



T.C.

MARMARA UNIVERSITY

FACULTY of ENGINEERING

COMPUTER ENGINEERING DEPARTMENT

CSE4288 Introduction to Machine Learning – Term Project

"Predicting Outcomes of Turkish Constitutional Court Decisions Using Explainable Artificial Intelligence"

Group Members

130319659 - Cihan Erdoğanyılmaz

150120998 - Abdelrahman Zahran

150120997 - Mohamad Nael Ayoubi

150120079 - Omar Sameh Belal

150120055 - Muhammed Talha Karagül

1. Abstract

This project aims to predict the outcomes of Turkish Constitutional Court decisions using machine learning and explainable AI methods. Various models, including Logistic Regression, Naive Bayes, Decision Tree, Random Forest, and Transformers, were employed to classify decisions as "violation" or "no violation." Data preprocessing, including stop word removal, lemmatization, and feature engineering, ensured robust text analysis. Results demonstrate the potential of combining traditional and modern machine learning approaches in the legal domain.

2. Introduction

The judicial system generates substantial textual data, presenting opportunities for automation and prediction. This project focuses on predicting Turkish Constitutional Court decision outcomes. We address challenges such as class imbalance and legal domain-specific preprocessing while evaluating various machine learning methods for explainability and performance.

3. Methodology

Data Collection

The dataset consists of 13,676 individual application decisions, cleaned to retain 13,286 consistent records focused on binary classification ("violation" and "no violation").

Data Preprocessing

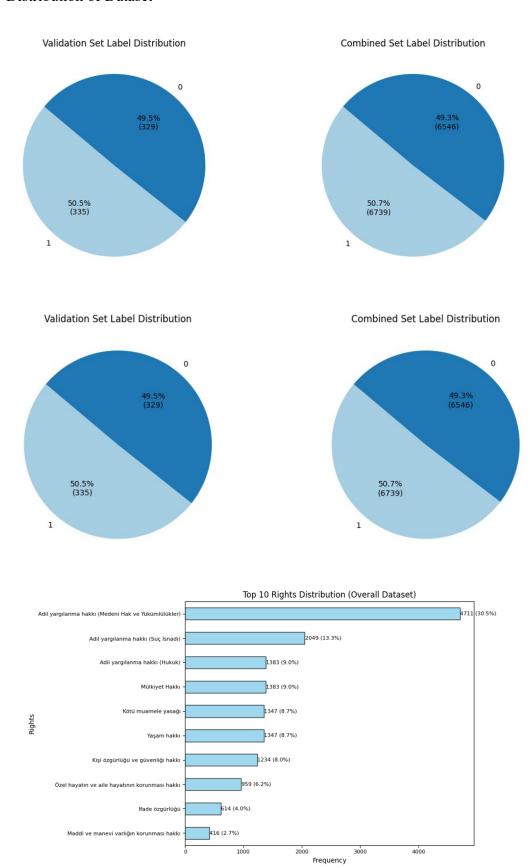
1. Cleaning Steps

- Excluded decisions with multiple inconsistent outcomes.
- o Retained only decisions relevant to the first two significant rights.

2. Text Preprocessing

- o Converted text to lowercase.
- Removed punctuation, irrelevant symbols, and domain-specific stop words (e.g., common Turkish and legal terms).

Distribution of Dataset



o Implemented lemmatization for root word normalization.

3. Feature Engineering

 Used TF-IDF for traditional models and pre-trained tokenizers for Transformers.

TF-IDF for Traditional Models

Term Frequency-Inverse Document Frequency (TF-IDF) is a statistical measure used to evaluate the importance of a word in a document relative to a collection (or corpus) of documents. It is a widely used feature extraction technique for traditional machine learning models in text classification tasks.

Steps Involved

1. Tokenization:

The text was split into individual tokens (words or phrases) to create a vocabulary for the dataset.

2. Term Frequency (TF):

Calculated as the frequency of a term appearing in a document.

$$TF = \frac{\text{number of times the term appears in the document}}{\text{total number of terms in the document}}$$

3. Inverse Document Frequency (IDF):

Assessed the rarity of a term across all documents. Rare terms were given higher weights.

$$IDF = log(\frac{\text{number of the documents in the corpus}}{\text{number of documents in the corpus contain the term}})$$

4. TF-IDF Score:

The final TF-IDF value for each term was calculated as: Product of its TF and IDF scores

5. Why TF-IDF?

• TF-IDF effectively balances frequent terms (e.g., "and," "the") and rare but significant terms in the context (e.g., legal terms like "violation," "admissibility").

