

Dönüştürücü Tabanlı Emsal Karar Bulma Yöntemi

A Transformer-based Prior Legal Case Retrieval Method

Ceyhun E. ÖZTÜRK

Elektrik ve Elektronik Mühendisliği Bölümü
Bilkent Üniversitesi
ASELSAN Research Center
Ankara, Türkiye
ceyhun.ozturk@bilkent.edu.tr

Ş. Barış ÖZÇELİK

Hukuk Fakültesi
Bilkent Üniversitesi
Ankara, Türkiye
bozcelik@bilkent.edu.tr

Aykut KOÇ

Elektrik ve Elektronik Mühendisliği Bölümü
Ulusal Manyetik Rezonans Araştırma Merkezi
Bilkent Üniversitesi
Ankara, Türkiye
aykut.koc@bilkent.edu.tr

Özetçe —Bu çalışmada BERTurk-Legal isimli dönüştürücü tabanlı model emsal karar bulma görevinde kullanılmak üzere önerilmektedir. BERTurk-Legal'in ön eğitimi Türkçe hukuk alanında bir veri kümesi ile yapılmıştır. Bu veri kümesi emsal kararlar ile ilgili herhangi bir etiket bulundurmamaktadır. BERTurk-Legal maskeli dil modelleme kullanılarak kendiliğinden denetimli bir şekilde eğitilmiştir. BERTurk-Legal sınıflandırma görevi üzerinde eğitilmeksizin Yargıtay davalarından oluşan bir veri kümesinde literatürdeki en iyi sonuçları vermiştir. Deney sonuçları Türkçe hukuk alanına özel dil modelleri geliştirme gerekliliğini göstermektedir.

Anahtar Kelimeler—Doğal dil işleme, hukuk teknolojileri, derin öğrenme, emsal karar bulma, hukukta NLP, Türkçe NLP.

Abstract—In this work, BERTurk-Legal, a transformer-based language model, is introduced to retrieve prior legal cases. BERTurk-Legal is pre-trained on a dataset from the Turkish legal domain. This dataset does not contain any labels related to the prior court case retrieval task. Masked language modeling is used to train BERTurk-Legal in a self-supervised manner. With zero-shot classification, BERTurk-Legal provides state-of-the-art results on the dataset consisting of legal cases of the Court of Cassation of Turkey. The results of the experiments show the necessity of developing language models specific to the Turkish law domain.

Keywords—Natural language processing, legal tech, deep learning, prior legal case retrieval, legal NLP, Turkish NLP.

I. GİRİŞ

Hukuk büyük miktarlarda metinlerin işlenmesini gerektiren sorumluluklar barındıran bir çalışma alanıdır. Doğal dil işleme (NLP) alanında yapılan çalışmalar sonucunda hukuk alanında metinleri insanlardan daha hızlı ve isabetli şekilde işleyebilecek yöntemler ortaya çıkmıştır. Hukukta NLP çalışmaları ilk günlerinden beri genellikle İngilizce metinler üzerine odaklanmıştır. [1]–[3] bu alandaki öncü çalışmalardır. [4] hukukta NLP çalışmalarını kapsamlı bir biçimde incelemiştir. Son yıllarda derin öğrenme yöntemlerinin kullanımı [5] ile birlikte hukukta NLP çalışmaları hız kazanmıştır. Emsal karar bulma [6]–[9], mahkeme karar tahmini [10], [11], varlık ismi tanıma (NER)

[12], [13] ve metinlerin ön yargıdan arındırılması [14] hukukta NLP'nin önemli uygulamalarındandır.

