Song Generator

Contents

- Pre-processing task for Taylor swift lyrics dataset
- GPT-Based transformer decoder model – Causal Language Modelling
- Greedy Decoding vs Minimum Bayes Risk Decoding Text Generation
- BLEU-Score based Performance metric
- Inference API
- Docker File
- Bonus Task:- Album Cover Generator (Conditional GAN)

Pre-processing Modules:-

- Convert to lowercase
- Remove numbers
- Remove first sentence
- Remove 'Embed' in last sentence
- Replace \n with "<newline>“ TAG
- Remove all extra spaces

Modelling Techniques:-

- Causal Language Modelling (GPT2)
- Transformer Decoder Mechanism
- performs well in natural language generation task
- Other very standard model BERT is more suited (still bit debatable) to a natural language understanding. That is based on Transformers encoder architecture, Sentiment Analysis.
- T5 – Transformer Encoder-Decoder could be tried out in current case with slight modification in dataset

Decoding Technique:-

Explored two decoding techniques

- **Greedy Decoding**

➢ Choose the best possible candidate generation (by choosing the best possible work at each iteration)

- **Minimum Bayes Risk Decoding**

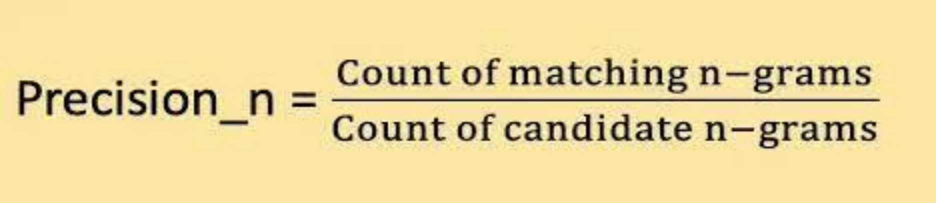
➢ A technique that can be used to generate more flexible and possibly more accurate results in tasks like machine translation and speech recognition by considering the probabilities of a range of potential outputs rather than just the single most likely one.

➢ Generate several candidates and choose the one which has highest average similarity (BLEU or ROGUE or edit distance) with all other candidate generations.

➢ It would increase the inference time, but overall performance will be improved

Performance Metric:-

- BLEU Score based technique
- N-gram overlap between machine translation output and reference translation (it’s usually BLEU Score modified where after each match in reference n-grams are removed from reference)



$$\text{Precision\_n} = \frac{\text{Count of matching n-grams}}{\text{Count of candidate n-grams}}$$

- Compute BLEU Score for n-grams of size 1 to 4
- Geometric mean of these 4 scores
- Add brevity penalty (for too short translations)

$$BLEU\ Score\ Final = m\left(1, \frac{\text{candidate} - \text{length}}{reference - length}\right)\left(\prod_{i=1}^{4} BLEU_i\right)^{1/4}$$

Results:-

Not clear to me:- what does 2-3 verses of song as input means? All these results are generated based on an assumption that 2-3 verses means 2-3 lines.

BLEU Score for 100 elements in Test dataset.

Greedy Decoding

Avg. BLEU :- 32.62 (38.9 with 10 initial lines as an input)

MBR Decoding (with 10 candidate translations)