 It provides a sparse, numerical representation of text, suitable for traditional machine learning models such as Logistic Regression, Naive Bayes, Decision Tree, and Random Forest.

Implementation

• Libraries Used: Scikit-learn was used for generating the TF-IDF matrix.

• Parameters Tuned:

o max features: Limited the number of features to 5000 to reduce dimensionality.

Modeling Techniques

We used a 90-5-5 train-dev-test split. The models were trained using 90% of the data, evaluated on 5% as the development set for hyperparameter tuning, and finally tested on 5% of the data for the final evaluation.

1. Logistic Regression (handled by Omer)

- o Baseline model for binary classification with weighted loss for imbalanced data.
- 2. Naive Bayes (handled by Muhammed Talha)
 - o Efficient probabilistic model for high-dimensional text data.
- 3. **Decision Tree** (handled by Abdelrahman)
 - o Captures decision-making logic, aiding explainability.
- 4. Random Forest (handled by Muhammed Nael)
 - o Ensemble model combining multiple decision trees for improved accuracy.

5. **Transformers** (handled by Cihan)

 Fine-tuned pre-trained BERT models for contextual understanding of legal texts.

Evaluation Metrics

Models were evaluated using accuracy, precision, recall, and F1 score..

4. Results

1. Model Performance

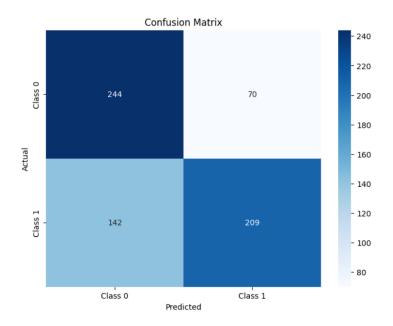
- Logistic Regression achieved baseline accuracy but struggled with class imbalance.
- Naive Bayes demonstrated speed and efficiency but lacked nuanced understanding.
- Decision Tree provided interpretable insights but showed lower generalization.
- Random Forest achieved high accuracy and robustness against overfitting.
- Transformers outperformed traditional models, excelling in contextual understanding and accuracy.

Model	Accuracy	Precision (Class 0)	Precision (Class 1)	Recall (Class 0)	Recall (Class 1)	F1- Score (Class 0)	F1- Score (Class 1)
Naive Bayes	0.67	0.63	0.73	0.76	0.60	0.69	0.66
Logistic Regression	0.78	0.74	0.82	0.82	0.74	0.78	0.78
Decision Tree	0.74	0.79	0.70	0.59	0.86	0.68	0.77
Random Forest	0.78	0.73	0.85	0.86	0.71	0.79	0.77
Transformer (Turkish BERT)	0.80	0.80	0.81	0.78	0.82	0.79	0.81

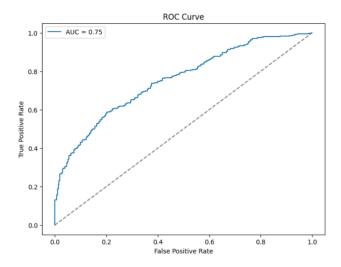
Why Transformers Outperform Traditional Models:

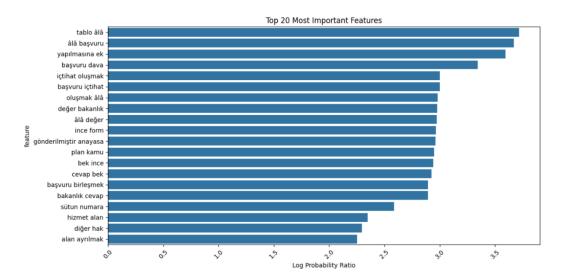
- Contextual Understanding: Transformers use self-attention mechanisms to understand relationships between all words in a sentence, allowing for nuanced comprehension of language.
- **Pre-trained Knowledge**: Models like BERT and GPT are pre-trained on vast corpora, equipping them with domain knowledge that traditional models lack.
- **Adaptability**: Fine-tuning on a specific task enables Transformers to capture subtle domain-specific patterns.

