Slide 1:

Good morning/afternoon everyone. Today, we'll be presenting our project titled 'Advance Deception Detection using Multimodal Analysis'. This work has been developed by our team at NMIMS, Navi Mumbai. Let's dive into how we combined AI, vision, and audio processing to detect deception in real-time.

Slide 2:

Humans aren't very good at detecting lies — even experts. Traditional methods like polygraphs are unreliable. So, we decided to build an AI system that uses multiple inputs — text, speech, and visuals — to detect deception accurately and objectively.

Slide 3:

We used two datasets. First, the Dolos dataset which includes 1,680 videos labeled as truthful or deceptive, with detailed facial and vocal behavior annotations. Second, the Politifact dataset which has over 11,000 factual statements labeled as true or false. These datasets helped us train and test our multimodal model.

Slide 4:

Our system works in three stages: Text is analyzed using BiLSTM to capture language patterns. Speech is processed to find pitch changes and filler words. And visual inputs capture facial expressions, blinks, and gestures in real-time. These modalities are fused early in the pipeline to ensure real-time prediction.

Slide 5:

We used Adam as the optimizer, with a learning rate of 0.001. Binary crossentropy was our loss function. We trained the model for 30 epochs with a batch size of 32, and used early stopping to prevent overfitting.

Slide 6:

This graph shows how our model's accuracy improved over 30 epochs. As you can see, validation accuracy followed training accuracy closely, indicating a good fit.

Slide 7:

Similarly, this graph shows the training and validation loss decreasing steadily — a sign of a well-trained model with no major overfitting.

Slide 8:

The model achieved a precision of 85.12%, a recall of 82.12%, and an F1 score of 83.98%. These results confirm the model’s ability to distinguish between truthful and deceptive behavior effectively.

Slide 9:

To enhance real-time accuracy, we introduced a dynamic thresholding mechanism. It adjusts the prediction cutoffs based on speech irregularities — like how many filler words or pitch changes were present during speech.

Slide 11:

Despite great results, there are limitations. The Dolos clips are short, making it hard to catch delayed deception. There’s also a gender imbalance, and performance drops under poor lighting or noisy environments. We plan to add new data types like heart rate and improve uncertainty handling.

Slide 12:

This technology has real-world applications — from hiring and border security to courtroom analysis. We envision this being deployed as a lightweight cloud API or embedded in mobile devices, ensuring ethical and accessible use.

Slide 13:

Thank you for your attention! We're excited about the future of real-time deception detection and are happy to answer any questions you may have.