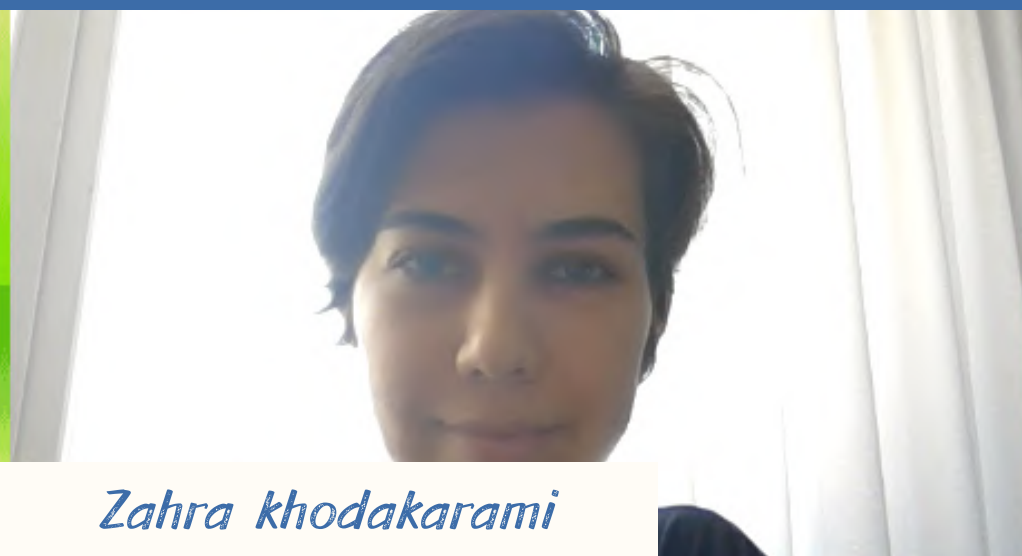


NEUROMATCH ACADEMY PARAPHRASE GENERATION



11TH JULY - 29TH JULY | BATCH 2022



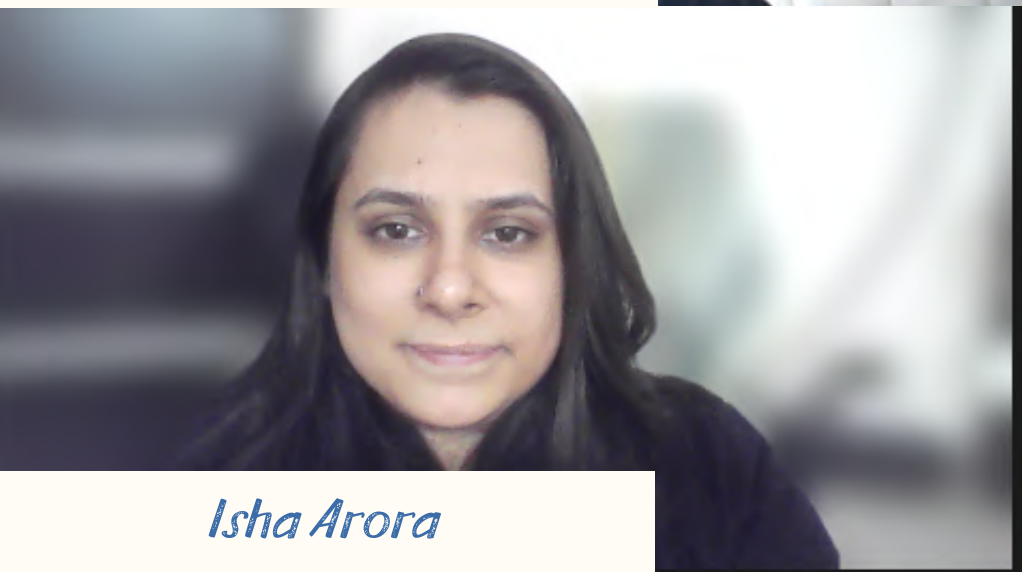
Zahra khodakarami



Vamshi Krishna Navulla



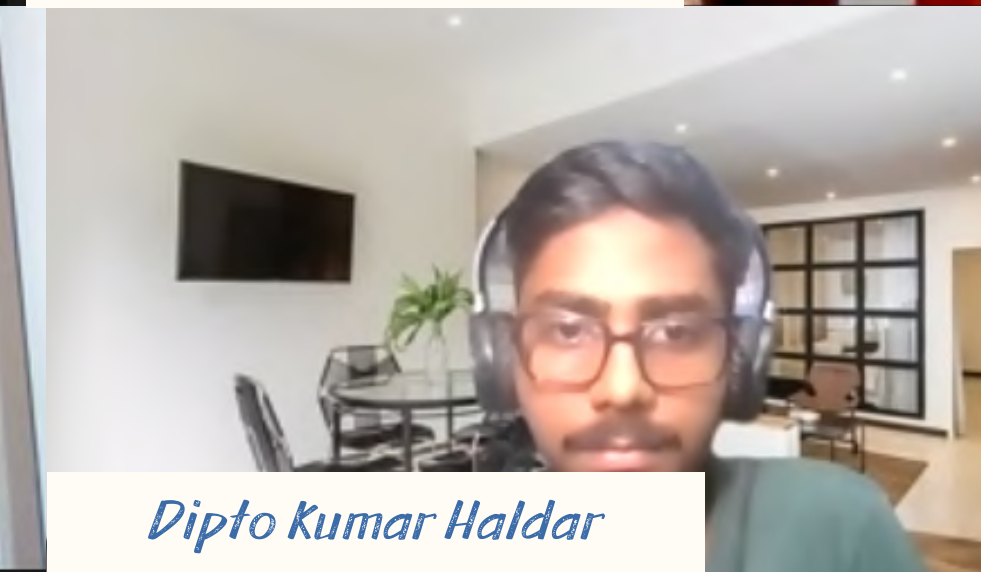
Akshit Singh



Isha Arora



Ranjit Patro