4.1 Naïve Bayes Results

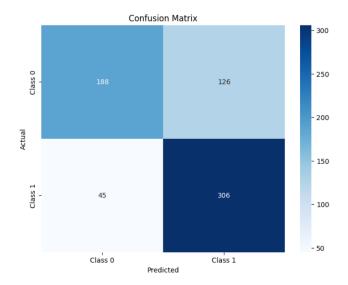


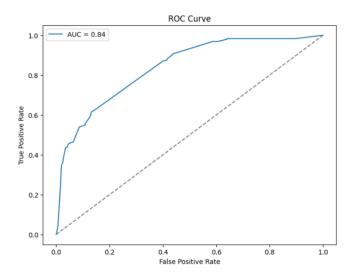
	Mean CV A		3	34 0.676307	45 0.67749604	0.69413629	0.68581616]
		precisi		f1-score	support		
١	Clas	s 0 0.6	63 0.78	0.70	314		
١	Clas	s 1 0.7	75 0.60	0.66	351		
١	accur	acy		0.68	665		
-	macro	avg 0.6	69 0.69	0.68	665		
	weighted	avg 0.6	69 0.68	0.68	665		
ı	ROC-AUC S	core: 0.75					
١	Top 10 Mo	st Discrimina	ative Feature	s:			
-		feature	importance				
-	3839	tablo âlâ	3.718489				
	4734	âlâ başvuru	3.670060				
		pılmasına ek					
1	461	başvuru dava	3.343185				
		ihat oluşmak					
	466 baş	vuru içtihat	3.000796				
	3272	oluşmak âlâ					
	1045 de	ğer bakanlık	2.977153				
	4735	âlâ değer	2.974209				
	2129	ince form	2.966180				





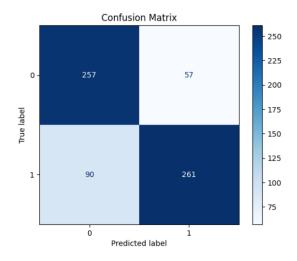
4.2 Decision Tree Results

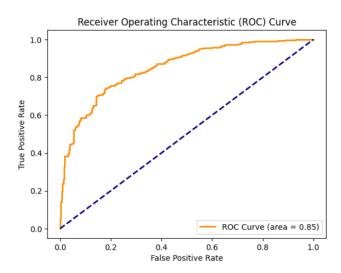




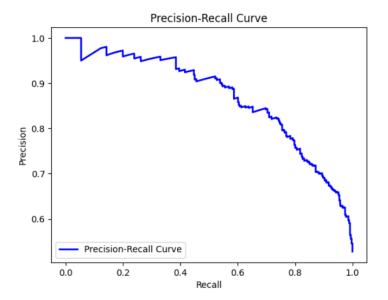
Accuracy: 0.74 Cross-Validati Mean CV Accura Classification	on Scores: [cy: 0.73	0.7227722	8 0.7274168	0.73137876	0.74167987	0.75158479]
	precision	recall	f1-score	support		
Class 0	0.81	0.60	0.69	314		
Class 1	0.71	0.87	0.78	351		
accuracy			0.74	665		
macro avg	0.76	0.74	0.73	665		
weighted avg	0.75	0.74	0.74	665		
ROC-AUC Score:	0.84					

4.3 Logistic Regression Results

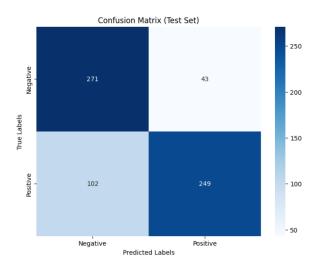


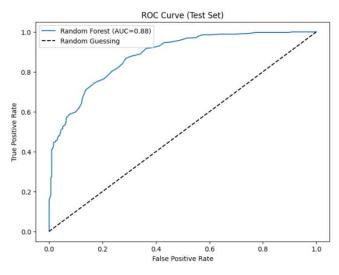


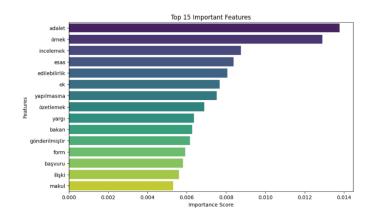
Classification	Report: precision	recall	f1-score	support
0	0.74	0.82	0.78	314
1	0.82	0.74	0.78	351
accuracy			0.78	665
macro avg	0.78	0.78	0.78	665
weighted avg	0.78	0.78	0.78	665

