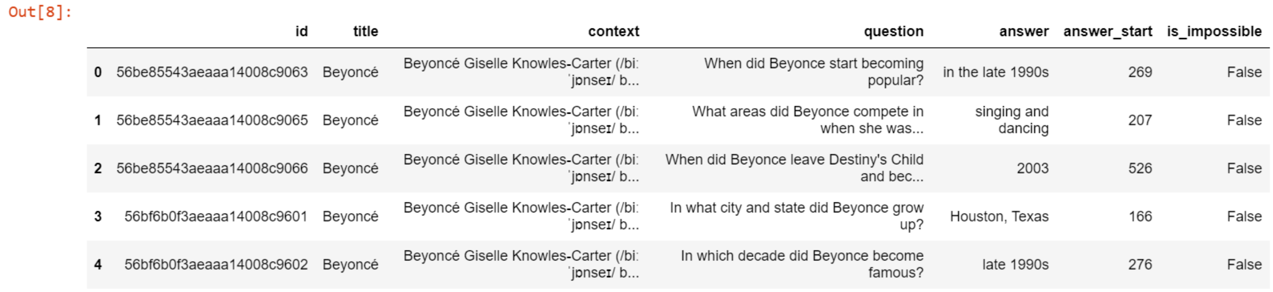
Question Answering model

Approach 2: Bert based models

Methodology

1.1 Dataset

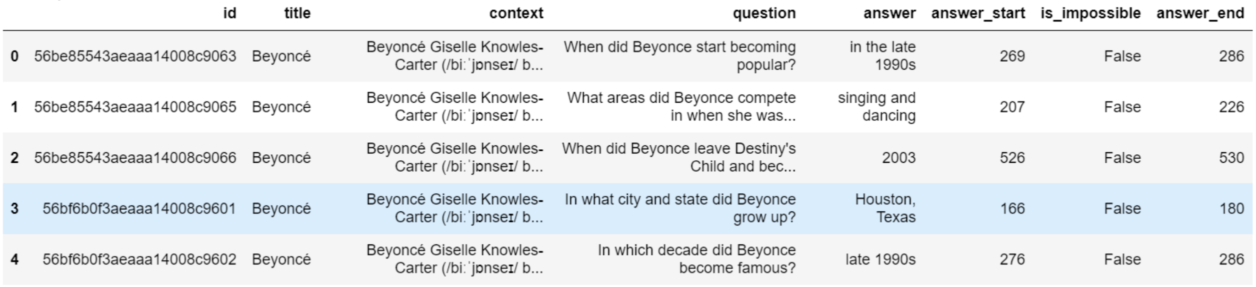
- Dataset snapshot:



Out[8]:

	id	title	context	question	answer	answer_start	is_impossible
0	56be85543aeaaa14008c9063	Beyoncé	Beyoncé Giselle Knowles-Carter (/biːˈjɒnseɪ/ b...	When did Beyonce start becoming popular?	in the late 1990s	269	False
1	56be85543aeaaa14008c9065	Beyoncé	Beyoncé Giselle Knowles-Carter (/biːˈjɒnseɪ/ b...	What areas did Beyonce compete in when she was...	singing and dancing	207	False
2	56be85543aeaaa14008c9066	Beyoncé	Beyoncé Giselle Knowles-Carter (/biːˈjɒnseɪ/ b...	When did Beyonce leave Destiny's Child and bec...	2003	526	False
3	56bf6b0f3aeaaa14008c9601	Beyoncé	Beyoncé Giselle Knowles-Carter (/biːˈjɒnseɪ/ b...	In what city and state did Beyonce grow up?	Houston, Texas	166	False
4	56bf6b0f3aeaaa14008c9602	Beyoncé	Beyoncé Giselle Knowles-Carter (/biːˈjɒnseɪ/ b...	In which decade did Beyonce become famous?	late 1990s	276	False

