Article

Deciphering Autonomous Vehicle Regulations with Machine Learning

Raj Bridgelall [1],[*] 🆔 **and Denver Tolliver** [2]

[1] Transportation, Logistics, & Finance, College of Business, North Dakota State University, P.O. Box 6050, Fargo, ND 58108-6050, USA

[2] Upper Great Plains Transportation Institute, North Dakota State University, P.O. Box 6050, Fargo, ND 58108-6050, USA; denver.tolliver@ndsu.edu

[*] Correspondence: raj@bridgelall.com or raj.bridgelall@ndsu.edu

Abstract: The emergence of autonomous vehicles (AVs) presents a transformative shift in transportation, promising enhanced safety and economic efficiency. However, a fragmented legislative landscape across the United States hampers AV deployment. This fragmentation creates significant challenges for AV manufacturers and stakeholders. This research contributes by employing advanced machine learning (ML) techniques to analyze state data, aiming to identify factors associated with the likelihood of passing AV-friendly legislation, particularly regarding the requirement for human backup drivers. The findings reveal a nuanced interplay of socio-economic, political, demographic, and safety-related factors influencing the nature of AV legislation. Key variables such as democratic electoral college votes per capita, port tons per capita, population density, road fatalities per capita, and transit agency needs significantly impact legislative outcomes. These insights suggest that a combination of political, economic, and safety considerations shape AV legislation, transcending traditional partisan divides. These findings offer a strategic perspective for developing a harmonized regulatory approach, potentially at the federal level, to foster a conducive environment for AV development and deployment.

Keywords: policy analysis; regulatory framework; data-driven policy making; federal and state vehicle regulations; legislative landscape; machine learning in legislation; political influence on policy; traffic safety; transportation innovation

1. Introduction

Economic and safety incentives are the primary motivators for autonomous vehicle (AV) deployments. The society of automotive engineers (SAE) categorized vehicles that can always operate without a human driver as level 5 automation [1]. Incentives include significant cost savings from eliminating the human driver, increased efficiency by operating at all hours, and enhanced safety by reducing accidents attributable to human factors such as intoxication and distracted driving [1]. Autonomous trucks can additionally benefit by eliminating driver-related costs like insurance, training, and onboard amenities for human comfort and entertainment. Despite these advantages, concerns about technology readiness and potential negative impacts on traffic safety and congestion have emerged. Such apprehensions have led to state-level legislative actions, including laws mandating a human backup driver, which counteract the core benefits of AVs [2].

Currently, the United States lacks comprehensive federal regulation to uniformly address the safety, design, and operation of AVs. The failure of the 2018 U.S. Senate bill, AV START, due to safety concerns, has left a void filled by individual states introducing their own AV legislation [3]. This has resulted in a fragmented regulatory environment, posing challenges for AV manufacturers and stakeholders who must navigate a patchwork of state-specific laws [4]. Furthermore, existing regulations often do not adequately differentiate

between the operation of autonomous trucks and other vehicle types, each with distinct sets of stakeholders.

This research aims to provide a comprehensive analysis of the varying state regulations in the United States and their alignment or conflict with one another. Specifically, this research seeks to identify factors associated with the propensity of an authority to pass laws supporting AV operation without a human backup driver. By applying machine learning (ML) techniques to state data, this study offers insights into the complex relationships between different factors influencing AV policy decisions. This approach bridges the understanding of disparate regulatory frameworks and highlights factors that significantly impact legislative outcomes for AVs.

The organization of the remainder of this paper is as follows. Section 2 presents a review of the literature on AV testing and deployment legislation. Section 3 describes the methodological workflow developed for this research. Section 4 discusses the analytical results and their implications for stakeholders, including investors and policymakers. Section 5 concludes the research and suggests directions for future work.

2. Literature Review

The literature on AVs reveals uncertainties about their impact on traffic, infrastructure, supply chain operations, the environment, and the economy. For instance, some predict that AVs will reduce congestion because they can smooth out traffic flows by eliminating the accordion effect [2]. Another optimistic expectation is that AVs will be safer because they will eliminate errors due to human factors [4]. Contrary views are that AVs will increase congestion because induced demand from lower costs and greater accessibility to non-drivers will add more vehicles to the road, including empty vehicles [5]. Some researchers posit that AVs will never fully eliminate crashes due to human factors so long as they coexist with human-driven vehicles [6]. Research has also highlighted that AVs can exacerbate inequities if policies do not address access for people with low income, people of color, and rural communities [7].

The regulatory landscape allowing AV operation on U.S. public roads is dynamic and involves a complex interplay between federal and state authorities [3]. There is a growing consensus that the federal government should oversee aspects such as accessibility, safety, design, and manufacturing, whereas state and local governments should retain their traditional roles in regulating titling, registration, traffic laws, and deployments [8]. Toward that end, the National Highway Traffic Safety Administration (NHTSA), a federal agency of the U.S. Department of Transportation (USDOT), updated its safety standard to include aspects of AVs such as not mandating steering wheels and driver's seats [9].

AVs cannot become an affordable and robust form of transportation until the legislative landscape matures [10]. However, ensuring safety, liability, and regulatory compliance in an evolving field often faces hurdles in legislative processes [11]. An exhaustive review of all the recent AV bills that failed revealed a complex interplay of factors. These include a lack of support from legislators, a lack of consensus, safety concerns, opposition from influential stakeholders such as labor unions, competing legislative priorities, and concerns about the practicality of enforcement, technology readiness, and policy implications such as potential economic impact. The literature does not currently offer a comparative analysis between states with differing legislative outcomes to provide more depth in understanding why certain regulations passed or failed, so future work could fill that gap.

There was a limited amount of research that recently explored the association of features in an authority with its propensity to pass supportive AV regulations. Alnajjar et al. (2023) analyzed panel data from 2011 to 2018 and found that an increase in a city's use of electric vehicles, gross domestic product (GDP) per capita, freeway vehicle miles traveled (VMT), and land use score was positively associated with the allowance of AV testing, whereas an increase in fatality cases had a negative association [12]. Bezai et al. (2021) noted that perceived safety issues and public acceptance are predominant barriers to AV adoption [13]. A survey by Freemark et al. (2020) revealed that population size and liberal

political ideologies are strongly associated with support for policies that regulate AV use to increase mobility for the low-income and disabled populations, reduce pollution, reduce traffic, increase transit use, and reduce the use of private cars [14]. The study also found no connection between population growth and support for general AV regulations. A survey by Mack et al. (2021) found that, compared with conservative ideologies, moderates and liberals reported higher AV adoption intention based on both higher perceived benefits and lower perceived concerns [15]. A survey by Othman (2021) revealed that the level of fear of AVs increased with an increase in the number of crashes involving AVs [1].

While the literature explores optimistic and pessimistic predictions about AVs, the present research provides a grounded analysis of how legislative actions are translating these expectations. This helps in understanding the practical challenges and considerations, which often go beyond theoretical discussions, that policymakers face. Whereas some research provided a high-level comparison of how other countries approach AV laws [3], there have been no comprehensive and recent studies focused on the dynamic U.S. regulatory landscape. Hence, the present research fills that gap by conducting an exhaustive review of recent AV bills in the United States to uncover the multifaceted factors influencing the success or failure of AV legislation.

While previous studies have explored associations between various socio-economic factors and AV regulation, the present research contributes by employing ML techniques to systematically analyze and rank these factors. This methodological rigor provides a more robust and data-driven foundation for understanding the dynamics of AV regulation. Furthermore, the present research aligns with and expands upon the findings of Freemark et al. (2020) and Mack et al. (2021) by offering empirical evidence on how political ideologies and public perceptions might influence AV legislation. This contribution is crucial in understanding the policy landscape, especially in terms of addressing public concerns and shaping future legislative strategies.

3. Methodology

Figure 1 illustrates the overall workflow of the methodology. The next subsections detail each procedure shown in the workflow. The Python programming language and its associated ML libraries implemented all procedures.
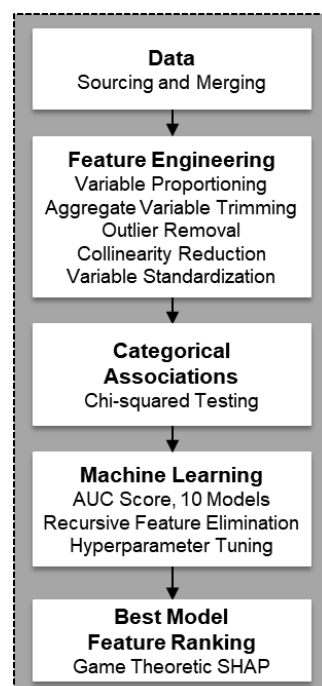

Figure 1. The methodological workflow.

