Maryam Esmalifalak

(832) 434-8731| [Maryam.e.falak@gmail.com](mailto:Maryam.e.falak@gmail.com) |

GitHub: <https://github.com/Esmali>|Linkdin: <https://www.linkedin.com/in/maryam-esmalifalak-a1098691/>

**Data Scientist & Quantitative Analyst**

Junior data scientist with 4+ years of experience in project work, professional and freelance jobs who takes pride at developing new forecasting models, and performing data management tasks by utilizing predictive modeling, and data mining algorithms to build models that translate data points into business insights.

**CORE PROFICIENCIES**

|  |  |  |
| --- | --- | --- |
| * Data and Quantitative Analysis * Data-Driven Personalization * Data mining and Visualization Tools * Statistics and Probabilities | * Machine Learning Algorithms * Business Intelligence (BI) * Database Management * Research, Reports and Forecasts | * Professional Ethics * Organization & Multitasking * Continuous Learning * KPI Dashboards |

**Professional Experience**

Institutional Research, HCC, Houston TX

***Data Scientist*** 12/2019 -Present

**Responsibilities:**

* Predicted student retention 15% better than traditional figures
* Created and presented models for student recruitment factors, achieving a 18% improvement of following cohort attainment for coming semester
* Mined large datasets in SQL/SAS to infer the driving factors of high grades
* Developed and delivered custom dashboards, reports, and portal views for monitoring at risk cohorts
* Identified, analyzed, and interpreted the trends that show the popularity of a courses
* Cleaned the data by reviewing departmental/institutional reports, printouts, and performance indicators to locate and correct the code problems

**Tools:** MS SQL Server, Pandas, GENSIM, Sklearn (PCA, DecisionTreeClassifier, SVM, Logistic Regression, Kmeans, KNN, MeanShift)

Housing and Residential Life, University of Houston, Houston TX

***Graduate Assistant, Data Analyst*** 07/2017 to Present

**Responsibilities:**

* Predicted whether students who resided on campus will churn 12% better than traditional figures
* Implemented K-mean clustering to learn the distribution of normal/abnormal behaviors of the residents in 7 residential halls
* Collaborated with IRB to gather all raw data from several tables in databases and cleaned the data for 7 residential halls
* Conducting descriptive analysis on the collected data to visualize important attributes for predictive modeling and found predicators for student retention
* Performed regression analysis on student housing and residential life database to estimate students’ GPA based on on-campus living experience
* Performing Ad-Hoc reporting of occupancy report with TABLEAU.

**Tools:** Oracle, Seaborn, Matplot, GENSIM, Pandas, Sklearn(Logistic Regression, SVM, Random Forest, K-Means ), Scipy , KPI Dashboard(TABLEAU)

**Technical skills**

* **Programming:** Python (Pandas, Numpy, Scipy, Matplot, Seaborn, Statsmodel, Sklearn , NLTK, GENSIM, Folium, Geopy), R, PHP, Java, Ruby on Rails
* **Business Intelligence:** TABLEAU, Google Data Studio, Power BI
* **Database:** MS SQL Server, MySQL, Oracle, SQLite
* **Development Tools:** Visual studio, Jupyter Notebook, Anaconda, Geany, RStudio, Jupiter, MySQL Work Bench, Linux Shell scripting in terminal
* **Machine Learning Algorithms:**
* *Supervised*: Linear Regression, Logistic Regression, Decision Trees, Random Forest, KNN, Support Vector Machines
* *Unsupervised*: K-Means, Fuzzy C-Means, PCA, t-SNE, Hierarchical clustering, Density-Based Clustering
* *Semi*-*supervised*: Deep Learning, Natural Language Processing

**Education**

**M.Ed., *Application of Machin Learning in Higher Education***

University of Houston, Houston, TX (2019-Present)

***Thesis:*** *Predicting post-secondary college enrollment at high school longitudinal study of 2009-2016 (hsls:09) to higher student college recruitment.*

**B.S., *Computer Information Systems***

University of Houston, Houston, TX (2016-2019)

***Capstone:*** *A Web-Based, multi user database application for automating the process of assigning jobs for employees based on their availably and location .*

**MOOC**

* **IBM Data Science Professional Certificate**

***Capstone:*** *Clustering neighborhoods in Pittsburgh, Pennsylvania metropolitan area and the pattern of housing prices in these clusters*

* **Customer Analytics (Wharton, University of Pennsylvania):** *Descriptive analytics, predictive analytics, prescriptive analytics, and their application to real-world business practices*
* **Machine Learning (By Andrew Ng):** Supervised/unsupervised machine learning, datamining, and statistical pattern recognition.
* **Neural Networks and Deep Learning (By Andrew Ng**): Build, train and apply fully connected deep neural networks.

**PUBLICATIONS**

* Ali Irannezhad Ajirlou , Hamidreza Esmalifalak, Maryam Esmalifalak, Sahar Pordeli Behrouz, “Market moods and network dynamics of stock return: The bipolar behavior”, January 2019, Journal of Behavioral Finance.
* Hamidreza Esmalifalak, Ali Irannezhad Ajirlou, Sahar Pordeli Behrouz, Maryam Esmalifalak, “Integration levels across global stock markets: A multidimensional scaling and cluster analysis”, Elsevier, Expert Systems with Applications Volume 42, Issue 22, 1, Pages 8393-8402, December 2015