



Aprendizaje Automático

1. Características generales

Nombre:	Aprendizaje Automático
Docentes:	Markus Eger, Jose Guevara Coto
Sigla:	CI-2600
Créditos:	4
Horas:	4 horas de teoría
Requisitos:	CI-1441 Paradigmas Computacionales
Correquisitos:	ninguno
Clasificación:	Curso propio, electivo
Ciclo:	VIII ciclo, 4to año
Horario:	Martes, Jueves 08:00-09:50hrs, Miércoles 08:00-11:50hrs
Consulta:	Markus Eger: Miércoles 13:30-15:00hrs, oficina 3-23 Jose Guevara Coto: Miércoles 13:30-15:00hrs, oficina 3-18

2. Descripción

Este curso es teórico-práctico, que busca introducir a los participantes a los temas del aprendizaje automático y profundo, con el fin de aplicar los conocimientos adquiridos en resolución de problemas y análisis de datos en diversos escenarios, desde juegos de video, datos económicos, e inclusive bioinformáticos. Durante el curso se cubrirán los fundamentos teóricos y prácticos de los algoritmos utilizados en el análisis de estos datos, su selección, e implementación. Asimismo, se incorporarán los conceptos de análisis de datos exploratorio, procesamiento de datos, entrenamiento de modelos, afinamiento mediante selección de variables, y validación de los modelos mediante métricas de desempeño. El curso está diseñado para estudiantes de Bachillerato en Computación e Informática, con conocimiento en Paradigmas Computacionales.

3. Objetivos

Objetivo general

El objetivo general es que los participantes conozcan los fundamentos teóricos y prácticos de los algoritmos de aprendizaje automático y cómo estos pueden ser aplicados para descubrir





conocimiento nuevo. Las y los estudiantes desarrollarán modelos y los utilizarán en tareas de clasificación y regresión, aplicando el conocimiento teórico adquirido.

Objetivos específicos

El objetivo general es que las y los estudiantes conozcan los datos utilizados en bioinformática y cómo pueden ser analizados mediante algoritmos de aprendizaje automático para descubrir conocimiento nuevo. Las y los estudiantes desarrollarán modelos y los utilizarán en tareas de clasificación y regresión, aplicando el conocimiento teórico adquirido.

Objetivos específicos

A lo largo del curso el o la estudiante desarrollará habilidades que le permitirán:

1. Familiarizarse con los algoritmos de aprendizaje automático y profundo de mayor aplicabilidad en ejes temáticos como análisis de datos.
2. Identificar las tendencias en los datos y el comportamiento en los mismos mediante el análisis exploratorio de datos.
3. Fomentar las buenas prácticas de reproducibilidad y proveniencia en experimentos de aprendizaje automático y profundo.
4. Aplicar flujos de procesos de aprendizaje automático y profundo para analizar de diversas fuentes (texto, video juegos, etc.).

4. Contenidos

Los ejes temáticos del curso y los objetivos a los que contribuyen se muestran en la tabla que sigue.

Eje temático	Desglose	Clases	Semana
	Lectura y discusión de la Carta al Estudiante	1	
Introducción a PyTorch y manejo de datos	Introducción al ambiente de trabajo con la librería PyTorch, a las buenas prácticas de trabajo (reproducibilidad en código). Introducción al Análisis Exploratorio de Datos.	3	1





Fundamentos del Aprendizaje Automático	Breve introducción a la clasificación y a la regresión. Se comparan y contrastan ambos métodos y los algoritmos usados en ellos. Se describen las métricas de desempeño	2	
Redes Neuronales	Descripción de las redes neuronales y sus fundamentos para realizar tareas de regresión y clasificación.	2	
Redes Neuronales Avanzadas	Conceptualización de redes neuronales avanzadas para resolución de tareas novedosas. Se introducen las GANs	3	
Máquinas de Soporte Vectorial	Introducción teórica a las máquinas de soporte vectorial. Aplicaciones de modelos construidos con estos algoritmos para resolución de tareas de clasificación y regresión	1	
Examen Parcial I		1	
Aprendizaje Semi-supervisado	Se introducen los conceptos de aprendizaje no supervisado en un marco teórico y sus aplicaciones en datos donde los conjuntos no poseen suficientes instancias etiquetadas. Se discuten avances en el aprendizaje semi-supervisado	1	
Aprendizaje No-supervisado	Introducción al aprendizaje no supervisado como alternativa cuando los datos no están etiquetados. Se introducen agrupamiento, K-medias y Expectation Maximization (EM).	1	