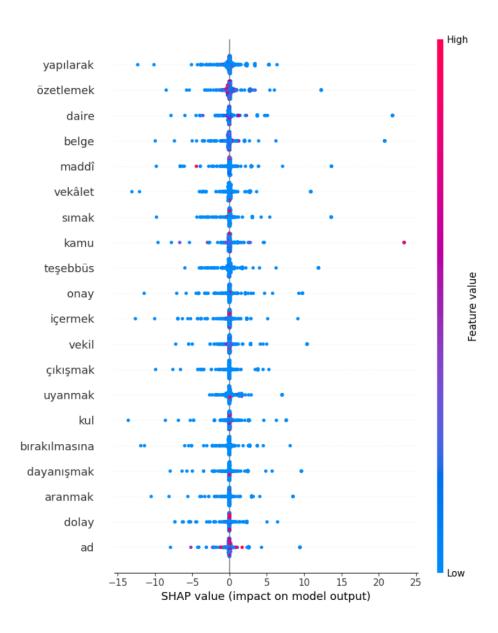


4.4 Random Forest Results





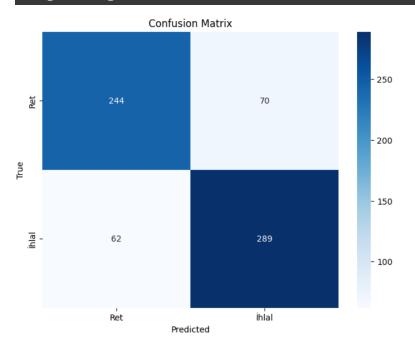




Cross-validat Average CV so CV score std:	ore: 0.76595	483839176		0.7733333	33 0.78544542	0.76178235]
Test Set Perf	ormance:					
	precision	recall	f1-score	support		
0	0.73	0.86	0.79	314		
1	0.85	0.71	0.77	351		
accuracy			0.78	665		
•	0.70	0.70				
macro avg	0.79	0.79	0.78	665		
weighted avg	0.79	0.78	0.78	665		

4.5 Transformer (Turkish BERT)

Accuracy: 0.80 Precision: 0.8 Recall: 0.8015 F1 Score: 0.80 Classification	301411990295 503759398496 313414271708	8007 3		
	precision	recall	f1-score	support
Ret İhlal	0.80 0.81	0.78 0.82	0.79 0.81	314 351
accuracy			0.80	665
macro avg weighted avg	0.80 0.80	0.80 0.80	0.80 0.80	665 665



XAI Visualizations

Input/Output - Heatmap Layout: Top/Bottom >

yargılama dosyası içeniğinden tespit edilen ilgili olaylar özelle şöyledir. Bartın Cumhuriyet Başşavıcılığının 3 / 4 / 2012 tarihli ve E. 2012 / 555 saylıl iddianamesiyle başvurucunun "izinsiz inşai ve fiziki müdahale yapına veya yaptırma" suçundarı oczalandırılması için Bartın Asliye Ceza Mahkemesi, Partın Asliye Ceza Mahkemesi, 21 / 5 / 2014 tarihli ve E. 2012 / 316, K.

