

Primer entrega de proyecto

Cristian Camilo Serna Betancur¹ Diego Alonso Herrera Ramírez² Sharid Samantha Madrid Ospina³

I. PROBLEMA PREDICTIVO

SE quiere predecir la tasa de desempleo trimestral o anual de una comuna de la ciudad de Medellín en función de diversas características demográficas y económicas de su población.

II. DATASET

Se usa el conjunto de datos *Gran Encuesta Integrada de Hogares* que proporciona el *Departamento Administrativo Nacional de Estadística* (DANE) en colaboración con el Municipio de Medellín. Esta encuesta provee información socioeconómica detallada con énfasis en la fuerza laboral y está desagregada para el Municipio y sus 16 comunas. El conjunto de datos contiene información demográfica y laboral de los hogares encuestados en Medellín.

III. MÉTRICAS DE DESEMPEÑO

La métrica de evaluación principal para el modelo es el Error Cuadrático Medio (MSE):

$$\mathbf{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_1)^2 \quad (1)$$