Capstone Project 1

Domain: Financial Services
Techniques: NLP, Machine Learning, Large Language Model (LLM)
Title: Personalized Financial Advisor using Large Language Model (LLM)

Overview and Problem Statement:

The field of finance can be complex and overwhelming for individuals seeking personalized financial advice. In order to make informed decisions regarding investments, retirement planning, budgeting, and financial products, individuals often require guidance from financial experts. The aim of this project is to develop an Intelligent Financial Advisor powered by a Large Language Model (LLM) to provide personalized financial advice and guidance to individuals. By leveraging NLP and machine learning techniques, the Intelligent Financial Advisor will assist users in making informed financial decisions and achieving their financial goals.

Specific Challenges:

1. Training and fine-tuning a Large Language Model (LLM) on financial datasets
2. Ensuring accurate understanding and response generation for financial queries.

Data Description:

Dataset: Alpha Vantage Financial Market Data Source: Alpha Vantage (https://www.alphavantage.co/). Alpha Vantage is a provider of financial market data APIs that offer real-time and historical data for various financial instruments such as stock prices, technical indicators, sector performance, exchange rates, and more. This data should be utilized to train and fine-tune the Large Language Model (LLM). Real-time financial data and market news and trends should be incorporated to provide current and accurate financial advice.

Applications:

1. Improved financial decision-making and goal achievement for individuals.
2. Offering personalized recommendations for investment portfolios, retirement plans, insurance policies, and other financial products based on individual goals and risk tolerance.

References:

1. Yang, Liu and Wang. 2023. FinGPT: Open-Source Financial Large Language Models. *ArXiv, abs/2306.06031*.

Capstone Project 2

Domain: Interior Design
Techniques: Stable Diffusion, Transformers
Title: AI-Driven Interior Design Using Stable Diffusion and Transformers

Problem Statement:

AI-driven innovation in creative industries such as design offers a very good potential. Be it architectural design, landscape design or interior design, the possibilities are endless. Such generated designs have the potential to drive rapid growth and profits in the interior design industry. The goal of this project is to generate realistic and novel interior room designs by leveraging Stable Diffusion and Transformer models, trained on the IKEA Interior Design Dataset. These generated designs can revolutionize the interior design industry by providing rapid, high-quality design options.

Specific Challenges:

1. Distinguishing a good design from a poor one or developing an evaluation metric is a complicated problem because good design is 'subjective'.
2. Ensuring the stable and reliable training of Stable Diffusion models to avoid issues like generating nonsensical designs.
3. Integrating Transformer models to effectively understand and apply design principles and context.

Dataset Description:

The dataset was collected from IKEA.com website for the purpose of building Style Search Engine (note: only for non-commercial use). It consists of: 2193 object (product) photos, 298 context (room scene) photos in which those objects appear, text descriptions for products, ground truth information on which items appear in which rooms.

Applications:

1. Web/app based designing tool for designers and real estate market - similar to Heavenly, Planner 5D
2. Automated interior design suggestions for enhancing user experience and creativity.

References:

1. Image to Image using Artificial Intelligence — Automating Room Interior Design, David Oluyale:
https://medium.com/@oluyaled/image-to-image-using-artificial-intelligence-automating-room-interior-design-5351bdfdd4bf

2. Integrating aesthetics and efficiency: AI-driven diffusion models for visually pleasing interior design generation:

https://www.researchgate.net/publication/378154158_Integrating_aesthetics_and_efficiency_AI-driven_diffusion_models_for_visually_pleasing_interior_design_generation
3. Tautkute et al., 2017, ACICS, IKEA: Interior Design Dataset
4. Denoising Diffusion Probabilistic Models, Ho et al., 2020: https://arxiv.org/abs/2006.11239
5. Attention is All You Need, Vaswani et al., 2017: https://arxiv.org/abs/1706.03762

