

# Harsh Kumar

[eharshkmr@gmail.com](mailto:eharshkmr@gmail.com) | +91 886 039 4423 | [in harsh-kumar-a7b98198](https://www.linkedin.com/in/harsh-kumar-a7b98198) | [gh harshcurious](https://www.github.com/harshcurious) | [Website: harshcurious.com](https://www.harshcurious.com)

## SKILLS

**Languages:** Python, C, Go

**ML Libraries:** Pandas, Numpy, scikit-learn, XGBoost, fbprophet, TBATS, statsmodels, SHAP, NLTK, spaCy, PyTorch, Transformers

**Data Technologies:** PostgreSQL, MySQL, Snowflake, MongoDB, PySpark, Polars, Metabase, Selenium

**Cloud Technologies:** AWS Sagemaker, S3, EC2, Step Functions, Lambda Functions

**Development Tools:** Git, Docker, Kubernetes, Jenkins, Github Actions, Bash, PowerShell

**Others:** HTML/CSS, Jira, Confluence,  $\LaTeX$ , Markdown

## WORK EXPERIENCE

### DECISION SCIENTIST | ONE

Bengaluru | 2024 – Current

- **Overdraft Default Risk**

- Created a probability scorecard to estimate the likelihood of a customer with overdraft limit to default on repayment.
- The scorecard helped shape the strategy which reduced the fraud rates by almost 20%.
- Built a feature library from scratch containing over 2000 features in PySpark.
- Productionized the model while building the code template for all machine learning models in the organization

- **ACH Transaction Risk**

- Contributed to the development of a real-time transaction model.
- Shaped the design of the feature calculation pipeline, including significant contribution to the feature calculation pipelines.
- The model is projected to reduce fraud loss by over 10% while simultaneously improving customer experience.

### DATA SCIENTIST | ZESTMONEY

Bengaluru, KA, IN | 2022 – 2023

- **Fraud likelihood estimation to reduce credit risk**

- Experimented with over 2000 features from 8 different data sources to produce an updated fraud detection model improving performance by up to 20%
- Created a new insight of segregating the customer base based on online vs in-store onboarding.
- Created a service API aggregating 8 different data sources with a latency of 200-400ms.

- **Volume projection of calls, emails and chats for workforce management**

- Designed a statistical time series using that reduced call disconnect rate from over 50% to under 20% during peak hours.
- Experimented with algorithms like ARIMA, SARIMA, fbprophet, and TBATS to account for multiple seasonality in data.

- **Model Monitoring**

- Helped create the model monitoring framework for all the deployed model at Zestmoney.
- Produced data visualization to improve visibility into model performance.

- **Collections models to prioritize loan recollection strategy**

- Took ownership of multiple models used to optimize collections.
- Created detailed workflows to manage loan collection strategies.

- **AutoML Platform**

- Helped build an internal ML platform for automatic modelling of internal financial data.
- Built the ELT pipeline in Polars for over 1000 CIBIL features while improving the latency by 40%.

### TEACHING ASSISTANT (TA) | UNIVERSITY OF MISSOURI-COLUMBIA

Columbia, MO, US | 2018 - 2019

- Taught problem solving skills to over 100 pre-calculus and calculus-I students including on-one mentoring.

## EDUCATION

### B.S. in Mathematics and Scientific computing

INDIAN INSTITUTE OF TECHNOLOGY KANPUR (IIT KANPUR)

Kanpur, India | Jul 2012 — Jun 2016

GPA 8.4

### M.S. in Mathematics and Scientific computing

INDIAN INSTITUTE OF TECHNOLOGY KANPUR (IIT KANPUR)

Kanpur, India | Jul 2016 — Jun 2017

GPA 9.1