AIML- PGCP : Capstone Project

The team-based Capstone Project completion is the objective of the PG Certification in AIML program. It will provide an opportunity for the participants to implement an end-to-end AIML project. The expectation from this Capstone Project is that the participants should be able to apply the learnings from the Advanced Certification program and demonstrate a full-fledged deployed AIML solution for the selected Problem Statement/Projects as given below.

PROJECT 1

Title: Multimodal Media Retrieval and Captioning System

Objective: Build a model capable of retrieving images and text based on cross-modal queries and generating descriptive captions for images lacking textual information.

Dataset Link: COCO Dataset, Conceptual Captions Dataset

Dataset Description:
A collection of diverse images with corresponding textual captions. The COCO dataset consists of 330k images, each annotated with five captions, capturing various objects and scenes. The Conceptual Captions dataset has millions of images with high-level, real-world context descriptions, ideal for training robust caption generation and retrieval tasks.

Project Overview:
Retrieving relevant visual and textual information based on cross-modal queries is a crucial challenge in AI, enhancing applications like content recommendation, digital archiving, and accessibility tools. This project involves creating a unified multimodal neural network that enables image-to-text and text-to-image retrieval, alongside generating accurate captions for images. The project will leverage both Computer Vision and Natural Language Processing techniques, embedding visual and textual data in a shared space for seamless multimodal interaction. The system's practical use cases include enhancing digital media archives, improving accessibility for visually impaired users, and providing content recommendations. The model will be evaluated on retrieval and captioning tasks, and the deployment will support real-time content curation applications.

Tools: Natural Language Toolkit, TensorFlow, PyTorch, Keras

Deployments: FastAPI, Cloud Application Platform | Heroku, Streamlit, Cloud Computing, Hosting Services, and APIs | Google Cloud

Final Submissions:

- GitHub Repository of the project
- Project Technical Report
- Project Presentation with desired outcomes
- Summary of 3 research papers