Engineering Research
3048-880X
Rademics Research Institute

https://doi.org/10.71443/er.ar18

A Real-Time Analysis of Predicting Titanic Survival Using Machine Learning Models

E. Elanchezhiyan[1]*, M. Sivaranjani*[2]*, R. Gobinisha*[3]*, F. Delphinaa*[4]

[1,2] *Assistant Professor, Department of Computer Science, Paavai Engineering College, Namakkal, Tamil Nadu, India.*

[3, 4] *Research Scholar, Department of Computer Science, Paavai Engineering College, Namakkal, Tamil Nadu, India.*

Article history

Accepted: 18 December 2025

***Keywords**:*
Machine Learning,
Survival Prediction,
Explainable AI (XAI),
Real-Time Prediction,
Feature Engineering,
Ensemble Learning.

Abstract

Survival prediction was a popular data mining task, providing insights in data-driven decision-making. The present study proposed a systematic approach for predicting survival using machine learning models, explainable artificial intelligence methods, and real-time inference. A robust preprocessing framework was developed to handle missing data, encode categorical variables, and address class imbalance, enhancing data integrity and quality. Data engineering methods, such as the creation of family size and encoding of age groups, were used to extract relevant insights from the data. Various classification algorithms, including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, and Gradient Boosting, were applied with cross-validation for optimal tuning. Models were evaluated based on accuracy, precision, recall, F1-score, and ROC-AUC, allowing for a comprehensive evaluation of their performance. Ensemble models were found to be most effective, as they model complex interactions between features. For improved transparency and understanding, explainable artificial intelligence techniques such as SHAP and LIME were employed to understand the contribution of features to predictions. We also developed a real-time predictive system by combining preprocessing with the trained model to enable fast prediction of new data. The framework was tested for accuracy and speed, showing its potential for real-world use. The results underline the need for a balance of accuracy, interpretability, and real-time prediction to build robust and scalable models.

1. Introduction

The RMS Titanic passenger survival prediction task has traditionally been used as a benchmark problem for supervised machine learning algorithms, especially classical classification algorithms. Initial work using the Titanic data has mostly centered on classical algorithms like logistic regression, decision trees, and Naïve Bayes classifiers due to their simplicity and interpretability. Logistic regression has been widely applied due to its ability to model probabilities and pinpoint statistically significant predictors, while decision trees provide easy-to-understand rule sets that can capture complex interactions between features such as age, gender, and ticket class. Although Naïve Bayes classifiers impose stringent independence assumptions, they have been shown to perform competitively in several studies. These traditional methods have shown the significance of socio-demographic variables, notably gender and ticket class, in predicting survival. But they have been found to exhibit model bias, limited ability to model complex feature interactions, and sensitivity to preprocessing methods. Furthermore, these models typically lack robustness with new patterns in data. As a result, while traditional machine learning approaches set a vital foundation for survival prediction, their limitations have prompted the investigation of more complex methods, such as ensemble methods and complex learning paradigms, to improve the performance and generalizability of survival prediction in more recent research.

Contact – elanchezhiyanelangopec@paavai.edu.in

Ensemble learning has become a prominent approach to boosting the prediction accuracy of classification problems, including survival prediction. This practice involves the combination of multiple base learners to reduce variance, bias, or both, leading to improved performance and generalization. Methods such as bagging and boosting have gained widespread attention, and algorithms such as Random Forest and Gradient Boosting have proven to be highly effective. Random Forest uses bootstrap aggregating to grow multiple decision trees and aggregates their predictions to enhance stability and prevent overfitting. On the other hand, boosting algorithms such as AdaBoost and Extreme Gradient Boosting (XGBoost) adaptively learn from misclassified samples. According to research, these techniques are especially adept at capturing complex non-linear relationships between features, common in structured data. The ensemble methods have shown robustness to noise, feature variability, and missing data, making them ideal for real-world data. These models can come with increased computational complexity and loss of interpretability, making it challenging to use them in real-time applications. As a result, there has been an increased emphasis on the design of ensemble models with improved efficiency. This emerging trend emphasizes the significance of ensemble learning as a key development in moving beyond traditional models and the necessity for efficient, scalable models for practical predictive applications.

Preprocessing and feature creation have been shown to play a critical role in survival prediction problems. Existing studies consistently show that feature engineering that results in meaningful and domain-specific features has often led to better improvements than simply using complex models. For the Titanic dataset, feature engineering strategies like deriving passenger titles from names, calculating family size, and discretizing continuous features such as age and fare have led to improvements in model performance. Further, handling missing data, especially for features like age and cabin, has been done with statistical imputation and predictive models. Categorical variable encoding techniques such as one-hot encoding and ordinal encoding have also been frequently used to enable machine learning algorithms. Data preprocessing was a crucial aspect of data cleaning and preparation, alongside feature engineering. In practice, the data to be modelled include noisy or inconsistent data and class imbalance, which can skew the results. We have used the Synthetic Minority Oversampling Technique (SMOTE), random under-sampling, and cost-sensitive learning to overcome class imbalance and enhance model generalization. The research demonstrates that effective preprocessing not only improves the accuracy of the models, but it also improves the stability and explainability of the model, and as such, it was an essential part of the process when developing predictive models.

The use of deep learning and real-time machine learning systems are new areas of research in survival prediction, with new opportunities and challenges. Artificial Neural Networks, in particular, multilayer perceptrons, have been used to model intricate interactions between features and to automatically extract features. But the small and structured nature of datasets like the Titanic dataset often leads to problems like overfitting and limited improvement in performance with deep learning models as compared to ensemble techniques. As a result, integrated systems integrating deep learning with conventional machine learning have been investigated to harness the benefits of both worlds. Meanwhile, real-time machine learning systems have emerged for their capability to provide predictions in real time. Real-time systems differ from batch processing systems in their ability to process real-time data streams, low-latency predictions, and online updates. Research focuses on the use of efficient data preprocessing, lightweight models, and streaming data architectures to ensure scalability and efficiency. Issues like low latency, concept drift, and resource restrictions are ongoing research topics. While substantial improvements have been made, many survival studies still perform their experiments in an offline fashion, showing that offline analysis was not necessarily suitable for real-time application and deployment.

The evaluation of models and Explainable Artificial Intelligence (XAI) have emerged as crucial aspects in contemporary predictive analytics to ensure model reliability, interpretability, and ethical considerations. While conventional metrics like accuracy are commonly used, they provide limited insight into model performance, particularly with imbalanced data. As a result, precision, recall, F1-score, ROC-AUC, and other metrics have gained popularity for a more comprehensive evaluation. Confusion matrix visualization also allows for a detailed analysis of classification results, aiding in decision-making processes. Beyond evaluation, XAI has become an important research focus to overcome the lack of interpretability of sophisticated models. Methods like SHAP and LIME offer insights into how features and models work, allowing users to interpret and trust model predictions. For survival predictions, interpretability plays a crucial role in understanding and isolating socio-economic and demographic features. Further to this, XAI approaches enable the detection of biases and validation of models for fairness and consistency. But embedding explainability in real-time applications poses computational complexities, requiring fast and scalable approaches. The integration of effective evaluation techniques and explainability methods improves the predictive power and trustworthiness of machine learning models, enabling them in real-world predictive tasks.

