Capstone Project 1

Project Title:

Hospital Emergency Prediction using ETS Data

Project Overview:

The objective of this project is to develop and implement a Machine Learning Operations (MLOps) framework for predicting emergent medical situations at the point of patient admission to a hospital. This predictive model will leverage historical patient data, particularly the Emergency Triage Score (ETS) dataset, to probabilistically anticipate instances where a significant influx of patients is likely to occur. This project seeks to harness data-driven insights to enhance the anticipatory and preparatory capacities of healthcare facilities in responding to emergent medical scenarios, thereby improving overall healthcare service efficiency and patient care outcomes.The goal of this project is to develop a Machine Learning Operations (MLOps) workflow to predict hospital emergencies at the time of patient admission. This predictive model will help healthcare providers allocate resources efficiently and prepare for potential surges in patient volume.

Dataset:

The project will utilize the Hospital Triage and Patient History Data from Kaggle: https://www.kaggle.com/datasets/maalona/hospital-triage-and-patient-history-data, specifically the ETS (Emergency Triage Score) dataset. This dataset contains valuable information about patient demographics, medical history, and triage scores, which can be used to predict the likelihood of an emergency upon admission.

Methodology:

- Data Preparation: Prepare and clean the data for analysis.
- Model Development: Create a predictive model using the cleaned data.
- MLOps Workflow: Set up a workflow for managing data and models on a Cloud platform such as Amazon Web Services (AWS).
- Model Monitoring and Updating: Continuously check and improve the model's performance.
- Documentation and Reporting: Keep records and share findings with others.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Challenges:

- Model Generalization: Creating a model that works in different hospitals.
- Data Privacy: Handling patient data while adhering to privacy laws.
- Model Interpretability: Making the model's decisions understandable.
- Real-time Data: Integrating real-time data for timely predictions.
- Model Monitoring: Continuously checking and adapting the model.
- Resource Constraints: Working within budget limitations.

Significance:

This project aims to leverage machine learning and MLOps principles to enhance emergency prediction in hospital admissions. By harnessing Cloud services for data management and deployment, the project seeks to provide healthcare providers with a valuable tool for optimizing resource allocation and improving patient care in emergency situations.

Reference:

1. Vântu A, Vasilescu A, Băicoianu A. Medical emergency department triage data processing using a machine-learning solution. Heliyon. 2023 Jul 22;9(8):e18402. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10412878/

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Capstone Project 2

Project Title:

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

Project Overview:

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

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

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.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Capstone Project 3

Project Title:

Customer Conversational Intelligence Platform Powered by an LLM Agent

Project Overview

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. Gather customer conversations from diverse sources like voice calls, chat transcripts, emails, and social media interactions.
2. Use LLM-Agent for:

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

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.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

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.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Capstone Project 4

Project Title:

Driver Demand Prediction for Optimal Food Delivery Charges

Project Overview:

The food delivery industry relies heavily on the efficiency of its delivery processes, where timely and accurate deliveries significantly impact customer satisfaction and overall experience. This project focuses on developing an effective predictive model that accurately estimates the time it takes for orders to reach customers. Delivery time is influenced by various factors, including the delivery person's age, ratings, geographic coordinates, and time-related variables. The successful implementation of this food delivery time prediction model will enable businesses to optimize their operational processes, leading to more accurate delivery time estimates and an improved overall delivery experience for customers. The aim is also to predict the demand for delivery drivers in specific regions and times, by analyzing order requests, driver activity, and related parameters, thereby optimizing delivery charges, ensuring consistency and minimizing customer drop-offs.

Dataset:

The Kaggle DeliverOnTime: Food Delivery Duration Predictor consists of the features such as
'ID', 'Delivery_person_ID', 'Delivery_person_Age', 'Delivery_person_Ratings',
'Restaurant_latitude', 'Restaurant_longitude',
'Delivery_location_latitude','Delivery_location_longitude', 'Order_Date',
'Time_Orderd','Time_Order_picked', 'Weatherconditions',
'Road_traffic_density','Vehicle_condition', 'Type_of_order',
'Type_of_vehicle','multiple_deliveries', 'Festival', 'City', 'Time_taken(min)'

Methodology:

- Data Workflow: Examine, preprocess, and engineer features from order and driver activity data.
- Model Development: Construct and train machine learning models to predict driver demand, emphasizing accuracy and minimization of error.
- MLOps Implementation: Employ the model using MLOps tools for real-time driver demand prediction, ensuring scalability and adaptability.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

- Monitoring & Maintenance: Regularly oversee model performance, retrain with fresh data, and establish alerts for significant demand spikes or unforeseen anomalies.
- Feedback: Procure feedback from users and the system to refine the model's predictions, making them more precise and relevant.
- Security & Compliance: Comply with data protection standards, fortifying security measures to protect the model and its data.

