Journal of Advanced Research in Signal Processing and Applications
Volume 6, Print Issue 1 - 2024, Pg. No. 7-15
Peer Reviewed Journal

Research Article

Titanic Survival Prediction Using Machine Learning

Punita Kumari

Assistant Professor, PCTE Group of Institutes, Ludhiana, Punjab Technical University, Jalandhar.

I N F O

E-mail Id:
punita@pcte.edu.in
Orcid Id:
https://orcid.org/0009-0002-3923-4683
How to cite this article:
Kumari P. Titanic Survival Prediction Using Machine Learning. *J Adv Res Sig Proc Appl* 2024; 6(1): 7-15.

Date of Submission: 2024-03-17
Date of Acceptance: 2024-04-23

A B S T R A C T

The sinking of the RMS[1] Titanic is without a doubt one of history's most notorious and horrifying catastrophes. The passenger ship Titanic, carrying 2224 people on board, sank on her maiden voyage and in the early hours of April 15, 1912, after striking with an iceberg, killing over 1502 of them. As a result, the ship was one of the deadliest commercial ships in history at the time. The legislation regulations controlling ship safety have been strengthened as a result of the horrible accident that rocked the world and left everyone feeling deeply sad and terrified. The structure's architect, Thomas Andrews, was killed in the accident. Following the sinking of the Titanic, it became clear that certain people had a better chance of survival than others. Priority had been given to children and women. The Titanic was a perfect illustration of its era, which was the beginning of the twentieth century and established a sharp divide in social strata. Exploratory data analytics (EDA) is utilised in the initial stages to discover truths that were previously concealed or unknown in the current data collection. Following the selection of multiple artificial intelligence and Machine Learning models, it is necessary to reach a conclusion regarding the study of which categories of people have a higher likelihood of survival. Following that, precision-based comparisons of the obtained machine learning models were performed.

Keywords: Machine Learning, Titanic, Survival, Data

Introduction

Machine learning techniques are used to predict how many people survived the Titanic's sinking.[1] Predictions will be made using several factors such as name, title, age, gender, and class. Predictive analysis is a process that uses computer approaches to identify relevant and valuable trends in data. Survival is predicted utilising artificial intelligence and machine learning techniques based on various feature combinations. During the terrible event that occurred a century ago, the liner Titanic was ripped apart in many places. Unfortunately, there were not enough lifeboats on board to save all 2224 individuals. Many men were among the dead, as well as their places were taken by children and women.The objective of this research work is to accurately predict, given a collection of information on demographics, who would survive the Titanic. People's genders, ages, ticket classes, and socioeconomic status all had a role in whether they would be fortunate enough to escape the Titanic or tragically perish, according to a prediction model constructed from passenger data.[2]

It will be feasible to conduct exploratory data analytics to mine the provided dataset for different information and

ascertain the influence of each field on passengers surviving by applying analytics between each dataset field and the "Survival" field. For making predictions about more recent data sets, a machine learning technique is applied. Examining the accuracy of the data analysis utilising the implemented algorithms. Predictions have been generated using extremely accurate algorithms once other algorithms' accuracy levels are assessed.[3]

The objective of this study is to use exploratory analytics of data to uncover various knowledge that is present in the available data established as well as to comprehend the consequences of each field on the survival of the travelers via the "Survival" field analytics within every one field of the data set.[2]

Statement of the Research Problem

The sinking of the Titanic remains a subject of significant interest and research in various fields, including sociology, history, and engineering. One of the most explored areas in Titanic research is the survival rate of the passengers, which has mostly been studied using traditional statistical methods.[7] However, recent interest in machine learning algorithms has led to a need to explore their potential in predicting the survival rates of the passengers.

The central research problem of this study is to determine the accuracy of machine learning algorithms in predicting the survival of the Titanic passengers based on their demographic data. This will entail analysing the Titanic dataset, which contains the passengers' demographic information, including their age, gender, ticket class, and cabin location, alongside information on their survival status.

