

Ejercicio practico (a)

Echeverry Luis (2420161006)
2420161006@estudiantesunibague.edu.co
Inteligencia Artificial
Programa de Ingeniería Electrónica
Facultad de Ingeniería
Universidad de Ibagué
Ibagué - Tolima
Semestre 2020A
Presentado al ingeniero Jose Armando Fernández Gallego

1. Planteamiento

Haciendo uso de los conocimientos adquiridos en la utilización y la implementación de la red neuronal del capítulo dos del libro guía, proponga una base de datos diferente a la usada previamente (Iris), grafique las variables, realice la interpretación de los datos obtenidos y concluya: ¿Es posible usar una red neuronal para este caso?, ¿se debe mejorar el sistema o el tipo de datos ingresados a la red?

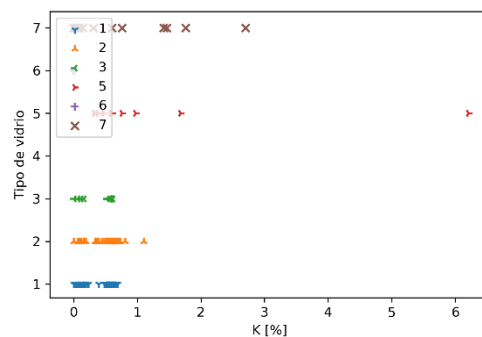
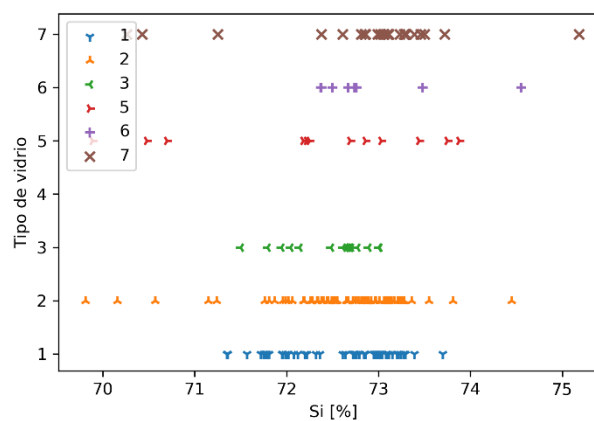
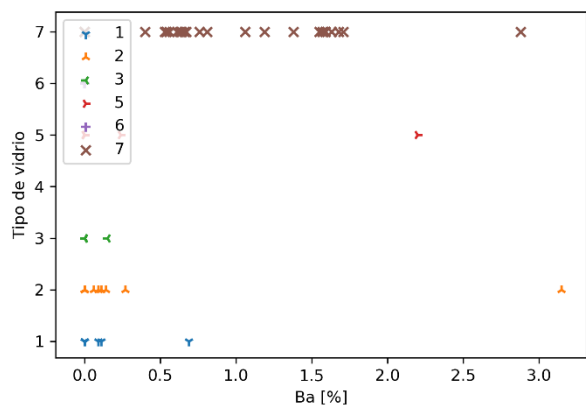
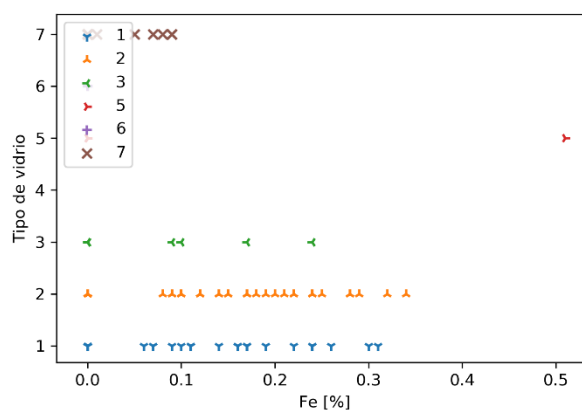
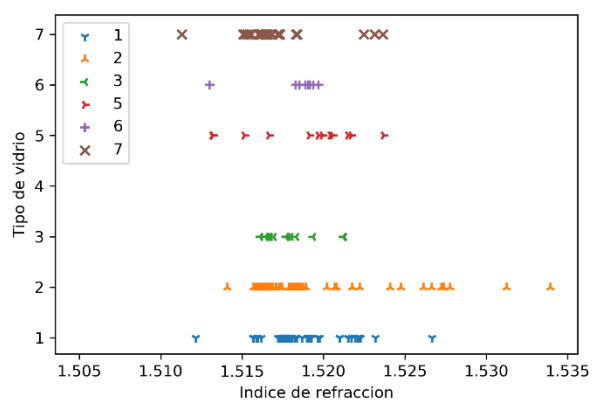
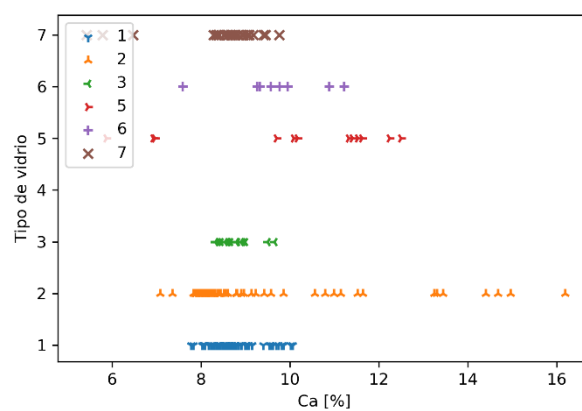
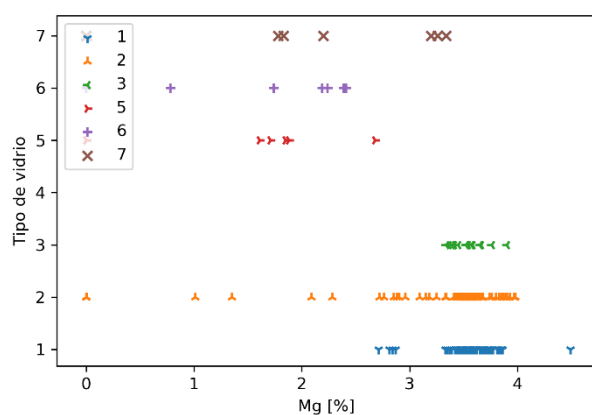
2. Desarrollo

