



Expte. n° FIQ-0955240-18

SANTA FE, 18 de Diciembre de 2018

VISTAS las actuaciones vinculadas con la elevación de la programación del curso de verano "Minería de Datos y Aprendizaje Automático", como curso extracurricular, efectuada por el Dr. Diego TOMASSI; y

CONSIDERANDO:

El aval del Director del Departamento de Matemática, lo expresado por la Secretaría Académica de la Facultad, como así también lo aconsejado por la Comisión de Enseñanza;

**EL CONSEJO DIRECTIVO
DE LA FACULTAD DE INGENIERIA QUIMICA**

ARTÍCULO 1º.- Aprobar el curso de verano "Minería de Datos y Aprendizaje Automático", como curso extracurricular de formación general.

ARTICULO 2º.- Aprobar el programa analítico, bibliografía y planificación del curso "Minería de Datos y Aprendizaje Automático", que como anexo forma parte integrante de la presente.

ARTÍCULO 3º.- Designar como Profesor Responsable del dictado del curso, al Dr. Diego TOMASSI.

ARTÍCULO 4º.- Designar el Tribunal Examinador con los siguientes docentes:

Titulares: Dr. Diego TOMASSI, Dra. Pamela LLOP y Dra. Liliana FORZANI

Suplentes: Dra. Sabrina DUARTE, Dra. Andrea BERGESIO y Dra. María Antonella GIEGO

ARTÍCULO 5º.- Inscribase, comuníquese, hágase saber en copia a Departamento Alumnado, Bedelía, Secretaría Académica y archívese.

RESOLUCION "C.D." N° 654

Abog. CECILIA BONGIOVANNI
PROSECRETARIA

Dr. ADRIAN BONIVARDI
DECANO

1918-
2018

En el año
del centenario
de la Reforma
Universitaria



Expte. n° FIQ-0955240-18
Resolución "C.D." n° 654/18

ANEXO

Departamento de Matemática

Asignatura: **"Minería de Datos y Aprendizaje Automático"**

Carreras: Optativa para las carreras de Licenciatura en Matemática Aplicada, Ingeniería Industrial, Ingeniería Química, Ingeniería en Alimentos, Ingeniería en Materiales, Licenciatura en Materiales

Carácter: Curso de Verano Extracurricular

Carga Horaria: 90 horas

Programa Analítico

UNIDAD I

Introducción al aprendizaje automático y la minería de datos. Aprendizaje supervisado vs no supervisado. Regresión, clasificación y clustering. Criterios de optimalidad. Riesgo de Bayes. Riesgo empírico y riesgo estructural.

UNIDAD II

Aprendizaje supervisado: modelos lineales. Regresión lineal múltiple. Regresión logística. Análisis discriminante lineal y cuadrático. Validación cruzada y métodos de remuestreo para selección y validación de modelos. Curvas ROC.

UNIDAD III

Aprendizaje supervisado: modelos lineales en alta dimensión. Selección por pasos vs regularización. Regresión ridge. Lasso y lasso por grupos. Reducción de dimensiones: regresión parcial y correlaciones canónicas.

UNIDAD IV

Aprendizaje supervisado: métodos basados en árboles. Particionado recursivo: Árboles de regresión y de clasificación. Bagging, boosting y random forests.

UNIDAD V

Aprendizaje supervisado: métodos basados en núcleos. Vecinos más cercanos. Criterio de margen máximo. Máquinas de vectores soporte para clasificación y regresión.

UNIDAD VI

Aprendizaje no supervisado: clustering. K-medias, Mean-shift, level sets. Clustering basado en modelos de mezclas y algoritmo EM. Clustering basado en grafos. Clustering jerárquico. Clustering de variables y biclustering.

UNIDAD VII

Aprendizaje no supervisado: reducción de dimensiones y modelos de variables latentes. Análisis de componentes principales (PCA). Análisis de componentes independientes. Análisis de factores. Métodos no-lineales de reducción dimensional.

UNIDAD VIII

Aprendizaje no supervisado: explorando dependencias entre variables. Medidas generalizadas de correlación y dependencia estadística. Modelos gráficos probabilísticos.