3.1. Data and Feature Engineering

The authors identified the 55 features summarized in Table 1 to test their association with the target class. This research defined the positive target class for ML prediction as regulations in which a human backup driver is optional. Hence, the target variable "optional" is a binary categorical variable that takes on the value one for states that stipulate an on-board human driver to be optional, and zero otherwise. The features spanned a broad realm of general statistics from each state to capture potential influences from diverse factors, including governance (gov.), economic (eco.), energy, environmental (env.), and social. The features also covered specific statistics, including infrastructure, transportation (transp.), and trucking economics. The values for all features are from the year 2021 for alignment with when the government officially published the results of the electoral college votes for the 2020 U.S. presidential election.

Table 1. Features selected and engineered for ML.

Category	Feature	Description	Source
Gov.	Votes_D	Democratic electoral college votes reported in 2021	[16]
	Votes_R	Republican electoral college votes reported in 2021	[16]
	VDpPOPM	Democratic electoral college votes per million population	Derived
	VRpPOPM	Republican electoral college votes per million population	Derived
	MPO	Metropolitan planning organizations (MPOs)	[17]
	POPMpMPO	Million population per MPO	Derived
Eco.	GDP_B	State gross domestic product (GDP) in billion dollars	[18]
	GDPpPOP	GDP_B/POP_M or GDP (USD B) per million population (POP)	Derived
	TM_B	Billion ton-miles (TM)	[19]
	TMpVMT	TM_B/VMT_HW_B or TM per vehicle miles traveled (VMT)	Derived
	Port_M_Ton	Million tons moved by seaports	[17]
	HW_Rev_M	State and local highway (HW) revenue (millions)	[20]
	REVpTM	HW_Rev_M/TM_B or USD M revenue per B ton-miles	Derived
	HW_Exp_M	State and local highway expenditures (millions)	[20]
	EXPpTM	HW_Exp_M/TM_B or USD M expense per B ton-miles	Derived
	REVpEXP	HW_Rev_M/HW_Exp_M or highway revenue/expense ratio	Derived
Energy	BTU_T	Trillion British thermal units (BTU) consumed by transport	[21]
	BTUpVMT	BTU_T/VMT_HW_B or trillion BTU per billion VMT	Derived
	Petrol_BBTU	Gasoline consumed by transportation (billion BTUs)	[21]
	GASpVMT	Petrol_BBTU/VMT_HW_B or BTUs of gasoline per VMT	Derived
Social	POP_M	State population (millions)	[22]
	POPpLA	POP_M × 1E6/LSqMi or population per sq mi of land	Derived
Env.	SqMi	State area in square miles	[23]
	LSqMi	Land area in square miles	[23]
	WSqMi	Water area in square miles	[23]
	WISqMi	Inland water in square miles	[23]
	WCSqMi	Coastal water in square miles	[23]
	CO2_MMT	Million metric tons of CO_2 emitted by transportation	[21]
	CO2pPOP	CO2_MMT/POP_M or annual CO_2 ton per capita	Derived
	Snow_Days	Annual days of snow on average	[24]
	Snow_Ins	Annual inches of snow on average	[24]
Infra.	Pub_Road_Mi_K	Miles (thousands) of public roads	[17]
	UI_Miles	Urban interstate miles	[17]
	RI_Miles	Rural interstate miles	[17]
	RoadpLand	Pub_Road_Mi_K per LSqMi	Derived
	UIpLand	UI_Miles per LSqMi	Derived
	RIpLand	RI_Miles per LSqMi	Derived
	PCT_Sys_Acc	Percentage of system (overall) acceptable IRI	[17]
	Bridges	Number of bridges in 2021	[17]
	BRpPOPK	Bridges/(POP_M × 1000) or bridges per thousand population	Derived
	Fr_Rail_Mi	Freight rail miles	[17]
	RailpLand	Freight rail miles per LSqMi	
	PORTpPOP	Port_M_Ton/POP_M or port tons per capita	Derived

Table 1. *Cont.*

Category	Feature	Description	Source
Transp.	VMT_HW_B	State VMT on highways (billion miles)	[25]
	VMTpLAN	VMT_HW_B/(LSqMi/1000) or B VMT per thousand sq. mi. land	Derived
	HW_Fatals	Highway fatalities	[17]
	FATLpPOP	HW_Fatals/POP_M or fatalities per million population	Derived
	Truck_Estab	Trucking business establishments	[17]
	Truck_Emp	Trucking company employees	[17]
	TRKEpCO	Truck_Emp/Truck_Estab or employees per company	Derived
	Truck_Pay_M	Trucking company annual payroll in million dollars	[17]
	TRKPpE	Truck_Pay_M * 1E6/Truck_Emp or truck pay per employee	Derived
	TRKpMPO	Truck_Estab/MPO	Derived
	Agencies	Transit agencies	[26]
	POPKpTRA	Population (thousands) per transit agency	Derived

As shown in Figure 1, the feature engineering process of the methodological workflow (Figure 1) included variable proportioning, aggregate variable trimming, outlier removal, collinearity reduction, and variable standardization. Aggregate variables tend to have heavily skewed distributions because they reflect the large variation in a dataset, such as state area and population, leading to distorted model performance. Proportional variables reduce skewness and provide control for confounding factors. For instance, a larger population might naturally have a higher total GDP, whereas GDP per capita controls for population size, giving a clearer picture of economic health. That is, proportional variables normalize the scale of the data to make them more meaningful for comparison across districts. Eliminating the aggregate variables encapsulated in the proportional variable helps to simplify model training without losing crucial information. The features labeled in Table 1 as "derived" are the proportional variables used for ML.

Outlier removal is important for improving the accuracy of the statistical analysis because it can significantly affect the mean and standard deviation of the dataset, resulting in misleading interpretations. Models based on feature distance metrics or normality assumptions are sensitive to outliers. This analysis removed Alaska, Hawaii, and the District of Columbia (DC) because of their geographical anomalies. Canada separates Alaska from the contiguous United States (CONUS), and Hawaii is an island state in the Pacific Ocean. DC is a federal district of the United States and has a unique status and demographic profile compared with states. These regions are also statistical outliers based on their landmass, with Alaska and DC being the largest and smallest, respectively. Furthermore, DC has no rural interstate miles. These differences can introduce variations that are misaligned with the general patterns of CONUS states.

Collinearity refers to the situation where two or more predictor variables of the ML model are highly correlated such that one is accurately predictable from the other. Hence, building ML models with highly correlated features can lead to overfitting, a condition where the model fits the training data well but fails to generalize in accurately predicting unseen data. Therefore, reducing multicollinearity by eliminating highly correlated variables increases the reliability, generalizability, and speed of training statistical models. This analysis used the Pearson correlation coefficient r as the measure of collinearity such that

$$r = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}} \tag{1}$$

where x_i and y_i are the individual samples, with $\overline{x}$ and $\overline{y}$ being their mean values, respectively.

The numerator represents the covariance between the two variables and the denominator is the product of the standard deviations of the two variables. Hence, r quantifies the degree to which a change in one variable associates with a change in another variable, assuming a linear relationship.

Standardizing the features to center the values around the mean with a unit standard deviation is a crucial preprocessing step for many ML algorithms. The transformation

involves subtracting the variable mean and dividing by the standard deviation. Standardization makes the variables more comparable on the same scale, resulting in a more efficient and effective training process. Other benefits of standardization are the elimination of non-uniformity in measurement scales due to different units such as miles and tons, and reducing the influence of large values that can dominate the model's learning process and lead to biased results.

3.2. Categorical Associations

The chi-squared test of independence is a statistical approach to determine if there is a significant relationship between two categorical variables. The test is based on comparing observed frequencies in each category of the variable with the expected frequencies if the two variables have no association or dependencies. Conducting the test requires a contingency table of the categories where E_{ij} is the expected frequency for the cell at row i and column j such that

$$E_{ij} = \frac{\sum_j O_{ij} \times \sum_i O_{ij}}{\sum_{ij} O_{ij}} \tag{2}$$

where O_{ij} is the observed frequency at the intersection of category i for one variable and category j for the other variable. A chi-squared statistic then measures how much the observed frequencies deviate from the expected frequencies such that

$$\chi^2 = \sum_{i,j} \frac{\left(E_{i,j} - O_{i,j}\right)^2}{E_{i,j}} \tag{3}$$

which is based on the sum of the squared differences between the entries of the expected and observed values. The null hypothesis is that there is no association between the two variables. Evaluating a p-value determines if one can reject the null hypothesis of no association. The p-value is the probability of observing a chi-squared statistic at least as large as the one observed when the null hypothesis is true. One can reject the null hypothesis at a low chance (typically a p-value ≤ 0.05) of observing such a large deviation from the expected frequencies, suggesting that there is an association between the two variables. The p-value is the area under the chi-squared distribution curve to the right of the chi-squared statistic. The general form of the probability density function (pdf) for the chi-squared distribution is

$$f(x;k) = \frac{2^{-k/2}e^{(-x/2)}x^{(k/2-1)}}{\Gamma(k/2)} \tag{4}$$

where x is a positive chi-squared statistic, k is the degree of freedom (DOF), and Γ is the Gamma function. The DOF is $(r-1) \times (c-1)$ where r and c are the number of rows and columns, respectively, of the contingency table. The Gamma function is

$$\Gamma(z) = \int_0^1 \left[\ln\left(\frac{1}{t}\right)\right]^{z-1} dt. \tag{5}$$

In practice, statistical software packages use numerical methods to compute the integral for finding the area under the pdf.

