Kiran **Kondisetti**

3645 Wellborn Rd, Apt No.1734

(+1) 520-406-3244 | n26111997@tamu.edu | medium.com/@kirankondisetti34 | kirankondisetti | kiran-kondisetti

Education

Texas A&M University (College Station), Master of Science   CGPA : 3.6/4   2019 - Present   *Texas, USA*

Focus: Statistical Machine Learning, Data Science

Skills

Languages	Python, R , SQL, C, GNU Octave, MATLAB, Pyspark, Databricks, ApacheSpark
Frameworks/Tools	Pandas, Numpy, SciPy, Scikit-Learn, TensorFlow, Pytorch,SymPy, Github, MS Office,Tableau
Background	Machine learning, Linear& Non-Linear Optimization, Computer Vision, Deep Learning, AWS
Software Packages	R Studio, Jupyter Notebooks, Spyder, Visual Studios

Work Experience

AARP   *Washington DC, USA*

Data Scientist { STACK-pyspark,spark.sql,Databricks,Tableau}   *FEB 2021 - Present*

- Analyzed a massive raw consumer web-traffic dataset using data mining techniques to transform and include critical dimensions by feature engineering using PySpark in Databricks Notebooks
- Building intricate ML prediction models to validate predictions through different validation techniques targeting 20% increased AARP membership adoption

Research & Publications

GRADUATE RESEARCH ASSISTANT   *College Station, USA*

DATA SCIENCE/MACHINE LEARNING {STACK-R,PYTHON,SQL }   *JULY 2020 - JAN 2021*

- Implemented NLP techniques and neural networks to find the impact of climatic changes on company's innovation, achieving 15% more returns to the company
- Utilized web scrapping techniques to extract and organize competitor data
- Publication- **Comparison between Traditional Perceptron Algorithm and Scikit learn Perceptron Algorithm (medium article)**
- Publication-**Bayesian Perspective of Regression and Classification Problems (medium article)**

Projects

Detecting Attack by Malicious Executables using different Machine Learning models   *College Station, Texas*

Github   *JAN. 2020 - MAY. 2020*

- Employed data processing techniques like validation, normalization and sorting of the given malicious software dataset
- Built and Compared two ensemble learning models constructed using decision trees, perceptron, and MLP using K-fold validation and plotted an ROC curve to choose the best threshold to select the most accuracte model which had an accuracy of 99.98%.

Image Classification Using CNN for E-Commerce Platform   *College Station, Texas*

Github   *JAN. 2020 - MAY. 2020*

- Achieved increase in volume of data through Data Augmentation post-visualization of the data to aid in better training of the model using rotation, width, height shift, shear transformation etc.
- Attained a classification accuracy of 99.98% using hyper-tuned CNN for optimized image search results in the e-Commerce site by image denoising using auto encoder in python

Clustering Geo-location data Intelligently to build the service stations in python(Unsupervised Learning)   *College Station, Texas*

Github   *AUG. 2020 - SEP. 2020*

- Visually estimated the number of clusters (k) by exploratory data analysis on the geographical data comprising of latitude and longitude coordinates
- Compared the Silhouette Scores of k-means clustering, DBSCAN and HDBSCAN for clustering data and concluded that HDBSCAN better clustered the data since it had a Silhouette score of 0.89.

COVID-19 Diagnosis using Neural Networks for Image Classification (Computer Vision)   *College Station, Texas*

Github   *OCT. 2020 - Present*

- Developing a fully automatic framework (CNN) for the diagnosis of respiratory infectious diseases, specifically a COVID19 infection, using Chest X-Ray Scan using the COVID-19 RADIOGRAPHY DATABASE in kaggle

Key Courses

Courses (Acedamic, Coursera & LinkedIn)	Engineering Data Analysis, Machine Learning, Python for Data Science, R Statistics Essential Training, SQL for Data Science, Non-Linear Algebra and Dynamic Programming (NLP),CNN, Time Series Analysis