



# IS

## Data Scientist

### SYNTHESE DE COMPETENCES

#### ❖ Compétences techniques :

- Python • R • SQL • Scala • Tableau • Excel • IBM Cloud (NLP)
- Apache Spark • PHP • Javascript • HTML/CSS • Pytorch • Tensorflow

#### ❖ Connaissances fonctionnelles :

- Environnement de R&D • Environnement de Recherche en laboratoire • Secteur de la Santé
- Secteur de l'environnement

#### ❖ Langues :

- Anglais : bilingue
- Espagnol: notions

### EXPERIENCES PROFESSIONNELLES

#### 1<sup>er</sup> sept. 2018 – 1<sup>er</sup> mars 2019 Data Scientist, Poclain

##### Objectif :

- Recherche de données en OpenData dans un but d'analyse et de croisement avec le dataset client pour y trouver corrélations et insights d'intérêts.
- Analyses et application de Machine Learning sur le dataset client afin de répondre à la problématique donnée.
- Processus itératif incrémental sur les données clientes et OpenData.
- Application de différents algorithmes de régression (Decision Tree, Random Forest Regressor, Linear Regressor, Logistic regression) et de classification (Linear Classification, Decision Tree, Random Forest Classifier).
- Usage d'un modèle statistique ARIMA utilisé pour l'analyse et la prédiction de TimeSeries.





## Environnement technique :

Python, ML Libraries (Numpy, Scikit learn, Pandas, Matplotlib, Seaborn etc...), Spark ML, Tableau

## 2 janv. 2018 – 26 juin 2018, Ingénieur en Machine Learning, Pacte Novation, Issy-les-Moulineaux

### 1<sup>er</sup> Objectif :

Création d'une plateforme de E-commerce utilisant du Machine Learning et d'autres technologies de l'Intelligence Artificielle pour permettre une meilleure expérience client et fournir une recommandation ciblée de produits.

- Création d'une bibliographie et étude de l'existant parmi les grands noms du E-commerce mondial.
- Prototypage des solutions existantes de Machine Learning du marché (IBM Cloud, Facebook ML, Microsoft Azure, AWS, etc).
- Etude de marché sur ces dernières et choix de la solution la plus adéquate aux besoins du projet.
- Arbitrage et test des différentes technologies les plus adéquates à la création du Front-End et Back-End le plus adapté aux demandes du projet.
- Analyse des différents algorithmes visant à fournir de la recommandation de produits aux différents clients et de permettre une meilleure compréhension du client.
- Début du développement du Front-End de la plateforme via Javascript, HTML/CSS et Angular pour la mise en place d'un agent conversationnel de type chatbot. Ainsi que du développement du Back-End de la plateforme via PHP relié à une base de données MySQL.
- Usage d'un algorithme de Collaborative Filtering dans le but de permettre la recommandation de produits et la compréhension du client.
- Réalisation d'un système de recommandation de produits statiques à l'aide d'Apache Mahout et de JAVA.

## Environnement technique :

HTML/CSS, JavaScript, Angular, PHP, MySQL, Apache Mahout

Projet de R&D. Environnement de développement

### 2<sup>nd</sup> Objectif :

Mission chez le client, Mana\_Vox (ONG luttant contre la déforestation), dans l'optique d'injecter du Machine Learning et user de technologies de Data science dans leur algorithme fonctionnel. Le but étant d'optimiser sa performance et sa précision ainsi que de réduire la volumétrie totale fournie en input dans cet algorithme.

- Analyse de l'existant et mise en place des axes d'action et de travail.
- Transformation de la problématique de cette entreprise en problématique adaptée à un projet de data science.





- Analyse des services de l'IBM Cloud dans l'objectif de voir les plus à même de profiter au projet.
- Enrichissement de l'algorithme avec différents services de l'IBM Cloud tels qu'IBM Watson Machine Learning, IBM Watson NLU, et l'IBM Watson Knowledge Studio.
- Mise en place d'une ébauche de module de Machine Learning visant à fournir un indice de pertinence via Python et ses différentes librairies adaptées au Data Science telles que Numpy, Pandas, Matplotlib, Seaborn et Scikit-learn.
- Réflexion sur la potentielle utilisation d'Apache Spark dans le contexte de projet du client.

#### **Environnement technique :**

- IBM Watson Machine Learning.
- IBM Watson Natural Language Understanding (NLU).
- IBM Watson Knowledge Studio.
- Python (+DataScience libs: sk-learn, numpy, pandas, matplotlib)
- Algorithme de clustering
- Apache Spark.

### **Avril. 2017 – juin 2017, Chercheur stagiaire, Laboratoire Sanpsy, CHU Pellegrin, Bordeaux, France**

#### Objectif :

Travail de recherche avec l'équipe CENS du laboratoire SANPSY, sous la direction du Dr. Ellemarije Altena. Le but étant d'étudier la réactivité émotionnelle, mesurée par les expressions faciales, en réponse à des stimuli émotionnels chez des patients souffrant d'insomnie chronique ainsi que des contrôles pour observer une potentielle différence.

- Inclusion et exclusion de patients
- Élaboration de la tâche de reconnaissance faciale à l'aide du logiciel Affectiva.
- Traitement et analyse des données des patients via des tests statistiques
- Test de simulateurs de conduite ainsi que vol pour étudier la représentation des symptômes des patients dans des conditions de stress données.
- Test d'Humains virtuels (docteur virtuel et patient dépressif virtuel, programmé via Unity) dans le but de vérifier la validité de leurs scénarios ainsi que leur efficacité.
- Mise en place de scénarios dans la CAVE (chambre de réalité virtuelle pour pallier aux phobies des patients à l'aide de nouvelles technologies immersives.

#### **Environnement technique :**





- Réalité Virtuelle
- Machine Learning (Decision Trees via Unity)
- Affectiva (logiciel de reconnaissance faciale)
- Analyse statistique (R et Excel)

### **Environnement fonctionnel :**

- Recherche en laboratoire

## **Mai 2016 – juin 2016, Chercheur Stagiaire, Inserm, Région de Bordeaux, France**

### Objectif :

Stage dans un laboratoire spécialisé dans les neurosciences dirigé par monsieur Stéphane Olié à l'INSERM, mai 2016 consistant en un travail de recherche sur la sclérose en plaques.

- Étude des symptômes sur des animaux modèles
- Mise en évidence de l'altération de la gaine de myéline des neurones.
- Quantification des neurones altérés.

## **FORMATION/ CERTIFICATIONS**

### **❖ 2016-2018, Master en Sciences Cognitives, Université de Bordeaux, France**

Projets de réalité virtuelle, réalité augmentée, neurosciences computationnelles ainsi qu'un projet de type Deep Learning avec la réalisation d'un réseau de neurones (réseau de Hopfield) en Python, sa modélisation, l'analyse de ses performances, l'usure naturelle dans ses performances et ensuite modélisation d'une pathologie dégénérative pour avoir un aperçu des pertes de performances du cerveau humain dans des conditions pathologiques.

### **❖ 2014-2016, Licence de Biologie, Université de Bordeaux, France**

Parcours Biologie moléculaire, cellulaire et physiologie.





❖ **2009-2010, Cambridge Proficiency Exam (C2 Level) ESOL examinations, University of Westminster, London, England.**

Diplôme qui à ce jour est le plus haut dans la totalité des examens ESOL de l'université de Cambridge. Diplôme de niveau C2 attestant de ma fluidité en Anglais et de mon bilinguisme dans cette langue.

