

Introducción a la inteligencia artificial

Entregable # 2 - World Energy Consumption



**UNIVERSIDAD
DE ANTIOQUIA**

1 8 0 3

Carlos Medina C

PROFESOR

RAUL RAMOS

INTRODUCCIÓN A LA INTELIGENCIA ARTIFICIAL

Introducción:

El objetivo de este informe es predecir si un país tiene una alta o baja proporción de energía renovable en su producción total de energía. Se utilizará un conjunto de datos sobre el consumo de energía en todo el mundo. El conjunto de datos contiene varias variables, incluyendo la producción y el consumo de energía renovable, así como la producción y el consumo de otras fuentes de energía.

Se ha cambiado el problema al predecir si un país tiene o no un alto porcentaje de energía renovable en su producción total de energía. Por lo tanto, la métrica de desempeño que se utilizará ahora será la precisión, el recall, la F1-score y el área bajo la curva ROC (AUC), en lugar del error medio absoluto (MAE) utilizado anteriormente para predecir la tendencia del crecimiento de la producción y el consumo de energía a nivel mundial.

Metodología

- Carga de datos:

Primero, se cargaron los datos del archivo CSV en un DataFrame de Pandas. El conjunto de datos contiene información sobre la producción y el consumo de energía en todo el mundo para diferentes años.

Limpieza de datos:

A continuación, se realizó una limpieza de datos para eliminar los valores faltantes y las variables innecesarias. También se creó una variable de destino que indica si un país tiene una alta o baja proporción de energía renovable en su producción total de energía. Se utilizó una proporción del 50% para la separación entre alta y baja proporción de energía renovable.

Análisis exploratorio de datos:

Luego, se realizó un análisis exploratorio de datos para examinar la distribución de las variables y la relación entre ellas. Se utilizaron gráficos de barras y gráficos de dispersión para visualizar los datos. Se observó una fuerte correlación entre la producción y el consumo de energía renovable, así como entre la producción y el consumo de otras fuentes de energía.

Preprocesamiento de datos:

Antes de aplicar los algoritmos de clasificación, se normalizaron los datos utilizando la técnica MinMaxScaler para asegurar que todas las características tuvieran la misma escala. Luego, se seleccionaron las siguientes características: 'biofuel_share_energy', 'hydro_share_energy', 'solar_share_energy', 'wind_share_energy', 'renewables_share_energy', como variables predictoras (X) y la variable objetivo 'alta_renovable' como la variable a predecir (y).

División de Datos

Se dividió el conjunto de datos en conjuntos de entrenamiento y prueba, utilizando una proporción del 80% para el conjunto de entrenamiento y el 20% para el conjunto de prueba.

Evaluación del modelo:

Se ha utilizado el modelo de árbol de decisión y la máquina de vectores de soporte para predecir si un país tiene un alto porcentaje de energía renovable. A continuación se muestran los resultados de ambos modelos:

Resultados del modelo de árbol de decisión:

Accuracy: 1.0

Precision: 1.0

Recall: 1.0

F1 Score: 1.0

ROC AUC Score: 1.0

Resultados del modelo de máquina de vectores de soporte:

Accuracy: 0.9979925437338687

Precision: 0.9032258064516129

Recall: 0.875

F1 Score: 0.8888888888888888

ROC AUC Score: 0.9370658465991316

Se observa que ambos modelos tienen un alto desempeño, con un 100% de precisión para el modelo de árbol de decisión y un 99.8% de precisión para el modelo de máquina de vectores de soporte. Además, se puede observar que el modelo de máquina de vectores de soporte tiene un buen desempeño en términos de precisión, recall, F1 score y ROC AUC score.

Conclusiones:

En conclusión, se puede utilizar el modelo de árbol de decisión o la máquina de vectores de soporte para predecir si un país tiene o no un alto porcentaje de energía renovable en su producción total de energía. Además, se puede utilizar la precisión, recall, F1 score y ROC AUC score para evaluar el desempeño de los modelos.