- But this dataset has only answer start index so to train Bert models we need the start and end indices.
- Once processed it looks like



	id	title	context	question	answer	answer_start	is_impossible	answer_end
0	56be85543aeaaa14008c9063	Beyoncé	Beyoncé Giselle Knowles-Carter (/biːˈjɒnseɪ/ b...	When did Beyonce start becoming popular?	in the late 1990s	269	False	286
1	56be85543aeaaa14008c9065	Beyoncé	Beyoncé Giselle Knowles-Carter (/biːˈjɒnseɪ/ b...	What areas did Beyonce compete in when she was...	singing and dancing	207	False	226
2	56be85543aeaaa14008c9066	Beyoncé	Beyoncé Giselle Knowles-Carter (/biːˈjɒnseɪ/ b...	When did Beyonce leave Destiny's Child and bec...	2003	526	False	530
3	56bf6b0f3aeaaa14008c9601	Beyoncé	Beyoncé Giselle Knowles-Carter (/biːˈjɒnseɪ/ b...	In what city and state did Beyonce grow up?	Houston, Texas	166	False	180
4	56bf6b0f3aeaaa14008c9602	Beyoncé	Beyoncé Giselle Knowles-Carter (/biːˈjɒnseɪ/ b...	In which decade did Beyonce become famous?	late 1990s	276	False	286

1.2 Creating the corpus

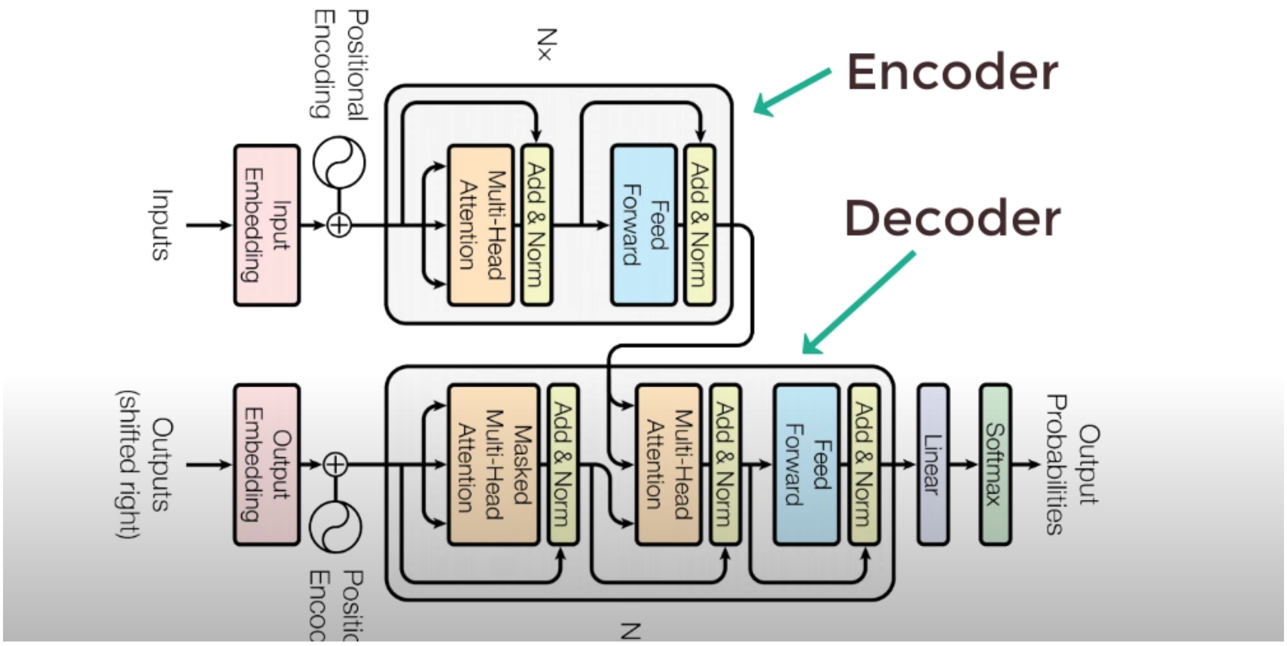
- The corpus for training was created by correcting the answer start index and adding the answer end indexes for each observation in the dataset.
- Then the question was concatenated to the context with a `[SEP]` token with the help of the tokenizer from huggingface.

