RETRIEVAL ENSEMBLE

Homework 3

BLM5109 Collective Learning

ÖNDER GÖRMEZ

21501035
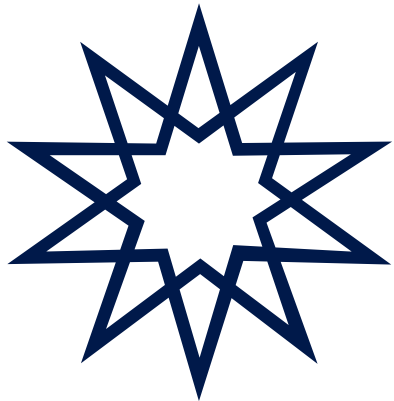
Agenda

Agenda

Embedding Elapsed Times for 50 Sample

Turkish Instructions Dataset
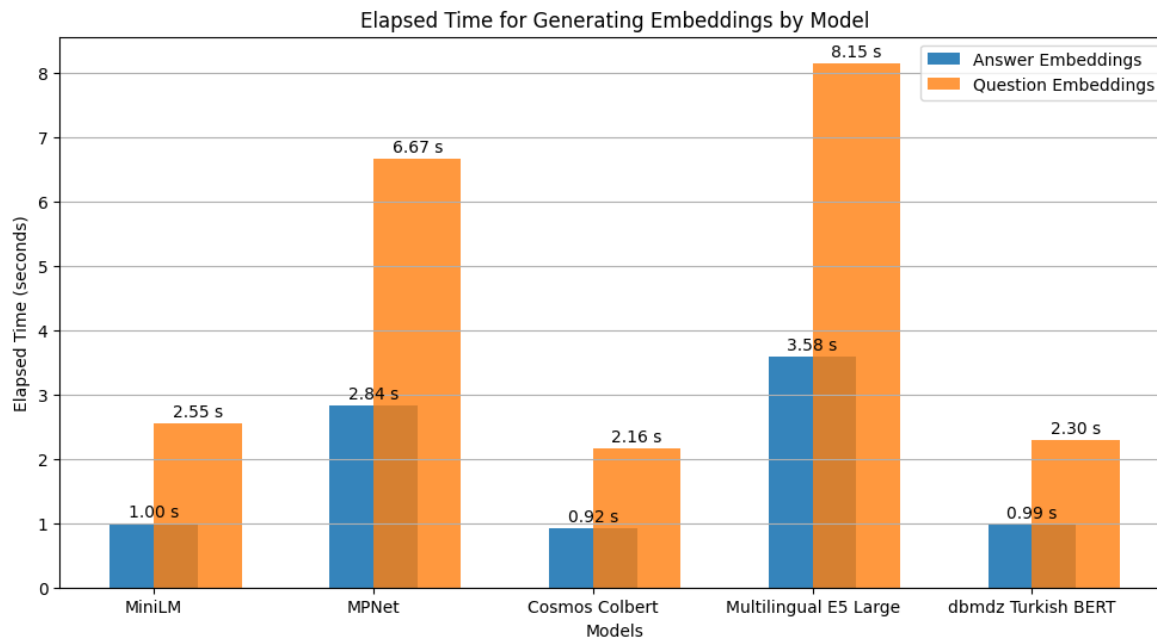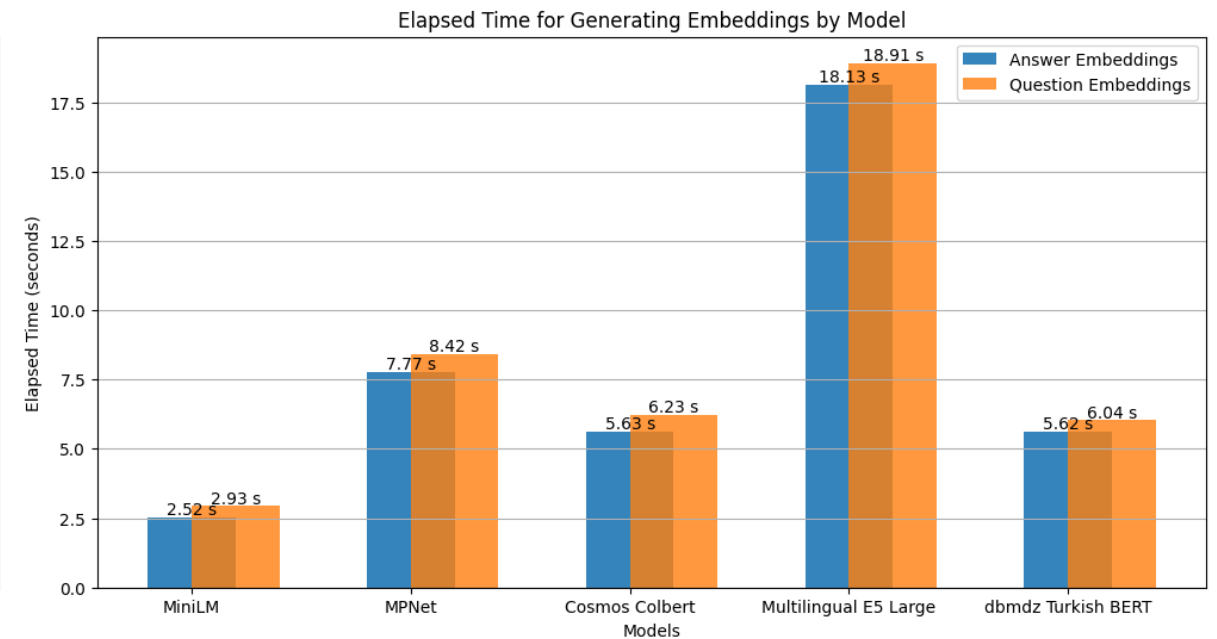
Trivia QA Dataset

