Piyush Jadhav

619-642-5973 | San Diego, CA, USA (Ready to relocate)
pjadhav5510@sdsu.edu | LinkedIn | Website | GitHub

SUMMARY

Highly motivated and innovative machine learning engineer with a strong background in natural language processing (NLP) and multimodal technologies. Demonstrated ability to leverage advanced technologies to solve real-world problems and impact millions of users. Proven track record of research and development in automated content generation, search enhancement, and chat experiences. Strong analytical and problem-solving skills with excellent communication and project management abilities.

WORK EXPERIENCE

AI4Business Labs, SDSU |*Graduate Research Assistant* | San Diego, USA **May 2023 – Present**
(Python, TensorFlow, Keras, deep learning, CNN, RNN, NLP, PowerBI)

- Developed multimodal content understanding algorithms on Large Language Models (LLMs), significantly improving user experience.
- Spearheaded the development of an innovative social media Image Captioning Bot, leveraging deep learning and NLP techniques to automatically generate descriptive captions for uploaded images.
- Achieved an impressive accuracy of 85% in generating contextually relevant captions, enhancing user engagement and accessibility.

FIS Global | Machine Learning Engineer | Bangalore, India **Mar 2021 – Aug 2022**
(Python3, PyTorch, Transformers, AWS ECS, AWS Lambda, Docker, NLP, TensorFlow)

- Achieved a high accuracy rate of 95%, leading to a reduction in fraudulent transactions, saving the company substantial monetary loss.
- Architected and implemented a feature engineering ETL pipeline, extracting relevant transactional and behavioral features.
- This enhanced model performance by 15% and reduced false positives by 20%, improving the overall efficiency of the model.
- Spearheaded the deployment of a scalable fraud detection model on AWS Lambda and ECS. Ensured real-time processing of high-volume transactions, achieving 99.9% uptime, and reducing latency by 40%, thereby enhancing fraud detection capabilities.
- Developed a proof-of-concept for deploying an architecture on AWS using AWS Lambda, API Gateway, MongoDB for fraud detection.

Analytics Quotient | Machine Learning Engineer Intern | Pune, India **Jun 2020 – Nov 2020**
(Python, Tableau, SQL, JavaScript)

- Developed automation scripts using Python and integrated them into data processing pipelines, reducing manual effort by 30%.
- Used machine learning techniques to build models for forecasting sales and customer behavior, impacting in 25% increase in revenue.
- Created interactive data visualizations using Matplotlib and Seaborn, enhancing understanding and interpretation of analytical results.

TECHNICAL SKILLS

- **Languages**: Python, R, SQL, Java
- **Databases**: MySQL, MongoDB, Redis
- **Libraries**: Pandas, NumPy, SciPy, Seaborn, Matplotlib, nltk, TensorFlow, Keras, scikit-Learn, LangChain, xgboost, PySpark
- **Tools**: Power BI, Tableau, Azure, Docker, AWS, GCP, Jupyter Notebook, GitHub, LLM, Llama, Dolly, Excel VBA, Google Sites

ACADEMIC PROJECTS

Russia Ukraine War Sentimental Analysis
(Python, BERT, Perceptron, LLMs, KNN, Logistic Regression, SVM, TF-IDF)

- Spearheaded comprehensive sentiment analysis on social media data regarding the Russo-Ukrainian War, employing advanced machine learning techniques including BERT.
- Conducted thorough model evaluations, revealing BERT's consistent outperformance over models Logistic Regression, KNN and SVM.
- Optimized text preprocessing and feature extraction techniques to enhance data quality and facilitate model performance.

Image Caption Generator using PySpark
(PySpark, nltk, Python, LLM, GCP, COCO, LSTM, TensorFlow, Keras)

- Developed and executed distributed algorithms using PySpark for scalable machine learning model training for image analytics & text analytics on Google Cloud Platform (GCP), efficiently managing vast datasets to enhance model convergence rates significantly.
- Successfully trained an advanced Machine Learning model by innovatively combining LSTM networks with innovative transformer-based architectures, resulting in a significant 5% increase in accuracy compared to the established baseline model.

US Homelessness Predictive Factors using Machine Learning | *Link*
(Python, R, lfe, Tableau, Google Sites, ArcGIS)

- Demonstrated proficiency in utilizing Tableau for comprehensive data visualization, alongside analytical techniques in R and Python.
- Successfully executed data reduction on a dataset with 332 variables to 13 variables, employing data cleaning and EDA techniques.
- Developed and implemented fixed-effect linear regression models using R's lm () and lfe package's felm () function, to analyze data across HUD-designated regions over an 8-year period, achieving a model with a 97% R-squared value.

EDUCATION

San Diego State University, San Diego, CA, USA | Master of Science in Big Data Analytics **Aug 2022 – May 2024**
University of Pune, Pune, MH, India | Bachelor of Engineering in Computer Science **Aug 2016 – May 2020**