3.3. Machine Learning

The limited dataset of 48 states restricts the complexity of the ML models. Complex models, such as deep neural networks, typically require large amounts of data to learn effectively [27]. Simpler models are more appropriate with small datasets, even though they may not capture the complexity of the relationships in the data as effectively as more complex models. To address these challenges, the workflow used 10 robust and mature ML models and employed cross-validation to help in assessing the model's ability to generalize.

The workflow then picked the best performing model to rank feature importance. The models learned the importance of features in providing the best predictive performance for the target class.

The 10 models were decision tree (DT), gradient boosting (GB), extreme GB (XGB), AdaBoost (ADB), random forest (RF), naïve Bayes (NB), k-nearest neighbor (kNN), support vector machine (SVM), logistic regression (LR), and artificial neural network (ANN). These traditional models are more data-efficient, and they provide more meaningful insights even with limited data. These models require fewer data points to learn the underlying patterns in a dataset. As the characteristics of the models themselves are not the focus of this research, Table 2 provides a summary of their theory of operations, advantages, and disadvantages. Numerous textbooks, including [27,28], provide more details about their operation.

Table 2. Machine learning models applied.

Model	Theory of Operations	Advantages	Disadvantages
DT	A flowchart-like structure in which each internal node represents a test on an attribute, each branch represents the outcome of the test, and each leaf node represents a class label.	Easy to interpret and visualize. Can handle both numerical and categorical data and requires little data preparation.	Prone to overfitting, especially with noisy or complex datasets. Can create biased trees if some classes dominate.
GB	Builds an additive model in a forward stage-wise fashion, allowing for the optimization of arbitrary differentiable loss functions.	Highly effective on various problems. Can handle mixed types of data, and often provides high predictive accuracy.	Requires careful tuning, is computationally expensive, and can overfit on small datasets.
XGB	Optimized distributed gradient boosting library utilizing decision trees with high performance and flexibility.	Handles sparse data well, provides built-in cross-validation, works well on large and complex datasets, and often outperforms other algorithms in predictive accuracy.	Computationally intensive, requires careful tuning of parameters, and may overfit if not properly regularized.
ADB	Short for adaptive boosting, is an ensemble boosting technique that adjusts the weights of weak classifiers to form a strong classifier.	Often achieves high accuracy, is less prone to overfitting, and improves the accuracy of weak classifiers.	Sensitive to noisy data and outliers, and performance can suffer if weak classifiers are too complex.
RF	An ensemble learning method that constructs multiple decision trees during training and outputs the mode of classes of the individual trees.	Effective in high-dimensional spaces, robust to overfitting, and handles both categorical and numerical data well.	Model interpretability can be challenging, and performance is dependent on the strength of individual trees.
NB	A simple probabilistic classifier based on applying Bayes' theorem with strong (naive) independence assumptions between the features.	Simple, fast, and efficient with large datasets. Performs well with categorical input variables compared to numerical variables.	The assumption of independent predictors is rarely true. Has a notable limitation with data scarcity for a category of a categorical variable.
kNN	A non-parametric method that classifies a sample by a plurality vote of its neighbors, assigning the class most common among its k nearest neighbors.	Simple, intuitive, and effective in practice, especially for low-dimensional datasets.	Computationally expensive, especially as dataset size grows. Sensitive to irrelevant or redundant features.

Table 2. *Cont.*

Model	Theory of Operations	Advantages	Disadvantages
SVM	Analyzes data for classification by maximizing the margin between decision boundaries and data points in high-dimensional feature space.	Effective in high-dimensional spaces, versatile with different kernel functions, and provides a unique solution since the optimization problem is convex.	Not suitable for large datasets, less effective on noisier datasets with overlapping classes, and requires a good kernel choice.
LR	Measures the relationship between a categorical dependent variable and independent variables by estimating probabilities using a logistic function.	Simple and efficient for binary classification problems. Outputs probabilities that are interpretable.	Assumes linearity between the dependent variable and the independent variables. Can struggle with complex relationships and is not flexible enough to capture more complex patterns.
ANN	Modeled loosely after the human brain to recognize patterns by interpreting sensory data through machine perception, labeling, and clustering.	Highly flexible and capable of learning nonlinear relationships. Performs well on a large variety of tasks and is scalable to large datasets.	Not appropriate for small datasets. Computationally intensive, prone to overfitting, and lacks interpretability.
Con.	The constant (Con.) model always predicts the same output regardless of the input.	Provides a baseline for model performance comparison.	Offers no real insight into data patterns.

3.3.1. Measure of Predictive Performance

The workflow used the AUC score as the primary measure of a model's predictive performance. The AUC score, which stands for "area under the curve", refers specifically to the area under the receiver operating characteristic (ROC) curve. The ROC curve is a graphical representation that illustrates the diagnostic ability of a binary classifier as a function of its discrimination threshold. The horizontal and vertical axes of the ROC plot the false positive and true positive rates, respectively. The false positive rate ($1 -$ specificity) is the ratio of false positives to the sum of false positives and true negatives. The true positive rate (sensitivity) is the ratio of the true positives to the sum of the true positives and the false negatives. Hence, the ROC captures the trade-off between the sensitivity and specificity of a model. The AUC score ranges from 0 to 1, providing an aggregate measure of the model's performance across all classification thresholds. A model that perfectly discriminates between the positive and negative class will have an ROC curve that passes through the upper left corner of the plot, which represents 100% sensitivity and 100% specificity with AUC = 1. An AUC score of 0.5 signifies that the model has no class separation capacity, which is no better than random guessing. The AUC score is particularly useful when dealing with imbalanced datasets, as it does not bias toward the majority class.

3.3.2. Recursive Feature Elimination (RFE)

For small datasets, the process of feature selection becomes critical yet challenging. Selecting the right features (variables) that have significant predictive power is crucial because a small dataset magnifies the impact of each feature. However, with limited data, it can be difficult to discern which features are truly relevant and which are coincidentally correlated with the target variable. Recursive feature elimination with cross-validation (RFECV) is a method in the ML workflow that identifies the most expressive features for a predictive model. RFECV begins by using cross-validation to train a model with the full set of features to obtain their level of importance either through coefficients or feature importance scores, and then iteratively eliminates the least prominent features. Cross-validation accounts for the variability in the dataset and reduces the risk of overfitting,

leading to a more robust feature selection. Cross-validation iteratively divides the dataset into a specified number of subsets called folds, trains the model on the union of all folds except one, and validates the predictive performance on the held-out fold. The cross-validation ends after involving each fold in the validation at least once.

RFECV keeps track of the model's performance based on the average AUC score for each set of features and reports the final selection that maximizes its predictive performance. The methods used to determine feature importance scores vary, depending on the model type. For example, linear models assign coefficients to the features based on the sensitivity of the target variable to changes in each feature, with all other features remaining constant. Decision tree-based models calculate feature importance based on the contribution of each feature to the purity of the node in a tree. A significant drop in model performance when there is a change in feature value indicates high importance. The scale of the features can significantly influence the variation of feature importance; hence, feature standardization as described earlier becomes necessary.

3.3.3. Hyperparameter Tuning

Most ML models have so-called hyperparameters, user settings that govern the model's overall behavior. The user must set hyperparameter values prior to the training process. Examples of hyperparameters include the learning rate that affects model convergence, the number of trees in a random forest, the depth of a decision tree, and the number of hidden layers in a neural network. Manually setting hyperparameters is not necessarily intuitive and can involve guessing and testing various combinations. This research used a more systematic grid search method that evaluated every hyperparameter combination to maximize model performance. The approach required defining a grid of parameter range and increment for evaluation, and then conducting an exhaustive search for the best performing parameter combination based on maximizing a performance metric such as the AUC. As the grid search method is computationally intensive, it is advantageous to apply RFECV prior to hyperparameter optimization.

3.4. Feature Ranking

This research adopted the Shapley values (SHAP) from cooperative game theory to fairly assign a value to each feature based on their contribution toward the prediction. The calculation accounts for the interaction effects between features by considering all combinations of features and how the addition of each feature contributes to the change in prediction. Hence, a SHAP value is the average marginal contribution of a feature value across all combinations such that

$$\phi_j = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} \left[f_S\left(X_{S \cup \{j\}}\right) - f_S(X_S) \right] \tag{6}$$

where ϕ_j is the SHAP value for feature j, F is the set of all features, S is a subset of features excluding j, $|S|$ is the number of features in subset S, $|F|$ is the total number of features, $f_S\left(X_{S \cup \{j\}}\right)$ is the model prediction with feature j included, and $f_S(X_S)$ is the model prediction without feature j. The sum is across all the subsets of features that exclude j. Features with higher absolute SHAP values have a greater impact on the model's output.