2014 / 476 sayılı kararı ile başvurucunun atılı suçtan 1 yıl 8 ay hapis ve 80, 00 TL adıl para cezası ile cezalandırılmasına ve hükmün açıklanmasının geri birakılmasına karar vermiştir. Kararın gerekçesi şöyledir: "
İddia, sanık savunması, katıları veklilini beyanı, mahallınde yapılan keşif, keşif neticesi alınan bilirkir japoru, nüfus ve adli sicil kayıtları ile tüm dosya kapsamı birlikle değerlendirildiğinde, Bartın ili Kirtepe mahallesi 13 pafta 130 ada 7 parsel numarada kayıtı ve Ankara Kültür ve Tabiat Varlıkları Kurulu tarafından 1 / 10 / 1991 tarihinde tescillenmiş ve sanık Süleyman Akbağ İ başvurucu. J tarafından kuyumcu olarak kullanılan iki adet dükkanın sanık tarafından birleştirilmek suretiyle izinsiz uygulamalara tabi tutulduğu, Karabük Kültür Varlıklarını Koruma Bölge Kurulunun 15 / 6 / 2011 tarih ve 2345 sayılı kararı ile yapılan uygulamaların basit onarımı aştiğinin ve restorasyon projesi kapsamında yapılması gerektiğinin, projeler koruma kurulunca onaylarımadan tescili taşımmazda herhangi bir inşai ve fiziki faaliyete bulunulmayacağının belirtildiği, mahallınde yapılan keşif neticesinde alınan inşaat bilirkişisi raporundan da anlaşıldığı üzere, 2.737219 sanık Süleyman Akbağ tarafından kuyumcu olarak kullanılan iki adet dükkanın birleştirilmek sureteyle tek bir mekan oluşturulduğu, bu durumun tescili taşınmazın özyününe aykırı olduğu tüm dosya kapsamından anlaşılmakla, sanık Süleyman Akbağ i nü üzerine atlı 2863 sayılı yasaya muhalefet suçı sabit olmakla eylemine uyan 2863 sayılı yasan 65 / maddesi uyarınca cezalandırılmasına, sanığın sabıkasız oluşu, ikşılığı, duruşmalardakı turum ve davranışları göz örune alınarak yeniden su

olmadiğindan CMK.

231 / 5 - 6 maddesi uyarınca sanık hakkındaki hapse ve adli para cezasına ilişkin hükmün açıklanmasınını geri bırakılmasına, sanığın CMK. 231 / maddesi uyarınca 5 yıl süre ile denetim süresine tabi tutulmasına, bu süre içerisinde sanık hakkında yükümlülük belirlenmesine yer olmadiğına dair hüküm kurulmuştur. "Başvurucu Bartın Asliye Ceza Mahkemesinin hükmün açıklanmasının geri bırakılması kararına karşı Bartın Ağır Ceza Mahkemesine itiraz etmiştir. Bartın Ağır Ceza Mahkemesine itiraz etmiştir. Bartın Ağır Ceza Mahkemesi, 10 / 7 / 2014 tarihin kararıyla başvurucunun itirazının reddine karar vermiştir. Ret kararı başvurucuya 12 / 8 / 2014 tarihinde tebliğ edilmiştir.

Input/Output - Heatmap Layou

Input Text

Başvurucu, hakkında düzenlenen trafik idari para cezası karar tutanağının bir örneğinin on iş günü içinde tebliğ edilmemesinin ve bu cezanın iptali talebinin reddine karar verilmesinin, Anayasa' nın maddesinde belirtilen adil yaryılanma hakkını ihlal ettiğini ileri sürmüştür. Başvuru, 21 / 11 / 2013 tarihinde İstanbul Asilye Hukuk Mahkemesi vastasıyla yaplımıştır. Dilekç ve eklerinin idari yönden yapılan ön incelemesi neticesinde Komisyona sunulmasına engel bir eksikliğin bulunmadığı tespit edilmiştir. Birinci Bölüm Birinci Komisyonunca, 28 / 2 / 2 / 2015 tarihinde kabul edilebilirlik incelemesinin Bölüm tarafından yapılmasına, dosyanın Bölüme gönderilmesine karar verilmiştir. A. Olaylar Başvuru formu ve eklerinde ifade edildiği şekliyle olaylar özetle şöyledir : İstanbul Emniyet Müdürlüğü Trafik Denetleme Şübe Müdürlüğü ekliplerince, 26 / 12 / 2012 tarihli ve 67 - 58873 seri numaralı trafik idari para cezası karar tutanağı ile 13 / 10 / 1983 tarihli ve 2918 sayılı Karayolları Trafik Kanunu' nun maddesine istinaden başvurucu hakkında 319, 00 TL idari para cezası kesilmiştir. Başvurucu tarafından, bahse konu idari para cezasının iptali için İstanbul Anadolu Sulh Ceza Mahkemesine başvuruda bulunulmuş, Mahkemece 18 / 9 / 2013 tarihli ve 2013 / 427 iş sayılı karar ile başvurunun reddine karar verilmiştir. Anılan karar, 30 / 3 / 2005 tarihli ve 5326 sayılı Kabahatler Kanunu' nun maddesinini (10) numaralı fıkrası uyarınca "Üçbin Türk Lirası dahlı İdari para cezalarına karşı başvuru üzerine verilen kararlar' kesin nitelikte olmasına rağmen tiraz yolu açık olmak üzere verilmiştir. Başvurucu, anılan karara karşı İstanbul Anadolu Asliye Ceza Mahkemesine tiraz yolu açık olmak üzere verilmiştir. Başvurucu, anılan karara karşı İstanbul Anadolu Asliye Ceza Mahkemesine tiraz yelinesine yer olmadığı' yönünde büküm kurulmuştur. Anılan karar, başvurucuya 22 / 10 / 2013 tarihinde tebliğ edilmiştir. Sireysel başvuru üzerine verilmiştir. Anılan karar, başvurucuya 22 / 10 / 2013 tarihinde tebliğ edilmiştir. Sireysel başv kılınmış personel taratından araçıarın tesci piakasılıa göre düzenienen traik loarı para cezasi karar tutanakları ... b) Tutanağın tebligat işlemlerine inlallı tespit edildiği tarither itibaren on iş gönü içinde başlanarak, Karayolları Trafik Kanununun 116 ncı maddesi hükümleri uyarınca araç sahibine, sahibi birden fazla ise tescil kaydının ilk sırasındaki sahibine posta yoluyla tebligat yapılır. "Yargıtay Ceza Dairesinin 29 / 4 / 2013 tarihli ve E. 2012 / 31671, K. 2013 / 10142 sayılı kararının ilgili kismı söyledir." Mahkemece. 27 / 01 / 1989 tarihli ve 20062 sayılı Resmi Gazete 'de yavımlanarak