Hukukta NLP çalışmalarında evrimsel sinir ağı (CNN) [7] ve yinlemeli sinir ağı (RNN) [10], [11] gibi derin öğrenme yöntemleri yoğun şekilde kullanılmıştır ancak daha büyük parametrelili bir derin öğrenme çeşidi olan dönüştürücü mimarisi [15] tabanlı yöntemler son yıllarda hukuk alanında birçok çalışmada [5], [16] literatürdeki en iyi sonuçları elde etmiştir. Dönüştürücü modeller word2Vec [17] ve GloVe [18] gibi algoritmalar aksine kelimelere sabit bir temsil vektörü atamak yerine kelimelere bulunduğu metne göre değişecek şekilde dinamik olarak temsil vektörü atamaktadır. BERT [19] ve GPT-4 [20] literatürde sıkça kullanılan dönüştürücü tabanlı mimarilerdir.

Türkçe hukuk metinleri için yapılan NLP çalışmaları [6], [11], [21] son yıllarda hız kazanmıştır. Türkçe emsal karar bulma alanında ise literatürde tek bir çalışma [6] bulunmaktadır. Bu çalışmada RNN türevi olan GRU ve LSTM modelleri kullanılarak emsal karar tahmini yapılmıştır. Bizim çalışmamızda ise RNN yerine günümüzdeki NLP alanında son dönemdeki en başarılı mimari olan dönüştürücü mimarisi kullanılmaktadır. Bildiride önerdiğimiz BERTurk-Legal modeli ile önceki çalışmalardaki en iyi model olan LSTM & ATIRE BM25 modeline [6] göre daha başarılı sonuçlar elde edilmiştir. BERTurk-Legal modelinin ön eğitimi için [22] çalışmasında yayınlanan Türk Hukuk Derlemi isimli derlem kullanılmıştır. Model maskeli dil modelleme yöntemi ile kendiliğinden denetimli şekilde eğitilmiştir. Modelin ilk ağırlıkları ise BERTurk [23] isimli genel maksatlı Türkçe dil modelinden alınmıştır. Bu şekilde hem geniş çaplı Türkçe eğitiminden hem de hukuk alanına özel bilgilerden yararlanan BERTurk-Legal modeli elde edilmiştir.

II. İLGİLİ ÇALIŞMALAR

İlgili çalışmalar emsal karar bulma ve Türkçe hukukta NLP hakkında olmak üzere iki alt başlıkta incelenmektedir.

A. Emsal Karar Bulma

Otomatik emsal karar bulma görevi bir bilgi bulgetir (Ing. information retrieval, kısaca IR) görevi türüdür. Her yıl düzenlenen COLIEE isimli bir hukukta NLP temalı yarışmada bir

emsal karar bulma görevi bulunmaktadır. Görev için sağlanan veriler Kanada Federal Mahkemesi verileridir [24]. 2019 yılındaki yarışmada JNLP takımı [7] emsal karar bulma alanında birinci olmuştur. Takım CNN tabanlı bir model ile doküman vektörleri üretmiştir. Ardından bu vektörler karşılaştırılarak sorgulanan her doküman için en benzer dokümanlar elde edilmiştir. Diğer bir deyişle sorgulanan her doküman için aday dokümanların bulgetir sıralamaları elde edilmiştir. Bu şekilde derin öğrenme ile doküman vektörleri elde edilmesi ve vektörlerin benzerlikleri ile sıralama yapılması IR literatüründe [25], [26] yaygın kullanılan bir yöntemdir. IR literatüründe derin öğrenme yöntemleri yaygın olsa da BM25 [27] ve LMIR gibi geleneksel IR yöntemleri bazı durumlarda derin öğrenme yöntemlerinden daha iyi performans göstermektedir. Örneğin, COLIEE 2021 yarışmasında emsal karar bulma alanında en iyi dört modelin üçü Okapi BM25 ya da LMIR yöntemini kullanmıştır [24].

[9] Hindistan Yüksek Mahkemesi için bir emsal karar bulma yöntemi geliştirmiştir. Yöntemde bulunan çift yönlü LSTM modeli ile kanıt ve tanıklık içeren cümleler işaretlenmektedir. Yöntem daha sonra dönüştürücü tabanlı bir model ile bu cümleleri işleyerek dokümanların bulgetir skorunu hesaplamaktadır. [8] karar metinlerini LEGAL-BERT modeli [5] ile işlemektedir. [8] LEGAL-BERT modeline metinleri sığdırabilmek için metinleri parçalara bölmüştür. Karar metinlerinin tamamını temsil eden doküman vektörleri elde etmek için LEGAL-BERT tarafından üretilmiş olan parçalara ait doküman vektörleri toplanmıştır. Doküman vektörlerinin benzerliğine bakılarak emsal kararlar elde edilmiştir.

