



Expte. N° 67.474.-

Santa Fe, 28 de marzo de 2016.-

VISTO el expediente de referencia mediante el cual el Directora de la Carrera Ingeniería en Informática, Ing.Horacio LOYARTE, propone la incorporación de la asignatura optativa “Minería de Datos y Aprendizaje Automático”, y

CONSIDERANDO:

QUE el peticionante presenta una propuesta de programa y planificación de la asignatura;

QUE se cuenta con la venia de Secretaría Académica y de la Comisión de Seguimiento Académico de la carrera;

POR ELLO, y teniendo en cuenta despacho de la Comisión de Enseñanza,

EL CONSEJO DIRECTIVO
De la Facultad de Ingeniería y Ciencias Hídricas
Resuelve:

ARTÍCULO 1º.- Incorporar la asignatura “Minería de Datos y Aprendizaje Automático”, la que se dicta para la carrera Licenciatura en Matemática Aplicada en la Facultad de Ingeniería Química, como Asignatura Optativa, para la carrera Ingeniería en Informática, a cargo del Dr. Diego TOMASSI, cuya planificación forma parte, como anexo, de la presente Resolución.

ARTÍCULO 2º.- Inscribase, comuníquese, dese a publicidad. Tome nota Secretaría Académica, Secretaría de Coordinación, Director de la carrera de Ingeniería en Informática, Director del Departamento Informática, Departamento Alumnado y Bedelía. Notifíquese al Dr. TOMASSI. Cumplido, archívese.-

RESOLUCIÓN CD N° 052/16



Expte. N° 67.474.-

ANEXO Resol CD N° 052/16

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.

Duración del curso: 15 semanas (4 horas semanales)

Conocimientos Requeridos: Fundamentos de Programación – Álgebra Lineal

Evaluación: dos exámenes parciales escritos – Trabajo Final

Cupo: 20 alumnos

Universidad Nacional del Litoral

Facultad de Ingeniería y
Ciencias Hídricas

Secretaría de Consejo Directivo

Ciudad Universitaria

C.C. 217

(3000) Santa Fe

Tel: (54) (0342) 4575 246 int. 213

Fax: (54) (0342) 4575 224

E-mail: consejo@fich.unl.edu.ar