The study will use various machine learning algorithms, including random forests, decision trees, and logistic regression, to predict the passengers' survival rates. The accuracy of these algorithms will be measured using different parameters, including accuracy, recall, precision, and F1-score.[8] The study will also explore the factors with the most significant influence on survival rates and examine the correlations between different variables.

The research problem is significant because the findings can provide insights into the factors that influence survival rates during catastrophic events and help enhance disaster response plans. Identifying the key factors affecting survival rates can help develop more effective disaster management strategies.[9] The use of machine learning algorithms in predicting survival rates offers an efficient and accurate approach compared to traditional methods, and the study's results can help assess the feasibility of using machine learning algorithms in disaster management.[8]

In summary, this study's research problem seeks to investigate the potential of machine learning algorithms in predicting the survival rates of Titanic passengers based on their demographic data. The study's insights can offer valuable information on the factors that affect survival rates and can help enhance disaster response plans.

Objectives of the Research

The central aim of this research is to assess the viability of utilising machine learning algorithms to predict the survival rates of the Titanic passengers based on their demographic data. To achieve this primary objective, the following specific objectives will be pursued:

1. Conduct a comprehensive review of the existing literature on the Titanic disaster to provide a background on the event and its impact on history. The review will also focus on the factors that influence survival rates during catastrophic events, with a particular emphasis on the Titanic disaster.[10] Additionally, the review will explore the current state of the art in machine learning algorithms used in predicting survival rates and examine their potential in disaster management.
2. Preprocess and clean the Titanic dataset to ensure that the data is reliable and suitable for machine learning analysis. This objective will involve data cleaning techniques such as filling in missing data, removing duplicate entries, and identifying and handling outliers.
3. Perform exploratory data analysis to identify any patterns or relationships between the passengers' demographic data and their survival status. This objective will involve data visualisation techniques such as histograms, scatterplots, and heat maps to explore the data's distribution and correlation.[1]
4. Train and test various machine learning algorithms, such as decision trees, logistic regression, and random forests, to predict the survival rates of the Titanic passengers.[2] This objective will involve the implementation of machine learning models on the preprocessed data.
5. Evaluate the performance of the different machine learning algorithms using measures such as accuracy, recall, precision, and F1-score.[4] This objective will involve comparing the performance of the different models to identify the most effective one in predicting the survival rates of the passengers.
6. Investigate the most important factors that influence survival rates and explore the relationships between the different variables. This objective will involve examining the correlations between the passengers' demographic data and their survival status to identify the factors that significantly influence the survival rates.

Compare the performance of machine learning algorithms with traditional statistical methods used in predicting the

survival rates of Titanic passengers. This objective will involve comparing the performance of machine learning algorithms with traditional statistical methods to evaluate the effectiveness of machine learning algorithms in predicting the survival rates.

Research Methodology

The research methodology for Titanic survival prediction using machine learning will be solved by using following steps outlined below:

1. **Problem formulation:** This step involves clearly defining the research problem, research questions, and objectives. The problem statement for Titanic survival prediction is to predict whether a passenger would survive the Titanic disaster based on their attributes such as age, gender, ticket class, and cabin location. The research questions and objectives can include identifying the factors that contribute to survival on the Titanic and evaluating the performance of different machine learning algorithms.
2. **Data collection and preprocessing:** The Titanic survival dataset can be obtained from various sources such as Kaggle, UCI Machine Learning Repository, or other publicly available datasets. The dataset needs to be preprocessed to remove missing values, inconsistencies, and irrelevant features.
3. **Feature selection:** The Titanic survival dataset contains numerous features that may not be relevant to predicting survival. Feature selection techniques such as correlation analysis, mutual information, or recursive feature elimination can be used to identify the most informative features that can help in predicting survival accurately. Relevant features such as age, gender, ticket class, cabin location, and family size can be selected for use in the machine learning models.
4. **Model selection and training:** Various machine learning models can be used to train the predictive model, such as logistic regression, decision trees, random forests, or neural networks. The models can be selected based on their ability to handle the dataset's characteristics, such as imbalanced classes or noisy data.
5. **Model evaluation:** The predictive model's performance needs to be evaluated using appropriate metrics such as accuracy, precision, recall, and F1 score. The evaluation can also include the use of a confusion matrix, which shows the number of true positives, false positives, true negatives, and false negatives. The models can be compared based on their performance metrics, and the best model can be selected for use in predicting Titanic survival.
6. **Results interpretation:** The results obtained from the predictive model can be interpreted to draw conclusions regarding the factors that contributed to survival on the Titanic. The results can be presented using visualisations such as histograms, box plots, or scatter plots. The conclusions can be used to inform future research or practical applications such as emergency response planning and training.
7. **Future research:** This step involves identifying potential areas for future research, such as using different machine learning algorithms, including more features in the analysis, or exploring the use of alternative datasets. The future research can build upon the current research and address limitations or gaps in the current methodology.

