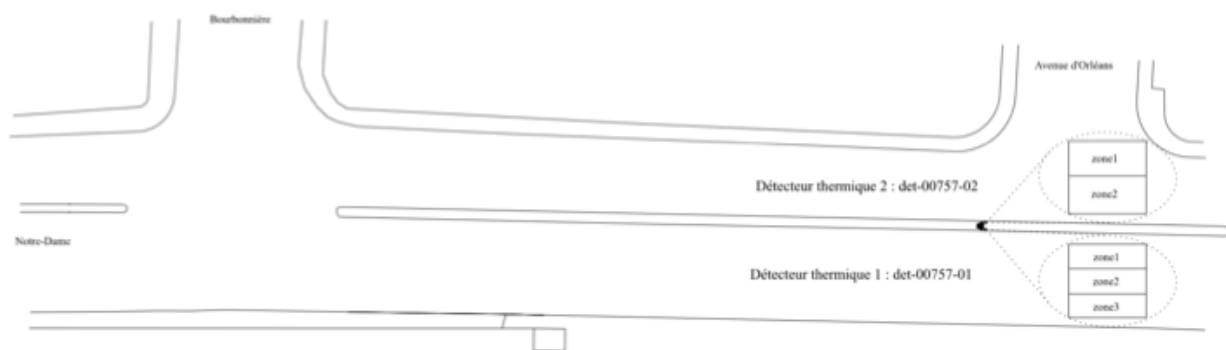


Figure 2: Exemple de la prise de vue de deux détecteurs thermiques.

Le tableau 2 montre les différentes périodes de collecte de données utilisées pour l'entraînement et la validation des algorithmes. Les périodes sont de 3 mois pour avoir

suffisamment d'exemples pour permettre l'apprentissage des algorithmes. De plus, elles sont sélectionnées à des moments de l'année où les mesures sont les plus nombreuses.

	det-00974-01 (Notre-Dame / Haig) direction Est	det-00773-01 Notre-Dame / Pie-IX direction Nord	det-00757-01 (Notre-Dame / Bourbonnière) direction Ouest
Période d'entraînement	2018-02-28 au 2018-04-30	2017-10-31 au 2017-12-31	2017-10-31 au 2017-12-31
Période de validation	2018-05-01 au 2018-05-31	2018-01-01 au 2018-01-31	2018-01-01 au 2018-01-31

Tableau 2 - Périodes de collecte de données utilisées pour l'entraînement et la validation des algorithmes. Une seule direction a été choisie par détecteur.

Les figures 3 et 4 montrent la distribution des données de chacune des paires de détecteur et de zone. Ce sont des représentations des données agrégées aux 15 minutes pour les périodes choisies. La majorité des mesures se situent entre 0 et 5% d'occupation et entre 0 et 250 véhicules par période de 15 minutes. De plus, on peut observer des valeurs aberrantes de nombre de véhicules. En effet, il serait surprenant que 1000 véhicules passent en 15 minutes devant le détecteur det-00757-01 de la zone 2.

Les données de circulation sont caractérisées par une saisonnalité journalière et hebdomadaire que l'on peut observer à la Figure 5.

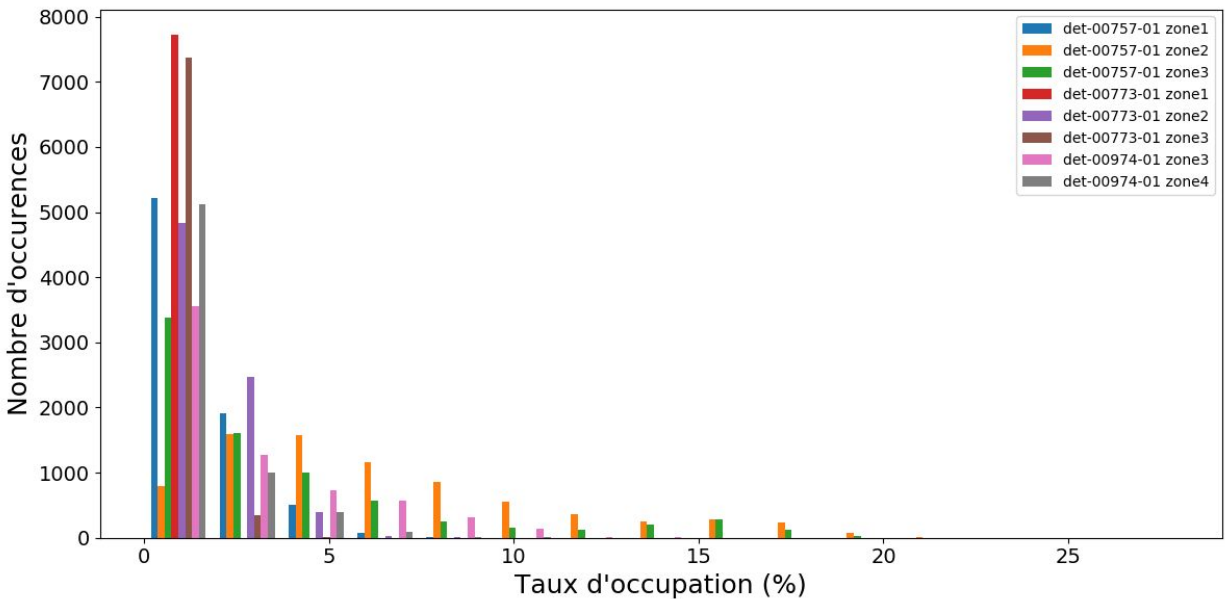


Figure 3: Distribution des données par détecteur et par zone. Les paires det-00757-01 zone 2 ainsi que det-00757-01 zone 3 sont celles qui ont fréquemment des valeurs de taux d'occupation élevés.

