

Kusuma Satya Sreeja Chalasani

kusumasatyasreeja@gmail.com | 940-629-3270 | www.linkedin.com/in/kusuma-chalasani | Denton, Texas, 76201

Professional Summary

Data Scientist and AI/ML Engineer with hands-on experience building data-driven and AI-powered applications using Python, SQL, and cloud platforms. Skilled in developing NLP-based solutions, working with Large Language Models (LLMs), and designing end-to-end pipelines from data processing to model evaluation and deployment. Experienced in building real-world projects including intelligent job matching systems, voice-enabled assistants, and GPT-powered applications. Strong foundation in data analysis, machine learning, and automation, with a growing focus on agentic AI systems and scalable AI solutions. Passionate about leveraging data and AI to solve real-world problems and improve user experiences.

Skills

- **Programming Languages:** Python (Pandas, NumPy), SQL
- **Artificial Intelligence & Machine Learning:** Natural Language Processing (NLP), Supervised & Unsupervised Learning (Regression, Classification, Clustering), Model Training & Evaluation, Prompt Engineering, OpenAI API, LLM-based Application Development, AI Agents, Retrieval-Augmented Generation (RAG)
- **Data Processing:** Data Cleaning, Feature Engineering, Data Transformation, Data Validation, Data Pipeline Fundamentals
- **Databases:** PostgreSQL, SQLite
- **Cloud & Automation:** Microsoft Azure, Google Cloud Platform, Cloud-based Data Processing
- **Tools & Development Environment:** Jupyter Notebook, Google Colab, GitHub, Visual Studio Code (VS Code)

Work Experience

University of North Texas - Student Assistant - IT / Departmental Operations

Dec 2024 - Present | Denton, TX

- Managed and maintained operational datasets for 240+ university devices, ensuring data accuracy and availability
- Analyzed recurring technical issues and generated data-driven insights to support process improvements
- Assisted in data tracking, reporting, and documentation workflows for IT operations
- Supported system deployment, updates, and maintenance following standardized procedures
- Collaborated with IT teams to ensure system reliability, minimal downtime, and efficient operations
- Contributed to organizing structured data for reporting and analytics use cases

Cloud4c Company - Cloud Support Engineer | Data Engineering & Operations

June 2022 - July 2024

Client: DHL - ServiceNow

- Built and supported data pipelines for cloud usage and performance data using Python and SQL, enabling analytics and machine learning workflows
- Processed large-scale datasets to support data-driven decision-making, analytics, and optimization
- Automated data workflows using Python and PowerShell, improving efficiency and reliability
- Performed trend analysis on cloud storage and performance metrics for infrastructure planning
- Worked with Azure and GCP environments for data processing and storage workflows
- Prepared structured datasets to support analytics and machine learning use cases
- Collaborated with cross-functional teams to ensure data consistency and system reliability

Project Experience

AI-Powered Data Scientist Agent & Explainable AI Dashboard

March 2026 - Present

- Developed an AI-driven system to automate data analysis, model selection, and prediction workflows using Python and machine learning techniques
 - Implemented automated EDA, feature engineering, model evaluation pipelines to streamline end-to-end data science tasks
 - Integrated explainable AI (XAI) techniques to provide model interpretability, actionable insights through an interactive dashboard
 - Designed a user-friendly interface for uploading datasets and generating real-time insights and visualizations
- Technologies:** Python, Scikit-learn, Pandas, Streamlit/React, OpenAI API

RAG-Based AI Interview Prep Assistant

March 2026

- Designed and developed an end-to-end Retrieval-Augmented Generation (RAG) system enabling document-based questions using dynamic PDF ingestion
 - Implemented document chunking (100-300 words), embedding generation, and vector-based retrieval using ChromaDB to improve answer relevance
 - Built an interactive Streamlit UI for document upload, querying, and display source-backed responses with file and page-level metadata
 - Engineered prompt strategies to ensure responses are grounded in retrieved context and handle out-of-scope queries without hallucination
 - Integrated full pipeline: ingestion -> chunking -> embedding -> retrieval -> generation using LangChain and Groq LLM
 - Improved system transparency and explainability by incorporating traceable source references in responses
- Technologies:** Python, LangChain, ChromaDB, Streamlit, Sentence Transformers, Groq API

