



**YILDIZ TEKNİK ÜNİVERSİTESİ
KİMYA-METALÜRJİ FAKÜLTESİ
MATEMATİK MÜHENDİSLİĞİ BÖLÜMÜ**

BİTİRME ÇALIŞMASI

**Mistral ve Llama Dil Modellerinin Müşteri Memnuniyeti Tespit
Etmedeki Başarılarının Karşılaştırılması**

Tez Yöneticisi: Dr. Mert BAL

20052021
Ömer Gürbüz

İstanbul, 2024

İÇİNDEKİLER

Özet	iv
Abstract	v
Giriş	1
1- Projenin Amacı ve İçeriği	2
2- Projede Kullanılan Modeller	3
2.1- Mistral	3
2.2- Llama	4
2.3- Translator	5
3- Projede Kullanılan Müşteri Yorumları	6
4- Proje Kodları	7
4.1 Kütüphaneler	7
4.2 Mistral	7
4.3 Llama	8
4.4 Llama Template	8
4.5 Mistral Template	9
4.6 Translator	9
4.7 Model	10
4.8 Örnek Çıktı Yorumları	11
4.8.1 Mistral Çıktı Örneği	12
4.8.2 Llama Çıktı Örneği	12
5- Proje Çıktıları	13
6- Sonuçlar ve Öneriler	22
Kaynakça	23
Ekler	24

ÖZET

Yapay zeka günümüzde gitgide daha yaygın şekilde kullanılmaya başlanmıştır. Yapay zekanın hayatımıza girmesindeki en önemli faktör büyük dil modelleridir. Tezde iki farklı dil modelinin müşteri yorumlarından hangi ölçütte çıkarım sağladıkları ve hangi sürede sonuç verdikleri işlenmiştir.

Projede kullanılan müşteri yorumları herhangi bir yasal problem çıkmaması amacıyla ChatGPT kullanılarak üretilmiştir. Üretilen yorumlar çelişki olmaması amaçlanarak tek tek kontrol edilmiş ve gerekli düzeltmeler yapılmıştır. Projede 10 farklı ürüne, 3'er pozitif 3'er negatif olmak üzere toplam 60 adet yorum kullanılmıştır. Amaç ilgili dil modellerinin müşterilerin bu ürünlerin hangi özelliklerinden memnun kaldıklarını, hangi özelliklerinden ise memnun kalmadıklarını tespit etmek ve ilgili dil modellerinden hangisinin daha tutarlı sonuç çıkardığını gözlemlemektir.

Kullanılan yorumlar ele alındığında Mistral modeli Llama'ya göre daha tutarlı ve detaylı sonuçlar çıkarmıştır. Llama'nın daha çok yorumları pozitif ve negatif olacak şekilde direkt çıktı olarak vermesine karşın Mistral, çıktı olarak yorumlarda bahsedilen özellikleri vermiştir. Yeniden kullanılan yorumlar ele alındığında Llama modelinin çok küçük bir farkla Mistral modelinden daha hızlı çalıştığı gözlemlenmiştir.

ABSTRACT

Artificial intelligence is increasingly being utilized in various fields today, and one of the most significant factors contributing to its integration into our lives is the development of large language models. In the thesis, the inference capabilities and processing speeds of two different language models from customer reviews were analyzed.

The customer reviews used in the project were generated using ChatGPT to avoid any legal issues. Each generated comment was individually checked to ensure consistency and necessary corrections were made. A total of 60 reviews were used, with 3 positive and 3 negative comments for each of the 10 different products. The aim was to determine which features of these products customers were satisfied with and which ones they were not, and to observe which language model produced more consistent results.

When considering the reviews used, the Mistral model provided more consistent and detailed results compared to Llama. While Llama tended to directly output the comments as positive or negative, Mistral provided the mentioned features in the comments as output. Upon reusing the comments, it was observed that the Llama model worked slightly faster than the Mistral model.

GİRİŞ

Günümüzde bir yapay zekâ devrimi yaşanıyor demek yanlış olmaz. Bu yapay zekâ devriminin en büyük öncüleri ise büyük dil modelleridir (LLM). Büyük dil modelleri gerek kodlama, gerek pazarlama, gerekse tıpta ve aklınıza gelebilecek her türlü alanda insanoğluna büyük katkılar ve kolaylıklar sağlamaktadır. Fakat bu teknolojinin yeni olması sebebiyle bu dil modellerinin nasıl daha verimli bir şekilde kullanılabileceği hakkında hala çalışmalar devam etmektedir. Bu dil modellerinden en ünlüsü Openai adlı şirketin geliştirdiği ChatGPT'dir.

Bu çalışmada ise iki farklı dil modelinin, gerçek müşteri yorumları kullanarak bu müşterilerin ilgili ürünün hangi özelliklerinden memnun kaldıklarını veya memnun kalmadıklarını nasıl bir başarı oranı ile tespit etmesi karşılaştırılacaktır. Ayrıca Mistral ve Llama dil modelleri hakkında bilgiler verilecektir. Projede kullanılan kodlar ve çıktı örneği tez yazısının sonuna eklenmiştir. Ayrıca ilgili çalışmanın başarı sonuçları Excel ortamında rapor haline getirilmiştir. Bu rapor ilerleyen sayfalarda detaylı bir şekilde incelenmiş ve sonuçlar olabildiğince anlaşılır şekilde açıklanmıştır.

1- PROJENİN AMACI VE İÇERİĞİ

Bu projede de iki adet LLM kullanılarak müşteri yorumlarından bir ürün hakkında çıkarımlar yapılmaya çalışılacak ve bu iki farklı model birbirleriyle çalışma hızları ve başarı konusunda karşılaştırılacaktır. Projede kullanılan modellerin hepsi “huggingface.co” sitesinden çekilmiştir. Totalde kullanılan modeller Llama, Mistral ve 2 adet çeviri modeli olmak üzere toplam 4 tanedir. Elde edilen sonuçlar Excel ortamında rapor haline getirilmiştir.