4. Results

Status is either enacted, executive order (EO), failed, no specific laws (none), pending, or vetoed by the governor. The specification for a human driver to be onboard is either optional, required, or unclear.

4.1. Legislative Status

By the end of 2023, all 48 states of the CONUS, except 4, had processed legislation relating to the operation of Avs in their district. Figure 2 shows the status of AV regulations at the end of 2023 for the CONUS. Figure 3 shows the requirements for a human backup

driver in these regulations. Laws requiring a human backup driver assume SAE level 3 or level 4 operation, expecting that a person in the vehicle will take over control as needed.
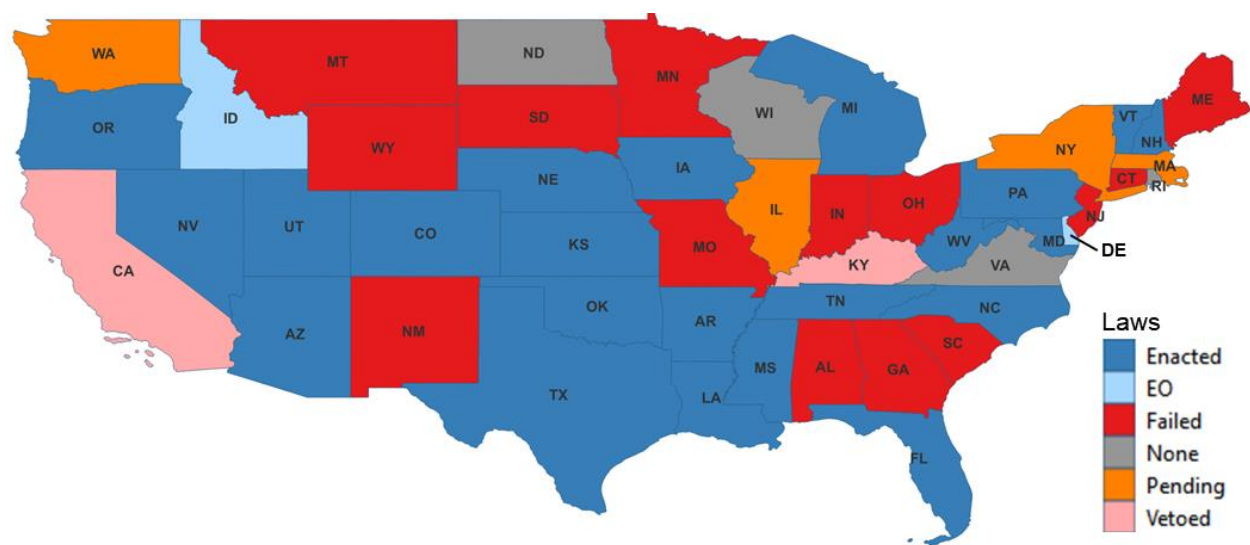

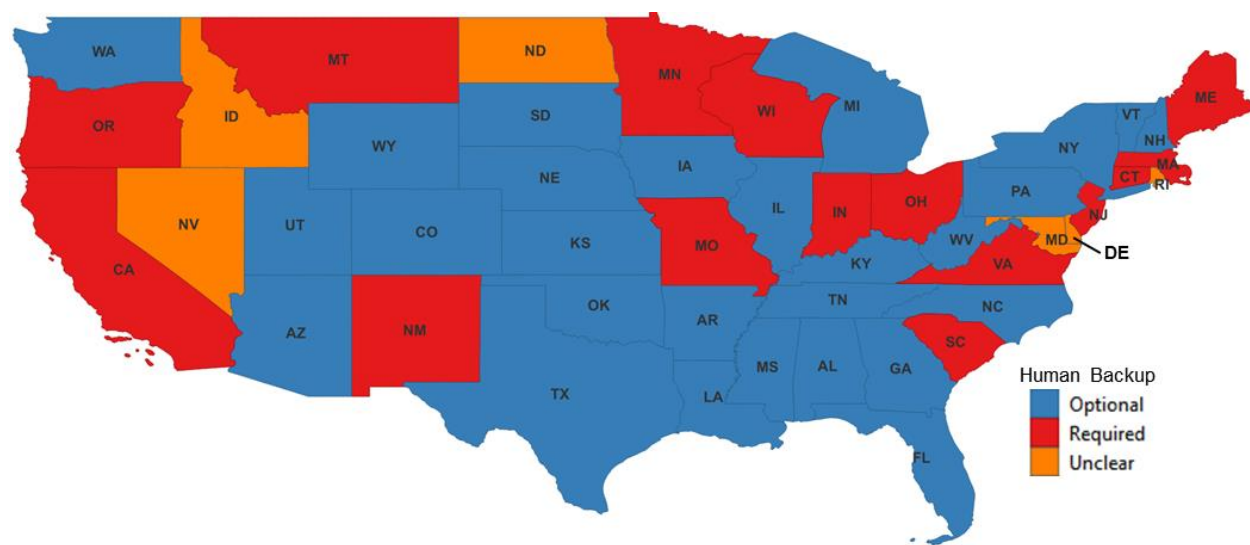**Figure 2.** Status of AV laws as of 2023 in states of the CONUS.

Figure 3. Requirements for a human backup driver as of 2023 in states of the CONUS.

Tables 3–6 summarize the status of AV legislation for the northeastern, southern, midwestern, and western U.S. states, respectively. The first column of each row of the table lists the state name and abbreviation, the bill identification, the year of last action, the status (enacted, pending, failed, vetoed, or none), and whether the legislation makes a human backup driver optional, required, or the law lacks clarity. The states with no pending legislation at the time were Virginia, Wisconsin, North Dakota, and Rhode Island. Nevertheless, those states engage in AV research as summarized in the legislative status tables below.

Table 3. Autonomous vehicle legislation in northeastern U.S. states.

State and Status	Special Conditions
Connecticut (CT) H 6486 (2021): Failed Human: Required	The operator must be an employee, a contractor, or trained personnel, holding a valid operator's license. AV testers must register each vehicle with the Commissioner of Motor Vehicles. Requires proof of liability insurance or a surety bond of at least USD 5 million. No testing on limited access highways. If involved in a crash, the vehicle must achieve a minimal risk condition and remain at the crash scene. The owner or a representative must report the crash immediately, remain until law enforcement arrives, and provide necessary information.
Maine (ME) H 1222 (2019): Failed Human: Required	Providers must have a license and maintain at least USD 25 million in insurance coverage. Providers must consent to state authority and have a registered office in Maine. If involved in an accident, the AV shall remain on the scene and the provider shall inform law enforcement that the system was engaged.
Massachusetts (MA) S 2257 (2023): Pending Human: Required	Clearly marked to indicate that it is operating autonomously. Data capture and storage systems must comply with the department's requirements. The registrar must certify manufacturers or operators.
New Hampshire (NH) S 216 (2020): Enacted Human: Optional	The AV must be capable of achieving a minimal risk condition if it malfunctions. The owner or manufacturer must submit proof of monetary responsibility.
New Jersey (NJ) A 1853 (2020): Failed Human: Required	Manufacturers must apply and receive commission approval. The operator must have the proper license for the vehicle type. Manufacturer's designated personnel must operate the vehicle. Manufacturer must obtain USD 5 million insurance or surety bond. Manufacturer must provide a written disclosure of information collected by the technology.
New York (NY) A 539 (2023): Pending Human: Optional	The AV must be capable of achieving a minimal risk condition if it malfunctions.
Pennsylvania (PA) H 2398 (2022): Enacted Human: Optional	Obtain certificate of safety and performance compliance from the Department of Motor Vehicles. Clearly marked to indicate that it is operating autonomously. A manufacturer-designated entity must monitor the vehicle's operation and be able to remotely intervene if necessary. Must operate within approved geofenced areas.
Rhode Island (RI) No bills Human: Unclear	No specific laws. However, the Rhode Island Department of Transportation (RIDOT) published a request for information in 2017 seeking advice relating to AV deployments [29].
Vermont (VT) S 149 (2019): Enacted Human: Optional	Exempts AV tester from the requirement to hold a commercial driver's license. AV tester must obtain a permit from the Commissioner of Motor Vehicles and notify the agency of any accidents involved.

Table 4. Autonomous vehicle legislation in southern U.S. states.

