Model for Data Training Judge's Decisions in Corruption Cases in Indonesia

Shidarta
Faculty Member, Business Law Department
Faculty of Humanities
Bina Nusantara University
Jakarta, Indonesia
shidarta@binus.edu

Teguh Prasandy
Faculty Member, Information Systems Department
School of Information Systems
Bina Nusantara University
Jakarta, Indonesia
teguh.prasandy@binus.edu

Yi Ying
Faculty Member, Chinese Literature Department
Faculty of Humanities
Bina Nusantara University
Jakarta, Indonesia
shidarta@binus.edu

Erick Fernando
Faculty Member, Information Systems Department
School of Information Systems
Bina Nusantara University
Jakarta, Indonesia
erick.fernando@binus.ac.id

Abstract

Data is one of the supporting elements for decision-making in all scientific fields, including in legal dogmatics. As a practical science, legal dogmatics is responsible for responding to legal problems in societies, particularly through court decisions. In the Indonesian legal tradition, which is heavily influenced by the civil law system, judges in charge of deciding cases do not always use references in the form of precedents drawn from similar case decisions. Situations like this often attract criticism as they are likely to cause inconsistencies in court decision-making and will ultimately reduce public confidence in the courts. This paper offers a way to overcome this tendency and aims to model judgment decisions from cases that are related to cases through data mining with the decision tree method. The methods used are Decision Tree ID3, with data derived from decisions on corruption cases decided by the Corruption Court in Indonesia over the past ten years. This data-training model has shown variables that can be used as data-training that really needs consistency in decision making in corruption cases in Indonesia. Considering that the judges' decisions are the result of factual as well as normative reasoning, there are many elements that must be anticipated. The judge's inconsistency in interpreting the facts results in the judge's interpretation of the elements of the crime. This disparity is the source of the weakness of the model that is currently presented in this study. One way to overcome this is to increase the data variance. If the researcher only uses duplicate data from the previous training data, the end result will not be very different.
Keywords
Judge's decision, Corruption, Decision tree, Data mining, Data-training.

1. Introduction
This research is a study of judges' decisions to build a training data model to create new initiatives on judge decisions in Indonesia. The results of the literature search did not let anyone see previous research on this subject. In fact, the judge's decision is a very challenging text to be researched through text mining studies. Understanding judicial decision-making requires consideration of the specific institutional framework in which judges operate. The options available to judges are determined not only by the law and the facts of the case, but also by the procedural context (Schlanger, et al, 2009).

This study has several obstacles that can be identified as the cause of the lack of initiative formation so far. One of them, legal documents in the form of judges' decisions have characteristics that are not easily analysed textually by researchers who do not have a legal background. The judge's decision generally starts from a real case which has different details in each case. On the other hand, the study of the text of the judge's decision requires multidisciplinary, interdisciplinary, or transdisciplinary cooperation. Therefore, the creation of a training data model as part of a text-mining study must involve several scientific fields together. This kind of collaboration is rarely carried out by legal researchers in Indonesia, especially in the field of study called meta-norms.

As technology continues to transform how we work and function, there are predictions that many aspects of human activity will be replaced or assisted by new technologies. While many human activities have changed over time because of human progress, recent changes in the context of technological change have forced some human functions, previously largely unaffected, to become more active. It can have far-reaching implications. In this respect, technology is already transforming legal practice and can reshape the judicial process, for example by replacing, supporting, or supplementing the role of the judiciary (Tania Sourdin, 2018; Rodriguez, 2021).

This article shows the initial efforts of interdisciplinary collaboration involving three scientific fields. First, this research gets valuable help from computer science, which is very reliable in any text mining study. Second, this research involves linguistics because it is related to text and cultural analysis. Judges' decisions in Indonesia must be clearly understood in Indonesian text format and include linguistic analysis. Third, this study also covers the legal dogmatic area because the texts of judges' decisions contain normative provisions using distinctive legal terminology. The relationship between the elements of the normative provisions used requires an explanation through the perspective of legal science. In addition, the novelty of this research is triggered by the realm of legal scholarship because the text to be studied is a legal product that has certain characteristics, which are different from text analysis in general.

