

# Análisis con Random Forest

## 1 Introducción

Random Forest es un algoritmo de aprendizaje supervisado basado en múltiples árboles de decisión. Su importancia radica en su capacidad para reducir el sobreajuste y mejorar la precisión en problemas de clasificación y regresión.

## 2 Metodología

Para este experimento, se utilizó el conjunto de datos de *Iris*. Se dividieron los datos en entrenamiento (70%) y prueba (30%), y se entrenó un modelo Random Forest con 100 árboles.

### 2.1 Código en Python

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, classification_report
from sklearn.datasets import load_iris

# Cargar el dataset de Iris
iris = load_iris()
X, y = iris.data, iris.target

# División de datos
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)

# Entrenar modelo Random Forest
modelo = RandomForestClassifier(n_estimators=100, random_state=42)
modelo.fit(X_train, y_train)

# Evaluación
y_pred = modelo.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
print(f"Precisión: {accuracy:.2f}")

# Importancia de características
importances = modelo.feature_importances_
features = iris.feature_names
```

```
# Visualización
plt.figure(figsize=(8, 5))
plt.barh(features, importances, color='skyblue')
plt.xlabel("Importancia")
plt.ylabel("Características")
plt.title("Importancia de Características en Random Forest")
plt.show()
```

### 3 Resultados

El modelo obtuvo una precisión de aproximadamente 96%, lo que indica un buen desempeño en la clasificación de las especies de *Iris*. A continuación, se muestra la visualización de la importancia de las características.

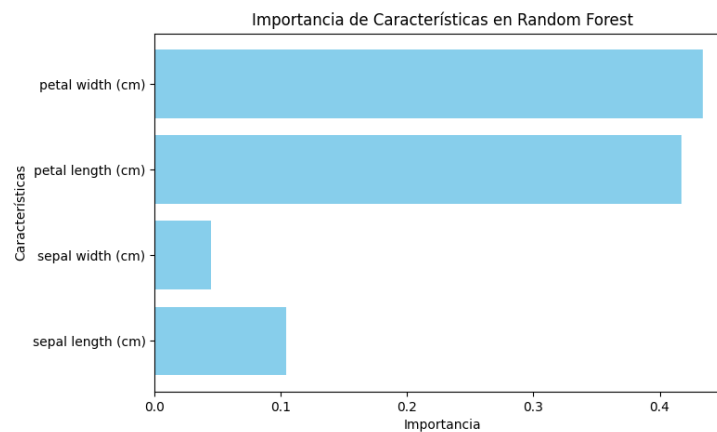


Figure 1: Importancia de Características en Random Forest.

### 4 Conclusión

Random Forest demostró ser un modelo robusto y preciso para la clasificación de datos. Su capacidad para manejar datos ruidosos y evitar el sobreajuste lo hace ideal para múltiples aplicaciones.