



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
**PROGRAMA DE POSGRADO EN CIENCIAS DE LA TIERRA**  
**PROGRAMA DE ACTIVIDAD ACADÉMICA**



Nombre de la asignatura: **MACHINE LEARNING**

Clave:	Semestre (s): 1 ( ) 2 ( ) (Semestre en el que se imparte)	Campo de Conocimiento: I Geofísica de la Tierra Sólida ( ) II Exploración, Aguas Subterráneas, Modelación y Percepción Remota (X) III Geología ( ) IV Ciencias Ambientales y Riesgos ( ) V Ciencias Atmosféricas, Espaciales y Planetarias. ( )	No. Créditos: <b>8</b>
Carácter: <b>Obligatoria de Elección ( )</b> <b>Optativa de Elección ( )</b>	<b>Horas</b>		<b>Horas por semana</b>
Tipo: <b>Teórico-Práctica ( X )</b> <b>Teórico ( )</b>	<b>Teoría: 2</b> (Número de horas)	<b>Práctica: 2</b> (Número de horas)	4
Modalidad: <b>Curso</b>	Duración del programa: <b>Semestral</b>		
<b>Horas al semestre</b>			
64			

**Seriación:** Sin seriación ( X )                      Obligatoria ( )                      Indicativa ( )

Actividad académica antecedente:

Actividad académica subsecuente:

**OBJETIVO GENERAL:**

El alumno aprenderá a entrenar y a usar modelos de machine learning con software relevante en la industria. Para un correcto uso de estos, el alumno adquirirá una base matemática y algorítmica con implementación en software de los métodos más populares de machine learning. El alumno podrá identificar problemas en los que los métodos aprendidos podrán ser aplicados. Además, se comparará el desempeño de cada algoritmo de manera individual y ensambles en diferentes aplicaciones.

**Índice Temático**

Unidad	Tema	Horas	
		Teóricas	Prácticas
1	Panorama general sobre machine learning	2	0
2	Proyecto aplicado de machine learning	4	4
3	Modelos lineales y regresión logística	3	3
4	Máquinas de soporte vectorial (Support Vector Machines)	3	3
5	Métodos basados en árboles de decisión	6	6
6	Modelos basados en teoría de gráficas	3	3
7	Aprendizaje no supervisado	3	3
8	Reducción de la dimensionalidad	3	3
9	Redes neuronales y aprendizaje profundo	6	6
Total de horas:			
Suma total de horas:		64	

## Contenido Temático

Unidad	Tema y Subtemas
1	1. Panorama general sobre machine learning 1.1. ¿Qué es el machine learning? 1.2. ¿Por qué y cuándo usar machine learning? 1.3. Clasificación de métodos de machine learning 1.4. Principales desafíos del machine learning 1.5. Prueba y validación de modelos de machine learning
2	2. Proyecto aplicado de machine learning (ciclo de vida) 2.1. Trabajar con datos reales 2.2. Panorama general de un proyecto de machine learning 2.3. Obtener lo datos 2.4. Análisis Exploratorio de Datos 2.5. Preprocesamiento de los datos 2.6. Seleccionar y entrenar un modelo 2.7. Mejora del modelo 2.8. Despliegue, monitoreo y mantenimiento del modelo
3	3. Modelos lineales y regresión logística 3.1. Regresión lineal 3.2. Optimización numérica 3.3. Modelos lineales 3.4. Regresión logística 3.5. Aplicación en geociencias
4	4. Máquinas de soporte vectorial (Support Vector Machines) 4.1. Clasificación lineal 4.2. Clasificación no lineal 4.3. Regresión 4.4. Aplicación en geociencias
5	5. Métodos basados en árboles de decisión 5.1. Árboles de decisión (Regresión y clasificación) 5.2. Bosques aleatorios, apilamiento y AdaBoost 5.3. Extreme Gradient Boosting (XGBoost) 5.4. Aplicación en geociencias
6	6. Modelos basados en teoría de gráficas 6.1. Redes bayesianas 6.2. Campos aleatorios markovianos 6.3. Inferencia
7	7. Aprendizaje no supervisado 7.1. Reglas de asociación 7.2. Análisis de conglomerados 7.3. mapas auto organizados 7.4. Curvas, superficies y componentes principales 7.5. Análisis de componentes independientes (ICA) 7.6. Clasificador Naive Bayes
8	8. Reducción de la dimensionalidad 8.1. Sobre dimensionalidad 8.2. Análisis de Componentes Principales 8.3. Otros métodos para reducir la dimensionalidad 8.4. Aplicación en geociencias
9	9. Redes neuronales y aprendizaje profundo 9.1. Redes neuronales biológicas y artificiales