The specificity of the analysis of legal texts is because this text refers to positive legal norms (Ashley, K. D., 2021; Watson, 2022; Nay, 2021). This standard is general and abstract. General, because it is addressed to everyone and not to certain people (ad hoc) (Michael and Joshua, 2011). Also, it can be called abstract because the object and its normative conditions do not lead to a particular context. The object of the norm is the behaviour that is commanded, prohibited, permitted or excluded. In addition, there are standard conditions that provide information about where, when, and when the standard applies.

The availability of court judgment documents in digital form opens up a wide range of possibilities for obtaining and using the information. Automatic summarization of these legal texts is an important and difficult task due to the unusual structure and high degree of complexity of these documents. Previous approaches in this direction rely on huge, labeled datasets, use hand-constructed functions, leverage domain knowledge, and pay attention to narrow subdomains to increase effectiveness (Deepa Anand & Rupali Wagh, 2019; Arpan Mandal, et al, 2021).

This study focuses on the text of the decisions of judges dealing with cases related to Article 12 paragraph (1) and Article 3 of Law No. 31 of 1999 in conjunction with Law No. 20 of 2001 in also conjunction with Law No. 19 of 2019 on the elimination of criminal acts of corruption. The formulation of Article 2 paragraph (1) reads as follows: "Everyone who unlawfully commits an act of enriching himself or another person or a corporation that can harm state finances or the state economy, shall be sentenced to imprisonment for life or imprisonment for a minimum of 4 (four) years and a maximum of 20 (twenty) years and a minimum fine of Rp. 200,000,000.00 (two hundred million rupiah)
and a maximum of Rp. 1,000,000,000.00 (one billion rupiah).” This formulation is slightly different from the formulation of Article 3, which in full reads as follows: "Every person who with the aim of benefiting himself or another person or a corporation, abuses the authority, opportunity or means available to him because of his position or status that can harm state finances or the state economy, shall be punished with life imprisonment or a minimum imprisonment of 1 (one) year and a maximum of 20 (twenty) years and or a minimum fine of Rp. 50,000,000.00 (fifty million rupiah)."

These two pieces of legislation are highly controversial but tend to curb corruption in Indonesia. All elements in these two articles are identified as to whether they correspond to the behaviour of legal subjects accused of specific cases. If it is fulfilled, the judge imposes legal consequences in the form of punishment; if it is not fulfilled, the judge pronounces acquittal or releases from all claims.

The choice of these two articles has a special reason, because the wording of these articles contains problems that are considered "defects of law". Article 3, which defines the subject-matter of the norm as abuse of office, turns out to be an opportunity to impose a milder sanction than Article 2 paragraph (1), which does not call for abuse of authority. This "legal defect" was found in several judges' decisions. Such things are interesting to observe when analysing whether the search for a series of judges' decisions dealing with corruption cases reveals the same judgment pattern. The models specified in the decision will be very open to being read differently by the training data (Amal Joby, 2021). The problem to be solved is how the training data model of the judges' decisions should be used in order to make meaningful suggestions from a legal dogmatic point of view. Since the purpose of this research is to create a training data model and analyse the model is still a pilot project, it is certain that the model will have many weaknesses. This article analyses these weaknesses so that the model can be improved from time to time.

2. Literature Review

2.1 Judge's Decisions

Judges are legal professionals that have an important influence on law enforcement. The law in Indonesia that regulates judicial power, namely Law no. 48 of 2009, does not provide a connotative definition. The law states that judges are judges at the Supreme Court and judges at judicial bodies under them in the general court environment, religious court environment, military court environment, state administrative court environment, and judges at special courts within the judicial environment.

