

Sujets pour la SEME Hauts de France
Lille, 29/01/18-02/02/18

Sujet 1

- **Entreprise** : Arcelor Mittal
- **Titre** : Automatisation du choix des paramètres dans un outil de séquençement
- **Descriptif** : Afin de pouvoir fabriquer des produits dans les meilleurs délais possibles, ArcelorMittal déploie actuellement un nouvel outil de séquençement sur l'ensemble de ses lignes de production, ici le laminoir qui transforme des brames d'acier en bobines. Celui-ci permet d'ordonner le passage de bobines de différentes caractéristiques (épaisseur de la tôle, largeur, rigidité,...), en fonction d'un jeu de contraintes propres aux machines, comme par exemple la nécessité de ne pas pouvoir faire varier fortement l'épaisseur de la tôle entre deux bobines successives. Ces champs de contraintes peuvent être assez complexes selon les lignes considérées. Pour se rapprocher d'un problème mathématique, on suppose que le vecteur $y \in Y$ contient l'ensemble des paramètres des brames/bobines en laminoir pendant une journée, et les composantes de $x \in X$ (à valeur continue, discrète ou bivalente) décrivent un séquençement possible. On dispose de deux fonctions qualité $f_1, f_2 : X \times Y \rightarrow \mathbb{R}$, la première mesurant la qualité protection outils de fabrication, et la deuxième la qualité retard de livraison. On souhaite donner un sens au problème mathématique multi-objectif (P1) $\min \{ f(x,y) : x \in X \}$, en cherchant des solutions $x \in X$ optimales au sens de Pareto, et en étudiant le lien avec le problème (P2) $\min \{ w_1 f_1(x,y) + w_2 f_2(x,y) : x \in X \}$ obtenu par « Scalarizing ». Le but du travail est de suggérer un outil d'aide à la décision de choix automatisé de paramètres w_1 et w_2 basé sur une boîte noire de résolution de (P2) fournie par ArcelorMittal de sorte que le compromis retenu entre les deux objectifs corresponde au mieux aux exigences de l'entreprise. En particulier, la valeur de w_j peut aider à mettre à l'échelle les différents objectifs, et donc de les rendre comparables, voir par exemple [Kaisa Miettinen, Nonlinear Multiobjective Optimization. Springer].

Sujet 2

- **Entreprise** : Delabie
- **Titre** : Loi de déplacement d'un piston immergé
- **Descriptif** : Le problème consiste à modéliser un système innovant de temporisation hydraulique. Il s'agit de déterminer un modèle mathématique simple et robuste permettant de simuler le déplacement d'un piston en fonction du temps. Le piston délimite deux chambres remplies d'un fluide (eau). La chambre supérieure est soumise à une pression du réseau d'alimentation. La chambre inférieure étanche communique avec la chambre supérieure au moyen d'un orifice calibré. Le piston est soumis à l'action d'un ressort et à des forces de frottement. Le groupe projet qui va travailler sur le sujet doit déterminer les fonctions mathématiques du déplacement, de la vitesse et de l'accélération du piston en fonction du temps. Les données du modèle proposé pourront être confrontées à des résultats mesurés sur une maquette.

Sujet 3

- **Entreprise** : Thirard
- **Titre** : Conception automatisée d'un système de clefs
- **Descriptif** : Comment concevoir un système de clefs qui ouvrent (ou n'ouvrent pas) un ensemble de serrures sur un grand bâtiment ? La conception d'un système de clefs requiert de discriminer par un jeu de filtres les différentes possibilités et impossibilités d'ouverture demandées par le client. Les filtres obéissent à des lois de correspondance simples. L'objectif est de concevoir un algorithme qui optimise la conception du système de clefs, i.e. qui fournisse une réponse correspondant au besoin du client tout en minimisant le coût, i.e. le nombre de filtres à produire par usinage. L'algorithme conçu sera confronté à un jeu de données réelles.

Sujet 4

- **Entreprise** : Crédit Agricole Consumer Finance (CACF)
- **Titre** : Définition du risque de « défaut » des clients au crédit à la consommation

- **Descriptif** : Le Credit Scoring est devenu une composante essentielle des institutions financières. Il s'agit d'évaluer la probabilité d'un client de faire défaut à un prêt. Plusieurs approches ont été proposées dans la littérature parmi lesquelles la régression logistique associée à une définition des bons et mauvais clients et qui a été adoptée par Crédit Agricole Consumer Finance (CACF dans la suite). Cette méthode souffre de plusieurs défauts, parmi lesquelles une performance relativement médiocre, sans doute partiellement due à la définition bons/mauvais clients retenue. C'est sur cette définition que nous proposons d'orienter la réflexion dans ce sujet à destination de la Semaine d'Etude Maths-Entreprises.

Sujet 5

- **Entreprise** : Crédit Mutuel Nord Europe (CMNE)
- **Titre** : Comment tirer partie des données bigdata pour bâtir de nouveaux modèles prédictifs
- **Descriptif** : Le CMNE dispose de nouvelles sources de données issues de l'activité digitale de ses clients via une base DMP, de données de satisfaction issues de campagnes récurrentes (les réponses étant en partie du texte libre). Ce gisement de données peut être croisé avec les données transactionnelles classiques du système d'information (données socio-démo, contrats, souscriptions, flux financiers, données de contact multi canal). L'approche classique du monde bancaire s'appuie sur l'utilisation exclusive des données structurées dans une démarche figée par problématique (appétence produit, attrition) mono-produit, mono-canal. L'objectif est d'utiliser la totalité des données mises à disposition (structurées-non structurées) et de construire des modèles prédictifs permettant de définir les parcours clients les plus adaptés tenant compte de l'axe temporel des données recueillies et de l'aspect multicanal de la relation client.

Sujet 6

- **Entreprise** : COFIDIS
- **Titre** : Prévion d'un évènement rare à partir de signaux de détection faibles sur données clients
- **Descriptif** : L'objectif est de proposer des méthodes de prévision adaptées à la détection d'un évènement rare, qui minimisent le taux d'erreur de classement. En présence d'évènement rare, des méthodes de type recherche d'anomalies en industrie pourraient être envisagées pour élargir le champ des possibles par rapport à la méthode standard de la régression logistique couramment utilisée en matière de crédit scoring. Au besoin, nous mettons à disposition les résultats de l'analyse croisée entre les variables candidates et la variable cible ainsi que la base de données (avec données brutes et données travaillées) pour d'éventuels essais ou preuve de concept.

Sujet 6b

- **Entreprise** : COFIDIS
- **Titre** : Optimisation combinatoire de variables candidates à l'estimation d'un modèle de credit scoring
- **Descriptif** : L'objectif de cette étude est de proposer une méthode automatique de test des croisements optimaux entre variables dans le cadre d'une régression logistique à cible binaire. Les croisements issus de cette phase de la modélisation permettront par définition d'aller au-delà des habitudes et des pratiques intuitives ou empiriques actuellement mises en œuvre et ainsi d'améliorer la performance de prédiction par de nouveaux usages de croisements d'informations.

NB : Il est possible qu'un seul des sujets entre 6 et 6b ne soit ouvert.