

# Nitish

Data scientist

nitishrohilla77624@gmail.com | +917206368408 | Gurugram Haryana

[PORTFOLIO](#) | [GitHub](#) | [LINKEDIN](#) |

## EDUCATION

J.C BOSE UNIVERSITY OF SCIENCE & TECHNOLOGY, YMCA FARIDABAD

1-08-2020 - 1-06-2024

B.TECH. - COMPUTER ENGINEERING WITH DATA SCIENCE

Faridabad, India

CGPA: 7.4

## EXPERIENCE

SKILL UP TECH | DATA SCIENTIST

Gurugram, Haryana | December 2025 – Present

I design and deploy solutions that extract meaningful insights from language. My core work involves utilizing **Python** and the **spaCy** framework to build and fine-tune **NLP** models. I also architect and implement critical data pipelines and workflow automations using **n8n** to ensure our AI services run efficiently and scale effectively.

## SKILLS

PROGRAMMING LANGUAGES

Python, JavaScript, C++, HTML, CSS

LIBRARIES/FRAMEWORKS

Node.js, Web Sockets, n8n, spaCy, NLP, LLM

TOOLS / PLATFORMS

Tableau, Power BI, Matplotlib, Pandas, Seaborn, NLTK, WordCloud

DATABASES

SQL, Data Warehousing, Excel, DataFrames

## PROJECTS / OPEN-SOURCE

INSURANCE CLAIM PREDICTION | [LINK](#)

Python, Pandas, Matplotlib, Regression, Decision Trees

Developed a machine learning web application to predict whether an insurance claim would be approved. Trained multiple classification models including **Decision Trees** and **Random Forest**. Used **Matplotlib** for data visualization and deployed the best-performing model. Achieved a model accuracy of 89%, leading to a 20% improvement over baseline predictions.

REAL-TIME CHAT APPLICATION | [LINK](#)

Flask, Web Sockets, Node.js

Created a live messaging platform to enable real-time user communication. Implemented bidirectional communication using **Web Sockets** with Socket.IO. Built a dynamic frontend with HTML/css and Js and backend using Flask. Supported message latency under **200ms**.

SENTIMENT ANALYSIS | [LINK](#)

Python, NLTK, Matplotlib, WordCloud, Logistic Regression

Developed a model to detect public sentiment in social media posts, focusing on identifying depressive or negative content. Cleaned and pre-processed tweets using **NLTK**, implemented **Logistic Regression** for sentiment classification, and visualized patterns with **WordCloud** and **Matplotlib**. Achieved an overall 82% accuracy with improved detection of depressive posts on 10,000+ tweets.

SALES ANALYSIS | [LINK](#)

Python, Pandas, Matplotlib, Seaborn, Tableau, Power BI

Analyzed historical sales data to uncover patterns and forecast future trends. Performed **EDA** using **Pandas, Seaborn, and Matplotlib**; built interactive dashboards with **Tableau**. Identified key trends and top-performing regions, resulting in a 15% improvement in forecast accuracy.