

COMPETENCIA GLOBAL

Al finalizar el curso el estudiante estará habilitado para entender los principales conceptos relacionados con los modelos de aprendizaje de máquina y tendrá la capacidad de comprender el funcionamiento de las redes neuronales profundas y su aplicación a la solución de problemáticas reales.

COMPETENCIAS ESPECÍFICAS

- a) Reconocer la naturaleza de un problema de análisis predictivo y ser capaz de crear una solución a través de un modelo optimo usando el lenguaje de programación python.
- b) Entender los elementos básicos de un problema de grandes datos y la necesidad de usar apache spark o dask para encontrar el modelo optimo que lo soluciones.
- c) Comprender los conceptos básicos del aprendizaje profundo en problemas de visión por computador y procesamiento de lenguaje natural.

MODALIDAD PRESENCIAL

Bajo el modelo de Presencialidad Asistida por Tecnología PAT

1. Curso de 54h presenciales (método online), 90h de trabajo autónomo. Total 144h.
2. Herramientas: Laptop o PC (mínimo de 8gb ram y procesador intel i3 o similar, 128 gb de disco duro) y una conexión a Internet (>10Mb).
3. Requisitos: Bases de álgebra lineal, estadística, métodos numéricos y programación.

RESUMEN DE CONTENIDOS

1. Introducción al curso. Configuración del sistema y bases de python.
2. Aprendizaje supervisado y no supervisado.
3. Representación de datos e ingeniería de características.
4. Evaluación y Mejoramiento de Modelos. Algoritmos en cadena y pipelines.
5. Métodos para Análisis de Grandes Datos.
6. Introducción al aprendizaje profundo.
7. Presentación de proyectos finales.

Sesiones

Actividades de aprendizaje

1 Bloque 1 (6h): Introducción al curso. Configuración del sistema y bases de python.

1. Programación en Python. Anaconda. Notebooks: Jupyter, Collabs y otros.
2. Librerías básicas: Numpy, Pandas y Scikit-Learn.

2 Bloque 2 (14h): Aprendizaje supervisado y no

1. Algoritmos para regresión y clasificación.
2. Preprocesamiento y escalamiento de datos.
3. Algoritmos para clustering.

3 Bloque 3 (8h): Representación de datos e ingeniería de características.

1. Variables categóricas.
2. Selección automática de características.

4 Bloque 4 (8h): Evaluación y Mejoramiento de Modelos. Algoritmos en cadena y pipelines.

1. Validación cruzada.
2. Búsqueda de parámetros con grillas.
3. Evaluación de métricas y scoring.
4. Construcción de pipelines para optimización de parámetros.

5 Bloque 5 (6h): Métodos para Análisis de Grandes Datos.

1. Elementos básicos de Apache Spark.
2. Elementos básicos de Dask.

6 Bloque 6 (8h). Introducción al aprendizaje profundo.

1. Aprendizaje profundo para visión en computador.
2. Aprendizaje profundo para procesamiento de lenguaje natural.

7 Bloque 7 (4h). Presentación de proyectos finales.

El curso se evaluara con la presentación de un proyecto final (15mins + 5 preguntas) y un reporte escrito usando un notebook.

El proyecto se hará en grupos de máximo 3 estudiantes. Habrá un día único para asesoría de los proyectos en la semana 29 (17 de Julio) las presentaciones en la semana 30 (24 de Julio).

RECURSOS TECNOLÓGICOS

Laptop o PC (mínimo de 8gb ram y procesador intel i3 o similar, 128 gb de disco duro) y una conexión a Internet (>10Mb).
Uso de la plataforma Webex. Todos los docentes y estudiantes, deberán tener un buen acceso a internet, sonido y cámara para poder facilitar las sesiones.

EVALUACIÓN

El curso se evaluara con la **presentación de un proyecto final** (15mins + 5 preguntas) y un reporte escrito usando un notebook. El proyecto se hará en grupos de máximo **3 estudiantes**. Habrá un día único para asesoría de los proyectos en la semana 29 (17 de Julio) las presentaciones en la semana 30 (24 de Julio).