B. Türkçe Hukukta NLP

[11] Türk hukuk sistemindeki ilk NLP yayınıdır. Çalışmada SVM ve karar ağacı gibi makine öğrenmesi yöntemleri ve RNN türevi derin öğrenme yöntemleri ile yüksek mahkemeler için karar tahmini yapılmıştır. Daha sonraki çalışmalarda mahkeme karar tahmini için [22] RNN türevi yöntemler, [28] ileri beslemeli sinir ağı (FFNN), [21] ise bir dönüştürücü modeli kullanmıştır. [13] LSTM ile koşullu rassal alan (CRF) mimarilerini birleştirerek Türkçe hukuk metinlerinde NER çalışması yapmıştır. [6] Türkçe emsal karar bulma alanındaki ilk akademik yayındır. Yayında RNN türevi modeller ve bu modellerin çeşitli BM25 algoritmalarıyla birleştirilmiş versiyonları doküman sıralama için kullanılmıştır. Yargıtay davaları için bir test veri kümesi oluşturulmuştur ve modeller bu kümede test edilmiştir. En iyi performans veren model ATIRE BM25 algoritmasıyla birleştirilen LSTM modelidir.

III. METODOLOJİ

Bildiride test kümesi olarak Turkish prior court case retrieval (TPCR) (*T*: Türkçe emsal karar bulma) veri kümesi¹ [6] kullanılmıştır. [6] tarafından tanımlanan emsal karar bulma görevinin amacı TPCR veri kümesindeki her karar metnini veri kümesindeki kendisine benzer diğer karar metinleriyle eşleştirmektir. TPCR veri kümesinde karar metinleri 26 gruba bölünmüştür. Her grup gruptaki karar metinlerinin konusunu özetleyen bir cümleyle ilişkilidir. TPCR 257 karar metninden oluşan ve grup başına yaklaşık 10 davadan oluşan dengeli sayılabilecek bir veri kümesidir.

Bildiride temel alınan çalışmadaki [6] gibi performans metrikleri olarak mikro-F1@5, mikro-keskinlik@5 ve mikro-hassasiyet@5 kullanılmıştır. Buna ek olarak her karar metni ile alakalı olan (emsal olan) yaklaşık 9 karar metni olduğu için mikro-F1@9, mikro-keskinlik@9 ve mikro-hassasiyet@9 metrikleri de kullanılmıştır. Mikro-keskinlik@k (mikro ortalama keskinlik) bulgetir algoritmasının bulduğu doğru eşleşme sayıları toplamının toplam aday doküman sayısına bölümüdür. Mikro-keskinlik@k metriği hesaplanırken bulgetir modelinin her sorgu dokümanı için k tane aday doküman bulduğu varsayılır. Mikro-keskinlik@k aday dokümanları benzerlik oranı çoktan aza olacak şekilde 1, 2, ... N şeklinde sıralayan bir IR modeli için aşağıdaki şekilde hesaplanmaktadır:

$$\text{Mikro-keskinlik@k (\%)} = 100 \times \frac{\sum_{i=1}^N \sum_{j=1}^k a_{ij}}{N \times k},$$

$$a_{ij} = \begin{cases} 1, & \text{eğer doküman } j \text{ sorgu dokümanı } i \text{ ile eşleşiyorsa,} \\ 0, & \text{aksi durumda.} \end{cases} \quad (1)$$

