Data Mining in Software Engineering

Course title:

Data Mining in Software Engineering

Course timing:	May 25, 2017
Mode of study:	Lectures 4hrs, Practice 6hrs, Total 10 hrs.
Study materials:	AnnouncedMay 5th on http://edu.susu.ru

Prerequisites for entering the course:

Basics in algorithms, mathematics, statistics, familiarity with some programming language.

Course summary:

Course devoted to methods, algorithms and software to discover hidden knowledge from data involved in Software Engineering. Course is lectured by *Assoc. Prof. Mikhail Zymbler* (South Ural State University, Chelyabinsk, Russia). Mikhail Zymbler is also Head of Data Mining Department of SUSU's Supercomputer Simulation Laboratory. His research interests include parallel algorithms for data mining, parallel database systems.

Course outline:

#	Title	Duration	Summary
Lectures			
1	Introduction to data mining	**1 hour**	Big Data phenomenon. Notion of Data Mining. Data Mining as a process. Applications of Data mining in Software Engineering.
2	Mining frequent patterns	**1 hour**	Market basket problem, support, confidence, association rules. Generating association rules from frequent itemsets.
3	Classification	**1 hour**	Learning step, classification step, training set, test set, classifier accuracy. Decision trees. k-Nearest-Neighbor Classification.
4	Clustering	**1 hour**	k-Means clustering. Agglomerative and divisive hierarchical clustering.
Practice			
1	KNIME basics	**1hours**	Basics of KNIME, open-source stand-alone package for data mining.
2	Data mining in Software Engineering	**5hours**	Solving typical data mining problems on given datasets from scope of Software Engineering using KNIMEpackage.

Reading:

Jiawei Han, Micheline Kamber and Jian Pei. Data Mining. Concepts and Techniques. 3rd ed.Morgan Kaufmann Publishers, 2012.

Software:

KNIME Analytics Platform, http://www.knime.org/

Course timetable:

Date	Time	Classes
May 25th	9:00-12:00	Lectures
	13:00-17:00	Practice