**gInvestigating the Legal-Social Justice Gap on Social Media: A Machine-learning-based Sentiment Approach**

**Hwa-Joon Choi**

Email:

**Tae Ik Kim**

Researcher

ModuLabs

324 Gangnam-daero

Gangnam-gu, Seoul 06252

South Korea

Email:

**Ji-Hwan Min**

Researcher

ModuLabs

324 Gangnam-daero

Gangnam-gu, Seoul 06252

South Korea

Email:

**Se-Hyun Lee**

Researcher

ModuLabs

324 Gangnam-daero

Gangnam-gu, Seoul 06252

South Korea

Email:

**Sang-Joon Kim**\*

Associate Professor

Ewha School of Business

Ewha Womans University

52 Ewhayeodae-gil

Seodaemun-gu, Seoul 03760

South Korea

Email: [s.kim@ewha.ac.kr](mailto:s.kim@ewha.ac.kr)

\* Corresponding author

**Investigating the Legal-Social Justice Gap on Social Media: A Machine-Learning-Based Sentiment Approach**

**Abstract**

In this study, we explore whether and how social justice can be deviated from legal justice. Acknowledging that justice is socially constructed, we attempt to empirically specifies how social evaluations on a situation, which, especially, are collectively mobilized on social media, can be differentiated from its official, legal judgment. The difference indicates that legal justice cannot perfectly reflect what the society actually pursues. To empirically compare between legal justice and social justice, we consider how car accidents are differently figured out in the perspectives of legal institutions and users on a social media platform. Using a YouTube channel which mainly deals with the legal judgments of car accidents, we analyze whether and how the social evaluations on the given car accidents from the YouTube users are different from what the legal judgment is actually made. Given that users’ comments on a video clip reflect their perspectives, values, and worldviews on the situation observed in the video clip, the collectively mobilized social evaluations on car accidents are measured by employing a machine-learning-based sentiment analysis of users’ comments. And we compare these social evaluations with the official, legal judgements to specify the legal-social justice gaps. Our findings present that under strong legality, little deviation of social justice from legal justice unfolds, whereas under weak legality, there are many opportunities that can expand the legal-social justice gap.

**Keywords: Justice, legal-social justice gap, machine learning, sentiment analysis, social media, YouTube**

**1. Introduction**

Law reflects society's values, based on changing reality (Barak, 2016). Dickson (2000) said that the traditional law we share must include living things in a broader view. This notion implies that even though legal institutions seem to be deeply involved in our lives, in reality, they do not perfectly reflect the everyday life of the public. Legal institutions have largely been dealt with in the realm of the groups who constitute, implement, and reinforce them, mostly the judiciary and the legislature. As such, those who are not related to legal institutions barely have tendency to pay attention to them (Slovak, 1981). Apparently, social actors in the legal fields are segregated, isolated, and fragmented.

With the advent of social media, such complexity of legal fields has been drastically changed. Over the past decades, various forms of social media have turned various social issues into online contents for the public, some of which have led to social movements and legislation (Harlow, 2012; Kidd & McIntosh, 2016; Lopes, 2014). The public's interests in the legal fields, such petition systems regarding various social issues and agendas, legalization processes, or social evaluations on the legal institutions, have been increased through social media. That is, social media have begun to actively reflect on social justice and express their opinions on certain laws for themselves. This naturally created a deviance between the scale of justice defined by the legal system and the scale of justice that collectively appeared in people's perception (Meyer & Rowan, 1977). Moreover, social media have reinforced the justice deviance, as they have made it clearer that human behavior, believed to be manifested in formal structures, is correspondingly influenced by informal structures (Baym, 202; Harrigan et al., 20211). This clearly shows that legal institutions may not always be the absolute or sole standard for justice. Further, social media enable people to re-define, re-frame, and re-visit the concept of justice from their own perspectives (e.g. Wæraas & Nielsen, 2016; Phillips & Malhotra, 2017). As a result, various definitions of justice are constructed across various social groups. And these socially constructed concepts of justice, called social justice, constitute the society as well (e.g. Koyama & Weasel, 2002; Degoey, 2000). This standpoint calls for research that needs to focus on how the difference between legal justice and social justice can become salient, especially through social media.

