Original scientific article / Izvorni znanstveni članak
Received / Zaprimljen: 21.10. 2023.
https://doi.org/10.21860/j.14.2.8

Ye-Chan Park*, Jaesung Lee**

Test Reliability in Case Law by Analysis of Korean Precedent Judgment Criteria on the Degree of Assault and Intimidation

SUMMARY

The application of criteria from previous judgments in subsequent cases is unclear. One such factor, the degree of assault and intimidation, influences the severity of a crime but is challenging for the public to assess due to its unclear derivation. To address this, numerous studies aim to identify core elements that empirically prove abstract concepts. This study introduces a method for analyzing judgments, extracting elements for each issue based on shared patterns in reference precedents. Unlike previous approaches, it can analyze numerous rulings through extensive data processing. The method relies on judgment criteria and establishes classification criteria according to theory, leading to linear regression results. The model demonstrates that the importance of elements varies for each issue, providing insight into the rationale behind judgments. The public can utilize this model to enhance judgment predictability by analyzing precedents. Moreover, it can be employed to assess judgment reliability in both the Anglo-American and civil law systems that recognize case law.

Keywords: Humanities data analysis, linear regression, precedents, importance analysis, feature selection.

INTRODUCTION

There are two main legal systems: the civil and common law systems. The common law system uses the "doctrine of stare decisis," which affects judgments as case law,

* Department of Public Service, Chung-Ang University, Seoul, Republic of Korea. ORCID: https://orcid.org/0009-0009-8428-2239.

** Department of Artificial Intelligence, Chung-Ang University, Seoul, Republic of Korea. ORCID: https://orcid.org/0000-0002-3757-3510.

Correspondence Address: Ye-Chan Park, Department of Public Service, Chung-Ang University, 84, Heukseok-ro, Dongjak-gu, 06974, Seoul, Republic of Korea. E-mail : qkrd1285@cau.ac.kr.

but the continental legal system does not use this doctrine. Korea has a civil law system, but precedents are affected by reflecting the elements of the common law system. Therefore, a judgment determines whether an act constitutes a crime by decomposing and systematically presenting the requirements for establishing a crime into certain elements. The court leaves the judgment of abstract measures, a concept not under criminal law, to the discretion of judges under the name of "objectively" social norms and "the consensus of judges," making it difficult for the public to predict the court judgment. This problem exists in various areas, such as whether causal relationships are established, mental and physical weakness, pornography, publicity, and independence; thus, studies have been conducted to judge and analyze abstract standards (Lee & Lee, 2008; Lyu, 2022).

If the abstract standards of the court, which are not created by the legislature, are identified and verified, they can be used to grasp the effect of the ruling and predict future rulings. Among them, the concept of *degree* in "degree of assault and intimidation," which is an indicator of various crimes, is also one of the abstract concepts as the Supreme Court of Korea (1993) stated, "It should objectively suppress or make the other person's resistance impossible in social norms." Therefore, the Supreme Court of Korea (2012) presents the judgment criteria, such as "all circumstances, such as the circumstances of the exercise of tangible power, the relationship with the victim, and the circumstances at the time and after that." Hence, in this study, the criteria are presented, but the judgment process is unknown, such as the degree of violence and intimidation and the degree to which rebellion is impossible. The case law is reorganized and analyzed through a dataset to determine whether judge-specific criteria are used.

It is possible to grasp the abstract criteria based on the Korean judgment because the ruling is made with a "reference precedent" in cases, such as the previous ruling, which acts as a critical factor. Reference precedents refer to the precedents referenced in the judgment of the trial, and the judgment presents the reference precedents. However, it is unclear whether the ruling is made with it. If it is not impossible to analyze a judgment with empirical data called reference precedents, it is meaningful because it can enhance the role of the reference precedents in actual judgments. Therefore, if an empirical analysis of the judgment is successful, the strength of each element in analyzing the judgment can be extracted, and the public can help predict the future as a kind of guideline. This study classifies the degree of assault and intimidation as a standard for judgment as a tool for case analysis. In addition, the degree of assault and intimidation in theories, such as those by Park (2011) and Bae (2015), can be classified, and the theory by Kim (2020) can be classified based on rebellion and can be arranged in order.

Baseline study

Many studies have been conducted to predict judicial decisions using rapidly developing machine learning technology. Katz, Bommarito and Blackman (2017) worked to create a model to predict the court's decision system. The machine learning technique called the random forest classifier was used to select court data. These efforts successfully constructed a model that could predict judgments in a single period and future and past judgments.

In line with these studies, research has been conducted to increase model accuracy, and Ahmad, Asghar, Alotaibi and Al-Otaibi (2022) applied hybrid models to predict court judgments accurately. A higher-rank feature was extracted through a hybrid decision support system based on deep learning, one of the convolutional neural networks (CNNs), and a function with higher accuracy was selected as a priority. Moreover, the case decision was predicted through a CNN with bidirectional long short-term memory, and this hybrid model achieved 91.52% accuracy. In addition, in Korea, Lee et al. (2021) suggested the possibility of establishing a judgment prediction system through machine learning, creating and establishing a corresponding model. In Japan, which already has a legal system similar to that of Korea, research has been conducted to learn and analyze judgments through various models, including Rzepka et al. (2008) and Tran et al. (2019).

Related research in Korea

Research has been conducted to present judgment criteria based on empirical research through the specific descriptions and technicalizations of the abstract and subjective concepts of precedents. Lyu (2022) attempted to define the concept by analyzing 17 precedents for the concepts of public and independent private schools. The definition of publicity and autonomy is not directly mentioned in the ruling, but private schools, emphasizing publicity and autonomy, are public educational institutions. The definition was derived based on what is mentioned in the precedent.

In addition, Lee and Lee (2008) categorized judgment sentences concerning pornography and extracted 13 judgment criteria. Although the concept and judgment criteria for pornography are presented in the precedent, they are very subjective and abstract as the subject of the final judgment of the degree of pornography. Thus, the sentences for judging pornography for the Supreme Court precedent were categorized, and the degree was determined according to the characteristics by examining the judgment process. Based on 22 precedents, the degree of pornography was determined according to the degree presented in the precedent after categorizing sentences for judging pornography. These attempts aimed to empirically analyze all

relevant judgments (i.e., precedents) distinguished from the microscopic approach to delve into the nature of individual cases.

As indicated by Kim (2008), it is common for most case analyses to deal with validity and judgment criteria for individual judgments. The reason is that most case analyses are easy to approach with a careful interpretation based on the researcher's rational and thoughtful arguments. The value-inherent concepts of assault and threat are very subjective, similar to the concept of medical negligence. Previous studies assessing this subjective concept of include Park et al. (2022). However, an analysis approach throughout the precedents was required to attempt empirical analyses to change the abstract and subjective concept of precedents and was successfully determined (Lyu, 2022). We aim to analyze the degree of assault and intimidation, such as through a study that categorized judgment sentences according to characteristics (referred to as factors) and a methodical research method (Lee & Lee, 2008) that analyzes precedents and evaluates judgment criteria. In particular, to deal with more precedents more efficiently than previous studies, a dataset is formed by preprocessing precedents to attempt an analysis through linear regression.

