



Diplomado en
**MACHINE
LEARNING**

CURSO 1
**INTRODUCCIÓN AL
MACHINE LEARNING**

CURSO 2
**MACHINE
LEARNING**

CURSO 3
**DEEP
LEARNING**



Curso

INTRODUCCIÓN A MACHINE LEARNING

El machine learning o aprendizaje automático es un conjunto de algoritmos que se entrena en un conjunto de datos para hacer predicciones o tomar acciones para optimizar algunos sistemas. Por ejemplo, los algoritmos de clasificación supervisados se utilizan para clasificar a los posibles clientes en prospectos buenos o malos, para fines de préstamo, con base en datos históricos.

La transformación digital está generando una ingente cantidad de datos que pueden ser utilizados para mejorar los procesos productivos, el desarrollo y venta de productos y servicios, entre otros. Las empresas y organizaciones tienen en los datos a uno de sus principales activos y requieren de un buen análisis para una óptima toma de decisiones para así mantener u obtener ventajas competitivas. Pero los datos tienen sólo valor potencial, es su análisis y sistematización el que permite incrementar la capacidad de innovar y obtener ventajas en las organizaciones. Lamentablemente los datos nunca están como uno quisiera. Los datos reales pueden ser impuros, pueden conducir a la extracción de patrones/reglas poco útiles. Datos malos nos llevarán a la mala toma de decisiones.

OBJETIVO

Este curso presentará las etapas del preprocesamiento de datos, exploración de datos y algoritmos clásicos de aprendizaje supervisado y no supervisado. Para cada uno de los temas revisados se muestran ejemplos de cómo usar los algoritmos ya sea con un código o haciendo uso de la biblioteca de paquetes de R/Python.

METODOLOGÍA

La metodología de UAI Online contempla clases en línea con diferentes elementos de aprendizaje individual que incluye controles por clase, lecturas, tareas y videos multimedia con el contenido del programa de estudio, por su parte, el aprendizaje colaborativo se fomenta a través de Foro de la Clases de discusión entre los participantes del curso.

El curso tiene una duración de 5 semanas para las clases online, que incluyen material de lectura obligatoria y lecturas complementarias, la participación en los Foro de la Clases y las actividades evaluativas a fin de desarrollar los conocimientos contemplados por el mismo. Trabajo semanal de 12 horas.

PROFESOR



Rolando De la Cruz

Magíster y Doctorado en Estadística, Pontificia Universidad Católica de Chile. Profesor regular de la Facultad de Ingeniería y Ciencias e investigador del Smart Center de la UAI. Director Académico del Magíster en Data Science. Su investigación se enfoca en el desarrollo de metodologías para el análisis de datos. Cuenta con más de 30 artículos científicos publicados en revistas ISI. Actualmente es editor asociado del ISI journal Statistical Modelling y evaluador de los programas FONIS y FONDECYT, para la National Security Agency Mathematical Sciences Program, EE.UU, y es par evaluador de la CNA para programas de postgrado. Además brinda asesoría en temas de Business Analytics, Big Data, Data Science y Machine Learning.

CONTENIDOS

Introducción al Machine Learning

- Introducción
- Descripción general del aprendizaje automático
- Tabla de datos y sus características
- Proceso de Machine Learning
- Instalar R/Python

Preprocesamiento de Datos

- Etapas del preprocesamiento de datos
- Limpieza de datos
- Tratamiento de missing data. Tratamiento de outliers
- Ingeniería de Atributos
- Aplicaciones con R/Python

Análisis Exploratorio de Datos

- Caracterizando variables de acuerdo a su tipo
- Medidas de tendencia central, dispersión, posición y forma
- Gráficos variables de acuerdo a su tipo
- Correlaciones entre atributos y gráficos asociados. Significancia
- Aplicaciones con R/Python

Aprendizaje Supervisado: Modelo de regresión lineal - Parte I

- Introducción. Formulación del modelo
- Métodos de Estimación: MCO, Gradiente Descendente, Máxima verosimilitud
- Uso de atributos categóricos. Interpretaciones
- Aplicaciones con R/Python

Aprendizaje Supervisado: Modelo de regresión lineal – Parte II

- Medidas de bondad de ajuste
- Identificando atributos estadísticamente significativos
- Multicolinealidad. Técnicas de selección de atributos
- Métricas de calidad predictiva
- Aplicaciones con R/Python

