



**Congreso Internacional sobre la
Enseñanza y Aplicación de las Matemáticas**

PREDICCIÓN DE ESPECIE DE PLANTA

"PLANT TOON"

Autores:

**Salinas Morales Diego Michel
Jiménez Gutiérrez Alberto
Maldonado Tapia Luis Felipe**

Asesor:

Judith Mayte Flores López

Artículo incluido en la publicación electrónica Memorias del Congreso ISSN 2448-7945 | Mayo 2024



Departamento de
Matemáticas





Objetivos:

- **Identificar** cómo variables de tamaño y anchura del pétalo y sépalo predicen la especie de las plantas.
- **Recopilar** datos sobre el crecimiento de las plantas según sus medidas.
- **Aplicar** técnicas de análisis estadístico para identificar patrones y relaciones entre estas variables y el crecimiento de las plantas.
- **Modelar y predecir** la especie de las plantas en función de las condiciones ambientales utilizando un enfoque de árbol de decisión
- **Contribuir** al avance de nuestro entendimiento sobre la biología de las plantas y su interacción con el medio ambiente.



Introducción:

En la intersección de la botánica y la tecnología, surge “Plant Toon”, un proyecto innovador diseñado para revolucionar la forma en que entendemos y clasificamos la diversidad vegetal. Utilizando un modelo avanzado de árbol de decisiones desarrollado en Python, Plant Toon no solo identifica especies vegetales con una precisión sin precedentes, sino que también abre un nuevo camino hacia la conservación y el estudio de la flora.

El corazón de Plant Toon es su algoritmo de aprendizaje automático, que se alimenta de una rica base de datos de características botánicas. Este modelo no es solo un testimonio de la programación sofisticada, sino también de una comprensión profunda de las complejidades de la vida vegetal. Al implementar técnicas de clasificación basadas en la entropía y la ganancia de información, Plant Toon distingue entre innumerables especies, ofreciendo a los investigadores y entusiastas una herramienta poderosa y accesible.





Metodología:

El proyecto “PLANT TOON” se basa en una metodología rigurosa para garantizar la precisión en la identificación de especies vegetales. La metodología se divide en varias etapas clave:

- **Recolección de Datos:** Se recopila un conjunto de datos exhaustivo que incluyen características botánicas de diversas especies vegetales. Este conjunto de datos sirve como base para el entrenamiento y la validación del modelo de árbol de decisiones.
- **Preprocesamiento de Datos:** Los datos recopilados se limpian y se normalizan para asegurar la consistencia. Las características relevantes se extraen y se codifican de manera adecuada para su uso en el modelo de clasificación.
- **Construcción del Modelo:** Utilizando la biblioteca scikit-learn de Python, se construye un modelo de árbol de decisiones. Se seleccionan los atributos más significativos basándose en la ganancia de información y la medida de impureza (índice Gini o entropía)¹².
- **Entrenamiento del Modelo:** El modelo se entrena con una parte del conjunto de datos. Durante esta fase, el algoritmo aprende a clasificar las especies basándose en las reglas de decisión que surgen de los datos.
- **Validación y Pruebas:** El modelo se valida y se prueba con una parte del conjunto de datos que no se utilizó durante el entrenamiento. Esto ayuda a evaluar la precisión y la generalización del modelo.
- **Optimización:** Se realizan ajustes en el modelo, como la poda del árbol, para evitar el sobreajuste y mejorar la precisión².
- **Implementación:** Una vez optimizado y validado, el modelo se implementa en una aplicación o servicio que puede ser utilizado por los usuarios para identificar especies vegetales.



Desarrollo:

- Importamos las librerías necesarias para el correcto funcionamiento de nuestro código

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score
import pickle
```



Desarrollo:

- Leemos el archivo “Iris.csv” para así, alimentar a nuestro modelo con los datos necesarios.

```
# Cargar los datos  
data = pd.read_csv('C:/Users/salin/Downloads/Iris.csv')
```



Desarrollo:

- Eliminamos la columna “Id” porque esta no es relevante en nuestro modelo.

```
# Eliminar la columna 'Id'  
data.drop(['Id'], axis=1, inplace=True)
```



Desarrollo:

- Separamos nuestras variables independientes (características) y nuestra variable dependiente (objetivo) del DataFrame 'data', lo que nos permite usarlas para entrenar nuestro modelo de machine learning.

