HARSH MAHESHWARI

Data Scientist, Automation & ReUsability Freak

in Harsh-Maheshwari Harsh-Maheshwari @ harshmaheshwari3110@gmail.com 9834470295

WORK EXPERIENCE

Sr. Vice President Data Science & Quantitative Research
Jan '24 - Present

Mastertrust Bangalore, India

- Leading the development of **mid-high frequency** options strategies for Indian markets using machine learning algorithms and statistical methods.
- Building a BERT model for **iv prediction** and understanding intraday sentiments of the market using historical price, iv and greeks information
- Building quantitative systems to understand market scenario and make predictions on the fly to influence decision making for deployment of trading strategies
- Building algorithms to make **similarity, directional and non directional** decisions to gain a competitive edge using option greeks and custom quantitative indicators resulting in intraday strategies with **sharpe** of more than **4** in index option markets
- Helping build an algorithmic trading system capable of managing investments exceeding **hundred crore** in Rupees.

Manager Data Science
March '23 - Oct '23

GoGlocal Mumbai, India

- Developed multi-channel e-commerce strategies for cross-border trade. Scaled the system to **1000+ SKUs** utilizing advanced data scraping and machine learning to ensure reduction in human effort.
- Leading the cloud deployment and scaling using **Ray** of a merchant facing product that generates actionable plans for clients, including **sales forecasts, inventory management, and expected profit margins** on a per SKU level, across major global markets such as Amazon, eBay, Walmart, and Lazada.
- Spearheaded the development & scaling of scraping pipelines for Amazon websites in multiple countries, extracting information to drive insights & inform decision-making.
- Empowered clients to improve their marketplace performance by developing **performance metrics** for merchant SKUs, comparing them against similar SKUs

Data Scientist
April '22 - March '23

Blue Yonder Bangalore, India

- Contributed to a **15+** member team responsible for delivering intelligent cloud-native SaaS ML solutions for all areas of **supply chain** in omnichannel commerce, including store fulfillment capacity, delivery date estimation, sales returns and replenishment forecasting, inventory estimation, markdown, and stockout avoidance.
- Worked on massive deep learning models on TensorFlow and Keras, alongside **TFX** pipelines on **Kubeflow**, to develop solutions that integrated data streaming and processing in **Apache Beam** on **Dataflow** and **Bigquery** databases.
- Developed extensive experience in time series, statistics & deep learning with large & noisy datasets exceeding 5TB. Collaborated with an auto-scaled Kubernetes platform, incorporating a multi-tenancy solution to share cloud resources for effective cost-cutting.

Enterprise Data Automation
June '21 - March '22

ANZ Bank Bangalore, India

- Responsible for processing APRA regulatory data & creating pipelines for 70+ source systems using ETL tools like IBM DataStage, Teradata & Control M
- Developed an end to end Data processing automation tool using Robot Framework & python for reducing the development & testing time by 50%
- Responsible for Creating & managing DataStage & Control M jobs in production to ensure correct loading of data from ascii, semi ascii or ebcdic files
- Lead an initiative **Grads4Tech** to improve technology culture & adoption in the bank

AI Research, Prof. Bryan Hooi
Nov '19 - Dec '19

School of Computing | NUS Singapore

- Developed Temporal Attention model for node classification and link prediction task. Analysed the dependency of graphs on time using dynamic models like Node2Vec, Temporal Node Embedding and Attention Walk
- Found inefficiencies of models like TMF, CTDNE, BANE on different data sets for link prediction task and improved them to get AUC of 86% for College Messages dataset

LIFE PHILOSOPHY

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

Persistence & Patience
"Keep On Keeping On"

Go Giver
Value, Compensation, Influence, Authenticity, Receptivity

EDUCATION

B.Tech. in Chemical Engineering

Indian Institute of Technology Mumbai

July'17 – May'21 CGPA : 7.7

Semester Exchange

Danmarks Tekniske Universitet

Jan 2019 – June 2019

- All India Rank of 735 among 1.3 million in JEE Main exam
- All India Rank of 1382 among 0.2 million in JEE Advanced exam

INTERESTS

Data Analysis & Modeling | Machine Learning | Big Data | Deep Learning | Computer Vision | Time Series Analysis | Graph & Network Analysis

