Research Report:

Ishrat Jahan Easha
ID: 23266038

Topic Name: Multimodal Sarcasm Detection in Videos

This report has worked on a pioneering research on "Multimodal Sarcasm Detection in Videos using Deep Learning Models" by Ananya Pandey and Dinesh Kumar Vishwakarma. In this social media era, sarcasm is a very common but complex way for communicating which is difficult to detect by automated tools. Existing research only worked with the textual data whereas, this paper worked with both textual and audio features which is the novelty for this paper. This paper aims to improve the accuracy of detecting sarcasm by incorporating minimal clues that may be present in the visual and audio data.

Proposed Methodology:
The proposed approach consists of three main steps:

Feature Extraction:
Textual features are extracted from video transcripts using a pre-trained language model. Simultaneously, audio features are derived from the audio track through a convolutional neural network.

Feature Fusion:
A novel attention-based mechanism is employed to effectively combine textual and audio features.

Sarcasm Classification:
The fused features are utilized to classify videos as either sarcastic or non-sarcastic.

Possible Shortcomings:
While reviewing some shortcomings were noticed. This paper only worked on one dataset and that is MUSTARD multimodal sarcasm detection benchmark dataset. This may limit its possibilities. If worked on a more varied and robust dataset,there may be more intriguing results.

Additionally, the paper did not consider episodic information, for instance, the identity of the speaker or the theme of the video. This may improve the accuracy of the sarcasm detection. This is important as in real life scenarios, context of any situation is very important.

Results:
The proposed approach was evaluated on the MUSTARD dataset, achieving an impressive F1-score of 0.82, outperforming other algorithms.

After applying methods on the MUMUSTARD dataset, they achieved a significant score of 0.82 which is quite impressive.

Results on all algorithms are given below
- Proposed approach: 0.82
- Algorithm 1: 0.75
- Algorithm 2: 0.68

In conclusion this paper has given a new pathway in detecting sarcasm based on both textual and audio data.