



UNIVERSIDAD DE GRANADA

INTELIGENCIA DE NEGOCIO
GRADO EN INGENIERÍA INFORMÁTICA

PRÁCTICA 1

RESOLUCIÓN DE PROBLEMAS DE CLASIFICACIÓN Y ANÁLISIS
EXPERIMENTAL.

Autor

José María Sánchez Guerrero

Rama

Computación y Sistemas Inteligentes



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍAS INFORMÁTICA Y DE
TELECOMUNICACIÓN

CURSO 2020-2021

Índice

1. Introducción	2
2. Procesado de datos	2
3. Configuración de algoritmos	2
4. Resultados obtenidos	2
5. Interpretación de resultados	2
6. Bibliografía	2

1. Introducción

En este trabajo vamos a analizar el comportamiento de distintos algoritmos de clasificación en el problema propuesto. Disponemos de un dataset, llamado "*Mammographic Mass dataset*", en el cual se desea predecir el tipo de tumor (benigno o maligno) en una serie de mamografías realizadas para un estudio sobre el cáncer de mama. Este estudio lo vamos a realizar gracias a los siguientes atributos proporcionados en el dataset:

- **BI-RADS.** Este parámetro representa un control de calidad de las mamografías. Consta de 7 categorías distintas, en las que, cuanto más alto sea el valor, hay una mayor probabilidad de que sea maligno.
- **Edad** del paciente.
- **Forma de la masa.** Dependiendo de como sea la masa anormal detectada, se clasifica como **Redondeada**, **Ovalada**, **Lobulada**, **Irregular** ó **No definida**.
- **Margen de masa.** Circumscribed = 1, microlobulated = 2, obscured = 3, ill-defined = 4, spiculated = 5 (nominal).
- **Densidad de la masa.** Valores entre 1 y 4, siendo 1 la más alta y 4 contenido graso (no tumoral).
- **Severidad.** Es el atributo que se desea predecir, es decir, si es un tumor benigno o maligno.

En el dataset hay datos de 961 pacientes, sin embargo, nos gustaría dejar un porcentaje para validar el modelo y así ver cómo va entrenando los datos.

2. Procesado de datos

3. Configuración de algoritmos

4. Resultados obtenidos

5. Interpretación de resultados

6. Bibliografía