# HARSH GUPTA

hgpt999@gmail.com | +91 987 065 0985 | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

## SKILLS

### PROGRAMMING LANGUAGES
Proficient
- Python

Familiar
- C/C++ • MATLAB

### DATABASE TECHNOLOGIES
Proficient
- SQL Server • MongoDB Server

### ML LIBRARIES
Proficient
- Pandas • Scikit Learn
- Keras • Plotly • NLTK • OpenCV

### VISUALIZATION TOOLS
Proficient
- Power BI • Excel

Familiar
- Tableau • DOMO

### Cloud Stack
Familiar
- AWS(machine learning services)

## EDUCATION

### NIT HAMIRPUR, HP
MSc Math's & Computing
CGPA: **9.21** | **2022**

### UNIVERSITY OF DELHI
BSc MATHEMATICS (Hons)
CGPA: **8.31** | **2020**

## COURSE WORK

- Linear Algebra • Machine Learning
- Soft Computing • Statistics • Database Management • OOP's

## POR

### CORE TEAM
**MEMBER** | DU 2018, 2019. Coordinated with a team of **15 Members** as a Bursar for the Conducting Mathematical Fest of **450+students**.

## ACTIVITIES

- **Qualified** GATE22 (Math's) with **921** rank.
- **Qualified** IITJAM (Math's) with **832** rank.
- Secured **2$^{nd}$ Position** in Inter Year Cricket Tournament(MSc)

## WORK EXPERIENCE

**SENIOR ANALYST @MERCADOS EMI** | JUN'22- PRESENT
- Successfully automated data pipelines using python, ensuring 100% availability of critical datasets.
- Developed robust web scraping code to extract data from multiple energy sector websites using Selenium and Beautiful soup.
- Implemented multi-processing in python to forecast all India state wise demand, leading to a 70% reduction in process execution time.
- Designed Power BI report, analyzed key metrics (Peak Demand, weather, supply), and presented findings to stakeholders.
- Automated data management tasks through Python, cutting down processing time by 35%.

**DATA SCIENCE INTERN @SPARK FOUNDATION** | JUN'21- JUL'21
- Employed Linear Regression to forecast student academic scores based on study hours with 94.2% accuracy and 4.87% mean absolute error.
- Visualized the relationship between study hours and academic scores using the Best Fit Line.
- Employed the Decision Tree Algorithm to predict the type of Iris flower using the features provided in the iris dataset.

## PROJECTS

**Automated Load Balancing System** | CLIENT | JUN'23- FEB'24
- Co-developed algorithm script in python for the power curtailment and revival by identifying non-priority based feeder for load shedding.
- Utilizing real-time data with frequency of 30 sec., for automating power substation curtailment/revival process, thus cutting DSM costs incurred by the state.
- Performed comprehensive data checks and validated ETL, ensuring constant flow of data.

**ELECTRICITY LOAD FORECASTING (REG.) |** CLIENT | JUL'22- DEC'22
- Developed end-to-end electricity demand forecasting model for Maharashtra, Uttar Pradesh, Bihar and Madhya Pradesh using Gradient Boosting, improved MAPE by 2%.
- Conducted detailed analysis to identify seasonal trends, holidays and anomalies in historical electricity demand patterns.
- Communicated forecasted results and insights to senior management and stakeholders, including recommendations for optimal resource allocation and risk management.

**DEEP LEARNING BASED AUTOMATED METER READING (CV) |** MASTER'S THESIS | JAN'22- MAY'22
- Constructed a cascading model architecture to achieve accurate object detection using dataset of 7000+ meter images.
- Utilized localized cropped meter images to annotate 10 digits and trained digit detection using deep learning object detection model YOLOv5.
- Achieved 99.7% accuracy in counter area identification and 87.8% accuracy in digit detection.
- Published Conference on International Conference on Big Data Mining and Information Processing 🔗

## CERTIFICATIONS

- Getting Started with AWS Machine Learning: Learned basis of AWS machine learning services.
- Successfully completed the SQL for Data Science course, mastering CRUD operations in SQL.
- Completed MongoDB course, specialize in CRUD operations.
- Deployment of Machine Learning Model: End-to-End deployment using Git, CircleCI, and Railway app on Udemy.