Rafaqat Ali Khan

khanrafaqat608@gmail.com | +92-330-6923830 | https://www.linkedin.com/in/rafaqatkhan/

EDUCATION

Bahauddin Zakariya University (BZU) **Sep 2020 – Jun 2024**
BS Information Technology

SKILLS & CERTIFICATIONS

Languages & Tools: Python (Pandas, Matplotlib, Numpy, Seaborn) | Scikit-Learn, TensorFlow | Keras, NLTK | Flask, Django, FastApi | Git, Github | OpenCV, BeautifulSoup | SQL | Jupyter Notebook
Data Science Skills: Data Visualization, Machine Learning, Natural Language Processing (NLP), Deep Learning, Data Cleaning, Exploratory Data Analysis, Generative AI, LLMs , Prompt Engineering
Certifications: Machine Learning Specialization (Deep learning.ai by Andrew Ng) | Coursera - What is Data Science, IBM

PROFESSIONAL EXPERIENCE

CognoRise InfoTech (Data Science Intern) **April 2024 – June 2024**
- Build a RAG (Retrieval Augmented Generation) pipeline from scratch and have it all run locally.
- Build custom Object detection model for detecting face mask wear in real time using Tensorflow Object Detection API.
- Worked on tasks including Iris Flower Classification, Credit Card Fraud Detection, Titanic Survival Prediction, Superstore Sales Prediction, Car Price Prediction, Cinema Tickets Sales Forecasting, Shopper Sentiments Analysis, and Fake News Prediction

PROJECTS

Automatic Image Captioning using Deep Learning and NLP
- Developed a model to recognize and generate accurate captions for images using neural network architectures and NLP.
- Implemented CNN and RNN architectures to extract features from images and generate descriptive captions
- Used Flikr_8k dataset, Glove for vector representation and Inception v3.
- Created front-end using React and backend using Fast-Api

N-Gram Modelling for E-Commerce Product Reviews & Classification with Transformers pipeline
- Generated a data set of reviews by scrapping over 5000 product reviews from an e-commerce website using Beautiful Soup (bs4) and Selenium WebDriver in Python.
- Conducted data cleaning, feature engineering, and text pre-processing techniques such as punctuation removal, stopwords removal, lemmatization, stemming, and word count.
- Constructed a corpus, vocabulary, and bigrams/trigrams, utilizing Bayesian rules to predict the next probable word.
- Utilized Transformers-pipeline for sentiment analysis, enabling labeling of positive and negative reviews.

AI Tool for Image-based Hashtag Generation
- Developed an AI tool that identifies trending hashtags for images based on the content of the image.
- Employed vit2-gpt2-image-captioning pre-trained model from hugging face repository to caption images given by the user.
- Leveraged openai's gpt-3.5-turbo model to generate hashtags for the captions generated.
- Built an interface using gradio, and made users to upload any images or URL, to generate hashtags for the given image and deployed in hugging face spaces huggingface.co/spaces/ImagetoHashtags