



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

Programa de

FUNDAMENTOS DE APRENDIZAJE AUTOMÁTICO Y RECONOCIMIENTO DE PATRONES

1. NOMBRE DE LA UNIDAD CURRICULAR

Fundamentos de aprendizaje automático y reconocimiento de patrones

2. CRÉDITOS

8 créditos

3. OBJETIVOS DE LA UNIDAD CURRICULAR

Este es un curso introductorio al aprendizaje automático, que busca un equilibrio entre la aplicación práctica y sus fundamentos teóricos. Se presenta la teoría que establece un marco conceptual para el aprendizaje automático, y que permite formalizar ciertos aspectos, tales como: en qué consiste aprender, cotas del aprendizaje, relaciones de compromiso, métodos de evaluación. Además, el curso intenta proporcionar los conocimientos necesarios para comprender e implementar algoritmos de aprendizaje automático.

La unidad curricular propone como objetivos que el estudiante sea capaz de desarrollar una sólida comprensión de los fundamentos del aprendizaje automático y manejar varias técnicas clásicas de aprendizaje automático y reconocimiento de patrones. Al finalizar la unidad curricular el estudiante será capaz de comprender e implementar algoritmos de aprendizaje automáticos con los cuales podrá resolver problemas prácticos de análisis de datos, clasificación y regresión. Se espera que esta unidad curricular brinde una base sólida sobre la cual continuar la formación en el área.

4. METODOLOGÍA DE ENSEÑANZA

El curso está organizado en dos clases semanales de dos horas, alternando clases teóricas y clases prácticas. En las clases teóricas se presentarán los temas de forma expositiva, conectando los contenidos con ejemplos prácticos y problemas reales, así como habilitando espacios activos para la reflexión e intercambio. Luego de cada bloque temático de unas tres

clases teóricas tendrá lugar una clase práctica. En las clases prácticas se formularán ejercicios que permitan mejorar la comprensión de los aspectos teóricos abordados en el curso. Los estudiantes desarrollarán habilidades prácticas relativas a la programación y aplicarán técnicas de aprendizaje automático sobre datos reales.

Horas de teórico: 46

Horas de práctico: 14

Horas de evaluación: 8

Horas estimadas de dedicación no presencial: 52

Horas de dedicación total: 120

5. TEMARIO

1. Introducción al aprendizaje automático: Componentes del problema de aprendizaje automático: conjunto de datos, función objetivo, conjunto de hipótesis, algoritmo de aprendizaje, función de costo. Tipos de aprendizaje: supervisado, no supervisado y por refuerzos. Otras perspectivas sobre el aprendizaje automático. Consideraciones éticas.
2. Teoría del aprendizaje: Factibilidad del aprendizaje. Desempeño frente a nuevos datos. Medidas de error, funciones de costo. Aprendizaje frente a datos ruidosos. Diferencia entre datos de entrenamiento y datos de prueba en términos matemáticos. Teoría de la generalización. Dimensión de Vapnik-Chevonenkis (VC). Cotas de generalización. Relación con el número de parámetros y los grados de libertad. Compromiso aproximación y generalización. Generalización vista desde el compromiso sesgo y varianza. Curvas de aprendizaje.
3. Sobreajuste a los datos: Sobreajuste y ajuste al ruido. Ruido determinístico versus ruido estocástico. Sobreajuste con polinomios. Regularización. Decaimiento de pesos. Elección del regularizador. Vínculo con dimensión VC. Conjunto de validación. Selección de modelos. Validación cruzada.
4. El modelo lineal: Clasificación lineal. El algoritmo de aprendizaje del perceptrón. Datos no linealmente separables. El algoritmo pocket. Regresión lineal. Regresión logística. Introducción a la optimización convexa. Descenso por el gradiente. Extendiendo el modelo lineal mediante transformaciones no lineales.
5. Métodos basados en similitud: Medidas de similitud. Vecinos más cercanos (k-NN). Funciones de base radial (RBF). Agrupamiento no supervisado con k-Means. Estimación de densidades de probabilidad. Ventanas de Parzen. Modelo de mezcla de Gaussianas (GMM). El algoritmo de máxima verosimilitud (EM).
6. Modelos no lineales: Combinación de perceptrones. El perceptrón multicapa (MLP). Redes neuronales (ANN). Función de activación diferenciable. Aproximación universal. Propagación hacia adelante. El algoritmo de propagación hacia atrás. Gradiente descendente estocástico. Máquinas de vectores de soporte (SVM). El hiperplano óptimo. Maximización del margen. Programación cuadrática. Vectores de soporte. Formulación del problema dual. Métodos de kernel. Soft margins SVMs.

7. Ingeniería de características y pre-procesado de datos: Estandarización de características: centrado, normalización y blanqueado. Reducción de la dimensionalidad y selección de características. Análisis de componentes principales. Limpieza de datos. Identificación y tratamiento de outliers.

