

Reconocimiento de Patrones

Ashley
Echevarria
Ruiz



01

Modelo de Clasificación

Un modelo de clasificación es un modelo de aprendizaje supervisado de Machine Learning. Esto significa que tanto las variables como las respuestas están presentes en el conjunto de datos. El objetivo de la clasificación es identificar patrones comunes en las variables para predecir a qué categoría o clase pertenecen.

El análisis de datos es fundamental en este tipo de proyectos, ya que se necesitan datos correctos y precisos para que los resultados de la clasificación sean exactos. Los clasificadores de Machine Learning funcionan mejor con números que con letras, por lo que primero convertí los datos. Identifiqué las variables categóricas y las variables continuas para realizar las agrupaciones respectivas.

Para convertir las variables en números, utilicé dos métodos: el método manual con la función de diccionario, donde especificaba qué datos transformar y en qué valores transformarlos, y el método de ``get_dummies``, que divide las columnas en varias columnas nuevas, agrupando los datos.

El balanceo de los datos es una parte crucial del tratamiento de datos. Si los datos no están balanceados, el modelo tiene dificultades para identificar patrones que generen ciertas respuestas. Para lograr un balance, es necesario añadir más datos, lo cual se realiza utilizando la función ``over_sampling``, que crea valores ficticios hasta alcanzar un balance adecuado. La función SMOTE también ayuda en este proceso al crear registros adicionales basados en los existentes, evitando duplicaciones.



02

Modelo Predictivo KNN

El modelo predictivo KNN nos ayuda a realizar predicciones utilizando como base los resultados más cercanos. El método funciona de la siguiente manera: recibe las variables independientes de cada observación, calcula la distancia entre ellas, ordena las distancias de menor a mayor, clasifica a partir del número de vecinos que elijamos, realiza un conteo de tipos de datos para cada clase y, finalmente, predice el resultado.