



Carrera de Posgrado: Doctorado en Ingeniería
Curso: TÓPICOS SELECTOS EN APRENDIZAJE MAQUINAL
Año Académico: 2020

1. Objetivos.

Que el alumno:

- Conozca los fundamentos teóricos de las técnicas más utilizadas y los avances recientes en el área del aprendizaje maquinal.
- Comprenda su significado a los efectos de la correcta implementación de los algoritmos.
- Identifique la utilidad de estas técnicas para su aplicación en problemas reales.
- Desarrolle habilidad para la lectura fluida y comprensiva de publicaciones científicas actuales sobre el tema.

2. Programa Analítico.

I- INTRODUCCIÓN.

Revisión de Probabilidad, Nociones de Teoría de la Información y Teoría de la Decisión. Clasificación estadística de patrones y regresión: aprendizaje supervisado paramétrico, no paramétrico y no supervisado. Minería de datos y agrupamiento de patrones.

II- ANÁLISIS ESTADÍSTICO DE DATOS.

Análisis de componentes principales, formulación probabilística. Análisis de componentes independientes: formulaciones alternativas, funciones objetivo y funciones de contraste. Modelos de mezclas no lineales y convolutivas. Métodos de proyección y reducción dimensional. Factorización de matrices nonegativas.

III- APRENDIZAJE BASADO EN REDES Y TÉCNICAS CLÁSICAS

Revisión de redes neuronales “clásicas”: Perceptrón simple y multicapa, redes con Funciones de Base Radial, Mapas Auto-organizativos. Naive Bayes, k -vecinos cercanos, Análisis Discriminante Lineal, Mezclas de gaussianas, k -medias. Aprendizaje por refuerzo. Ensamble de clasificadores, Bagging, Boosting. Aprendizaje profundo.

IV- APRENDIZAJE BASADO EN ÁRBOLES Y REGLAS DE DECISIÓN.

Métodos que no utilizan métricas. Generación de árboles de decisión: CART, ID3, C4.5. Reglas para separación, crecimiento y podado. Tratamiento de atributos con valores faltantes. Relación y equivalencias con redes neuronales.



V- APRENDIZAJE BASADO EN NÚCLEOS.

Construcción de núcleos y aprendizaje basado en núcleos. Teoría estadística del aprendizaje: clasificadores de riesgo empírico mínimo. Máquinas de soporte vectorial. Máquinas multi-clase.

VI- APRENDIZAJE DE DATOS SECUENCIALES

Revisión de redes neuronales dinámicas: redes de Hopfield, redes neuronales con retardos temporales, redes de Elman y Jordan. Modelos ocultos de Markov discretos y continuos. Algoritmos hacia adelante y hacia atrás. Algoritmo de Viterbi. Entrenamiento por maximización de la esperanza.

VII- VALIDACIÓN Y SELECCIÓN DE CARACTERÍSTICAS

Figuras de mérito en aplicaciones de clasificación y regresión. Capacidad de generalización y sobre-entrenamiento. Métodos de estimación del error: partición simple, validación cruzada, particiones múltiples, Bootstrap, 0.632-bootstrap. Análisis ROC. Métodos de selección de características.

VIII- APLICACIONES

Aplicaciones de aprendizaje maquina a problemas con datos reales.

3. Bibliografía.

- [1] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, MIT Press, 2012.
- [2] Zhi-Hua Zhou, *Ensemble Methods: Foundation and Algorithms*, CRC Press, 2012.
- [3] Ethem Alpaydin, *Introduction to Machine Learning*. MIT Press: Adaptive Computation and Machine Learning series, 2010.
- [4] Stephen Marsland, *Machine Learning: An Algorithmic Perspective*. Chapman & Hall/CRC: Machine Learning & Pattern Recognition Series, 2009.
- [5] Y. Bengio, *Learning Deep Architectures for AI*, Now Publishers, Canadá, 2009.
- [6] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, Academic Press, Elsevier, 2009.
- [7] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, Cambridge MA, 1998.
- [8] Christopher M. Bishop, *Pattern Recognition and Machine Learning*. Springer: Information Science and Statistics, 2006.
- [9] V. Cherkassky, F. Mulier, *Learning from Data: Concepts, Theory and Methods*. Wiley-International Science, 1998.
- [10] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing*. John Wiley & Sons, 2002.
- [11] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification*. Wiley-Interscience, 2001.
- [12] X. D. Huang, Y. Ariki, M. A. Jack, *Hidden Markov models for speech recognition*. Edinburgh University Press, 1990.
- [13] Hyvärinen, J. Karhunen, E. Oja, *Independent Component Analysis*. John Wiley & Sons, 2001.
- [14] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.
- [15] B. D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge University Press, 1999.
- [16] J. R. Quinlan, *C4.5: Programs for Machine Learning*. 1993.
- [17] R. P. N. Rao, B. A. Olshausen, M. S. Lewicki (Eds.), *Probabilistic Models of the Brain: Perception and Neural Function*. MIT Press, 2002.
- [18] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, 2000.
- [19] Yoshua Bengio, "Learning Deep Architectures for AI", Foundations and Trends in Machine Learning archive, Vol. 2 Issue 1, pp 1-127, 2009.