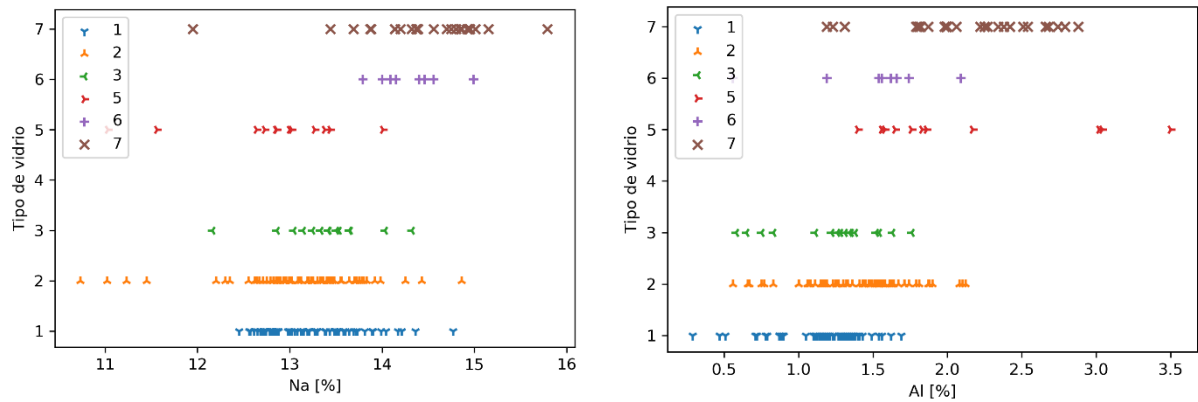
Para comenzar, se inició realizando la búsqueda de una base de datos que se ajustara correctamente con los requerimientos planteados. Para esto, fue suministrado por parte del docente un repositorio perteneciente a la universidad de california, Irvine. El cual contiene un gran número de bases de datos, utilizadas especialmente en aplicaciones de machine learning e inteligencia artificial. Además, se considero que, dentro de todas las bases de datos presentes en el repositorio, se seleccionaran únicamente las clasificadas como predeterminadas para la clasificación, ya que este es el objetivo principal que se tiene.

Considerando todos los requerimientos anteriormente planteados, se seleccionó la base de datos denominada glass, la cual contiene un número total de 214 muestras de 7 tipos diferentes de vidrios. Las cuales, a su vez, tienen 10 parámetros que fueron evaluados al momento de realizar las pruebas necesarias para crear cada una de las muestras. Dentro de los parámetros que contiene la base de datos, algunos de los mas relevantes son el índice de refracción de cada uno de los fragmentos de vidrio (RI), porcentaje de magnesio (Mg) y el porcentaje de potasio (K). estos dos últimos fueron los utilizados en este caso para realizar la grafica de los valores que se verán más a delante.

