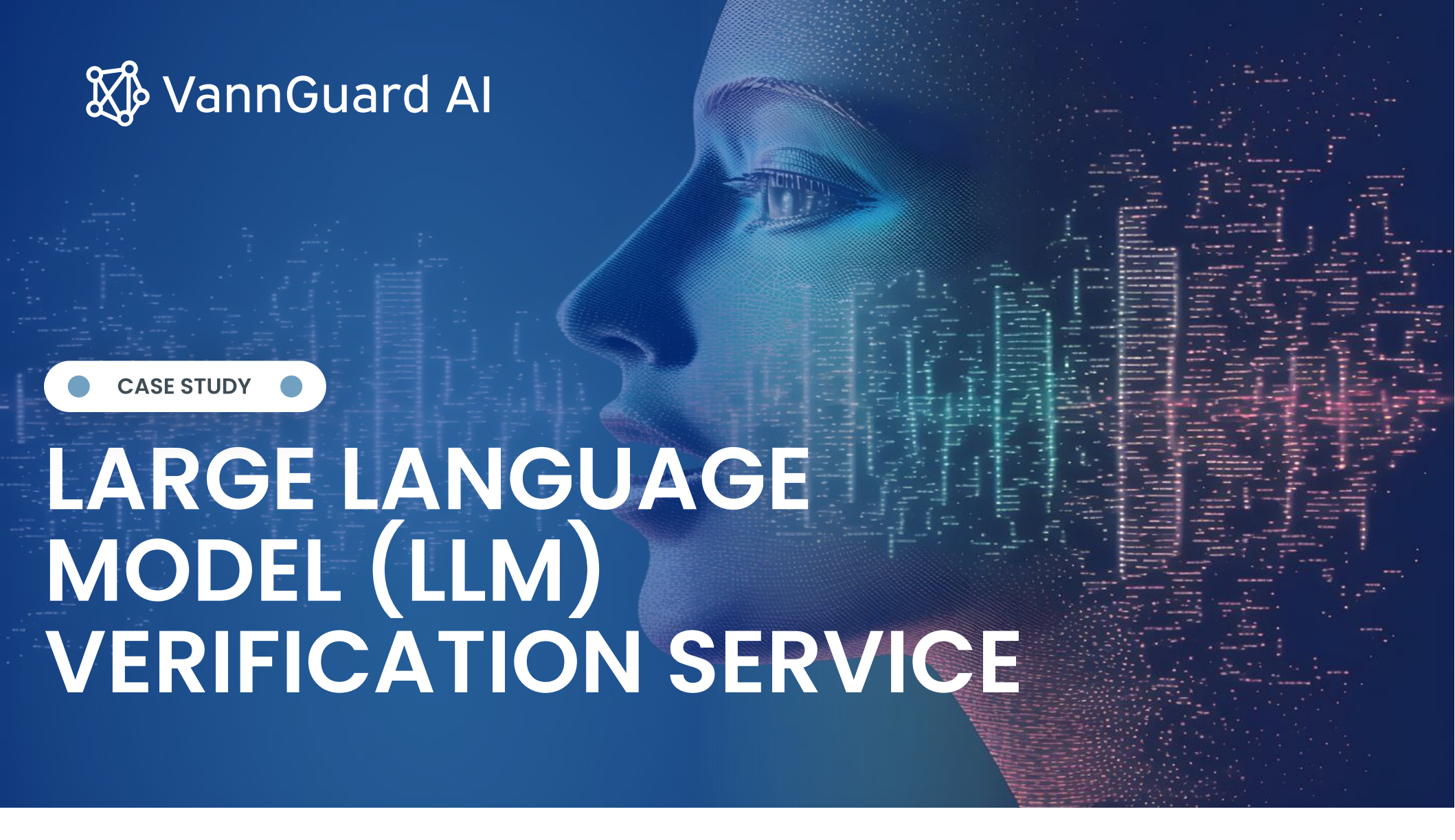
VannGuard AI



CASE STUDY

LARGE LANGUAGE MODEL (LLM) VERIFICATION SERVICE

Overview

Paradigm Networks, a pioneering entity in the generative AI sector, aims to enhance the adoption of AI in enterprises by addressing critical data governance challenges. Their comprehensive approach ensures policy-based access, compliance, and information validation. By focusing on essential use cases such as code generation, legal research, and data privacy, Paradigm Networks provides robust security controls, observability, and verification mechanisms for generated content.

Customer

Paradigm Networks, USA

Country: USA

Customer Size: 1000+

Industry: B2B

Publish Date: 11/04/2024

Problem Statement

With the growing reliance on generative AI, enterprises face significant challenges in validating the accuracy and compliance of AI-generated content. Ensuring that responses from various large language models (LLMs) such as Claude, LLama2, Bard, Gemini, and ChatGPTs meet corporate standards is crucial for fostering trust and reliability in AI solutions. Paradigm Networks needed a robust verification service capable of evaluating and validating the responses generated by these LLMs to uphold the integrity and security of corporate data.

Technical Solution

Our team developed a sophisticated Verification Service integrated as a microservice within the Paradigm Networks application. Initially utilising existing tools like Factool for response verification, we soon embarked on an extensive exploration of the latest LLM Verification and Evaluation research. This led to the creation of a hybrid verification system that combines traditional natural language processing (NLP) techniques with advanced LLMs.

The verification service was meticulously benchmarked against diverse vertical datasets, including healthcare, sports, legal, and financial domains. This rigorous testing ensured the verifiers' robustness and accuracy. The service employed technologies such as OpenAI GPT-3.5, GPT-4, Amazon Bedrock, Bard, Claude, Gemini, Django, Djongo, MongoDB, Huggingface Sentence Transformers, Guardrails-ai, Pandas, Spacy, and NLTK. Upon achieving stable and reliable results, the verifiers were transitioned to production, ready to be deployed for Paradigm Networks' corporate customers.

Results

The implementation of the LLM Verification Service significantly bolstered end-users' trust in using generative AI with their corporate data. By ensuring the validity of the responses generated by various LLMs, Paradigm Networks could confidently offer their AI solutions to enterprise clients. The service's effectiveness was reflected in improved Retrieval Scores, Generation Scores, BLEU Scores, RAGAS Scores, and ELO Ratings, affirming its capability to deliver accurate and compliant AI-generated content. This project exemplifies our commitment to advancing AI technology while addressing critical enterprise needs for data governance and validation.

Technologies

- OpenAI GPT-3.5
- OpenAI GPT-4
- Bard
- MongoDB
- Spacy
- Sentence Transformers
- Claude
- Djongo
- Pandas
- Github
- Guardrails-ai
- Gemini
- Django
- NLTK
- AWS EC2

Domains

- Generative AI
- Natural Language Processing (NLP)
- Natural Language Generation (NLG)
- Reterival Augmented Generation (RAG)
- Corporate LLMs
- LLM Response Verification
- LLM Response Validation
- LLM Response Evaluation