SAIKUMARREDDY POCHIREDDYGARI

(315) 742-4440 • saikumarreddy.pochireddygari@gmail.com • linkedin.com/in/saikumarreddyp

EDUCATION

Syracuse University, School of Information Studies, Syracuse, NY
May 2024
M.S. Applied Data Science
Relevant Coursework: Natural Language Processing, Database Management, Machine Learning, Visual Analytics, Deep Learning

SRM University, Chennai, India
May 2018
B.E. Electronics and Communication Engineering
Relevant Coursework: Math, Statistics, Probability, MATLAB, Inferential Statistics, Descriptive Statistics, C

WORK EXPERIENCE

Graduate Data Science Research Assistant, School of Information Studies, Syracuse, NY
Nov 2022 – Present
- Improving process for better navigation of Geo-data Wiki for better user experience using Python and Docker
- Scrapped 1Million Reddit and 5 Million Twitter data pertaining to climate and constructed dependency tree using Python
- Feature Engineered 25 features for Reproducibility ML Research and performed manual coding of the features using Excel
- Conducted thorough requirements gathering with manager to understand the data structure, schema, and transformation needs for migrating data from MongoDB to Azure Data Warehouse
- Developed robust Extract, Transform, Load (ETL) pipelines using Azure Data Factory to efficiently extract data from MongoDB, perform necessary data transformations and load the data of 2TB to Azure Synapse warehouse
- Designed and implemented an incremental data load strategy to handle the migration of 2TB data in batches minizing the downtime and ensuring continuous updates until migration is complete

Machine Learning Engineer, TATA Consultancy Services, Hyderabad, IND
June 2018 – Aug 2022
- Achieved customer satisfaction index of 95.44% for smooth migration of premium models on AWS using AWS stack
- Designed and developed hybrid machine learning operations architectural framework using Python
- Revised process flow by building custom XGBoost images, 3 pipelines, and 15 airflow scripts using Python
- Decreased billable time by 10% to optimize cost on large scale data by proposing a custom image solution using Python
- Improved process improvement by 60% by devising a filter for product identification to aid downstream teams using Python

PROJECTS

Crypto jacking Detection Challenge - *Python, Stats, Machine Learning*
Mar 2023 – Apr 2023
- Predicted a networking activity using Machine Learning to enable better security to a user
- Feature Engineered 18 features which helped in better identification of bad network activity
- Improved F1 Score by 58% in identifying bad network activities by using machine learning techniques

Facebook Friend Case Study - *Python, Machine Learning, EDA, Visualization, Statistics, Graph*
Mar 2022 –Apr 2022
- Assessed potential link between two users to enable better user experience by using machine learning
- Performed feature engineering to create 52 features for each user in a graph and modelled best fit algorithm using Python
- Improved F1 Score by 7% in identifying link between two user nodes in a graph network using machine learning

Analyzing Motor Vehicle Accidents – *Tableau, Dashboard, Story, Data Engineering*
Mar 2023 – Apr 2023
- Built a ETL pipeline for data transformations using Azure ML and Azure Databricks
- Performed data transformations on 30GB of data using Spark and connected o/p data to Tableau using connector
- Performed data analysis to analyze the root cause of accidents in NYC and built a story dashboard using Tableau

TECHNICAL SKILLS/CERTIFICATIONS

Programming Languages: Python, R, SQL
Certifications: Google Tensorflow Certified Developer, Microsoft Azure Certified Data Scientist, Azure Certified Data Engineer Associate, IBM Data Science Professional
Frameworks: MLFlow, Flask, Django, Dask, DVC, GIT, Keras, Agile, TensorFlow, API, Pytorch, Apache Spark, Hadoop, kubeflow
Database Management/ETL: SQL, Mongo DB, Cassandra DB, NoSQL, MS SQL Server, Spark
Machine Learning: Regression, Random Forest, XGBoost, Bayesian Analysis, Decision Tres, Clustering Analysis, SVM
Cloud: Azure ML Studio, AWS Sagemaker, Azure Databricks, Azure Devops, Synapse, Azure Machine Learning, Azure Data Factory
Software Tools: PyCharm, VSCode, Docker, Azure Data Studio, MS Excel, MS PowerPoint, MS Word, JIRA, Tableau, Gitlab

LEADERSHIP EXPERIENCE

Nexis Technology Lab, Syracuse University, Syracuse, NY
Oct 2022 – Present
- Leading a data science team of 25 in solving business problems pertaining to data science