

SoSe 2014: M-TANI: Big Data Analytics

Lecture 1 – 23/04/2014

Dr. Nikolaos Korfiatis

Welcome to this course

- Course title: M-TANI: Big Data Analytics
- Instructor
 - Dr. Nikolaos Korfiatis
 - Todor Ivanov, Sead Izberovic (Assistants)
- Date und Time:
 - Wednesday(18.04-11.07), 16 to 18 o'clock, Room **612** (Rob.Mayer-Str. 10 / Gräfstr. 38 – 6th Floor)
 - Exam method: TBD

Course Lecturers

- Dr. Nikolaos Korfiatis

- PhD in Information Systems and Economics, Copenhagen Business School (CBS) Denmark
- Dipl. Ing. and MSc in Information Retrieval and HCI, Royal Institute of Technology / KTH Stockholm Sweden
- BSc Econ, Athens University of Economics
- More on: <http://www.bigdata.uni-frankfurt.de/~nkorf>



Contributors

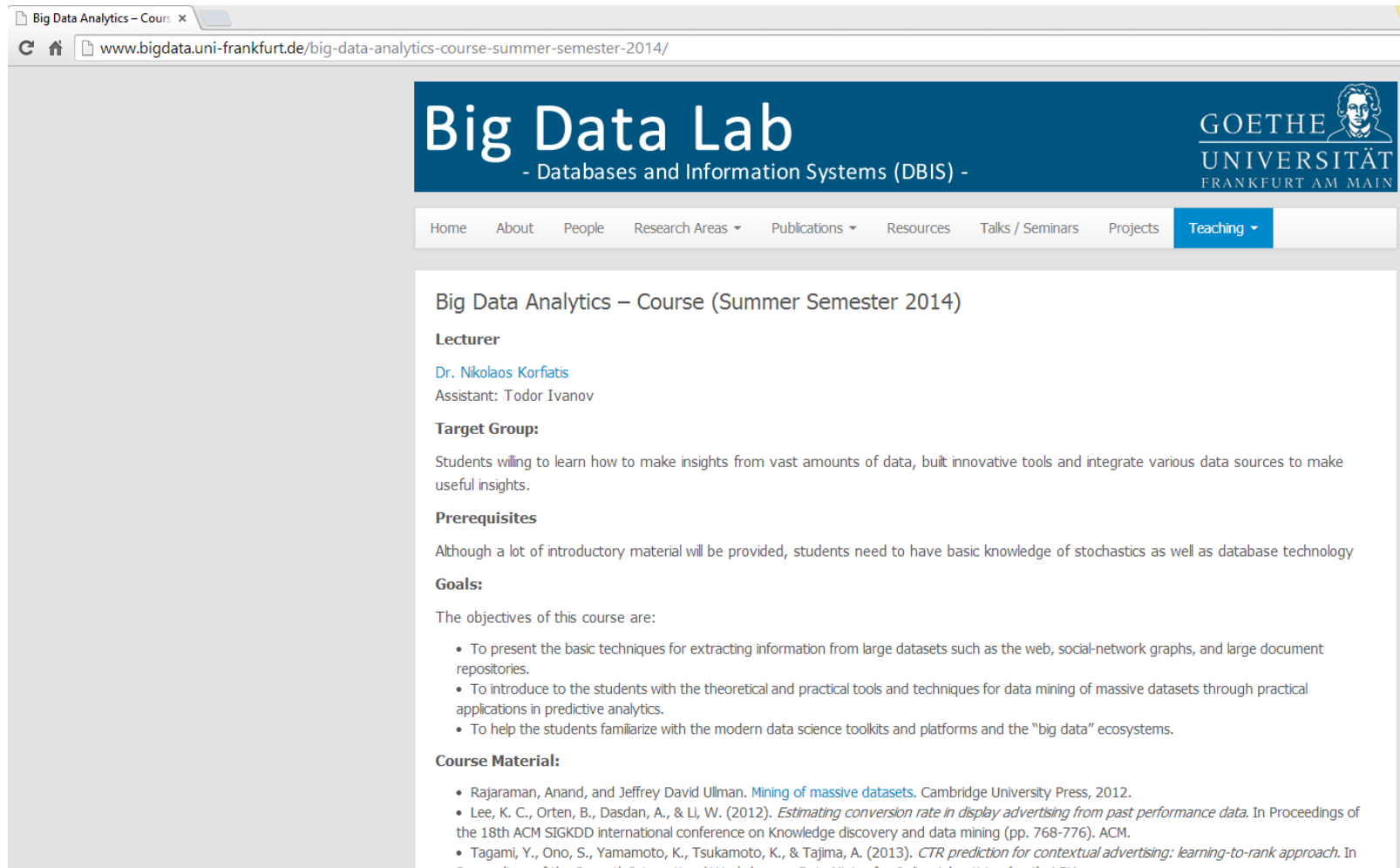
- Todor Ivanov, MSc
 - BSc in Computational Engineering (TU Darmstadt)
 - MSc in Distributed Software Systems (TU Darmstadt)
 - PhD Student at the Chair of Databases and Information Systems. Goethe University Frankfurt