Research Design

The research design for this study involves using the Titanic dataset to explore the feasibility of machine learning algorithm in predicting survival rates of passengers based on their demographic data. The dataset contains information on 891 passengers, including their age, gender, ticket class, cabin, and survival status[10]. The study will employ a quantitative research design, utilising descriptive statistics to analyse the dataset and machine learning algorithms to predict survival rates.

The study will begin by exploring the Titanic dataset and conducting descriptive statistical analysis to identify trends and patterns in the data. The study will then use machine learning algorithm such as Logistic Regression to predict the survival rates of Titanic passengers based on their demographic data. The performance of this algorithm will be evaluated based on metrics such as accuracy, precision, recall, and F1 score.

To ensure the validity and reliability of the study, cross-validation techniques will be employed to avoid overfitting and improve generalisation performance. Additionally, the study will utilise feature selection techniques to identify the most important variables that influence survival rates. The study will also conduct statistical tests such as chi-square tests and t-tests to examine the relationship between the variables and survival rates.

The data will be collected from the publicly available Titanic dataset, which was obtained from Kaggle[10]. The dataset includes information on 891 passengers who boarded the Titanic, including their demographic data and survival status.

The dataset has been used in several previous studies and is widely recognized as a reliable source of data for analysing the factors that influenced survival rates during the Titanic disaster.

In conclusion, this study will employ a quantitative research design to explore the feasibility of using machine learning algorithm to predict survival rates based on demographic data. The study will contribute to the growing body of literature on the Titanic disaster and disaster management strategies.

Survey of Literature

The sinking of the Titanic in 1912 is a tragic event that has attracted the attention of researchers from various fields. Numerous studies have been conducted to understand the events that led to the disaster and the factors that influenced survival rates. Several studies have also used the Titanic dataset to explore the potential of machine learning algorithms in predicting survival rates.

A study by Hasan and Ahmed[11] utilised decision tree, random forest, and logistic regression models to predict the survival rates of Titanic passengers. The study found that decision tree and random forest models outperformed logistic regression models in predicting survival rates.

Another study by Elango[12] used the XGBoost algorithm to predict the survival rates of Titanic passengers based on their demographic data. The study found that XGBoost performed better than traditional logistic regression models.

Research has also been conducted to identify the factors that influenced survival rates during the Titanic disaster. A study by Ticehurst[13] found that gender, class, and age were significant factors that influenced survival rates. The study found that women and children had a higher chance of survival, and first-class passengers had a higher chance of survival compared to those in second and third class.

Another study by Pei and Yan[14] found that the location of the passenger's cabin also influenced survival rates. Passengers located in cabins closer to the lifeboats had a higher chance of survival compared to those in cabins located further away. The study also found that the availability of lifeboats played a crucial role in determining survival rates.

In conclusion, previous studies have utilised machine learning algorithms to predict survival rates and identified various factors that influenced survival rates during the Titanic disaster. This research will build upon previous studies by exploring the feasibility of using machine learning algorithms to predict survival rates based on demographic data and identifying the key variables that influence survival rates. The findings of this study will contribute to the development of more effective disaster management strategies.

