Capteur avec sa propre incertitude

1 ligne par camion

**Dictionnaire des données**

1. Timestamp : heure de passage du camion
2. Offset : rien à en tirer à priori

3 . Site\_ID : le pont considérer

1. Stage trace : à ne pas étudier (à retirer du tableau)

5 - Warning flag : typologie des écarts

5 -24 Véhiculer surchargés

5 - 23 incertitude sur la mesure

5 - 22 : Essieu surchargé

légal en France : 13 tonnes pour chaque essieu

5 - 21 à 6.15 : recalage réalisé par le système

5 - 11 système de capteur KO

5 - 9 véhicule changé de catégorie

6 - Lane : analyse dans le sens Honfleur - Le Havre. 1 la voie la plus a droite, 2 la voie à gauche.

Le camion doit rester à gauche sauf pour doubler.

7 - vitesse (m / s) : le capteur est en haut d'une pente

(la pente étante importante, les camions arrivent à l'entrée du pont à 30 km.h)

8 - nombre d'essieux :donne une info sur le type de camion généralement varie

* 1. - subclass id : classification réalisée par le système.

10 -Groupe d'essieu : ex : 113 (simple / simple / tridem)

Minimum 2

11 - poids global du véhicule (devrait etre la somme du poids de chaque essieu) en kilo newton (KN)

12 à 28 : poids sur chaque essieu

12 + N : distance totale entre les essieu : donne par approximation la longueur du camion

13 + N : distances entre chaque axe (en m)

12 + 2N : température utilisée pour la compensation (T°C) : utilisé pour évaluer la dilatation du matériau

13+2N: legacy value a ne pas utiliser

14+12N : chi2 réduit : chi2 calculé entre le signal mesuré et le signal recalculé pour évaluer le poids

15 + 12N : non remonté

Idées brainstorming:

- Site\_id : accès Millau et Saint Avold : étudier l'effet site

Légalité :

Vitesse limitée à 90 km.h sur le pont (la pente étante importante, les camions arrivent à l'entrée du pont à 30 km.h)

* Somme des amendes que l'on peut tirer des écarts (mauvaise circulation, vitesse, mauvaise voie)
* Poids toléré pour chaque essieu en France : 13 tonnes => étudier la distribution du poids sur les essieu ( ont-ils tendance de surchargé à l'arrière ?)

Typologie warning flag / analyse fiabilité du système

Idée comparer le warning flag relevé avec une prévision réalisée avec notre propre modèle

Est-ce que des warnings flags manquants et pourraient etre proposés ?

Température

Analyse de la corrélation entre la température et la détection d'anomalie.

Entre la température l'heure de la journée, la saison…

Classification des poids lourds

Clustering d'après Michelin : poids sur chaque essieu et les distance entre essieu et le nombre d'essieu

Proposer une classification et la comparer par exemple a celle du système (subclass-id)

Essieu

Utile pour classifier les poids lourds. Analyser la présence d'outliers.

Vitesse : analyse de la vitesse en fct du poids. Possibilité de retrouver des groupes qui pourraient etre les typologies de moteurs ?

Poids :

Vérification que le système calcul bien le poids global (somme des poids de chaque essiau)

Contexte physique du pont :

20 % de pente

A traiter

* **Feature enginneering** : retyper certaines données et transformer certaines variables (unités) (ex : poids a recalculer en tonnes, vitesse en km/h, certaines variables a passer en factor (qualitative) ex : Subclass ou group class

* **Statistiques descriptives / exploratoire**s : analyse monodimensionnelle, recherche d'invidus aberrants, découverte des distribution, analyses factorielles qui permettent d'avoir un résumé visuel des données

ACP / ACM sur la base du boosting qui ressort (via $importance) les variables les plus souvent appelées (les plus discriminantes) **Tout le monde : objectif éditer un notebook**

* **Clustering** (analyse non supervisée) via Kmeans / CAH des camions : définir une typologie (trouver un profil type du camion)

Retrouver les individus mal définis **Jérémy**

* **Détection d'anomalies** (analyse supervisée) sur le camion qui passe ou le système lui même

Le niveau anomalie oui / non :

La probabilité ou le quantile de l'anomalie (prévoir autre chose que l'espérance, un intervalle ou un quantile, utiliser le package "quantreg") **Franziska**

Tester l'approche "Anomaly Detection" de Frédéric (cf documentation) alternative au supervisé : **Jérémy**

* **Prévision du flux de camion** (saisonnalité) :

Prévision du flux de camion (modèle de prévision, suivant par exemple le jour)

Facteurs à étudier : profil de passage nuit / jour, jour, période de l'année  **Frédéric**

Agenda

* Prochain point : 21/12 à 19h
* Congés :

Frédéric (en vacances du 23/12 au 01/01)

Franziska (18-19/12 pas dispo du tout, en vacances du 23/12 au 01/01)

Jérémy (23/12 au 01/01 peu disponible sur cette période)

Bonus

* Proposer un modèle de pricing en fonction de la classification des camions

Présentation du travail

* Jupyter R notebook / R notebook Rstudio ou Rmarkdown : la restitution documentaire
* Shiny restitution maquettes pour la présentation