

**Danışman
Hoca**

**DR. KADRIYE FİLİZ
BALBAL**



DEÜ FEN FAKÜLTESİ BİLGİSAYAR BİL BİTİRME PROJESİ

ÖĞRENCİLER

Alaaeddin KHEİTİ
Ali AKKOYUN
Esat YUSUF DALBUDAK
Hüseyin ELİBOL

Taklitçi Ai





01

PROJE TANIMI

02

LİTERATÜR TARAMASI VE SONUCU

03

PROJE ÖNERMESİ VE TASARIMI

04

UYGULAMA

05

SONUÇLAR

06

PROJE SUNUM ŞEKLİ

07

ENGELLER VE SORUNLAR

08

KAYNAKÇA



**TABLE OF
CONTENTS**

PROJE TANIMI

Projenin amacı:

- yapay zeka ve doğal dil işleme tekniklerini kullanarak ünlü şairlerin yazım tarzlarını taklit ederek benzersiz şiirler üretebilen akıllı model geliştirmektir.

LİTERATÜR TARAMASI VE SONUCU

KONU İLE İLGİLİ İNCELEDİĞİMİZ MAKALELER:

- 7 MAKALE

PROJEMİZE BENZEYEN ÇALIŞMA:

- On the power of special-purpose GPT models to create and evaluate new poetry in old styles.

LİTERATÜRE TARAMA ÖZETİ

1- Şiir üretmek ile ilgilenen çalışmalar çoğu ingilizce dil ile şiir üretir

2- Şairlerin yazma tarzı taklit etmeye çalışan çalışmalar çok az ve olursa bir yazar yada bir şair üzerinde çalışılmış aynı modelde birden fazla şair tarzı taklit edenler çok az.

3- şiir üretiminde çeşitli teknikler kullanıldı ve projemiz için kısıtlarımız ve imkanlarımız göz önünde bulundurarak en uygun teknikleri seçmeye çalıştık.

bitirme projesi için literatür taraması																			
genel bilgiler			şir hangi dil ile üretti	kullanılan teknikler										kullanılan değerlendirme metrikler					
çalışma başlığı	çalışmanın tarihi	hazırlayan	türkiye / ingilizce / diğer	RNN	LSTM	OC2Vec	BERT	GPT2	GANs	GPT	GPT1	GPT3.5	pre-trained / transfer	word cloze	F1 score	BLEU	GLAS	SIF	cosine similarity
ng writing style through transfer learning	3 yıl önce	Martin Slomczykowski	text not poetry - ingilizce	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
J: A Deep-learning-Based System for Automatic Arabic Poem Generation	14-09-2023	Eshrag Ali Refaee - Jazan University - ksa	Diğer - arabça	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
ed Poetry Generation: Controlling Personal style with Recurrent Neural Networks	2018	Jia Wei, Qiang Zhou, Yici Cai Tsinghua University, Beijing, China	Diğer - çince	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
neration using Transformers and Doc2Vec Embeddings	2020	Marvin C. Santillan, Arnulfo P. Azcarra	ingilizce	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
idem Turkish Poet: Fine-Tuned GPT-2	2023	Uygur Kurt Aykut Çayır	türkiye	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
power of special-purpose GPT models to write and evaluate new poetry in old styles	2023	Piotr Śawicki1, Marek Grzes1, Edoardo Grefe2	ingilizce	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Literatür tarama detaylı özetine bakmak için buraya [tıklayınız..](#)

Projede Kullanabileceğimiz teknikler

1- **RNN** ve gelişmiş versyonları

2- Açık kaynak pre-trained modeller
GPT2, BERT gibi..

3- Büyük dil modelleri **GPT3.5, GPT4, Gemini**

Projede Kullanabileceğimiz metrikler

1- **Uzman** kişilere sormak

2- Bazı yaygın metrikler **cosine similarity, BLUE, Word cloud**

3- Taklit başarısı ölçmek için **sınıflandırma modeli** kullanmak.

Neden RNN kullanmadık ?

- 1- Transformere dayalı modellere göre daha **düşük performans**
- 2- Uzun yazı ile **bağlam sorun** yaşanabilir
- 3- İyi performans elde etmek için çok **fazla veri gerektirir**
- 4- Bir deneme yaptık kayıt edilecek sonuç elde edemedik. **GRU**

Neden Açık kaynaklı modeller GPT2.. kullanmadık

- 1- **Türkçe** destekli iyi **model yok**
- 2- Yeniden eğitmek çok **hesaplama kaynağı** ister
- 3- **İki defa eğitmesi gerekiyor** önce şiir veriler ile sonra özel olarak taklit etmek istediğimiz şairlerin verileri.
- 4- Türkeç destekli bazı GPT2 ile denedik ancak **iyi sonuca varmadık**.