Accuracy Comparison for 50 Sample

Turkish Instructions Dataset
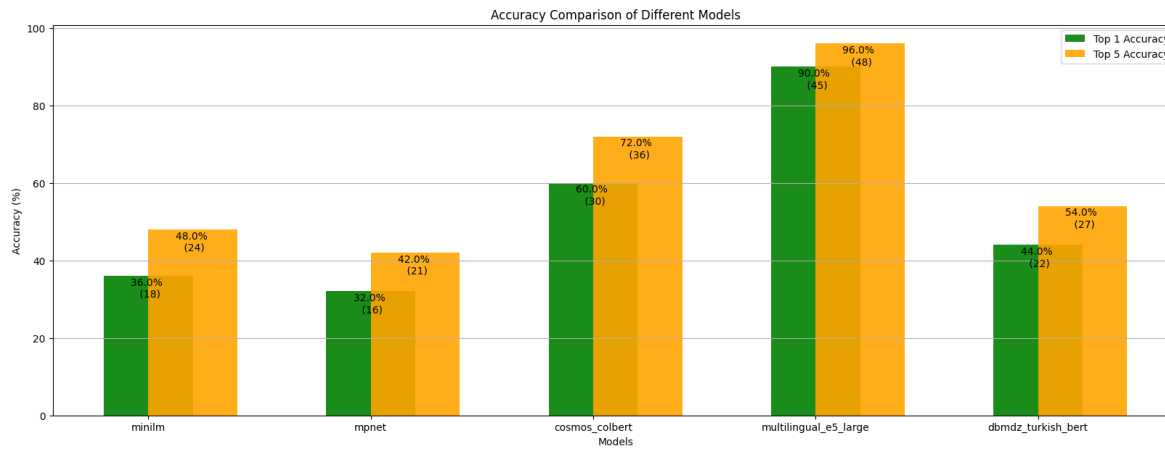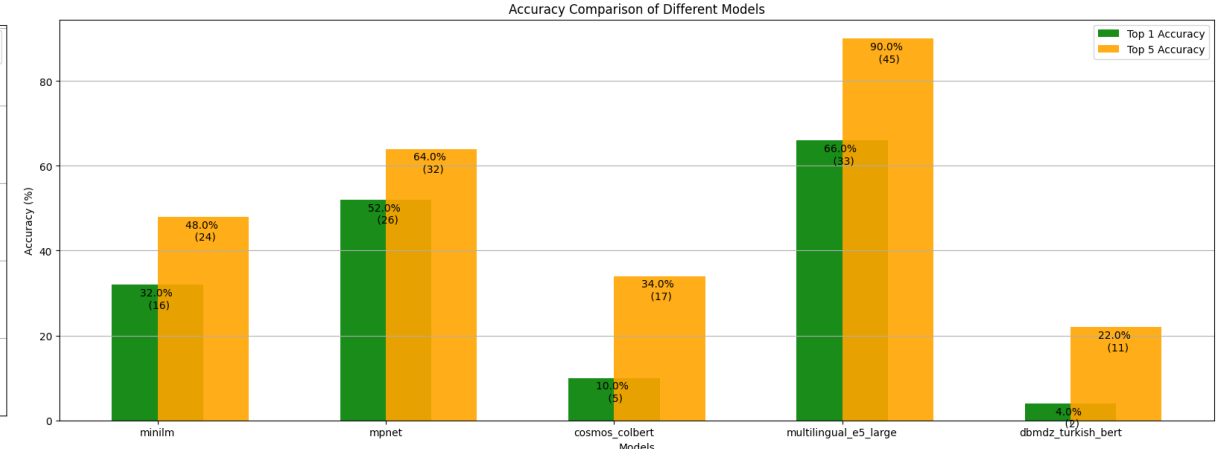
Trivia QA Dataset

Accuracy Comparison with Ensemble for 50 Sample

Turkish Instructions Dataset
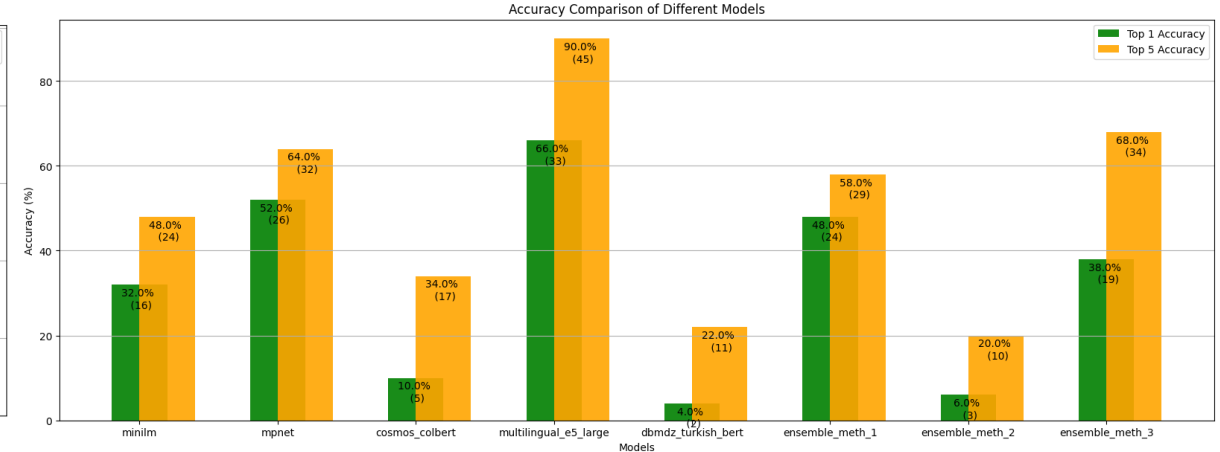
Trivia QA Dataset

