MAKÜ MMF
MÜHENDİSLİK-MİMARLIK FAKÜLTESİ

Uluslararası Mühendislik, Tasarım ve Teknoloji Dergisi, 2024, 6(2), Sayfa: 53-63
International Journal of Engineering, Design and Technology, 2024, 6(2), Page: 53-63

Araştırma Makalesi
Research Article

Legal text classification in Turkey: A machine learning approach to divorce and zoning decisions

Türkiye'de hukuki metin sınıflandırması: Boşanma ve imar kararlarına makine öğrenmesi yaklaşımı

Tülay Turan[1*], Ecir Uğur Küçüksille[2]

[1]Department of Computer Technologies, Burdur Mehmet Akif Ersoy University, Burdur, Türkiye
[2]Department of Computer Engineering, Süleyman Demirel University, Isparta, Türkiye

Abstract: The increasing volume of legal data in recent years requires integrating artificial intelligence (AI) techniques for efficient management and use. Critical challenges include classifying legal texts into specific fields or topics. This is crucial to advancing legal research and practice. This article aims to categorically classify Turkish court decisions, an area that has yet to be adequately researched before, compared to classification studies in international law texts. The study aims to contribute significantly to developing artificial intelligence-supported solutions to guide Turkish legal decisions by dividing legal texts into specific areas, thus increasing the efficiency and accessibility of the legal system. The study first created a data set consisting of divorce and zoning cases. Then, basic models were established with K-Nearest Neighbor (KNN), Support Vector Machines (SVM), Decision Trees (DT), and Random Forests (RF) algorithms to determine the algorithm that would classify the cases most effectively. Hyperparameter optimization was performed for each model to increase the Base Model performances. This process was supported by the 10-fold cross-validation method. Improved models were established with the hyperparameter values obtained as a result of optimization. As a result of comparative analysis, the SVM model had an impressive 90% accuracy rate in classifying legal texts. This result will significantly contribute to the development of intelligent legal systems by achieving significant success in classifying legal texts in Turkey.

Keywords: Legal Text Classification, Turkish Court Decisions, Machine Learning Algorithms, Hyperparameter Optimization, SVM

Özet: Son yıllarda artan hukuki veri hacmi, verimli yönetim ve kullanım için Yapay Zeka (AI) tekniklerinin entegrasyonunu gerektirmektedir. Kritik zorluklar arasında hukuki metinlerin belirli alanlara veya konulara göre sınıflandırılması yer almaktadır; bu, hukuki araştırma ve uygulamanın ilerletilmesi açısından çok önemlidir. Bu makale, uluslararası hukuk metinlerindeki sınıflandırma çalışmalarına kıyasla daha önce yeterince araştırılmamış bir alan olan Türk mahkeme kararlarını kategorik olarak sınıflandırmayı amaçlamaktadır. Çalışma, hukuki metinleri belirli alanlara ayırarak, Türk hukuki kararlarında yönlendirmeye yönelik yapay zeka destekli çözümlerin geliştirilmesine önemli ölçüde katkıda bulunmayı ve böylece hukuk sisteminin verimliliğini ve erişilebilirliğini artırmayı amaçlamaktadır. Çalışmada ilk olarak boşanma ve imar davalarından oluşan bir veri seti oluşturulmuştur. Daha sonra davaları en etkin şekilde sınıflandıracak algoritmayı belirlemek için K-En Yakın Komşu (KNN), Destek Vektör Makineleri (SVM), Karar Ağaçları (DT) ve Rastgele Ormanlar (RF) algoritmaları ile temel modeller kurulmuştur. Temel Model performanslarını arttırmak için her bir model için hiperparametre optimizasyonu gerçekleştirilmiştir. Bu süreç, 10 katlı çapraz doğrulama yöntemi ile desteklenmiştir. Optimizasyon sonucunda elde edilen hiperparametre değerleri ile iyileştirilmiş modeller kurulmuştur. Karşılaştırmalı analiz sonucunda, SVM modeli hukuki metinlerin sınıflandırılmasında %90 gibi etkileyici bir doğruluk oranına sahip olmuştur. Bu sonuç, Türkiye'deki hukuki metinlerin sınıflandırılmasında önemli bir başarıyı elde ederek, akıllı hukuk sistemlerinin gelişimine önemli katkılar sağlayacaktır.

Anahtar kelimeler: Hukuki Metin Sınıflandırması, Türk Mahkeme Kararları, Makine Öğrenmesi Algoritmaları, Hiperparametre Optimizasyonu, SVM

1. Introduction

Legal text classification is a critical process in legal informatics. It aims to systematically classify legal documents into predefined classes or categories according to their content and thematic relevance (Boella et al., 2011). This process facilitates the effective management, access, and analysis of legal texts, thereby increasing the accessibility of legal information and supporting legal research and decision-making processes. Legal text classification leverages advances in Natural Language Processing (NLP) and Machine Learning (ML) to automate the analysis of complex legal language and document structures (Li et. al., 2020).

* İletişim Yazarı / *Corresponding author*. Eposta/*Email* : tulayturan@mehmetakif.edu.tr
Geliş / *Received*: 05.03.2024, Revizyon / *Revised*: 14.04.2024
Kabul / *Accepted*: 16.04.2024

Burdur Mehmet Akif Ersoy University Press

Content of this journal is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Studies in the literature have focused primarily on legal texts from various jurisdictions and languages, reflecting the global interest in legal informatics and the universal challenges posed by legal text classification. Much research has been conducted on English-language legal texts focusing on courts' decisions in the United States, the United Kingdom, and international tribunals such as the European Court of Human Rights (ECtHR). Aletras et al. (2016) have established a precedent for predictive legal analytics using NLP and ML by predicting the outcomes of ECHR cases by analyzing case texts. Chalkidis, Androutsopoulos, et al. (2019) focused on predicting the legal verdict on ECHR cases. Chen et al. (2022) focused on text classification for US legal texts and used pre-trained word embedding-based deep learning algorithms. Studies of Chinese legal texts have analyzed decisions from various levels of the Chinese judiciary, highlighting the challenges and strategies specific to the Chinese legal language. In their study called IFlyLegal, Wang et al. (2019) introduced an integrated system that performs legal document analysis by using deep contextual representations and various attention mechanism techniques. Lei et al. (2017) automatically classified Chinese decision documents using machine learning algorithms. Their study stated that they created a vector space model using TF-IDF after segmenting the words. Zhang et al. (2022) proposed a Chinese legal element identification method based on BERT's contextual relationship capture mechanism to identify elements by measuring the similarity between legal elements and case descriptions. In their study, Benedetto et al. (2023) focused on Italian legal document classification. Their study used the BERT pre-trained Language model for legal document classification. Tagarelli and Simeri (2022) used Italian BERT pre-trained on the Italian civil code or its sections for the classification task. Nineesha and Deepalakshmi (2022) used Indian legal documents to classify legal texts in their study. In their research, they compared the performance of deep learning techniques. Kalia et al. (2022) developed models with a Support Vector Machine, Naive Bayes, K-Nearest Neighbor, and Decision Tree to classify case events of the Central Information Commission of India.

In the Republic of Turkey, which has a rich legal heritage, applying artificial intelligence techniques in legal text classification is essential and valuable. Despite growing research on the classification of legal texts in various jurisdictions, studies focusing on the Turkish legal system are scarce (Sert et. al., 2022; Aydemir, 2023; Görentaş et. al., 2023; Turan et. al., 2023). This article aims to fill this gap by developing a methodology for automatically classifying Turkish court decisions, focusing on two common cases (zoning and divorce). For this purpose, a data set containing divorce and zoning cases was first created in the study. The data set was put through natural language processing processes. After the DDI steps, the data set was divided into 80% training and 20% test sets using the holdout method. Then, basic models were established using KNN, SVM, DT, and RF machine learning algorithms. Hyperparameter optimization was performed to increase the performance of the models. This process was supported by the 10-fold cross-validation method. As a result, the SVM model achieved an impressive 90% accuracy rate in classifying legal texts.

This study aims to contribute to the development of intelligent legal systems by providing a concrete methodology and application example for studies in the field of legal text classification in Turkey. In addition, it will serve as an essential reference point for developing classification models appropriate to the characteristics of local legal systems.

2. Materials and Methods

2.1. Dataset

In order to develop natural language processing applications in the field of law, it is important to first examine the online decision sharing of these courts and determine in which court field the study data set will be created. Table 1 shows online decision sharing information of the courts of the Republic of Turkey.

In sharing the decision contents, the first instance court decisions of the Judicial Judiciary and Administrative Jurisdiction were examined. The first point that draws attention in these reviews is that there needs to be a platform where the state shares decisions. As a result, commercial web pages sharing legal decisions were searched and decision examples were accessed on the Lexpera

Table 1. Online decision pages of the courts of the Republic of Turkey (Turan, 2023)

	Court	Link
Judiciary	First Instance Courts	https://www.lexpera.com.tr/
	Regional Courts of Justice	https://emsal.uyap.gov.tr/ https://www.lexpera.com.tr/
	Supreme Court	https://karararama.yargitay.gov.tr/
Administrative jurisdiction	First Instance Courts	https://www.lexpera.com.tr/
	Regional Administrative Courts	https://emsal.uyap.gov.tr/ https://www.lexpera.com.tr/
	Council of state	https://karararama.danistay.gov.tr/
Constitutional Judiciary	Constitutional Court	https://www.anayasa.gov.tr/tr/kararlar-bilgi-bankasi/

website. A data set was created by downloading the first instance court decisions, including zoning and divorce case decisions, from the relevant page to be used in the study. The representation of the most frequently encountered words in the texts of zoning and divorce decisions in the data set is given in Figure 1.