ANALYSIS METHOD

Database limit

All baseline studies use data from the Supreme Court Database (SCDB) (Bailey & Maltzman, 2008; Benjamin, Desmarais, 2012; Epstein & Martin et al., 2007; Lee, Broedersz, & Bialek, 2015; Martin & Quinn, 2002; Segal & Spaeth, 2002) because it is high-quality, expertly coded data with 240 variables, such as chronological variables, case background variables, justice-specific variables, and outcome variables. Professor Spaeth and many others have helped to make these datasets. The SCDB features 200 years of data on the Court's behavior but has some critical limits. Shapiro (2008) pointed out that SCDB omitted substantial and vital information. The SCDB does not indicate how the Supreme Court works and appears in precedents because this dataset is designed to be used from a public policy perspective, so it rarely deals with legal issues. For example, a variable related to an issue is often unrelated to legal issues because it describes political and social issues. In addition, datasets are coded for only one problem, and legal problems are often intertwined with various issues. Finally, legal regulations and problems indicate different things, but they are not adequately analyzed. Various legal provisions can be intertwined in a problem, and one legal provision can be used in several problems. Shapiro (2008) noted these problems lead to errors in the classification of Schenck, the First Amendment case involving abortion protesters, which is a problem for abortion protesters.

Classifying assault and intimidation and setting independent and dependent variables

The severity of assault, as a determinant in legal judgments, not only influences the categorization of various crimes within the Korean legal system but also significantly shapes the subsequent legal consequences, including sentencing and penalties. In particular, the severity of assault affects the classification and determination of rape, robbery, blackmail, intimidation, and assault (Bae, 2015; Kim, 2016; Kim & Seo, 2016; Lee, Jang & Kang, 2016; O, 2014; Park, 2011). Depending on the classification, the concept can be classified into a wide range of intimidation types in the crimes of obstructing the execution of official duties, coercion of duty, special escape, rebellion, and disturbance. In addition, the concept can be classified in the crime of coercion or blackmail as a threat to the extent of inducing fear or, in the crime of robbery or rape, as an intimidation level that makes it impossible or incredibly difficult to rebel against the other party according to the prevailing social norms.

Because the keyword "rebellion" is included, the degree of assault and intimidation in a crime depends on how much rebellion is possible as a classification standard to determine the degree of intimidation that can be classified (Kim, 2020). The degree of rebellion is used in precedents sequentially according to the crime, making it easy to collect data; therefore, it is appropriate to analyze the precedents to determine the importance of each variable. Therefore, the degree of assault and intimidation is determined according to the classification criteria. Based on the precedents, the court deemed the degree of assault, which is selected as the primary dependent variable, to be the level of assault and intimidation that is a) impossible or significantly difficult to defy, b) quite difficult to defy, c) difficult to defy, or d) does not reach the level of resistance. Judgments are based on precedents stating the degree of assault and intimidation, including precedents mentioning assault and intimidation presumed to be of the same degree. Thus, the dependent variable includes the state in which it is significantly challenging to resist. A value is assigned to the judgment factor corresponding to each dependent variable according to the classified criteria because the rebellion level, which requires the highest level of violence and intimidation, varies from *impossible* to *not reaching the level of resistance*.

In addition, as suggested by various theories, no objective standard defines the extent to which the degree of difference exists, which is at the researcher's discretion. Arbitrary decisions are not appropriate because they can undermine validity. Therefore, the analysis proceeds with the assumption that it is equal according to the degree. In the decision tree machine learning process, the degree of difficulty of rebellion, which suggests a high error in classification, was combined with the degree of *impossible*. The number of variables was reduced because there was no benefit to distinguishing

them, as they were not differentiated or used together in the actual judgment. No errors occurred during the subsequent research process. No significant problems arose from the regression analysis, so the simple uniform dependent variable was adopted.

In conclusion, although the dependent variable is presented equally from 1 to 0, most precedents do not make a difference in cases, such as the degree to which rebellion is impossible or significantly difficult. Thus, the variable is reflected and classified as 1. As a result, assault and intimidation to the extent that rebellion is impossible or significantly difficult to resist are assigned a value of 1. Violence and intimidation, which are very difficult to rebel against, are set at 0.5. Insurmountable violence and intimidation are set at 0.25, and unsophisticated assault and intimidation are set at 0.

How to score variables

Variables are those that the court considered important based on the level of assault and intimidation in the parts mentioned in the precedents. The background refers to the reason for exercising tangible power and whether a motive also exists for the victim. Coping refers to whether the victim coped with the crime after the exercise of tangible force (fleeing from the accused, reporting the crime, or committing suicide), and damage refers to how it affected the victim (physical and property damage to the victim). Avoidance refers to whether the victim can escape by doing his or her best (i.e. when the accused oppresses the victim with violence or cannot escape due to the specificity of the space). Relationship refers to whether the victim and the accused are in a special relationship (e.g., friends, lovers, family, etc.), and the characteristics of the accused are compared to the victim (e.g., they have more stature, are older, or are unwilling to give orders), which refers to a case that corresponds to a relationship that can be used a weapon or leverage. Demand and harm refer to whether the demand obtained through intimidation and assault is illegal or whether harm occurs if the demand is not fulfilled.

Data collection

Judgments comprise matters of judgment, the judgment summary, articles of reference, precedents referenced, the preamble, the order, and reasons. A double judgment presents the critical issues in making a judgment in the trial through the judgment and summary of the judgment. In addition, judgment matters have a specific form, such as "for or not," and have high consistency for judgment analysis. Therefore, extracting judgments and collecting data according to the research topic is easy if the researcher first checks the judgments that specify the research topic.

Data for analysis were collected from the environment under the study, and data collection focused on case law, a situation in which factors were controlled to maintain a constant level. When collecting data, it is difficult to analyze the precedent as it is; that is, it is challenging to preprocess natural language, so the data were collected using a standardized format, such as the judgment of the precedent (degree of assault and intimidation). If the "degree of assault and intimidation" is included in the judgment, and the judgment provides a detailed explanation, the content after "whether or not about" was analyzed to determine the degree of assault and intimidation.

Whether or not each collected variable is qualitative can be determined, and each variable is divided into consequences (robbery, rape, blackmail, coercion, and innocence) according to the degree of assault and intimidation. Cases that deal with the degree of assault and intimidation are judged only on the severity of the crime rather than on whether it is a crime. Cases of coercion, blackmail, rape, and robbery were collected, primarily dealing with the degree of violence and intimidation among the issues. Data were collected primarily by extracting texts from all judgment items that met the classification criteria presented above or had other expressions in the precedents or similar content in the preamble. Some of the judgments briefly mentioned the process of the judgment, and others explained the process of the judgment in detail. However, parts that do not explain the process that led to the judgment do not affect the judgment (Recommended Guidelines on the Writing Style of Legal Judgments, 1998). Therefore, even if the judgments do not mention it in detail and the cases are few, this study collected it as data. Based on 100 precedents well cited as double reference precedents, a precedent mentions the degree of violence and intimidation in each case. Most of the rulings collected in this way fall under the Supreme Court ruling. Judgment of an appropriate proportion of lower courts' cases is appropriate for analyzing the judgment. However, in this study, it is acceptable to use these. Because higher court cases and lower court cases are not likely to be judged differently, lower court cases are more likely to cite Supreme Court Judgment and Supreme Court Judgment hardly changes. In particular, when the sentence was finalized, the judgment of the second trial was also collected and cited. It is possible to judge the victim's statement, police report, and case investigation document presented in precedents. Appendix 1 provides the details of each case considered in this study.