Output Text



Input Text

Başvuru, disiplin cezasına karşı yapılan itirazın reddine dair işlemin iptali istemiyle açılan davanın süre aşımı yönünden reddedilmesi nedeniyle mahkemeye erişim hakkının ihlal edildiği iddiasına ilişkindir. Başvuru 10 / 12 / 2014 tarihinde yapılmıştır. Başvuru, başvuru formu ve eklerindi nidari yönden yapılan ön incelemesinden sonra Komisyona sunulmuştur. Komisyonca başvurunun kabul edilebilirlik incelemesinin Bölüm tarafından yapılmasına karar verilmiştir. Bolum Başkanı tarafından başvurunun kabul edilebilirlik ve esas incelemesinin birlikte yapılmasına karar verilmiştir. Başvuru tormu ve eklerinde if alda edildiği şekliyle ilgili olaya özelte şöyledir. İzmir PTT Başmıddırdiğü Dağıtımı ve Toplama Merkezi Müdurlüğünde şef olarak görev yapan başvurucu, disiplin amirinin 12 / 9 / 2012 tarihli işlemi ile 14 / 7 / 1965tarihli ve 657 sayılı Devlet Memurları Kanunu 'nun maddesinin B - (a) bendi uyarınca kınama cezası ile cezalandırılmıştır. Başvurucunun söz konusu disiplin cezasına karşı 28 / 9 / 2012 tarihlinde yapıtığı itirazı Disiplin Kuruluk nararınını pitali istemiyle 25 / 3 / 2013 tarihinde yapıtığı itirazı Disiplin Kuruluk nararınını pitali istemiyle 25 / 3 / 2013 tarihinde zimir İdare Mahkemesinde (Mahkeme) dava açmıştır. Mahkemenin 27 / 11 / 2013 tarihinde yapıtığı itirazı Disiplin Kuruluk nararınını pitali istemiyle 25 / 3 / 2013 tarihinde izmir İdare Mahkemesinde (Mahkeme) dava açmıştır. Mahkemenin 27 / 11 / 2013 tarihinde yapıtığı itirazı Disiplin cezasının gerekçesinde özetle şu hususlara yer verilmiştir: Öncelikle 657 sayılı Kanun'un maddesi ile disiplin cezalarına karşı yapılacak itiraz, İtiraz mercii ve itirazın tabi olacağı süre yonlerinden özel bir düzenleme yapıldığı, düzenlemenin özel niteliğinden dolayı söz konusu itirazın 6 / 1 / 1 / 1982 tarihli ve 2577 sayılı İdari Yargılama Usulu Kanunu'nun maddesi kapsamında değerlendirilemeyeceği vurgulamıştır. Disiplin cezası aldığına dair tebligat yapılanı iğilinin bu işleme karşı döğrudan dava açmaşvarakı tiraz yoluna başvurusu halik a