To figure out the legal-social justice gap on social media, we employ text-mining approaches (e.g. Luzano & Nakayama, 2021; Wang et al., 2022). Given that the values of institution unfold through actors’ verbalized frames such as discourse, narratives, or other semiotic frames (Phillips & Malhotra, 2017), various forms of verbal expressions on social media can bring broader understandings on how the institutional values can be understood, interpreted, translated, and embodied by institutional agencies (Pries-Heje & Baskerville, 2017; Tracey, Dalpiaz, & Phillips, 2018). In this sense, texts on social media can be a critical and fruitful source to capture how individuals embody the values of different institutions, namely legal institutions and social institutions for this study. Accordingly, we have two empirical strategies to discern legal institutions and social institutions. First, given that law dogmatics are specific by the authoritative source of law (Aarnio, 1986), we use official documents published by judicial agencies to capture legal justice. This is because the public documents of the judiciary, such as rulings and verdicts, have been thought to be sufficient to represent the judiciary's position on social issues. On the other hand, according to Novak (2000), social justice represents issues on social system and focuses on social order. Accordingly, this study collected social justice data from online social media where the public freely expresses their opinions (Jost et al., 2018). In this study, among several social media, we consider YouTube, which provides both video and text as free and open source, as a research setting. YouTube, a video-based social media service, can deliver and store information in real time. And since it is a popular platform, pieces of information related to users' reactions, such as comments, have been studied to look into how social discourses have been formed (Benson, 2015; Liew & Hassan, 2021).

To capture the legal-social justice gaps in YouTube, two contingencies are considered. One contingency is strong legality. The strong legality is found when legal judgments are decisively made as well as authoritative interpretations are converged. Even in cases under the strong legality, if the users’ opinions are not consistent to the legal judgment, we assume that there might be a legal-social justice gap. The other contingency is weak legality. The weak legality is found that legal judgments are equivocal and authoritative interpretations are divergent. Under the weak legality, the legal-social justice gap can unfold in distinctive ways, compared to the strong legality. Based on these two contingencies, this study seeks to derive implications on how social media construct social evaluations vis-à-vis legal judgments and what factors play an important role in the process through the difference between the definition collectively formed in social media and the definition stipulated in the actual legal system.

In this regard, this research is supposed to answer the following two main research questions:

*Research question1: Under the strong legality where messages of legal justice are unequivocally delivered and interpretations of the public on it are limited and convergent, how does the society recognize the artifacts of social deviance between legal justice and social justice?*

*Research question2: Under the weak legality where messages of legal justice are equivocally delivered and interpretations of the public on it are limitless and divergent, how does the society recognize the artifacts of social deviance between legal justice and social justice?*

**2. Theoretical background**

2.1. Justice and social justice

Justice is an abstract concept, but it has been studied and interpreted from very diverse perspectives for a long time because of its great influence on human daily life. Ancient Egyptian texts and the Old Testament view justice as moral uprightness, goodness, faultlessness, and perfection ​​(Udoudom & Bassey, 2018). Justice was also a major topic of discussion among ancient philosophers. The Greek philosopher, Plato said that justice is a thing fundamentally aimed at social harmony and well-being in the structural relationship between individuals and societies (Cornford, 1976). This is also the theoretical basis for social justice established by Plato. Thomas Aquinas said that justice is a strong will to fulfill individual responsibilities (O’Callaghan, 2017). Beyond the philosophical point of view, implications for justice comply with governing laws (Udoudom & Bassey, 2018). The modern Oxford dictionary defines justice as “the fair treatment of people or the quality of being fair or reasonable” (Hornby & Cowie, 1977).