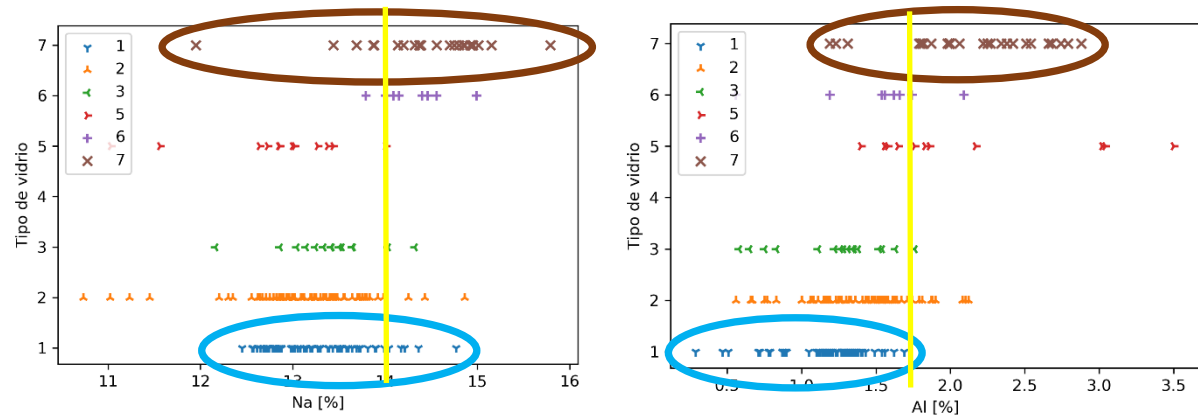
2.1 Grafica de las variables de interés

Para realizar la grafica de las variables con las cuales se desea hacer el análisis y la separación de las muestras, primero de necesito realizar una selección de dos variables que fueran linealmente separables, ya que para este caso en particular este requerimiento es esencial para realizar la clasificación de las muestras. Las diferentes graficas realizadas se muestran a continuación.





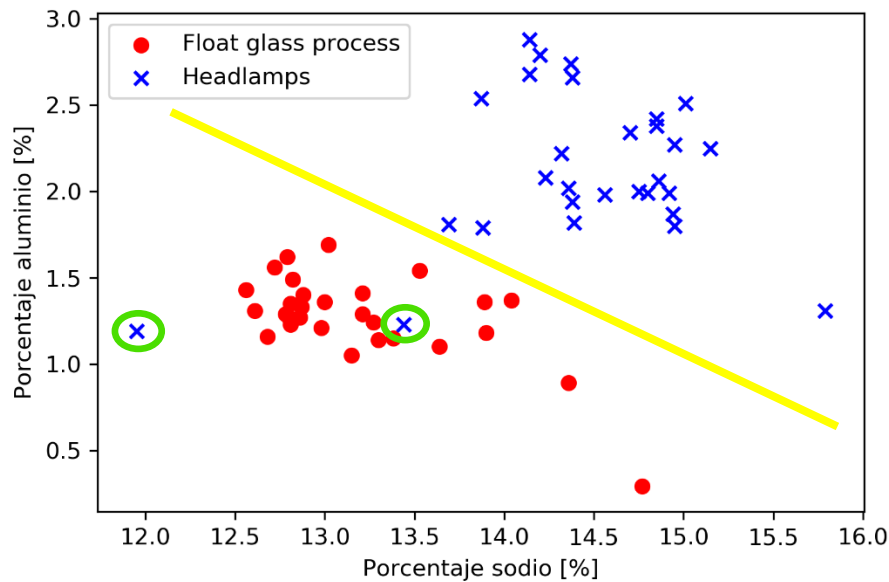
De las gráficas anteriores podemos analizar que existen 2 que juegan un papel fundamental en el objetivo de separar dos clases de vidrios, que son la del porcentaje de aluminio (Al) y el porcentaje de sodio (Na), con lo cual podemos observar que para un gran número de datos de la clase 1 (Float glass process) y de la clase 7 (Headlamps), existe la posibilidad de establecer una separación acertada de estos dos tipos de vidrios basados en estos parámetros. Para aclarar más la separación de las dos clases se decidió delimitar la misma con una línea de color amarillo y encerrar ambas clases seleccionadas con un ovalo del color respectivo de cada grupo de datos como se muestra a continuación.



Aunque la separación de los datos no es perfecta, se puede analizar que el numero de muestras que no entran dentro de la separación es muy reducido, por esta razón se procedió a realizar la grafica de las dos variables para poder realizar las conclusiones pertinentes.

2.2 Análisis de datos

Al graficar las dos variables obtenidas anteriormente se obtuvo lo siguiente:



Como se puede evidenciar, los dos conjuntos de datos están separados por una línea de color amarillo, de esta manera se puede afirmar que el conjunto de datos es linealmente separable. Por otra parte, existen algunos datos que se encuentran fuera de la limitación establecida, encerrados con un óvalo color verde, los cuales podrían llegar a traer problemas más adelante al momento de utilizar alguno de los algoritmos de separación vistos en clase.

3. Conclusiones

- Para este caso en particular, se podría considerar utilizar una red neuronal para realizar la separación de estos dos tipos de vidrio, teniendo en cuenta que existen algunas muestras que se podrían llegar a considerar ruido de medición y de esta manera obviarlas durante el proceso de entrenamiento de la red.
- Si existiera la posibilidad de mejorar la base de datos presente, se debería comenzar por ampliar el número de muestras tomadas. Ya que a pesar de que la base de datos tiene un número total de 213 muestras, como existen 7 tipos de vidrios diferentes, el número de las mismas que le pertenece a cada uno de estos tipos es más reducido.