



Universidad de Puerto Rico
Recinto Universitario de Mayagüez
Colegio de Artes y Ciencias
Departamento de Ciencias Matematicas



ESMA 4016 MINERIA DE DATOS Y APENDIZAJE AUTOMATICO

Horas crédito: 3	Horas contacto: 3 horas de conferencia semanales
Requisitos previos: (ESMA 3016 y MATE 4031) o Permiso Director	Requisitos concurrentes:
Descripción del curso (español): Introducción a las principales técnicas de minería de datos y aprendizaje automático aplicadas al aprendizaje supervisado y no supervisado. Entre los tópicos que se incluyen están regresión y clasificación, métodos de vecinos más cercanos, arboles de decisión, repotenciar, redes neurales y maquinas de vectores de soporte.	
Descripción del curso (inglés): Introduction to the major techniques for data mining and machine learning applied to both supervised and unsupervised learning. Topics include regression and classification, nearest neighbors methods, decision trees, boosting, neural networks and support vector machines	
Objetivos del curso: Entender los conceptos fundamentales para llevar a cabo minería de datos y descubrimiento de conocimiento en base de datos usando métodos de Machine Learning. Experimentar algunos algoritmos más usados en minería de datos y Machine Learning en conjuntos de datos reales.	

Profesor: Edgar Acuña Oficina: OP307 Extension: Horas de Oficina: Martes y Jueves de 9.00 a 10.30pm y M de 12.30 a 3.30pm.

Contenido del curso:

<i>Temas a cubrir</i>	<i>Horas contacto</i>
Introduction to data mining and Machine Learning	2
Data Cleaning	3
Data Visualizacion	3
Dimension Reduction: Feature Selection and Principal Component Analysis	4
Supervised Learning: Regression	4
Supervised Learning: Logistic Regression	3

Review of Probability	1										
Supervised learning: Linear Discriminant Analysis	3										
Estimaton of Prediction Error	2										
Supervised Learning: Naïve Bayes	2										
Exam	2										
K Nearest Neighbors	2										
Decision Trees	3										
Random Forest	2										
Neural Networks	3										
Unsupervised Learning: Clustering	6										
Total de horas: (deben ser equivalentes a las horas crédito del curso)	45										
Estrategias instruccionales: <input checked="" type="checkbox"/> conferencia <input type="checkbox"/> discusión <input type="checkbox"/> cómputos <input type="checkbox"/> laboratorio <input type="checkbox"/> investigación <input type="checkbox"/> otros, especifique:											
Recursos mínimos disponibles: materiales y equipo necesarios para cumplir los objetivos del curso											
Estrategias de evaluación y su peso relativo: <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>Por ciento</th> </tr> </thead> <tbody> <tr> <td><input type="checkbox"/> examen</td> <td>30%</td> </tr> <tr> <td><input type="checkbox"/> tareas(3)</td> <td>40%</td> </tr> <tr> <td><input type="checkbox"/> proyecto</td> <td>30%</td> </tr> <tr> <td>TOTAL: 100%</td> <td></td> </tr> </tbody> </table>			Por ciento	<input type="checkbox"/> examen	30%	<input type="checkbox"/> tareas(3)	40%	<input type="checkbox"/> proyecto	30%	TOTAL: 100%	
	Por ciento										
<input type="checkbox"/> examen	30%										
<input type="checkbox"/> tareas(3)	40%										
<input type="checkbox"/> proyecto	30%										
TOTAL: 100%											
Texto: James, Witten, Tibshirani & Hastie. Introduction to Statistical Learning Springer, 2014 http://www-bcf.usc.edu/~gareth/ISL/ISLR%20Seventh%20Printing.pdf Laython, R. Learning Data Mining with Python (2015). Disponible el el github del curso https://github.com/eacunafer/Data-Mining-Machine-Learning-subgraduado-/blob/master/RLDMPy.pdf											
Referencias: 1- Torgo, L. Data Mining with R. CRC Press, 2011. http://www.bagualu.net/wordpress/wp-content/uploads/2015/10/Data_Mining_with_R_Learning_with_Case_Studies.pdf 2- Hastie, T., Tibshirani, R. and Friedman, J. The Elements of Statistical Learning: Data Mining, Inference and Prediction, 2nd Edition. Springer, 2009. https://web.stanford.edu/~hastie/ElemStatLearn/printings/ESLII_print12.pdf 3- Alpaydin, E. Introduction to Machine Learning. Third Edition. MiT Press, 2014											

[http://cs.du.edu/~mitchell/mario_books/Introduction to Machine Learning - 2e -
Ethem Alpaydin.pdf](http://cs.du.edu/~mitchell/mario_books/Introduction_to_Machine_Learning_-_2e_-_Ethem_Alpaydin.pdf)

4-A Complete Tutorial to Learn Data Science with Python from Scratch

<https://www.analyticsvidhya.com/blog/2016/01/complete-tutorial-learn-data-science-python-scratch-2/>

Videos:

Clase del Profesor Nando Freitas (British Columbia University)

<http://www.cs.ubc.ca/~nando/340-2012/lectures.php>

Clase del Profesor Andrew Ng (Stanford-Coursera)

<http://robotics.stanford.edu/~ang/courses.html>

Sistema de calificación:

☒cuantificable (de letra) ☐no cuantificable

Curva estándar

100-90 A; 89-80 B; 79-70 C; 69-60 D; 59-0 F

Después de identificarse con el profesor y la institución, los estudiantes con impedimento recibirán acomodo razonable en sus cursos y evaluaciones. Para más información comuníquese con Servicios a Estudiantes con Impedimentos en la Oficina del Decano de Estudiantes (Oficina #1) o a los teléfonos 787-265-3864 ó 787-832-4040 x 3772, 2040 y 3864.

EAF/Febrero 2018