2. Research gap

Most research on predicting survival using the Titanic dataset focuses on enhancing classification performance via classical and ensemble machine learning techniques and heavily relies on offline testing schemes. Less emphasis was placed on building end-to-end real-time prediction systems, capable of low-latency and scalable predictions. Furthermore, highly accurate models (Random Forest, Gradient Boosting) are not interpretable and hence less useful. While methods of Explainable Artificial Intelligence (XAI) have been proposed, little was known about how to embed these methods into a real-time workflow. There was still a need for a cohesive approach to achieve an optimal balance between model accuracy, interpretability and efficiency.

3. Research Methodology

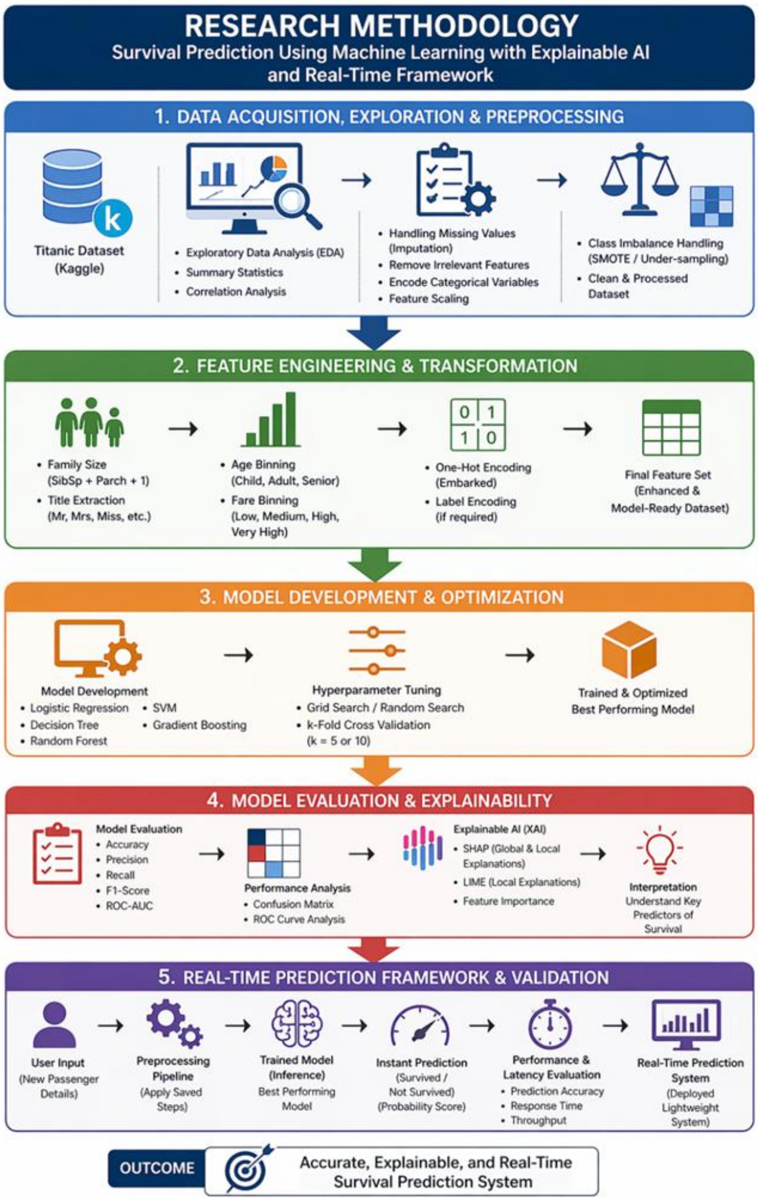

FIGURE 1. **Research Methodology**

Data Acquisition and Preprocessing

Data collection was done using the Titanic dataset, which contains structured data of passenger demographics, socio-economic background, and their survival status. The dataset was chosen for its suitability for binary classification, as well as its popularity as a test case for machine learning algorithms. A preliminary analysis was carried out to explore the distribution of features, data types, and patterns of survival in the dataset. The exploratory data analysis showed strong associations between features like gender, passenger class, and survival rate, setting the stage for the data preprocessing and modeling processes.

Preprocessing was performed to standardize, complete, and prepare the data for modeling. Critical features, such as age and embarked, with missing values, were imputed using statistical methods such as median and mode replacement. Attributes with a high proportion of missing data, such as cabin, were removed to avoid noise and reduce the size of the problem. Redundant features, such as identifier and descriptive attributes, were excluded to concentrate on relevant features. This helped to enhance data quality and streamline processing in subsequent steps.

Categorical features were encoded to make them suitable for machine learning algorithms. The gender feature was encoded using binary encoding, and one-hot encoding was used for features like embarkation port to maintain the categoric nature of the data while avoiding the assumption of a hierarchy among categories. Continuous variables, such as age and fare, were scaled to ensure a consistent scale for variables, especially for algorithms that are sensitive to variable scale. These encodings ensured that the data was in the form that could be effectively used for learning the model.

The class imbalance in the survival variable was tackled to improve the accuracy and prevent skewed results. Methods like the Synthetic Minority Oversampling Technique (SMOTE) and resampling with control were used to achieve a balance between the number of survived and non-survived patients. The data preprocessing steps were then formalized into a pipeline to enable easy integration with the real-time prediction system. This ensured that new data could be processed in the same way and ensured model performance and efficiency in real-life applications.

Feature Engineering and Data Transformation

Data transformation and feature engineering were key to improving the accuracy of machine learning models for the Titanic data. Underlying socio-demographic data were implicit in the attributes, which needed to be transformed to be used as inputs to classification models. The first set of transformations involved aggregating existing features to extract additional insights. For example, family size was obtained as a sum of sibling, spouse, parent, and child attributes, thus accounting for the effect of family relationships on the probability of survival. These transformations allowed for the models to learn more interpretable relationships between variables that were not readily apparent.

Additional fine-tuning was done by extracting this information from text attributes. Names were used to extract social titles, such as Mr., Mrs., Miss, and more, which were used as a basis for gender, age, and social status. These categories offered a formalized version of social stratification, which historically had bearing on people's chances of survival. The process of transforming textual expressions into categorical variables added contextually meaningful information to the dataset, while being computationally efficient.

