



INSTITUTO/S: Tecnología e Ingeniería

CARRERA/S: Licenciatura en Informática

MATERIA: Fundamentos de Aprendizaje Automático

NOMBRE DEL RESPONSABLE DE LA ASIGNATURA: Santos, Juan Miguel

EQUIPO DOCENTE: Santos, Juan Miguel

CUATRIMESTRE: 1^{ro}

AÑO: 5^{to}

PROGRAMA N°: 38

(Aprob. Por Cons.Directivo fecha XX)

Instituto/s: Tecnología e Ingeniería

Carrera/s: Licenciatura en Informática

Nombre de la materia: Fundamentos de aprendizaje automático

Responsable de la asignatura y equipo docente: Santos, Juan Miguel

Cuatrimestre y año: 1^{er} cuat. del 5^{to} año

Carga horaria semanal: 4 hs

Programa N°: 38

Código de la materia en SIU: ING_COMP13

Fundamentos de Aprendizaje Automático

1. Fundamentación

El objeto de estudio de Fundamentos de aprendizaje automático son los métodos y algoritmos que son implementados en programas que aprenden. Aprender, en el contexto de este tema y esta materia refiere a la capacidad que tiene un programa de mejorar una medida de desempeño, previamente definida, a través de la experiencia (cantidad y veces que los datos son leídos por el programa).

Los métodos estudiados en Fundamentos de aprendizaje automático permiten al licenciado tratar con problemas asociados a grandes volúmenes de datos, tales como clasificar, asociar, inferir y modelar.

El graduado de la licenciatura podrá evaluar diferentes métodos para tratar con un problema e implementar computacionalmente una solución, lo cual le otorga a su perfil flexibilidad para enfrentarse a diferentes situaciones su vida profesional.

2. Propósitos y/u objetivos

Propósitos

- Promover la comprensión de la naturaleza estadística o analítica o algorítmica de los métodos estudiados.
- Contribuir al análisis del problema desde el punto de vista del diseño de la solución: ¿cómo elegir los datos?, ¿cómo elegir la medida de desempeño del programa? ¿cómo elegir el método que mejor se adecúa a la solución del problema?
- Identificar las métricas que permitan evaluar los resultados de acuerdo al problema que se desea resolver.

Objetivos

Son objetivos de esta materia que los/as estudiantes:

- Reflexionen acerca de las posibilidades que tienen los métodos estudiados para abordar cada clase de problema.
- Identifiquen las características de un problema en concreto y las vinculen con el diseño de la solución.

3. Programa sintético:

1. Introducción y métricas de evaluación
2. Aprendizaje supervisado
3. Regresión
4. Aprendizaje no supervisado

4. Programa analítico

1. Introducción y métricas de evaluación

Aprendizaje, inferencia y predicción.

Conjunto de datos, atributos explicativos y atributo objetivo.

Definición de programa que aprende.

Tipos de problemas que resuelve el aprendizaje automático.

Ejemplo de método de aprendizaje: Find-S

Precisión, Accuracy, Recall y F1-score.

Matriz de confusión.

Histograma y gráficos box-plot.

Curva ROC.

Validación cruzada y Bootstrap.

2. Aprendizaje supervisado

Probabilidad condicional.

Teorema de Bayes.

Aprendizaje Bayesiano.

Clasificador ingenuo de Bayes.

Ejemplos de aplicación.

Árboles de decisión. Random Forest.

K-vecinos más cercanos y K-vecinos más cercanos pesado.

Concepto de separabilidad lineal.

Perceptron simple.

SVM (máquinas con vectores soportes).

3. Regresión.

Regresión lineal simple.

Regresión lineal múltiple.

Regresión logística como clasificador.

4. Aprendizaje no supervisado

K-medias. Medida de variación.
Agrupamiento jerárquico.
Dendrograma.

4.1 Organización del contenido:

Describir brevemente el contenido de cada bloque, eje y/o unidad.

El contenido se organiza a través de cuatro trabajos prácticos.

EL primer trabajo práctico corresponde a Aprendizaje Bayesiano. En este trabajo los alumnos tratarán con un conjunto de datos que deben analizar antes de usarlo para implementar el clasificador de Bayes. También usarán métricas de evaluación para evaluar los resultados obtenidos. La idea fundamental del primer bloque es que los alumnos traten con datos reales y reflexiones sobre el significado de inferencia estadística.

El segundo trabajo práctico corresponde a Árboles de decisión y K-vecinos-más-cercanos que son métodos supervisados de clasificación. En este caso, se buscará que los alumnos interioricen la importancia de separar el conjunto de datos en, datos de entrenamiento y datos de prueba. Adicionalmente, los alumnos explorarán sobre el concepto de sobreajuste usando como recurso la poda de los árboles de decisión.

El tercer trabajo práctico está orientado a máquinas con vectores soportes (SVM). La expectativa es que los alumnos puedan ver en términos prácticos el concepto de separabilidad lineal mediante la implementación del perceptrón. Por otro lado, a diferencia de los trabajos prácticos anteriores, podrán usar una librería (en vez de programarlo ellos) para evaluar los resultados del método SVM usando núcleos lineales, radiales y polinómicos. En este trabajo práctico, por primera vez, deberán los alumnos construir su conjunto de datos a partir de imágenes y luego usar SVM para segmentarla en regiones.

El cuarto trabajo práctico apunta a regresión logística y aprendizaje no supervisado (k-medias y agrupación jerárquica). Para la parte de regresión logística, se usarán datos de pacientes que sufren una enfermedad. Luego, se pedirá que, al mismo conjunto de datos, ignorando el atributo objetivo, implementen k-medias y agrupamiento jerárquico. De esta forma, al conocer el valor de atributo objetivo y no usarlo, se espera que los alumnos puedan reflexionar acerca de los métodos usados a la luz de los resultados obtenidos.

