SAI CHEEDEPUDI

DATA SCIENTIST • Email • (630) 699-4860 • LinkedIn • Github • Portfolio

PROFESSIONAL SUMMARY

Versatile and results-driven Data Scientist with a proven track record in leveraging advanced analytics and machine learning techniques to drive impactful business outcomes. Experienced in designing and implementing NLP applications, developing predictive models for vehicle recommendations, and optimizing data workflows for enhanced efficiency. Adept at leading end-to-end data science projects, from concept and modeling to deployment and continuous improvement. Proficient in utilizing cutting-edge technologies and cloud platforms to deliver data-driven insights and solutions.

EDUCATION

SOUTHERN ILLINOIS UNIVERSITY | Carbondale, IL
***Master of Science; Major in Computer Science* | 4.0 | 2021-2022**

- Thesis: Cause and Effect of Digital Distress in Illinois State

Data Collection, Pre-processing, ML Modeling, Sentiment Analysis, Data Analytics, Statistical Analysis.

ANDHRA UNIVERSITY | INDIA
***B.Tech; Major in Computer Science* | 8.8 | 2018**

- Final Year Project: Fuzzy Preference Model

TECHNICAL SKILLS

Programming Languages	Python, Java, R, C, C++, SQL
Web Development Frameworks and Libraries	Spring, Django, LangChain, Flask, Fast API, Angular
Data Science	Numpy, Pandas, Scikit-learn, Tensorflow, Pytorch, Pyspark, Hugging Face Transformers, CNN, NLP, LLM, OpenAI, LLaMa, LSTM, Chat GPT, BERT, Linear Regression, Logistic Regression, SVM, Random Forest, Decision Tree, K-Means, Forecasting, Reinforcement, KNN, XGBoost, Bagging
Visualization Tools And Libraries	Tableau, Power BI, Seaborn, Matplotlib
Development Environments (IDEs)	Eclipse, PyCharm, Jupyter, Spider, Visual Studio
Tools and Technologies for Deployment and DevOps	Git, Redis, Google Kubernetes engine (GKE), Docker, Heroku, Streamlit, GitLab, GitHub

PROFESSIONAL EXPERIENCE

Data Scientist | Carmax REMOTE **| May, 2022-Present**
Python, MySQL, Jupyter, Tensorflow, Pytorch, Pyspark, Keras, Seaborn, Matplotlib, Power BI, Tableau, ML Models, Deep Learning, Huggingface Transformers, LLM, Gen AI, NLP, Git, FAST API, AWS, Azure, Databricks, Django

- Led the development and implementation of cutting-edge recommendation models, pattern recognition and clustering of users, improved precision by 20% using advanced models (**LSTM, BERT, SVD, MLPRegressor**) for personalized user interactions, resulting in 15% increase in revenue generated from vehicle sales.
- Implemented natural language processing techniques for text generation tasks using models like OpenAI's GPT (Generative Pre-trained Transformer), LSTM , and BERT models to optimize customer interactions, resulting in a 15% increase in conversion rates and a 25% decrease in customer churn.
- Spearheaded the enablement of CI/CD through code versioning control, repository maintenance, and automated deployments; leveraged GitLab, Jenkins, Docker, and GKE on AWS, Azure, and GCP to streamline software delivery, resulting in increased development team productivity by 25%.

- Conducted comprehensive evaluation of recommendation algorithms, leveraging performance metrics such as Click-Through Rates (CTR) and employing **hypothesis testing, A/B testing, and t-tests**.
- Contributed to data-driven decision-making and continuous improvement of algorithms, resulting in a 20% increase in user engagement and a 15% boost in conversion rates.

