

Formulario para la presentación de Cursos de Posgrado/Doctorado – Res. CD2819/18 - ANEXO 1**Información académica**

Año de presentación (*)

2021

1-a-

Departamento docente que inicia el trámite:
Computación
Nombre del curso:
Reglas de asociación y patrones
Nombre, Cargo y Título del docente responsable:
Mg. Cecilia Ana Ruz, Profesora adjunta, Departamento de Computación
En caso de dictarse en paralelo con una materia de grado, nombre de la misma:
Reglas de asociación y patrones secuenciales
Nombre y Título de los docentes que colaboran con el dictado del curso (*) (*):
-
Fecha propuesta para el primer dictado luego de la aprobación:
Segundo cuatrimestre/ cuarto bimestre

Duración:

Duración total en horas	32
Duración en semanas	8

Distribución carga horaria:

Número de horas de clases teóricas	12
Número de horas de clases de problemas	10
Número de horas de trabajos de laboratorio	10
Número de horas de trabajo de campo	
Número de horas de seminarios	

Forma de evaluación:

Presentación de trabajos grupales y examen final

Lugar propuesto para el dictado (departamento, laboratorio, campo, etc.):

Departamento / Zoom

Puntaje propuesto para la carrera de doctorado:

1.5 (uno y medio)

Número de alumnos:

Mínimo: 5

Máximo:20

Audiencia a quien está dirigido el curso:

Cualquier persona que esté interesada en las técnicas de minería de datos, en particular las técnicas no supervisadas.

Necesidades materiales del curso:

Ninguna

1-b-

Programa analítico del curso con Bibliografía (puede adjuntarse en hojas separadas):

La materia estará orientada a conseguir que el alumno sea capaz de:

- a) Explicar el funcionamiento de los algoritmos de reglas de asociación y patrones secuenciales
- b) Analizar las mejoras / variantes propuestas a los algoritmos de reglas de asociación / patrones secuenciales.
- c) Resolver problemas de data Mining aplicando reglas de asociación/ patrones secuenciales

Unidad 1: Introducción

Definición de data Mining. Su integración con el proceso de Descubrimiento del conocimiento. Métodos supervisados y no supervisados. Principales técnicas. Problemas más comunes que se resuelven aplicando de DM

Unidad 2: Reglas de Asociación

Reglas de asociación: conceptos básicos: itemset, itemset frecuente, itemset máximo e itemset cerrado. Algoritmo básico de a priori. Medidas de "calidad" de las reglas de asociación. Reglas de asociación multilevel. Reglas generalizadas. Reglas de asociación temporales.

Unidad 3: Análisis de Secuencias

Definición de análisis de secuencias. Algoritmo básico. Análisis generalizado de secuencias. Restricciones temporales al análisis de secuencias. Análisis de trayectorias.

Bibliografía

- Pang-Ning tan, Michael Steinbach, Vipin Kumar. Introduction to Data Mining –2^{da} edición – Adisson Wesley
- Jiawei Han, Micheline Kamber. [Data Mining, Concepts and Techniques- 2^{da} edición- The Morgan Kaufmann Series in Data Management Systems](#)
- Anand Rajaraman, Jeffrey Ullman. Mining of masive datasets, <http://www.mmds.org/>
- Pang-Ning Tan*, Vipin Kumar, Jaideep Srivastava. Selecting the right objective measure for association analysis. Information Systems, 29(4), 293-313 (2004).
- Bing Liu, Wynne Hsu, and Yiming Ma. 1999. Mining association rules with multiple minimum supports. In Proceedings of the fifth ACM SIGKDD international conference on

Knowledge discovery and data mining (KDD '99). Association for Computing Machinery, New York, NY, USA, 337–341. DOI:<https://doi.org/10.1145/312129.312274>