For corruption cases, the courts that handle them are special courts within the general judiciary, namely the courts for criminal acts of corruption which have been established since 2009 in every provincial capital in Indonesia. Courts at the first level are authorized to hear corruption cases that are prosecuted by prosecutors from the Corruption Eradication Commission and the Prosecutor's Office. The panel of judges consists of three or five people, not all of whom are career judges.

Against the decisions of the judges at the court of first instance, the prosecutor and/or the defendant has the right to file an appeal to the high court (Bench-Capon, T,2020). District courts and high courts are essentially fact-checking institutions, so they can present defendants, witnesses, and experts at trial. This is different from the court of cassation at the Supreme Court which only examines documents containing written texts (McCarty, 2018).

Court decisions in the procedural law system in Indonesia have had a relatively standardized and consistent format since before Indonesia's independence. The verdict consists of the head of the decision, the identity of the subject involved in the case, consideration of facts, legal considerations, sentencing, and closing. In a corruption case court decision, the identity referred to is the name, place of birth, age, gender, nationality, residential address, religion, and occupation of the defendant. The subjective factors of these judges may relate to the decisions they make, considering that everyone has a tendency towards certain values. Family background, education, social status, and religion have long been suspected as contributors to the birth of this subjectivity. Jerome Frank (1931) once even said that a judge's decision could be influenced by trivial things like what he had eaten for breakfast. However, such factors will not be easily traced by simply highlighting the text of the verdict. Grajzl et al, (2012) say, that the effectiveness of a judicial system depends crucially on the ability of the courts to resolve cases and to resolve them in a prompt manner. This is especially true of legal systems characterized by court backlogs and delays (Greenleaf, 2018). From a policy perspective, therefore, it is essential to understand the various determinants of court output.
The text of decisions in corruption cases that attract attention is usually in the amount of financial or economic damage to the state (Hovell, 2013). These losses must correlate with the type and severity of the punishment. In cases of corruption, the penalty can include imprisonment, a fine, and money exchange. If the fine is not paid, the convicted person must serve a prison sentence as a substitute for the fine. In Indonesian criminal law, imprisonment and imprisonment have different meanings. The difference lies in the rights granted to the convict. Convicts undergoing internment receive more rights than convicts in prison. Replacement payment is a form of punishment that requires the convict to pay a sum of money to compensate the state for losses that his crime has caused. The payment must be made no later than one month after the decision on the case becomes final. If the payment is not made, the prison term will be added.

2.2 Model for Data Training
It was stated that, in relation to legal reasoning, the process which is worth studying is the process of argumentation is a process of justification (MacCormick, 1994). The data-training model, thus, illustrates a justification of legal reasoning as shown in the judges' decisions.

Decision tree induction teaches the decision tree by training tuples with class tags (Schmidhuber, 2015). A decision tree is a flowchart-like tree structure in which each internal node (not a leaf node) defines a test for an attribute, each branch represents the result of the test, and each leaf node (or leaf node) contains a test label class is. The top node of the tree is the root node (Han and Pei, 2012). Decision tree classifications make predictions using multiple "if...then..." conditions (LeCun, 2015). This is like operating instructions in various programming languages that you have learned. A decision tree consists of a root node, branches, and leaf nodes. Each inner node represents an input attribute condition, each branch specifies the result of the condition, and each leaf node contains a class label. The root node is the top node of the tree. This requires a collection of sample datasets that are used to train a rule so that a particular decision classifier is called training data (Bhatia, 2019; Joby, 2021).

Decision tree guided algorithms are widely used in various knowledge discovery and pattern recognition fields. They have the advantage of producing understandable classification/regression models and satisfactory levels of accuracy in multiple application areas such as medical diagnostics and credit risk assessment (Rodrigo C. Barros, et al, 2015; Soniya, 2015).

3. Methods
The text used as the object of investigation in this research was the decisions of judges in corruption cases. Researchers took decisions that were on the official website of the Supreme Court which contained decisions at the cassation level in 2018 and 2019. All these decisions were categorized as important decisions. This category indicated that each decision contained legal considerations that had a contribution to the interpretation of criminal law regulations on eradicating corruption. The articles that directly related to these cases were Article 2 paragraph (1) and Article 3 of the Corruption Eradication Law.