Graduate Research Assistant | Southern Illinois University **Carbondale, IL | Aug, 2021-May, 2022**

Python, R, MySQL, Jupyter, Tensorflow, Pytorch, Pyspark, Keras, Seaborn, Matplotlib, Power BI, Tableau, ML Models, Deep Learning, Federated Learning, Differential Privacy, TimeSeries Forecasting, Huggingface Transformers, ARIMA, Alexnet, Squeezenet, Resnet, Gen AI, LLM, NLP, Git, AWS, Django

- Enhanced data management processes by implementing feature engineering and handling missing values techniques, leading to a 30% improvement in data consistency and reliability for model development.
- Developed and implemented generative models for image synthesis and text generation tasks, resulting in a 25% improvement in data augmentation capabilities and increased diversity in generated content.
- Implemented a deep learning model utilizing multi-GPU parallel processing for accelerated training and enhanced model performance, resulting in a 50% reduction in training time and a 20% increase in model accuracy.
- Mentored and guided 100+ students in Computational Statistics, providing expert guidance on assignments, subject quizzes, and diverse projects; boosted student performance.

Software Product Developer - Data Science | Opsramp acquired by HPE **Hyderabad, IN | May, 2018-July, 2021**

Python, Java, Jupyter Notebook, Visual Studio, Tensorflow, Pytorch, Keras, Seaborn, Matplotlib, ML Models, Deep Learning, LLM, NLP, Django, Microservices, Spring, Eclipse, MySQL, Cassandra, Elasticsearch, Kafka, Redis, Memchache, HTML/CSS, JS, jQuery, Vue, React, GCP, AWS, Git, Jira, Docker, Confluence, Kubernetes, Linux

- Engineered sophisticated models for anomaly detection in IT infrastructure monitoring, leveraging statistical analysis and machine learning techniques for template recommendation; achieved a 50% decrease in FPR, enhancing system reliability.
- Spearheaded the implementation of new functionalities, including real-time monitoring and availability of resources, leveraging metric and availability graphs to optimize system performance by 25%.
- Pioneered in OpsRamp migration from monolith application to microservices. Restructured 25% functionalities to microservices.
- Infrastructure Tab Revamping: Created a new UI and backend for various functionalities in the infrastructure tab.
- Facilitated in Data Migration from MySQL to ElasticSearch; built APIs in Java and dealt with ElasticSearch queries; and improvised Advanced search feature by enhancing the performance and user experience by 25%.
- Written over 500 Unit Test cases at service layer and database layer using JUnit and Mockito. Used Selenium to perform UI automation testing.
- Implemented two-factor authentication using Spring Security for an additional layer of user verification to enhance security.
- Leveraged Jenkin pipelines to manage the build, test, and deployment procedures, used Apache Maven to create and maintain build scripts, streamlining the building procedure and analyzed the quality of code using SonarQube.
- Managed real-time production issues of OpsRamp increasing customer satisfaction by 10% and performance of the overall product.

PROJECTS

Diabetic Retinopathy | *Python, Pandas, Machine Learning and Deep Learning Models, GPU, Jupyter, Visual Studio*

- Developed a deep learning model that classified Retinopathy in diabetic patients by analyzing high-resolution retina images, achieving an average accuracy of 89%.
- Implemented advanced differential privacy techniques to protect sensitive user data, resulting in a 50% reduction in data breaches and ensuring compliance with privacy regulations.

Grading System | *Java, Multi-Threading, Client-Server Architecture, Eclipse*

- A Client-Server Architecture that can act as a Classroom Grading System.
- Orchestrated the implementation of an innovative server-based grading system, revolutionizing the assignment process and enhancing efficiency by automating grading procedures, resulting in a 40% decrease in grading turnaround time.

Pallet Forecasting | *Python, ARIMA, FAST API, Evidently AI, Jupyter, Mysql*

- To minimize the inventory cost of Pallet production by a time series forecasting model, ARIMA.
- Used Fast API for deployment, Evidently AI for Monitoring and maintenance for time series analysis and forecasting.

Generative AI Projects| *Python, Transformers, Hugging Face, GPT, Fast API, Gradio, Google Colab,Visual Studio*

- Developed a generative chatbot using state-of-the-art language models such as GPT, providing users with dynamic and contextually relevant responses. Achieved a 30% increase in user engagement and satisfaction
- Implemented fine-tuning strategies to tailor the chatbot's responses to domain-specific requirements, text-to-image generation, text generation, document summarization models, resulting in more accurate outcomes.