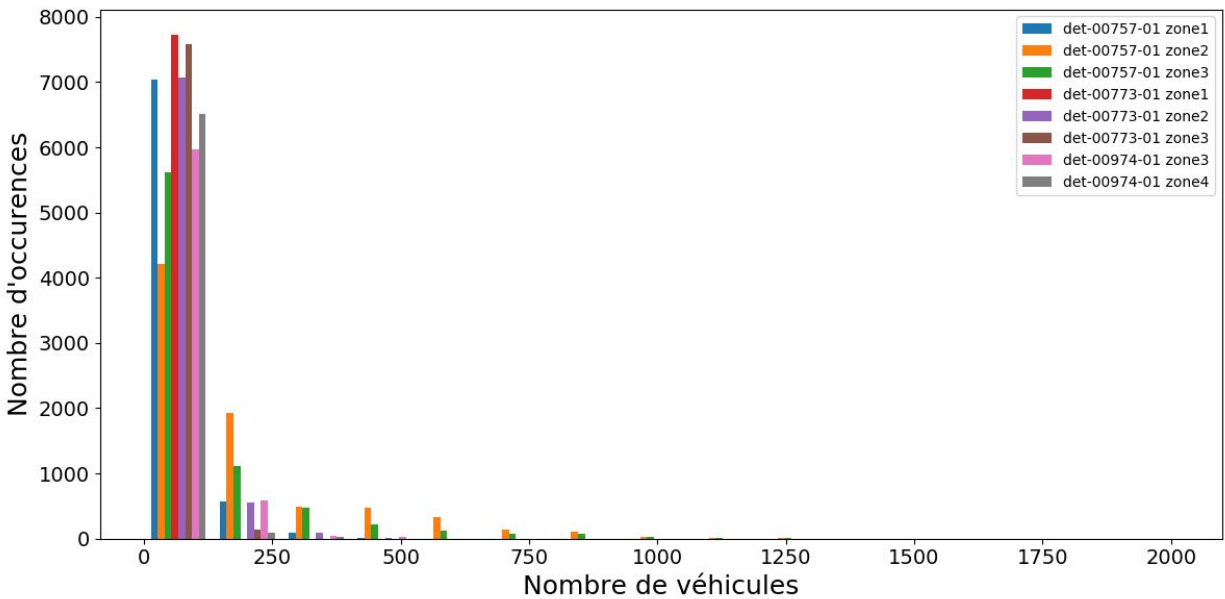


Figure 4: Distribution des données par détecteur et par zone. Les paires det-00757-01 zone 2 ainsi que det-00757-01 zone 3 sont celles qui ont fréquemment des valeurs de nombre de véhicules élevés.

Traitement des données

Les données brutes représentent un grand volume et il a été nécessaire d'effectuer un prétraitement pour en diminuer la fréquence et permettre un temps de réponse plus rapide. Pour ce faire, nous avons effectué une agrégation par moyenne pour le taux d'occupation et par somme pour le volume de véhicules.

Comme entrée au modèle, nous avons calculé les valeurs suivantes chaque minute:

- Les valeurs récentes passées
- Les tendances récentes
- L'heure de la journée ainsi que le jour de la semaine
- Des agrégations sur des fenêtres glissantes passées

Les données de chaque zone d'un détecteur sont utilisées pour la prédiction des zones avoisinantes.

Pour utiliser un point de donnée lors de l'entraînement du modèle, il est nécessaire que toutes les valeurs de composantes soient non nulles, sinon ce point ne sera pas utilisé. Cette condition s'applique également au moment de la prédiction.

Une analyse subséquente a permis d'isoler les facteurs les plus importants pour augmenter la performance du modèle.

L'algorithme d'apprentissage automatique choisi est XGBoost. Il offre l'avantage d'être robuste aux composantes moins utiles ce qui est bénéfique dans notre cas puisque nous souhaitons déployer notre modèle sur plusieurs tronçons et donc l'importance de chaque composante peut varier d'un tronçon à l'autre.

Méthode d'évaluation

Pour évaluer la performance des modèles, la MAE a été choisie. Cette mesure de performance offre l'avantage d'être facilement interprétable. De plus, c'est une mesure qui est robuste aux valeurs nulles qui sont inhérentes au jeu de données courant. En effet, la nuit lorsque la circulation est faible, il est fréquent que la valeur agrégée à la minute soit nulle si aucune voiture n'est passée dans une zone.

Pour comparer la plus-value apportée par le modèle, nous avons comparé les mesures d'erreur avec celles provenant d'une méthode naïve qui ne fait que répéter la dernière valeur observée.

Résultats

Dans cette section sont présentées les mesures d'erreurs des prédictions des modèles sur leur période de validation. À la Figure 5, nous observons les allures des courbes sur une semaine de l'ensemble de données de validation. On remarque que les saisonnalités sont suivies par le modèle. Cependant, les valeurs extrêmes ne sont pas toutes bien apprises par le modèle.