4.2 Bibliografía y recursos obligatorios:

- Tom M Mitchell and Tom M Mitchell. (1997). *Machine learning, volume 1*. McGraw-hill New York.

4.3 Bibliografía optativa:

- Deisenroth, M. P., Faisal, A. A., y Ong, C. S. (2020). *Mathematics for machine learning*. Cambridge University Press.

- Kubat, M., y Kubat. (2017). *An introduction to machine learning* (Vol. 2). Cham, Switzerland: Springer International Publishing.

5. Metodología de enseñanza:

La materia prevé cuatro tipos de metodologías: clases teóricas, clases de consulta, evaluación y presentación de los trabajos.

En las clases teóricas se presentan los métodos a utilizar, los fundamentos teóricos de los mismos, definiciones, algoritmos y ejemplos de aplicación. En las clases teóricas se promueve la participación de los alumnos mediante consultas o intervenciones vinculadas al contenido de la clase.

En las clases de consulta, los/as alumnos/as asisten con dudas sobre los contenidos vistos en las clases teóricas o con dudas concretas acerca de la resolución del trabajo práctico, tales como en la toma de una decisión, descripción de un error en particular o dificultades en la implementación (codificación en un lenguaje en particular).

En las clases de evaluación, a los alumnos se le entregará una serie de preguntas a responder y que se corresponden con los temas vistos en las teóricas. La evaluación es individual.

En las clases de presentación, los grupos (entre 2 y 3 alumnos) hacen la presentación del trabajo práctico. Se le sugiere que lo hagan mediante una presentación en power point, page-beamer o cualquier otra herramienta útil para tal fin.

Plan de trabajo en el campus:

El campus permitirá compartir el material de clase (diapositivas usadas durante las teóricas), enunciados de los trabajos prácticos, enunciados de ejercicios para realizar, ejemplos de código para que los alumnos se familiaricen con la programación de los métodos y archivos con conjuntos de datos necesarios para los trabajos prácticos y ejercicios.

6. Actividades de investigación y extensión (si hubiera)

No aplica.

7. Evaluación y régimen de aprobación

7.1 Aprobación de la cursada

Para aprobar la cursada y obtener la condición de regular, el régimen académico establece que debe obtenerse una nota no inferior a cuatro (4) puntos. Todas las instancias evaluativas deberán tener una instancia de recuperatorio. Podrán acceder a la administración de esta modalidad solo aquellos y aquellas estudiantes que hayan obtenido una nota inferior o igual a 6 (seis) puntos en el examen parcial.

Siempre que se realice una evaluación de carácter recuperatorio, la calificación que los/as estudiantes obtengan reemplazará la calificación obtenida en el examen que se ha recuperado y será la considerada definitiva a los efectos de la aprobación.

El/La alumno/a deberá poseer una asistencia no inferior al 75% en las clases presenciales. En cuanto a las cursadas de materias virtuales se requerirá que el estudiante ingrese al aula virtual como mínimo una vez por semana.

7.2 Aprobación de la materia

La materia puede aprobarse por promoción, evaluación integradora, examen final o libre.

Promoción directa: tal como lo establece el art°17 del Régimen Académico, para acceder a esta modalidad, el/la estudiante deberá aprobar la cursada de la materia con una nota no inferior a siete (7) puntos, no obteniendo en ninguna de las instancias de evaluación parcial menos de seis (6) puntos, sean evaluaciones parciales o recuperatorios. El promedio estricto resultante deberá ser una nota igual o superior a siete (7) sin mediar ningún redondeo.

Evaluación integradora: tal como lo establece el art°18 del Régimen Académico, podrán acceder a esta evaluación aquellos estudiantes que hayan aprobado la cursada con una nota de entre cuatro (4) y seis (6) puntos.

La evaluación integradora tendrá lugar por única vez en el primer llamado a exámenes finales posterior al término de la cursada. Deberá tener lugar en el mismo día y horario de la cursada y será administrado, preferentemente, por el/la docente a cargo de la comisión. Se aprobará tal instancia con una nota igual o superior a cuatro (4) puntos, significando la aprobación de la materia.

La nota obtenida se promediará con la nota de la cursada.

Examen final: Instancia destinada a quienes opten por no rendir la evaluación integradora o hayan regularizado la materia en cuatrimestres anteriores. Se evalúa la totalidad de los contenidos del programa de la materia y se aprueba con una calificación igual o superior a cuatro (4) puntos. Esta nota no se promedia con la cursada.

7.3 Criterios de calificación

La calificación de cada evaluación se determinará en la escala 0 a10, con los siguientes valores: 0, 1, 2 y 3: insuficientes; 4 y 5 regular; 6 y 7 bueno; 8 y 9 distinguido; 10 sobresaliente.

8. Cronograma

Semana	Tema	Modalidad
1	Introducción. Preliminares.	Presencial
2	Método de Bayes	Virtual
3	Métricas de evaluación. Consultas TP.	Presencial/Virtual
4	Evaluación 1. Presentación TP1.	Presencial
5	Árboles de decisión.	Virtual
6	Random forest. K-vecinos-más-cercanos.	Virtual
7	Consultas TP.	Presencial/Virtual
8	Evaluación 2. Presentación TP2.	Presencial

9	Regresión, SVM.	Virtual
10	SVM (continuación)	Virtual
11	Consultas TP.	Presencial/Virtual
12	Evaluación 3. Presentación TP3.	Presencial
13	K-medias. Agrupamiento jerárquico.	Virtual
14	Consultas TP.	Presencial/Virtual
15	Consultas TP.	Presencial/Virtual
16	Evaluación 4. Presentación TP4.	Presencial