

Liberty Mutual Insurance es una de las mayores empresas del mundo de los seguros.

Dentro de la industria de seguros de empresas, las pérdidas por incendios representan una parte significativa de las pérdidas totales de propiedad. Alta gravedad y baja frecuencia son las características fundamentales

Las pérdidas por incendios son volátiles por naturaleza, lo cual hace difícil modelar.

La tarea consiste en predecir el ratio transformado de la pérdida total por fuego("target"), usando la información proporcionada. Esto permitirá una identificación más precisa de la exposición al riesgo de cada asegurado y la capacidad de adaptar la cobertura de seguro para su operación específica.

El dataset consiste en casi medio millón de registros de clientes de seguro, las variables proporcionadas son relativas a la póliza, criminalidad, geodemográficas y meteorológicas.

Información detallada de variables:

id : identificador único

target : ratio transformado de la pérdida total por fuego

var1 : categórica com posibles valores : 1, 2, 3, 4, 5, Z*

var2 : categórica com posibles valores : A, B, C, Z*

var3 : categórica com posibles valores : 1, 2, 3, 4, 5, 6, Z*

var4 : categórica com posibles valores en orden jerárquico : A1, B1, C1, D1, D2, D3, D4, E1, E2, E3, E4, E5, E6, F1, G1, G2, H1, H2, H3, I1, J1, J2, J3, J4, J5, J6, K1, L1, M1, N1, O1, O2, P1, R1, R2, R3, R4, R5, R6, R7, R8, S1, Z*

var5 : categórica com posibles valores : A, B, C, D, E, F, Z*

var6 : categórica com posibles valores : A, B, C, Z*

var7 : categórica com posibles valores : 1, 2, 3, 4, 5, 6, 7, 8, Z*

var8 : categórica com posibles valores : 1, 2, 3, 4, 5, 6, Z*

var9 : categórica com posibles valores : A, B, Z*

var10 – var17 : conjunto de variables numéricas normalizadas representando características de la póliza (note: var11 es el peso usado en el cálculo del índice de Gini)

dummy : categórica com posibles valores : A, B

crimeVar1 – crimeVar9: conjunto de variables numéricas
normalizadas con datos de criminalidad

geodemVar1 – geodemVar37 : conjunto de variables numéricas
normalizadas con datos geodemográficos

weatherVar1 – weatherVar236 : conjunto de variables numéricas
normalizadas meteorológicas