Social justice is a situation which social consensus is not yet reached than justice. Since no dictionary or thesaurus refers to a common and clear definition of social justice (Buettner-Schmidt & Lobo, 2012), there are differences in interpreting social justice. However, since social justice is interpreted individually in various areas of society, it may be summed up and only guessed to be the broad meaning and implications of today's social justice. For social justice in philosophy, there is an opinion that inherits the idea of ​​John Rawls who presented justice as the first virtue of social institutions. They said that a fair society is “one in which free citizens possess equal basic rights in an equal system”, and Wenar (2008) stated that the opinion refers to today’s social justice. The opinion of the legal community is more specific and modern. Kennedy (2005) stated that today's social justice includes relations between race, gender, environment, and nations, and Rand (2006) mentioned minorities who are not fully entitled to social empowerment for the scope of social justice. In psychology, social justice has been discussed in the viewpoint of authority, power, and peer pressure, particularly paying attention to the connection between individuals and others (Hatfield & Rapson, 2005). Sociological literature deals with social and economic inequality and the allocation of the goods in the realm of social justice, in their discussion replacing it with the term 'distributive justice' (Scott & Marshall, 2009).

The focus of interpretation of social justice differs from area to area in a modern sense. In order to sharpen the purpose of the study and to answer the research questions, therefore, this study is required to properly set the research scope of social justice. Given that the purpose of the study is to find out and measure deviance between legal justice and social justice and that the research setting is on social media, sensibly it is assumed that social justice in this study is a combination of interpretations of social justice in social domains excluding legal justice and that social media represents voices of the social domains.

2.2. Social Media and Social Engagement

Galbraith (1974) stated that lack of information increases uncertainty because disparity in information between the two is important for a task with the amount of information required to perform. Ishii, Lyons, and Carr (2019) argued that uncertainty is caused by lack of information and can be solved by increasing the amount of information. In brief, sufficiency of information has a negative relationship with uncertainty.

The level of information richness matters to the media. It is an important premise in explaining the properties of media that deliver information. Media richness theory, comprehensively explaining the properties of media, highlights the importance of the richness of information as well as the amount of information (Daft, Lengel, & Trevino, 1987; Daft & Lengel, 1986). Subsequent studies suggested four factors that determine the richness of information: immediate feedback, multiple cues, language variety, and personal focus.

In this respect, today's social media has increased both the amount and richness of information, so it is suitable for sufficiently collecting and in-depth analysis of public opinions. The public participating in social media freely express their opinions and sentiments. They actively participate by directly sharing contents. Others can react to them in multiple manners: by instantly commenting, clicking sentiment icons and so on. As such, as large amount of information and reaction to it are rapidly circulated, many recent social science studies have chosen social media as a medium for information collection.

The epoch-making development of communication technology triggered by the advent of social media has brought about changes in social and cultural aspects. Stucke (2009) said that social media-based technologies make it possible to form users' personal and cultural networks, and Poster (2018) said that communication technology is not a mere exchange of ideas or information, but a culture, that is, the formation of individual identities. A change in the social and cultural aspects also includes political changes. Kapor (1991) argued that the convergence of old and new communication technologies will transform the media into a freer, more rational and democratic space, providing new political possibilities. Schejter and Tirosh (2015) said that as individual users immediately access and distribute a large amount of information beyond the limits of place and time, users have discovered the potential and democratic environment in which social media happen to contribute to public purposes. In this social trend, social media is an important medium for redefining social justice.

Today social media is catalyzing a paradigm shift in shaping social agenda. Dutton(2008) pointed out that an action that occurred on social media have achieved the democratization of information and culture through the process of sharing social issues and attracting the public attention. For example, in some conservative Arab countries, citizens encountered mass-circulated information on social media and raised social movements aspiring for change (Benkler, 2002; Howard & Hussain, 2011). For those who seek justice, social media may be viewed as an efficient tool to convey and exchange their messages, as well as a space to elevate their presence. In this respect, social media is a good medium to observe today's social justice.

2.3. Text Mining and Sentiment Analysis

Text mining refers to extracting various written sources and then connecting them to find new information with the help of machine (Hearst, 2003). Hearst (2003) pointed out that text mining differs from traditional web search in that it aims to find new information. Tan (1999) defined text mining as a process of finding interesting patterns or knowledge in unstructured text, adding that this could be regarded as an extension of data mining that operates on existing standardized databases.

Although text mining is complicated, in general they may be made up of two main steps (Tan, 1999). The first step is to convert free-form unstructured text into an intermediate form. The second is to turn the intermediate form obtained in the previous step through the process of knowledge distillation into concrete form or knowledge. Examples of knowledge extraction include clustering, classification, visualization, and predictive models. Recently, more sophisticated natural language processing algorithms and solutions have come out with the development of machine learning, so the time and technical steps required for text mining has been drastically shortened.

