ABDALLAH ELSHARAWY

Fairfield, IA 52556 • +1-641-233-9933 • aelsharawy65@gmail.com
https://www.linkedin.com/in/abdallahgaber/

MACHINE LEARNING ENGINEER

AI specialist with 3 years of experience in data analysis, visualization, modelling, and training using TensorFlow, Keras, Spark, and Sklearn. Expertly managed the AI lifecycle, from advanced data preprocessing to model development, evaluation, and deployment, enhancing predictive accuracy and efficiency. Proficient in Python, Data Engineering, NiFi, Kafka, MapReduce, Hadoop.

Statistical Analysis • Predictive Modeling • Ensemble Methods • Experimentation and A/B Testing • Deep Learning
Computer Vision • Unsupervised Modeling • Data Preprocessing and Feature Engineering
Data Visualization • Data Warehousing • Hadoop ecosystem

Languages: Python, Java, Scala
Frameworks: TensorFlow, Keras, Scikit-Learn, TensorFlow, XGBoost, OpenCV, Numpy, Pandas, Matplotib, Seaborn
Databases: HiveQL, Impala, Oracle, MS SQL, PostgreSQL
Big Data: Hadoop, Spark, Hive, NiFi, Impala, MapReduce, Pig, Airflow

PROFESSIONAL EXPERIENCE

*CAREER NOTE: Completed on-campus studies and currently taking distance education courses to complete a **Master's Degree in Computer Science** (Available for full-time, W-2 employment).*

EJADA SYSTEMS LTD., Cairo, Egypt • 06/2022 – 09/2023
IT services provider specializing in systems integration, software development, IT consulting and big data solutions.
Big Data Engineer
Data Warehouse migration to Big Data

- Migrated the Enterprise Data Warehouse (EDW) to a big data architecture using Hadoop and Spark enhancing scalability and performance.
- Translated business requirements into optimized data models using Hive and Spark, improving data management efficiency.
- Documented and delivered ETL jobs using Spark, Airflow, and Atlas, ensuring efficient data processing and transformation.
- Developed Proofs of Concept (POCs) using NiFi to showcase the feasibility of proposed data flow alternative approaches.
- Collaborated on integrating advanced analytics and machine learning models Python, Sklearn, and Keras, enhancing data-driven decision-making.
- Monitored system stability using AirFlow, YARN, and Hadoop, ensuring uninterrupted data flow and system reliability.

Technologies Used: *Spark, Apache NiFi, Oracle, Hive, Impala, Apache Airflow, IBM DataStage.*

egabi Solutions., Cairo, Egypt • 02/2020 – 03/2021
egabi Solutions is a leading software integrator and IT solutions provider in Egypt and United Arab Emirates.
Data Analyst
Collaborated on building Python data pipelines, developing predictive machine learning models and Proof of Concepts.

- Conducted data analysis using Python libraries such as Pandas, NumPy, and Matplotlib to showcase trends, patterns and insights.
- Extracted and transformed data from various databases to prepare for analysis ensuring data integrity and compatibility.
- Developed and implemented machine learning models to generate insights from data using scikit-learn resulting in predictive models that improved accuracy.
- Collaborated with cross-functional teams to understand data needs and deliver actionable insights promoting teamwork that optimized project outcomes and client satisfaction.
- Developed a Proof of Concept (POC) demonstrating the potential of low-cost classical machine learning in predictive analytics highlighting the effectiveness of simpler models over high-cost deep neural networks for straightforward problems.

Technologies Used: *Python, Scikit-Learn, Pandas, NumPy, Matplotlib, SQL, Jupyter Notebooks.*

INTERNSHIP EXPERIENCE

INFORMATION TECHNOLOGY INSTITUTE, Cairo, Egypt • 04/2021 – 01/2022
Provider of professional industrial training for graduates.
AI Trainee
AI Trainee program powered by EPITA - School of Engineering and Computer Science in France. Gain hands-on experience in cutting-edge AI technologies like machine learning and deep learning, preparing you for AI-driven roles in diverse industries.

