Aman Satya

9986428963 | satyaaman97@gmail.com | Website: satyaaman97.github.io

Education

University of Colorado Boulder Aug. 2019 – May 2021

Masters of Science in Computer Science *Subplan: Intelligent Systems, Boulder, CO*

- **Key Courses:** Machine Learning, Data Mining, Natural Language Processing, Probability in Computer Science, Neural Nets and Deep Learning

BMS College of Engineering May. 2019

Bachelors of Engineering in Computer Science *Bengaluru, India*

Technical Skills

Languages: Python,Java, Scala, SQL, C, JavaScript, HTML/CSS, R
Frameworks: Apache Spark, Pyspark, DBT, AWS EMR, AWS EC2, Kafka, Hadoop, Data Vault 2.0
Developer Tools: Airflow, AWS DynamoDB, Cassandra, Redis
Libraries: Scikit-learn, Pandas, NumPy, Tensorflow, Keras, PyTorch, Matplotlib, Seaborn

Experience

Network 18 Aug 2022 - Present

Software Development Engineer-1: Data Products *Noida, India*

- Building data gathering platform handling both Product and Technology perspectives for Network 18 digital properties like moneycontrol, CNBC, News18 etc. Building datamarts on top of the platform for business intelligence and data science use cases. Tech Stack: Python, Pyspark, SQL, DBT, jinja, Data Vault 2.0, Postgres

ZEE5 July 2021 - Aug 2022

Software Development Engineer- Big Data Engineer *Bangalore, India*

- Built data pipeline to prepare an analytical record of advertisers at Zee5 platform, aggregated 50+ attributes using pyspark. Built ETL pipelines for CDP and developed data pipelines for audience creation, used for Ad campaigns. Automated workflow of spark jobs using Airflow. Tech Stack: Python, Scala, Pyspark, Apache Spark, AWS, Airflow, Postgres.

Times Club June. 2020 - Aug. 2020

Data Scientist Intern *Redwood City,CA, USA*

- Developed a ranking model using scikit-learn Python library for prioritizing merchant-onboarding on Credit Card bills payments app to increase user retention. Stack: Python, Excel, Scikit library.

Status Solutions Aug 2020- May 2021

Research Staff- Capstone Project *Ohio, USA*

- Developed an algorithm to identify patterns leading to PTSD. Created data pipeline and underlying features using supervised and cluster algorithms via scikit-learn Python library. Eventually pushed the predictions to MySQL database. Tools and algorithms: Pytohn, scikit-learn, KNN, XGBoost, Random Forest, Decison Trees , SVM Classifier

Lumata Digital June. 2017 - July. 2017

Software Engineer Intern *Noida, India*

- Developed a sentiment classification of social media posts to determine user's views towards brands.Used Convolutional Neural Network on Tensorflow to classify posts into sentiment buckets.

Projects

Stock Trading Interface using data from Reddit:
Developed a stock trading web application which enables user to trade on the basis of sentiments of live streaming textual data coming from /r/WallStreetBets using Reddit API. Stack:DynamoDB, Kubernetes, Kafka and React

Class imbalance problem in image classification:
Evaluated and resolved class imbalance problem in German Traffic Sign Image Data Sets via Borderline SMOTE oversampling technique, thereby increasing test data accuracy under ROC-AUC curve from 91.8

Fake News Detection With Enriched Attributes Using NLP:
Synthesized a new fake news detection model by adding new attributes of Click-baitness and Content Similarity into existing LSTM model using Keras Library; Resulted in improvement of 0.1% in accuracy