Figure 1. Display of the most frequently encountered words in the court texts included

2.2. Natural Language Processing

NLP is a subfield of artificial intelligence (AI) and linguistics that focuses on the interaction between computers and human (natural) languages. It aims to enable computers to understand, interpret, and produce human language in a valuable and meaningful way. NLP combines computational linguistics (rule-based modeling of human language) with statistical, machine learning, and deep learning models. These technologies allow computers to process human language in the form of text or audio data and 'understand' the whole meaning of the speaker or writer, together with their intent and feelings (Hirschberg and Manning, 2015).

NLP is crucial to data analytics; It enables automatic analysis of large volumes of text data to gain insight, identify trends, and detect sentiment. This is particularly relevant in marketing, customer service, healthcare, and legal informatics, where understanding human communication can provide a competitive advantage and operational efficiency (Zhong et.al., 2020). In law, NLP facilitates the analysis of legal documents, case law, and legislation, helping legal professionals quickly find relevant information and make more informed decisions (Ashley and Brüninghaus, 2009; Alarie et. al., 2018). Figure 2 shows the classification of legal texts into divorce and zoning cases using natural language processing.

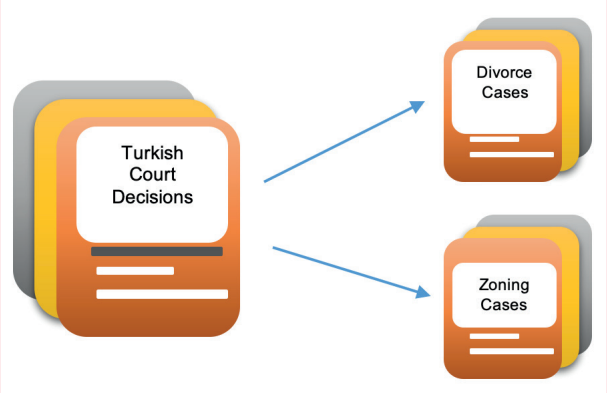


Figure 2. Classification of legal texts into divorce and zoning cases using Natural Language Processing

2.3. Classification Algorithms

K-Nearest Neighbor

The KNN algorithm is a simple yet powerful non-parametric method for classification. It operates on a very intuitive principle: an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors (k is a positive integer, typically small). If k=1, then the object is assigned to the class of its nearest neighbor (Laaksonen and Oja, 1996; Mucherino et. al., 2009).

In the basic equation, the distance between two points (e.g., x and y) is typically calculated using the Euclidean distance, though other distances (Manhattan, Minkowski, etc.) can be used depending on the context:

$$d(x,y) = \sqrt{\sum_{i}^{n}(x_i - y_i)^2} \quad (1)$$

where n is the number of dimensions (features) and x_i, y_i are the values of the ith feature for points x and y, respectively.

Support Vector Machines

SVM is a set of supervised learning methods used for classification, regression, and outliers detection. The basic idea behind SVM is to find the hyperplane that best divides a dataset into two classes. The hyperplane's equation can be written as follows (Gunn, 1998; Hearst et.al., 1998; Awad et.al., 2015):

$$w.x - b = 0 \quad (2)$$

where w represents the weight vector, x is the input features, and b is the bias. The weight vector w determines the orientation of the hyperplane, while the bias b determines the distance of the hyperplane from the origin. Together, they define the decision boundary: points for which $w \cdot x - b > 0$ fall into one class, while those for which $w \cdot x - b < 0$ fall into the other. SVMs aim to maximize the margin between the classes, which is inversely proportional to the norm of w.

Decision Trees

DT is a nonparametric supervised learning method used for classification and regression. A decision tree builds classification or regression models in the form of a tree structure, breaking down a dataset into smaller subsets while simultaneously developing an associated decision tree incrementally. The final result is a tree with decision nodes and leaf nodes (Gupta et.al., 2017; Patel and Prajapati, 2018; Charbuty and Abdulazeez, 2021).

The decision at each node is made based on the feature x_i that maximizes or minimizes a certain criterion, such as the Gini impurity or the information gain. The information gain is calculated as follows:

$$IG(D_p, x_i) = I(D_p) - \sum_j^m \frac{N_j}{N_p} I(D_j) \quad (3)$$

where $IG(Dp,xi)$ is the information gain of parent dataset Dp split on feature xi, I is the impurity measure (e.g., entropy for information gain), Np is the number of samples in the parent dataset, Nj is the number of samples in the jth child dataset, and m is the number of child datasets. The equation calculates the reduction in impurity or "entropy" achieved by splitting the parent dataset Dp on feature xi. The goal is to maximize this reduction, thereby making the data subsets as pure as possible at each step of the tree.

Random Forests

RF is an ensemble learning method for classification, regression, and other tasks. It operates by constructing a multitude of decision trees at training time and outputting the class, that is, the mode of the classes (classification) or mean prediction (regression) of the individual trees (Breiman, 2001; Cutler et.al., 2007; Ali et.al., 2012).

Random Forests do not have a simple equation like the algorithms mentioned above, as they are based on the aggregation of the results of multiple decision trees. However, the general principle of operation can be described as:

$$RF(x) = \frac{1}{B} \sum_{b=1}^{B} T_b(x) \quad (4)$$

where $RF(x)$ is the prediction of the Random Forest for input B is the number of trees in the forest, and $Tb(x)$ is the prediction of the bth decision tree. This formula represents the aggregation process in Random Forests, where the outputs of multiple decision trees are combined to produce a final result. For classification tasks, this typically involves selecting the most frequent prediction (mode) among all trees in the forest.

2.4. Hyperparameter Optimization with GridSearchCv

Optimizing machine learning algorithms is a critical step in the model development, ensuring that models achieve the best possible performance on given tasks (Feurer and Hutter, 2019; Yang and Shami, 2020). This process involves setting the model's hyperparameters, the configuration settings used to configure machine learning models.

GridSearchCV is a hyperparameter tuning technique that comprehensively searches a specific subset of hyperparameters, evaluating and comparing the performance of models trained with each combination through cross-validation (Kartini et.al., 2021; Alhakeem et.al., 2022). The goal is to find the optimal set of hyperparameters that results in the best model performance.

The equation of GridSearchCV for selecting the optimal hyperparameter set (H*) can be abstractly formulated as:

$$H^* = arg_H max CV_k(M(H, D_{train}), D_{val} \quad (5)$$

- H* is the optimal set of hyperparameters.
- H ranges over the grid of possible hyperparameter combinations.
- CVk represents the K-fold cross-validation process.
- $M(H,Dtrain)$ is the model trained with hyperparameters H on the training dataset $Dtrain$.
- $Dval$ is the validation dataset used in cross-validation.

2.5. Evaluation metrics

Evaluation metrics are crucial for evaluating the performance of classification models and provide information about how well a model's predictions match actual observed results. These measurements help understand the strengths and weaknesses of the model, guiding improvements and adjustments (Hossin and Sulaiman, 2015; Vujović, 2021). Some evaluation metrics commonly used in classification: Accuracy, Precision, Recall (Sensitivity), F1 Score, Area Under the ROC Curve (AUC-ROC), and Confusion Matrix (Fatourechi et.al. 2008; Liu et.al., 2014).

Accuracy is one of the most straightforward metrics used in evaluating classification models. It measures the proportion of correct predictions (both true positives and true negatives) among the total number of cases examined.

The accuracy of a model is calculated as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

where:

- TP (True Positives) is the number of correct positive predictions,
- TN (True Negatives) is the number of correct negative predictions,
- FP (False Positives) is the number of incorrect positive predictions, and
- FN (False Negatives) is the number of incorrect negative predictions.

3. Experiments

This study consists of six main chapters focusing on classifying Turkish legal texts. First, the Turkish legal dataset was pre-processed by applying natural language processing (NLP) techniques. This pre-processing phase involves making the texts ready for analysis. Secondly, a weighting process was applied to determine the importance of each term in the vector space model using the Term Frequency Inverse Document Frequency (tf-idf) method. This method assigns importance to words in the text based on their distribution in the document collection. In the third section, four different classification models based on supervised learning were developed with KNN, SVM, DT, and RF. These models are designed to be used to classify Turkish legal texts. In the fourth stage, hyperparameter optimization was performed to improve the performance of these classification models. This optimization helped determine the necessary model parameters to ensure the best performance. The fifth part includes the installation of tuned models after optimization processes. This phase involves creating final versions of the developed classification models by integrating the optimization process results. In the last section, the classification performances of the developed models were compared using the accuracy evaluation metric. This evaluation reveals the effectiveness and reliability of the models in the legal text classification task. The study's methodology is visualized with the system diagram presented in Figure 3.

Firstly, legal experts were interviewed to analyze the divorce and zoning decision texts to be used in the study. As a result of these interviews, information not necessary for the research, such as the subject of the application and the application process, was removed from the decision texts. Figure 4 shows the sample structure of the decision text, and Figure 5 shows the sample structure obtained after the data preprocessing process.

