GAURI TRIPATHI

gauritripathi2002@gmail.com LinkedIn Github

EDUCATION

SRM University, KTR **Sept '20 – June'24**
B.Tech in Computer Science Engineering *9.02 CGPA*

RESEARCH PAPER

Enhanced Medical Decision Support using GPT | **Link** **Jul '23 – Nov'23**
FCOM-FINTECH Conference *Dept. of Computing Technologies, SRM KTR*

EXPERIENCE

Surya Consultants **Jun '23 – Dec'23**
Data Scientist - Intern

- Developed and implemented a dynamic client data analysis system using Python and SQL, leading to a 30 percent improvement in data processing efficiency.
- Employed ML models like Logistic Regression for predictive modeling and trend analysis, resulting in a better analysis accurate market forecasts for clients.
- Collaborated in the creation of an interactive dashboard using Plotly, Seaborn for real-time data visualization, enhancing client presentations and decision-making processes.
- Conducted comprehensive data cleansing using SpacyNLP, ensuring a increase in data readability.

PROJECTS

Enhanced Medical Decision Support using GPT | *Transformers, Bitsandbytes, Torch* | Github

- Advanced AI Model: Built on Mistral-7B, a powerful model that can generate thoughtful and accurate medical advice.
- Quantized for Efficiency: Future enhancements include quantizing the model to 2, 4, and 8 bits for task-specific use cases, optimizing performance without compromising accuracy.
- High-Quality Interaction: Ensures factual and nuanced communication, prioritizing user well-being and understanding.
- Logical / IQ Tasks: Perform roleplays, and chain-of-thought reasoning.

COVID-19 Data Analysis Project | *SQL, Apache Superset* | Github

- Comprehensive analysis of COVID-19 data, focusing on various aspects like death rates, infection rates, and vaccination progress across different countries
- Apache Superset Integration: these queries will be implemented on SQL Apache Superset to enhance data visualization and dashboarding capabilities

Starbucks Nutritional Analysis | *Pandas, Matplotlib, Seaborn* | GitHub

- Data Loading, Preprocessing, and Exploration: using Pandas, loading and exploring the Starbucks dataset. Conducted thorough initial data analysis and optimized dataset
- Statistical Analysis and Data Cleaning: extracting key insights like mean, standard deviation, and quartiles
- Matplotlib and Seaborn for creating insightful visualizations including scatter plots, bar plots, and histograms to elucidate relationships between various nutritional elements.

TECHNICAL SKILLS

Languages: Python, SQL
Technologies: Anaconda, Apache, Linux, Jupyter, MySQL, Azure, WordPress
Developer Tools: Git, Docker, Kubernetes, VS Code, AWS Sagemaker, AWS EC2
Libraries: PyTorch, Numpy, Pandas, Scikit-Learn, Seaborn, FastAPI, Matplotlib
MLOps: WeightsBiases

EXTRA-CURRICULAR

- AMVMUN 2018, 2019: Special Mention in the AMVMUN for the All India Political Party Meet.
- Sub Editor Photographer: Held the post for the School Photography Committee