

IES Pere Maria Orts

Sistemas de Aprendizaje Automático

Práctica 2_1: Introducción

Autor:

Kenny Berrones

Profesor:

David Campoy Miñarro



iesperemariaorts



GENERALITAT
VALENCIANA

Índice

1. Introducción	2
2. Diferencias	2
3. Conclusiones	4

1. Introducción

En esta parte se pretende realizar una pequeña investigación sobre las diferencias a grandes rasgos entre la programación tradicional y la programación empleando Inteligencia Artificial.

Para ello vamos a emplear estos enfoques para lo que es tratar datos de cáncer de mama.

2. Diferencias

Entendemos a la programación tradicional como un conjunto de pasos ordenados secuencialmente con el fin de resolver un problema. Por su lado la IA, se basa en aprender a partir de datos ya existentes con el fin de poder resolver un problema concreto.

En este caso estamos tratando datos sobre cáncer de mama, en el enfoque de programación tradicional lo que hacemos es obtener distintas gráficas en base a los datos que tenemos, en este caso en concreto obtenemos dos gráficas, la del **mean radius** que indica la cantidad de casos (positivo o negativo) dado el radio medio del tumor y la del **mean texture** que mide la rugosidad de la textura del tumor. Al código de la programación tradicional le hemos hecho este cambio:

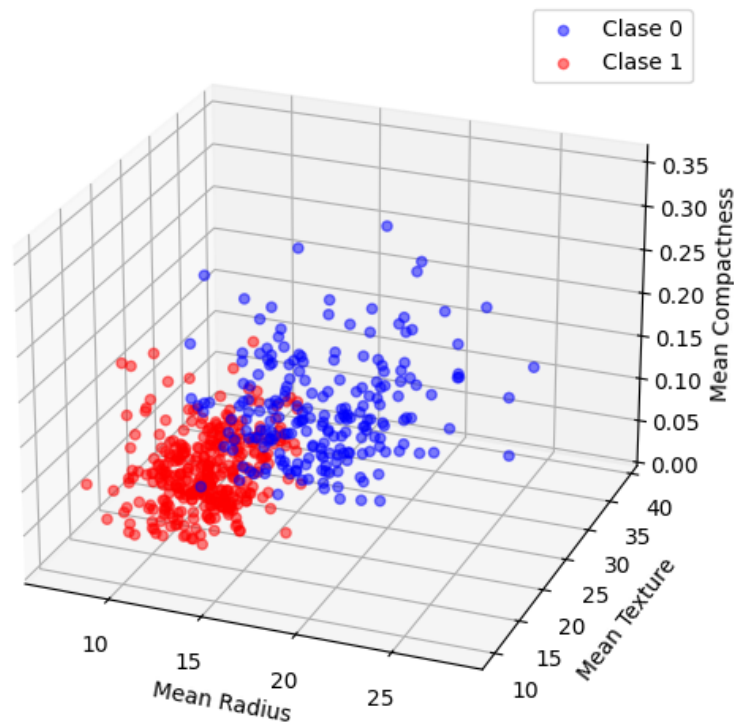


Figura 1: Diagrama de dispersión en 3D que relaciona las tres características

Por su lado, el código proporcionado para la versión IA no realiza esto, si no que crea un clasificador para poder predecir en base a los datos con los que ha entrenado si se trata de un caso de cáncer de mama o no. Para poder conseguir su cometido emplea el

RandomForestClassifier, se trata de un clasificador que usa los Trees para clasificar las entradas, finalmente se hace una especie de votación y se elige lo que la mayoría haya votado. Este enfoque funciona muy bien ya que tiene una buena tasa de acierto, hemos hecho pruebas con distintos números de árboles, estos han sido los resultados probando con un distinto número de árboles:

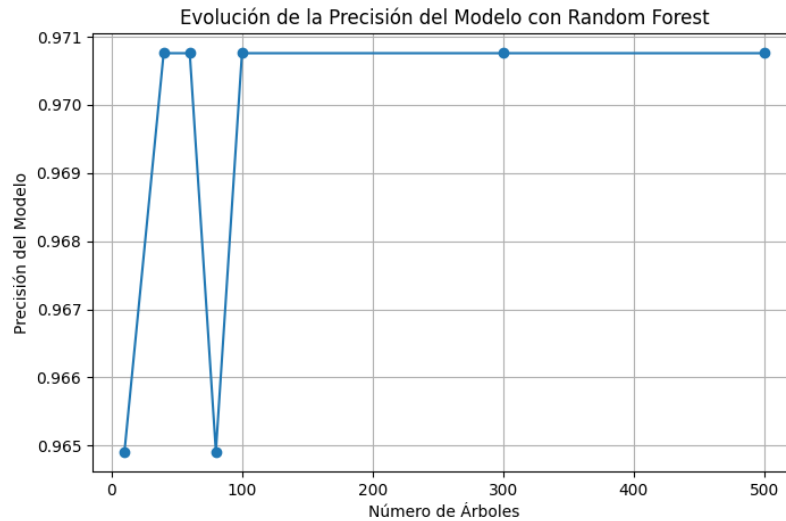


Figura 2: Tasa de acierto de RandomTreeClassifier para distinta cantidad de estimadores

