



Universidad Nacional de San Luis  
Rectorado

"2016 - Año del Bicentenario de la Declaración de la Independencia Nacional"

**ES COPIA**  
OSCAR GUILLERMO SEGURA  
Director de Despacho  
UNSL

SAN LUIS, - 1 JUL. 2016

**VISTO:**

El Expediente EXP-USL: 6177/2016 mediante el cual se solicita la protocolización del Curso de Posgrado: **MINERÍA DE DATOS**; y

**CONSIDERANDO:**

Que el Curso de Posgrado se propone dictar en el ámbito de la Facultad de Ciencias Físico Matemáticas y Naturales del 27 de junio al 8 de julio de 2016, con un crédito horario de 60 horas presenciales y bajo la coordinación del Dr. Marcelo Luis **ERRECALDE**.

Que la Comisión Asesora de Posgrado de la Facultad de Ciencias Físico Matemáticas y Naturales recomienda aprobar el curso de referencia.

Que el Consejo de Posgrado de la Universidad Nacional de San Luis en su reunión del 7 de junio de 2016, analizó la propuesta y observa que el programa del curso, bibliografía, metodología de evaluación y docentes a cargo, constituyen una propuesta de formación de posgrado de calidad en su campo específico de estudio.

Que, por lo expuesto, el Consejo de Posgrado aprueba la propuesta como Curso de Posgrado, según lo establecido en Ordenanza CS N° 23/09.

Que corresponde su protocolización.

Por ello y en uso de sus atribuciones

**EL RECTOR DE LA UNIVERSIDAD NACIONAL DE SAN LUIS**

**RESUELVE:**

**ARTÍCULO 1°.-** Protocolizar el dictado del Curso de Posgrado: **MINERÍA DE DATOS**, en el ámbito de la Facultad de Ciencias Físico Matemáticas y Naturales del 27 de junio al 8 de julio de 2016, con un crédito horario de 60 horas presenciales.

**ARTÍCULO 2°.-** Protocolizar el cuerpo docente constituido por: Responsable: Dr. Marcelo Luis **ERRECALDE** (DNI N° 17.023.968), Colaboradores: Dra. Leticia **CAGNINA** (DNI N° 26.214.269), Dr. Edgardo **FERRETTI** (DNI N° 26.450.715), Auxiliar: Dr. Luis **ÁVILA** (DNI N° 31.342.776) todos de esta Casa de Altos Estudios.

*[Handwritten signature]*  
D. Jorge Raúl Colmán  
Vice Rector - UNSL  
c/c Rectorado RR N° 24/16

*[Handwritten signature]*  
Dra. Alicia Marcela Primitista  
Secretaria de Posgrado  
U.N.S.L.

Cpde RESOLUCIÓN R N° **880**



Universidad Nacional de San Luis  
Rectorado

"2016 - Año del Bicentenario de la Declaración de la Independencia Nacional"

**ES COPIA**  
OSCAR GUILLERMO SEGURA  
Director de Despacho  
UNSL

**ARTÍCULO 3°.-** Aprobar el programa del Curso de referencia, de acuerdo al **ANEXO** de la presente disposición.-

**ARTÍCULO 4°.-** Comuníquese, insértese en el Libro de Resoluciones, publíquese en el Digesto Electrónico de la UNSL y archívese.-

**RESOLUCIÓN R N° 880**  
mav

Dra. Alicia Marcela Printista  
Secretaria de Posgrado  
U.N.S.L.

Ing. Jorge Raúl Olguín  
Vicerrector - UNSL  
a/c Rectorado RR N° 840/16



Universidad Nacional de San Luis  
Rectorado

"2016 - Año del Bicentenario de la Declaración de la Independencia Nacional"

ES COPIA  
OSCAR GUILLERMO SEGURA  
Director de Despacho  
UNSL

## ANEXO

**DENOMINACIÓN DEL CURSO:** MINERÍA DE DATOS

**UNIDAD ACADÉMICA RESPONSABLE:** Facultad de Ciencias Físico Matemáticas y Naturales

**CATEGORIZACIÓN:** Perfeccionamiento

**RESPONSABLE:** Dr. Marcelo Luis ERRECALDE

**COLABORADORES:** Dra. Leticia CAGNINA, Dr. Edgardo FERRETTI,

**AUXILIAR:** Dr. Luis ÁVILA

**COORDINADOR:** Dr. Marcelo Luis ERRECALDE

**CRÉDITO HORARIO:** 60 horas

**MODALIDAD DE DICTADO:** Presencial

**FECHA DE DICTADO DEL CURSO:** 27 de junio al 8 de julio de 2016

**FECHA PREVISTA PARA ELEVAR LA NÓMINA DE ALUMNOS APROBADOS:** 15 de agosto de 2016

**DESTINATARIOS:** Egresados con título de grado universitario de 4 años o más en Cs. Informáticas y en disciplinas afines a la temática del curso.

