# Titanic Survival Prediction: Enhancing Cruise Ship Safety

- Industry: Tourism & Hospitality | Domain: Maritime Safety & Risk Management
- By: Sai Prakhya

# Use Case Summary

# Objective

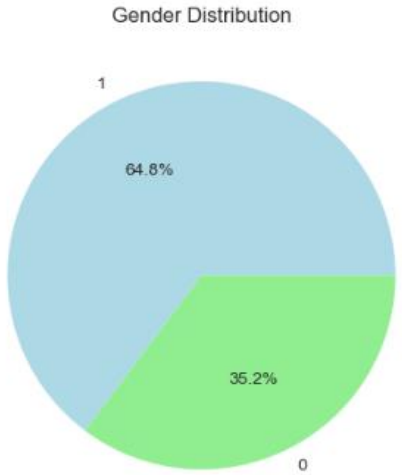
- Develop and compare machine learning models to predict passenger survival probability, forming the foundation of the 'Safety Companion' system.

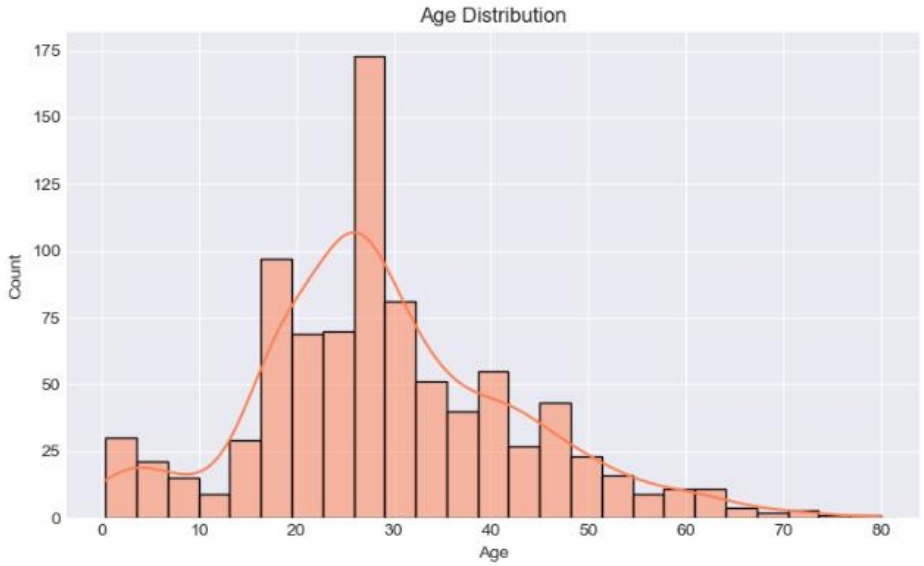
# Exploratory Data Analysis (Gender & Age Distributions)

- We analyzed the gender and age distributions to understand the demographics of passengers:
- 64.8% male, 35.2% female.
- Majority of passengers between the ages of 20-40.

# EDA Image: Gender & Age Distribution

.

