

# GOURANG PATIDAR

Data Scientist

☎ +91-8839006621

✉ gourangpatidar2003@gmail.com

🔗 LinkedIn Profile

🐙 GitHub Profile

• Leetcode

## EDUCATION

### •Jabalpur Engineering College

Bachelor of Technology in Artificial Intelligence And Data Science

2021-25

CGPA: 7.6

## EXPERIENCE

### •Founder

TheQuickAI.com

July 2024- Present

Remote

- Launched a **Generative AI SaaS** for creating quizzes from four input types (Text , Video url , Blog url , Pdf) , increasing user engagement and content generation **efficiency by 40%**.
- Architected a Cassandra DB solution for the Quiz on Hostinger VPS, achieving 99.9% uptime and enhancing reliability for **1,500+ monthly interactions**.

### •ML Engineer (Intern)

Ineuron.ai

Aug 2023 - Sep 2023

Remote

- Achieved **94%** accuracy in sentiment analysis on student feedback data using **NLTK**.
- Built predictive models using TensorFlow, reducing forecasting errors by **21%**.
- Conducted **A/B testing** on machine learning models to identify optimal model configurations, leading to an 11% increase in user retention.

### •Data Analyst (Intern)

Ineuron.ai

Oct 2022 - Dec 2022

Remote

- Used Python **clustering methods** to determine groups of states where courses were underperforming, leading to improvements that increased profit by 4%.
- Assisted in developing a **recommendation engine** using collaborative filtering techniques, resulting in a 22% rise in user interaction with recommended listings.

## PROJECTS

### –Portfolio

A **GEN-AI** based Streamlit-hosted website showcasing a collection of various products.

- \* Created a **YouTube Video Summarizer**: Converted subtitles to text and summarized using Langchain.
- \* Extracted and stored large datasets as vectors in **Cassandra DB**, enabling efficient querying and analysis of PDF-related information.
- \* Implemented a **live search** integration for ChatGPT using APIs, enhancing user accessibility to real-time information.

### –Text Summarizer

A trained NLP model that summarizes text using a dataset of 27,000 rows.

- \* Processed raw text from sources using Python and libraries such as **transformers** and **torch**.
- \* Engineered a summarization model with AutoModelForSeq2SeqLM and AutoTokenizer, reducing manual summarization time **by 50 hours/month** and streamlining workflows.
- \* Optimized the summarization model using TrainingArguments, Trainer, and DataCollatorForSeq2Seq, boosting accuracy by **40%** and reducing processing time by **25%**.

### –Insurance Premium Prediction

A End-to-End ML project deployed on **Elasticbeanstalk** and **CodePipeline** of AWS

- \* Estimated individual healthcare costs to help customers focus on the health benefits of insurance plans.
- \* Developed APIs using Flask and optimized model performance with **GridSearchCV** and **Cross-Validation**.
- \* Implemented data preprocessing, feature engineering, and visualization to enhance accuracy and provide actionable insights.

### –Google Play Store Data Analysis

A dataset containing 10,000 rows and 14 columns was analyzed in a Jupyter notebook.

- \* Identified approximately **22 insights** using NumPy, Pandas, Seaborn, and Matplotlib.
- \* Created a **Power BI** dashboard that demonstrated data intuitively.

## TECHNICAL SKILLS AND INTERESTS

**Machine Learning:** Python, Scikit-learn, AWS, Azure, ML Algorithms, Statistics & Probability

**Data Analytics:** NumPy, Power BI, MS Excel, MySQL, Pandas, Matplotlib, Seaborn

**Deep Learning:** Natural Language Processing, TensorFlow, Keras, Neural Networks, Mathematics, Linear Algebra

**GenAI:** Langchain, Llama Index, OpenAI, Gemini Pro