Yukarıdaki eşitlikte N toplam aday doküman sayısıdır. Bu metrikte veri kümesinin her örneği eşit değerdedir. Mikro-hassasiyet@k bulgetir algoritmasının bulduğu doğru eşleşme sayıları toplamının toplam alakalı doküman sayısına bölümüdür. Mikro-hassasiyet@k aşağıdaki şekilde hesaplanmaktadır:

$$\text{Mikro-hassasiyet@k (\%)} = 100 \times \frac{\sum_{i=1}^N \sum_{j=1}^k a_{ij}}{\sum_{i=1}^N r_i}. \quad (2)$$

Yukarıdaki eşitlikte r_i , doküman i için alakalı metin sayısını ifade etmektedir ve a_{ij} Eşitlik 1'de tanımlandığı şekilde kullanılmıştır. Mikro-F1@k ise aşağıdaki şekilde tanımlanmıştır:

$$\text{Mikro-F1@k (\%)} = \frac{2 \times \text{Mikro-kes.}@k \times \text{Mikro-has.}@k}{\text{Mikro-kes.}@k + \text{Mikro-has.}@k}. \quad (3)$$

Yukarıdaki eşitlikle mikro-kes.@k ve mikro-has.@k sırasıyla mikro-keskinlik@k ve mikro-hassasiyet@k'nin kısaltmalarıdır.

Bildiride karar metinlerine ait doküman vektörleri oluşturmak için BERTurk-Legal modeli kullanılmıştır. BERTurk-Legal ve BERTurk modelleri [23] BERT-BASE [19] modeline mimari olarak denktir. NLP alanında token'lar metinlerin anlamlı en küçük birimleridir. BERTurk-Legal'in 12 katmanı, 512 maksimum token sayısı ve 768 temsil (*İng.* embedding) boyutu vardır. Bu yüzden 512 token'den daha uzun metinlerin sadece ilk 512 token'i işlenmektedir. BERTurk-Legal'in her katmanı seri bağlanmış iki alt katmandan oluşmaktadır. İlk alt katmanın çıktısı aşağıdaki şekilde bulunmaktadır:

$$\mathbf{O}_1 = \text{Katman Normu}(\mathbf{X} + \text{Kendine Dikkat}(\mathbf{X})). \quad (4)$$

Yukarıdaki eşitlikte \mathbf{O}_1 çıktıdır, *Katman Normu* katman normalizasyon katmanıdır ve \mathbf{X} girdi metninin temsilleridir [15]. *Kendine Dikkat* katmanının h tane başı vardır ve katmanın çıktısı aşağıdaki şekilde bulunmaktadır:

$$\begin{aligned} \text{Kendine Dikkat}(\mathbf{X}) &= \text{Birleştir}(\text{Baş}_1, \dots, \text{Baş}_h) \mathbf{W}^O, \\ \text{Baş}_i &= \text{Dikkat}(\mathbf{X} \mathbf{W}_i^Q, \mathbf{X} \mathbf{W}_i^K, \mathbf{X} \mathbf{W}_i^V), \\ \text{Dikkat}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) &= \text{Softmax}\left(\frac{\mathbf{Q} \mathbf{K}^T}{\sqrt{d_k}}\right) \mathbf{V}. \end{aligned} \quad (5)$$

Yukarıdaki eşitliklerde *Birleştir* dikkat başlarını birleştiren fonksiyondur. $h = 12$ 'dir. \mathbf{W}_i^O , \mathbf{W}_i^Q , \mathbf{W}_i^K ve \mathbf{W}_i^V ağırlık

¹https://github.com/koc-lab/yargitay_retrieval_dataset

TABLO I: TPCR veri kümesi ile alınan ortalama test skorları. Skorlar yüzde ile verilmiştir. En iyi skorlar kalın gösterilmektedir.

