

Ingénierie Mathématique  
Master mention Mathématiques et applications

UNIVERSITÉ PIERRE ET MARIE CURIE (PARIS 6)



Responsables : Edwige Godlewski  
Laboratoire Jacques-Louis Lions  
Barre 15-25 bureau 311  
Téléphone : 01 44 27 42 99  
Télécopie : 01 44 27 72 00  
Email : edwige.godlewski@upmc.fr

Marie Postel  
Laboratoire Jacques-Louis Lions  
Barre 15-25 bureau 313  
Téléphone : 01 44 27 54 08  
Télécopie : 01 44 27 72 00  
Email : marie.postel@upmc.fr

## **Résumés des stages en entreprise**

effectués par les étudiants du Master IM  
filière Mathématiques Pour l'Entreprise (MPE)

**année 2016-2017**

<http://www.ljll.math.upmc.fr/MPE>

Secrétariat : Barre 15-25 bureau 107, 4 place Jussieu, 75005 Paris.

Université Pierre et Marie Curie (Paris 6), BC 187, 4 place Jussieu, 75252 Paris cedex 05



Cette brochure constitue la treizième édition de la publication des résumés des stages en entreprise effectués par les étudiants de M2 du Master spécialité Ingénierie Mathématique (parcours MPE) de Sorbonne Université\*.

Elle fait suite à la publication pendant 16 années consécutives d'une brochure des résumés des stages en entreprise effectués par les étudiants du DESS de Mathématiques Appliquées. Cette formation a changé de nom depuis la rentrée universitaire 2004 à la faveur de la réforme LMD (Licence-Master-Doctorat). Elle est maintenant, sous le nom Ingénierie mathématique, majeure Mathématiques pour l'Entreprise (MPE), un parcours professionnel du Master Sciences et Technologies, mention Mathématiques et applications, de Sorbonne Université.

Ce document est destiné aux anciens ou futurs étudiants du Master Ingénierie mathématique, aux responsables de stages en entreprise et aux employeurs. Nous espérons qu'il illustrera, par la diversité et la qualité des sujets abordés, le dynamisme de notre formation et la variété des carrières auxquelles elle conduit.

Signalons qu'à la rentrée de septembre 2018, le parcours Ingénierie Mathématique sera ouvert à l'apprentissage. A cette occasion, il s'associe avec l'ISUP et proposera trois filières

- Ingénierie Financière et Modèles Aléatoires (IFMA)
- Ingénierie Mathématique Pour l'Entreprise (IMPE)
- Ingénierie Financière et Data Science (IFDS) †

Pour conclure, signalons les activités de l'Association des Anciens du MAster d'Ingénierie Mathématique, dont le site internet se trouve à l'adresse

<http://a2maim.free.fr/>

Cette association créée en 2006 à l'initiative d'anciens du DESS, favorise les échanges entre les étudiants des différentes promotions et contribue à leur insertion professionnelle, par exemple en diffusant des offres de stages et d'emploi.

---

\*. Cette nouvelle université résulte de la fusion de l'Université Paris Sorbonne-Paris 4 et de l'Université Pierre et Marie Curie - Paris 6. La formation Master IM est rattachée à la Faculté des Sciences et Ingénierie de SU

†. Anciennement parcours GRIE de l'ISUP

## **Responsables pédagogiques dans les différentes spécialisations en 2016-2017**

Probabilités - statistiques : V. Lemaire, L. Abbas-Turki

Laboratoire de Probabilités, Statistiques et Modélisation<sup>‡</sup>, site Jussieu, couloir 16-26 1<sup>er</sup> étage

Téléphone : 01 44 27 70 47, Télécopie : 01 44 27 72 23

Autres intervenants : J.-P. Baudry, E. Aidekon, N. Akakpo, O. Bardou, B. Guedj, T. Abdellatif, M. Thieullen,

Analyse numérique - calcul scientifique : E. Godlewski

Laboratoire Jacques-Louis Lions, site Jussieu, couloir 15-25, 3<sup>e</sup> étage

Téléphone : 01 44 27 42 99, Télécopie : 01 44 27 72 00

Autres intervenants : M. Cerf, B. Despres, P. Frey, F. Hecht, X. Juvigny, M. Postel, F.-X. Roux

Mécanique : P. Druault

Institut Jean Le Rond d'Alembert, site Jussieu, couloir 55-65

Téléphone : 01 44 27 37 90

Autres intervenants M. Abbas, E. Boyère, J. Delmas, J. Waeytens

Options :

C++ : F. Hecht

Code\_Saturne : R. Chakir

Excel VBA : J. Rameaux

GPU : L. Abbas-Turki

Python : N. Lantos

**Responsable des stages** : M. Postel

Marie.Postel@upmc.fr

Campus Jussieu, couloir 15-25, 3<sup>e</sup> étage, 313

Téléphone : 01 44 27 54 08

**Secrétariat du Master Ingénierie Mathématique** : F. Hardoyal

Francelise.Hardoyal@upmc.fr

Campus Jussieu, couloir 15-25, 1<sup>er</sup> étage, 107

Téléphone : 01 44 27 51 14

**Responsable du Master Ingénierie mathématique** : E. Godlewski

Edwige.Godlewski@upmc.fr

Campus Jussieu, couloir 15-25, 3<sup>e</sup> étage, 311

Téléphone : 01 44 27 42 99

---

<sup>‡</sup>. Fusion des laboratoires de "Probabilités et Modèles Aléatoires" et de "Statistique Théorique et Appliquée"

**ArianeGroup**  
Analyse de Mission\*  
**Responsable : Monsieur Max CERF**  
**Étudiant : Sélim ZIDI**

## **Sujet : Support à l'optimisation d'un scénario de phase balistique**

### **Résumé :**

Le service Analyse de Mission, dans lequel a été effectué ce stage, est impliqué dans la phase d'avant-projet du projet OneWeb. OneWeb est un projet d'envoi d'une constellation de 650 micro-satellites de télécommunication (150 kg) en orbite basse (1200 km d'altitude) contre 36 000 km pour les satellites géostationnaires. Ils ont pour but de pouvoir fournir des connexions internet haut débit pour tous et à moindre coût notamment dans les pays émergents et dans les zones rurales. En effet, les connexions par fibre ou par câble ont été conçues pour les pays riches mais ne sont pas rentables économiquement pour les pays en développement.

