Collision Warning System Using Naïve Bayes Classifier

Ahmed Tijani¹
The University of Toledo College of Engineering
ahmed.tijani@rockets.utoledo.edu

Richard Molyet
The University of Toledo College of Engineering
richard.molyet@utoledo.edu

Mansoor Alam
Northern Illinois University College of Engineering
malam1@niu.edu

Abstract. Motor vehicle crashes can lead to traumatic experiences. High-impact collisions usually cause severe injuries or fatalities. A collision warning system that analyzes driving behaviors and warns drivers of impending crashes can prevent road collisions and save lives, so increasing traffic safety. An application of the Naïve Bayes classifier model to determine the potential for rear-end collisions between highway vehicles is presented. The Naïve Bayes classifier is a supervised machine-learning model based on Bayes’s theorem. Two vehicles are utilized, with one vehicle following the other. The parameters studied are speed, distance, acceleration, and deceleration. Training examples involving over 100 potential collision scenarios have been evaluated. Simulation results show that the model successfully responds to and correctly predicts potential collisions.

Keywords. Bayes’s theorem, Naïve Bayes, machine learning, rear-end collision, collision warning system, driver assistance systems.
1. Introduction

According to a recent report, more than 1.3 million people worldwide die in car accidents every year. In 2019, over 38,000 fatal car accidents occurred in the United States. Approximately 38% of all fatal car accidents result from a collision with another vehicle. The most common collisions are rear-end or side collisions. Every day, more than nine people are killed and 1,000 injured due to distracted driving [1]. These collisions could be avoided if drivers were given warnings in advance. Because of human limitations, drivers often cannot correctly predict dangerous situations. Several factors contribute to road accidents. Speeding, acceleration, deceleration, and distance are related to most fatal car crashes [2,3,4]. According to the National Highway Traffic Safety Administration (NHTSA), one of the most common reasons for road accidents is speeding. Driving above the speed limit increases the risk of losing control of a vehicle, which can cause an accident. Around one-third of all fatal car accidents occurred as a result of speeding [5]. Also, acceleration in rear-end collisions is one of the major causes of serious injuries. NHTSA found that the risk of permanent injuries in rear-end collisions increases when vehicles were accelerating [4]. Injuries such as whiplash are related to acceleration and deceleration in car crashes [6]. A short distance or tailgating is the fifth leading cause of vehicle collisions according to the US Department of Public Safety [7]. Keeping a safe distance between vehicles gives drivers enough time to stop safely. A driver needs 0.7s–1.5s to react to an emergency situation [8]. Incorporating intelligent transportation systems (ITS) into the transportation infrastructure and into vehicles can improve and increase transportation safety by reducing the number of accidents each year [9].

One of the current types of ITS is Automatic Emergency Braking or Advanced Emergency Braking (AEB), which determines a possible forward collision with another vehicle in time to avoid a crash. The system at first notifies the driver of an impending collision. If the driver does not react to the situation, the AEB system applies the brakes automatically to prevent the collision [10]. This system is designed to prevent and reduce the frequency of rear-end collisions, and relies on several sensors and devices, such as radar, LiDAR, cameras, and GPS [10,11]. These devices detect the movement of vehicles and help drivers react to dangerous situations in advance. The European Parliament has adopted a proposal to make the AEB system mandatory in commercial vehicles manufactured after October 2013, and for all new vehicles manufactured after October 2015 [12]. According to NHTSA, ten automakers in the US equipped more than half of their vehicles designed between September 2017 and August 2018 with AEB [13]. This kind of collision warning system assists drivers in assessing whether a vehicle is in a dangerous situation based on the data collected.

A state-of-the-art collision warning system notifies the driver of a potential incoming hazard situation if a distraction occurs momentarily. It uses visual and audible warnings that alert drivers to a potential crash with the detected front vehicle. If the distance between the two vehicles is decreasing, the Forward Collision Warning (FCW) system determines a possibility of a collision based on the vehicles’ speeds. The visual and audible warnings are issued to alert the driver to apply the brakes. To prevent inaccurate results, the system is only activated when the speed of the vehicle is over 3mph [14].

This type of collision system is almost exclusive to modern and luxury vehicles. To make such systems more accessible, a collision warning system using the Naïve Bayes (NB) classifier offers a simple and economical alternative for lower-priced and older vehicles.
This study applied the Naïve Bayes theorem from machine learning to a collision warning system. The Naïve Bayes classifier is a practical learning method, one of the commonly used supervised machine learning algorithms based on Bayes’s theorem with strong independence assumptions between the features. The performance of the method shows effective results in multi-class predictions [15].

One advantage of implementing this method is that the Naïve Bayes is highly accurate when applied to large data. It is fast, as compared to other classification models such as the k-nearest neighbors (k-NN), which does not require iterations since the probabilities can be directly computed. Also, it can make real-time predictions, making the conditional probabilities easy to evaluate [16].

Another advantage is simplicity. Some forward collision systems rely on mathematical equation models to calculate potential collisions. The parameter modeling of frontal crash, for instance, needs to be investigated and analyzed in several degrees, then used to analyze the response of the driver during the impact. An optimization procedure is required to assist multibody vehicle model development along with validation and evaluation [17]. The Naïve Bayes classifier is simple to implement and could provide valuable results when the conditional independence assumption holds.

This study discusses a rear-end collision between highway vehicles using the NB classifier. The main challenge of implementing this method involves different driving conditions that rely on dependent variables, which could affect the conditional independence assumption as well as the performance of the system. An improved Naïve Bayes algorithm that combines machine-learning algorithms, such as a random forest classifier, could expand the model’s capability.

The main purpose of this paper is to present a collision-warning method that is simple, economical, and an available alternative for unequipped vehicles. It also provides information about impending crashes that can prevent road collisions and reduce driver errors.

The method utilizes a machine learning technique that compares hypothetical driving conditions to the data from the training set examples. Therefore, explicit high-level mathematical models and analysis of the problems are not required. Using this approach, it is possible to equip more vehicles with the collision system, which can contribute to reducing road accidents.

The main contribution of this paper is to utilize a machine-learning algorithm to assist drivers in potentially dangerous situations, and to create a collision-warning model for vehicles unequipped with the system to help increase road safety.
2. Related Work

In one study, a collision warning system model was used to analyze a driver’s braking response time and to study the influence of expectancy and automation complacency on real-life emergency braking. The research demonstrated the significance of the Brake Reaction Time (BRT) parameter and its contribution to improving road safety. Some other studies have proposed that reaction time (RT) can be affected by the driver’s expectations when facing dangerous situations. However, using advanced driver assistance systems can change these considerations. A collision warning system could offer quicker and more efficient responses, but it also requires an evaluation process and monitoring task that may cause automation complacency. The objectives of this study are to examine two aspects: the ability of automation compliance to generate a collision warning system in a real-life scenario, and the type of component expectancy that can affect the different tasks involved in an assisted BRT process. Presence/absence of anticipatory information, previous direct experience, reliability of the device, and predictability of the hazard determined by repeated use of the warning system are the four components of expectancy that were evaluated. The results of the study provided an indication of perception time and the mental elaboration process of the collision warning system alerts. Specifically, reliable warning made the decision-making process faster; misleading warning led to autonomous complacency, reducing the visual search for hazard detection; little direct experience reduced the total response; and unpredicted failure of the device led to inattentional blindness and potential pseudo-accidents with surprise obstacle intrusion [18].

Addressing the significant issue of front-to-rear collisions, another study discusses the effectiveness of forward collision warning (FCW) and autonomous emergency braking systems in decreasing the risk of front-to-rear collision rates. The study aim is to analyze the effectiveness of FCW alone, the operation of a low-speed autonomous emergency braking (AEB) system at speeds up to 19mph that does not alert the driver before the braking, and the contributions of FCW that uses AEB to operate at higher speed in reducing front-to-rear collisions and injuries. Poisson regression was used to compare rates of police-reported crash involvements per insured vehicle year in 22 US states during the four years 2010–14 between passenger vehicle models with FCW alone or with AEB, and the same models where the optional systems were not acquired, controlling for other factors affecting crash risk. Related evaluations compared rates between several Volvo vehicle models that use standard low-speed AEB systems and other luxury vehicle models not equipped with AEB systems. The results of the study show that the reduction rate for rear-end striking crash involvements was 27% for FCW alone, 43% for low-speed AEB, and 50% for FCW with AEB. The rates that involved injuries were slightly different, at 20%, 45%, and 56% respectively. The rates were reduced after involving third-party injuries by 18%, 44%, and 59% respectively. In the rear-end striking crashes with third-party injuries, the reductions were marginally significant for FCW alone. Other reductions were statistically significant. FCW alone reduced rates of rear-end striking crash involvements by 13% and low-speed AEB by 12%. However, FCW with AEB increased rates of rear-end striking crash involvements by 20%. Based on these evaluations, rear-end crashes could have been prevented if all vehicles were equipped with FCW and AEB [19].

A related study examines adaptive forward collision warnings and their impact on behavioral adaptation. Adaptive Advanced Driver Assistance Systems (ADAS) that change warnings based on the driver’s demand for assistance provide great potential to increase safety. However, it is difficult to obtain the behavior of drivers when they deal with dynamically adapting technologies, especially in scenarios when the system shows unexpected behavior due to failure of driver-state
monitoring. A driving simulator research with N=48 participants was conducted to illustrate the consequences of unreliable adaptive ADAS on safety and to determine failures of an adaptive FCW impact driving behavior. Participants were involved in critical brake events in scenarios that may or may not include a distracting secondary task. An adaptive FCW offered visual warnings to undistracted drivers and highly supportive visuo-haptic warnings (brake jerks or vibration) to distracted drivers. The system unexpectedly provided inaccurate adapted warnings that present the combination of warning type and distraction was reversed in 20% of the brake scenarios. The adaptive FCW and a non-adaptive standard FCW that only offered visual warnings were compared. The results of this study indicated that there are two resources for incorrect warnings that impaired driver reactions and safety in distracted drivers and that these resources had adverse behavioral effects. The first consists of violations of the drivers’ expectations about the warning and behavioral adaptation. The second resource is the absence of the compensatory effect of the highly supportive warning in case of distraction. On the other hand, correctly adapted warnings decreased decrements in brake reaction times and fully offset safety deficits associated with driver distraction. An effectiveness examination of the adaptive system’s potential to assist drivers when correct warnings were elicited failed to illustrate an advantage for the adaptive FCW over the non-adaptive FCW. The study findings emphasize that high reliability is crucial for adaptive ADAS to improve safety and to avoid the adverse effects of behavioral adaptation [20].
3. Background

3.1. Bayes Theorem

Bayesian learning methods are among the most practical approaches to some types of learning problems. In machine learning, the objective is to define the best hypothesis that is the most probable hypothesis from the space \( H \) giving the observed training data \( D \). The approach of the Bayes theorem is to provide a way to compute probabilities. It computes the probability of a hypothesis based on its prior probability. Bayes theorem is an essential part of Bayesian learning methods because it provides a way to compute the posterior probability from the prior probability.

Bayes Theorem: 
\[
P(h|D) = \frac{P(D|h)P(h)}{P(D)}
\]  
(1)

Where, \( P(h) \) is the initial probability that hypothesis \( h \) holds before observing the training data. \( P(D) \) is the prior probability of the observing training data \( D \). \( P(D|h) \) is the probability of observing data \( D \) in which the hypothesis \( h \) holds. \( P(h/D) \) is posterior probability of \( h \). \( P(h/D) \) reflects the training data \( D \) in contrast to the prior probability \( P(h) \) that is independent of \( D \). According to the Bayes theorem, \( P(h/D) \) increases with \( P(h) \) and \( P(D|h) \). Also, \( P(h/D) \) decreases as \( P(D) \) increases because the \( D \) will be observed independent of \( h \). To determine the MAP hypostasis, Bayes theorem can compute the posterior probability of every candidate hypothesis \([15]\).

Therefore, the \( h_{\text{MAP}} \) provides 
\[
h_{\text{MAP}} = \arg \max P(h|D)
\]

\[
= \arg \max_{h \in H} \frac{P(D|h)P(h)}{P(D)}
\]  
(2)

3.2. Naïve Bayes Classifier

The Naïve Bayes classifier, based on Bayes’s theorem, is one of the commonly used supervised machine learning algorithms. The performance of the Naïve Bayes classifier is comparable to some types of supervised machine learning, such as the decision tree and neural networks. The Naïve Bayes classifier is applicable to learning tasks in which each conjunction of parameter value \( x \) is defined by a set of attribute values where the target function \( f(x) \) can assume any value from some finite set \( V \). To classify the next instance, the Bayesian approach is applied to specify the most probable target value \( (V_{\text{MAP}}) \), giving the attribute values \(<a_1, a_2, a_3, ... a_n>\) that show:

\[
V_{\text{MAP}} = \arg \max_{V_j \in V} P(v_j|a_1, a_2, ... a_n)
\]

Using Bayes’s theorem, the above equation becomes:

\[
V_{\text{MAP}} = \arg \max_{v_j \in V} P(v_j|a_1, a_2, ... a_n)P(v_j)
\]  
(3)
The Naïve Bayes classifier is based on the hypothesis that the attribute values are conditionally independent given the target value, which is
\[ P(a_1, a_2 \ldots a_n | v_j) = \prod_i P(a_i | v_j) \]
Substituting the equation into (3), we get the Naïve Bayes approach. Therefore, the Naïve Bayes classifier is:
\[ V_{NB} = \arg \max_{v_i \in v_i} P(v_j) \prod P(a_i | v_j) \]
(4)
where \((V_{NB})\) is the target value output by the classifier. The learning steps of the Naïve Bayes method include the estimation of different \(P(v_j)\) and \(P(a_i | v_j)\) terms based on their frequencies over the training data. The estimations are related to the learned hypothesis, which is applied to classify every new instance by using equation 4 [15].

3.3. Validation

It is important to evaluate the model prediction performance. A confusion matrix was used to measure the performance of the classification model. The classification accuracy, sensitivity, and precision are shown in equations 5,6,7 [21].

\[
\text{Accuracy} = \frac{\text{TruePositive} + \text{TrueNegative}}{\text{TruePositive} + \text{TrueNegative} + \text{FalsePositive} + \text{FalseNegative}}
\]
(5)
\[
\text{Sensitivity} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalseNegative}}
\]
(6)
\[
\text{Precision} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalsePositive}}
\]
(7)
4. Methods and Procedures

This study investigated a method of determining the potential for a rear-end collision between highway vehicles. A variety of training example scenarios have been investigated. The variables of the training data were generated based on the information provided by NHTSA and the Federal Motor Carrier Safety Administration. Table 1 shows an example of the training data. Training examples were generated to be performed under ideal weather conditions between two vehicles, \( V_1 \) & \( V_2 \). The study investigated the hypothesis of potential collisions between the two vehicles.

The first vehicle (\( V_1 \)) was pursued by the second vehicle (\( V_2 \)). The probabilities of potential collisions were calculated by applying the Naïve Bayes classifier. The system is designed to assist drivers operating the pursuing vehicle (\( V_2 \)). An audible alarm would activate, and a text message would display on the dashboard of \( V_2 \) if there were potential for a collision. The training examples consisted of conjunctions of three variable parameters: speed, distance, and acceleration–deceleration. According to NHTSA, the standard speed limit on US highways is between 55 and 75 MPH [22]. Therefore, the speed was considered high if it was between 76 and 85 MPH, and low if it was between 35 and 54 MPH. The three-second rule is the rule of thumb that allows drivers to keep a safe trailing distance between vehicles at any speed. According to Federal Motor Carrier Safety Administration, the recommended safe distance between two vehicles under ideal weather conditions is three seconds [3]. Hence, the distance between \( V_1 \) and \( V_2 \) was “safe” when it was three seconds; “far”, when it was longer than three seconds; and “close”, when it was shorter than three seconds.

The acceleration and deceleration of both vehicles were also determined. “True” indicated that the vehicle was accelerating, and “false” indicated that it was decelerating or not accelerating. The collision system calculated whether there was a potential for a collision between the two vehicles on a highway. The calculations were based on the data that the system received, and the system determined whether there was potential for a collision by applying the Naïve Bayes theorem. The collision-warning system is designed to assist drivers in avoiding potentially dangerous driving situations. The results showed that the system successfully predicted and responded to different driving scenarios. Also, the collision-warning system has an 86% accuracy of correctly predicting collisions. Table 1 contains a sample of the training examples.
Table 1. A partial training set for the Naïve Bayes classifier

<table>
<thead>
<tr>
<th>#</th>
<th>Direction</th>
<th>Speed V1</th>
<th>Speed V2</th>
<th>Distance</th>
<th>Acceleration V1</th>
<th>Acceleration V2</th>
<th>PC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>One Way</td>
<td>High Speed</td>
<td>Speed Limit</td>
<td>Safe</td>
<td>T</td>
<td>T</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>One Way</td>
<td>High Speed</td>
<td>Speed Limit</td>
<td>Close</td>
<td>T</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>One Way</td>
<td>High Speed</td>
<td>Speed Limit</td>
<td>Far</td>
<td>T</td>
<td>T</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>One Way</td>
<td>High Speed</td>
<td>High Speed</td>
<td>Safe</td>
<td>T</td>
<td>T</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>One Way</td>
<td>High Speed</td>
<td>Speed Limit</td>
<td>Close</td>
<td>T</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>One Way</td>
<td>High Speed</td>
<td>High Speed</td>
<td>Far</td>
<td>T</td>
<td>T</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>One Way</td>
<td>Speed Limit</td>
<td>Speed Limit</td>
<td>Safe</td>
<td>F</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>8</td>
<td>One Way</td>
<td>Speed Limit</td>
<td>Speed Limit</td>
<td>Close</td>
<td>T</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>9</td>
<td>One Way</td>
<td>Speed Limit</td>
<td>Speed Limit</td>
<td>Far</td>
<td>T</td>
<td>T</td>
<td>No</td>
</tr>
<tr>
<td>10</td>
<td>One Way</td>
<td>Speed Limit</td>
<td>High Speed</td>
<td>Safe</td>
<td>T</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>11</td>
<td>One Way</td>
<td>Speed Limit</td>
<td>High Speed</td>
<td>Close</td>
<td>F</td>
<td>F</td>
<td>Yes</td>
</tr>
<tr>
<td>12</td>
<td>One Way</td>
<td>Speed Limit</td>
<td>High Speed</td>
<td>Far</td>
<td>T</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>13</td>
<td>One Way</td>
<td>Speed Limit</td>
<td>Low Speed</td>
<td>Safe</td>
<td>F</td>
<td>F</td>
<td>No</td>
</tr>
<tr>
<td>14</td>
<td>One Way</td>
<td>Speed Limit</td>
<td>Low Speed</td>
<td>Far</td>
<td>T</td>
<td>T</td>
<td>No</td>
</tr>
<tr>
<td>15</td>
<td>One Way</td>
<td>Low Speed</td>
<td>Speed Limit</td>
<td>Safe</td>
<td>F</td>
<td>F</td>
<td>Yes</td>
</tr>
<tr>
<td>16</td>
<td>One Way</td>
<td>Low Speed</td>
<td>Speed Limit</td>
<td>Close</td>
<td>T</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>17</td>
<td>One Way</td>
<td>Low Speed</td>
<td>Speed Limit</td>
<td>Far</td>
<td>F</td>
<td>F</td>
<td>Yes</td>
</tr>
<tr>
<td>18</td>
<td>One Way</td>
<td>Low Speed</td>
<td>High Speed</td>
<td>Safe</td>
<td>T</td>
<td>T</td>
<td>Yes</td>
</tr>
<tr>
<td>19</td>
<td>One Way</td>
<td>Low Speed</td>
<td>High Speed</td>
<td>Close</td>
<td>F</td>
<td>F</td>
<td>Yes</td>
</tr>
<tr>
<td>20</td>
<td>One Way</td>
<td>Low Speed</td>
<td>High Speed</td>
<td>Far</td>
<td>F</td>
<td>F</td>
<td>Yes</td>
</tr>
<tr>
<td>21</td>
<td>One Way</td>
<td>Low Speed</td>
<td>Low Speed</td>
<td>Safe</td>
<td>T</td>
<td>T</td>
<td>No</td>
</tr>
<tr>
<td>22</td>
<td>One Way</td>
<td>Low Speed</td>
<td>Low Speed</td>
<td>Close</td>
<td>F</td>
<td>F</td>
<td>Yes</td>
</tr>
<tr>
<td>23</td>
<td>One Way</td>
<td>Low Speed</td>
<td>Low Speed</td>
<td>Far</td>
<td>F</td>
<td>F</td>
<td>No</td>
</tr>
</tbody>
</table>
A confusion matrix is used to validate the performance of the classification model. It enables the visualization of the performance of an algorithm. Fig. 1 shows the model classification. The confusion matrix shows that in group 2 of the “True” class there are 13 data points misclassified into group 1 of the predicted class. Group 2 contains 48 data points that are correctly classified, while 35 data points are correctly classified in group 1.

![Fig. 1: Model Classification](image)

Fig. 2 shows the performance of the classification model using confusion matrix terminology. The accuracy of the collision-warning system is 86%, the error rate is 0.13, the false positive rate is 0.27, the precision is 0.79, sensitivity is 0.89, prevalence is 0.5, and true negative rate is 0.73.
Table 2 illustrates the three scenarios that were examined for $V_1$ and $V_2$. These three scenarios are depicted in Fig. 3, Fig. 6, and Fig. 9, respectively.

<table>
<thead>
<tr>
<th>Table 2. Three Possible Training Example Scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scenarios</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td><strong>Scenario 1</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Scenario 2</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Scenario 3</strong></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
4.1. Scenario 1

In Fig. 3, $V_1$ is moving at a high speed of 90 MPH with acceleration, and $V_2$ is moving at a low speed of 40 MPH with no acceleration. The distance between the two vehicles is 8 seconds or 469 ft, which is "far". The speed introduced in this scenario is considered as new data to the system.

![Fig. 3: A safe scenario for $V_1$ and $V_2$](image)

Fig. 3: A safe scenario for $V_1$ and $V_2$

Fig. 4 shows the results of applying the Naïve Bayes theorem. After entering the values of speed, distance, and acceleration for scenario 1, the probability of a potential collision is $P(\text{Yes}) = 0.0004692$, and the probability of no potential collision is $P(\text{No}) = 0.01959$. Therefore, $P(\text{No})$ has a higher probability than $P(\text{Yes})$, and there will be no collision warning. The collision-warning system provided the correct prediction even though new data were given to the system outside the training data. The collision system will send a message to the driver dashboard stating, Safe, there is no potential for collision, because the distance between the two vehicles meets the road safety requirement, and $V_2$ is not accelerating and is moving at a lower speed than the $V_1$.

![The Probabilities of a Potential Collision](image)

Fig. 4: The probability for scenario 1
Fig. 5: Conditional Probabilities

Fig. 5 shows the conditional probabilities of speed, distance, and acceleration for the two vehicles: $P(\text{High Speed}|\text{Yes})$, $P(\text{Low Speed}|\text{Yes})$, $P(\text{Acceleration}|\text{Yes})$, and $P(\text{Distance}|\text{Yes})$. For example, given that $P(\text{Yes}) = 0.0004692$ and $P(\text{No}) = 0.01959$, the conditional probability of the speed for $V_1$ is $P(\text{High Speed}|\text{Yes}) = 0.2982$, and the conditional probability of the speed for $V_2$ is $P(\text{Low Speed}|\text{Yes}) = 0.070$. Other conditional probabilities follow in the same manner.

4.2. Scenario 2

Fig. 6: An unsafe scenario for $V_1$ and $V_2$

In Fig. 6, $V_1$ is not accelerating and is moving within the speed limit of 70 MPH. However, $V_2$ is accelerating and moving at a high speed of 77 MPH. The distance between the two vehicles is “close” – 1 second or 113ft.
Fig. 7 shows there is a potential collision. The value of $P(\text{Yes}) = 0.0137$, and the value of $P(\text{No}) = 0.000730$. An alarm will activate, and a message will be sent to the dashboard stating, *Caution, there is a potential for collision*, because the distance between the two vehicles is close, and $V_2$ is accelerating and moving at a higher speed than $V_1$.

Fig. 8: Conditional probabilities
Fig. 8 shows the conditional probabilities of speed, distance, and acceleration for the two vehicles: 
P (High Speed| Yes), P (Speed limit| Yes), and P (Distance| Yes), etc. For example, given the 
probability P (Yes)= 0.0137 and P (No) = 0.000729, the conditional probability of the speed for 
V₁ is P (Speed Limit| Yes) = 0.2982, and the conditional probability of the speed for V₂ is P (High 
Speed| Yes) = 0.5263. Other conditional probabilities follow in the same manner.

4.3. Scenario 3

![Diagram of two cars: V₂ and V₁](image)

Fig. 9: A safe scenario for V₁ and V₂

In Fig. 9, V₁ & V₂ are not accelerating and are moving within the speed limit of 65 MPH and 60 
MPH, respectively. The distance between the two vehicles is “safe” at 3 seconds or 264 ft.

![Bar chart showing probabilities](image)

Fig. 10: The probability for scenario 3

Fig. 10 shows there is not a potential collision. The value of P (Yes) = 0.003463, and the value of 
P (No) = 0.005268.
Fig. 11 shows the conditional probabilities of speed, distance, and acceleration for the two vehicles. P (Speed Limit| Yes), P (Accelerations| Yes), and P (Distance| Yes), etc. For example, P (Yes) = 0.003463 and P (No) = 0.005268, the conditional probability of the speed for V₁ and V₂ is P (Speed Limit| Yes) = 0.2982, the conditional probability of the speed for V₁ is P (Speed Limit| No)= 0.385, and the conditional probability of the speed for V₂ is P (Speed Limit| No)= 0.333. Other conditional probabilities follow in the same manner.

**Conclusion**

This paper investigated a rear-end collision warning system for highway vehicles using the Naïve Bayes classifier technique from supervised machine learning. The collision-warning system is designed to assist drivers to avoid potentially dangerous driving situations. The parameters used in this study were speed, acceleration–deceleration, and the distance between two vehicles. The results showed that the system successfully predicted and responded correctly to different driving scenarios with 86% accuracy. Future work in this area can improve the capabilities of collision systems so that they can respond to more complex scenarios, such as variable weather conditions, changing lanes, passing vehicles in two-way traffic, and vehicles equipped with an automatic braking system.
References