Agenda

Embedding Elapsed Times for 200 Sample

Turkish Instructions Dataset
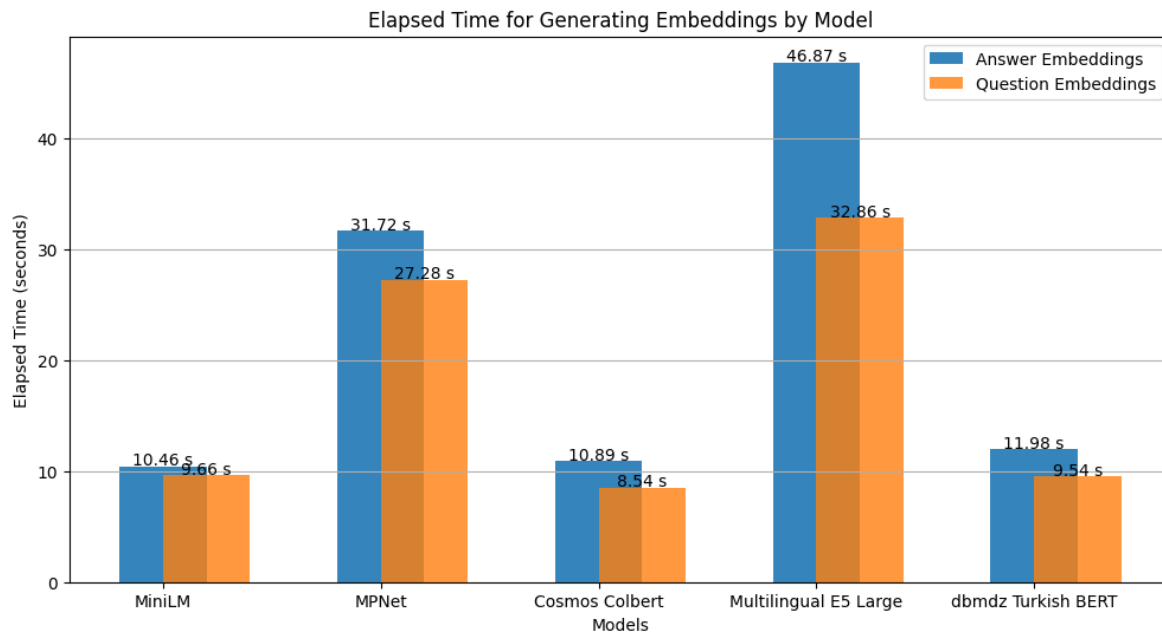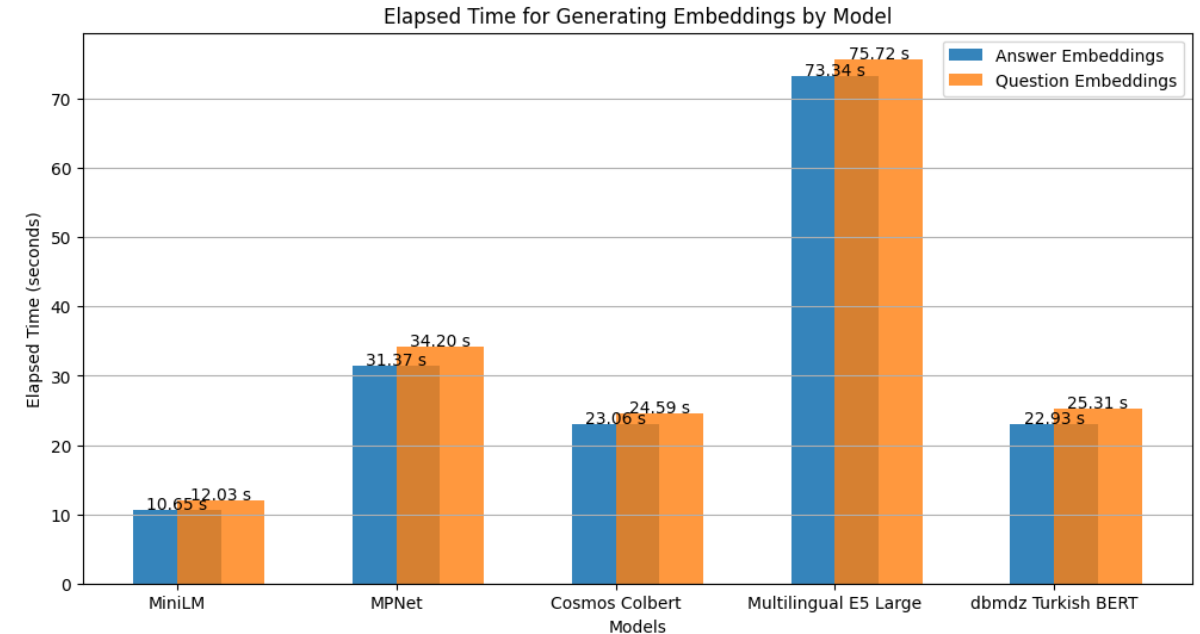
Trivia QA Dataset

Accuracy Comparison for 200 Sample
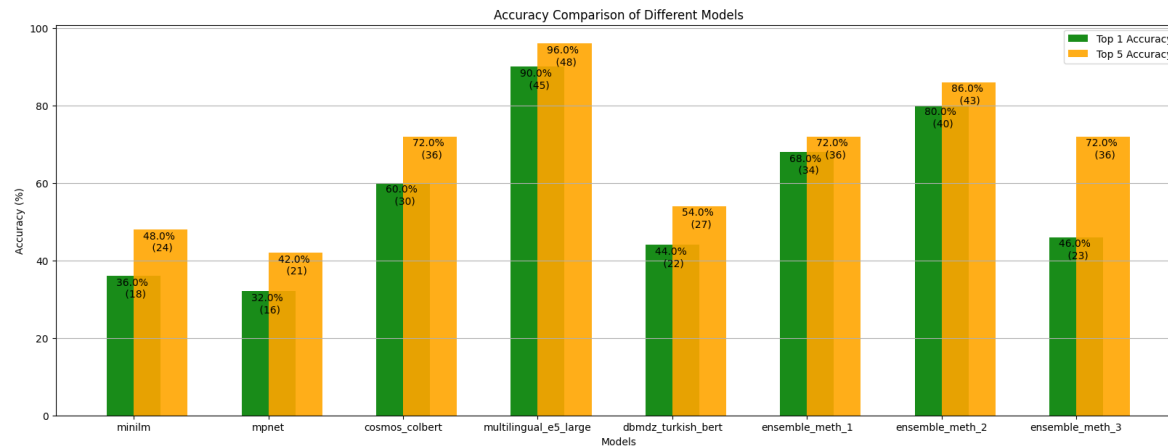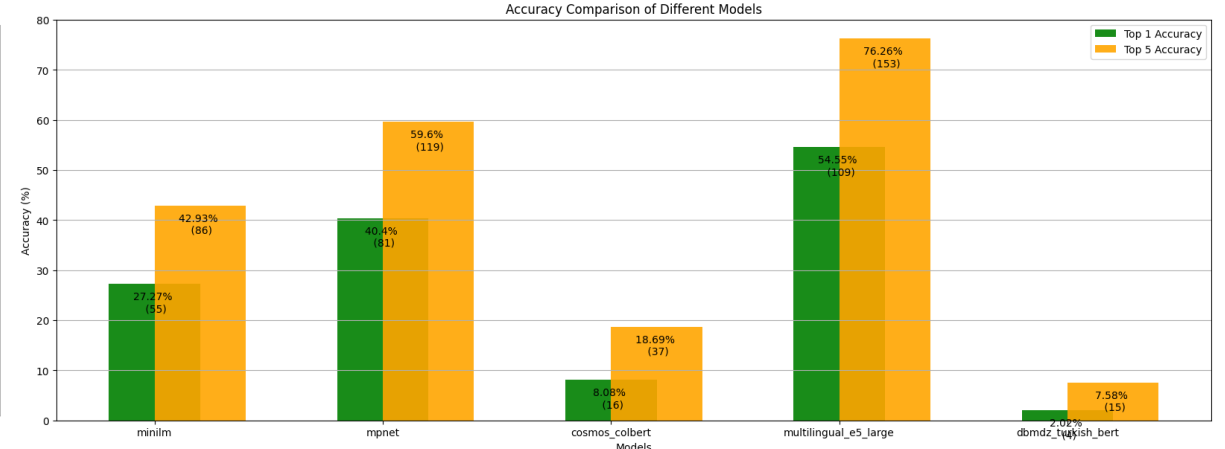
Turkish Instructions Dataset
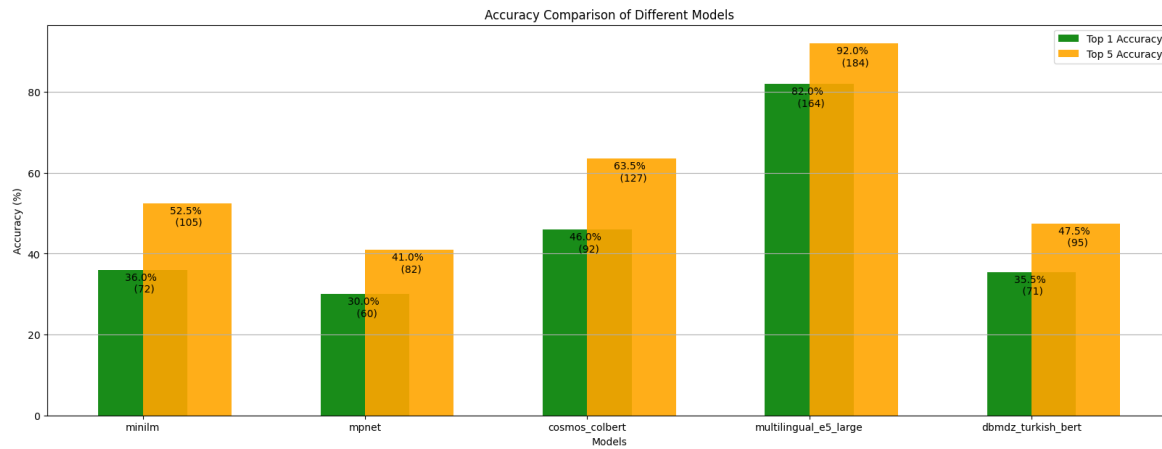
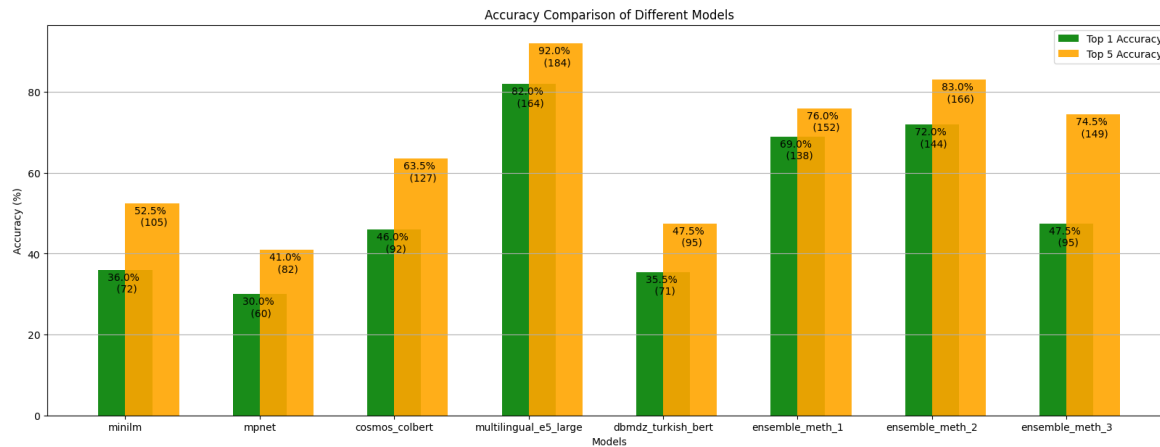## Trivia QA Dataset

