Sudarshan Paranjape

Boston, MA, 02215
857-200-4575 paranjape.su@northeastern.edu linkedin.com/in/sudarshanp4 github.com/sudarshan4120

Education

Northeastern University **Expected May 2025**
Masters of Science in Data Science *GPA: **3.8/4.0***
Relevant Courses: Intro to Data Management and Processing, Natural Language Processing

Mumbai University **May 2022**
Bachelor of Engineering in Electronics and Telecommunications *CGPA: **9.44/10***
Relevant Courses: Database Management, OOPM, Applied Mathematics

Technical Skills

Programming Languages: SQL, Python, R
Visualization / Developer tools: VS Code, Power BI, Tableau, Git
Libraries: Matplotlib, Seaborn, NumPy, Pandas, TensorFlow, PyTorch
DevOps: AWS cloud (EC2, S3, RDS, EFS), Snowflake
Databases: MySQL, Oracle DB, MongoDB, Scikit-Learn
Data Science/Analytics Techniques: Data collection, Regression, classification, Probability, Product Analysis, Statistical Analysis, Data Modeling, Bayesian inference

Experience

TryCatch Group **Jan 2023 - Jul 2023**
Data Science Intern *Mumbai, India*

- Orchestrated comprehensive data analysis of a sales dataset comprising **5M+** records using Python, encompassing data quality assessment, ETL and multivariate analysis
- Composed an interactive **Power BI** & **Tableau** dashboard to provide stakeholders with business strategy & actionable insights to enhance growth and sales, leading to a **15%** increase in product sales conversion rates
- Incorporated Power BI to perform data cleansing and Exploratory Data Analysis, creating compelling visualizations that improved identifying trends & data accessibility and reduced report generation time by 30%, enhancing user experience

Pacific Coatings **Feb 2022 - Dec 2022**
Data Science Intern *Mumbai, India*

- Spearheaded inventory management, streamlining operations and reducing stockpiling time of **250+** products by **15** days, resulting in enhanced inventory efficiency
- Teamed up with business stakeholders and subject matter experts in the cement industry to fine-tune time series forecasting models, adapting them to unique industry challenges, leading to a **25%** rise in monthly revenue.

Projects

Hotel Recommendation System | *Supervised ML* – Link **Nov 2023**

- Devised a comprehensive Hotel Recommendation system with Python, leveraging NLP, semantic Parsing on dataset scraped from *Booking.com* API comprising of **515,738** hotel reviews
- Conducted comprehensive factor analysis for detailed hotel executing sentiment analysis using **VADER** (NLTK), categorizing customer sentiments in 5 categories
- Developed predictive models like Decision Tree, Random Forest & Adaboost achieving Mean Squared Error of only **2%**

Fraud Detection in Financial Transactions | *Python, SQL* – Link **Oct 2023**

- Executed Credit Card fraud detection on a heterogeneous dataset comprising of **10,000+** transactions & conducted feature engineering, data cleaning using **SQL** Database& python libraries like Numpy, Pandas & Spacy
- Implemented Cuckoo filters with a random forest classifier, establishing a robust two-step mechanism for effective fraud detection in financial transactions with a test accuracy of **86%**

Optical Character Recognition for Sanskrit Language | *Python* – Link **Jun 2022**

- Built an accessibility-enhancing Deep Learning project for Sanskrit manuscripts in Devanagari Script while segmenting **32** words as input a robust Convolutional Neural Network (CNN) model
- Assembled a diverse custom dataset of both simple and intricate Sanskrit characters of over **100,000** images with more than **40** classes to facilitate robust model training and evaluation
- Spearheaded feature extraction and character classification achieving a testing accuracy of **94%** and training accuracy of **98%** on a custom-made dataset