En phase d'avant-projet, OneWeb nécessite l'optimisation d'un grand nombre de paramètres afin de limiter le coût de cette mise en orbite complexe et les risques de collision entre satellites. L'objectif principal de ce stage est ainsi d'améliorer l'outil utilisé pour simuler et optimiser la mise à poste des constellations. En effet, un grand nombre d'optimiseurs sont disponibles au sein de la plate-forme de simulation et de nombreuses variables peuvent être prises en compte en tant que contraintes ou valeurs à optimiser. Il s'agit donc de déterminer les optimisations les plus pertinentes et, si possible, de fournir une procédure à suivre pour l'utilisateur de l'outil.

Ce stage s'est décomposé en plusieurs étapes. La première a été la prise en main des outils et du code industriel très complexe de la plate-forme d'optimisation. Il a ensuite fallu ajouter de nouvelles fonctionnalités notamment des vérifications de distance entre les satellites avant le lancement des optimisations afin de s'assurer que celles-ci peuvent aboutir à une solution réaliste. Ensuite, il s'est agi de définir des tests permettant de vérifier le bon fonctionnement des optimiseurs et d'analyser les résultats obtenus. Enfin, une dernière partie s'est articulée autour d'un code Monte-Carlo permettant de tester un grand nombre de trajectoires afin de s'assurer que les marges d'erreurs sont acceptables.

Cette contribution a permis de compléter certaines parties du code et de commencer le travail d'optimisation du scénario OneWeb. Les premiers cas tests mis en place vont pouvoir être réutilisés et complexifiés afin de s'approcher de plus en plus d'optimisations avec des données réelles à terme.

---

\*. Adresse de l'entreprise : 66 Route de Verneuil, 78130 Les Mureaux

## **Capgemini**

Pôle Insights and Data \*

**Responsables : Messieurs Jean-Baptiste NIEPCERON et Matthieu GUIET**

**Étudiante : Charlotte KERNY**

### **Sujet : Analyse du comportement de consommation de la clientèle d'Auchan Retail France**

#### **Résumé :**

J'ai effectué mon stage de 6 mois au sein de l'Entreprise Capgemini, première entreprise de services du numérique en France ainsi que le numéro six mondial de ce secteur en 2016. Capgemini étant une entreprise de prestations, j'ai travaillé chez un de ses clients, Auchan, au service Marketing Clients and Data au sein du département de la Direction de l'Innovation.

L'objectif de ce stage était d'analyser le comportement de consommation de la clientèle d'Auchan Retail France.

J'ai pu, lors de cette expérience, effectuer différentes missions :

- analyse de performance d'animations commerciales menées, ciblés (SMS, email, etc.) et non ciblés (tracts, radio, etc.) dans une "cross-canal" (auchan.fr, hyper, drive, auchan direct, simply, ...)
  - ★ visualiser et communiquer les résultats des analyses auprès du service Connaissance Client
  - ★ exploiter toutes les sources de données (on / off line, interne / externe), explorer de nouveaux usages, mener des analyses pour identifier de nouveaux axes ou leviers de croissance
- suivi du lancement d'un nouveau concept : la livraison à domicile
  - ★ analyser la performance
  - ★ créer un score d'appétence afin de détecter les clients susceptibles d'utiliser la livraison à domicile.

Ce stage m'a permis de me familiariser davantage avec les études statistiques et d'acquérir de nouvelles compétences avec la connaissance d'un nouveau langage de programmation le SQL, d'un nouveau logiciel Teradata et d'un nouveau domaine "le Marketing". Grâce à l'élaboration d'un score et au fait d'avoir suivi un Data Analyst, j'ai pu également monter en compétences sur le logiciel R studio. Cette expérience fut très enrichissante car j'ai pu appliquer le cours théorique suivi au premier semestre dans un contexte concret, travailler en équipe sur différents projets variés et conforter mon choix de continuer ma carrière dans le domaine de la Data au sein de cette Entreprise.

---

\*. Adresse de l'entreprise : 165 Avenue de Bretagne, 59000 Lille

**Crédit Agricole S.A.**

DRG/MRU/GRO \*

**Responsable : Monsieur Nicolas DAMAY**

**Étudiante : Marie-Suzanne SARRE**

**Sujet : Analyse Automatique de Données Textuelles**

**Résumé :**

Mon projet de fin d'études s'est déroulé au sein du Crédit Agricole S.A., organe central du Groupe Crédit Agricole. J'ai intégré la Direction Risque Groupe et rejoint l'équipe du Groupe de Recherche Opérationnelle.

L'objectif de mon stage était de mettre en place un algorithme de classification de textes.

Dans un premier temps, il a fallu faire un important travail de bibliographie. J'ai donc étudié l'état de l'art concernant notre problématique et cherché les modèles de classification qui s'appliquaient au mieux à notre sujet. Nous nous sommes principalement intéressés aux Réseaux de Neurones Récurrents. Nous avons également étudié des algorithmes de classification plus classiques tels que les K Plus Proches Voisins.