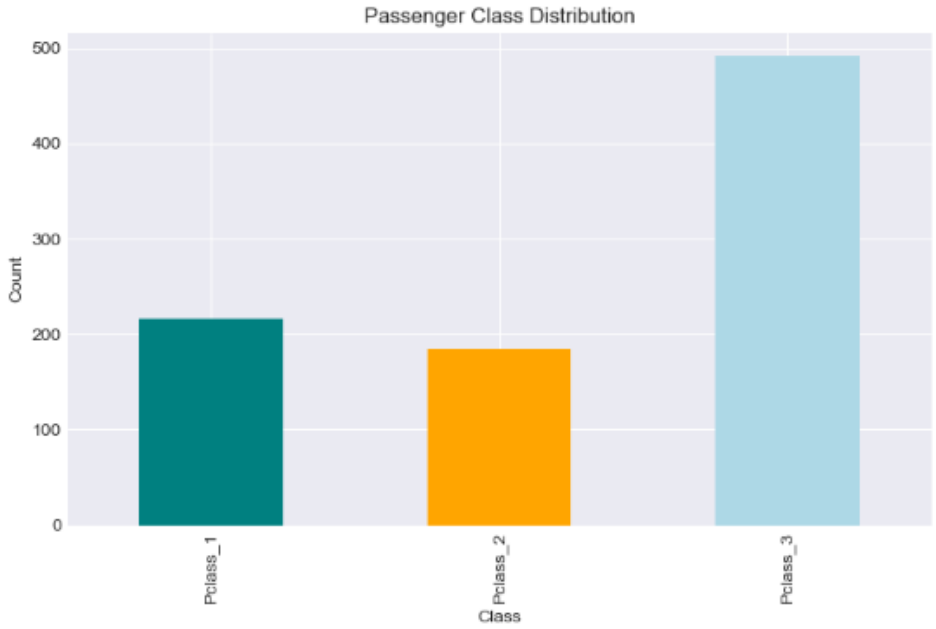


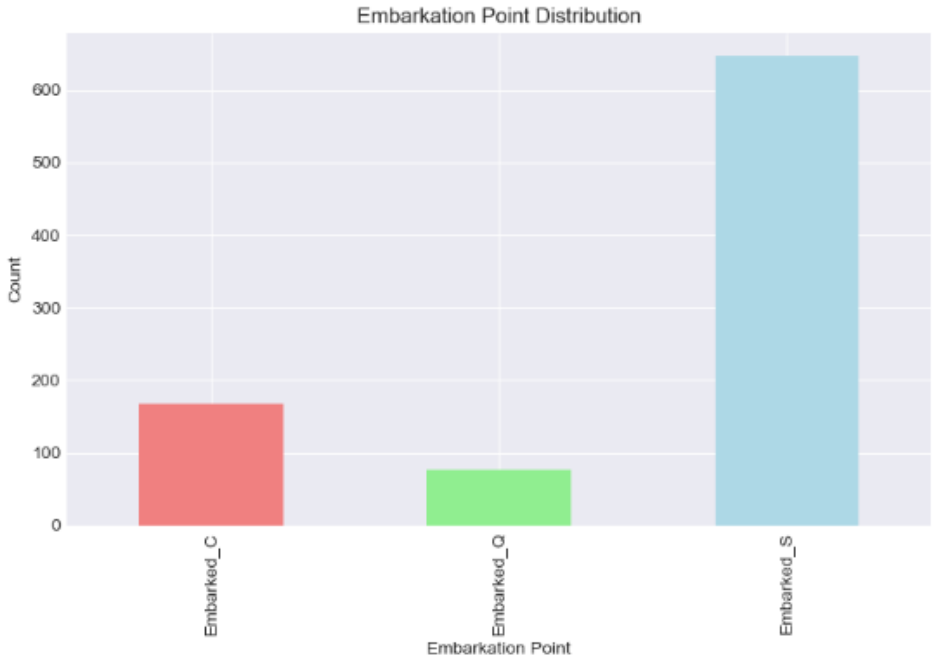



# Exploratory Data Analysis (Passenger Class & Embarkation Points)

- Passenger class and embarkation point distribution reveal socio-economic status:
- Majority in 3rd class, embarked at Southampton.