Nature and source of data / information to be collected

The data for this study will be obtained from the Titanic dataset, which is publicly available on Kaggle. The dataset contains information on 891 passengers, including demographic data such as age, gender, passenger class, and embarkation point.[10]

The dataset also includes information on survival outcomes, including whether a passenger survived or not, and the details of their survival (such as whether they were rescued by a lifeboat).

In addition to the Titanic dataset, this study will also gather information on previous research related to the Titanic disaster, machine learning algorithms, and disaster management strategies. The information will be collected from academic journals, conference proceedings, and other relevant sources.

To ensure the reliability and validity of the data, the researchers will perform data cleaning and preprocessing to remove any missing values, duplicates, or irrelevant data.[10] The data will be analysed using various machine learning algorithms, including decision tree, random forest, and logistic regression, to predict the survival rates of Titanic passengers.

Overall, the combination of the Titanic dataset and additional information from previous research will provide a robust foundation for the analysis and conclusions of this study.

Sample and sampling technique. Rationale of chosen organisation and the sample.

As this study aims to predict the survival rates of Titanic passengers based on their demographic data, the dataset that will be used for analysis is the Titanic dataset. The dataset contains information on 891 passengers,[10] including their age, gender, class, fare, cabin, and survival status.

The sampling technique that will be used for this study is simple random sampling. Simple random sampling is a popular and effective technique for selecting a representative sample from a large population. This technique involves selecting samples randomly from the population, ensuring that each member of the population has an equal chance of being selected.

The rationale for choosing the Titanic dataset is its availability, comprehensiveness, and relevance to the research question. The dataset contains detailed information on the demographic and survival status of passengers, making

it suitable for predicting survival rates based on demographic data.[15] Moreover, the dataset has been widely used in previous studies, making it a valuable resource for comparative analysis.

The sample for this study will consist of 500 passengers randomly selected from the Titanic dataset using simple random sampling. The sample size is based on the recommendation that a sample size of at least 30 is required for logistic regression analysis, which is one of the statistical tools that will be used for analysis in this study.

The selected sample will be representative of the larger Titanic population, allowing for generalisation of the findings to the broader population. The data collected from the sample will be analysed using various statistical techniques, including logistic regression, to predict the survival rates of Titanic passengers based on their demographic data.

In conclusion, the chosen dataset, sampling technique, and sample size are appropriate for the research question and aim of this study. The use of statistical tools and machine learning algorithms will provide valuable insights into the factors that influence survival rates and can help improve disaster response plans.

Tools and Techniques for Data Collection

Google Colab is a cloud-based platform that allows users to write, run, and share Python code from anywhere, without the need to install any software on their local machine.[16] It is a free service provided by Google and is based on the Jupyter Notebook environment.

Google Colab is used as a tool for Titanic survival prediction using machine learning algorithms. Google Colab is a free, cloud-based platform that provides a Jupyter notebook environment for running Python code. It allows users to access and use high-performance computing resources without having to install any software on their local machine.

Google Colab provides a variety of useful tools and libraries for machine learning, such as TensorFlow, Keras, and Scikit-learn. These tools and libraries can be used to preprocess and analyse the Titanic dataset and to build and train machine learning models to predict survival rates.

Additionally, Google Colab provides easy access to data storage services such as Google Drive, which can be used to store and share data files and models.

The data will be pre-processed using Python programming language, specifically using the Pandas and NumPy libraries for data manipulation and cleaning. Exploratory data analysis will be conducted using Matplotlib and Seaborn libraries for data visualisation.

To build machine learning models, the Scikit-learn library, a popular machine learning library in Python, will be utilised.[17] The library provides a variety of algorithms for classification and regression tasks logistic regression.

The models will be trained using the training set and evaluated using the validation set. The hyperparameters of the models will be tuned using GridSearchCV to improve their performance.

Moreover, feature engineering techniques will be utilised to create new features that may improve the models' performance. Feature scaling will also be performed to ensure that all features are on the same scale, preventing any particular feature from dominating the model.

