Ronak Sankaranarayanan

Houston, TX - 77084 | ronaksankar.it@gmail.com | 774-418-9871 | linkedin.com/in/ronaksitds/ | Personal Website

SUMMARY: Data analyst with a Master's degree in Data Science and 1-year professional experience in delivering end-to-end data solutions. Thoroughly proficient in Python and SQL programming, ETL tools, warehouses and databases, and data visualization tools - Tableau and Power BI.

EDUCATION:

Master of Science, Data Science, 4.0/4.0 ***Aug '20 - May '22***
Worcester Polytechnic Institute, Worcester, MA, USA

Bachelor of Technology, Information Technology, 3.3 / 4.0 ***Aug '15 - May '19***
Anna University, Chennai, TN, India

WORK EXPERIENCE:

Research Assistant, Worcester Polytechnic Institute, Massachusetts, USA ***May '21 - May '22***

- Implemented *DOSGAN, GMM-UNIT, and STARGAN-V2* generative models for Synthetic Time Series data generation for Human Context Recognition (HCR) using Pytorch
- Evaluated the quality and diversity of the synthetic data generated by the models with Kullback–Leibler (KL) Divergence and Frechet Inception Distance (FID) score. The best model attained a low KL divergence and FID Score of **0.41 and 2.11**.
- Analyzed the impact of synthetic data on the performance of the Human Context Recognition Classifier model using CNN and LSTM layers. Generated data has improved the performance of Deep Learning Classifiers by **24%.**

Data Analyst, Market Simplified, Chennai, India ***Jun '19 - Jul '20***

- **Crashlytics for Mobile Application**
  - Captured application performance issues using *Firebase Crashlytics* and developed Tableau dashboards to keep track of said issues.
  - Collaborated with software developers to fix said issues and acted as a liaison between developers and business stakeholders
- **MSF Shield - Customer analytics for banking application**
  - Developed targets and designed promotional campaigns for Banking applications as a part of the in-house campaign management team.
  - Performed *A/B testing* for banking campaign on student loans towards the target audience and measured the success with a number of visits to the websites and appointments booked for counseling. Observed a **15% increase** in website visits and appointments requested regarding the specific student loan campaign.

ACADEMIC PROJECTS:

Dollar Tree: Case Study Analysis, *Business Intelligence* ***Sept '21 - Dec '21***

- Designed three detailed dashboards on Sales, Cost, and Macroeconomics analysis on Dollar Tree and Family Dollar stores across the United States using Tableau.
- Embedded 5 filters (Brand, Region, Time period, and Category) to drill down the hierarchy and filter analysis.
- Examined external environmental trends from an array of sources such as truck utilization, fuel prices, consumer price index, and their impact on retail and sales.

Scalable COVID-19 analysis using Big Data, *Big Data Management* ***Jan '21 - May '21***

- Constructed multiple distributed and scalable map-reduce jobs in Hadoop for a dataset with 10 million data points containing personID and coordinates of seats.
- Identified and displayed personIDs close contact with infected people using distance-based computation using SparkSQL.
- Used SparkContext to retrieve data from the HDFS and implemented a clustering algorithm to find the number of clusters in the concert using PySpark MLlib.

Movie Prediction Rating System, *Statistical Learning* ***Sep '20 - Dec '20***

- Extracted 450,000 records from multiple sources in JSON format and merged them into a single holistic dataset.
- Preprocessed the data with the steps of feature engineering, Synthetic Minority Oversampling Technique (SMOTE), and PCA on the data.
- Implemented, Trained, and Tested *Support Vector Machine (SVM), Random Forest, and XGBoost* using 3-fold Cross Validation in R.
- Attained an **F1 score of 0.75** on the XGBoost as the best performing and **0.41 F1 score** on SVM as least performing.

SKILL:

Languages:	Python, R, SQL, SAS, Java, C++, React-Native
Visualization:	Tableau, Power BI, Matplotlib, Seaborn
Database:	MySQL, SparkSQL, PostgreSQL, Oracle SQL, MongoDB, SQLAlchemy
Technologies:	Firebase Crashlytics, Hadoop, Apache Spark, Snowflake, Git, Anaconda, Tensorflow, Pytorch, PySpark, Pandas, Numpy, AirFlow, Docker, Azure, AWS, Kubernetes, MLFlow.
Analytics & ML:	Classification, Regression, Classification, Boosting algorithm, Time Series Forecasting, CNN, RNN, NLP, Predictive modeling, Web Analytics, Hypothesis Testing, Data Augmentation, Big Data Management, Big Data Analytics