Ensuite, nous sommes passés à l'étape de modélisation des données. En effet, comme un algorithme peut seulement traiter des données numériques et que le travail se faisait sur des textes, nous avons donc dû donner une représentation vectorielle à nos documents. Pour cela, deux méthodes ont été utilisées : TD-IDF (term frequency-inverse document frequency), qui permet de transformer un texte en vecteur et Word2vec, qui crée des vecteurs pour chacun des mots du corpus.

Enfin, nous avons pu construire et appliquer les différents modèles retenus lors de l'état de l'art et les comparer. Les résultats obtenus sont très encourageants. L'algorithme le plus performant atteint un score proche de 96 %.

---

\*. Adresse de l'entreprise : 12 Place des Etats-Unis, 92120 Montrouge

**Dassault Systèmes**  
Service SIMULIA \*  
**Responsable : Monsieur Julien GEWINNER**  
**Étudiant : Arezki AIT MOULOUD**

**Sujet : Reconstruction de contraintes sur maillage**

**Résumé :**

J'ai réalisé mon stage de fin d'études au sein de la société Dassault Systèmes et plus précisément au sein de son département de simulation SIMULIA. Ce département ou « brand » a notamment pour but de fournir la meilleure qualité de maillage sur des pièces quelconques.

De manière générale, l'objectif de mon stage a été de produire un maillage de bonne qualité pour la simulation numérique en se basant sur un maillage déjà réalisé sur la pièce.

La difficulté était de retrouver des contraintes sur ce maillage pour pouvoir régénérer une meilleure discrétisation. Pour ce faire on a commencé par introduire des critères géométriques qui nous ont permis de retrouver des contraintes sur le maillage. Ensuite il a fallu construire ces contraintes avant de pouvoir régénérer un nouveau maillage.

---

\*. Adresse de l'entreprise : 10 Rue Marcel Dassault, 78140 Vélizy-Villacoublay



**Deepki**

Département développement produit \*

**Responsable : Monsieur Antoine WOLFF**

**Étudiant : Thomas FONTAINE**

**Sujet : Analyse et exploration statistique de données appliquées à la gestion des performances énergétiques de parcs bâtiments**

**Résumé :**

Dans l'analyse d'un parc de bâtiments, les moyens utilisés pour répondre aux problématiques d'efficacité énergétique sont souvent techniques (pose d'instruments de mesures comme des compteurs, ou encore audits énergétiques) ce qui peut être coûteux lorsque l'on s'intéresse à un parc de bâtiments de taille conséquente.

L'objectif de ce stage était de participer à une toute autre approche en développant et en appliquant des méthodes et algorithmes statistiques afin de répondre à ces questions d'efficacité énergétique.

Au sein de l'équipe développement produit, j'ai pu collaborer avec une équipe d'ingénieurs statisticiens et d'ingénieurs en efficacité énergétique afin de mettre en place des plans d'actions de réduction de consommations concrets qui faisaient suite aux différentes analyses statistiques que j'ai pu mener.

La question à laquelle j'ai du répondre tout au long de mon stage peut se formuler de la façon suivante :

Comment l'exploration statistique des données de consommations énergétiques peut aider au ciblage de sites problématiques et à la compréhension/caractérisation des problèmes ?

En premier lieu, j'ai participé à l'application de méthodes d'apprentissage statistique non supervisé pour la détection d'anomalies dans les consommations de chaque bâtiment.

La seconde partie de mon stage a consisté à utiliser des méthodes d'exploration statistiques de données afin de cibler les potentielles économies d'énergie.

Ce stage m'a permis de développer mes connaissances en statistiques et particulièrement celles autour des modèles d'apprentissage non supervisé. Mes compétences en développement informatique ont également été renforcées.

---

\*. Adresse de l'entreprise : 48 Rue René Clair, 75018 Paris

**EDF R&D**

LHSV \*

**Responsables : Madame Nicole GOUTAL et Monsieur Minh-Hoang LE**

**Étudiant : Virgile DUBOS**

**Sujet : Prise en compte de la végétation dans TELEMAC-2D à l'aide d'un modèle de porosité**

**Résumé :**

L'objectif du stage est d'intégrer et valider un modèle de porosité au sein du système TELEMAC-2D. Nous devons évaluer le plus rigoureusement possible les incertitudes de ce modèle par rapport aux autres modèles existants et aux données expérimentales.

La végétation sur les berges ou sur les plaines d'inondation d'une rivière ou dans le cours d'eau lui-même a une grande incidence sur la distribution horizontale et verticale de la vitesse et de la turbulence. Que ce soit pour simuler une crue ou des transferts de polluant dans la rivière, la prise en compte des effets de la végétation dans la simulation est capitale pour aboutir à une représentation / prévision optimale de l'écoulement. Du point de vue de l'opérationnel, le gain en précision de la prise en compte de la végétation ne doit pas se faire au détriment du temps de calcul.

Une représentation implicite de la végétation est ici testée et consiste à représenter la végétation par une porosité isotrope. L'écoulement est décrit par le modèle de Saint-Venant avec porosité et prend en compte la bathymétrie, les frottements et la force de traînée induite par la végétation. Ces équations sont résolues par la méthode des volumes finis centrés en cellules sur un maillage triangulaire non-structuré. Les termes faisant intervenir la porosité nécessitant des traitements numériques particuliers, nous adaptons un flux HLLC étendu utilisé pour les équations de Saint-Venant classiques.