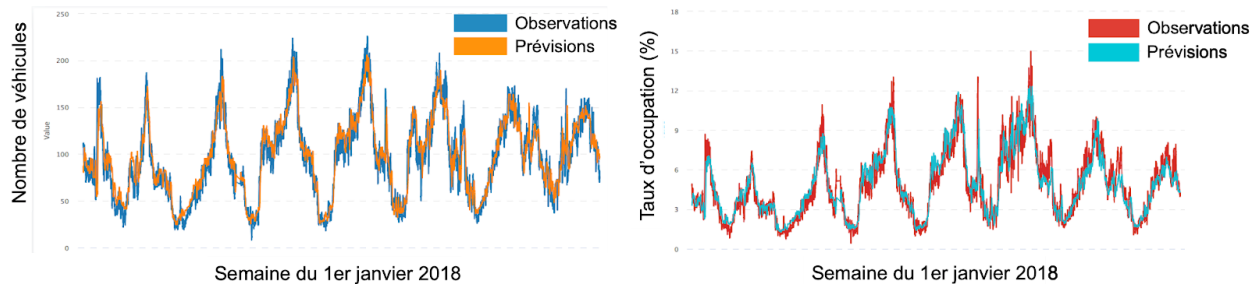
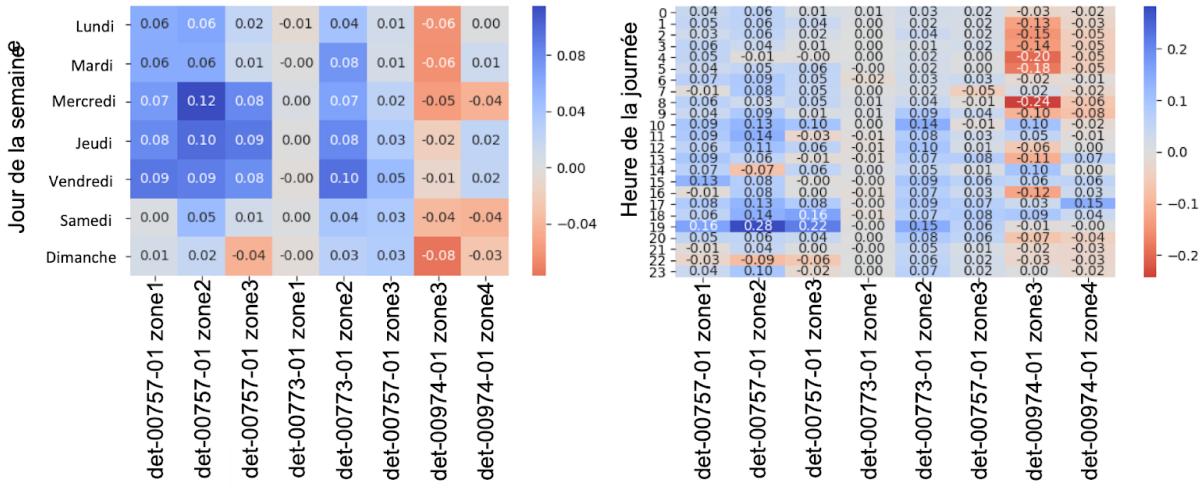
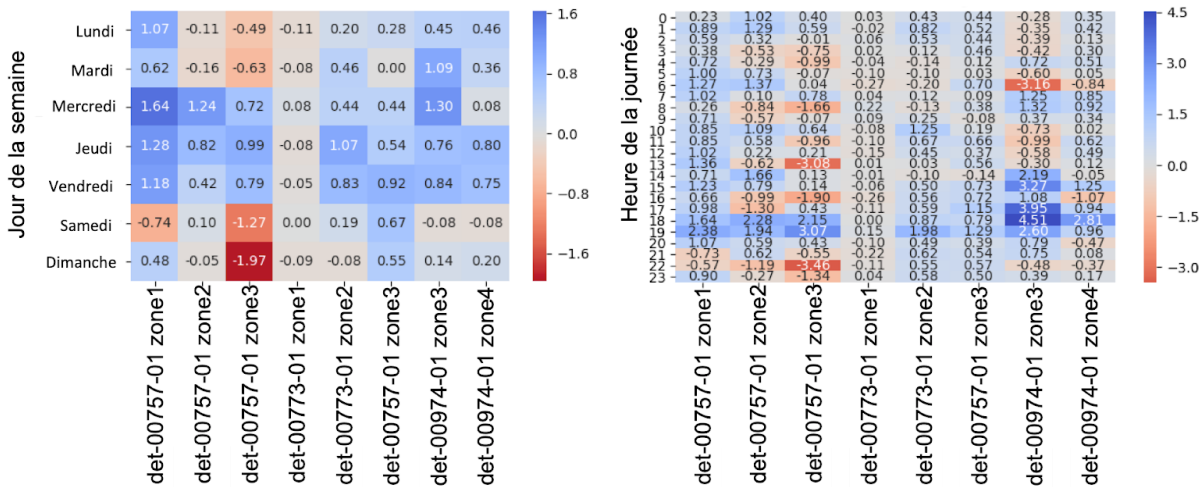


Figure 5: Allure des observations de nombre de véhicules et du taux d'occupation de la zone 2 du détecteur det-00757-01. À haut niveau le modèle suit les saisonnalités des observations.

La figure 6 montre les résultats en termes de différence de MAE entre le modèle d'apprentissage automatique et la méthode naïve. La couleur donne une indication sur le modèle qui prédit le mieux. Lorsqu'une case est rouge, cela signifie que la méthode naïve est supérieure alors que lorsque la case est bleue, cela signifie que le modèle appris est supérieur et donc que son erreur est plus faible. Les résultats sont montrés par heure de la journée ainsi que par jour de la semaine, et ce, pour chaque zone de chaque détecteur.



(a-b)



(c-d)

Figure 6: Comparaison des modèles développés versus une approche naïve par heure de la journée et par journée de la semaine pour le taux d'occupation (a-b) et pour le volume (c-d).

On remarque à la Figure 6 (a,c) que les mercredis, jeudis et vendredis le modèle est supérieur à la méthode naïve, pour 6 des 8 zones considérées et ce pour les deux tâches de prédiction du taux d'occupation et de prédiction du volume de véhicules. Pour la zone 1 du détecteur det-00773-01, on remarque que la différence entre les 2 approches est très faible. Lors des journées de fin de semaine, l'approche naïve est souvent plus précise que le modèle, plus spécialement pour la zone 3 du détecteur det-00757-01.

La Figure 6 (b,d) montre que le modèle performe mieux entre 14h et 19h comparativement à l'approche naïve. La plus grande différence est à la zone 3 du détecteur det-00974-01. La méthode naïve performe encore une fois mieux que le modèle pour la zone 3 du détecteur det-00757-01 et la différence est très faible pour la zone 1 du détecteur det-00773-01.

