Niraj Kumar

neeraj2681.github.io | neeraj2681@gmail.com | linkedIn/neerajbhatt2 | github/neeraj2681

EDUCATION

M.Tech in Computer Science   Kolkata, India | 2020-22
Indian Statistical Institute

B.Tech in Computer Science and Engineering   Pauri, Uttarakhand | 2015-19
Govind Ballabh Pant Engineering College

WORK EXPERIENCE

Data Scientist   Gurgaon, Haryana | Since Aug 2022
@American Express

- Designed **Fraud Detection Models** for a variety of Industries like Money Transfer, Retail, Travel etc.
- Designed **Auto-Review** model with **past-lookup variables** leading to **50% reduction in review** and **only 4% Fraud $ Acceptance**
- Designed Strategy Rules over and above the base models to enhance the model performance and capture seasonal fraud trends.
- Includes end to end Machine Learning Model Development and Deployment
- Skills: **Python, Pyspark, SQL, Data Analysis, Machine Learning Algorithms, Strategy Designing**

PROJECTS

DETECTING ANOMALIES IN SURVEILLANCE VIDEOS   Tensorflow, Deep Neural Networks, 3D CNNs

- Utilising Encoder-Decoder based reconstruction and prediction techniques
- Utilisation of different **Deep Neural Networks**
- Real-time applicable solution @30FPS

QUESTION ANSWERING MODEL   NLP, Transformer Neural Network, Deep Learning

- A **BERT** based NLP model for answering the questions based on some given context
- Web Crawling was used to fetch Wikipedia pages based on user the query, which are then used as context
- Model then returns the starting and ending position of the answer from the relevant Wikipedia page.

OPEN-SOURCE BASED CUSTOM LLM   Langchain, LLMs, Pinecone, Giskard

- Different open source LLMs to train on custom dataset using Retrieval Augmented Generation
- Answering the user queries based on a given Youtube video - using transcript of the video as context
- **End to End model development and deployment on the cloud**
- **Evaluation using Giskard**

SKILLS

Languages: C, Python, Java, Bash, SQL, Rust

Domains: Classical Machine Learning, Deep Learning, Image Processing and Computer Vision, NLP, LLMs

Tools and Libraries: Pandas, MATLAB, Git, Pytorch, Tensorflow, LaTeX, Pyspark, Langchain

PERSONAL INTERESTS

Testing different investing strategies, Mathematics, Learning about different cultures around the globe(By maintaining my Duolingo streak), Hiking.