Aprendizaje Supervisado: Regresión logística

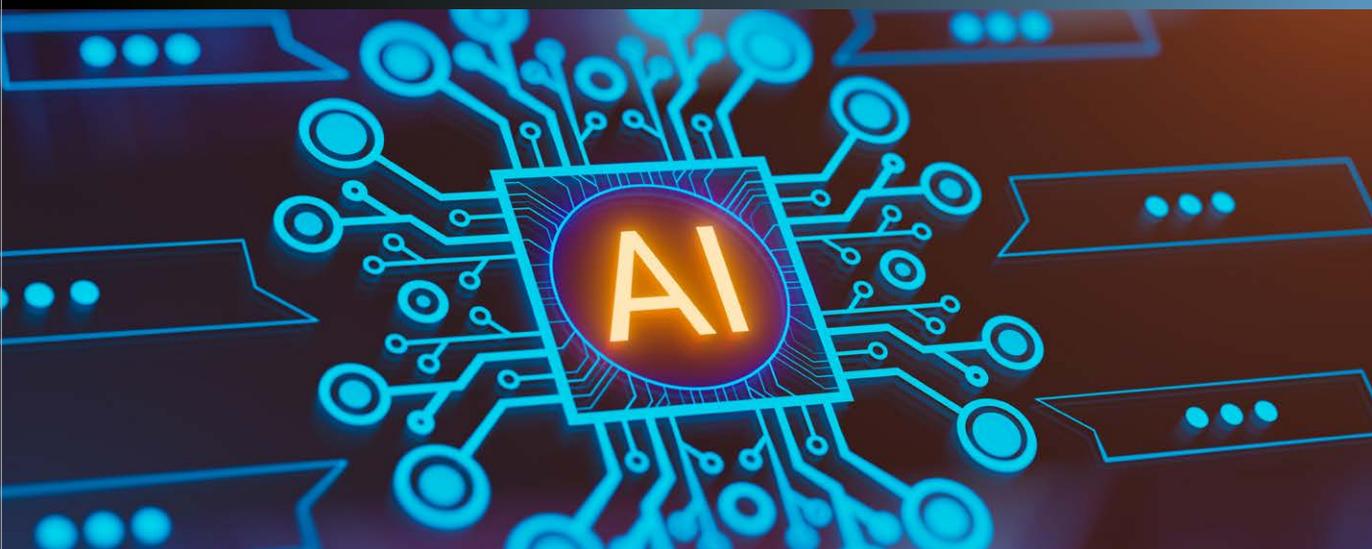
- Formulación de un modelo de regresión logística. Medidas de dependencia. Estimación. Interpretaciones
- WOE y Valor de la Información
- Técnicas de selección de variables
- Aplicaciones con R/Python

Aprendizaje Supervisado: Validación de la Regresión logística

- Matriz de confusión y Métricas
- Curva ROC, AUC, K-S, Gini, Briers
- Gain y Lift chart
- Caso desequilibrio de clases. Soluciones
- Aplicaciones con R/Python

Algoritmos de Aprendizaje No Supervisados

- Algoritmos jerárquicos
- Dendograma. Interpretación
- Algoritmo K-medias
- Elección del número de grupos
- Aplicaciones con R/Python



Curso

MACHINE LEARNING

La transformación digital está generando una ingente cantidad de datos que pueden ser utilizados para mejorar los procesos productivos, el desarrollo y venta de productos y servicios, entre otros. Las empresas y organizaciones tienen en los datos a uno de sus principales activos y requieren de un buen análisis para una óptima toma de decisiones para así mantener u obtener ventajas competitivas.

El machine learning o aprendizaje automático es un conjunto de algoritmos que se entrena en un conjunto de datos para hacer predicciones o tomar acciones para optimizar algunos sistemas. Por ejemplo, los algoritmos de clasificación supervisados se utilizan para clasificar a los posibles clientes en prospectos buenos o malos, para fines de préstamo, con base en datos históricos. Las técnicas del machine learning, junto a reglas de negocio se pueden utilizar para reducir el riesgo y calificar objetivamente la calidad de un cliente.

OBJETIVO

Este curso presentará algoritmos de aprendizaje supervisados y no supervisados tales como: métodos de regularización, naive Bayes, k-vecinos más cercanos, árboles de clasificación/regresión, algoritmos ensamblados, Support Vector Machine, Algoritmos jerárquicos, K-medias, DBScan y reglas de asociación. Se explica cómo validar modelos supervisados de machine learning para evitar sobreajustes, así como evaluar su performance mediante métricas de calidad predictiva. Para cada uno de los temas revisados se muestran ejemplos de cómo usar los algoritmos en el lenguaje de programación R/Python.

