Vikas Kumar

orvikas05@gmail.com — New Delhi, India

vikas24.github.io — github.com/Vikas24 — linkedin.com/in/vikas-kumaror05

PROFESSIONAL SUMMARY

Computer Science undergraduate with hands-on experience in Machine Learning and Data Science, focused on building end-to-end ML pipelines with emphasis on robust evaluation beyond accuracy. Experienced in handling real-world challenges such as class imbalance, feature engineering, and model optimization. Currently expanding expertise in Natural Language Processing, Deep Learning, and Transformer-based architectures.

EDUCATION

Bachelor of Technology in Computer Science   2023 – 2027
Guru Gobind Singh Indraprastha University, New Delhi, India

SKILLS

Programming & Tools: Python, Git, GitHub

Data Analysis & Machine Learning: Pandas, NumPy, Scikit-learn, Exploratory Data Analysis (EDA), Feature Engineering, Model Evaluation

Machine Learning Techniques: Logistic Regression, Class Imbalance Handling (SMOTE), Hyperparameter Tuning (GridSearchCV), Cross-Validation (k-fold), ROC-AUC and Precision–Recall Analysis, Threshold Tuning

Currently Learning: Natural Language Processing, Deep Learning, Transformer Architectures

PROJECTS

Titanic Survival Prediction — End-to-End Machine Learning Project
Portfolio: vikas24.github.io — *GitHub:* github.com/Vikas24/titanic-survival-prediction

- Built an end-to-end machine learning pipeline for binary classification on the Titanic dataset, covering data preprocessing, model training, and evaluation
- Performed detailed exploratory data analysis to identify key survival drivers such as passenger class, gender, age, and fare
- Implemented structured preprocessing including missing value imputation, categorical encoding, and feature preparation
- Addressed class imbalance using SMOTE applied exclusively to training data to avoid data leakage
- Optimized Logistic Regression using GridSearchCV to tune regularization strength and solver configuration
- Evaluated model performance using accuracy, ROC-AUC, Precision–Recall curves, and k-fold cross-validation
- Achieved approximately 80% test accuracy and a mean cross-validated ROC-AUC of 0.85, demonstrating strong generalization
- Developed environment-independent Python scripts runnable on both local systems and Google Colab to ensure reproducibility

EXPERIENCE

Machine Learning Trainee — Self-Directed Projects — Remote   2026 – Present

- Designed and implemented multiple machine learning projects spanning the full ML lifecycle, from data preprocessing to model evaluation
- Focused on real-world ML challenges including imbalanced datasets, metric selection, and robust evaluation strategies
- Developed modular, reusable Python scripts for preprocessing, training, hyperparameter tuning, and evaluation
- Iteratively improved model performance through error analysis, metric-driven optimization, and incorporation of feedback