

| Fadhila Yousuf  Data Scientist    **LinkedIn:** <https://linkedin.com/in/fadhila-yosof>  **GitHub:**<https://github.com/fadhilayosof>  **Portfolio:** <https://fadhilayosof.github.io/Portfolio/> | 4429 count ct. sterling Heights MI, 48310**|**  Cell no 586-718-9379 [fmyosof@mtu.edu](mailto:fmyosof@mtu.edu) **|** |
| --- | --- |
| PROJECTS **Identifying Symptoms of Orthopedic Patients as Normal or Abnormal(Classification Problem)**   * **Skills used:** Python, Pandas, sklearn, Matplotlib * **Project Objective:** Classifying Biomechanical Features of Orthopedic Patients based on provided features to decrease the time required for diagnosis. * **Quantifiable result:** We could classify the type of tumour resulting in 85% accuracy using the K-means algorithm with K=5 * We could classify the same result with 82% accuracy by using the Naive Bayes model * **Github:**<https://github.com/fadhilayosof/Knn_Nb_Project1/blob/main/Knn_Nb_Project.ipynb>   **Implementing Deep Neural Network With Keras For Handwriting Classification and Recognition (MNIST Dataset)**   * **Skills Used:** Matplotlib, NumPy, Seaborn, Sklearn, Python, Neural Networking, Keras, Tensorflow * **Project Objective:** Implementing Deep Neural Network with Keras for handwriting classification and recognition * **Quantifiable Results:** We could classify the type of number resulting in 97% accuracy. * **Github:**<https://github.com/fadhilayosof/hand_writting_recognition_project.>   **TalkingData Project on Bagging and Boosting Ensemble Model**   * **Skilled Used**: Python.Pandas and sklearn * **Project Objective**: The classification goal is to The data contains observations of about 240 million clicks, and whether a given click resulted in a download or not (1/0): * **Quantifiable Result**: We could predict whether a given click resulted in a download or not. The accuracy of the XGBoostClassifier is 94.87% Accuracy of the Bagging classifier accuracy is 99%. * Used the XGBoostClassifier & BaggingClassifier whether a 240 million click resulted in a download or not. Compared predictive performance by fitting an XGBoostClassifier & BaggingClassifier model to the data * Selected best model based on the train and test performance * **Github:**<https://github.com/fadhilayosof/Bagging_Boosting_Project/blob/main/Bagging_Boosting_Project.ipynb>   **Examining the effect of environmental factors and weather on Bike rentals (Linear Regression)**   * **Skills used:** Python, Pandas, sklearn, Linear Regression * **Project Objective:** Predicting Bike rental demand on a basis of weather and seasonal factors in advance to take appropriate measures which finally will result in bike utilization. * **Quantifiable result:** We could predict the rental bike count resulting in 78% Accuracy.   Used Linear Regression to predict the number of bikes rented in the city of Seoul   * **Github:**<https://github.com/fadhilayosof/linear-regression-project>   **Predicting Employee Attrition**   * **Skills used**: Python, Pandas, SKlearn, Matplotlib,DecisionTreeRegressor,LinearRegression * **Project Objective**: Using machine learning to predict employee attrition in Python, We use linear Regression,and Decision Tree Regressor as classifier for employee attrition and measure the accuracy of models that are built.   **Quantifiable Result:** We could predict Employee Attrition using linear regression model.   * The accuracy of linear regression is 97% * Encoded categorical variables to numeric using Sklearn due to the presence of many string columns. * The selected best model is based on train and test performance. * **Github:**<https://github.com/fadhilayosof/Suicide-Rates-Overview-1985-to-2016-project>   **Clustering of San Fransisco Employees based on salary (Unsupervised Learning- K means Cluster)**   * **Skilled Used**: Python.Pandas and sklearn * **Project Objective**: The classification goal is to find the number of clusters for the data. Policymakers to understand how they can improve on their policies for employee benefits * **Quantifiable Result:** We could predict a number of clusters. Used the clustering algorithms, K-Means and Hierarchical clustering. * **Github:**<https://github.com/fadhilayosof/k-means-project/blob/main/Copy_of_k_means_project.ipynb> | **EXPERIENCE**  **Graduate Assistant**  **Department of Mathematics and Statistics.**  **Michigan Tech University**  **2015-2018**   * Extensive experience working with large data sets. * Exceptional ability in analyses, and model and data interpretation. * Prepared data analysis plans and produced charts/graphs and statistical reports. * Outstanding problem-solving skills.   **Teacher Assistant**  **Department of Statistics.**  **University of Tripoli Libya,**  **2008-2013**   * .. * .. * .. * …  SKILLS  * **Programming Languages:** Python. * **Library:** NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, NLTK, TensorFlow, Keras * **Data Science Skills:** Data Visualization, Data Cleaning, Exploratory Data Analysis, Probabilities & Statistics, Machine Learning, Predictive Modeling, Model Optimization, Deep Learning, NLP, Model Deployment, Computer Vision. * **Database:** MySQL, SQL server * **Other Skills:** Git/GitHub, Auto CAD * **Soft Skills:** Problem Solving, Collaboration, Critical Thinking * **CERTIFICATION**   + **Machine Learning A-Z™: Hands-On Python & R in Data Science, 2020**   + **The Data Visualization: Excel, Tableau, Python, R Certificate, 2020.**   + **The Complete SQL Bootcamp, 2021.**  EDUCATION **Bootcamp - Data Science**  **TECH I.S., Santa Clara,**  **October 2021-January 2022**  Graduated as a Data Scientist in an accelerated program with an immersive hands-on project work experience in Python, ML, AI, NLP, Analysis, Data Visualization, Probabilities & Statistics, Learning, Intelligence, etc.,  **Michigan Tech University**  **Houghton MI**  **M.S. in Statistics (2018)**  **Tripoli University, Libya**  **M.S. in Statistics (2008)**  **Al-Jabal Al Gharbi University**  **Gharyan, Libya**  **B.A. in Mathematics**  **2001** |