

Rutas generadas para el sistema de transporte masivo Supervisado

Capítulo 1

Introducción

En este documento se describe el **Dataset** que fue generado como parte del proyecto de inteligencia artificial, el cual se encuentra basado en un sistema de transporte masivo, el **Dataset** que generamos en la Actividad 2 contiene información sobre las diferentes rutas entre estaciones, incluyendo los detalles como el tiempo del viaje, la línea que fue utilizada y si la ruta es óptima o no.

Estructura del Dataset

El **Dataset** contiene las siguientes columnas:

Columna	Descripción
Estación Origen	Nombre de la estación de origen.
Estación Destino	Nombre de la estación de destino.
Línea	Nombre de la línea de transporte que es utilizada para conectar las estaciones.
Tiempo (min)	Tiempo de viaje entre las estaciones, expresado en minutos.
Ruta Optima	Indica si la ruta es óptima (Si) o no (No) en función de un tiempo límite.

Descripción de los datos

- **Número total de las rutas generadas:** El número total de las rutas generadas incluye tanto las conexiones directas como las indirectas, incluyendo las rutas con múltiples líneas.

- **Criterio para determinar si una ruta es óptima:** Una ruta se considera óptima si el tiempo total de viaje es menor o igual a 60 minutos.
- **Conexiones indirectas:** El sistema genera las conexiones indirectas basadas en las reglas descritas en el código. Por ejemplo, si no hay una conexión directa entre Tatooine y Yavin IV, se evalúa si ya existe una conexión a través de una estación intermedia como Alderaan.

Dataset

Estación Origen	Estación Destino	Línea	Tiempo (min)	Ruta Óptima
Tatooine	Alderaan	Ruta 1	10	Si
Alderaan	Yavin IV	Ruta 2	20	Si
Yavin IV	Hoth	Ruta 3	30	Si
Hoth	Dagobah	Ruta 4	40	Si
Dagobah	Bespin	Ruta 5	50	Si
Bespin	Endor	Ruta 6	60	Si
Endor	Naboo	Ruta 7	70	No
Naboo	Coruscant	Ruta 8	80	No
Coruscant	Kamino	Ruta 9	90	No
Kamino	Mandalore	Ruta 10	100	No
Tatooine	Yavin IV	Ruta 1 + Ruta 2	30	Si
Alderaan	Hoth	Ruta 2 + Ruta 3	50	Si
Yavin IV	Dagobah	Ruta 3 + Ruta 4	70	No
Hoth	Bespin	Ruta 4 + Ruta 5	90	No
Dagobah	Endor	Ruta 5 + Ruta 6	110	No
Tatooine	Hoth	Ruta 1 + Ruta 2 + Ruta 3	60	Si
Alderaan	Dagobah	Ruta 2 + Ruta 3 + Ruta 4	90	No
Yavin IV	Bespin	Ruta 3 + Ruta 4 + Ruta 5	120	No
Tatooine	Dagobah	Ruta 1 + Ruta 2 + Ruta 3 + Ruta 4	100	No

Propósito del Dataset

El **Dataset** se utilizó para entrenar un modelo de aprendizaje supervisado basado en árboles de decisión. Los datos fueron transformados y divididos en conjuntos de entrenamiento y prueba para evaluar la precisión del modelo.

Capítulo 2

Rutas generadas para el sistema de transporte masivo no Supervisado

Introducción

En el presente documento se describe en detalle el **Dataset** generado como parte del proyecto de aprendizaje no supervisado. Este **Dataset** se utiliza para aplicar el algoritmo **K-Means**, con el objetivo de agrupar rutas similares en clústeres en función de sus características.

El propósito principal es analizar el sistema de transporte masivo y descubrir patrones ocultos en los datos, tales como agrupaciones de rutas con tiempos de viaje similares o conexiones recurrentes entre estaciones.

Estructura del Dataset

El **Dataset** fue generado a partir de las conexiones específicas definidas en el sistema de transporte masivo. Estas conexiones incluyen información sobre las estaciones conectadas, la línea utilizada y el tiempo de viaje entre ellas. Las características principales del dataset son:

Columna	Descripción
Estación Origen	Nombre de la estación de origen.
Estación Destino	Nombre de la estación de destino.
Línea	Nombre de la línea de transporte que es utilizada para conectar las estaciones.
Tiempo (min)	Tiempo de viaje entre las estaciones, expresado en minutos.
Ruta Optima	Indica si la ruta es óptima (Si) o no (No) en función de un tiempo límite.

Fuentes de Datos

Las rutas se generaron automáticamente a partir de un conjunto de estaciones y sus conexiones.

Estas conexiones fueron definidas manualmente y modelan un sistema de transporte masivo ficticio. Cada conexión incluye:

- **Estaciones:** Un conjunto de 11 estaciones ficticias que representan puntos geográficos importantes.
- **Líneas:** Cada conexión entre estaciones está asociada a una línea específica, lo que permite identificar las rutas.
- **Tiempo de viaje:** Cada conexión tiene asignado un tiempo de viaje en minutos, que varía entre 10 y 100 minutos.