Text mining helps convergence research a lot. Text mining techniques have been in great demand for various social science areas beyond the realm of computer science. They have made it easy for researchers to access, collect, and analyze unstructured data. Good examples include analysis of online posts through text mining and analysis of comments on social network services. Further, researchers become able to analyze patterns between words and context of sentences beyond semantic analysis of individual words (Jeon & Seo, 2013).

Sentiment analysis refers to all the analysis in the minds of people showing various sentiments and opinions about an object (Liu, 2012). Kim and Song(2016) defines sentiment analysis as a specific process of classifying or quantifying sentiments found in unstructured data in an objective way and transforming them into structured data. Since humans specifically express various sentiments through language, sentiment analysis is also an area of ​​text mining. Accordingly, many sentiment analysis studies are being conducted using text mining techniques (Medhat, Hassan, & Korashy, 2014).

Online social media, where opinions can be freely exchanged and saved, can be the most convenient research subject for sentiment analysis (Feldman, 2013). In the social media world where various social events are shared and spread in real time, users actively show their sentiments. At the same time, the influence of social media communication is increasing, which is beneficial for academic research with sentiment analysis. Therefore, this study took the social media platform YouTube as the subject of sentiment analysis.

**3. Research Methods**

3.1. Research Settings

To examine the discrepancies between social and legal justice, it is assumed that both legal and social justice are equivalently observed and captured and that the operational definitions of social and legal justice are comparable and consistent in terms of their scales and meanings. That is, this study aims to empirically capture and compare the legal and social justice based on the common measurement metrics of verbalized frames. For this, YouTube was selected as our research setting, as it provides an open space in which users can share their thoughts on a particular legal judgment. YouTube, as a global social media, has been used in various studies because it is easy to collect users' sentiment and opinions (Severyn, Uryupina, Plank, Moschitti, & Filippova, 2014), and this study also utilized these characteristics.

This study chooses a YouTube channel of South Korea as a specific research setting. It is a channel run by a lawyer, an expert in car accident litigation. It has gained public popularity due to its common topic and active channel users. Channel users self-report their car accident videos and the lawyer, the operator of the channel, gives advice on them. As for each car accident video, users make a number of comments. As of October 2022, more than 18,000 accident report videos have been gathered, and it has over 1.6 million subscribers. There are a few reasons for selecting this channel as the research setting. First, since the vehicle accident videos on the channel are shared by the driver involved in the accident without editing, it makes it possible for one to objectively judge the on-site situation. Second, car accidents can bring up various social agendas under which different individual perspectives can be exchanged and sometimes grievances can be collectively mobilized, such as drunk driving, speeding, child safety etc. Third, car accidents are accompanied by court rulings, making it easy to obtain legal opinions and data on the vehicle accidents. Thus, the channel is suitable for observing public opinions on legal decisions. On top, the channel has a space for description on videos, and some car accident videos uploaded to the channel include the contents of the court ruling in the description part. The channel structure lets the channel visitors see the legal judgment on the accident video and in turn make comments on it. Since legal rulings and opinions of unspecified users coexist in one space, it is useful to compare legal and social justice.

To explore the legal-social justice gaps, we sampled video contents from the channel. The focal channel for this study has a content format in which a dashcam video clip is provided, facilitated by a lawyer; underneath such formatted video contents, various comments from the users are posted on the YouTube platform. In the channel, assumed that the judgments on car accidents are oftentimes controversial and vary depending on what aspects one pays more attention to, accordingly, three discursive components are found: official legal judgments, a lawyer’s interpretations, and users’ evaluations. Among these discursive components, we assume that the official judgments represent legal justice and the users’ evaluations are considered as social justice. As each video content includes these three discursive components, the video contents on the channel are considered as the unit of analysis in this study.