Yöntem	Mikro-F1@5 (%)	Mikro-Keskinlik@5 (%)	Mikro-Hassasiyet@5 (%)	Mikro-F1@9 (%)	Mikro-Keskinlik@9 (%)	Mikro-Hassasiyet@9 (%)
LSTM & ATIRE BM25	33,96	47,21	26,52	41,02	40,80	41,24
BERTurk	18,87	26,17	14,76	19,46	19,31	19,60
BERTurk-Legal	38,14	52,89	29,82	42,85	42,53	43,17

matrisleridir. d_k eğitilebilir bir parametredir ve K^T , K 'nin transpozudur [15], [19]. İkinci alt katmanın çıktısı aşağıdaki şekilde bulunmaktadır:

$$O_2 = \text{Katman Normu}(O_1 + \text{İleri Besleme}(O_1)). \quad (6)$$

Yukarıdaki eşitlikte O_2 BERTurk-Legal katmanının çıktısıdır. *İleri Besleme* iki lineer katmanı olan ve bu katmanlar arasında ReLU aktivasyonu bulunan bir FFNN ağıdır [15].

BERTurk-Legal'in parametrelerinin başlangıç değerleri BERTurk modelinden alınmıştır. Bu şekilde BERTurk-Legal modeline genel Türkçe bilgisi transfer edilmiştir. BERTurk'un eğitiminde kullanılan Türkçe derlemi günlük dilde yazılmış metinlerden oluşmaktadır ve derlemelerin toplam boyutu 35 GB'tır [23]. BERTurk-Legal'in ön eğitimi için eğitim kümesi olarak [22] çalışmasında yayınlanan Türk Hukuk Derlemi [22] isimli derlemdeki Yargıtay karar metinleri kullanılmıştır. Derlem Türk yüksek mahkemelerine ait karar metinlerinden oluşmaktadır. Derlemde 332.662 Yargıtay karar metni bulunmaktadır. Bu metinlerin toplam boyutu 1.198 MB'tır. BERTurk'un eğitimi tamamı küçük harfli olan metinlerle yapıldığı için Türk Hukuk Derlemi de BERTurk-Legal'in eğitiminde kullanılmadan önce tamamen küçük harfli hâle getirilmiştir. Ön eğitim yöntemi olarak maskeli dil modelleme kullanılmıştır. Bu kapsamda eğitim kümesindeki kelimelerin %20'si maskelenmiştir ve BERTurk-Legal modeli maskelenmiş kelimeleri tahmin etmeye çalışarak eğitilmiştir.

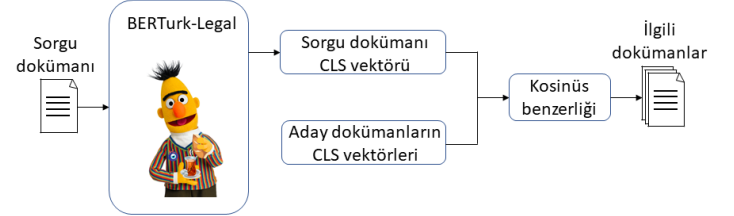
Eğitimi tamamlanan BERTurk-Legal modeli kullanılarak TPCR veri kümesindeki karar metinlerinin doküman vektörleri elde edilmiştir. BERTurk-Legal çıktısındaki ilk token vektörü doküman vektörü olarak kullanılmıştır. İlk token CLS diye adlandırılan, her girdi metnin başına eklenen ve temsil vektörü girdi metnin tamamıyla ilgili bilgi içeren özel bir token'dır [19]. Emsal kararı bulunmak istenen sorgu dokümanı için aday dokümanlar aşağıdaki formüle göre sıralanmaktadır:

$$R_i = \text{argsort}\{S\} \\ S = \{sim(d_i, d_1), sim(d_i, d_2), \dots, sim(d_i, d_N)\} \quad (7) \\ \wedge sim(d_i, d_i) \notin S\}[0 : k].$$

Yukarıda eşitlikte R_i i 'nci doküman sorgulandığında benzerliğine göre sıralanmış dokümanların ilk k tanesinin indisleridir. $sim(d_i, d_j)$ i ve j endeksli iki dokümanın benzerlik skorudur. Benzerlik hesabı için doküman vektörlerinin kosinüs benzerliği alınmıştır. N , TPCR veri kümesindeki doküman sayısıdır. argsort fonksiyonu kendisine verilen listedeki değerleri çoktan aza olacak şekilde sıralamaktadır. BERTurk-Legal'in kullanımı Şekil 1'de gösterilmiştir.

