Mahmoud Saeed

+201113936587| mahmoud.saeed3690@gmail.com | GitHub | LinkedIn| Kaggle | codingNingas | HackerRank |
Military Status: Completed. | Nationality: Egyptian

Experience

AUGUST 2023 PRESENT

Data Track Headway Program (Data science & ML) | Giza Systems.

- Implemented point anomaly detection algorithms to identify outliers in the dataset, enhancing data quality and ensuring the reliability of analytics results
- Resolved critical issues in the ETL pipeline using Apache NiFi, ensuring seamless data flow and minimizing downtime.
- Updated and optimized the pipeline for asset group management, incorporating features for efficient asset group deletion and addition.
- Used automated machine learning (AutoML) techniques to optimize model selection and hyperparameter tuning.

JUNE 2022 – JULY 2022

Data Science and Business Analytics virtual internship| The Sparks foundation.

- **I have used Python for making cleaning, exploratory, and making deep analysis of data.**
- **Also used Sklearn for Kmean-cluster, Linear Regression, and Decision tree.**
- **Used NLP to predict sentiment Using Random Forest Classifier and nltk.**

Education

JULY 2021

Faculty of Computers and Artificial Intelligence | IS – DS departments | Cairo University.

Projects

- **Sentiment analysis (Graduation Project),** *[Machine learning, python, Amazon scrapping, twitter API,]* Grade: **A+**
  - make data cleaning, exploratory, and making deep analyses using python
  - Used nltk and spacy to get aspects from the data.
  - predicted sentiment Using Decision tree Classifier.
- heart-failure-prediction-Notebook *[Machine learning, python]*
  - *Predict heart failure using different machine learning algorithms*
  - *get deep insights from data using pandas*
- Fake News Detection*[Python, machine learning, nlp]*
  - *Clean and make Tf-Idf technique to get the count of words for each sentence*
  - *Implement machine learning algorithm to predict the fake news*
- Movie Recommendation*[Python, nlp]*
  - Implement 3 different techniques to recommend the movies based on Vote, title, overview, and genres
- Meta-Database-Capstone-Project *[MySQL, stored procedure, join, sub query, view].*
  - Implement stored procedures and sub-queries to make analysis using MySQL
- Netflix Analysis Python pyspark*[Python, pyspark, analysis]*

- Analysis Netflix movies dataset using pyspark

- Implemented Naïve Bayes, Apriori algorithm, and K-mean clustering from scratch.

- Azure projects [Azure, dataflow, pipeline, azure SQL database, SC-dimension type (1,2,3)].
- EDA Supermarket Sales Pyspark [Python, pyspark, analysis]
  - Analysis supermarket sales dataset using pyspark

- Netflix Analysis Python pyspark [Python, pyspark, analysis]
  - Analysis Netflix movies dataset using pyspark

- Sentiment using Logistic Regression [Machine learning, python, NLP]
  - Implement Logistic Regression algorithm to predict sentiment

Awards, Accomplishments

· Data Analysis Professional and Advanced Track in Udacity (FWD) · Applied text mining with python · Natural language processing with classification · intermediate SQL · Meta database Engineering ·

Skills

Python, Machine learning, NLP, Classification, Regression, Database modeling, Pyspark, MySQL, ETL, Java, OOP, Solid principles, Design patterns, DW concepts,, Basics mongoDB, Basics PostgreSQL.