SUPERVISED LEARNING

TITANIC

AI - A2_55
01/05/2024

GONÇALO SANTOS
GONÇALO PINTO
RUI CARVALHO

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

PROBLEM DESCRIPTION

The goal of this project is to predict whether a passenger aboard the Titanic survived or not during the tragic sinking on April 15, 1912. This is a binary classification problem where the target variable is "Survived" (1 for survived, 0 for not survived).

Goal: The goal is to develop a machine learning model that accurately predicts whether a passenger survived or not during the Titanic disaster based on available attributes such as age, gender, and socio-economic class. The aim is to achieve a high level of prediction accuracy to understand the factors influencing survival outcomes.

DATASET

The dataset contains information about passengers, including socio-economic status, age, gender, number of siblings/spouses aboard, number of parents/children aboard, ticket details, fare, cabin, port of embarkation, and other related attributes.

	pclass	survived	name	sex	age	sibsp	parch	ticket	fare	cabin	embarked	boat	body	home.dest
0	1	1	Allen, Miss. Elisabeth Walton	female	29.00	0	0	24160	211.3375	B5	S	2	NaN	St Louis, MO
1	1	1	Allison, Master. Hudson Trevor	male	0.92	1	2	113781	151.5500	C22 C26	S	11	NaN	Montreal, PQ / Chesterville, ON
2	1	0	Allison, Miss. Helen Loraine	female	2.00	1	2	113781	151.5500	C22 C26	S	NaN	NaN	Montreal, PQ / Chesterville, ON
3	1	0	Allison, Mr. Hudson Joshua Creighton	male	30.00	1	2	113781	151.5500	C22 C26	S	NaN	135.0	Montreal, PQ / Chesterville, ON
4	1	0	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)	female	25.00	1	2	113781	151.5500	C22 C26	S	NaN	NaN	Montreal, PQ / Chesterville, ON

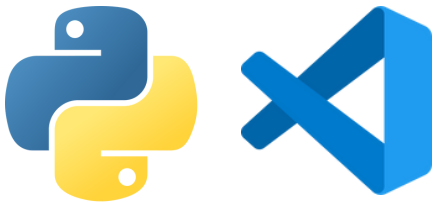
TOOLS / ALGORITHMS

Algorithms

Decision Tree Classifier

Neural Networks

K-Nearest Neighbour



Libraries

- Pandas
- Seaborn
- Scikit-Learn
- NumPy
- Anaconda
- SciPy
- MatPlotLib

WORK ALREADY IMPLEMENTED

- Read all the dataset
- Pre-process data
- Check for missing values or outliers

REFERENCES

- https://www.kaggle.com/datasets/sakshisatre/titanic-dataset
- IA slides and exercises