Given this, for analysis, this study considered two contingencies, aforementioned in introduction: strong legality and weak legality. For the strong legality, this study considered a sample of car accidents with verdicts in which either the offender or the victim was 100% at fault. Most judgments in vehicle accidents tend to end in partial negligence on both sides. A judgment of complete negligence on one side is a rare case and can be seen as a clear judgment in terms of legal justice. As for the first research question, it is possible to find out the deviance from the aspect of social justice by examining the case that ended with complete negligence on one side and the opinions of social media users about it. For the weak legality, on the contrary, this study considered a sample of car accidents where a part of the victim's negligence was acknowledged but the degree of the negligence was not clearly determined. As such, there are various opinions and interpretations of the negligence in the video of the accident judged as partial negligence on social media. When multiple authoritative interpretations are possible, the second research question can be answered by analyzing social media users' sentiments about the ruling related to the weak legality.

3.2 Research Design

To analyze the comments on Youtube, we employed a supervised learning-based sentiment analysis approach (Pannala et al., 2016; Zhao et al. 2021; Bhagat al., 2020). The sentiment analysis performed through various machine learning techniques are useful to go through the labeling process of positive and negative comments about sentiment (Bhagat al., 2020).

In this study, to figure out the users’ reactions to each trial through comments, we proceeded in five steps: (1) data collection, (2) data preprocessing, (3) information extraction and semanteme classification, (4) labeling, and (5) Analysis. It is illustrated in Figure 1.

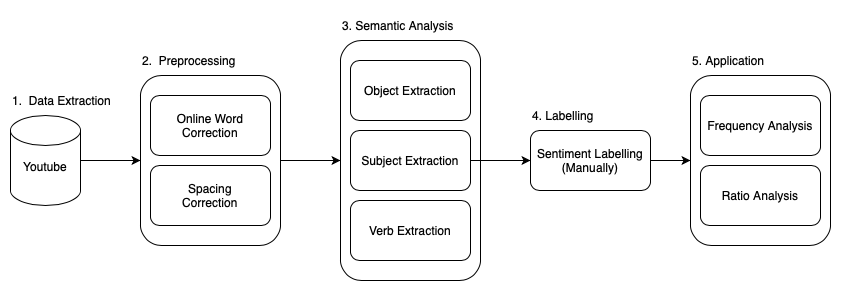


Figure 1. Sentiment analysis procedure

3.2.1. Data Collection

Step1. Data Extraction

In the data collection step, the video addresses and metadata of the focal YouTube channel were collected with the technical support of the YouTube API. Among the posted car accident videos, only the videos including the actual verdict in the metadata were sorted out together and all comments corresponding to the video were collected as well. Meanwhile, the subject words for the sentiment analysis were limited to 'judge' or 'judgment'. As ‘judge’ and ‘judgement’ delivers messages of legal justice, it seems that users are likely to make comments on them. Considering that this study is a comparison between legal justice and social justice, it is necessary to extract words representing legal justice from comments. Thus, it is assumed that the comments containing 'judge' and 'judgment' are appropriate for the sentiment analysis.

On an additional note, in this study, if there are multiple legal judgments in one video, it is specified that the most recent judgment was referred to. The Korean judicial system consists of a three-instance trial system. Most traffic accident judgments are finalized in the first trial, but some are finalized in the second and third trials. Most of the accident videos used for the study sample were also concluded in the first trial, but some were concluded in the second trial. There was no video of the third trial. In other words, some videos contain both the first and second trial judgments, but in the case of multiple judgments, only the final judgment, that is, the second trial judgment was used as a research sample.

Among more than 15,000 videos at the onset of this study, May 2022, on the focal channel, 154 videos containing judgments and 50,320 comments on them in total were extracted. A description of the sample videos is briefed in Exhibit 2, and the sample comment briefed in Exhibit 3.

|  | Judgements by fault ratio | | | | | Total  (Videos) |
| --- | --- | --- | --- | --- | --- | --- |
| 100:0 | 90:10 | 80:20 | 70:30 | others |
| 1st trial | 32 | 17 | 32 | 15 | 17 | 113 |
| 2nd trial | 17 | 3 | 9 | 5 | 7 | 41 |
| Total | 49 | 20 | 41 | 20 | 24 | 154 |