Aprendizaje Reforzado	Introducción teórica en Markov Decision Processes, la base de aprendizaje reforzado. Discusión de algoritmos Q-Learning y SARSA, y aplicaciones de redes neuronales para aprendizaje reforzado (Deep Q-Learning).	3	
Afinamiento de Modelos de Aprendizaje Automático	Introducción teórica a los conceptos de ingeniería de variables, donde se incluye la construcción de nuevos atributos a partir de conocimiento de experto. Se introducen los conceptos de selección y extracción de variables (reducción de dimensionalidad), y afinamiento. Se explora el efecto del afinamiento de modelo en el sobreajuste de modelos	1	
Planteamientos éticos	Se presentan planteamientos éticos y dilemas de esta índole asociados al uso indebido del aprendizaje automático	1	
	Examen parcial II	1	

5. Metodología

El curso tendrá una duración de 8 semanas, durante el período del III ciclo del 2019 (verano). El curso CI-2600 “**Aprendizaje Automático**” es teórico práctico y requiere de un nivel de involucramiento y participación por parte de las y los estudiantes significativo. Por consiguiente, es requerido que los participantes del curso lean la bibliografía del curso.

Las clases teóricas serán complementadas por sesiones de laboratorio donde las personas participantes del curso pondrán en ejecución los conceptos transmitidos en las sesiones teóricas. El total de las sesiones de laboratorio son 8, y serán evaluados mediante entregables, cada uno con un valor de 7.5%, para un valor total de 60% de la nota final del curso. Estos deberán ser sometidos en las fechas indicadas por los instructores. La no entrega o incumplimiento de la entrega equivale a un cero en la nota para ese laboratorio.

El curso consta de dos exámenes, los cuales evaluarán la materia correspondiente a los módulos estudiados una semana antes de la evaluación. Las evaluaciones son de carácter individual, de carácter presencial (no podrán ser realizados virtualmente), y no se podrá





contar con ningún tipo de material de apoyo para su realización. Los exámenes se realizarán el 4/2 y el 4/3. Cada evaluación o examen tendrá un valor de 20% de la nota final.

Durante la realización de cualquier trabajo y/o evaluación debe recordar que usted es estudiante de la Universidad de Costa Rica, y de la Escuela de Ciencias de la Computación e Informática, por consiguiente, está sujeto a los más altos estándares de ética definidos en el “REGLAMENTO DE ORDEN Y DISCIPLINA DE LOS ESTUDIANTES DE LA UNIVERSIDAD DE COSTA RICA”

6. Evaluación

A continuación, se presenta el resumen de acuerdo con los criterios citados anteriormente:

Asignación	Valor
Laboratorios (8 unidades, de 7.5% cada uno)	60%
Examen 1	20%
Examen 2	20%
TOTAL	100%

Para aprobar el curso el estudiante debe tener una nota igual o superior a 6.75. Si la nota final está entre 5.75 y 6.74 tendrá derecho a realizar un examen de ampliación, el cual incluye toda la materia del curso; en dicho examen deberá tener una nota mínima de 7.0 para aprobar el curso, siendo la nota final 7.0. En caso de que el estudiante tenga una nota igual o inferior a 5.74, o bien en caso de presentar el examen de ampliación con una nota inferior a 7.0 reprobará el curso.

Observaciones:

- Por cada de las laboratorios los estudiantes tienen que presentar un reporte con sus resultados, y también tienen que enviar el código fuente usado para producir estos resultados. La reproducibilidad de los resultados es de gran importancia.
- Todo trabajo de los laboratorios debe ser entregado de forma digital por medio de correo electrónico a los dos profesores y el asistente.
- Los quices se harán en todas las lecciones y en cualquier momento durante el transcurso de la lección, y solo se repondrán en los casos que establece el Reglamento de Régimen Académico Estudiantil en su artículo 24. En los quices se puede utilizar material de apoyo; pueden ser individuales o grupales.
- La investigación consiste en la presentación al grupo, en grupo de cuatro personas como máximo, de una aplicación sobre temas vistos por el profesor en clases. Se formarán los grupos y se asignarán los temas el primer día de clases.
- Los criterios de evaluación de cada trabajo asignado se entregarán oportunamente.





