# Project Title

ASLINAKLI: Unmasking Deepfakes One Click at a Time

# Abstract:

This project addresses the growing threat of deepfake-based misinformation by introducing ASLINAKLI, a robust multimodal fake news detection system. It combines advanced NLP and computer vision using transformer-based text encoders and deep CNNs, fused via a gated mechanism. The project emphasizes explainability, robustness, and accessibility, offering a practical web interface for real-time detection.

# Project Objectives:

* Detect deepfake-based fake news through a multimodal learning approach.
* Combine textual and visual inputs using transformer and CNN-based encoders.
* Introduce gated fusion and modality dropout for adaptability and robustness.
* Deploy an intuitive web interface for real-time deepfake detection.

# Methodology:

## Dataset Acquisition and Pre-processing:

Dataset: Fakeddit, containing 20,000 text-image labeled posts across six classes. Text preprocessed using MPNet and DistilRoBERTa; images resized, normalized, and augmented.

## Model Selection and Training:

Text: MPNet-base-v2 and DistilRoBERTa. Image: ResNet-152 and ViT. Gated fusion combines embeddings. Training via PyTorch Lightning using Adam optimizer and CrossEntropyLoss.

## Testing and Evaluation:

ASLINAKLI (MPNet + ResNet152) achieved 97.6% accuracy on Fakeddit.

## Step 1: Dataset Collection and Pre-processing

Applied embedding, resizing, augmentation. Created paired inputs: (text, image, label).

## Step 2: Model Selection and Training

Integrated transformer and CNN models. Gated fusion aligns and weights modalities. Trained using PyTorch Lightning.

## Step 3: Inference and Future Deployment

Deployed as a web tool with drag-and-drop and URL input. Plans to scale via real-world deepfake response integration.

# Key Findings:

* Gated fusion significantly outperforms traditional fusion techniques.
* Modality dropout improves robustness in noisy/missing data.

# References:

Selected from Zhang et al. (2022), Kumar & West (2022), Chen et al. (2020), Cheng et al. (2022), Karduni et al. (2019)

# Team Members:

* Satvik Karan
* Sargam Tyagi