2 - PROJEDE KULLANILAN MODELLER

2.1– Mistral

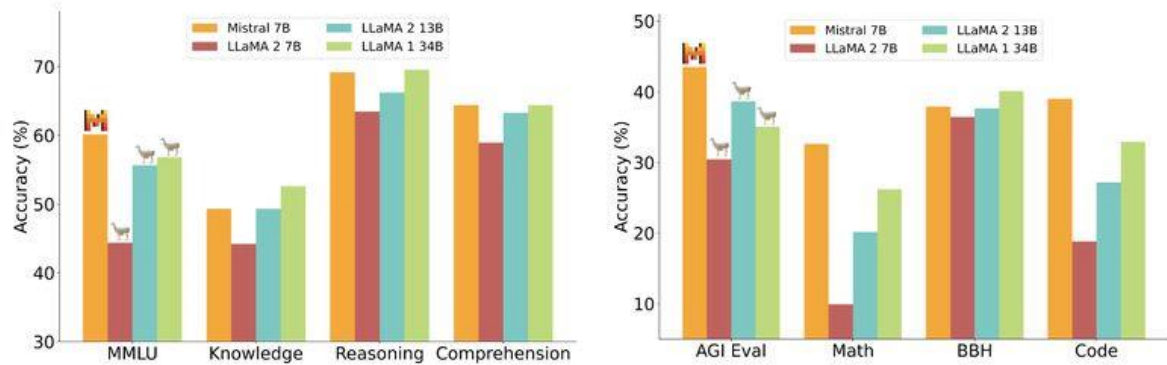
Projede kullanılan model Mistral7B (7B 7 milyar parametrelili bir model olduğu anlamına geliyor) büyük dil modelinin quantize edilmiş versiyonudur. Bu dil modelinin quantize edilmiş versiyonunu kullanmaktaki amaç GPU'ya binen yükü olabildiğince azaltmaktır. Proje Google-Colab ortamında yazıldığından ve ücretsiz sürümde GPU 15GB ile sınırlı olduğundan ötürü Mistral7B büyük dil modeli quantize edilmeden kullanılamamaktadır. Projede kullanılan model Mistral7B-OpenOrca-GPTQ'dur. Bu model yaklaşık 1,2 milyar parametreye sahiptir ve GPU'ya yükü yaklaşık 4GB kadardır. Python'da Transformers kütüphanesi altında kullanılabilir.

2.2 - LLaMA

Projede kullanılan model LLaMA7B büyük dil modelinin quantize edilmiş versiyonudur. 1.1 bölümünde anlatılan sebeplerden ötürü LLaMA büyük dil modelinin quantize edilmiş versiyonu kullanılmaktadır. Bu model yaklaşık 1,13 milyar parametreye sahiptir ve GPU'ya yükü yaklaşık 4GB kadardır. Python'da Transformers kütüphanesi altında kullanılabilir.

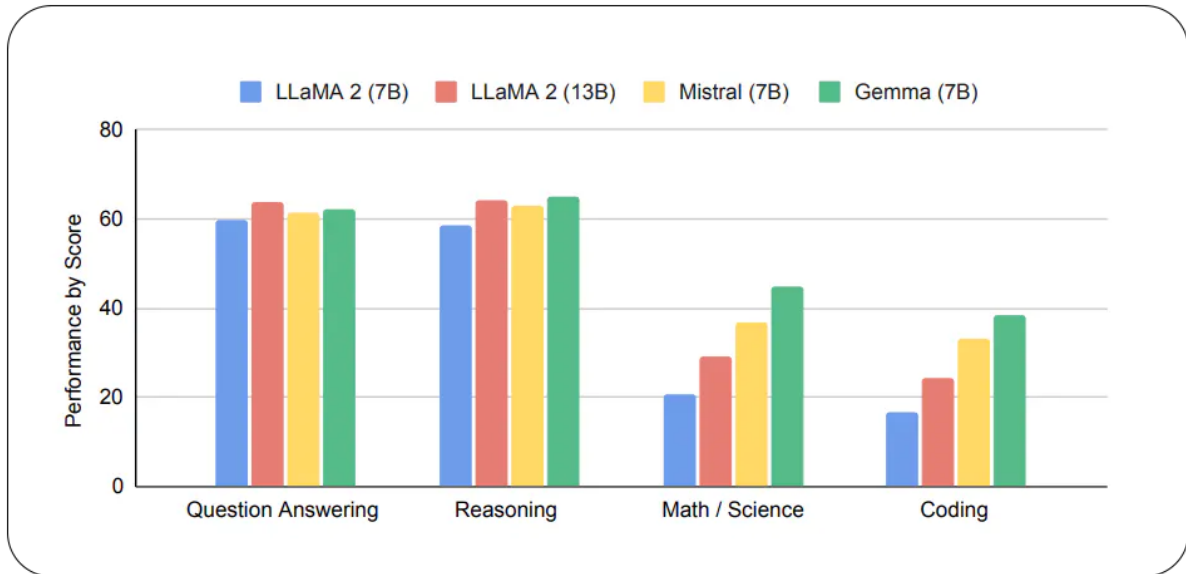
Performance in details

We compared Mistral 7B to the Llama 2 family, and re-run all model evaluations ourselves for fair comparison.



Performance of Mistral 7B and different Llama models on a wide range of benchmarks. For all metrics, all models were re-evaluated with our evaluation pipeline for accurate comparison. Mistral 7B significantly outperforms Llama 2 13B on all metrics, and is on par with Llama 34B (since Llama 2 34B was not released, we report results on Llama 34B). It is also vastly superior in code and reasoning benchmarks.

Görsel 2.2.1



Görsel 2.2.2

BigBench-Hard	Llama2 7B	Mistral-7B	Mistral-7B-Instruct	OpenOrca Preview2 (13B)	Mistral-7B-OpenOrca	% of Mistral-7B-Instruct Performance	% of Llama2 7B Performance
causal_judgement	0.521	0.5263	0.5263	0.590	0.6	114.0%	115.1%
date_understanding	0.629	0.664	0.664	0.724	0.6531	98.4%	103.9%
disambiguation_qa	0.310	0.3566	0.3411	0.306	0.5233	153.4%	168.8%
geometric_shapes	0.089	0.1003	0.1003	0.137	0.2006	200.0%	225.1%
logical_deduction_five_objects	0.216	0.24	0.278	0.310	0.258	92.8%	119.4%
logical_deduction_seven_objects	0.150	0.1714	0.2114	0.233	0.1943	91.9%	129.5%
logical_deduction_three_objects	0.363	0.3633	0.46	0.527	0.4167	90.6%	114.7%
movie_recommendation	0.236	0.326	0.344	0.344	0.346	100.6%	146.6%
navigate	0.511	0.5	0.534	0.522	0.527	98.7%	103.1%
reasoning_about_colored_objects	0.338	0.452	0.652	0.670	0.664	101.8%	196.7%
ruin_names	0.250	0.3281	0.3214	0.357	0.4174	129.9%	167.0%
salient_translation_error_detection	0.189	0.2044	0.2385	0.305	0.1593	66.8%	84.1%
snarks	0.530	0.5414	0.7293	0.674	0.7017	96.2%	132.3%
sports_understanding	0.509	0.497	0.6197	0.588	0.7089	114.4%	139.2%
temporal_sequences	0.258	0.272	0.235	0.442	0.327	139.1%	126.7%
tracking_shuffled_objects_five_objects	0.201	0.228	0.2192	0.196	0.2144	97.8%	106.8%
tracking_shuffled_objects_seven_objects	0.155	0.1514	0.1509	0.138	0.152	100.7%	97.8%
tracking_shuffled_objects_three_objects	0.363	0.3633	0.46	0.527	0.4167	90.6%	114.7%
Average	0.32329	0.349	0.394	0.422	0.416	105.6%	128.5%

