

LECTURE NOTES

COMP8014

Knowledge Data Discovery

Data Mining Tools

LEARNING OUTCOMES

LO 2: Mahasiswa diharapkan mampu menggunakan teknik eksplorasi data dan preprocessing

OUTLINE MATERI :

1. Data Mining Tools: Gambaran Umum
2. Membandingkan Data Mining Tools
3. RapidMiner
4. Weka

ISI MATERI

Acknowledgements

These slides are based on the paper of RapidMiner Studio Manual v6, WEKA Manual for Version 3-7-12, and book of Kotu and Deshpande “Predictive Analytics and Data Mining, Concepts and Practice with RapidMiner”

Gambaran Umum?

- Penemuan pengetahuan di database adalah bidang yang berkembang pesat, yang pengembangannya didorong oleh minat penelitian yang kuat serta kebutuhan praktis, sosial, dan ekonomis yang mendesak.
- Beberapa tahun terakhir, alat penemuan pengetahuan telah digunakan terutama di lingkungan penelitian, produk perangkat lunak yang canggih kini berkembang dengan cepat.
- Dalam kursus ini, kami memberikan gambaran umum alat data mining dan membahas alat sumber terbuka untuk mempelajari metode pengumpulan data dasar, termasuk data preprocessing dan metode peraturan asosiasi:
 - RapidMiner
 - WEKA

Perbandingan Data Mining Tools?

- Lihat Table pada halaman 7.
- Most popular open source tools:
 - RapidMiner, 44.2% share (39.2% in 2013)
 - R, 38.5% (37.4% in 2013)
 - Python, 19.5% (13.3% in 2013)
 - Weka, 17.0% (14.3% in 2013)
 - KNIME, 15.0% (5.9% in 2013)

- Most popular commercial tools:
 - SAS Enterprise Miner
 - MATLAB
 - IBM SPSS Modeler

Rapid Miner?

Lihat penjelasan singkat di Slide atau mencari literature lainnya.

Weka?

Lihat penjelasan singkat di Slide atau mencari literature lainnya.

SIMPULAN

1. Alat DM Populer, Open Source, Pemrograman / Statistik telah disajikan.
2. Dua alat populer di DM, yaitu Rapid Miner dan WEKA, diperkenalkan secara singkat

DAFTAR PUSTAKA

1. Han, J., Kamber, M., & Pei, Y. (2006). "Data Mining: Concepts and Technique". Edisi 3. Morgan Kaufman. San Francisco
2. Tan, P.N., Steinbach, M., & Kumar, V. (2006). "Introduction to Data Mining". Addison-Wesley. Michigan
3. Witten, I. H., & Frank, E. (2005). "Data Mining : Practical Machine Learning Tools and Techniques". Second edition. Morgan Kaufmann. San Francisco