The researcher analysed these decisions and presented them into columns that mostly followed the systematics of the decisions. After the data was transformed, the researchers then analysed it using a rapid miner. The next researcher compared the results of the logic design shown by the judges with the results of the data training.

3. Result and Analysis
The first thing that must be done in making this data-training model is the construction of a decision tree. For legal scholars, the basis for the construction must start from the formulation of the rules. This requires caution because each prescriptive rule must first be analysed, which consists of the components of the norm subject, the norm operator, the norm object, and the norm condition. If these components are applied to Article 2 paragraph (1) and Article 3 of the Corruption Eradication Law, the results will be shown in Table 1 as follows:
Table 1: Components of Primary Norm of Article 2 Paragraph (1) and Article 3

<table>
<thead>
<tr>
<th>COMPONENT</th>
<th>ARTICLE 2 PARAGRAPH (1)</th>
<th>ARTICLE 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norm Subject</td>
<td>Every one</td>
<td>Every one</td>
</tr>
<tr>
<td>Norm Operator</td>
<td>prohibited</td>
<td>prohibited</td>
</tr>
<tr>
<td>Norm Object</td>
<td>to commit an act of enriching:</td>
<td>to abuse:</td>
</tr>
<tr>
<td></td>
<td>- himself, or</td>
<td>- the authority, or</td>
</tr>
<tr>
<td></td>
<td>- another person; or</td>
<td>- opportunity, or</td>
</tr>
<tr>
<td></td>
<td>- a corporation</td>
<td>- means</td>
</tr>
<tr>
<td>Norm Condition</td>
<td>unlawfully</td>
<td>with the aim of benefitting:</td>
</tr>
<tr>
<td></td>
<td>that can harm[s]</td>
<td>- himself, or</td>
</tr>
<tr>
<td></td>
<td>- the state finances, or</td>
<td>- another person, or</td>
</tr>
<tr>
<td></td>
<td>- the state economy</td>
<td>- a corporation</td>
</tr>
<tr>
<td></td>
<td>that can harm[s]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- the state finances, or</td>
<td>- the state economy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*) The word "can" in this article has been declared by the Constitutional Court through its decision number 25/PUU-XV/2016 as contrary to the constitution and has no binding force.

All the articles mentioned above are rules that contain prescriptive norms. This norm will have consequences for the applicability of the rule. The applicability is sentencing in the form of imprisonment and fines. Only these two types of punishment are mentioned in these articles. It turned out that in the decisions studied, the judges handed down more than two types of sentences. In the Indonesian Criminal Code, the two types of punishment are the main punishments that allow additional penalties to be given. Article 18 of the Law on the Eradication of Corruption Crimes provides for several additional types of crimes. In the judges' decisions studied, it was found that two types were applied, namely replacement payment and confirmation of goods.

Frederick Scahuer (1994) says that a prescriptive norm will be formed into a rule by taking the form of a hypothetical proposition. The rule has two parts. The first part is called protasis or operative facts which specify the scope of the rule, the factual conditions triggering the application of the rule. Schauer prefers to call it a factual predicate.

If the factual predicate in Article 2 paragraph (1) and Article 3 of the Law on the Eradication of Corruption Crimes is described in a decision tree, it will be seen that the judge's decisions have four prostrations. They all must deal with facts that have to be proven. If the judge decides that these elements are proven, then a logically valid argument of the form 'if p then q; p;∴ q'.

The two formulations of this article are confusing because they are essentially the same. The object of the norm in Article 2 paragraph (1) changes to the condition of the norm in Article 3. The difference briefly can be seen in Article 3 which states that the object of the norm is the abuse of authority, opportunity, and means, so it is interpreted that the subject of the norm of that article is different from the subject of the norm of Article 2 paragraph (1). That is, the subject of this norm of Article 3 should not be every person, but every state official. The new problem is confusing because the secondary norm, namely the sanctions for violating Article 3 is actually lighter than the violation of Article 2 paragraph (1), even though a person's status as a state official should be an aggravating factor.