Exhibit 2. A summary of sample videos

|  | Judgements by fault ratio | | | | | Total  (Comments) |
| --- | --- | --- | --- | --- | --- | --- |
| 70:30 | 90:10 | 80:20 | 70:30 | others |
| 1st trial | 9,685 | 2,828 | 11,700 | 2,809 | 5,120 | 32,142 |
| 2nd trial | 7,955 | 1,683 | 4,692 | 1,644 | 2,204 | 18,178 |
| Total | 17,640 | 4,511 | 16,392 | 4,453 | 7,324 | 50,320 |

Exhibit 3. A summary of sample comments on the sample videos

Step2. Preprocessing

The preprocessing step is to process the collected comments into appropriate units for the sentiment analysis of semantemes, the ultimate purpose of this study. YouTube comments, which is commonly colloquial, appear with various text variations. These variants must be properly processed in advance for the following steps.

First, it went through the process of converting the online colloquial comments into written forms. Various forms of emojis generated in an online environment were processed too. Meaningless repeated single vowels/consonants were also removed and modified to be as close to written language to facilitate the next step. Additionally incorrect spacing and sentence patterns were reviewed. Only the comments going through this process moved to the next step.

Step3. Semantic Analysis

In this step the collected comments were divided into semanteme units. For the word separation, the language analysis technology API operated by a national institution ETRI (Electronics and Telecommunications Research Institute) was hired. The semanteme analysis API is an add-on natural language process software application that analyzes a semanteme, the minimum unit in a natural language sentence, based on a tag set.

Due to the nature of Korean language structure, the linguistic part projected with sentiment is noun and semantemes are subjects, so this study has paid attention to the subject of comments. As a result of the API-operated process, subject units split out of the comments. The processed subjects in the descending order of frequency are summarized in Exhibit 4.

| Topic | # of comments | | |
| --- | --- | --- | --- |
| 1st trial | 2nd trail | Total |
| Judge | 3,311 (10.3%) | 2,786 (5.3%) | 6,97 (12.1%) |
| Insurance company | 3,137 (9.8%) | 535 (2.9%) | 3,672 (7.3%) |
| Dashcam vehicles | 1,309 (4.1%) | 635 (3.5%) | 1,944 (3.9%) |
| Accident | 1,097 (3.4%) | 529 (2.9%) | 1,626 (3.25%) |
| Verdict | 770 (2.4%) | 567 (3.1%) | 1,337 (2.7%) |
| Unprotected | 936 (2.9%) | 363 (2.0%) | 1,299 (2.6%) |
| Left-turn | 667 (2.1%) | 366 (2.0%) | 1,033 (2.1%) |
| Negligent | 713 (2.2%) | 311 (1.7%) | 1,024 (2.0%) |
| Driver | 650 (2.0%) | 329 (1.8%) | 979 (1.9%) |

Exhibit 4. Main topics of comments

There are a couple of intriguing points worth pointing out about the API result. First, a number of comments are made on the actors involved in the judgement not in the car accident situation. About 20 percent of the comments interpret about the judge and the insurance company. Second, not all the topics with high frequency are suitable for sentiment analysis. Some words that are difficult to project sentiments into. For example, ‘insurance company’, ‘dashcam car’, and ‘left-turn’ are not the objects to which social media users’ sentiments are directly projected. Therefore, we decided to narrow the set of subject units with high frequency down to the two units to which users project direct sentiments about the car accident video: judgment and judge. The users’ comments, which came out with the adjectives and complements connected the sentiments about the judgment, were considered as the ones onto whom social media users often projected their sentiments. Out of the 50,320 comments obtained in the first step, only 12,912 comments that contained judgment or judge were used for the remaining steps.