Accuracy Comparison with Ensemble for 200 Sample

Turkish Instructions Dataset
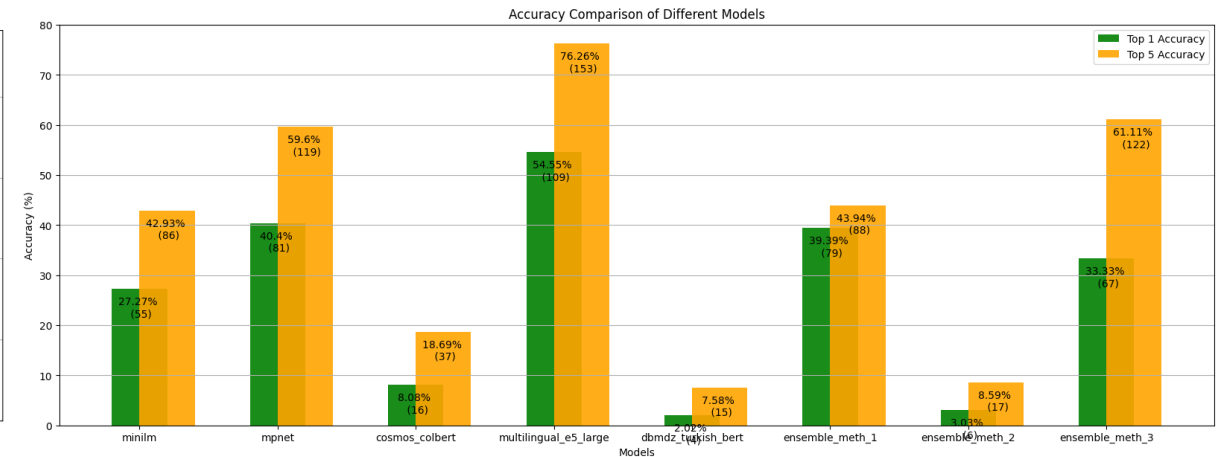
Trivia QA Dataset