1.3 Training

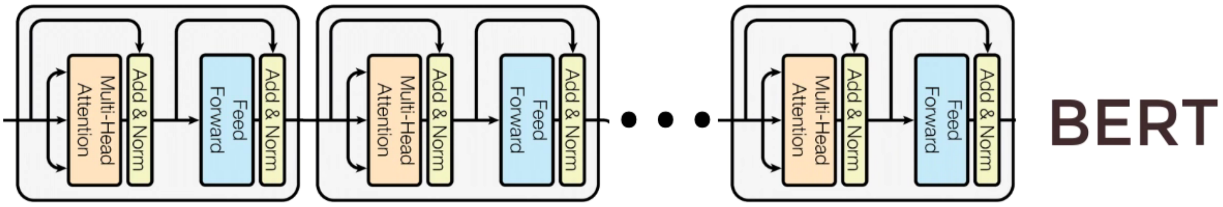
- For training, I used the pre-trained bert models to initialize the weights for Bert layers, with the appropriate head for `Question Answering modeling`.
- Then, I trained the models for 3-4 epochs (cause the running time with gpu takes 4-5hrs).
- I trained 2 models:
    1. Bert
    2. DistilBert

1.4 Model explaination:

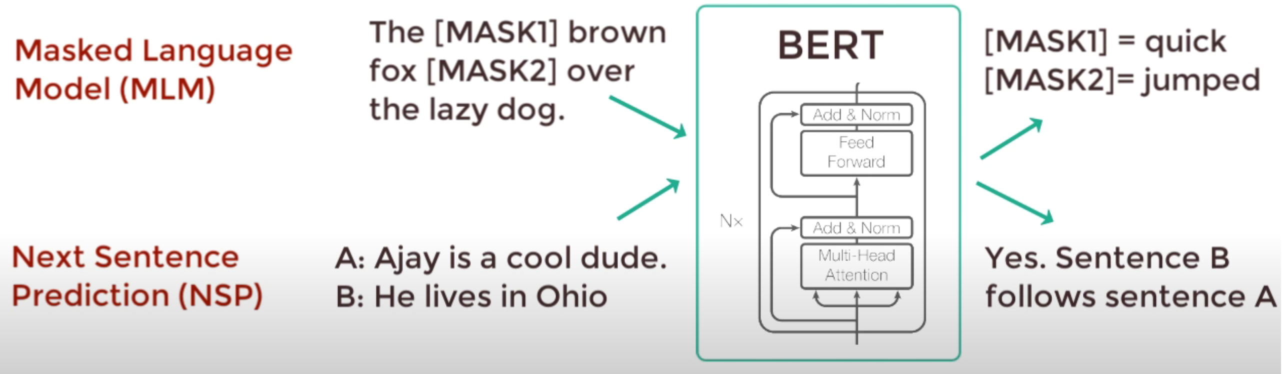
- Transformer:



