Raj Das

rajdas2001.github.io/rajdas/
rajdas.2001@gmail.com | +91 7682956963

EDUCATION

VSSUT, BURLA
B.Tech (MME) | 2018-2022
CGPA: 8.18

D.A.V. PUBLIC SCHOOL, PKT
Intermediate | 2018
Percentage: 77.6%

D.A.V. PUBLIC SCHOOL, PKT
High School | 2016
CGPA: 9.4

LINKS

LinkedIn:// **rajdas293**
Website:// **bit.ly/rajdas293**
Github:// **rajdas2001**

COURSEWORK

- Introduction to Statistics
- Machine Learning Specialization
- Deep Learning Specialization
- Natural Language Processing
- LLMs and Generative AI
- Introduction to Git and GitHub

CERTIFICATIONS

- AWS Certified Cloud Practitioner
- Salesforce Certified AI Associate

SKILLS

- Python
- SQL
- Data Science
- Data Cleaning
- Exploratory Data Analysis (EDA)
- Data Visualization
- Machine Learning
- Deep Learning
- Microsoft Power BI
- Microsoft Excel
- AWS
- Git and GitHub

EXPERIENCE

WIPRO | DATA SCIENTIST
Mar 2023 – Present | Bengaluru, India

- Developing an NLP powered chatbot using Nuance Mix and Microsoft Azure for an insurance company to help users with their queries, change policy details, initiate claims, and provide general support.
- Working in cross-functional teams, overseeing prototype to production, and communicating results with stakeholders.
- Developed a sentiment analysis model using Transformers architecture to analyze customer feedback on products, resulting in a 20% improvement in customer satisfaction score.

PROJECTS

RETRIEVAL AUGMENTED GENERATION (RAG) WITH LLAMA 2

- **Tech Stack:** Llama 2 7B Chat GGML, LangChain, C Transformers, Chroma DB, Streamlit
- Developed and implemented Retrieval Augmented Generation (RAG) using Llama 2 7B Chat GGML model to integrate relevant knowledge bases, thus improving responses' factual accuracy and completeness.
- Increased the model's ability to answer complex and open-ended questions by 30%.
- Reduced memory footprint of the LLM by 46.96% using quantized GGML model.

HUMAN SENTIMENT ANALYZER

- **Tech Stack:** Python, Hugging Face - BERT, NLTK, spaCy
- Developed a text classification model based on Transformers architecture.
- The model achieved an accuracy of 90.8% in identifying text polarity.
- The project can be used for preventing cyberbullying, brand/product analysis and feedback analysis.

IPL DATA ANALYSIS

- **Tech Stack:** Python, NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn
- Analyzed historical match data (2008-2017) to identify player statistics and key factors influencing match outcomes.
- Utilized machine learning models (Logistic Regression, Random Forest) to achieve an accuracy of 88% in predicting match winners.

CRACK DETECTION ON SURFACES

- **Tech Stack:** Python, TensorFlow, Keras
- Developed a model to identify cracks on concrete surfaces for early crack detection and maintenance of the Hirakud Dam, Odisha.
- The model is built on top of a pre-trained CNN model (InceptionV3).