Challenges:

- Handling potential data inconsistencies, such as missing data or outliers.
- Predicting sudden surges in order demand, which may not always follow historical patterns.
- Ensuring real-time driver demand prediction without compromising the accuracy or speed of the system.

Significance:

This project will benefit the food delivery sector by providing consistent delivery charges, leading to enhanced customer satisfaction and loyalty. By ensuring the right number of drivers are allocated based on demand, it will also improve operational efficiency and potentially reduce costs for delivery services.

Additional Reading:
Urban distribution of short-term food delivery demand

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Capstone Project 5

Project 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.

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. Personalized recommendations for investment portfolios, retirement plans, insurance policies, and other financial products based on individual goals.

References:

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

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

Capstone Project 6

Project Title:

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

Project Overview:

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

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Methodology:

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

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

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

4. Automation & Interaction:
    - Automate real-time log analysis.
    - Enable AI queries for root cause insights.
    - Integrate with ITSM platforms for incident management.

Challenges:

1. 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.
2. 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.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Capstone Project 7

Project Title:

Automated Answer Validation for Science Question Answering

Overview and Problem Statement:

In the domain of scientific question answering, validating answers and providing accurate feedback is critical for effective learning. The goal of this capstone project is to develop an automated answer validation system using techniques like Siamese text similarity model, etc. The system will compare student responses with the correct answer and distractors to determine the level of correctness and provide appropriate feedback. The automated answer validation system for science question answering will benefit educators and students in science-related subjects. It will streamline the assessment process, reduce manual effort, and ensure consistent evaluations, leading to improved learning outcomes in the science domain.

Challenges:

1. Implementing a robust Siamese text similarity model for answer validation in the science domain.
2. Preprocessing and structuring the dataset from the provided SCIQ dataset to create meaningful pairs of student responses and correct answers.
3. Handling varying lengths of text data during model training and inference.
4. Addressing potential semantic ambiguities and domain-specific challenges in science questions and responses.

Data Description:

The SciQ dataset contains 13,679 crowdsourced science exam questions about Physics, Chemistry and Biology, among others. The questions are in multiple-choice format with 4 answer options each. For the majority of the questions, an additional paragraph with supporting evidence for the correct answer is provided. (string).

- answer1: One of the candidate answers for the question (string).
- answer2: One of the candidate answers for the question (string).

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

- answer3: One of the candidate answers for the question (string).
- answer4: One of the candidate answers for the question (string).
- correct_answer: The correct answer for the question (string).
- support: The supporting text for the question (string).

Methodology:

- Combine the question and supporting text to create meaningful context for answer validation in the science domain. Structure the data to form pairs of student responses and correct answers, along with corresponding labels (0 for incorrect and 1 for correct).
- Implement a Siamese neural network architecture using NLP libraries and deep learning frameworks. Train the model on the pairs of student responses and correct answers to learn the underlying text representations
- MLOps Implementation: Develop an interface for educators to input student responses and query the similarity model for validation in the science domain.
- Monitoring & Maintenance: Fine-tune the model and update the answer validation system based on the evaluation results.

References:

1. SCIQ (Science Question Answering) dataset from Kaggle: https://www.kaggle.com/datasets/thedevastator/sciq-a-dataset-for-science-question-answering
2. Siamese network for image and text similarity: https://medium.com/@prabhnoor0212/siamese-network-keras-31a3a8f37d04

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Capstone Project 8

Project 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.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

- 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.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Capstone Project 9

Project Title:

Image-based plant disease identification

Overview and Problem Statement:

Crop losses due to diseases are a major threat to food security every year, across countries. Conventionally, plant diseases were detected through a visual examination of the affected plants by plant pathology experts. This was often possible only after major damage had already occurred, so treatments were of limited or no use. Recently, access to smartphone based image capturing has highly increased amongst farmers and agriculturists. This has led to the successful adoption of plant disease diagnostic applications based on deep learning techniques. This is of immense value in the field of agriculture and an excellent tool for faster identification and treatment of crop diseases. It holds key importance in preventing crop based food and economic losses**.** The goal of this project is to build a convolutional neural network or to use transfer learning and develop a plant disease identification tool.