3.3.2. Data Analysis

Step 4. Labeling by sentiment

For labeling, random sampling was performed because it was time-consuming and hard to manually label all the comments. Alternatively, some 5% of the comments obtained in the previous step, that is, 600 comments, were randomly selected for sentiment analysis. The 600 comments were divided into two clusters of 300 each according to the level of legality to answer the research questions. The first cluster represents the strong legality, with 300 comments from the videos with a verdict with its fault ratio of 100:0, to answer the first research question. The second cluster represents the weak legality to answer the second research question. It also delves into 300 comments from the video with the various fault rates. Finally, the two clusters were ready for labeling

The 600 comments of the two clusters were manually labelled by four researchers, and classified into three categories: positive, negative, and neutral. That is to say, every comment was labelled one of the three sentiment types. Each researcher read the same comment directly and decided positive, negative, or neutral based on the contents superficially revealed in the comment in a blind situation without knowing the classification value of the other researcher. As for the classification value of the comments, if all the values ​​were matched, that value was assigned, and if not, the most frequent value was assigned. If the final classification value did not reach an agreement, it was judged as an error value. Out of the 600 comments, there was only one error value.

Step5. Application

After classifying and labeling the selected comments, descriptive statistics such as frequency analysis and ratio analysis were performed. The analysis results juxtaposed the numbers and figures of the strong legality and those of the weak legality. Finally they were to be compared in different dimensions

**4. Result**

The result of the sentiment analysis depicts users’ comments in two dimensions: sentiment and the level of legality. It is briefed in Exhibit. 5 as below.

|  | Weak Legality | Strong legality |
| --- | --- | --- |
| Positive | 23 (7.67%) | 112 (37.45%) |
| Negative | 237 (79.0%) | 78 (26.08%) |
| Neutral | 40 (13.33%) | 109 (36.54%) |
| Uncategorized | 0 | 1 |
| Total | 300 | 300 |

Exhibit 5. Result of sentiment analysis

The two groups showed a difference in the number of positive, negative, and neutral comments. To see if the difference is statistically significant, this study conducted a significance test.

While the rate of the positive comments for the group under the strong legality was 37.46% and its standard error was 0.028, and that of the positive comments under the weak legality was 7.67% and its standard error was 0.015. It indicated that the difference was statistically significant with 99.9% reliability. The rates of negative comments for each group were 26.09% and 79.00%, and the standard errors were 0.025 and 0.024 respectively. It also came to a significant difference with 99.9% reliability. Last, the rates of neutral comments under the strong legality and the weak legality were 36.45% and 13.33%, and their standard errors were 0.028 and 0.020, respectively, showing a significant difference with 99.9% reliability.

The results indicate that for the weak legality the number of negative comments almost ten times more than that of positive comments whereas for the strong legality the number of negative comments is lesser than that of positive comments. Moreover, some 36 percent of the strong legality group keeps neutral sentiment, which is quite contrary to the fact that only 13 percent of the weak legality group keeps a neutral stance. As for the weak legality it may be considered that not all the reduced negative reactions are converted to positive reactions, clearly a fairly large proportion remain in a neutral position. These numbers imply that the weak legality group is more likely to be deviant from the judgement, and less interpretation is available for the strong legality.

When the result of the sentiment analysis is divided by the level of trial and then by the degree of legality, some intriguing findings come out as shown in Exhibit 6. For reference, there is no third trial judgment in this study, because the Korean judiciary concludes most vehicle accident judgments at the second trial.

|  | Legality | Sentiment | | | |
| --- | --- | --- | --- | --- | --- |
| Positive | Negative | Neutral | Total |
| 1st trial | Strong | 95 (44.39%) | 56 (26.17%) | 63 (29.44%) | 214 (100%) |
|  | Weak | 4 (2.96%) | 116 (85.92%) | 15 (11.1%) | 135 (100%) |
| 2nd trial | Strong | 34 (25.0%) | 40 (29.41%) | 62 (45.59%) | 136 (100%) |
|  | Weak | 2 (1.80%) | 97 (87.39%) | 12 (10.81%) | 111 (100%) |

Exhibit 6. Result of sentiment analysis by the level of trial

The rate of positive response in the strong legality in the first trial and in the second trial was 44.39% and 25%, and the standard errors were 0.034 and 0.037. With 99% statistical confidence, it means that in the case of the strong legality, users respond more positively in the first trial responds more positively than in the second trial. In contrast, there is little difference in the strong legality the between the first trial and the second trial. Their rates are 26.17% and 29.41%, and the standard errors were 0.030 and 0.039 respectively. In statistical words, they are not hard to tell from each other. Last, the neutral response rates were 29.44% and 45.59%, and the standard errors were 0.031 and 0.042. With 99% confidence more users under strong legality show neutral emotions in the second trial than in the first trial. In brief, under the strong legality the proportion of the positive responses decreases while that of the neural responses increases with no significant chance in the negative emotion as the trial goes by from the first to the second.

