

HARSH MAHESHWARI

Senior Data Scientist | Quantitative Researcher | Algorithmic Trader | ML Engineer | Automation Specialist

Email: harshmaheshwari3110@gmail.com **Phone:** +91 9834470295 **LinkedIn:** linkedin.com/in/Harsh-Maheshwari **GitHub:** github.com/Harsh-Maheshwari **Portfolio / Home Page:** Decide With ML

PROFESSIONAL SUMMARY

Innovative and results-driven **Quantitative Trader, Data Scientist, and Machine Learning Engineer** with 5+ years of experience spanning algorithmic trading, supply chain ML, e-commerce intelligence, regulatory data automation, and academic AI research. Proven track record of **building and deploying production-grade ML and deep learning systems** at scale, designing **alpha-generating quantitative trading strategies** with Sharpe ratios exceeding 4 in live index option markets, and leading cross-functional teams to deliver high-impact data products.

Deep expertise in **mid-to-high frequency options trading, time series forecasting, computer vision, NLP, distributed computing, and MLOps**. Adept at translating complex statistical models and analytical insights into actionable strategies for both technical teams and business stakeholders. Passionate about applying AI/ML techniques to financial markets and real-world systems for a sustainable, measurable competitive edge.

Core identity: **Data Scientist, Automation & ReUsability Freak** — obsessive about building systems that generalize, scale, and compound value over time.

WORK EXPERIENCE

Sr. Vice President — Data Science & Quantitative Research

Mastertrust | Bangalore, India | January 2024 – Present

Leading quantitative research and data science for a multi-strategy algorithmic trading desk managing portfolios exceeding ₹100 crore in Indian equity and index option markets.

- Designed and deployed **mid-to-high frequency options trading strategies** for Indian index option markets using machine learning algorithms and advanced statistical modeling, achieving an **intraday Sharpe ratio greater than 4**.
- Engineered deep learning models — including **LSTMs, Informer, BERT, and Random Forest** — to **predict implied volatility (IV)**, forecast price direction, and extract **intraday market sentiment** from historical price action, volatility surfaces, and option Greeks.
- Built a **BERT-based IV prediction model** for understanding intraday sentiment of the market using historical price, IV, and Greeks information.
- Developed **real-time market regime classification systems** enabling adaptive deployment of strategies based on risk-reward setups, volatility conditions, and liquidity dynamics.
- Architected **automated decision-making algorithms** for both **directional and non-directional trading**, incorporating volatility skew, open interest dynamics, and custom quantitative indicators to gain competitive edge.

- Applied **reinforcement learning and Bayesian optimization** for dynamic portfolio allocation, optimizing hedging strategies, position sizing, and risk management across multiple trading books.
- Led the development of a **scalable backtesting framework** incorporating expectancy analysis, execution charges, slippage modeling, and capital utilization efficiency metrics.
- Spearheaded deployment of **high-performance data pipelines** for processing multi-terabyte options market data, leveraging parallel computing, feature engineering, and anomaly detection techniques.
- Helped build a **multi-strategy algorithmic trading system** handling investment portfolios exceeding **₹100 crore**, ensuring robustness, compliance, continuous alpha generation, and alpha degradation monitoring.
- Mentored and led a team of quant researchers, data scientists, and developers, fostering a culture of innovation in AI-driven and advanced statistical trading strategies.

Key Technologies: Python, LSTMs, Informer, BERT, Random Forest, Options Greeks, IV Modelling, Volatility Surface, Bayesian Optimization, Reinforcement Learning, Parallel Computing, Backtesting Frameworks

Manager — Data Science

GoGlocal | Mumbai, India | March 2023 – October 2023

Led data science and ML engineering for a cross-border e-commerce intelligence platform serving global marketplaces.

