**Projet**

**Prédiction du trafic à partir de données IoT dans une smart city**

# **Introduction:**

Notre dataset comprend des données cruciales sur le trafic collecté à différents points de jonction. Nous avons commencé par comprendre l’ensemble de données, et puis nous avons procédé au prétraitement de celui-ci, en nous débarrassant de toute donnée incorrecte et en normant nos valeurs, en rendant nos données prêtes pour l’analyse.

Nous avons ensuite effectué l’analyse exploratoire des données et la modélisation, ce qui nous a permis de mieux comprendre nos données.

# **Prétraitement des Données :**

Les différences notables dans notre ensemble de données après le prétraitement sont les nouvelles colonnes ajoutées pour distinguer les jours les uns des autres, de l’identification de certains jours fériés à la séparation des jours de la semaine des week-ends.

Nous avons également ajouté une nouvelle colonne pour les véhicules, où les valeurs de l’original sont normalisées dans un intervalle de [0,1].

A screenshot of a computer

Description automatically generated

Figure 1 - Le résultat de notre prétraitement.

# **Analyse Exploratoire des Données (AED) :**

Pour cette étape, nous avons travaillé avec les bibliothèques python matplotlib et seaborn. Les graphiques que nous avons créés et leurs interprétations sont les suivants :

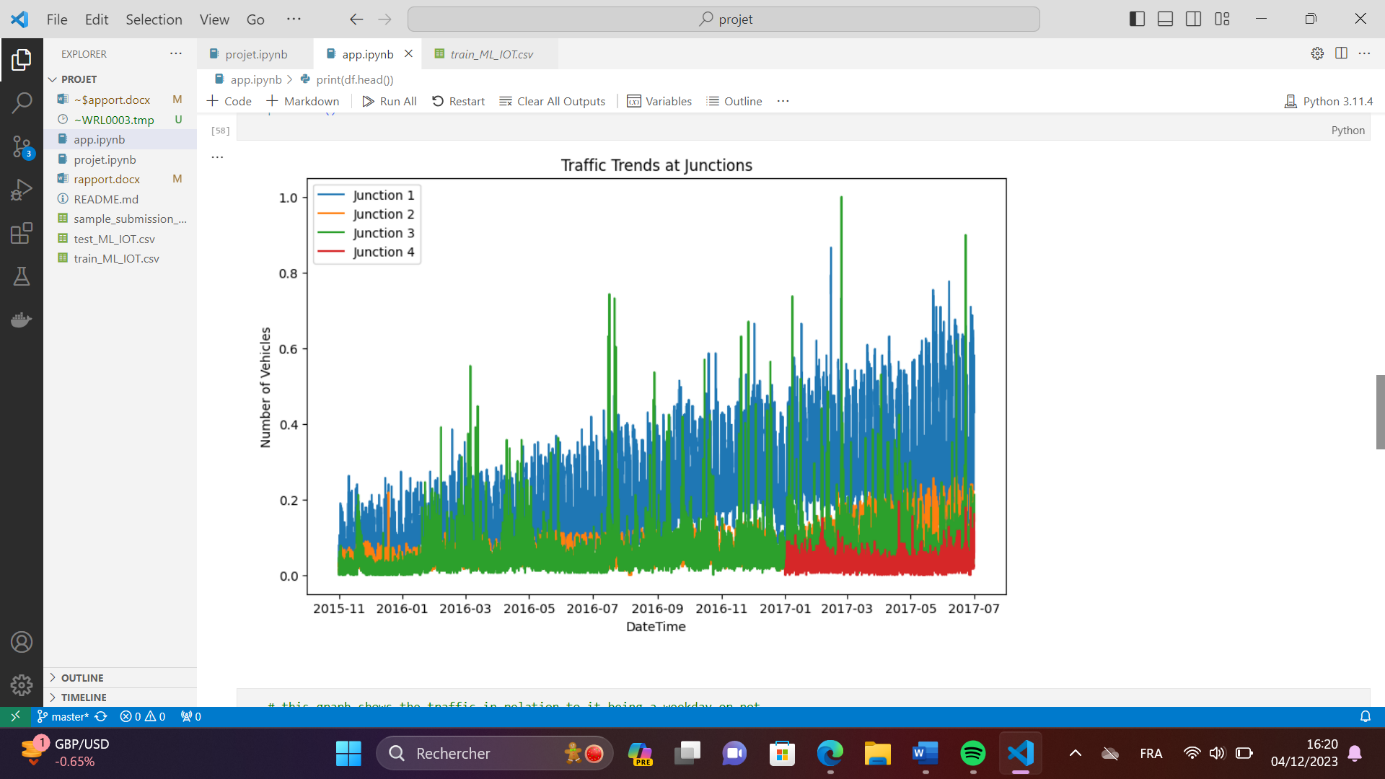


Figure - Temporal Trend Visualization

Ce graphique montre les tendances du trafic à différents points de jonction au fil du temps. Il utilise la colonne 'DateTime' sur l’axe des x, représentant le temps, et la colonne 'Normalized\_Vehicles' sur l’axe des y, représentant le nombre de véhicules normalisés à des fins de comparaison.

Nous remarquons que le premier carrefour (bleu) est le plus cohérent avec le trafic élevé et augmente progressivement avec les années. . Alors que le troisième carrefour (vert) a les pointes les plus élevées à des dates particulières, que nous pourrions supposer sont des vacances annuelles ou des événements spéciaux qui attirent les foules.

La deuxième jonction (jaune) a très peu de trafic tout au long de l’année, atteignant son sommet à peine 0,2, mais commence à augmenter lentement au cours des quatre derniers mois. Et la dernière jonction (rouge) semble être très cohérente avec aucune augmentation linéaire du trafic pendant le temps limité dont nous disposons.

A screenshot of a computer

Description automatically generated

Figure 3 - bar graph

Ce diagramme à barres illustre la quantité moyenne de trafic pendant les jours de semaine et les fins de semaine. Il est nettement évident que le volume de trafic automobile pendant les week-ends est inférieur à celui des jours de semaine.

Cette situation sera causée par le fait que les travailleurs et les étudiants devront déménager plus souvent pendant certains jours de la semaine que d’autres et doivent être pris en considération par le système avant de faire toute prédiction.

A screenshot of a computer screen

Description automatically generated

Figure 4 - Hourly Traffic Trend Visualization for weekends and weekdays.

Ces graphiques nous montrent comment le nombre moyen de véhicules, normalisé, varie à chaque heure de la journée pendant les jours de travail et les fins de semaine. Nous remarquons une grosse baisse aux premières heures de la journée, jusqu’à ce qu’il atteigne sa valeur la plus basse (qui est d’environ 12 voitures) à cinq heures du matin, puis elle commence à augmenter progressivement avec le début de la journée et reste élevée jusqu’à minuit.

Pendant les jours de la semaine, il augmente considérablement par rapport aux week-ends, les travailleurs et les étudiants transportant un facteur important dans cela.