IV. DENEYLER VE SONUÇLARI

Bildiride BERTurk-Legal ile yapılan çalışmalara ek olarak [6]'daki en iyi model olan LSTM & ATIRE BM25 de yeniden



Şekil 1: BERTurk-Legal'in emsal karar bulma için kullanımı. BERTurk-Legal kullanılarak karşılaştırılacak dokümanlar için CLS vektörleri elde edilir. Ardından bu vektörlerin kosinüs benzerliğine bakılarak birbirine benzer dokümanlar bulunur.

eğitilmiştir. Daha stabil sonuçlar elde etmek için BERTurk-Legal ve LSTM & ATIRE BM25 modellerinin üç farklı tohum (*İng. seed*) değeri ile eğitimi yapılmış ve üçer model versiyonu üretilmiştir. TPCR veri kümesi ile yapılan testlerin sonuçları Tablo I'de verilmiştir. Tablodaki sonuçlar üç farklı tohum değeri ile üretilen model versiyonları ile alınan sonuçların ortalamasıdır. BERTurk-Legal modeline yapılan ön eğitimin faydasını göstermek için BERTurk modeli ile de deney yapılmıştır. Yapılan deneyler sonucunda BERTurk-Legal'in tüm deneylerde en iyi sonuçları verdiği görülmektedir. Beklendiği üzere BERTurk-Legal BERTurk'ten çok daha iyi sonuçlar vermiştir. İki model arasında tüm metriklerde %20 civarlarında skor farkı gözlenmiştir. Bu durum dil modellerinin yalnızca günlük dilde yazılmış metinlerle eğitildiğinde hukuksal jargonu anlayamadığını göstermektedir.

BERTurk-Legal deneylerde ikinci en iyi model olan LSTM & ATIRE BM25 modelini tüm metriklerde geride bırakmıştır. BERTurk-Legal ve LSTM & ATIRE BM25 modellerinin eğitimi Türk Hukuk Derlemi ile yapılmıştır. İki model arasındaki başarı farkının önemli bir sebebi ise parametre ilklen-dirmesi için herhangi bir başka modelden yararlanılmamıştır [6]. BERTurk-Legal modeli ise BERTurk modelinin parametrelerini kullanarak Türkçe bilgi transferi yapmıştır. Bu sayede BERTurk-Legal hem genel Türkçeye hem de Türkçe hukuk jargonuna hakim bir model olmuştur. LSTM & ATIRE BM25 modeli ATIRE BM25 isimli kelimelerin bire bir ne kadar uyduğunu hesaplayan bir algoritmadan ve bu algoritmanın oluşturduğu sıralamayı yeniden sıralayan LSTM tabanlı bir otokodlayıcıdan oluşmaktadır [6]. BERTurk-Legal'in bu modele göre daha fazla parametrelili bir dil modeli olması da daha başarılı olmasında önemli bir nedendir.

V. SONUÇ

Günümüzde dönüştürücü bazlı modeller NLP çalışmaları ön plana çıkmaktadır. Bu sebeple bildiride Türk hukuk sisteminde emsal karar bulma görevi için dönüştürücü tabanlı

bir model olan BERTurk-Legal modeli eğitilmiştir ve kullanılmıştır. BERTurk-Legal'in ön eğitimi Türk Hukuk Derlemi [22] isimli Türkçe hukuk için üretilmiş bir veri kümesi ile yapılmıştır. Bu veri kümesi emsal kararlar ile ilgili herhangi bir etiket bulundurmamaktadır. BERTurk-Legal maskeli dil modelleme kullanılarak kendiliğinden denetimli bir şekilde eğitilmiştir. BERTurk-Legal sınıflandırma görevi üzerinde eğitilmeksizin -diğer bir deyişle literatürde zero-shot sınıflandırma olarak geçen yöntem ile- Yargıtay davalarından oluşan TPCR isimli veri kümesinde literatürdeki en iyi sonuçları vermiştir. Türkçe hukuk alanında yapılacak gelecek NLP çalışmalarında dönüştürücü modellerinin kullanımının daha da artması beklenmektedir. Bilgisayar donanımları geliştikçe 512 token'dan daha uzun metinlerin işlenmesine imkan veren dönüştürücü modelleri uzun hukuk metinlerini işlemek için daha sık kullanılabilir.