- Developed and scaled **multi-channel e-commerce strategies for cross-border trade**, automating workflows for **1,000+ SKUs** using advanced data scraping, NLP, and machine learning — reducing manual effort by **50%** and improving operational efficiency.
- Led the cloud deployment and scaling of a **merchant-facing analytics product** using **Ray and AWS** for distributed processing, generating real-time **sales forecasts, inventory management insights, and expected profit margins per SKU** across Amazon, eBay, Walmart, and Lazada — increasing revenue estimation efficiency by **30%**.
- Spearheaded development and scaling of **large-scale web scraping pipelines** for Amazon websites across multiple countries, ensuring high data quality and real-time updates for competitive intelligence.
- Developed **time series forecasting models** for demand prediction and inventory planning using machine learning and **Bayesian inference**, leading to a **20% reduction in inventory holding costs** and improved SKU-level replenishment strategies.
- Designed and implemented **NLP-driven algorithms for SKU performance ranking**, leveraging semantic similarity, attribute extraction, and keyword optimization — improving product visibility and conversion rates.
- Built and deployed **AI-driven product listing enhancement tools** incorporating **LLMs, Stable Diffusion**, and fine-tuned OpenAI models for background generation, keyword extraction, and natural language generation, enabling automated content creation and faster product listings.
- Developed **trend-tracking models** to monitor marketplace price fluctuations and demand shifts, integrating unsupervised learning and pattern recognition to improve pricing strategies.
- Empowered clients with **performance metrics for merchant SKUs**, benchmarked against similar SKUs on various global marketplaces, resulting in optimized listings and improved sales performance.

Key Technologies: Python, Ray, AWS, Distributed Computing, Web Scraping, Selectolax, NLP, Stable Diffusion, OpenAI API, Time Series Forecasting, Bayesian Inference, XGBoost, LightGBM, Scikit-Learn, FastAPI, Streamlit

Data Scientist

Blue Yonder | Bangalore, India | April 2022 – March 2023

Part of a 15+ member team delivering intelligent, cloud-native SaaS ML solutions for supply chain management across omnichannel commerce.

- Contributed to a **15+ member team** delivering cloud-native SaaS ML solutions across all areas of **supply chain** in omnichannel commerce: logistic optimization, store fulfillment capacity, delivery date estimation (pick, pack, ship, transit), sales returns & replenishment forecasting, inventory estimation, markdown, and stockout avoidance.
- Developed and deployed **large-scale deep learning models** for demand forecasting, inventory optimization, and order fulfillment using **TensorFlow, Keras, and TFX pipelines on Kubeflow**.
- Built and optimized **time series forecasting models** for supply chain applications, processing **5TB+ of high-dimensional, noisy data** to improve delivery date estimates, stockout prediction, and replenishment planning.
- Used **Apache Beam on Dataflow** for bulk data ingestion, feature engineering, and predictive modeling across omnichannel commerce operations, with **BigQuery** as the primary data warehouse.
- Collaborated on an **auto-scaled Kubernetes platform** with a multi-tenancy solution to share cloud resources for effective cost-cutting and improved margins.
- Integrated statistical modeling and deep learning techniques to enhance decision-making across fulfillment capacity, **markdown pricing**, and inventory allocation — improving forecasting accuracy.
- Developed **interpretable ML solutions and business metrics**, ensuring complex analytical insights were translated into actionable strategies for non-specialized stakeholders.
- Self-motivated team player with extensive international experience and proven ability to take charge of ambiguous, large-scale ML problems.

Key Technologies: Python, TensorFlow, Keras, TFX, Kubeflow, Apache Beam, Dataflow, BigQuery, Kubernetes, Docker, Time Series Analysis, MLOps, kedro, Airbyte, DBT, Feast, Airflow, mlflow, Prefect

Enterprise Data & Intelligent Automation Engineer

ANZ Bank | Bangalore, India | June 2021 – March 2022

Responsible for regulatory data pipelines and end-to-end process automation for one of Australia's largest banks.