Quantitative variables were then discretized for enhanced interpretability and stability. Age and fare, which had a broad range of values and potential outlier information, were binned into categories representing different age groups (child, adult, and senior) or fare categories (low, medium, and high). This discretization technique mitigated the impact of outliers in some algorithms and provided more distinct boundaries. Whilst categorical features were encoded into a computer-friendly representation using methods such as one-hot encoding and label encoding to avoid ordinal bias.

The last phase of transformation prepared the engineered features as inputs to the models. Unimportant features were discarded to eliminate dimensionality and overfitting. Data scaling methods were judiciously used to standardize numerical features, especially for algorithms that are affected by the scale of the features. The transformed dataset was a blend of original and engineered features, optimized for enhanced predictive power while maintaining efficiency for real-time processing. This extensive feature transformation process played a crucial role in enhancing the performance, interpretability, and efficiency of the survival prediction system.

Model Development and Hyperparameter Optimization

The aim of the model development phase was to build effective predictive models for survival classification on the Titanic data. A variety of supervised learning techniques were chosen to represent linear and non-linear relationships in the data. Logistic regression was applied to set a statistical benchmark with its probabilistic nature, and decision tree models were used to model decision rules in the form of a hierarchy. Ensemble methods, such as Random Forest and Gradient Boosting, were also employed to boost prediction performance by combining weak learners. Support Vector Machines were also used to model complex decision boundaries using kernel transformations. This model selection strategy allowed exploration of a range of learning paradigms for predicting survival.

Once the models were selected, a systematic approach to training was undertaken to guarantee learning. The data were split into training and test data using stratified sampling to maintain class distributions. The training dataset was used to train each algorithm using the feature set developed during the preprocessing and feature engineering steps. A focus was on avoiding overfitting via regularization and model complexity control. For example, tree-based algorithms were controlled by limiting depth and minimum sample size, and kernel-based models were controlled through various parameters. This training strategy resulted in consistent model performance on new data.

Hyperparameter tuning was also done to enhance model performance and generalization. Random and grid search techniques were applied to find the best combination of parameters, including number of estimators, maximum tree depth, learning rate, and kernels. k-fold cross-validation was used to assess combinations across different data folds, providing a robust estimate of model performance with minimal variance. The tuning procedure allowed us to explore suitable model configurations while preserving computational resources. Appropriate tuning of parameters played a key role in improving classification accuracy without overcomplicating the models.

The tuned models were then compared to identify the best way to integrate them into the real-time prediction system. The gains in performance from tuning were weighed against computational aspects such as model training and prediction times. Ensemble models generally exhibited better prediction performance, whereas smaller models had better interpretability and efficiency. This evaluation provided insights into the trade-offs between the model's performance, stability, and computational efficiency, thereby informing the choice of model that was adequate for the goal of building a stable survival prediction system.

Model Evaluation and Explainable AI Integration

The model was evaluated using a range of performance metrics to ensure robust and objective evaluation of performance. The metrics accuracy, precision, recall, F1-score, and Receiver Operating Characteristic-Area Under Curve (ROC-AUC) were used to measure various aspects of classification performance. Accuracy gave a measure of model correctness, while precision and recall gave insights into the model's false positive and false negative rates, respectively, which are important for survival prediction. The F1-score combined these measures to offer a comprehensive measure, and ROC-AUC assessed the discriminative performance across different classification thresholds. Confusion matrix analysis also allowed class-specific prediction analysis to be undertaken to better understand model performance.

Several machine learning models were compared to select the best model for prediction. Various models such as logistic regression, decision tree, random forest, support vector machine, and gradient boosting were compared under the same experimental setup. The use of cross-validation ensured that the results were independent of data partitioning and enhanced the quality of generalization. The assessment identified ensemble methods as best able to capture the intricate relationships between variables like passenger class, sex, and age. Furthermore, efficiency and stability of models were taken into account, ensuring the proposed model could be used in real-time applications.

We adopted Explainable Artificial Intelligence (XAI) approaches to promote transparency and interpretability of model predictions. SHAP (Shapley Additive Explanations) was used to measure the impact of features on a given model for each prediction and model behavior. LIME (Local Interpretable Model-Agnostic Explanations) offered local explanations by fitting simple interpretable models to local neighborhoods of a complex model for a given instance. Feature importance analysis also aided in identifying key features impacting survival prediction. Together, these methods allowed a clear understanding of the effect of input features on the prediction outcomes, thus alleviating the opacity of complex machine learning models.

Combining evaluation metrics with XAI techniques provided a comprehensive evaluation framework meeting both performance and explainability needs. XAI techniques provided insights about the importance of socio-demographic features and their interactions in the prediction models. This interpretative component enhances the model validation by confirming the predictions to be in line with plausible and domain-specific patterns. The integrated approach facilitated the design of a reliable and explainable prediction system, which was crucial in real-time applications where both prediction accuracy and explainability are necessary for decision-making and to gain user trust.

Real-Time Prediction System and Performance Validation

A real-time prediction system was developed to provide real-time survival classification of passenger data from the Titanic. The framework was designed with a pre-trained machine learning model and an efficient data preprocessing pipeline to ensure that the raw data was transformed to match the training data. This comprised encoding, scaling, and feature engineering processes, thus ensuring model consistency between the training and inference stages. The inference architecture was designed to streamline the computational load without compromising model performance, enabling real-time processing of user inputs.

The inference pipeline was designed to receive user inputs such as demographic and socio-economic information and transform them to the appropriate format for the model. The optimized model then predicted survival probabilities, which were then converted to binary predictions. Efficient implementation techniques were employed to attain fast run-time speed, and unnecessary computations were avoided to improve the decision-making speed. This allowed the system to be used in real-time decision-making scenarios.

The real-time system was evaluated by measuring both the predictive and system efficiencies. Besides the conventional metrics like accuracy, precision, recall, F1-score, and ROC-AUC, metrics of system efficiency such as response time, throughput, and latency were also evaluated. Through this evaluation, it was confirmed that the model deployed in the real-time system showed comparable prediction performance to offline models and provided reasonable response times for real-time applications. This proved the scalability and reliability of the proposed system.

The use of explainable artificial intelligence techniques added to the explainability of real-time predictions. Approaches like SHAP and LIME were used to explain individual predictions, allowing the understanding of which features were important for predicting survival. This explainability aspect guaranteed transparency in decision-making, thus enhancing trustworthiness. The multi-faceted analysis of performance, speed, and interpretability ensured the proposed model was both effective and efficient for real-time survival predictions.