Course Schedule

Date	Title
23.04.2014	Introductory concepts: data mining, statistical techniques, predictive analytics, text mining
30.04.2014	
07.05.2014	
	Map-Reduce and Distributed file systems
14.05.2014	Similarity Mining
21.05.2014	Link analysis and Pagerank
28.05.2014	
04.06.2014	
11.06.2014	
18.06.2014	Frequent Item sets
25.06.2014	Application workshop 1: Attribution in Advertising
02.07.2014	
09.07.2014	
16.07.2014	Application workshop 2: Recommender Systems

Course website



Big Data Analytics – Cours x

www.bigdata.uni-frankfurt.de/big-data-analytics-course-summer-semester-2014/

Big Data Lab

- Databases and Information Systems (DBIS) -

GOETHE
UNIVERSITÄT
FRANKFURT AM MAIN

Home About People Research Areas Publications Resources Talks / Seminars Projects Teaching

Big Data Analytics – Course (Summer Semester 2014)

Lecturer

[Dr. Nikolaos Korfiatis](#)
Assistant: Todor Ivanov

Target Group:

Students willing to learn how to make insights from vast amounts of data, built innovative tools and integrate various data sources to make useful insights.

Prerequisites

Although a lot of introductory material will be provided, students need to have basic knowledge of stochastics as well as database technology

Goals:

The objectives of this course are:

- To present the basic techniques for extracting information from large datasets such as the web, social-network graphs, and large document repositories.
- To introduce to the students with the theoretical and practical tools and techniques for data mining of massive datasets through practical applications in predictive analytics.
- To help the students familiarize with the modern data science toolkits and platforms and the “big data” ecosystems.

Course Material:

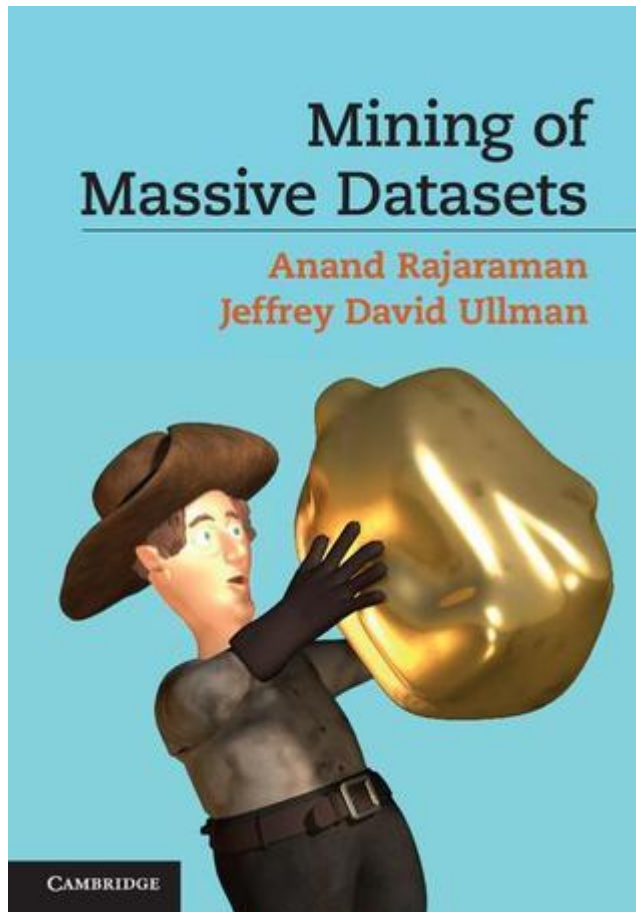
- Rajaraman, Anand, and Jeffrey David Ullman. [Mining of massive datasets](#). Cambridge University Press, 2012.
- Lee, K. C., Orten, B., Dasdan, A., & Li, W. (2012). *Estimating conversion rate in display advertising from past performance data*. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 768-776). ACM.
- Tagami, Y., Ono, S., Yamamoto, K., Tsukamoto, K., & Tajima, A. (2013). *CTR prediction for contextual advertising: learning-to-rank approach*. In Proceedings of the Seventh International Workshop on Data Mining for Online Advertising (pp. 1-10).

Office Hours

- Every Thursday 16.00 -17.00
- By appointment through email.

Course Credits

- Credit points
 - You will receive 3 credit points for this course upon successful examination



Mining of Massive Datasets

- Available for free download at:
 - <http://infolab.stanford.edu/~ullman/mmds.html>
- Course slides will be adopted from the course book
 - <http://www.stanford.edu/class/cs246/handouts.html>