State and Status	Special Conditions
Alabama (AL) SB311 (2023): Failed Human: Optional	Vehicle must be capable of achieving a minimal risk condition if it malfunctions. If involved in an accident, the AV shall remain on the scene, and the owner or representative must report. State, federal, or common law determines liability.
Arkansas (AR) H 1562 (2021): Enacted Human: Optional	Remote operators must have a valid driver's license for the vehicle class. State or federal law determines liability for accidents. The original manufacturer of a vehicle converted into an AV by a third party is not liable in legal actions.
Delaware (DE) EO 14 (2017): EO Human: Unclear	Does not provide specific conditions related to AV operations on public roads. Primarily focuses on establishing an advisory council on AVs.
Florida (FL) H 1289 (2021): Enacted Human: Optional	Restrictions apply to the operation of AVs on roads with certain speed limits.

Table 4. *Cont.*

State and Status	Special Conditions
Georgia (GA) H 249 (2021): Failed Human: Optional	The AV must be capable of achieving a minimal risk condition if it malfunctions.
Kentucky (KY) H 135 (2023): Vetoed Human: Optional	Achieve a minimal risk condition upon a malfunction. If involved in an accident, the AV owner or a representative must notify law enforcement and file a written report within ten days of the occurrence.
Louisiana (LA) S 453 (2022): Enacted Human: Optional	Exempt AVs from state motor vehicle equipment laws or regulations that relate to or support operation by a human driver and are not relevant for an automated driving system.
Maryland (MD) S 685 (2023): Enacted Human: Unclear	Does not provide specific conditions related to AV operations on public roads. Primarily focuses on authorizing AV converters to sell or lease.
Mississippi (MS) H 1003 (2023): Enacted Human: Optional	The AV must be capable of achieving a minimal risk condition if it malfunctions.
North Carolina (NC) H 814 (2022): Enacted Human: Optional	Low speed AVs can operate on streets and highways where the posted speed limit is 45 mph or less. On a highway with two travel lanes, AVs must turn off the roadway to allow faster moving vehicles to pass when it is safe to do so.
Oklahoma (OK) S 1541 (2022): Enacted Human: Optional	The owner must submit a law enforcement interaction plan to the Department of Public Safety. The owner must provide proof of insurance coverage equal to at least USD 1 million. The owner must report an accident while the AV remains on the scene.
South Carolina (SC) H 4015 (2013): Failed Human: Required	Have a designated or remote operator to assume control if needed. Have a mechanism to capture and store data in case of an accident.
Tennessee (TN) S 0151 (2017): Enacted Human: Optional	The AV must be capable of achieving a minimal risk condition if it malfunctions. Owner maintains insurance providing of at least USD 5 million for death, bodily injury, and property damage.
Texas (TX) S 2205 (2017): Enacted Human: Optional	Clearly signal autonomous operation. Have a licensed entity monitor and remotely operate if needed. Should not operate when weather or road conditions pose unreasonable risks. Full deployment requires federal approval and a regulatory framework established by the Department of Public Safety.
Virginia (VA): No Bills Human: Required	No specific laws. However, Virginia Tech Transportation Institute (VTTI) leads an initiative aimed at supporting AV testing and certification [30].
West Virginia (WV) H 4787 (2022): Enacted Human: Optional	Submit a law enforcement interaction plan to the department of motor vehicles. The AV must be capable of achieving a minimal risk condition if it malfunctions and request a human driver to intervene.

Table 5. Autonomous vehicle legislation in midwestern U.S. states.

State and Status	Special Conditions
Illinois (IL) H 2913 (2023): Pending Human: Optional	The manufacturer must notify the Secretary of State and self-certify under certain conditions, define geographical boundaries for the operation, maintain incident records, and provide summaries to the Secretary of State and the National Highway Traffic Safety Administration.
Indiana (IN) S 141 (2023): Failed Human: Required	The human safety operator must meet all state and federal qualifications to operate both a motor vehicle and an automated vehicle.
Iowa (IA) S 302 (2019): Enacted Human: Optional	The AV must be capable of achieving a minimal risk condition if it malfunctions. If involved in an accident, the AV shall remain on the scene. The owner or a representative must promptly report any accident.

Table 5. *Cont.*

State and Status	Special Conditions
Kansas (KS) S 313 (2022): Enacted Human: Optional	The AV must be capable of achieving a minimal risk condition if it malfunctions. If involved in an accident, the AV shall remain on the scene. The owner or a representative must promptly report any accident.
Michigan (MI) S 706 (2022): Enacted Human: Optional	Must comply with user fee requirements on designated roadways or lanes.
Minnesota (MN) SF 0214 (2022): Failed Human: Required	Prohibits AV operation on Minnesota roads and highways. Requires a human operator to perform all dynamic driving tasks.
Missouri (MO) S 452 (2021): Failed Human: Required	Prohibits AVs from operating without a human driver and imposes a fine of up to USD 1000 if violated. AVs are subject to the same laws as conventional vehicles.
Nebraska (NE) LB 989 (2018): Enacted Human: Optional	The AV must be capable of achieving a minimal risk condition if it malfunctions. If involved in an accident, the AV shall remain on the scene. The owner or a designated person shall report any crash or collision. The AV must safely negotiate railroad crossings unless granted an exemption.
North Dakota (ND) No bills Human: Unclear	No specific laws about operating on public roads. However, HB 1519 allocated funding for grants and studies on multiple autonomous technologies, including uncrewed aircraft systems.
Ohio (OH) HB699 (2018): Failed Human: Required	A licensed commercial driver must be physically present in the vehicle, monitoring its operation and ready to take control in emergencies. Operator must maintain proof of monetary responsibility throughout the registration period.
South Dakota (SD) H 1120 (2023): Failed Human: Optional	Have a designated supervisor. Have a fallback plan for human takeover in case of system failure. Operate within a limited design domain.
Wisconsin (WI) No bills Human: Required	No specific laws. However, the state has an AV proving ground [31].

Table 6. Autonomous vehicle legislation in western U.S. states.

State and Status	Special Conditions
Arizona (AZ) S 1333 (2022): Enacted Human: Optional	The AV owner or a representative must report accidents. The AV must stop at accident scenes or as close as possible. Liability determined by state or federal law.
California (CA) A 316 (2023): Vetoed Human: Required	AVs over 10,001 pounds must have a trained human safety operator to operate and shut off the vehicle. The operator must meet all federal and state qualifications for the vehicle type, in both automated and non-automated modes. Manufacturers must report collisions involving AVs to the Department of Motor Vehicles within 10 days. Owners must submit annual submission of disengagement reports.
Colorado (CO) S 213 (2017): Enacted Human: Optional	Human driver laws do not apply to automated driving systems, except those regulating the physical driving of a vehicle. Prohibits testing if the system cannot comply with laws, unless approved by the Colorado State Patrol and Colorado Department of Transportation. State, federal, or common law determines liability.
Hawaii (HI) S 1008 (2021): Pending Human: Unclear	Does not provide specific conditions related to AV operations on public roads. Primarily focuses on establishing an autonomous transportation task force within the Department of Transportation.
Idaho (ID) EO 2018-01 (2018): EO Human: Unclear	Does not provide specific conditions related to AV operations on public roads. Primarily focuses on creating a committee to study and support the testing and deployment of autonomous and connected vehicles.

Table 6. *Cont.*

State and Status	Special Conditions
Montana (MT) H 339 (2023): Failed Human: Required	AVs cannot operate on public highways until the Department of Transportation completes specific rulemaking.
Nevada (NV) S 182 (2023): Enacted Human: Unclear	Does not provide specific conditions related to AV operations on public roads. Primarily addresses regulatory aspects concerning the manufacturing and selling of AVs.
New Mexico (NM) H 378 (2023): Failed Human: Required	An operator must be able to monitor vehicle performance, intervene, operate the vehicle, and turn off the engine.
Oregon (OR) H 4063 (2018): Enacted Human: Required	The operator must possess the proper class of driver's license for the type of vehicle. The owner must obtain a general liability insurance policy that covers no less than USD 5 million for death or bodily injury for each person covered and for property damage.
Utah (UT) H 101 (2019): Enacted Human: Optional	The AV must be capable of achieving a minimal risk condition if it malfunctions or it must request human intervention. Owner or operator must report accidents and provide registration and insurance information.
Washington (WA) S 5594 (2023): Pending Human: Optional	The AV must be capable of achieving a minimal risk condition if it malfunctions. If involved in an accident, the AV shall remain on the scene. The owner must report the accident. Municipalities can require crash data sharing.
Wyoming (WY) S 16 (2022): Failed Human: Optional	Clearly marked to indicate that it is operating autonomously.

The general conditions for allowing AV operation with or without a backup driver were the following:

- Vehicles must have obtained a title and have a valid manufacturer's certificate.
- Vehicles must have registration and insurance.
- Vehicles must comply with all applicable state and federal regulations or have an exemption.
- Occupants must comply with safety belt and child passenger restraining system requirements.
- Operators must possess the proper class of driver's license for the type of vehicle they operate unless specifically exempted by the law.