Challenges:

1. Extending the applicability of the plant disease identification tool for farmers through the use of Indian regional languages [Optional challenge]

Dataset:

The PlantVillage dataset consists of 54303 healthy and unhealthy leaf images divided into 38 categories by species and disease. Dataset link:
https://data.mendeley.com/datasets/tywbtsjrjv/1

Methodology:

- Data Workflow: Examine, preprocess, and associate labels for images from PlantVillage dataset.
- Model Development: Construct and train machine learning models to identify plant disease, emphasizing accuracy and minimization of error.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

- MLOps Implementation: Employ the model using MLOps tools for real-time plant disease identification, ensuring scalability and adaptability.
- Monitoring & Maintenance: Regularly oversee model performance, retrain with fresh data, and establish alerts for significant disease spread.
- Feedback: Procure feedback from users and the system to refine the model's predictions, making them more precise and relevant.

References:

1. Dataset link: https://data.mendeley.com/datasets/tywbtsjrjv/1
2. Dataset source details: https://www.tensorflow.org/datasets/catalog/plant_village
3. Liu and Wang (2021): https://plantmethods.biomedcentral.com/articles/10.1186/s13007-021-00722-9
4. Mohanty et al., (2016): https://www.frontiersin.org/articles/10.3389/fpls.2016.01419/full

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Capstone Project 10

Project Title:

Food Image Segmentation

Overview and Problem Statement:

Worldwide, obesity has nearly tripled since 1975. In 2016, more than 1.9 billion adults, 18 years and older, were overweight (WHO sources). In such a situation, documenting dietary caloric intake is crucial to manage weight loss. Food image segmentation is a critical and indispensable task for developing health-related applications such as automated estimation of food calories and nutrients as a means for dietary monitoring. One of the challenges in this area is the improvement of accuracy in dietary assessment by food image analysis. However, how to derive the food information (e.g., food type and portion size) from food images effectively is a challenging task and an open research problem. In this project, participants are expected to make a model that can segment the food components present in an input food image and build an application that can predict the food class and the food portions from it.

Data Description:

Food recognition dataset (Ciocca et. al., 2017) is one of the few publicly available, pixel segmented datasets on food. It contains 1,027 images of food on trays, with 73 classes of food and 3,616 labeled instances of food. The tray images have been manually segmented using carefully drawn polygonal boundaries.
Dataset link: http://www.ivl.disco.unimib.it/activities/food-recognition/

Methodology:

- Data Workflow: Explore, preprocess, and associate labels for Food recognition dataset.
- Model Development: Construct and train machine learning models to segment and predict food class, emphasizing accuracy and minimization of error.
- MLOps Implementation: Develop an interface for users to input food tray images and get the food class and the food portions from it.
- Monitoring & Maintenance: Regularly oversee model performance, retrain with fresh data, and establish alerts for significant high calorie food portions.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

- Feedback: Procure feedback from users and the system to refine the model's predictions, making them more precise and relevant.

Challenges:

1. The complex appearance of food makes it difficult to localize and recognize ingredients in food images, e.g. the components may overlap with one another in the same image, and the identical ingredient may appear distinctly in different food images.

Applications:

It has many applications such as calorie estimation, diet management, industrial compliance and customer satisfaction.

References:

1. Ciocca et al., 2017: Food Recognition: A New Dataset, Experiments, and Results
2. Yang, 2020: Food Image Segmentation with fast.ai
3. Liu et al., 2016: DeepFood: Deep Learning-based Food Image Recognition for Computer-aided Dietary Assessment

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Capstone Project 11

Project Title:

Fine-Tuning the IndicTrans NMT Model for Healthcare Conversations

Overview and Problem Statement:

This project aims to enhance the performance of the IndicTrans Neural Machine Translation (NMT) model specifically for healthcare conversations between patients and hospital staff (doctors/nurses). The current model has very poor accuracy on medical terminology translations from Indian languages to English. The objective is to fine-tune the model using a dataset of healthcare conversations to improve its accuracy and fluency in translating medical terminology and patient queries from Indian languages to English. By fine-tuning the model on domain-specific data, we aim to optimize its performance for real-world healthcare communication scenarios.