Abog. CECILIA BONGIOVANNI
PROSECRETARIA

Dr. ADRIAN BONIVARDI
DECANO



Expte. n° FIQ-0955240-18
Resolución "C.D." n° 654/18

Departamento: de Matemática

Asignatura: **"Minería de Datos y Aprendizaje Automático"**

Carreras: Optativa para las carreras de Licenciatura en Matemática Aplicada, Ingeniería Industrial, Ingeniería Química, Ingeniería en Alimentos, Ingeniería en Materiales, Licenciatura en Materiales

Carácter: Curso de Verano Extracurricular

Carga Horaria: 90 horas

Bibliografía

- James, G., Witten, D., Hastie, T. y Tibshirani, R. An Introduction to Statistical Learning, with Applications in R. Springer, 2013.
- Bishop, C. Pattern Recognition and Machine Learning. Springer, 2006.
- Izenman, A.J., Modern Multivariate Statistical Techniques: Regression, Classification and Manifold Learning. Springer, 2008.
- DasGupta, A. Probability for Statistics and Machine Learning. Fundamentals and Advanced Topics. Springer, 2011

Abog. Cecilia BONGIOVANNI
PROSECRETARIA

Dr. Adrian BONIVARDI
DECANO



Expte. n° FIQ-0955240-18
Resolución "C.D." n° 654/18

Departamento: de Matemática

Asignatura: **"Minería de Datos y Aprendizaje Automático"**

Carreras: Optativa para las carreras de Licenciatura en Matemática Aplicada, Ingeniería Industrial, Ingeniería Química, Ingeniería en Alimentos, Ingeniería en Materiales, Licenciatura en Materiales

Carácter: Curso de Verano Extracurricular

Carga Horaria: 90 horas

Planificación

• Correlatividades

Licenciatura en Matemática Aplicada: tener aprobadas Estadística, Programación y Álgebra Lineal I.

Ingeniería Industrial: tener aprobadas Probabilidad y Estadística, Matemática C y Computación. Tener regularizada Administración de Operaciones.

Ingeniería Química, Ingeniería en Alimentos, Ingeniería en Materiales y Licenciatura en Materiales: tener aprobadas Probabilidad y Estadística, Matemática C e Informática.

• Periodo de dictado y número de alumnos estimado

Curso intensivo de verano.

Cupo mínimo: 4 alumnos. Cupo máximo: 16 alumnos

• Carga horaria semanal y total de la asignatura

22.5 horas semanales, durante cuatro semanas. 90 horas totales.

• Profesor Responsable

Diego Tomassi (Res. CD n° 047/16)

• Plantel docente:

Tomassi, Diego. Profesor Adjunto contratado, dedicación simple.

Gieco, María Antonella. Jefe de Trabajos Prácticos interino, dedicación simple.

• Tribunal Examinador:

Titulares: Dr. Diego Tomassi, Dra. Pamela Llop y Dra. Liliana Forzani

Suplentes: Dra. Sabrina Duarte, Dra. Andrea Bergesio y Dra. Ma. Antonella Gieco

• Requisitos para obtener la regularidad:

Asistir al 75% de las clases y obtener por lo menos el 50% en cada una de las dos evaluaciones de regularidad previstas, con al menos 58% de promedio entre dichas



evaluaciones. Las evaluaciones de regularidad se tomarán durante el cursado, serán escritas y de una hora de duración e incluirán conocimientos prácticos de los temas desarrollados de acuerdo al cronograma.

Régimen de promoción de la asignatura

Mediante examen final integrador:

Quienes no promocionen la asignatura mediante evaluación continua podrán rendir un examen final como alumnos regulares o libres según corresponda, en los turnos ordinarios habilitados por la facultad. El examen para alumnos regulares será escrito, teórico-práctico, de 3 horas de duración, y con uso de computadora. El examen se evaluará sobre un total de 100 puntos y la nota final será la que corresponda según Resolución CD(FIQ) 611/09. El examen de alumnos libres se registrará de acuerdo al artículo 43 del régimen de enseñanza.

Mediante evaluación continua:

Para promocionar la materia, los alumnos podrán realizar 2 evaluaciones parciales, las que incluirán la resolución de problemas que abarcarán todos los contenidos estudiados en la asignatura. Para acceder a la promoción es necesario que los alumnos mantengan su condición de regular. Aquel alumno que obtenga un promedio no menor a 58% en las evaluaciones parciales se considerará promocionado y estará eximido de rendir el examen final. No se propondrán instancias de recuperación a las evaluaciones parciales. La calificación final resultará del promedio de los dos parciales y se utilizará la escala de calificaciones dispuesta en la resolución CD 611/09.