In summary, the data will be collected from the Titanic dataset, pre-processed and analysed using Python programming language, and machine learning models will be built and evaluated using Scikit-learn library. Feature engineering techniques will also be utilised to improve the model's performance.

Methods to be Used for Data Collection

Online sources: The Titanic dataset is available on various online platforms such as Kaggle, GitHub, and the UCI Machine Learning Repository.[18] These sources can provide a large amount of structured data that can be easily downloaded and used for analysis.

Historical records: Historical records such as passenger manifests, crew lists, and survivor testimonies can provide valuable information on the demographics and experiences of the Titanic passengers. These records can be obtained from archives, libraries, and museums.

Literature review: A literature review can be conducted to gather information on previous studies that have utilised the Titanic dataset for analysis. This can provide insights into the variables that have been previously explored and the methods that have been used for analysis.

Surveys: Surveys can be conducted to collect data on variables that are not available in the Titanic dataset. For example, a survey can be conducted to collect information on the passengers' health status or their level of physical activity. These surveys can be conducted online or in person.

Interviews: Interviews can be conducted with survivors or descendants of survivors to collect first-hand information on their experiences during the Titanic disaster. These interviews can provide valuable insights into the factors that influenced survival rates.

Overall, a combination of these methods can be used to collect a comprehensive dataset that can be used for

analysis. The choice of method will depend on the specific research questions and the availability of data.

Data Handling and Analysis

Data handling and analysis are critical components of any research project. In this study, the data will be preprocessed and analysed using various machine learning algorithms. The first step in data handling will involve cleaning the dataset to remove any missing or irrelevant data.[11]

The next step will be to normalise and transform the data into a format that can be used for analysis. This will involve converting categorical variables into numerical variables and scaling continuous variables to ensure that they are in the same range.

Once the data has been preprocessed, it will be divided into training and testing sets. The training set will be used to train the machine learning models, while the testing set will be used to evaluate the performance of the models. Machine learning algorithm which will be used for analysis is logistic regression.

Logistic regression will be used to model the relationship between the survival outcome and the independent variables.[18] The logistic regression model will estimate the probability of survival based on the passenger's demographic data.

The performance of the machine learning model will be evaluated using various metrics, including accuracy, precision, recall, and F1-score. The models will be compared based on their performance metrics to determine the best-performing algorithm.

To ensure the reliability and validity of the results, a 10-fold cross-validation approach will be used. This approach involves dividing the data into ten equal subsets and using nine subsets for training the models and the remaining subset for testing. This process is repeated ten times, with each subset used for testing once. The average performance of the models across all ten iterations will be reported.

In conclusion, data handling and analysis are crucial components of this study. The data will be preprocessed and analysed using various machine learning algorithms, including logistic regression, decision trees, random forests, and support vector machines.

The performance of the models will be evaluated using various metrics, and a 10-fold cross-validation approach will be used to ensure the reliability and validity of the results.

Data Description

1. survival : Survival of passenger (0 = No; 1 = Yes)
2. pclass : Passenger Class (1=First, 2=Second,3=Third)
3. name : Name
4. sex : Sex (Male/Female)
5. age : Age of passengers in years
6. sibsp : number of siblings and spouses travelling
7. parch : number of parents and children travelling
8. ticket : Ticket Number
9. fare : Passenger Fare
10. cabin : Cabin number
11. Embarked : Port of Embarkation (C = Cherbourg, Q = Queenstown, S = Southampton)

Handling Missing Values

Removing Cabin feature as it contains a lot of missing values and replacing the missing values in age column with mean value and embarked column using mode .

