# Manoj Kumar Bashaboina

+1(240)4380867  manoj7@umd.edu  linkedin.com/in/manojkumarbashaboina  bashaboinamanojkumar

## Education

**University of Maryland, College Park, USA**  **Aug 2024 – May 2026**
*Master of Data Science*  *GPA: 3.9/4.0*

**Vasavi College of Engineering, Hyderabad, India**  **Aug 2019 – May 2023**
*Bachelor of Technology in Information Technology*  *GPA: 8.35/10*

## Work Experience

**University of Maryland, College Park**  **March 2025 – Current**
*Do Quantum (Graduate Researcher)*  *Maryland, US*

- Developed portfolio optimization system for S&P 500 stocks, processing 750k+ rows of data with custom reshaping and Parquet caching, reducing reload times by 5x using Python, pandas, and yfinance APIs.
- Optimized expected returns via LightGBM ML on momentum, fundamentals, and NLP sentiment, integrating Black-Litterman for stable $\hat{\mu}$ estimates across 500+ tickers.
- Implemented VaR/CVaR risk metrics and Mean-CVaR optimization with cvxpy, achieving 20% downside risk reduction in backtests and Sharpe ratio ¿1 through walk-forward validation.
- Extended to hybrid quantum-classical framework using CVaR, QAE, VQE, and QUBO for accelerated estimation, incorporating full pipeline from data collection to quantum circuit design and backtesting.

## Skills

**Languages & Frameworks:**  Python, Java, C, SQL, JavaScript, HTML/CSS, Tableau, Git, Docker, Selenium
**Infrastructure:** AWS (EC2, S3, Load Balancers, SQS), CI/CD
**AI & Data Systems:** LLMs, RAG

## Projects

**Custom Bitcoin Chatbot with Ollama Python**  **May 2025**

- Developed an advanced Bitcoin analysis chatbot using a RAG framework with Ollama and LLaMA3, integrating real-time market data, technical indicators, and sentiment signals for actionable insights.
- Engineered robust data pipelines to aggregate real-time Bitcoin prices from CoinGecko API and financial news from NewsAPI, applying NLTK sentiment analysis, caching, and computing RSI, MACD, and Bollinger Bands.
- Integrated a TensorFlow/Keras LSTM for short-term Bitcoin forecasting and implemented a FAISS-based vector retrieval system using nomic-embed-text embeddings, with optimized chunking for low-latency semantic search.
- Deployed a fully containerized Streamlit web application using Docker, featuring interactive dashboards, real-time Plotly visualizations, and automated trading signal generation.

**Vehicle Sales Analysis**  **Dec 2024**

- Conducted end-to-end analysis of vehicle sales data, including data preprocessing, outlier removal (IQR and Isolation Forest), and hypothesis testing.
- Applied regression models (XGBoost, Random Forest, LightGBM) with GPU acceleration, achieving an $R^2$ score of 95%. Evaluated model performance using metrics such as RMSE and MSE.
- Performed hyperparameter tuning using GridSearchCV and K-fold cross-validation.
- Conducted correlation analysis and ensemble modeling to extract actionable insights and enhance predictive performance.

**Automated Grape Plant Leaf Disease Detection**  **May 2023**

- Developed a deep learning-based pipeline for grape plant disease detection, employing YOLOv8 for initial disease localization to identify bounding boxes around affected areas on grape leaves.
- Utilized the Segment Anything Model (SAM) for precise segmentation of these regions, isolating the diseased portions for further analysis.
- Performed self-labeling of the PlantVillage dataset using labeling tools to annotate the leaves, optimizing the YOLOv8 model for improved disease localization accuracy.
- Built an end-to-end automated system integrating YOLOv8, SAM, and CNN for disease classification, enabling early-stage disease detection and supporting effective farm management and prevention strategies.