SHUBHAM GOYAL

📞 (857) 370-1627 | Boston, MA | ✈ goyal.shubh15@gmail.com | 🔗 Linkedin | 🌐 Portfolio website

EDUCATION

Northeastern University, Boston, USA — GPA: 3.6/4.0
Masters of Science in Information Systems (Computer Software Engineering) — *Sep 2022 – May 2024*
Relevant Courses: Big Data & Intelligent Analytics, Neural Modeling (Deep Learning), Network Science

Indian Institute of Technology (IIT) Kharagpur, India — GPA: 7.5/10.0
Integrated Bachelor of Science, Master of Science in Geophysics — *Jul 2014 – May 2019*

WORK EXPERIENCE

Verato Inc — McLean, VA
Data Science Intern - Bayesian Modeling — *Jun 2023 – Aug 2023*
- Developed and evaluated a probabilistic model for imputing missing data values in personally identifiable information
- Enhanced individual profile-matching algorithm leveraging bipartite graph networks and gradient boosting models

American Express — Bengaluru, India
Assistant Manager (Data Science), Global Collections — *Nov 2021 – Aug 2022*
- **Xgboost for Defaulter Collections:** 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:** Collaboratively launched a Platform Capability for corporate portfolio with an incremental annual PTI impact of ∼$2MM, by systematically integrating bank statements with underwriting platforms
- **Bank-based Risk-Index:** Formulated a risk index using attributes emerging from 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 dashboards of customers setup for financial gathering

SKILLS/TECH STACK

Programming: Python | SQL | PySpark | R | C++ | AWS | GCP | Docker | Airflow | Tensorflow | FastAPI | NetworkX
Data Science: Machine Learning | Deep Learning | Statistical Inference | NLP | Time Series | Transformers | Tableau

PROJECTS

Founder - GrocerEase: AI-Based Digitization of Groceries (Pitch Deck) *Husky Startup Challenge, NEU*
- Addressed a prevalent problem of forgetfulness and provided an efficient way to digitize purchases by just scanning receipts
- Ranked in Top-5 finalists in the startup incubator competing among 20 other startups for $10K funding

LES Weather Forecasting: Multi-Modal Architecture (Git) *Neural Modeling Project, NEU*
- Developed a classic Encoder-Decoder architecture using ConvLSTM and LSTM for weather forecasting at Lake Michigan
- Used satellite images and meteorological data to create the model attaining a recall of 0.85 on highly imbalanced data

DIME Network Analysis: Study of US Political Contributions (Git) *Network Science Project, NEU*
- Researched DIME data to formulate a network among the monetary contributions in US Federal Elections
- Leveraged network science principles to showcase scale-free properties, and ideology-based communities among contributors

TravelBud: AI Tool for Travel planning (Git) *Big Data Project, NEU*
- Developed a travel itinerary-generating application optimized based on user areas of interest, budget, and destination
- Leveraged RapidAPI for flight, hotels data, and OPENAI GPT3 for chat completion coupled in a cloud ecosystem

VaniVerse (Git) *Big Data Project, NEU*
- VaniVerse is a conversational AI platform that can be used to generate prompts for a given audio file
- OPENAI Whisper & GPT API are leveraged to process audio files and generate prompts simultaneously using Airflow