Actividad: Clasificación con máquina de vectores de soporte y random forest

**Objetivos**

Mediante esta actividad se pretende que ponga en práctica la creación de modelos basados en máquinas de vector de soporte y random forest. El objetivo es comprender de forma práctica con un problema determinado las diferencias que existen a la hora de entrenar los diferentes modelos.

* Comprender el tratamiento de datos que hay que realizar para los modelos de clasificación.
* Entender y aplicar el algoritmo de Random Forest a un problema de clasificación.
* Entender y aplicar el algoritmo de Support Vector Machine a un problema de clasificación.
* Evaluar y analizar los resultados de los clasificadores.
* Investigar la aplicación de los modelos de clasificación a problemas reales.

**Descripción de la actividad**

Debes realizar la actividad en el Notebook adjunto. Debes ir completando el código solicitado, así como responder a las preguntas que se van realizando.

**Tareas que se deben realizar**

* Análisis descriptivo de los datos:
  + Debe completarse el código solicitado y responder a las preguntas. Todo ello en el notebook dado como base.
* Clasificación:
  + Debe completarse el código solicitado y responder a las preguntas. Todo ello en el notebook dado como base.
* Investigación:
  + Buscar un artículo científico (https://scholar.google.es/) con un caso de uso de clasificación empleando una de las dos técnicas (o ambas) vistas en la actividad. Los artículos deben estar en revistas científicas, y deben ser posteriores a 2015. No debe utilizar técnicas de Deep Learning.
  + Para el artículo indicar:
    - Objetivo: cuál es el objetivo de la investigación, es decir a qué problema real está aplicando la clasificación.
    - Cómo utilizan las técnicas de clasificación, si realizan alguna adaptación de los algoritmos indicarse.
    - Principales resultados de la aplicación y de la investigación.

**Rúbrica**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Descripción | Puntuación máxima  (puntos) | Peso  % |
| Criterio 1 | Realiza el análisis descriptivo de los datos de manera adecuada | 2 | 20% |
| Criterio 2 | Realiza adecuadamente la aplicación del algoritmo de Support Vector Machine | 3 | 30% |
| Criterio 3 | Realiza adecuadamente la aplicación del algoritmo de Random Forest | 3 | 30% |
| Criterio 4 | El artículo de investigación cumple los criterios solicitados y analiza correctamente el trabajo. | 2 | 20% |
|  |  | **10** | **100 %** |

**Entrega:** Notebook Python.