Pretreatment

Pretreatment is a variable attribute regarding the degree to which related words are cultivated in the data. As mentioned in setting the variables, preprocessing concerns whether a related situation exists in the explanation of the situation in the text. The

relationship between the victim and suspect corresponds to the debt borrowed and received, the creditor's relationship, the description of the victim's damage in the police investigation, and whether the damage includes words regarding a preposition, debtor, and creditor, for example. If the word recognized in the precedent is included as a variable, it is set to 1 (otherwise, it is 0) to form a dataset. This determination is based only on whether the issue set exists as a variable because it contradicts the study purpose if there is controversy regarding this.

Many subjective opinions of the researcher are added when a stochastic value is proposed to be allocated to the input variable. The same result value could not be derived when the same research method was applied to another case. In addition, it is impossible to explain why a variable is set to a stochastic value. Further, there was no significant difference in the degree of explanation or accuracy analyzed through software (WEKA) in the case of judging with or without. Therefore, this was chosen because the most objective case was judged with or without. The judgment is first made based on the objective facts presented by the judgment, but if the objective facts do not exist and depend on the victim's statement, the judgment is dismissed and is not cited as data.

Each rebellion is marked as 0.25 if difficult, 0.5 if quite difficult, and 1 if remarkably difficult or impossible, and the reasons are described above. This value is determined based on the judgment. If it is an old judgment or is not indicated in the judgment, the dataset is created based on the order, reason, or summary of the judgment afterward.

Multicollinearity test

The elements considered in the analysis were designated as variables, and the delineation between these variables is not explicitly defined, raising the possibility of collinearity. To assess the appropriateness of variable settings, an examination of collinearity is conducted. Initially, the interrelationships among all variables are investigated, and the presence of collinearity for each variable is assessed based on their Pearson correlation coefficients. The reason for choosing the Pearson correlation coefficient was to quantitatively assess the linear relationships between variables in the given dataset. This method is effective in measuring strong linear relationships in the data and provides an accurate understanding of the correlation levels between each pair of variables through the correlation matrix. This coefficient is calculated using the following formula:

$$\left[r = \frac{\sum \left(X_i - \underline{X}\right)\left(Y_i - \underline{Y}\right)}{\sqrt{\sum \left(X_i - \underline{X}\right)^2 \sum \left(Y_i - \underline{Y}\right)^2}} \right]$$

where *Xi Yi* is a separate observation of each of the two variables, $\underline{X}\ \underline{Y}$ represents the average of each variable. As depicted in Table 1, the correlation matrix summarizes the pairwise Pearson correlation coefficients for each variable. The analysis of the correlation matrix indicates minimal correlation between variables.

Table 1. The pairwise Pearson correlation coefficients for each variable.

	Coping	**Avoidance**	**Rebellion**	**Circum-stances**	**Demand/ harm**	**Relation-ships**	**Charact-eristics**	**Damage**
Coping	1.0000	0.4741	0.2769	0.3693	0.3282	-0.0269	0.2680	0.1788
Avoidance	0.4741	1.0000	0.3452	0.5136	0.5136	-0.1307	0.2380	0.3142
Rebellion	0.2769	0.3452	1.0000	0.2452	0.1184	0.1550	0.0580	0.3453
Circum-stances	0.3693	0.5136	0.2452	1.0000	0.5833	0.0364	0.2969	0.3632
Demand/ Harm	0.3282	0.5136	0.1184	0.5833	1.0000	-0.0546	0.1619	0.1721
Relation-ships	-0.0269	-0.1307	0.1550	0.0364	-0.0546	1.0000	-0.1414	0.1418
Charact-eristics	0.2680	0.2380	0.0580	0.2969	0.1619	-0.1414	1.0000	0.3633
Damage	0.1788	0.3142	0.3453	0.3632	0.1721	0.1418	0.3633	1.0000

Therefore, although some association exists between each variable, no linear dependency exists (a negligible level).

Model analysis result

The machine learning software was developed by Weka at Waikato University and is based on Java. Classification and regression were primarily used among the provided algorithms, along with data preparation, classification, regression, clustering, association rule mining, and visualization. The judgment process of the trial could be analyzed by recognizing and classifying patterns through the decision tree (j48, random tree). Judging the dependent variable to determine whether each variable could contribute was possible. Classification j48 submitted results of 89% accuracy, and 93% accuracy was achieved for the random tree.

However, the regression analysis, which can perceive the importance of elements, is more suitable for extracting information in the judgment process. Therefore, the results were analyzed based on the linear regression results. First, the results of the data construction were analyzed through a linear regression model. In addition, a hypothesis test was conducted regarding whether the model for this is appropriate.

LINEAR REGRESSION AND MODEL TEST

A linear regression circle is defined as follows:

$$\hat{y} = 0.0601 * \text{Circumstances} + 0.2885 * \text{coping} + 0.0689 * \text{damage} + 0.2110 * \text{avoidance} + 0.1639 * \text{rebellion} + 0.1286 * \text{characteristics} + 0.2816 * \text{demand, harm-} 0.1767.$$

The model provides a 93.08% association performance, and the root mean square error is 0.17. Based on the derived linear regression model, the predicted values are applied to the precedent to analyze the values assigned to the judgments. For variable setting, the value derived for the elements can be predicted to be severe depending on whether the element is present, so it would be appropriate to assume that the value is higher than the reference point. A *t*-test is performed to test the hypothesis for individual variables to analyze whether the values through the regression parameters of each variable have significant results. In other words, the variables and individual linear regression coefficients are inferred as an index to analyze how much the predicted value with each parameter affects the result regarding whether each *t*-value is significant at a certain confidence level. It is selected by adopting it as an indicator of how much linear relevance exists.

In addition, the analysis results for the primary analysis are presented through standard errors and standard deviations derived from linear regression values. A greater *t*-value indicates that the result is more meaningful at a higher confidence level; thus, the importance of individual variables can also be analyzed through the *t*-value. In addition, the difference between the sample and linear regression value can also be confirmed by analyzing the standard error and standard deviation. Each variable is significantly related to the result, even at a confidence level of 99%. The relationship between the variables *path* and *damage* is minimal (reliable at the 95% level) compared to other variables but not at a low level. The influence is strong in the order of demand/harm, > coping > avoidance > rebellion > characteristics (*t*-value, parameter). The parameter analysis (93.1% of the explanation ratio of the predictors) and the *t*-value (meaningful at 99%) are sufficiently large (99% indicates significance at the 99% level).