Le tableau 3 montre la MAE agrégée par zone et par tronçon de chaque détecteur choisi. Le modèle entraîné est légèrement supérieur à la méthode naïve une fois que les résultats sont agrégés.

détecteur / Zone	MAE modèle entraîné (% d'occupation)	MAE méthode naïve (% d'occupation)	MAE modèle entraîné (nombre de véhicules)	MAE méthode naïve (nombre de véhicules)
det-00757-01 / zone 1	0.42	0.47	7.19	8.01
det-00757-01 / zone 2	0.91	0.98	14.88	15.23
det-00757-01 / zone 3	0.70	0.74	11.57	11.35
det-00773-01 / zone 1	0.03	0.03	0.75	0.71
det-00773-01 / zone 2	0.38	0.44	6.07	6.52
det-00773-01 / zone 3	0.24	0.27	4.93	5.41
det-00974-01 / zone 3	0.57	0.52	8.14	8.79
det-00974-01 / zone 4	0.34	0.34	5.65	6.01
Moyenne	0.45	0.47	7.40	7.75

Tableau 3: Erreur moyenne de prédiction par détecteur et zone de prédiction

Discussion des résultats

Il est important de comparer la performance des approches dans les moments où la prise de décision peut réduire la congestion dans la Ville. C'est pour cette raison qu'il y aura discussion sur la performance des approches pendant l'heure de pointe, les différences entre les données acquises à Lyon et à Montréal ainsi que les cas spécifiques où la méthode naïve performe mieux que le modèle entraîné.

La figure 7 montre la relation entre le taux d'occupation et le volume de véhicules dans la ville de Lyon (Wynter et al., 2015). Les auteurs décrivent à l'aide d'une courbe de second degré qu'il est possible de déterminer les situations de circulation problématiques. En effet, lorsque le point se trouve à gauche du maximum de la courbe, le trafic est fluide et lorsqu'il se trouve à sa droite, la circulation est congestionnée.

Nous avons tenté de réaliser la même analyse avec les données disponibles à la Ville de Montréal. À la figure 8, on observe la relation entre le taux d'occupation et le volume de véhicules passant sur les zones des tronçons sélectionnés de la Ville de Montréal. La courbe orange montre une tentative de caractériser les données à l'aide d'une courbe de second degré. On remarque que dans aucune des zones on observe la même tendance qu'à Lyon. Voyant cette différence entre les données, il serait difficile de tirer des conclusions pertinentes

en comparant les mesures d'erreur obtenues dans l'étude réalisée sur les données de la ville de Lyon et celles exposées ici. Toutefois, il serait intéressant de pouvoir comparer les performances des deux approches sur le même jeu de données.

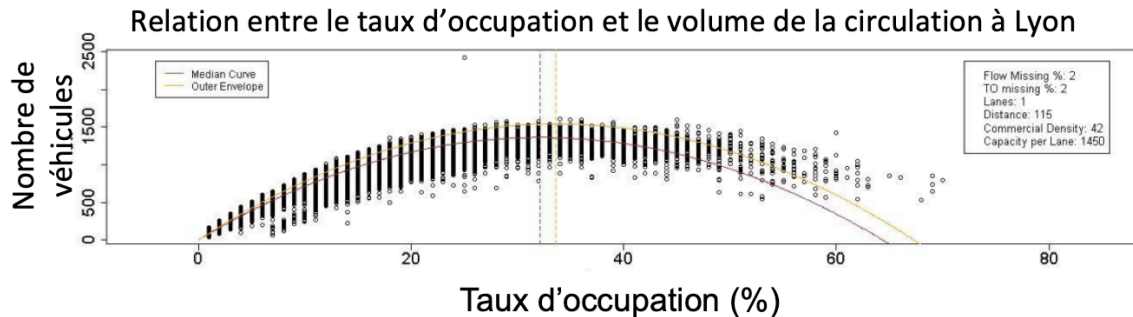


Figure 7: Relation entre le taux d'occupation et le volume de véhicules dans les données de la ville de Lyon. On remarque une tendance quadratique inverse entre les variables cibles (Wynter et al., 2015).