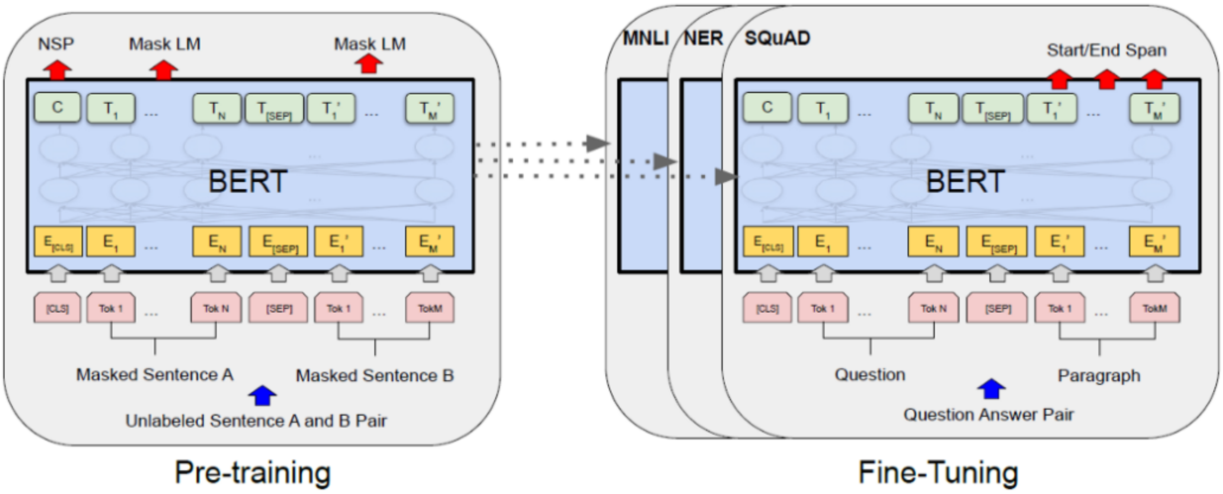
- Bert:



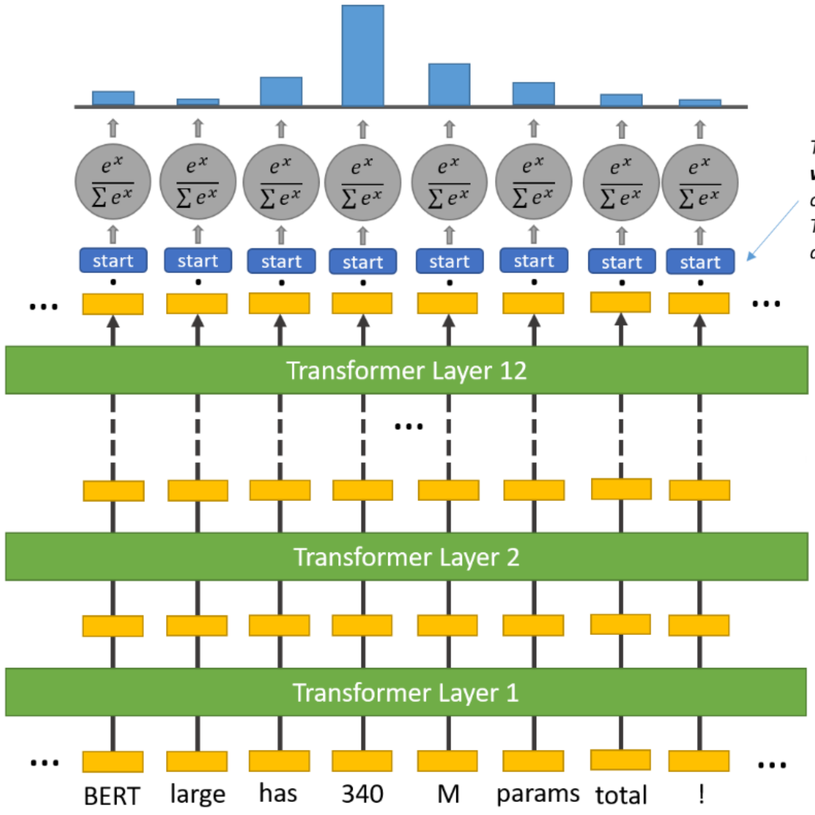
- Bert learns NSP, MLM:


Masked Language Model (MLM)
The [MASK1] brown fox [MASK2] over the lazy dog.
BERT
Add & Norm
Feed Forward
Nx
Add & Norm
Multi-Head Attention
[MASK1] = quick
[MASK2]= jumped
Next Sentence Prediction (NSP)
A: Ajay is a cool dude.
B: He lives in Ohio
Yes. Sentence B follows sentence A

- Fine tuning for question answering:

