NLP PROJECT

COREFERENCE RESOLUTION IN HINDI AND ENGLISH

Team Name: Delayed Response

Team Members:

1. Harshit Gupta (2020114017)
2. Vamshi Krishna Bonagiri (2020114011)
3. Vanshpreet Singh Kohli (2020114014)

Introduction

The task of finding mentions that refer to the same real-world entity is known as coreference resolution. Majority of recent research has focused on handcrafted characteristics.

Furthermore, the majority of coreference resolution work has been done in English, with little to no work done in Hindi. On English and Hindi datasets, we investigate neural-network based approaches. We also try to use Bert to generate contextual word embeddings that help the models perform better. In Hindi, we also develop a baseline model for neural coreference resolution.

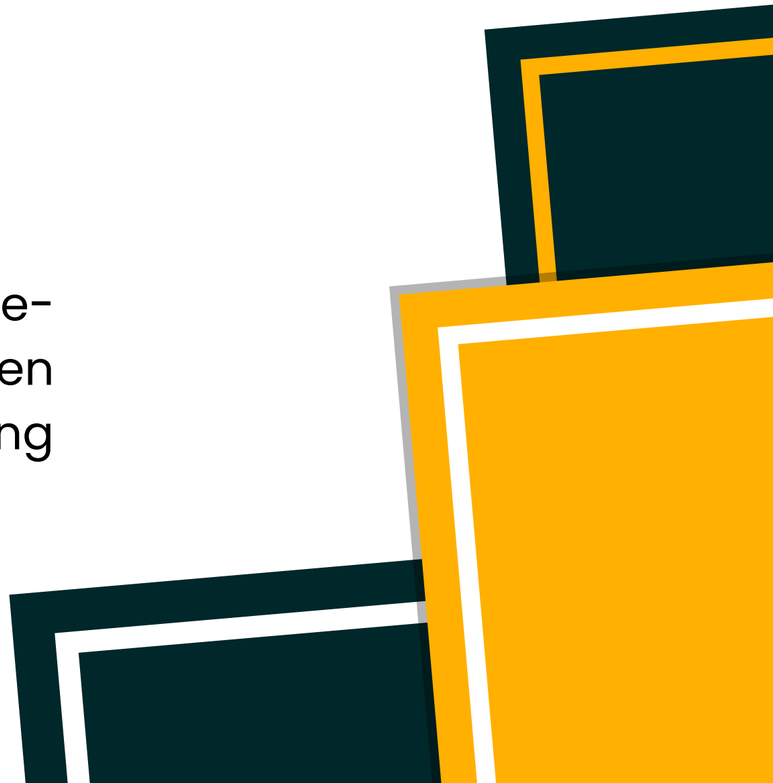
Datasets Used

Hindi Coreference Annotated Data

Dataset from our very own FC Kohli Center on Intelligent Systems (KCIS), IIIT-H, India. We tried applying the concepts learned and our models to the dataset and tried creating our very own baseline for the same.

GAP Coreference Dataset

GAP is a gender-balanced dataset with 8,908 coreference-labelled pairs of (ambiguous pronoun, antecedent name) taken from Wikipedia and distributed by Google AI Language for testing coreference resolution in practical applications.



Models And Results

Model	Accuracy	Precision	Recall	F1 Score
Model 1	0.575250	0.569376	0.575250	0.569171
Model 2	0.730750	0.731322	0.730750	0.730986
Model 3	0.808405	0.908681	0.808405	0.846419
Model 4	0.830500	0.830590	0.830500	0.828175
Model 5	0.821787	0.813485	0.821787	0.817592

Model 1: English Baseline
Model 2: English + BERT Embeddings
Model 3: Hindi + BERT Embeddings
Model 4: English + BERT Model (GAP Dataset)
Model 5: Hindi + BERT Embeddings + LSTM

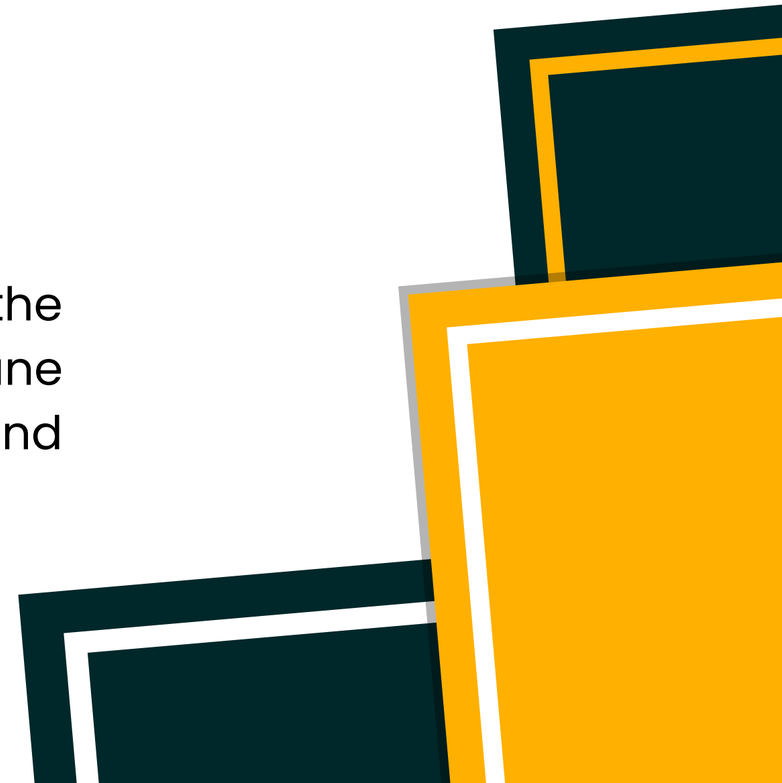
Challenges faced

Scarce Resources

here were no pre-exisiting deep learning methods implemented on the Hindi Dataset. As a result of which, we had to implement everything from scratch,

Performance limitations

Using pretrained BERT models and ensuring that the RAM of the system doesn't seize was extremely important. We had to tune the siz of the batches, the frequency of back propagation and the size of the dataset.



Conclusions

We explored the task of Coreference Resolution using Mention Pair models with contextual word embeddings and Deep Learning Approaches. We also propose a baseline model for the neural Hindi Coreference Resolution task, which has an F1 score of 0.85. Future work can be done by using Mention Ranking and Clustering methods instead of Mention Pair models.