Capstone Project 3

Domain: Human Resources and Technology
Techniques: Large Language Model (LLM) Fine-Tuning, Natural Language Processing (NLP), Text-to-Speech, Video Generation
Title: Automated Video Creation from Resumes Using GPT

Overview and Problem Statement:

Creating a concise and engaging video summary of a resume can significantly enhance the recruitment process. Traditional methods of resume screening are time-consuming and may overlook key aspects of a candidate's profile. This project addresses the need for an automated solution to convert text-based resumes into 1-minute video summaries, effectively showcasing the candidate's skills, experiences, and achievements. The project involves fine-tuning a GPT model to generate a script from a resume and utilizing text-to-speech and video generation technologies to create the video.

Dataset:

The primary dataset used for this project will be the [Resume Dataset]() from Kaggle. This dataset contains a diverse collection of resumes, covering various industries and roles. The resumes are annotated to highlight key sections such as contact information, summary, skills, work experience, education, and achievements. This comprehensive dataset will be instrumental in training the model to accurately interpret and summarize resume content.

Specific Challenges:

- Ensuring the model accurately interprets the context and importance of various resume sections to create a coherent and relevant script.
- Tailoring the video script to reflect the unique strengths and experiences of each candidate.
- Generating human-like and engaging scripts that effectively summarize the resume content.
- Integrating text-to-speech technology to produce natural-sounding narration and combining it with relevant visual elements to create a professional video.

References:

1. Text-to-Video Generative AI Models: The Definitive List: https://aibusiness.com/nlp/ai-video-generation-the-supreme-list
2. Everything to Know About OpenAI’s New Text-to-Video Generator, Sora: https://www.scientificamerican.com/article/sora-openai-text-video-generator/
3. Resume Video Maker: https://www.steve.ai/resume-video-maker

Capstone Project 4

Domain: E-commerce and Fashion
Techniques: Deep Learning (Stable Diffusion, Transformer)
Title: Fashion Compatibility Prediction

Overview and Problem Statement:

The fashion domain is a very important and lucrative application of computer vision. According to a recent study by Statista, the fashion industry's worth was estimated to be $1.5 trillion in 2020 and it keeps growing, representing a huge market for garment companies, designers, and e-commerce entities. Fashion image retrieval and fashion image attribute learning have been the two main areas of study in this domain. The goal of this project is to compose or predict fashion outfits automatically, working to address challenges in compatibility and aesthetics using advanced deep learning models like Stable Diffusion and Transformers.

Specific Challenges:

1. Learn compatibility relationships among fashion items to facilitate effective fashion recommendation using a Transformer model. Transformers are highly effective in capturing contextual relationships in sequential data.
2. Utilize Stable Diffusion for generating high-quality, realistic images of fashion outfits that follow compatibility criteria.
3. Note that this project is compute intensive. This might require the use of AWS Sagemaker free/purchased account or equivalent high-performance computing resources.

Data Description:

Polyvore is a popular fashion website, where users create and upload outfit data. These fashion outfits contain rich multimodal information like images and descriptions of fashion items, number of likes of the outfit, hash tags of the outfit, etc. Researchers have utilized this information for various fashion tasks. Therefore, here, a curated part of the Polyvore dataset (made available for research purpose through the ACM MM'17 paper "Learning Fashion Compatibility with Bidirectional LSTMs" [paper] [code]) will be used. It contains 164,379 items (each item contains a pair - product image and a corresponding text description). The average number of fashion items in an outfit is 6.5. The fashion-compatibility-prediction.txt contains ~7,000 outfits, where 4,000 are incompatible and 3,000 are compatible. In each line the first number indicates the compatibility (1 is compatible, 0 is not) followed by a sequence of fashion items consisting the outfit.

