KEVIN PATEL

Phone: (517) 505-9591
patelk69@msu.edu   East Lansing, MI 48823

I am an aspiring Data Scientist. During my under-graduation, I have gained knowledge about computer science. I am proficient in building solution to solve real world problem with the help of machine learning and Artificial Intelligence . My area of interest are Machine Learning and Data Science.

LinkedIn : https://www.linkedin.com/in/kevin-patel-50b18a195/
Github : https://github.com/kevin200010

EDUCATION

M.S. Michigan State University, USA   Sept 2022 - currently
Master's in Data Science
Subjects : Introduction to Data Science, Mathematical Statistics for Data Science, Big Data Analysis, Data Mining, Applied Machine Learning
GPA : 4/4

B.Tech Pandit Deendayal Energy University, India   Aug 2018 -July 2022
Bachelor of Computer Engineering
Subjects : Design Analysis and Algorithm, Database Management System, Object Oriented Programming, Web Technologies, Artificial Intelligence, Operating System, Compiler Design
GPA : 9.59 /10

RESEARCH/WORK EXPERIENCE

Data Science Intern
VERN.AI (Tucknologies Holdings Inc.), East Lansing, USA)   May 2023 – Aug 2023
Menager : Craig Tucker

- Developed a user-centric offline Question-Answer bot using the **Vicuna(7B) model** and **Langchain**. Leveraged **sentence transformers all-MiniLM-L6-v2** to generate embeddings for efficient searching within **Vector DB** and ensuring targeted information retrieval using semantic search and after summarizing using Vicuna-7B it further processed by **Langchain** for accurate question-answer sequences with managing context within the session.
- Build an interactive bot tailored for symptom checking and providing user-centric recommendations. The bot integrates **Dialogflow NLU** for accurate query processing and rich response rendering. Utilized GCP's Translate and Speech API for supporting multilingual interactions, ensuring accessibility for diverse users. The backend leverages Azure storage for symptom documentation, used **Azure OpenAI's GPT model** for **data summarization** and **response generation**.

Research and Data Science Intern
Indian Meteorological Department, Bhopal-India   2021 - 2022
Advisor : Ved Prakash Singh
- EDA, Preprocessing, data cleaning and Feature engineering of Meteorological observation such as Radar data and RSRW data
- Build model to predict impactful low-altitude wind speed accurately, that provide prediction to overcome the issue of cone of silence for standard radar observation.
- Worked on Deep Clustering for cloud data clustering
- Build Unsupervised Multi-level Clustering for Segmentation of Remote Sensing Imageries using DB-Scan and K-means with adaptive K.
- The unsupervised algorithm which can perform the multi-level clustering of images on the basis of different disjoint feature-sets. Then it merges the different clusters by validating similarity conditions.
- The results of clusters generated at each level were evaluated through DB index, Dunn index.
- Work of Deep Clustering has been proposed in research paper which is under the review phase for ICONIP conference

Machine Learning Intern
Pragnakalp Techlabs, Ahmedabad-India   2021 - 2021
Mentor : Mitul Patel
- It is NLP based product provider start-up, which use Working on appointment booking system using Dialogflow platform
- Build Docsaar- the website for document parsing using
- Hand gesture digit-detection system using CNN using self generated dataset by detecting ROI for each data

PUBLICATIONS

"*Song playlist generator system based on Facial Expression and Song Mood*"

• Presented in 1st International Conference of Artificial Intelligence and Machine Vision AIMV-21,
• link : https://ieeexplore.ieee.org/document/9670976
• Research paper is based on emotion detection and song mood prediction using different ML and DL Algorithms.

PROJECTS

Natural Language Query Interface with GPT-3 and MySQL:
- Designed and implemented a Natural Language Interface to interact with a **MySQL database** using OpenAI's GPT-3 API. Leveraged the **text-davinci-003** model of **GPT-3** to translate natural language queries into SQL statements. Post data retrieval utilized the text-DaVinci model from GPT-3 to craft a user-friendly response based on the SQL query and its results. The entire backend was seamlessly integrated with a Streamlit-based user interface for enhanced user interaction.

Cricket World-Cup Data Analysis

- Created Power BI report to identify top players for a cricket team on the scrapped data from espncricinfo with Brightdata website tool, clean and transform data with pandas and evaluate various performance metrics for players.
- Used the resulting Power BI Dashboard to select players for various categories (Fast-Bowler, all-rounder, openers/finisher Batsman) and ultimately choose the best 11 for game. This can be used by team-management to select team and it reduce time by 5-6 hour per match.

Hospital Stay length Prediction

- Build web-dashboard using PowerBI as well as Streamlit and deployed on Heroku to come-up with conclusive story and business insights and build predictive model to predict approximate stay length of new patient.
- Perform data cleaning, data engineering as well as missing value imputation before performing EDA.

Deep Clustering on multiscale data

- Distinguishing different objects on the basis of its features which are heterogeneous and multi-dimensional in nature.
- The unsupervised algorithm which can perform the multi-level clustering of images on the basis of different disjoint feature-sets using DB-Scan and adaptive K-means algorithms. Then it merges the different clusters by validating similarity conditions.

Mood Song

• A system which can recommend the song based on user mood.
 • Real time analyzing the user mood based on his/her facial expression and after that find most appropriate song make uses mood energetic from his /her past listened history with the help of Spotify API.

Virtual Mouse

• The main idea behind this project is to enable the curser moment with hand fingers gestures. Used OpenCV for image processing and apply different filters to locate fingers.

Reviews Classifier

• Build text-summarization(Abstractive as well as Extractive summarization for classification of customer review.)

Bharat HandiCraft Website (E-commerce)

- Developed a website for Handicraft Item dealer to take order for different products.
- Used ReactJS framework to develop the frontend UI, NodeJS to handle backend queries and Used MongoDB for database.

Disk-Scheduling Algorithm Visualizer

• website to visualize the working of disk scheduling algorithm like FCFS, LOOK , C-LOOK, SCAN & C-SCAN.

PROFESSIONAL ACHIEVEMENTS

Finalist of SSIP Hackathon 2020
• State-Level hackathon (TEAM)
• Garbage-collection system automation using AI

Winner of HACKFACT in 2020
• PDPU University In-House Hackathons(TEAM)
• Restaurant management system

Hacktober-Fest Completer
• Open-Source Contribution (SOLO)
• Select in top 10000 from 169000

PROFESSIONAL CERTIFICATES AND COURSES

- **Joy of Computing with Python (NPTEL) , IIT Madras**

- **SQL, Stanford-Online**

- **What is Data-Science?(IBM-Coursera)**

PROFESSIONAL SERVICE

Pypower Project (Youtube channel)
Co-founder of python tutoring channel (link : https://www.youtube.com/c/PyPowerProjects)

LANGUAGES

English : Frequent speaker
Gujrati : Mother-tongue
Hindi : frequent

COMPUTER SKILLS

Programming Languages & Framework : Python, R, C++, C, JavaScript, HTML, CSS, NodeJS, REST API

Data Science Skills : Data Engineering, Predictive Modeling, Power BI, Data-Visualization, Time-Series Analysis, Decision Analysis, Big Data Analysis (Hadoop)

Database and hosting platforms : SQL, MySQL, MongoDB, Git, Heroku Model

Development Algorithms : Linear Regression, Logistic Regression, SVM, K-means, DB-Scan, KNN, Decision Tree, ANN, RNN, CNN, Deep Learning, Machine Learning