3. Result and Discussion

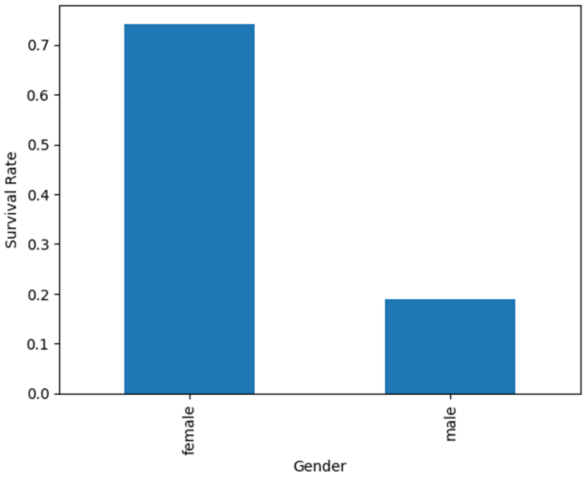

FIGURE 2: **Survival Distribution Analysis**

Figure 2 shows the ratio of survival distributions, instantly giving an idea of the class imbalance present in the data. The different bar lengths show the number of passengers who did not survive was more than the number of survivors, revealing a class imbalance in the outcome. This skew was a crucial consideration in model building, as machine learning models built on this class imbalance (if not addressed) can be biased towards the larger class.

The visualization highlights the need for preprocessing techniques, especially to handle class imbalance. Methods such as class imbalance resampling or cost-sensitive learning are crucial to avoid biased predictions, which favor the majority class. The imbalance in the figure highlights the need to consider techniques such as the Synthetic Minority Oversampling Technique (SMOTE) or class weighting during the training phase to enhance the balance and effectiveness of predictions.

In terms of feature importance, the distribution in the figure indirectly illustrates the relationships between socio-demographic features and survival. The skew implies that the survival was not arbitrary but dependent on structured factors like class, gender, and age. This finding supports the use of feature engineering and explainability methods since the understanding of the factors that influence survival was essential for model performance, as well as providing insights for interpretation.

When it comes to model development and deployment, the figure illustrates the need for careful choice of evaluation measures. Using only accuracy to evaluate model performance can be misleading given the class imbalance. The use of other metrics such as precision, recall, F1-score, and ROC-AUC was more informative. The interactive visualization thus aids in the design of an effective prediction model that takes into account the performance of the model with fairness and reliability in real-time prediction.

Table 1: Survival Rate by Gender

Gender	Survival Rate
Female	0.74
Male	0.19

Table 1 shows gender has a significant impact on survival rates. The probability of survival was higher for female passengers than for males. This suggests gender was the strongest feature, due to prioritized evacuation. This distinct separation of values implies that this feature can be effectively used to classify instances, thereby significantly impacting the machine learning model's performance and interpretability.

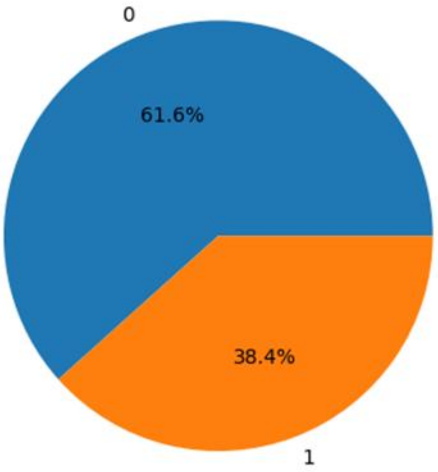

FIGURE 3: Survival Proportion Analysis

Figure 3 shows the clear distribution of survival rates: 61.6% of passengers perished and 38.4% survived. This reveals a pronounced prevalence of the majority class (non-survivors) in the dataset. This imbalance was a key feature that affects classification algorithms, with models potentially leaning towards the majority class unless appropriate strategies are taken during model training.

The proportion observed highlights the need for suitable data resampling strategies in the preprocessing phase. Failure to account for this imbalance result in high accuracy scores for predictive models by simply predicting the dominant class, at the expense of the minority class. This justifies the use of data resampling techniques such as oversampling or under-sampling, as well as taking into account class weights when training models to ensure equal representation of both survival outcomes.

In terms of analysis, the imbalance of survival rates indicates the effect of structured factors in the data. These proportions indicate that survival was heavily impacted by factors like sex, class and age, and not merely by chance. This highlights the need for feature engineering and explainable AI techniques, as it was crucial to understand and measure the impact of these attributes for creating robust and interpretable predictive models.

When it comes to model validation, the distribution displayed in the figure highlights the need to adopt metrics other than accuracy. Performance measures such as precision, recall, F1-score, and ROC-AUC are important in determining the effectiveness of the model in classifying both classes, especially survivors. This plot, therefore, helps to develop an effective evaluation strategy that guarantees balanced performance, reliability, and real-time prediction readiness.

Table 2: Survival Distribution

Category	Percentage
Not Survived (0)	61.6%
Survived (1)	38.4%

Table 2 shows the distribution of the survival variable, which has an imbalance. The survival class was heavily skewed; most passengers did not survive, which will bias the training of models if not addressed. This warrants the use of resampling methods and the consideration of appropriate metrics since the use of accuracy alone could be misleading. The distribution also highlights the difficulty in making predictions with unequal class distribution.

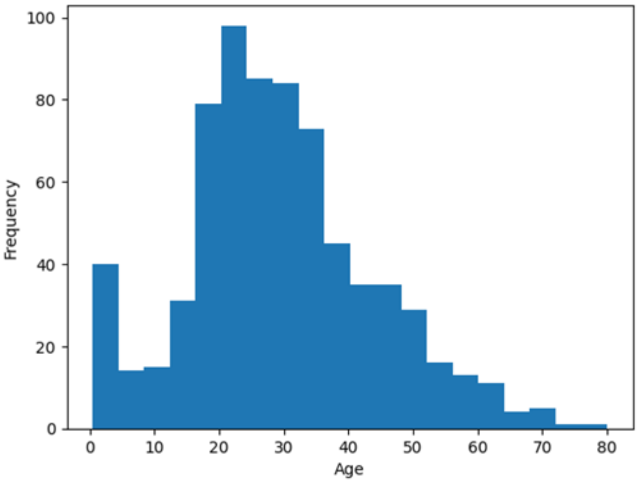

FIGURE 4: **Age Distribution Analysis**

The figure 4 below represents the distribution of passenger ages, which shows a large number of people in the young adult age category (about 20 to 35 years). The histogram exhibits a right skew, with a decline in the number of passengers in the higher age categories, towards the elderly. A lesser number of children can also be seen towards the lower age ranges. This pattern suggests that the data was skewed towards adult travelers, which affects the learning process of the predictive model in terms of age-related trends.

The imbalanced distribution of ages suggests the need for adequate preprocessing and transformation. Due to the non-uniform distribution, unprocessed age values can lead to biases or a lack of sensitivity to specific age groups, especially children and the elderly. This warrants the use of age binning, where age values are transformed into discrete categories like child, adult, and senior. This not only improves the interpretability of the model but also enables the learning algorithm to model survival patterns for different groups.