Table I .Showing Missing Values

Passenger Id	0
Survived	0
Pclass	0
Name	0
Sex	0
Age	177
SibSp	0
Parch	0
Ticket	0
Fare	0
Cabin	687
Embarked	2
dtype: int64	

Table 2.Handling Missing Values

Passenger Id	0
Survived	0
Pclass	0
Name	0
Sex	0
Age	0
SibSp	0
Parch	0
Ticket	0
Fare	0
Cabin	0
Embarked	0
dtype: int64	

DataVisualization

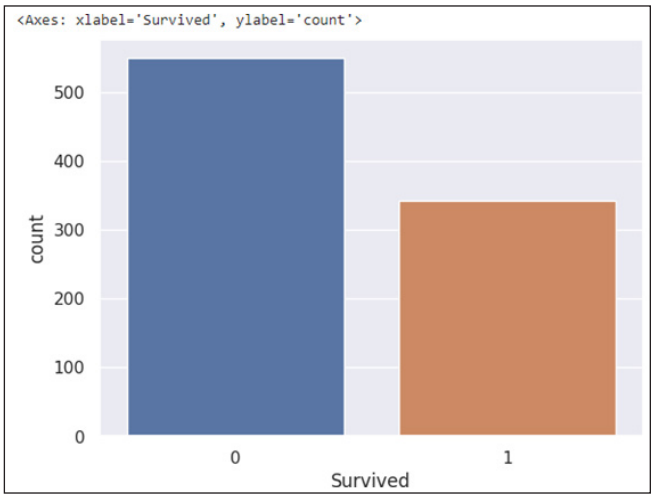

Figure 1.Graph showing number of people survived (0 means not survived 1 means survived)