Applications:

1. Outfit match recommendation
2. Accessories recommendation
3. Realistic generation of fashion outfits

4. Personalized fashion recommendations

References:

1. Polyvore dataset: https://github.com/xthan/polyvore-dataset
2. High-Resolution Image Synthesis with Latent Diffusion Models, Rombach et al., 2021: https://arxiv.org/abs/2112.10752
3. Virtual Fashion Designer: https://www.e2enetworks.com/blog/fine-tuning-stable-diffusion-to-create-a-virtual-fashion-designer-for-customers

Capstone Project 5

Domain: R & D
Techniques: Large Language Model (LLM) Fine Tuning, Computer Vision, NLP
Title: Gemini-SCICAP: Enhancing Scientific Figure Captioning with a Language Model

Overview and Problem Statement:

Scientific figures often contain crucial information, and providing accurate captions is essential for better comprehension. Existing generic captioning models may not capture the specialized terminology and context found in scientific literature. This project addresses the need for a dedicated model for scientific image captioning. This project involves fine tuning the Gemini Large Language Model (LLM) to generate accurate and contextually relevant captions for scientific figures. A model capable of understanding and describing complex scientific visuals will be created, combining the power of NLP with computer vision

Dataset

SCICAP (Scientific Captioning) is a large-scale image captioning dataset containing real-world scientific figures and captions. The dataset is constructed using over two million images from more than 290,000 papers collected and released by arXiv. It covers a wide range of scientific domains, making it a comprehensive resource for training and evaluating the model.

Specific Challenges

- Scientific Terminology: Adapting the language model to understand and generate captions with domain-specific scientific terminology.
- Complex Visuals: Handling intricate scientific visuals that may include charts, graphs, and diagrams.
- Contextual Understanding: Ensuring the model captures the context of the scientific content to provide informative and coherent captions.
- Multi-Modal Learning: Integrating both text and visual information for effective image captioning.

References

1. Gemini: A Family of Highly Capable Multimodal Models
2. Gemini: An Overview of Multimodal Use Cases
3. SCICAP Dataset: A Large-Scale Scientific Image Captioning Benchmark

Capstone Project 6

Domain: Customer Service and Conversational AI
Techniques: Natural Language Processing (NLP), Large Language Models (LLMs), Sentiment Analysis, Intent Recognition, Topic Modeling
Title: Customer Conversational Intelligence Platform Powered by an LLM Agent

Overview and Problem Statement:

This project aims to develop a state-of-the-art Customer Conversational Intelligence Platform powered by a Large Language Model (LLM) agent. The LLM's advanced language understanding will drive the analysis of customer interactions across diverse channels (chatbots, call centers, email, social media). The platform will extract actionable insights from this data, enabling businesses to optimize customer service processes and significantly enhance the overall customer experience.

Datasets:

Name: Relational Strategies in Customer Interactions (RSiCS)
- Description: This dataset contains a corpus for improving the quality and relational abilities of Intelligent Virtual Agents.
- Link: Link to the dataset

Name: 3K Conversations Dataset for ChatBot from Kaggle
- Description: The dataset includes various types of conversations such as casual or formal discussions, interviews, customer service interactions, and social media conversations.
- Link: Link to the dataset

Name: Customer Support on Twitter Dataset from Kaggle
- Description: This is a large corpus of tweets and replies that can aid in natural language understanding and conversational models.
- Link: Link to the dataset

Challenges:

1. Data Collection: Gather customer conversations from diverse sources like voice calls, chat transcripts, emails, and social media interactions.
2. Use LLM-Agent for:
   a. Sentiment Analysis – accurate detection of customer emotions (positive, negative, neutral) and granular sentiment categories (frustration, satisfaction, inquiry, etc.) throughout conversations.

b. Intent Recognition - understanding the underlying purpose behind customers' queries, enabling tailored responses and resolutions
c. Topic Modeling - discovering recurring themes and patterns within conversations, highlighting trending issues, feedback topics, and potential areas for improvement.
d. Agent Performance Evaluation - Analyzing agent interactions to provide constructive feedback, identifying training needs, and recognizing exceptional service.
e. LLM-Driven Real-time Recommendations - Empowering agents with suggestions for next-best actions or responses during active conversations, optimizing outcomes.