Görsel 2.2.3

2.3 - Translator (Helsinki-NLP)

Projede translator kullanılmasının sebebi Llama ve Mistral modellerinin Türkçe dilini desteklememesinden kaynaklanmaktadır. Bu yüzden Türkçe yorumlar önce İngilizce 'ye, modelin çıkardığı İngilizce cevaplar ise Türkçe diline çevrilmiştir. Yani sonuçta toplamda 2 adet çeviri modeli kullanılmıştır. Her ne kadar çeviri hatalarından kaynaklı çeşitli hatalar meydana gelse de ilgili çeviri modeli iyi çalışmaktadır. İki model yaklaşık 2-4GB GPU ram kullanmaktadır.



Görsel 2.3.1

3 – PROJEDE KULANILAN MÜŞTERİ YORUMLARI

Projede kullanılan müşteri yorumları herhangi bir şirketle problem yaşanmaması adına ChatGPT kullanılarak oluşturulmuştur. Oluşturulan her yorum içeriği bakımından kontrol edilmiş gereken yerde el ile düzeltmeler yapılmıştır. Projede, 10 farklı ürüne 3 olumlu 3 olumsuz olmak üzere toplam 60 adet yorum kullanılmıştır.

```
[[['Harika bir ses kalitesi ve kablosuz özgürlük! Kesinlikle tavsiye ederim.',  
'Çok rahat, hafif ve pratik. Spor yaparken kullanmak için mükemmel bir seçim.',  
'Bağlantı bazen kesiliyor, özellikle yoğun yerlerde kullanırken sıkıntı yaşanabiliyor.',  
'Şarj süresi beklediğimden kısa, sık sık şarj etmek zorunda kalıyorum.',  
'Kolay bağlantı ve kullanım, mükemmel gürültü engelleme. Bu fiyata çok iyi bir ürün.',  
'Kulak içine yerleştirme konusunda biraz rahatsız edici, uzun süre kullanımda ağrı yapabiliyor.'],  
'Hızlı işlemci, etkileyici kamera ve şık tasarımı. En iyi telefonlardan biri!',  
'Yüksek performans ve dayanıklılık, her türlü uygulama ve oyunu sorunsuz çalıştırıyor.',  
'Fiyatı biraz yüksek, daha uygun fiyatlı alternatifler tercih edilebilir.',  
'Mükemmel ekran kalitesi, uzun pil ömrü ve akıcı kullanıcı arayüzü. Kesinlikle öneririm.',  
'Bazı kullanıcılar arasında yazılım güncellemeleriyle ilgili sorunlar yaşanabiliyor.',  
'Parmak izi okuyucu bazen çalışmıyor, bu da güvenlik açısından endişe verici olabiliyor.'],  
'Takılıp kaldığı yerler olabiliyor, bazen manuel müdahale gerektirebiliyor.',  
'Ses seviyesi biraz yüksek, sessiz çalışan bir model tercih edilebilir.',  
'Evinizi temizlemek için harika bir yardımcı, akıllı navigasyon ve güçlü emiş gücü.',  
'Programlanabilir ve uzaktan kontrol edilebilir özellikleriyle hayatı kolaylaştırıyor.',  
'Köşe ve kenarları temizlemekte çok etkili, evcil hayvan sahipleri için ideal bir seçim.',  
'Pil ömrü beklediğimden kısa, büyük alanları temizlemek için birden fazla şarj gerekebiliyor.'],  
'Bazı kullanıcılar arasında Bluetooth bağlantı sorunları yaşanabiliyor, cihazlar arasında iletişim kopuklukları olabiliyor.',  
'Ekran boyutu biraz küçük, bazı kullanıcılar için bilgi okunması zor olabilir.',  
'Yüzeyi çabuk çiziliyor, ekranı koruyucu kullanmak gerekebiliyor.',  
'Sağlık takibi, spor aktiviteleri ve bildirimler için mükemmel bir yardımcı, şık tasarımıyla da dikkat çekiyor.',  
'Uzun pil ömrü, suya dayanıklılık ve geniş uygulama desteği. Herkesin sahip olması gereken bir aksesuar.',  
'Uygun fiyatlı olmasına rağmen kaliteli malzeme ve fonksiyonlara sahip, performansıyla şaşırtıyor.'],  
'Fiyatı biraz yüksek, bütçesi kısıtlı olanlar için alternatif arayışı gerekebilir.',  
'Yüksek performans, uzun pil ömrü ve şık tasarımı. Hem iş hem de eğlence için mükemmel bir seçim.',  
'Fan sesi bazen rahatsız edici olabiliyor, sessiz çalışan bir model tercih edilebilir.',  
'Hızlı açılış ve tepki verme süresi, çoklu görevleri rahatça yönetmek için ideal.',  
'Isınma sorunu yaşayan kullanıcılar var, uzun süreli kullanımlarda cihazın sıcaklığı artabiliyor.',  
'Taşınabilirlik ve güçlü donanım kombinasyonu en iyi dizüstü bilgisayarlar arasında yer alıyor.'],  
'Mükemmel kahve demleme performansı, kolay temizlik ve şık tasarımı. Sabahları benim için vazgeçilmez!',  
'Farklı kahve çeşitlerini kolayca hazırlama imkanı, kullanımı da oldukça pratik.',  
'Büyük su haznesi ve otomatik kapanma özelliği ile güvenli ve kullanışlı bir ürün.',  
'Bazı modellerde su sızıntısı problemi yaşanabiliyor, dikkatli olmak gerekiyor.',  
'Filtre değişim maliyetleri zamanla artabiliyor, bu da bakım masraflarını artırıyor.',  
'Sıcaklık kontrolü biraz zor, bazen istenilen kahve sıcaklığına ulaşmak mümkün olmayabiliyor.'],  
'Şarj süresi biraz uzun, beklemek gerekebiliyor. Daha hızlı şarj eden modeller tercih edilebilir.',  
'Bazı modellerde bağlantı sorunları yaşanabiliyor, cihazlar arasında iletişim kopuklukları olabiliyor.',  
'Hızlı şarj özelliği ve yüksek kapasitesi ile hayat kurtarıcı bir ürün. Her zaman yanımda taşıyorum!',  
'Kompakt tasarımı ve hafif yapısıyla taşınması kolay, seyahatlerde büyük bir avantaj sağlıyor.',  
'Birden fazla cihazı aynı anda şarj edebilme özelliği, çoklu kullanıcılar için ideal.',  
'Dayanıklılık konusunda sorunlar yaşanabiliyor, düşmelere karşı daha sağlam bir yapı tercih edilebilir.'],  
'Mükemmel ses kalitesi, şık tasarım ve kablosuz bağlantı özelliği ile her yerde kullanılabilir.',  
'Pil ömrü biraz kısa, uzun süreli kullanımlarda sık sık şarj etmek gerekiyor.',  
'Bazı kullanıcılar arasında bağlantı kopmaları yaşanabiliyor, stabil bir bağlantı için daha güçlü sinyal gerekiyor.',  
'Bass performansı biraz zayıf, daha güçlü bir bas isteyenler için yeterli olmayabilir.',  
'Taşınabilirlik açısından ideal, pikniklerde, plajda veya ev partilerinde vazgeçilmezimiz haline geldi.',  
'Su geçirmez özelliği sayesinde dış mekanlarda kullanımı sorunsuz, dayanıklı bir yapıya sahip.'],  
'Bazı modellerde ses seviyesi biraz yüksek, çalışırken rahatsız edici olabiliyor.',  
'Hızlı pişirme özelliği, geniş iç hacim ve kullanım kolaylığı ile mutfağın vazgeçilmezi oldu.',  
'Farklı pişirme programları ve hassas ayarlar sayesinde her yemeği mükemmel şekilde pişirebiliyorum.',  
'Kolay temizlenebilir iç yüzeyi ve şık dış tasarımı ile mutfakta modern bir hava katıyor.',  
'Kapı kolu veya düğmelerde zamanla aşınma ve arıza sorunları yaşanabiliyor.',  
'Isı dağılımı konusunda bazı kullanıcılar arasında şikayetler var, homojen pişirme sağlamak için dikkatli olmak gerekiyor.'],  
'Kusursuz temizlik, enerji ve su tasarrufu ile hayatı kolaylaştıran bir ürün. Mutlaka edinilmeli!',  
'Farklı program seçenekleri ve sessiz çalışma özelliği ile kullanımı çok pratik.',  
'Bazı modellerde su damlatma problemi yaşanabiliyor, bu da mutfakta ekstra temizlik gerektirebiliyor.',  
'Geniş iç hacmi sayesinde büyük tabak ve tencereleri bile sorunsuz temizleyebiliyor.'],
```

