SHUBHAM GOYAL

☎ (857) 370-1627 | Boston, MA

✈ goyal.shubh15@gmail.com | 🔗 Linkedin | ⊙ Github/@shu3hamiitkgp | Ⓜ Medium/@goyal.shubh15

EDUCATION

Northeastern University   Boston, MA
Masters of Science in Information Systems (Computer Software Engineering)   *Sep 2022 – May 2024*
Relevant Courses: Data Science, Big Data & Intelligent Analytics, Neural Modeling, Natural Language Processing

Indian Institute of Technology (IIT) Kharagpur   Kharagpur, India
Integrated Bachelor of Science, Master of Science in Geophysics   *2014 – 2019*

WORK EXPERIENCE

American Express   Bengaluru, India
Assistant Manager, Global Collections Data Science   *Nov 2021 – Aug 2022*

- **Icollect model revamp:** Spearheaded re-development of collections model to rank order highly collectible defaulters using Xgboost algorithm with Bayesian hyperparameter optimization
- Deployed model in production attaining ∼15% Gini uplift in benchmark metrics with ∼$1.2MM benefit in collections
- **High-Balance default case study:** Adjudged best presenter for in-depth analysis and default preventing recommendations, resulting in saving ∼$400K exposure at default

Analyst (Credit Risk Strategy), US Commercial limits   *Nov 2019 – Nov 2021*

- **Proactive Limit Increase Strategy:** Recognized as Analyst of the Quarter for leading first corporate portfolio campaign- 400+ clients received a total increment of ∼$73MM resulting in 11% increase in Y-o-Y spending
- **Bank-based Underwriting:** Launched a Global Capability for corporate portfolio with an incremental annual PTI impact of ∼$2MM, by systematically integrating bank statements with underwriting platforms. Formulated a risk index using attributes emerging from NLP-based intelligence of transactions in bank statements
- **Low-Risk Underwriting Strategy:** Re-designed strategy for small businesses setup for risk assessment, with an expected incremental spend impact of ∼$20MM annually and loss save of $0.2MM. Executed a 6-month A/B test pilot to formalize strategy, partnering with engineering and compliance teams to execute a phased solution
- **MIS for Commercial Business Portfolio:** Derived key success metrics to track customer performance for end-to-end MIS tracking of customers setup for financial gathering

Voksedigital Consultancy   Pune, India
Data Science Intern - Billboard Viewership Intelligence   *May 2017 – July 2017*

- Built a data pipeline utilizing Google Maps API to pull traffic and location data. Devised an algorithm that linked changes in journey duration to traffic density, to estimate the number of people on road
- Designed location-specific dashboard with round-the-clock population estimate, neighborhood type, and peak traffic hours

SKILLS/TECH STACK

Programming: Python | SQL | PySpark | SAS | Hive | Excel | R | C++ | Java | Bash | OOPs | Tableau
Data Science: Machine Learning | Deep Learning | Transfer Learning | Statistical Inference | NLP | Time Series
Cloud Tools: AWS | GCP | Snowflake | Docker | HuggingFace | Airflow | Docker | Git | FastAPI

PROJECTS

SEVIR Data Exploration Tool   *Big Data Project, NEU*

- Developed a data exploration platform to provide data retrieval service for satellite weather imagery data
- The application had decoupled microservices of FastAPI and Streamlit, which were dockerized and hosted on AWS EC2

Facial Age Classification And Sentiment Analysis   *Data Science Project, NEU*

- Leveraged CNN and ResNet50 models to classify facial images in 3 age segments achieving over 86% validation accuracy
- Scrapped ∼1K Amazon reviews with BeautifulSoup and got 96% sentiment classification accuracy leveraging Naive Bayes

Industrial Accident Classification   *Summer Research Project, IIT KGP*

- Developed a classification model on accident description text data using SVM, Random Forest & Maximum Entropy
- Achieved 87% prediction accuracy with the research aiding towards enhancing strategies to minimize accident occurrences

CERTIFICATIONS AND ACHIEVEMENTS

- Achieved an **All India Rank within Top 0.5%** in the JEE 2014, among the 1.3 million registered applicants
- **Certifications (Coursera)** : Machine Learning with Python, Data Analysis with Python, Data Visualization with Python, Databases and SQL for Data Science with Python