

MÉTODOS DE REGRESIÓN

MACHINE LEARNING CON REGRESIÓN LINEAL MÚLTIPLE

MACHINE LEARNING CON RANDOM FOREST

RESUMEN

Objetivo: Los métodos de Machine Learning con regresión lineal múltiple y Machine Learning con Random Forest. Nos pueden ayudar a la toma de decisiones que condiciona la evolución de cualquier empresa y los responsables deben ser capaces de tomar decisiones de la forma más segura posible, eliminando el azar del proceso. Estas técnicas de *machine learning*, que trabajan con una vertiente de la inteligencia artificial, ha nacido para responder a esa necesidad, ya que se está convirtiendo en la herramienta fundamental para la toma de decisiones fiables a través del análisis de grandes cantidades de datos y hechos.

Material y metodología: Usamos técnicas de métodos de regresión para obtener la importancia de las variables que pudieran ayudar a crear modelos de predicción para diferentes situaciones en los que tengamos datos de un cierto fenómeno que está inicialmente determinado por un conjunto de variables. Esta metodología es una parte de *machine learning*. La base de datos fue creada con registros *MATRICULA OFICIAL EN LOS MUNICIPIOS DEL MAGDALENA SEGUN ZONA, JORNADA E INSTITUCIONES EDUCATIVAS EN LOS AÑOS AÑOS 2020 Y 2021*, Alumnos matriculados en el departamento del Magdalena según zona geográfica, jornada escolar e instituciones educativas (**variable predictora**) durante los años 2020 y 2021.

Resultado: la mejor técnica fue Machine Learning con Random Forest que como podemos evidenciar en la imagen, nos han permitido elevar el Score de 0.13 a 0.87 aproximadamente y reducir el RMS de 42.99 a 16.96

score_SM	is	0.13647124750767126
score_LR	is	0.0905651438620394
score_LR2	is	0.04044475661842106
score_RF	is	0.8655599308931311
RMS2_SM	is	42.9930796332781
RMS2_LR	is	44.12106140578592
RMS2_LR2	is	45.3205472413669
RMS2_RF	is	16.963853660505258

Conclusiones: Estos métodos no podrían ayudar a predecir el numero de estudiantes que aprobaran y/o reprobaran una o varias asignaturas, ayudándonos a definir el posible número de cursos a ofertar para esas asignaturas