Agenda

Embedding Elapsed Times for 2000 Sample

Turkish Instructions Dataset

Elapsed Time for Generating Embeddings by Model
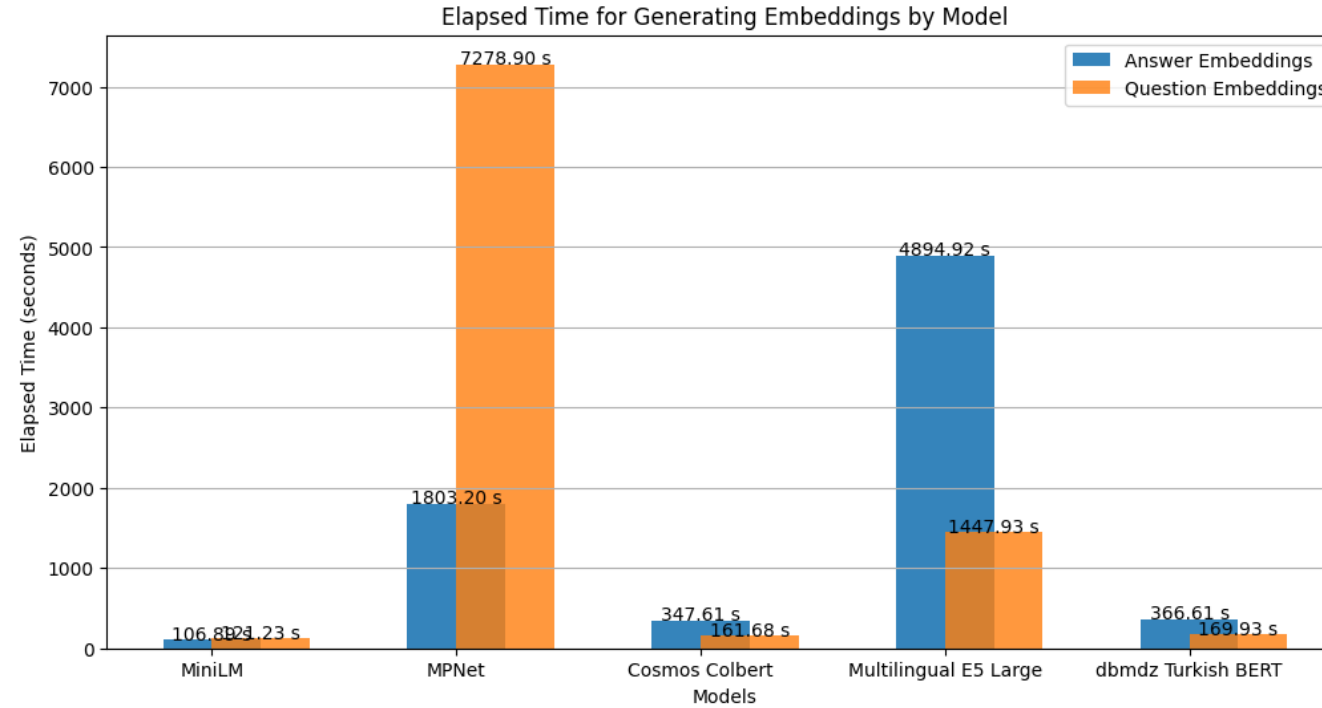
Model	Question (sec.)	Question (min.)	Answer (sec.)	Answer Süresi (min.)
all-MiniLM-L12-v2	121	2	106	1
all-mpnet-base-v2	7278	121	1803	30
ytu-ce-cosmos/turkish-colbert	161	2	347	6
Multilingual-E5-large-instruct	1447	24	4894	81
bert-base-turkish-cased	169	2	366	6
Total	**9176**	**153**	**7516**	**124**

Accuracy Comparison for 2000 Sample
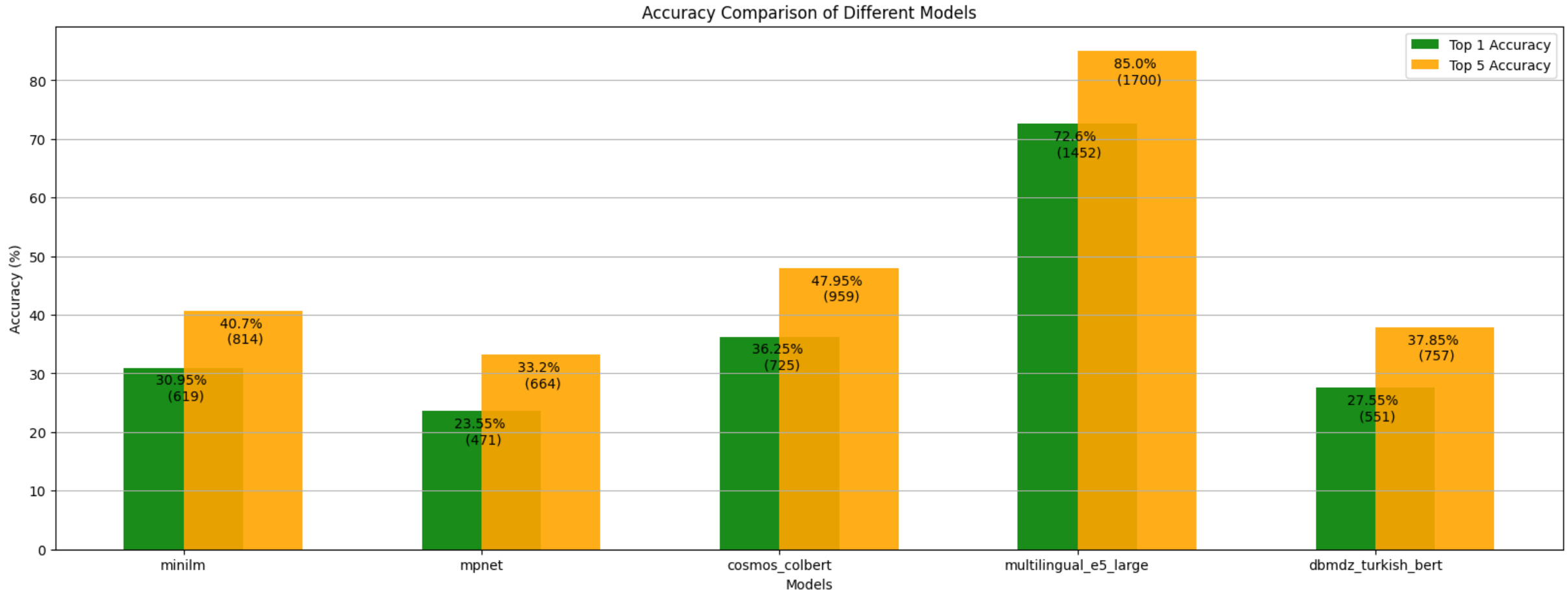
Turkish Instructions Dataset