Universidad Nacional del Litoral Ciudad Universitaria
Facultad de Ingeniería y C.C. 217
Ciencias Hídricas Ruta Nacional Nº 168 – Km. 472,4
(3000) Santa Fe
Tel: (54) (0342) 4575-234/244 – int. 103
Dirección de Posgrado Fax: (54) (0342) 4575 224
E-mail: posgrado@fich.unl.edu.ar



4. Docentes.

4.1 Docente responsable: Dr. Leandro Vignolo.

4.2 Docente(s) corresponsable(s): Dr. Hugo L. Rufiner, Dr. Diego H. Milone.

4.3 Docente(s) colaborador(es): Dra. Georgina Stegmayer, Dr. Gastón Schlotthauer,
Dr. Leandro Di Persia, Dr. C. Martinez.

5. Conocimientos previos requeridos.

Nociones de inteligencia computacional y optimización. Álgebra lineal. Fundamentos de sistemas y señales. Probabilidad y estadística. Lenguajes de programación.

6. Carga horaria.

6.1 Teoría: 60 horas (de dictado online y/o clases grabadas con coloquio online)

6.2 Coloquio y/o Práctica: 30 horas.

6.3 Total: 90 horas.

7. Forma de evaluación.

7.1 Cantidad y tipo de exámenes parciales:

El alumno deberá presentar el 100% de las guías de trabajos prácticos y laboratorio resueltas y defenderlas en un coloquio con los docentes en las fechas previstas en el cronograma. La presentación deberá realizarse a través del correspondiente informe, acompañado de los códigos desarrollados en formato digital. Se recomendará la discusión y colaboración entre los alumnos del curso para la resolución de los distintos ejercicios planteados. Sin embargo, por ser esta una instancia de evaluación, el material entregado deberá ser de autoría individual.

7.2 Tipo y duración del examen final:

El examen final consistirá en la defensa oral de un trabajo, en el cual el alumno deberá realizar un análisis crítico y eventual reproducción de un artículo publicado en una revista científica internacional de la especialidad, oportunamente asignado por los profesores. Deberá entregarse el correspondiente informe escrito, completando las deducciones teóricas, comentando aspectos omitidos de la implementación, puntos positivos, negativos y posibles mejoras. Durante la defensa oral se evaluará también el dominio de los conceptos desarrollados en la asignatura y se requerirán las justificaciones y desarrollos teóricos correspondientes en el pizarrón.

Universidad Nacional del Litoral
Facultad de Ingeniería y
Ciencias Hídricas

Dirección de Posgrado

Ciudad Universitaria
C.C. 217

Ruta Nacional Nº 168 - Km. 472,4
(3000) Santa Fe

Tel: (54) (0342) 4575-234/244 - int. 103

Fax: (54) (0342) 4575 224

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



8. Fecha tentativa de inicio del dictado y duración del Curso (en semanas).

Fecha de inicio: 04/09/2020 (todos los viernes a las 10:30hs)

Duración: 15 semanas, (clases teóricas virtuales, pre grabadas u online, de 4 horas)

9. Cupo de alumnos

Diez alumnos (10)

10. Lugar, día(s) y horario(s) de dictado y requerimientos.

10.1 Horario tentativo para las clases online: viernes de 10:30 a 14:30 hs.

10.2 Horario para coloquios y discusión de clases grabadas: viernes de 10:30 a 12:00 hs.

- En el caso de las clases grabadas, los alumnos deben ver los videos antes de la clase de coloquio correspondiente a ese tema. La clase consistirá, en ese caso, en respuesta a consultas y discusión sobre el contenido de los videos de la clase.

10.3 Conexión a internet, acceso a Google Meet. PCs: una por alumno con Matlab y/o Octave instalado, compiladores de C/C++ Standard y herramientas para Phyton. Otros: acceso a publicaciones periódicas.

10.4 Página Web del curso: <http://tsam-fich.wikidot.com/>

11. Adjuntar breve CV de los docentes.

Se adjunta CV de: Dr. Hugo Leonardo Rufiner, Dr. Diego H. Milone, Dr. Leandro Di Persia, Dr. Gastón Schlotthauer, Dr. César Martinez, Dr. Leandro Vignolo, Dra. Georgina Stegmayer

Santa Fé, 07/08/20

Firma del docente responsable.

Universidad Nacional del Litoral	Ciudad Universitaria
Facultad de Ingeniería y	C.C. 217
Ciencias Hídricas	Ruta Nacional Nº 168 - Km. 472,4
	(3000) Santa Fe
Dirección de Posgrado	Tel: (54) (0342) 4575-234/244 - int. 103
	Fax: (54) (0342) 4575 224
	E-mail: posgrado@fich.unl.edu.ar



CRONOGRAMA TENTATIVO DE CLASES TSAM (VIERNES de 10:30 a 14:30 hs)

Semana N° 1 (04/09/20)

Tema/s: Clasificación estadística de patrones y regresión: aprendizaje supervisado paramétrico, no paramétrico y no supervisado. (clase grabada y coloquio)

Docente: C. Martínez.