YÖNELDİĞİMİZ ÇÖZÜM:

GPT 3.5 modeli kullanıp ve taklit etmek istediğimiz şairlerin verileri ile ince ayar(**Fine tuning**) yapmaya karar verdik.

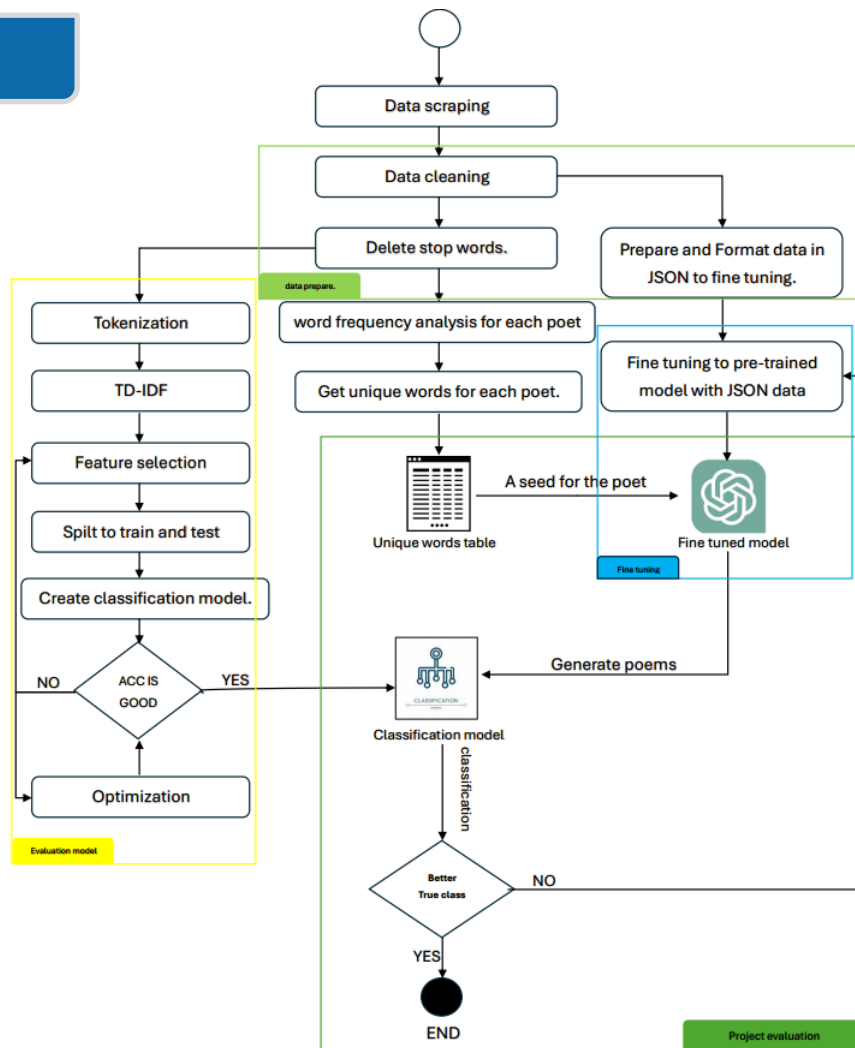
Fine tuned GPT 3.5 modeli neden kullandık?

- 1- **Türkçe** destekliyor
- 2- **Az veriler** ile **iyi performans**
- 3- **Özel durumlar için** ince ayar yapmak standart modelin performansından **daha iyi** olabilir
- 4- **Hesaplama kaynağı** sağlıyor.
- 5- Fine tuning **kolaylığı**.
- 6- Piyasada kullanımı **çok yaygın**.