Accuracy Comparison of Different Models

Accuracy Comparison with Ensemble for 2000 Sample

Turkish Instructions Dataset

Accuracy Comparison of Different Models

Embedding Elapsed Times for 2000 Sample

Trivia QA Dataset

Elapsed Time for Generating Embeddings by Model
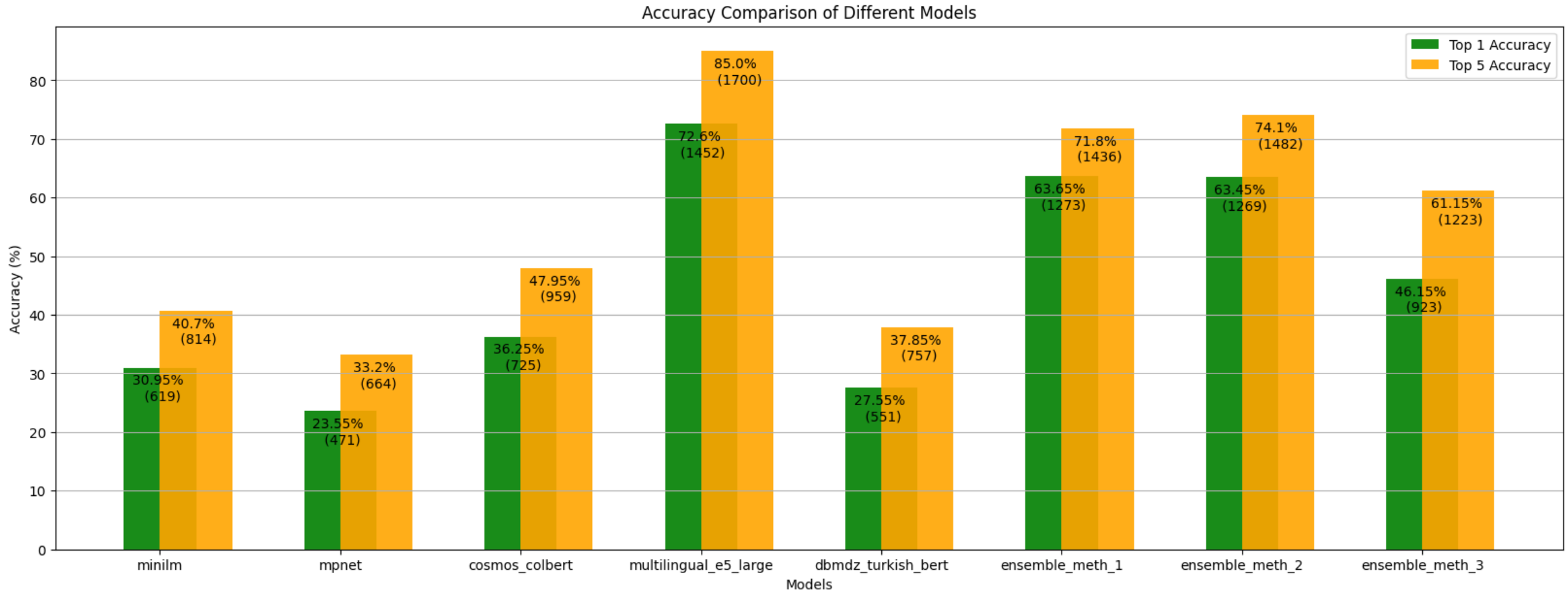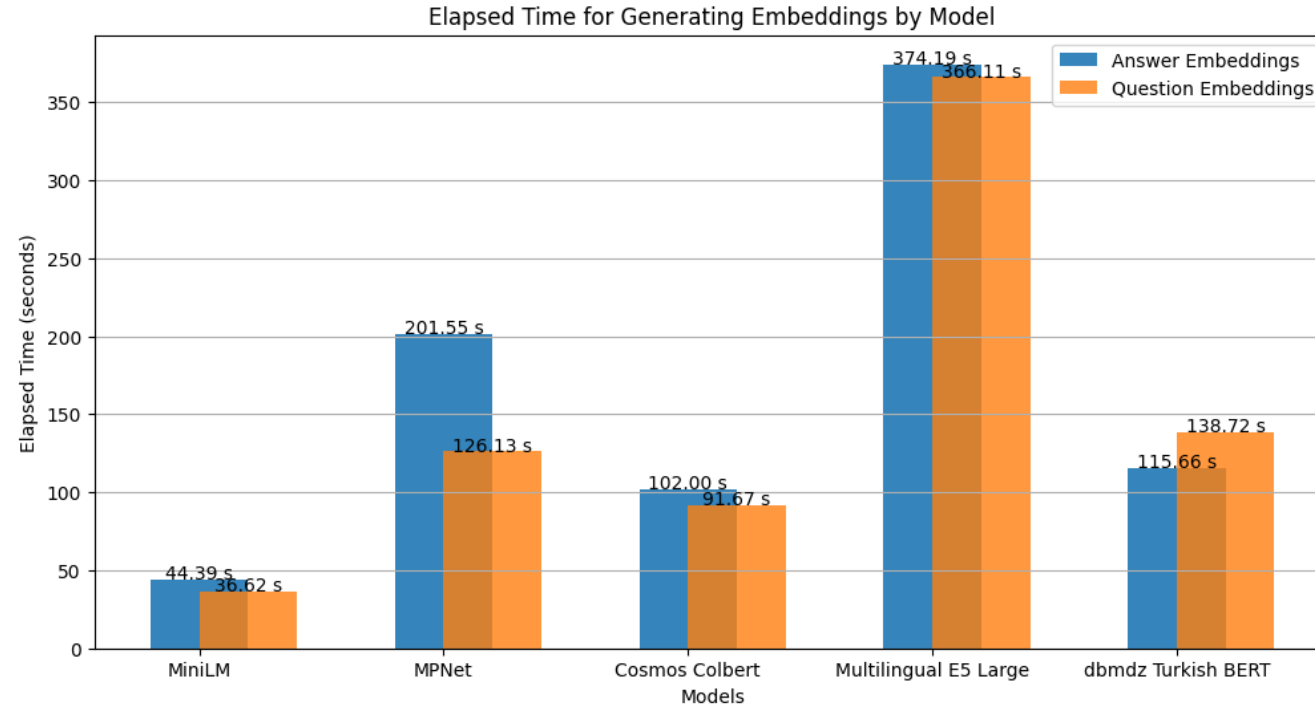
Model	Question (sec.)	Question (min.)	Answer (sec.)	Answer Süresi (min.)
all-MiniLM-L12-v2	36	0	44	0
all-mpnet-base-v2	126	2	201	3
ytu-ce-cosmos/turkish-colbert	91	1	102	1
Multilingual-E5-large-instruct	366	6	374	6
bert-base-turkish-cased	138	2	115	1
Total	**757**	**12**	**836**	**13**