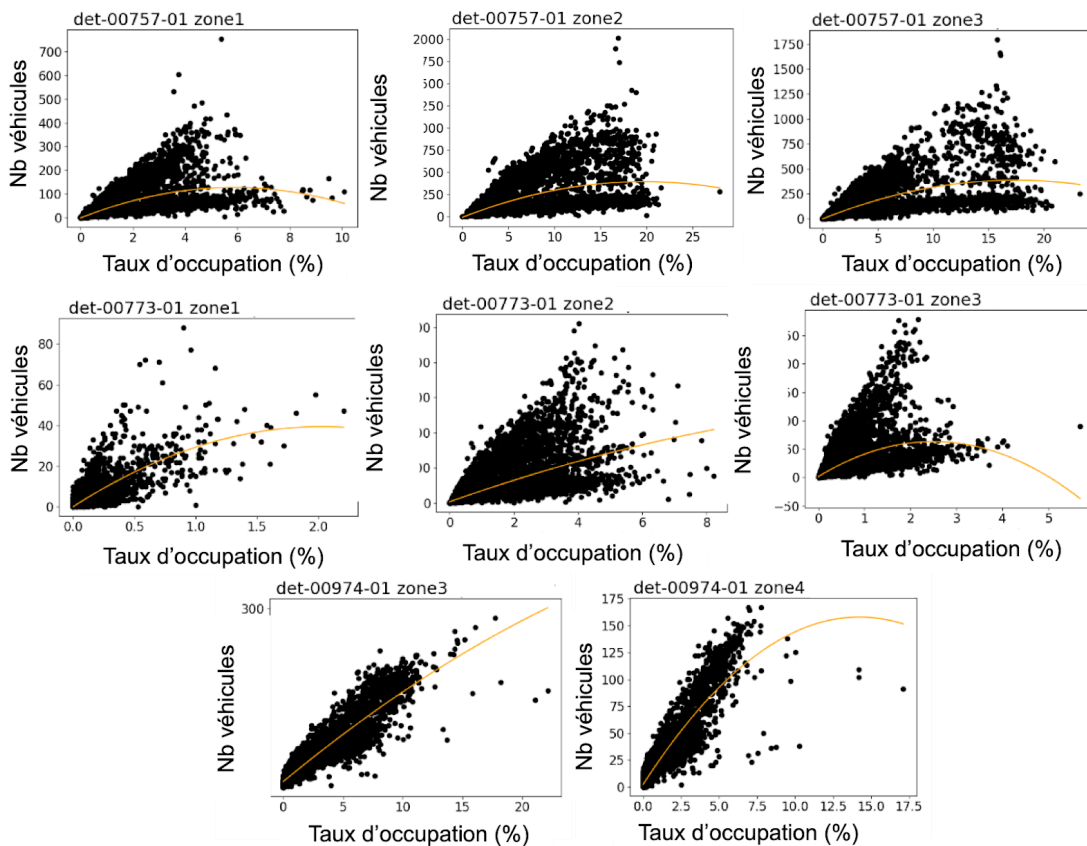


Figure 8: Relation entre le taux d'occupation et le volume de véhicules sur les zones sélectionnées de la Ville de Montréal.

Nous pouvons toutefois remarquer qu'à la Figure 8, pour certaines intersections tel que le détecteur det-00974-01, la tendance est linéaire.

Suite aux résultats obtenus à la Figure 6, on remarque que le modèle entraîné performe mieux par rapport à la méthode naïve lors de l'heure de pointe de fin de journée (16h à 19h) et lors des jours de semaine. C'est à ces moments qu'il est le plus utile pour la Ville de prendre des décisions éclairées puisque c'est pendant ces périodes que la congestion est la plus grande sur le réseau routier. De plus, il est attendu que la méthode naïve fonctionne bien lorsqu'il y a peu de fluctuation dans les données. C'est pour cette raison que la différence de performance est moins marquée le soir et les fins de semaine. Les valeurs de MAE présenté au tableau 3 ont été jugées satisfaisantes par l'équipe d'ingénierie de la Ville pour aller de l'avant avec les prochaines phases du projet.

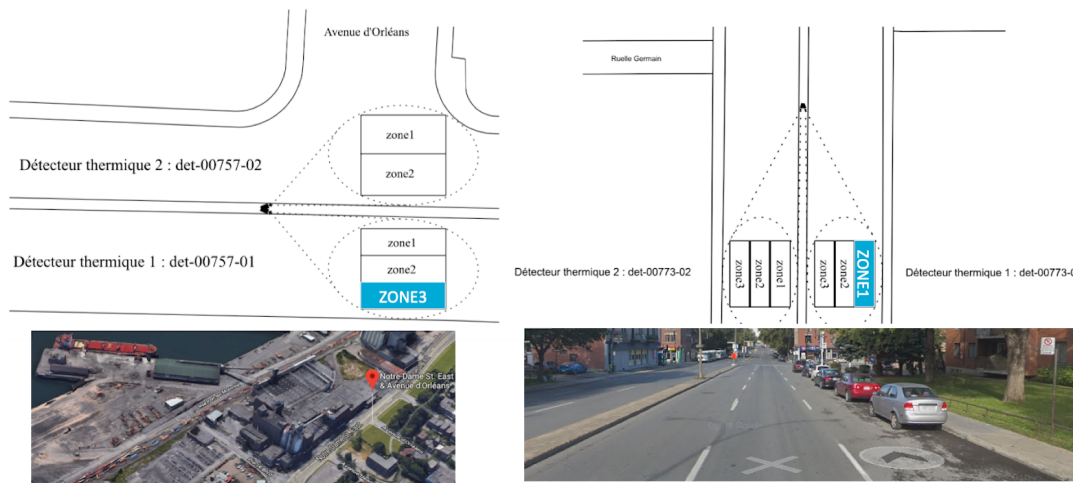


Figure 9: (a) Vue en contexte de la zone 3 du détecteur det-00757-01. Elle donne accès à une entreprise. (b) Vue en contexte de la zone 1 du det-00773-01. On remarque que la voie est utilisée comme voie de stationnement.

La Figure 9 (a) montre que la zone 3 du détecteur det-00757-01 peut être utilisée pour tourner à droite vers une entreprise. Cette activité de livraison pourrait rendre la tâche d'apprentissage du modèle plus complexe puisque l'activité de cette zone peut être fortement différente des autres zones de ce tronçon. En effet, on peut imaginer que cette zone est bloqué en pleine semaine parce que la livraison fait attendre les camions et non pas parce qu'il suit le flux périodique habituel de la circulation (pointe en soirée, en semaine). Il serait alors intéressant d'étudier les cycles d'achalandage de l'entreprise et étudier l'effet que cela peut avoir sur le modèle. La Figure 9 (b) montre que la zone 1 du détecteur det-00773-01 est parfois utilisée comme zone de stationnement. Son flux de véhicule en est donc fortement affecté. Un ajout intéressant au modèle pourrait être les horaires de stationnement des différentes voies de circulation.

En plus des activités de livraison et des horaires de stationnement, il serait intéressant d'étudier l'impact d'autres sources de données dans le modèle. Un avantage du modèle choisi est qu'il apprend sur différentes composantes numériques et possiblement catégorielles si correctement transformées. Cela signifie qu'au fur et à mesure que la Ville choisi d'enrichir le flux de données avec d'autres sources il sera possible d'étudier l'impact de l'ajout des nouvelles composantes au modèle et les intégrer si elles ont un impact positif. On peut imaginer d'ajouter les prédictions météorologiques, les événements d'accidents ou bien le calendrier des événements spéciaux dans chaque quartier de la métropole, par exemple l'horaire des spectacles et des événements sportifs du Centre Bell au centre-ville.

Phases de déploiement du projet

Afin de faciliter le travail en parallèle de chaque équipe, soit l'équipe de données et l'équipe d'ingénierie de circulation, le projet a été pensé pour être réalisé en plusieurs phases. Cette évolution progressive permet également de valider les différents concepts requis pour le fonctionnement optimal de la solution sans dépenser des sommes inutilement et d'impliquer les bonnes personnes au bon moment afin d'être plus efficace. Ce fonctionnement permet également aux équipes de valider le travail de l'autre et d'apprendre graduellement sur les techniques mises en place par l'autre équipe.

