AIML Recruitment Task Instructions

This task is common for all three batches, but the expectations differ based on your year. You will be working on the IMDb Movie Review Dataset (sentiment classification: positive/negative). The dataset has two columns: **review** (text) and **sentiment** (label).

General Instructions:
- All participants must submit their solutions along with a short README explaining their approach.
- Code/scripts should be uploaded to a **GitHub Repository** or an **open access Google Drive folder**.
- The link should be shared through the Google Form (to be provided).

First Years (2025 Batch)

- Train a simple Logistic Regression OR Naive Bayes model.
- Show the accuracy of your model.
- Write a short README explaining your preprocessing steps and model used.
- Resources:
- - Logistic Regression: https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
- - Naive Bayes: https://scikit-learn.org/stable/modules/naive_bayes.html
- - Text Vectorization (CountVectorizer): https://scikit-learn.org/stable/modules/feature_extraction.html#text-feature-extraction

Second Years (2024 Batch)

- Try at least TWO models (e.g., Logistic Regression, SVM, Random Forest).
- Compare results using multiple metrics (Accuracy, Precision, Recall, F1-score).
- Include visualization plots (Confusion Matrix, training/test performance plots).
- Write a README explaining your approach, metrics, and results.
- Resources:
- - Model Evaluation: https://scikit-learn.org/stable/modules/model_evaluation.html
- - Support Vector Machines: https://scikit-learn.org/stable/modules/svm.html
- - Confusion Matrix Visualization: https://scikit-learn.org/stable/auto_examples/model_selection/plot_confusion_matrix.html

Third Years (2023 Batch)

- Use pretrained models (e.g., BERT, DistilBERT, or other transformer-based models).
- Provide a user input interface (CLI / Streamlit / Flask) to type a review and get the sentiment prediction.
- Compare results with simpler models if possible.
- Write a README explaining your approach and additional features.
- Resources:
- - HuggingFace Transformers: https://huggingface.co/transformers/
- - DistilBERT: https://huggingface.co/docs/transformers/model_doc/distilbert
- - Streamlit for UI: https://docs.streamlit.io/
- - Flask for UI: https://flask.palletsprojects.com/