BİLGİLENDİRME

A. Koç'un çalışmaları BAGEP 2023 Bilim Ödülü tarafından desteklenmektedir. Bu çalışma Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) (1001-120E346) fonuyla desteklenmiştir.

KAYNAKLAR

- [1] F. Kort, "Predicting Supreme Court Decisions Mathematically: A Quantitative Analysis of the "Right to Counsel" Cases," *American Political Science Review*, vol. 51, no. 1, p. 1–12, 1957.
- [2] K. D. Ashley, "Case-based reasoning and its implications for legal expert systems," *Artificial Intelligence and Law*, vol. 1, pp. 113–208, 1992.
- [3] V. Aleven, "Using background knowledge in case-based legal reasoning: A computational model and an intelligent learning environment," *Artificial Intelligence*, vol. 150, pp. 183–237, 11 2003.
- [4] T. Bench-Capon, A. M. Araszkiwicz, A. K. Ashley, K. Atkinson, F. Bex, and et al., "A history of AI and Law in 50 papers: 25 years of the int. conf. on AI and Law," *Artificial Intelligence and Law*, vol. 20, pp. 215–319, 2012.
- [5] I. Chalkidis, M. Fergadiotis, P. Malakasiotis, N. Aletras, and I. Androutsopoulos, "LEGAL-BERT: The Muppets straight out of Law School," in *Findings of the Association for Computational Linguistics: EMNLP 2020*. Online: ACL, Nov. 2020, pp. 2898–2904.
- [6] C. E. Öztürk, O. Köksal, Ş. B. Özçelik, and A. Koç, "Prior Case Retrieval for the Court of Cassation of Turkey," in *2022 16th Int. Conf. on Signal-Image Technology & Internet-Based Systems (SITIS)*, 2022, pp. 539–544.
- [7] V. Tran, M. L. Nguyen, and K. Satoh, "Building legal case retrieval systems with lexical matching and summarization using a pre-trained phrase scoring model," in *Proc. of the Seventeenth Int. Conf. on Artificial Intelligence and Law*, ser. ICAIL '19. New York, NY, USA: ACM, 2019, p. 275–282.
- [8] S. Althammer, S. Hofstätter, M. Sertkan, S. Verberne, and A. Hanbury, "PARM: A Paragraph Aggregation Retrieval Model for Dense Document-to-Documents Retrieval," in *Advances in Information Retrieval: 44th European Conf. on IR Research, ECIR 2022, Stavanger, Norway, April 10–14, 2022, Proceedings, Part I*. Berlin, Heidelberg: Springer-Verlag, 2022, p. 19–34.
- [9] B. Ali, R. More, S. Pawar, and G. K. Palshikar, "Prior Case Retrieval using Evidence Extraction from Court Judgements," in *Joint Proc. of the Workshops on Automated Semantic Analysis of Information in Legal Text (ASAIL 2021) & AI and Intelligent Assistance for Legal Professionals in the Digital Workplace (LegalAIIA 2021)*. São Paulo, Brazil: CEUS-WS.org, 2021, pp. 1–11.
- [10] I. Chalkidis, I. Androutsopoulos, and N. Aletras, "Neural Legal Judgment Prediction in English," in *Proc. of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: ACL, Jul. 2019, pp. 4317–4323.
- [11] E. Mumcuoğlu, C. E. Öztürk, H. M. Ozaktas, and A. Koç, "Natural language processing in law: Prediction of outcomes in the higher courts of Turkey," *Information Processing & Management*, vol. 58, no. 5, 2021.
- [12] E. Leitner, G. Rehm, and J. Moreno-Schneider, "Fine-grained named entity recognition in legal documents," in *Int. Conf. on Semantic Systems*. Springer, 2019, pp. 272–287.