Les différentes phases permettent de:

- Mettre en place le mode TRP sur le secteur sélectionné
- Valider la prédiction de la circulation
- Mettre en place un mode TRP utilisant la prédiction qui optimise la gestion de la circulation, soit le mode *TPP*.

Phase 1: Prédiction sur historique et implantation TRP

C'est à cette phase que l'équipe d'ingénierie de circulation met en place le mode TRP sur le secteur sélectionné. Ceci implique d'utiliser des données temps-réel (volume et occupation) récupérées par les contrôleurs de feux de circulation comme intrant pour la fonction TRP, fonction configurée dans le système de contrôle des feux de circulation. C'est la fonction TRP qui sélectionne à l'aide des données reçues, le plan de feux de circulation approprié (période d'affichage vert, jaune et rouge) pour la combinaison volume/occupation reçue.

Les experts de données utilisent à cette phase des données sauvegardées (historique volume et occupation) afin de concevoir et valider le modèle de prédiction des données de circulation.

À cette phase l'équipe d'ingénierie de circulation valide si la prédiction de circulation est assez précise (MAE du taux d'occupation < 1%, MAE du volume < 20 véhicules par 15 minutes en heure de pointe) et l'équipe de données apprend à comprendre le mode TRP.

Voir la figure 10 pour le détail du processus de données et de la responsabilité de chaque équipe.

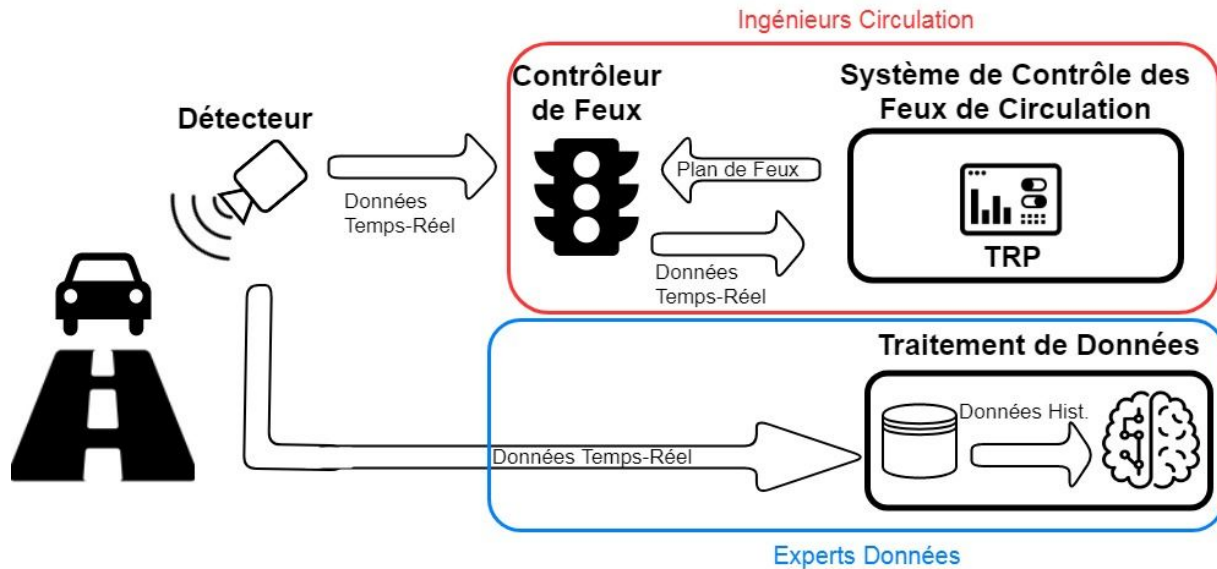


Figure 10 : Phase 1 de prédiction avec historique de données et implantation du mode TRP.

Au moment d'écrire cet article, nous avons maintenant complété cette phase du projet.

Phase 2: Prédiction temps-réel et donnée interface pour TRP

À cette phase l'équipe d'ingénierie de circulation utilise des données temps-réel (volume et occupation) récupérées par le module de traitement des données comme intrant pour la fonction TRP. C'est toujours la fonction TRP qui sélectionne, à l'aide des données reçues, le plan de feux de circulation approprié (période d'affichage vert, jaune et rouge) pour la combinaison volume/occupation reçue.

Les experts de données utilisent à cette phase les données temps-réel (volume et occupation) afin d'alimenter le modèle de prédiction des données de circulation.

Le but de cette phase est de valider que le module de traitement de données est en mesure d'acheminer des données au système de contrôle des feux de circulation. Les deux équipes, ingénierie de circulation et expert de données, travaillent ensemble à mettre en place et valider ce nouveau lien.

Voir la figure 11 pour le détail du processus de données et de la responsabilité de chaque équipe.