Table 2. *T*-test result and linear regression summary.

	Standard deviation	**Standard error**	**Parameter**	**t-value**	**t-test**
Circumstances	0.05	0.5	0.0601	1.202	<
Coping	0.047937	0.479372	0.2885	6.018285	99%
Damage	0.048524	0.485237	0.0689	1.419926	<
Avoidance	0.049237	0.492366	0.211	4.28543	99%
Rebellion	0.049237	0.492366	0.1639	3.328825	99%

Relationship	0.045126	0.451261		0	<
Characteristics	0.045605	0.456048	0.1286	2.819878	99%
Demand/harm	0.042923	0.429235	0.2816	6.560513	99%
Probability error			-0.1767		

Tree algorithm

The software was created in orange3 by Ferenc Borondics, a principal beamline scientist at SMIS SOLEIL synchrotron, using the Python language. Katz, Bommarito and Blackman (2017) also applied random trees to datasets, so the random forest was applied in this study. The tree algorithm provided by orange3 was also applied. The result of applying the algorithm is provided in Table 3.

Table 3. Experimental case prediction results using the tree and random forest (target indicates ground truth).

Case	Target	Tree	Random forest
93do901	0.25	0.23709803	0.222222
2018do19493	0.5	0.58477988	0.5
2001do7095	0	0.19742697	0
87do690	0	0.19742697	0
2004do2611	0	0.44716205	0.35
2000do1914	0.75	0.68115342	0.75
93do428	1	0.82218532	1
92do2884	0	0.16867551	0.125
90do2224	0	0.23474599	0
91do546	0	0.30351522	0.25
92do259	0	0.09686527	0.125
99do2608	0	0.20136626	0.222222
2010do9633	0	0.16867551	0.125
2005do3071	0.75	0.75337239	0.666667
2006do5979	0.75	0.81877125	0.833333
2012do4031	0.75	0.88521479	0.833333

56do50	1	0.81877125	0.833333
81do409	1	0.88521479	0.833333
79no323	0	0.23474599	0
77no293	0.5	0.51061761	0.666667
75no1028	1	0.95702503	1
81no527	0	-0.1444795	0
76do1932	0	0.33924698	0.25
2000do1253	0.75	0.88521479	0.833333
2006do4449	0	-0.1141407	0
88gohab114	0	0.12861403	0
2001do359	0	0.47682915	0.3125
2016do16948	0.75	0.92129327	0.916667
2001do230	0	-0.1498725	0
88do1628	0.75	0.75337239	0.666667
2005do4799	0	-0.0419574	0
86do931	1	0.95702503	1
85gohab907	0	0.30770339	0
73no239	1	0.92129327	0.916667
98do70	0	0.19780935	0
91do1184	0.25	0.20440727	0.25
2004do4437	1	0.95702503	1
2001do4462	0	-0.0726692	0
2000do5395	0	0.23709803	0.222222
99do519	0.75	0.89095456	0.875
2012do14788	0.75	0.64542166	0.75
2012jeondo252	0.75	0.64542166	0.75
95do91	1	0.95702503	1
84do2732	1	0.95702503	1
90do2102	0.25	0.23170498	0.375

2007do7064	0.25	0.23709803	0.222222
2011do8805	0	-0.0726692	0
2016no69	0	-0.0726692	0
2000do4415	0.5	0.37232816	0.3125
2015do16696	0.25	0.30890827	0.35
2018do7709	0.75	0.81877125	0.833333
98no988	0.5	0.61508294	0.5
99do4305	0.5	0.57900446	0.5
2011no1393	0.25	0.33085663	0.3125
2018do13792	0	0.06217049	0
2018do2236	0	0.06217049	0
84do2083	0.5	0.51296964	0.5
84do573	0.5	0.30890827	0.35
91do80	0.5	0.30890827	0.35
93do915	0.5	0.30890827	0.35
96do1959	0.5	0.37535181	0.222222
2003do709	0.5	0.44075068	0.5
2004do1565	0.5	0.44075068	0.5
2005do4738	0.5	0.26743674	0.375
81do3202	0	-0.1802112	0
2012do6157	0	-0.0096398	0
85do1687	0.5	0.54672244	0.5
92no176	0.5	0.3069293	0.125
82do2714	0	-0.0399785	0
87do1656	0.5	0.47682915	0.3125
89do2036	0.5	0.51256092	0.5
91do1824	0.5	0.44454713	0.5
96do2151	0.5	0.2731765	0.5
99do4305	0.5	0.2731765	0.5

79do1660	0	-0.0726692	0
79do2565	0	-0.0062257	0
90do114	0	-0.0062257	0
93do2339	0	-0.0399785	0
2013do6809	0	-0.0399785	0
96do1728	0.5	0.30351522	0.25
2007do1375	0.5	0.30351522	0.25
2003do763	0.25	0.23709803	0.222222
2010do13774	0.25	0.23709803	0.222222
2004do227	0.25	0.20440727	0.25
73do2578	0.25	0.23709803	0.222222
2003do763	0.25	0.20440727	0.25
2003do4151	0.25	0.23709803	0.222222
96do3411	1	0.95702503	1
99do242	1	0.95702503	1
2007do7601	1	0.95702503	1
95no1686	1	0.89095456	0.875
69do927	1	0.95702503	1
71do441	1	0.95702503	1
82do2838	1	0.95702503	1
85do2115	1	0.95702503	1
91do2267	1	0.95702503	1
86do2203	1	0.95702503	1
98no308	0	-0.0399785	0
2013do11899	1	0.92129327	0.916667
85no2225	0	-0.0399785	0

Whether the result falls within the range of greater or less than 0.15 was successfully predicted to determine whether it meets the target. As shown in the Figure 1, Pythagoras trees are presented to explain the decision tree and random forest shapes, which have various tree shapes (Beck et al., 2014). In the Pythagoras tree, the branch

extends in one direction, and if specific variables are selected, the result can be easily derived. Figure 2 presents the decision tree created based on given data. Our experiments demonstrated that 83% and 82% of accuracy values are observed by the random forest method and the tree, respectively.

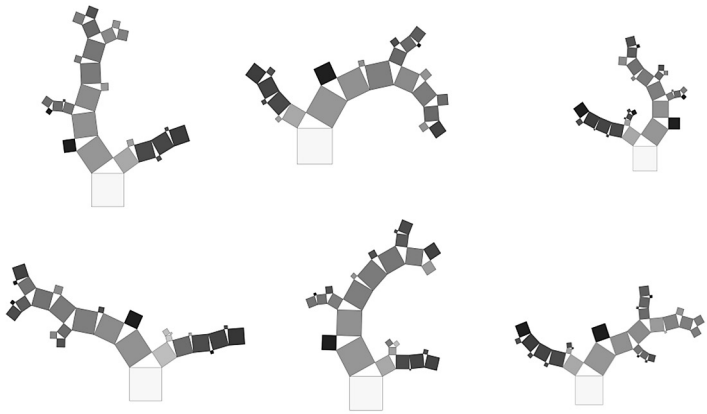

Figure 1. Examples of random Pythagoras trees.

