



Proyecto Final

Universidad de Guadalajara, Centro Universitario de Ciencias exactas e Ingenierías

Alumno:

Cárdenas Pérez Calvin Cristopher Código: 214798706

Carrera:

Ingeniería en Computación (INCO)

Materia:

Seminario de Solución de Problemas de Inteligencia artificial II

Maestro:

Dr. Diego Oliva

Ciclo:

2024 A

Sección:

D04

Feca de entrega:

15 de mayo del 2024

Introducción

En un mundo donde la conservación de la biodiversidad y la gestión de los ecosistemas son temas de gran relevancia, el uso de técnicas de aprendizaje automático se ha vuelto cada vez más importante. Este proyecto tiene como objetivo aplicar estas técnicas para clasificar especies en categorías predefinidas. Utilizando conjuntos de datos que abarcan una amplia gama de características biológicas y ecológicas, buscamos desarrollar modelos predictivos precisos que puedan contribuir significativamente a la comprensión y gestión de la biodiversidad.

El enfoque principal es desarrollar modelos que puedan identificar y clasificar correctamente las especies en función de sus características. Estos modelos pueden ser fundamentales para una variedad de aplicaciones, desde la conservación de la biodiversidad hasta la gestión de los recursos naturales. Al utilizar algoritmos de aprendizaje automático, como la regresión logística y las redes neuronales, esperamos obtener resultados que sean tanto precisos como prácticos para la conservación y gestión de los ecosistemas.

Resultados y Análisis

Zoo Dataset

```
=== Resultados para Zoo Dataset: ===

Modelo: Regresion Logistica
Metrics:
Accuracy: 95.24%
Precision: 92.86%
Recall: 95.24%
Specificity: 100.00%
F1 Score: 93.65%

Modelo: K-Vecinos Cercanos
Metrics:
Accuracy: 95.24%
Precision: 92.86%
Recall: 95.24%
Specificity: 100.00%
F1 Score: 93.33%

Modelo: Maquinas de Vectores de Soporte (SVM)
Metrics:
Accuracy: 95.24%
Precision: 92.86%
Recall: 95.24%
Specificity: 100.00%
F1 Score: 93.33%

Modelo: Naive Bayes
Metrics:
Accuracy: 85.71%
Precision: 91.67%
Recall: 85.71%
Specificity: 100.00%
F1 Score: 85.24%

Modelo: Red Neuronal
Metrics:
Accuracy: 95.24%
Precision: 92.86%
Recall: 95.24%
Specificity: 100.00%
F1 Score: 93.33%
```

Los resultados para el Zoo Dataset muestran un rendimiento general sólido en todos los modelos. Se observa una consistencia en la capacidad de los modelos para clasificar las instancias correctamente. Los modelos de Regresión Logística, K-Vecinos Cercanos, Máquinas de Vectores de Soporte (SVM) y Red Neuronal exhiben niveles altos y similares de precisión, recall y especificidad, indicando una capacidad robusta para clasificar tanto las instancias positivas como negativas en el conjunto de datos del zoo. Por otro lado, el modelo Naive Bayes

muestra una precisión ligeramente menor, lo que sugiere que podría tener dificultades para identificar algunas instancias positivas, aunque sigue demostrando una buena capacidad para identificar instancias negativas. En conjunto, estos resultados indican que los modelos son capaces de clasificar adecuadamente las instancias en el conjunto de datos del zoo, proporcionando una base sólida para la aplicación de estos modelos en escenarios de clasificación similares.

Zoo2 Dataset

```
=== Resultados para Zoo2 Dataset: ===  
  
Modelo: Regresión Logística  
Metrics:  
Accuracy: 100.00%  
Precision: 100.00%  
Recall: 100.00%  
Specificity: 100.00%  
F1 Score: 100.00%  
  
Modelo: K-Vecinos Cercanos  
Metrics:  
Accuracy: 88.89%  
Precision: 91.67%  
Recall: 88.89%  
Specificity: 100.00%  
F1 Score: 87.83%  
  
Modelo: Máquinas de Vectores de Soporte (SVM)  
Metrics:  
Accuracy: 100.00%  
Precision: 100.00%  
Recall: 100.00%  
Specificity: 100.00%  
F1 Score: 100.00%  
  
Modelo: Naïve Bayes  
Metrics:  
Accuracy: 100.00%  
Precision: 100.00%  
Recall: 100.00%  
Specificity: 100.00%  
F1 Score: 100.00%  
  
Modelo: Red Neuronal  
Metrics:  
Accuracy: 100.00%  
Precision: 100.00%  
Recall: 100.00%  
Specificity: 100.00%  
F1 Score: 100.00%
```

El conjunto de datos Zoo2 presenta un escenario en el que los modelos logran una clasificación casi perfecta de las especies animales. Todos los modelos, incluyendo Regresión Logística, K-Vecinos Cercanos, Máquinas de Vectores de Soporte (SVM), Naive Bayes y Redes Neuronales, alcanzan una precisión del 100% en todas las métricas evaluadas. Este resultado sugiere que las características presentes en el Zoo2 Dataset son altamente distintivas y fácilmente separables para cada especie animal en particular. Es importante tener en cuenta que, aunque estos resultados pueden parecer excepcionales, también podrían indicar un posible sobreajuste de los modelos a los datos de entrenamiento.

