

Ravi kumar Chavva

Portfolio : <https://ravikumarchavva.com>

LinkedIn : <https://www.linkedin.com/in/ravikumar-chavva/>

Github : <https://github.com/ravikumarchavva/>

Email : ravikumarchavva@outlook.com

Phone : +91 6304424091

Experience

AI Intern - Aavaaz

Dec 2024 – present

- Working on NLP tasks, including **Sentiment Analysis** using PyTorch.
- Assisting in data preprocessing and model evaluation for text classification **projects**.
- Actively participating in **learning sessions** to enhance AI and **speech recognition** skills.

Projects

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

([GitHub](#) 2024)

- Business Problem:** Enable broadcasters, analysts, and teams to **enhance fan engagement** and make informed strategic decisions with real-time win probability insights.
- Solution:**
 - Gathered data from cricsheet.org with 570k+ rows and scraped additional statistics from espn.cricinfo.com.
 - Built a robust data pipeline using Apache Airflow, Spark, and HDFS for efficient real-time data handling. Implemented cumulative aggregated statistics with window functions and transformed data into a machine learning-ready format.
 - Designed a multi-model architecture integrating LSTMs, CNNs, and DNNs as encoders, followed by a decoder network for comprehensive predictions, utilizing PyTorch for model training, Weights & Biases (WandB) for performance tracking and hyperparameter tuning, and ONNX for optimized inference.
- Impact:** Achieved **86% accuracy** in final overs, delivering actionable insights for strategic decision-making and significantly **enhancing fan engagement** during live matches.

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

([GitHub](#) 2024)

- Business Problem:** Minimize churn by **identifying at-risk customers** early for IBM’s subscription-based businesses, ensuring improved revenue stability and long-term growth.
- Solution:**
 - Conducted Exploratory Data Analysis (EDA) on 7,000+ customer records from the [IBM telecom dataset](#) to identify key churn factors.
 - Built a baseline Logistic Regression model, addressing class imbalance with SMOTE, achieving an 81% accuracy.
 - Enhanced performance using ensemble methods like CatBoost and XGBoost through Bayesian optimization.
 - Deployed the solution via FastAPI in a Dockerized serverless environment on GCP, delivering real-time predictions with an average response time of 130 ms.
- Impact:** Achieved **86% churn detection accuracy**, identified charges and internet services as major factors contributing to churn. The analysis highlighted the need for improving internet service quality and offering personalized plans to enhance customer satisfaction and retention.

Skills

- | | |
|--|---|
| • Programming Languages | • Python, SQL |
| • Data Visualization and Manipulation | • Matplotlib, Seaborn, Numpy, Pandas, PySpark, Polars |
| • Statistical Modeling | • Scikit-learn |
| • Deep Learning Architectures | • Neural networks and Transformers (CNN, RNN, VIT) |
| • Deep Learning Frameworks | • PyTorch, TensorFlow, ONNX |
| • Deployment and Version Control Systems | • Fastapi, Next.js, Docker, Git |
| • Machine Learning Operations (MLOps) | • Weights and Biases |
| • Familiar with | • Apache airflow, hdfs, GCP, GitHub Actions |

Education

Bachelor of Technology (B. Tech) in Computer Science and Engineering
Sri Venkateswara College of Engineering, Tirupati, India

Nov 2021 – Present | CGPA: 8.4

Certifications

- Machine Learning for Engineering and Science Applications (silver medal) – NPTEL
- Machine Learning Specialization – Andrew Ng - Coursera

Extracurricular Activities

Data Nexus: Founded a data science community at college, engaging 50+ students in workshops, hands-on sessions, and resource-sharing to accelerate learning and project development.