- Designed and managed **ETL pipelines for APRA regulatory data** across **70+ source systems**, leveraging **IBM DataStage, Teradata SQL, and Control-M**.
- Developed an **end-to-end data processing automation tool** using **Robot Framework and Python**, reducing development and testing time by **50%**.
- Responsible for creating and managing **DataStage and Control-M jobs in production**, ensuring correct and reliable data ingestion from ASCII, semi-ASCII, and EBCDIC file formats.
- Led the **Grads4Tech initiative**, driving technology adoption, innovation culture, and ML/DevOps/automation-first approaches within the bank.
- Applied **machine learning, DevOps, and automation-first thinking** to modernize data engineering workflows in a large regulated institution.

Key Technologies: IBM DataStage, Teradata, Control-M, Robot Framework, Python, SQL, ETL, Unix Shell Scripting, DevOps

Business Intelligence Analyst (Internship)

Axis Bank | WFH, India | April 2020 – June 2020

- Built a **micro-service for predicting customer income** using nominatim API, Scikit-Learn, Docker, and Kubernetes, achieving prediction accuracy of **over 90%**.
- Engineered **new geospatial features** from customer address data using the **K-Nearest Neighbor algorithm** and geocoding services.
- Built an **XGBoost regression model** leveraging customer history, transaction data, and engineered geographic features.
- Developed a **Django web interface** containerized in Docker for convenient deployment and stakeholder access.

Key Technologies: Python, XGBoost, Scikit-Learn, Django, Docker, Kubernetes, nominatim API, KNN, Geocoding

AI Research Scholar

School of Computing, National University of Singapore (NUS) | Singapore | November 2019 – December 2019

Research under Prof. Bryan Hooi on dynamic graph neural networks and temporal link prediction.

- Developed a **Temporal Attention model** for node classification and link prediction, leveraging dynamic graph structures to improve predictive accuracy.
- Analyzed and optimized models including **Node2Vec, Temporal Node Embedding, Attention Walk, TMF, CTDNE, and BANE** — identifying inefficiencies and enhancing performance.
- Improved **link prediction** capabilities, achieving an **AUC of 86%** on the College Messages dataset by refining embedding techniques and integrating novel temporal attention mechanisms.
- Analyzed the temporal dependency of graphs on time using dynamic models, providing insights into how graph structure evolves with time.

Key Technologies: Python, Graph Neural Networks, Node2Vec, Temporal Node Embedding, Attention Walk, TMF, CTDNE, BANE, NetworkX, Deep Learning

PROJECTS & INTERNSHIPS

Computer Vision Lead — Autonomous Drone (IARC / Barcelona Smart Drone Challenge)

Innovation Cell, IIT Bombay | India | October 2019 – March 2020

Under guidance of Prof. Dhwani Shukla. Team of 15 students from different departments innovating at the intersection of robotics, CV, and autonomous systems.

- Designed and developed an **autonomous fixed-wing drone** for search and rescue missions customized for the IARC (International Aerial Robotics Competition) and Barcelona Smart Drone Challenge.
- Integrated a **ROS-enabled camera system** for real-time detection and recognition of characters from a height of **50 meters** on a flying drone.
- Achieved **94% accuracy** using **transfer learning on MobileNetV2** architecture trained on a custom aerial dataset.
- Developed a **multi-core zoom-in characters recognition model** based on the Char74 dataset using **OpenCV & Keras**.

- Led computer vision development within a multidisciplinary team combining mechanical, electrical, and software engineering.