The court decision texts were later added to the Excel work file. In the study, decisions regarding "zoning" are defined as "1," and decisions regarding "divorce" are defined as "0" in the Excel file. Figure 6 shows the Excel data set's first 20 decision texts and class information.

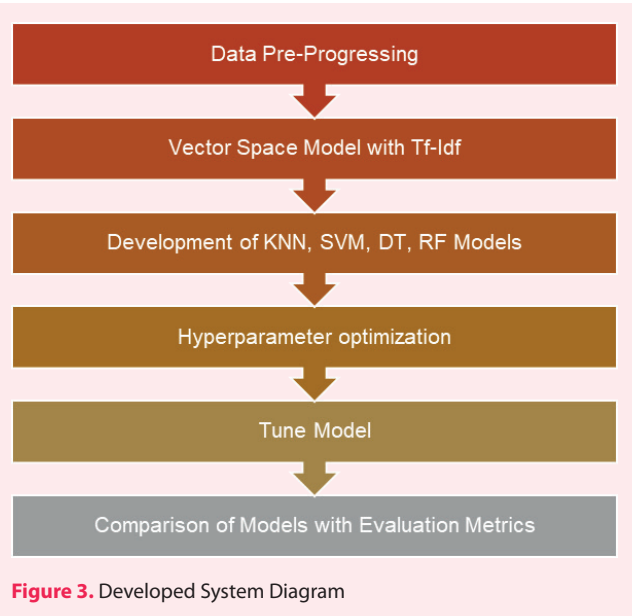


Figure 3. Developed System Diagram

In this research, the classification study of Turkish legal texts begins with applying natural language processing techniques. First, a comprehensive data preprocessing process was applied to make the legal texts suitable for analysis. This process involves converting texts into a standard form and cleaning out elements unnecessary for analysis. The data preprocessing steps performed on the data in the study are shown in Figure 7.

Figure 8 shows the data's appearance before and after applying stopwords, noisy entity removal, and lowercase operations, which are data pre-processing steps. After these processes, frequently recurring words and noisy data were removed from the data, and uppercase letters were converted to lowercase letters.

The study used the Term Frequency - Inverse Document Frequency (TF-IDF) method for vector representation and classification of Turkish legal texts. TF-IDF calculates the importance of a word in a document by evaluating the frequency of that word in the document (TF) and the rarity of the word in the entire document collection (IDF) (Yun-tao at.al., 2005; Bafna et.al., 2016). This is a practical feature extraction method frequently used in text mining and natural language processing. The TF-IDF method used in classifying Turkish legal texts is a powerful tool to reveal semantic differences between documents and represent texts. This method allows the creation of high-quality features necessary for training classification models.

When calculating the *tf* text frequency value, the ratio of the number of terms in the sentence to the total number of words in the sentence is considered. The calculation method is shown in Equation 7.

$$tf(t,d) = \frac{f_{t,d}}{\sum t' \epsilon\, d\, ft',d} \tag{7}$$

The *idf* importance value is the base two logarithms of the ratio of the total number of sentences to the total number of the selected term in all sentences. The calculation method is shown in Equation 8.

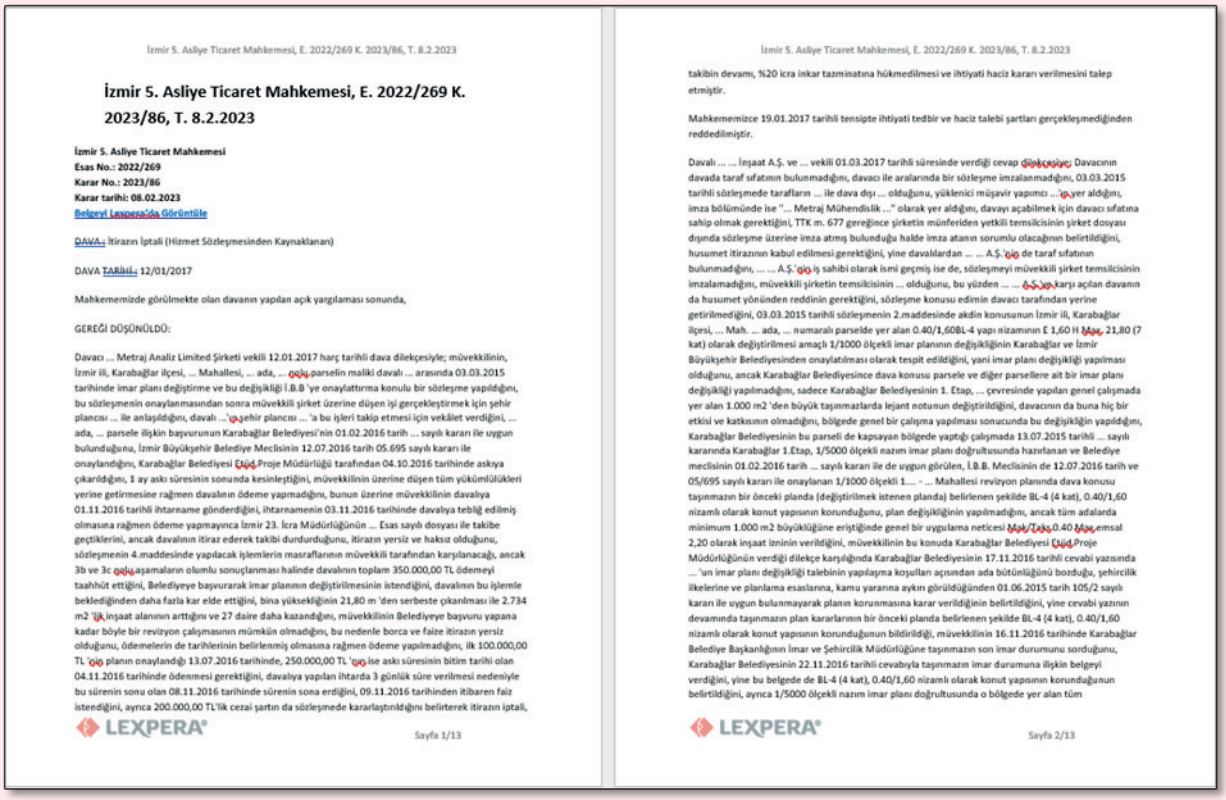
$$idf = (t,D) = \log \frac{N}{|\{d \epsilon D : t \epsilon d\}|} \tag{8}$$

After calculating the tf and idf values, the tf-idf value of each word is obtained by multiplying the two values. The calculation method is shown in Equation 9.

$$tfidf = (t,d,D) = tf(t,d).idf(t,D) \tag{9}$$

In the study, tf-idf vector calculations were made for all words in the data set. Figure 9 shows an example of the calculated tf-idf values of words.

In the third part of the research, models were developed



İzmir 5. Asliye Ticaret Mahkemesi, E. 2022/269 K. 2023/86, T. 8.2.2023

İzmir 5. Asliye Ticaret Mahkemesi, E. 2022/269 K. 2023/86, T. 8.2.2023

İzmir 5. Asliye Ticaret Mahkemesi
Esas No.: 2022/269
Karar No.: 2023/86
Karar tarihi: 08.02.2023
Belgeyi Lexpera'da Görüntüle

DAVA: İtirazın İptali (Hizmet Sözleşmesinden Kaynaklanan)

DAVA TARİHİ: 12/01/2017

Mahkememizde görülmekte olan davanın yapılan açık yargılaması sonunda,

GEREĞİ DÜŞÜNÜLDÜ:

Davacı ... Metraj Analiz Limited Şirketi vekili 12.01.2017 harç tarihli dava dilekçesiyle; müvekkilinin, İzmir ili, Karabağlar ilçesi, ... Mahallesi, ... ada, ... parselin maliki davalı ... arasında 03.03.2015 tarihinde imar planı değiştirme ve bu değişikliği İ.B.B 'ye onaylattırma konulu bir sözleşme yapıldığını, bu sözleşmenin onaylanmasından sonra müvekkili şirket üzerine düşen işi gerçekleştirmek için şehir plancısı ... ile anlaşıldığını, davalı ...'ın şehir plancısı ... 'a bu işleri takip etmesi için vekâlet verdiğini, ... ada, ... parsele ilişkin başvurunun Karabağlar Belediyesi'nin 01.02.2016 tarih ... sayılı kararı ile uygun bulunduğunu, İzmir Büyükşehir Belediye Meclisinin 12.07.2016 tarih 05.695 sayılı kararı ile onaylandığını, Karabağlar Belediyesi Etüd Proje Müdürlüğü tarafından 04.10.2016 tarihinde askıya çıkarıldığını, 1 ay askı süresinin sonunda kesinleştiğini, müvekkilinin üzerine düşen tüm yükümlülükleri yerine getirmesine rağmen davalının ödeme yapmadığını, bunun üzerine müvekkilinin davaliya 01.11.2016 tarihli ihtarname gönderdiğini, ihtarnamenin 03.11.2016 tarihinde davaliya tebliğ edilmiş olmasına rağmen ödeme yapmayınca İzmir 23. İcra Müdürlüğünün ... Esas sayılı dosyası ile takibe geçtiklerini, ancak davalının itiraz ederek takibi durdurduğunu, itirazın yersiz ve haksız olduğunu, sözleşmenin 4.maddesinde yapılacak işlemlerin masraflarının müvekkili tarafından karşılanacağı, ancak 3b ve 3c aşamalarının olumlu sonuçlanması halinde davalının toplam 350.000,00 TL ödemeyi taahhüt ettiğini, Belediyeye başvurarak imar planının değiştirilmesinin istendiğini, davalının bu işlemle beklediğinden daha fazla kar elde ettiğini, bina yüksekliğinin 21,80 m 'den serbeste çıkarılması ile 2.734 m2 'lik inşaat alanının arttığını ve 27 daire daha kazandığını, müvekkilinin Belediyeye başvuru yapana kadar böyle bir revizyon çalışmasının mümkün olmadığını, bu nedenle borca ve faize itirazın yersiz olduğunu, ödemelerin de tarihlerinin belirlenmiş olmasına rağmen ödeme yapılmadığını, ilk 100.000,00 TL 'lik planın onaylandığı 13.07.2016 tarihinde, 250.000,00 TL 'lik ise askı süresinin bitim tarihi olan 04.11.2016 tarihinde ödenmesi gerektiğini, davaliya yapılan ihtarda 3 günlük süre verilmesi nedeniyle bu sürenin sonu olan 08.11.2016 tarihinde sürenin sona erdiğini, 09.11.2016 tarihinden itibaren faiz istendiğini, ayrıca 200.000,00 TL'lik cezai şartın da sözleşmede kararlaştırıldığını belirterek itirazın iptali,