Enfin, pour valider notre approche, un code de calcul a été écrit en Fortran. Les simulations se font avec un maillage "grossier" dont chaque élément est attribué à un niveau de porosité. Le modèle de Saint Venant avec porosité permet d'accélérer les simulations pour un résultat équivalent aux autres approches implicites. Davantage de cas tests sont nécessaires avant de pouvoir intégrer ce code au système TELEMAC mais l'approche avec porosité est prometteuse.

---

\*. Adresse de l'entreprise : 6 Quai Watier, 78400 Chatou

**EDF R&D**

Département PRISME \*

**Responsable : Monsieur Sofiane HENDILI****Étudiant : Pablo PEREIRA ÁLVAREZ****Sujet : Méthode d'hyper-réduction pour la simulation numérique du soudage****Résumé :**

La simulation numérique du soudage est un problème complexe. En effet, ce type de simulation doit prendre en compte un nombre important de paramètres, des phénomènes thermomécaniques non-linéaires et des maillages de grandes tailles, avec un raffinement important dans les zones soudées. Par conséquent, la résolution du problème thermomécanique non linéaire du soudage nécessite un temps de calcul élevé.

Des études paramétriques sont souvent réalisées dans le cadre de la simulation numérique du soudage. Les temps de calcul rencontrés en simulation numérique du soudage rendent ces études paramétriques très coûteuses voire irréalisables. D'autre part, dans certains cas industriels, comme ceux rencontrés dans le projet SPAR-2, le procédé de soudage induit un nombre important de passes de soudage. Ce type procédé est appelé le soudage multi-passes.

On propose une méthode, dite d'hyper-réduction, qui permet de réduire les temps de calcul de la simulation numérique du soudage, mais ayant toujours des résultats qui nous permettent de prédire efficacement les défauts du soudage. Dans le cadre du projet SPAR-2, une thèse a été lancée en partenariat avec l'École de Mines de Paris et le département ERMES d'EDF R&D pour étudier cette méthode et ses possibles applications industrielles.

Le but de ce stage était de tester la méthode d'hyper-réduction appliquée à la simulation numérique du soudage sur plusieurs cas, en particulier pour le soudage multipasses.

---

\*. Adresse de l'entreprise : 6 Quai Watier, 78400 Chatou

**EDF Énergies Nouvelles**  
Bureau d'Études Éolien \*  
**Responsable : Monsieur Clément RUMEN**  
**Étudiant : Réda YAHOU**

**Sujet : Étude de l'impact des bridages acoustiques sur la fatigue des composants structurels des éoliennes**

**Résumé :**

Mon stage a eu lieu au sein de l'équipe calcul de chargements du Bureau d'Étude. Actuellement, une éolienne est certifiée 20 ans lors de sa conception et dans des conditions de fonctionnement standards définis par la norme industrielle. Cependant, les études récentes ont montré que les nombres d'arrêts machine étaient plus importants que ceux pris en compte dans les calculs standards. C'est particulièrement le cas des arrêts acoustiques, arrêts qui ont été mis en place afin de réduire le bruit subit par les riverains habitants près des parcs éoliens, dans la mesure où ces structures émettent beaucoup de bruit sous certaines conditions et dans le respect des normes acoustiques imposées par la loi. Des effets de fatigue provoqués par ces arrêts ont été observés au niveau des composants structurels des éoliennes. Ces effets sont significatifs quant à l'estimation de la durée de vie.

Le but du stage a donc été la prise en compte de ces nombres d'arrêts réels dans le calcul de chargement habituel, et l'analyse de l'impact résultant sur la fatigue des composants structurels et sur l'estimation de durée de vie. Pour ce faire, il a fallu définir une caractérisation mécanique de ces arrêts et établir un moyen de les extraire, afin de les inclure dans l'analyse finale de fatigue.

Dans un premier temps a eu lieu une phase de documentation, dans laquelle j'ai pu me familiariser avec la théorie aérodynamique des éoliennes, la mécanique de la fatigue et le processus et les outils du calcul de chargement éolien. Par la suite, les données nécessaires pour l'étude ont été extraites via un système de télégestion stockant tous les paramètres des éoliennes dans chaque parc, puis celles-ci ont été filtrées selon les besoins grâce à un code que j'ai implémenté en VBA. Ce code a notamment permis de compter le nombre d'arrêts réel et de retourner les inputs nécessaires pour l'analyse de fatigue. Le calcul de chargements a été réalisé sur BLADED, un code aéro-élastique permettant la conception des éoliennes et la simulation de leur comportement mécanique. La comparaison des résultats de l'étude réalisée avec ceux de l'étude classique (nombre de stops standards) a révélé que les éoliennes subissaient bien plus de fatigue dans le cas du nombre d'arrêts réel. Les composants les plus impactés sont en particulier le moyeu, la tour et les pales. Les résultats obtenus sont venus confirmer les prédictions faites par l'équipe avant le début du stage.

---

\*. Adresse de l'entreprise : 100 Esplanade Général de Gaulle, Coeur Défense, Tour B, 92400 COURBEVOIE

**Enersun Technology**

Département d'application de données \*

**Responsable : Monsieur Xiaobo ZHANG****Étudiant : Xiaofan LI****Sujet : Intégration de données****Résumé :**

1. Intégration de données hétérogènes multidimensionnelles.
2. Avertissement de foudre pour les tours de transmission.
3. Prédiction de l'épaisseur de givrage des lignes de transmission.
4. Prévisions de charge d'électricité.