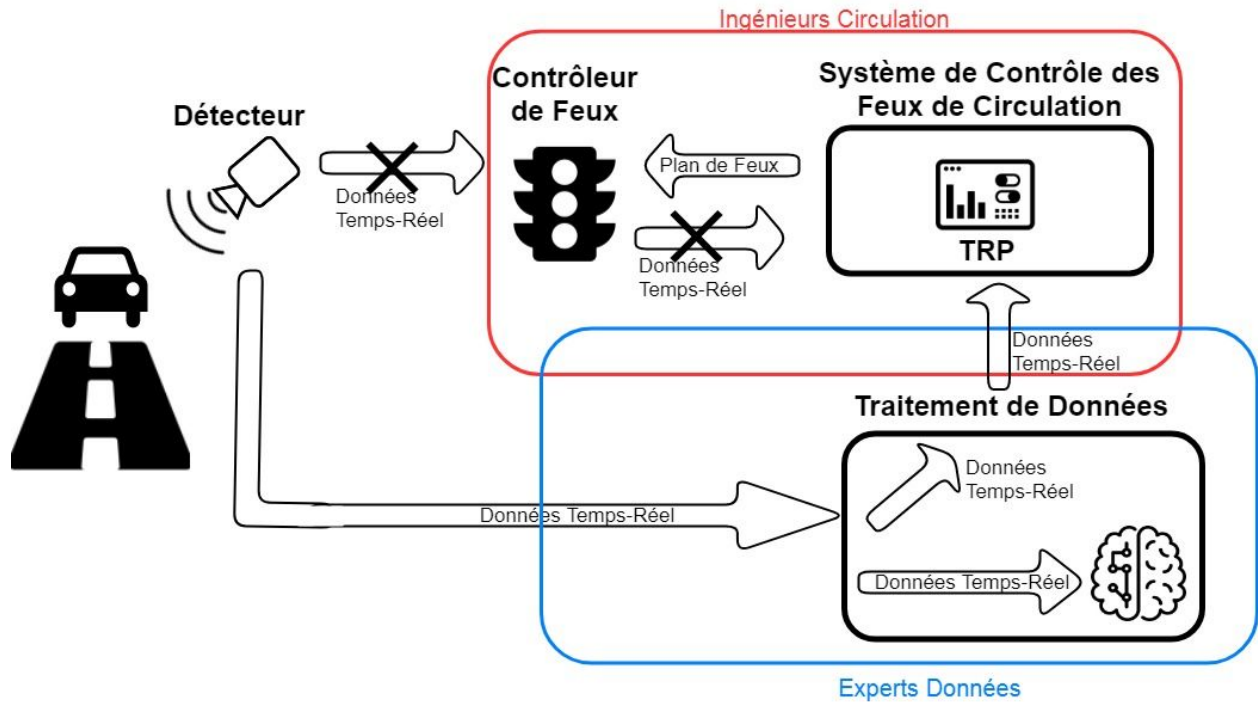


Figure 11 : Phase 2 de prédiction avec données temps-réel et utilisation pour le mode TRP.

Phase 3: Données prédiction pour TRP

À cette phase l'équipe d'ingénierie de circulation utilise les données prédites (volume et occupation) récupérées par le module de traitement de données comme intrant pour la fonction TRP. C'est toujours la fonction TRP qui sélectionne à l'aide des données reçues, le plan de feux de circulation approprié (période d'affichage vert, jaune et rouge) pour la combinaison volume/occupation reçue.

Les experts de données s'assurent à cette phase que des données de prédiction de la circulation de qualité sont partagées au système de contrôle des feux de circulation.

Les deux équipes, ingénierie de circulation et expert de données, travaillent ensemble à mettre en place et valider cette phase afin que les données prédites soient adéquates en tout temps pour la fonction TRP. Le module de traitement doit tenir compte des événements imprévus, par exemple un accident ou une panne d'un détecteur.

Voir la figure 12 pour le détail du processus de données et de la responsabilité de chaque équipe.

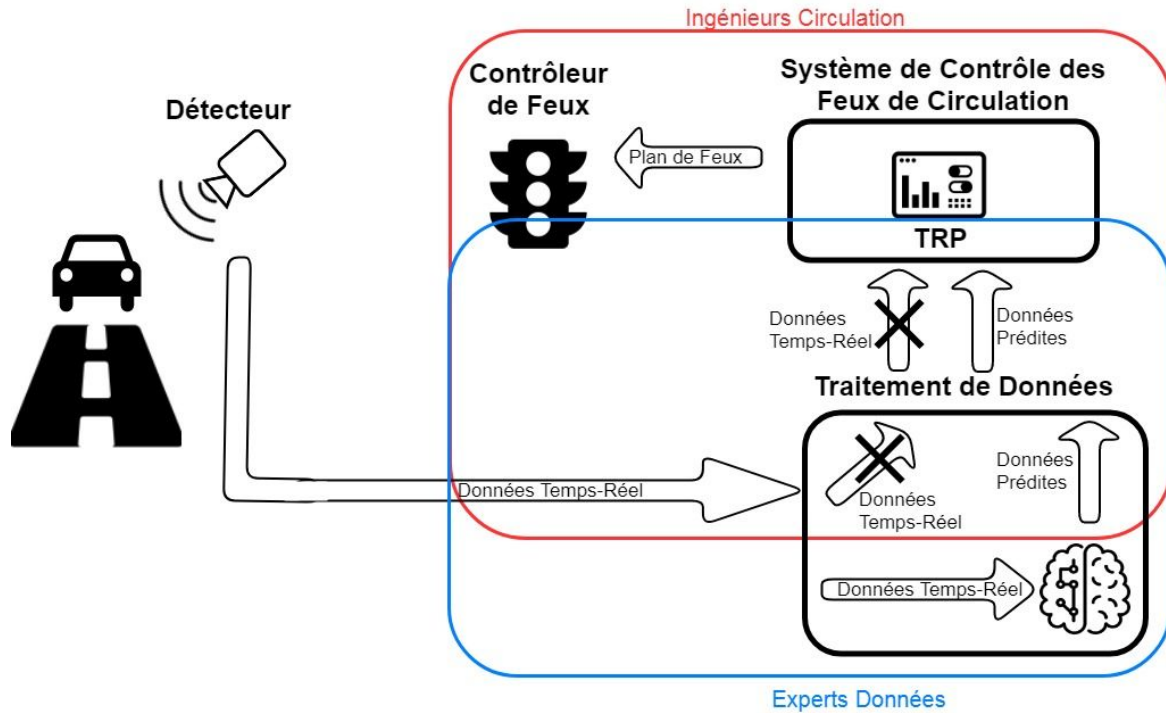


Figure 12 : Phase 3; utilisation des prédictions pour le mode TRP.