AI Guardian - Fraudulent Job Detection System

Jan 2026 - Present

- Developed an AI-based system to detect fraudulent job postings using machine learning and NLP techniques
 - Performed text preprocessing, feature extraction, and classification to identify suspicious job descriptions
 - Applied models such as Logistic Regression and Random Forest to improve fraud detection accuracy
 - Evaluated model performance using accuracy, precision, recall, and F1-score
 - Leveraged dataset from Kaggle and implemented end-to-end data pipeline including cleaning, training, and evaluation
- Technologies:** Python, Scikit-learn, NLP, Pandas, Machine Learning

CMHT-IT Voice Assistant

June 2025 - Aug 2025

- Developed a full-stack, voice-enabled campus navigation web app using Flask REST API, SQLite, and JavaScript, integrating Web Speech API for real-time recognition, gTTS for text-to-speech, NLP-based intent recognition, and a responsive SPA frontend with cross-device compatibility.
- Technologies:** Python, Flask, SQLite, JavaScript, Web Speech API, gTTS, NLP

Intelli AI - Intelligent Job Matcher & Email Notifier

June 2025 - Aug 2025

- Built an AI-driven job-matching system using NLP techniques to match candidate skills with job descriptions
 - Implemented text preprocessing, feature extraction, and similarity scoring to improve matching accuracy (>90%)
 - Integrated automated notifications using Gmail API and logged results for evaluation and monitoring
 - Evaluated model performance and iterated on feature selection to improve accuracy
- Technologies:** Python, PostgreSQL, Gmail API, NLP, AI

Dish Genie (AI-Powered Recipe Assistant)

May 2025 - June 2025

- Developed an AI-powered application using OpenAI GPT API to generate recipe recommendations based on user-provided ingredients
 - Designed prompt workflows to improve response relevance and consistency
 - Integrated AI inference into a backend service and connected it to a responsive frontend
 - Planned enhancements including personalization, dietary constraints, and voice input
- Technologies:** OpenAI GPT API, Node.js, ReactJS, Tailwind CSS, JavaScript

Real-Time Crime Data Streaming & Visualization - Google Cloud Platform

UNT | April 2025 - May 2025

- Developed a real-time crime data streaming and analysis pipeline using PySpark Structured Streaming on GCP Dataproc, integrating data from Cloud Storage. Processed, aggregated, and visualized emergency call data to identify top crime types using Google Colab and Matplotlib.
- Technologies:** Hive, Big Query, PySpark, Google Cloud Dataproc, Cloud Storage, Python, Pandas, Google Colab, Matplotlib

Machine Learning Project- Energy Consumption Forecasting

UNT | Dec 2024

- Implemented machine learning models to forecast household energy consumption using historical time-series data
- Performed data preprocessing, feature engineering, and model evaluation to improve prediction accuracy by 20%
- Compared multiple models and selected optimal approaches based on performance metrics

Technologies: Python, Machine Learning, Data Preprocessing

Machine Learning Project- Emotion Based Music Recommendation System

Loyola Academy | March 2022

- Developed a machine learning system using CNN-based emotion detection from physiological signals and facial expressions
- Processed sensor and image data to classify emotional states and generate personalized music recommendations
- Evaluated model performance and improved accuracy through feature tuning

Technologies: Python, Machine Learning, CNN (Convolutional Neural Networks), Signal Processing

Education

Master of Science, Data Science

University of North Texas | Aug 2024 - May 2026

Bachelor of Science, Data Science and Data Analytics

Loyola Academy Degree & Pg College | 2019 - 2022

Certifications

- Introduction to Model Context Protocol (MCP)
- AI Fluency: Framework & Foundations
- Oracle Cloud Infrastructure Generative AI
- Azure Fundamentals AZ - 900
- AI For Everyone

Anthropic | March 2026

Anthropic | March 2026

Oracle | Oct 2025

Microsoft Certified | Feb 2023

DeepLearning.AI | June 2021

Extracurricular Activities

- Passionate about arts
- Active member of the ISA Cultural Team, contributing to event planning and creative activities