LEXPERA

Sayfa 1/13

İzmir 5. Asliye Ticaret Mahkemesi, E. 2022/269 K. 2023/86, T. 8.2.2023

takibin devamı, %20 icra inkar tazminatına hükmedilmesi ve ihtiyati haciz kararı verilmesini talep etmiştir.

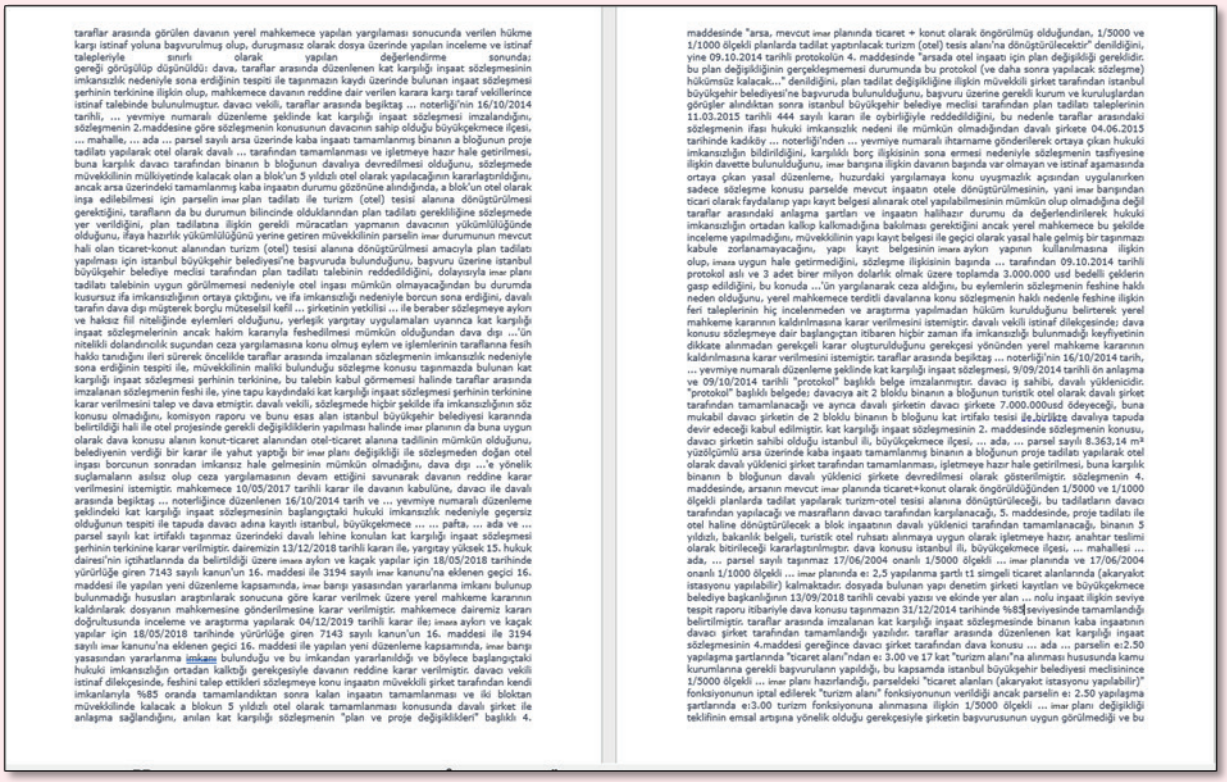
Mahkememizce 19.01.2017 tarihli tensipte ihtiyati tedbir ve haciz talebi şartları gerçekleşmediğinden reddedilmiştir.

Davalı ... ... İnşaat A.Ş. ve ... vekili 01.03.2017 tarihli süresinde verdiği cevap dilekçesiyle; Davacının davada taraf sıfatının bulunmadığını, davacı ile aralarında bir sözleşme imzalanmadığını, 03.03.2015 tarihli sözleşmede taraflarını ... ile dava dışı ... olduğunu, yüklenici müşavir yapımcı ...'ın yer aldığını, imza bölümünde ise "... Metraj Mühendislik ..." olarak yer aldığını, davayı açabilmek için davacı sıfatına sahip olmak gerektiğini, TTK m. 677 gereğince şirketin münferiden yetkili temsilcisinin şirket dosyası dışında sözleşme üzerine imza atmış bulunduğu halde imza atanın sorumlu olacağının belirtildiğini, husumet itirazının kabul edilmesi gerektiğini, yine davalılardan ... ... A.Ş.'nin de taraf sıfatının bulunmadığını, ... A.Ş.'nin iş sahibi olarak ismi geçmiş ise de, sözleşmeyi müvekkili şirket temsilcisinin imzalamadığını, müvekkili şirketin temsilcisinin ... olduğunu, bu yüzden ... ... A.Ş.'nin karşı açılan davanın da husumet yönünden reddinin gerektiğini, sözleşme konusu edimin davacı tarafından yerine getirilmediğini, 03.03.2015 tarihli sözleşmenin 2.maddesinde akdin konusunun İzmir ili, Karabağlar ilçesi, ... Mah. ... ada, ... numaralı parselde yer alan 0.40/1,60BL-4 yapı nizamının E 1,60 H Max. 21,80 (7 kat) olarak değiştirilmesi amaçlı 1/1000 ölçekli imar planının değişikliğinin Karabağlar ve İzmir Büyükşehir Belediyesinden onaylatılması olarak tespit edildiğini, yani imar planı değişikliği yapılması olduğunu, ancak Karabağlar Belediyesince dava konusu parsele ve diğer parsellere ait bir imar planı değişikliği yapılmadığını, sadece Karabağlar Belediyesinin 1. Etap, ... çevresinde yapılan genel çalışmada yer alan 1.000 m2 'den büyük taşınmazlarda lejant notunun değiştirildiğini, davacının da buna hiç bir etkisi ve katkısının olmadığını, bölgede genel bir çalışma yapılması sonucunda bu değişikliğin yapıldığını, Karabağlar Belediyesinin bu parseli de kapsayan bölgede yaptığı çalışmada 13.07.2015 tarihli ... sayılı kararında Karabağlar 1.Etap, 1/5000 ölçekli nazım imar planı doğrultusunda hazırlanan ve Belediye meclisinin 01.02.2016 tarih ... sayılı kararı ile de uygun görülen, İ.B.B. Meclisinin de 12.07.2016 tarih ve 05/695 sayılı kararı ile onaylanan 1/1000 ölçekli 1.... - ... Mahallesi revizyon planında dava konusu taşınmazın bir önceki planda (değiştirilmek istenen planda) belirlenen şekilde BL-4 (4 kat), 0.40/1,60 nizamlı olarak konut yapısının korunduğunu, plan değişikliğinin yapılmadığını, ancak tüm adalarda minimum 1.000 m2 büyüklüğüne eriştiğinde genel bir uygulama neticesi Max/Taks 0.40 Max emsal 2,20 olarak inşaat izninin verildiğini, müvekkilinin bu konuda Karabağlar Belediyesi Etüd Proje Müdürlüğünün verdiği dilekçe karşılığında Karabağlar Belediyesinin 17.11.2016 tarihli cevabi yazısında ... 'un imar planı değişikliği talebinin yapılaşma koşulları açısından ada bütünlüğünü bozduğu, şehircilik ilkelerine ve planlama esaslarına, kamu yararına aykırı görüldüğünden 01.06.2015 tarih 105/2 sayılı kararı ile uygun bulunmayarak planın korunmasına karar verildiğinin belirtildiğini, yine cevabi yazının devamında taşınmazın plan kararlarının bir önceki planda belirlenen şekilde BL-4 (4 kat), 0.40/1,60 nizamlı olarak konut yapısının korunduğunun bildirildiği, müvekkilinin 16.11.2016 tarihinde Karabağlar Belediye Başkanlığının İmar ve Şehircilik Müdürlüğüne taşınmazın son durumunu sorduğunu, Karabağlar Belediyesinin 22.11.2016 tarihli cevabıyla taşınmazın imar durumuna ilişkin belgeyi verdiğini, yine bu belgede de BL-4 (4 kat), 0.40/1,60 nizamlı olarak konut yapısının korunduğunun belirtildiğini, ayrıca 1/5000 ölçekli nazım imar planı doğrultusunda o bölgede yer alan tüm