TECHNICAL SKILLS

AWS Ecosystem | Ray (Distributed Computing) | Tensorflow Extended | Kubeflow | Apache Beam | Docker | Kubernetes | kedro | Airbyte | DBT | Feast | mlflow | Prefect | SQL | PyTorch | keras | Tensorflow | Numpy | Pandas | ScikitLearn | Matplotlib | Opencv | Dask | Django | Robot Framework | IBM DataStage | Teradata | Control M | Apache Spark

PROJECTS AND INTERNSHIPS

Computer Vision Lead
Oct '19 - Dec '20

Innovation Cell IIT Bombay, India

- 15 students from different departments come together to innovate reality.
- Designed and developed an autonomous fixed wing for search and rescue missions customized for the IARC challenge. Integrated a ROS enabled camera system for detection of characters from a height of 50 meters on a flying drone
- Obtained accuracy of 94 % using transfer learning on MobileNetV2 architecture on custom aerial data set Also, Developed a multi-core zoom-in characters recognition model based on Char74 dataset using OpenCV & Keras

Modelling of Gas sensor data
Aug '19 - Nov '19

Prof. Sharad Bhartiya IIT Bomaby, India

- Analysed Concentration of CO and Ethylene based on sampling time, Auto correlation, cross correlation and DFT analysis on time varying sensor data.
- Analysed the system by introducing 1%, 10% & coloured noises using Impulse estimate, Step estimate Plots & ETFE and Predicted concentration of dynamic gas mixture using ARX, ARMAX, non-linear ARX models.
- Achieved best accuracy of 92.96% & 92.75%for concentration of CO and Ethylene resp. in the 1 step ahead prediction

Data Analyst
Nov '18 - Dec '18

PV-Diagnostics Mumbai, India

- Designed an **SQLite Database** for various projects with multiple parameters having large number of solar I-V, power data-points. Populated the data and performed **Data Sorting, Cleaning and Summary creation using Python**
- Developed new metrics with physical significance to a solar plant based on the translation equation and business information
- Derived new insights and detected outliers through data cleaning, innovative feature engineering and EDA. Classified different types of defects in solar panels using the logistic regression model, achieved 85% accuracy

Hate Speech Detection
Oct 2018 – Nov 2018

Guide - Prof. Sunita Sarawagi IIT Bombay , India

- Twitter tweets data-set with labels as 'Hate Speech' , 'Offensive Language' or 'Neither' was used for training a feed forward network in classification task
- The data-set was thoroughly cleaned and analysis was done on what kinds of words are used in each class and with what frequency. This analysis was represented pictorially as a **word Cloud**.
- Glove pre-trained **Word Embedding** was modified for the given data-set and mapped for each word in the text vocabulary
- A 8 layer deep feed forward network was used to train and training accuracy of **84.30%** and testing accuracy of **82.30%** was achieved in 30 epochs

Image Classification
Aug 2018 – Oct 2018

Guide - Prof. P. Balamurugan IIT Bombay, India

- Studied the working of a multi-layer perceptron and the architectures of Convolution Neural Networks and their applications.
- Deployed different network architectures like DenseNet, Dual Path, GoogleNet, MobileNet, ShuffleNet, ResNet, ResNext, SENet on CIFAR10 dataset using PyTorch.
- Deployed FFN with 7 and 5 layer depths on MNIST database of handwritten digits and achieved accuracy of **98.4 %** and average loss of **0.0465**

Chemical Unit Operations
January 2019 – June 2019

Guide - Prof. Ivan Hundebøl DTU, Denmark

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

CERTIFICATES

COURSERA : R Programming
Data Science Toolbox
Getting and Cleaning Data
Exploratory Data Analysis

COURSEWORK

Machine Learning
Optimization Models
Differential Equations
Linear Algebra
Chemical Engineering Thermodynamics
Computational Methods Lab
Numerical Analysis
Transport Phenomenon

VOLUNTEER EXP.

Blog Series Exchange Semester
- Published a Blog Series of 8 posts about my time in Europe, travel, food, and Denmark Technical University

Mentor E-Cell
- Mentored three teams each of three freshmen for EnB-Buzz an Innovate Start up Idea competition. Helped them to come up with solution to a problem and then developing a Business model canvas for their Start up.

Organizer Mood Indigo
- Worked with a team of volunteers to enhance the experience of student guests from different colleges of India at Mood Indigo the annual cultural fest at IIT Mumbai and helped in the smooth function of other events as well

STRENGTHS

- Strong communication & listening skills.
- Excellent interpersonal skills.
- Flexible and able to work in a team .
- Self-confident with professional behavior .
- High sense of time and co-operation .
- Excellent attention to details and fast learner.