

Minería de Datos

Objetivo

Conocer de manera general las técnicas y enfoques del proceso general de Minería de Datos. Se conocen los fundamentos y conceptos necesarios de cada una de las etapas del proceso. Se explora el uso de fuentes de datos para análisis y toma de decisiones resultantes de tareas de clasificación, predicción o agrupamiento.

Contenido

- a. Introducción a minería de datos
 - a. Objetivos de la minería de datos
 - b. Problemas aptos para minería de datos
 - c. Aplicaciones comerciales
 - d. Aplicaciones no-comerciales
 - e. Técnicas generales para análisis de los datos (predicción, clasificación, clustering).
- b. Selección de Fuentes de Datos y Calidad de Datos
 - a. Datos esperados para minería de datos
 - b. Calidad de los datos
 - c. Fuentes posibles de datos.
- c. Preprocesamiento y Preparación de Datos
 - a. Operaciones sobre los datos
 - b. Problemas en el manejo de datos reales
 - c. Selección de variables
 - d. Muestreo, selección de registros
 - e. Análisis de correlación
 - f. Creación de nuevas variables, agregación de variables
- d. Técnicas de Análisis
 - 1.1. Principales técnicas para el análisis de datos.
 - 1.2. Aplicación de las técnicas de análisis de datos
 - 1.3. Técnicas para la identificación de características, tendencias y relaciones en los datos
 - 1.4. Visualización
 - 1.5. Técnicas estadísticas (correlación, análisis factorial)
- e. Creación de modelos de datos: clasificación y predicción
 - a. Principales técnicas para clasificación y predicción de datos
 - b. Aplicación de las técnicas de clasificación y predicción de datos
 - c. Inducción de reglas: C4.5.
 - d. Redes neuronales.
 - e. Técnicas estadísticas: regresión.
- f. Creación de modelos de datos: clustering
 - a. Conceptos fundamentales de clustering
 - b. Principales técnicas para el clustering
 - c. Aplicación de técnicas de clustering

- d. Redes neuronales: Kohonen SOM
- e. Técnicas estadísticas: K-means
- f. Clustering difuso: Fuzzy c-Means
- g. Evaluación de modelos
- a. Conceptos fundamentales
- b. Técnicas de evaluación de modelos
- c. Aplicación de las técnicas de evaluación de modelos

Bibliografía

1. Data Mining: Concepts and Techniques, 2nd ed. Jiawei Han and Micheline Kamber Morgan Kaufmann, 2006
2. Data Mining: Practical Machine Learning Tools and Techniques I.H. Witten and E. Frank Morgan Kaufmann, 1999
3. Machine Learning Tom M. Mitchell Prentice Hall, 2003
4. Inteligencia Artificial Russell y Norvig Prentice Hall, 2009
5. Pang-Ning Tan, Michael Steinbach and Vipin Kumar. Introduction to Data Mining. Addison-Wesley. 2006. ISBN: 0321321367.
6. Anil K. Jain, Richard C. Dubes. Algorithms for Clustering Data. Prentice Hall. 1988. ISBN: 013022278X.
7. T. Hastie, R. Tibshirani and J. Friedman. Elements of Statistical Learning: Data Mining, Inference and Prediction. Springer-Verlag. 2001. ISBN: 0387952845.