Dataset:

Custom data is required for this project, wherein the patient's statement in a non-English Indian language (eg. Telugu) and its corresponding English translation are generated. These conversations may cover various medical scenarios, including patient inquiries, symptom descriptions, treatment discussions, and medical advice exchanges. [See a related (paid) dataset comprised of a healthcare discussion in Telugu here: https://www.futurebeeai.com/dataset/speech-dataset/healthcare-call-center-conversation-telugu-india#contact-form]. Other possibilities are to create synthetic medical conversation dataset (Telugu (or another Indian language) and English translation) using an LLM. For an example, see Fig. 1.

Methodology:

The IndicTrans NMT model, a pre-trained machine translation model tailored for Indian languages, will be selected as the base model for fine-tuning.

- Preprocess the healthcare conversation dataset and format it for fine-tuning. This involves cleaning the data, tokenizing the text, and organizing it into a suitable format for training.
- Fine-Tune the pre-trained IndicTrans NMT model with the healthcare conversation dataset.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

- Evaluate the fine-tuned model's performance in translating healthcare conversations accurately and fluently. Evaluation metrics such as BLEU score, accuracy, and fluency will be used to measure the model's effectiveness.
- Once the fine-tuning process is complete and the model demonstrates satisfactory performance on the validation set, deploy it for use as a Whatsapp based app in healthcare communication scenarios.

By fine-tuning the IndicTrans NMT model on healthcare conversation data, this project will address the specific challenges and nuances of medical communication, thereby improving the accessibility and quality of healthcare services for Indian language speakers

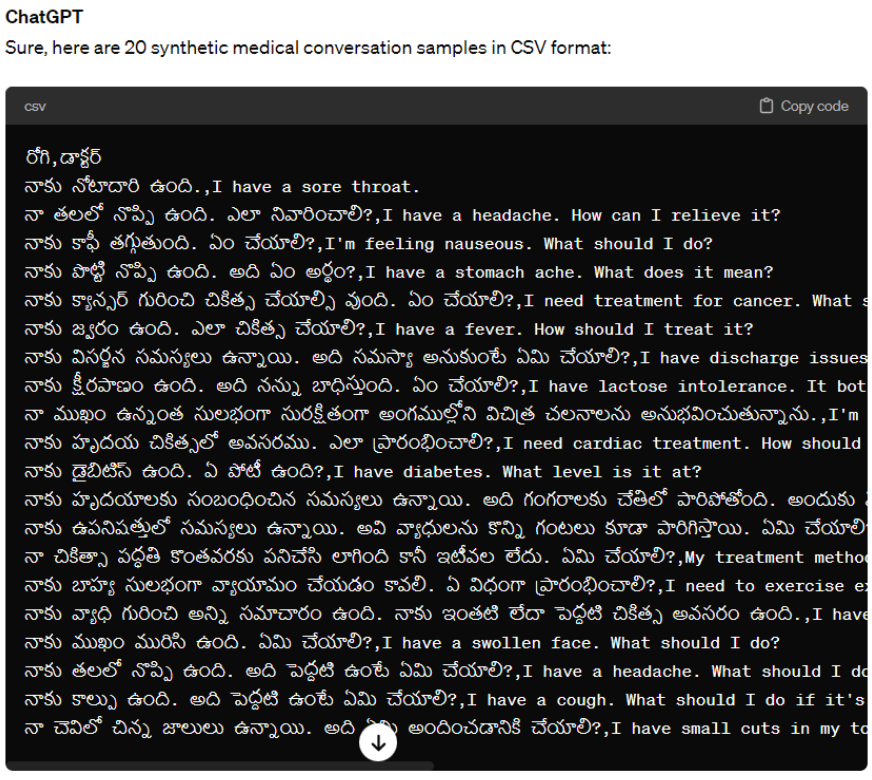

Fig. 1 Example of synthetic data creation using ChatGPT (Telugu->English medical conversation samples)

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Capstone Project 12

Project Title:

Automatic Image Captioning

Overview and Problem Statement:

Captioning the images with proper description is a popular research area of Artificial Intelligence. A good description of an image is often said as “Visualizing a picture in the mind”. The generation of descriptions from the image is a challenging task that can help and have a great impact in various applications such as usage in virtual assistants, image indexing, a recommendation in editing applications, helping visually impaired persons, and several other natural language processing applications. In this project, we need to create a multimodal neural network that involves the concept of Computer Vision and Natural Language Process in recognizing the context of images and describing them in natural languages (English, etc). Deploy the model and evaluate the model on 10 different real-time images. The goal of this project is to build an image captioning model to generate captions of an image.