Dipto Kumar Haldar



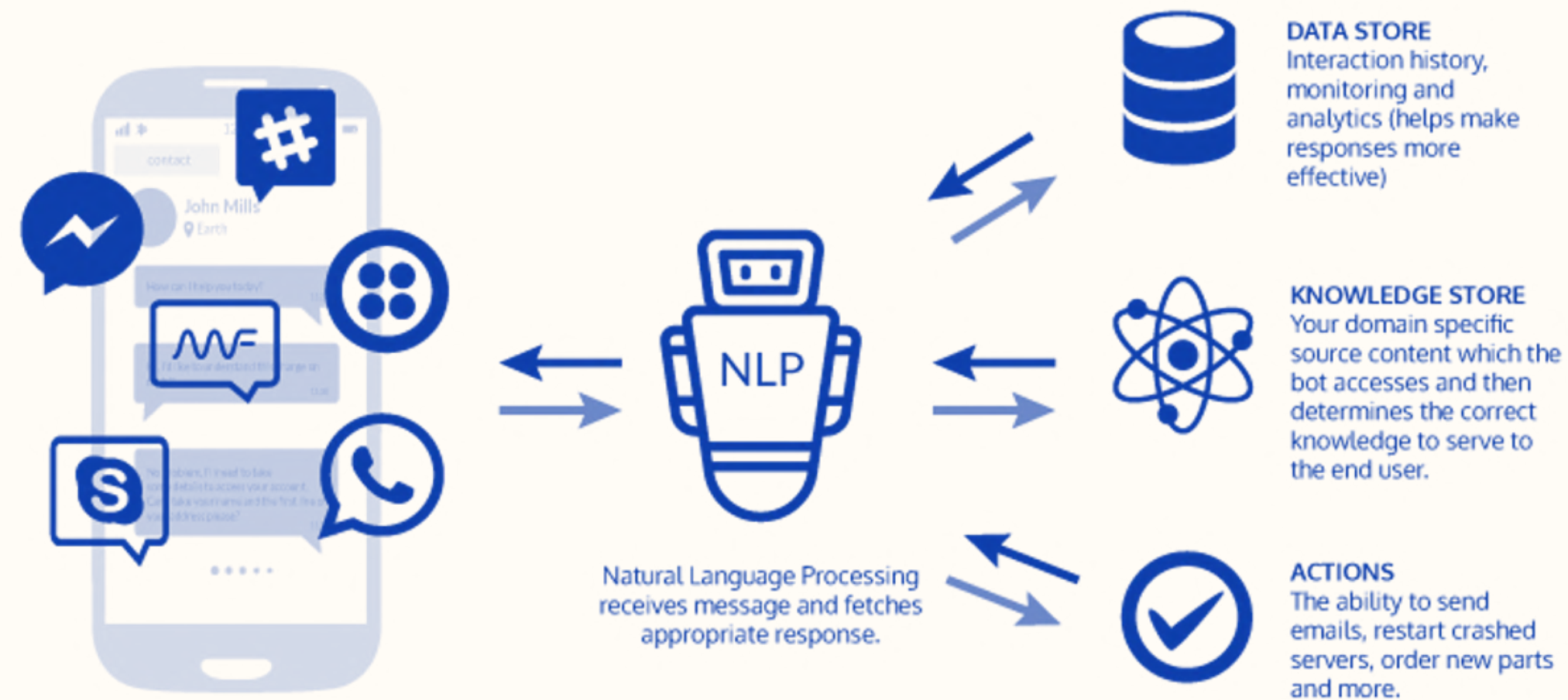
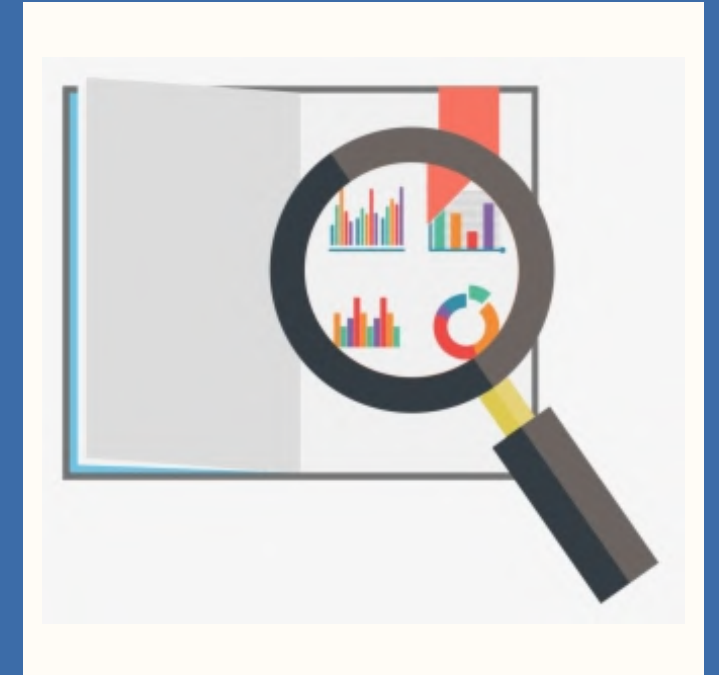
Sneha Kumari

KIBBEH THE SECOND

RESEARCH QUESTIONS

How can changes in the learning rate (LR) affect the accuracy of the model?

If two transformers are implemented in a row, would they perform better in terms of accuracy/result?



DATASET USED

PARAPHRASE ADVERSARIES FROM WORD SCRAMBLING (PAWS)

This dataset contains 108,463 human-labeled and 656k noisily labeled pairs that feature the importance of modeling structure, context, and word order information for the problem of paraphrase generation.