Görsel 2.1

4 - PROJE KODLARI

4.1 – Kütüphaneler

```
!pip install accelerate
!pip install transformers optimum
!pip install auto-gptq
!pip install transformers
!pip install sacremoses

import time
import torch
import torch.nn as nn
from transformers import AutoModelForCausalLM, AutoTokenizer, pipeline
```

4.2 - Mistral

```
class GenerativeModel_Mistral:
    def __init__(self):
        self.path = "TheBloke/Mistral-7B-OpenOrca-GPTQ"
        self.generative_model = AutoModelForCausalLM.from_pretrained(self.path,
                                                                    device_map="auto")

        self.tokenizer = AutoTokenizer.from_pretrained(self.path)

    def generate(self, input_text, max_new_tokens=256):

        s = time.time_ns()
        inputs = self.tokenizer(input_text, max_length=512, truncation=True, return_tensors='pt').to('cuda')

        with torch.no_grad():
            out = self.generative_model.generate(**inputs,
                                                max_new_tokens=max_new_tokens)
        out_str = self.tokenizer.decode(out[0], skip_special_tokens = True)

        e = time.time_ns()
        return {'elapsed_time': ((e-s)//1000000)/1000, 'output': out_str}
```

4.3 - Llama

```
class GenerativeModel_Llama:
    def __init__(self):
        self.path = "TheBloke/Llama-2-7b-Chat-GPTQ"
        self.generative_model = AutoModelForCausalLM.from_pretrained(self.path,
                                                                    device_map="auto")

        self.tokenizer = AutoTokenizer.from_pretrained(self.path)

    def generate(self, input_text, max_new_tokens=256):

        s = time.time_ns()
        inputs = self.tokenizer(input_text, max_length=512, truncation=True, return_tensors='pt').to('cuda')

        with torch.no_grad():
            out = self.generative_model.generate(**inputs,
                                                max_new_tokens=max_new_tokens)

        out_str = self.tokenizer.decode(out[0], skip_special_tokens = True)

        e = time.time_ns()
        return{'elapsed_time': ((e-s)//1000000)/1000, 'output': out_str}
```

4.4 Llama Template

```
class Template_llama:
    template = \
    """
    <<[INST] <<SYS>>
    Find out what the customers are satisfied and dissatisfied with based on the content of the text group.
    Just separate positive and negative things based on the text group.
    Each comment in the text group belongs to a different customer.
    Don't explain anything.

    You should write like that:
    Positive
    * <|positive_thing|>
    * <|positive_thing|>

    Negative
    * <|negative_thing|>
    * <|negative_thing|>

    <</SYS>>
    Tell me what is negative and what is positive in the text group.
    Comments: {}

    [/INST]
    Answer:
    """
```

4.5 – Mistral Template

```
class Template_mistral:
    template = \
    """
    <s><|im_start|>system
    Find out what the customers are satisfied and dissatisfied with based on the content of the text group.
    Just separate positive and negative things based on the text group.
    Each comment in the text group belongs to a different customer.
    Don't explain anything.

    Here is an example for answer format:
    "
    Positive
    * <|positive_thing|>
    * <|positive_thing|>

    Negative
    * <|negative_thing|>
    * <|negative_thing|>
    "

    <|im_start|>user
    Tell me what is negative and what is positive in the text group.
    Comments: {}<|im_end|>

    <|im_start|>assistant
    Answer: ""
```