Ejemplo del Dataset Generado

El **Dataset** incluye las siguientes rutas:

Estación Origen	Estación Destino	Línea	Tiempo (min)
Tatooine	Alderaan	Ruta 1	10
Alderaan	Yavin IV	Ruta 2	20
Yavin IV	Hoth	Ruta 3	30
Hoth	Dagobah	Ruta 4	40
Dagobah	Bespin	Ruta 5	50
Bespin	Endor	Ruta 6	60
Endor	Naboo	Ruta 7	70
Naboo	Coruscant	Ruta 8	80
Coruscant	Kamino	Ruta 9	90
Kamino	Mandalore	Ruta 10	100
Tatooine	Yavin IV	Ruta 1 + Ruta 2	30
Alderaan	Hoth	Ruta 2 + Ruta 3	50
Yavin IV	Dagobah	Ruta 3 + Ruta 4	70
Hoth	Bespin	Ruta 4 + Ruta 5	90
Dagobah	Endor	Ruta 5 + Ruta 6	110
Tatooine	Hoth	Ruta 1 + Ruta 2 + Ruta 3	60
Alderaan	Dagobah	Ruta 2 + Ruta 3 + Ruta 4	90
Yavin IV	Bespin	Ruta 3 + Ruta 4 + Ruta 5	120
Tatooine	Dagobah	Ruta 1 + Ruta 2 + Ruta 3 + Ruta 4	100

Preparación de los Datos

Antes de aplicar el modelo de agrupamiento (K-Means), los datos fueron preparados mediante los siguientes pasos:

Transformación de datos categóricos

Las columnas Estación Origen, Estación Destino y Línea fueron transformadas en valores numéricos mediante el uso de la biblioteca LabelEncoder.

Estandarización de los datos

Las características numéricas fueron escaladas utilizando StandardScaler para garantizar que todas las columnas tengan una media de 0 y una desviación estándar de 1. Esto mejora el rendimiento del algoritmo K-Means al evitar que una característica domine a las demás debido a su escala.

Datos Preparados

El dataset preparado para el modelo tiene las siguientes columnas y los valores escalados:

Estacion Origen	Estacion Destino	Linea	Tiempo (min)
-1.566.699	-1.566.699	-1.566.699	-1.566.699
-1.218.544	-1.218.544	-1.218.544	-1.218.544
-0.870388	-0.870388	-0.870388	-0.870388
-0.522233	-0.522233	-0.522233	-0.522233
-0.174078	-0.174078	-0.174078	-0.174078

Propósito del Dataset

El dataset fue diseñado para:

1. Aplicar el modelo K-Means:

- Agrupar las rutas en clústeres según sus características.
- Descubrir los patrones ocultos en las rutas de transporte masivo.

2. Análisis y optimización del sistema:

- Identificar los clústeres de las rutas similares para optimizar los recursos.
- Detectar las rutas con tiempos del viaje similares o las estaciones recurrentes.

3. Visualización de resultados:

- Generar los gráficos que representen los clústeres y permitan analizar cómo se agrupan las rutas.

Resultados del Agrupamiento

Después de aplicar el modelo K-Means, cada ruta fue asignada a un clúster. Un ejemplo de los resultados es el siguiente:

Estacion Origen	Estacion Destino	Linea	Tiempo (min)
Tatooine	Alderaan	Ruta 1	10
Alderaan	Yavin IV	Ruta 2	20
Yavin IV	Hoth	Ruta 3	30
Hoth	Dagobah	Ruta 4	40
Dagobah	Bespin	Ruta 5	50
Bespin	Endor	Ruta 6	60
Endor	Naboo	Ruta 7	70
Naboo	Coruscant	Ruta 8	80
Coruscant	Kamino	Ruta 9	90
Kamino	Mandalore	Ruta 10	100

Conclusión

El uso combinado de métodos supervisados y no supervisados demuestra cómo la inteligencia artificial puede ser aplicada de manera efectiva en sistemas complejos como el transporte masivo. Mientras que el aprendizaje supervisado se enfoca en la predicción de resultados específicos, el aprendizaje no supervisado explora patrones y relaciones ocultas en los datos.

Ambos enfoques generaron información valiosa que puede ser utilizada para:

1. **Optimizar recursos:** Identificar rutas prioritarias para mejorar su eficiencia.
2. **Diseñar estrategias:** Planificar nuevas conexiones basadas en agrupamientos detectados.
3. **Mejorar la experiencia del usuario:** Reducir tiempos de viaje y hacer el transporte más eficiente.

Por tanto, es evidente que el proyecto demuestra cómo los métodos de inteligencia artificial pueden contribuir significativamente a la gestión y optimización de sistemas de transporte masivo, proporcionando herramientas analíticas y predictivas para la toma de decisiones estratégicas.