```
# Definir las variables  
X = data[['SepalLengthCm', 'SepalWidthCm', 'PetalLengthCm', 'PetalWidthCm']]  
y = data['Species']
```



Desarrollo:

- Dividimos nuestro conjunto de datos en conjuntos de entrenamiento y prueba, lo que nos permite entrenar nuestro modelo en una parte de los datos y evaluar su rendimiento en otra parte independiente.

```
# Dividir los datos en entrenamiento y prueba  
X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.20, random_state=0)
```



Desarrollo:

- Establecemos un modelo de clasificación basado en árboles de decisiones y lo entrenamos con el conjunto de datos de entrenamiento, utilizando la entropía como criterio para medir la calidad de las divisiones en el árbol de decisiones.

```
# Establecer el modelo de clasificación árbol de decisiones  
dc_tree = DecisionTreeClassifier(criterion='entropy')  
dc_tree.fit(X_train, y_train)
```



Desarrollo:

- Generamos predicciones utilizando el modelo de árbol de decisiones entrenado, y las almacenamos en la variable `y_pred` para su posterior evaluación.

```
# Realizar predicciones  
y_pred = dc_tree.predict(X_test)
```



Desarrollo:

- Generamos predicciones utilizando el modelo de árbol de decisiones entrenado, y las almacenamos en la variable `y_pred` para su posterior evaluación.

```
# Evaluar el modelo
accuracy = accuracy_score(y_test, y_pred)
print("Accuracy:", accuracy)
```



Desarrollo:

- Calculamos la precisión del modelo de clasificación y la imprimimos en la consola para que así evaluar el rendimiento del modelo. La precisión nos dice qué tan bien nuestro modelo puede predecir las etiquetas correctas en el conjunto de datos de prueba.

```
# Evaluar el modelo  
accuracy = accuracy_score(y_test, y_pred)  
print("Accuracy:", accuracy)
```



Desarrollo:

- Guardamos nuestro modelo entrenado en un archivo binario utilizando la biblioteca pickle, lo que nos permite cargar y utilizar el modelo en el futuro sin necesidad de volver a entrenarlo desde cero.

```
#Generamos el .pkl  
with open('flores.pkl', 'wb') as file:  
    pickle.dump(dc_tree, file)
```



Desarrollo:

- Guardamos nuestro modelo entrenado en un archivo binario utilizando la biblioteca pickle, lo que nos permite cargar y utilizar el modelo en el futuro sin necesidad de volver a entrenarlo desde cero.

```
#Generamos el .pkl  
with open('flores.pkl', 'wb') as file:  
    pickle.dump(dc_tree, file)
```



Resultados

Desarrollamos un sistema el cuál garantiza con una efectividad del 0.9166, es decir, capaz de predecir con una precisión alta la especie que es nuestra flor, únicamente utilizando datos como la altura y grosor de nuestros pétalos y sépalo.

Conclusiones

Este proyecto ha desarrollado un modelo de aprendizaje automático capaz de predecir la especie de una flor con alta precisión. El modelo se puede utilizar para identificar flores desconocidas o para clasificar grandes colecciones de flores. El modelo también se puede utilizar para desarrollar nuevas aplicaciones, como una herramienta para la identificación de flores en dispositivos móviles.



Bibliografía

- A, D. B., & Tognetti, J. (s. f.). *Técnicas de análisis de crecimiento de plantas: su aplicación a cultivos intensivos*. http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1669-23142016000300007
- *Técnicas de análisis de crecimiento de plantas: su aplicación a cultivos intensivos*. (2017, 11 mayo). <https://digital.cic.gba.gob.ar/items/95e0d9d8-9009-4d34-9176-1dac2df3aa1e>
- CK-12 Foundation. (s. f.). CK-12 Foundation. <https://flexbooks.ck12.org/cbook/ck-12-conceptos-biologia/section/9.13/primary/lesson/crecimiento-de-las-plantas/>
- Cherlinka, V. (2023, 31 octubre). Etapas de crecimiento de las plantas: rol y funciones. *EOS Data Analytics*. <https://eos.com/es/blog/etapas-de-crecimiento-de-una-planta/>

Gracias por su atención

Plant Toon ©



Departamento de
Matemáticas