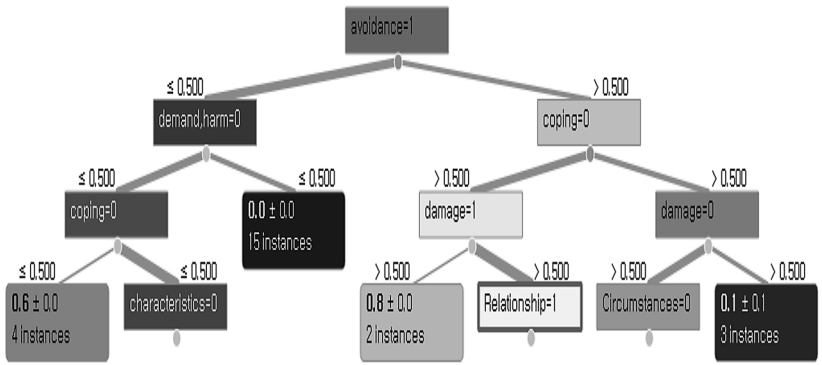

Figure 2. An illustration of decision tree for experimental dataset.

Neural network

The neural network is a multilayer perceptron algorithm with backpropagation. The neural network removes instances with unknown target values and continues with the categorical variables. If any of these columns are empty, they are removed, and the mean value is attributed to the missing value. The data are normalized by centering them on the mean and adjusting to the standard deviation of 1. This process is exemplified in Figure 3.

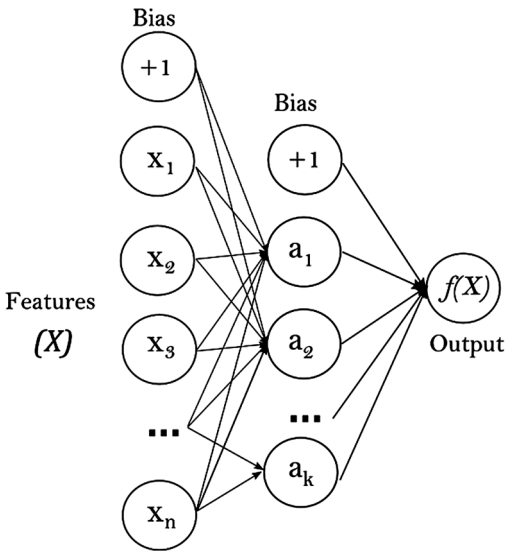

Figure 3. Operating mechanisms of the hidden layer.

The neural network employed in this paper is Sklearn's multilayer perceptron algorithm. Parameter 0.1 for regularization (L2 regularization) is helped by penalizing weights with large magnitudes to avoid overfitting. Figure 4 is a decision boundary plot that appears with lesser curvatures showing how much this model penalizes. The activation function for the hidden layer is the logistic sigmoid function, and the solver is used for weight optimization. The stochastic gradient-based optimizer is called Adam.

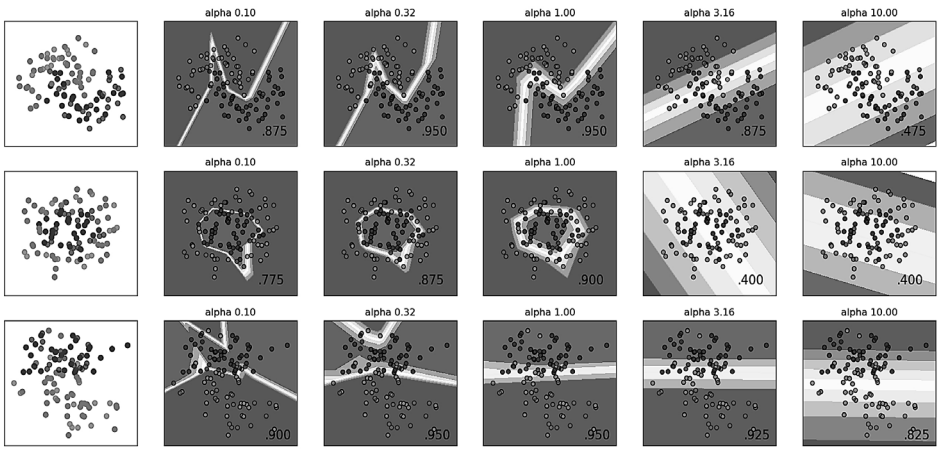

Figure 4. Example of penalizing degree of 0.1 L2 regularization.

The mean square error is 0.026. The mean absolute error is 0.119, and the coefficient of determination is 0.825. Table 4 compares the results of the neural network with the random forest results.

Table 4. Comparing the result summary.

	Mean square error	**Mean absolute error**	**Coefficient of determination**
Neural network	0.025	0.119	0.825
Random forest	0.028	0.114	0.807

Test analysis result

Even if the judgment is made, because the judgment process does not provide a black box or process, there is difficulty in empirically analyzing the judgment due to the lack of "reliability" regarding whether it repeatedly results in the same or similar results. The judgment presents the critical issues in making the judgment in the trial through the matters of judgment and the summary of the judgment.

The subject studied by the researcher dealt with one of the critical issues affecting the judgment of the trial, the degree of assault and intimidation, and the judgment of assault, intimidation, rape, and robbery was analyzed to demonstrate the judgment process. Through the research, the ruling process was derived using data analysis, which has high reliability and succeeded in empirically analyzing the judgment. Therefore, this analysis can be used as a guideline because it can be the basis of a system for predicting and analyzing judgments through artificial intelligence or the judgment process of other core issues, especially those likely to be subjectively judged.

This research method contributes digitized data to the existing research methods. Compared to the previous studies by Lee and Lee (2008) and Lyu (2022) used 17 cases and 22 cases, respectively. Learning more than 100 rulings through programming was implemented; thus, the possibility of exception cases was blocked. This study presented similar judgments based on work by Kim (2008), who extracted the elements of medical negligence and indicated how critical the factors were by expressing the importance as regression parameters. The double dilemma of preparing subjective standards again to objectify the subjective elements shows that this study succeeded in objectifying and securing the reliability to produce the same results when the subjective standards were reused as research methods. We analyzed whether there were valid reasons or examples for values judged to be important and those that

were not and examined whether they were consistent with reality. As a result of the analysis, each value can present a valid case and reason.

Next, if the model becomes inappropriate or unsuitable for analysis, there is a possibility that the model should be rejected. A case in which the linear model result was unsuitable for the actual result was analyzed in detail, and the reasoning was examined. In summary, the result of inconsistency due to model nonconformity was only 5% of the total regression result, and the case of the most inconsistency was caused by not rendering a high result value because of the degree of assault and intimidation, which are some of the analyzed criminal components. Therefore, the degree of assault and intimidation corresponds to the subjective area of judgment; thus, it is easy to analyze precedents in terms of objectivity, providing a meaningful result.