Methodology:

Select GPT2/GPT3, fine-tune the LLM agent extensively on a large dataset of customer conversations annotated for sentiment, intent, topics, etc. Develop ML algorithms to support the LLM agent. The primary focus will be on the LLM's ability to perform sentiment analysis, intent recognition, topic modeling, and agent performance assessment. Utilize platforms like SageMaker or equivalent to automate the ML workflow.

Example:
Customer: Hello, I ordered a laptop from your website, and it's been a week, but I haven't received it yet. Can you help me track my order?

Platform Analysis:

Categorization: Inquiry about order tracking.
Sentiment Analysis: Neutral sentiment.
Resolution Status: Unresolved.
Support Agent: Hi there! I apologize for the delay in your order. Could you please provide me with your order number? I'll check the status for you.

Customer: My order number is 123456789.

Platform Analysis:

Categorization: Providing order information.
Sentiment Analysis: Neutral sentiment.
Resolution Status: In progress.

Support Agent: Thank you for providing the order number. Let me check that for you. [Platform sends a real-time request to the order tracking system]

Platform Analysis:

Real-time Analysis: The platform receives updated order tracking information. The laptop is currently in transit and is expected to arrive in two days.
Support Agent: Good news! Your laptop is on its way and should be delivered within the next two days. Here's your tracking number: ABC123XYZ. You can use this number to monitor its progress.

Customer: Thank you for the information. I appreciate your help.

Platform Analysis:

Sentiment Analysis: Positive sentiment.
Resolution Status: Resolved.
Support Agent: You're welcome! If you have any more questions or need further assistance, feel free to ask. Have a great day!

Significance:

As companies accumulate immense volumes of customer interaction data, the ability to unlock meaningful insights and streamline customer service processes becomes a competitive advantage. The envisioned platform, with its real-time analysis capabilities, has the potential to revolutionize customer service, ultimately translating into greater customer satisfaction, increased operational efficiency, and a strengthened market position for businesses.

Reference:

1. Conversational Health Agents: A Personalized LLM-Powered Agent Framework, Mahyar Abbasian, Iman Azimi, Amir M. Rahmani, Ramesh Jain: https://arxiv.org/html/2310.02374v4
2. Building a Conversational AI Agent with Long-Term Memory Using LangChain and Milvus, Zilliz:
https://medium.com/@zilliz_learn/building-a-conversational-ai-agent-with-long-term-memory-using-langchain-and-milvus-0c4120ad7426

Capstone Project 7

Domain: Search Engine Optimization, Generative AI, LLMs
Techniques: Generative pre-trained transformers (GPTs)
Title: Automated SEO using ChatGPT

Overview and Problem Statement:

The goal of this project is to create an automated Search Engine Optimization (SEO) tool using ChatGPT, an AI-based chatbot system. The tool will use natural language processing (NLP) and machine learning (ML) algorithms to analyze website content, identify SEO issues, and provide recommendations for improvement. The tool will help website owners and SEO professionals to optimize their website's content and improve search engine rankings more efficiently and effectively.

Methodology:

1. Design a chatbot interface that can interact with website owners and SEO professionals and collect website information such as website URL, keywords, and target audience.
2. Train the ChatGPT model with SEO-specific data and vocabulary, including industry keywords, search engine algorithms, and best practices for on-page and off-page optimization. *Note: Dataset for this project is not readily available and should be collected/generated*
3. Implement an NLP algorithm to analyze website content, including meta tags, headings, images, and links, and identify SEO issues such as duplicate content, missing tags, and broken links.
4. Develop an ML algorithm that can learn from the SEO analysis data and provide personalized recommendations for optimization based on the website's content and target audience.
5. Integrate the automated SEO tool with popular SEO platforms and tools, such as Google Analytics and Google Search Console, to provide more comprehensive data and insights.