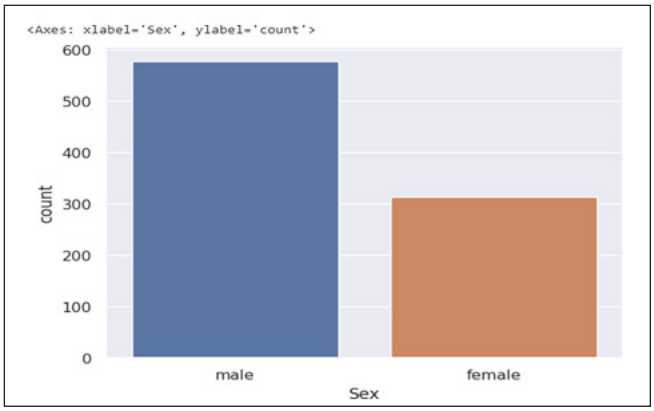

Figure 2.Graph showing number of male and female

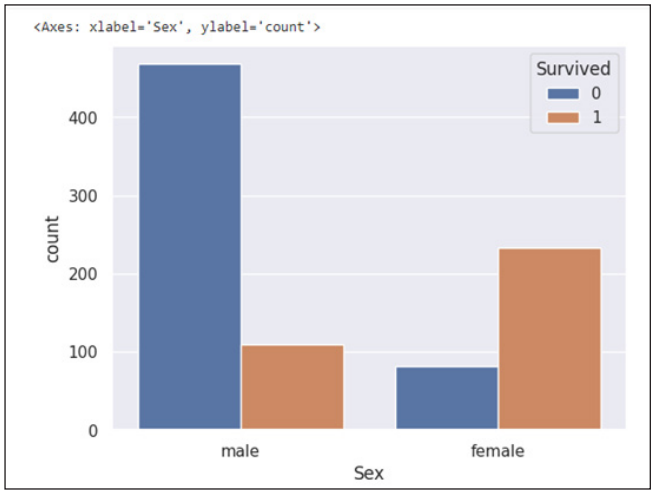

Figure 3.Graph showing survival of male and female

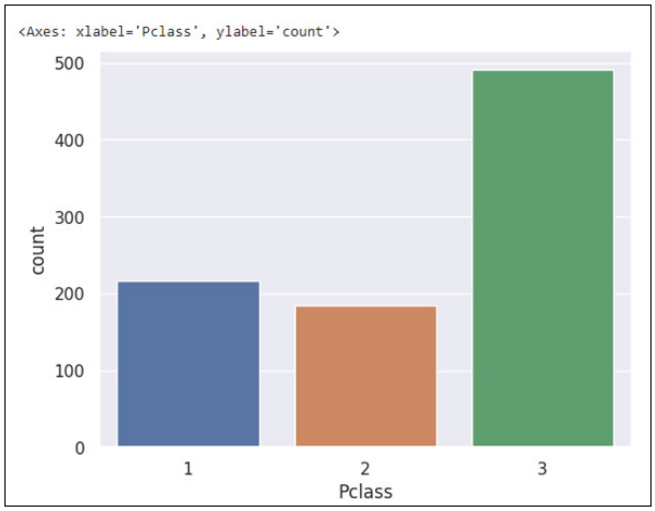

Figure 4.Graph showing passenger class of passengers

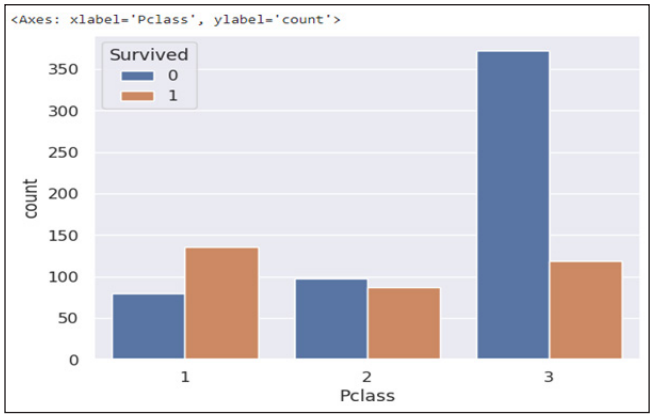

Table 7.Graph showing passenger survival belonging to their respective class

Statistical Tools to be Used for Analysis

The statistical tool that is used to analyse the Titanic dataset and predict survival rates. Logistic regression is a statistical technique used to model the relationship between a binary dependent variable and one or more independent variables.[18]

In the case of the Titanic dataset, the dependent variable is the survival of the passenger, which can take two values: 0 for not survived and 1 for survived. The independent variables, also known as predictors or features, are the demographic information of the passenger such as age, gender, class, etc.

Logistic regression estimates the probability of the dependent variable given the values of the independent variables. The output of logistic regression is a probability score ranging from 0 to 1.

This probability score can be interpreted as the likelihood of survival given the demographic information of the passenger. The logistic regression model can also be used to identify the most significant predictors of survival by examining the coefficients of the independent variables.[18]

Logistic regression is a widely used statistical tool in various fields such as healthcare, marketing, and social sciences. It is easy to implement and interpret, and it can handle both categorical and continuous predictors.

In the case of the Titanic dataset, logistic regression has been widely used by researchers to predict survival rates and identify the key factors that influence survival.

The statistical tool that is used to analyse the Titanic dataset and predict survival rates. Logistic regression is a statistical technique used to model the relationship between a binary dependent variable and one or more independent variables.[18]

In the case of the Titanic dataset, the dependent variable is the survival of the passenger, which can take two values: 0 for not survived and 1 for survived. The independent variables, also known as predictors or features, are the demographic information of the passenger such as age, gender, class, etc.

Logistic regression estimates the probability of the dependent variable given the values of the independent variables. The output of logistic regression is a probability score ranging from 0 to 1.

This probability score can be interpreted as the likelihood of survival given the demographic information of the passenger. The logistic regression model can also be used to identify the most significant predictors of survival by examining the coefficients of the independent variables.[18]

Logistic regression is a widely used statistical tool in various fields such as healthcare, marketing, and social sciences. It is easy to implement and interpret, and it can handle both categorical and continuous predictors.

In the case of the Titanic dataset, logistic regression has been widely used by researchers to predict survival rates and identify the key factors that influence survival.

Limitation of the proposed research work

While there have been many studies on the factors associated with survival on the Titanic, there are several limitations to consider when predicting survival rates using machine learning models:

Incomplete data: While there are passenger and survivor lists available, there may be missing or incomplete data on some individuals, such as their age, occupation, or socio-economic status. This missing data can limit the accuracy of the model.