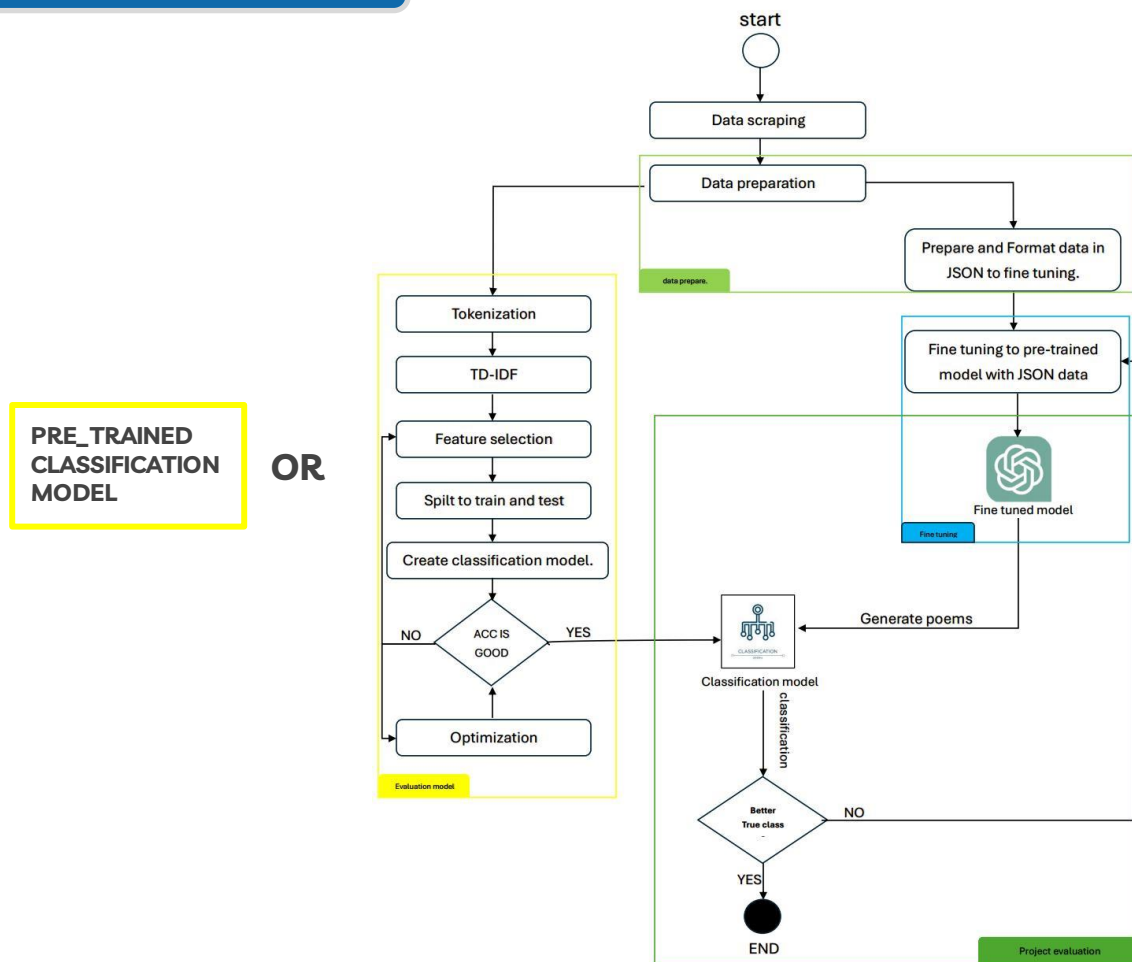
PROJE ÖNERMESİ VE TASARIMI

**PRE_TRAINED
CLASSIFICATION
MODEL**

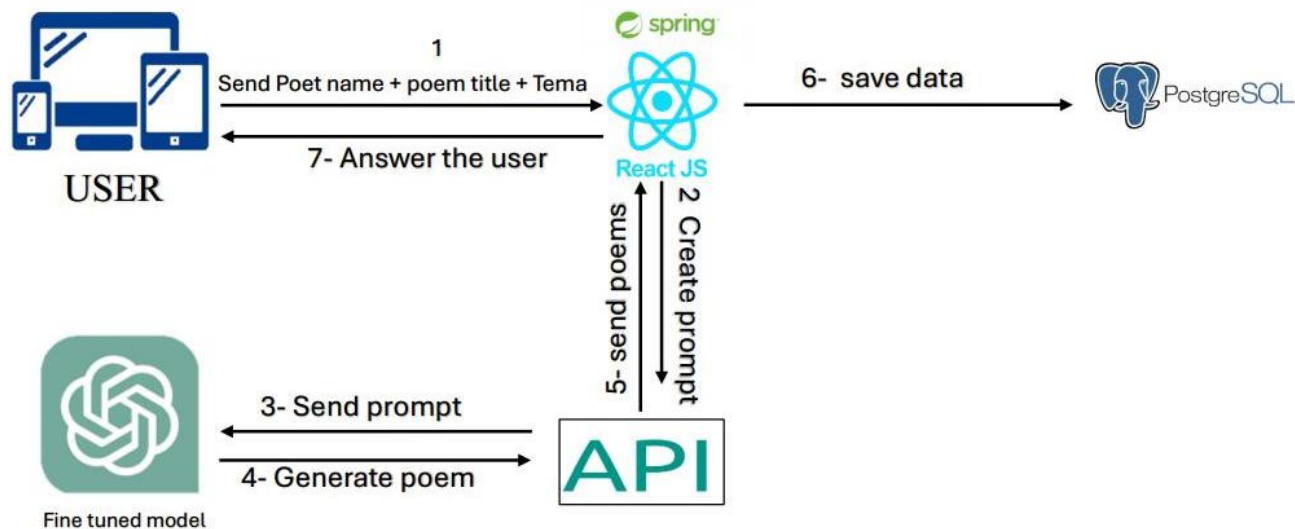
OR



Flowchart of project



Project architecture



UYGULAMA

UYGULAMA – 1. Aşama

Veri kaynağı:

- Antoloji şiir sitesi

Seçtiğimiz şairler:

- Orhan veli kanık
- Attila İlhan
- Behçet Necatigil
- Abdurrahim karaköç
- Nazim hikmet Ran

Veri boyutu:

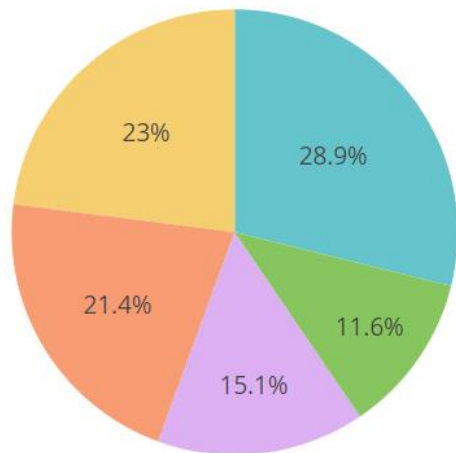
- (1273, 3)

Scraping kütüphaneler:



Şairlerin bazında Veri dağılımı:

Abdurrahim Karakoç	368
Behçet Necatigil	293
Nâzım Hikmet	273
Attilâ İlhan	192
Orhan Veli Kanık	148



Abdurrahim Karakoç
Behçet Necatigil
Nâzım Hikmet
Attilâ İlhan
Orhan Veli Kanık

Şiirlerin kategorilerine sınıflandırmak:

- Kullanıcıya daha fazla seçenek sunabilmek için ve modele yardımcı olacak bir feature daha eklemek istedik o da «şiirin konusu» ve bunu çıkartmak için **Gemini** kullandık



poem_title	theme
Anlatamıyorum	Hüzün
İstanbul'u Dinliyorum	İstanbul
Vatan İçin	Vatan
Bedava	Acı
Açsam Rüzgara	Deniz

Verileri uygun formatta dönüştürmek

- Fine tuning yapabilmek için verileri seçtiğimiz modele uygun bir formatta (JSONL) hazırladık **örnek:**

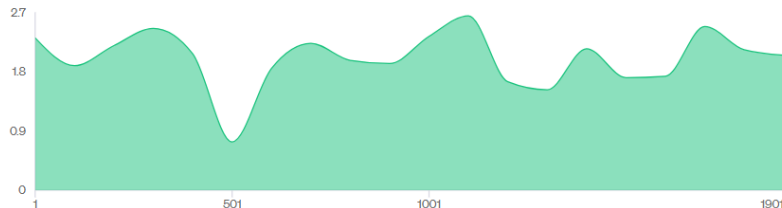
```
