Mustafa Sayli

Data Scientist

Nottingham, England
07552846461
msayliuk@gmail.com

Profile

A dedicated data scientist with a PhD in applied mathematics, possessing excellent problem-solving and numeracy skills that facilitate the understanding of complex methods to interpret data. Skilled in Data analysis, Machine Learning, Python, Matlab, SQL, Tableau, Julia, and hands-on experience with various tools including Scikit-learn, TensorFlow/Keras, NumPy, Pandas, Seaborn, Matplotlib, and SciPy. Passionate about working on collaborative projects, developing new machine learning products, and expanding knowledge on understanding data science best practice.

Work Experience

Data Scientist-*trainee*, Clarusway IT Services and Consulting **May 2023–Present**

- *Customer Segmentation and Cohort Analysis:* A transnational online retail company dataset analysed to understand customer behaviours and increase retention/response rate. EDA, data visualisation, and data quality issues addressed. Recency, Frequency and Monetary (RFM) analysis is performed and customers are segmented. Data preprocessing, K-means algorithm (by utilising Elbow method, Silhouette analysis, Hopkins test, ARI score etc) studied for segmentation. Cohort analysis performed to find key insights.
- *Customer Retention Analysis:* An e-commerce organization dataset analysed with SQL and Tableau. Customer segmentation, retention rate and churn analysis performed.
- *Car Price Prediction:* A messy dataset with 50+ features scraped from online car trading company to build car price prediction models. EDA process, data visualisation, data cleansing, future engineering, univariate/multivariate analysis, statistical tests, outlier handling, and missing value imputation applied using Numpy, Pandas, Matplotlib, Seaborn, Scipy. Data pre-processing, Regression, Random Forest, XGBoost algorithms are implemented. Hyperparameter tuning and future importance examined. AWS for data pipeline management (EC2, S3, SageMaker, Lambda, API Gateway) utilised for model development and deployment.
- *Fraud Detection:* To predict whether a credit card transaction is fraudulent or not a highly unbalanced dataset (positive class (frauds) accounts for 0.17% of all transactions) with mostly unknown features handled. Logistic Regression, Random Forest, XGBoost, and Neural Network algorithms are utilised along with *Unbalanced Data Techniques* (e.g. SMOTE techniques, threshold tuning, class weight modification). To monitor data and models *Deepchecks* tools (e.g. Data Integrity, Train-Test Validation, and Model Evaluation) are used. After hyperparameter tuning (via GridSearchCV, Optuna) and future importance (SHAP) examination, final model deployed via Streamlit API.
- *Customer Sentiment Analysis:* Natural Language Processing (NLP) techniques are used to find broad trends in the written thoughts of the customers. The goal was to predict whether customers recommend the product that they purchased using the information in their review text. After implementing text preprocessing and vectorization via nltk library, ML(Logistic regression, SVM, RF, AdaBoost)/ANNs algorithms and BERT modelling are applied.

Associate Lecturer (Data Analysis & Economics), Nottingham Trent University **Oct 2020–Present**

Teaching the *Economics and Data Analysis for Managers* module to large groups of students with diverse backgrounds. Covered topics: *Data visualisation methods, Descriptive statistics, Covariance/Correlation, Regression analysis, Inferential statistics, Hypothesis testing, Time series analysis, Macro/Micro economics topics.* Presenting both technical concepts and implementations of them on lab sessions with 10+ different data sets.

Research Fellow (Mathematical Modelling), University of Nottingham **May 2021–May 2023**

Developed neuroscience models and wrote efficient programs (in Python, MATLAB and Julia) to simulate them. Processed model output data in MATLAB and Python to develop succinct visuals summarising the data and analysis. Created automated pipelines allowing multiple model simulations and analyses to occur autonomously. Effectively conveyed my analysis to a range of audiences through presentations at multiple conferences and seminars.

Research Assistant, METU, Turkey **Oct 2011–July 2017**

Studied various artificial neural networks (ANNs) models by implementing dynamical system approach. Simulated different models and analysed/visualised model output data. Wrote 11+ research papers/reports, and presented final outputs. Taught seminars classes for various mathematics and engineering modules to multiple student groups.

Education

PhD in Computational Applied Mathematics, University of Nottingham, UK **2017–2021**

Thesis: *Piecewise Linear Dynamical Systems: From Nodes to Networks*. Studied various network models using dynamical system tools and analysed/visualised simulated data.

BSc (Hons) Mathematics, METU, Turkey **2006–2011**

Publications, Awards, Prizes and Outreach

- Global robust asymptotic stability of variable-time impulsive **BAM neural networks**, *Neural Networks, 67-73*. See my 10+ publications list at *Google scholar*.
- Vice-Chancellor's Scholarship for Research Excellence, The University of Nottingham, 2017-2021.
- Tri Campus Postgraduate Award, Dean Moore Scholarship, 2020.
- Science in the Park, British Science Association, 2019 (presented my experiment to public audience).

Skills and Additional Information

- **Technical skills:** Python, NumPy, Pandas, Matplotlib, Seaborn, ML algorithms, Scikit-learn, Deep Learning, ANNs, CNNs, RNNs, TensorFlow/Keras, NLP fundamentals, Cloud for ML and Model Deployment, Cloud Computing, AWS, EC2, Notebook Instance, S3, SageMaker, AWS Lambda, API Gateway, Streamlit, Statistical foundations, MATLAB, Julia, SQL and basic competency in R, Tableau, MS Office Suite, Excel, Google Sheets, Minitab, XPP, Git/Github, mathematical modelling and algorithmic approach.
- **Soft skills:** Multi-disciplinary team working skills and ability to work autonomously, able to learn new concepts quickly, a curious mindset, paying particular attention to details and ensuring the accurate results, experience in carrying out research and producing accurately written research papers/reports, excellent problem-solving/numeracy skills,able to developing solutions to complex business problems, lead analytical approaches, data-driven decision-making, experience in working to deadlines and tight timescales within Agile, organising own workload, and a good communicator.
- **Online Certifications e.g.** Google Cloud Big Data and Machine Learning Fundamentals, Using Python to Access Web Data, Python Data Structures, Data Analysis with Python (IBM), Introduction to ChatGPT, Stanford Machine Learning/Deep Learning Courses, SQL for Data Science, Introduction to R Programming for Data Science.
- **Languages:** English, Turkish.
- References available on request.
- Permitted to work in the UK (have Indefinite Leave to Remain) and not require any sponsorship for the role.