In general, the laws stipulated that no other local governing body, including cities, may prohibit AV operation, or impose taxes, fees, or other requirements upon them.

The second column of the tables describes any special conditions of the legislation beyond the general requirements described above. Some special conditions require the vehicle to operate within a limited design domain or designated geofenced areas, avoid operation when weather or road conditions pose unreasonable risks, and comply with user fees on designated roadways or lanes. Some special conditions require the vehicle to clearly indicate when it is operating autonomously, and that it must be capable of achieving a "minimal risk condition" if the system fails. Such a condition reduces the risk of an accident by having the system move the vehicle to the nearest shoulder of the road if capable, stopping the vehicle, and activating its emergency signal lamps.

For liability purposes, several state laws define the automated driving system (ADS) to be the vehicle operator. A common requirement is for owners to maintain minimum insurance levels covering death, bodily injury, and property damage. A related requirement is that owners or operators must submit a law enforcement interaction plan to the department of motor vehicles. The plan must describe how to communicate with a fleet support specialist, how to safely remove the vehicle from the roadway, and the steps needed to safely tow the vehicle. A frequent stipulation for vehicles involved in an accident is that it shall remain on the scene while the owner or a representative reports the incident to law

enforcement. A few of the legislations require the system to have a designated or remote operator to assume control if needed.

Other special conditions require the owners to obtain a certificate of safety and performance compliance from the department of motor vehicles. Related requirements stipulate that AVs must have the capability to capture and store data in case of a crash, and that municipalities can require crash data sharing. A related requirement for some legislation is that manufacturers must provide a written disclosure of information that the technology collects.

4.2. Feature Engineering

Figure 4 shows a heatmap of the Pearson correlation coefficient among the selected variables (Table 1). Each cell of the matrix is a color code of the correlation coefficient for all pairwise combinations of the variables. The legend to the right of the heatmap shows the color–value association. The matrix is symmetric, with the diagonal cells showing the highest coefficient of 1.0 (red) because those cells show the variable correlated with itself.
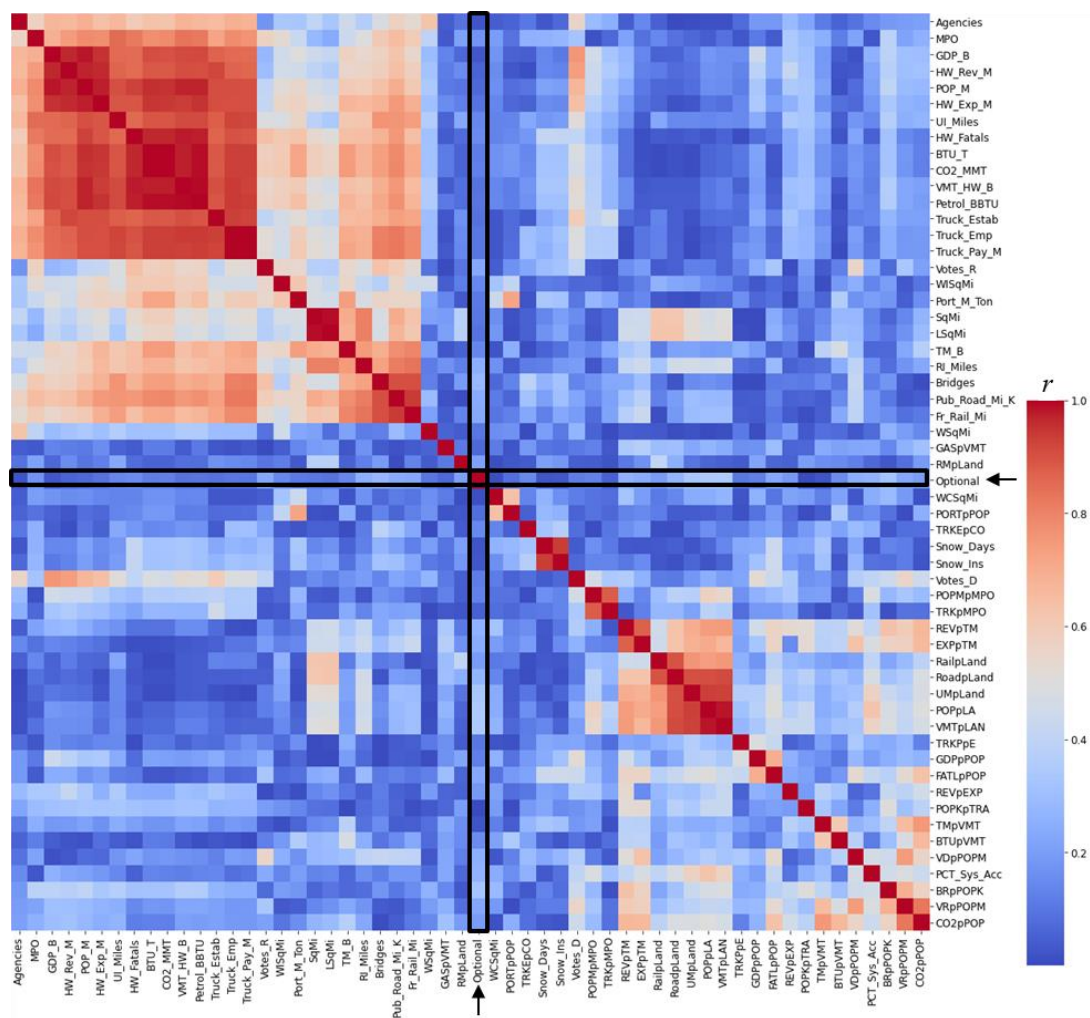

Figure 4. Heatmap of the correlation matrix of the ML features selected.

The heatmap sorted the variables by their correlation level to easily identify the cluster of collinear variables near the top left. The highly correlated variables are aggregate variables such as GDP, population, vehicle miles traveled, highway revenue, highway fatalities, urban interstate miles, energy utilization, fuel consumption, carbon dioxide emissions, the number of transit agencies, trucking establishments, trucking employment, trucking pay-

roll, and the number of MPOs. The procedure eliminated 28 aggregate variables, resulting in 27 variables to further evaluate for collinearity reduction. Based on a customary Pearson correlation coefficient threshold of 0.75, collinearity reduction eliminated 8 variables, leaving 19 for ML. The heatmap also highlights the binary target variable "optional" with an arrow and a bounding box. The heatmap does not show a clear correlation with the target feature, suggesting that they are independent of the target and suitable for the ML process.

4.3. Categorical Associations

Figure 5 shows the observed frequencies of the two categorical variables reflecting the bill status (Figure 2) and the stipulation for a human backup driver (Figure 3). For statistical stability, the analysis consolidated executive orders into enacted laws, vetoed into failed legislation, and no legislation into the pending category. Table 7 summarizes the results of the chi-squared statistical tests for association. The p-value for status and human backup requirement, being significantly less than 0.05, indicates there is a statistically significant association between the status of AV legislation and their requirement for a human backup driver. Figure 5a shows high values for enacted legislation that stipulates a human driver to be optional and failed legislation that requires an on-board human driver. While this test shows a significant association, it does not indicate causation.
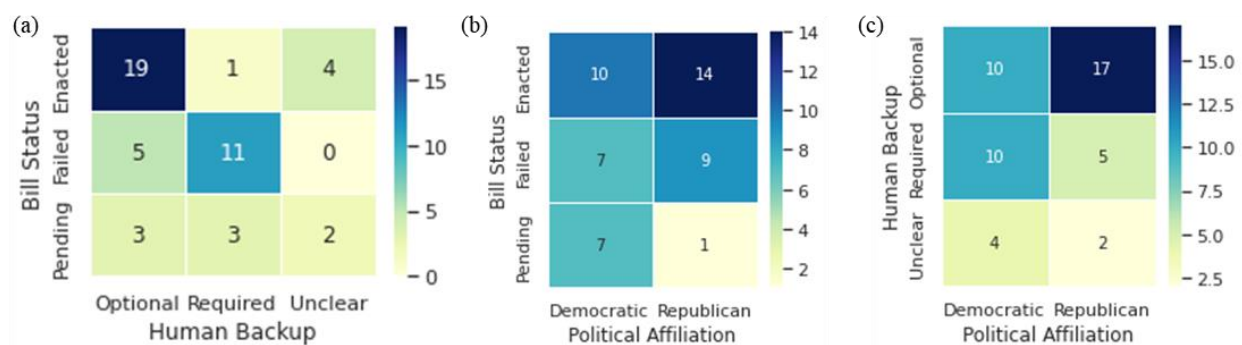
Figure 5. Contingency matrices to test the statistical associations of the categorical variables for (**a**) status vs. human requirement, (**b**) status vs. political affiliation, and (**c**) human requirement vs. political affiliation.

Table 7. Results of the chi-squared statistical test for association.