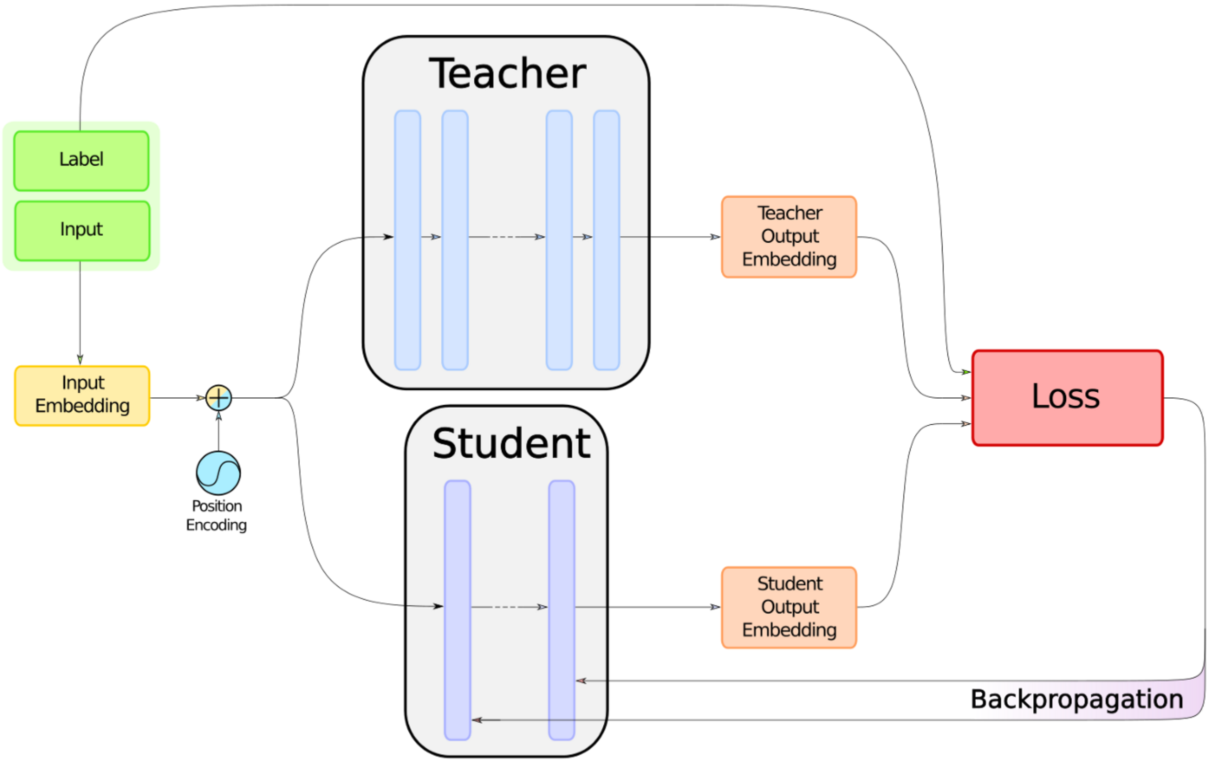


- In more detail: Similarly for end index



This length 768 vector is the ***weights*** *for the start token classifier.*
The ***same weights*** *are applied to* ***every position***.

- Distilbert: Made by distiling bert with a Teacher (Trained Bert) to teach the student (distilbert) with half the number of transformer layers, to learn to perform nearly the same as BERT.