LEXPERA

Sayfa 2/13

Figure 4. Decision text example structure



taraflar arasında görülen davanın yerel mahkemece yapılan yargılaması sonucunda verilen hükme karşı istinaf yoluna başvurulmuş olup, duruşmasız olarak dosya üzerinde yapılan inceleme ve istinaf talepleriyle sınırlı olarak yapılan değerlendirme sonunda; gereği görüşülüp düşünüldü: dava, taraflar arasında düzenlenen kat karşılığı inşaat sözleşmesinin imkansızlık nedeniyle sona erdiğinin tespiti ile tasınmazın kaydı üzerinde bulunan inşaat sözleşmesi şerhinin terkinine ilişkin olup, mahkemece davanın reddine dair verilen karara karşı taraf vekillerince istinaf talebinde bulunulmuştur. davacı vekili, taraflar arasında beşiktaş ... noterliği'nin 16/10/2014 tarihli, ... yevmiye numaralı düzenleme şeklinde kat karşılığı inşaat sözleşmesi imzalandığını, sözleşmenin 2.maddesine göre sözleşmenin konusunun davacının sahip olduğu büyükçekmece ilçesi, ... mahalle, ... ada ... parsel sayılı arsa üzerinde kaba inşaatı tamamlanmış binanın a bloğunun proje tadilatı yapılarak otel olarak davalı ... tarafından tamamlanması ve işletmeye hazır hale getirilmesi, buna karşılık davacı tarafından binanın b bloğunun davaliya devredilmesi olduğunu, sözleşmede müvekkilinin mülkiyetinde kalacak olan a blok'un 5 yıldızlı otel olarak yapılacağının kararlaştırıldığını, ancak arsa üzerindeki tamamlanmış kaba inşaatın durumu gözönüne alındığında, a blok'un otel olarak inşa edilebilmesi için parselin imar planı tadilatı ile turizm (otel) tesisi alanına dönüştürülmesi gerektiğini, tarafların da bu durumun bilincinde olduklarından plan tadilatı gerekliliğine sözleşmede yer verildiğini, plan tadilatına ilişkin gerekli müracatları yapmanın davacının yükümlülüğünde olduğunu, ifaya hazırlık yükümlülüğünü yerine getiren müvekkilinin parselin imar durumunun mevcut hali olan ticaret-konut alanından turizm (otel) tesisi alanına dönüştürülmesi amacıyla plan tadilatı yapılması için istanbul büyükşehir belediyesi'ne başvuruda bulunduğunu, başvuru üzerine istanbul büyükşehir belediye meclisi tarafından plan tadilatı talebinin reddedildiğini, dolayısıyla imar planı tadilatı talebinin uygun görülmemesi nedeniyle otel inşası mümkün olmayacağından bu durumda kusursuz ifa imkansızlığının ortaya çıktığını, ve ifa imkansızlığı nedeniyle borcun sona erdiğini, davalı tarafın dava dışı müşterek borçlu müteselsil kefil ... şirketinin yetkilisi ... ile beraber sözleşmeye aykırı ve haksız fiil niteliğinde eylemleri olduğunu, yerleşik yargıtay uygulamaları uyarınca kat karşılığı inşaat sözleşmelerinin ancak hakim kararıyla feshedilmesi mümkün olduğundan dava dışı ...'ün nitelikli dolandırıcılık suçundan ceza yargılamasına konu olmuş eylem ve işlemlerinin taraflarına fesih hakkı tanıdığını ileri sürerek öncelikle taraflar arasında imzalanan sözleşmenin imkansızlık imar planı nedeniyle sona erdiğinin tespiti ile, müvekkilinin maliki bulunduğu sözleşme konusu taşınmazda bulunan kat karşılığı inşaat sözleşmesi şerhinin terkinine, bu talebin kabul görmemesi halinde taraflar arasında imzalanan sözleşmenin feshi ile, yine tapu kaydındaki kat karşılığı inşaat sözleşmesi şerhinin terkinine karar verilmesini talep ve dava etmiştir. davalı vekili, sözlesmede hiçbir şekilde ifa imkansızlığının söz konusu olmadığını, komisyon raporu ve bunu esas alan istanbul büyükşehir belediyesi kararında belirtildiği hali ile otel projesinde gerekli değişikliklerin yapılması halinde imar planının da buna uygun olarak dava konusu alanın konut-ticaret alanından otel-ticaret alanına tadilinin mümkün olduğunu, belediyenin verdiği bir karar ile yahut yaptığı bir imar planı değişikliği ile sözleşmeden doğan otel inşası borcunun sonradan imkansız hale gelmesinin mümkün olmadığını, dava dışı ...'e yönelik suçlamaların asılsız olup ceza yargılamasının devam ettiğini savunarak davanın reddine karar verilmesini istemiştir. mahkemece 10/05/2017 tarihli karar ile davanın kabulüne, davacı ile davalı arasında beşiktaş ... noterliğince düzenlenen 16/10/2014 tarih ve ... yevmiye numaralı düzenleme şeklindeki kat karşılığı inşaat sözleşmesinin başlangıçtaki hukuki imkansızlık nedeniyle geçersiz olduğunun tespiti ile tapuda davacı adına kayıtlı istanbul, büyükçekmece ... ... pafta, ... ada ve ... parsel sayılı kat irtifaklı taşınmaz üzerindeki davalı lehine konulan kat karşılığı inşaat sözleşmesi şerhinin terkinine karar verilmiştir. dairemizin 13/12/2018 tarihli kararı ile, yargıtay yüksek 15. hukuk dairesi'nin içtihatlarında da belirtildiği üzere imara aykırı ve kaçak yapılar için 18/05/2018 tarihinde yürürlüğe giren 7143 sayılı kanun'un 16. maddesi ile 3194 sayılı imar kanunu'na eklenen geçici 16. maddesi ile yapılan yeni düzenleme kapsamında, imar barışı yasasından yararlanma imkanı bulunup bulunmadığı hususları araştırılarak sonucuna göre karar verilmek üzere yerel mahkeme kararının kaldırılarak dosyanın mahkemesine gönderilmesine karar verilmiştir. mahkemece dairemiz kararı doğrultusunda inceleme ve araştırma yapılarak 04/12/2019 tarihli karar ile imara aykırı ve kaçak yapılar için 18/05/2018 tarihinde yürürlüğe giren 7143 sayılı kanun'un 16. maddesi ile 3194 sayılı imar kanunu'na eklenen geçici 16. maddesi ile yapılan yeni düzenleme kapsamında, imar barışı yasasından yararlanma imkanı bulunduğu ve bu imkandan yararlanıldığı ve böylece başlangıçtaki hukuki imkansızlığın ortadan kalktığı gerekçesiyle davanın reddine karar verilmiştir. davacı vekili istinaf dilekçesinde, feshini talep ettikleri sözleşmeye konu inşaatın müvekkili şirket tarafından kendi imkanlarıyla %85 oranda tamamlandıktan sonra kalan inşaatın tamamlanması ve iki bloktan müvekkilinde kalacak a blokun 5 yıldızlı otel olarak tamamlanması konusunda davalı şirket ile anlaşma sağlandığını, anılan kat karşılığı sözleşmenin "plan ve proje değişiklikleri" başlıklı 4. maddesinde "arsa, mevcut imar planında ticaret + konut olarak öngörülmüş olduğundan, 1/5000 ve 1/1000 ölçekli planlarda tadilat yaptırılacak turizm (otel) tesis alanı'na dönüştürülecektir" denildiğini, yine 09.10.2014 tarihli protokolün 4. maddesinde "arsada otel inşaatı için plan değişikliği gereklidir. bu plan değişikliğinin gerçekleşmemesi durumunda bu protokol (ve daha sonra yapılacak sözleşme) hükümsüz kalacak..." denildiğini, plan tadilat değişikliğine ilişkin müvekkili şirket tarafından istanbul büyükşehir belediyesi'ne başvuruda bulunulduğunu, başvuru üzerine gerekli kurum ve kuruluşlardan görüşler alındıktan sonra istanbul büyükşehir belediye meclisi tarafından plan tadilatı taleplerinin 11.03.2015 tarihli 444 sayılı kararı ile oybirliğiyle reddedildiğini, bu nedenle taraflar arasındaki sözleşmenin ifası hukuki imkansızlık nedeni ile mümkün olmadığından davalı şirkete 04.06.2015 tarihinde kadıköy ... noterliği'nden ... yevmiye numaralı ihtarname gönderilerek ortaya çıkan hukuki imkansızlığın bildirildiğini, karşılıklı borç ilişkisinin sona ermesi nedeniyle sözleşmenin tasfiyesine ilişkin davette bulunulduğunu, imar barışına ilişkin davanın başında var olmayan ve olmayan ve istinaf aşamasında ortaya çıkan yasal düzenleme, huzurdaki yargılamaya konu uyuşmazlık açısından uygulanırken sadece sözleşme konusu parselde mevcut inşaatın otele dönüştürülmesinin, yani imar barışından ticari olarak faydalanıp yapı kayıt belgesi alınarak otel yapılabilmesinin mümkün olup olmadığına değil taraflar arasındaki anlaşma şartları ve inşaatın halihazır durumu da değerlendirilerek hukuki imkansızlığın ortadan kalkıp kalkmadığına bakılması gerektiğini ancak yerel mahkemece bu şekilde inceleme yapılmadığını, müvekkilinin yapı kayıt belgesi ile geçici olarak yasal hale gelmiş bir taşınmazı kabule zorlanamayacağını, yapı kayıt belgesinin imara aykırı yapının kullanılmasına ilişkin olup, imara uygun hale getirmediğini, sözleşme ilişkisinin başında ... tarafından 09.10.2014 tarihli protokol aslı ve 3 adet birer milyon dolarlık olmak üzere toplamda 3.000.000 usd bedelli çeklerin gasp edildiğini, bu konuda ...'ün yargılanarak ceza aldığını, bu eylemlerin sözleşmenin feshine hakkı neden olduğunu, yerel mahkemece terditli davalarına konu sözlesmenin haklı nedenle feshine ilişkin feri taleplerinin hiç incelenmeden ve araştırma yapılmadan hüküm kurulduğunu belirterek yerel mahkeme kararının kaldırılmasına karar verilmesini istemiştir. davalı vekili istinaf dilekçesinde; dava konusu sözleşmeye dair başlangıçtan itibaren hiçbir zaman ifa imkansızlığı bulunmadığı keyfiyetinin dikkate alınmadan gerekçeli karar oluşturulduğunu gerekçesi yönünden yerel mahkeme kararının kaldırılmasına karar verilmesini istemiştir. taraflar arasında beşiktaş ... noterliği'nin 16/10/2014 tarih, ... yevmiye numaralı düzenleme şeklinde kat karşılığı inşaat sözleşmesi. 9/09/2014 tarihli ön anlaşma ve 09/10/2014 tarihli "protokol" başlıklı belge imzalanmıştır. davacı iş sahibi, davalı yüklenicidir. "protokol" başlıklı belgede; davacıya ait 2 bloklu binanın a bloğunun turistik otel olarak davalı şirket tarafından tamamlanacağı ve ayrıca davalı şirketin davacı şirkete 7.000.000usd ödeyeceği, buna mukabil davacı şirketin de 2 bloklu binanın b bloğunu kat irtifakı tesisi ile birlikte davaliya tapuda devir edeceği kabul edilmiştir. kat karşılığı inşaat sözleşmesinin 2. maddesinde sözleşmenin konusu, davacı şirketin sahibi olduğu istanbul ili, büyükçekmece ilçesi, ... ada, ... parsel sayılı 8.363,14 m² yüzölçümlü arsa üzerinde kaba inşaatı tamamlanmış binanın a bloğunun proje tadilatı yapılarak otel olarak davalı yüklenici şirket tarafından tamamlanması, işletmeye hazır hale getirilmesi, buna karşılık binanın b bloğunun davalı yüklenici şirkete devredilmesi olarak gösterilmiştir. sözleşmenin 4. maddesinde, arsanın mevcut imar planında ticaret+konut olarak öngörüldüğünden 1/5000 ve 1/1000 ölçekli planlarda tadilat yapılarak turizm-otel tesisi alanına dönüştürüleceği, bu tadilatların davacı tarafından yapılacağı ve masrafların davacı tarafından karşılanacağı, 5. maddesinde, proje tadilatı ile otel haline dönüştürülecek a blok inşaatının davalı yüklenici tarafından tamamlanacağı, binanın 5 yıldızlı, bakanlık belgeli, turistik otel ruhsatı alınmaya uygun olarak işletmeye hazır, anahtar teslimi olarak bitirileceği kararlaştırılmıştır. dava konusu istanbul ili, büyükçekmece ilçesi, ... mahallesi ... ada, ... parsel sayılı taşınmaz 17/06/2004 onanlı 1/5000 ölçekli ... imar planında ve 17/06/2004 onanlı 1/1000 ölçekli ... imar planında e: 2,5 yapılanma şartlı t1 simgeli ticaret alanlarında (akaryakıt istasyonu yapılabilir) kalmaktadır. dosyada bulunan yapı denetim şirketi kayıtları ve büyükçekmece belediye başkanlığının 13/09/2018 tarihli cevabi yazısı ve ekinde yer alan ... nolu inşaat ilişkin seviye tespit raporu itibariyle dava konusu taşınmazın 31/12/2014 tarihinde %85 seviyesinde tamamlandığı belirtilmiştir. taraflar arasında imzalanan kat karşılığı inşaat sözleşmesinde binanın kaba inşaatının davacı şirket tarafından tamamlandığı yazılıdır. taraflar arasında düzenlenen kat karşılığı inşaat sözleşmesinin 4.maddesi gereğince davacı şirket tarafından dava konusu ... ada ... parselin e:2.50 yapılaşma şartlarında "ticaret alanı"ndan e: 3.00 ve 17 kat "turizm alanı"na alınması hususunda kamu kurumlarına gerekli başvuruların yapıldığı, bu kapsamda istanbul büyükşehir belediyesi meclisinince 1/5000 ölçekli ... imar planı hazırlandığı, parseldeki "ticaret alanları (akaryakıt istasyonu yapılabilir)" fonksiyonunun iptal edilerek "turizm alanı" fonksiyonunun verildiği ancak parselin e: 2.50 yapılaşma şartlarında e:3.00 turizm fonksiyonuna alınmasına ilişkin 1/5000 ölçekli ... imar planı değişikliği teklifinin emsal artışına yönelik olduğu gerekçesiyle şirketin başvurusunun uygun görülmediği ve bu