Test Statistic	Status vs. Human	Status vs. Affiliation	Human vs. Affiliation
Chi-squared	20.79	5.42	4.15
DOF	4	2	2
p-value	3.5×10^{-4}	0.067	0.126
Reject Independence	Yes	No	No

The p-value for status and political affiliation, being greater than 0.05, indicates that there is no statistically significant association between the status of AV legislation and the aggregate political leaning of the state. Similarly, the p-value for human backup requirement and the political leaning of a state, being greater than 0.05, indicates that there is no statistically significant association between requiring a human backup driver and the state's aggregate political leaning. This suggests that the matter is bipartisan, and that other factors may have influenced the proposed stipulation for human requirements and the outcome of the legislation. Therefore, turning to more complex ML models can offer more nuanced insights.

4.4. Machine Learning

Table 8 summarizes the outcome of the ML, ranked by the AUC performance metric for each model. In general, the tree-based models provided better AUC performance than the others, with GB providing the best performance. Hence, the workflow applied RFE and hyperparameter tuning to the GB model prior to ranking the features using the SHAP method.

Table 8. ML predictive performance metrics.

Model	AUC	Accuracy
GB	0.757	0.688
RF	0.721	0.625
XGB	0.674	0.604
DT	0.631	0.604
NB	0.614	0.562
ADB	0.595	0.604
LR	0.578	0.583
kNN	0.550	0.542
SVM	0.540	0.646
ANN	0.501	0.479
Constant	0.500	0.562

The focus of the ML models was on the interpretative value they provide in understanding the complex factors influencing AV legislation rather than on predicting outcomes with high accuracy. Hence, despite the low AUC and accuracy scores, the models enable hypothesis generation because they encapsulate relative feature importance. The low scores reflect the inherently complex and intertwined factors that influence legislative outcome. Hence, low scores do not detract from the model's ability to identify and rank variables based on their relative importance in the dataset. The next subsection presents these features and their ranking.

4.5. Feature Ranking

Figure 6 is the output of the SHAP feature ranking method explained above. The features on the left are in the order of their level of contribution in predicting the positive target class, which is that a human backup driver is optional. Each point on the plot expresses two values associated with each data instance. The color code as shown in the legend is the relative value for that feature, with red and blue representing higher and lower feature values, respectively. The horizontal position is the SHAP value indicating the level of impact of that feature in predicting the target class. The positive and negative SHAP values indicate the impact level in target class prediction as indicated for that direction, with the positive target class being that a human backup driver is optional. Data instances with similar SHAP values will appear as a spatial cluster at that position.

The observation from Figure 6 is that every feature has a mix of instances with high and low values that are associated with the positive target class. However, there are some general trends in terms of most instances with high or low values. For example, most of the instances with low feature values for democratic electoral college votes per million population (VDpPOPM), port tons per capita (PORTpPOP), and population per land area (POPpLA) were associated with the positive target class. In contrast, most of the instances with high feature values for fatalities per million population (FATLpPOP) and population (thousands) per transit agency (POPKpTRA) were associated with the positive target class. The discussion section that follows offers some interpretation of these results.
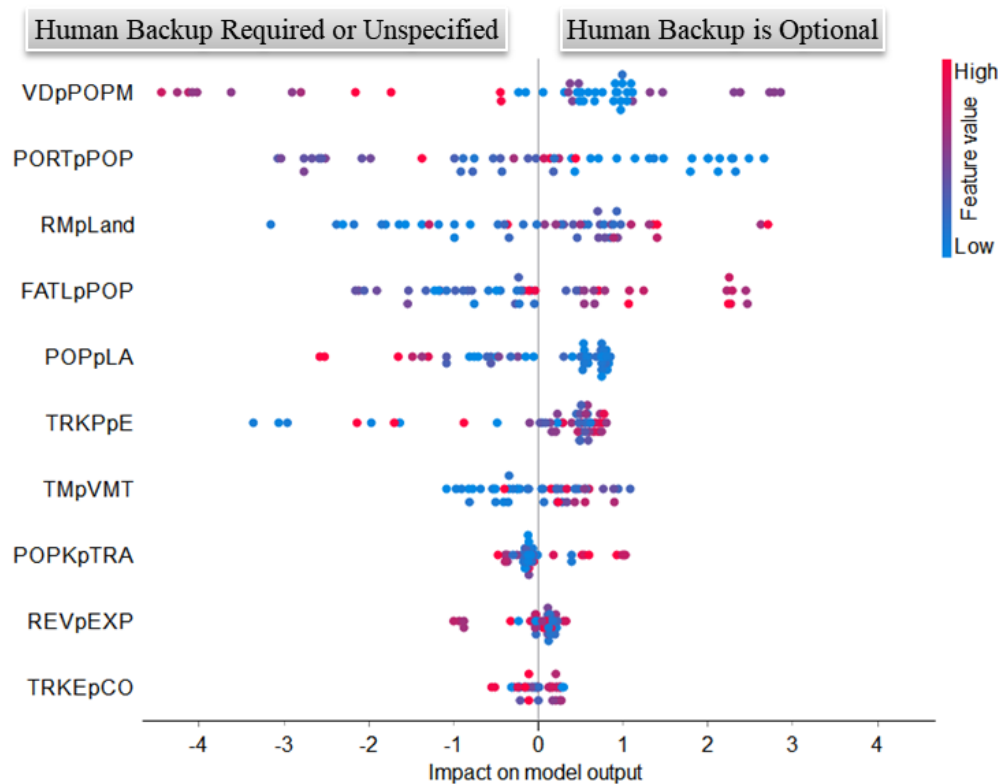
Figure 6. Feature ranking.

5. Discussion

The analysis reveals that a significant majority of CONUS states have engaged in AV-related legislation, underscoring the increasing recognition of AVs' role in the future of transportation. The diverse state-level approaches, particularly regarding the requirement for a human backup driver, reflect the complexity of AV regulation. This variance indicates both the challenges and the opportunities in harmonizing AV policies across different jurisdictions.

The application of machine learning (ML) models, particularly the feature engineering process, has uncovered several factors influencing AV legislation. A key finding is the lack of a strong correlation between the target feature (the requirement for a human backup driver) and other variables. This suggests a multifaceted interplay of factors driving legislative decisions, which goes beyond simple linear associations. The chi-squared test reveals a significant association between legislative status and the requirement for a human backup driver. States with enacted legislation are more inclined to consider human drivers optional, reflecting either a growing confidence in AV technology or a legislative move toward advanced AV deployment stages. Interestingly, the chi-squared test also shows that AV legislation transcends traditional political divides, indicating a broader, non-partisan approach to AV regulation.

The SHAP analysis highlights critical features influencing legislative outcomes. For instance, the association of features like democratic electoral college votes per million population (VDpPOPM) and port tons per capita (PORTpPOP) with the optional status of a human backup driver hints at the role of economic and demographic factors in shaping AV policies. This could indicate that states with less intense political partisanship and lower dependence on port-related economic activities might be more receptive to innovative AV policies. The association of lower population density (POPpLA) with the positive target class could imply that less densely populated states might be more inclined to adopt AV-friendly legislation. This could be due to fewer traffic complexities or a greater willingness to experiment with new technologies in less congested areas. Interestingly,

higher values for fatalities per capita (FATLpPOP) and potential service demand per transit agency (POPKpTRA) are associated with the positive target class. This might indicate that states with higher road fatality rates or higher populations per transit agency are potentially more motivated to seek AV solutions, considering them to enhance road safety and improve public transit efficiency.

The associations between these variables and AV legislation underscore the multifaceted nature of policy making in this domain. This finding suggests that a blend of political, economic, demographic, and safety considerations influence the legislative decisions of states. The influence of political factors (VDpPOPM) and economic activities (PORTpPOP) on AV legislation decisions suggests that broader political and economic contexts play a significant role in shaping the openness of a state to AV technology. The impact of population density (POPpLA) indicates that demographic factors, particularly urbanization levels, may affect the readiness or hesitance of a state to embrace AV technology. The correlation with fatalities (FATLpPOP) and public transit demands (POPKpTRA) suggests that states with more pressing road safety concerns or strained public transit systems might view AVs as a potential solution to these challenges.

These findings emphasize the need for policymakers to consider broader socio-economic contexts when drafting AV legislation. The diversity in state-level AV legislation suggests an urgent need for a more unified regulatory framework, potentially at the federal level. Such harmonization could prevent a disjointed regulatory landscape that might impede AV development and deployment. Emphasizing safety provisions and striving for a minimal risk condition demonstrates a universal focus on public safety and liability, crucial for public acceptance and the success of AV technology. The insights gained from this research can guide stakeholders, including policymakers and industry leaders, in shaping future AV regulations that balance innovation with safety, equity, and public acceptance.

