Chaitanya Kumar Battula

P: (774) 251- 8434 **LinkedIn | GitHub | Portfolio** **E:** chaitanyabattula7@gmail.com

Machine Learning | Generative AI | Data Science | Deep Learning | Computer Vision | NLP | LLM | A/B testing | Data Validation | Data Analytics | Models | Statistics | Business Analytics | Cloud Architecture

AI & ML Engineer | Gen-AI Engineer | Data Scientist | Data Analyst

Education:

- **Purdue University, USA:** Master of Computer Science **| GPA: 3.80** **May 2024**
- **California Institute of Technology, USA:** Post Graduation in Cloud Computing **Nov 2022**
- **Purdue University, USA:** Post Graduation in Artificial Learning and Machine Learning **Sep 2021**
- **Indian Institute of Quant Finance, India:** Post Graduation in Algorithmic Trading **Feb 2024**
- **JSS Science & Technology University, India:** Bachelor of Computer Science **| GPA: 3.70** **Jun 2021**

Technical Skills:

- **Programming Language:** Python, Julia, R, C, C++, HTML/CSS,
- **Python Frameworks for ML:** PyTorch, TensorFlow, Pandas, Numpy, Matplotlib, sklearn, Keras, PySpark, Open CV, NLTK
- **Advanced Machine Learning:** ML, DNN, Generative AI, CNN, RNN, Advanced CV & Signal Processing, NLP, LLM, Reinforcement Learning
- **Data Analytics:** Spark, SAS, Tableau, Power BI, Excel
- **Databases:** SQL (SQL Server, MySQL, Postgre SQL) and NoSQL (MongoDB)
- **Business Skills:** MS office suite (Excel, Word, PowerPoint), Time Series Analysis, Statistics, Google WS.
- **Cloud Infrastructure Expertise:** AWS,GCP etc. | **Misc:** Git

Work Experience

Machine learning Engineer- Internship | Radical AI, New York, NY, USA **Jun 2024 - Present**

- As an AI Engineer at Radical AI, developed REX, an AI tutor utilizing Google Gemini and Vertex AI, achieving a 90% user satisfaction rate. Additionally, engineered and deployed end-to-end solutions for the AI Worksheet Generator, Quiz Generation AI, AI Flashcard Generator, and Flight Manager AI Chatbot.

Machine learning Engineer- Internship | Finland Labs, Delhi, India **Oct 2020**

- Led the development of advanced deep learning models for computer vision, focusing on feature extraction, object detection, and image segmentation. Integrated proprietary machine learning algorithms with pre-trained models and IoT via AWS, optimizing real-time data processing and decision-making efficiency.

Project Experience

Flight Manager AI Chatbot Prediction

- Engineered a cutting-edge "Flight Reservation Chatbot" leveraging Google's Gemini language model, integrating advanced NLP for search and booking functionalities. Implemented Google Cloud SDK with Vertex AI, ensuring scalable, real-time performance with response times under 200ms and system uptime exceeding 99.9%. Employed Gemini's advanced function calling to create customized tools for precise flight search and booking, and developed a user-centric interface with Streamlit, enhancing real-time interaction and user experience.

You tube Transcript-to-Flashcard Generator

- Developed "DynamoCard," an AI-powered, open-source tool that revolutionizes digital education methods by integrating advanced frontend and backend technologies for efficient parsing and structuring of extensive YouTube transcripts. Utilized a Semantic Extraction Algorithm (SEA) to autonomously extract and synthesize key concepts from academic lectures and lengthy videos.

Quiz Generation AI

- Engineered an advanced "Quiz Generation System" utilizing Google's Gemini Pro model, seamlessly integrated with Vertex AI and LangChain for optimal NLP performance. Implemented sophisticated embedding techniques and contextual analysis to dynamically generate quizzes from user-provided documents. Deployed a FastAPI backend for high-throughput data processing with latency reduced to under 50ms and employed a Streamlit-based UI, enhancing user interaction and intuitive quiz management.

Real time Speech Translator, Spanish-To-English

- Developed an advanced real-time Spanish-to-English speech translation system utilizing state-of-the-art Generative AI techniques. Integrated automatic speech recognition (ASR) for speech-to-text conversion, followed by Transformer-based machine translation for high-accuracy language conversion, and deliver voice output. Achieved a Translation Quality Metric (TQM) score of 92 and a BLEU score of 0.85, demonstrating exceptional accuracy and fluidity in real-world applications.

Speech Emotion Recognition System

- Employed Mel-frequency cepstral coefficients (MFCCs) to capture nuanced audio features and engineered a robust Recurrent Neural Network (RNN) model equipped with a Connectionist Temporal Classification (CTC) loss function, enabling seamless alignment of audio and text data. Performance evaluation was conducted using Word Error Rate (WER) metrics, showcasing the model's efficacy in accurately transcribing speech with high fidelity. Additionally, developed a secure desktop application (an exe file), that receives audio files as input and outputs the detected emotion in a user-friendly interface, further enhancing the system's accessibility and practical utility.

Demand and Price Forecasting

- Engineered an advanced multi-model demand forecasting framework incorporating ARIMA, SARIMA, Facebook Prophet, LSTM and Temporal Convolutional Network. Each model was strategically selected to address critical forecasting challenges such as stationarity, seasonality, and intricate nonlinear relationships, achieving an average forecasting accuracy of 92%. The system's robust design ensured superior predictive performance and adaptability across diverse applications, enhancing strategic decision-making and operational efficiency.

Customer Life Value Prediction

- Developed an advanced suite of predictive models for Customer Lifetime Value (CLV) using sophisticated ensemble techniques, including Gradient Boosting Machines (GBM) and Random Forests. Conducted extensive feature engineering, dealt outliers with Isolation forest, obtained best features based on feature importance, hyper tuned and evaluated model performance with regression metrics such as Mean Absolute Error (MAE) of 4 and an R-squared (R^2) score of 0.99, demonstrating high precision and predictive accuracy. Obtained LIME and SHAP values to explain the advanced interpretability of the model. This system enhanced client segmentation and marketing strategies.

Anomaly Detection

- Engineered an advanced financial fraud detection system using a Generative Adversarial Network (GAN) with an integrated anomaly detection mechanism. Trained the GAN on normal transaction datasets to identify fraudulent activity through reconstruction errors and anomaly scores. Optimized the threshold dynamically using statistical and adaptive methods, achieving a Precision of 82%, Recall of 89%, and an AUC-ROC of 0.74.

Brain Tumor Detection and Classification using VGG16

- Engineered This project focuses on building a deep learning model to detect and classify brain tumors using MRI images. The model leverages the VGG16 architecture, which is pre-trained on the ImageNet dataset. The objective is to accurately classify MRI images with accuracy of 73% into different categories, such as glioma tumor, meningioma tumor, no tumor, and pituitary tumor.

Customer Segmentation

- Engineered a comprehensive suite of advanced clustering algorithms for segmenting bank customers based on behavior, demographics, and transaction history. Implemented K-means, Hierarchical, DBSCAN, Gaussian Mixture Models, Mean Shift, Agglomerative, and Affinity Propagation, optimizing each for specific data characteristics. Achieved a best Silhouette Score of 0.42 and a Davies-Bouldin Index of 0.83, indicating high cluster quality and separation. These models facilitated highly personalized marketing strategies.