Mahmoud Saeed

+201113936587| mahmoud.saeed3690@gmail.com | [GitHub](https://github.com/mahmoudsaeed99) | [LinkedIn](https://www.linkedin.com/in/mahmoud-saeed-3b218a18b/)| [Kaggle](https://www.kaggle.com/mahmoudsaeed99) | [codingNingas](https://www.codingninjas.com/codestudio/profile/906eb14b-0945-4857-8596-8c3d7d3d6b5e) | [HackerRank](https://www.hackerrank.com/saeedmahmoud389) | 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](https://www.kaggle.com/code/mahmoudsaeed99/predict-using-different-models-for-beginners)-Notebook *[Machine learning, python]*
* *Predict heart failure using different machine learning algorithms*
* *get deep insights from data using pandas*
* [Fake News Detection](https://www.kaggle.com/code/mahmoudsaeed99/fake-news-detection-by-lr-nb-dt-rfc)*[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](https://www.kaggle.com/code/mahmoudsaeed99/movies-recommendationsystem)*[Python, nlp]*
* Implement 3 different techniques to recommend the movies based on Vote, title, overview, and genres
* [Meta-Database-Capstone-Project](https://github.com/mahmoudsaeed99/db-capstone-project)*[MySQL, stored procedure, join, sub query, view].*
* Implement stored procedures and sub-queries to make analysis using MySQL
* [Netflix Analysis Python pyspark](https://www.kaggle.com/code/mahmoudsaeed99/netfilx-analysis-eda-insights-using-pyspark)*[Python, pyspark, analysis]*
* Analysis Netflix movies dataset using pyspark
* Implemented [Naïve Bayes, Apriori algorithm, and K-mean clustering](https://github.com/mahmoudsaeed99/Data-Mining) from scratch.
* [Azure projects](https://github.com/mahmoudsaeed99/Azure-Project) [Azure, dataflow, pipeline, azure SQL database, SC-dimension type (1,2,3)].
* [EDA Supermarket Sales Pyspark](https://www.kaggle.com/code/mahmoudsaeed99/eda-supermarket-sales-by-using-pyspark) [Python, pyspark, analysis]
* Analysis supermarket sales dataset using pyspark
* [Netflix Analysis Python pyspark](https://www.kaggle.com/code/mahmoudsaeed99/netfilx-analysis-eda-insights-using-pyspark) [Python, pyspark, analysis]
* Analysis Netflix movies dataset using pyspark
* [Sentiment using Logistic Regression](https://www.kaggle.com/code/mahmoudsaeed99/sentiment-by-lr-implementation-code) [Machine learning, python, NLP]
* Implement Logistic Regression algorithm to predict sentiment

# Awards, Accomplishments

∙ Data Analysis [Professional](https://confirm.udacity.com/QS3CDVTM) and [Advanced](https://confirm.udacity.com/HYUNDKQG) Track in Udacity (FWD) ∙ [Applied text mining with python](https://coursera.org/share/ee714aa52fcbaab0f729cf3f748c9b30) ∙ [Natural language processing with classification](https://coursera.org/share/42bfe6485d00005bb3cc4f27579536e0) ∙ [intermediate SQL](https://www.datacamp.com/statement-of-accomplishment/course/50fa91dee139b5b45c560a9ab5405476d9dc4410) ∙ [Meta database Engineering](https://coursera.org/share/af6d80763041cc70d5e84205b3e610f5) ∙

# Skills

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