Challenges:

1. Dataset images are of different shapes and sizes
2. Input images can be of different formats JPEG, PNG, etc

Data Description:

The Flickr8k dataset is a collection of sentence-based image descriptions. Dataset consists of 8k images in JPEG format with different shapes and sizes. Images are paired with five different captions which provide clear descriptions of the salient entities and events. The images were chosen from six different Flickr groups and included a variety of scenes and situations. Dataset link: https://github.com/goodwillyoga/Flickr8k_dataset

Methodology:

- Data Workflow: Explore, preprocess, and associate labels for the dataset.
- Model Development: Construct and train machine learning models for image captioning.
- MLOps Implementation: Develop an interface for users to input images for captioning.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

- Monitoring & Maintenance: Regularly oversee model performance, retrain with fresh data.
- Feedback: Procure feedback from users and the system to refine the model's predictions, making them more precise and relevant.

Applications:

1. Usage in virtual assistants
2. Image indexing
3. Recommendation in editing applications
4. Helping visually impaired persons

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Capstone Project 13

Project 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.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

- 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.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Capstone Project 14

Project Title:
SareeGen: AI-Driven Saree Design Generator

Project Overview:

The objective of this project is to develop an innovative AI-driven system, SareeGen, which harnesses the power of generative models to create novel saree designs based on textual descriptions. It integrates generative models to enable users to generate custom saree designs and provides design recommendations. Fine-tuning will be performed using the "Saree & Handloom Clothes Dataset" to capture diverse patterns and styles inherent to sarees.

Methodology:

- Collect and preprocess the "Saree & Handloom Clothes Dataset" at https://neural-loom.github.io/ to create a suitable dataset for saree design generation.
- Implement a VAE-GAN architecture for saree design synthesis, enabling the generation of diverse and creative designs
- Implement a ViT-based model for text-to-image translation, allowing users to input textual descriptions of desired saree designs.
- Use diffusion models to further enhance the generation of novel saree designs
- Fine-tune the generative models on the Saree & Handloom Clothes Dataset to ensure culturally and stylistically appropriate designs.
- Develop a user-friendly web interface for users to input text descriptions, view generated saree designs, and receive recommendations.
- Create an intuitive web interface where users can input text descriptions of desired saree designs. Display generated designs and provide options for users to modify and customize them.

Significance:

This project aims to revolutionize the saree design industry by leveraging advanced generative models, including VAEs, GANs, and ViT, to empower users to explore and create custom saree designs. The integration of the Saree & Handloom Clothes Dataset ensures that the generated designs respect Indian cultural richness and diversity, making SareeGen a valuable tool for saree enthusiasts, designers, and fashion enthusiasts alike.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Capstone Project 15

Project Title:

Automated Car Exterior Damage Assessment Using Object Segmentation

Overview and Problem Statement:

Automated detection and assessment of car exterior damages are essential for car dealers and insurance providers to streamline the damage assessment process. This capstone project aims to develop a comprehensive MLOps pipeline for the automated detection of car damage and estimation of damage severity using the Mask R-CNN computer vision technique. By deploying this model in a production-ready environment, we eliminate manual assessments and improve efficiency.

Methodology:

- Collect damaged and undamaged car images from various sources, and incorporate the COCO car damage dataset for diversity.
- Use the Mask R-CNN model to accurately detect and segment damaged areas on cars with high precision.
- Develop an MLOps pipeline that can scale to accommodate a large volume of image data and efficiently serve predictions.
- Utilize cloud based computational resources optimally managing costs under given limit
- Create a production-ready deployment pipeline using cloud-based services, allowing for easy scaling and robust handling of inference requests.
- Set up a CI/CD pipeline for automated model updates and deployments. Implement monitoring and logging solutions to track model performance and identify issues in real-time.

Significance:

The outcome of this capstone project will be a fully operational and scalable MLOps pipeline for automated car exterior damage assessment, enabling efficient and accurate damage detection and severity estimation, which can have significant practical applications for the automotive and insurance industries.

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE

Capstone Project 16

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

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

COPYRIGHT © TALENTSPRINT, 2025. ALL RIGHTS RESERVED.
CONFIDENTIAL. DO NOT SHARE