

Machine Learning Immersion

Se entrega el siguiente set de datos de una empresa del sector financiero, CreditScoring.csv, el cual es un dataset que cuenta con la información del área de riesgos de una entidad financiera y necesita su ayuda para construir un modelo para predecir **la propensión o probabilidad a default de un prestatario**, es decir predecir la probabilidad de que alguien experimente dificultades financieras en los próximos dos años.

Los bancos juegan un papel crucial en las economías de mercado. Deciden quién puede obtener financiación y en qué términos y pueden tomar o deshacer decisiones de inversión. Para que los mercados y la sociedad funcionen, los individuos y las empresas necesitan acceso al crédito. Los algoritmos de calificación crediticia, que adivinan la probabilidad de incumplimiento, son el método que utilizan los bancos para determinar si se debe otorgar o no un préstamo. No olvidemos que por cada cliente que nos deja de pagar, los bancos deben provisionar ese monto y por ende perder capital que podrían usar en otros negocios de inversión.

- Las variables que se disponibilizan son:

Variable	Descripción de la Variable
SeriousDlqin2yrs	Persona experimenta 90 días de morosidad o peor
RevolvingUtilizationOfUnsecuredLines	Saldo total de tarjetas de crédito y líneas de crédito personales, excepto bienes inmuebles y ninguna deuda a plazos, como préstamos para automóviles, dividido por la suma de los límites de crédito
age	Edad del prestatario en años
NumberOfTime30-59DaysPastDueNotWorse	Número de veces que el prestatario ha estado atrasado entre 30 y 59 días, pero no ha empeorado en los últimos 2 años.
DebtRatio	Pagos mensuales de la deuda, pensión alimenticia, costos de vida divididos por el ingreso bruto mensual
MonthlyIncome	Ingreso mensual
NumberOfOpenCreditLinesAndLoans	Número de préstamos abiertos (cuotas como préstamos para automóviles o hipotecas) y líneas de crédito (por ejemplo, tarjetas de crédito)
NumberOfTimes90DaysLate	Número de veces que el prestatario ha estado atrasado 90 días o más.
NumberRealEstateLoansOrLines	Número de préstamos hipotecarios e inmobiliarios, incluidas líneas de crédito sobre el valor neto de la vivienda
NumberOfTime60-89DaysPastDueNotWorse	Número de veces que el prestatario ha vencido 60-89 días, pero no ha empeorado en los últimos 2 años.
NumberOfDependents	Número de dependientes en la familia excluyéndose a sí mismos (cónyuge, hijos, etc.)

Entregables del TA:

- 1.- Definir el problema de la naturaleza que se tiene a continuación, además de los objetivos de negocio bien definidos.
- 2.- ¿Qué tipo de variables se utilizan en el problema de negocio?