Résumé : La première étape de tous les projets est la collecte de données, après la première phase d'acquisition et d'intégration des données, j'ai appris à utiliser PL SQL simplement. Ensuite, selon les données obtenues, commence la phase d'intégration des données, qui nécessite leur exploration et la compréhension globale de leur structure. Arrive ensuite la phase de sélection d'un algorithme d'apprentissage automatique adapté aux besoins exprimés dans le projet, et également à la nature des données. C'est également à ce stade qu'on choisit d'utiliser R ou Python pour la programmation algorithmique. Il faut en dernier lieu tester les sorties, et compléter le rapport de projet. Le troisième projet traité pendant ce stage m'a permis d'apprendre à utiliser un algorithme de réseau neuronal. Pour cela les données non supervisées ont d'abord dû être transformées en données supervisées. Par ailleurs les choix du réseau neuronal, de sa largeur et de sa profondeur, ont nécessité beaucoup de recherches. Enfin le quatrième et dernier projet a duré plus longtemps que prévu, en raison de mauvaises prédictions initiales, et de la nécessité de consulter spécifiquement les experts métiers pour obtenir les caractéristiques d'entrée, et concevoir une prévision de roulement, avant d'obtenir des résultats relativement fiables.

**Remarques :**

Ying REN : Responsable de groupe d'analyse de données

Jian QIU : Partenaire de Alibaba Cloud (projet Prévisions de charge d'électricité)

---

\*. Adresse de l'entreprise : Kunming Enersun Technology Co. Ltd, Kunming, Yunnan, China

**Eukleia**

ERP Software Solution \*

**Responsable : Monsieur Olivier VERDENET****Étudiant : Sofiane HADDADI****Sujet : Construction d'un modèle de prévision****Résumé :**

L'objectif de ce stage était de construire un modèle de prévision de l'activité de vente des produits pétroliers pour un des clients potentiels d'Eukleia spécialisé dans le domaine de l'énergie.

Au début de mon stage, j'ai travaillé sur la conception générale de ce modèle de prévision et j'ai présenté mon travail au client.

Dans la deuxième partie de mon stage, j'ai étudié les travaux réalisés par d'autres chercheurs sur ce sujet, dans le domaine de l'énergie, et j'ai testé différents outils d'intelligence artificielle afin de trouver celui ou ceux qui seraient les plus appropriés au contexte du client. A la fin de cette partie, j'ai choisi la méthode Machines Vecteurs Support Régression (SVR) comme méthode d'intelligence artificielle. Ce choix est validé par mon tuteur de stage.

La dernière partie de mon stage a consisté à mettre en œuvre la solution SVR dans le langage choisi et à faire des tests sur les données réelles.

---

\*. Adresse de l'entreprise : 12 Boulevard Poissonnière 75009 Paris.

**Fotonower**

Service Recherche et Développement \*

**Responsable : Monsieur Victor REUTENAUER**

**Étudiant : Wei GAO**

**Sujet : Utilisation du deep learning pour la classification de média et autres applications****Résumé :**

L'objectif de ce stage a été d'utiliser les outils de reconnaissance d'images pour développer les modèles selon notre besoin.

Nous avons 3 types de besoin : classier les images en différents clusters, détecter un objet dans une image, prédire une valeur pour une image.

Nous avons utilisé des réseaux de neurones pour atteindre les résultats souhaités.

Dans un premier temps, un important travail de bibliographie à été effectué pour comprendre les différentes architectures de réseau de neurones. Il a fallu choisir un outil qui nous permettait d'entraîner les modèles avec les architectures bien choisies. Suite à ce travail, nous avons choisi d'utiliser l'outil *Caffe*. Dans un deuxième temps, nous avons calibré les différents modèles de réseau de neurones déjà existants pour les adapter à nos données et à nos besoins.

Pendant mon stage, j'ai réalisé plusieurs projets, et plus particulièrement sur les traitements des photos de voitures.

Les projets sont : classier les photos de voitures par modèle, classier les photos de voiture par orientation, détecter la plaque d'immatriculation d'une voiture, déterminer la perte totale de voiture, prédire le prix de réparation d'une voiture.

Pour la classification, nous avons utilisé les réseaux de neurones de type *AlexNet*, *VGG*, *ResNet*.

Pour la détection, nous avons utilisé l'architecture de réseau de neurones *Faster R-CNN*.

Nous avons utilisé une architecture *LSTM* (Long short-term memory) pour classier une série d'images ordonnées.

**MOTS-CLÉS :** Réseau de neurones, *CAFFE*, *Pycaffe*, *AlexNet*, *VGGnet*, *ResNet*, *FASTER R-CNN*, *LSTM*, Réseau de neurones récurrent, Apprentissage par transfert

---

\*. Adresse de l'entreprise : 30 Rue Charlot 75003 Paris

**INRIA (Institut National de Recherche en Informatique et en Automatique)**  
Équipe ACUMES \*  
**Responsables : Messieurs Régis DUVIGNEAU et Abderrahmane HABBAL**  
**Étudiant : Noam ABETTAN**

**Sujet : Optimisation aérodynamique d'un véhicule de course de type Formula-E**

**Résumé :**

J'ai effectué mon stage au sein de l'Équipe-Projet ACUMES du centre de Recherche INRIA Sophia-Antipolis. Les activités de cette équipe concernent la modélisation et l'optimisation de systèmes complexes régis par des équations aux dérivées partielles, avec notamment des applications en mécanique des fluides, des structures, au trafic routier et piétonnier.

L'objectif de mon stage était de réaliser une étude d'optimisation aérodynamique du système de refroidissement d'un véhicule de course de type Formula-E, dans le cadre d'une collaboration avec l'écurie monégasque Venturi. La première partie a consisté en la simulation effective de l'écoulement de l'air au sein de l'écope de refroidissement. Après quelques résultats sur le modèle de conduit 3D, on a pu s'apercevoir de la complexité de l'écoulement et donc on a ré-envisagé une approche 2D.

