

Primera Entrega de Proyecto

Inteligencia Artificial

Código: 2508401

Grupo: 01

Fecha: 05/ 07/ 2022

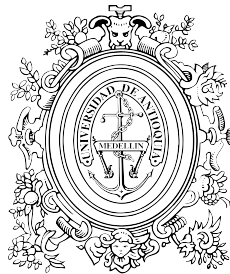
Elaborado por

Sofía Gutiérrez Tangarife, CC: 1001576929

Melissa Galeano Ruiz, CC: 1000416463

Profesor

Raúl Ramos Pollán



**UNIVERSIDAD
DE ANTIOQUIA**

1 8 0 3

Facultad de Ingeniería

Departamento de Ingeniería Mecánica

Universidad de Antioquia

Contents

1	Planteamiento del problema	2
2	Dataset	2
3	Métricas de desempeño	3
4	Desempeño	4
5	Referencias	4

1 Planteamiento del problema

Una nave espacial llamada "Titanic" fue una nave interestelar de pasajeros lanzada hace un mes. El propósito principal de esta nave era transportar a 13000 emigrantes a tres exoplanetas que orbitaban estrellas cercanas. En su ruta sideral a su primer destino (el exoplaneta 55 Cancri E), al rodear a Alpha Centauri esta nave colisionó con una anomalía espacio-temporal escondida en una nube de polvo estelar. Desafortunadamente, aunque la nave permaneció intacta, casi la mitad de los pasajeros fueron transportados a una dimensión alterna. Para rescatar a la tripulación y recuperar a los pasajeros perdidos se debe diseñar un modelo que permita predecir cuáles pasajeros fueron enviados a la dimensión alterna por la anomalía usando los registros recuperados del sistema de computación dañado de la nave espacial.



Figure 1: Nube de Polvo Estelar tomado de [1]

2 Dataset

El dataset a utilizar proviene de una competencia de Kaggle y la cual está dividida en datos de entrenamiento y datos de validación. Los datos de entrenamiento está conformado por 2/3 de los pasajeros totales de la nave y los datos de validación están conformados por 1/3 de los pasajeros.

En el archivo que contiene los datos de entrenamiento nombrado como *train* se pueden encontrar las siguientes variables:

- **PassengerId:** Identificación para cada pasajero donde cada ID toma la forma *gggg-pp* donde *gggg* indica el grupo con el que viaja el pasajero y *pp* es su número correspondiente dentro del grupo.
- **HomePlanet:** Planeta del que partió o residía el pasajero.
- **CryoSleep:** Indica si el pasajero eligió ser puesto en animación suspendida (cabina) durante el viaje.
- **Cabin:** Número de cabina donde permenece el pasajero.

- **Destination:** Destino de desembarque del pasajero.
- **Age:** Edad del pasajero.
- **VIP:** Si el pasajero pagó o no por el servicio VIP en el viaje.
- **RoomService:** Variable binaria categórica que indica si el pasajero contrató o no el servicio.
- **FoodCourt:** Variable binaria categórica que indica si el pasajero contrató o no el servicio.
- **ShoppingMall:** Variable binaria categórica que indica si el pasajero contrató o no el servicio.
- **Spa:** Variable binaria categórica que indica si el pasajero contrató o no el servicio.
- **VRDeck:** Variable binaria categórica que indica si el pasajero contrató o no el servicio.
- **Name:** El nombre y apellido del pasajero.
- **Transported:** Variable que indica si el pasajero fue o no transportado a otra dimensión.

En el archivo donde se encuentra los datos de validación nombrado como *test* se pueden encontrar las variables anteriormente mencionadas, pero sin incluir los pasajeros transportados que es la variable que se desea predecir.

3 Métricas de desempeño

La métrica de evaluación principal para este modelo es *Accuracy* o métrica de precisión la cual es usada para evaluar modelos de clasificación y se basa en la cantidad de verdaderos positivos (true positives) o el porcentaje de predicciones correctas.

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \quad (1)$$

$$Accuracy = \frac{TN + TP}{TP + TN + FP + FN} \quad (2)$$

Donde:

- *TP* : True Positives
- *TN* : True Negatives
- *FP* : False Positives
- *FN* : False Negatives

Por otro lado, en cuanto a la métrica de negocio, se tiene interés en que las predicciones sean lo suficientemente confiables para aumentar la confiabilidad de cada uno de los viajes interestelares.

4 Desempeño

En los viajes interestelares son muy comunes las anomalías espacio-temporales y debido a esto se desea implementar en las nuevas agencias de transporte interestelar un método de predicción que nos permita brindarle con seguridad a cada pasajero información acerca de su probabilidad de ser transportado a otra dimensión, y así, este pueda ser recuperado fácilmente. Con esta información, se espera determinar qué tanto se pueden reducir los costos en función a la probabilidad.

5 Referencias

[1] *Spaceship Titanic* | *Kaggle*. (n.d.). Retrieved July 1, 2022, from <https://www.kaggle.com/competitions/spaceship-titanic/overview>