



Nadia OUHSSAINE

Data Scientist

🏠 1 rue Léa Maury
94380 Bonneuil-sur-Marne
☎ 06.50.31.79.66
✉ ouhssaine.nadia@gmail.com

👤 Projets académiques

- Estimation du coefficient de consanguinité sur des données génomiques
R, Expectation-Maximisation Algorithm, Hidden Markov Model
- Estimation du génotype sur des données NGS (séquençage haut débit)
R, Genotype-calling, Expectation Maximisation Algorithm
- Analyse de données génomiques : Apprentissage par pénalisation et étude de stabilité du cancer ovarien
R, Bioconductor, Biobase, glmnet
- Analyse des transactions d'horodateurs parisien sur l'année 2014
Python, seaborn, pandas, pymongo (MongoDB)
- Analyse de données de restaurants dans le monde
Python, scikit-learn, mca, seaborn
- Représentation et classification de visages après réduction par ACP
Python, scikit-learn, pca, k-nn
- Analyse de données des accidents routiers en France sur l'année 2017
R, tidyverse, ggplot

🔗 Projets consultables sur :
<https://github.com/NadiaOuhssaine/Projets>



Expérience professionnelle

- **2020** **Data Scientist stagiaire**
Groupe PSA, Carrières-sous-Poissy



Maintenance prédictive sur des véhicules d'essais du groupe avec la mise en place d'un outil de détection d'anomalies sur des données issues de boîtiers embarqués dans les véhicules.

Machine et Deep learning : DBSCAN, Isolation Forest, LSTM Autoencoder, etc.

- Technique d'interprétation des algorithmes par la valeur de Shapley (théorie des jeux)
- Requêtage sur les bases de données relationnelles
- Langage : Python – SQL
- Librairies : spark – pandas – scikit learn – tensorflow/keras – shap – matplotlib – seaborn – hadoop – Hbase

- **2017 - 2018** **Tutorat des étudiants de Licence scientifique**
Université Paris Est Créteil, CRETEIL



Accompagnement et soutien en mathématiques, informatique et physique



Formation

- **2018 – 2020** **Master Mathématiques et applications**
spécialité Ingénierie Mathématique



Université Paris Descartes, PARIS

Major de promotion - Mention Bien

- **2015 – 2018** **Licence Mathématiques et interactions**
Université Paris Est Créteil, CRETEIL



- **2015** **Baccalauréat général série Scientifique**
Lycée Gutenberg, CRETEIL



Compétences professionnelles

- **Machine Learning** : Apprentissage supervisée / semi-supervisée / non supervisée – Deep Learning - Analyse multidimensionnelle / Réduction de dimensionnalité - Algorithme d'estimation – Performance (ROC,AUC,...) – Interprétation (SHAP, LIME)
- **Statistiques prédictives** : Construction d'estimateurs : moments, quantiles, maximum de vraisemblance, à noyau, etc - Modèles statistiques : linéaire généralisé / gaussien / logistique / logistique pénalisé / Hidden Markov Model – Mesure de qualité : Critère d'information d'Akaike (AIC)/ Bayésien (BIC) – Série temporelle : Processus AR, MA, ARMA
- **Statistiques décisionnelles** : Tests paramétriques/non paramétriques : Student, Wilcoxon, Fisher, ANOVA, Kruskal Wallis, etc. - Tests multiples : FDR, FWER
- **Analyse de données** : Data visualization : ggplot – matplotlib – seaborn - Exploratory Data Analysis : pandas – spark – SQL – Web Scraping
- **Base de données** : Modèle conceptuel de données à partir d'un cahier des charges – SGBD : MySQL, ORACLE, PostgreSQL – NoSQL : MongoDB
- **Langage informatique** : Python, R, SQL, java, LaTeX, Markdown
- **Logiciel statistique, environnement et Framework** : Tableau, PowerBI, SAS, Matlab, Scilab, Jupyter, Rstudio, Hadoop, Atom, Eclipse, MySQL, SQL Developer
- **Système d'exploitation / Bureautique** : Windows, Linux (Ubuntu) / Microsoft Office, Libre Office
- **Méthodologie de gestion de projet** : AGILE (Scrum), Cycle en V
- **Langage** : Anglais B1 - Allemand (notions)