Figure 5. Sample structure of the decision text obtained as a result of data preprocessing

using KNN, SVM, DT, and RF algorithms. To improve the performance of these models, a comprehensive hyperparameter optimization process was conducted using the GridSearchCV object of the scikit-learn library. In GridSearchCV, separate models were built for each of the determined hyperparameter values, and the hyperparameter combination that provided the highest accuracy rate was determined. Figure 10 shows the accuracy rates of each model depending on parameter changes. Table 2 shows the hyperparameter values that provide the most successful results.

After the hyperparameter optimization process was completed, the performances of the final models were analyzed using the accuracy evaluation criterion. According to these evaluation results, the Support Vector

	Cases	Class
2	müşterek çocuğun yurt dışında bulunan mutat meskenine iade edilmesine kar	0
3	aile mahkemesi tarafından verilen tedbir kararında geçen ifadeler nedeniyle n	0
4	taşınmazın planda önce kamu hizmeti alanına ayrılması sonra da ticari rekrea	1
5	belediyenin dere yatağını imara açmasından kaynaklı çarpık yapılaşma ve DSİ	1
6	aile birliğinin temelden sarsılması nedenine dayalı olarak karar verildi	0
7	taşınmazın koruma amaçlı uygulama planda doğal sit alanı olarak belirlenmes	1
8	aile hayatına saygı hakkının ihlal edildiğine ilişkin iddianın açıkça dayanaktan y	0
9	kişinin maddi ve manevi varlığını koruma ve geliştirme hakkının ihlal edildiğine	0
10	şiddete maruz kalan kadının koruyucu tedbirler kapsamında işyeri değişikliği t	0
11	taşınmazın orman olduğu olgusuna ve orman niteliğindeki taşınmazların zilyetl	1
12	üzerinde aile konutu şerhi bulunan taşınmazla ilgili olarak ortaklığın satış yolu	1
13	taraflar arasında görülen nafaka davasının yapılan muhakemesi sonunda mah	0
14	mülkiyet hakkının ihlalinin sonuçlarının ortadan kaldırılması için yeniden yargıl	1
15	belediyece uygulamaya geçirilen planının kanunda öngörülen şekil şartlarına u	1
16	park ve dinlenme alanı olarak belirlenmesi yönündeki müdahalenin kamu yara	1
17	başvurucu ve eşinin dava konusunda anlaştıklarını tespit ederek tarafların ma	0
18	şeref ve itibar hakkının ihlal edildiğine ilişkin iddianın açıkça dayanaktan yoksu	0
19	özel mülkiyette bulunan taşınmazların kamu hizmeti alanı olarak ayrılmasında	1
20	müşterek çocuğun velayetinin annesi yerine babasına verilmesi nedeniyle ebe	0
21	özel parselasyon sonucu yol olarak terk edilen taşınmaza kamulaştırma işlem	1