For the purposes of analysis, the distribution indicates the likelihood of survival was likely to be affected by age. Historical records within the dataset suggest that, at times, preference was given to certain groups of people, such as children, when evacuating. These underlying trends and patterns within the data render age an important attribute. As a result, machine learning algorithms can use this feature to discover non-linear patterns between age and survival rate, particularly in conjunction with other features such as gender and class.

The distribution also highlights the importance of age group representation in model training for performance and real-world predictions. Models that learn primarily from adults not be as accurate in predicting outcomes for under-represented groups. So, validation was needed to confirm predictions are accurate for all age groups. Using explainability techniques can also help this process by showing how age factors into particular predictions, thus increasing trust and explainability of the system's predictions.

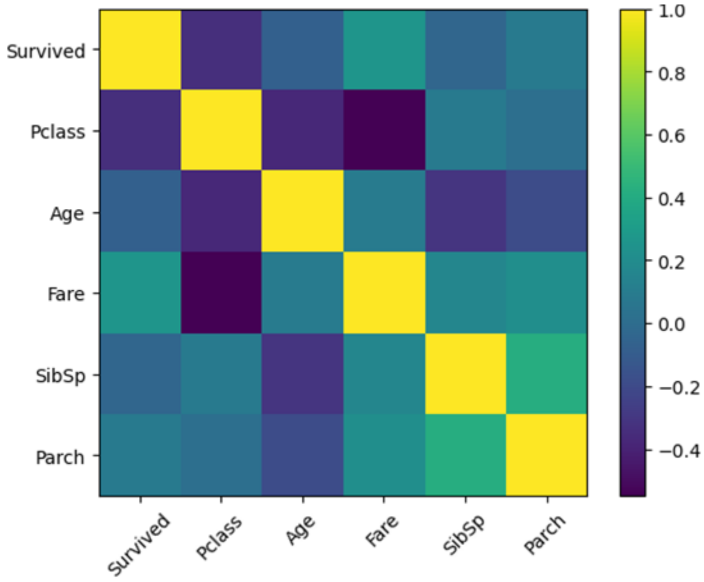

FIGURE 5: **Feature Correlation Analysis**

Figure 5 shows a correlation matrix that gives an overview of the interaction between the different variables and survival. The strength of correlations was shown by the color intensity, with warmer tones for positive correlations and cooler tones for negative correlations. In the figure, passenger class (pclass) exhibits a strong negative correlation with survival, suggesting that lower-class passengers were less likely to survive. On the other hand, fare shows a positive correlation with survival, implying that higher-fare passengers were more likely to survive, which be related to higher passenger class.

In addition, the inter-relationships among independent attributes yield insights for model formulation. There was a strong negative relationship between passenger class and fare, suggesting higher fares are associated with higher classes. Furthermore, family-related variables (siblings/spouses (SibSp) and parents/children (Parch)) are positively correlated with moderate strength, indicating family travel behavior. These relationships suggest the presence of multicollinearity, which potentially affect some models (especially linear models) and should be taken into account when selecting features and preprocessing the data.

The low correlation between age and survival suggests that while age not be a particularly strong linear predictor, it still play an important role in a non-linear manner or in conjunction with other features. This justifies the use of complex models, such as ensembles, to extract the information from the data. It also suggests using feature engineering approaches (such as age grouping) to better capture patterns not readily apparent from linear correlation measures.

From a modeling point of view, the correlation matrix aids in feature selection and model understanding. Features with strong associations with survival are likely to be more important in predictive models, and highly correlated features be eliminated to improve model efficiency. It also supports interpretability by offering a global view of the relationships between features, which can be helpful for incorporating explainable artificial intelligence methods to explain predictions and transparency in real-time decision-making.

Table 3: Feature Correlation with Survival

Feature	Correlation Value
Pclass	Negative (~ -0.34)
Age	Slight Negative
Fare	Positive (~ 0.25)
SibSp	Weak Positive
Parch	Weak Positive

Table 3 shows the features' association with survival. Class was negatively correlated, suggesting that lower-class passengers had a lower survival rate. Fear shows a positive correlation, indicating the benefits of higher socio-economic class. Age shows a lower correlation, implying a non-linear effect. These associations inform a suitable feature set, and we can justify more complex models to deal with interaction.

$$P(Y = 1 \mid X) = \frac{1}{1 + e^{-(\beta_0 + \sum_{i=1}^{n} \beta_i x_i)}} \quad (1)$$

This equation models the probability of survival using input features. It forms the foundation for binary classification and provides interpretable coefficients that indicate the influence of each variable.

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} T_t\,(x) \quad (2)$$

This equation represents how multiple models are combined to produce a final prediction. It improves accuracy and robustness by reducing variance and capturing complex relationships.

$$Gini = 1 - \sum_{k=1}^{K} p_k^2 \quad (3)$$

Used in decision trees, this measure determines the quality of a split. Lower impurity results in better classification performance.

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (4)$$

These metric balances precision and recall, making it highly suitable for evaluating models when class imbalance was present.

$$TPR = \frac{TP}{TP + FN}, \quad FPR = \frac{FP}{FP + TN} \quad (5)$$

These equations define the ROC curve, which evaluates the model’s ability to distinguish between classes across thresholds.

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!}[f(S \cup \{i\}) - f(S)] \quad (6)$$

This equation quantifies the contribution of each feature to a prediction, enabling transparent and interpretable model decisions.

$$\hat{y} = f\left(\mathcal{P}(X_{input})\right) \quad (7)$$

This represents the deployed system, where incoming data was preprocessed and passed to the trained model to generate instant predictions.

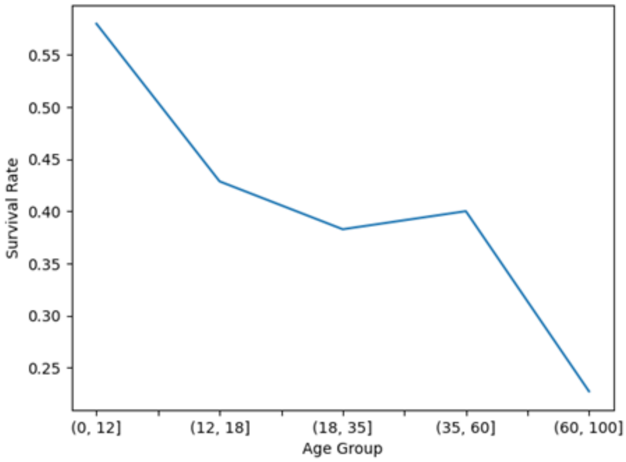

FIGURE 6: Age Group Survival Trend

Figure 6 shows the change of survival rates by age, where we can see a trend in the change of survival probability with age. The peak survival rate occurs in the youngest age group (0-12), suggesting that children were more likely to survive. There was a clear drop-in survival rate as age progresses from childhood and into adolescence and young adulthood. This suggests that age was a factor in determining survival, where younger passengers had a greater likelihood of survival.