4.6 - Translator

```
from transformers import MarianMTModel, MarianTokenizer

class Translator:

    def __init__(self):

        access_token = "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"

        self.tr_en_model_name = "Helsinki-NLP/opus-mt-tc-big-tr-en"
        self.tr_en_tokenizer = MarianTokenizer.from_pretrained(self.tr_en_model_name, token=access_token)
        self.tr_en_model = MarianMTModel.from_pretrained(self.tr_en_model_name, token=access_token)
        self.tr_en_model.to('cuda')

        self.en_tr_model_name = "Helsinki-NLP/opus-mt-tc-big-en-tr"
        self.en_tr_tokenizer = MarianTokenizer.from_pretrained(self.en_tr_model_name, token=access_token)

        self.en_tr_model = MarianMTModel.from_pretrained(self.en_tr_model_name, token=access_token)
        self.en_tr_model.to('cuda')

    def tr_to_en(self, tr_str):
        s = time.time_ns()
        tokenized = self.tr_en_tokenizer(tr_str, return_tensors="pt", padding=True, truncation=True).to('cuda')
        out = self.tr_en_model.generate(**tokenized)
        translated = self.tr_en_tokenizer.decode(out[0], skip_special_tokens=True)

        e = time.time_ns()
        return {'input_translation_time': ((e-s)//1000000)/1000, 'output': translated}

    def en_to_tr(self, en_str):
        s = time.time_ns()
        tokenized = self.en_tr_tokenizer(en_str, return_tensors="pt", padding=True, truncation=True).to('cuda')
        out = self.en_tr_model.generate(**tokenized)
        translated = self.en_tr_tokenizer.decode(out[0], skip_special_tokens=True)

        e = time.time_ns()
        return {'elapsed_time_translator': ((e-s)//1000000)/1000, 'output': translated}
```

4.7 - Model

```
class LLMEngine:
    def __init__(self, model=None):
        self.k = 0
        self.model = model
        if not model:
            self.model = GenerativeModel_Mistral()
            self.k = 1
        self.translator = Translator()

    def positive_negative(self, text_list, lang='tr'):

        if lang == 'tr':
            text_list = [self.translator.tr_to_en(text)['output'] for text in text_list]
            ep_list = [self.translator.tr_to_en(text)['input_translation_time'] for text in text_list]
            ep_translator = sum(ep_list)
            lines = '\n* ' + '\n* '.join(text_list)
            lines = lines[:400]
            if self.k == 1:
                template = Template_mistral.template.format(lines)
            else:
                template = Template_llama.template.format(lines)
            out = self.model.generate(template, max_new_tokens=100)
            elapsed_time = out['elapsed_time']
            raw_out = out['output']
            raw_out = raw_out[raw_out.find(' [/INST]')+8:]
            raw_out = raw_out[raw_out.find('Answer: ')+8:]

            if lang == 'tr':
                out_str = self.translator.en_to_tr(raw_out)
            else:
                out_str = raw_out

            return {'elapsed_time_model': elapsed_time,
                    'input_translation_time': ep_translator,
                    'out_translation_time': out_str['elapsed_time_translator'],
                    'total_translation_time': out_str['elapsed_time_translator'] + ep_translator,
                    'output': out_str['output'],
                    'raw_out': raw_out}
```

4.8 - Örnek Çıktı Yorumları

'Bu ürünü aldığım gerçekten memnunum! Kaliteli malzemelerden yapılmış, dayanıklı ve kullanımı çok kolay. Fiyat performans açısından da gayet başarılı.'

'Ürünün beklediğim performansı vermedi maalesef. Kullanımı oldukça karmaşık ve bazı fonksiyonları tam olarak çalışmıyor. Biraz daha iyileştirme yapılması gerektiğini düşünüyorum.'

'Harika bir ürün! Hem tasarımı hem de işlevselliğiyle beni etkiledi. Kullanımı çok kolay ve sağlam bir yapıya sahip. Kesinlikle tavsiye ederim.'

'Ürünün kalitesi beklediğimden düşük çıktı. İlk kullanımdan sonra bazı sorunlar yaşamaya başladım ve garanti kapsamında değişim talep ettim. Daha dikkatli bir şekilde üretim yapılmalı.'

'Bu ürünü almadan önce birçok alternatif araştırdım ve sonunda en iyisi olduğuna karar verdim. Gerçekten doğru bir seçim yapmışım, performansı ve dayanıklılığıyla beni şaşırttı.'

'Ürünü ilk aldığımda gayet memnundum ancak kısa süre sonra bazı sorunlar yaşamaya başladım. Müşteri hizmetleriyle iletişime geçtim ancak sorunum hala çözülmedi. Hayal kırıklığına uğradım.'

'Bu ürün gerçekten kaliteli bir tasarıma sahip. İşlevselliği mükemmel ve kullanımı oldukça kolay. Herkese rahatlıkla tavsiye edebilirim.'

'Ürünün fiyatına göre performansı oldukça düşük. Beklediğim verimi alamadım ve birkaç haftalık kullanımdan sonra bazı fonksiyonları tamamen işlevsiz hale geldi. Dezavantajları göz önünde bulundurarak almanızı önermem.'

'Bu ürünü kullanmaya başladıktan sonra hayatım çok daha kolaylaştı. İhtiyacım olan her şeyi tek bir yerde bulabiliyorum ve performansından oldukça memnunum. Kesinlikle alınması gereken bir ürün.'

'Ürünü aldığımdan beri sürekli sorun yaşıyorum. Hem tasarımı hem de performansı beklediğimden çok daha kötü çıktı. Bu fiyata alınabilecek daha iyi alternatifler olduğunu düşünüyorum.'

4.8.1 - Mistral Çıktı Örneği

```
{'elapsed_time_model': 2.671,  
'input_translation_time': 3.6329999999999996,  
'out_translation_time': 0.264,  
'total_translation_time': 3.8969999999999994,  
'output': 'Pozitif * kaliteli malzemeler * dayanıklı * kullanımı kolay * performans açısından başarılı fiyat Negatif * beklenen performansı vermedi * kullanımı karmaşık * bazı fonksiyonlar tam olarak çalışmıyor * iyileştirme gerekli',  
'raw_out': '\n Positive\n * quality materials\n * durable\n * easy to use\n * successful price in terms of performance\n\n Negative\n * did not give the expected performance\n * complex to use\n * some functions not fully functioning\n * improvement needed'}
```

4.8.2 – Llama Çıktı Örneği

```
{'elapsed_time_model': 2.659,  
'input_translation_time': 3.652,  
'out_translation_time': 0.32,  
'total_translation_time': 3.972,  
'output': 'r: Pozitif: * Kaliteli malzemelerden yapılmış * Dayanıklı * Kullanımı kolay * İyi fiyat-performans oranı Negatif: * Performans açısından beklentileri karşılamıyor * Kullanımı karmaşık * Bazı fonksiyonlar tam işlevsel değil',  
'raw_out': 'r:\n \nPositive:\n\n* Made of quality materials\n* Durable\n* Easy to use\n* Good price-performance ratio\n\nNegative:\n\n* Not meeting expectations in terms of performance\n* Complex to use\n* Some functions not fully functional'}
```


5 – PROJE ÇIKTILARI

Proje çıktıları çeviri hatalarından olabildiğince kaçınmak amacıyla rapora İngilizce olarak geçirilmiştir.

