Ashish Dhiman

+1(404) 509-0254 | Atlanta, GA | ashish1610dhiman | ashish1610dhiman@gmail.com | Portfolio website

Education

Georgia Institute of Technology, Atlanta, USA **GPA: 4.0/4.0**
Master of Science in Analytics: Computational Data Analysis & Analytical Tools Track *Aug 2022 — Dec 2023*

Coursework: Machine Learning, Deep Learning, Regression, Computational Statistics, Uncertainty Quantification, Deterministic Optimization, Non Linear Programming, Optimisation methods in Finance, Financial Risk Management

Indian Institute of Technology, Kharagpur, INDIA **GPA: 8.32/10.0 | Class Rank 2**
B.Tech(Honours) in Aerospace Engineering, with Specialisation in Optimization Theory *Jul 2015 — May 2019*

Work Experience

American Express, US Consumer Credit Risk — Gurgaon, India

Assistant Manager - Data Science *Aug 2021 — Jul 2022*

- **Graph Network**: Created a Directed Cyclic Graph with consumers as nodes and shared trades as edges, using 250M rows of bureau trade-line data on AWS. This improved defaulter capture rate & saved **$2.5M** in annual credit defaults.
- **Xgboost for Covid deferrals**: Awarded **Sr. VP award** for developing a model to identify customers enrolled in Covid-19 relief programs of external creditors, by using tradeline level history of monthly balances and payments. Collaborated with 5 colleagues in Experian to implement the pipeline on Experian's infrastructure.
- **Feature Selection Research**: Implemented and tested Gradient Boosted Feature Selection & min-Redundancy Max-Relevance methods on big data of 30M records from US & Canada market using Spark & MapReduce.
- **Delinquency Index**: Used balance & delinquency time series to improve capture of high balance defaulters by 1.1%.

Analyst - Data Science *Aug 2019 — Jul 2021*

- **Customer Segmentation**: Predicted the external credit card with the highest card spend, using transfer learning and disambiguation logic on new accounts data. The predictions were then used to identify potential growth buckets.
- Awarded **Analyst of Quarter** for automation of 3 dashboards from bureau trigger data using UNIX crontab.
- **Resume Parsing**: Slashed resume screening time by 30% using using a combination of regex rules, zero-shot classifier and **Named Entity recognition** (NER) model and other Natural Language Processing techniques.
- Improved accuracy of an external payment prediction algorithm by 7%, using SMOTE to treat data imbalance.

Technical Skills

- **Programming languages:** C | Python | R | SQL | Scala | MATLAB | SAS | LaTeX | Bash | Excel | Hive | D3
- **Software/Frameworks:** AWS (EMR/S3) | PySpark | Keras | TensorFlow | Tableau | Hadoop | MapReduce | Yarn
- **Machine Learning:** XGBoost | A/B Testing | Optimization | Forecasting | Statistical Modelling | Bayesian Modelling

Projects & Publications

Deep Evidence Regression for Credit Risk *Uncertainty Quantification class project* (GitHub)
- Enhanced Deep Evidence Regression, an uncertainty-aware deep learning model, for credit risk applications.
- Broadened the model's capabilities to effectively handle Weibull distributed target variables like Loss Given Default.

Food Recognition and Recommender System *Machine Learning class project* (GitHub)
- Used ResNet50 CNN with transfer Learning to achieve a Top-5 classification accuracy of 91% on food images.
- Applied Collaborative Filtering and SVD to recommend the best recipes to a user with an overall NDCG of 90%.

Forecasting Brent futures price structure with HMM *Computational Statistics class project* (GitHub)
- Developed a predictive model using Principal component analysis and Hidden Markov Gaussian Mixture model, enabling accurate detection of shifts in price structure of Brent futures in the commodity market with 78% accuracy.

Time Series Clustering on S&P 500 stocks (Data Driven Investor, Medium article)
- Analyzed Covid-era rebound trajectories by clustering SP 500 securities using Dynamic Time Warping distance.

Index tracking portfolio optimisation *Bachelors Thesis, IIT KGP* (GitHub)
- Solved Enhanced Index tracking (Mixed Integer) with Heuristic kernel search to improve the performance by 12%.
- Used Nonnegative Principal Components and Nonnegative Matrix factorization to remove short-term volatility by 8%.

Ranking of Mutual Fund Houses *Inter Hall Data Science, IIT KGP* (GitHub)
- Led a 20-member team in ranking Mutual Fund Houses, employing LSTM and VAR for net asset value forecasting.