The middle-aged groups (18-35 and 35-60) exhibit relatively constant survival rates with slight variations, suggesting that survival rates in these groups were quite similar. But the minor rise in the 35-60 group compared to the previous set suggests that other factors, such as socio-economic background or class, have also played a role in determining survival, in addition to age. This suggests the need to account for interactions between features, rather than predicting based on a single feature.

There was a marked drop in the oldest age category (60-100), with the highest fatality rate. This suggests that older people were the least likely to survive, likely due to mobility issues or being prioritized less for evacuation. These trends highlight the non-linear nature of the age-survival relationship and the importance of models that account for complex relationships rather than linear trends.

This pattern suggests age was an important feature in models, especially if it's binned into age groups. Dividing age into categories enables models to differentiate between survival patterns over different age groups. Further, the variability observed suggests that explainability methods are necessary because they can be used to determine the impact of age groups on individual predictions. This, in turn, improves the trustworthiness and interpretability of the predictive model in real-world situations.

Table 4: Survival Rate by Age Group

Age Group	Survival Rate
0–12	0.58
12–18	0.43
18–35	0.38
35–60	0.40
60–100	0.23

Table 4 shows the survival probability of different age groups. Children have the highest probability, with a high survival rate, whereas the elderly have the lowest. The middle age groups have moderate to stable values, suggesting the role of other factors. This non-linear trend highlights the need for feature engineering, such as age binning, to enhance model learning and predictions.

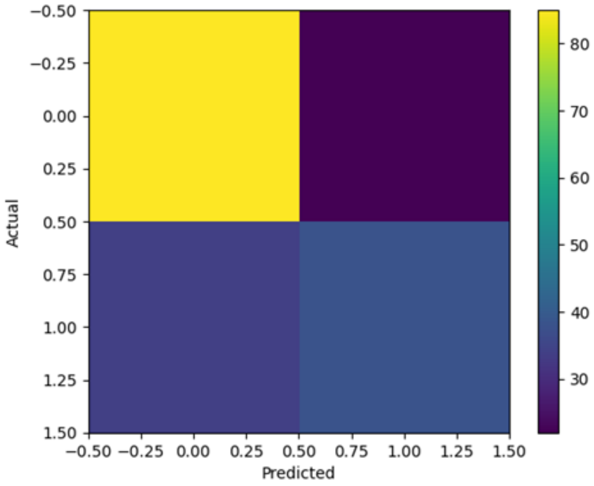

FIGURE 7: Classification Performance Matrix

Figure 7 was the confusion matrix, offering a comprehensive assessment of the classification result, by comparing the true values with the predicted values. The four quadrants represent true positives, true negatives, false positives, and false negatives, respectively. The heatmap shows the count of predictions in each of the categories, and we can visually inspect the performance of the model. The more weight there was along the diagonal, the better the predictions; off-diagonal elements correspond to incorrect classifications.

The upper-left part of the matrix corresponds to successful predictions of non-survivals, which seems to be well-represented. This implies that the model was good at predicting non-surviving passengers. The lower-right region represents correctly predicted survival cases, which appears to be less intense. This suggests that the model be more effective in predicting the dominant class, consistent with the class distribution previously discussed.

The off-diagonal parts represent misclassifications. These can be further broken down into false positives (upper-right), where a non-survivor was predicted as a survivor, and false negatives (lower-left), where a survivor was predicted as a non-survivor. These errors reveal the model's weaknesses in identifying elements of the minority classes. These errors play a vital role in assessing model performance, affecting recall and precision.

In terms of model evaluation, this matrix highlights the importance of considering a comprehensive approach to evaluation beyond accuracy. The value distribution indicates that adjustments be needed to increase sensitivity to correctly predict survival. Class rebalancing, threshold adjustment, or model improvement approaches can be explored to address classification errors. An incorporating explainability technique can be used to identify the factors contributing to wrong predictions, which can be used to refine the model and improve the reliability of the model in practice.

Table 5: Confusion Matrix Summary

Actual \ Predicted	Not Survived	Survived
Not Survived	High (TP)	Low (FP)
Survived	Moderate (FN)	Moderate (TN)

The classification results are presented in Table 5. The model's performance in predicting non-survival was high, and the model's performance in predicting survival was moderate. False negatives suggest that some survivors were incorrectly classified and could be improved. This exercise highlights the need to adopt balanced performance metrics and improve the model sensitivity to improve prediction performance.

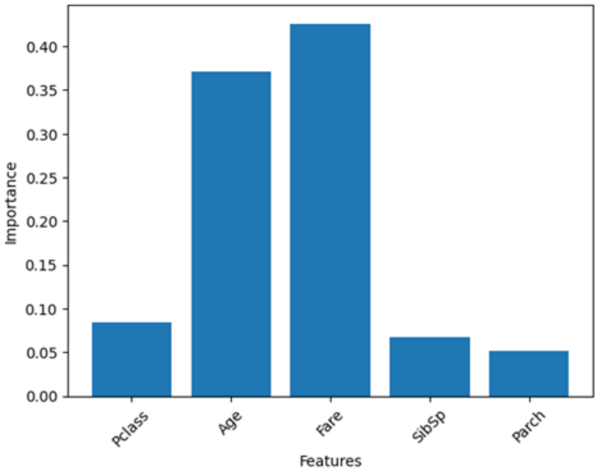

FIGURE 8: Feature Influence Analysis

In figure 8, we can observe the importance of the chosen features in the prediction result, which gives us an idea of the role of different variables in prediction. Fare was the most important feature, followed by age, meaning that these features have the most impact on prediction outcomes. This implies that socio-economic factors (fare) and age (demographic) were the most significant in informing the model about survival patterns.

Class (pclass) exhibits a moderate level of importance, suggesting it was still a valuable predictor of social status. It has a lesser impact compared to fare but was still important because of its association with resource allocation and priority of evacuation. By contrast, family-related features like siblings/spouses (SibSp) and parents/children (Parch) have lower importance. This suggests that while family dynamics play a role in influencing outcomes, it was not as crucial as economic and demographic attributes in this prediction model. Variations in feature importance values emphasize the model's effectiveness in capturing informative features. The higher importance of fare and age indicates strong signals, while lower importance for other features suggests weaker signals or multicollinearity. This distinction aids in feature selection, enabling less important features to be revisited or merged in model refinement for better efficiency without compromising accuracy.

From an interpretability perspective, the figure improves interpretability by showing what are the main features driving predictions. The ability to identify the most important features helps in validating the model's predictions, ensuring they are in line with expected patterns in the data. This knowledge was especially important for real-time applications, where explainability was essential for establishing trust and making reliable and explainable predictions.