# EDA Image: Passenger Class & Embarkation Point Distribution

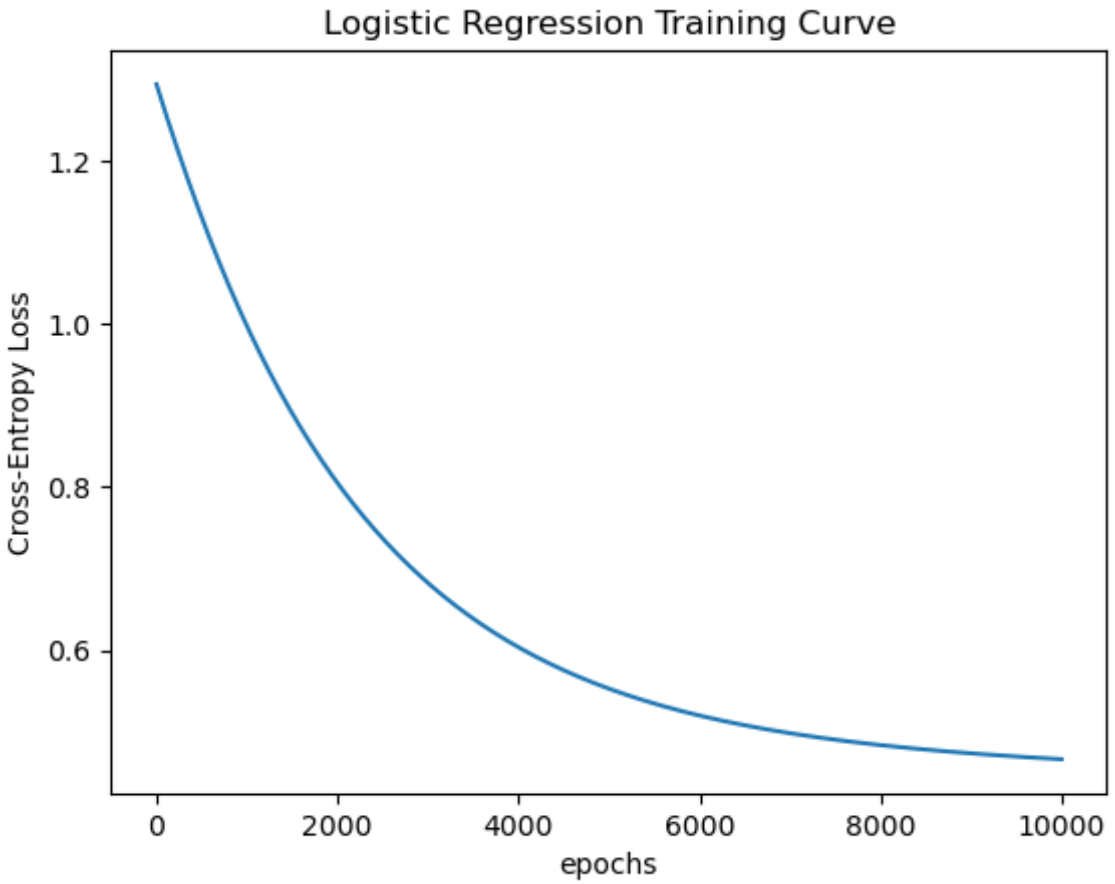


# Model: Logistic Regression

- Logistic Regression was used as a baseline model with 79.80% accuracy on training data. Feature normalization was applied to improve performance.

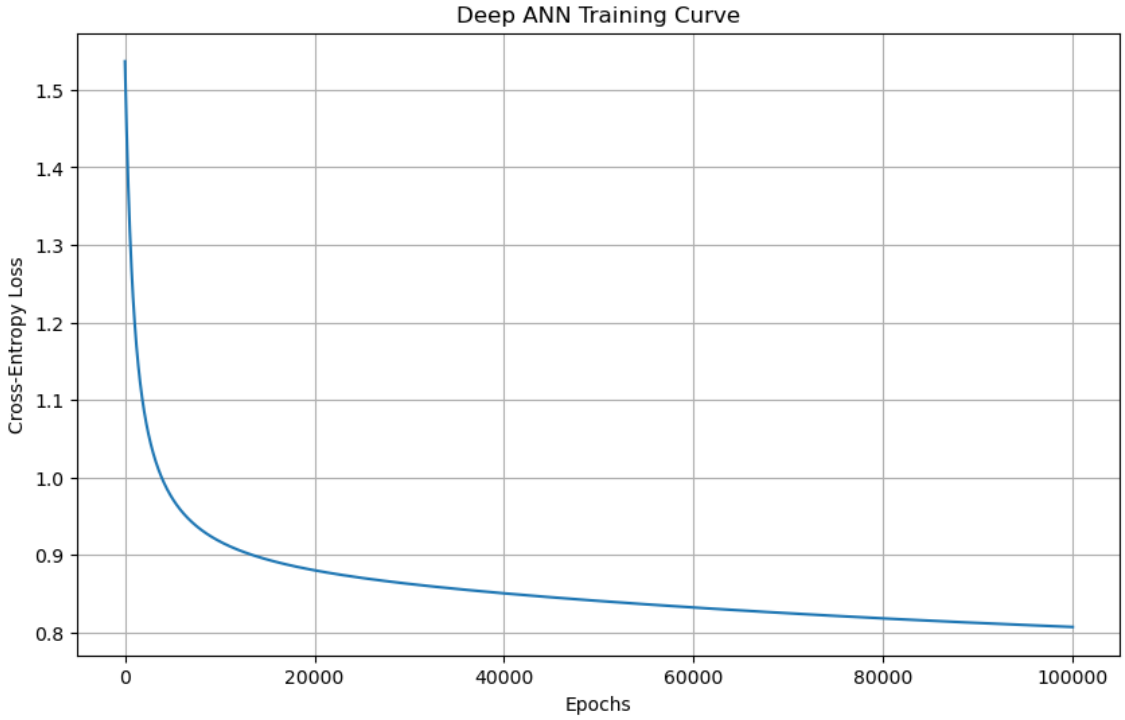
# Logistic Regression Training Curve



# Model: Deep Artificial Neural Network (ANN)

- The Deep ANN model captured non-linear relationships, achieving 82.49% accuracy. The architecture had [128, 64, 32] hidden layers with a ‘np.tanh' activation function.

# Deep ANN Training Curve



# Feature Importance Explanation

- Feature importance analysis shows that Passenger Class and Sex are the strongest predictors, followed by Age and Family Size.

# Feature Importance

13

| Feature | Importance |
| --- | --- |
| Pclass_3 | 1.064671 |
| Sex | 0.984643 |
| Pclass_2 | 0.417891 |
| Age | 0.226299 |
| Pclass_1 | 0.224947 |
| FamilySize | 0.110296 |

# Conclusion & Future Steps

- Deep ANN outperformed Logistic Regression. Future steps include improving risk assessment, collecting more real-time data, and implementing simulation tests.