SHREYA VONTELA

Tel: +1 (510)-988-1801 | **Email** : shreyavontelaberkeley@gmail.com | LinkedIn | Portfolio | GitHub

PROFESSIONAL SUMMARY

Data Scientist with 7 years of experience driving analytics, modeling, and AI initiatives across financial service sector. Expertise in Portfolio Analytics, Tableau Visualizations, Predictive Modeling, NLP, GenAI and Agentic AI, with a strong track record of building end-to-end modeling pipelines for borrower risk, delinquency, complaint analysis, and control automation. Passionate about translating complex data into actionable intelligence.

EXPERIENCE

Capital One (Discover Financial Services), *Principal Data Scientist* **Jan 2023 – Oct 2025** *(WA, USA)*

- Built *Predictive ML models* for borrower delinquency and early payoff risk using LightGBM and SHAP interpretability; enabled proactive servicing campaigns and reduced delinquency progression by 22%.
- Developed an *NLP model* to automatically read and categorize borrower complaints, routing potential compliance issues to the correct servicing teams; reduced manual review time by 70%
- Built a GenAI-powered *Loss Mitigation Agent using RAG + LLM* (LangChain) to summarize borrower cases, propose modification strategies, and draft compliant communications; cut analyst handling time by 45%.
- Developed *GenAI COPILOTS* leveraging RAG, LangChain, and Snowflake to interpret regulatory controls and enable *natural language–to–SQL querying* for business users; automated exception monitoring and self-service analytics, reducing control-testing and ad-hoc analysis time by 50%.
- Achieved a 95% reduction in fraudulent applications and sharply mitigated synthetic identity fraud risk by implementing a XGBoost anomaly detection model utilizing feature engineering on Adobe clickstream data.

Leadership

- Managed a distributed team of *eight data scientists and offshore analysts*, overseeing project design, model governance, and executive-level reporting.
- Subject-matter expert for Servicing Analytics, leading cross-functional initiatives across Credit Risk, Compliance, and Data Engineering to unify data sources and modeling pipelines.
- Partnered with Model Risk Management and Regulatory Policy to define AI adoption roadmaps, ensuring interpretability, documentation, and compliance alignment.

Aspen Capital, *Data Scientist* **June 2022 – Jan 2023** *(OR, USA)*

- Working on designing a strategy to offer second mortgages and credit cards to existing borrower pool by leveraging TransUnion and CoreLogic datasets. *(Modeling Default probability and Lending rates)*

PennyMac, *VP of Data Science* **May 2021 – May 2022** *(CA, USA)*

Loan Status Disposition Model

- Built an end-to-end framework of Light GBM classification models using real time credit variables to predict the loan disposition (six possible outcomes) over next 24 months. (~$25M revenue in the GNMA EBO loan space)

Prepay Default Model

- Built a neural network on a large dataset (~1 TB) to predict the conditional probabilities of prepayment and various delinquencies to manage interest and credit risk of the mortgage portfolio.
- Performed extensive EDA to uncover how different loan specific and macroeconomic estimators (from Moody's) have a nonlinear relationship with the borrower prepayment behavior. (Ex. Unemployment, Hurricanes)

CalPERS, *Investment Associate Intern* **Sept 2020 – Nov 2020** *(CA, USA)*

- Built a *linear programming model* to design an index future roll strategy that minimizes the roll cost. *[Code]*
- Calculated incremental financing costs for stressed liquidity portfolio by comparing collateral REPO & OIS rates.

JPMorgan Chase & Co, *Quant Researcher (Credit Risk)* **April 2019 – March 2020** *(India)*

- Developed a *quadratic regression model* to capture the non-linearity between discount factor and discount rate.
- Clustered the loan portfolio into 13 segments by using *K means algorithm and EDA techniques.*

Goldman Sachs, *Asset Management Intern* **April 2018 – June 2018** *(India)*

- Interned as a pension fund advisor for a $2B AUM client in a multi-asset class macro investment strategy team
- Built *excel based tool* to spot investment opportunities across 26 country indices using earnings & sales values

Bank of America Merrill Lynch, *Developer, Best employee Award* **July 2015 – June 2017** *(India)*

- Developed a framework for real-time trade data settlement & reconciliation of global Equity & Forex derivatives
- Automated the reconciliation of end of day transactions of banks nostro accounts *(Saved 300+ man hours)*

FREELANCE CONSULTING PROJECTS

AI-Assisted Endoscopic Guidance System *[Code]*

- Led development of an AI-assisted endoscopic guidance system combining a *YOLOv8-based anatomical detection model* and a CNN "True North" orientation classifier to label throat structures and guide scope positioning in real time; achieved mAP 0.78, 92 % orientation accuracy and optimized models for mobile device deployment.

Trading strategy for tourism industry in USA, *Data Science Team at Kayrros [Code]*

- Built a neural network with adaptive LR & time series cross validation to predict casino stock price (9.6% return)

Academic/ Personal Projects/ Blogs

- Performed *Geospatial analysis* & built a house price prediction model to find top investable zips in USA *[Code]*
- Modeled the probability of default of *Lending Club loans* and calibrated the XGBoost model *[Code]*
- Performed *PCA on 130 features & used XG Boost Classifier* to model trading for Jane St Kaggle Project *[Code]*
- Built a convex portfolio optimization model to perform asset allocation based on Sharpe ratio (*87% return*)

EDUCATION

University of California, Berkeley – Haas School of Business **Expected March 2021**
Master of Financial Engineering – Specialization in Machine Learning *Berkeley, CA, USA*

Indian Institute of Management, Lucknow, India *(Institute Ranker)* **2017 – 2019**
MBA: Master of Business Administration (Exchange at EMLYON Business School, Lyon) *India and France*

Birla Institute of Technology and Science (BITS), Pilani, India **2011 – 2015**
Bachelor of Technology in Electronics and Instrumentation Engineering *India*

CORE SKILLS

Machine Learning and AI:
Regression & Classification · Ensemble Methods (Random Forest, XGBoost, Bagging, Boosting) · SVM · KNN · Clustering · CNNs & RNNs · Deep Learning · Time Series Forecasting · A/B Testing · NLP (Text Classification, Tokenization, Embeddings, Transformers) · Generative AI (RAG, Prompt Engineering, Agentic AI Systems)

Programming and Tools:
Python · TensorFlow · Scikit-learn · PyTorch · Amazon SageMaker (AWS) · SQL · Git · Linux · Excel · Tableau

Model Lifecycle and ML Ops:
Feature Engineering · Model Evaluation & Monitoring · API Integration (Flask/FastAPI) · Model Optimization (ONNX, TorchScript) · Workflow Automation (Airflow) · Versioning & Reproducibility (MLflow)

Certifications:
Deep Learning (Coursera), TensorFlow in practice (Coursera), R programming (Udemy)

Domain Expertise:
Credit Risk Modeling · Loan Portfolio Analytics · Fraud Detection · Loss Mitigation Modeling · Regulatory Compliance Automation