Phase 4: Optimisation et autres sources de données

Cette phase n'est pas clairement définie, mais nous savons qu'avec l'architecture mise en place de multiples possibilités d'optimisation et de validation du mode TRP sont possibles. Par exemple en comparant avec un outil d'intelligence artificielle la sélection des plans de feux et les données de circulation temps-réel, nous pourrions probablement créer notre propre algorithme pour remplacer le mode TRP. Nous pensons également que d'autres intrants pourraient nous aider à mieux prédire la circulation tel que la météo, des entraves routières (travaux, accidents, incendies) ou même des routes fermées par des événements spéciaux (manifestation, festival ou événement sportif).

Voir la figure 13 pour le détail des exemples d'optimisations futurs.

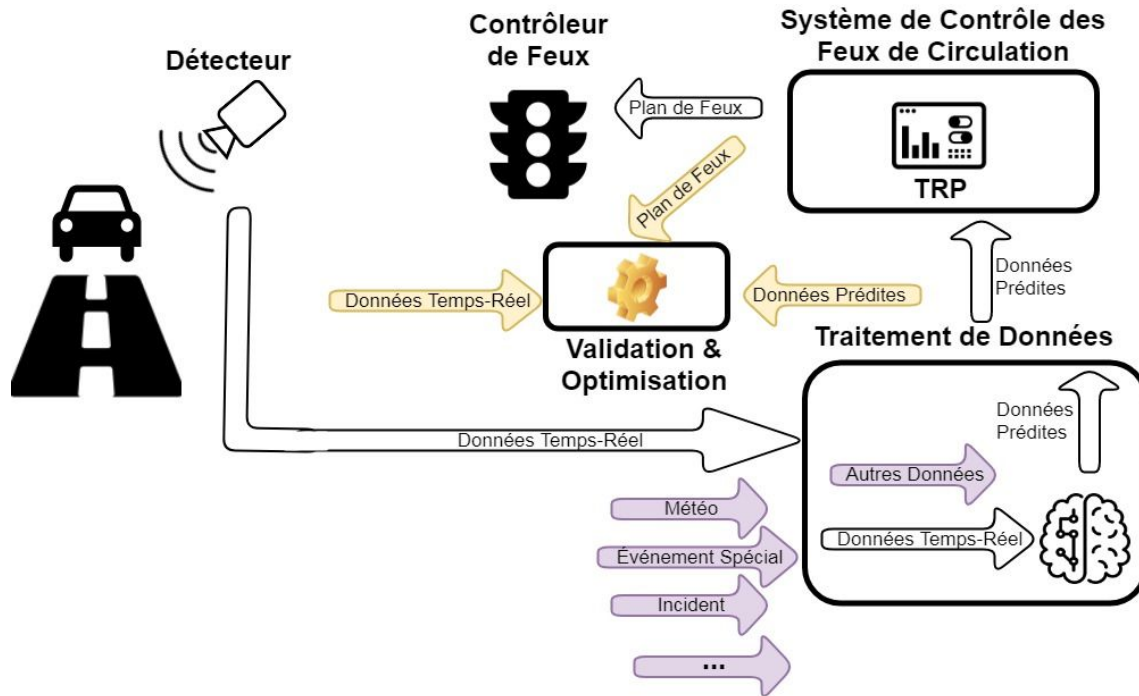


Figure 13 : Phase 4; optimisations possibles pour le mode TRP.

Conclusion

Suite à la réalisation de la phase 1 du projet, nous avons pu confirmer que le modèle d'apprentissage machine est meilleur que la méthode naïve (dernier 15 minutes) et qu'il satisfait l'équipe d'ingénierie de circulation. Le modèle est flexible et pourra être entraîné avec d'autres sources de données afin d'améliorer sa performance en toutes circonstances. De plus, il est indépendant du tronçon sur lequel il peut être implanté et pourra être facilement déployé sur les secteurs clés de la Ville sans adaptation.

La phase 2 est en cours et permettra de s'approcher davantage de la mise en production complète du mode gestion TRP avec prédiction (TRP). Le but ultime de ce projet. En parallèle également nos équipes suivent de près l'arrivée des véhicules connectés et autonomes et évaluent l'impact de ces technologies sur notre façon de gérer les feux de circulation.

Références

- Castro, P. S., Zhang, D., & Li, S. (2012). Urban Traffic Modelling and Prediction Using Large Scale Taxi GPS Traces. Dans J. Kay, P. Lukowicz, H. Tokuda, P. Olivier, & A. Krüger (Éd.), *Pervasive Computing* (p. 57-72). Springer Berlin Heidelberg.
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Min, W., & Wynter, L. (2011). Real-time road traffic prediction with spatio-temporal correlations. *Transportation Research Part C: Emerging Technologies*, 19(4), 606-616. <https://doi.org/10.1016/j.trc.2010.10.002>
- Pan, B., Demiryurek, U., & Shahabi, C. (2012). Utilizing Real-World Transportation Data for Accurate Traffic Prediction. *2012 IEEE 12th International Conference on Data Mining*, 595-604. <https://doi.org/10.1109/ICDM.2012.52>
- Traffic light. (2019). Dans *Wikipedia*. Consulté à l'adresse https://en.wikipedia.org/w/index.php?title=Traffic_light&oldid=896680815
- Ville de Montréal. (2019). Informations de circulation en temps réel - CKAN. Consulté 9 mai 2019, à l'adresse <http://donnees.ville.montreal.qc.ca/dataset/circulation-mobilite-temps-reel>
- Wynter, L., Trager, B. M., Yu, Y., Kararianakis, Y., Jabari, S., Coldefy, J., ... Baudel, T. (2015). Traffic estimation and prediction for urban road networks, application to grandlyon. *ITS World Congress, Bordeaux. ERTICO*. Citeseer.
- Xie, J., & Choi, Y.-K. (2017). Hybrid traffic prediction scheme for intelligent transportation systems based on historical and real-time data. *International Journal of Distributed Sensor Networks*, 13(11), 1550147717745009. <https://doi.org/10.1177/1550147717745009>