1.5 Evaluating

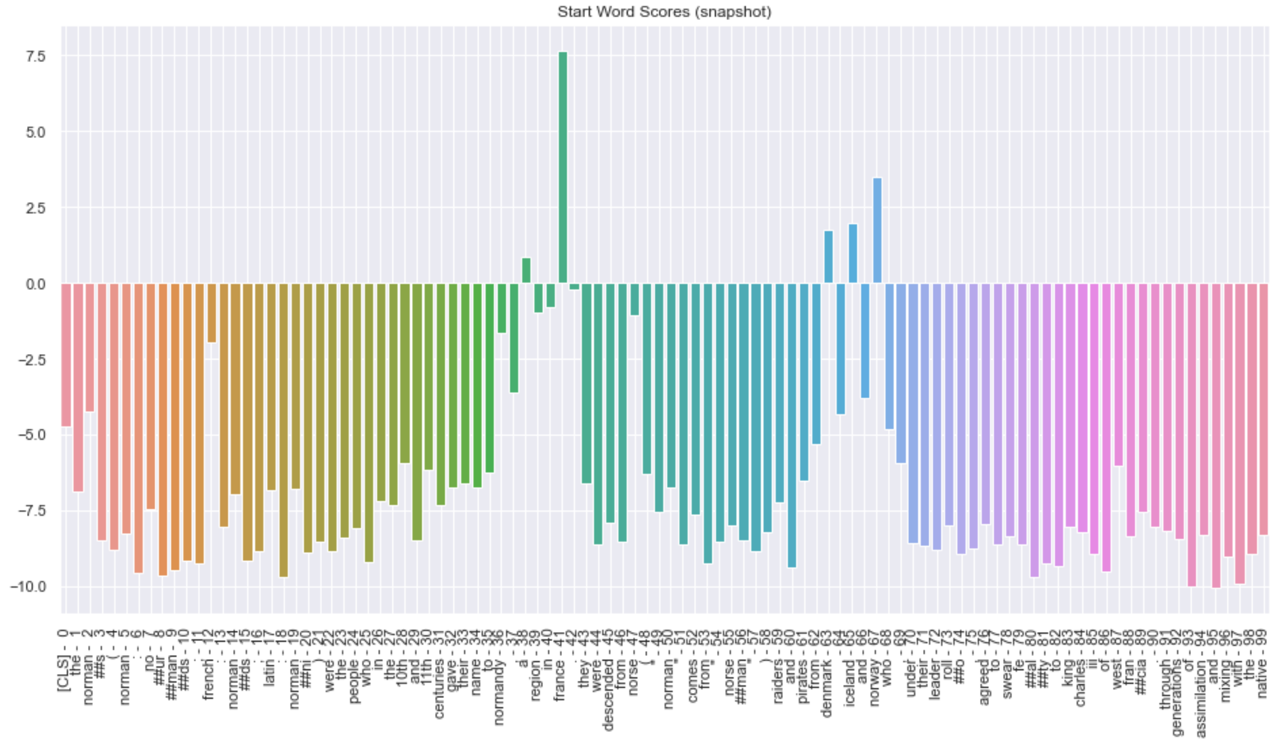
- To evaluate the performance of these models I used the evaluation script provided by SQuAD, that calculates the exact match (`EM`) scores and `F1-scores`.

