

Inteligencia Artificial

Practica 5: Machine Learning

Profesor: Enrique Francisco Soto Astorga
Ayudante de Laboratorio: Ricardo Enrique Pérez Villanueva

April 4, 2025

1 Proyecto

En este proyecto vamos a crear un modelo de predicción del clima.

1.1 Predicción del Clima

Vamos a utilizar un dataset con datos climáticos para crear 2 modelos, un clasificador y un modelo de predicción. El dataset contiene 5 datos: Temperatura, Humedad, Velocidad del Viento, Cobertura de Nubes y Presión. Finalmente, tiene una columna de etiquetas, la cual nos dice que dependiendo del valor de las condiciones anteriormente mencionadas, va llover o no.

2 Implementación

Vamos a implementar 2 modelos de Machine Learning, uno de regresión para predecir valores, y otro de clasificación. Para el modelo de clasificaciones, vamos a crear un modelo que, dado las 5 características, haga una clasificación si va a llover o no. La parte mas importante es ver la forma de usar las 5 características para obtener una clasificación. El modelo de clasificación puede ser o K-Nearest Neighbors o Decision Trees, después de realizar la clasificación, veremos cual es la característica mas importante, calculando la "Feature Importance"¹, es recomendable graficar la Feature Importance para tener una mejor comprensión de los datos.

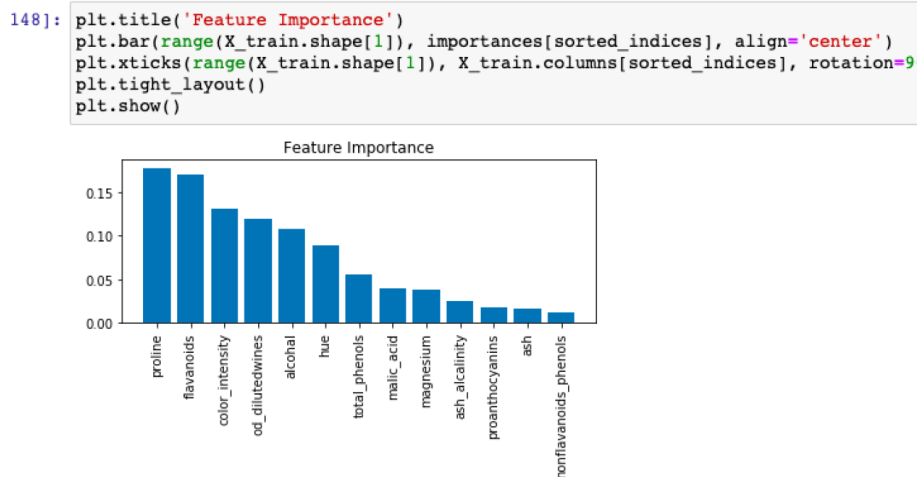


Figure 1: Ejemplo de Feature Importance, noten como se ven las gráficas.

Después, vamos a crear un modelo de regresión para predecir si lloverá o no. Esto se logra usando las 5 características como X y Y, como un valor de 0 y 1, es decir, si lloverá o no. Esto se logra fácilmente con la función `replace`.

3 Entrega

- Se recomienda que se haga mediante un Notebook en Google Collab

¹La forma para calcular la Feature Importance es diferente para cada modelo.

- Se recomienda que utilicen la implementación que se vio en clase, es decir, compartir con un Link en el Classroom su archivo de Google Collab donde hicieron la Red Neuronal, recuerden dar permisos de ejecución.
- Se va a calificar la precisión del modelo, como se menciona arriba, que la precisión de ambos modelos sea al menos del 0.80.
- Pueden usar regresión Lineal o Polinomial para el regresor, y para la clasificación pueden usar K-Nearest Neighbors o Decision Trees.

4 Documentación y materiales útiles:

A continuación se les ofrece fuentes de información que les van a ayudar a realizar su practica

- Como calcular la Feature Importance
<https://machinelearningmastery.com/calculate-feature-importance-with-python/8>
- Documentación de K-Nearest Neighbors en sci-kit-learn:
<https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html>
- Documentación de Decision Trees:
<https://scikit-learn.org/stable/modules/tree.html>
- Ejemplo de Decision Tree:
<https://www.datacamp.com/tutorial/decision-tree-classification-python>
- Ejemplo de Regresión Logística en 3 Dimensiones. <https://www.kaggle.com/code/bryceschultz26/3d-regression>