Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

*Digital Object Identifier 10.1109/ACCESS.2022.Doi Number*

Predecir el origen de un vino a partir de sus componentes químicos

**Daniel Felipe Ospina Pérez1, Sebastián Castellanos Mejía1**

1Departamento de Ingeniería de Sistemas, Universidad de Antioquia, Medellín, Colombia

Daniel Felipe Ospina Pérez (e-mail: daniel.ospinap@udea.edu.co). Sebastián Castellanos Mejía (e-mail: sebastian.castellanosm@udea.edu.co).

**RESUMEN** En este informe, se aborda el problema de clasificación de diferentes tipos de vinos utilizando la base de datos de vinos de la Universidad de California en Irvine - UCI. Se describen las variables de entrada, la variable a predecir y se exploran las características del conjunto de datos. Además, se realiza una revisión de la literatura sobre trabajos relacionados con el mismo problema de clasificación.

1. **INTRODUCCIÓN**

El problema de clasificación de vinos tiene aplicaciones significativas en la industria vitivinícola y en la enología. El objetivo es desarrollar un modelo de aprendizaje automático que pueda clasificar los vinos en diferentes categorías basadas en sus características químicas. En este informe, se analizarán las variables de entrada, la variable a predecir y se revisarán trabajos relacionados.

* 1. ***DESCRIPCIÓN DEL PROBLEMA DE PREDICCIÓN*** El problema de predicción que se aborda implica clasificar diferentes tipos de vino en función de sus propiedades químicas.Esta tarea consiste en identificar el tipo de vino al que pertenece la muestra en función de sus propiedades químicas específicas. Los tres tipos de vino en cuestión se denominan “Tipo 1”, “Tipo 2” y “Tipo 3”.
  2. ***CAMPO DE APLICACIÓN***

~~Esta~~ ~~cuestión~~ de clasificación es relevante para el campo de la enología (el estudio del vino) y la industria del vino. Clasificar los vinos en función de sus propiedades químicas podría tener aplicaciones prácticas en la producción y comercialización del vino. Por ejemplo, un enólogo o una bodega puede utilizar un modelo de clasificación para identificar y etiquetar automáticamente los vinos en función de su calidad o características distintivas.

* 1. ***CLASIFICACIÓN O REGRESIÓN***

En este caso, se resuelve un problema de clasificación. La razón es que se asigna cada muestra (cada botella de vino en el conjunto de datos) a una de tres categorías distintas: 'Tipo 1', 'Tipo 2' o 'Tipo 3'. Las tareas de clasificación implican predecir la pertenencia a una

categoría o clase particular, a diferencia de las tareas de regresión, en las que se predice un valor numérico continuo.

* 1. ***VARIABLES***

En la tabla 1 se describen las variables de las muestras.

TABLA 1

la tabla proporciona información sobre las variables en el conjunto de datos, incluyendo su rol y tipo. Los roles de "Objetivo" y "Característica" son importantes para entender cómo se utilizarán estas variables en un problema de clasificación

|  |  |  |
| --- | --- | --- |
| Nombre | Rol | Tipo |
| class | Objetivo | Categórica |
| Alcohol | Característica | Continuo |
| Malicacid | Característica | Continuo |
| Ash | Característica | Continuo |
| Alcalinity\_of\_ash | Característica | Continuo |
| Magnesium | Característica | Entero |
| Total\_phenols | Característica | Continuo |
| Flavanoids | Característica | Continuo |
| Nonflavanoid\_phenols | Característica | Continuo |
| Proanthocyanins | Característica | Continuo |
| Color\_intensity | Característica | Continuo |
| Hue | Característica | Continuo |
| 0D280\_0D315\_of\_diluted\_ wines | Característica | Continuo |
| Proline | Característica | Entero |

La base de datos de vinos de UCI no contiene valores

faltantes, lo que significa que no se requiere ningún proceso de imputación de datos o eliminación de instancias con datos faltantes en este caso.

* 1. ***BÚSQUEDA DE ARTÍCULOS RELACIONADOS***

Se realizó una búsqueda de artículos relacionados con el problema de clasificación de vinos utilizando la base de datos de vinos de UCI. En la tabla 2, se presentan resúmenes de los trabajos encontrados.

TABLA 2

Resúmenes de Trabajos Relacionados

|  |  |
| --- | --- |
| Titulo | Resumen |
| Wine Quality Prediction Using Machine Learning | En este artículo, se aborda el problema de predecir la calidad del  vino utilizando la base de datos de vinos de UCI |
|  |  |