6. BIBLIOGRAFÍA

6.1 Básica

1. Yaser S. Abu-Mostafa, Malik Magdon-Ismail, Hsuan-Tien Lin (2012). Learning from data. USA. AMLBook.

6.2 Complementaria

2. Bishop, Christopher M. (2011). Pattern Recognition and Machine Learning. USA. Springer.

7. CONOCIMIENTOS PREVIOS EXIGIDOS Y RECOMENDADOS

7.1 Conocimientos Previos Exigidos: Probabilidad y estadística, álgebra lineal, programación, aplicación de transformaciones a señales, análisis de sistemas, procesamiento de señales.

7.2 Conocimientos Previos Recomendados:

ANEXO A
Para todas las Carreras

A1) INSTITUTO

Instituto de Ingeniería Eléctrica

A2) CRONOGRAMA TENTATIVO

Consiste en un cronograma de avance semanal con detalle de las horas de clase asignadas a cada tema.

Semana 1	Tema 1 <u>Introducción</u> : Componentes del problema de aprendizaje automático. El algoritmo de aprendizaje del perceptrón. Datos linealmente separables. Tipos de aprendizaje. Otras perspectivas sobre el aprendizaje automático. Consideraciones éticas. (2 hs de clase). Tema 2 <u>Teoría del aprendizaje</u> : Factibilidad del aprendizaje. Modelo PAC. Cota dada por la desigualdad de Hoeffding. (2 hs de clase).
Semana 2	<i>Práctico 1 (Algoritmo de aprendizaje del perceptrón. La desigualdad de Hoeffding. Datos no separables, algoritmo pocket.)</i> . (2 hs de clase). Tema 4: <u>El modelo lineal</u> : Clasificación lineal. Extracción de características. Datos no linealmente separables. Regresión lineal. Extendiendo el modelo lineal mediante transformaciones no lineales. (2 hs de clase).
Semana 3	Tema 2 <u>Teoría del aprendizaje</u> : Incorporar al problema las nociones de error y de ruido. Medidas de error, funciones de costo. Aprendizaje frente a datos ruidosos. Motivación de la teoría del aprendizaje. (2 hs de clase) Tema 2 <u>Teoría del aprendizaje</u> : Implicaciones conceptuales y prácticas del contraste entre entrenamiento y prueba. Número efectivo de hipótesis. El concepto de punto de quiebre. (2 hs de clase)
Semana 4	<i>Práctico 2 (Clasificación lineal con datos reales. Train y test. Extracción de características. Regresión lineal. Error y ruido.)</i> . (2 hs de clase). Tema 2 <u>Teoría del aprendizaje</u> : Teoría de la generalización. Cota polinómica de la función de crecimiento: La dimensión de Vapnik-Chervonenkis (VC). Interpretación de la dimensión VC y cota de generalización. (2 hs de clase)
Semana 5	Tema 2 <u>Teoría del aprendizaje</u> : Generalización vista desde el compromiso Sesgo-Varianza. Curvas de aprendizaje. (2 hs de clase) Tema 4 <u>El modelo lineal</u> : Regresión logística. Introducción a la optimización convexa. Gradiente descendente. (2 hs de clase)
Semana 6	<i>Práctico 3 (Teoría de la generalización. Regresión logística. Gradiente descendente. Optimización convexa)</i> . (2 hs de clase). Tema 3 <u>Sobreajuste a los datos</u> : Sobreajuste y ajuste al ruido. Ruido determinístico y ruido estocástico. (2 hs de clase)

