Wassim   Ben youssef

Apartment 402, Al Noon Residence, Al Barsha 1, Al Barsha, Dubai, UAE

+971 58 59 50 189| wassim.benyoussef4@gmail.com| https://www.linkedin.com/in/wassim-ben-youssef-54532910b/

Senior Data Scientist with 6 years of experience in machine learning and data analysis. Expertise in developing and implementing models for various industries, including construction, real estate, hospitality, and fintech. Skilled in leading teams, coordinating projects, and delivering actionable insights. Experienced in teaching and authoring online courses.

WORK EXPERIENCE

Senior Data Scientist at Artefact MENA, Dubai   November 2022 – Present
Artefact MENA is a data consultancy company aiding corporations in their digital transformation initiatives.
- Led data analysis and machine learning projects for clients in the construction, real estate, and hospitality sectors.
- Managed a team of 3 data analysts and data scientists.
- Projects implemented:
  - Developed a **clustering-based approach** to estimate the cost of future construction projects.
  - Created a **rule-based rating system** to rank apartments.
  - Built machine learning models to **forecast** the occupancy rate of golf courses.
  - Built multiple **P&L reporting dashboards** for Golf and Hospitality sectors, aggregating multiple data sources
- Coordinated with data scientists and analysts to build and maintain PowerBI dashboards.
- Engaged in client-facing roles, conducting meetings to understand clients' needs and gather feedback.
- Strategy:
  - Assisted data consultants and clients in planning future **data-driven use cases**, ensuring alignment with business objective
  - Automated processes to **evaluate use cases impact and revenue growth**, emphasizing data's role in driving results
  - Translated data insights into **actionable recommendations**, facilitating better decision-making
  - Cultivated **client relationships** through meetings and feedback, **refining project scopes**

Data Scientist at Xelix, London   January 2020 – November 2022
Xelix is a fintech company that helps identify errors and frauds in invoices and provides analytics and reports for Accounts Payable.
- Joined as the first Data Scientist and **automated** the invoice error and fraud detection process.
- Projects accomplished:
  - Developed **machine learning models** to identify **errors and frauds** in invoice fields.
  - Created a model to identify invoices associated with incorrect suppliers.
  - Implemented **rule-based algorithms and machine learning models** to detect duplicate invoices and reduce overpayment.
- Implemented **scalable production code** to process millions of invoices using Python libraries (e.g., Pandas, Numpy, Django, Scikit-Learn).
- Utilized AWS and GCP for databases, file storage, and cloud computing.
- Designed and implemented the **monitoring and retraining lifecycle** of models based on client feedback.
- Developed standardized processes for the Data Science team, including **labelling helper, feature selection, hyperparameter tuning, and retraining frameworks**.
- Trained an email classification model using **NLP** technologies (NLTK, Spacy, TF-IDF).
- Participated in recruitment and mentored interns and new hires.

Course Author for EC-Council   Contract job, January 2021 – May 2021
Authoring an online course of 4 hours about **Advanced Machine Learning for Business Professionals.** Creating all the content and the slides and recording the videos.
The course is available on the online learning platform CodeRed.

Teacher assistant at Le Wagon, Berlin   Contract job, August 2020 - Present
Le Wagon is a coding bootcamp where I work as a Data Engineering and Deep Learning teacher assistant.
- Assisted students in understanding concepts, business applications, and completing technical exercises.
- Topics covered:
  - Recurrent Neural Network
  - ML basics
  - Natural Language Processing
  - Deploying ML models
  - Training at scale

Data Scientist at Digital Fineprint, London   June 2017 - January 2020
Digital Fineprint is an award-winning insurtech company that assists the insurance community in engaging and understanding SME clients.
- Played a key role in **developing products**, delivering client work, and building the data science team.
- Identified new data assets, **APIs**, and data collection methods.
- Conducted **data cleansing and feature engineering.**
- Built end-to-end **machine learning solutions** using Python and AWS.
- Expertise in NLP for text classification, sentimental analysis, named entity extraction, and topic modeling.
- Utilized **supervised and unsupervised learning algorithms**, statistical modeling, and neural networks with libraries like **scikit-learn**, **LightGBM**, and **Keras**.
- Implemented **production code** and deployed solutions in AWS with **Docker**.
- Collaborated with the product and development teams for scalable data collectors and model implementation.
- Built **ETL pipelines** in AWS using AWS Lambda, S3, AWS Glue, and AWS RDS SQL.
- Interacted with sales teams and clients to understand business requirements and translate them into product features.
- Participated in **recruitment** and **mentored interns**.

TECHNICAL SKILLS

Programming Languages: Python (6 years), R (2 years)
Relation Databases: MySQL, PostgreSQL
Cloud platforms: AWS (EC2, ECR, S3, ECS, Lambda, Glue), GCP (CloudSQL, Cloud Storage, Compute Engine, Firewall rules), Azure (DataBricks)
Version control: Github, Bitbucket

EDUCATION

Master of Science in Data Science, London   September 2016 - June 2017
City, University of London
Thesis subject: Building a model to predict job salary from job description
Master of Science in Computer Science and Applied Mathematics, Paris   September 2014 – August 2016
 ENSIIE, a French engineering school

LANGUAGES

French: Native language
English: Fluent
Arabic: Beginner