



# ASLINAKLI: Unmasking Deepfakes One Click At a Time

**Authors:** Satvik Karan, Sargam Tyagi

**Affiliation:** Department of Computer Science and Engineering, Bennett University

An advanced approach to detect deepfakes by integrating text and image analysis.

# Motivation & Problem Statement

## Threat to Society

Fake news and deepfakes undermine democracy

## Limitations

Traditional methods focus on single modality only

## Goal

Develop multimodal system combining text and image







# Dataset: Fakeddit



## Size

20,000 labeled posts  
combining text + images



## Modalities

Both textual and visual  
content per post



## Labels

6 categories for multi-class classification

# Text Preprocessing

## Models

- MPNet-base-v2
- DistilRoBERTa-v1

## Embedding

- 768-dimensional vectors
- Mean pooling or [CLS] token

## Objective

Capture deep semantic text relationships

# Image Preprocessing

## Primary Model

ResNet-152 pretrained on ImageNet

## Techniques

- Resizing & Normalization
- Random cropping
- Horizontal flipping

## Alternative

Vision Transformer (ViT) captures global context

# Model Selection Rationale

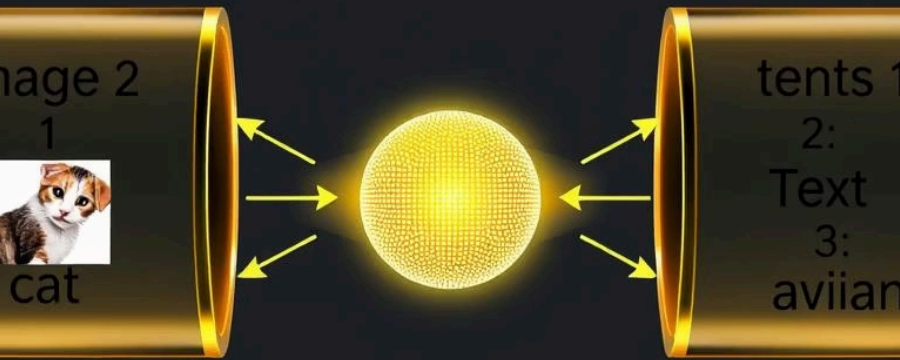
## Text Encoders

- MPNet: rich semantics, high accuracy
- DistilRoBERTa: lightweight, efficient

## Image Encoders

- ResNet-152: detailed CNN feature extraction
- ViT: attention-based, models long-range

# Gated Fusion



## Core Innovation – Gated Fusion Mechanism



### Challenge

Simple concatenation ignores input quality variability



### Approach

Gated fusion learns importance weights per modality dynamically



### Mechanism

Sigmoid gate combines embeddings into unified vector

Tor surge tun·line yountlobosf des b.3hi gliup,  
spean au watical dy. dupedsint fay tyour orrpap.



# Full Model Architecture

## Encoders

- Text: MPNet/DistilRoBERTa → 768D
- Image: ResNet-152/ViT → 300-512D

## Fusion & Projection

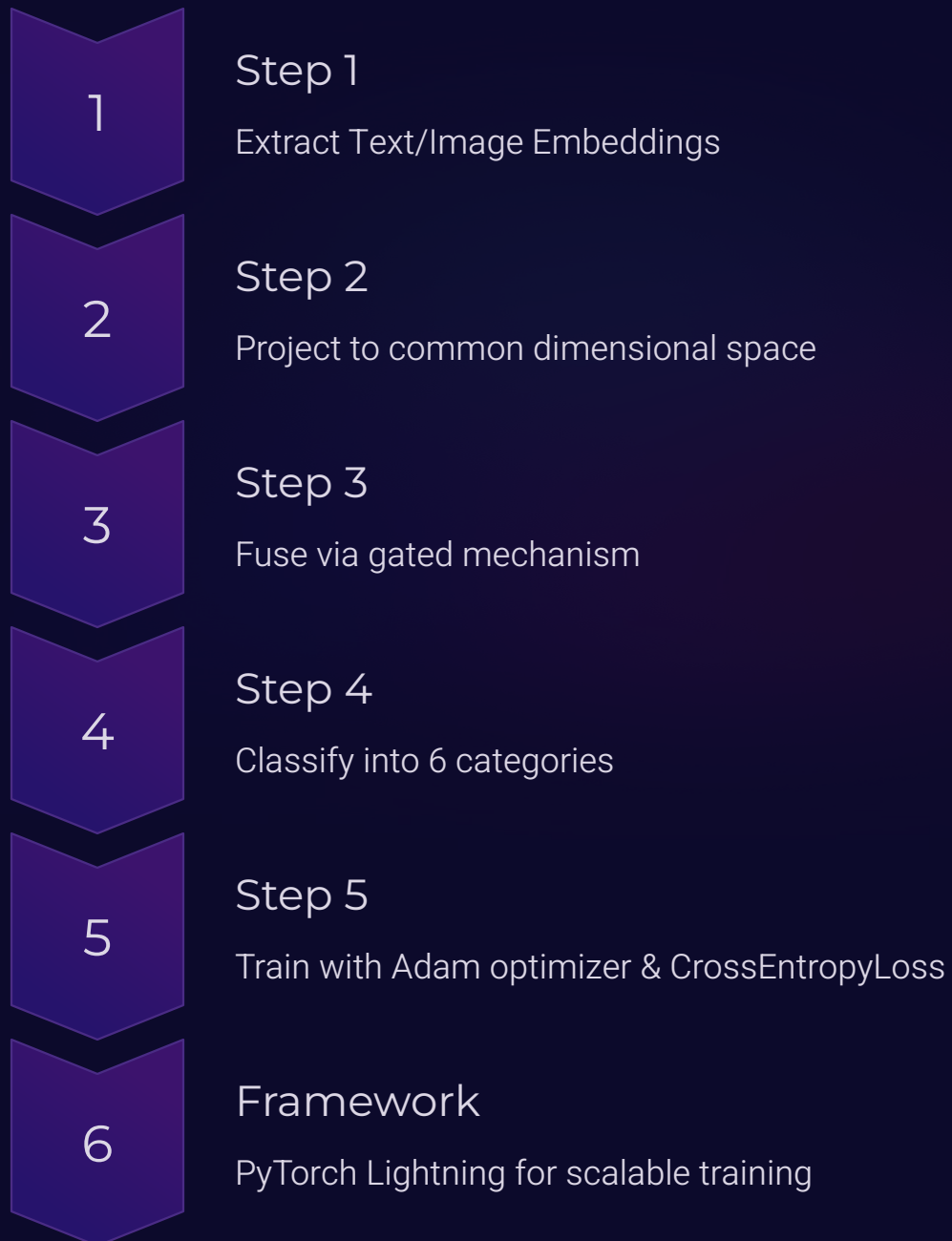
- Projection layer aligns modalities
- Gated fusion module adapts weights

## Classifier

Dense layers with dropout for 6 classes



# Execution Pipeline





## Results & Observations

### Accuracy

Outperforms unimodal baselines significantly

### Robustness

Effective across diverse fake news conditions

### Adaptability

Real-time, deployable in lightweight & high-fidelity settings