Key Technologies: Python, OpenCV, Keras, MobileNetV2, Transfer Learning, ROS, Char74 Dataset, PyTorch

Time Series Analysis — Gas Sensor Modelling

Guide: Prof. Sharad Bhartiya, IIT Bombay | August 2019 – November 2019

- Analyzed concentration of **CO and Ethylene** using sampling time analysis, autocorrelation, cross-correlation, PSD analysis, and DFT analysis on time-varying sensor data.
- Predicted concentration of dynamic gas mixture using **ARX, ARMAX, ARIMA, and non-linear ARX models**.
- Achieved accuracy of **92.96% for CO** and **92.75% for Ethylene** in 1-step ahead prediction.
- Analyzed system behavior under **1%, 10%, and colored noise** using Impulse estimate, Step estimate plots, and ETFE plots.

Key Technologies: Python, MATLAB, ARX, ARMAX, ARIMA, Time Series Analysis, Signal Processing, DFT, PSD

Data Analyst — Solar PV Diagnostics (Internship)

PV-Diagnostics | Mumbai, India | November 2018 – December 2018

- Designed an **SQLite database** for solar project data with multiple parameters and large volumes of solar I-V and power data-points using Python.
- Developed **new physical metrics** for solar plant performance based on the translation equation and business information.
- Performed data cleaning, innovative **feature engineering**, and EDA to derive new insights and detect outliers.
- Classified different **types of defects in solar panels** using logistic regression, achieving **85% accuracy**.

Key Technologies: Python, SQLite, Pandas, Scikit-Learn, Logistic Regression, EDA, Feature Engineering

Natural Language Processing — Hate & Offensive Speech Detection

Guide: Prof. Sunita Sarawagi, IIT Bombay | October 2018 – November 2018

Team of 4 students.

- Built a classification system for Twitter tweets labeled as 'Hate Speech', 'Offensive Language', or 'Neither' using a deep feed-forward network.
- Cleaned data using **NLTK** and represented word frequency distributions as **Word Clouds**.
- Modified **GloVe pre-trained Word Embeddings** for the dataset; mapped all vocabulary to 50-dimensional vector space.
- Trained an **8-layer deep feed-forward network**, achieving training accuracy of **84.30%** and testing accuracy of **82.30%** in 30 epochs.

Key Technologies: Python, NLTK, GloVe, Word Embeddings, Deep Learning, NLP, TensorFlow/Keras

Computer Vision — Image Classification & Object Detection

Guide: Prof. P. Balamurugan, IIT Bombay | August 2018 – February 2020

- Studied and implemented multi-layer perceptrons (MLP) and Convolutional Neural Network (CNN) architectures.
- Deployed **DenseNet, Dual Path, GoogleNet, MobileNet, ShuffleNet, ResNet, ResNext, SEnet** on CIFAR10 dataset using **PyTorch**.
- Deployed FFN with 7 and 5 layer depths on MNIST, achieving accuracy of **98.4%** and average loss of **0.0465**.
- Developed a **Parking Spot Detection** model from real-time camera feed using **OpenCV and RetinaNet**.
- Built **Emotion Recognition, Face Detection, and Face Recognition** models using the face recognition library.

Key Technologies: PyTorch, OpenCV, DenseNet, ResNet, MobileNet, RetinaNet, CIFAR10, MNIST, Face Recognition

Chemical Unit Operations (Exchange Semester)

Guide: Prof. Ivan Hundebøl, Danmarks Tekniske Universitet (DTU), Denmark | January 2019 – June 2019

- **Liquid-Liquid Extraction:** Determined efficiency of the mixer settler across 4 conditions based on 2 emulsion types and 2 rotational speeds. Analyzed mass balance, mass transfer coefficients, and compared funnel test results with mixer settler.
 - **Gas Flow in Pipes:** Investigated fan properties including characteristic curve, power consumption, and efficiency. Verified Bernoulli's equation; measured gas flow rate using venturi, orifice, and pitot tube.
 - **Batch Distillation:** Performed separation of water and ethanol with constant reflux ratio. Conducted heat analysis on the system using Rayleigh's modified equation.
-

EDUCATION

B.Tech. in Chemical Engineering

Indian Institute of Technology Bombay (IIT Bombay / IIT Mumbai) July 2017 – May 2021 | CGPA: 7.7