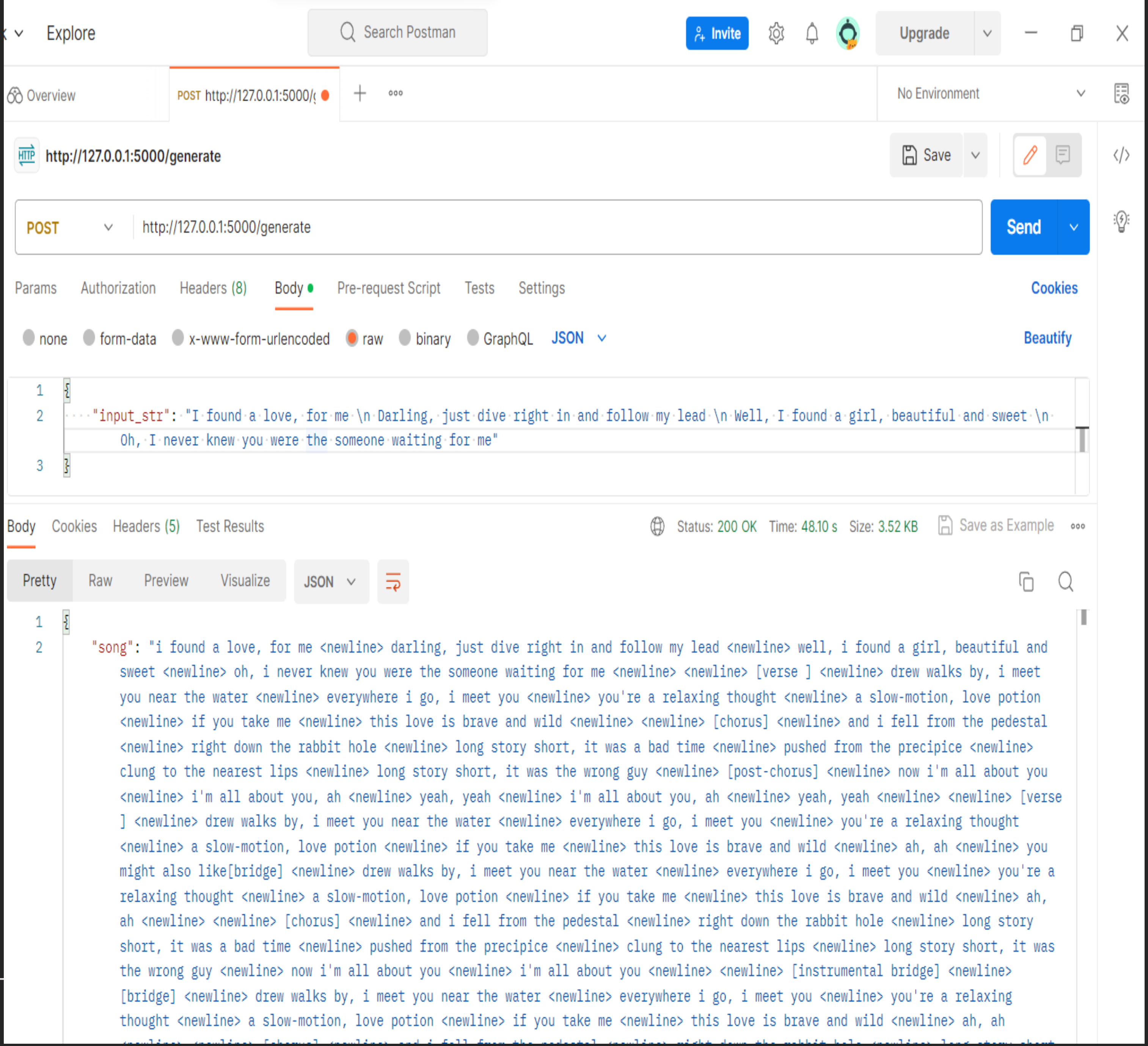
Avg. BLEU :- 35.02 (40.4 with 10 initial lines as an input)

- Also, we need to fine-tune or explore several values for temperature variable in MBR decoding which decides how much randomness are we allowing, overall performance will depend on this

Note:- See Appendix for general understanding of BLEU Score

Inference API

- Used Flask to expose the model as an API and then containerised the code into a Docker container.
- POST method option to take input and generate output
- GET method option to test the working of model
- Used Postman to test the API to generate the “Ed Sheeran Perfect Lyrics” by our model.
- Attached the Dockerfile as well as part of the deliverable.
- Couldn’t actually run that Dockerfile in my system due to unavailability of the software.