Il a fallu par la suite mettre en oeuvre toute la chaîne de calcul (CAO - maillage - simulation - optimisation) développée par l'équipe de recherche. Plus précisément, il a fallu construire plusieurs maillages non-structurés 2D à partir des données CAO du véhicule. J'ai ensuite implémenté un modèle aéro-thermique de radiateur dans la plate-forme de simulation d'écoulements turbulents (<http://num3sis.inria.fr>). Enfin, on a effectué une analyse de sensibilité vis-à-vis de différents paramètres, avant de réaliser une optimisation complète de la forme du conduit de refroidissement et d'un système de déflecteurs à l'aide d'une méthode à base d'apprentissage (surface de réponse). Le critère d'optimisation principal étant de maximiser le débit dans le conduit de refroidissement.

Les travaux de ce stage constituent un point de départ pour une approche 3D puisqu'ils nous ont permis de comprendre quels étaient les paramètres influents sur l'augmentation du débit dans l'écope.

---

\*. Adresse de l'entreprise : INRIA Sophia Antipolis Méditerranée, 2004 Route des Lucioles, 06902 Sophia Antipolis



**My Money Bank**

Département Risque \*

**Responsable : Monsieur Abdessalem ABDERRAHIM****Étudiant : Hai JING****Sujet : Chargé d'Etudes Statistiques Risque****Résumé :**

Le principal risque de portefeuille des prêts du secteur bancaire est la défaillance possible des emprunteurs. Du point de vue réglementaire, chaque institution financière doit préparer différentes méthodes applicables en fonction des circonstances. Il s'agit de demander aux banques non seulement de contrôler la performance actuelle du portefeuille, mais aussi d'envisager les effets sur les différents scénarios possibles dans le futur.

Dans ce contexte, nous avons étudié les deux types de processus de gestion du risque pour répondre aux demandes : le Triggers Reporting et le stress test, deuxième point que nous avons développé en détail dans le rapport.

Nous avons développé deux modèles de stress testing, le Modèle à Correction d'erreurs (MCE) et le modèle de Credit-PortfolioView (MCPV). A partir de ces modèles, la probabilité de défaut de portefeuille est simulée sur deux scénarios macro-économiques. Le scénario Baseline consiste à prévoir les conditions macroéconomiques futures les plus probables. Le scénario Adverse consiste à prévoir la probabilité de défaut en cas de situation défavorable comme la crise économique de 2008.

**MOTS-CLÉS :** Risque de crédit, Probabilité de défaut, Stress Testing, Stress scénario, Trigger Reporting, Modèle à Correction d'erreurs, Modèle de CreditPortfolioView

---

\*. Adresse de l'entreprise : Tour Europlaza - 20 avenue André Prothin - 92063 La Défense

**ONERA**

Département de Traitement de l'Information et Systèmes (DTIS) \*

**Responsables : Madame Juliette RYAN et Monsieur Ludomir OTESKI**

**Étudiant : Daniel VANZO**

**Sujet : Calcul haute performance pour l'optique adaptative****Résumé :**

L'unité Calcul Haute Performance (CHP) dans laquelle j'ai effectué mon stage, a pour mission de développer des algorithmes et des codes haute performance en collaboration avec les autres départements de l'ONERA et les partenaires comme l'Institut d'Optique Graduate School (IOGS). L'objectif de ce stage a été de développer un code de calcul sur GPU pour le problème de reconstruction de la commande en Optique Adaptative (OA) en collaboration avec l'IOGS.

Ces travaux ont pour but d'être appliqués à l'European Extremely Large Telescope (E-ELT), télescope actuellement (2017) en construction au Chili. L'OA pour les télescopes consiste à déformer un miroir pour compenser les défauts de l'image dûs aux turbulences atmosphériques. Il en résulte une meilleure résolution des images reçues par le télescope. Le problème de reconstruction de la commande consiste à résoudre un système linéaire creux de grande taille (jusqu'à  $3\,960\,100 \times 28\,868$ ) en temps limité ( $< 1$  ms).

Les GPU, de part leur architecture, sont adaptés aux calculs massivement parallèles et ont été choisis pour faire les calculs de l'OA. Les travaux du stage ont consisté à développer et tester différents algorithmes de résolution du système linéaire sur GPU afin de mettre en oeuvre un algorithme efficace.

Le développement a été fait avec l'API CUDA et les bibliothèques CUBLAS, CUSPARSE et CUSOLVER.

Des résultats satisfaisants ont été obtenus avec la carte NVIDIA Tesla P100 pour des tailles de matrice "petites" (jusqu'à  $640\,000 \times 4\,744$ ), avec des temps de calcul inférieurs à 1 ms.

Pour les tailles de matrices les plus grandes ( $3\,960\,100 \times 28\,868$ ) des temps de calcul de 8,2 à 24,9 ms ont été obtenus. Bien qu'encore supérieur à 1 ms ces résultats sont encourageants car notablement plus bas que les temps de calcul avec la méthode directe ( $> 1016$  ms).

---

\*. Adresse de l'entreprise : 8 Chemin de la Hunière et des Joncherettes , Palaiseau

## **Orange Labs**

IMT/OLN/RNM/WTE/MADE \*

**Responsable : Monsieur Mohamed Kadhém KARRAY**

**Étudiante : Erlinda MENÉNDEZ PÉREZ**

**Sujet : Étude de la distribution de QoS entre les différentes localisations des utilisateurs dans les réseaux cellulaires 4G**

### **Résumé :**

Un réseau cellulaire est constitué d'un ensemble de stations de base qui assurent chacune le service dans une certaine zone de couverture appelée cellule. On a considéré un modèle où la localisation des stations de base est un processus ponctuel de Poisson sur le plan.