**LUGAR DE DICTADO:** Departamento de Informática – San Luis.

**CUPO:** 20 personas.

**FUNDAMENTACIÓN:** A partir de la popularización de Internet y el crecimiento del número de bases de datos accesibles, se ha producido en los últimos años un gran crecimiento en los volúmenes y disponibilidad de datos alrededor del mundo. Se ha estimado que la cantidad de datos almacenados en las bases de datos del mundo se duplica cada veinte meses. Este crecimiento cuantitativo de los datos no se ha reflejado en un crecimiento cualitativo de la información disponible. Existe un "gap" creciente entre la generación de los datos y nuestro entendimiento de los mismos. Una alternativa para solucionar este problema que ha recibido un interés creciente es el área de Minería de datos. La Minería de datos (en inglés *Data Mining* y de ahora en más MD) es el proceso de descubrir patrones en los datos. Este proceso debe ser automático (o más usualmente) semi-automático. Los patrones descubiertos deben ser significativos en el sentido que deben conducir a alguna ventaja (generalmente económica).

### OBJETIVOS:

Introducir al alumno en los principales conceptos vinculados a la Minería de datos. Al finalizar el curso los alumnos deberán ser capaces de identificar las principales etapas del proceso KDD, tareas relevantes de la Minería de datos, modelos y formas de evaluación del conocimiento extraído. Se espera además que los mismos sean capaces de detectar situaciones propicias para abordar con técnicas de MD y aplicar estas técnicas en aplicaciones concretas del mundo real.

Cpde RESOLUCIÓN R N° 880

Ing. Jorge Raúl Olguin  
Vicepresidente - UNSL  
a/c Rectorado RR N°  
876/16

Dra. Alicia Marcela Priinista  
Secretaria de Posgrado  
U.N.S.L.



Universidad Nacional de San Luis  
Rectorado

"2016 - Año del Bicentenario de la Declaración de la Independencia Nacional"

**ES COPIA**  
OSCAR GUILLERMO SEGURA  
Director de Despacho  
UNSL

### CONTENIDOS MÍNIMOS:

Aspectos Generales. Definición de Minería de Datos (MD). Relevancia. El proceso de KDD. Tareas de la Minería de Datos. Aprendizaje de Conceptos. Clasificación, Análisis de Clusters, Regresión. Evaluación de modelos. Algoritmos de aprendizaje. Plataformas de Minería de datos. Aplicaciones.

### PROGRAMA:

#### Unidad 1. Aspectos Generales

Definición de Minería de Datos (MD). Origen. Relación con otras disciplinas. Áreas de aplicación. Aprendizaje de Máquina (AM). Etapas del proceso KDD. AM versus MD. MD en el contexto del proceso KDD. Principales tareas de la minería de datos. Clasificación, Aprendizaje de conceptos, Categorización, Regresión, Análisis de grupos (cluster analysis), Análisis de asociaciones (association analysis), Análisis de correlaciones.

#### Unidad 2. Aprendizaje Automático

Definición de Aprendizaje Automático (AA). Importancia y Aplicaciones del AA. Tarea de aprendizaje bien definida. Tipos de "feedback" para el aprendiz. Diseño de un sistema de aprendizaje. Aspectos de estudio relevantes en AA. Introducción a Weka.

#### Unidad 3. Aprendizaje Supervisado

Etapas del aprendizaje Supervisado. Aprendizaje de árboles de decisión. Algoritmo ID3. Medidas de entropía y ganancia de información. Sesgo inductivo en ID3. Métodos para evitar el sobre-entrenamiento.

#### Unidad 4. Máquinas de Vectores Soporte (SVM)

Historia de SVM. Clasificadores lineales separables y no separables. Funciones Núcleo en SVM. Software. Ejemplos de uso.

