DEPARTMENT OF COMPUTING
SCIENCE & ENGINEERING
LAB -1 FILE

Course Name : Agentic_AI

Course Code: CSCR3215

Joy Kumar Paul

2023816450

SEMESTER: **6th**

Section: **CSH**

Group: **G1**

FACULTY-IN CHARGE

Ayush Kumar Singh

ASSISTANT PROFESSOR

CSE DEPARTMENT

SHARDA UNIVERSITY

GREATER NOIDA

Ducumentation

Title: Fine- Tuning BLIP for Image Captioning

This code fine-tunes a pretrained BLIP (Bootstrapping Language-Image Pretraining) model for image captioning using a football image dataset from Hugging Face. First, it installs and imports required libraries, then loads the dataset containing image–caption pairs. A BLIP processor is used to convert images into pixel tensors and captions into token IDs, which are fed into the BLIP conditional generation model. During training, images and their captions are passed through the model, which automatically computes the loss between predicted and actual captions; this loss is minimized using the AdamW optimizer through backpropagation over several epochs. After training, the model is switched to evaluation mode and used to generate captions for sample images using the `generate()` method, and the predicted captions are decoded back into text and displayed with the corresponding images, demonstrating that the fine-tuned model can produce meaningful descriptions for football-related images.