

빅데이터분석

전 찬 준

cjchun@chosun.ac.kr

- Spaceship Titanic –

Spaceship Titanic

Predict which passengers are transported to an alternate dimension



Overview

Data

Code

Models

Discussion

Leaderboard


Rules

Overview

 This competition runs indefinitely with a rolling leaderboard. [Learn more.](#)

Description



 **Recommended Competition**

We highly recommend [Titanic - Machine Learning from Disaster](#) to get familiar with the basics of machine learning and Kaggle competitions.

Welcome to the year 2912, where your data science skills are needed to solve a cosmic mystery. We've received a transmission from four lightyears away and things aren't looking good.

The *Spaceship Titanic* was an interstellar passenger liner launched a month ago. With almost 13,000 passengers on board, the vessel set out on its maiden voyage transporting emigrants from our solar system to three newly habitable exoplanets orbiting nearby stars.

Competition Host

Kaggle



Prizes & Awards

Knowledge
Does not award Points or Medals

Participation

3,024 Competitors
2,504 Teams
15,272 Entries

Tags

- Beginner
- Tabular
- Binary Classification

Table of Contents



- Description
- Evaluation

학습 데이터(Train)

- 승객에 대한 개인 정보 (약 8700명)
 - PassengerId : 승객 ID (ggggg_pp)
 - HomePlanet : 출발 행성(거주지)
 - CryoSleep : CryoSleep 여부
 - Cabin : 객실 번호(Deck/Num/Side) → Side는 Port(P) 혹은 Starboard(S)
 - Destination : 승객의 목적지
 - Age : 승객의 나이
 - VIP : VIP 서비스 이용 여부
 - RoomService, FoodCourt, ShoppingMall, Spa, VRDeck : 해당 서비스를 위해 지불한 비용
 - Name : 승객의 이름
 - **Transported : 다른 차원으로 이동했는지 여부(True / False) → Target**

평가 데이터(Test)

- 승객에 대한 개인 정보 (약 4300명)
 - PassengerId : 승객 ID (ggggg_pp)
 - HomePlanet : 출발 행성(거주지)
 - CryoSleep : CryoSleep 여부
 - Cabin : 객실 번호(Deck/Num/Side) → Side는 Port(P) 혹은 Starboard(S)
 - Destination : 승객의 목적지
 - Age : 승객의 나이
 - VIP : VIP 서비스 이용 여부
 - RoomService, FoodCourt, ShoppingMall, Spa, VRDeck : 해당 서비스를 위해 지불한 비용
 - Name : 승객의 이름
 - **Transported : 다른 차원으로 이동했는지 여부(True / False) → 예측 대상**

제출 양식

- PassengerId : 평가 데이터에 해당되는 승객의 ID
- Transported : True or False.

과제#1

- 배점 : 5점
- 기한 : 3월 31일 (일) 자정까지
- 제출 파일 :
 - 주피터(ipynb) 혹은 파이썬(py) 파일
 - 제출한 submission.csv 파일
 - 레포트(hwp / pdf / docx 중 하나) 파일
 - 어떠한 입력 특징들을 활용했는지?
 - 어떠한 모델을 학습했는지? (반드시 Neural Network를 활용할 것! PyTorch를 반드시 활용할 것!)
 - 하이퍼파라메타는 어떻게 설정했는지?
 - 그래서 결과는 어느정도 나왔는지?
 - 소견
- 지금까지의 코드를 반드시 활용하자.
- Kaggle에 있는 공식 Code 및 Discussion을 활용하면 0점 처리함

과제#1

- 평가의 착안점
 - 딥러닝을 활용했는가?
 - PyTorch를 활용했는가?
 - 평가 데이터에 대한 정확도가 너무 낮게 나오지는 않았는가?
 - 강의자료에 있는 코드를 적극 활용했는가?
 - 레포트를 성실하게 작성하였는가?