Elucidation of Article 2 paragraph (1) seeks to resolve this confusion by saying that what is meant by "unlawfully" includes acts against the law in a formal sense as well as in a material sense. That is, even though the act is not regulated in laws and regulations, it is still an act that can be punished if the act is considered disgraceful because it is not in accordance with the sense of justice or the norms of social life in society. In this provision, the word "can" before the phrase "damaging the state's finances or economy" indicates that a criminal act of corruption is a formal offense, i.e., the existence of a criminal act of corruption is sufficient if the elements of the act that have been formulated are fulfilled, not with the consequences. The Constitutional Court immediately refuted this way of thinking through its decisions number 003/PUU-IV/2006 and 25/PUU-XV/2016. This change in the way of grammatical interpretation presented by the Constitutional Court has made judges' decisions bolder to use these two articles more
pragmatically. If the judges want a lighter sentence, the judge will tend to use Article 3 instead of Article 2 paragraph (1).

All decisions in this research were identified thanks to the data mining process, which then presented a description relating to: (1) sex of the accused, (2) his/her age, (3) working status, (4) the applicant in cassation, (5) the elements of criminal offences serving as the legal basis for the indictment, (6) aggravating factors, (7) mitigating factors, and (8) penalties imposed by judges. Sanctions for acts of corruption in Indonesia have several variants, both cumulative and alternative. In addition to imprisonment, the accused is usually also required to pay a fine of a specified nominal amount. If this fine is not paid, the defendant is obliged to replace it with a prison term of up to six months. If judges determine a loss to state finances, they usually fine the defendant to pay the damages. If the judges find that the sums cannot be paid, the judge can order that the defendant's property be confiscated and turned over to the state. A textual study of these decisions then finds data showing the relationship between one text and another.

<table>
<thead>
<tr>
<th>Decision Number</th>
<th>Sex of the Accused</th>
<th>Working Status</th>
<th>the applicant in cassation</th>
<th>Abuse of authority, unlawfully</th>
<th>Sentencing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>Official</td>
<td>Prosecutor</td>
<td>Yes</td>
<td>Guilty of doing corruption together.</td>
</tr>
<tr>
<td>2</td>
<td>Male</td>
<td>Official</td>
<td>Prosecutor</td>
<td>Yes</td>
<td>Guilty of doing corruption together.</td>
</tr>
<tr>
<td>3</td>
<td>Male</td>
<td>Businessman</td>
<td>The accused</td>
<td>No</td>
<td>proven to have committed but not a crime</td>
</tr>
<tr>
<td>4</td>
<td>Male</td>
<td>Businessman</td>
<td>The accused</td>
<td>No</td>
<td>Guilty of doing corruption.</td>
</tr>
<tr>
<td>5</td>
<td>Female</td>
<td>Businessman</td>
<td>Prosecutor</td>
<td>Yes</td>
<td>Guilty of doing corruption together.</td>
</tr>
<tr>
<td>6</td>
<td>Male</td>
<td>Ex-official</td>
<td>The accused</td>
<td>Yes</td>
<td>proven to have committed corruption continuously</td>
</tr>
<tr>
<td>7</td>
<td>Male</td>
<td>Ex-Official</td>
<td>Prosecutor</td>
<td>No</td>
<td>Guilty of doing corruption together.</td>
</tr>
<tr>
<td>8</td>
<td>Male</td>
<td>Official</td>
<td>Prosecutor</td>
<td>Yes</td>
<td>proven to have committed corruption continuously</td>
</tr>
<tr>
<td>9</td>
<td>Male</td>
<td>Businessman</td>
<td>The accused</td>
<td>No</td>
<td>Guilty of doing corruption together.</td>
</tr>
<tr>
<td>10</td>
<td>Female</td>
<td>Official</td>
<td>The accused</td>
<td>No</td>
<td>Guilty of doing corruption.</td>
</tr>
<tr>
<td>11</td>
<td>Female</td>
<td>Official</td>
<td>The accused</td>
<td>Yes</td>
<td>Guilty of doing corruption together.</td>
</tr>
<tr>
<td>12</td>
<td>Male</td>
<td>Businessman</td>
<td>Prosecutor</td>
<td>No</td>
<td>Guilty of doing corruption together.</td>
</tr>
<tr>
<td>13</td>
<td>Male</td>
<td>Businessman</td>
<td>The accused</td>
<td>Yes</td>
<td>Guilty of doing corruption together.</td>
</tr>
<tr>
<td>14</td>
<td>Male</td>
<td>Businessman</td>
<td>Prosecutor</td>
<td>No</td>
<td>Guilty of doing corruption together.</td>
</tr>
<tr>
<td>15</td>
<td>Male</td>
<td>Businessman</td>
<td>Prosecutor</td>
<td>No</td>
<td>Guilty of doing corruption together.</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 contains examples of several corruption case decisions that have been investigated. Several elements are not listed due to the consideration that they have uniformity. So, only elements that have different answers are placed in the table.

