Joumana Mohamed Youssef El-Sayed

New Cairo, Egypt • s-joumana.yousef@zewailcity.edu.eg • +20 1142660579
• www.linkedin.com/in/joumana-mohamed-857951267 • github.com/Joumana7

Professional Summary

Junior Data Science and AI student with a solid background in machine learning, data engineering, and predictive analytics.

• Experienced in using Python, C++, and SQL for building and deploying ML models, supported by solid mathematical skills.

• Developing machine learning models and optimizing deep learning pipelines for real-time use.

• Completed multiple academic projects in healthcare, finance,NLP and various fields.

• Eager to apply technical and analytical skills in innovative AI-driven environments.

Education

B.Sc. Computational Science & AI
Zewail City of Science and Technology – Giza, Egypt
Major: Data Science & Artificial Intelligence | Expected Graduation: 2026
GPA: 2.98/4.0
Relevant Coursework:

- Machine Learning, Deep Learning, Information Retrieval, Database Systems, Software Development
- Statistics, Optimization, Data Visualization, Project Management

Certifications

Supervised Machine Learning: Regression and Classification *September 17, 2024*
Coursera (authorized by DeepLearning.AI & Stanford University)

- Instructed by Andrew Ng
- Covered linear regression, logistic regression, and supervised ML algorithms.
- Verified Certificate

Data Science & Analytics *July 27, 2024*
HP LIFE Online Course – HP Foundation

- Gained knowledge of key data science practices, tools, and methodologies.
- Explored the impact of data-driven decision-making in business environments.
- View Certificate: Certificate Link

Research Experience

Credit Card Fraud Detection Using Machine Learning and Explainable AI
Academic Research Project | 2024 – 2025

- Conducted a comparative study on various machine learning models for detecting fraudulent credit card transactions using the IEEE-CIS dataset (over 590,000 records).
- Evaluated the performance of algorithms including Logistic Regression, Random Forest, Decision Tree, Naive Bayes, MLP, SVM, and XGBoost.
- Addressed class imbalance using SMOTE and class weighting to improve model sensitivity.
- Applied Explainable AI techniques such as SHAP, LIME, PDP, ICE, and Permutation Feature Importance to enhance model transparency and interpretability.
- Found that ensemble models (XGBoost and Random Forest) achieved the best performance in terms of AUC and F1-score.
- Highlighted the importance of explainability in deploying fraud detection systems in real-world financial settings.
- Developed and tested the complete machine learning pipeline using Python in Google Colab.
- View GitHub repository

Projects Experience

AI-Based Pathfinding and Reinforcement Learning

Academic Project | 2024–2025

- Implemented and compared various search algorithms (BFS, DFS, UCS, IDS, A*, Greedy, Hill Climbing, Simulated Annealing, Genetic Algorithm) in a grid-based environment to find optimal delivery paths.
- Evaluated algorithms based on time, memory, path length, and success rate. Found A* to be the most balanced in terms of performance and efficiency.
- Built a Q-Learning agent to navigate an 100*100 grid with obstacles and random goals.
- Trained the agent using reward-based learning over 500 episodes with a decaying exploration rate.
- Achieved consistent improvement in performance and learned an optimal path through reinforcement learning.
- Gained hands-on experience in classical AI, **Q-Learning,** environment modeling, and performance analysis.
- View GitHub Repository

Parkinson's Disease Detection and Severity Estimation Using Deep Learning

Academic Project | 2024–2025

- Developed a deep learning system to classify individuals as Healthy or Parkinson's patients and to estimate disease severity using UPDRS scores.
- Used speech signal features from the UCI Parkinson's Speech Dataset with 26 extracted features per sample across various speech types.
- Preprocessed and reshaped data into 3D sequences suitable for time-series modeling.
- Implemented and trained five deep learning models: LSTM, GRU, BiLSTM, SimpleRNN, and DenseNet1D using TensorFlow/Keras.
- Evaluated performance using accuracy (classification) and MAE (regression).
- Achieved highest classification accuracy with SimpleRNN (~87.5%) and best regression performance with GRU (MAE ~8.81).
- Demonstrated the effectiveness of temporal models in speech-based medical diagnostics.
- **Tools & Libraries:** Python, Pandas**, TensorFlow, Keras,** Scikit-learn, Matplotlib
- View GitHub Repository

Information Retrieval with Query Expansion and Evaluation

Academic Project – 2024

- Developed a full-featured Information Retrieval (IR) system using the Vaswani dataset.
- Performed text preprocessing (cleaning, tokenization, stemming, stopword removal) and built an inverted index for document retrieval.
- Implemented TF-IDF-based document ranking and query preprocessing pipeline.
- Enhanced query results using semantic query expansion with GloVe, ELMo, and BERT embeddings, as well as RM3 relevance feedback.
- Evaluated system performance using Precision, Recall, and Mean Average Precision (MAP).
- Designed a simple search interface using Flask for user interaction and testing.
- **Tools & Libraries:** Python, Pandas, scikit-learn, **GloVe, ELMo, BERT,** Flask.
- View GitHub Repository

Interactive Dashboard for Product Cost and Quantity Analysis

Tableau Visualization Project – 2024

- Designed and developed an interactive Tableau dashboard to analyze product costs and quantities across multiple dimensions including country, category, quantity intervals, and time.
- Created visualizations such as pie charts, bar charts, line graphs, tree maps, and geographical maps to answer specific business questions related to cost distribution and product performance.
- Implemented effective dashboard layout with clear labeling, consistent spacing, and user-friendly visual design to enhance interpretability.
- Performed cost aggregation and trend analysis across countries and months, and visualized product quantity distribution per category.
- Ensured readability and user engagement through responsive design, color contrast optimization, and interactive filtering.
- Tools & Skills: **Tableau**, Data Visualization, Dashboard Design, Exploratory Data Analysis (EDA)
- View GitHub Repository

Technical Skills

Programming Languages: Python, C++, SQL

Machine Learning & Deep Learning: PyTorch, TensorFlow, Keras, Scikit-learn, XGBoost

NLP & Information Retrieval: BERT, ELMo, Word2Vec (Gensim), PyTerrier

Data Analysis & Processing: Pandas, NumPy, SciPy, SMOTE, Matplotlib, Seaborn

Web & API Development: Flask, Google Search API, BeautifulSoup

Databases: Microsoft SQL Server, MongoDB

Visualization Tools: Tableau, Plotly, Power B

Algorithms: Sorting, Searching, Graph

Soft Skills

– Teamwork, Leadership, Communication, Presentation, Project Planning, Problem-solving

– Adaptability, Organization, Flexibility, Time management, Attention to Details

– Trustworthy and dependable in collaborative environments

Languages

– English: Fluent – Full professional proficiency
– Arabic: Native speaker