- Ramakrishnan Srikant and Rakesh Agrawal. 1996. Mining quantitative association rules in large relational tables. In *Proceedings of the 1996 ACM SIGMOD international conference on Management of data (SIGMOD '96)*. Association for Computing Machinery, New York, NY, USA, 1–12. DOI:<https://doi.org/10.1145/233269.233311>
- Kaushik, M., Sharma, R., Peious, S.A. et al. A Systematic Assessment of Numerical Association Rule Mining Methods. *SN COMPUT. SCI.* 2, 348 (2021). <https://doi.org/10.1007/s42979-021-00725-2>
- J. Ale, G. Rossi.: The Itemset's Lifespan Approach to Discovering General Temporal Association Rules. In *Proc. ACM 2nd Workshop on Temporal Data Mining, held in KDD-2002*, July 2002
- Brin, S. –Motwani, R-Ullman, J.- Tsur, S. : Dynamic Itemset Counting and Implications Rules for Market Basket Data. *Proc. ACM SIGMOD* :pages 225-264, 1997.
- Chang-Hung Lee, Cheng-Ru Lin, Ming-Syan Chen. On Mining General Temporal Association Rules in a Publication Database. In *Proc. of the 2001 IEEE International Conference on Data Mining (ICDM'01)*, pages 337-344, 2001.
- Banu Özden, Sridhar Ramaswamy, , Avi Silberschatz. Cyclic Association Rules. In *Proc. of the 14th International Conference on Data Engineering (ICDE 98)*, pages 412-421, 1998.
- Sridhar Ramaswamy, Sameer Mahajan, Avi Silberschatz. On the Discovery of Interesting Patterns in Association Rules. In *Proc. of the 24th VLDB Conference*, pages 368-379, 1998.
- Tzung-Pei Hong, Guo-Cheng Lan, Ja-Hwung Su, Pei-Shan Wu, Shyue-Liang Wang, Discovery of temporal association rules with hierarchical granular framework, *Applied Computing and Informatics*, Volume 12, Issue 2, 2016, (<https://doi.org/10.1016/j.aci.2016.01.003>.)
- Balderas, Berzal, Cubero, Eisman & Marín “Discovering Hidden Association Rules ” KDD'2005, Chicago, Illinois, USA
- Yang, X. Y., Liu, Z., & Fu, Y. (2010, June). MapReduce as a programming model for association rules algorithm on Hadoop. In *Information Sciences and Interaction Sciences (ICIS), 2010 3rd International Conference on* (pp. 99-102). IEEE.
- Jeffrey Dean and Sanjay Ghemawat. 2008. MapReduce: simplified data processing on large clusters. *Commun. ACM* 51, 1 (January 2008), 107–113. DOI:<https://doi.org/10.1145/1327452.1327492>
- Q. Zhao and S. S. Bhowmick, “Sequential Pattern Mining: A Survey”, Technical Report, CAIS, Nanyang Technological University, Singapore, No. 2003118, 2003.
- Wensheng Gan, Jerry Chun-Wei Lin, Philippe Fournier-Viger, Han-Chieh Chao, and Philip S. Yu. 2019. A Survey of Parallel Sequential Pattern Mining. *ACM Trans. Knowl. Discov. Data* 13, 3, Article 25 (July 2019), 34 pages. DOI:<https://doi.org/10.1145/3314107>
- Yu Zheng. 2015. Trajectory Data Mining: An Overview. *ACM Trans. Intell. Syst. Technol.* 6, 3, Article 29 (May 2015), 41 pages. <http://research.microsoft.com/apps/pubs/?id=241453>

Ejercicios aplicando el algoritmo a priori en sus diversas formas
Tp integrador con aplicación de reglas de asociación a un problema de Open Data.
Para el desarrollo de las actividades practicas se puede utilizar tanto R como Phytton

(*) Todos los cursos tendrán una validez de 5 años

(*)(*) Las actualizaciones de los docentes colaboradores son informados por la Dirección departamental al inicio de cada dictado del curso

Firma Subcomisión
Doctorado

Firma del docente
responsable



E-mail y teléfono del docente responsable

Cecilia Ana Ruz, tel 11-4-147-5823, cruz@dc.uba.ar

Formulario para la presentación de Cursos de Posgrado/Doctorado - Res. CD2819/18 - ANEXO 2

Solicitud de Financiación

Año de presentación (*)

Departamento docente que inicia el trámite:

Nombre del curso:

Nombre y Título del docente responsable:

Costo propuesto del curso por alumno (*):

Justificación del monto propuesto:

(*) Las excepciones aplicables para cada alumno serán consistentes con la reglamentación del Consejo Directivo que regula los aranceles y excepciones (Res. CD 484/13). El docente responsable del curso solicitará las excepciones por nota al consejo directivo a través de Mesa de Entradas.