Misclassified instance

Input/Output - Heatmap	Layout : <u>Left/Right</u> ✓
Input Text	Output Text
Başvurucu. uzun süre devam eden tutukluluk baskı altında ve müdafi olmaksızın alınan ifadelerin hükme esas alınması ve hükuka aykırı arama yapılması nedeniyle Anayasa "ını ve maddelerinde düzenlenen kişi hürnyeti ve güvenliği edili yapılmam hakkının hilad elidiğini ileri sirmiştir. Regiveniyləri ile adili yapılmam hakkının hilad elidiğini ileri sirmiştir. Regiveniyləri ile adili yapılmam hakkının hilad elidiğini ileri sirmiştir. Regiveniyləri ileri kabil delileriliri incelemesini bölüm tararlındar. Birini Bölüm Qüroni Komiyoruna. 25 / 12/ 2012 tarihinde, kabil delileriliri incelemesini bölüm tararlındar şirmiştir. Birini Bölüm Qüroni komiyoruna. 25 / 12/ 2012 tarihinde, kabil delileriliri incelemesini bölüm tararlındar yapılmasına, dosyanın Bölüme gürderilmesine karar verilmiştir. Birini tararlındar 12 / 12/ 2013 tarihinde, başvurunanı kabılı edilerilirik ve esası hakkındaki incelemesinin biriliriliriliri. Başvuru konusu olay ve olgular. 15 / 12/ 2013 tarihinde Adalet Bakanlığına gönderilmesine karar verilmiştir. 30.72/17. Başvuru konusu olay ve olgular. 15 / 12/ 2013 tarihinde Adalet Bakanlığın gönderilmesine karar verilmiştir. 30.74/12/12 Başvuru konusu olay ve olgular. 15 / 12/ 2013 tarihinde Adalet Bakanlığın göndərilmesine karar verilmiştir. 30.74/12/12 Başvuru konusu olay ve olgular. 15 / 12/ 2013 tarihinde Adalet Bakanlığın göndərilmiştir. Adalet Bakanlığın göndərilmiştir. Adalet Bakanlığın göndərilmiştir. 30.75/12/13 tarihinde görüş başvurucun yalışılı süresinden sorra 30.75/12/13 tarihinde görüş başvurucun yalışılı süresinden sorra 30.75/12/13 tarihinde görüş başvurucun yalışılı süresinden sorra 30.75/12/13 tarihinde görüş başvurucun yalışılı süresinden sorra 30.75/12/13 tarihinde görüş başvurucun yalışılı (10.75/12/13) tarihinde deriliriliriliriliriliriliriliriliriliril	Rec Intal

5. Discussion

- Challenges

• **Legal-specific preprocessing**: Processing legal texts, including tasks such as Turkish stopword removal, lemmatization, and domain-specific tokenization, was time-intensive but necessary for improving model accuracy.

- Interpretation

- Traditional models: Logistic Regression, Naive Bayes, Decision Tree, and Random
 Forest models performed adequately. Among these, Random Forest provided the
 best balance of explainability and performance, offering a strong trade-off between
 interpretability and predictive power.
- Transformers (Turkish BERT): While the transformer model excelled in terms of
 accuracy and demonstrated superior contextual understanding, it was
 computationally intensive and required significant hardware resources for training
 and inference.

- Solutions

Domain-specific preprocessing: The application of legal-specific preprocessing, including Turkish stopword removal and lemmatization, improved the models' relevance and performance in legal text analysis.

6. Conclusion

This project demonstrates the potential of machine learning in legal text analysis. Combining traditional and modern approaches improved prediction accuracy and interpretability. Future work could explore ensemble strategies combining Transformers with Random Forest or other traditional models. Further domain adaptation and model fine-tuning will enhance practical application.

