

<b>PROGRAMA</b>		
<b>MATERIA</b>	<b><i>Aprendizaje Automático</i></b>	
<b>CARRERAS</b>	<b><i>Todas</i></b>	
<b>SEDE</b>	<b><i>Miguelete</i></b>	
<b>CARÁCTER DE LA MATERIA</b>	<b><i>Electiva de grado y de posgrado</i></b>	
<b>PERÍODO DE VIGENCIA</b>	<b><i>1<sup>er</sup> y 2do cuatrimestre 2020</i></b>	<b><i>DURACION: cuatrimestral</i></b>
<b>DOCENTES</b>		
<b>MATERIAS CORRELATIVAS</b>	<b><i>Conocimientos elementales de matemática y programación</i></b>	
<b>CARGA HORARIA</b>	<b><i>Clases Teórico-prácticas y de Laboratorio: 96</i></b>	<b><i>Total de horas semanales: 6 Total de horas de la materia: 96</i></b>

### **OBJETIVOS:**

El objetivo general del curso es introducir al alumno en los principios y conceptos fundamentales del aprendizaje automático, la ciencia de datos y la inteligencia artificial.

### **CONTENIDOS MÍNIMOS**

Conceptos elementales de programación, probabilidad y estadística. Algoritmos de aprendizaje supervisado y no supervisado. Aplicaciones generales de la Inteligencia Artificial.

### **FUNDAMENTACIÓN**

Vivimos en un mundo donde cada día se generan miles de millones de gigabytes de datos sobre diferentes aspectos de nuestras vidas y que resultan en un gran activo en nuestra búsqueda de soluciones para los grandes desafíos de la ciencia, la ingeniería, la medicina y los negocios. En los últimos años han surgido nuevas tecnologías que, a partir del análisis de grandes cantidades de datos, permiten la generación de conocimiento con un impacto notable en la ciencia, pero también economía y en el mercado como lo muestra la revolución producida en las ventas on-line y en los negocios digitales. Este curso procura preparar a los estudiantes para poder realizar análisis de datos y aplicar algoritmos de aprendizaje automático, una tarea cada vez más requerida en nuestra sociedad. Los conocimientos adquiridos serán fundamentales para poder profundizar en la inteligencia artificial y seleccionar las herramientas adecuadas en diferentes contextos.

### **METODOLOGÍA DE EVALUACIÓN**

Dos exámenes parciales. Los alumnos de posgrado deberán rendir un examen final obligatorio. La materia es promocionable para alumnos de grado.

### **METODOLOGÍA DIDÁCTICA**

Los contenidos se organizan en clases teórico-prácticas y trabajos prácticos de laboratorio. En cada tema se incluyen actividades prácticas que promueven la participación de los estudiantes. Las diferentes técnicas estudiadas serán introducidas con una breve descripción teórica y con énfasis en su aplicación a diversos conjuntos de datos reales. Las aplicaciones generales serán presentadas a partir de seminarios de expertos en el tema. El curso se brindará con un formato completamente interdisciplinario para todas las carreras de la UNSAM.

## PROGRAMA ANALÍTICO.

1. Programación y análisis exploratorio de datos:
  - 1.1 Python: estructuras de datos y control.
  - 1.2 Librerías para la manipulación de datos: Pandas, Seaborn y Scikit-learn.
  - 1.3 Visualización de datos: Histogramas, box-plots, qq-plots, scatter-plots.
  - 1.4 Limpieza e imputación de datos faltantes.
2. Probabilidad y estadística:
  - 2.1 Distribuciones: Binomial, Poisson, Gaussiana.
  - 2.2 Estimadores, correlaciones, máxima verosimilitud.
  - 2.3 Conceptos elementales de la teoría de la información.
  - 2.4 Enfoques frecuentista y Bayesiano.
3. Aprendizaje supervisado:
  - 3.1 Regresión lineal. Cuadrados mínimos. Interpretación estadística.
  - 3.2 Algoritmos de Clasificación: Naive Bayes, Árboles, R. Forests, SVM.
  - 3.3 Sobreajuste y validación cruzada. K-folding.
  - 3.4 Redes neuronales.
4. Aprendizaje no supervisado:
  - 4.1 Algoritmos de clustering: K-means, mezcla de Gaussianas.
  - 4.2 Máxima verosimilitud y algoritmo EM.
  - 4.3 Criterios de selección de modelos.
  - 4.4 El problema de la dimensión y la selección de atributos.
5. Aplicaciones generales de Inteligencia Artificial:
  - 5.1 Procesamiento de texto y lenguaje natural.
  - 5.2 Aplicaciones científicas.
  - 5.3 Aplicaciones de negocios.

## BIBLIOGRAFÍA.

- Ch. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- G. Boente y V. Yohai, *Notas de Estadística*, Notas online FCEN-UBA, 2010.
- A. Gelman et al., *Bayesian Data Analysis*, CRC press, 2013.
- J. Guttag, *Introduction to Computation and Programming Using Python*, MIT, 2013.
- K. Murphy, *Machine Learning, a Probabilistic Perspective*, MIT press, 2013.
- A. Ng, *Machine Learning, CS229 Lecture Notes*, Online, Stanford, 2017.
- J. VanderPlas, *Python Data Science Handbook: Essential Tools for Working with Data*, O'Reilly, 2017.
- Géron, A., *Hands-On Machine Learning with Scikit-Learn and TensorFlow*, O'Reilly, 2017.