NIKHIL SOMNATH BORADE

Pune, Maharashtra, India

+91 94215 12632 | nikhilsborade3@gmail.com | LinkedIn | GitHub | Portfolio

PROFESSIONAL SUMMARY

AI/ML Engineer with hands-on experience designing, training, evaluating, and deploying end-to-end machine learning and NLP solutions across classification, regression, and predictive analytics use cases. Proficient in Python, data preprocessing, feature engineering, and model optimisation using Scikit-learn with a proven track record of tuning hyperparameters via cross-validation. Experienced in integrating ML models into production-ready web applications using Streamlit, translating raw data into actionable business insights. Eager to contribute to scalable, practical, and deployable AI systems that solve real-world business problems.

TECHNICAL SKILLS

Languages	Python (primary), SQL (MySQL)
AI / ML	Supervised Learning, Classification, Regression, Model Training & Evaluation, Hyperparameter Tuning, Cross-Validation, ML Pipelines, ColumnTransformer
Libraries	NumPy, Pandas, Scikit-learn (KNN, Naive Bayes, Random Forest), Matplotlib, Seaborn, NLTK
NLP	Text Preprocessing, Tokenization, TF-IDF Vectorization, Text Classification,Feature Extraction
Data Analysis	Data Cleaning, Exploratory Data Analysis (EDA), Statistical Analysis, Feature Engineering, Correlation Analysis
Deployment	Streamlit (web apps), Pickle (model serialisation), ML Pipelines
Tools	Git, GitHub, Power BI, Microsoft Excel

PROJECTS

Fake News Detection using NLP | *Oct 2025*

Python | NLTK | Scikit-learn | TF-IDF | Naive Bayes

- **Problem:** Misinformation spreads rapidly across online platforms and is extremely difficult to identify manually at scale, de- manding an automated classification solution
- **Approach:** Designed an end-to-end NLP pipeline — raw text was cleaned, tokenised, and stripped of stopwords using NLTK; TF-IDF vectorization was applied to convert text into numerical features; a Multinomial Naive Bayes classifier was trained and evaluated using accuracy, precision, and recall metrics
- **Insight:** Recurring word patterns and high-frequency term distributions serve as reliable discriminators between fake and real news articles
- **Result:** The model achieved **93.7% classification accuracy**, demonstrating the viability of automated, scalable NLP-based content moderation systems

Telco Customer Churn Prediction | *Nov 2025*

Python | KNN | Cross-Validation | ROC-AUC | Scikit-learn

- **Problem:** Telecom providers lose a significant customer base each quarter with no reliable early warning mechanism, making proactive retention strategies difficult to implement
- **Approach:** Preprocessed the dataset through missing-value handling and feature scaling (StandardScaler); implemented a K- Nearest Neighbours classifier and performed systematic hyperparameter tuning of K using stratified cross-validation to prevent overfitting
- **Insight:** Customer usage frequency, contract type, and service tenure are the strongest behavioural signals that predict churn risk ahead of actual cancellation
- **Result:** Achieved **77.7% accuracy** and an ROC-AUC of **0.73**, providing the retention team with a ranked list of at-risk customers for targeted intervention

Travel Package Purchase Prediction — Streamlit App | *Dec 2025*

Python | Random Forest | Streamlit | Feature Engineering | Scikit-learn

- **Problem:** Sales teams lacked a data-driven method to identify customers most likely to purchase travel packages, leading to inefficient outreach and lower conversion rates
- **Approach:** Performed feature engineering to extract high-signal variables, trained and tuned a Random Forest classifier using cross-validation, and evaluated performance using accuracy, precision, recall, and F1-score
- **Insight:** Customer income, number of follow-ups by the sales team, and pitch satisfaction score emerged as the three strongest predictors of purchase intent
- **Deployment:** Built and deployed a production-ready interactive web application using Streamlit that takes live customer input and returns a real-time purchase probability prediction
- **Impact:** The tool enables sales teams to programmatically prioritise high-probability leads, directly improving outreach efficiency and conversion rates

EDUCATION

Bachelor of Computer Science — *July 2022 – May 2025*

Savitribai Phule Pune University — *Nashik, Maharashtra*