

# Tanuj Saxena

**Data Analyst Intern | Data Science Enthusiast | Skilled in Data Analysis & Visualization**

+91-8126560686/ [tanuj.saxena.rks@gmail.com](mailto:tanuj.saxena.rks@gmail.com) / <https://www.linkedin.com/in/tanuj-saxena-970271252/> , Greater Noida, India

## Key Skills

- Python | Java | C | React.js
- Pandas | Numpy | TensorFlow | OpenCV | Flask | Scikit-Learn
- MongoDB | MySQL
- Machine Learning | Statistical Modeling | OOPs

## Tools

Power BI | Tableau | MS Excel | MySQL Workbench | MYSQL | MS Office | Microsoft Azure

## Professional Experience

### CodSoft , India

#### Role- Machine Learning Intern

Sep 2023- Oct 2023

- Improved the classification accuracy of a Python-based machine learning model from 70% to 80% by optimizing data preprocessing techniques and model selection using Scikit-learn.
- Designed and implemented an SMS Spam Detection model using Support Vector Machine (SVM), achieving an accuracy increase from 92% to 99% through feature extraction and hyperparameter tuning.
- Enhanced customer churn prediction by conducting comprehensive data analysis and model optimization for Logistic Regression, KNN, Decision Tree, and Random Forest, resulting in an accuracy boost from 81% to 86% with Random Forest.
- Utilized data analysis techniques such as cross-validation and grid search to ensure robust model performance and data-driven insights.

## PROJECTS

### Adaptive AI Questioning

- The Adaptive AI Questioning system uses Python and Scikit-learn to dynamically alter question difficulty based on individual student performance which is more efficient than other system.
- Applied Item Response Theory (IRT) and Cognitive Diagnostic Models (CDM) to optimize question sequences, resulting in a 10% improvement in model accuracy.
- It uses Object-Oriented Programming principles for scalability and analyses real-time student reactions to improve engagement and educational efficacy.

### Sentiment Analysis Web Application

- Developed a machine learning-based sentiment analysis application using Streamlit, incorporating Logistic Regression with TF-IDF vectorization and BERT for accurate sentiment detection results, 10% increase in accuracy.
- Implemented robust text preprocessing techniques, enabling users to input text or upload CSV/Excel files, with real-time sentiment predictions and downloadable results.
- Optimized model performance, achieving an F1-score improvement of 12%, and boosting overall model accuracy to 92% for reliable and actionable sentiment

## Certifications

- Neural Network and Deep Learning- Deeplearning.Ai
- Advance Learning Algorithms- Stanford Online
- Machine Learning Specialization- Stanford Online
- Data Analytics Essential- Cisco
- Python for Data Science- IBM
- Machine Learning- IIT Kharagpur
- Project Management- IIT Kharagpur

## Education

**Bachelor of Technology** || Computer Science and Engineering  
Sharda University

2022-2026

**Senior Secondary** || Science Education  
St. Peter's Sr. Sec. School

2020-2022