Çıktı 1’de Mistral, Llama’ya göre daha detaylı bir sonuç çıkarmıştır. Llama çıktısında belirli olumlu ve olumsuz özellikler fark edilememiştir. Kulak içine yerleştirme konusundaki olumsuz eleştiriyi iki modelde fark edememiştir.

Yorumlar	
<p>Harika bir ses kalitesi ve kablosuz özgürlük! Kesinlikle tavsiye ederim., Çok rahat, hafif ve pratik. Spor yaparken kullanmak için mükemmel bir seçim., Bağlantı bazen kesiliyor, özellikle yoğun yerlerde kullanırken sıkıntı yaşanabiliyor., Şarj süresi beklediğimden kısa, sık sık şarj etmek zorunda kalıyorum., Kolay bağlantı ve kullanım, mükemmel gürültü engelleme. Bu fiyata çok iyi bir ürün., Kulak içine yerleştirme konusunda biraz rahatsız edici, uzun süre kullanımda ağrı yapabiliyor.</p>	
<p>Mistral_Çıktı</p> <p>Positive</p> <ul style="list-style-type: none">* Great sound quality and wireless freedom!* Its very comfortable, light and practical, its an excellent choice to use when playing sports.* Easy connection and use, excellent noise blocking, a very good product at t <p>Negative</p> <ul style="list-style-type: none">* The connection is sometimes interrupted, especially when using it in busy places.* The charging time is shorter than I expected, I have to charge often.,	<p>Llama_Çıktı</p> <p>Positive:</p> <ul style="list-style-type: none">* Great sound quality and wireless freedom! I highly recommend it.* Its very comfortable, light and practical, its an excellent choice to use when playing sports. <p>Negative:</p> <ul style="list-style-type: none">* The connection is sometimes interrupted, especially when using it in busy places.* The charging time is shorter than I expected, I have to charge often.,
Mistral_HS	Llama_HS

4.764

3.590

Çıktı 2’de Mistral Llama’ya göre daha detaylı ve doğru çıktı vermiştir. Parmak izi okuyucu hakkındaki olumsuz yorum iki model tarafından da fark edilememiştir.

Hızlı işlemci, etkileyici kamera ve şık tasarım. En iyi telefonlardan biri!,
Yüksek performans ve dayanıklılık, her türlü uygulama ve oyunu sorunsuz çalıştırıyor.,
Fiyatı biraz yüksek, daha uygun fiyatlı alternatifler tercih edilebilir.,
Mükemmel ekran kalitesi, uzun pil ömrü ve akıcı kullanıcı arayüzü. Kesinlikle öneririm.,
Bazı kullanıcılar arasında yazılım güncellemeleriyle ilgili sorunlar yaşanabiliyor.,
Parmak izi okuyucu bazen çalışmıyor, bu da güvenlik açısından endişe verici olabiliyor.

Positive	Positive:
* Fast processor	* Fast processor
* Impressive camera	* Impressive camera
* Stylish design	* Stylish design
* High performance	* One of the best phones!
* Durability	
* Excellent screen quality	Negative:
* Long battery life	* Price is slightly higher
* Smooth user interface	* Some users may have problems with software updates.,
* Recommended	
Negative	
* Slightly higher price	
* More affordable alternatives may be preferred	
* Problems with software updates,	



Çıktı 3’de Mistral Llama’dan daha iyi performans göstermiştir. Pil ömrü konusundaki olumsuz yorum iki model tarafından da fark edilememiştir.

Takılıp kaldığı yerler olabiliyor, bazen manuel müdahale gerektirebiliyor.,
Ses seviyesi biraz yüksek, sessiz çalışan bir model tercih edilebilir.,
Evinizi temizlemek için harika bir yardımcı, akıllı navigasyon ve güçlü emiş gücü.,
Programlanabilir ve uzaktan kontrol edilebilir özellikleriyle hayatı kolaylaştırıyor.,
Köşe ve kenarları temizlemekte çok etkili, evcil hayvan sahipleri için ideal bir seçim.,
Pil ömrü beklediğimden kısa, büyük alanları temizlemek için birden fazla şarj gerekebiliyor.

Positive

- * smart navigation
- * powerful suction power
- * programmable and remotely controllable features
- * effective in cleaning corners and edges
- * ideal choice for pet owners

Negative

- * places requiring manual intervention
- * sound level is a bit high.

Positive:

- * Great helper to clean up your home
- * Programmable and remotely controllable features make life easier
- * Very effective in cleaning corners and edges

Negative:

- * There can be places where he hangs out, sometimes requiring manual intervention
- * Sound level is a bit high
- * A quiet working model may be preferred,

3.658

3.468

Çıktı 4'te Mistral Llama'dan daha iyi performans göstermiştir. Llama uzun pil ömrünü negatif olarak algılamıştır. İki model de fiyat-performans hakkındaki olumlu yorumu algılayamamıştır.

Bazı kullanıcılar arasında Bluetooth bağlantı sorunları yaşanabiliyor, cihazlar arasında iletişim kopuklukları olabiliyor.,

Ekran boyutu biraz küçük, bazı kullanıcılar için bilgi okunması zor olabilir.,

Yüzeyi çabuk çiziliyor, ekranı koruyucu kullanmak gerekebiliyor.,

Sağlık takibi, spor aktiviteleri ve bildirimler için mükemmel bir yardımcı, şık tasarımıyla da dikkat çekiyor.,

Uzun pil ömrü, suya dayanıklılık ve geniş uygulama desteği. Herkesin sahip olması gereken bir aksesuar.,

Uygun fiyatlı olmasına rağmen kaliteli malzeme ve fonksiyonlara sahip, performansııyla şaşırtıyor.

Positive

- * Health tracking is an excellent assistant for sports activities and notifications.
- * Stylish design.
- * Long battery life.