{ "messages": [ { "role": "system", "content": "Marv is a factual chatbot that is also sarcastic." }, { "role": "user", "content": "What's the capital of France?" }, { "role": "assistant", "content": "Paris, as if everyone doesn't know that already." } ] }
```

Seçmiş olduğumuz model

- GPT 3.5 TURBO 0125

Fine tuning parametreler:

- epoch: 3
- Batch size: 2 auto
- LR: 2 auto

Training Loss:**- 2.0488**

Step	Training loss
1901	2.0488
1801	2.1317
1701	2.4859
1601	1.7312
1501	1.7097
1401	2.1485
1301	1.5251
1201	1.6459
1101	2.6482
1001	2.3362
901	1.9230
801	1.9724
701	2.2309
601	1.8489
501	0.7325
401	2.0724
301	2.4581
201	2.1958
101	1.8902
1	2.3123

Rastgele çıktılar:

Üzgünüm hayat deyip bana bakma
 Hüznümü anlatacak kelime yoktur.
 Belâ çökmüş üstüme, karşıma bakma
 Kahrımı içime atacak kelime yoktur. Dostlar, günlük
 sevinçleriniz varmış
 Parıltılı derdinizi yazmış
 Bilin ki benim yüreğim ağrımış
 Acımı anlatacak kelime yoktur. Ağlayan gözlerimi
 güneşte yıkamam
 Kapandı bir kez pencereler, açamam
 Bu feryadımı seste gizleyemem
 Zulmümü akacak kelime yoktur.Eser mazi rahatlığı,
 benlere eziyet
 Hisler zindan olmuş, dilde bir kezzap vebâ
 Semalarda ne bayram ne de bayram selâset
 Cânımı uçuracak kelime yoktur.