http://127.0.0.1:5000/generate

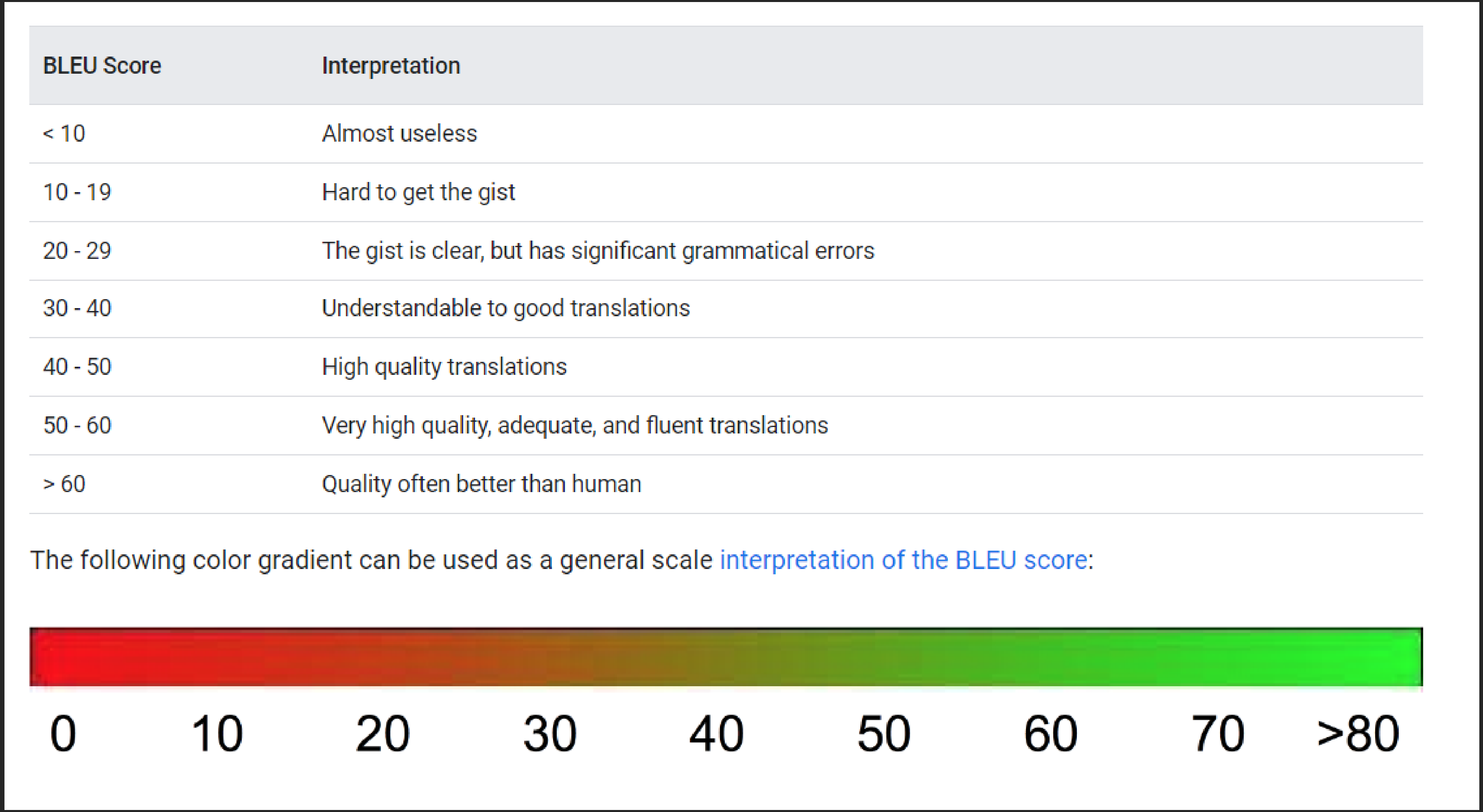
POST http://127.0.0.1:5000/generate

```
{
    "input_str": "I found a love, for me \n Darling, just dive right in and follow my lead \n Well, I found a girl, beautiful and sweet \n Oh, I never knew you were the someone waiting for me"
}
```

Status: 200 OK Time: 48.10 s Size: 3.52 KB

```
{
    "song": "i found a love, for me <newline> darling, just dive right in and follow my lead <newline> well, i found a girl, beautiful and sweet <newline> oh, i never knew you were the someone waiting for me <newline> <newline> [verse ] <newline> drew walks by, i meet you near the water <newline> everywhere i go, i meet you <newline> you're a relaxing thought <newline> a slow-motion, love potion <newline> if you take me <newline> this love is brave and wild <newline> <newline> [chorus] <newline> and i fell from the pedestal <newline> right down the rabbit hole <newline> long story short, it was a bad time <newline> pushed from the precipice <newline> clung to the nearest lips <newline> long story short, it was the wrong guy <newline> [post-chorus] <newline> now i'm all about you <newline> i'm all about you, ah <newline> yeah, yeah <newline> i'm all about you, ah <newline> yeah, yeah <newline> <newline> [verse ] <newline> drew walks by, i meet you near the water <newline> everywhere i go, i meet you <newline> you're a relaxing thought <newline> a slow-motion, love potion <newline> if you take me <newline> this love is brave and wild <newline> ah, ah <newline> you might also like[bridge] <newline> drew walks by, i meet you near the water <newline> everywhere i go, i meet you <newline> you're a relaxing thought <newline> a slow-motion, love potion <newline> if you take me <newline> this love is brave and wild <newline> ah, ah <newline> <newline> [chorus] <newline> and i fell from the pedestal <newline> right down the rabbit hole <newline> long story short, it was a bad time <newline> pushed from the precipice <newline> clung to the nearest lips <newline> long story short, it was the wrong guy <newline> now i'm all about you <newline> i'm all about you <newline> <newline> [instrumental bridge] <newline> [bridge] <newline> drew walks by, i meet you near the water <newline> everywhere i go, i meet you <newline> you're a relaxing thought <newline> a slow-motion, love potion <newline> if you take me <newline> this love is brave and wild <newline> ah, ah <newline> <newline> [chorus] <newline> and i fell from the pedestal <newline> right down the rabbit hole <newline> long story short
```

Appendix 1 – BLEU Score Rough understanding (Google Cloud)



BLEU Score	Interpretation
< 10	Almost useless
10 - 19	Hard to get the gist
20 - 29	The gist is clear, but has significant grammatical errors
30 - 40	Understandable to good translations
40 - 50	High quality translations
50 - 60	Very high quality, adequate, and fluent translations
> 60	Quality often better than human

The following color gradient can be used as a general scale interpretation of the BLEU score: