Proceedings of CONF-MLA 2025 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2025.LD25515

Logistic Regression Model-based Analysis of Self-driving Car Disengagement Influencing Factors

Yuhan Gu

College of Mathematics and Finance, Hunan University of Humanities, Science and Technology, Loudi, China
2510630411@qq.com

Abstract. Accelerated by breakthroughs in sensor fusion and edge computing, autonomous vehicle (AV) technologies now face stringent regulatory requirements for disengagement event logging. As a leading global center for AV testing, California's Department of Motor Vehicles (DMV) data highlights that disengagement incidents remain a critical measure of technological maturity. Utilizing the 2019 Autonomous Vehicle Disengagement Reports, this study constructs a binary logistic regression model to systematically examine the key determinants of disengagement events. This study highlights the key contributing elements through meticulous feature selection and model interpretation, providing practical solutions to mitigate disengagement risks. Python analysis identifies three primary factors: sudden halts caused by unintended manoeuvres, inconsistencies in perception between vehicle sensors and human assessment, and scenarios requiring manual intervention to avoid stationary vehicles in the lane.

Keywords: Logistic regression, Automatic driving, Machine learning, Binary classification

1. Introduction

Fundamental breakthroughs in autonomous driving(AD) emerged during the 1980s through advances in computational architectures. Notably, in 1989, Carnegie Mellon University (CMU) was among the first to successfully utilize neural networks for the control of intelligent vehicles (IVs), establishing a foundational framework for the development of intelligent control systems. Moving into the early 21st century, international competitions played a vital role in advancing AD research. From 2004 onwards, the Defense Advanced Research Projects Agency (DARPA) organized a series of high-profile challenges, testing the capabilities of IVs under demanding and unpredictable environments. The 2010s witnessed operational domain expansion from controlled environments to complex urban roads, driven by algorithmic advances and compute scaling. In 2016, Drive.ai received permission to conduct IV trials in California, while nuTonomy launched a fleet of autonomous taxis in Singapore during the same period [1]. Disengagement events-defined as human interventions or autonomous mode deactivations-constitute key safety and maturity indicators for self-driving systems. This research draws on the 2019 California Autonomous Driving Vehicle Disengagement Report as its primary data source, employing a binary logistic regression approach to systematically explore the correlations between a range of influencing variables and the

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.LD25515

likelihood of disengagement occurrences. The methodology involved several steps: first, the necessary datasets were acquired from Kaggle, after which the raw data underwent thorough preprocessing. A binary logistic regression model was then developed and utilised on the dataset. This method facilitates accurate detection and measurement of the impacts of various influencing factors on disengagement events, thereby providing essential information for technology developers seeking to improve the performance and reliability of autonomous driving algorithms.

2. Literature review

The triennial surge in disengagement research (2023-2025) has yielded multifactorial analytical frameworks. For example, Wu et al. conducted a comprehensive analysis of disengagement records collected by California DMVs, leveraging a suite of statistical models to evaluate how different factors influence disengagement behavior. Their findings highlighted that the leading cause of disengagement was the inconsistent performance of autonomous vehicles during real-world road tests [2]. In a distinct 2023 study, Chengula et al. employed Latent Dirichlet Allocation (LDA) alongside Natural Language Processing (NLP) to identify prevailing themes in disengagement data, indicating that common triggers comprised erratic road user behaviour, inaccurate system assessments of traffic conditions, and sensor-related complications [3]. In 2024, Tu Huizhao and colleagues applied the Tobit model to measure the avoidance-related disengagement rates in autonomous vehicles [4]. Kohanpour et al. utilized a mix of statistical methods, including Pearson and Spearman correlation coefficients, regression analysis, and the chi-square test, ultimately demonstrating that disengagement frequency is not merely influenced by environmental factors but also by interactions with other traffic participants [5]. Additionally, a study by Skokan and Mareček in 2025 examined the progression of disengagement occurrences alongside developments in automated driving technology utilising linear regression and the Shapiro-Wilk test. Their findings revealed that although the total instances of disengagement have increased, the annual frequency of these events has been decreasing as the overall test mileage increases, indicating continuous enhancements in system reliability [6]. Regarding data handling, earlier research has often prioritized modeling over the detailed processing of raw data. In contrast, this paper emphasizes meticulous data preparation: the original datasets span both 2018-2019 and 2019, with the latter featuring greater completeness, minimal missing values, and almost no outliers, resulting in a more standardized and reliable data foundation. When constructing the analytical model, this study selects a binary logistic regression framework, which demonstrates a strong compatibility with the characteristics of the available data.

3. Data

3.1. Data sources and descriptions

This study employs the publicly available 2019 Autonomous Vehicle Disengagement Reports dataset from Kaggle [7]. These files document both the instances of disengagement and the corresponding autonomous miles driven by covering all California public road tests conducted by permitted entities during the 2018-2019 reporting year (December 1 to November 30). The dataset is partitioned into separate files, with dedicated records for companies receiving first-time permits in 2018 and reported testing activities in California for the first time—extending beyond the typical annual reporting cycle. Each entry in these reports contains information on the manufacturer, permit number, date of the event, vehicle identification number (VIN), whether the vehicle was capable of

DOI: 10.54254/2755-2721/2025.LD25515

operating without a human driver, the presence of a driver, the initiator of the disengagement (including AV system, test driver, remote operator, or passenger), the type of location where disengagement occurred (such as interstate, freeway, highway, rural road, street, or parking facility), and a descriptive account of the circumstances that led to the disengagement.

3.2. Data cleansing and pre-processing

The first step in preparing the data involved the manual removal of all 2018 records from the combined tables, ensuring that only the relevant 2019 information remained. Additional columns were incorporated as necessary, including the "DATE," "VEHICLE IS CAPABLE OF OPERATING WITHOUT A DRIVER (Yes or No)," "DISENGAGEMENT LOCATION (Interstate, Freeway, Highway, Rural Road, Street, or Parking Facility) et al." During this process, particular care was taken to standardize the categorical fields to maintain consistency across all records and prevent errors due to discrepancies or inconsistent naming conventions. Subsequent to these preparatory steps, all entries containing missing values in key columns were eliminated from the dataset to ensure data quality and completeness. For duplicate detection, the dataset was systematically examined: the first occurrence of each unique record was preserved, while subsequent identical entries were flagged as TRUE in a designated marker column. All rows marked as duplicates (TRUE) were then removed from the dataset. After completing this rigorous cleaning, out of an initial 8,537 rows, a total of 5,285 were identified as duplicates and deleted, resulting in a refined set of unique records suitable for analysis. It is essential to acknowledge that specific anomalies, such as the undulating formatting included in the "VIN NUMBER" column for Zoox, Inc., are intrinsic to the original source data. This formatting anomaly does not affect the precision or results of this research.

4. Model

4.1. Model definitions and assessment indicators

Binary logistic regression is a prevalent statistical technique employed for addressing classification tasks where the outcome variable is dichotomous, such as "yes/no" or "success/failure" scenarios. This model functions by estimating the probability that a particular event occurs, given a set of predictor variables, through the use of the logit (log-odds) transformation. In essence, it establishes a mathematical relationship between the explanatory variables and the likelihood of the target variable falling into one of two possible categories by applying the logit function to model the probability distribution [8].

Indicators commonly used in assessment models include: (1)Accuracy: percentage of correct predictions;

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

Where TP is predicted positive category, actual positive category; TN is predicted negative category, actual negative category; FP is predicted positive category, actual negative category (misreporting); and FN is predicted negative category, actual positive category (underreporting).

(2)Precision: the proportion of samples predicted to be in the positive category that are actually in the positive category;

DOI: 10.54254/2755-2721/2025.LD25515

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

(3)Recall: the proportion of samples that are actually in the positive category that are predicted to be in the positive category; A high recall rate indicates a reduced number of overlooked reports and is appropriate for contexts where missed detections are critical.

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

(4)F1-score, the reconciled mean of precision and recall.

$$F1-score = \frac{2\times \text{Precision}\times \text{Recall}}{\text{Precision}+\text{Recall}} = \frac{2TP}{2TP+FP+FN} \tag{4}$$

$F1$ scores are particularly important when categories are unbalanced (e.g., few positive categories).

4.2. Model construction and analysis

(1)Model Construction

In logistic regression, it is common to map a feature vector $X = [x_1, x_2, \ldots, x_n]$ to a probability value $P(y = 1X)$, which represents the probability that the target variable y will take the value of 1 given the feature X. Assume that the relationship between the target variable y and the input feature X is expressed through a logistic function as:

$$P(y = 1X) = \sigma(w_0 + w_1x_1 + w_2x_2 + \cdots + w_nx_n) \tag{5}$$

where $\sigma(z)$ is a sigmoid function, viz:

$$\sigma(z) = \frac{1}{1+e^{-z}} \tag{6}$$

such that the value of the model's output $P(y = 1X)$ is between 0 and 1.

To train a logistic regression model, it is essential to construct a loss function, typically employing a log-likelihood function, sometimes referred to as a cross-entropy loss function, i.e.:

$$J(w) = -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}\log\left(h\left(x^{(i)}\right)\right) + \left(1-y^{(i)}\right)\log\left(1-h\left(x^{(i)}\right)\right)\right] \tag{7}$$

where m is the number of training samples, $y^{(i)}$ is the true label of the i th sample (0 or 1), and $h\left(x^{(i)}\right)$ is the predicted value of the model, i.e. $h\left(x^{(i)}\right) = \sigma\left(w_0 + w_1x_1^{(i)} + w_2x_2^{(i)} + \cdots + w_nx_n^{(i)}\right)$ The goal is to minimize this loss function and find the optimal parameters $w_0, w_1, \ldots, w_n$.

To find the parameters that minimize the loss function, a gradient descent algorithm is usually used. The update rule for gradient descent is:

$$w_j := w_j - \alpha \cdot \frac{\partial J(w)}{\partial w_j} \tag{8}$$

DOI: 10.54254/2755-2721/2025.LD25515

where w_j is the parameter j, α is the learning rate, and $\frac{\partial J(w)}{\partial w_j}$ is the partial derivative of the cost function with respect to the parameter w_j

For each parameter w_j, the partial derivatives are given by:

$$\frac{\partial J(w)}{\partial w_j} = \frac{1}{m}\sum_{i=1}^{m}\left(h\left(x^{(i)}\right) - y^{(i)}\right)x_j^{(i)} \tag{9}$$

This represents the adjustment of the corresponding weight parameter w_j by the prediction error (i.e., $h\left(x^{(i)}\right) - y^{(i)}$) of each training sample.

Finally, according to Python, the five factors with the largest and smallest features are selected. Since there are 525 factors in total, the first five feature factors are replaced by 1-5 and the last five factors are replaced by 521-525. The results of the made visualization are shown in Fig. 1 and Fig. 2, where the dominant feature is unwanted maneuver or stop by vehicle — the vehicle had to stop abruptly, has the highest coefficient of 1.223434, which means that it has a significant impact on the disengagement situation.

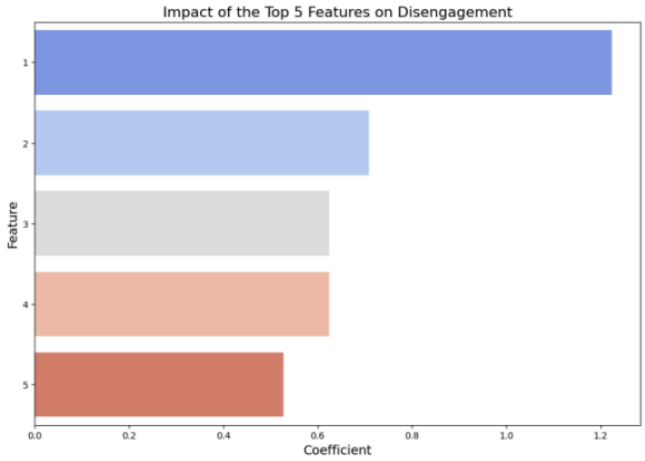

Figure 1. Impact of the top 5 features on disengagement

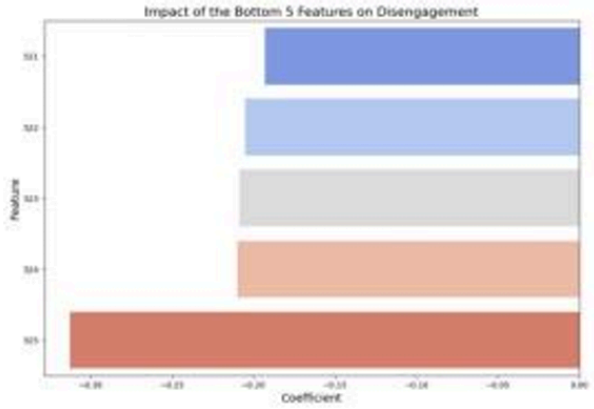

Figure 2. Impact of the bottom 5 features on disengagement

(2)Model Solution

DOI: 10.54254/2755-2721/2025.LD25515

The top five factors with high impact are listed, of which the most influential disengagement factor is characterised by a coefficient of 1.223434, followed by perception discrepancy —a discrepancy arose between how the vehicle's perception sensor (radar, lidar, and camera) data perceived the vehicle's environment and how a human driver perceived the environment, with a coefficient of 0.708116, indicating that when there is a discrepancy between a vehicle's perception system and the actual situation, the likelihood of a disengagement event increases; then Need to manually drive around car stopped in lane, with a coefficient of 0.624676, i.e., this feature indicates that the need to manually drive around a vehicle stopped in lane may increase the probability of disengagement; and then Driver took over to make manual lane change, with a coefficient of 0.624676, i.e., this feature indicates that the driver intervened in order to complete the lane change, which may lead to disengagement of the autopilot system; and Disengage for undesired motion planning behavior, did not slow down when meeting a cut-in vehicle also had a large impact with a coefficient of 0.528097, indicating that undesired motion planning behavior occurred when the vehicle was planning its motion behavior, which could also lead to disengagement.

Conversely, the characteristics with minimal or adverse effects, which may diminish the probability of a disengaged takeover, were selected as the top three in descending order of absolute value as Disengage due to an undesired vehicle manoeuvre resulting from a planning discrepancy during trajectory generation, disengage for a software malfunction related to a potential performance issue with a component of the autonomous driving system (including third-party software components), and disengage when the driver applies the brake pedal because the vehicle was exceeding the expected speed. Vehicle not located within an active construction zone. No emergency vehicles or accidents are present in the area. Weather and road conditions are dry in the vicinity. The coefficients are -0.312555, 0.209890, and -0.208478, respectively. The coefficients are -0.312555, 0.209890, and -0.208478, respectively. This suggests that the occurrence of these elements diminishes the likelihood of disengagement.

(3)Model Training and Testing

In the first step, the model parameters $w_0, w_1, \ldots, w_n$ are randomly initialized; in the second step, the parameters w_j are updated using gradient descent; in the third step, the second step is repeated until the loss function converges, i.e., the model parameters are no longer significantly changed.

(4)Model Validation

Usually models can be validated according to the following three methods:

1) Cross-validation

K-fold cross-validation: the dataset is divided into K subsets, which are sequentially trained with K-1 subsets, and the remaining one subset is used for testing, which is repeated K times. The final model performance is determined by the average result of all K experiments. It helps to reduce overfitting and improve the generalization ability of the model. The cross-validation accuracy is found to be 0.9545 according to Pythonn and has a small standard deviation (±0.0045), indicating stable model performance.

2)Evaluating model performance

The model accuracy was found to be 0.9831 according to Python, which indicates that the overall accuracy of the model is high.

The confusion matrix is: $\begin{bmatrix}[617 & 0] \\ [11 & 23]\end{bmatrix}$, where 617 is True Negative (TN), i.e., the number of correctly predicted No; 23 is True Positive (TP), i.e., the number of correctly predicted Yes; 11 is False Negative (FN), i.e., the number of incorrectly predicted No that were actually Yes; and 0 is

DOI: 10.54254/2755-2721/2025.LD25515

False Positive (FP), i.e., the number of incorrectly False Positive (FP), i.e., the number of Yeses that were incorrectly predicted to be Yes.

The classification report is shown in Table 1:

Table 1. Classification report

	precision	recall	f1-score	support
0	0.98	1.00	0.99	617
1	1.00	0.68	0.81	34
accuracy			0.98	651
macro avg	0.99	0.84	0.90	651
weighted avg	0.98	0.98	0.98	651

The precision is 0.98 for the "No" category and 1.00 for the "Yes" category. The recall is 1.00 for the "No" category and 0.68 for the "Yes" category. The F1-score is 0.99 for the "No" category and 0.81 for the "Yes" category. It shows that the model is very accurate in predicting the No category, but the recall is a little lower for the Yes category, which means the model may miss some of them. shows that the model is very accurate in predicting the No category, but the recall for the Yes category is a little lower, which means that the model may miss some of the Yes categories when it recognizes them.

3)ROC curves and AUC values

ROC Curve (Receiver Operating Characteristic Curve): plots the relationship between the true rate (TPR) and the false positive rate (FPR), which is used to evaluate the performance of the model under different thresholds [9]. AUC (Area Under the Curve): the area under the ROC curve; the larger the AUC value, the better the model performance.

Derived from Python:

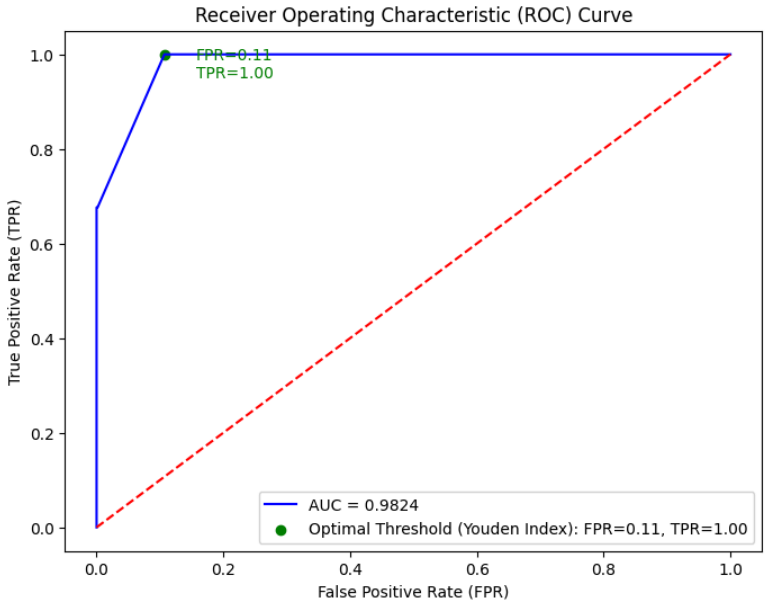

Figure 3. Receiver Operating Characteristic(ROC) curve

The ROC curve is shown in Fig. 3 above, indicating that the blue curve exhibits good classification performance as it is close to the upper left corner. The AUC value is 0.9824, which is a

DOI: 10.54254/2755-2721/2025.LD25515

very high AUC value close to 1 indicating that the model has a very good classification ability.

5. Discussion

Firstly, the range of features utilized in this analysis is somewhat restricted; the current independent variable, which describes the cause of disengagement, may not fully account for all relevant factors influencing disengagement events—such as environmental variables (e.g., weather conditions, road surface status), driver behavior, and other operational elements. In future studies, it would be beneficial to incorporate additional attributes like environmental influences, driver response times, and specific road characteristics to enrich the explanatory power of the model. Secondly, binary logistic regression inherently assumes a linear relationship between the predictors and the log odds of the outcome. Consequently, if the fundamental interactions are nonlinear, this method may fail to adequately represent the data's complexity. To mitigate these constraints, more advanced algorithms —such as decision trees or random forest models—should be investigated for improved modelling of nonlinear effects. Additionally, some disengagements may be attributed to prolonged continuous driving; neglecting this temporal information could limit the model's predictive performance. Introducing time-dependent analysis methods, such as time series modeling or recurrent neural networks, might further enhance prediction accuracy by accounting for sequential dependencies. The dataset may possess intrinsic limitations; for example, it can be predominantly derived from a restricted array of manufacturers or car types, hence omitting wider vehicle demographics. Given that disengagement patterns can vary substantially across different brands, vehicle types, and operational environments, expanding the dataset to encompass greater diversity would help improve the generalizability and robustness of future research findings [10].

6. Conclusion

The regression coefficients indicate that the primary factors contributing to disengagement are instances of the vehicle performing unintended manoeuvres or abruptly halting, along with scenarios characterised by perceptual discrepancies—particularly when there is a divergence between the environment as perceived by the vehicle’s sensor array (including radar, lidar, and cameras) and the human driver's interpretation. The presence of either of these conditions significantly raises the probability of disengagement events. In contrast, certain factors are associated with a decreased or even negative impact on disengagement likelihood. For example, "Disengage for unwanted maneuver of the vehicle caused by a planning discrepancy while generating an appropriate trajectory" indicates that if there is a divergence between the intended path and the vehicle’s executed trajectory, the chance of disengagement may actually be lower, likely because the driver can proactively intervene and correct the vehicle’s motion. Similarly, disengagements triggered by software faults—possibly due to performance issues within either native or third-party components of the autonomous system—suggest that these incidents can prompt timely driver intervention, thereby averting more severe system failures. Another situation involves the driver pressing the brake pedal when the vehicle travels faster than expected, especially under normal environmental conditions—such as when the vehicle is not in an active construction zone, there are no emergency vehicles or accidents nearby, and the road and weather conditions are favorable. In such cases, the human driver’s active intervention to slow the vehicle ensures that control is promptly restored, thereby reducing the likelihood of an adverse disengagement event.

DOI: 10.54254/2755-2721/2025.LD25515

References

[1] Chen, L., Li, Y., Huang, C., Li, B., Xing, Y., Tian, D., ... & Wang, F. Y. (2022). Milestones in autonomous driving and intelligent vehicles: Survey of surveys. IEEE Transactions on Intelligent Vehicles, 8(2), 1046-1056.
[2] Wu, K. W., Wu, W. F., Liao, C. C., & Lin, W. A. (2023). Risk assessment and enhancement suggestions for automated driving systems through examining testing collision and disengagement reports. Journal of Advanced Transportation, 2023(1), 3215817.
[3] Chengula, T. J., Mwakalonge, J., Comert, G., Perkins, J., & Ngeni, F. (2023). The use of topic modeling and categorical correlations to unveil autonomous vehicle disengagement patterns. Available at SSRN 4666730.
[4] Huizhao Tu, Chang Lu, Miaojia Lu, & Hao Li. (2024). Safety Influences on Autonomous Driving Road Tests Based on Hazard Avoidance Disengagement. Journal of Jilin University(Engineering and Technology Edition), 54(7), 1935-1943.
[5] Kohanpour, E., Davoodi, S. R., & Shaaban, K. (2025). Trends in Autonomous Vehicle Performance: A Comprehensive Study of Disengagements and Mileage. Future Transportation, 5(2), 38.
[6] Skokan, A., & Mareček, J. (2025). Could Disengagement Reports Indicate Evolution of Autonomous Vehicles?. Vehicles, 7(2), 32.
[7] https: //www.kaggle.com/datasets/art12400/2019-autonomous-vehicle-disengagement-reports
[8] Harris, J. K. (2021). Primer on binary logistic regression. Family medicine and community health, 9(Suppl 1), e001290.
[9] Nahm F. S. (2022). Receiver operating characteristic curve: overview and practical use for clinicians. Korean journal of anesthesiology, 75(1), 25-36. https: //doi.org/10.4097/kja.21209
[10] Sinha, A., Chand, S., Vu, V., Chen, H., & Dixit, V. (2021). Crash and disengagement data of autonomous vehicles on public roads in California. Scientific data, 8(1), 298.