MUGUNDHAN MURUGESAN

470-815-3047 | mmurugesan3@gatech.edu | linkedin.com/in/mugndhn | github.com/mugndhn

SUMMARY

- Highly motivated graduate student actively seeking full-time opportunities in data science & ML starting May 2023.
- Adept in Python, SQL, PySpark, PyTorch and cloud technologies tools like AWS and databricks, with hands-on experience in AI/ML libraries such as Pandas, NumPy, scikit-learn for statistical analysis and data storytelling.
- Skilled in solving complex business problems using analytical tools through internship and coursework experience.

EDUCATION

Georgia Institute of Technology, Atlanta, GA | GRE - 327/340   August 2021 - May 2023
Master of Science (M.S), Computational Science and Engineering (CSE)   GPA - 3.9/4.0

- Relevant Courses - Machine Learning, Numerical Linear Algebra, Statistics, Data & Visual Analytics, Algorithms, Computer Vision, Simulations, Business Intelligence & Analytics, Price Analytics & Revenue Management.
- Graduate Teaching Assistant for CS 4460 - Introduction to Information Visualization (D3.js, HTML, CSS, Tableau)

National Institute of Technology, Tiruchirappalli, India   June 2016 - June 2020
Bachelor of Technology (B.Tech), Civil Engineering (Minor in Management Studies)   CGPA - 8.16/10

PROFESSIONAL EXPERIENCE

Data Scientist Intern, Revenue Team | FLEETCOR, Atlanta, GA   May 2022 - August 2022

- Built an algorithm to fuzzy match ~400k customer records from ETL data pipelines of different business services using NLP with optimized memory usage and improved the data quality by 50% compared to the previous Alteryx method.
- Recalibrated existing customer attrition model using gradient boost algorithm and market segmentation, improving the recall by 20% and identified attrition key factors for recommending sales team on the customer retention activities.
- Developed best practices for AWS tools like SageMaker, S3 bucket and GitLab for automation and model deployment.
- Coordinated with internal teams to automate existing revenue improvement products saving 50% in time, and effectively communicated the results of machine learning models to stakeholders from a business perspective.

Site Supervisor | Konstructs Designers, India   August 2020 - March 2021

- Handled day-to-day progress of building construction (~$2.5M in total project costs) and effectively coordinated and scheduled construction activities across different contractors like masons, electricians, plumbers, and painters.
- Identifying key areas of construction project expenditure by visualizing project financial data on Microsoft Excel and proposed alternative methods by cost analysis of materials and contracts, improving the profit margin by 10%.

PROJECTS

Graduate School Projects | Georgia Institute of Technology   August 2021 - Present

- Computer Vision CS6476 - Developed SIFT local feature matching algorithm, camera calibration using RANSAC, scene recognition & semantic segmentation with deep learning (ResNet architecture) using PyTorch & Numpy.
- Forecasted the crop yield with long short-term memory (LSTM) neural networks on a time-series data with features such as temperature, precipitation, soil moisture, and humidity to anticipate the demand for agricultural products.
- Credit Risk Modelling - Predicted the probability of car loan repayment using logistic regression model on a dataset with ~200 features by feature engineering and, identified key features to reduce the risk from a business standpoint.
- Developed an application using simpy and pygame python packages to run discrete event simulation for studying the effect of passenger flow congestion on the delay of flight departure in airports and recommend process improvements.

Detecting Fake reviews using Natural Language Processing (NLP) | CSE 6242   August 2021 - December 2021

- Processed and analysed ~7M text data (~33 GB) of Amazon reviews using apache spark on the databricks platform.
- Engineered the text data into feature vectors using Bidirectional Encoder Representations from Transformers (BERT).
- Built a neural network classification model to classify the Amazon reviews as fake or real with an accuracy of 76%.
- Deployed into production on a flask web application which uses link of the amazon review or the review text to predict.

SKILLS

- **Programming Languages -** Python, SQL, R, PyTorch, Scala, PySpark, Hadoop (MapReduce), Matlab, Julia.
- **Analytics Tools -** Numpy, Pandas, Scipy, scikit-learn, AWS, Flask, matplotlib, ggplot, D3.js, Tableau, Excel.
- **ML concepts -** Linear/Logistic Regression, ANOVA, Naïve Bayes, Gradient Boosting, AdaBoosting, Random Forest, Support Vector Machines, K-Means Clustering, NLP, Computer Vision, Deep Learning, Neural Networks.

ACHIEVEMENTS

Winner of - Best ML Hack Using DeepNote - at Hacklytics 2022 | Georgia Tech | Link   February 2022

- Developed an NLP based tool useful for performing sentiment analysis on the replies to tweets about Mental Health.
- Trained an NLP neural network to predict sentiment scores using ~10k replies for tweets collected by web scraping.
- Automated a chatbot on selenium which replies a positive message to users based on predicted sentiment scores.