The noted factors are demand and harm, which have the highest regression parameter values. The order of discussion afterward also depends on the value of the regression parameter. All crimes are established only when they are intended to be committed by intimidation and assault, and the harm of committing them occurs if the victims do not submit. It is also the first factor to judge. In other words, it is a constituent requirement to deal with the threat itself before considering the degree of the threat. An example of demand that is a natural demand or harm is a newspaper manager forcing someone to subscribe to the newspaper. Therefore, it is reasonable to analyze demand as the most crucial factor.

The second item is related to coping. If not dealt with, coping does not lead to a crime, but when judging, if it is judged that the coping level of the victim is insignificant, the degree of the crime is also judged lightly. An example is when the victim did not run away and stayed with the perpetrator. Next, variables with low regression parameter values were analyzed. In the above case, even if a criminal act affects the victim, the basis for the act is not recognized solely for such reasons as reducing the criminal's sentence. There are cases in which the victim caused the reason for the act, such as when the victim was assaulted and threatened due to not paying back the money. Therefore, although mentioned in precedents, this situation does not affect the judgment.

Relationships are dismissed for the same reason. It does not matter what relationship the victim and defendant have during the crime. In particular, there are cases where the relationship is between the creditor and debtor or the employee and customer. The next item is related to damage. Damage exists in all crimes because crime is established on the premise that damage occurs. Therefore, it is difficult to consider it a variable that determines the degree. For example, an individual was acquitted despite being assaulted and threatened with a hammer for fraudulent gambling.

Finally, errors corresponding to discrepancies were analyzed. In total, 25% of the errors correspond to the mismatch value where the model was unsuitable due to analysis failure, which can be considered errors due to the limitations of the analysis method. The types of errors can be classified into four categories: 1) the data have noise, 2) the degree of assault and intimidation is not smooth, 3) the study objectives are similar, and 4) the error corresponds to a failure to fit the model. Four cases correspond to the fourth error type: 91 do 1824, 96 do 1959, 2000 do 1914, 2012 do 14788, and 2012 do 252.

CONCLUSION

This study showed that presenting the possibility of the court's judgment process using a linear regression model, neural network, and random forest. Because there can be a hidden factor that affects to the result of the judgment, doubts about the reliability of the trial can arise, assuming that the judgment can be changed due to the tendency of the judge and the political situation due to outside factors. However, unlike these predictions, it is possible to make consistent judgments and extract the judgment process. Therefore, this research method can be presented as a guideline for areas that should focus on predicting or analyzing judgments in other cases by empirically proving the judgment process.

While this study has successfully delved into the judgment process regarding assault, intimidation, rape, and robbery, it is imperative to extend the discussion to the practical implications for legal practitioners and judges. The insights derived from this analysis hold the potential to enhance decision-making processes in real-world legal scenarios. Legal practitioners can benefit from a deeper understanding of the nuanced factors influencing judgments in cases involving assault and intimidation.

Furthermore, this research prompts considerations about the broader applicability of the model developed. Beyond its specific focus on assault-related cases, there is potential for this model to be applied to other core legal issues that are subject to subjective judgment. The empirical basis and reliability of this data analysis make it a strong candidate for serving as a foundation for predictive and analytical systems, including those powered by artificial intelligence.

In considering this role in legal decision-making, this study can serve as a viable alternative to stare decisis in common law systems. Because the judge's judgment process is not known in the common law system, it is possible to make different judgments depending on the judge, but the research method of this study can be applied to distinguish differences between different judgments and block the

possibility of making "misjudgment." This avenue of exploration is particularly relevant as this navigates evolving landscapes in legal technology and methodologies.

This study analyzed the linear regression of the degree of assault and intimidation, which is a component of crime that can seem subjective and cause doubts about the arbitrariness of the judgment leading to the results (the degree to which resistance is remarkably difficult, to suppress resistance, etc.). In particular, this method has a 93% association due to suitability, and 5% of the errors categorize the judgments, which can serve as sufficient grounds for use as a tool to apply and judge new judgments and cases. The other models (random forest and neural network) provide the classification with 83% accuracy and under 0.1% errors. In addition, the elements considered important were identified through data mining. Even people without legal knowledge can schematize and quantify the data to determine whether an act has a criminal element and how much crime was committed. The regression inference of individual elements succeeded with a confidence level of 99% or higher, demonstrated through success in the hypothesis inference in the linear model. There is also a fundamental difference among the factors presented in the judgment.

In practice, precedents make judgments based on standards that can be proved. However, in the case of the relationship variable, inference failed because it was excluded from the linear model due to a lower correlation than expected. Judgment is merely a reference point presented through a list, but which factors focus on the judgment can be determined. Based on the regressed value, the cases of "demand/harm, avoidance," and "response," in which the value of the parameter is offset by 0.2, affect the judgment through considerable values, and "damage (0.601)" and "background (0.0689)," which have the lowest parameter values, also explain this effect well. Since this study mainly analyzed Supreme Court judgments, we will add lower court cases to subsequent studies.

ACKNOWLEDGEMENT

This work was supported by the Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) (2021-0- 01341, Artificial Intelligence Graduate School Program (Chung-Ang University)).

REFERENCES

Ahmad, S., Asghar, M. Z., Alotaibi, F. M. & Al-Otaibi, Y. D. (2022). A hybrid CNN + BILSTM deep learning-based DSS for efficient prediction of judicial case decisions. *Expert Systems with Applications*, *209*, 118318. https://doi.org/10.1016/j.eswa.2022.118318

Bae, J. (2015). *The Criminal Code Competition* (9th ed.). Hong Mun-sa.

Bailey, M. A., & Maltzman, F. (2008). Does legal doctrine matter? Unpacking law and policy preferences on the US Supreme Court. *American Political Science Review*, *102*(3), 369–384. https://doi.org/10.1017/S0003055408080283

Beck, F., Burch, M., Munz, T., Di Silvestro, L. & Weiskopf, D. (2014). Generalized Pythagoras trees for visualizing hierarchies. In *IVAPP '14 Proceedings of the 5th International Conference on Information Visualization Theory and Applications* (pp. 17-28).

Benjamin, S. M. & Desmarais, B. A. (2012). Standing the test of time: The breadth of majority coalitions and the fate of US Supreme Court precedents. *Journal of Legal Analysis*, *4*(2), 445-469. https://doi.org/10.1093/jla/las016

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5-32.

Chatterjee, S. & Hadi, A. S. (2015). *Regression analysis by sample* (K.-Y. Kim, H.-C. Kang, M.-S., Jeon, & S.-G., Lee). Free Academy.

Epstein, L, Martin, A. D., Quinn, K. M., & Segal, J. A. (2007). Ideological drift among supreme court justices: Who, when, and how important. *Northwestern University Law Review*, *101*(4), 1483-1542.

Katz, D. M., Bommarito II, M. J. & Blackman, J. (2017). A general approach for predicting the behavior of the Supreme Court of the United States. *PLOS ONE*, *12*(4), e0174698. https://doi.org/10.1371/journal.pone.0174698

Kim, I.-S. & Seo, B.-H. (2016). *The newly written Criminal Code Competition* (8th Jeungbo Ed.). Park Young-sa.

