Kevin Patel

East Lansing, MI | +1 517-505-9591 | kevin18patel@gmail.com | LinkedIn | GitHub | Portfolio

Results-driven Data Scientist with 2+ years of experience in developing Data Science solutions for enterprise purposes. Proficient in Python, Machine Learning techniques, deep learning techniques, and Natural Language Processing. Skilled in data analysis, ETL, and cloud deployment. Excited to apply skills in a new industry and learn from experienced professionals.

EDUCATION

Master of Science in Data Science

Michigan State University | Aug 2022 - April 2024 GPA: 4.0/4.0

Bachelor of Technology in Computer Engineering

Pandit Deendayal Energy University | July 2018 - May 2022 GPA: 9.59/10.0

PROFESSIONAL EXPERIENCE

Graduate Research Assistant

CMSE Department, Michigan State University | Oct 2023 – Present

- Working with Attention-based GNN and DNA-BERT.
- Research objective: Finding Gene-expression using 3D structural data.

Data Science Intern

VERN.AI (Tucknologies Holdings Inc.), East Lansing, USA | May 2023 - Aug 2023

- Developed a user-centric offline Question-Answer bot using the Vicuna(7B) LLM and Langchain.
- Built an interactive bot utilizing LLMs to facilitate symptom checking and deliver personalized recommendations based on the symptoms.

Data Scientist

India Meteorological Department, Bhopal, India | Aug 2021 - July 2022

- Translated unlabeled RADAR data into actionable insights using Python libraries (scikit-learn, OpenCV), & rectified cloud height measurement errors for a 25% accuracy improvement, enhancing weather data precision
- Curated & classified 5 years of historical data from media weather reports to build a training dataset, employing Convolutional Neural Networks, Text Analytics & Natural Language Processing
- Enhanced data analysis through dimensionality reduction, feature extraction & interpolation. Achieved a 20% accuracy boost in Central India's weather forecasting while executing Long Short-Term Memory modeling
- Developed a predictive model for wind data accuracy in harsh weather forecasting, revolutionizing harsh weather forecasting.
- Implemented Unsupervised Approach for clustering scattered cloud data by merging clustering techniques. Improved clustering accuracy by 51% using a new approach of Deep Clustering for Multi-level Image Segmentation.

Machine Learning Intern

Pragnakalp Techlabs. Pvt. Ltd, Ahmedabad, India | May 2021 - July 2021

- Developed and managed the product Docsaar - a website for document parsing.
- Gathered and managed clients incoming features and design change requirements. Designed question-answer chatbot according to the requirements of the client.
- Used Dialogflow NLU, implemented multilingual speech and text support through a Flask application, and seamlessly integrated the chatbot with an interactive and user-friendly UI.

PROJECTS

Relation Extraction [link]

- Technologies: Bi-LSTM, CNN, BERT, ADAM Optimizer
- Description: Built and compared the performance of three different approaches (Bi-LSTM, CNN, and BERT) for sentence entities relation extraction.
    - Improved F1 Score per entity by 4.2% and Macro F1-Score by 3.6% .
    - Implemented feature fusion attention based architecture with CNN to enhance accuracy.
    - Utilized the transfer learning capability of the BERT model to see how well they can perform on RE task.

Potential Client Profiling using LLM and Client Recommendation (poster)

- Technologies: GPT 3.5 Turbo, Langchain
- Description: Focused on finding clusters of similar content websites using LLM models.
- Created a retrieval-augmented generation (RAG) architecture.
- Generated embeddings for large documents using GPT 3.5 Turbo.
- Currently working on a summary generator model to enhance RAG accuracy and overall clustering.
- Engineered a Knowledge Graph (KG) from scrapped text and built a bot for question-answering using Langchain GraphCypherQAChain.

Natural Language Query Interface with GPT-3 and MySQL

- Technologies: GPT-3, MySQL
- Description: Designed an interface to interact with MySQL database using OpenAI's GPT-3 API.
- Leveraged GPT-3 to translate natural language queries into SQL statements.
- Crafted user-friendly responses based on SQL query results.

Cricket World-Cup Data Analysis

- Technologies: Pandas, SQL, Power BI
- Description: Created a Power BI report to identify top cricket players, reducing selection time by 5-6 hours per match.
    - Performed ETL on scrape data from espncricinfo with Brightdata.
    - Evaluated various performance metrics for players. [link]

Hourly Energy Demand/Load Forecast

- Technologies: PySpark, Spark MLlib, AWS RDS, Seasonal ARIMA
- Description: Set up a PySpark streaming ETL pipeline for real-time energy demand.
    - Applied statistical tests and performed grid search for forecasting models.
    - SARIMAX outperformed some deep learning models (Facebook Prophet, and GRU) (R2: 0.83).

Hospital Data Analysis [dashboard link]

- Technologies: Streamlit, Heroku, Regression Analysis
- Description: Conducted data cleaning, analysis, and missing value imputation.
    - Developed a web-dashboard for conclusive story and business insights.
    - Built a predictive model for predicting approximate stay length of new patients.

PUBLICATION

Song Playlist Generator System Based on Facial Expression and Song Mood

Presented in AIMV-21

- K. Patel and R. K. Gupta, "Song Playlist Generator System Based on Facial Expression and Song Mood," 2021 International Conference on Artificial Intelligence and Machine Vision (AIMV), Gandhinagar, India, 2021.(link)

SKILLS

Expert : C, C++, Python, Statistics, probabilistic modeling, EDA, Machine Learning, Tableau, Scikit-learn, PyTorch, Prompt Engineering, LLM, SQL, Data Structure and Algorithm , NLP, Deep Learning, ReactJS, Firebase. Recommendation System,

Intermediate : R, GAN, GNN, Tensorflow, Knowledge-Graph, Torch, AWS Sagemaker, MongoDB, Node JS, Azure, GCP

Achievements and Extracurricular

- **Finalist of SSIP Hackathon 2020 (State-Level hackathon)**
- **Winner of HACKFACT in 2020**
- **Hacktoberfest Completer (in top 10000 from 169000)**
- **"PyPower Projects" YouTube Channel - Link:**PyPower Projects **,Facilitated 4+ seminars to help 250+ students at university.**