Data Instances

Below are two examples from the dataset:

	Sentence 1	Sentence 2	Label
(1)	Although interchangeable, the body pieces on the 2 cars are not similar.	Although similar, the body parts are not interchangeable on the 2 cars.	0
(2)	Katz was born in Sweden in 1947 and moved to New York City at the age of 1.	Katz was born in 1947 in Sweden and moved to New York at the age of one.	1

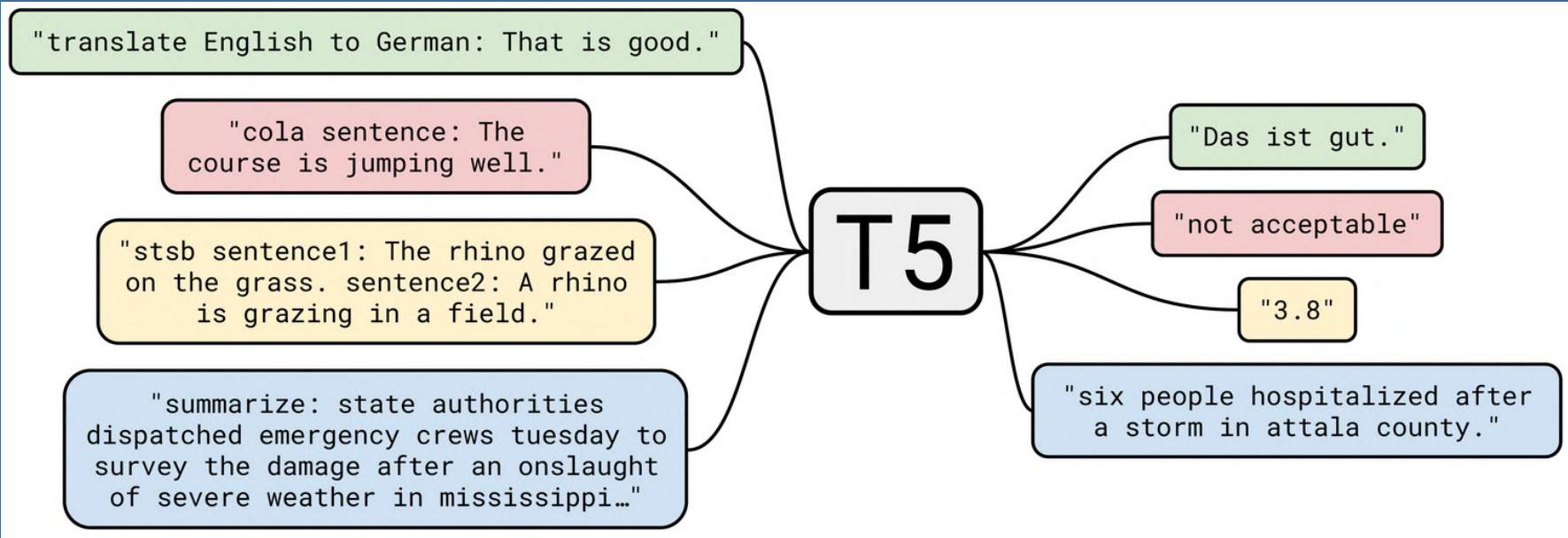


MODELS USED

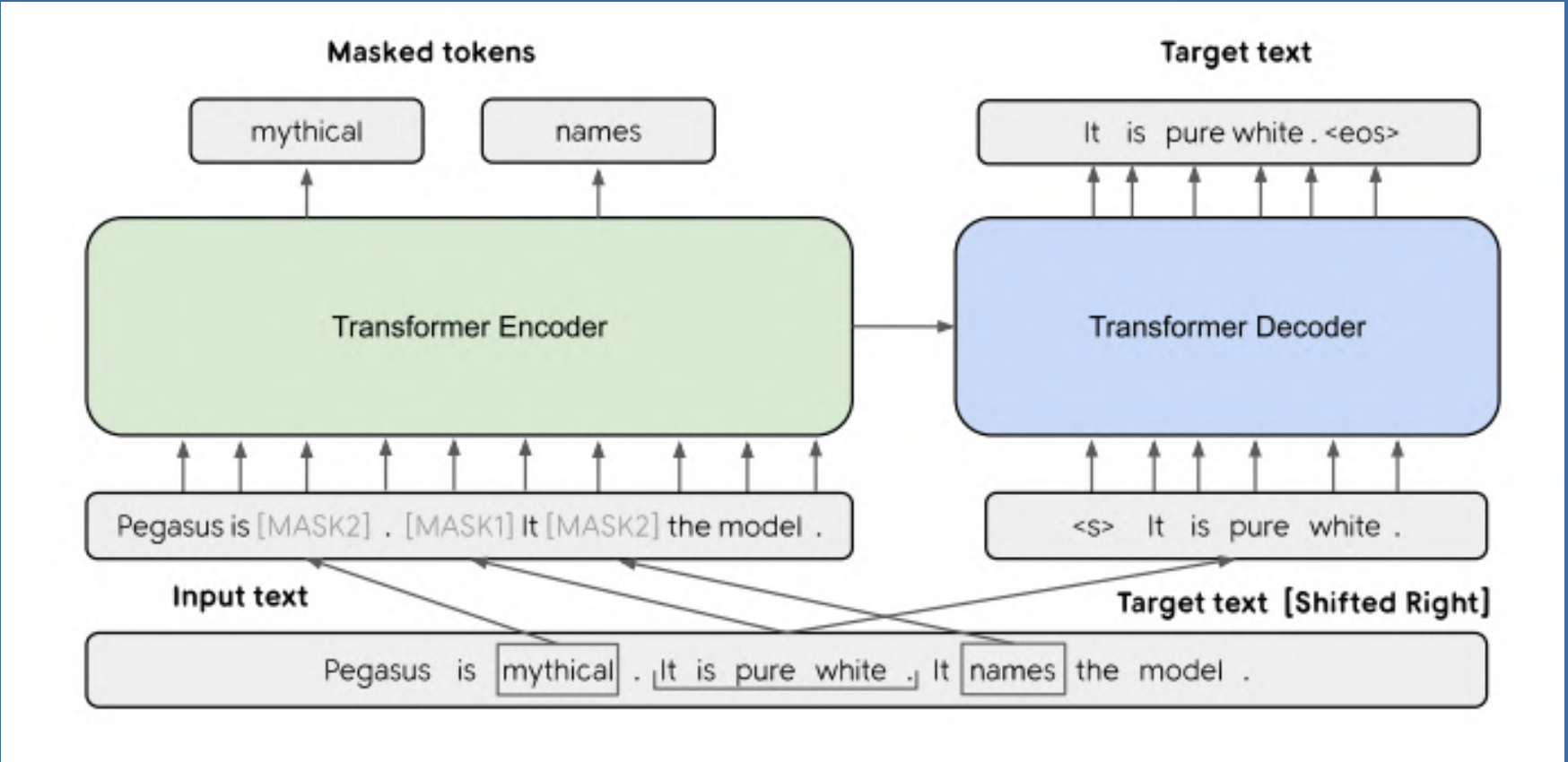


Transformers are proven models in generating paraphrases out of text. In this regard two transformer models were used

T5-base Transformer finetuned for Paraphrase generation.



PEGASUS Transformer. An Encoder-Decoder model with 2-pretrained objectives: GSG and MLM. Mainly used for Extractive Summary tasks. Can also be used for paraphrasing.



EVALUATION METRIC

To calculate the accuracy of Paraphrase Generation, we used Evaluation metrics: **METEOR** (Metric for Evaluation of Translation with Explicit ORdering).

It is an extended and improved version of the more popular **BLUE SCORE**

It gives a correlation of upto **0.96** with **human judgement**

DISADVANTAGE

In METEOR score calculation, "computer" may match "computer" with a score of 1, "computers" with a score of 0.8 and "workstation" with a score of 0.3.