Project Deliverables:

1. An automated SEO tool that uses ChatGPT to analyze website content and provide recommendations for optimization.
2. A user-friendly chatbot interface that can interact with website owners and SEO professionals and collect website information.

Capstone Project 8

Domain: Legal Studies
Techniques: Natural Language Processing, Generative AI
Title: AI Patent Advisor: Leveraging Large Language Models for Patent Analysis and Technology Transfer Facilitation

Overview:

With the exponential growth of patent databases, extracting valuable insights and facilitating technology transfer has become increasingly challenging. This project aims to develop an AI-powered advisor that can analyze patents, provide comprehensive summaries, and recommend potential commercial applications and licensing opportunities. By fine-tuning the model on the AI-Growth-Lab patents and claims dataset, sourced from various patent repositories, the AI Patent Advisor will empower inventors, businesses, and legal professionals to navigate the complex landscape of patent law and innovation.

Dataset:

The AI-Growth-Lab patents and claims dataset (https://huggingface.co/datasets/AI-Growth-Lab/patents_claims_1.5m_traim_test) will serve as the foundation for training the AI Patent Advisor. This dataset comprises a vast collection of patents and their associated claims, covering diverse technical fields and industries. With over 1.5 million patent documents, the dataset provides a rich source of information for training and fine-tuning the LLM model. Each patent document includes detailed descriptions, claims, and metadata, enabling comprehensive analysis and understanding of patented technologies.

Methodology:

- Data Collection and Preprocessing: Gather AI-Growth-Lab patents and claims, preprocess to clean noise, standardize formatting, and tokenize text
- Fine-tuning LLM: Utilize state-of-the-art LLM architecture, fine-tune on AI-Growth-Lab dataset for patent-specific language adaptation.
- Analyze patents, extract key concepts, and generate concise summaries for efficient knowledge extraction.
- (Optional) Implement semantic matching algorithms to map patented inventions to potential commercial applications and licensing opportunities.
- Design a user-friendly interface enabling users to input patents, explore summaries, and receive technology transfer recommendations.

- (Optional) Testing and Validation: Assess AI Patent Advisor's performance, including accuracy of summaries, relevance of transfer recommendations, and overall usability.
- (Optional) Deployment and Maintenance: Deploy the AI Patent Advisor, ensuring scalability and reliability, with protocols for regular maintenance and updates to align with evolving patent language and industry trends.

Challenges:

- Handling the complexity and variability of patent language and terminology; Ensuring the accuracy and relevance of patent summaries and technology transfer recommendations

Significance:

The AI Patent Advisor project holds immense potential to transform the landscape of patent analysis and technology transfer facilitation: It will enable inventors, businesses, and legal professionals to efficiently analyze patent documents and extract valuable insights; bridge the gap between patented inventions and commercial applications by identifying potential licensing opportunities and strategic partnerships Overall it will promoter innovation by accelerating the pace of innovation by facilitating knowledge dissemination and collaboration within the patent ecosystem.

Capstone Project 9

Domain: Fashion and Textile
Techniques: Stable Diffusion
Title: SareeGen: AI-Driven Saree Design Generator Using Stable Diffusion and Prompt Engineering

Overview and Problem Statement:

The objective of this project is to develop SareeGen, an innovative AI-driven system that leverages Stable Diffusion and prompt engineering to create novel saree designs based on textual descriptions. By fine-tuning the model on the "Saree & Handloom Clothes Dataset," SareeGen will generate culturally rich and stylistically diverse designs. This project will also incorporate prompt engineering techniques to refine the input descriptions, ensuring the model produces high-quality, detailed saree designs aligned with user expectations.