- Applied supervised, unsupervised, and deep learning techniques using TensorFlow and Scikit-Learn to derive insights and predictions from diverse datasets, achieving high accuracy.
- Utilized fundamental mathematical concepts including linear algebra, calculus, and statistics to inform and optimize machine learning models.
- Enhanced model performance through numerical optimization techniques using methods like RMSProp, and Adam, leading to improved predictive accuracy.
- Leveraged Apache Spark and Hadoop for scalable data processing, enabling the efficient analysis of large datasets.
- Administered Linux systems using Bash scripting and system tools to ensure secure and reliable computing environments.
- Conducted comprehensive data preparation and visualization using Pandas, Matplotlib, and Seaborn to generate actionable insights and interpret predictive models effectively.
- Utilized Python programming extensively across all project phases, from data preprocessing and model development to deployment and monitoring.
- Completed a successful graduation project on sign language translation using AI, demonstrating proficiency in computer vision and natural language processing techniques.

Technologies Used: *Python, Keras, Tensorflow, Sklearn, Spark, Kafka, Numpy, Pandas, Matplotlib, Seaborn, dash, Linux-Shell.*

ACADEMIC PROJECTS

Maharishi International University (2024) ASL Translator - Letters: Designed and developed a system to translate American Sign Language letters using Tensorflow, Sklearn, Mediapipe, and CV2. Providing high precision over the testing methods and live experiments. Utilized Deep Learning, Computer Vision, Keras, Tensorflow, Sklearn, CV2

Maharishi International University (2024) Big Data - MapReducce: Demonstrated and implemented key MapReduce and Spark approaches, including local aggregations, pairs and stripes methods, order inversion, and inverted indexing. Demonstrating the capabilities of modern MapReduce approaches with the Hadoop eco-system. Utilized Hadoop, Mapreduce, Spark, Scala, Java

Maharishi International University (2024) Big Data - Hadoop-Based Pipeline: Designed and developed a data pipeline using Spark Stream with a streaming API. Integrated Kafka consumer for local file production and HBase for data storage. Created an external table in Hive from HBase data and utilized PyHive for data access and insights visualization. Demonstrating the seamless interaction between Hadoop components to provide scalable and persistent data flow. Utilized Hadoop, Kafka, Spark, HDFS, Hbase, Hive, Python

Information Technology Institute (2022) ASL Translator - Words: Developed a system to convert American Sign Language words to text and speech, and vice versa. Created and preprocessed a custom dataset, built a predictive model, and processed outputs for sentence construction and speech generation. Proofing the high reliability of ML solutions for such problems. Utilized Tensorflow, MediaPipe, CV2, POS Tagging, Transformers,BERT

Information Technology Institute (2021) Arabic Handwritten Characters Recognition: Developed a system to recognize and classify Arabic handwritten characters. Preprocessed images, built and modelled them, and monitored model performance and overfitting. Demonstrating the capabilities of image processing algorithms and machine learning models. Utilized Tensorflow, Sklearn, Seaborn, CV2, Pandas, Numpy, Matplotlib

Information Technology Institute (2021) Credit Card Fraud Detection: Classified credit card transactions to detect fraud. Addressed data imbalance while maintaining high precision and recall of the model. Analysing and tackling common issues with data modelling and evaluation. Utilized Machine Learning, Sklearn, SMOTE

EDUCATION

Master of Science in Computer Science
(In progress via distance education; expected completion 06/2026)
Maharishi International University, Fairfield, Iowa

<u>Key Courses</u>: *Big Data, Machine Learning, Enterprise Architecture, Algorithms*

Bachelor of Science in Informatics and Computer Science
The British University in Egypt, Cairo, Egypt (in partnership with London South Bank University), 2019
Received dual certificates from both institutions