

SGN-43006 Knowledge Mining and Big Data (5 cr)

Examination 2.12.2014 / Ari Visa

A: No Calculator

The questions are returned.

Please, do not forget the Kaiku-feedback!

1. Define terms: Data Mart, On Line Analytical Processing, and Big Data	6p
2. What is HADOOP? What is the relation between HADOOP and MapReduce algorithms? Are there other techniques to the same purpose? If answer is yes then give an example.	6p
3. How to define association rule mining? List the types of associations! How to look after associations?	6p
4. What is the difference between classification and prediction? How do you define an outlier? What are the main approaches to detect an outlier?	6p
5. How do you define cluster analysis? What are the main players that define the quality of clustering. What are the main categories of clustering techniques?	6p