Abdurrahim karaköç

Fine tuning sonucu

Daha bir yel alıp boydan
 Dallar doldu dallardan
 Ormandan yel alıp boydan
 Dostlar doldu dostlardan.

Behçet Necatigil

Fine tuning sonucu

Modeli performansını değerlendirmek için iki yöntem geliştirdik:

1- genel değerlendirme (kişisel):

- konu tutarlı (0-5 puan)
- başlık etkisi (0-5 puan)
- yazı uzunluğu (0-5 puan)
- yazı doğruluğu (0-5 puan)

2- Taklit performansını ölçmek için **sınıflandırma modeli** kullandık

konu tutarlı	başlık etkisi	yazı doğruluğu	uzunluk uygunluğu
4	3.56	4.62	3.76

Rasladığımız zayıf yönler:

1- şiirde **kıtalar arasında** boşluk bırakılmıyor yerine yapışık harfler olur.

2- bazı şiirler çok **kısa**.

3- 25 tane şiir içinde bir tane şiir **formatında değildi**.

Ne çıkış kaldıYorumu üzerinde yapraklar.
Gün uykudan güne uyandığında
Yanar alev alev deniz.
Bahçede kızıl bir palmiye
Şimdi saat yedi
Üstüme kötü bir gece çökecek.

dostlarım
bırakın imha planlarını
gizli hesapları

Arif Damar (Poetika, 9 Temmuz 2019) Hayatta bir tek yeminimi
hatırlıyorum: "Ecel beni nasıl yaralarsa öyle yara

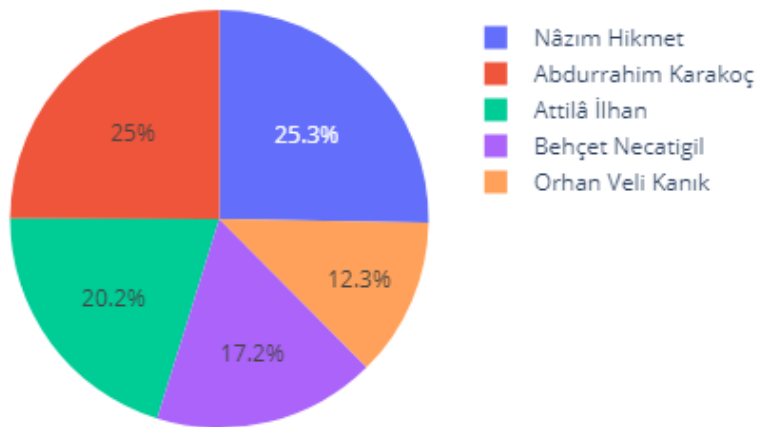
UYGULAMA – 2. Aşama(iyileştirme)

Veri iyileştirmesi:

- 1- 5 farklı siteden yeni veri toplayıp eksik şiir ile değiştirmek
- 2- Şiirlerin uzunluğu analizi edip çok kısa olanları veri setinden atmak.
- 3- şiirlerin uzunluğu analizi edip uzunluk dağılımı dengelemek.
- 4- şiirlerin sayısı şairler bazında dengelemek
- 5- kıtalar arasındaki oluşan yapışık harflar yerine boşluk bırakmak

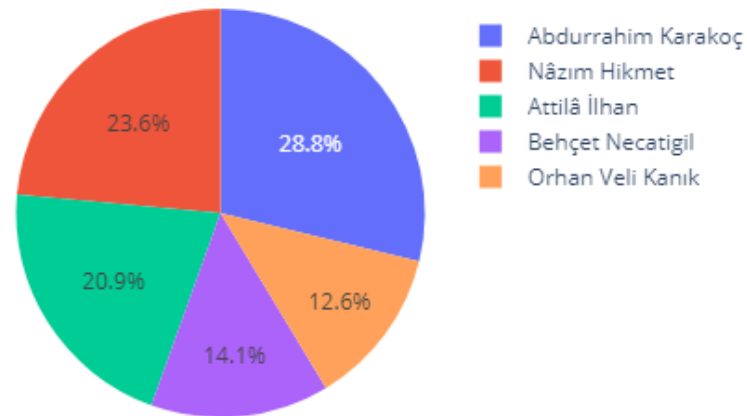
Veri iyileştirmesi:

distribution train data



Veriyi dengeledikten sonra dağılımı

distribution test data



Veri iyileştirmesi:

Şiir uzunluğu hakkında bir örnekler

Behçet Necatigil	ideal(40-75]	48
	long(75-125]	33
	max(>300)	2
	small(10-20]	56
	standart(20-40]	50
	very long(125-300]	12

Attilâ İlhan	ideal(40-75]	45
	long(75-125]	64
	max(>300)	11
	small(10-20]	9
	standart(20-40]	55
	very long(125-300]	60

Orhan Veli Kanık	ideal(40-75]	36
	long(75-125]	22
	max(>300)	1
	small(10-20]	28
	standart(20-40]	46
	very long(125-300]	10
	very small(<=10)	5

Nâzım Hikmet	ideal(40-75]	85
	long(75-125]	41
	max(>300)	22
	small(10-20]	7
	standart(20-40]	85
	very long(125-300]	60

Veri iyileştirmesi:

- Yeni Veri boyutu:

- (1273, 3)

Benim oralarda hiçbir işim yoktu
Şeytana uydum
Aç ahtapotlar kaynaşırken dipte
Kaypak kalabalıkta sürükleni.....



**Yaklaşık eksik verilerinin yüzde 50'i
tamamlayabildik**

Verinin eksikliği tamamlamak - örnek

Benim oralarda hiçbir işim yoktu
Şeytana uydum,
Aç ahtapotlar kaynaşırken dipte
Kaypak kalabalıkta sürükleniyordum.

İnce yüzünüzde üzgünce bir bakış
Birden sizi gördüm,
Açtı arı doruklarda bir safran
Durdum.

İlk sevgili güldü yitik anılardan
Mutsuz, yalnız
Sessiz kınamanızı, utançlarda küçülmüş
Aldım, geri döndüm.

Gelsem,
Siz yine orada mısınız?

2. Fine tuning parametreler:

- epoch: 1
- Batch size: 2 auto
- LR: 2 auto

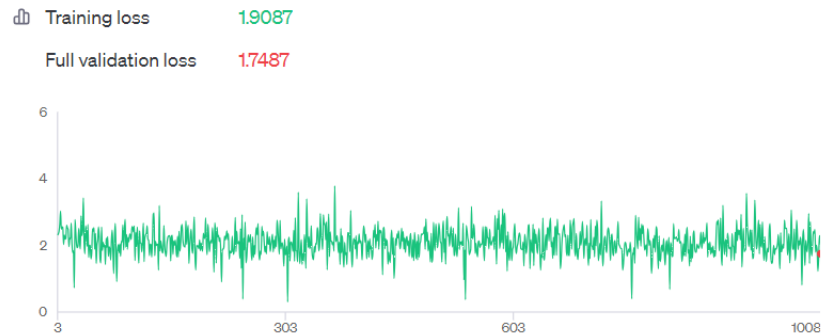
2. Modelin Training Loss:

- 1.9087

2. Modelin validation Loss:

- 1.7487

2. Fine tuning model(base model: daha önce eğittğimiz model)



SONUÇLAR

Rastgele çıktılar:

- ŞAİR: **Abdurrahim karaköç**
- KONU: **Vatan**
- BAŞLIK: **“Ah vatanım”**

Bir zindandır vatanım, bir mezar
Kan, gözyaşı, gamlı iklim iklim...
Ecdadımın maziden aldığı gan suyunu
Kurt sürüsü içmektedir sürülerce uşak!

Kavgayı ve korkuyu yemesini bilmeyen ruhlar
Yemek yemeye çalışmakta sıra
Üç kurşuluktan milyon dolarlık cüceler
Vahsî uzuvlarla beslenmektedir bekçi uşak!

Yanlışlığı, kirliği görmezlikten geledim
Üstelik taşırmayla kablâ değil teyemmüm
Bir yanda, avuçlarının içinde yılan kılınc
Dudak büküp dalga geçmekte, öte yanda uşak!

- ŞAİR: **Nâzım Hikmet**
- KONU: **İnanç**
- BAŞLIK: **“İnançlı Yolculuk”**

Ey, dostum!
Bak;
On adımda bir düşüyoruz
ve onbirincide cansız,
kaldırıyorlar üzerinden bizi.
Yok oluyoruz.
Öyleyse,
İnançlı Yolculuk!
Göğsümüze kara toprak alamet yakıp soğuktan.
Ayakların altında
'avuçlarını sırtına vermiş
erimemiş kar topları' gibi
yıldızlar.
Yıldızlar, dostum!
İnançlı yolculuk.