Figure 6. Excel representation of data set content

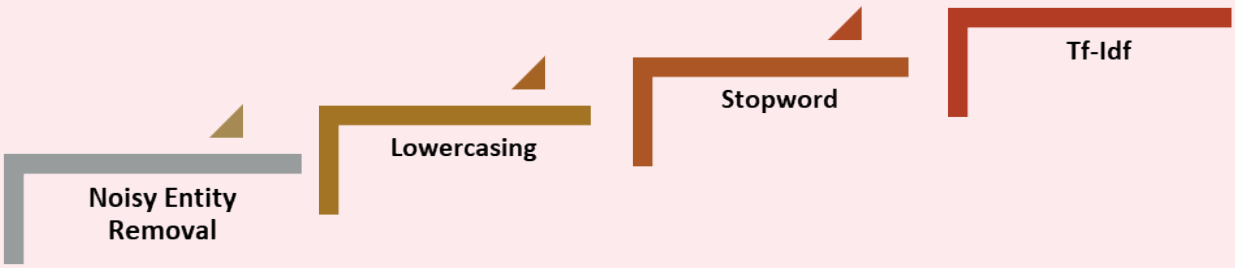


Figure 7. Data preprocessing steps performed in the study

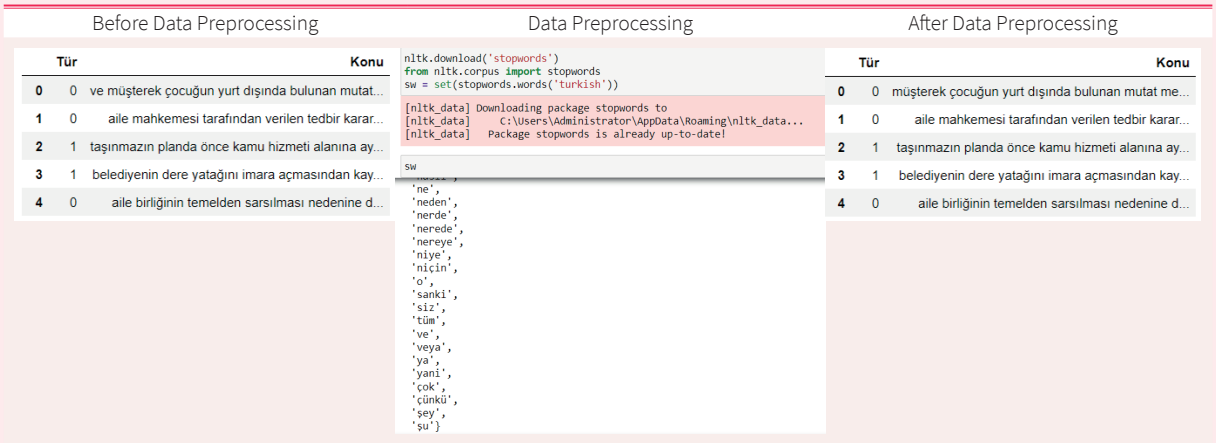


Figure 8. Data appearance before and after applying data preprocessing

Machines model was determined to have the highest performance, with an average accuracy value of 90%. The performance results of all models included in the study are shown in Figure 11.

These findings comparatively evaluate the effectiveness of these classification models on the legal text classification task and show that the SVM model achieves a significantly higher accuracy rate than other models in this particular context. These results underline that using the SVM model in classifying legal texts can produce high-performance results under specific hyperparameter settings.

4. Results and Discussion

This research addresses how Artificial Intelligence (AI) techniques can effectively classify legal texts in the face of increasing volumes of legal data. In particular, a study was conducted to classify court decisions in Turkey categorically. In this context, a data set containing divorce and zoning situations was created, and basic classification models were developed using K-nearest neighbors (KNN), Support Vector Machines (SVM), Decision Trees (DT), and Random Forests (RF) algorithms. Hyperparameter optimization was performed to increase the performance of the models, and improved models were created as a result of this process, which was supported by the 10-fold cross-validation method. The main finding of the research is that the SVM model performs best after hyperparameter optimization with an impressive average accuracy rate of 90%. This result shows that the SVM model is superior to other traditional classification methods in the legal text classification task, especially under specific hyperparameter settings.

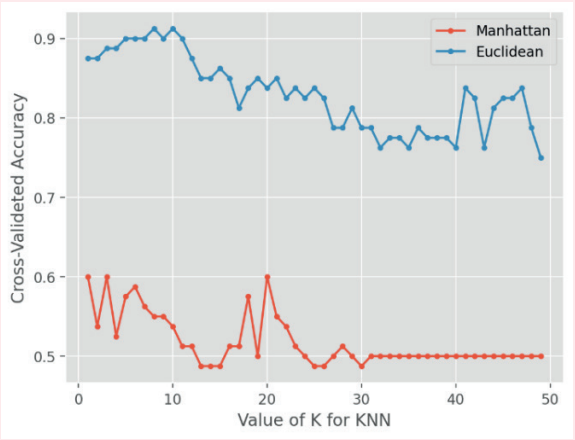
Table 2. Hyperparameter results of models

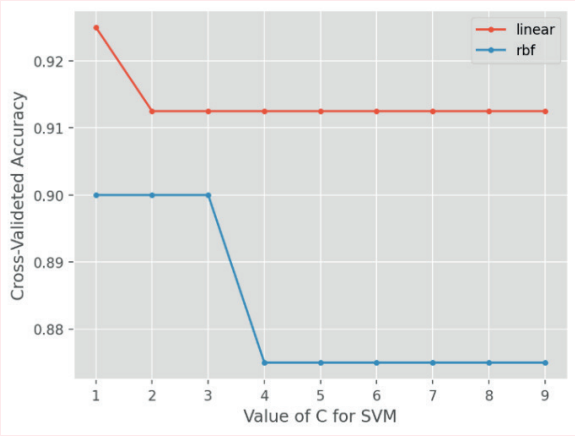
Models	HyperParameters	Value
K-Neaest Neighbors Model	n_neighbors	8
	p(metric)	Euclidean
Support Vector Machines Model	c	1
	kernel	linear
Random Forest Model	max_features	8
	n_estimators	1000
Decision Tree Model	criterion	gini
	max_depth	2
	min_samples_split	3

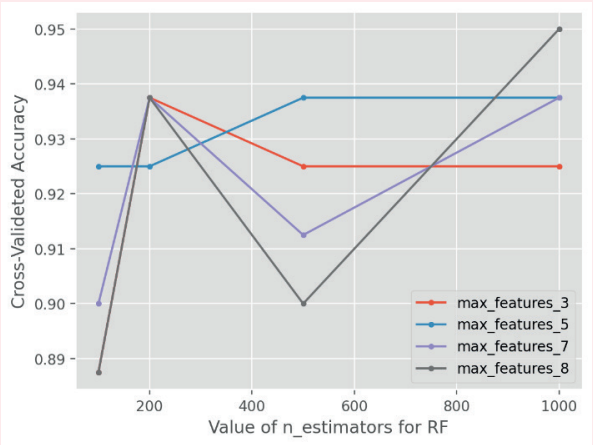
```
from sklearn.feature_extraction.text import TfidfVectorizer
tf=TfidfVectorizer()
text_tf= tf.fit_transform(df['Konu'])
print(text_tf)
```

```
  (0, 652)      0.33707466117274276
  (0, 360)      0.2221086378643965
  (0, 190)      0.33707466117274276
  (0, 284)      0.33707466117274276
  (0, 439)      0.33707466117274276
  (0, 447)      0.3156044606996708
```

Figure 9. Example of calculated TF-IDF values for words in the dataset.







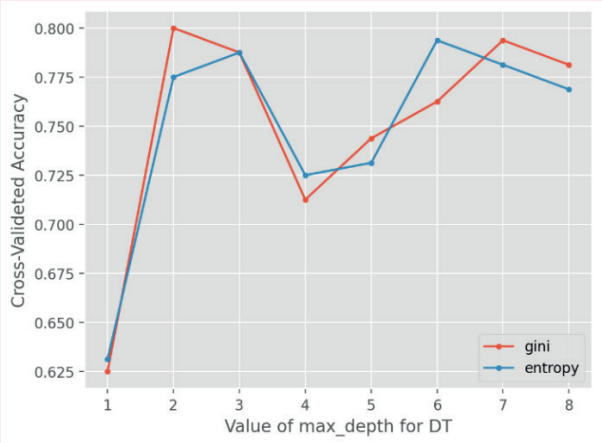


Figure 10. Accuracy rates of models depending on parameter changes

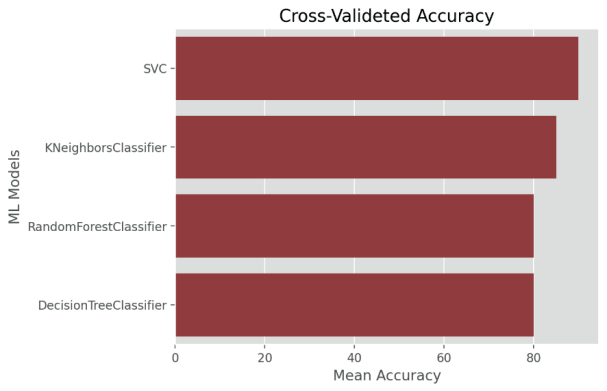


Figure 11. Accuracy results of the models

The research revealed how natural language processing and machine learning techniques can effectively classify Turkish legal texts. This work also provided valuable insights into optimizing the hyperparameters of specific machine-learning models. This optimization process is a critical step that significantly improves the model's performance. Additionally, this study has significantly contributed to legal technology applications by revealing the applicability of machine learning models in legal text classification.

This study fills a significant gap compared to the existing literature by focusing on the classification of Turkish court decisions, which is an under-researched area according to international law texts. Previous studies have mainly focused on legal documents from other jurisdictions, such as European (Kaur and Bozic, 2019; Cui et. al., 2023) and American (Lİ et. al., 2019; Eliot, 2020) case law. The focus of this study on Turkish legal texts brings a new dimension to the field of legal informatics. It reveals the applicability and effectiveness of artificial intelligence in a different legal system.

The originality of this research lies in its application of artificial intelligence techniques to classify Turkish legal texts that are unique in terms of their linguistic and legal characteristics, especially divorce and zoning cases. This study highlights the potential of machine learning in improving the efficiency and accuracy of legal document classification, especially by achieving a high accuracy rate with the SVM model. In addition to contributing to legal informatics, this research also provides a practical tool for lawyers in Turkey.

