

Projet Prédiction du Trafic Urbain à l'aide de Données IoT pour une Ville Intelligente

Contexte :

Dans le cadre d'une initiative de transformation urbaine, votre ville s'engage dans un ambitieux projet pour devenir une **ville intelligente** (smart city) en intégrant des technologies avancées de gestion et de prévision du trafic. Cette évolution vise à améliorer la qualité de vie des citoyens en optimisant les services publics, en réduisant les embouteillages, et en planifiant de façon proactive les infrastructures futures.

Vous travaillez en tant que **data scientist** au sein de l'équipe chargée de la gestion du trafic, en collaboration avec le gouvernement local. Votre mission consiste à développer un modèle de prédiction du trafic urbain capable de prévoir les volumes de circulation aux principaux carrefours de la ville. L'objectif final est de fournir une solution qui permette de gérer les pics de trafic, notamment durant les heures de pointe, les jours fériés et les événements spéciaux. Ces prévisions contribueront à la conception d'un système de circulation plus robuste et adaptable, soutenant ainsi l'essor de la ville intelligente.

Problématique :

Les schémas de circulation varient considérablement selon les périodes (jours ouvrés, week-ends, jours fériés). Le gouvernement souhaite anticiper ces fluctuations pour mieux gérer le trafic et répondre aux besoins d'une population urbaine en croissance. À partir de données de trafic captées par des capteurs IoT placés sur quatre carrefours clés de la ville, vous devrez élaborer un modèle capable de prévoir l'affluence de véhicules pour les quatre prochains mois, sur la base d'un historique de 20 mois.

Enjeux :

1. **Complexité des Données** : Les données sont hétérogènes, certaines périodes sont sous-représentées, et des données manquent parfois pour certains carrefours.
2. **Temporalité et Saisonnalité** : La prédiction doit tenir compte des variations saisonnières, des jours de semaine et des jours exceptionnels.
3. **Préparation pour l'Avenir** : Les prévisions de trafic serviront de base pour des décisions opérationnelles et stratégiques, visant à faire de la ville un environnement intelligent et réactif aux besoins des citoyens.

Objectifs :

- Prévoir le volume de trafic dans les quatre carrefours pour les quatre mois à venir.
 - Analyser et visualiser les tendances de circulation pour identifier les périodes de trafic intense.
 - Proposer des recommandations pour l'optimisation de la gestion du trafic urbain.
-

Description des Données

Vous disposerez d'un jeu de données comprenant :

- **DateTime** : Date et heure de chaque mesure.
- **Junction** : Identifiant unique de chaque carrefour (valeurs de 1 à 4).
- **Vehicles** : Nombre de véhicules détectés par les capteurs IoT.
- **ID** : Identifiant de chaque mesure.

Instructions et Phases du Projet

Phase 1 : Exploration et Compréhension de l'Ensemble de Données

- Analyser les variables pour identifier les tendances temporelles et comprendre la structure des données.
- Identifier les valeurs manquantes, les incohérences et la périodicité des mesures.
- Questions à se poser : Quelles périodes sont bien couvertes ? Quels schémas de circulation semblent émerger ? Quelles sont les sources de variation ?

Phase 2 : Prétraitement des Données

- **Nettoyage des Données** : Traiter les valeurs manquantes en utilisant des techniques comme l'interpolation ou les moyennes par périodes similaires.
- **Transformation des Données** : Créer des variables pour capturer les caractéristiques temporelles, telles que le jour de la semaine, le type de jour (ouvré, week-end, férié) et la saison.
- **Normalisation et Structuration** : Préparer les données sous une forme exploitable par les modèles, notamment pour les séries temporelles.

Phase 3 : Analyse Exploratoire des Données (AED)

- **Visualisation des tendances** : Générer des graphiques pour observer l'évolution du trafic par carrefour, jour et heure.
- **Identification des Périodes Clés** : Mettre en évidence les périodes de forte affluence et les comportements de trafic spécifiques aux jours fériés ou aux événements.
- **Corrélations et Répartition** : Analyser les corrélations entre le trafic aux différents carrefours et les tendances saisonnières.

Phase 4 : Modélisation et Prédiction

- **Choix des Modèles** : Utiliser des modèles adaptés aux séries temporelles (comme ARIMA, Prophet) et des modèles basés sur les réseaux neuronaux (comme LSTM) pour les données chronologiques.
- **Évaluation et Validation** : Appliquer une validation croisée temporelle pour évaluer la précision de chaque modèle et choisir celui offrant les meilleures prévisions.
- **Prédiction** : Déployer le modèle retenu pour générer les prévisions sur quatre mois, en analysant les scénarios possibles et en identifiant les périodes critiques de trafic.

Phase 5 : Interprétation des Résultats et Recommandations

Analyse des Résultats

- **Précision et Fiabilité des Prédictions** : Évaluer la précision du modèle en comparant les prévisions aux données historiques . Identifier les points où les prédictions sont particulièrement fiables et les zones d'incertitude, notamment pour les périodes de trafic intense ou les jours exceptionnels (jours fériés, événements locaux).
- **Identification des Périodes Critiques** : Analyser les résultats pour mettre en évidence les périodes de pic de trafic aux différents carrefours. Ces informations pourront guider des décisions opérationnelles, telles que l'optimisation des signalisations ou l'anticipation de la gestion des flux.