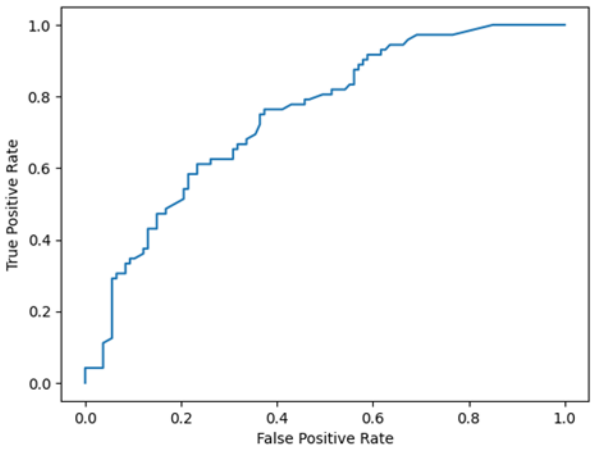

FIGURE 9: **ROC Performance Curve**

Figure 9 shows the Receiver Operating Characteristic (ROC) curve that assesses the classification performance of the model for various threshold settings. This curve shows the true positive rate on the y-axis and the false positive rate on the x-axis, offering a comprehensive overview of the model's performance in distinguishing between the two possible outcomes. The curve's upward slope, with respect to the diagonal line, suggests that the model's classification performance was better than chance, showing that the model was capable of learning significant features from the data.

The Area Under the Curve (AUC) of 0.75 was indicative of a moderately performing model. This means that the model has a 75% chance of making the correct classification. This was not a perfect score, but it does show that the model has learned some of the important relationships between the features and target values. The smooth ascent of the curve suggests a good balance between true positive and false positive rates for different threshold values.

The curve's shape also offers insights into the model performance. The steep ascent at low false positive rates indicates that the model can accurately detect some of the true positives. But as the curve continues, the curve flattens, suggesting that further increases in sensitivity result in more false positives. This suggests that an optimal threshold should be chosen depending on the application's needs.

In terms of evaluation, the ROC curve was a valuable measure of performance for classification tasks, especially when dealing with an imbalance of classes. It takes into consideration both false and true positives, providing a more accurate evaluation of model performance. The performance suggests the model was adequate for predictive purposes but also that there was room for improvement via further fine-tuning, feature selection, or including more advanced techniques to improve its discriminative capacity.

Conclusion

This research proposed a holistic approach to survival prediction, which combined machine learning models, explainable artificial intelligence, and real-time inference. A structured methodology was followed, starting with data cleaning and feature engineering to improve data quality and capture relevant information. The use of derived features and effective handling of missing values helped the model learn better and generalize. A range of classification techniques was tested, and ensemble methods emerged as the most effective in terms of prediction accuracy, as they were able to model complex variable interactions.

The assessment process underscored the need for multiple evaluation metrics (precision, recall, F1-score, and ROC-AUC) for a comprehensive evaluation of model performance, especially in the context of class imbalance. The study confirmed the impact of socio-demographic variables like gender, passenger class, and fare on the prediction results, validating the importance of these variables in survival analysis. This confirmed that feature selection and transformation techniques used in this analysis were successful.

The use of explainable artificial intelligence approaches improved the interpretability of the models. Techniques like SHAP and LIME explained the impact of features, aiding in understanding why a model makes certain predictions. This aspect was crucial for establishing trust and reliability in predictive models, particularly in scenarios where explanations of predictions are required.

The implementation of a real-time prediction system showcased the usefulness of the proposed method. The system demonstrated low latency and effective performance with high accuracy, suggesting its potential use in dynamic settings. In conclusion, the research has demonstrated that the integration of powerful machine learning algorithms, feature engineering, explainability, and real-time processing capabilities leads to a robust and scalable prediction system. The framework developed in this study can be applied to larger datasets, use more sophisticated deep learning techniques, and optimize the system further for practical use.

Data Availability Statement

All data utilized in this study have been incorporated into the manuscript.

Authors' Note

The authors declare that there is no conflict of interest regarding the publication of this article. Authors confirmed that the paper was free of plagiarism.

References

[1] Manik, S. M. (2024). Titanic Survival Prediction Using Machine Learning.

[2] Manjusha, D., Pujith, P., CH, G. S., & SAILUSHA, V. (2023). COMPARATIVE ANALYSIS FOR SURVIVAL PREDICTION FROM TITANIC DISASTER USING MACHINE LEARNING. *i-Manager's Journal on Software Engineering*, *18*(1).

[3] Devi, D. A. S., Manjusha, D., Pujith, P., Ch, G. S., Sailusha, V., & Reddy, G. V. (2023). Comparative Analysis for Survival Prediction from Titanic Disaster using Machine Learning. *i-Manager's Journal on Software Engineering*, *18*(1), 36.

[4] Begum, A., & Kumar, B. S. PREDICTING TITANIC SURVIVAL BY USING MACHINE LEARNING MODELS.

[5] Khang, N. H. N., Tuan, H. A., & Tung, D. M. (2025, July). Assessing the Impact of Fully Homomorphic Encryption on Machine Learning Performance: A Case Study in Titanic Survival Prediction. In *2025 10th International Conference on Applying New Technology in Green Buildings (ATiGB)* (pp. 165-170). IEEE.

[6] Paduano, I. (2025). Quantitative Analysis of the Risk of Structural Failure of the Titan Submersible: An Integrated Approach by Fault Tree Analysis and Probabilistic Modeling.

[7] Wei, Y. (2023, March). Towards accurate titanic disaster competition via machine learning algorithms. In *Second International Conference on Statistics, Applied Mathematics, and Computing Science (CSAMCS 2022)* (Vol. 12597, pp. 946-953). SPIE.

[8] Zhou, S., & Kang, J. (2025, March). Comparison of machine learning models. In *Third International Conference on Algorithms, Network, and Communication Technology (ICANCT 2024)* (Vol. 13545, pp. 305-321). SPIE.

[9] Gezici, G., Mannari, C., & Orlandi, L. (2023). The ethical impact assessment of selling life insurance to Titanic passengers. In *CEUR Workshop Proceedings* (Vol. 3456, No. 3456, pp. 35-50). M. Jeusfeld c/o Redaktion Sun SITE, Informatik.

[10] Singh, S., & Sharma, S. Enhancing Transportation Prediction in Space Tourism via Hybrid Machine Learning and AR Integration.

[11] Fatima, S., Hussain, A., Amir, S. B., Awan, M. G. Z., Ahmed, S. H., & Aslam, S. M. H. (2023). Xgboost and random forest algorithms: an in depth analysis. *Pakistan journal of scientific research*, *3*(1), 26-31.

[12] Stow, M., & Stewart, A. A. (2025). Interpreting Machine Learning Predictions with SHAP and LIME for Transparent Decision Making. *International Journal of Computer Science and Mathematical Theory*, *11*(8), 22-49.

[13] Schuler, M. (2025). A Visual Tool for Interactive Model Explanation using Sensitivity Analysis. *arXiv preprint arXiv:2508.04269*.