Future research should focus on expanding the types of legal documents analyzed and increasing the dataset size to increase the robustness and generalizability of the models. Additionally, integrating pre-trained language models such as BERT and GPT-4 can further improve classification accuracy and provide deeper insights into the semantic structures of legal texts. These advances can significantly benefit legal technology practices and assist legal professionals in their research and decision-making processes.

As a result, this research has made significant contributions to developing effective machine learning models for Turkish legal text classification. The findings obtained may guide future studies on model selection and hyperparameter optimization. In addition, this study will contribute to the development of legal technology applications by providing a solid basis for legal text analysis and classification research. Finally, for future research to proceed successfully and expand the applications of artificial intelligence in the field of law, case texts in the Republic of Turkey need to be made more widely available online. This will expand the scope of legal text classification studies and allow more effective use of artificial intelligence technologies in legal research and applications.

Orcid

Tülay Turan https://orcid.org/0000-0002-0888-0343
Ecir Uğur Küçüksille https://orcid.org/0000-0002-3293-9878

References

Alarie, B., Niblett, A., Yoon, A. H. (2018). How artificial intelligence will affect the practice of law. *University of Toronto Law Journal*, 68(1): 106-124.

Aletras, N., Tsarapatsanis, D., Preoţiuc-Pietro, D., Lampos, V. (2016). Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective. *PeerJ computer science*, 2, e93.

Alhakeem, Z. M., Jebur, Y. M., Henedy, S. N., Imran, H., Bernardo, L. F., Hussein, H. M. (2022). Prediction of ecofriendly concrete compressive strength using gradient boosting regression tree combined with GridSearchCV hyperparameter-optimization techniques. *Materials*, 15(21): 7432.

Ali, J., Khan, R., Ahmad, N., Maqsood, I. (2012). Random forests and decision trees. I*nternational Journal of Computer Science Issues (IJCSI)*, 9(5): 272.

Ashley, K. D., Brüninghaus, S. (2009). Automatically classifying case texts and predicting outcomes. *Artificial Intelligence and Law*, 17:125-165.

Awad, M., Khanna, R., Awad, M., Khanna, R. (2015). Support vector machines for classification. *Efficient learning machines: Theories, concepts, and applications for engineers and system designers*, 39-66.

Aydemir, E. (2023). Estimation of Turkish Constitutional Court Decisions in Terms of Admissibility with NLP. In 2023 IV International Conference on Neural Networks and Neurotechnologies (NeuroNT), IEEE, pp. 17-20.

Bafna, P., Pramod, D., Vaidya, A. (2016). Document clustering: TF-IDF approach. In 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), IEEE, pp. 61-66.

Benedetto, I., Sportelli, G., Bertoldo, S., Tarasconi, F., Cagliero, L., Giacalone, G. (2023). On the use of Pretrained Language Models for Legal Italian Document Classification. *Procedia Computer Science*, 225: 2244-2253.

Boella, G., Di Caro, L., Humphreys, L. (2011). Using classification to

support legal knowledge engineers in the Eunomos legal document management system. In Fifth international workshop on Juris-informatics (JURISIN).

Breiman, L. (2001). Random forests. *Machine learning*, 45: 5-32.

Chalkidis, I., Androutsopoulos, I., Aletras, N. (2019). Neural legal judgment prediction in English. arXiv preprint arXiv:1906.02059.

Charbuty, B., Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends*, 2(01): 20-28.

Chen, H., Wu, L., Chen, J., Lu, W., Ding, J. (2022). A comparative study of automated legal text classification using random forests and deep learning. *Information Processing & Management*, 59(2), 102798.

Cui, J., Shen, X., Wen, S. (2023). A survey on legal judgment prediction: Datasets, metrics, models and challenges. *IEEE Access*.

Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, 88(11):2783-2792.

Eliot, L. (2020). Legal judgment prediction (ljp) amid the advent of autonomous ai legal reasoning. *arXiv preprint arXiv:2009.14620*.

Fatourechi, M., Ward, R. K., Mason, S. G., Huggins, J., Schlögl, A., Birch, G. E. (2008). Comparison of evaluation metrics in classification applications with imbalanced datasets. In 2008 seventh international conference on machine learning and applications, IEEE, pp. 777-782.

Feurer, M., Hutter, F. (2019). Hyperparameter optimization. *Automated machine learning: Methods, systems, challenges*, 3-33.

Görentaş, M. B., Uçkan, T., Arli, N. B. (2023). Uyuşmazlık Mahkemesi Kararlarının Makine Öğrenmesi Yöntemleri ile Sınıflandırılması. *Yüzüncü Yıl Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 28(3): 947-961.

Gunn, S. R. (1998). Support vector machines for classification and regression. *ISIS technical report*, 14(1): 5-16.

Gupta, B., Rawat, A., Jain, A., Arora, A., Dhami, N. (2017). Analysis of various decision tree algorithms for classification in data mining. *International Journal of Computer Applications*, 163(8): 15-19.

Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and their applications*, 13(4): 18-28.

Hirschberg, J., Manning, C. D. (2015). Advances in natural language processing. *Science*, 349(6245): 261-266.

Hossin, M., Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International journal of data mining & knowledge management process*, 5(2): 1.

Kalia, A., Kumar, N., Namdev, N. (2022). Classifying case facts and predicting legal decisions of the indian central information commission: a natural language processing approach. In Advances in Deep Learning, Artificial Intelligence and Robotics: Proceedings of the 2nd International Conference on Deep Learning, Artificial Intelligence and Robotics, (ICDLAIR), pp. 35-45.

Kartini, D., Nugrahadi, D. T., Farmadi, A. (2021). Hyperparameter tuning using GridsearchCV on the comparison of the activation function of the ELM method to the classification of pneumonia in toddlers. In 2021 4th International Conference of Computer and Informatics Engineering (IC2IE), IEEE., pp. 390-395.

Kaur, A., Bozic, B. (2019). Convolutional Neural Network-based Automatic Prediction of Judgments of the European Court of Human Rights. In *AICS*, pp. 458-469.

Laaksonen, J., Oja, E. (1996). Classification with learning k-nearest neighbors. In Proceedings of international conference on neural networks (ICNN'96) IEEE, pp. 1480-1483.

Lei, M., Ge, J., Li, Z., Li, C., Zhou, Y., Zhou, X., Luo, B. (2017). Automatically classify chinese judgment documents utilizing machine learning algorithms. In Database Systems for Advanced Applications: DASFAA 2017 International Workshops: BDMS, BDQM, SeCoP, and DMMOOC, Suzhou, China, March 27-30, 2017, Proceedings 22, pp. 3-17.

Li, Q., Peng, H., Li, J., Xia, C., Yang, R., Sun, L., He, L. (2020). A survey on text classification: From shallow to deep learning. arXiv preprint arXiv:2008.00364.

Li, S., Zhang, H., Ye, L., Guo, X., Fang, B. (2019). Mann: A multichannel attentive neural network for legal judgment prediction. *IEEE Access*, *7*, 151144-151155.

Liu, Y., Zhou, Y., Wen, S., Tang, C. (2014). A strategy on selecting performance metrics for classifier evaluation. *International Journal of Mobile Computing and Multimedia Communications* (IJMCMC), 6(4): 20-35.

Mucherino, A., Papajorgji, P. J., Pardalos, P. M., Mucherino, A., Papajorgji, P. J., Pardalos, P. M. (2009). K-nearest neighbor classification. *Data mining in agriculture*, 83-106.

Nineesha, P., Deepalakshmi, P. (2022). Automated Techniques on Indian Legal documents: A Review. In 2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT), IEEE, pp. 172-178.

Patel, H. H., Prajapati, P. (2018). Study and analysis of decision tree based classification algorithms. I*nternational Journal of Computer Sciences and Engineering*, 6(10): 74-78.

Sert, M. F., Yıldırım, E., Haşlak, İ. (2022). Using Artificial Intelligence to Predict Decisions of The Turkish Constitutional Court. *Social Science Computer Review*, 40(6):1416-1435.

Turan, T. (2023). "Açıklanabilir Yapay Zeka İle Hukuki Metin Analizi", (Phd Thesis). https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Tagarelli, A., Simeri, A. (2022). Unsupervised law article mining based on deep pre-trained language representation models with application to the Italian civil code. *Artificial Intelligence and Law*, 30(3): 417-473.

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.

Vujović, Ž. (2021). Classification model evaluation metrics. *International Journal of Advanced Computer Science and Applications*, 12(6): 599-606.

Wang, Z., Wang, B., Duan, X., Wu, D., Wang, S., Hu, G., Liu, T. (2019). IFlyLegal: a Chinese legal system for consultation, law searching, and document analysis. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP): System Demonstrations, pp. 97-102.

Yang, L., Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415: 295-316.

Yun-Tao, Z., Ling, G., Yong-cheng, W. (2005). An improved TF-IDF approach for text classification. *Journal of Zhejiang University-Science*, 6(1): 49-55.

Zhang, D., Zhang, H., Wang, L., Cui, J., Zheng, W. (2022). Recognition of Chinese legal elements based on transfer learning and se-

mantic relevance. *Wireless Communications and Mobile Computing*

Zhong, H., Xiao, C., Tu, C., Zhang, T., Liu, Z., Sun, M. (2020). How does NLP benefit legal system: A summary of legal artificial intelligence, arXiv preprint arXiv:2004.12158