Negative

- * Bluetooth connectivity issues, communication disconnects between devices.
- * Small screen size, difficult to read information for some users.
- * The surface is drawn quickly, the screen may need to be used as a protector..

Positive:

- * Some users may have Bluetooth connectivity issues, communication disconnects between devices.
- * Health tracking is also an excellent assistant for sports activities and notifications, with its stylish design.

Negative:

- * The screen size is a bit small, for some users the information can be difficult to read.
- * The surface is drawn quickly, the screen may need to be used as a protector.
- * Long battery life,



Çıktı 5’te iki modelde birbirine benzer çıktılar verse de Llama daha iyi çıktılar vermiştir.
Taşınabilirlik ve güçlü donanım konusundaki yorum iki model tarafından da algılanamamıştır.

Fiyatı biraz yüksek, bütçesi kısıtlı olanlar için alternatif arayışı gerekebilir.,

Yüksek performans, uzun pil ömrü ve şık tasarım. Hem iş hem de eğlence için mükemmel bir seçim.,

Fan sesi bazen rahatsız edici olabiliyor, sessiz çalışan bir model tercih edilebilir.,

Hızlı açılış ve tepki verme süresi, çoklu görevleri rahatça yönetmek için ideal.,

Isınma sorunu yaşayan kullanıcılar var, uzun süreli kullanımlarda cihazın sıcaklığı artabiliyor.,

Taşınabilirlik ve güçlü donanım kombinasyonu en iyi dizüstü bilgisayarlar arasında yer alıyor.

Positive

- * High performance
- * Long battery life
- * Stylish design
- * Fast opening and response time

Negative

- * Price is a little high
- * Fan sound can sometimes be uncomfortable
- * There are users who have a wa,

Positive:

- * High performance
- * Long battery life
- * Stylish design
- * Fast opening and response time

Negative:

- * Price is a little high
- * Fan sound can sometimes be uncomfortable
- * Limited budget may make the alternative more appealing
- * Some users prefer a quiet-running model,



Çıktı 6'da iki model de birbirine benzer çıktılar da verse Mistral daha detaylı ve doğru bir çıktı vermiştir. Sıcaklık kontrolü konusundaki olumsuz yorum iki model tarafından da fark edilememiştir.

Mükemmel kahve demleme performansı, kolay temizlik ve şık tasarım. Sabahları benim için vazgeçilmez!,

Farklı kahve çeşitlerini kolayca hazırlama imkanı, kullanımı da oldukça pratik.,
Büyük su haznesi ve otomatik kapanma özelliği ile güvenli ve kullanışlı bir ürün.,
Bazı modellerde su sızıntısı problemi yaşanabiliyor, dikkatli olmak gerekiyor.,
Filtre değişim maliyetleri zamanla artabiliyor, bu da bakım masraflarını artırıyor.,
Sıcaklık kontrolü biraz zor, bazen istenilen kahve sıcaklığına ulaşmak mümkün olmayabiliyor.

Positive

- * Excellent coffee brewing performance
- * Easy cleaning
- * Stylish design
- * Easily preparing different types of coffee
- * Safe and useful product
- * Large water chamber
- * Automatic shutdown feature

Negative

- * Water leak problem in some models
- * Increased filter replacement costs.

Positive:

- * Excellent coffee brewing performance
- * Easy cleaning
- * Stylish design
- * Mornings are indispensable to me!

Negative:

- * Some models may have a water leak problem
- * Filter replacement costs can increase
- * Need to be careful,



Çıktı 7’de iki model de birbirine benzer çıktılar vermiştir. Dayanıklılık konusundaki olumsuz yorum iki model tarafından da farkedilememiştir.

Şarj süresi biraz uzun, beklemek gerekebiliyor. Daha hızlı şarj eden modeller tercih edilebilir.,
Bazı modellerde bağlantı sorunları yaşanabiliyor, cihazlar arasında iletişim kopuklukları olabiliyor.,

Hızlı şarj özelliği ve yüksek kapasitesi ile hayat kurtarıcı bir ürün. Her zaman yanımda taşıyorum!,
Kompakt tasarımı ve hafif yapısıyla taşınması kolay, seyahatlerde büyük bir avantaj sağlıyor.,
Birden fazla cihazı aynı anda şarj edebilme özelliği, çoklu kullanıcılar için ideal.,

Dayanıklılık konusunda sorunlar yaşanabiliyor, düşmelere karşı daha sağlam bir yapı tercih edilebilir.

Positive

- * fast charging capability
- * high capacity
- * compact design
- * lightweight structure
- * great advantage in travel

Negative

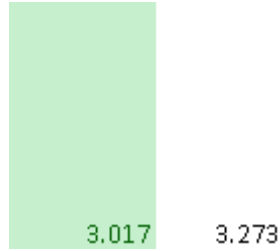
- * long charging time
- * connectivity issues
- * communication disconnects between devices,

Positive:

- * Fast charging capability
- * High capacity
- * Life-saving product
- * Compact design
- * Lightweight structure
- * Easy to carry

Negative:

- * Charging time is a bit long
- * Some models may have connectivity issues
- * Faster charging models are preferable,



Çıktı 8’de Pozitif özellikleri Mistral, negatif özellikleri ise Llama daha iyi şekilde tespit etmiştir. Su geçirmezlik özelliği hakkındaki olumlu yorum iki model tarafından da tespit edilememiştir.

Mükemmel ses kalitesi, şık tasarım ve kablosuz bağlantı özelliği ile her yerde kullanılabilir,
Pil ömrü biraz kısa, uzun süreli kullanımlarda sık sık şarj etmek gerekiyor.,
Bazı kullanıcılar arasında bağlantı kopmaları yaşanabiliyor, stabil bir bağlantı için daha güçlü sinyal gerekiyor.,
Bass performansı biraz zayıf, daha güçlü bir bas isteyenler için yeterli olmayabilir.,
Taşınabilirlik açısından ideal, pikniklerde, plajda veya ev partilerinde vazgeçilmezimiz haline geldi.,
Su geçirmez özelliği sayesinde dış mekanlarda kullanımı sorunsuz, dayanıklı bir yapıya sahip.