Tema/s: Presentación del plan del curso. Introducción general. **Presentación TP N°1.**

Docente: H. L. Rufiner.

Semana N° 2 (11/09/20)

Tema/s: Análisis de componentes principales, formulación probabilística. Métodos de proyección y reducción dimensional.

Docente: L. Di Persia.

Semana N° 3 (18/09/20)

Tema/s: Minería de datos y agrupamiento de patrones. Revisión de redes neuronales: perceptrón simple y multicapa, redes con funciones de base radial, mapas auto-organizativos. Naive Bayes, k -vecinos más cercanos, análisis discriminante lineal, mezclas de gaussianas, k medias. Ensamble de clasificadores, bagging, boosting. (clase grabada y coloquio)

Docente: D. Milone.

Semana N° 4 (25/09/20)

Tema/s: Análisis de componentes independientes: formulaciones alternativas, funciones objetivo y funciones de contraste. Modelos de mezclas no lineales y convolutivas. Factorización de matrices nonegativas.

Docente: L. Di Persia.

Semana N° 5 (02/10/20)

Tema/s: Modelos ocultos de Markov discretos y continuos. Algoritmos hacia adelante y hacia atrás. Algoritmo de Viterbi. Entrenamiento por maximización de la esperanza. (clase grabada y coloquio)

Docente: D. Milone.

Semana N° 6 (09/10/20)

Tema/s: Métodos que no utilizan métricas. Generación de árboles de decisión: CART, ID3, C4.5. Reglas para separación, crecimiento y podado. Tratamiento de atributos con valores faltantes. Relación y equivalencias con redes neuronales. (clase grabada y coloquio)

Docente: H. L. Rufiner.

Semana N° 7 (16/10/20)

Tema/s: Revisión de redes neuronales dinámicas: redes de Hopfield, redes neuronales con retardos temporales, redes de Elman y Jordan. Otras redes dinámicas. (clase grabada y coloquio)

Presentación TP N°2. Propuesta de artículos para trabajo final (durante la semana).

Docente: H. L. Rufiner.

Universidad Nacional del Litoral
Facultad de Ingeniería y
Ciencias Hídricas

Dirección de Posgrado

Ciudad Universitaria
C.C. 217

Ruta Nacional N° 168 - Km. 472,4
(3000) Santa Fe

Tel: (54) (0342) 4575-234/244 - int. 103

Fax: (54) (0342) 4575 224

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



Semana N° 8 (23/10/20)

Tema/s: Introducción al Aprendizaje Profundo: Maquinas de Boltzmann (RBM) Restringidas. Redes de Creencias profundas (DBN). Autocodificadores apilados profundos (SAE). Algoritmos de aprendizaje. (clase grabada y coloquio)

Entrega TPN°1 resuelto (informe en formato digital).

Elección de artículos para trabajo final.

Docente: H. L. Rufiner.

Semana N° 9 (30/10/20)

Tema/s: Construcción de núcleos y aprendizaje basado en núcleos. Teoría estadística del aprendizaje: clasificadores de riesgo empírico mínimo. Máquinas de soporte vectorial. Máquinas multiclase. Máquinas de núcleos malos. (clase online)

Docente: G. Schlotthauer.

Semana N° 10 (06/11/20)

Tema/s: Aprendizaje por refuerzo (clase online).

Docente: G. Stegmayer.

Selección de características (clase grabada y coloquio).

Docente: L. Vignolo.

Semana N° 11 (13/11/20)

Tema/s: Figuras de mérito en aplicaciones de clasificación y regresión. Capacidad de generalización y sobre-entrenamiento. Métodos de estimación del error: partición simple, validación cruzada, particiones múltiples, bootstrap, 0.632-bootstrap. Análisis ROC. (clase grabada y coloquio)

Docente: L. Vignolo.

Semana N° 12 (20/11/20)

Tema/s: Aplicaciones de aprendizaje maquina a problemas con datos reales.

Docentes: G. Stegmayer.

Semana N° 13 (27/11/20)

Tema/s: Aplicaciones de aprendizaje maquina a problemas con datos reales.

Docentes: L. Di Persia.

Semana N° 14 (04/12/20)

Tema/s: **Entrega de informe del TP N°2 resuelto.**

Semana N° 15 (11/12/20)

Tema/s: **Defensa de trabajo final (entrega de informe y formulario de revisor del trabajo).**

Docente: Todos.

LISTADO TENTATIVO DE TRABAJOS PRÁCTICOS TSAM

TP N°1: Desarrollo e implementación de algoritmos selectos de aprendizaje maquina

TP N°2: Resolución de problemas de clasificación y regresión con datos reales.

Universidad Nacional del Litoral	Ciudad Universitaria
Facultad de Ingeniería y	C.C. 217
Ciencias Hídricas	Ruta Nacional N° 168 - Km. 472,4
	(3000) Santa Fe
Dirección de Posgrado	Tel: (54) (0342) 4575-234/244 - int. 103
	Fax: (54) (0342) 4575 224
	E-mail: posgrado@fich.unl.edu.ar