Abog. Cecilia BONGIOVANNI
PROSECRETARIA

Dr. Adrian BONIVARDI
DECANO



Expte. n° FIQ-0955240-18
Resolución "C.D." n° 654/18

Departamento: de Matemática

Asignatura: **"Minería de Datos y Aprendizaje Automático"**

Carreras: Optativa para las carreras de Licenciatura en Matemática Aplicada, Ingeniería Industrial, Ingeniería Química, Ingeniería en Alimentos, Ingeniería en Materiales, Licenciatura en Materiales

Carácter: Curso de Verano Extracurricular

Carga Horaria: 90 horas

Cronograma de desarrollo de actividades - temas

Tipo de Actividad	Temas	Tiempo asignado (horas)	Docentes a cargo (Apellido)
Teoría 1	Introducción al aprendizaje automático y la minería de datos. Aprendizaje supervisado vs no supervisado. Regresión, clasificación y clustering. Criterios de optimalidad. Riesgo de Bayes. Riesgo empírico y riesgo estructural.	3	Tomassi
Coloquio	Overfitting. Regularización.	1.5	Tomassi
Teoría 2	Aprendizaje supervisado: modelos lineales. Regresión lineal múltiple.	3	Tomassi
Resolución de problemas	Implementación computacional	1.5	Gieco
Teoría 3	Aprendizaje supervisado: validación cruzada y métodos de remuestreo para selección y validación de modelos.	2	Tomassi
Coloquio	Suavizado no paramétrico	1	Tomassi
Resolución de problemas	Implementación computacional.	1.5	Gieco
Teoría 4	Aprendizaje supervisado: modelos lineales. Regresión logística. Análisis discriminante lineal. Curvas ROC.	3	Tomassi
Resolución de problemas	Implementación computacional	1.5	Tomassi
Teoría 5	Aprendizaje supervisado: modelos lineales en alta dimensión. Selección por pasos vs regularización. Regresión ridge. Lasso y lasso por grupos.	3	Tomassi
Resolución de problemas	Implementación computacional	1.5	Gieco
Teoría 6	Aprendizaje supervisado: Reducción de dimensiones: regresión parcial y correlaciones canónicas.	3	Tomassi
Resolución de problemas	Implementación computacional	1.5	Gieco
Teoría 7	Aprendizaje supervisado: métodos basados en árboles. Particionado recursivo. Árboles de regresión y de clasificación.	3	Tomassi
Resolución de problemas	Implementación computacional	1.5	Tomassi
Teoría 8	Aprendizaje supervisado: Bagging, boosting y randomforests.	3	Tomassi
Resolución de problemas	Implementación computacional	1.5	Tomassi
Teoría 9	Aprendizaje supervisado: métodos basados en núcleos. RKHS.	4.5	Tomassi
Teoría 10	Aprendizaje supervisado: métodos basados en núcleos. Criterio de margen máximo. Máquinas de vectores soporte para clasificación y regresión.	3	Tomassi
Resolución de problemas	Implementación computacional	1.5	Gieco
Teoría 11	Aprendizaje supervisado: métodos basados en núcleos. Máquinas de vectores soporte para detección de instancias anómalas.	2	Tomassi
Resolución de problemas	Implementación computacional	2.5	Gieco
Teoría 12	Aprendizaje supervisado: métodos basados en núcleos. Máquinas de vectores soporte para datos secuenciales.	2	Tomassi

6

1918-
2018En el año
del centenario
de la Reforma
Universitaria

Resolución de problemas	Implementación computacional	2.5	Tomassi
Teoría 13	Aprendizaje no supervisado: estimación de densidades.	3	Tomassi
Resolución de problemas	Implementación computacional	1.5	Gieco
Teoría 14	Aprendizaje no supervisado: clustering por k-medias y centroides. Medidas de calidad	3	Tomassi
Resolución de problemas	Implementación computacional	1.5	Gieco
Teoría 15	Aprendizaje supervisado: clustering por modelos de mezclas. Algoritmo EM.	3	Tomassi
Resolución de problemas	Implementación computacional	1.5	Tomassi
Teoría 16	Aprendizaje no supervisado: clustering jerárquico. Mapas auto-organizativos (redes SOM)	3	Tomassi
Resolución de problemas	Implementación computacional	1.5	Tomassi
Teoría 17	Aprendizaje no supervisado: reducción de dimensiones. PCA. Análisis de factores.	3	Tomassi
Resolución de problemas	Implementación computacional	1.5	Gieco
Teoría 18	Aprendizaje no supervisado: análisis de componentes independientes.	3	Tomassi
Resolución de problemas	Implementación computacional	1.5	Tomassi
Teoría 19	Aprendizaje no supervisado: reducción de dimensiones no lineal y aprendizaje de variedades. Clustering espectral.	3	Tomassi
Resolución de problemas	Implementación computacional	1.5	Gieco
Teoría 20	Aprendizaje no supervisado: medidas generalizadas de correlación. Inferencia de modelos gráficos.	4.5	Tomassi

Abog. Cecilia BONGIOVANNI
PROSECRETARIA

Dr. Adrian BONIVARDI
DECANO