1.6 Results

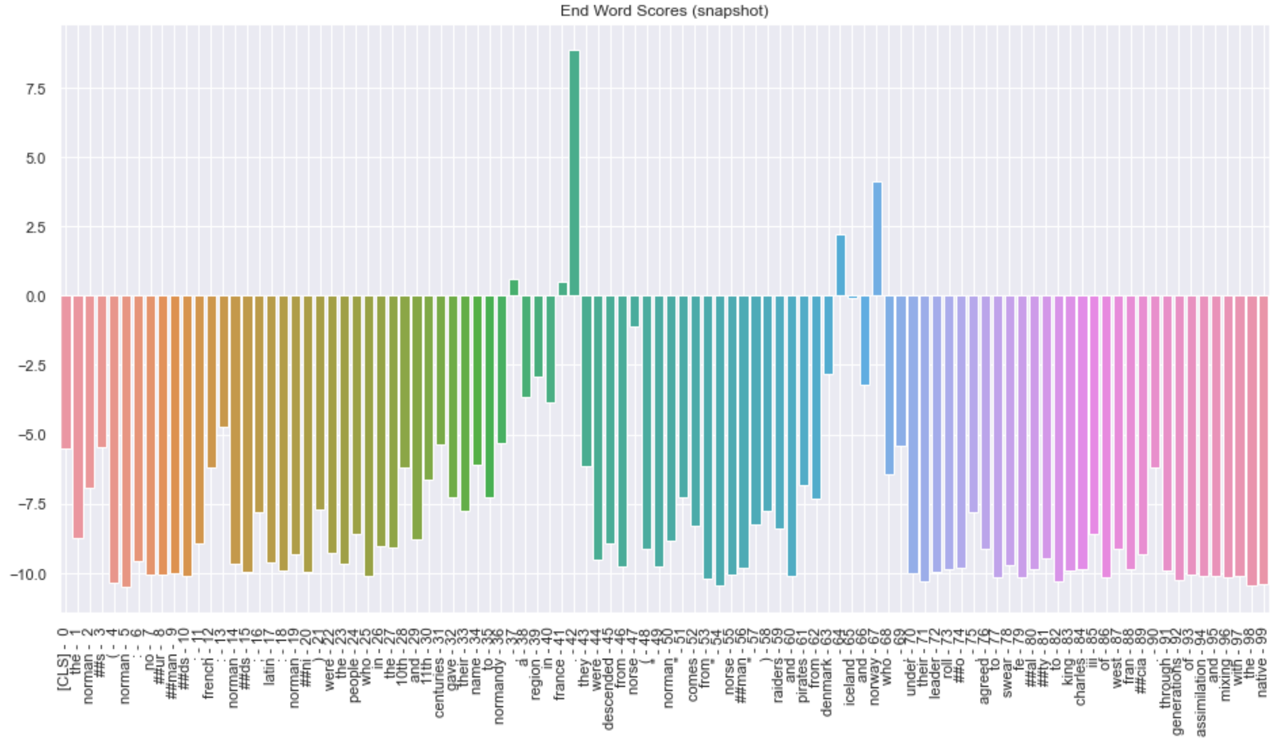
- Sample Question Answering output

```
Context The Normans (Norman: Nourmands; French: Normands; Latin: Normanni) were the people who in the 10th and 11th centuries g
ave their name to Normandy, a region in France. They were descended from Norse ("Norman" comes from "Norseman") raiders and pir
ates from Denmark, Iceland and Norway who, under their leader Rollo, agreed to swear fealty to King Charles III of West Franci
a. Through generations of assimilation and mixing with the native Frankish and Roman-Gaulish populations, their descendants wou
ld gradually merge with the Carolingian-based cultures of West Francia. The distinct cultural and ethnic identity of the Norman
s emerged initially in the first half of the 10th century, and it continued to evolve over the succeeding centuries.
question In what country is Normandy located?
answer france .
------------------------------
Context The Normans (Norman: Nourmands; French: Normands; Latin: Normanni) were the people who in the 10th and 11th centuries g
ave their name to Normandy, a region in France. They were descended from Norse ("Norman" comes from "Norseman") raiders and pir
ates from Denmark, Iceland and Norway who, under their leader Rollo, agreed to swear fealty to King Charles III of West Franci
a. Through generations of assimilation and mixing with the native Frankish and Roman-Gaulish populations, their descendants wou
ld gradually merge with the Carolingian-based cultures of West Francia. The distinct cultural and ethnic identity of the Norman
s emerged initially in the first half of the 10th century, and it continued to evolve over the succeeding centuries.
question When were the Normans in Normandy?
answer 10th and 11th centuries
------------------------------
Context The Normans (Norman: Nourmands; French: Normands; Latin: Normanni) were the people who in the 10th and 11th centuries g
ave their name to Normandy, a region in France. They were descended from Norse ("Norman" comes from "Norseman") raiders and pir
ates from Denmark, Iceland and Norway who, under their leader Rollo, agreed to swear fealty to King Charles III of West Franci
a. Through generations of assimilation and mixing with the native Frankish and Roman-Gaulish populations, their descendants wou
ld gradually merge with the Carolingian-based cultures of West Francia. The distinct cultural and ethnic identity of the Norman
s emerged initially in the first half of the 10th century, and it continued to evolve over the succeeding centuries.
question When were the Normans in Normandy?
answer 10th and 11th centuries
------------------------------
```