Semana 7	<p>Tema 3 <u>Sobreajuste a los datos</u>: Regularización. Regularización de orden suave. Decaimiento de pesos. Elección del regularizador. Vínculo con dimensión VC. (2 hs de clase)</p> <p>Tema 3 <u>Sobreajuste a los datos</u>: Conjunto de validación. Selección de modelos. Validación Cruzada. Teoría versus práctica en el sobreajuste. (2 hs de clase)</p>
Semana 8	<p><i>Práctico 4 (Sobreajuste, regularización y validación, usando regresión lineal y regresión logística.). (2 hs de clase)</i></p> <p><i>Repaso para el parcial (2 hs de clase)</i></p>
Semana 9	<p>Tema 2 <u>Teoría del aprendizaje</u> y Tema 3 <u>Sobreajuste a los datos</u>: Epílogo de los fundamentos teóricos: navaja de Occam, sesgo del muestreo y data snooping. (2 hs de clase)</p> <p>Tema 5 <u>Métodos basados en similitud</u>: Medidas de similitud. Vecino(s) más cercano(s) (k-NN). Rol y selección de k. Eficiencia. (2 hs de clase)</p>
Semana 10	<p>Tema 5 <u>Métodos basados en similitud</u>: Funciones de base radial (RBF). Agrupamiento no supervisado con k-Means. (2 hs de clase)</p> <p><i>Práctico 5 (Clasificación con k-NN. Funciones de base radial. Agrupamiento con k-Means.). (2 hs de clase)</i></p>
Semana 11	<p>Tema 5 <u>Métodos basados en similitud</u>: Estimación de densidades de probabilidad. Ventanas de Parzen. Modelo de mezcla de Gaussianas (GMM). El algoritmo de máxima verosimilitud (EM). (2 hs de clase)</p> <p>Tema 6 <u>Modelos no lineales</u>: El perceptrón multicapa. Propagación hacia adelante. Función de activación diferenciable. Aproximación universal. (2 hs de clase)</p>
Semana 12	<p>Tema 6 <u>Modelos no lineales</u>: Redes neuronales. El algoritmo de propagación hacia atrás. Gradiente descendente estocástico. Inicialización, terminación, selección de la tasa de aprendizaje. Aproximación versus generalización. Regularización y validación. (2 hs de clase)</p> <p><i>Práctico 6 (Perceptrón multicapa. Propagación hacia atrás. Gradiente descendente estocástico. Regularización y validación.). (2 hs de clase)</i></p>
Semana 13	<p>Tema 6 <u>Modelos no lineales</u>: Máquinas de vectores de soporte. Hiperplano óptimo. Maximización del margen. Programación cuadrática. Vectores de soporte. Formulación del problema dual. (2 hs de clase)</p> <p>Tema 6 <u>Modelos no lineales</u>: Métodos de kernel. Transformaciones no lineales. El truco del kernel. SVM de margen suave. (2 hs de clase)</p>
Semana 14	<p><i>Práctico 7 (Máquinas de vectores de soporte y métodos de kernel. Maximización del margen y margen suave.). (2 hs de clase)</i></p> <p>Tema 7 <u>Ingeniería de características y limpieza de datos</u>: Normalización, centrado y blanqueado. Reducción de la dimensionalidad y selección de características. Análisis de componentes principales (PCA). (2 hs de clase)</p>
Semana 15	<p>Tema 7 <u>Ingeniería de características y limpieza de datos</u>: Detección y tratamiento de outliers. (2 hs de clase)</p> <p><i>Repaso para el parcial (2 hs de clase)</i></p>

A3) MODALIDAD DEL CURSO Y PROCEDIMIENTO DE EVALUACIÓN

El proceso de evaluación incluye la realización individual de 6 controles de lectura y de dos parciales.

Controles de lectura:

- se realizarán al finalizar cada unidad temática y cada control consistirá en un conjunto de preguntas que el estudiante debe responder
- se dispondrán a través del EVA y el estudiante deberá realizar el control dentro de un plazo de tiempo previamente establecido
- los controles corresponden al 20% del puntaje total de la unidad curricular
- para aprobar la unidad curricular el estudiante deberá realizar al menos 5 de los 6 controles y alcanzar como mínimo el 50% de los puntos del total de los controles, no existen mínimo de puntos para cada control

Parciales:

- se realizarán en los períodos de parciales, en forma presencial
- consistirán en la resolución de problemas y ejercicios
- los controles corresponden al 80% del puntaje total de la unidad curricular
- para aprobar la unidad curricular el estudiante deberá realizar los dos parciales y alcanzar un mínimo de 25% en cada uno de ellos, y además deberá alcanzar un mínimo de 60% en la suma de los puntos de los dos parciales

Aquellos estudiantes que no alcancen el puntaje de aprobación, pero que cumplan las siguientes dos condiciones, podrán rendir un tercer parcial de recuperación.

Condiciones:

- superar el 50% de la suma de los puntos de los parciales
- superar las exigencias mínimas para la aprobación (50% de los controles y 25% de cada parcial, al menos 5 de los 6 controles)

Tercer parcial:

- tendrá lugar entre el segundo período de parciales y el primer período de exámenes
- para aprobar la unidad curricular el estudiante deberá alcanzar un mínimo de 60% de los puntos del parcial

A4) CALIDAD DE LIBRE

Los estudiantes no podrán acceder a la Calidad de Libre.

A5) CUPOS DE LA UNIDAD CURRICULAR

No se definen cupos.

ANEXO B para la carrera Ingeniería Eléctrica .

B1) ÁREA DE FORMACIÓN

Telecomunicaciones

B2) UNIDADES CURRICULARES PREVIAS

Curso: los exámenes de probabilidad y Estadística y Programación 1, el curso de Señales y Sistemas y 50 créditos en Matemáticas.

Alternativamente podrán cursar quienes tengan el curso de Muestreo y Procesamiento Digital, el curso de Probabilidad y Estadística y el examen de Programación 1.

Examen:

APROB. RES. CONSEJO DE FAC. ING.

Fecha 11/06/2019 Exp. 060180-001105-19