L'objectif du stage est d'étudier la distribution de la qualité du service (QoS) perçue par les utilisateurs dans les différentes localisations de ceux-ci dans le réseau. La QoS dans une localisation donnée est liée aux paramètres radio tels que l'affaiblissement de propagation, la puissance totale reçue, l'interférence et le rapport interférence sur signal (ISR).

Le stage a comporté plusieurs étapes. Dans un premier temps, on a étudié l'effet de la hauteur de l'équipement de l'utilisateur sur les moyennes spatiales des différents paramètres radio grâce aux résultats de la théorie de la géométrie stochastique. Dans ce cadre on a établi des expressions analytiques de ces moyennes dans deux cas particuliers : lorsqu'il n'y a pas de shadowing et lorsque la hauteur de l'équipement d'utilisateur et de l'antenne de la station de base sont égales. La seconde étape a consisté à étudier la distribution spatiale de la capacité du lien et la capacité cellulaire, au sens de la théorie de l'information, d'abord dans le cas où il n'y a que du bruit ou que de l'interférence et ensuite dans le cas où il y a simultanément ces deux effets. Finalement, on a pris en compte la dynamique des arrivées et des départs des utilisateurs grâce à la théorie des files d'attente. Cela nous a conduits à étudier l'équation de charge dans le cas du bruit blanc gaussien additif (AWGN) et dans le cas MIMO (antennes multiples à l'émission et à la réception). La résolution de l'équation de charge nous a permis de calculer la distribution des débits dans les différentes localisations des utilisateurs dans le réseau. Le résultat le plus important et le plus innovant de ce stage est de caractériser explicitement l'impact du nombre d'antennes sur la distribution du débit de l'utilisateur entre les différentes localisations dans le réseau.

Ce stage m'a permis d'acquérir des connaissances dans le domaine des télécommunications et en particulier dans l'application de la statistique et des probabilités dans ce secteur. L'étude bibliographique tout au long de ce stage m'a permis d'approfondir mes connaissances en théorie des files d'attente et en géométrie stochastique et de découvrir la théorie de l'information. D'autre part, il m'a permis de voir comment articuler les trois théories de l'information, des files d'attente et de géométrie aléatoire pour résoudre le problème d'ingénierie considéré, de développer mes compétences en programmation Matlab et en travail d'équipe. Enfin, la découverte de la vie en entreprise est passée par la participation aux réunions d'équipe où j'ai fait une présentation de certains résultats et à un événement organisé en l'honneur des stagiaires par la direction de RNM (Radio Access Networks and Microwaves).

---

\*. Adresse de l'entreprise : 44 Avenue de la République, 92320 Châtillon

## **SNCF Transilien**

Département Tarification Vente Validation (TVV) \*

**Responsables : Madame Nathalie ALHINC et Monsieur Fabien FRIMONT**

**Étudiante : Claire LE MEHAUTE**

**Sujet : Analyse de données en vue de définir les critères de performance des automates de vente des titres de transport Transilien**

### **Résumé :**

SNCF (Société Nationale des Chemins de Fer Français) est une entreprise ferroviaire publique française. C'est l'un des premiers groupes mondiaux de mobilité de personnes.

Transilien est une des branches de SNCF Voyageurs qui gère une partie du transport en Île-de-France.

J'ai effectué mon stage au sein de la division Excellence Opérationnelle du département TVV. Ce département gère et développe les services liés à la tarification, la vente des titres de transport et leur validation. Ses principales activités sont la création de nouveaux services et l'exploitation du système et des services Vente et Validation. Au sein de ces services, se trouvent d'une part les outils de vente appelés ART (Automate Rapide Transilien) et d'autre part, les bornes de validation appelées CAB (Contrôle Automatique de Billets).

J'ai réalisé les missions de stage au sein du projet KPI-ART (Key Performance Indicator - ART). Ce projet a été créé afin de définir les indicateurs de performance des automates et de mettre en place les plans d'action pour améliorer leur performance.

L'objectif de mon stage était d'améliorer la performance des automates de vente Transilien et de proposer des indicateurs pour suivre l'état de fonctionnement de ces automates.

J'ai analysé les données techniques des ART pour définir les critères de performance de ces automates, détecter leurs dysfonctionnements et améliorer leurs taux de disponibilité. Pour cela, j'ai mis en exergue la présence ou l'absence de corrélations entre certaines données techniques des automates (le nombre de redémarrages et le nombre de transactions par exemple).

J'ai également déterminé les facteurs qui influencent le nombre de redémarrages des ART en établissant un modèle statistique. Je suis à l'origine de la création de nouveaux tableaux de bord sur une plate-forme d'analyse et de visualisation de données appelée Kibana et ayant pour but de mieux comprendre le fonctionnement des ART.

Au cours de ce stage, je me suis appuyée sur les outils suivants :

- La méthode de régression linéaire que j'ai utilisée à l'aide du logiciel statistique R.
- Les outils internes SNCF m'ont également servi.
  - Logstash via Elasticsearch pour Kibana
  - La lecture pure des logs dans Transfolio, outil de supervision en temps réel des automates de vente
- Les données chiffrées sur le taux de disponibilité des ART et sur la satisfaction clientèle m'ont été fournies par les rapports internes appelés Etoile et S Mesure.

