



Guía docente

230337 - AIPRAC - Resolución de Problemas con Inteligencia Artificial: un Enfoque Práctico

Última modificación: 17/06/2024

Unidad responsable: Escuela Técnica Superior de Ingeniería de Telecomunicación de Barcelona

Unidad que imparte: 739 - TSC - Departamento de Teoría de la Señal y Comunicaciones.

Titulación: GRADO EN INGENIERÍA DE TECNOLOGÍAS Y SERVICIOS DE TELECOMUNICACIÓN (Plan 2015). (Asignatura optativa).

GRADO EN INGENIERÍA ELECTRÓNICA DE TELECOMUNICACIÓN (Plan 2018). (Asignatura optativa).

Curso: 2024

Créditos ECTS: 2.0

Idiomas: Catalán

PROFESORADO

Profesorado responsable: Enric Monte Moreno

Otros:

CAPACIDADES PREVIAS

Programación Python y buenos conocimientos en las asignaturas de "Cálculo Vectorial", "Señales y Sistemas" y "Probabilidad y Estadística"

REQUISITOS

"Cálculo Vectorial" y "Probabilidad y Estadística"

METODOLOGÍAS DOCENTES

En cada sesión de clase se dedicará una hora a explicar la base de datos con la que se trabajará, y las propiedades estadísticas y geométricas de la misma. A continuación, considerando las propiedades de la base de datos, se justificará la elección de un método de inteligencia artificial adecuado para la misma. Las dos horas siguientes se dedicarán a trabajar con la base de datos, utilizando un 'notebook' de Jupyter en lenguaje Python.

OBJETIVOS DE APRENDIZAJE DE LA ASIGNATURA

El objetivo de la asignatura es dotar a los alumnos de herramientas para analizar las propiedades de una base de datos relacionada con una aplicación real, y decidir cuál es la técnica de inteligencia artificial más adecuada para resolver el problema.

HORAS TOTALES DE DEDICACIÓN DEL ESTUDIANTADO

Tipo	Horas	Porcentaje
Horas grupo grande	6,0	12.00
Horas grupo pequeño	14,0	28.00
Horas aprendizaje autónomo	30,0	60.00

Dedicación total: 50 h



CONTENIDOS

Análisis Exploratorio de Datos (EDA) y Regresión Logística

Descripción:

Base de Datos: Enfermedad Cardíaca en Sudáfrica
Técnica: EDA y Regresión Logística

Objetivos específicos:

Análisis del conjunto de datos de Enfermedad Cardíaca en Sudáfrica, que contiene información sobre pacientes masculinos, incluidas características como edad, presión arterial, niveles de colesterol, historial de tabaquismo y presencia de enfermedad cardíaca. Utilice técnicas de EDA (histogramas, diagramas de dispersión, diagramas de caja) para comprender la distribución de estas características y explorar posibles relaciones entre ellas y la enfermedad cardíaca. Aplique la regresión logística para predecir la probabilidad de que un paciente tenga una enfermedad cardíaca en función de estas características. Evalúe el rendimiento del modelo utilizando métricas como precisión, exactitud y recuperación. Analice los coeficientes del modelo para comprender qué factores tienen la influencia más significativa en el riesgo de enfermedad cardíaca.

Actividades vinculadas:

Jupyter Notebook en python

Dedicación: 3h

Grupo mediano/Prácticas: 3h

Clasificación con K-Nearest Neighbors (KNN) y Árboles de Decisión

Descripción:

Base de Datos: Predicción de abandono del cliente
Técnica: K-Nearest Neighbors y Árboles de Decisión

Objetivos específicos:

Análisis de una base de datos que contenga información del cliente. Use estos datos para predecir qué clientes son propensos a marcharse (dejar de usar el servicio). Aplique KNN y árboles de decisión para clasificar a los clientes como potenciales desertores o no. Compare el rendimiento de ambos modelos y analice los factores que influyen en la deserción.

Actividades vinculadas:

Jupyter Notebook en python

Dedicación: 3h

Grupo mediano/Prácticas: 3h

Reducción de Dimensionalidad con Análisis de Componentes Principales (PCA)

Descripción:

Base de Datos: Datos del mercado de valores
Técnica: Análisis de Componentes Principales (PCA)

Objetivos específicos:

Análisis de una base de datos de precios históricos de acciones de varias empresas. Use PCA para reducir la dimensionalidad de los datos mientras conserva la mayor parte de la información. Visualice los datos reducidos e interprete los componentes principales. Analice cómo se relacionan estos componentes con los movimientos de los precios de las acciones.

Actividades vinculadas:

Jupyter Notebook en python

Dedicación: 3h

Grupo mediano/Prácticas: 3h



Clústering con K-Means

Descripción:

Base de Datos: Segmentación de clientes
Técnica: Clustering K-Means

Objetivos específicos:

Análisis de una base de datos que contenga información del cliente e historial de compras. Use el clustering K-means para segmentar a los clientes en grupos diferenciados según sus comportamientos de compra. Analice las características de cada segmento de cliente e identifique posibles estrategias de marketing para cada grupo.

Actividades vinculadas:

Jupyter notebook en python.

Dedicación: 3h

Grupo mediano/Prácticas: 3h

Pronóstico de Series Temporales con ARIMA

Descripción:

Base de Datos: Datos de temperatura diaria
Técnica: Promedio Móvil Integrado Autorregresivo (ARIMA)

Objetivos específicos:

Análisis de una base de datos de temperaturas diarias. Use modelos ARIMA para pronosticar temperaturas futuras. Evalúe la precisión del pronóstico utilizando métricas como el error cuadrático medio. Analice las tendencias estacionales y los patrones de los datos e incorpórellos al modelo si es necesario.

Actividades vinculadas:

Jupyter notebook en python.

Dedicación: 3h

Grupo mediano/Prácticas: 3h

Clasificación de Texto con Máquinas de Vectores de Soporte (SVM)

Descripción:

Base de Datos: Reseñas de películas
Técnica: Máquinas de Vectores de Soporte (SVM)

Objetivos específicos:

Análisis de una base de datos de reseñas de películas. Preprocesa los datos de texto (por ejemplo, elimina palabras vacías, realiza stemming o lematización). Utiliza Máquinas de Vectores de Soporte (SVM) para clasificar las reseñas de películas como positivas, negativas o neutras. Explora diferentes técnicas de procesamiento de texto y evalúa su impacto en el rendimiento del modelo.

Actividades vinculadas:

Jupyter notebook en python

Dedicación: 3h

Grupo mediano/Prácticas: 3h



Última Sesión

Descripción:

Recapitulación de lo aprendido

Dedicación: 2h

Grupo mediano/Prácticas: 2h

SISTEMA DE CALIFICACIÓN

La calificación se basará en el 'notebook' entregado al final de cada sesión de trabajo.

BIBLIOGRAFÍA

Básica:

- James, G., Witten, D., Hastie, T., & Tibshirani, R.. An introduction to statistical learning [en línea]. New York: springer, 2013 [Consulta: 12/06/2024]. Disponible a: <https://www.statlearning.com/>.
- Richert, Willi.. Building Machine Learning Systems with Python [en línea]. 2013. New York: Packt Publishing, 2013 [Consulta: 17/06/2024]. Disponible a: <https://www.packtpub.com/product/building-machine-learning-systems-with-python/9781782161400>.