- **All India Rank (AIR) 735** among 1.3 million candidates in JEE Main examination (May 2017)
- **All India Rank (AIR) 1382** among 0.2 million candidates in JEE Advanced examination (May 2017)
- Received **Letter of Appreciation for exceptional academic performance** from HRD Minister Smriti Irani in CBSE Class 10

Semester Exchange — Chemical Engineering

Danmarks Tekniske Universitet (DTU), Denmark January 2019 – June 2019

- Represented IIT Bombay as a **Chemical Engineering Exchange Scholar** at one of Europe's leading technical universities.
 - Published a **Blog Series of 8 posts** documenting time in Europe, travel, food, and studies at Denmark Technical University.
-

TECHNICAL SKILLS

Programming Languages

Python | SQL | R | MATLAB | Julia | Unix Shell Scripting | Make

Data Manipulation & Analysis

Numpy | Pandas | Modin

Machine Learning & Deep Learning

TensorFlow | TensorFlow Extended (TFX) | PyTorch | Keras | Scikit-Learn | XGBoost | LightGBM | CatBoost | TSFresh | Statsmodels | Feature Engine | Walkforward Analysis | Bayesian Optimization | Graph & Network Analysis | NetworkX

Quantitative Trading & Finance

Options Pricing Models | Options Greeks Analysis | Implied Volatility (IV) Modelling | Volatility Surface | Backtesting Frameworks | Expectancy Analysis | Slippage Modelling | Position Sizing | Risk Management | Directional & Non-Directional Strategies | Mid-to-High Frequency Trading | Reinforcement Learning for Portfolio Optimization

Time Series Analysis

ARIMA | ARX | ARMAX | Non-linear ARX | Autocorrelation | Cross-correlation | DFT | PSD | Signal Processing | Spectral Analysis

Computer Vision

OpenCV | MobileNetV2 | RetinaNet | DenseNet | ResNet | ResNext | GoogleNet | ShuffleNet | SNet | ROS (Robot Operating System) | Transfer Learning | Face Recognition | Object Detection

Natural Language Processing

GloVe Word Embeddings | BERT | LSTM | Semantic Similarity | Keyword Extraction | Attribute Extraction | Stable Diffusion | OpenAI API | LLMs

Big Data, Cloud & Infrastructure

AWS Ecosystem | GCP | Apache Beam | Dataflow | Apache Spark | SparkML | Dask | Vaex | Ray (Distributed Computing) | Kubernetes | Docker | Kubeflow | Gunicorn | Uvicorn

Databases

BigQuery | PostgreSQL | MySQL | MongoDB | Redis | SQLite | Teradata

MLOps & Data Pipeline Tools

kedro | Airbyte | DBT | Feast | Airflow | mlflow | Prefect | Jenkins | GitHub Actions | Git

ETL & Legacy Data Systems

IBM DataStage | IBM InfoSphere Information Server | Teradata | Control-M | Robot Framework | n8n Automation

API & Web Development

FastAPI | Streamlit | Django | Selectolax

Data Visualization & Business Intelligence

Grafana | Matplotlib | Seaborn | Holoviews | HVPlot | Google Analytics | Google Data Studio | Google AdWords | Atlassian Suite | Knime | MS Excel | Jupyter

DevOps & Version Control

Git | Docker | Kubernetes | Jenkins | GitHub Actions | CI/CD Pipelines

CERTIFICATIONS & COURSES

Professional Certifications

- **The Data Scientist's Toolbox** — Johns Hopkins University (Coursera)
- **R Programming** — Johns Hopkins University (Coursera)
- **Getting and Cleaning Data** — Johns Hopkins University (Coursera)
- **Exploratory Data Analysis** — Johns Hopkins University (Coursera)
- **Data Science Foundations: Data Engineering** — Coursera
- **Google Analytics Individual Certification** — Google (August 2020)
- **Big Data with Hadoop** — Eckovation