Zoo3 Dataset

```
=== Resultados para Zoo3 Dataset: ===  
  
Modelo: Regresion Logistica  
Metrics:  
Accuracy: 92.86%  
Precision: 92.86%  
Recall: 92.86%  
Specificity: 100.00%  
F1 Score: 92.86%  
  
Modelo: K-Vecinos Cercanos  
Metrics:  
Accuracy: 85.71%  
Precision: 85.71%  
Recall: 85.71%  
Specificity: 100.00%  
F1 Score: 85.71%  
  
Modelo: Maquinas de Vectores de Soporte (SVM)  
Metrics:  
Accuracy: 92.86%  
Precision: 92.86%  
Recall: 92.86%  
Specificity: 100.00%  
F1 Score: 92.86%  
  
Modelo: Naive Bayes  
Metrics:  
Accuracy: 92.86%  
Precision: 92.86%  
Recall: 92.86%  
Specificity: 100.00%  
F1 Score: 92.86%  
  
Modelo: Red Neuronal  
Metrics:  
Accuracy: 92.86%  
Precision: 92.86%  
Recall: 92.86%  
Specificity: 100.00%  
F1 Score: 92.86%
```

En el caso del conjunto de datos Zoo3, observamos un rendimiento sólido pero no perfecto por parte de los modelos. Todos los modelos muestran una precisión bastante alta en todas las métricas evaluadas, con una precisión generalmente superior al 90%. Sin embargo, ninguno de los modelos logra una clasificación perfecta, lo que sugiere que las especies presentes en el Zoo3 Dataset pueden tener características que son más difíciles de distinguir en comparación con el Zoo2 Dataset. A pesar de esto, los modelos todavía muestran un rendimiento considerablemente bueno, lo que indica que aún son capaces de capturar patrones importantes en los datos y realizar una clasificación efectiva de las especies animales.

Dataset combinado

```
=== Resultados para Dataset Combinado ===  
  
Modelo: Regresión Logística  
Metrics:  
Accuracy: 97.67%  
Precision: 98.26%  
Recall: 97.67%  
Specificity: 100.00%  
F1 Score: 97.71%  
  
Modelo: K-Vecinos Cercanos  
Metrics:  
Accuracy: 93.02%  
Precision: 93.49%  
Recall: 93.02%  
Specificity: 100.00%  
F1 Score: 92.92%  
  
Modelo: Maquinas de Vectores de Soporte (SVM)  
Metrics:  
Accuracy: 97.67%  
Precision: 98.26%  
Recall: 97.67%  
Specificity: 100.00%  
F1 Score: 97.71%  
  
Modelo: Naive Bayes  
Metrics:  
Accuracy: 90.70%  
Precision: 94.35%  
Recall: 90.70%  
Specificity: 100.00%  
F1 Score: 90.33%  
  
Modelo: Red Neuronal  
Metrics:  
Accuracy: 97.67%  
Precision: 98.06%  
Recall: 97.67%  
Specificity: 100.00%  
F1 Score: 97.70%
```

El Dataset Combinado muestra un rendimiento general sólido en todos los modelos, con altos niveles de precisión, recall y especificidad. Sin embargo, el modelo de Naive Bayes tiene un rendimiento ligeramente inferior en términos de precisión y F1 Score en comparación con los otros modelos.

Conclusión

Este proyecto me ha gustado es algo que ya llevamos como practicando con las practicas anteriores y a sido algo distinto en cuanto al potencial del aprendizaje automático para abordar desafíos en la clasificación de especies animales. A través de la experimentación con diferentes modelos y conjuntos de datos, he observado cómo la precisión de los modelos puede variar según la complejidad de las características de los datos. Este proceso me ha permitido apreciar la importancia de la selección adecuada de datos y la evaluación cuidadosa del rendimiento del modelo. A pesar de los desafíos, creo firmemente en el potencial de estas herramientas para contribuir significativamente a la conservación de la biodiversidad y la gestión de la vida animal en el futuro.

En cuanto al curso me gustaron las practicas unas más difíciles que otras tengo que admitirlo pero todas me fueron ayudando a desarrollarme más en diferentes campos y también aprendiendo de las mismas prácticas y esto me ayudo para poder terminar con este proyecto y con el curso en general.