All the elements contained in the primary norms of the two articles (if p) must be sufficient conditions to bring the perpetrator as a norm subject to be punished (then q). In this data-training model, the element of harming state finances or the state economy is an interesting element to be placed as an operative fact or a factual predicate. Figure 1 shows that the element of the magnitude of the state's financial or economic loss will be a determining factor for the severity of the punishment. This element must be accompanied by the condition of the norm "unlawfully" which does not become a determining factor for the severity of the punishment, but whether there is a punishment. If the element "unlawfully" is not fulfilled, the punishment will be lost.

The element of abuse of authority, opportunity, or means is also a determining factor, but this factor only applies to perpetrators of corruption who violate Article 3. Therefore, the two determining factors, if any, should also influence the severity of the punishment. If not, then the defendant cannot be convicted.
This data-training model can entail several propositions based on the ID3 decision tree (see Figure 1).

1. If an action has damaged the finances or the economy of the state, whether supported by aggravating or mitigating circumstances, the consequence of this action is imprisonment and a fine, with the possibility of the addition of penalties for compensatory money and confiscation of property; noting that failure to pay the fine will be replaced by imprisonment and failure to pay the replacement money by imprisonment.

2. If an action constitutes an abuse of power, opportunity, and means, whether supported by aggravating and mitigating circumstances, the action is punishable by imprisonment and a fine, with the possibility of additional penalties being imposed. Substitute currency and confiscation of property; noting that failure to pay the fine will be replaced by imprisonment and failure to pay the replacement money by imprisonment.

3. Jika suatu perbuatan adalah perbuatan melawan hukum, terlepas apakah perbuatan itu didukung oleh faktor pemberat atau peringan, maka perbuatan itu akan dipidana dengan penjara dan denda; noting that failure to pay the fine will be replaced by imprisonment and failure to pay the replacement money by imprisonment.

4. If an action is not unlawful even if the means are misused, the action is free from any claim for sanctions.

The first three propositions are statements that can be obtained directly through grammatical interpretation of the articles in criminal law and the law on eradicating corruption. There is nothing special about the propositions. The downside of these three propositions is that they often cannot be read individually. The data-training model must pay attention to the formulas in the hypothetical syllogism regarding necessary and sufficient conditions. An antecedent placed as an “if” proposition must be considered carefully whether it is a necessary condition or a sufficient condition. For example, is the condition for state financial loss the only condition for causing legal consequences, or are there other conditions? The researcher, therefore, must anticipate this situation. In Indonesian legislation and regulations, the phrase "if and only if" is never found, indicating that this phrase is a sufficient condition. This means that legal scholars must pay close attention to the mistakes in making the model.