METODOLOGÍA

La metodología de UAI Online contempla clases en línea con diferentes elementos de aprendizaje individual que incluye controles por clase, lecturas, tareas y videos multimedia con el contenido del programa de estudio, por su parte, el aprendizaje colaborativo se fomenta a través de Foro de la Clases de discusión entre los participantes del curso.

El curso tiene una duración de 5 semanas para las clases online, que incluyen material de lectura obligatoria y lecturas complementarias, la participación en los Foro de la Clases y las actividades evaluativas a fin de desarrollar los conocimientos contemplados por el mismo. Trabajo semanal de 12 horas.

CONTENIDOS

- Métodos de Regularización
- Modelos supervisados: Naive Bayes y K-vecinos más cercanos
- Modelos supervisados: Árboles de clasificación/regresión
- Modelos supervisados: Algoritmos ensamblados – bagging/boosting
- Support Vector Machine
- Algoritmos de Aprendizaje No Supervisados: Algoritmos jerárquicos
- Modelos no supervisados: Algoritmos no jerárquicos
- Modelos no supervisados: Reglas de Asociación

PROFESOR



Rolando De la Cruz

Magíster y Doctorado en Estadística, Pontificia Universidad Católica de Chile. Profesor regular de la Facultad de Ingeniería y Ciencias e investigador del Smart Center de la UAI. Director Académico del Magíster en Data Science. Su investigación se enfoca en el desarrollo de metodologías para el análisis de datos. Cuenta con más de 30 artículos científicos publicados en revistas ISI. Actualmente es editor asociado del ISI journal Statistical Modelling y evaluador de los programas FONIS y FONDECYT, para la National Security Agency Mathematical Sciences Program, EE.UU, y es par evaluador de la CNA para programas de postgrado. Además brinda asesoría en temas de Business Analytics, Big Data, Data Science y Machine Learning.



Curso
**DEEP
 LEARNING**

Deep learning o aprendizaje profundo es un tipo de machine learning destinado principalmente para resolver problemas de percepción, tales como reconocimiento de imágenes, video, audio, etc. Típicamente se ha focalizado en aprendizaje supervisado, pero también existen modelos de deep learning para los otros tipos de aprendizaje. Esencialmente, también se busca una función de mapeo que relacione datos de entrada con una salida deseada. Las principales diferencias con machine learning tradicional, es que la entrada es la fuente directamente, una imagen, por ejemplo, sin la necesidad de extraer atributos que caractericen la imagen, y que luego sean la entrada al algoritmo de machine learning. Lo otro es que el aprendizaje se realiza mediante muchas capas, típicamente, mediante una estructura de redes neuronales.

OBJETIVOS

- Entender los fundamentos y teoría de los algoritmos existentes en Deep learning.
- La capacidad de utilizar redes neuronales para deep learning vía el framework Keras usando Python.
- Ver diferentes aplicaciones que utilizan un enfoque de deep learning.

METODOLOGÍA

La metodología de UAI Online contempla clases en línea con diferentes elementos de aprendizaje individual que incluye controles por clase, lecturas, tareas y videos multimedia con el contenido del programa de estudio, por su parte, el aprendizaje colaborativo se fomenta a través de Foro de la Clases de discusión entre los participantes del curso.

El curso tiene una duración de 5 semanas para las clases online, que incluyen material de lectura obligatoria y lecturas complementarias, la participación en los Foro de la Clases y las actividades evaluativas a fin de desarrollar los conocimientos contemplados por el mismo. Trabajo semanal de 12 horas.

CONTENIDOS

- Introducción
- Redes Neuronales
- Redes neuronales en la práctica
- Redes neuronales convolucionales
- Redes neuronales convolucionales en la práctica
- Redes neuronales recurrentes (RNN)
- Redes neuronales recurrentes en la práctica
- Aplicaciones

PROFESOR



Gonzalo Ruz

Doctorado en Machine Learning, Universidad Cardiff, Reino Unido. Ingeniero Civil Electricista y Magíster en Ciencias de la Ingeniería con mención Ingeniería Eléctrica, Universidad de Chile. Profesor titular de la Facultad de Ingeniería y Ciencias de la Universidad Adolfo Ibáñez. Trabajó en el departamento de Inteligencia de Negocios de Sonda S.A., dedicándose a diseñar y ejecutar proyectos de Data mining, en particular, diseño de sistemas de detección de fraude y modelos predictivos.