Bias in the data: The data available may be biassed, for example, if there are more records available for certain classes of passengers, such as first-class passengers, than for others. This can lead to biassed results and inaccurate predictions.

Causality vs correlation: While machine learning models can identify correlations between variables, they cannot determine causality. For example, a model may find that passengers who were female were more likely to survive, but it cannot determine whether this was due to their gender or other factors, such as their age or social class.

Limited scope: Predictive models can only be built based on the available data, which may have limitations in scope or depth. For example, the available data may not include information on the crew members or the rescue efforts, which could have had an impact on the survival rates.

Lack of generalizability: Models trained on data from the Titanic may not be generalizable to other contexts or time periods, as the factors associated with survival may differ based on the specific circumstances.

Overall, while machine learning models can provide insights into the factors associated with survival on the Titanic, it is important to consider the limitations of the available data and the models themselves when making predictions.

References

1. Survival Prediction For Titanic Data Using Machine Learning Algorithms 1p. Ravindra, 2dr. P. Vijayapal Reddy,
2. Predicting The Survival Of Titanic Passengers Using Different Machine Learning Models Hitesh Pujari1, Rohit Singh Adhikari2, Plaban Patra3
3. Titanic Machine Learning From Disaster M.A.D.-Python Team: Dylan Kenny, Matthew Kiggans, Aleksandr Smirnov
4. Titanic Survival Prediction Kavya N C, Mr. Srinivasulu M
5. Survival Prediction For Titanic Data Using Machine Learning Algorithms 1p. Ravindra, 2dr. P. Vijayapal Reddy,
6. Predicting The Survival Of Titanic Passengers Using Different Machine Learning Models Hitesh Pujari1, Rohit Singh Adhikari2, Plaban Patra3
7. Titanic Machine Learning From Disaster M.A.D.-Python Team: Dylan Kenny, Matthew Kiggans, Aleksandr Smirnov
8. Berg, J. (2012). Titanic: Machine Learning From Disaster. Kaggle. Retrieved From Https://Www.Kaggle.Com/C/Titanic/Overview
9. Leite, C., & Vieira, S. M. (2018). Machine Learning Applied To Titanic Disaster: Survivors Prediction. 2018 Ieee 10th International Conference On Intelligent Data Acquisition And Advanced Computing Systems:

Technology And Applications (Idaacs). Doi: 10.1109/Idaacs.2018.8527435
10. Chen, J., & Li, Y. (2020). Prediction Of Titanic Passenger Survival Using Machine Learning Techniques. Journal Of Data Analysis And Information Processing, 8(3), 93-104.
11. Hall, M. (2020). Machine Learning In Disaster Response: A Review. Ieee Access, 8, 142933-142948.
12. Salazar, C., & Castillo, O. (2019). Machine Learning In The Analysis Of Titanic Disaster. Journal Of Big Data, 6(1), 13.
13. Kaggel. (2021). Titanic: Machine Learning From Disaster. Retrieved From Https://Www.Kaggle.Com/C/Titanic
14. Hasan, M. T., & Ahmed, A. (2018). Prediction Of Titanic Passenger Survival Using Machine Learning Techniques. Arxiv Preprint Arxiv:1805.03060.
15. Elango, V., Priya, A., & Rajam, B. (2020). Predicting The Survival Of Titanic Passengers Using Machine Learning Algorithms. In Proceedings Of The Second International Conference On Advances In Computing, Communication And Information Science (Pp. 543-550). Springer.
16. Ticehurst, R. (2012). 'women And Children First': Gender And Class In The Evacuation Of The Titanic. Journal Of Men's Studies, 20(2), 104-119.
17. Pei, Y., & Yan, H. (2018). Analysis Of Survival Factors Of Titanic Passengers Based On Data Mining. Destech Transactions On Computer Science And Engineering, (Cste).
18. Kaggle. (2018). Titanic: Machine Learning From Disaster. Retrieved From Https://Www.Kaggle.Com/C/Titanic/Data.