Methodology:

1. **Dataset Collection and Preprocessing**: Collect and preprocess the "Saree & Handloom Clothes Dataset" from Neural Loom, which includes diverse saree designs, patterns, and styles. The dataset will consist of annotated images and descriptions to train the text-to-image model effectively.
2. **Model Architecture**: Stable Diffusion will serve as the core model for saree design generation. The model will be fine-tuned on the Saree & Handloom Clothes Dataset to ensure the generated saree designs respect cultural integrity and stylistic elements, such as traditional patterns, colors, and textures.
3. **Prompt Engineering**: Develop and implement prompt engineering techniques to enhance text-to-image generation. Users will input textual descriptions (e.g., "modern pink silk saree with floral motifs"), and prompt engineering will refine these inputs for optimal results.
   - Fine-tune prompts for better coherence, ensuring that the system captures details like fabric type, regional styles, embroidery, and pattern intricacies. Experiment with prompt structures and variations to improve the model's ability to interpret and translate user inputs into high-quality saree designs.
4. **Fine-Tuning Stable Diffusion**: Fine-tune the Stable Diffusion model using the Saree & Handloom Clothes Dataset, ensuring that it can generate intricate and culturally appropriate saree designs.
   - Focus on capturing essential design aspects such as fabric textures, traditional and contemporary patterns, color schemes, and embroidery details.

5. **Text-to-Image Translation**: Leverage Stable Diffusion to translate refined prompts into visually detailed saree designs. This process will involve translating user inputs directly into high-resolution, contextually accurate saree images.
6. **Web Interface Development**: Build an intuitive, user-friendly web interface where users can input textual descriptions of their desired saree designs.
    - Allow users to view, modify, and customize generated saree designs through interactive tools such as sliders for color selection, pattern modification, and fabric customization.

Significance:

This project will revolutionize the saree design industry by integrating Stable Diffusion and prompt engineering. Users will be able to generate custom saree designs that are culturally rich and visually appealing. By refining user inputs and guiding the model through prompt engineering, SareeGen will generate designs that accurately reflect the intricacies of Indian sarees, making it a valuable tool for fashion designers, saree enthusiasts, and the broader fashion industry.

Capstone Project 10

Title: Enhancing Customer Service with AgentLite-powered Chatbots

Overview and Problem Statement:

This project aims to develop an advanced customer service chatbot system for e-commerce platforms using AgentLite, a lightweight library for building and advancing task-oriented LLM agent systems. The project will focus on creating a more efficient, accurate, and personalized customer support experience by leveraging AgentLite's capabilities in task decomposition, multi-agent systems, and flexible reasoning strategies.

Dataset:

The project will utilize a combination of publicly available and synthetic datasets:

1. Customer Service Conversations:
    - Source: Kaggle's Customer Support on Twitter dataset
    - Content: Real customer service interactions from various companies
2. E-commerce Product Catalog:
    - Source: Amazon product dataset (a subset focused on electronics and home appliances)
    - Content: Product descriptions, specifications, prices, and customer reviews
3. Synthetic E-commerce Queries:
    - Source: Generated using GPT-3 or similar language models
    - Content: Simulated customer queries specific to e-commerce scenarios

Methodology:

1. Data Preparation: Clean, preprocess, and augment datasets for training.
2. AgentLite Implementation: Design multi-agent architecture and implement task decomposition.
3. Training and Fine-tuning: Develop and train agents with various reasoning strategies.
4. Integration and Testing: Develop web interface and conduct A/B testing with different agent configurations.
5. Evaluation: Measure performance metrics and gather user feedback on multi-agent interactions.

Challenges:

1. Task Decomposition: Break complex queries into manageable subtasks.
2. Multi-Agent Coordination: Ensure smooth agent collaboration in customer service.

3. Reasoning Strategy Selection: Implement and evaluate strategies for various scenarios.
4. Product Knowledge Integration: Efficiently manage e-commerce catalog in knowledge base.
5. Scalability: Handle high-volume concurrent users without latency.
6. User Experience: Design intuitive interface for multi-agent system interactions.

