# To get the embeddings, we have tried two versions of DNABERT.

## DNABERT - 1

```
file: aakash -> embeddings.ipynb
Python version: 3.9.21
```

## DNABERT - 2

```
file: Laavanya -> embeddings2.ipynb
Python version: 3.8.20
```

Packages were installed individually referring to:

Requirements for DNABERT2

DNABERT - 2 was tried on one of the parquet files (batch 0)

```
time taken to process 1,00,000 sequences: 112m 8.7s
```

Next task: Compare the embeddings obtained from DNABERT - 1 with DNABERT - 2 using cosine similarity & choose.