Rastgele çıktılar:

- ŞAİR: **Behçet Necatigil**
- KONU: **Hastalık**
- BAŞLIK: **“Hastalığın Acısı”**

Beni ben halim siz benden görünen böyle
Hani bıyıklı çanaklar hani evler çevrende
Biraz daha sandım dönmeli dönmeli batıya
Yine de kış geldi tuz doldu

Ölülerim yanı başımda
Yaşayanları saymıyorum
İşte

- ŞAİR: **Orhan Veli Kanık**
- KONU: **Rüya**
- BAŞLIK: **“Rüyanın İzleri”**

Kalk, kendi üzerinde seneler geçmişsin
Bir yere benzer, bir yere benzemezsin
Gece uzaklarına kaçmış hâlâ rüyanın izleri
Sana benzemez, sana benzerdin.
Birden bire yutkunursun, omuzlarını silkeler
Kendi sesin olma, kendi sesin olma
Gözlerini açtığında gece tavsiye ederim
Loş bir ışık yak, loş bir ışık yak şeklinde devam ediyor.

Model:

- TURKISH BERT

Model parametresi:

- Epoch: 61
- Training batch size: 64
- Label count: 5

Traning loss:

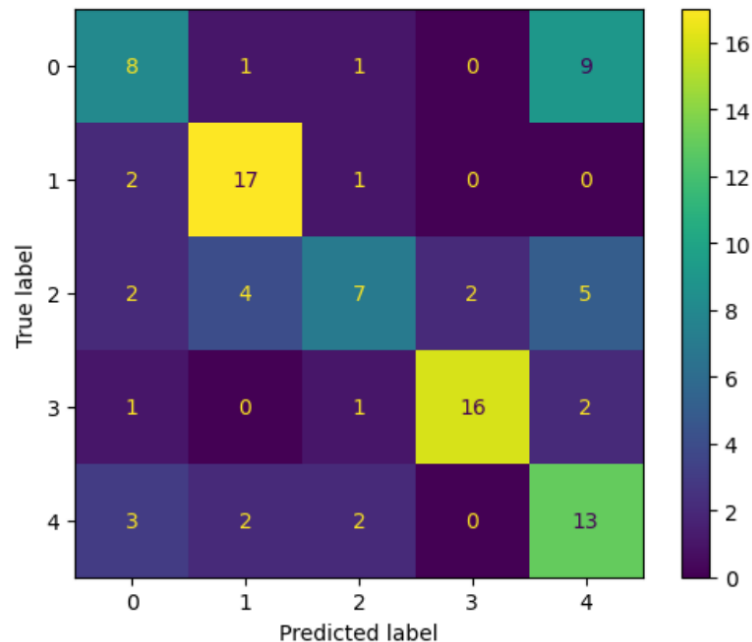
- 0.1864

Test sonucu

	precision	recall	f1-score	support
0	0.85	0.78	0.82	37
1	0.91	0.79	0.85	63
2	0.88	0.78	0.83	46
3	0.98	0.99	0.99	103
4	0.74	0.91	0.82	67
accuracy			0.88	316
macro avg	0.87	0.85	0.86	316
weighted avg	0.89	0.88	0.88	316

Ürettiğimiz 100 adet şiir sınıflandırma model ile test etme sonucu

	precision	recall	f1-score	support
0	0.50	0.42	0.46	19
1	0.71	0.85	0.77	20
2	0.58	0.35	0.44	20
3	0.89	0.80	0.84	20
4	0.45	0.65	0.53	20
accuracy			0.62	99
macro avg	0.63	0.61	0.61	99
weighted avg	0.63	0.62	0.61	99



Normal GPT 3.5 turbo ve Fine tuned modelimiz karşılaştırmak

Karşılaştırma sonucu:

Fine tuned modelimiz

	precision	recall	f1-score	support
0	0.50	0.42	0.46	19
1	0.71	0.85	0.77	20
2	0.58	0.35	0.44	20
3	0.89	0.80	0.84	20
4	0.45	0.65	0.53	20
accuracy			0.62	99
macro avg	0.63	0.61	0.61	99
weighted avg	0.63	0.62	0.61	99

Normal GPT 3.5 turbo

	precision	recall	f1-score	support
0	0.00	0.00	0.00	5
1	0.00	0.00	0.00	5
2	0.00	0.00	0.00	5
3	0.00	0.00	0.00	5
4	0.19	0.60	0.29	5
accuracy			0.12	25
macro avg	0.04	0.12	0.06	25
weighted avg	0.04	0.12	0.06	25

Fine tuning yapmaya neden ihtiyaç duyduk ?