Accuracy Comparison for 2000 Sample

Trivia QA Dataset
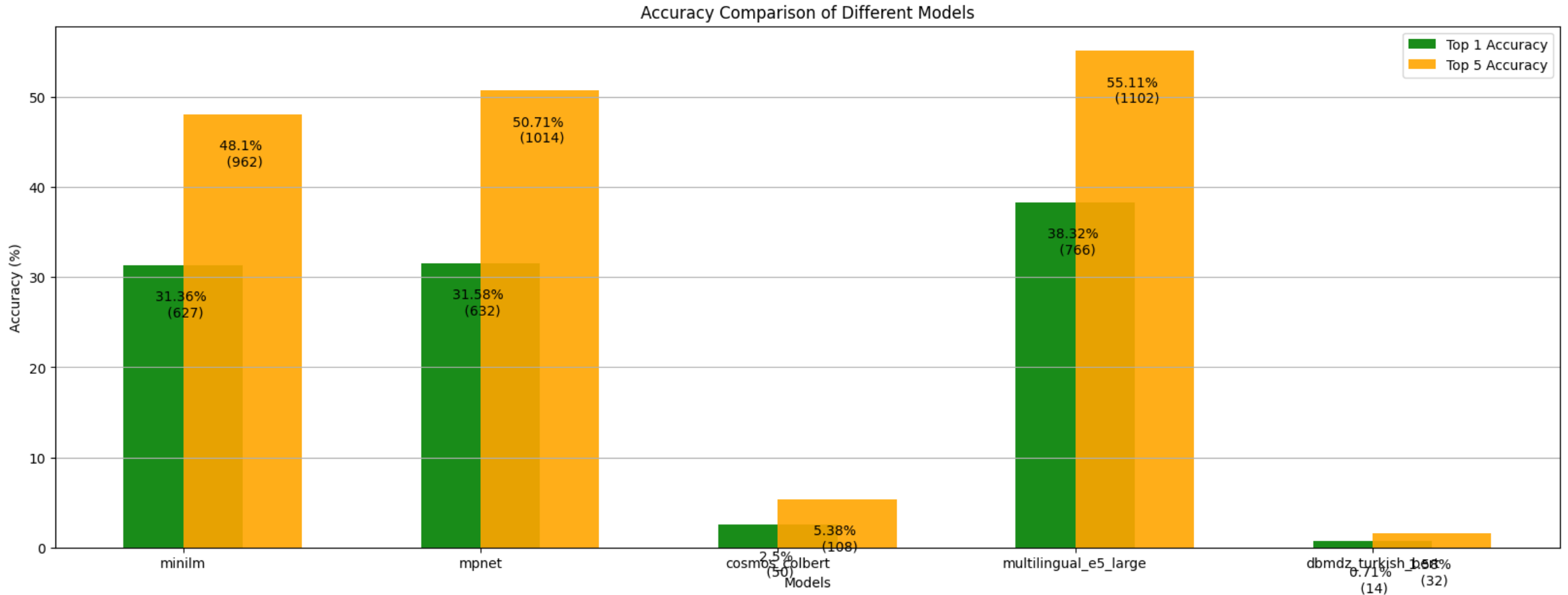
Accuracy Comparison of Different Models

Accuracy Comparison with Ensemble for 2000 Sample

Trivia QA Dataset

Accuracy Comparison of Different Models
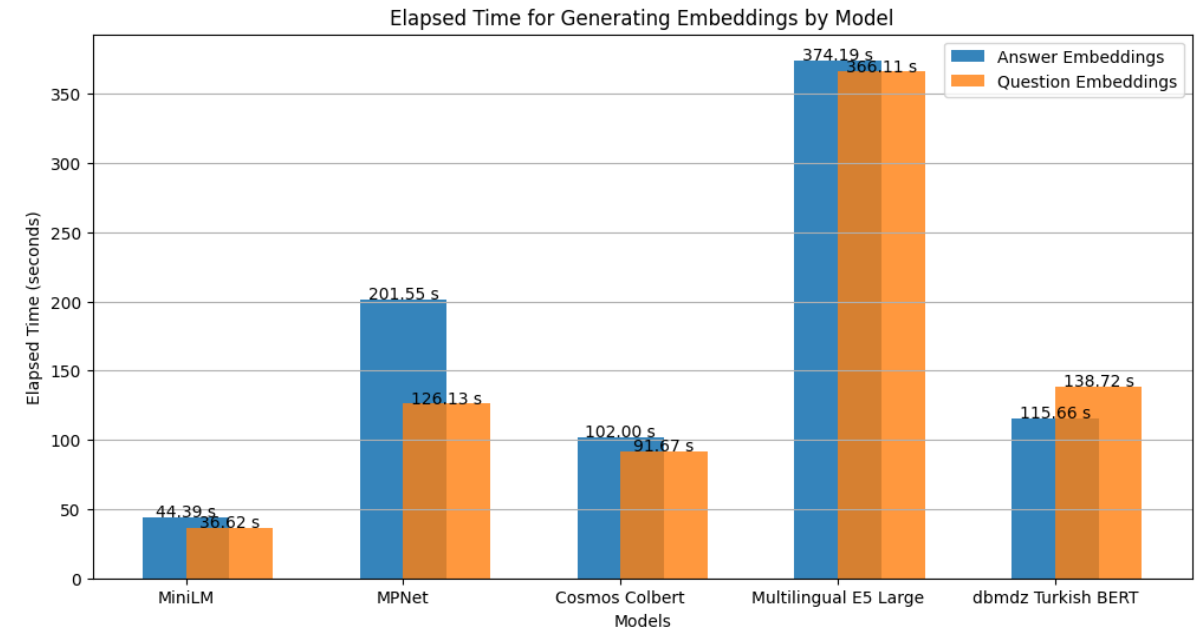
Embedding Elapsed Times for 2000 Sample
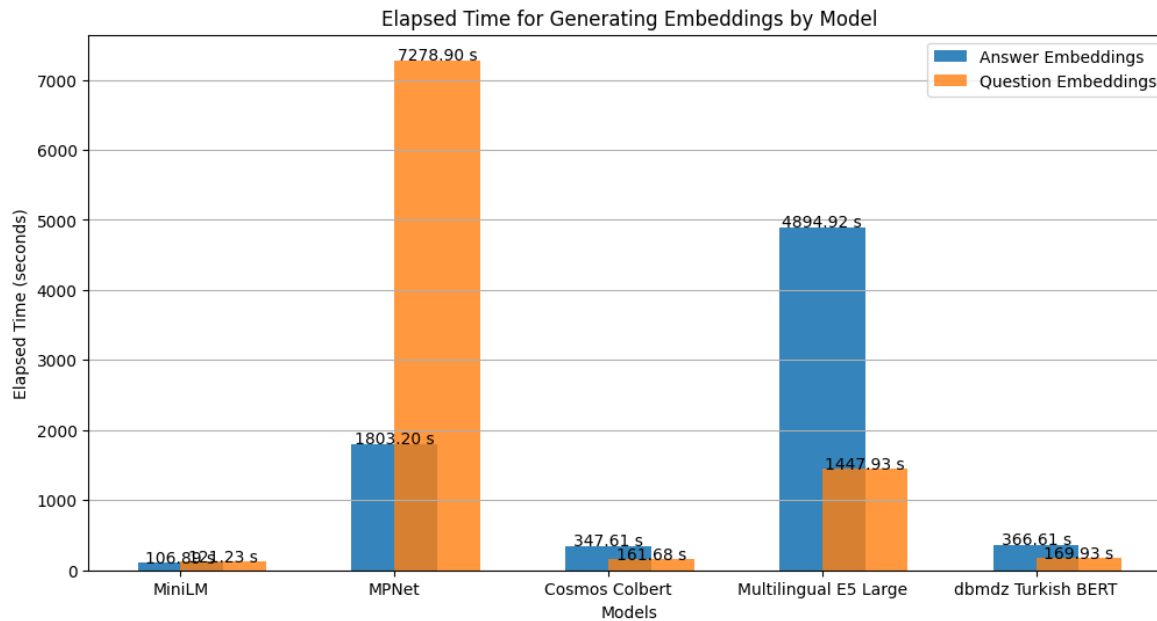
Turkish Instructions Dataset
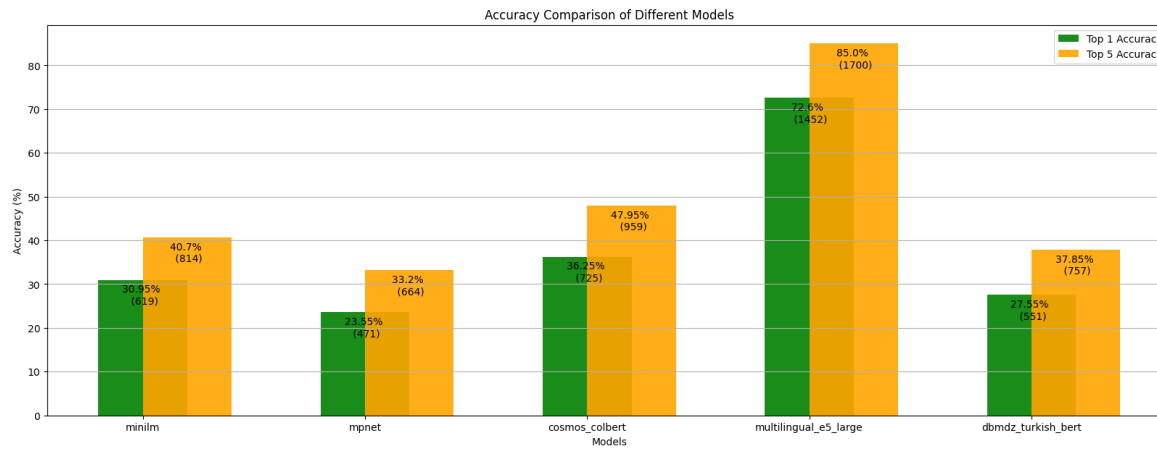