The fourth proposition also contains weaknesses since if an act is not against the law, the sanction should be non-existent, in the sense that the act of corruption is not proven to have occurred. The proposition says that his actions are indeed corruption, but the perpetrators cannot be punished. For that, there are two possibilities. First, there was an error in the legal reasoning made by the judge in the decision. Second, this data-training model may not contain one addition, namely regarding the presence or absence of a valid-defence in the criminal act. The weakness of a judge's decision can be traced from an illogical proposition according to the standards of criminal law.
The weakness that is very likely to occur is when the researcher conducts an analysis of Table 2 that is too specific to the perpetrator of the crime. For example, the gender of the defendant. Some of the propositions below demonstrate these weaknesses. Among other things, the propositions read as follows: (1) if an action is an abuse of authority, and the perpetrator is a woman, then she is guilty of committing a crime of corruption together; or (2) if an action is an abuse of authority, and a perpetrator is a man, then he is guilty of committing a crime of corruption together. Propositions like this clearly have no meaning because regardless of the gender of the perpetrator, the verdict is the same.

From the design of the data-training model and the results of the rapid miner, it should be seen that the decision tree reflects the design thinking of the judges in their decisions. When the researchers investigated the fines in these decisions, it turned out that the fines were not related to the amount of state losses. The same is true for employment status. In fact, even though there are state losses that are quite large, not all decisions provide penalties for the convict to compensate for the loss by requiring him to pay compensation.

This data-training model, therefore, provides a lesson about the difficulties that must be faced by legal text-mining researchers who study judges' decisions. Because each case presents different facts, it is necessary to map the variables and attributes of the facts that must appear in the table first. All variables and attributes must have the potential to correspond to elements of the normative provisions that form the basis of the indictment law. If this is not done, then the weaknesses due to illogicality in the meaning of the proposition will appear. This is reasonable because the propositions as data training will follow the reading rules of the resulting model. If the researcher wants something different, one way is to add variations to the training data. The more data and variations of the data will produce a variety of decision variations. On the other hand, if the researcher only uses duplicated data from the previous training data, the result will not very much.

4. Conclusion
This research shows that the judge's decision is a very complex document, but it is still open for analysis using the data-training model. The difficulties in transforming all the data to be analysed by rapid miners are due to the variations in each of these facts, both concerning the subject of the perpetrator and the actions that correspond to the object of the norm. Such diverse norm conditions also give rise to multiple interpretations, thus affecting the punishment. The aggravating and mitigating factors are also not all explicitly stated in the decisions. All these variations raise weaknesses that every compiler of the data-training model must anticipate. The authors note that this study requires more training data with many variations. With the limited variations, there seems to be leaps in the judge's way of thinking compared to the propositions displayed by the machine.

Acknowledgments
The authors would like to show their gratitude to Bina Nusantara University for supporting the legal text-mining research.

References
Arpan Mandal; Kripabandhu Ghosh; Saptarshi Ghosh; Sekhar Manda, Unsupervised approaches for measuring textual similarity between legal court case reports. Artificial Intelligence and Law, 2021, doi:10.1007/s10506-020-09280-2
Bhatia, P. Data Mining and Data Warehousing Principles and Practical Techniques. Cambridge University Press, Cambridge, 2019


Biographies

Shidarta is a faculty member of law at Bina Nusantara University. He is the author of several books and articles on legal reasoning, legal philosophy, legal research method, competition law, and consumer protection. He is one of the founders of the Association of Indonesian Legal Philosophy (AFHI) and the Lecturer Forum of Business Competition (FDPU) of which he used to be the chairman of both.
Teguh Pransandy is a ...

Yi Ying Dr. Yi Ying, M.Lit., M.Pd., is an Associate Professor in Curriculum Development at Bina Nusantara University. She obtained her master’s degree in Applied Chinese Linguistics from Xiamen University and Education Administration from Christian University of Indonesia, Jakarta, Indonesia. She obtained her Ph.D. in Curriculum Development from the Indonesia University of Education. Her research interests include Chinese language teaching and learning, mobile-assisted learning, China Studies, and tourism.

Erick Fernando is a ....