Kim, S. C. (2020). *Criminal law review*. Sojin.

Kim, S.-D. (2016). *The Criminal Code Competition* (4th ed.). Sungkyunkwan University Press.

Kim, Y.-B. (2008). The meaning and criterion of medical malpractice (negligence) from moderating the burden of proof in a medical malpractice suit. *The Korean Society of Law and Medicine*, *9*(1), 57-127.

Lee, E. D., Broedersz, C. P. & Bialek, W. (2015). Statistical Mechanics of the US Supreme Court. *Journal of Statistical Physics*, *160*(2), 275-301. https://doi.org/10.1007/s10955-015-1253-6 23.

Lee, J.-J. & Lee, H.-Y. (2008). Understanding the standard of the Korean Supreme Court's decisions in the obscenity cases. *Journal of Media Economics & Culture*, *6*(3), 79-132.

Lee, J.-S., Jang, Y.-M., & Kang D.-B. (2016). *Criminal law introduction* (10th ed.). Park Young-sa.

Lee, Y.-K., Park, K.-R., & Lee, H.-Y. (2021). Tax Judgment Analysis and Prediction using NLP and BiLSTM. *Journal of Digital Convergence*, *19*(9), 181-188.

Lyu, J. S. (2022). *An analysis of the meaning of publicity and the independence of private schools shown among case laws related to private schools* (master's thesis). Korea National University of Education.

Martin, A. D. & Quinn, K. M. (2002). Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, *Political Analysis*, *10*(2) 134–153. https://doi.org/10.1093/pan/10.2.134

O, Y. (2014). *Criminal Law Competition* (3rd ed.). Park Young-sa.

Park, S.-K. (2011). *Criminal Law Competition* (8th ed.). Park Young-sa.

Park, Y.-C & Lee J.S. (2022). Analysis of Supreme Court of Korea's Judgment Criteria on the Degree of Assault and Intimidation - For Empirical Verification of the Court's Judgment Process. *Journal of AI Humanities (JAIH), 10*(0), 77-114. 10.46397/JAIH.10.3

Pyo, H. S., Young, I. S., Tteum, G.E., Cho, H. S., Yun, J. Z. & Yeon, L. J. (2020). A study on the possibility of the use of criminal big data through the analysis of sentencing. *The Korean Association of Police Science Review, 22*(6), 209-230.

Recommended Guidelines on the Writing Style of Legal Judgments. (1998, August 20). *Judicial Precedent Rules No. 625-1*, Enforced on August 20, 1998.

Rzepka, R., Shibuki, H., Kimura, Y., Takamaru, K., Matsuhara, M. & Murakami, K. (2008). Judicial precedents processing project for supporting Japanese lay judge system. In *Workshop on Semantic Processing of Legal Texts*, LREC2008, 33-41.

Segal, J. A. & Spaeth, H. J. (1996). The influence of stare decisis on the votes of United States Supreme Court justices. *American Journal of Political Science*, *40*(4), 971-1003. https://doi.org/10.2307/2111738

Segal, J. A. & Spaeth, H. J. (2002). *The Supreme Court and the attitudinal model revisited*. Cambridge: Cambridge University Press.

Segal, J. A., Epstein, L., Cameron, C. M. & Spaeth, H. J. (1995). Ideological values and the votes of US Supreme Court justices revisited. *The Journal of Politics*, *57*(3), 812–823. https://doi.org/10.2307/2960194

Shapiro, C. (2008). Coding complexity: Bringing law to the empirical analysis of the Supreme Court. *Hastings Law Journal*, *160*(2), 60, 477.

Supreme Court of Korea. (1993, March 9). Decision in 92do2884 case.

Supreme Court of Korea. (2012, July 12). Decision in 2012do4031 case.

Tran, V. D., Nguyen, M. L., Shirai, K. & Satoh, K. (2019). An approach of rhetorical status recognition for judgments in court documents using deep learning models. In *2019 11th International Conference on Knowledge and Systems Engineering* (KSE) (pp. 1-6). IEEE.

APPENDIX

Appendix 1. Cases in the recoding sample (Korean Court cases).

The cases presented in this appendix, titled "Cases in Recording Sample (Korean Court Cases)," are directly sourced from judgments in South Korea. The decision to include these cases in an appendix is due to their extensive number (over 100), making it impractical to list them in the main bibliography.

Case Names
Daegu High Court. (1981, December 4). Decision in 81 No 527 case.
Daegu High Court. (1979, July 19). Decision in 79 No 323 case.
Daejeon High Court (Cheongju). (2016, August 18). Decision in 2016 No h69 case.
Incheon District Court. (1992, May 21). Decision in 92 No 176 case.
Jeju District Court. (1988, December 23). Decision in 88 Gohap 114 case.
Seoul High Court. (1973, April 23). Decision in 73 No 239 case.
Seoul High Court. (2011, July 21). Decision in 2011 No 1393 case.
Seoul High Court. (1977, May 6). Decision in 77 No 293 case.
Seoul High Court. (1985, October 30). Decision in 85 No 2225 case.
Seoul High Court. (1975, September 26). Decision in 75 No 1028 case.
Seoul Southern District Court. (1985, November 29). Decision in 85 Gohap 907 case.
Supreme Court of Korea. (1971, April 20). Decision in 71 do 441 case.
Supreme Court of Korea. (2010, April 29). Decision in 2007 do 7064 case.
Supreme Court of Korea. (1984, February 14). Decision in 81 do 3202 case.

Case Names
Supreme Court of Korea. (2001, February 23). Decision in 2000 do 4415 case.
Supreme Court of Korea. (2000, February 25). Decision in 99 do 4305 case.
Supreme Court of Korea. (1985, February 26). Decision in 84 do 2732 case.
Supreme Court of Korea. (1983, February 8). Decision in 82 do 2714 case.
Supreme Court of Korea. (2007, January 25). Decision in 2006 do 5979 case.
Supreme Court of Korea. (2020, January 30). Decision in 2018 do 2236 en banc case.
Supreme Court of Korea. (1985, June 25). Decision in 84 do 2083 case.
Supreme Court of Korea. (2004, June 25). Decision in 2004 do 2611 case.
Supreme Court of Korea. (1981, March 24). Decision in 81 do 409 case.
Supreme Court of Korea. (1983, March 8). Decision in 82 do 2838 case.
Supreme Court of Korea. (1974, May 14). Decision in 73 do 2578 case.
Supreme Court of Korea. (1991, May 28). Decision in 91 do 80 case.
Supreme Court of Korea. (1956, May 8). Decision in 56 do 50 case (judgment).
Supreme Court of Korea. (1980, November 25). Decision in 79 do 2565 case.
Supreme Court of Korea. (2018, October 25). Decision in 2018 do 7709 case.
Supreme Court of Korea. (1979, October 30). Decision in 79 do 1660 case.
Supreme Court of Korea. (2001, October 9). Decision in 2001 do 3594 case.
Daegu High Court. (1998, September 15). Decision in 98 No 308 case.
Seoul High Court. (1998, August 8). Decision in 98 No 988 case.
Seoul High Court. (1995, September 19). Decision in 95 No 1686 case.
Supreme Court of Korea. (2013, April 11). Decision in 2010 do 13774 case.
Supreme Court of Korea. (2001, April 27). Decision in 2001 do 230 case (judgment).
Supreme Court of Korea. (2006, April 27). Decision in 2003 do 4151 case.
Supreme Court of Korea. (1999, April 9). Decision in 99 do 519 case.
Supreme Court of Korea. (1991, August 13). Decision in 91 do 1184 case.
Supreme Court of Korea. (1990, August 14). Decision in 90 do 114 case.
Supreme Court of Korea. (2000, August 18). Decision in 2000 do 1914 case.
Supreme Court of Korea. (1976, August 24). Decision in 76 do 1932 case.
Supreme Court of Korea. (2019, August 29). Decision in 2018 do 13792 en banc case.
Supreme Court of Korea. (2012, August 30). Decision in 2012 do 6157 case.