---

\*. Adresse de l'entreprise : 34 Rue du Commandant Mouchotte, 75014 PARIS

## **Weborama**

Activation d'Audience Programmatique \*

**Responsable : Monsieur Thomas ALLEMAND**

**Étudiant : Samuel BERGANZ**

### **Sujet : Assistant Trader Média et Data Analyst**

#### **Résumé :**

Créée en 1998, Weborama est une « **data company** » spécialiste de l'optimisation des actions marketing et média par la data. L'entreprise, d'origine française, conçoit et commercialise des actifs 100% propriétaires, notamment une Data Management Platform (DMP) ainsi qu'une base de données de profils d'internautes. Weborama opère en Europe de l'Ouest et de l'Est, en Amérique du nord et du sud. Le Groupe compte 250 collaborateurs entre Paris, Londres, Moscou, Madrid, Milan, Lisbonne, Amsterdam et Mexico.

C'est dans ce contexte que j'ai rejoint l'équipe d'Activation d'Audience Programmatique pour effectuer mon stage de fin d'études en tant qu'Assistant Trader Média et Data Analyst. Le marketing digital consiste à réaliser un certain nombre d'actions de conquêtes de clients et de fidélisation de ces derniers. Le "programmatisation" est quant à lui un ensemble de procédés algorithmiques, car à grande échelle, d'achats d'espaces publicitaires ciblés. Ces méthodes d'achats se font via des technologies d'achats, appelées D.S.P. (Demand Side Platform), qui se connectent sur des places de marchés virtuelles (les Ad-Exchanges) basées sur les mêmes principes que les places de marchés financiers, et diffusent des impressions ciblées, à la bonne personne, et avec la bonne publicité. L'équipe que j'ai rejointe est ainsi composée de Data Activation Managers qui se chargent d'actions destinées à la conquête de prospects et de fidélisation des clients, par l'activation des audiences cibles et des procédés automatisés d'achats d'espaces publicitaires.

La campagne publicitaire a ainsi pour but de marquer l'esprit de l'internaute en faisant la promotion d'une marque, d'un produit ou d'un service, et éventuellement de l'inciter à acheter en ligne le produit ciblé par la publicité délivrée. La campagne est basée à la fois sur les données clients qui sont gérées dans Weborama Audience Manager, la Data Management Platform de Weborama ainsi que la base de données propriétaire (200 millions de profils). Une fois la segmentation d'audience faite, les segments d'audience sont transférés vers Appnexus l'une des DSP utilisées par Weborama.

En tant qu'assistant trader média, j'ai ainsi dû découvrir le monde du programmatisation, m'habituer à l'ensemble du vocabulaire de ce nouvel écosystème, participer à des réunions avec nos partenaires, comprendre le fonctionnement des différentes plateformes et les méthodes d'activation d'audiences. J'ai ensuite pu gérer un portefeuille d'annonceurs via la mise en place de campagnes au travers des étapes suivantes : répondre aux "briefs" du client, mettre en place les scénarios de campagne, suivre la bonne diffusion des dispositifs publicitaires médias dans l'objectif d'une bonne diffusion visant à la satisfaction du client, et créer des bilans de fin de campagne. Une fois l'univers du programmatisation mieux appréhendé, cela m'a grandement aidé dans mes missions de Data Analyst.

En tant que Data Analyst, j'ai pour charge d'effectuer des analyses et méta analyses permettant d'améliorer les statistiques de nos campagnes publicitaires. Cela a nécessité, dans un premier temps, la bonne maîtrise des reportings et des nombreuses métriques liées à nos campagnes, qui représentent toutes les informations. A partir de cela, il est possible d'utiliser différents outils statistiques (dendogrammes, corrélations, ACP, Data Visualisations) dans un but d'amélioration. De plus, j'ai également travaillé sur un projet VBA, qui a été la traduction opérationnelle du travail de l'équipe d'Activation d'Audience Programmatique, pour automatiser différents tris de reporting et ainsi obtenir un gain de temps important pour l'ensemble des membres de mon équipe.

---

\*. Adresse de l'entreprise : 37 Boulevard des Capucines 75 002 Paris

## Les étudiants de l'année 2016-2017

<b>Nom</b>	<b>Prénom</b>	<b>Entreprise</b>	<b>Parcours</b>	<b>Page</b>
ABETTAN	Noam	INRIA	m	14
AIT MOULOU	Arezki	Dassault Systèmes	m	6
BERGANZ	Samuel	Weborama	m	19
DUBOS	Virgile	EDF R&D	m	8
FONTAINE	Thomas	Deepki	m	7
GAO	Wei	Fotonower	m	13
HADDADI	Sofiane	Eukleia	ps	12
JING	Hai	My Money Bank	ps	15
KERNY	Charlotte	Capgemini	ps	4
LE MEHAUTE	Claire	SNCF	ps	18
LI	Xiaofan	Enersun	ps	11
MENÉNDEZ PÉREZ	Erlinda	Orange Labs	ps	17
PEREIRA ÁLVAREZ	Pablo	EDF R&D	m	9
SARRE	Marie	Crédit Agricole S.A.	ps	5
VANZO	Daniel	ONERA	m	16
YAHOU	Reda	EDF Energie Nouvelle	m	10
ZIDI	Selim	Ariane Group	ps	3

Les résumés sont classés dans la brochure par ordre alphabétique de l'entreprise.

Sigles des parcours :

MPE ps : Mathématiques Pour l'Entreprise - analyse numérique, calcul scientifique et probabilités, statistiques

MPE m : Mathématiques Pour l'Entreprise - analyse numérique, calcul scientifique et mécanique