Diğer benzer çalışmanın performansı VS Fine tuned modelimiz taklit performansı

Karşılaştırma sonucu:

Fine tuned modelimiz performansı

	precision	recall	f1-score	support
0	0.50	0.42	0.46	19
1	0.71	0.85	0.77	20
2	0.58	0.35	0.44	20
3	0.89	0.80	0.84	20
4	0.45	0.65	0.53	20
accuracy			0.62	99
macro avg	0.63	0.61	0.61	99
weighted avg	0.63	0.62	0.61	99

Diğer çalışmanın performansı

Walt Whitman GPT-3 vs Walt Whitman original			
Model	Correct	Incorrect	Accuracy
Ada 4e	127	73	63.5%
Ada 7A 4e	140	60	70%
Babbage 4e	131	69	65.5%
Babbage 7A 4e	134	66	67%
Curie 1e	150	50	75%
Curie 4e	123	77	61.5%
Curie 7A 4e	131	69	65.5%
Davinci 1e	144	56	72%
Davinci 4e	174	26	87%
Davinci 7A 4e	137	63	68.5%

Bu tablo sadece
bir yazar taklidi
değerlendirmesi

Taklit performansı diğer çalışm..

Sonuç olarak:

- Geliştirdiğimiz modelin taklit performansı ortalama olarak **62%dir**, verileri yeterince yeterli ve verilerinde eksik olmayan yazarlar için **80-85%** e kadar ulaşmaktadır.
- **100** adet şiir den oluşan çektiğimiz test verleri (sample) içinde çok kısa olan şiirler sadece **5%** teşkil ediyor ve rasladığımız her hangi bir yazı yanlışlığı yoktur.

Proje sunum şekli

Kullanıcı tarafından yapılan işler:

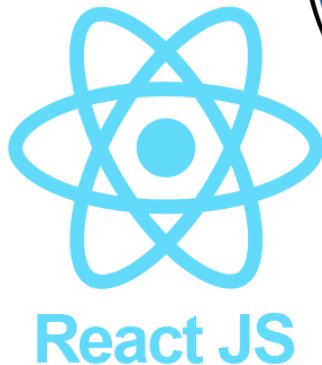
- 1- Şair seçmek
- 2- Şiirin başlığı önermek
- 3- Şiirin konusu seçmek
- 4- Generate e tıklamak

Web sitede kullanılan teknikler:

- Java spring
- React

Veri tabanı:

- postgresql



Şiir Oluşturmak İçin Artık Şiir TaklitçiAI

Tek Tuşla Şiir Üretin!

Dokuz Eylül üniversitesi öğrencileri olarak tasarladığımız ŞiirTaklitçiAI test etmek ister misiniz ?

KULLANMAYA BAŞLA



[ŞİİRİNİ ÜRETMEK İSTEDİĞİNİZ ŞAIRIN 'ŞİİR ÜRET' BUTONUNA BASMANIZ YETERLİ]

Tek tuşla şiir üretmeye başla

Şiirinizin Olası Başlığını giriniz.

Şiirinizin Temasını giriniz.

**Orhan
Veli Kanık**

AI chatbot, personalized
recommendations

**Toplam oluşturulmuş
şiir : 0**

ŞİİR ÜRET

**Atilla
İlhan**

AI chatbot, personalized
recommendations

**Toplam oluşturulmuş
şiir : 0**

ŞİİR ÜRET

**Behçet
Necatigil**

AI chatbot, personalized
recommendations

**Toplam oluşturulmuş
şiir : 0**

ŞİİR ÜRET

**Nazım
Hikmet Ran**

Advanced AI chatbot, priority
support, analytics dashboard

**Toplam oluşturulmuş şiir :
9.99**

ŞİİR ÜRET

**Abdurrahim
Karakoç**

Custom AI chatbot, advanced
analytics, dedicated account

Toplam oluşturulmuş şiir : 111

ŞİİR ÜRET


Çıktı: Burada çıktı metni yer alacak

Engeller ve Sorunlar

Engeller ve sorunlar:

- 1- Veri eksikliği tamamını giderilememesi.
- 2- Fine tuning ücretli olması epoch sayısı rahat bir şekilde arttıramıyoruz.
- 4- Şiirin kalitesi göreceli olması.

KAYNAKÇA



**AI olmam bana
her zaman ilham
geleceğini
anlamında
gelmiyor canım!**