Mohamed Sherif El-Boraie

Data Scientist Enthusiast | Machine Learning Engineer

Cairo, EG | mohamed.sherif.elboraie@gmail.com | Exemption from military service| (+20)1090342885 |
Single | 8th Dec 2000 | LinkedIn URL | GitHub URL

PROFESSIONAL SUMMARY

Highly motivated recent graduate with a BSc in Software Engineering from Helwan University, pursuing a career in Data Science and ultimately becoming a Machine Learning Engineer. Possesses hands-on Experience in analyzing complex data sets, designing predictive models, and employing data engineering techniques to prepare and clean data for analysis. Proficient in programming languages such as Python, R, SQL, and visualization tools such as Tableau, Power BI, and Excel. Demonstrated expertise in implementing statistical, machine learning techniques to build predictive models and algorithms, and model deployment as well as web scraping, A/B Testing, and data visualization.

EDUCATION

Bachelor's Degree in Software Engineering, Helwan University, Helwan, Egypt | 2018 - 2022

Faculty of Computer and Artificial Intelligence

- *GPA*: 3.49/4.00 (Excellent with honor degree)
- *Relevant Courses:* Software Engineering, Database Systems, Information Systems Security, Software Documenting, Data Structures, Artificial Intelligence, Data Warehouses, and Data Mining.

American Diploma Degree, Heliopolis Modern Language School, Cairo, Egypt | 2015 - 2018

High School Diploma

- *GPA*: 2.89/4.00
- *Relevant Courses:* Sociology, Modern History, Psychology, Economics, Critical Thinking, Business Modules Management (Project Management, Strategic Management, Human Resources, Strategic Management, Accounting), and Philosophy.

CERTIFICATES & INTERNSHIPS

RadicalX | Artificial Intelligence Engineering Intern | Oct – Nov 2023	**Technocolabs Softwares** | Data Science and Machine Learning Internship | Sep – Nov 2023
The Sparks Foundation | Data Science and Data Analysis Internship | Jun 2023	**Oasis InfoByte** | Data Science Internship | July 2023
Data Scientist Career Track | 365 Data Science (2023)	A/B Testing in Python | 365 Data Science (2023)
Web Scraping and API Fundamentals in Python | 365 Data Science (2023)	Machine Learning with Naïve Bayes & K-Nearest Neighbors | 365 Data Science (2023)
Convolutional Neural Networks with TensorFlow in Python | 365 Data Science (2023)	Deep Learning with TensorFlow 2 | 365 Data Science (2023)
The Machine Learning Process A-Z | 365 Data Science (2023)	Introduction to R Programming | 365 Data Science (2023)
Statistics & Probability & Mathematics | 365 Data Science (2023)	Introduction to Tableau and Advanced Microsoft Excel | 365 Data Science (2023)
SQL | 365 Data Science (2023)	Python Programmer Bootcamp | 365 Data Science (2023)
Information Technology Institute | Front-End Internship (2019)	**RMS for IT Solutions** | Full-Stack Web Development Internship (2019)

PROJECT EXPERIENCE

AI Paired Programming with Autogen Framework| 2023

- Utilized Autogen framework with two agent types (User and Assistant) allowing multiple instances. Implemented Task Master and Junior Bot assistants alongside a user agent. Conducted fine-tuning through prompt engineering and automated agent switching using a chat manager.

- Implemented a chat manager for seamless agent switching between Task Master and Junior Bot, addressing limitations stemming from ChatGPT-4 API key constraints. Deployed a user-friendly web app UI using *Streamlit* for efficient interaction, emphasizing prompt engineering for specific AI responses.
- Improved understanding of the roles of different agents in the collaborative coding process. Achieved optimal results iteratively, finding the right balance between specificity and general use.

Bike Rental Analysis: Predicting Demand and Factors Influencing Usage in the Capital Bikeshare System | 2023

- Analyzed bike rental patterns, Washington D.C., USA, utilizing two years of historical data from 2011 to 2012.
- Applied KNN-regression, linear regression, and random forest algorithms to predict rental demand based on factors such as weather conditions, day of the week, and seasonal variations, achieving high R-squared scores of 0.97 for KNN, 0.99 for linear regression, and 0.997 for random forest.
- Conducted thorough data preprocessing, including visualization, handling outliers and null values, and implementing feature scaling for KNN and linear regression models, while employing grid search with cross-validation for hyperparameter tuning in the random forest model.
- Demonstrated the practical application of bike sharing data in understanding travel behavior, environmental impacts, and urban mobility, contributing to the field of transportation research and its implications for traffic, environmental sustainability, and public health.

Text Classification using Naive Bayes Algorithm | 2023

- Applied Naive Bayes algorithm for text classification on a dataset to categorize posts as positive or negative.
- Extracted relevant columns and performed data preprocessing, including dropping unnecessary columns.
- Split the data into training and testing sets for model evaluation.
- Utilized CountVectorizer() to transform the text data into numerical features.
- Created and trained a Naive Bayes model (Multinomial/ComplementNB) for classification.
- Evaluated the model's performance on the test set, achieving accuracy scores of 95% for Multinomial and 93% for ComplementNB.
- Presented visual metrics, including a confusion matrix and classification report, to assess the model's performance.

Adaptative Speech APP - SIGN LANGUAGE TRANSLATOR (Graduation Project) | FCAI – Helwan University | 2022

- Developed an AI-based mobile application for deaf individuals to improve their communication with the world by detecting sign language from mobile camera images and converting them into readable phrases, using a convolutional neural network trained on a preprocessed dataset of hand signs. Achieved an impressive accuracy of 99.3% on the validation dataset, demonstrating strong expertise in computer vision and deep learning techniques.
- Led data preprocessing and augmentation efforts, improving the model's robustness and generalization. Collaborated with a team to split the dataset into training, validation, and testing sets, effectively ensuring the model's accuracy and reliability.
- Employed hyperparameter tuning techniques to optimize the CNN architecture, achieving a model with exceptional performance. Trained and evaluated the best model, leading to a remarkable 97.67% accuracy on unseen test data.

TECHNICAL PROFICIENCES

Programming: Python, R Programming, C, C++, SQL.

Data Science Packages: Pandas, Numpy, Sklearn, TensorFlow, Pytorch, Keras, Matplotlib, Seaborn, Web Scraping.

Data Science Algorithms: Classification (Naïve Bayes, KNN, SVM, Decision Tree, Random Forest), Regression (Linear, Multiple Linear, Ridge, Lasso, Random Forest, SVR), Classification (K-Means, Hierarchical), Deep Learning (CNN, Transfer Learning), Reinforcement Learning.

Concepts: OOP, AI, Machine Learning, Deep Learning, Network Foundations (CCNA) , Model Deployment

Frameworks: Microsoft SQL Server, MySQL, Jupyter, Spyder, R Studio, Tableau, Microsoft Power BI, Microsoft Office (Excel, Word, PowerPoint), Graphic Design (Photoshop, Illustrator, Adobe XD)

OTHER RELEVANT INFORMATION

Languages:

- **Native Language:** Arabic (Native)
- **Intermediate Language:** English (B2)
  Planning to take IELTS exam for language proficiency.
- **Additional Languages:** French (A2), Deutsch (A1)