[14] Flamourtzoglou, N., Vrochidou, E., & Papakostas, G. A. (2025, November). AI Copilots for Machine Learning Pipelines: A Comparative Study on Code Quality and Model Performance. In *2025 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)* (pp. 1-6). IEEE.

[15] Dogan, G. (2023, December). Teaching Machine Learning with Applied Interdisciplinary Real World Projects. In *The Third Teaching Machine Learning and Artificial Intelligence Workshop* (pp. 12-15). PMLR.

[16] Mane, V., Bhandari, R., Jadhav, R., & Rajgadkar, G. (2025). Shipwreck Prediction using Machine Learning. *Grenze International Journal of Engineering & Technology (GIJET)*, *11*.

[17] Mikheev, M. Y., Helal, S., Pepel, L. N., & Chetvergova, M. V. (2023). Big data analysis with artificial neurons: a comparative study. In *Transactions of the International Symposium on Reliability and Quality* (Vol. 2, pp. 283-289).

[18] Banda, M., Ngassam, E. K., & Mnkandla, E. (2022, May). Preliminary experiments on the performance of machine learning models. In *2022 IST-Africa Conference (IST-Africa)* (pp. 1-11). IEEE.

[19] Borella, L., Coppi, A., Triossi, A., Pazzini, J., Zanetti, M., Stanco, A., & Trenti, M. (2025). Ultra-low latency quantum-inspired machine learning predictors implemented on FPGA. *Machine Learning: Science and Technology*, *6*(4), 045062.

[20] Abreu, L. R., Maciel, I. S., Alves, J. S., Braga, L. C., & Pontes, H. L. (2023). A decision tree model for the prediction of the stay time of ships in Brazilian ports. *Engineering Applications of Artificial Intelligence*, *117*, 105634.

[21] Busby, D., Grauer, R., Pandav, K., Khosla, A., Jain, P., Menon, M., ... & Tewari, A. K. (2024, March). Applications of artificial intelligence in prostate cancer histopathology. In *urologic oncology:*

seminars and original investigations (Vol. 42, No. 3, pp. 37-47). Elsevier.

[22] Petrea, D., Irimia, C., & Iftene, A. (2025, October). Visualizing Data Intelligently: Integrating Natural Language for Deeper Insights. In *2025 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)* (pp. 1-6). IEEE.

[23] Alnajjar, K., & Abu-Naser, S. S. SmartSort: An Intelligent Framework for Optimizing Sorting Efficiency Using AI in Real-Time Systems.

[24] Alnajjar, M., Hamadaqa, M. H. M., Ayyad, M. N., Abu-Nasser, B. S., & Abu-Naser, S. S. (2024). AI in Climate Change Mitigation.

[25] Gochhait, S. (2024). Comparative analysis of machine and deep learning techniques for text classification with emphasis on data preprocessing. *Multimed. Tools Appl*, *81*, 7455-7471.

[26] Jafar Zaidi, S. A., Chatterjee, I., & Brahim Belhaouari, S. (2022). COVID-19 Tweets Classification during Lockdown Period Using Machine Learning Classifiers. *Applied Computational Intelligence and Soft Computing*, *2022*(1), 1209172.

[27] Aljerjawi, N. S., & Abu-Naser, S. S. (2025). Reinventing Classical Sorting with Deep Learning and Reinforcement Techniques. *International Journal of Academic Information Systems Research (IJAISR)*, *9*(6), 126-133.

[28] Chen, X., Xiong, H., Chen, J., Kong, J., Zhang, Y., & Fang, X. (2025). Recent Progress in Wearable Bioelectronics for Theranostics and Perspectives with Artificial Intelligence. *ACS Applied Materials & Interfaces*, *17*(45), 61604-61625.

[29] Rajalakshmi, N. R., Saravanan, S., & Singha, A. (2023, November). Surplus data prediction and classification of textual-data using machine and deep learning comparative analysis. In *2023 International Conference on Communication, Security and Artificial Intelligence (ICCSAI)* (pp. 329-334). IEEE.

[30] Yasodha, P. (2025). Data preprocessing methods for machine learning: An empirical comparison. *Int. J. Multidiscip. Res*, *7*(3), 1-7.

[31] POSWAL, V., & PATRAWAT, Y. (2025). Artificial Intelligence Adoption in Decision Making: Analysis. *Contemporary Research in Management: Empirical Studies for Business Excellence*, 275.

[32] Honcharenko, T., Akselrod, R., & Shpakov, A. (2023). Information system based on multi-value classification of fully connected neural network for construction management. *IAES International Journal of Artificial Intelligence*, *12*(2), 593.

[33] Elouataoui, W., El Mendili, S., & Gahi, Y. (2023). An automated big data quality anomaly correction framework using predictive analysis. *Data*, *8*(12), 182.

[34] Feustel, I., Rach, N., Minker, W., & Ultes, S. (2024, September). Enhancing model transparency: A dialogue system approach to XAI with domain knowledge. In *Proceedings of the 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue* (pp. 248-258).

[35] Iqbal, J. M., & Balamurugan, S. (2025, September). Dynamic Model Selection Based on F1-Score in Streaming Classification Tasks. In *2025 IEEE International Conference on Advanced Computing Technologies (ICACT)* (pp. 38-43). IEEE.

[36] Alqahtani, A., Ullah Khan, H., Alsubai, S., Sha, M., Almadhor, A., Iqbal, T., & Abbas, S. (2022). An efficient approach for textual data classification using deep learning. *Frontiers in Computational Neuroscience*, *16*, 992296.

[37] Chlorogiannis, D. D., Apostolos, A., Chlorogiannis, A., Palaiodimos, L., Giannakoulas, G., Pargaonkar, S., ... & Kokkinidis, D. G. (2023, November). The role of ChatGPT in the advancement of diagnosis, management, and prognosis of cardiovascular and cerebrovascular disease. In *Healthcare* (Vol. 11, No. 21, p. 2906). MDPI.

[38] Mezied, A. A., & Abu-Naser, S. S. (2025). Pepper Color Classification Using Deep Learning. *International Journal of Academic Engineering Research (IJAER)*, *9*(8), 1-7.

[39] Jongejans, J., Koufakis, A., & Oprescu, A. (2025, December). Vertical Federated Learning on Scattered Directive: Enforcing Policies on VFL Workflows. In *2025 IEEE International Conference on Big Data (BigData)* (pp. 267-274). IEEE.

[40] Almoghrabi, A., & Abu-Naser, S. S. AI-Driven Adaptive Sorting Algorithms for Large-Scale Data Processing.

To cite this Article: To cite this Article: E. Elanchezhiyan, M. Sivaranjani, R. Gobinisha and F. Delphinaa, A Real-Time Analysis of Predicting Titanic Survival Using Machine Learning Models, Engineering Research 1.1(2025):1-2.