Transportation Research Part F 81 (2021) 41–54

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

Autonomous vehicle safety: Understanding perceptions of pedestrians and bicyclists

Subasish Das

Texas A&M Transportation Institute, 3500 NW Loop 410, San Antonio, TX 78229, United States

ARTICLE INFO

Article history:
Received 13 November 2020
Received in revised form 18 March 2021
Accepted 25 April 2021

Keywords:
Autonomous vehicles
Pedestrians
Bicyclists
Perception
Association rules mining
Topic modeling

ABSTRACT

Autonomous vehicle (AV) technologies have been rapidly advancing. One benefit of AVs is that the technology could eliminate many driver errors and also mitigate many pedestrian and bicyclist collisions. Real-world AVs have been tested in many cities. Five companies are running around 50 AVs in Pittsburgh, following the autonomous testing guidelines. BikePGH, a non-profit organization located in Pittsburgh, Pennsylvania conducted a follow-up survey in 2019 (the first survey was conducted in 2017) to understand non-motorists' opinions of AVs. This study examined how pedestrians and bicyclists perceived AV safety based on their understanding and experiences. At first, this study performed a comparison group test to determine which questions vary by participants' AV safety rating. The responses were later analyzed with a data mining method known as 'association rules mining.' A new performance measure, known as the rule power factor, was then used to identify the significant patterns in the form of rules. The participants also provided their thoughts in responses to the open-ended questions. Using Latent Dirichlet Allocation (LDA), a topic modeling algorithm, 40 topic models were developed based on five open-ended questions. The findings show that the non-motorists showed comparatively fewer negative opinions towards AVs than positive assessments. The results also show that perception patterns vary by the participant's rating on AV safety. Findings of this study would be beneficial for the AV stakeholders in making AVs and roadways safer for non-motorists.

1. Introduction

Non-motorized traffic fatalities are increasing greatly. In 2018, 7,140 pedestrians and bicyclists were killed in road traffic crashes (a 3.75% increase from 2017, and a 50% increase from 2009) (NHTSA, 2020a, NHTSA, 2020b). Furthermore, in collisions involving pedestrian-vehicle and bicyclist-vehicle, non-motorized users are more likely to be killed or injured.

Full automation level AVs can handle various roadway traffic scenarios without human input, which reduces crashes caused by human errors. An extensive implementation of AVs will also lessen traffic congestion and air pollution. It is essential to comprehend end-user's understanding and mindset towards AVs to make sure that consumers are aware of its many advantages (Das et al., 2019). In March 2018, a pedestrian walking her bicycle was crossing a four-lane arterial road around 10p.m. and was struck by a 2017 Volvo XC90 that was operating autonomously under modification from Uber's sensors and software. Since the occurrence of this fatal collision incident, safety concerns associated with AV deployment have increased. A non-profit organization named BikePGH, located in Pittsburgh, Pennsylvania, conducted a follow-up survey in 2019 (first

E-mail address: s-das@tti.tamu.edu

https://doi.org/10.1016/j.trf.2021.04.018
1369-8478/

survey conducted in 2017) to understand non-motorists' perception of AVs. It is important to note that BikePGH changed some of the earlier questions to focus more on safety perception about AVs (BikePGH, 2020). It is also interesting to note that they included a question about the participants' change of opinion following the Uber AV crash incident. The participants were asked to rate AV safety. This study took this unique question design as an opportunity to examine the non-motorist's perception towards AV safety based on their real-life interactions and knowledge on AVs. The research questions are as follows: (1) Does perception on AV safety differ by participant knowledge, prior interactions with AVs, and other key factors? (2) What are the key traits of the participants based on their AV safety rating?

To answer the research questions, this study applied association rules mining (ARM) to the recent 2019 survey data conducted by BikePGH to provide rule-based insights regarding non-motorists' perception towards AV safety. Conventional survey analysis methods are usually limited to perform extended categorical analysis from a survey. In many cases, the interpretations and conclusions are restricted to broad generalizations. Additionally, a majority of statistical methods are limited to the single variable effect, thus the clustering or group effect is often not considered. The ARM method can provide measures such as lift and rule power factor in a way to provide contexts of the proportion measures for various sub-clusters in the dataset. This study showed the value in restricting the analysis to smaller clusters to gain understanding about participants' response patterns towards AVs while contemplating whether they in the past had any interaction with AVs as pedestrians and bicyclists.

2. Literature review

Understanding customer perception of AVs is a critical issue that many studies have explored. The current literature review comprises AV safety and issues associated with AVs and non-motorists.

Howard and Dai (2014) assessed public attitudes regarding AVs by utilizing the responses of 107 adopters in Berkeley, California. An evaluation determined the vehicle characteristics that people liked and disliked, and the participants also expressed how they envisioned the inclusion of the technology. The positive attitudes were associated with finding right parking spots, potential safety improvements from AVs, and multitasking while driving. On the other hand, people were also concerned with the cost of the vehicles, liability, and losing control. Men have been found to be less concerned with control and more concerned with liability than women. Kyriakidis et al. (2015) investigated the willingness, attitudes, and acceptance of users to buy partially, highly, and fully automated vehicles. An Internet-based survey with 63 questions led to 5,000 responses from 109 countries. Overall, the results indicated that the most enjoyable mode of driving for respondents was manual driving. Additionally, respondents were mostly concerned with safety, legal issues, and software hacking/misuse.

Pillai (2017) simulated a virtual scene that facilitates interactions between pedestrians and driverless vehicles. This study designed a framework for vehicle–pedestrian interactions in an AV, which can model driverless vehicle behavior before AV technology is deployed on a large scale.

Nearly 1,000 individuals were interviewed by Hulse et al. (2018) on their perceptions in regard to the acceptance and safety of self-driving cars. The findings showed that little opposition was indicated to the use of AVs on public roads and AVs are perceived as a "somewhat low risk" form of transportation. Vehicles like existing autonomous trains were perceived as riskier. Age, gender, and risk-taking features were found to contribute to the general attitudes towards these cars and perceived risk of different vehicle types. For example, younger adults and men expressed greater acceptance. Canis (2018) determined that 30 percent of survey participants conveyed reluctance in buying an AV. Also, this study reported that more than half of U.S. drivers "feel less safe at the prospect of sharing the road with a self-driving vehicle." Promoting pedestrians and bicyclists' belief in AVs is a critical part of easing public distrust. Jayaraman et al. (2018) attempted to address this shortcoming of non-motorist and AV interactions by examining AV and pedestrian interactions via the uncertainty reduction theory. The results of this study suggest that the type of crosswalk affects the impact of aggressive driving on trust in AVs. The AV's driving behavior had little impact on trust at signalized crosswalks, but it had a major impact at unsignalized crosswalks.

With two natural language processing (NLP) tools, Das et al. (2019) evaluated people's attitudes toward AVs and the existing polarities regarding the content and automation level. Patterns of top key words and topics were identified using approximately seven million words from a large number of YouTube videos. The study found that the engagement with AVs technologies from the public is more pronounced than in the past and that the automation level directly increases with the possible perception of safety. Finally, they concluded that sentiments towards AVs were positive more often than uncertain, negative, or litigious sentiments. Deb et al. (2017) investigated pedestrian attitudes toward fully autonomous vehicles (FAVs) by creating and distributing a pedestrian receptivity questionnaire for FAVs. The findings showed that the safety and interaction scores significantly predicted the pedestrians' intention to cross the road in front of FAVs, but the compatibility score did not. All three subscale scores predicted the acceptance of FAVs in the existing traffic system. In another study, Deb et al. (2018) examined the external features of FAVs to determine the potential features that could help pedestrians understand an FAV's intended behavior at a crosswalk, affect their crossing behavior, and improve their receptivity toward FAVs. The study found that the inclusion of external features significantly increased pedestrians' receptivity toward FAVs.

After examining 2017 survey data collected by BikePGH, Penmetsa et al. (2019) revealed that respondents with direct experience or interaction with AVs reported significantly higher safety advantage expectations than respondents lacking any AV interaction experience. Das et al. (2020a) used 2017 BikePGH survey data to examine whether the perception towards AVs differ in non-motorists based on real-life AV interactions. The results also show that participants with real

AV interactions have higher interest and expectations opposed to those without any interactions with AVs. Rasouli and Tsotsos (2020) explored the patterns of pedestrians' attitudes by surveying pedestrian behavior studies, including conventional studies on pedestrian-driver interaction and more recent ones that involve AVs.

The current state-of-the-art literature review has several limitations. First, the perception indications are not dependent on the potentially safe adoption of AVs. Second, the interaction between participants and AVs are unknown in most of the studies. Third, very few studies have investigated the perception of AVs in regard to non-motorists such as bicyclists and pedestrians. Fourth, conventional survey analysis methods are limited in their interpretation of the cooccurrences of the responses. To understand the perception of roadway users about the adoption of AVs, the current study presents a suitable data analysis tool with the ability to identify associations and patterns from a survey.

3. Methodology

To answer research question 1, Chi-square test has been performed to select the questions that vary by the participant's AV safety rating. Association rules mining was performed to respond to research question 2. Additionally, open-ended responses were analyzed using topic modeling technique. Fig. 1 shows the flowchart of the research design.

3.1. Association rules mining

In transportation safety analysis, association rules mining has been widely used (Montella, 2011; Montella et al., 2012; Das et al., 2018, 2019a, 2019b, 2019c; Montella et al., 2020; Das et al., 2020b; Kong et al., 2020). Association rules mining has been capability of identifying intuitive results from a large complex dataset. Due to the nature of the algorithm, a large unsupervised dataset (in which response and explanatory variables are not defined) can be easily examined by developing important decision rules. Consider $I = \{i_1, i_2, i_3, ...i_n, \}$be a set of N distinct items. Consider D be a set of transactions where each transaction T comprises of a set of items, such that $T \in I$. Each transaction is can be related to individual identifier. An association rules is shown in the form of ***Antecedent*** $\rightarrow$ ***Consequent*** or $\boldsymbol{A} \rightarrow \boldsymbol{B}$, where $A \in I$ and $B \in I$. The parameters used for the association rule algorithm are support, confidence, and lift. Support indicates probability of transactions having both antecedent and consequent. Confidence indicates association between antecedent and precedence part of the rule. Many researchers have proposed different performance measures for rule mining to extract intuitive rules. Lift is the most common interest measure. Lift measures how many times more often A and B occur together than expected if they were statistically independent. A lift value 1 indicates independence between A and B. If the value of lift is greater than 1, it indicates that A and B appear more frequently together in the data and are considered to be positively dependent on each other. In actual, the lift inclines to be higher for large itemsets as compare of small itemsets. This study used rules power factor (RPF), an interest measure developed by Ochin and Joshi (2016), which provides more informative regarding importance of rules. If the relative count of A and B is greater in a transaction than count of A and B in another transaction by keeping total transaction constant, confidence is not sufficient to determine the important rule. RPF, a recently developed performance measure, focuses on the importance of association between antecedent and consequent of rules. When antecedent and consequent association increases, rule importance increases and hence RPF. The equations of the related measures are shown in Table 1. Three example cases are also provided to provide the context of using RPF. The first four rows provide count for each of antecedent, consequent, and their combinations. The third column of the table provides equation of each of performance measures (support, confidence, lift, and RPF). For example, the frequencies in Case 3 are higher than the frequencies of Case 2.

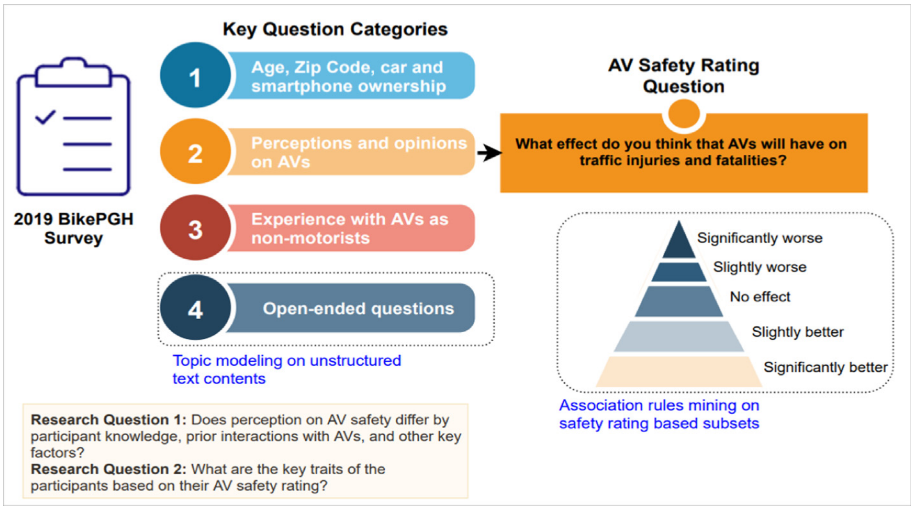

Fig. 1. Flowchart of the research design.

Table 1
Association rules formula with examples.

Desc	Notation	Equation	Case 1	Case 2	Case 3
Count of A	$n(A)$	–	20	40	60
Count of B	$n(B)$	–	30	60	80
Count of AB	$n(AB)$	–	6	25	49
All Counts	n	–	100	100	100
Support of A	$S(A)$	$S(A) = \frac{n(A)}{n}$	$S(A) = \frac{20}{100} = 0.20$	$S(A) = \frac{40}{100} = 0.40$	$S(A) = \frac{60}{100} = 0.60$
Support of B	$S(B)$	$S(B) = \frac{n(A)}{n}$	$S(B) = \frac{30}{100} = 0.30$	$S(B) = \frac{60}{100} = 0.60$	$S(B) = \frac{80}{100} = 0.80$
Support of rule A → B	$S(\mathrm{A} \rightarrow \mathrm{B})$	$S(\mathrm{A} \rightarrow \mathrm{B}) = \frac{n(AB)}{n} = S(\mathrm{B} \rightarrow \mathrm{A})$	$S(\mathrm{A} \rightarrow \mathrm{B}) = S(\mathrm{B} \rightarrow \mathrm{A}) = \frac{6}{100} = 0.06$	$S(\mathrm{A} \rightarrow \mathrm{B}) = S(\mathrm{B} \rightarrow \mathrm{A}) = \frac{25}{100} = 0.25$	$S(\mathrm{A} \rightarrow \mathrm{B}) = S(\mathrm{B} \rightarrow \mathrm{A}) = \frac{49}{100} = 0.49$
Confidence of rule A → B	$C(\mathrm{A} \rightarrow \mathrm{B})$	$C(\mathrm{A} \rightarrow \mathrm{B}) = \frac{S(\mathrm{A} \rightarrow \mathrm{B})}{S(A)}$	$C(\mathrm{A} \rightarrow \mathrm{B}) = \frac{0.06}{0.20} = 0.30$	$C(\mathrm{A} \rightarrow \mathrm{B}) = \frac{0.25}{0.40} = 0.625$	$C(\mathrm{A} \rightarrow \mathrm{B}) = \frac{0.49}{0.60} = 0.817$
Confidence of rule B → A	$C(\mathrm{B} \rightarrow \mathrm{A})$	$C(\mathrm{B} \rightarrow \mathrm{A}) = \frac{S(\mathrm{A} \rightarrow \mathrm{B})}{S(B)}$	$C(\mathrm{B} \rightarrow \mathrm{A}) = \frac{0.06}{0.30} = 0.20$	$C(\mathrm{B} \rightarrow \mathrm{A}) = \frac{0.25}{0.60} = 0.417$	$C(\mathrm{B} \rightarrow \mathrm{A}) = \frac{0.49}{0.80} = 0.612$
Lift of A → B	$L(\mathrm{A} \rightarrow \mathrm{B})$	$L(\mathrm{A} \rightarrow \mathrm{B}) = \frac{C(\mathrm{A} \rightarrow \mathrm{B})}{S(B)} = \frac{C(\mathrm{B} \rightarrow \mathrm{A})}{S(A)}$	$L(\mathrm{A} \rightarrow \mathrm{B}) = \frac{0.06}{0.20 \times 0.30} = 1.0$	$L(\mathrm{A} \rightarrow \mathrm{B}) = \frac{0.25}{0.40 \times 0.60} = 1.042$	$L(\mathrm{A} \rightarrow \mathrm{B}) = \frac{0.49}{0.60 \times 0.80} = 1.021$
RPF of A → B	$RPF(\mathrm{A} \rightarrow \mathrm{B})$	$RPF(\mathrm{A} \rightarrow \mathrm{B}) = C(\mathrm{A} \rightarrow \mathrm{B}) \times S(\mathrm{A} \rightarrow \mathrm{B})$	$RPF(\mathrm{A} \rightarrow \mathrm{B}) = 0.30 \times 0.06 = 0.018$	$RPF(\mathrm{A} \rightarrow \mathrm{B}) = 0.625 \times 0.25 = 0.15625$	$RPF(\mathrm{A} \rightarrow \mathrm{B}) = 0.817 \times 0.49 = 0.40$

However, the lift measure does not consider this increase. Case 2 shows higher lift value than Case 3. On the other hand, RPF captures the increases of the frequencies. Thus, RPF value is higher in Case 3.

3.2. Latent Dirichlet allocation

It is shown in Table 2 that respondents can provide ‘free texts’ for the following five questions:

- Any other suggested laws or regulations?

Table 2
2019 BikePGH questions.

Questions	Type	Code	Options
What is your age?	Personal	age	select age
Zip Code	Personal	zipcode	select Zip code
Do you (or someone in your household) own an automobile?	Personal	car ownership	yes/no
Do you own a smartphone?	Personal	smartphone ownership	yes/no
How familiar are you with the technology behind AVs?	Personal	familiarity with AV technology	choose from five options 1–5 *(1 being not at all and 5 being extremely familiar)*
To what extent have you been paying attention to the subject of AVs in the news?	Personal	paying attention to AV news	choose from five options 1–5 *(1 being not at all and 5 being paying very attention)*
In March of 2018, an AV struck and killed Elaine Herzberg, a pedestrian, in Tempe, AZ. As a pedestrian and/or bicyclist, how did this event and its outcome change your opinion about sharing the road with AVs?	Personal	change opinion for Herzberg death	choose from six options 1–5 *(1 being significantly negative and 5 being significantly positive)*
Have you shared the road with an AV while riding your bicycle on the streets of Pittsburgh?	Personal	interact with AV while biking	choose from three options: *yes, no, not sure*
On a typical day, how safe do you feel sharing the road with AVs? *(1 being very unsafe and 5 being very safe)*	Opinion	feel safe with AVs	choose from six options: 1–5 *(1 being very unsafe and 5 being very safe)*
On a typical day, how safe do you feel sharing the road with human-driven cars?	Opinion	feel safe with cars	choose from six options: 1–5 *(1 being very unsafe and 5 being very safe)*
Do you think that it is appropriate to use Pittsburgh’s public streets as a proving ground for AVs?	Opinion	thoughts on Pittsburg as AV testing	choose from three options: *yes, no, not sure*
On City of Pittsburgh public streets, should AV companies be required to report all safety-related incidents with the proper authorities, even if a police report isn’t required?	Opinion	AVs should report safety	choose from three options: *yes, no, not sure*
What effect do you think that AVs will have on traffic injuries and fatalities?	Opinion	AVs reduce injuries and fatalities	choose from five options: *significantly worse, slightly worse, no effect, slightly better, significantly better*
On City of Pittsburgh public streets, do you think that AVs should operate in “manual mode” while in an active school zone?	Opinion	AVs in school zone	choose from three options: *yes, no, not sure*
On City of Pittsburgh public streets, should AV companies be required to share some non-personal data (e.g., number of trips, pickup/drop off locations, number of miles driven) with the proper authorities (e.g. Department of Mobility, PennDOT, Public Safety)?	Opinion	AVs should share data	choose from three options: *yes, no, not sure*
On City of Pittsburgh public streets, should AV speeds be capped at 25 mph when operating in “autonomous mode”?	Opinion	AVs on limited speed roadways	choose from three options: *yes, no, not sure*
On City of Pittsburgh public streets, should AVs have two full-time employees (pilot and co-pilot) at all times?	Opinion	AVs with two drivers	choose from three options: *yes, no, not sure*
On City of Pittsburgh public streets, should AV companies be required to disclose information and data as to the limitations, capabilities, and real-world performance of their cars with the proper authorities?	Opinion	AVS should disclose information	choose from three options: *yes, no, not sure*
Any other suggested laws or regulations?	Opinion	other AV regulations	narrative
Anything else you’d like to share regarding AVs?	Opinion	other comments	narrative
Are you currently an active member of BikePGH?	Membership	member of BikePGH	choose from two options: *yes, no*
Have you been near an AV while walking or using a mobility device (wheelchair, etc.) in Pittsburgh?	Interaction	interact with AVs while walking	choose from three options: *yes, no, not sure*
If you answered YES to either question, were there notable circumstances? What were your observations?	Interaction	Interaction details	narrative
Please share any positive experiences that you have had with an AV.	Interaction	positive AV experience	narrative
Please share any negative experiences that you have had with an AV.	Interaction	negative AV experience	narrative

- Anything else you'd like to share regarding AVs?
- What were your observations with your interaction with AVs?
- Please share any positive experiences that you have had with an AV.
- Please share any negative experiences that you have had with an AV.

As free text is unstructured data, it is important to investigate the insights in these opinions. Topic modeling is an NLP tool that is suitable for exploring the unknowns from unstructured textual contents. Latent Dirichlet allocation (LDA) is a probabilistic topic model that is widely used in topic modeling (Blei et al., 2003). This modeling technique has an effective and efficient probability inference algorithm, and it can generate highly interpretable topics from unsupervised, complex, unstructured textual data (Blei et al., 2003; Blei, 2012; Feuerriegel and Ratku, 2016). The basic idea is simple: each topic is a distribution over words in the vocabulary, and every document is modeled as a distribution over topics. Accordingly, every document is assumed to have been generated by the following process (Blei et al., 2003; Blei, 2012):

1. *Document-topic association*. For every document d in corpus $\mathscr{D}$, pull a random variable $\theta_d \in \mathbb{R}^K$ from the Dirichlet distribution given by θ_d $Dir(\alpha)$. Here θ_d is the relative proportion of K topics that appear in a given document.
2. *Word Count*. For each topic k, pull a random variable β_k $Dir(\eta)$. It identifies the distribution of terms in that specific topic.
3. *Topic-word association*. For every word t in document d, pull a topic $\mathscr{z}_t$ $Mult(\theta_d)$ from a multinomial distribution with θ_d prior and a scaled word frequency tf_t $Mult(\beta_{\mathscr{z}_t})$ from the multinomial distribution.

3.3. 2019 BikePGH survey

In September 2016, testing of semi-autonomous vehicles on Pittsburgh's streets was started by ride sharing service company Uber. Shortly after, Bike Pittsburgh (BikePGH) initiated a survey to capture the perception about AVs from the viewpoint of non-motorists. In 2019, a similar survey was conducted with modifications to the 2017 questions. BikePGH promoted this survey in their website, social media channels, and a few news articles. The respondent location patterns show representativeness of Pittsburgh. It is anticipated that prior public perception may have been changed due to the Uber incident in Arizona (Uber AV struck and killed a pedestrian named Elaine Herzberg). During 2017–2019, rapid changes were made in the AV industry, especially in Pittsburgh area. Since Uber's AV launch in 2016, several companies began testing in Pittsburgh. The state of Pennsylvania has already passed an AV Testing Guidance, and the federal government is near the passing of the American Vision for Safer Transportation Through Advancement of Revolutionary Technologies (AV START) Act. Additionally, AV technology has presumably improved. The follow-up 2019 survey determined any landscape alterations and the sentiments of Pittsburghers on bike and on foot about sharing the road with AVs. To ensure the safest introduction humanly possible, this preparation is necessary in order to deal with the new reality (BikePGH, 2020). Table 2 list the questions included in 2019 survey. The respondents are instructed to respond the questions based on their real-life encounters with the AVs and personal option based on their experience as non-motorists.

3.4. Comparison based AV safety related responses

This study used the final survey dataset with responses from 795 respondents. To answer the first research question, it is important to develop cross tabular based on the participant's safety perception of AVs. The current study is limited to understanding the safety related issues associated with AVs. Statistical analyses were performed using R (version 3.6.0) using the package 'compareGroups' (Salvador, 2020) for descriptive tables. This study defined statistical significance as p-value < 0.05 (statistically significant values are shown with asterisk mark in Table 3). One of the critical safety related question (coded as '*AVs reduce injuries and fatalities*') is '*What effect do you think that AVs will have on traffic injuries and fatalities?*' Respondents can choose any from the five choices: *significantly worse, slightly worse, no effect, slightly better, and significantly better.* Table 3 lists the Chi-square test (or Fisher's exact test if n < 5) cross-tabulation results. The results show that two questions (question on car and smartphone ownership) are not statistically significant based on their respondents' opinions on AV safety. It is also interesting to note that majority of the participants are in support of AV safety (around 73% of the responses are in favor of AV safety). Note that age is an integer in the collected data. This study categorized age by considering the age ranges into four major groups.

4. Results

4.1. Rules mining results

Based on the findings of the Chi-squared test results, the responses that are significantly different by AV safety related questions are considered for association rules mining. This study aimed to perform a supervised rule mining by keeping the consequent fixed for five different responses, respectively. Top rules for each of response criteria are described below. The top rules are selected are based on insightful rules and RPF measures.

Table 3
Chi-squared tests and descriptive statistics for selected questions by opinions on AV safety (*coded as 'AVs reduce injuries and fatalities'*).

Variable Category	Significantly Worse (N = 39)	Slightly Worse (N = 76)	No effect (N = 100)	Slightly Better (N = 270)	Significantly Better (N = 303)	p value
age						<0.001*
< 24	7 (5.69%)	2 (2.63%)	3 (3.00%)	15 (5.55%)	18 (5.94%)	
25–44	12 (30.8%)	29 (38.2%)	43 (43.0%)	128 (47.4%)	166 (54.8%)	
44–54	8 (20.5%)	13 (17.1%)	17 (17.0%)	42 (15.6%)	49 (16.2%)	
54+	16 (41.0%)	32 (42.1%)	37 (37.0%)	85 (31.5%)	70 (23.1%)	
AVs in school zone						<0.001*
No/Not sure	7 (17.9%)	17 (22.4%)	39 (39.0%)	130 (48.1%)	203 (67.0%)	
Yes	32 (82.1%)	59 (77.6%)	61 (61.0%)	140 (51.9%)	100 (33.0%)	
AVs should share data						<0.001*
No/Not sure	5 (12.8%)	14 (18.43%)	28 (28.0%)	58 (21.5%)	97 (32.0%)	
Yes	34 (87.2%)	62 (81.6%)	72 (72.0%)	212 (78.5%)	206 (68.0%)	
AVs on limited speed roadways						<0.001*
No/Not sure	10 (25.6%)	27 (35.5%)	49 (49.0%)	158 (58.5%)	243 (80.2%)	
Yes	29 (74.4%)	49 (64.5%)	51 (51.0%)	112 (41.5%)	60 (19.8%)	
familiarity with AV technology						<0.001*
Extremely familiar	7 (17.9%)	7 (9.21%)	14 (14.0%)	35 (13.0%)	99 (32.7%)	
Mostly familiar	13 (33.3%)	20 (26.3%)	22 (22.0%)	85 (31.5%)	111 (36.6%)	
Not familiar at all	12 (30.8%)	11 (14.5%)	12 (12.0%)	28 (10.4%)	6 (1.98%)	
Somewhat familiar	7 (17.9%)	38 (50.0%)	52 (52.0%)	122 (45.2%)	87 (28.7%)	
change opinion for Herzberg death						0.005*
No change	6 (15.8%)	21 (27.6%)	49 (49.0%)	165 (61.1%)	234 (77.7%)	
Significantly more negative opinion	30 (78.9%)	26 (34.2%)	17 (17.0%)	23 (8.52%)	2 (0.66%)	
Significantly more positive opinion	0 (0.00%)	0 (0.00%)	0 (0.00%)	1 (0.37%)	7 (2.33%)	
Somewhat more negative opinion	2 (5.26%)	29 (38.2%)	33 (33.0%)	78 (28.9%)	48 (15.9%)	
Somewhat more positive opinion	0 (0.00%)	0 (0.00%)	1 (1.00%)	3 (1.11%)	10 (3.32%)	
interact with AV while biking						0.008*
No	16 (41.0%)	36 (48.0%)	34 (34.0%)	82 (30.5%)	84 (27.8%)	
Not sure	9(23.1%)	16 (21.3%)	12 (12.0%)	43 (16.0%)	36 (11.9%)	
Yes	14 (35.9%)	23 (30.7%)	54 (54.0%)	144 (53.5%)	182 (60.3%)	
interact with AVs while walking						0.013*
No	18 (46.2%)	28 (36.8%)	30 (30.0%)	87 (32.2%)	78 (25.8%)	
Not sure	6 (15.4%)	16 (21.1%)	9 (9.00%)	19 (7.04%)	19 (6.29%)	
Yes	15 (38.5%)	32 (42.1%)	61 (61.0%)	164 (60.7%)	205 (67.9%)	
own_car						0.485
No	4 (10.3%)	6 (8.11%)	5 (5.00%)	14 (5.19%)	15 (4.97%)	
Yes	35 (89.7%)	68 (91.9%)	95 (95.0%)	256 (94.8%)	287 (95.0%)	
own_smartphone						0.073
No	5 (12.8%)	8 (11.0%)	5 (5.05%)	8 (2.96%)	5 (1.65%)	
Yes	34 (87.2%)	65 (89.0%)	94 (94.9%)	262 (97.0%)	298 (98.3%)	
paying attention to AV news						0.001*
Not at all	2 (5.13%)	0 (0.00%)	3 (3.00%)	3 (1.11%)	0 (0.00%)	
To a large extent	16 (41.0%)	15 (19.7%)	21 (21.0%)	72 (26.7%)	160 (53.0%)	
To a moderate extent	15 (38.5%)	29 (38.2%)	35 (35.0%)	112 (41.5%)	100 (33.1%)	
To little extent	0 (0.00%)	10 (13.2%)	7 (7.00%)	21 (7.78%)	5 (1.66%)	
To some extent	6 (15.4%)	22 (28.9%)	34 (34.0%)	62 (23.0%)	37 (12.3%)	

*Note: p-values with * indicate statistically significant at 95% confidence level.*

4.1.1. Rules for AV safety improvement as Significantly worse

Table 4 lists top 20 rules in responses in responses to AV safety as 'significantly worse' consequents. Each outcome is listed from the most important to least important rule based on RPF measurements. The most important rule listed on the table, with an RPF of 0.012, is 'change opinion for herzberg death = Significantly more negative opinion.' From the list of the rules, it is found that 'change opinion for herzberg death = Significantly more negative opinion' is present in all 15 rules. A few frequently present responses are 'AVs in school zone = Yes' (4 rules), 'AVs should share data = Yes' (5 rules), and 'AVs on limited speed roadways = Yes' (4 rules). It is also interesting that the responses are associated with the respondents with 'paying attention to AV news = To a large extent' (4 rules). Another finding is that the respondents have no real-life experience with AV interactions as a pedestrian.

4.1.2. Rules for AV safety improvement as slightly worse

Table 5 lists top 20 rules in responses in responses to AV safety as 'slightly worse' consequents. Each outcome is listed from the most important to least important rule based on RPF measurements. The most important rule listed on the table, with an RPF of 0.011, is '*AVs in school zone = Yes*.' From the list of rules, it is found that '*AVs in school zone = Yes*' is present in 9 rules. Some recurring responses on the list include '*AVs on limited speed roadways = Yes*' (8 rules) and '*change opinion for herzberg death = Significantly more negative opinion*' (8 rules). It is also found that these respondents are more like to have real-life interaction with AVs as non-motorists. Similarly, it is found that these respondents are less familiar with AV technology.

Table 4
Top rules when 'opinion on AVs reduce injuries and fatalities is significantly worse'

No.	Antecedent	Support	Confidence	RPF	Lift	Count
A01	change opinion for herzberg death = Significantly more negative opinion	0.038	0.309	0.012	6.420	29
A02	AVs should share data = Yes + change opinion for herzberg death = Significantly more negative opinion	0.035	0.325	0.011	6.770	27
A03	AVs in school zone = Yes + change opinion for herzberg death = Significantly more negative opinion	0.034	0.333	0.011	6.937	26
A04	AVs in school zone = Yes + AVs should share data = Yes + change opinion for herzberg death = Significantly more negative opinion	0.031	0.338	0.011	7.035	24
A05	AVs on limited speed roadways = Yes + change opinion for herzberg death = Significantly more negative opinion	0.032	0.316	0.010	6.586	25
A06	AVs in school zone = Yes + AVs on limited speed roadways = Yes + change opinion for herzberg death = Significantly more negative opinion	0.030	0.343	0.010	7.144	23
A07	AVs should share data = Yes + AVs on limited speed roadways = Yes + change opinion for herzberg death = Significantly more negative opinion	0.030	0.329	0.010	6.838	23
A08	AVs on limited speed roadways = Yes + change opinion for herzberg death = Significantly more negative opinion + paying attention to AV news = To a large extent	0.018	0.483	0.009	10.047	14
A09	AVs in school zone = Yes + change opinion for herzberg death = Significantly more negative opinion + paying attention to AV news = To a large extent	0.017	0.500	0.008	10.405	13
A10	change opinion for herzberg death = Significantly more negative opinion + paying attention to AV news = To a large extent	0.018	0.438	0.008	9.105	14
A11	familiarity with AV technology = Mostly familiar + change opinion for herzberg death = Significantly more negative opinion + paying attention to AV news = To a large extent	0.010	0.667	0.007	13.874	8
A12	change opinion for herzberg death = Significantly more negative opinion + interact with AV while biking = No + interact with AVs while walking = No	0.014	0.478	0.007	9.953	11
A13	AVs should share data = Yes + change opinion for herzberg death = Significantly more negative opinion + interact with AVs while walking = No	0.016	0.429	0.007	8.919	12
A14	change opinion for herzberg death = Significantly more negative opinion + interact with AVs while walking = No	0.017	0.394	0.007	8.198	13
A15	AVs should share data = Yes + change opinion for herzberg death = Significantly more negative opinion + paying attention to AV news = To a large extent	0.016	0.414	0.006	8.611	12

Table 5
Top rules when 'opinion on AVs reduce injuries and fatalities is slightly worse'

No.	Antecedent	Support	Confidence	RPF	Lift	Count
B01	AVs in school zone = Yes	0.073	0.146	0.011	1.564	56
B02	AVs on limited speed roadways = Yes	0.062	0.162	0.01	1.728	48
B03	AVs in school zone = Yes + AVs on limited speed roadways = Yes	0.052	0.196	0.01	2.097	40
B04	AVs on limited speed roadways = Yes + familiarity with AV technology = Somewhat familiar + change opinion for herzberg death = Significantly more negative opinion	0.019	0.5	0.01	5.347	15
B05	familiarity with AV technology = Somewhat familiar + change opinion for herzberg death = Significantly more negative opinion	0.019	0.455	0.009	4.861	15
B06	AVs in school zone = Yes + AVs should share data = Yes	0.061	0.153	0.009	1.632	47
B07	AVs should share data = Yes	0.078	0.104	0.008	1.114	60
B08	AVs on limited speed roadways = Yes + change opinion for herzberg death = Significantly more negative opinion	0.029	0.278	0.008	2.978	22
B09	AVs should share data = Yes + AVs on limited speed roadways = Yes	0.052	0.156	0.008	1.671	40
B10	AVs in school zone = Yes + familiarity with AV technology = Somewhat familiar + change opinion for herzberg death = Significantly more negative opinion	0.017	0.464	0.008	4.965	13
B11	AVs in school zone = Yes + AVs should share data = Yes + AVs on limited speed roadways = Yes	0.043	0.185	0.008	1.983	33
B12	change opinion for herzberg death = Significantly more negative opinion	0.03	0.245	0.007	2.617	23
B13	AVs in school zone = Yes + interact with AV while biking = No	0.036	0.2	0.007	2.139	28
B14	AVs should share data = Yes + familiarity with AV technology = Somewhat familiar + change opinion for herzberg death = Significantly more negative opinion	0.017	0.433	0.007	4.634	13
B15	AVs in school zone = Yes + AVs on limited speed roadways = Yes + change opinion for herzberg death = Significantly more negative opinion	0.025	0.284	0.007	3.033	19
B16	AVs in school zone = Yes + AVs on limited speed roadways = Yes + familiarity with AV technology = Somewhat familiar	0.027	0.239	0.007	2.552	21
B17	interact with AV while biking = No	0.045	0.141	0.006	1.509	35
B18	familiarity with AV technology = Somewhat familiar	0.047	0.121	0.006	1.296	36
B19	AVs in school zone = Yes + change opinion for herzberg death = Significantly more negative opinion	0.025	0.244	0.006	2.605	19
B20	AVs should share data = Yes + change opinion for herzberg death = Significantly more negative opinion	0.025	0.229	0.006	2.448	19

4.1.3. Rules for AV safety improvement as No effect

Table 6 lists top 20 rules in responses in responses to AV safety as 'no effect' consequents. Each outcome is listed from the most important to least important rule based on RPF measurements. The most important rule listed on the table, with an RPF of 0.012, is '*AVs in school zone = Yes*.' From the list of rules, it is found that '*AVs in school zone = Yes*' is present in 6 rules. Some recurring responses on the list include '*AVs should share data = Yes*' (7 rules), '*AVs on limited speed roadways = Yes*' (7 rules), and '*interact with AVs while walking = Yes*' (7 rules). The respondents of this particular group are more familiar with Av technology compared to the first two groups. It is also found that these respondents are more likely to have real-life experience with AVs as non-motorists.

4.1.4. Rules for AV safety improvement as slightly better

Table 7 lists top 20 rules in responses in responses to AV safety as 'slightly better' consequents. Each outcome is listed from the most important to least important rule based on RPF measurements. The RPFs in this table are significantly higher

Table 6
Top rules when 'opinion on AVs reduce injuries and fatalities is no effect'.

No.	Antecedent	Support	Confidence	RPF	Lift	Count
C01	AVs in school zone = Yes	0.078	0.157	0.012	1.218	60
C02	familiarity with AV technology = Somewhat familiar	0.068	0.175	0.012	1.362	52
C03	AVs should share data = Yes	0.092	0.123	0.011	0.959	71
C04	AVs on limited speed roadways = Yes	0.065	0.168	0.011	1.309	50
C05	AVs in school zone = Yes + AVs should share data = Yes	0.065	0.162	0.011	1.263	50
C06	interact with AVs while walking = Yes	0.078	0.129	0.010	1.001	60
C07	paying attention to AV news = To some extent	0.044	0.217	0.010	1.684	34
C08	AVs should share data = Yes + interact with AVs while walking = Yes	0.065	0.142	0.009	1.102	50
C09	AVs on limited speed roadways = Yes + interact with AV while biking = Yes	0.043	0.212	0.009	1.645	33
C10	interact with AV while biking = Yes	0.069	0.130	0.009	1.008	53
C11	AVs on limited speed roadways = Yes + interact with AV while biking = Yes + interact with AVs while walking = Yes	0.036	0.241	0.009	1.877	28
C12	AVs in school zone = Yes + AVs on limited speed roadways = Yes	0.048	0.181	0.009	1.411	37
C13	AVs in school zone = Yes + AVs should share data = Yes + interact with AVs while walking = Yes	0.045	0.191	0.009	1.488	35
C14	AVs in school zone = Yes + interact with AVs while walking = Yes	0.051	0.171	0.009	1.330	39
C15	AVs should share data = Yes + familiarity with AV technology = Somewhat familiar	0.051	0.170	0.009	1.325	39
C16	AVs in school zone = Yes + familiarity with AV technology = Somewhat familiar	0.043	0.201	0.009	1.565	33
C17	AVs on limited speed roadways = Yes + interact with AVs while walking = Yes	0.043	0.196	0.008	1.528	33
C18	AVs should share data = Yes + AVs on limited speed roadways = Yes	0.052	0.156	0.008	1.215	40
C19	familiarity with AV technology = Somewhat familiar + paying attention to AV news = To some extent	0.032	0.248	0.008	1.925	25
C20	AVs should share data = Yes + AVs on limited speed roadways = Yes + interact with AVs while walking = Yes	0.039	0.201	0.008	1.566	30

Table 7
Top rules when 'opinion on AVs reduce injuries and fatalities is slightly better'.

No.	Antecedent	Support	Confidence	RPF	Lift	Count
D01	AVs should share data = Yes	0.270	0.361	0.098	1.045	208
D02	change opinion for herzberg death = No change	0.212	0.349	0.074	1.010	163
D03	interact with AVs while walking = Yes	0.210	0.348	0.073	1.006	162
D04	interact with AV while biking = Yes	0.186	0.350	0.065	1.012	143
D05	AVs in school zone = Yes	0.179	0.360	0.065	1.043	138
D06	familiarity with AV technology = Somewhat familiar	0.156	0.404	0.063	1.170	120
D07	AVs should share data = Yes + interact with AVs while walking = Yes	0.168	0.365	0.061	1.058	129
D08	paying attention to AV news = To a moderate extent	0.143	0.390	0.056	1.129	110
D09	AVs should share data = Yes + change opinion for herzberg death = No change	0.153	0.362	0.055	1.048	118
D10	AVs on limited speed roadways = Yes	0.144	0.374	0.054	1.082	111
D11	AVs in school zone = Yes + AVs should share data = Yes	0.143	0.357	0.051	1.034	110
D12	AVs should share data = Yes + AVs on limited speed roadways = Yes	0.130	0.391	0.051	1.131	100
D13	AVs should share data = Yes + interact with AV while biking = Yes	0.139	0.352	0.049	1.019	107
D14	AVs on limited speed roadways = Not sure	0.114	0.421	0.048	1.219	88
D15	AVs should share data = Yes + familiarity with AV technology = Somewhat familiar	0.119	0.402	0.048	1.163	92
D16	AVs should share data = Yes + paying attention to AV news = To a moderate extent	0.112	0.402	0.045	1.163	86
D17	familiarity with AV technology = Somewhat familiar + interact with AV while biking = Yes	0.095	0.465	0.044	1.346	73
D18	interact with AV while biking = Yes + interact with AVs while walking = Yes	0.130	0.337	0.044	0.975	100
D19	change opinion for herzberg death = No change + interact with AVs while walking = Yes	0.130	0.336	0.044	0.971	100
D20	change opinion for herzberg death = No change + interact with AV while biking = Yes	0.117	0.356	0.042	1.030	90

than the previous tables. The most important rule on the table, with an RPF of 0.098, is ‘AVs in school zone = Yes.’ From the list of rules, it is found that ‘*AVs in school zone = Yes*’ is present in 6 rules. Some recurring responses on the list include ‘AVs should share data = Yes’ (7 rules), ‘*AVs on limited speed roadways = Yes*’ (7 rules), and ‘*interact with AVs while walking = Yes*’ (7 rules). This group is more likely to be knowledgeable about AV technology and they have real-life experience with AVs as non-motorists.

4.1.5. Rules for AV safety improvement as Significantly better

Table 8 lists top 20 rules in responses in responses to AV safety as ‘significantly better’ consequents. Each outcome is listed from the most important to least important rule based on RPF measurements. The RPFs in this table are significantly higher than the previous tables. The most important rule on the table, with an RPF of 0.146, is ‘*change opinion for herzberg death = No change.*’ From the list of rules, it is found that ‘*change opinion for herzberg death = No change*’ is present in 8 rules. Some recurring responses on the list include ‘*AVs on limited speed roadways = No*’ (7 rules) and ‘*paying attention to AV news = To a large extent*’ (5 rules). This group is more likely to be knowledgeable about AV technology and they have real-life experience with AVs as non-motorists.

4.2. Topic modeling results

It is shown in Table 2 that respondents can provide ‘free texts’ for the following five questions:

- Any other suggested laws or regulations?
- Anything else you’d like to share regarding AVs?
- What were your observations with your interaction with AVs?
- Please share any positive experiences that you have had with an AV.
- Please share any negative experiences that you have had with an AV.

As free text is unstructured data, it is important to investigate the insights in these opinions. Topic modeling is a suitable NLP tool to explore the unknowns from unstructured textual contents. Before starting the topic modeling technique, several basic steps of data cleaning were performed. Stop words, redundant words, numbers, and punctuations are removed initially. Besides, this study used additional steps influenced by Zipf’s law, including removing words that occur once. = Additionally, lemmatization was performed, which uses lexicon entries and morphological evaluations to eliminate the inflectional part of a word by converting it into a dictionary based short form of the word. The topic models were developed using ‘tidytext’ package *(*Silge and Robinson, 2020*).* Fig. 2 displays five categories of open-ended responses with eight-word clouds in each category generated from open ended questions. The size of the word depends on the frequency of that word in the topic. In each, top 30 most frequent words are shown. Fig. 2(a) displays words to describe the details of the interactions. The word “stop” or a variation of the word was one of the most prominent words in four of the topics (Topic 1, Topic 2, Topic 4, and Topic 7). The word “bike” or “bicycle” is prominent in four topics as well (Topic 1, Topic 3, Topic 7, and Topic 8), and the

Table 8
Top rules when ‘opinion on AVs reduce injuries and fatalities is significantly better.’

No.	Antecedent	support	confidence	RPF	lift	count
E01	change opinion for herzberg death = No change	0.297	0.490	0.146	1.276	229
E02	AVs on limited speed roadways = No	0.214	0.625	0.134	1.626	165
E03	paying attention to AV news = To a large extent	0.205	0.564	0.116	1.468	158
E04	AVs on limited speed roadways = No + change opinion for herzberg death = No change	0.177	0.651	0.115	1.693	136
E05	interact with AVs while walking = Yes	0.260	0.429	0.111	1.116	200
E06	AVs on limited speed roadways = No + interact with AVs while walking = Yes	0.156	0.710	0.111	1.847	120
E07	change opinion for herzberg death = No change + interact with AVs while walking = Yes	0.205	0.530	0.109	1.379	158
E08	AVs in school zone = No	0.165	0.623	0.103	1.619	127
E09	interact with AV while biking = Yes	0.231	0.435	0.101	1.132	178
E10	change opinion for herzberg death = No change + paying attention to AV news = To a large extent	0.153	0.638	0.098	1.659	118
E11	AVs should share data = Yes + change opinion for herzberg death = No change	0.203	0.479	0.097	1.245	156
E12	AVs should share data = Yes	0.264	0.352	0.093	0.917	203
E13	AVs on limited speed roadways = No + paying attention to AV news = To a large extent	0.121	0.769	0.093	1.999	93
E14	AVs on limited speed roadways = No + change opinion for herzberg death = No change + interact with AVs while walking = Yes	0.130	0.714	0.093	1.858	100
E15	change opinion for herzberg death = No change + interact with AV while biking = Yes	0.173	0.526	0.091	1.368	133
E16	AVs in school zone = No + change opinion for herzberg death = No change	0.132	0.658	0.087	1.712	102
E17	interact with AVs while walking = Yes + paying attention to AV news = To a large extent	0.147	0.585	0.086	1.523	113
E18	AVs on limited speed roadways = No + interact with AV while biking = Yes	0.125	0.676	0.084	1.759	96
E19	AVs in school zone = No + AVs on limited speed roadways = No	0.117	0.698	0.082	1.815	90
E20	interact with AV while biking = Yes + paying attention to AV news = To a large extent	0.132	0.604	0.080	1.570	102

Fig. 2. Word clouds generated from five open ended questions.

words "car(s)" or "vehicle(s)" are also featured in four topics (Topic 3, Topic 4, Topic 6, and Topic 7). Fig. 2(b) illustrates word cloud topics based on positive interactions with AVs. The word "driver" or another variation on the word "drive" is prominent in four of the topics (Topic 1, Topic 4, Topic 7, and Topic 8). There are also many words that seem to describe experiences where AVs have driven safely such as "safer" (Topic 1 and Topic 3), "polite," "cautiously" (Topic 4), "accordance," "expected" (Topic 5), "normal," "normally," and "conservatively" (Topic 6). These words indicated that positive interactions with AVs are characterized by cautious and predictable driving. Fig. 2(c) displays word cloud topics based on negative interactions with AVs. The word "negative" is prominent in two topics (Topic 1 and Topic 5), and the word "slow" is featured in two topics as well (Topic 7 and Topic 8). The word "safe" is shown in one topic (Topic 3); this indicates that some negative interactions with AVs may be caused by them being overly cautious which causes them to drive at slower speeds. On the other hand, words such as "unpredictable" (Topic 2), "idiot," "rush," "slammed," "tragic," and "incident" (Topic 6) indicate potentially unsafe driving which may have led to these negative interactions.

Fig. 2(d) displays word cloud topics based on other AV regulations. The words "public" or "people" are featured in four topics (Topic 1, Topic 2, Topic 3, and Topic 8), while words like "limit" or "limitation," "regulate" or "regulation," "ban," or "government" are featured in six topics (Topic 2, Topic 4, Topic 5, Topic 6, Topic 7, and Topic 8). This shows that AV use is regulated or limited by the government. The words "25," "mph" (Topic 2), and "speed" (Topic 4) are also featured, which could demonstrate how the government restricts the speed of AVs on a stricter basis than other vehicles. Fig. 2(e) is based on other comments. The words "autonomous" or "av(s)" are mentioned in five of the topics (Topic 2, Topic 5, Topic 6, Topic 7, and Topic 8). Some prominent words that were featured from these comments as well as previous word clouds include "stop" (Topic 1), "speed" (Topic 2), and "incident" (Topic 6). Some of the unique words featured from these comments include "Herzberg" (Topic 2), "nope," "hope" (Topic 3), "shit" (Topic 3 and Topic 4), "oh," (Topic 4), and "uber" (Topic 6).

It is found that the topic model outcomes are very generalized. It provided some topic-based trends from the unistructural text contents. There is a need for a stand-alone study which will focus all of the responses by performing inductive and deductive data analysis method. This particular analysis is currently out of the scope of this study. To provide some contexts of the response patterns, some of the key responses are listed below:

- "I usually encounter them on my bicycle and it immediately makes me self-conscious. I realize that my movements and behavior will be recorded by the system and integrated in the machine learning algorithms for bike/ped identification and motion forecasting. Anyway, most of the time they're driving ultra-conservatively and apparently without direction. They make turns which are infrequently used by normal traffic because they're just milling around. You can intimidate them into cautious behavior fairly easily by feigning jumping in front of them (similar to adopting defensive positioning when taking the lane)."
- "After the crash that killed a pedestrian in AZ, it felt like I saw a huge decrease in AVs on Pittsburgh's roadways. I definitely see them and experience them less than I did before that time and feel even less aware of them. With this overall decrease, I also get the sense that drivers are constantly behind the wheel and that more people are driving the cars like they normally would versus using the AV technology."
- "As a cyclist, I've been passed by them. They leave decent space to the left. As a pedestrian, I've encountered them in crosswalks. I'm never sure whether it's safe to step out in front of them, unless the driver motions to me. I know how to tell if a human driver can see me, but how do I know if artificial intelligence (AI) sees me? I wish they had some sort of pedestrian friendly light or signal. I haven't run across this yet, but I would be similarly unsure about making a left turn on my bike in front of a driverless car."
- "Entirely benign. The AV behaved better than 75% of motorists."
- "Hard to tell which vehicles are autonomous, and which are simply used to map the streets to gather data to train the systems that drive the AVs. Both types of vehicles have a camera on them, and the mapping vehicles have a person driving, while the AVs have a person in driver seat that can take over in event of emergency. Would be best if there were some way to differentiate between the two types of vehicles."
- "AV testing should benefit the city, not just the AV companies. AV companies should be strictly limited in how they can use data they collect about Pittsburgh residents. E.g., there should be strict controls about with which other companies' data can be shared, as well as specifications for internal processes for safeguarding and managing the data."
- "AV'S will be a positive, helpful part of the future. In all likelihood they will be safer than direct human driven; but there will be mishaps and even depth. We need prudent regulations that will support and embrace the development of this forward leaning technology in Pittsburgh. It could be a worthy economic element in our future, and a source of civic pride."
- "Accidents happen all the time, but AV takes out human error, distracted driving, aggressive driving, and driving under the influence. As someone who walks to work every day, I experience a lot of unsafe drivers, but have never experienced this with AV. I am in full support of incorporating them in this city."

5. Findings and discussions

The study is based on two major research questions that are described in the introduction section. The study has been carefully designed to answer the key research questions. To summarize this study, the key findings are listed as follows:

- The respondents are needed to choose one from the five options on AV safety (significant worse to significantly better). The findings show that response patterns vary based on participant's knowledge and prior interactions with AVs. Twenty top rules are described in this study to explore the perception of the respondent's on AV safety.
- Respondents who think AV safety is significantly worse are more likely to respond 'change opinion for herzberg death = Significantly more negative opinion,' 'AVs in school zone = Yes,' 'AVs should share data = Yes,' and 'AVs on limited speed roadways = Yes.' It is also found that these respondents do not have real-life interaction with AVs as non-motorist.
- Respondents who think AV safety is slightly worse are more likely to respond 'AVs in school zone = Yes,' 'AVs should share data = Yes,' and 'AVs on limited speed roadways = Yes,' and 'change opinion for herzberg death = Significantly more negative opinion.' It is also found that these respondents do not have real-life interaction with AVs as non-motorist. These respondents are also likely to somewhat familiar with the AV technology.
- Respondents who think AV safety has no effect are more likely to respond 'AVs in school zone = Yes,' 'AVs should share data = Yes,' and 'AVs on limited speed roadways = Yes.' It is also found that these respondents have real-life interaction with AVs as non-motorist. These respondents are also likely to somewhat or to some extent familiar with the AV technology.
- Respondents who think AV safety is significantly worse are more likely to respond 'change opinion for herzberg death = No change,' 'AVs in school zone = Yes,' 'AVs should share data = Yes,' and 'AVs on limited speed roadways = Ye s.' It is also found that these respondents have real-life interaction with AVs as non-motorist. These respondents are also likely to somewhat or to a moderate extent familiar with the AV technology.
- Respondents who think AV safety is significantly better are more likely to respond 'change opinion for herzberg death = No change,' 'AVs in school zone = No,' 'AVs should share data = Yes,' and 'AVs on limited speed roadways = N o.' It is also found that these respondents have real-life interaction with AVs as non-motorist. These respondents are also likely to a large extent familiar with the AV technology.
- Respondents with real-life interactions with AVs as non-motorist and their familiarity with AV technology consider AVs safe compared to the group with no real-life AV experience and non-familiarity of AV technology. Similarly, the first group does not change their opinion due to the recent Uber fatal crash. On the other hand, the second group feels that this event has changed their view on AV safety significantly.
- All respondents are in favor of AV companies share their data.
- The topic models on five free text opinions show some of the key topics such as regulation, traffic law, speed limit, uber crash, safety concerns, and interaction with AVs. The generalized finding from LDA requires additional investigation.

6. Conclusions

The AV industry has been in the process of developing AVs that do not require drivers to steer, either at all or in a set of predetermined situations. Nowadays, nearly all large car manufacturers as well as some operators in various fields of technology are running their own development projects for AVs. Many cities are allowing the testing of AVs on their roads. The general conception is that AV will increase safety as there will be almost no human errors. However, safety concerns regarding AVs still exist. Walking and biking have become popular non-motorized modes of transport due to their health benefits and environment protection. High pedestrian and bicycle exposure are also associated with higher non-motorist related collisions. As both non-motorized mode and AVs are gaining paces in recent years, it is important to understand AV safety measures from the viewpoint of pedestrians and bicyclists.

This study evaluated the opinions of participants about AV safety from the perspective of different measures. The results show that AV safety ratings have been widely varied based on the type of traits in the responses. Participants who favor AVs as safe transport are not worried about the Uber fatal crash. These groups remain informed of the advancement AV industries. The results also show that pedestrians with previous AV interaction consider AVs safer than human drivers and recognize the safety potential for AVs. Participants, who think the safety of AVs is significantly worse, have a strong negative opinion on Uber fatal crash. The topic models generated from the open-ended question responses show different clustered thoughts and opinions based on the nature of the questions. The results of the study provide evidence that experiences and knowledge with AVs are associated with positive perceptions. This finding supports the value of demonstration projects that provide the opportunity for pedestrians and bicyclists to interact with AVs.

The current study has several limitations. First, the survey area is limited to Pittsburgh. Second, the current algorithm only tests two performance measures (lift and rule power factor) to produce insightful knowledge. There is a need to examine additional performance measures, such as leverage, to further examine the performance of the rules. Third, site characteristics of the participants have not been examined. Fourth, majority of the AVs are operational on low-speed limit roadways (25 mph or less). AV movement on high-speed roadways would require further investigation. Fifth, the open-ended question analysis requires additional investigation such as inductive and deductive data analysis. Future studies can mitigate these research gaps by using qualitative analysis on a redesigned and improved survey questionnaire.

Acknowledgments

The author likes to thank two anonymous reviewers for their excellent suggestions.

References

BikePGH. 'Sharing the Road' with Autonomous vehicles: Survey Results 2019. https://www.bikepgh.org/our-work/advocacy/save/av-survey-results-2019/ Accessed: July 31, 2020.

Blei, D., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res., 3*, 993–1022.

Blei, D. (2012). Probabilistic topic models. *CACM, 55*, 77–84.

Canis, B. Issues in Autonomous Vehicle Deployment. No. R44940, 2018.

Das, S., Sun, X., Goel, S., Sun, M., Rahman, A., & Dutta, A. (2020). Flooding Related Traffic Crashes: Findings from Association Rules. *Journal of Transportation Safety & Security*.

Das, S., Dutta, A., Lindheimer, T., Jalayer, M., & Elgart, Z. (2019a). YouTube as a Source of Information in Understanding Autonomous Vehicle Consumers: Natural Language Processing Study. *Transportation Research Record, 0361198119842110*.

Das, S., Dutta, A., & Fitzpatrick, K. (2020). Technological Perception on Autonomous Vehicles: Perspectives of the Non-Motorists. *Technology Analysis & Strategic Management*.

Das, S., Kong, X., & Tsapakis, I. (2019b). Hit and Run Crash Analysis Using Association Rules Mining. *Journal of Transportation Safety & Security*.

Das, S., Dutta, A., & Sun, X. (2019c). Patterns of Rainy Weather Crashes: Applying Rules Mining. *Journal of Transportation Safety & Security*.

Das, S., Dutta, A., Avelar, R., Dixon, K., Sun, X., & Jalayer, M. (2019). Supervised Association Rules Mining on Pedestrian Crashes in Urban Areas: Identifying Patterns for Appropriate Countermeasures. *International Journal of Urban Sciences, 23*, 30–48.

Das, S., Dutta, A., Jalayer, M., Bibeka, A., & Wu, L. (2018). Factors Influencing the Patterns of Wrong-Way Driving Crashes on Freeway Exit Ramps and Median Crossovers: Exploration Using 'Eclat' Association Rules to Promote Safety. *International journal of transportation science and technology, 7*, 114–123.

Deb, S., Strawderman, L., Carruth, D. W., DuBien, J., Smith, B., & Garrison, T. M. (2017). Development and validation of a questionnaire to assess pedestrian receptivity toward fully autonomous vehicles. *Transportation Research Part C: Emerging Technologies, 84*, 178–195.

Deb, S., Strawderman, L. J., & Carruth, D. W. (2018). Investigating pedestrian suggestions for external features on fully autonomous vehicles: A virtual reality experiment. *Transportation Research Part F: Traffic Psychology and Behaviour, 59*, 135–149.

Feuerriegel, S., & Ratku, D. N. (2016). In *Analysis of how underlying topics in financial news affect stock prices using latent Dirichlet allocation* (pp. 1072–1081). IEEE Computer Society.

Jayaraman, S. K., Creech, C., Robert, L. P., Jr., Tilbury, D. M., Yang, X. J., Pradhan, A. K., & Tsui, K. M. (2018). Trust in AV: An Uncertainty Reduction Model of AV-Pedestrian Interactions. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, HRI '18* (pp. 133–134). New York, NY, USA: Association for Computing Machinery.

Howard, D., & Dai, D. (2014). Public Perceptions of Self Driving Cars: The case of Berkeley. *California. Transportation Research Board 93rd Annual Meeting, 14* (4502), 1–16.

Hulse, L. M., Xie, H., & Galea, E. R. (2018). Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age. *Safety Science, 102*, 1–13.

Kong, X., Das, S., Jha, K., & Zhang, Y. (2020). Understanding Speeding Behavior from Naturalistic Driving Data: Applying Classification Based Association Rule Mining. *Accident Analysis & Prevention*.

Kyriakidis, M., Happee, R., & de Winter, J. C. F. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation Research Part F: Traffic Psychology and Behaviour, 32*, 127–140.

Montella, A. (2011). Identifying crash contributory factors at urban roundabouts and using association rules to explore their relationships to different crash types. *Accident Analysis and Prevention, 43*(4), 1451–1463.

Montella, A., Aria, M., D'Ambrosio, A., & Mauriello, F. (2012). Analysis of powered two-wheeler crashes in Italy by classification trees and rules discovery. *Accident Analysis and Prevention, 49*, 58–72.

Montella, A., de Oña, R., Mauriello, F., Riccardi, M. R., & Silvestro, G. (2020). A data mining approach to investigate patterns of powered two-wheeler crashes in Spain. *Accident Analysis and Prevention, 134* 105251.

National Highway Traffic Safety Administration (NHTSA). Traffic Safety Facts: Bicyclists and Other Cyclists. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812884 Accessed: August 1, 2020.

National Highway Traffic Safety Administration (NHTSA). Traffic Safety Facts: Pedestrians. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812850 Accessed: August 1, 2020.

Ochin, S., & Joshi, N. (2016). Rule Power Factor: A New Interest Measure in Association Classification. *Procedia Computer Science, 93*, 12–18.

Penmetsa, P., Adanua, E., Wood, D., Wang, T., & Jones, S. (2019). Perceptions and expectations of autonomous vehicles – A snapshot of vulnerable road user opinion. *Technological Forecasting & Social Change., 143*, 9–13.

Pillai, A.K., 2017. Virtual Reality based Study to Analyse Pedestrian attitude towards Autonomous Vehicles. Master's Thesis. School of Computer Science and Communication, Department of Media Technology and Interaction Design, Stockholm, Sweden.

Rasouli, A., & Tsotsos, J. K. (2020). Autonomous Vehicles That Interact With Pedestrians: A Survey of Theory and Practice. *IEEE Transactions on Intelligent Transportation Systems, 21*, 900–918.

Salvador, I. R Package 'compareGroups'. https://cran.r-project.org/web/ packages/compareGroups/compareGroups.pdf Accessed August 1, 2020.

Silge, J., and D. Robinson. Text Mining with R. O'Reilly, 2020. https://www.tidytextmining.com/ Accessed: August 1, 2020.