**Taller 2.**

**Predecir la pobreza.**

* El documento no debe tener más de 10 (diez) páginas e incluir, como máximo, 8 (ocho) pruebas (tablas y/o figuras). La bibliografía y las exhibiciones no cuentan para el límite de páginas.
* Describir en el texto el código utilizado, la forma de replicarlo y todas las tablas e imágenes que se agreguen.
* README

**Integrantes:** Isabella Mendez Pedraza. Cód.: 201814239

Manuela Ojeda Ojeda. Cód.: 201814476

Juan Sebastian Tellez Melo. Cód.: 201513710

Andres Mauricio Palacio Lugo. Cód.: 201618843

**Link del repositorio:** <https://github.com/AndresMPL/Repositorio_PS2.git>

**Modelo predictivo de la pobreza de los hogares.**

I = función de indicador que toma 1 si el ingreso familiar está por debajo de una cierta línea de pobreza.

Predecir la pobreza de dos maneras:

* Problema de clasificación: 0 (no pobre) y 1 (pobre).
* Problema de predicción de ingresos. Con el ingreso previsto, puede usar la línea de pobreza y obtener la clasificación.

1. **Introducción.**

* La introducción expone brevemente el problema y si hay antecedentes (Literatura).
* Describir brevemente los datos y su idoneidad para abordar la cuestión del conjunto de problemas.
* Vista previa de los resultados y las principales conclusiones.

1. **Datos.**

* Describir la adecuación de los datos para resolver la pregunta predictiva, el proceso de construcción de la muestra, incluyendo cómo se limpiaron y combinaron los datos y cómo se crearon nuevas variables.
* Incluir un análisis descriptivo de los datos. Como mínimo, incluir una tabla de estadísticas descriptivas con su interpretación. Un análisis profundo que ayude al lector a comprender los datos, su variación y la justificación de las elecciones de datos.

1. **Modelos y resultados.**

Especificaciones y modelos utilizados para las tareas predictivas.

Subsecciones:

1. Modelos de clasificación. Describir el enfoque de clasificación, es decir, su intento de predecir directamente ceros (no pobre) y unos (pobre).
2. Modelos de regresión de ingresos. Describir el enfoque de predicción de ingresos, es decir, su intento de predecir primero los ingresos y luego predecir indirectamente ceros (no pobre) y unos (pobre).
3. Modelo final. Describir los modelos que seleccionó como su presentación final en la competencia. Hasta 2 envíos contarán para la puntuación de la tabla de clasificación. Si se seleccionan menos de 2, Kaggle seleccionará automáticamente entre los envíos con la mejor puntuación. Esta subsección debe incluir:
4. Una explicación detallada de los modelos finales elegidos para la evaluación en Kaggle. La explicación debe incluir cómo se entrenó el modelo, la selección de hiperparámetros y cualquier otra información relevante.
5. Una comparación con al menos otras 2 especificaciones, para cada enfoque.
6. Una descripción de las variables utilizadas en el modelo y discutir su importancia relativa en la predicción.
7. Una descripción de cualquier estrategia de submuestreo utilizada para abordar los desequilibrios de clase.
8. **Conclusiones y recomendaciones.**

Principales conclusiones del trabajo.

1. **Bibliografía.**
2. **Apéndice.**