CCSEL1-18 Professional ELective

[Spaceship Titanic]

MARASIGAN, VEM AIENSI A.
3BSCS-I

Documentation and Leaderboard
January 2, 2024

SpaceShip Titanic-Code

Python · dataset, Spaceship Titanic

Notebook Input Output Logs Comments (0) Settings

Competition Notebook	Run	Public Score	Best Score	
Spaceship Titanic	30.1s	0.49310	0.4931 V2	Version 2 of 2

Add Tags

In [1]:
```
# Vem Aiensi Marasigan
```

In [2]:
```
# This Python 3 environment comes with many helpful analytics libraries installed
# It is defined by the kaggle/python Docker image: https://github.com/kaggle/docker-python
# For example, here's several helpful packages to load

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

# Input data files are available in the read-only "../input/" directory
# For example, running this (by clicking run or pressing Shift+Enter) will list all files under the in
put directory

import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))
```

Table of Contents

CCSEL1-18 Professional ELective

[Spaceship Titanic]

MARASIGAN, VEM AIENSI A.
3BSCS-I

Documentation and Leaderboard
January 2, 2024

SpaceShip Titanic-Code

0 Edit

Notebook Input Output Logs Comments (0) Settings

Exploring Datasets

Table of Contents
- Exploring Datasets
- Importing py libraries
- Solution
- Submitting the prediction

In [3]:
```
test_data = pd.read_csv("/kaggle/input/spaceship-titanic/test.csv")
test_data.head()
```

Out[3]:

	PassengerId	HomePlanet	CryoSleep	Cabin	Destination	Age	VIP	RoomService	FoodCourt	ShoppingMall	Spa	VR
0	0013_01	Earth	True	G/3/S	TRAPPIST-1e	27.0	False	0.0	0.0	0.0	0.0	0.0
1	0018_01	Earth	False	F/4/S	TRAPPIST-1e	19.0	False	0.0	9.0	0.0	2823.0	0.0
2	0019_01	Europa	True	C/0/S	55 Cancri e	31.0	False	0.0	0.0	0.0	0.0	0.0
3	0021_01	Europa	False	C/1/S	TRAPPIST-1e	38.0	False	0.0	6652.0	0.0	181.0	58
4	0023_01	Earth	False	F/5/S	TRAPPIST-1e	20.0	False	10.0	0.0	635.0	0.0	0.0

In [4]:
```
ss_data = pd.read_csv("/kaggle/input/spaceship-titanic/sample_submission.csv")
ss_data.head()
```

Out[4]:

	PassengerId	Transported
0	0013_01	False
1	0018_01	False
2	0019_01	False
3	0021_01	False
4	0023_01	False

SpaceShip Titanic-Code

0 Edit

Notebook Input Output Logs Comments (0) Settings

Importing py libraries

Table of Contents
- Exploring Datasets
- Importing py libraries
- Solution
- Submitting the prediction

In [5]:
```
import tensorflow as tf
import tensorflow_decision_forests as tfdf
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

CCSEL1-18 Professional ELective

[Spaceship Titanic]

MARASIGAN, VEM AIENSI A.
3BSCS-I

Documentation and Leaderboard
January 2, 2024

SpaceShip Titanic-Code

0 Edit

Notebook Input Output Logs Comments (0) Settings

Table of Contents

Solution

In [6]:

```
rf = tfdf.keras.RandomForestModel()
rf.compile(metrics=["accuracy"]) # Optional, you can use this to include a list of eval metrics

# Load the test dataset
test_df = pd.read_csv('/kaggle/input/spaceship-titanic/test.csv')
submission_id = test_df.PassengerId

# Replace NaN values with zero
test_df[['VIP', 'CryoSleep']] = test_df[['VIP', 'CryoSleep']].fillna(value=0)

# Creating New Features - Deck, Cabin_num and Side from the column Cabin and remove Cabin
test_df[["Deck", "Cabin_num", "Side"]] = test_df["Cabin"].str.split("/", expand=True)
test_df = test_df.drop('Cabin', axis=1)