Case Names
Supreme Court of Korea. (2002, December 10). Decision in 2001 do 7095 case.
Supreme Court of Korea. (1990, December 11). Decision in 90 do 2224 case.
Supreme Court of Korea. (2013, December 12). Decision in 2013 do 11899 case.
Supreme Court of Korea. (2007, December 13). Decision in 2007 do 7601 case.
Supreme Court of Korea. (1992, December 22). Decision in 92 do 2596 case.
Supreme Court of Korea. (1986, December 23). Decision in 86 do 2203 case.
Supreme Court of Korea. (1993, December 24). Decision in 93 do 2339 case.
Supreme Court of Korea. (1997, February 14). Decision in 96 do 1959 case.
Supreme Court of Korea. (2019, February 14). Decision in 2018 do 19493 case.
Supreme Court of Korea. (1994, February 22). Decision in 93 do 428 case.
Supreme Court of Korea. (2001, February 23). Decision in 2000 do 5395 case.
Supreme Court of Korea. (1997, February 25). Decision in 96 do 3411 case.
Supreme Court of Korea. (1989, February 28). Decision in 87 do 690 case.
Supreme Court of Korea. (2006, January 13). Decision in 2005 do 4799 case.
Supreme Court of Korea. (2005, January 28). Decision in 2004 do 227 case.
Supreme Court of Korea. (2012, July 12). Decision in 2012 do 4031 case.
Supreme Court of Korea. (2005, July 15). Decision in 2004 do 1565 case.
Supreme Court of Korea. (1969, July 22). Decision in 69 do 927 case.
Supreme Court of Korea. (2012, July 26). Decision in 2011 do 8805 case.
Supreme Court of Korea. (1993, July 27). Decision in 93 do 901 case.
Supreme Court of Korea. (2005, July 28). Decision in 2005 do 3071 case.
Supreme Court of Korea. (1986, July 8). Decision in 86 do 93 case 1.
Supreme Court of Korea. (2007, June 1). Decision in 2006 do 4449 case.
Supreme Court of Korea. (2000, June 9). Decision in 2000 do 1253 case.
Supreme Court of Korea. (1998, March 10). Decision in 98 do 70 case.
Supreme Court of Korea. (1990, March 27). Decision in 89 do 2036 case.
Supreme Court of Korea. (1995, March 28). Decision in 95 do 91 case.
Supreme Court of Korea. (1993, March 9). Decision in 92 do 2884 case.
Supreme Court of Korea. (1999, March 9). Decision in 99 do 242 case.
Supreme Court of Korea. (1991, May 10). Decision in 90 do 2102 case.

Case Names
Supreme Court of Korea. (2007, May 10). Decision in 2007Do1375 case.
Supreme Court of Korea. (2003, May 13). Decision in 2003Do709 case.
Supreme Court of Korea. (2013, May 16). Decision in 2012 jeondo 252 (conjugated) en banc case.
Supreme Court of Korea. (1991, May 28). Decision in 91 do 546 case.
Supreme Court of Korea. (1984, May 9). Decision in 84 do 573 case.
Supreme Court of Korea. (2010, November 11). Decision in 2010 do 9633 case.
Supreme Court of Korea. (1985, November 12). Decision in 85 do 2115 case.
Supreme Court of Korea. (1991, November 26). Decision in 91 do 2267 case.
Supreme Court of Korea. (1988, November 8). Decision in 88 do 1628 case.
Supreme Court of Korea. (2017, October 12). Decision in 2016 do 16948 case.
Supreme Court of Korea. (1987, October 26). Decision in 87 do 1656 case.
Supreme Court of Korea. (2017, October 26). Decision in 2015 do 16696 case.
Supreme Court of Korea. (2004, October 28). Decision in 2004 do 4437 case.
Supreme Court of Korea. (2001, October 30). Decision in 2001 do 4462 case.
Supreme Court of Korea. (2013, September 13). Decision in 2013 do 6809 case.
Supreme Court of Korea. (1993, September 14). Decision in 93 do 915 case.
Supreme Court of Korea. (1996, September 20). Decision in 95 do 1728 case.
Supreme Court of Korea. (1999, September 21). Decision in 99 do 2608 case.
Supreme Court of Korea. (1985, September 24). Decision in 85 do 1687 case.
Supreme Court of Korea. (1991, September 24). Decision in 91 do 1824 case.
Supreme Court of Korea. (1996, September 24). Decision in 96 do 2151 case.
Supreme Court of Korea. (2003, September 26). Decision in 2003 do 763 case.
Supreme Court of Korea. (2005, September 29). Decision in 2005 do 4738 case.

Pouzdanost testova u sudskoj praksi utvrđena analiziranjem kriterija za donošenje presuda korejskog presedana o stupnju napada i zastrašivanja

SAŽETAK

Primjena kriterija iz prijašnjih presuda u potonjim slučajevima je nejasna. Jedan od tih faktora, stupanj napada i zastrašivanja, utječe na ozbiljnost zločina, ali ga javnost može teško procijeniti jer nije jasno otkud potiče. U svrhu rješavanja ovog problema, brojne studije nastoje utvrditi temeljne elemente koji empirijski dokazuju apstraktne koncepte. Studija uvodi metodu za analiziranje presuda, izvlačeći elemente svakog problema prema istim uzorcima iz prethodnih sličnih slučajeva. Za razliku od prijašnjih pristupa, ovaj pristup može analizirati brojne presude kroz detaljnu obradu podataka. Ova se metoda oslanja na kriterije za donošenje presuda i utvrđuje kriterije razvrstavanja prema teoriji, što dovodi do rezultata linearne regresije. Model pokazuje kako važnost elemenata razlikuje kod svakog problema, čime dobivamo uvid u logiku donošenja presuda. Javnost može koristiti ovaj model za unaprjeđivanje predvidljivosti presuda analiziranjem presedana. Štoviše, može se koristiti za procjenjivanje pouzdanosti u angloameričkom sustavu i sustavu građanskog prava koji prepoznaju sudsku praksu.

Ključne riječi: analiza humanističkih podataka, linearna regresija, presedani, analiza važnosti, izbor obilježja.