7. References

- [1] E. Mumcuoglu, 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, p. 102684, 2021.
- [2] A. C. Aras, C. E. Öztürk and A. Koç, "Feedforward Neural Network Based Case Prediction in Turkish Higher Courts," *2022 30th Signal Processing and Communications Applications Conference (SIU)*, Safranbolu, Turkey, 2022, pp. 1-4.
- [3] M. F. Sert, E. Yıldırım, and İ. Haşlak, "Using artificial intelligence to predict decisions of the Turkish Constitutional Court," *Social Science Computer Review*, vol. 40, no. 6, pp. 1416–1435, 2021.
- [4] E. Aydemir, "Estimation of Turkish Constitutional Court Decisions in Terms of Admissibility with NLP," 2023 IV International Conference on Neural Networks and Neurotechnologies (NeuroNT), Saint Petersburg, Russian Federation, 2023, pp. 17-20.
- [5] 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," 2022 International Conference on Innovations in Intelligent SysTems and Applications (INISTA), Biarritz, France, 2022, pp. 1-6.
- [6] C. E. Öztürk, Ş. B. Özçelik and A. Koç, "Predicting Outcomes of the Court of Cassation of Turkey with Recurrent Neural Networks," 2022 30th Signal Processing and Communications Applications Conference (SIU), Safranbolu, Turkey, 2022, pp. 1-4.
- [7] M. B. L. Virtucio et al., "Predicting Decisions of the Philippine Supreme Court Using Natural Language Processing and Machine Learning," 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), Tokyo, Japan, 2018, pp. 130-135.
- [8] O.-M. Sulea, M. Zampieri, M. Vela, and J. van Genabith, "Predicting the law area and decisions of French Supreme Court cases," in *Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017*, Varna, Bulgaria, Sep. 2017, pp. 716–722.
- [9] C. Erdoğanyımaz and B. Mengünoğul, "An Original Natural Language Processing Approach to Language Modeling in Turkish Legal Corpus: Improving Model Performance with Domain Classification by Using Recurrent Neural Networks," 2022 Innovations in Intelligent Systems and Applications Conference (ASYU), Antalya, Turkey, 2022, pp. 1-6.

- [10] Erdoğanyılmaz, C., Mengünoğul, B., & Balci, M. (2023, September). Unveiling the Black Box: Investigating the Interplay between AI Technologies, Explainability, and Legal Implications. In 2023 8th International Conference on Computer Science and Engineering (UBMK) (pp. 569-574). IEEE.
- [11] Turan, T., Küçüksille, E., & Alagöz, N. K. (2023). Prediction of Turkish Constitutional Court Decisions with Explainable Artificial Intelligence. Bilge International Journal of Science and Technology Research, 7(2), 128-141.
- [12] A.Vaswani et al., "Attention is all you need," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.*, 2017, pp. 6000–6010.
- [13] E. Roth, *The Shapley Value: Essays in Honor of Lloyd S. Shapley*. Cambridge, U.K.: Cambridge Univ. Press, 1988.
- [14] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 4768–4777.

8. Appendices

Code Snippets

Code example of using shap library for extracting features that have the higher impact

Model training, performing cross validation and lastly predicting code example:

```
# Load data
train_data = load_data('train_processed.json')
dev_data = load_data('dev_processed.json')
test_data = load_data('test_processed.json')

# Combine train and dev data
train_dev_data = pd.concat([train_data, dev_data])
X_train_dev = train_dev_data['text']
y_train_dev = train_dev_data['labels']
X_test = test_data['text']
y_test = test_data['labels']

# Vectorize text data using TF-IDF
vectorizer = I_fidfVectorizer(max_features=5000)
X_train_dev_vectorized = vectorizer.fit_transform(X_train_dev).toarray()
X_test_vectorized = vectorizer.transform(X_test).toarray()

# Initialize and train Random Forest
clf = RandomForestClassifier(n_estimators=100, random_state=42)

cv_scores = cross_val_score(clf, X_train_dev_vectorized, y_train_dev, cv=5, scoring='f1')
print('Ncross-validation scores:", cv_scores)
print("Average CV score:", cv_scores.mean())
print("CV score std:", cv_scores.std())

# Train final model on full training data
clf.fit(X_train_dev_vectorized, y_train_dev)

# Evaluate on test set
y_pred = clf.predict(X_test_vectorized)
y_pred_proba = clf.predict(_proba(X_test_vectorized))[:, 1]
```