# Smart paraphraser: Paraphrasing + Grammar correction.

Focused on paraphrasing. Grammar correction can be added later

## Task

To create a smart paraphraser which does well in adverserial paraphrasing task.

## Models

1. **T5-transformer** for paraphrase generation. Three of them perform well viz, tuner007, parrot-paraphraser and vamsi
2. **HRQ-VAE** performs poorly for out of domain tasks. Probably not a good choice for our use case
3. **Quality controlled paraphrase generation**: Does not perform well for out of domain taask
4. **APT**: This model is T5transformer finetuned on APT dataset. Does not perform accurate predictions(adds random names starting with @)Trained on google paws dataset., but provides senetence with different syntactical representations
5. **BART for paraphrasing**: Claimed to perform well Could not check myself as colab kept running out of memory
6. **Joint Copying and Restricted Generation for Paraphrase**: Code not available. Need to train from scratch.
7. **Pushing Paraphrase Away from Original Sentence:A Multi-Round Paraphrase Generation Approach:BTmPG**: From the predictions presented in paper it gives some good paraphrased sentences and somtimes sentences with different semantics. Could be ensembled with semantics similarity model e.g. an encoder to get good predictions. Cuda out of memory error. Could not verify on colab

## Datasets

1. PAWS
2. Quora Question Pairs
3. PIT
4. Paranmt-50
5. PPDB
6. MSCOCO
7. APT
8. Parabank
9. CLIC
10. There is a also DIRT paraphrase dataset, but i am not able to find the download link.