

# Ravi kumar Chavva

+91 6304424091   [LinkedIn](#)   [GitHub](#)   [Portfolio](#)   [ravikumarchavva@outlook.com](mailto:ravikumarchavva@outlook.com)

## Professional Experience

Data Analyst – FLR Spectron

**Jan 2025 –Present**

- Currently working on the development of a forecasting pipeline to predict fruit quality, using time series analysis and supervised ML models to support supply chain optimization.
- Building an automated document data extraction pipeline using Generative AI, aiming to process unstructured PDFs and text documents into structured insights.

AI Intern - Aavaaz

**Dec 2024 – Jan 2025**

- Assisted in the NLP research project and gained hands-on experience with PyTorch for sentiment analysis and text classification tasks.
- Participated in collaborative learning sessions to build expertise in speech recognition technologies and core applied AI concepts to solve real-world problems.

## Projects

T20 Cricket Win Prediction ([project link](#))

**([GitHub](#) 2024)**

- Designed to provide broadcasters, analysts, and sports teams with real-time win probability insights, enhancing strategic decision-making and fan engagement in cricket matches.
- Built a robust data pipeline with Apache Airflow, Spark, and HDFS, processing 570k+ rows of historical cricket data from [cricsheet.org](https://cricsheet.org) and scraped additional statistics from [espncricinfo.com](https://espncricinfo.com).
- Engineered a multi-model architecture combining LSTMs, CNNs, and DNNs to predict match outcomes, trained using PyTorch and optimized with Weights & Biases.
- Simulated real-time match scenarios, achieving 85% accuracy in predicting match outcomes during critical final overs.

Customer Churn Prediction ([project link](#))

**([GitHub](#) 2024)**

- Developed a customer churn prediction model to identify at-risk customers, enhancing retention strategies and revenue stability.

- Analysed a dataset of 7,000+ customer records from [ibm-telecom-churn](#) to uncover key churn drivers and actionable insights.
- Achieved an F1 score of 86% after applying SMOTE and optimizing with CatBoost & XGBoost, enabling accurate identification of high-risk churn customers while keeping false positives low—improving targeting for retention campaigns while minimizing false alarms.
- Deployed the model as a serverless API on GCP using FastAPI and Docker, achieving real-time predictions with a 130ms response time under simulated load.
- Identified charges and internet service quality as key churn drivers, recommending personalized plans and targeted improvements in internet service to reduce churn rates.

## Skills

- |  |   |
|--|---|
| • <b>Programming Languages</b>               | Python, SQL   |
| • <b>Data Visualization and Manipulation</b> | Matplotlib, Seaborn, Numpy, Pandas, PySpark, Polars |
| • <b>Machine Learning</b>                    | Scikit-Learn, PyTorch, TensorFlow, ONNX             |
| • <b>Deployment, DevOps &amp; MLOPs</b>      | FastAPI, Next.js, Docker, Git, Weights and Biases   |
| • <b>Familiar with</b>                       | Apache airflow, HDFS, GCP, GitHub Actions           |

## Education

Bachelor of Technology (B. Tech) in Computer Science and Engineering (Data Science)

Sri Venkateswara College of Engineering, Tirupati, India,  
CGPA:8.5

**Nov 2021 – Present**

## Certifications

Machine Learning Specialization – Andrew Ng – Coursera

**AUG-2022**

Machine Learning for Engineering and Science Applications (silver medal) – NPTEL

**MAR-2024**

Introduction to large language models – NPTEL

**June-2025**

## Summary

Final-year student and full-time Data Analyst with hands-on experience in developing AI-driven solutions, ML pipelines, and predictive modeling. Proficient in PyTorch, FastAPI, and Docker for real-time deployments. Skilled in Computer Vision, model optimization, and delivering data insights to drive informed decision-making and business impact.