

Trabajo Encargado 1 (TE1)

Dados Estructurados

Escuela de Posgrado - Universidad Nacional del Altiplano de Puno
Doctorado en Ciencias de la Computación
Prof. Dr. Ivar Vargas Belizario

1 Descripción

Seleccionar uno de los siguientes conjuntos de datos:

- (i) Shelter Animal Outcomes
<https://www.kaggle.com/competitions/shelter-animal-outcomes/>
- (ii) Binary Classification of Machine Failures
<https://www.kaggle.com/competitions/playground-series-s3e17/>
- (iii) Predict Health Outcomes of Horses
<https://www.kaggle.com/competitions/playground-series-s3e22/>

Proponer una solución computacional completando las siguientes tareas¹:

- (i) Implementar tareas de pre-procesamiento y transformación de datos. Apoyar la solución con visualizaciones de atributos que considere necesarias. Verificar si el conjunto de datos está balanceado.
- (ii) Dividir el conjunto en entrenamiento (*Train*) y validación (*Validation*). Entrenar uno ó varios modelos de *machine learning* (*Random Forest*, *XGBoost*, etc.) con las instancias del conjunto *Train*.
- (iii) Validar el modelo con el conjunto *Validation*. Para la validación considere una o varias métricas: *accuracy*, *precision*, *recall* o *f1-score*.
- (iv) Determinar por predicción las etiquetas del conjunto *Test*. Opcionalmente, realizar el envío de la solución del conjunto *Test* en el portal de kaggle.
- (v) Redactar un informe (no más de 4 páginas) describiendo la solución propuesta. Incluya gráficos para apoyar la descripción.

2 Forma de envío

Considerar las siguientes indicaciones:

- (i) Enviar el **notebook** (indicando la versión compilada, si fuera el caso) y el **informe** de su trabajo al correo **ivargasbelizario@gmail.com** con el asunto “**TE1-DCC302-25I Trabajo Encargado**”.
- (ii) Si optan por compartir el notebook en el portal Kaggle, compartirlo directamente con la cuenta del profesor: “**Ivar Vargas Belizario (ivarvb)**”. En ambos casos es necesario mantener la privacidad del notebook.
- (iii) Fecha de entrega: hasta el **25/05/2025**.

¹opcionalmente utilizar el siguiente código base: <https://www.kaggle.com/code/ivarvb/dcc25te1/>