References:

1. Liu, Z., Yao, W., Zhang, J., Yang, L., Liu, Z., Tan, J., ... & Savarese, S. (2023). AgentLite: A Lightweight Library for Building and Advancing Task-Oriented LLM Agent System. arXiv preprint arXiv:2308.12515.
2. Henderson, Matthew, et al. 2019. "A repository of conversational datasets." arXiv preprint arXiv:1904.06472.
3. Serban, Iulian Vlad, et al. 2017. "A deep reinforcement learning chatbot." arXiv preprint arXiv:1709.02349.

Capstone Project 11

Title: A Versatile LLM-Based Code Generation Platform for Varied Complexity Tasks

Overview and Problem Statement:

This project aims to develop an advanced code generation system powered by Large Language Models (LLMs) that can interpret various input types and generate appropriate code solutions for tasks of varying complexity. The system will be designed to understand the nuances of different programming tasks, break them down into manageable components, and provide a comprehensive solution that includes flow charts, task lists, test cases, and functional code.

Methodology:

1. Data Preparation: Create multi-modal dataset (text, audio, images)
2. LLM Setup: Select and fine-tune LLM for code generation
3. Input Processing: Develop multi-modal input handling (text, audio, image)
4. Complexity Analysis: Design task complexity assessment algorithm
5. Clarification System: Implement LLM-driven question generation
6. Task Visualization: Create flow chart and task list generator
7. Code Generation: Develop core functionality for multiple languages
8. Testing Module: Automate test case generation and execution
9. Refinement Process: Design feedback loop for code improvement
10. User Interface: Create adaptable, user-friendly interface
11. Evaluation: Conduct extensive testing and user studies

Challenges:

1. Input Processing: Handle diverse input types accurately
2. Complexity Analysis: Assess task complexity reliably
3. Context Comprehension: Ensure LLM understands complex programming tasks
4. Code Quality: Generate high-quality, industry-standard code
5. Scalability: Design for wide range of tasks and languages
6. Ethics: Address code ownership and responsible AI use
7. User Experience: Create intuitive interface for varied expertise levels

References:

1. Chen, M., Tworek, J., Jun, H., Yuan, Q., Pinto, H. P. D. O., Kaplan, J., ... & Zaremba, W. (2021). Evaluating large language models trained on code. arXiv preprint arXiv:2107.03374.

2. Li, Y., Choi, D., Chung, J., Kushman, N., Schrittwieser, J., Leblond, R., ... & Sifre, L. (2022). Competition-level code generation with AlphaCode. Science, 378(6624), 1092-1097.
3. Jiang, N., Shen, Y., Chen, Z., Yang, C., Zhang, L., & Cao, Y. (2023). InstructCoder: A generative model for code infilling and rewriting based on fine-grained code instructions. arXiv preprint arXiv:2305.11766.

Capstone Project 12

Title: Interactive Root Cause Analysis in IT Operations Using Generative AI: A Case Study with OpenStack Infrastructure Logs

Overview and Problem Statement:

This project aims to develop an interactive root cause analysis (RCA) system using Generative AI to understand and explain incidents in IT departments. RCA is a critical task for identifying the underlying causes of incidents, system failures, and performance bottlenecks. Traditional RCA is a time-consuming process often dependent on expert input.

Generative AI can automate and enhance this process by analyzing system logs and performance data, detecting patterns, and explaining potential root causes interactively. A key aspect of this project is to make the AI model amenable to queries, allowing IT teams to ask the AI for detailed insights about incidents, such as:

- "What caused the server outage?"
- "Explain the root cause of the storage failure."
- "What factors contributed to this anomaly?"

The project will explore how transformer-based models can be used for interactive troubleshooting in a cloud infrastructure environment, focusing on the OpenStack ecosystem.

Dataset:

1. **OpenStack Infrastructure Logs (LogHub):**
    - **Description:** The dataset includes logs from OpenStack cloud infrastructure, covering key services such as Nova (compute), Neutron (networking), and Cinder (storage). These logs capture events, system activities, and error messages, providing a rich dataset for RCA in cloud environments.This dataset can be used to generate and explain potential root causes for service failures, anomalies, and performance issues in OpenStack infrastructure.
    - **Source:** OpenStack Logs on LogHub

Methodology:

Data Preprocessing:

- Clean and normalize OpenStack logs.
- Use NLP to structure log data.
- Apply anomaly detection for incidents.

