



La Rectora Magnífica de la Universidad San Jorge,

conforme a la legislación y normativa vigentes, certifica que

Don Miguel Ángel Croche Muñoz

ha cursado y superado satisfactoriamente, con una calificación final de Sobresaliente, los estudios correspondientes al

Curso Técnico en Machine Learning en Python

con una duración de 150 horas y una carga de 6 ECTS

Villanueva de Gállego (Zaragoza), 25 de Abril de 2023

M. Berta Sáez Gutiérrez
Rectora

Alberto Martín García
Responsable del área de Calidad, Validaciones y Medio Ambiente

- **Machine Learning en Python** - 150 horas - 6 ECTS

- 1. LOS DATOS, FUENTES DE DATOS Y EXTRACCIÓN**

Tipos básicos de datos. Tipos de datos compuestos. Fuentes de datos

- 2. TEORÍA DE PROBABILIDADES Y ESTADÍSTICA**

Bases de la teoría de la probabilidad. Parámetros centrales, de dispersión y de forma de las variables aleatorias discretas. Distribuciones. Variables aleatorias multidimensionales. Introducción a la estadística. Apéndice: Otras distribuciones

- 3. REPRESENTACIÓN DE DATOS CON SEABORN**

Tipos de datos en aprendizaje de máquinas. Repaso de conceptos de Matplotlib. Combinación de funciones de Matplotlib y de Seaborn. Funciones a nivel de ejes y a nivel de figura en Seaborn. Gráficas de dispersión. Gráficas de funciones del tiempo. Agregación y representación de la incertidumbre. Trazado de subconjuntos de datos con mapeos semánticos. Figuras con gráficas múltiples relacionadas con facetas. Datos categóricos. Diagramas de caja usando Seaborn. Representaciones en violín. Diagramas de barra. Representación de la distribución de una variable. Estimación de la densidad del núcleo. Representación de distribuciones de dos variables. Representación de relaciones entre pares de variables. Representación de relaciones entre variables

- 4. PROCESADO DE DATOS**

Librerías de interés en preprocesado de datos. La librería scikit-learn. Limpieza de datos. Normalización y escalado. Extracción y selección de características

- 5. BASES DEL MODELADO DE DATOS Y DEL APRENDIZAJE**

Aprendizaje de máquinas. El modelado de los datos. Clasificación de los algoritmos de aprendizaje. Conjuntos de entrenamiento, validación y prueba. Análisis de errores. Función de pérdida. Validación cruzada. Sobre-ajuste, infra-ajuste y regularización. Métricas en scikit-learn. Clases desequilibradas. Matriz de confusión. Tipos de modelos. La librería scikit-learn en aprendizaje

- 6. ALGORITMOS SUPERVISADOS (I). REGRESIÓN Y MÁQUINAS DE VECTORES SOPORTE**

Parte I. Regresión. Parte II. Máquinas de vectores soporte

- 7. APRENDIZAJE SUPERVISADO (II). REDES NEURONALES**

Las neuronas biológicas. Aplicaciones de las redes neuronales. Las neuronas artificiales. El perceptrón lineal. Transformación de un problema no lineal en uno lineal. Redes neuronales artificiales. El perceptrón multicapa. Teoría de aproximación universal. Implementación de las redes neuronales con scikit-learn. Clasificación multiclase (cuatro clases) con perceptrón multicapa

- 8. SISTEMAS DE APRENDIZAJE NO SUPERVISADOS. AGRUPAMIENTOS**

Tipos de clustering. El algoritmo k-means. El algoritmo kmeans de la librería sklearn. K vecinos más próximos. Variantes del algoritmo básico. DBSCAN. Uso del módulo scipy en clustering