

# Predictor de bicicletas prestadas

## Análisis Avanzado de Datos - Práctica Final

Alejo Martín, Arias Filippo (NIA: 100487858)

2023

### I. Introducción

Los sistemas de bicicletas compartidas se han vuelto muy populares en todo el mundo debido a su capacidad para mejorar la movilidad urbana, la salud y el medio ambiente. Estos sistemas generan grandes cantidades de datos que pueden ayudarnos a comprender mejor cómo se mueven las personas en las ciudades y qué eventos importantes ocurren en ellas. En este proyecto, vamos a utilizar un conjunto de datos de un sistema de bicicletas compartidas para construir un modelo de aprendizaje supervisado que prediga la cantidad de bicicletas alquiladas en función de diferentes factores, como la hora del día, la temperatura, la humedad y las condiciones meteorológicas.

El objetivo principal de este proyecto es mostrar cómo el aprendizaje supervisado puede ser útil para predecir la demanda de bicicletas compartidas según las condiciones ambientales y temporales, lo que podría ayudar a los operadores del sistema a mejorar la disponibilidad de bicicletas y a satisfacer mejor las necesidades de los clientes. Para lograr este objetivo, vamos a seguir un enfoque de aprendizaje supervisado, utilizando un conjunto de datos etiquetados para entrenar un modelo que pueda predecir la cantidad de bicicletas alquiladas según los diferentes factores que hemos mencionado.

### II. Descripción del Dataset

El dataset utilizado en este proyecto es un conjunto de registros de préstamos de bicicletas compartidas en un sistema automatizado. Contiene información sobre más de 17,000 préstamos realizados cada hora durante un período de dos años, desde enero de 2011 hasta diciembre de 2012. Cada registro contiene detalles importantes, como la fecha y hora del préstamo, las condiciones climáticas, y la cantidad de bicicletas alquiladas por usuarios registrados y ocasionales.

El dataset consta de 15 columnas que proporcionan información útil sobre los patrones de préstamo de bicicletas. Las primeras columnas incluyen detalles sobre el día y la hora del préstamo, así como la temporada, el año, el mes, el día de la semana y si es un día laborable o festivo. También se registran la temperatura, la sensación térmica, la humedad y la velocidad del viento. Además, la columna "weathersit" indica las condiciones climáticas generales en la hora del préstamo. Finalmente, se incluyen tres columnas que registran el número de bicicletas alquiladas por usuarios ocasionales, usuarios registrados y el número total de bicicletas alquiladas.

- **Instant:** Este es el índice de registro y no tiene un valor informativo por sí mismo.
- **Dteday:** Esta variable indica la fecha en que se realizó el préstamo. La fecha está en formato año-mes-día.
- **Hr:** Esta variable indica la hora del día en que se realizó el préstamo.
- **Season:** Esta variable indica la temporada en la que se realizó el préstamo. Los valores son 1 (primavera), 2 (verano), 3 (otoño) y 4 (invierno).
- **Year:** Esta variable indica el año en que se realizó el préstamo.
- **Mnth:** Esta variable indica el mes en que se realizó el préstamo.

- **Holiday:** Esta variable indica si el día en que se realizó el préstamo era un día festivo (1) o no (0).
- **Weekday:** Esta variable indica el día de la semana en que se realizó el préstamo. Los valores son 0 (domingo) a 6 (sábado).
- **Workingday:** Esta variable indica si el día en que se realizó el préstamo era un día laborable (1) o no (0).
- **Weathersit:** Esta variable indica las condiciones climáticas generales en la hora del préstamo. Los valores son 1 (despejado), 2 (nublado), 3 (lluvia ligera/nieve ligera) y 4 (lluvia intensa/nieve intensa).
- **Temp:** Esta variable indica la temperatura en grados Celsius en el momento del préstamo.
- **Atemp:** Esta variable indica la sensación térmica en grados Celsius en el momento del préstamo.
- **Hum:** Esta variable indica la humedad relativa en el momento del préstamo.
- **Windspeed:** Esta variable indica la velocidad del viento en km/h en el momento del préstamo.
- **Casual:** Esta variable indica el número de bicicletas alquiladas por usuarios ocasionales en la hora del préstamo.
- **Registered:** Esta variable indica el número de bicicletas alquiladas por usuarios registrados en la hora del préstamo.
- **Count:** Esta variable indica el número total de bicicletas alquiladas en la hora del préstamo (es decir, la suma de los valores de las variables Casual y Registered).

### III. Descripción del problema

En este proyecto, nuestro objetivo es predecir el número de bicicletas que se alquilarán en una hora determinada utilizando un modelo de aprendizaje supervisado. Para lograr esto, vamos a considerar varias variables que podrían influir en la cantidad de bicicletas alquiladas.

Una de las variables más importantes es la hora del día, ya que se espera que la cantidad de bicicletas alquiladas varíe a lo largo del día debido a los patrones de movilidad de las personas. Es posible que haya más alquileres de bicicletas durante las horas pico de la mañana y la tarde, cuando la gente se dirige al trabajo o regresa a casa.

Otra variable importante es el clima. Las condiciones climáticas pueden tener un gran impacto en la cantidad de bicicletas alquiladas. Es posible que haya menos alquileres de bicicletas en días lluviosos o fríos, mientras que en días soleados y cálidos se pueden alquilar más bicicletas.

Además, las variables de temperatura, humedad y velocidad del viento también pueden tener un efecto en la cantidad de bicicletas alquiladas. Es posible que las personas estén menos dispuestas a alquilar bicicletas en días extremadamente calurosos o húmedos.

Finalmente, es importante tener en cuenta las variables relacionadas con los usuarios, como la cantidad de usuarios registrados y ocasionales. Es posible que los patrones de alquiler difieran según el tipo de usuario y que los usuarios registrados alquilen bicicletas con más frecuencia que los usuarios ocasionales.

#### IV. Análisis de los datos

#### V. Método

#### VI. Resultados

##### 1. Limitaciones

#### VII. Conclusiones y Trabajo Futuro

#### Apéndice A

Appendix uno va aquí.

#### Apéndice B

Appendix dos va aquí.

#### Referencias

---