#### Unidad 5. Regresión

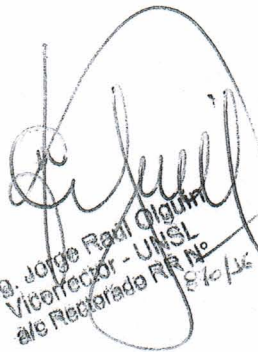
Formulación del problema. De los modelos lineales paramétricos de regresión a los procesos Gaussianos. Incerteza en los datos e información a priori. Inferencia Bayesiana con procesos Gaussianos.

#### Unidad 6. Análisis de Agrupamientos (Clustering)

Conceptos básicos y Algoritmos. Grupos significativos vs. Grupos útiles. Tipos de agrupamientos. Particional vs Jerárquico. Difuso vs. No Difuso. Tipos de clusters. Medidas de similitud. Validación de grupos. Algoritmos de agrupamiento.

#### Unidad 7. Aplicaciones, casos de estudio y recursos.

Aspectos prácticos del proceso de MD. Preparación de los datos. Evaluación, interpretación y visualización de patrones. Casos de estudio. MD en el Agro. Minería de Textos y de la Web. Herramientas de AA y MD. Weka. RapidMiner. scykit-learn. Bases de Datos públicas.

  
D. Jorge Raúl Olgún  
Vice Rector - UNSL  
del Rectorado R.º N.º  
8/10/16

  
Dra. Alicia Mercedes Primista  
Secretaría de Posgrado  
U.N.S.L.

Cpde RESOLUCIÓN R N° 880



Universidad Nacional de San Luis  
Rectorado

"2016 - Año del Bicentenario de la Declaración de la Independencia Nacional"

ES COPIA  
OSCAR GUILLERMO SEGURA  
Director de Despacho  
UNSL

### SISTEMA DE EVALUACIÓN:

La evaluación consistirá en la entrega de un informe correspondiente a un proyecto relacionado a las temáticas del curso. Este informe, podrá reportar los resultados de la revisión bibliográfica sobre un problema específico de AA o MD, o los resultados de una experiencia de implementación/uso de algunas de las técnicas/herramientas consideradas a lo largo del curso. La nota final del curso, resultará de la obtenida en la elaboración de este informe promediada con las notas obtenidas en las actividades evaluadas en las distintas unidades de este programa.

### BIBLIOGRAFÍA:

Obligatoria:

- [1] Ian H. Witten and Eibe Frank. Data mining: practical machine learning tools and techniques. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, third edition, 2011.
- [2] José Hernández Orallo, José Ramírez Quintana, and Cesar Ferri Ramírez. Introducción a la Minería de Datos. Departamento de Sistemas Informáticos y Computación, Universidad Politécnica de Valencia, 2004.
- [3] Tan, Steinbach y Kumar. "Introduction to Data Mining". Pearson. 2006.
- [4] C. E. Rasmussen. Gaussian Processes for Machine Learning. MIT Press. 2006
- [5] Tom Mitchell. "Machine Learning". McGraw-Hill Series in Computer Science, 1997.

Complementaria:

- [1] Shavlik, J. y Dietterich, Thomas. "Readings in Machine Learning". Morgan Kaufmann, 1990.
- [2] T. Hastie, R. Tibshirani, J. H. Friedman. "The Elements of Statistical Learning". Springer, 2002.
- [3] David J. Hand, Heikki Mannila, Padhraic Smyth. "Principles of Data Mining (Adaptive Computation and Machine Learning)". MIT Press, 2001.

Publicaciones y apuntes complementarios sobre los distintos tópicos del curso que estarán disponibles en la página del curso.

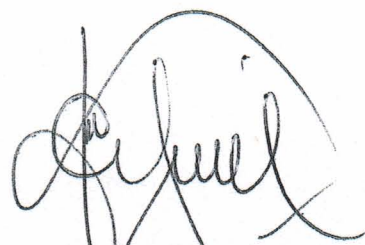
**ARANCEL:** Gratuito.

**COSTOS Y FUENTE DE FINANCIAMIENTO:** Se realiza como parte de las tareas docentes del profesor responsable y los colaboradores.

Cpde RESOLUCIÓN R N°  
may

880

  
Dra. Alicia Marcela Printista  
Secretaría de Posgrado  
U.N.S.L.

  
Ing. Jorge Raúl Olguín  
Vicerrector - UNSL  
a/c Rectorado RR N° 840/16