Positive

- * Stylish design
- * Excellent sound quality
- * Wireless connectivity
- * Portability

Negative

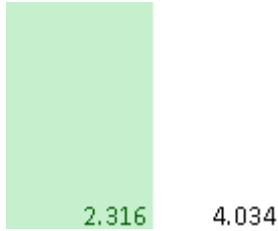
- * Short battery life
- * Disconnections
- * Weak bass performance,

Positive:

- * Excellent sound quality
- * Stylish design
- * Wireless connectivity

Negative:

- * Battery life is short
- * Disconnections may occur
- * Bass performance may not be enough for some users
- * Requires a stronger signal for a stable connection,



Çıktı 9’da her ne kadar Mistral Llama’dan daha iyi çıktı vermiş olsa da iki model de başarılı sonuçlar verememiştir.. Mistral iki, Llama ise hiçbir negatif özelliği tespit edememiştir.

Bazı modellerde ses seviyesi biraz yüksek, çalışırken rahatsız edici olabiliyor.,
Hızlı pişirme özelliği, geniş iç hacim ve kullanım kolaylığı ile mutfağın vazgeçilmezi oldu.,
Farklı pişirme programları ve hassas ayarlar sayesinde her yemeği mükemmel şekilde pişirebiliyorum.,
Kolay temizlenebilir iç yüzeyi ve şık dış tasarımı ile mutfağa modern bir hava katıyor.,
Kapı kolu veya düğmelerde zamanla aşınma ve arıza sorunları yaşanabiliyor.,
Isı dağılımı konusunda bazı kullanıcılar arasında şikayetler var, homojen pişirme sağlamak için dikkatli olmak gerekiyor.

Positive

- * fast cooking feature
- * wide internal volume
- * ease of use
- * different cooking programs and delicate settings
- * easy-to-clean interior
- * stylish exterior design

Negative

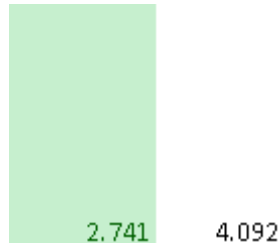
- * sound level is a little high and uncomfortable when working,

Positive:

- * The sound level is a little high in some models.
- * It became indispensable to the kitchen with its fast cooking feature, wide internal volume, and ease of use.
- * I can cook every dish perfectly thanks to different cooking programs and delicate settings.

Negative:

- * Nothing negative mentioned in the text group.,



Çıktı 10'da iki model benzer çıktılar vermiş ve iki model de iki adet olumsuz özelliği tespit edememiştir.

Kusursuz temizlik, enerji ve su tasarrufu ile hayatı kolaylaştıran bir ürün. Mutlaka edinilmeli!,
Farklı program seçenekleri ve sessiz çalışma özelliği ile kullanımı çok pratik.,
Bazı modellerde su damlatma problemi yaşanabiliyor, bu da mutfakta ekstra temizlik gerektirebiliyor.,
Geniş iç hacmi sayesinde büyük tabak ve tencereleri bile sorunsuz temizleyebiliyor.,
Bazı deterjanların köpürme problemi nedeniyle programı tamamlamadan durabiliyor.,
Bazı kullanıcılar arasında kurutma performansı konusunda şikayetler var, tam olarak kurutmuyor.

Positive

- * perfect cleaning
- * energy and water saving
- * practical to use
- * different program options
- * quiet working feature
- * large internal volume

Negative

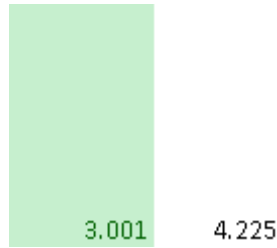
- * water drip problem
- * extra cleaning in the kitchen

Positive:

- * It makes life easier with perfect cleaning, energy, and water saving.
- * It is very practical to use with different program options and quiet working features.
- * It is able to clean even large plates and pots without problems.

Negative:

- * Some models may have a water drip problem, which may require extra cleaning in the kitchen.



Mistral LLM modelinin ortalama hesaplama süresi 3.60, Llama'nınki ise 3.56'dır. Yani bu örnekler göz önüne alındığında Llama, Mistral'den bir miktar daha hızlıdır.

Ortalama Hesaplama Süreleri	3.6002	3.5661
-----------------------------	--------	--------

SONUÇLAR VE ÖNERİLER

Alınan sonuçlar kullanılan yorumlar bağlamında ele alındığında Mistral modelinin Llama dil modeline göre daha iyi performans gösterdiği ve daha tutarlı sonuçlar verdiği gözlemlenmiştir. Llama'nın ilgili yorumları daha çok pozitif ve negatif olarak ayırıp direkt çıktı olarak vermesine karşın Mistral ürünlerin özelliklerini çıktı olarak vermiştir. İki modelde her yorum grubu içinde en az bir yoruma ait özellikleri tespit edememiştir. Özellikle her yorum grubunun en sonundaki yorum iki model tarafından da genele bakıldığında fark edilememiş ve sebebi tam olarak anlaşılamamıştır. Modellerin ortalama çıktı verme süreleri incelendiğinde Llama modelinin Mistral dil modeline göre oldukça ufak bir miktar daha hızlı çalıştığı söylenilebilir.

İlgili projenin hem daha gerçekçi çıktılar vermesi hem de daha iyi optimize edilebilmesi açısından gerekli yasal izinler alınarak daha büyük ve gerçek veri grubu üzerinde çalışılarak projenin olduğundan daha ileriye götürülmesi mümkündür. Ayrıca gerçek hayattaki ürün yorumları göz önüne alındığında bu yorumlarda dil bilgisi, noktalama işareti vb. kurallara dikkat edilmediği ve bolca yazım yanlışı yapıldığı görülecektir. Bu sebeple projenin hayata geçirilmesi için gerçek ve bol miktarda veriyle çalışılıp, ilgili modellerin bu verilerden doğru çıkarımları yapabilecek hale getirilmesi elzemdir. Ayrıca proje, daha güçlü GPU ve CPU'lar kullanılarak daha kompleks modellerle tekrar inşa edilebilir ve daha mütevazı dil modellerine göre sonuçların hangi derece iyileştiği karşılaştırılabilir.

KAYNAKÇA

1-) Vaswani, Ashish, Noam M. Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser and Illia Polosukhin. “Attention is All you Need.” *Neural Information Processing Systems* (2017).

2-) <https://huggingface.co/TheBloke/Mistral-7B-OpenOrca-GPTQ>

3-) <https://huggingface.co/TheBloke/LLaMa-7B-GPTQ>

4-) <https://huggingface.co/Helsinki-NLP/opus-mt-tc-big-en-tr>

5-) <https://huggingface.co/Helsinki-NLP/opus-mt-tc-big-tr-en>

EKLER

1- [Bitirme Projesi.xlsx](#)

2- [final_project.py](#)