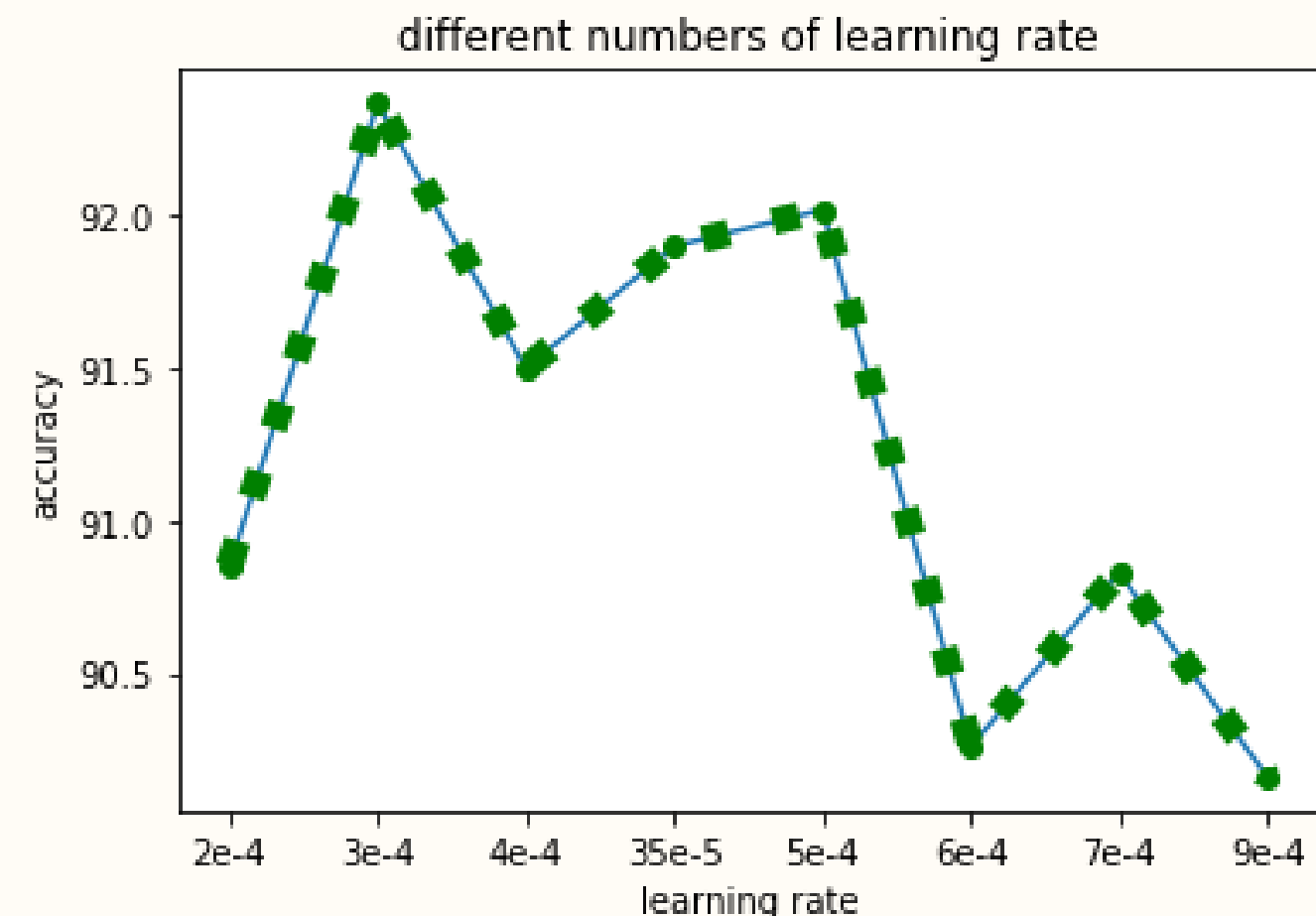
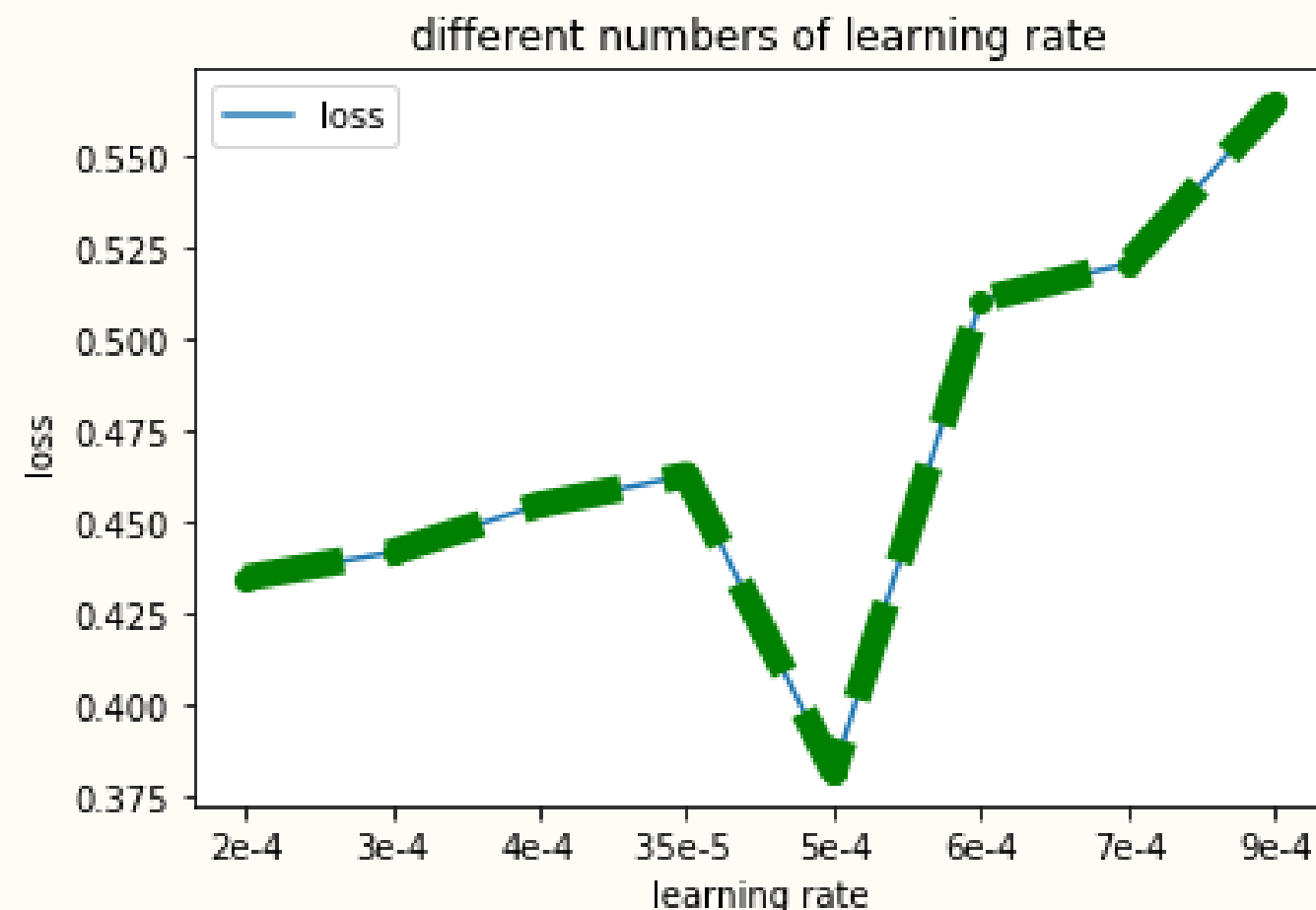




RESULTS & CONCLUSION

- Increasing the LR worsened the loss
- Accuracy fluctuated
- The overall trend was decreasing.

- Pegasus: 60%
- T5 trained model: 92%
- T5 trained model + Pegasus: 56%
- Using a trained T5 model and the pre-trained Pegasus transformer for paraphrasing cannot increase the accuracy of the task



CHALLENGES



- Asking the right question & deciding on the right hypothesis
- Started with the COCO dataset but later switched to PAWS, which worked out a lot better
- Appropriate model selection - shifted from Seq2Seq to Transformer
- Computational limits - Reached the GPU limit in Colab
- Selected a new & advanced evaluation metric but were unable to implement

FUTURE WORK

- Further research can be conducted on investigating different models, such as using trained models and placing them in pipeline to increase overall accuracy of paraphrase generation.
Each model could have a different emphasis on paraphrase generation
- Finding a good Evaluation metrics in the Area of Paraphrase Generation.



TEAM MEMBERS

KIBBEH THE SECOND

ANY QUESTIONS?



THANK YOU!

NEUROMATCH ACADEMY