Benoit de Menthiere

Data scientist

18 October 1994

b.sdementhiere@gmail.com

+33 6 32 13 67 04

linkedin.com/in/benoit-de-menthiere

github.com/bensdm

Education

2018-2019
Telecom ParisTech
Post graduate Master in Big Data and Machine learning|Paris

2015-2018
Arts et Metiers ParisTech
Specialization in information and knowledge system, data mining and machine learning|Lille Paris

Computer skills

- Expert:
Python, Pandas, Scikit-learn, Plotly, Keras
- Proficient:
Tensorflow, SQL, Git, OpenCV, D3js
- Beginner:
Spark, Docker, AWS / GCP, Java, C++, NOSQL, Hadoop, Scala, Tableau, R

Languages

French: native
English: fluent
Spanish: basic

MOOC

- Data Science A-Z: Real-Life Data Science, Udemy
- Deep Learning A-Z: Hands-On Artificial Neural Networks, Udemy
- Fundamentals for Big Data, IMT

Interests

-Sports: tennis (ball boy at Roland Garros), football, bike (1300km trip)
-Travels: Middle East (5 years expatriation in Abu Dhabi), Haiti (humanitarian aid), USA, Mexico, Morocco, Israel, Europe
-Games: chess, poker

Summary

Data scientist, passionate about cutting-edge technology and solving real-world problems. Analytical, intellectual and scientifically-minded with academic background in general engineering. Eager to deploy tools that translate data into business insights.

Experiences

Jul-Apr 2020 Data scientist — Schlumberger, Houston USA

Cement thickening time prediction:
- Optimization of the preprocessing pipeline: 8 times faster
- Build dashboard to control data quality
- Clustering of experiments and outlier detection
- Build regression model with confidence estimation: reduce errors by 25%
- Deploy user-friendly webapp dividing the time needed to find the optimal cement recipe by 30

Maintenance cost prediction:
- Build a constrained linear regression model

Subsurface fiber optics responses prediction:
- Unsupervised pretraining using infoVAE on 30M data points
- Expected savings $1M/well

Sep-Jun 2019 Data scientist student — Bearing Point - Telecom ParistTech, Paris

Optimization of the energy consumption of a silicon furnace:
- Data analysis and visualisation to identify causes of decline in silicon production
- ARMA model to anticipate these drops

Feb-Aug 2018 Data scientist intern — Sopra Steria, Paris, Luxemburg

Trajectory clustering via deep learning representation using auto-encoder:
- Fine tuning parameters
- Visualization

Optical character recognition and entities extractions on invoices and medical prescriptions:
- Computer Vision with OpenCV
- Natural language processing with NLTK
- Deployment with docker

Jul-Sep 2017 Data scientist summer intern — Ignation, Paris

Implementation of an algorithm based on pedestrian tracking in order to improve passenger flow management in public transport (POC):
- Transfer learning using pre-trained YOLO model
- Optimization and test

Realization of a preprocessing pipeline for automating

2016-17 Manager of student restaurant — Arts et Metiers ParisTech, Lille

Inventory and supply-chain management:
- Management of a €20k budget

Projects

Feb 2020 Google hashcode — Houston

Optimization competition: Top 5%

Nov 2018 Hackathon: Huawei big data challenge — Paris

Malware detection using software log files:
- 99% accuracy
- Ranked 6th out of 25 teams

2017 Master projects — Paris

Relevant projects:
- Inclass Kaggle competition: cover type prediction of forests: 3/64
- Build a tools classifier using CNN: 95% accuracy
- Lyrics generator using LSTM and markov chain