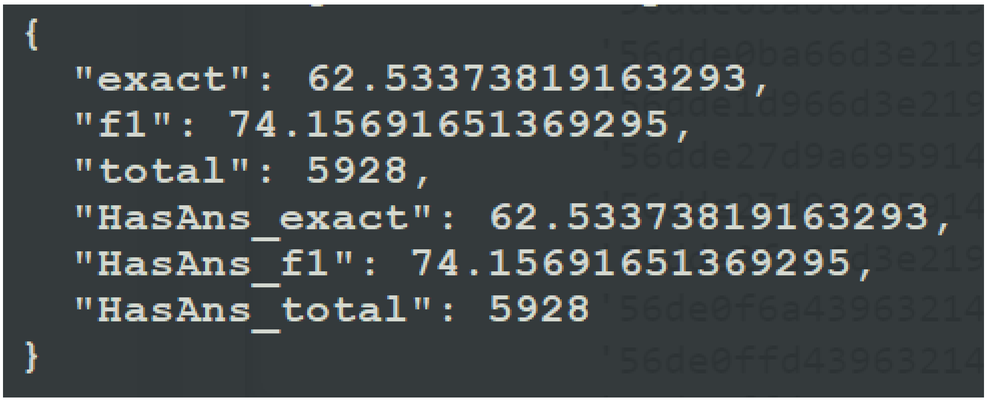
- Visualization of the start index scores: To make the visualization appealing, just a subset of the words are shown.



- Visualization of the answer scores:To make the visualization appealing, just a subset of the words are shown.

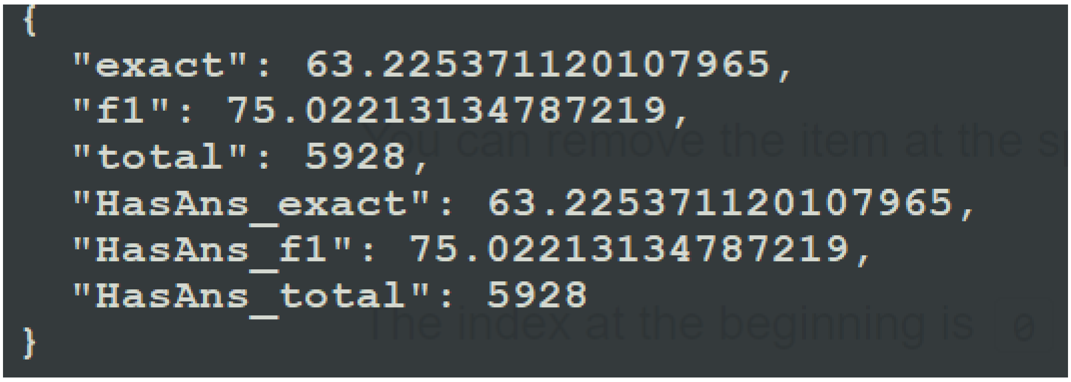


- Bert Model Results:



```
{
  "exact": 62.53373819163293,
  "f1": 74.15691651369295,
  "total": 5928,
  "HasAns_exact": 62.53373819163293,
  "HasAns_f1": 74.15691651369295,
  "HasAns_total": 5928
}
```

- DistilBert Results:



```
{
  "exact": 63.225371120107965,
  "f1": 75.02213134787219,
  "total": 5928,
  "HasAns_exact": 63.225371120107965,
  "HasAns_f1": 75.02213134787219,
  "HasAns_total": 5928
}
```

Questions

1. I intend to develop an ensemble out of Bert, DistilBert, and most likely Alberta/Roberta, so my question is whether it makes sense to apply a weighted averaging on the start and end index scores.
2. Next im going to experiment with ALBERTA and RoBertA, can you please send me some reading materials on it, if possible?
3. Also, please let me know if im going in the right track or I need to address any major concerns?
4. Also, if possible, could you kindly recommend anything else you'd like me to implement?