CHANDINI SAISRI UPPUGANTI

📞 669-467-3738 ✉ chandinisaisri.uppuganti@sjsu.edu 🔗 linkedin.com/in/chandini/ ⬛ github.com/chandini2595

Education

San Jose State University, San Jose, CA   **Aug 2024 – May 2026**
Master of Science in Computer Software Engineering in Data Science   *GPA - 3.8/4*

Technical Skills

Languages: Python, SQL, C, Java, R, PL/SQL, PySpark, Shell
Operating Systems: Linux, Windows
Cloud & Databases: MySQL, Oracle, GBQ, SAP ECC, Snowflake, AWS (Lambda, QuickSight, EC2, S3, RDS, SQS, IAM, ECS, Glue, Bedrock), Azure (Data Lake Storage, Synapse), NoSQL (MongoDB), Informatica, Trello, Jira, Databricks
Machine Learning: scikit-learn, TensorFlow, Keras, PyTorch, Pandas, NumPy, Matplotlib, Seaborn, Pycaret, Autogluon, Colab, OpenCV, GANs, Transformers, LSTM, Computer Vision, Responsible AI, NLP, CoreFlow , Deep learning, Neural Networks
Other Tools: SSIS, IICS, Informatica Powercenter, Docker, Tableau, Power BI, Github, Ansible, Google Colab, Kubernetes, Jupyter Notebook, PowerBI

Projects

AI based Legal Brief Argument Linker | *Python, MiniLM, BERT, HuggingFace Transformers, Unsloth*   **April 2025**
- Developed a pairwise classification model using fine-tuned **MiniLM** and **BERT** via **HuggingFace Transformers** and **Unsloth** to link argument–counterargument sections in legal briefs; built preprocessing pipelines for JSON datasets and deployed an interactive **Streamlit** UI to visualize and explore predictions for legal research assistance.

PersonaRank: AProduct Scoring Engine | *Python, TensorFlow, PyTorch, NLP, Graph Analysis, scikit-learn*   **August 2024**
- Designed a hybrid product scoring system using **NLP embeddings** (BERT, Word2Vec) and **graph-based similarity** to personalize search relevance; built and evaluated ranking models (Logistic Regression, SVM, Gradient Boosting), and optimized preprocessing pipelines to enhance model training by 30% and user engagement by 25%.

LLM-RAG (Retrieval-Augmented Generation) | *LLM (OpenAI, Gemini), Python*   **January 2024**
- Pioneered development of a sophisticated Retrieval-Augmented Generation (RAG) model, integrating LLM with external databases, boosting poem retrieval precision by 45% and generating highly meaningful and contextually rich poetic content.

Sales Forecasting and Optimization System | *Python, scikit-learn, Pandas, TensorFlow*   **March 2022**
- Developed a Sales Forecasting and Optimization System using ARIMA, SARIMA, and LSTM, improving forecast accuracy by 20% through EDA, trend analysis, and feature engineering (lag variables, moving averages, economic indicators), while optimizing models with cross-validation and hyperparameter tuning, reducing RMSE and MAE for precise sales predictions.

Experience

Senior Data Engineer: Informatica, Hyderabad, India   **Jul 2022 – Jul 2024**
- Architected and automated robust ETL pipelines to collect, clean, and validate data from diverse on-premise and **cloud databases** using **Informatica** and **AWS Glue**, boosting processing efficiency by 30% and ensuring high data accuracy.
- Performed in-depth data analysis on petabyte-scale datasets using complex SQL queries to identify trends and patterns. Built and maintained dynamic **Tableau** dashboards to effectively communicate insights to stakeholders.
- Developed reliable, automated data migration workflows using Informatica PowerCenter and custom **Python** ETL scripts (version controlled in **GitHub**), orchestrated with Apache Airflow for scalable scheduling and monitoring, achieving 99% data quality accuracy.

Data Engineer: Wipro, Bangalore, India   **Jul 2020 – Jul 2022**
- Designed and implemented a 5-zone architecture for migrating data from SAP ECC & SOAP to **Azure** Data Lake Storage and Synapse Analytics, supporting efficient database design and **performance optimization**.
- Built and maintained secure, structured ETL pipelines using IICS, improving **data integrity** by 25% and ensuring seamless integration into analytical databases.
- Implemented automated data quality checks and validation processes to uphold data accuracy and consistency across storage layers, while processing large volumes of **big data** using **Hadoop**.

Achievements and Extracurricular Activities

- **Emerging Consultant of the Year** at Informatica for accelerating project delivery through rapid mastery of adv analytics.
- **Cloud Data Quality R41** Professional Certification
- **Cloud Data Integration R41** Professional Certification.
- **AWS Cloud** Practitioner Certification.
- Participated in the **Stanford Law Hackathon**, developing AI-powered solutions for legal document analysis and contract intelligence using LLMs and retrieval-based systems.