Relevant Academic Coursework (IIT Bombay)

Machine Learning | Optimization Models | Differential Equations | Linear Algebra | Data Analysis | Calculus | Numerical Analysis | Computational Methods Lab | Computer Programming | Process Modelling & Identification | Process Control | A.I. in Process Engineering | Chemical Engineering Thermodynamics | Transport Phenomenon

SOFT SKILLS & COMPETENCIES

- **Self-Motivation & Initiative:** Proactively takes ownership of tasks and drives projects independently from ambiguous starting points.
 - **Analytical Thinking:** Strong ability to break down complex, multi-dimensional problems and implement data-driven solutions balancing computational complexity, business constraints, and interpretability.
 - **Adaptability:** Skilled at managing multiple stakeholders and adapting to evolving project requirements in dynamic environments.
 - **Effective Communication:** Capable of articulating technical insights to both specialized and non-specialized audiences — translating model outputs into business value.
 - **Collaboration & Leadership:** Proven experience leading cross-functional, cross-geographic teams. Track record of leading and mentoring quant researchers, data scientists, and developers.
 - **Agile Team Management & Storytelling**
 - **Critical Thinking & Hypothesis Generation**
 - **Quantitative Research**
 - **Attention to Detail & Fast Learner**
 - **Interpersonal Skills:** Strong communication, active listening, and ability to build relationships across teams and seniority levels.
 - **Team Player:** Flexible and collaborative; comfortable working across cross-functional and cross-geographic teams.
 - **Self-Confidence & Professionalism:** Consistent professional conduct with confidence in presenting and defending ideas.
 - **Time Management & Cooperation:** High sense of prioritization and reliability in delivering on commitments.
-

AREAS OF INTEREST

Data Analysis & Modeling | Machine Learning | Deep Learning | Big Data | Computer Vision | Time Series Analysis | Graph & Network Analysis | Quantitative Finance | Natural Language Processing | Algorithmic Trading | Distributed Systems | MLOps

EXTRACURRICULAR ACTIVITIES

Hackathons

- **Top 10 team** at the **Intel Python Hack Fury AI Hackathon** — IIT Bombay (2019)
- **2nd Runner Up** among 30+ teams from IIT Bombay in the **Andro-NG Android App Development** competition (2018)
- **2nd Position** in the **Chemical Process Simulation Hackathon** using ASPEN software (2018)

Volunteer & Leadership

- **Mentor, E-Cell (IIT Bombay):** Mentored three teams of freshmen for **EnB-Buzz**, an Innovate Start-up Idea competition. Helped teams identify problems, develop solutions, and build Business Model Canvas for their startups.
- **Organizer, Mood Indigo (IIT Bombay):** Worked with a team of volunteers to enhance the experience of student guests from colleges across India at Mood Indigo, the annual cultural festival of IIT Mumbai.
- **Blog Series — Exchange Semester:** Published a blog series of **8 posts** documenting European travel, food, culture, and life at Denmark Technical University.

Cultural & Sports

- **1st Position** in Literary Arts General Championship for Hostel 2, IIT Bombay (2017)
 - **3rd Position** in Street Play Arcade, IIT Bombay (2017)
 - **Gold Medal** in the Tug of War team game, General Championship for Hostel 2, IIT Bombay (2018)
 - Successfully completed **yearlong Volleyball training** under the NSO program of IIT Bombay (2017)
 - **Planned and travelled across 10+ countries and 15+ cities in Europe** on a tight budget during exchange semester (2019)
-

PHILOSOPHY & BELIEFS

"Invest in the truly important things in life. Remember, everything that is worth something costs something."

- **Persistence & Patience:** "Keep On Keeping On"
 - **Go Giver Philosophy:** Value | Compensation | Influence | Authenticity | Receptivity
-

References and LinkedIn recommendations available upon request.