



# BARECHE Yassine

CANDIDATURE POUR STAGE DE  
FIN D'ETUDE DATA SCIENCE  
6 MOIS

## COORDONNÉES

- **Tel :** 07 67 73 08 96
- **E-mail :** barecheyassine@gmail.com
- **Linekdln:** [Yassine-Bareche](#)
- **GitHub :** [github.com/yassino01](#)
- **Adresse :** Paris, France

## ACCOMPLISSEMENTS

- Certifié "IBM Cognitive Class" :
  - [Data Science Orientation](#)
  - [Python for data Science and AI](#)
  - [Data bases and SQL for Data Science](#)
- Troisième place au hackathon, "NASA Space Apps Challenge 2018- Paris".
- GitHub : Arctic Code Vault Contributor Badge
- MPM: [Ms project & Project management](#)

## LANGUES

- Français (Bilingue)
- Anglais (Niveau B2)
- Arabe (Bilingue)

## CENTRES D'INTERET

- Hackathons.
- Contribution à l'Open source.
- Clubs scientifiques.
- Sport en compétition (Licencié FFF).
- Industrie automobile.
- Voyage.
- Cuisine.

## ETUDES

- [2019-2021] **Master DAC - Données, Apprentissage et Connaissances**  
Sorbonne Université (ex UPMC), Paris VI
- [2018-2019] **Licence Informatique - mention " Très Bien "**  
Université Paris Diderot, Paris VII

## EXPÉRIENCES

- [fevrier-Juin] **Buisness Intelligence avec Dataiku**  
2020 Sorbonne Université, master 1 DAC
- Étude sur le niveau de bonheur des populations mondiales.
  - Collecte des données (Kaggle, Data World ..etc).
  - Conception/création d'un **Data Warehouse** avec les données obtenues (schéma en étoile).
  - **Data Visualisation et Machine Learning** (Prédiction) à l'aide de «Dataiku».
- [fevrier-Juin] **Classification et génération du texte en Deep Learning**  
2020 Sorbonne Université, master 1 DAC ,
- Classification des propos par locuteurs Chirac/Mitterrand et une étude sur les sentiments détectés dans les avis des internautes. Génération des tweets de Trump (NLP).
  - **Prétraitement** du texte (Stemming, stopwords) et dictionnaire de mots (TF-IDF).
  - Echantillonnage (Up/downsampling), apprentissage, test et validation des modèles.
  - Génération des séquences de texte à l'aide des **réseaux de neurones récurrents (RNNs)**.
  - Outils : **SVM, Naïve Bayes, CRU, LSTM, ( Pytorch, Nltk, numpy, Sk-learn et matplotlib)**
- [Janvier-Juin] **Diffusion d'information dans les réseaux sociaux**  
2020 Sorbonne Université, master 1 DAC ,
- Modélisation des processus de diffusion avec des données incomplètes.
- **Prédiction des probabilités** de diffusion pour un modèle en cascade indépendant.
  - Apprentissage des paramètres pour la diffusion de réseau latent.
  - **Inférence bayésienne** des réseaux de diffusion avec des temps d'infection inconnus.
  - Applications: **recommandation, maximisation d'influence, prédiction de comportements.**
  - Outils : Python, Numpy, Scipy, Scikit-learn, EM, NaiveBayes,
- [Juin-Aout] **Numérisation et exploration des données**  
2017 Ecole Supérieur d'Informatique - ESI Alger
- Développement et déploiement d'un système d'information automatisant le traitement des données relatives aux sessions des concours d'accès au doctorat LMD.
- **Conception et création** des bases de données **multidimensionnelles**.
  - **Indexation** et développement d'un **moteur de recherche**.
  - Elaboration **des tableaux de bord** et traitement des données en arrière plan.
  - Outils : UML, WAMPS, **SQL**,

## COMPÉTENCES TECHNIQUES

- Programmation** C | Java(FX) | Python | Ruby | Ocaml | Bash |
- Machine Learning** Scikit-learn | Numpy | Scipy | NLP (Nltk )
- Deep learning** Pytorch | Keras | TensorFlow | CNN, RNN, LSTM, GRU
- Data mining** Matplotlib | Seaborn | Bokeh | Pandas |
- Bases de données** SQL | Map reduce | Spark | Hadoop | JDBC |
- Web Sémantique** XQuery | RDF | OWL | SPARQL |
- Computer Vision** Feature Extraction | Pattern Recognition |
- Business Intelligence** Dataiku | Ms Access | Ms Excel |
- Divers** Jupyter-notebook | IBM-Cloud | Merise | UML | Scrum