This study, while comprehensive, faces limitations due to the limited data scope and the rapidly evolving nature of AV technology. Hence, future research will incorporate more recent data and consider the dynamic nature of AV legislation through longitudinal assessment. The ML models used might not account for unrecorded variables or emerging trends that could influence legislative outcomes. Moreover, the interpretation of these models requires caution, as they provide correlations that should not be misconstrued as causation. The study simplifies the complexity of legislative processes into quantifiable data. This simplification might overlook the nonlinear nature of how people develop, debate, and enact policies. Hence, future studies will explore the potential of large language models (LLMs) and qualitative analyses applied to a larger corpus to provide deeper insights. Additionally, future work will examine international data to provide a broader perspective on global trends in AV legislation, offering comparative insights that could further inform U.S. policy development.

6. Conclusions

This study aimed to provide insights into the complex legislative landscape of autonomous vehicles (AVs) in the United States. Through the deployment of machine learning (ML) techniques, this study highlighted the multifaceted factors influencing state-level AV legislation, with a particular focus on regulations pertaining to human backup drivers. This study demonstrated that AV legislation is not merely a technological or safety issue, but it encapsulates socio-economic factors, transcending traditional partisan lines. A significant finding is the correlation between states with higher road fatality rates or potential for greater public transit demands and the propensity to adopt AV-friendly policies. This aligns with recognizing AVs as a solution to enhance road safety and public transit efficiency.

There were several implications for policy harmonization and stakeholder engagement. The varied legislative approaches across states underscore the need for a more harmonized regulatory framework, potentially at the federal level. This could streamline the path for AV development and deployment, providing clarity for manufacturers and stakeholders. The findings also advocate for a more inclusive dialogue among policymakers, industry players,

and the public. Understanding the diverse perspectives can aid in crafting regulations that balance innovation with public welfare and safety.

There is a need for ongoing research to track the evolution of AV policies and their impacts. Longitudinal studies can provide insights into the effectiveness of different legislative approaches and inform future policy adjustments. Further research comparing the U.S. legislative framework with international counterparts could yield valuable insights, offering a global perspective on AV regulation strategies. Investigating public perception and the economic impact of AV deployment could offer a more holistic understanding of the societal implications of AVs. Overall, this study illuminates the current state of AV legislation and sets the stage for future explorations in this evolving field. As AV technology continues to advance, it is imperative that our legislative frameworks adapt in tandem, ensuring a safe and efficient integration of AVs into our transportation ecosystem.

References

1. Othman, K. Public acceptance and perception of autonomous vehicles: A comprehensive review. *AI Ethics* **2021**, *1*, 355–387. [CrossRef] [PubMed]
2. Ahmed, H.U.; Huang, Y.; Lu, P.; Bridgelall, R. Technology Developments and Impacts of Connected and Autonomous Vehicles: An Overview. *Smart Cities* **2022**, *5*, 382–404. [CrossRef]
3. Ziyan, C.; Shiguo, L. China's self-driving car legislation study. *Comput. Law Secur. Rev.* **2021**, *41*, 105555. [CrossRef]
4. Hemphill, T.A. Autonomous vehicles: US regulatory policy challenges. *Technol. Soc.* **2020**, *61*, 101232. [CrossRef]
5. Dimitrakopoulos, G.; Tsakanikas, A.; Panagiotopoulos, E. *Autonomous Vehicles: Technologies, Regulations, and Societal Impacts*; Elsevier: Amsterdam, The Netherlands, 2021.
6. Grindsted, T.S.; Christensen, T.H.; Freudendal-Pedersen, M.; Friis, F.; Hartmann-Petersen, K. The urban governance of autonomous vehicles–In love with AVs or critical sustainability risks to future mobility transitions. *Cities* **2022**, *120*, 103504. [CrossRef]
7. Emory, K.; Douma, F.; Cao, J. Autonomous vehicle policies with equity implications: Patterns and gaps. *Transp. Res. Interdiscip. Perspect.* **2022**, *13*, 100521. [CrossRef]
8. USLC. H.R.3711—SELF DRIVE Act. United States Library of Congress. 2022. Available online: https://www.congress.gov/bill/117th-congress/house-bill/3711 (accessed on 11 January 2024).
9. Uhlemann, E. Legislation supports autonomous vehicles but not connected ones [connected and automated vehicles]. *IEEE Veh. Technol. Mag.* **2022**, *17*, 112–115. [CrossRef]
10. Sjoberg, K. Activities on legislation for autonomous vehicles take off [Connected and Automated Vehicles]. *IEEE Veh. Technol. Mag.* **2021**, *16*, 149–152. [CrossRef]
11. Alawadhi, M.; Almazrouie, J.; Kamil, M.; Khalil, K.A. Review and analysis of the importance of autonomous vehicles liability: A systematic literature review. *Int. J. Syst. Assur. Eng. Manag.* **2020**, *11*, 1227–1249. [CrossRef]
12. Alnajjar, H.; Ozbay, K.; Iftekhar, L. An exploratory analysis on city characteristics likely to affect autonomous vehicle legislation enactment across the United States. *Transp. Policy* **2023**, *142*, 37–45. [CrossRef]
13. Bezai, N.E.; Medjdoub, B.; Al-Habaibeh, A.; Chalal, M.L.; Fadli, F. Future cities and autonomous vehicles: Analysis of the barriers to full adoption. *Energy Built Environ.* **2021**, *2*, 65–81. [CrossRef]
14. Freemark, Y.; Hudson, A.; Zhao, J. Policies for autonomy: How American cities envision regulating automated vehicles. *Urban Sci.* **2020**, *4*, 55. [CrossRef]
15. Mack, E.A.; Miller, S.R.; Chang, C.-H.; Fossen, J.A.V.; Cotten, S.R.; Savolainen, P.T.; Mann, J. The politics of new driving technologies: Political ideology and autonomous vehicle adoption. *Telemat. Inform.* **2021**, *61*, 101604. [CrossRef]

16. NARA. Electoral College. 2021. Available online: https://www.archives.gov/electoral-college/2020 (accessed on 6 January 2024).
17. BTS. *State Transportation Statistics*; United States Department of Transportation: Washington, DC, USA, 2024. Available online: https://www.bts.gov/product/state-transportation-statistics (accessed on 7 January 2024).
18. BEA. GDP by State. 2024. Available online: https://www.bea.gov/data/gdp/gdp-state (accessed on 7 January 2024).
19. BTS. *Freight Flows by State*; United States Department of Transportation: Washington, DC, USA, 2024. Available online: https://www.bts.gov/browse-statistical-products-and-data/state-transportation-statistics/freight-flows-state (accessed on 7 January 2024).
20. USCB. *Selected Monthly State Tax Collections*; United States Department of Commerce: Washington, DC, USA, 2024. Available online: https://www.census.gov/data/experimental-data-products/selected-monthly-state-sales-tax-collections.html (accessed on 7 January 2024).
21. BTS. *Energy Consumption and CO2 Emissions*; United States Department of Transportation: Washington, DC, USA, 2024. Available online: https://www.bts.gov/browse-statistical-products-and-data/state-transportation-statistics/energy-consumption-and-co2. (accessed on 7 January 2024).
22. USCB. *State Population Totals and Components of Change: 2020–2023*; United States Department of Commerce: Washington, DC, USA, 2023. Available online: https://www.census.gov/data/tables/time-series/demo/popest/2020s-state-total.html (accessed on 7 January 2024).
23. USCB. *State Area Measurements and Internal Point Coordinates*; United States Department of Commerce: Washington, DC, USA, 2021. Available online: https://www.census.gov/geographies/reference-files/2010/geo/state-area.html (accessed on 7 January 2024).
24. CRPL. Average Yearly Snowfall by American State. Current Results Publishing Ltd. (CRPL). 2024. Available online: https://www.currentresults.com/Weather/US/average-snowfall-by-state.php (accessed on 7 January 2024).
25. BTS. *State Highway Travel*; United States Department of Transportation: Washington, DC, USA, 2024. Available online: https://www.bts.gov/browse-statistical-products-and-data/state-transportation-statistics/state-highway-travel (accessed on 7 January 2024).
26. FTA. National Transit Database. 2021. Available online: https://www.transit.dot.gov/ntd/data-product/2021-database-files (accessed on 11 January 2024).
27. Aggarwal, C.C. *Data Mining*; Springer International Publishing: New York, NY, USA, 2015; p. 734.
28. Burkov, A. *The Hundred-Page Machine Learning Book*; Andriy Burkov: Quebec City, QC, Canada, 2019.
29. RIDOT. *Connected and Autonomous Vehicles and Other Innovative Transport System Technologies Framework for Implementation and Integration*; Rhode Island Department of Transportation (RIDOT): Providence, RI, USA, 2017.
30. VTTI. *Virginia Automated Corridors*; Virginia Tech Transportation Institute (VTTI): Blacksburg, VA, USA, 2024; Available online: https://www.vtti.vt.edu/facilities/vac.html (accessed on 10 January 2024).
31. WAVPG. Automated Vehicles. 2024. Available online: https://wiscav.org/ (accessed on 10 January 2024).