Model Selection:

- Use transformers for interactive root cause generation.
- Implement LSTM/autoencoders/logBERT for anomaly detection.
- Build an interactive UI for querying the AI.

Training & Validation:

- Train on OpenStack logs for accurate root cause insights.
- Validate with real and synthetic incidents.
- Test query responses for clarity.

Automation & Interaction:

1. Automate real-time log analysis.
2. Enable AI queries for root cause insights.
3. Integrate with ITSM platforms for incident management.

Challenges:

Data Complexity and Noise:

- OpenStack logs are complex and can be noisy, with incomplete data or inconsistencies across different services (e.g., compute, networking). Effective log parsing and filtering are crucial for ensuring meaningful RCA.

Interactive Explainability:

- One of the main challenges is ensuring that the AI can explain root causes in a human-understandable way. The model needs to provide not only predictions but also justifications, ensuring the results are interpretable and actionable.

References:

1. LogHub - OpenStack Dataset:https://github.com/logpai/loghub.
2. LogBERT: Log Anomaly Detection via BERT - https://arxiv.org/abs/2103.04475

Capstone Project 13

Title: AI-Powered Plant Disease Detection and Farmer Assistance Using Generative AI

Overview and Problem Statement:

This project focuses on helping farmers identify plant diseases through image recognition and providing actionable advice. Farmers will upload images of their crops, and the system will use the MobileNetV2 model to detect diseases and Generative AI to generate clear, helpful explanations.

The primary objective is to build a system that provides accurate disease detection and recommendations. An optional challenge is to enhance the system with multilingual support, offering explanations in regional languages (e.g., Hindi, Tamil, Telugu), but this is not mandatory.

Dataset:

1. MobileNetV2 Plant Disease Dataset (Hugging Face):
    - Pretrained MobileNetV2 model trained on a large collection of plant disease images.
2. Agricultural Dataset:
    - Agricultural terms, disease descriptions, and treatment recommendations for training the Generative AI model to provide disease explanations and treatment advice.

Methodology:

1. Data Preprocessing:
    - Preprocess plant images for uniformity (resizing, noise removal).
    - Prepare agricultural phrases and disease explanations for the model.
2. Model Selection:
    - MobileNetV2 for Plant Disease Detection:
        - Use the pretrained MobileNetV2 model for identifying diseases from plant images.
    - Generative AI (GPT-3.5/4) for Disease Explanations:
        - Use GPT-3.5/4 to generate disease explanations and treatment suggestions, tailored to the detected disease.
    - Optional Multilingual Support (Challenge):
        - Participants may choose to integrate multilingual capabilities, translating disease explanations into regional languages, but this remains optional.

3. Training & Validation:
    ○ Train the models to deliver accurate disease identification and explanations.
    ○ Validate the system using real-world plant disease images and ensure clarity of explanations.
4. User Interaction & Query Support:
    ○ Build a user-friendly interface where farmers can upload images and receive instant disease diagnoses and treatment suggestions.
    ○ Allow users to ask additional queries, like "How to treat this disease?" and get detailed responses.

Challenges:

1. Image Quality:
    ○ Ensure the system handles varying image quality effectively (e.g., blurred or low-light images).
2. Optional Multilingual Support:
    ○ Accurately translating agricultural terms and maintaining the clarity of explanations across different languages.
3. Digital Literacy:
    ○ Create an intuitive and simple user interface that accommodates varying levels of digital literacy among farmers.

References:

1. MobileNetV2 for Plant Disease Detection (Hugging Face): Available at MobileNetV2 Plant Disease Model.
2. Optional Translation Models (IndicTrans, mBART): Indic NLP Project. Available at IndicTrans.