

Course APPLIED DATA ANALYSIS

Basic Information

This is a course, which contributes to MSc award in Mathematics

Duration	144 hours (4 ECTS)
Starting date	September 1st
Study credits	4 ECTS credits
Language of instruction	English level B1 (European Framework of Reference of Communicative Skills)
Academic requirements	<ul style="list-style-type: none">– BSc degree in Mathematics, Physics, Computer Science, Engineering or equivalent (a copy of your diplomas from previous university studies and transcripts of completed courses and grades)– Skype interview

Course Description

“Applied Data Analysis” is the course, which develops students’ practical skills for solving real-world data problems. This course is for students, who already have foundational skills (which one can get by doing previous courses on data analysis – The Basics of Machine Learning and Advanced Methods of Data Analysis)

Through building their projects on data analysis in Engineering, Medicine, Education and Business student will also learn new methods and techniques.

Special Features of the Course

Although students will study new methods and their mathematical justification, the main feature of the course is working with real data and problems, provided by our industry and academic partners.

Course Aim

- To introduce methods and instrument of data analysis specific to different fields of knowledge;
- To develop skills of building analytical projects for real data

Course Objectives

- To introduce features of data analysis methods and their implementation to different fields;
- To develop skills of stating tasks and estimating information content of data;

- To equip student with abilities to make decisions and change methods and design of analysis if needed;
- To develop skills of leadership and working in groups.

Learning Outcomes of the Course

By the end of the course, students will be able to:

- Develop a research question;
- Estimate data quality and make decisions to collect more information or another sort of information if needed;
- Choose appropriate methods, metrics of model quality, taking into account features of the research field;
- Interpret results of analysis;
- Report their findings.

Course (module) Structure

Learning Activities	Hours
Lectures	18
Practice sessions / Seminars	36
Self-study Assignments	54
Final Exam (including preparation)	36
Total study hours	144

Course Outline

Week	Lectures	Practice session / Assignments	Hours ¹
1-3	Data analysis for Engineering equipment monitoring	<ul style="list-style-type: none"> - Anomaly detection - One-class support vector machines - Local Outlier Factor - Isolating forest - Project No 1 	20
4-7	Data analysis for business	<ul style="list-style-type: none"> - Customer behavior analysis - Sales network segmentation - Project No 2 	28
8-13	Learning analytics	<ul style="list-style-type: none"> - Prediction of learning success - Student simulation - Social nets analysis - Project No 3 	32
14-18	Data analysis in medicine	<ul style="list-style-type: none"> - Pattern recognition - Project No 4 	28
19-20	Final exam		36

¹ Hours designed for Classroom sessions, Web-sessions, Home Assignments etc.

Assessment

Grade policy for both practical projects and the final exam is:

- A (excellent work) 91–100 points
- B (above average work) 81–90 points
- C (average work) 71–80 points
- D (below average work) 50–70 points
- F (failed work) < 50 points

The exam is taken orally. Each exam ticket consists of 2 theoretical questions from the list of exam questions.

The final grade includes grade for projects and exam in a ratio of 60:40.

Attendance Policy

Students are expected to attend classes regularly. However, occasional skipping classes is permissible if a student does all necessary in-class work at home.

Lecturer(s) and Tutors, Contact Information



Tatiana KUSTITSKAYA

Ph.D., Associated Professor at School of Space and Information Technologies, Siberian Federal University

(room 311) 26-1, Kirensky st, Krasnoyarsk, Russia

Tel: +7 391 291 2790,

tkustitskaya@sfu-kras.ru, <http://ikit.sfu-kras.ru/e/204>

Core Reading

All the books are available on-line in Elsevier at SFU library.

1. **John Kruschke**, Doing Bayesian Data Analysis, 2nd Edition, A Tutorial with R, JAGS, and Stan, eBook ISBN: 9780124059160, Hardcover ISBN: 9780124058880, Imprint: Academic Press, Published Date: 3rd November 2014, Page Count: 776
2. **William Menke**, Geophysical Data Analysis, 4th Edition, Discrete Inverse Theory, Paperback ISBN: 9780128135556, eBook ISBN: 9780128135563, Imprint: Academic Press, Published Date: 12th April 2018, Page Count: 352