- El uso de lápiz en cualquier evaluación se permite, pero no se acepta reclamos. Por lo que, el uso del lapicero es recomendado.
- Por cada día natural de retraso en la entrega de cualquier trabajo del curso se rebajará un punto en la escala de 1 a 10.
- Si envía por correo electrónico con uno o más días de retraso, se aplicará la regla de rebajo de puntos expuesta arriba con base en la fecha de envío.
- Cuando el estudiante sepa que tendrá que faltar un día particular en el cual debe entregar algún trabajo, se recomienda que lo comunique a la profesora antes de ese día, para coordinar la entrega de alguna forma.
- En todos los trabajos y las evaluaciones de los estudiantes, se calificará la redacción y ortografía; por lo que, se rebajará de la nota obtenida un punto por cada falta de ortografía y mala redacción.
- En todos los trabajos y las evaluaciones de los estudiantes, se debe entregar una autoevaluación y coevaluación, con el fin de evaluar el aporte de sus compañeros de equipo y el propio. La calificación final del trabajo o evaluación será el promedio entre la calificación obtenida y las evaluaciones (autoevaluación y coevaluación).
- En todos los trabajos y las evaluaciones, se le solicitará al estudiante que firme una lista de entrega, para que el estudiante tenga un documento que compruebe que entregó y realizó lo solicitado por el profesor.
- Los estudiantes pueden discutir los trabajos con quien sea. Esto incluye hablar sobre interpretaciones del ejercicio asignado, por donde se podría atacar el problema, inclusive la estrategia completa de solución (si alguno de los que discuten ya lo resolvió); además, se puede sugerir y/o buscar material complementario, etc. Lo que no se puede es copiar la solución de ninguna fuente, ya sea un(a) compañero(a), un libro, Internet, etc.
- Los trabajos serán revisados por los profesores y/o el asistente, si se encuentra evidencia de cualquier tipo de copia, y es la primera vez, los estudiantes involucrados tendrán un cero de nota. Si ocurre una segunda vez, el caso se remitirá a la comisión disciplinaria de la ECCI para aplicar el reglamento.
- En cuanto a reportes y presentaciones se castigará el plagio, el cual se considera copia y será castigado de la misma manera. Se considera plagio la copia literal de segmentos (texto, figuras, tablas u otros datos no textuales) de otra fuente, sin comillas y sin referencia, aunque sea traducido, así como el parafraseo sin aportes de ningún tipo.
- Cuando un(a) estudiante no pueda asistir a efectuar una evaluación por alguna razón de fuerza mayor: la muerte de un pariente hasta de segundo grado, la enfermedad del estudiante u otra situación de fuerza mayor o caso fortuito; se seguirá con lo normado en el artículo 24 del Reglamento de Régimen Académico Estudiantil.

7. Cronograma

Clases

N°	Fecha Inicio	Fecha Fin	Duración (clases)	Tema
----	--------------	-----------	-------------------	------





1	07/01	09/01	2	Introducción
2	14/01	14/01	1	Aprendizaje estadístico
3	16/01	28/01	4	Principios básicos
4	30/01	30/01	1	Support Vector Machines
5	06/02	11/2	2	Aprendizaje semi supervisado Aprendizaje no supervisado
6	13/02	18/02	2	Aprendizaje por refuerzo
7	20/02	25/02	2	Aplicaciones
8	27/02	27/02	1	Resumen y Conclusión

Laboratorios

N°	Fecha	Duración (clases)	Tema
1	08/01	1	Data management
2	15/01	1	Aprendizaje estadístico
3	22/01	1	Redes neuronales
4	29/01	1	Generative Adversarial Networks
5	05/02	1	Support Vector Machines
6	12/02	1	Clustering
7	19/02	1	Aprendizaje por refuerzo
8	26/02	1	Planteamientos éticas

8. Software

Para llevar a cabo las asignaciones del curso es necesario el siguiente software

Software

Descripción

Python 3 y librerías

Lenguaje de programación y módulos para análisis y construcción de modelos de aprendizaje automático/profundo





9. Bibliografía

1. Stevens, E. and Antiga, L. *Deep Learning with PyTorch: Essential Excerpts*, 2019