As for the weak legality, the distribution of the sentiments is similar regardless of the level of trial. Commonly the negative response outnumbers the others a lot, accounting for around 85 % of the entire responses both in the first trial and the second trial. Even the proportion of the neural response, which is comparatively high in the strong legality, is low.

**5. Conclusion**

This study investigates how social justice can be deviated from legal justice under the strong and weak legality in addition to the answers to the research questions. Under the strong legality, where the car accident has been judged by the judiciary to be 100:0 negligence, the rate of positive comments (37.45%) is higher than the rate of negative comments (26.08%). The neutral rate (36.54%) is also high, which means that only about 1 out of 4 users who commented complain about the car accident verdict. In other words, when the judiciary determines that the cause of an accident is obvious, and the public generally tends to agree. In this case, since the gap between legal and social justice is not wide, the tendency of social deviance is not expected to be high. On the other hand, under the weak legality, where the judiciary acknowledged mutual negligence, the negative rate (79%) overwhelms the positive (7.67%) and neutral (13.33%). On average, 4 out of 5 comments written after watching the video are dissatisfied with the ruling of the judiciary. Comments on the judgement that went through the appeal process and were finally judged as partial negligence in the second trial reached 87.3% of negative comments as shown in Exhibit 6. This particular high rate of negative sentiment implies that users are more likely to get upset at the final verdict with partial negligence because they see such cases as the point where legal justice is most deviant from social justice.

In the case of a car accident that ended with partial negligence in the judiciary, it tends to have a high possibility of more authoritative interpretations. In this case, the difference between the legal definition of the judiciary and the social justice gathered from social media is very large. In other words, social media users, representative of the collective definition of social justice in our research setting, rarely agree with legal definitions.

This study provides the following implications. First, this study sets a popular interest as a research topic to find out the difference between legal justice and social justice. Transportation is a means used by many people in today's daily life, and there are large and small accidents frequently. According to statistics from the Traffic Accident Analysis System (TAAS), 229,600 incidents occurred as of 2019. Also, this study takes a YouTube channel with the cumulative number of views exceeded 2 billion views at the time of this study as a research sample. It is meaningful in that it transforms the public concern of traffic accidents into a measure of social justice and compares it with legal justice.

Second, this study has quantified the deviance between legal and social justice. Since there could be dissatisfaction with the judgment of the judiciary in many social areas, there will definitely be a deviance between legal justice and social justice, found in several previous studies. However, studies to quantify the deviance difference were not common. In this study, the opinions of social media platforms were analyzed and quantified to estimate the deviance between social justice and legal justice in numbers. Third, a trend of social deviance was found. Specifically, it was found that the negative sentiment rate for vehicle accidents judged as partial negligence was much higher than as total negligence. In addition, it was found that the negative rate was higher in the 2nd trial judgment than in the 1st trial judgment even for the same partial negligence vehicle accident. It can be guessed that people's expectation psychology is reflected according to the rate of accident negligence and the level of the judiciary, which can be interpreted as a factor leading to social deviance of social justice (Oliver, 1977).

Along with the various implications of the study, this study has the following limitations. First, the research sample of this study may have inherent bias. The car accident video sharing YouTube channel, a research sample, is run by the lawyer specializing in traffic accidents. Along with the public interest of delivering correct information related to traffic accidents, the purpose of video sharing is to raise issues with the judgment of the judiciary. In this environment, video viewers, which are social media uses at the same time, can easily express their opinions of criticism on the judgment of the judiciary. In addition, a media condition that people easily sympathize with other comments also encourages bias. The video titles or popular comments can easily affect comments of social media users. For the reasons mentioned above, this study sample may have bias. Third, it is a difference in the method used in the research. In this study, a Korean semanteme analysis API was used for the sentiment analysis, and the authors of this study directly read the sentences and conducted sentiment analysis. If semantemes are analyzed with other APIs, or if others conduct sentiment analysis, there would be differences in the research results.