# Convert boolean to 1's and 0's
test_df['VIP'] = test_df['VIP'].astype(int)
test_df['CryoSleep'] = test_df['CryoSleep'].astype(int)

# Convert pd dataframe to tf dataset
test_ds = tfdf.keras.pd_dataframe_to_tf_dataset(test_df)

# Get the predictions for testdata
predictions = rf.predict(test_ds)
n_predictions = (predictions > 0.4).astype(bool)
output = pd.DataFrame({'PassengerId': submission_id,
                       'Transported': n_predictions.squeeze()})
output.head()
```

SpaceShip Titanic-Code

0 Edit

Notebook Input Output Logs Comments (0) Settings

Submitting the prediction

In [7]:

```
sample_submission_df = pd.read_csv('/kaggle/input/spaceship-titanic/sample_submission.csv')
sample_submission_df['Transported'] = n_predictions
sample_submission_df.to_csv('/kaggle/working/submission.csv', index=False)
sample_submission_df.head()
```

Out[7]:

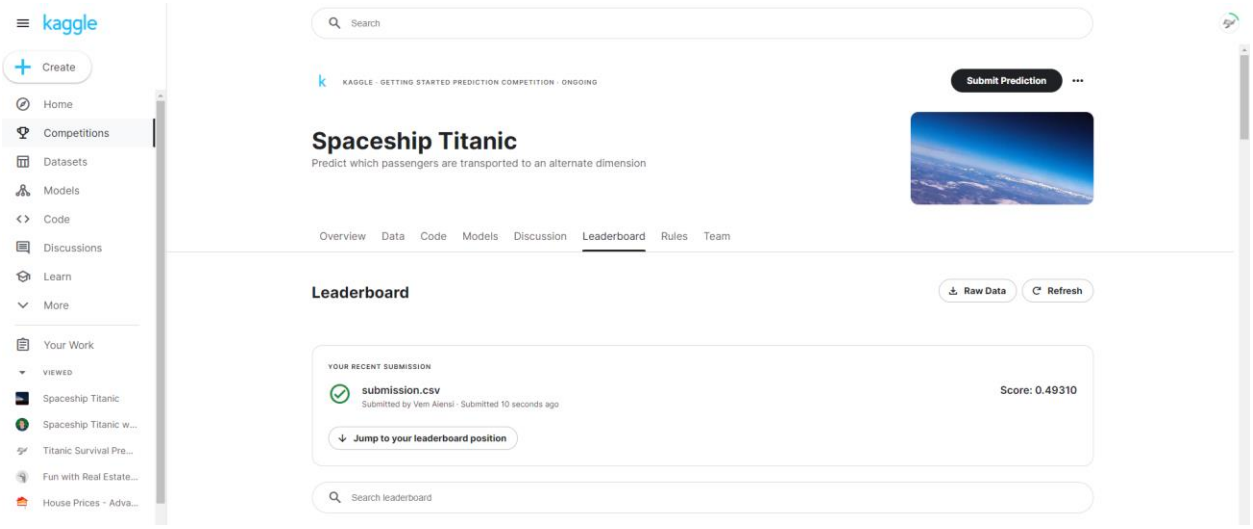
	PassengerId	Transported
0	0013_01	False
1	0018_01	False
2	0019_01	False
3	0021_01	False
4	0023_01	False

CCSEL1-18 Professional ELective

[Spaceship Titanic]

MARASIGAN, VEM AIENSI A.
3BSCS-I

Documentation and Leaderboard
January 2, 2024



Files for this challenge

https://github.com/VemAiensi/Professional-Elective-Course/tree/main/Kaggle-Competiton/Space-Titanic

Other Competition

https://github.com/VemAiensi/Professional-Elective-Course/tree/main/Kaggle-Competiton