Vemos que con pocos árboles ya da una buena tasa de acierto, y a partir de cierto número se mantiene estable, por lo que es inconveniente seguir aumentando.

Por otro lado, tenemos la matriz de confusión, se trata de una tabla que nos permite ver visualmente para ver como se ha comportado el modelo frente a ciertos datos, en nuestro caso es la siguiente:

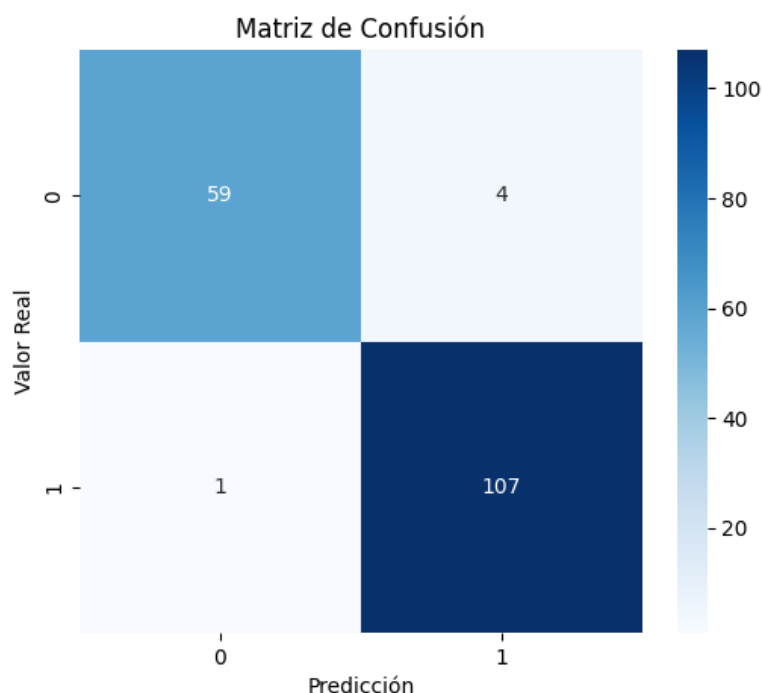


Figura 3: Matriz de confusión para el dataset de test

Para explicar el funcionamiento de la matriz de confusión emplearemos la imagen anterior, tenemos dos ejes, el “eje x” que se corresponde con lo que el modelo predice y el “eje y”, además, tanto el valor real como el predicho será 0 o 1, es decir, pertenecerá a una clase u otra. En la siguiente tabla vemos una explicación de esto:

Real	Predicción		Descripción
	Clase 0	Clase 1	
Clase 0	Verdadero Negativo (TN)	Falso Positivo (FP)	Correcto: 0 real y 0 predicho
Clase 1	Falso Negativo (FN)	Verdadero Positivo (TP)	Correcto: 1 real y 1 predicho

Cuadro 1: Matriz de confusión con explicación de cada celda.

Entonces, lo ideal es optimizar los falso positivos como negativos, ya que si nos fijamos en la imagen hay un caso donde el valor real es 1 (tiene cáncer de mama) y el modelo ha indicado 0 (no tiene cáncer de mama).

3. Conclusiones

En mi opinión, no veo aún el día en el que la IA quite el trabajo a un programador, ingeniero, etc. Entiendo que al final será más como una herramienta que nos facilite las cosas, además, creo que en ciertos campos será muy complicado que su uso se extienda, por ejemplo, cuando estamos hablando de campos donde hay vidas humanas implicadas puede ser muy útil, pero claro, en caso de que el algoritmo de IA falle también puede

resultar muy costoso, o también en campos donde se trata con dinero real también creo que es muy poco probable, en dichos campos creo que el enfoque será más de programación tradicional pero sin dejar de lado el uso de la IA para facilitar distintas labores.

Por lo que, a los “programadores” tradicionales aún les queda mucho tiempo, pero es importante que aprendamos a usar estas herramientas para facilitarnos el trabajo