- [13] C. Çetindağ, B. Yazıcıoğlu, and A. Koç, "Named-entity recognition in Turkish legal texts," *Natural Language Engineering*, vol. 29, no. 3, p. 615–642, 2023.
- [14] N. Sevim, F. Şahinuç, and A. Koç, "Gender bias in legal corpora and debiasing it," *Natural Language Engineering*, vol. 29, no. 2, p. 449–482, 2023.
- [15] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, and et al., "Attention is All you Need," in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., vol. 30. Curran Associates, Inc., 2017.
- [16] J. Niklaus, V. Matoshi, P. Rani, A. Galassi, M. Stürmer, and I. Chalkidis, "LEXTREME: A Multi-Lingual and Multi-Task Benchmark for the Legal Domain," 2023.
- [17] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," in *1st Int. Conf. on Learning Representations (ICLR), Workshop Track Proceedings*, Y. Bengio and Y. LeCun, Eds., Scottsdale, Arizona, USA, 2013.
- [18] J. Pennington, R. Socher, and C. Manning, "GloVe: Global Vectors for Word Representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: ACL, Oct. 2014, pp. 1532–1543.
- [19] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proc. of the 2019 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis, Minnesota: ACL, Jun. 2019, pp. 4171–4186.
- [20] OpenAI, "GPT-4 Technical Report," 2023.
- [21] O. Akça, G. Bayrak, A. M. Issifu, and M. C. Ganiz, "Traditional Machine Learning and Deep Learning-based Text Classification for Turkish Law Documents using Transformers and Domain Adaptation," in *2022 Int. Conf. on Innovations in Intelligent Systems and Applications (INISTA)*, 2022, pp. 1–6.
- [22] C. E. Öztürk, Ş. B. Özçelik, and A. Koç, "Predicting Outcomes of the Court of Cassation of Turkey with Recurrent Neural Networks," in *2022 30th Signal Processing and Communications Applications Conf. (SIU)*. Safranbolu, Turkey: IEEE, 2022.
- [23] S. Schweter, "BERTurk - BERT models for Turkish," <https://doi.org/10.5281/zenodo.3770924>, 2020, accessed:2022-06-19.
- [24] J. Rabelo, R. Goebel, Y. Kano, M.-Y. Kim, M. Yoshioka, and K. Satoh, "Summary of the Competition on Legal Information Extraction/Entailment (COLIEE) 2021," in *Proc. of the Eighth Int. Competition on Legal Information Extraction/Entailment (COLIEE 2021)*, ser. COLIEE '21. Sao Paulo, Brazil: University of Alberta, 2021, pp. 1–7.
- [25] P.-S. Huang, X. He, J. Gao, I. Deng, A. Acero, and L. Heck, "Learning deep structured semantic models for web search using clickthrough data," in *Proc. of the 22nd ACM Int. Conf. on Information & Knowledge Management*, 10 2013, pp. 2333–2338.
- [26] H. Palangi, L. Deng, Y. Shen, J. Gao, X. He, and et al., "Deep Sentence Embedding Using Long Short-Term Memory Networks: Analysis and Application to Information Retrieval," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 4, p. 694–707, Apr 2016.
- [27] S. E. Robertson, S. Walker, S. Jones, M. Hancock-Beaulieu, and M. Gatford, "Okapi at TREC-3," in *TREC*, 1994.
- [28] A. C. Aras, C. E. Öztürk, and A. Koç, "Feedforward Neural Network Based Case Prediction in Turkish Higher Courts," in *2022 30th Signal Processing and Communications Applications Conf. (SIU)*. Safranbolu, Turkey: IEEE, 2022, pp. 1–4.