	<p>9.2. Perceptrón multicapa</p> <p>9.3. Imágenes y visión computarizada (arquitecturas CNN)</p> <p>9.4. Aprendizaje profundo para textos, secuencias y series de tiempo</p> <p>9.5. Aplicación en geociencias</p>
--	--

**Bibliografía Básica:**

- Géron, Aurélien. 2019. **Hands-On Machine Learning with Scikit-Learn, Keras and TensorFlow**. <https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/>.
- Shi, Guangren. 2014. **Data Mining and Knowledge Discovery for Geoscientists**. ScienceDirect.
- Bishop, Christopher. 2006. **Pattern Recognition and Machine Learning**. Information Science and Statistics. New York: Springer-Verlag. <https://www.springer.com/us/book/9780387310732>.
- Hastie, Trevor, Robert Tibshirani, and J. H. Friedman. 2009. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. Springer Series in Statistics. New York, NY: Springer. <https://www.springer.com/gp/book/9780387848570>.

**Bibliografía Complementaria:**

- Chollet, François. 2018. **Deep Learning with Python**. Shelter Island, New York: Manning Publications Co.
- Mohaghegh, Shahab D. 2017. **Shale Analytics: Data-Driven Analytics in Unconventional Resources**. <http://site.ebrary.com/id/11343490>.
- Poulton, Mary M. 2001. **Computational Neural Networks for Geophysical Data Processing**. Amsterdam; New York: Pergamon. <http://www.knovel.com/knovel2/Toc.jsp?BookID=1902>.
- Cranganu, Constantin, Henri Luchian, and Mihaela Elena Breaban, eds. 2015. **Artificial Intelligent Approaches in Petroleum Geosciences**. Springer International Publishing. <https://www.springer.com/us/book/9783319165301>.
- Holdaway, Keith R. 2014. **Harness Oil and Gas Big Data with Analytics: Optimize Exploration and Production with Data Driven Models**. Wiley & SAS Business Series. Hoboken, New Jersey: Wiley.
- Holdaway, Keith R., and Duncan H. B. Irving. 2018. **Enhance Oil & Gas Exploration with Data-Driven Geophysical and Petrophysical Models**. Hoboken, New Jersey: Wiley.
- Perrin, Michel. 2013. **Shared Earth Modeling: Knowledge Driven Solutions for Building and Managing Subsurface 3D Geological Models**. Paris: Editions Technips.
- Ma, Y. Z. 2019. **Quantitative Geosciences: Data Analytics, Geostatistics, Reservoir Characterization and Modeling**. <https://doi.org/10.1007/978-3-030-17860-4>.
- Kanevski, Mikhail et al 2009. **Machine Learning for Spatial Environmental Data: Theory, Applications, and Software**. CRC Press.
- Kanevski, Mikhail, ed. 2008. **Advanced Mapping of Environmental Data: Geostatistics, Machine Learning and Bayesian Maximum Entropy**. Geographical Information Systems Series. London : Hoboken, NJ: ISTE ; Wiley.

<b>Sugerencias didácticas:</b>		<b>Mecanismos de evaluación de aprendizaje de los alumnos:</b>	
Exposición oral	( X )	Exámenes Parciales	( X )
Exposición audiovisual	( )	Examen final escrito	( X )
Ejercicios dentro de clase	( X )	Trabajos y tareas fuera del aula	( )
Ejercicios fuera del aula	( X )	Exposición de seminarios por los alumnos	( X )
Seminarios	( )	Participación en clase	( X )
Lecturas obligatorias	( X )	Asistencia	( X )
Trabajo de Investigación	( X )	Seminario	( )
Prácticas de taller o laboratorio	( X )	Otras: Tareas, seminarios, reportes de prácticas, exámenes parciales, examen final.	( )
Prácticas de campo	( )		
<b>Línea de investigación:</b>			
Cualquiera de los cinco campos del conocimiento.			