

# Ujjwal Thapliyal

## Data Scientist

ujjwalthapliyal4@gmail.com | 8439653056 | Dehradun Uttarakhand | <https://www.linkedin.com/in/ujjwal-thapliyal/>

<https://portfolio-ujjwal-thapliyal.netlify.app/> | [github.com/UjjwalThapliyal123](https://github.com/UjjwalThapliyal123)

### PROFILE

Data Scientist experienced in **analyzing large datasets**, developing **machine learning algorithms**, and delivering data-driven insights that influence **business strategy**, pricing decisions, and **personalized user experiences**. Strong background in statistical analysis, algorithm development, and data visualization, with hands-on experience deploying and monitoring production-ready ML systems using **Python and SQL**.

### PROJECTS

#### Flight Price & Customer Satisfaction Prediction System,

June 2025 – July 2025

Machine Learning | Python | Scikit-learn | XGBoost | MLflow | Streamlit

- Analyzed **110K+ flight and passenger records** to identify demand trends, service quality drivers, and pricing patterns supporting dynamic pricing strategy.
- Performed **data cleaning, exploratory analysis, and feature engineering** to improve signal quality and model stability.
- Developed and evaluated regression models (Linear Regression, Random Forest, XGBoost), achieving  **$R^2 = 0.85$**  and reducing RMSE by **23%** versus baseline.
- Built classification models on **10K+ labeled samples** to predict customer satisfaction, achieving **80% accuracy and 0.80 F1-score**.
- Applied statistical evaluation techniques (RMSE, MAE, precision, recall) to guide model selection and optimization.
- Delivered visual analytics dashboards enabling real-time performance monitoring and insight communication.

#### FlipSatisfy – Customer Satisfaction Prediction, NLP | FastAPI | Docker

Jul 2025 – Aug 2025

- Designed a **machine learning pipeline** on **85,907** customer support interactions, transforming unstructured and categorical data into analytical features.
- Engineered **NLP preprocessing workflows**, including TF-IDF vectorization and encoding strategies, to improve text signal extraction.
- Trained and benchmarked multiple algorithms, improving test accuracy from **95.6% to 96.9%** through **hyperparameter optimization**.
- Deployed the optimized model as a containerized **FastAPI** service, enabling **real-time inference** and system monitoring.
- Integrated predictions into a **visual reporting layer** to support operational decision-making

#### Book Recommendation System, TF-IDF | K-Means | Cosine Similarity

Jun 2025 – July 2025

- Built a **content-based recommendation system** on **5,000+** items using TF-IDF and cosine similarity.
- Delivered personalized rankings with **sub-second response time**, reducing user search effort by **~40%**.
- Applied clustering techniques to improve recommendation diversity and relevance.

### SKILLS

Programming & Data Analysis: Python, SQL, SciPy

Machine Learning & Deep Learning: Classification, Regression, Clustering, NLP, CNN, RNN

Statistics & Quantitative Methods: Probability, Inferential Statistics, Optimization

Data Visualization: Matplotlib, Seaborn, Analytical Dashboards

Data Handling: Data Cleaning, Exploratory Data Analysis (EDA), insight generation, Data interpretation

Tools & Platforms: Scikit-learn, XGBoost, MLflow, FastAPI, Docker, Streamlit

### EDUCATION

#### Master of Computer Applications (MCA), SGRR University

July 2024 – Present | Dehradun Uttarakhand

#### Master Data Science Program, IIT Madras Guvi HCL

November 2025 – June 2025 | India

#### Bachelor's Degree 2021–2024, SGRR University

July 2021 – July 2024 | Dehradun Uttarakhand

### CERTIFICATES

#### Data Science Certification

Guvi HCL : IIT Madras

#### Power Bi Certificate

Guvi HCL