

A decorative geometric pattern on the left side of the slide. It features a large light blue circle in the upper left, a dark blue square with concentric circles below it, a purple triangle to the right of the circle, a solid magenta square below the purple triangle, and a grey square with concentric lines in the bottom left corner. A small dark blue circle is positioned at the intersection of the purple triangle and the magenta square.

MID TERM PROGRESS

- Presented by: Sai Chandra Sri Ramula
- Course/Project: IMAGE CAPTIONING SYSTEM
- Overview: A walkthrough of our Image Captioning project, including progress, challenges.

PROJECT OVERVIEW

- Problem Statement: Automatically generating textual descriptions for images using AI.
- Dataset: Flickr8k (8,000 images with multiple captions).
- Objective: Build a Data Science model that can generate captions based on image content.
- Approach:
- Extracted image features using CNN (InceptionV3, ResNet50).
- Used RNN (LSTM) with Attention to generate captions.

METHODOLOGY & PROGRESS

1.Data Preparation & Text Preprocessing

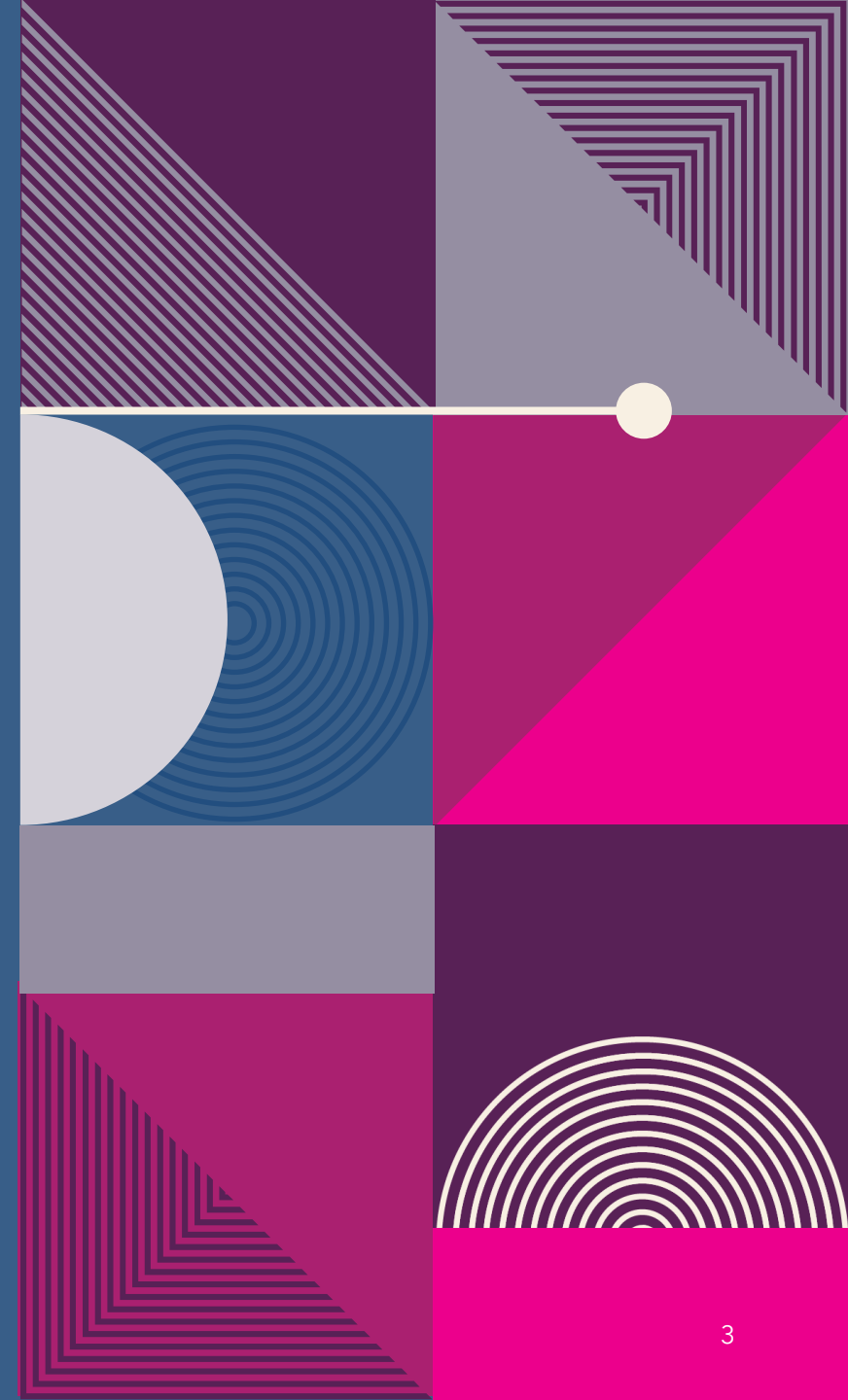
- Loaded dataset and split into Train, Validation, and Test sets.
- Cleaned captions by removing punctuation and converting to lowercase.
- Tokenized captions and created word-to-index mappings.
- Applied padding for uniform sequence length.

2.Feature Extraction & Analysis

- Used InceptionV3 & ResNet50 for feature extraction.
- Saved extracted features (Features.pkl) for efficient training.
- Conducted PCA & t-SNE analysis to compare extracted features.

3.Model Definition & Training

- Implemented CNN Encoder + RNN Decoder (LSTM + Attention).
- Trained the model successfully.
- Faced Overfitting Issues (accuracy reached 1.0000 too fast).



SCATTER PLOT VISUALIZATION USING INCEPTIONV3 AND RESNET50



ADDRESSING OVERFITTING

OVERFITTING IN MODEL TRAINING & FIXES

•Problem:

- Loss dropped too quickly.
- Validation accuracy was unrealistically high.

•Fixes Applied:

- Added Dropout (reduces model overfitting).
- Applied L2 Regularization.
- Checked dataset split to ensure correct validation.

•Next Steps:

- Evaluate model on unseen images.
- Fine-tune model with different hyperparameters.

CONCLUSION

- **Summary:**
 - Explained dataset, text preprocessing, methodology, training, and overfitting issue.
 - Discussed how we addressed overfitting.
 - Next steps: Improve captioning performance and optimize model inference.
- **Future Work:**
 - Implement different decoding strategies (Beam Search, Greedy decoding).
 - Experiment with different pre-trained CNN architectures for better feature extraction.(Like Resnet50)
 - Further improve text preprocessing and feature Extraction to handle complex captions.
- **Open for Questions!**

An abstract geometric design on the left side of the slide. It features a dark blue background with various geometric shapes and patterns. A white circle is positioned near the top left. Below it, a light blue semi-circle is visible. To the right of the semi-circle, there is a pink triangle with diagonal lines. Further down, there is a pink square with a pattern of concentric lines. At the bottom, there is a pink triangle with a pattern of concentric lines. The overall design is modern and minimalist.

THANK YOU