

Suvaleena Paul

+1 (470)-815-2822 | Atlanta, Georgia | suvaleenapaul.sp@gmail.com |  

TECHNICAL SKILLS

Programming: Python, R, SQL, SAS, Java (RestAPI, Angular Js, Polymer Js), Javascript, HTML5

Databases: MS SQL, MySQL, Oracle DB, NoSQL (MongoDB), AWS Redshift, Microsoft Azure, Kafka

Core Competencies: Data Mining, Data Analysis

Tools: Alteryx, Tableau, PowerBI, CSS 3, Web-Scraping using Selenium, Scrapy, Maven, RapidMiner

Environments: Anaconda, RStudio, Jupyter, Apache Hadoop, Apache Spark, Teradata, Matlab, VS Code

Project Management: SCRUM, Agile

PROFESSIONAL EXPERIENCE

Bank of America

Kennesaw, GA, USA

Assistant Vice President - Consumer Product Strategy Analyst III

December 2022 - Present

- Identifying opportunity for new fraud strategy development using statistical and analytical techniques
- Building and optimizing data-driven, analytically based fraud rule strategies which could encompass application fraud, digital money movement products, device authentication strategy

Ascantt Business Systems Inc.

Plano, TX, USA

Data Analyst - Digital Strategy Asset and Portfolio Management

May 2022 – December 2022

- Transformed raw critical business data in AWS Redshift using SQL queries for further analysis
- Analyzed financial data related to budget and expense and provided meaningful solutions to Toyota Motor North America through Tableau/PowerBI expense analysis reports for purchase order of vehicles

Institute of Insight, Georgia State University

Atlanta, GA, USA

Graduate Research Assistant – Data Scientist

August 2021 – April 2022

- Built a machine learning pipeline for tweet extraction from Twitter for changes in stock prices followed by streaming the data in batches through Kafka topics and running an ML model script in Apache Spark
- Developed a Tableau dashboard to visualize the analysis of the predicted test dataset stored in MongoDB
- Scraped Twitter using Twint and labeled extracted tweets on Health Technology Discourse during COVID for Sentiment and Emotional analysis using Vader and designed monthly topic models

Infosys Ltd.

Bengaluru, India

Full Stack Developer - FINACLE

November 2019 - July 2021

- Handled front-end customization with software development in java for designing a web application
- Developed back-end integration of banking functionalities and customized APIs in Java for deployment
- Received a Performance Based Incentive of USD 800 for exceptional performance in project delivery

Goldman Sachs Inc.

Bengaluru, India

Data Analyst – Realty Management Division

February 2019 - May 2019

- Automated processes and optimized SQL queries utilizing Alteryx and visualized the data in Tableau
- Revamped business processes using Business Intelligence tools and automation to increase efficiency by saving 40 hours/month of human engagement on data analysis and database structuring
- With a team of 7, organized the first ever Alteryx based Hackathon hosted by the organization

Indian Oil Corporation Ltd.

Kolkata, India

Data Analyst – Additives and Lubricants

May 2017 - July 2017

- Designed a model using ABC Analysis to maintain an optimum level of inventory per month
- Identified excess inventory of additives blocking a capital of approximately USD 10 million and proposed efficient raw material distribution options that led to an overall profit for the firm

EDUCATION

J. Mack Robinson College of Business, Georgia State University

Atlanta, GA

Master of Science in Data Science and Analytics, *GPA – 4.06*

August 2021 – December 2022

Symbiosis Centre for Distance Learning

Pune, India

Post Graduate Diploma in Data Science and Statistics, *GPA – 3.48*

January 2020 – July 2021

Manipal Institute of Technology, Manipal Academy of Higher Education

Manipal, India

Bachelor of Technology in Chemical Engineering, *GPA – 3.18*

July 2015 - June 2019

PROJECTS (CODE ON GITHUB LINK)

Behind the Mask - Image Analytics using GANs*Atlanta, GA*J. Mack Robinson College of Business, GSU

April 2022

- Trained masked and unmasked images using Attention GAN for unpaired image-to-image translation
- Transformed images for data augmentation, ran the model in batches of 4, for 200 epochs and optimized the model using Adam to get unmasked images with expressions

Predicting Smart Grid Stability - Deep Learning Analytics (ANN)*Atlanta, GA*J. Mack Robinson College of Business, GSU

April 2022

- Predicted the stability of a 4-node star Smart Grid based on Artificial Neural Network model
- Designed a 12-node input layer, 24-24-24-12 node hidden layer and a single node output layer model architecture and achieved an accuracy of 97.87%

Topic Analysis - BERT Tokenizer*Atlanta, GA*J. Mack Robinson College of Business, GSU

April 2022

- Tokenized 80% of the data using pre-trained 'BERT Based Uncased' model and trained the model using 'Bert for Sequence Classification'
- Achieved an accuracy of 50.84% on classifying the texts into the given 40 categories by running the model in batches of 3 for 5 epochs, optimized the model using Adam and tuned a scheduler to linearly adapt the learning rate

Walmart Sales Prediction – Predictive Analytics*Atlanta, GA*J. Mack Robinson College of Business, GSU

March 2022

- Pre-processed and visualized the ACF and PACF of the Walmart monthly sales data for further modeling
- Fitted the data using Auto-ARIMA, SARIMA and Ljung Box plot statistics for determining the model
- Predicted the monthly sales for 2012 using One-Step Ahead prediction and FB Prophet model

Web Scraping – Selenium*Atlanta, GA*J. Mack Robinson College of Business, GSU

October 2021 – December 2021

- Automated browser activity and scraped multiple pages for the required information using Selenium
- Created database by exporting the data using python to MySQL and analyzed the same through queries

Machine Learning Model Building - Python*Atlanta, GA*J. Mack Robinson College of Business, GSU

October 2021 – December 2021

- Cleaned, analyzed, and visualized a dataset for Loan Defaulter Prediction using Applicants' information
- Used feature engineering and iterated over hyper-parameters for improving performance of models
- Compared ROC curves, AUC, F1 scores and accuracies, achieved a maximum F1 score of 0.89 with Random Forest Classifier and trained the model to detect potential loan defaulters in the test data

CERTIFICATIONS

- **DataCamp: Certified Professional Data Analyst** December 2022
- **Microsoft Certified: Azure Data Fundamentals** September 2022
- **Microsoft Certified: Azure AI Fundamentals** September 2022
- **Certified Peer Educator – NASPA, USA** March 2022
- **Infosys Ltd. Foundation Training Program (FTP) Graduate** November 2019 – March 2020
- **Goldman Sachs Certified Data School Graduate** March 2019 – April 2019
- **Alteryx Core Certification** April 2019

HONORS AND AWARDS

- **Member of Beta Gamma Sigma – International Business Honor Society** December 2022
- **Member of Golden Key International Honor Society** November 2022