**ENE4062-01 Fall 2023  
Machine Learning Theory Project Proposal**

**Prediction of Survival Probability of Titanic Passengers**

조윤지, 강민지, Titouan MILLET

**Division of Electronics & Electrical Engineering, Dongguk University, Seoul, Republic of Korea**

**Description**

The Titanic, a magnificent ocean liner, left an indelible mark on history through its tragic sinking in 1912. On its maiden voyage, a dream of grandeur turned into a nightmare as the ship struck an iceberg and descended into the frigid waters of the Atlantic. The Titanic's sinking remains one of the most iconic maritime disasters, claiming many lives. This event triggered significant reforms in maritime safety and left an enduring impact on history.

Within the Titanic's historical context, a crucial question emerges: Can we predict the probability of a passenger's survival based on their personal characteristics? This question is of great significance as it allows us to explore how factors such as ticket class, gender, age, and more influence the likelihood of survival. Using a logistic regression model, this project aims to comprehend the dynamics behind critical decisions made on board the Titanic and predict survival probabilities.

Logistic regression is the most suitable method for this problem as it is specifically designed for solving binary classification problems, such as predicting survival (1 for survivor, 0 for non-survivor). It enables effective modeling and analysis of the relationships between the probability of survival and passenger characteristics. Logistic regression is widely employed in studies of this nature and will provide an in-depth understanding of the factors that influenced survival aboard the Titanic.

Key Steps for Resolution:

1. Initial data preparation will be essential, involving cleaning and preparing the data for rigorous analysis.
2. The careful selection of relevant features will be a focal point, enabling the identification of the most influential variables affecting survival probability.
3. Implementation of logistic regression, our prediction tool, will be executed to create a robust model.
4. Model evaluation will employ metrics like accuracy, sensitivity, specificity, among others, to measure its performance.
5. Lastly, interpretation of model coefficients will shed light on the impact of each characteristic on survival probability, offering a comprehensive understanding of the factors at play in this historic tragedy.

Keywords: Titanic, Survival Prediction, Logistic Regression, Machine learning, Data Analysis

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Project schedule** | | | | |
|  | **November**  **1-9** | **November**  **9-30** | **Décembre**  **1-7** | **Décembre**  **8-14** |
| **Writing Proposal** |  |  |  |  |
| **Code** |  |  |  |  |
| **Project presentation PowerPoint** |  |  |  |  |
| **Project report** |  |  |  |  |