Trivia QA Dataset
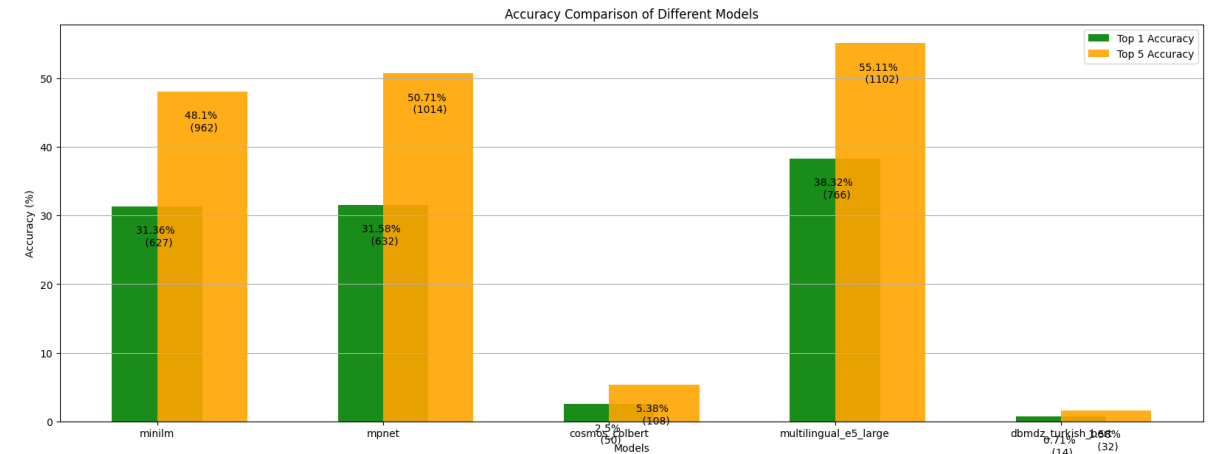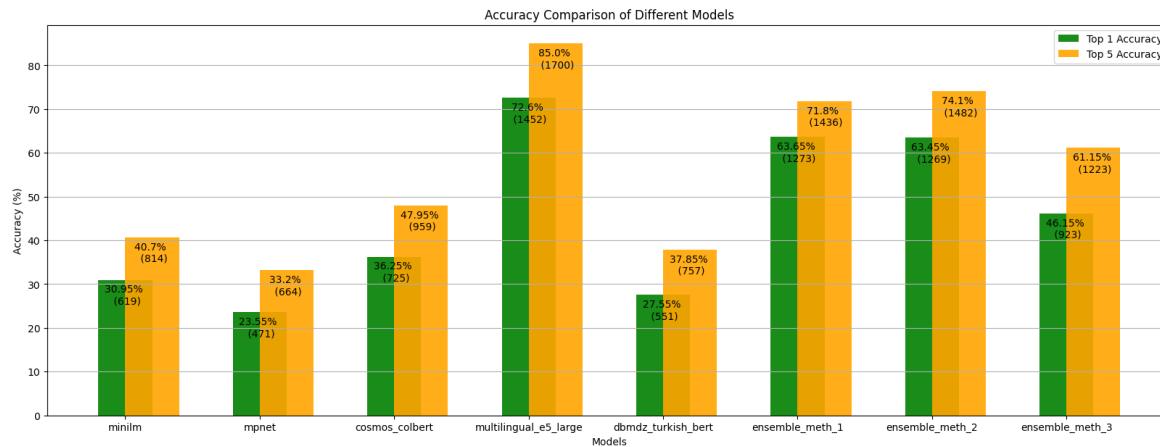
Accuracy Comparison for 2000 Sample

Turkish Instructions Dataset

Trivia QA Dataset

Accuracy Comparison with Ensemble for 2000 Sample

Turkish Instructions Dataset

Trivia QA Dataset

Agenda

Conclusions

- En başarılı model multilingual-e5-large

- Her model her dataset için uygun değildir.

  - cosmos/turkish-colbert

  - turkish-bert

- Veri seti ön işleme modelin çalışabilmesi / performansı için önemlidir.

- Ensemble yönteminin başarısı veri seti ve modellere bağlıdır.

Agenda

References

[1] turkish_instructions

[2] trivia_qa

[3] all-MiniLM-L12-v2

[4] all-mpnet-base-v2

[5] turkish-colbert

[6] multilingual-e5-large-instruct

[7] bert-base-turkish-cased

Agenda

Q & A

THANK YOU FOR LISTENING...