COURSE DESCRIPTION CARD - SYLLABUS

Course name
Statistics and data analysis

Course

Field of study
Computing
Area of study (specialization)
-
Level of study
First-cycle studies
Form of study
full-time
Year/semester
2/4
Profile of study
general academic
Course offered in
Polish
Requirements
compulsatory

Number of hours

Lecture
30
Laboratory classes
30
Other (e.g. online)
-
Tutorials
-
Projects/seminars
-

Number of credit points

4

Lecturers

Responsible for the course/lecturer:
Jerzy Stefanowski, Inst. Informatyki
Wydział Informatyki i Telekomunikacji
ul. Piotrowo 2, 60-965 Poznań
tel: 61 665-2933

Responsible for the course/lecturer:

Prerequisites

Students should have knowledge concerning probability and mathematical algebraic (discrete and continuous probability, expected values, variance, distributions of basic random variables, standardized normal distribution, Lindberg-Levy central theorem, point and interval estimation; vector and matrixes operations). He/she should have skills allowing solving basic problems related to the above concepts of the probability calculus and practical skills of using basic software tools. He/she should be able to acquire information from given sources of information. Finally student should understand the needs to extend their knowledge and competences.

Course objective

Provide students knowledge regarding statistical data analysis and its applications to computer and engineering sciences. In particular, it is planned to:

1. Present basic statistical knowledge on descriptive statistics (statistical measures as well as graphical exploratory tools); shifting from probability to statistical inference, testing hypotheses, modeling correlation and linear regression, basic non-parametric tests.
2. Develop students' skills in solving problems related to applying statistical or data mining methods to computer science problems, collecting data, choosing appropriate statistical tests and interpreting their results.
3. Acquire such skills by solving practical tests and using statistical software during laboratory classes (mainly R project).

Course-related learning outcomes

Knowledge

Students should have:
1. well founded knowledge on statistics useful for understanding fields of computer science and other engineering disciplines (K1st_W4-W5)
2. knows basic methods of descriptive statistics, data presentations, statistical inference constructing histograms, correlation and regression, parametric and non-parametric tests (z-test, t-Student test, F-test, chi-test) which are particularly useful to solve problems within computer science, in particular analysis algorithms, computer systems and artificial intelligence (K1st_W7)

Skills

1.is able to calculate basic statistical measures (K1st_U3)
2. is able to use descriptive statistics measures and hypothesis testing to solve simple experimental tasks of computer science (K2st_U4)
3. can verify statistical hypotheses for computer sciences and related ones (K1st_U4)
4. is able to analyse simple problems of data analysis, choose an appropriate statistical method, present and interpret results of the statistical analysis

Social competencies

1. understands that knowledge and skills related to computer science and data mining quickly becomes non relevant and it is necessary to self-learn and extend his/her knowledge and skills (K1st_K1-2)

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

Lecture: based on a written test including several questions on the scope of course

Laboratory: evaluation of student's knowledge necessary to carry out steps of solving a lab task; monitoring students' activities during classes; evaluation of lab reports (partly started during classes, finished after them). Additionally short written or online tests, oral answers to questions

Programme content

Lecture covers the following topics:
An introduction to selected topics of statistics and data analysis. Its scope includes: measurement scales, graphical tools, descriptive measures (mean, median, mode, variance, standard deviation), selected reminding of binominal and standarized normal distributions, introduction to hypothesis testing, error analysis, significance levels, Z-test for a single population, t-test for a single population, tests for

comparing means in two populations, F-test, studying relationships between numerical variables, Pearson correlation and linear regression, non-parametric tests (chi-test, Wilcoxon test,) rang Spearman correlation, basic steps of statistical data collection and sampling schemas, selected topics of graphical presentation and exploratory analysis.

Laboratories –working on assigned competition tasks. The tasks are divided into steps and milestones – student should regularly report their activity. Simple written tests, oral answers to questions

Teaching methods

Lecture - learning methods based on multimedia presentation, illustrated with examples or case studies, solving together tasks

Laboratories – tasks, practical exercises also with some programming, discussion, teamwork.

Bibliography

Basic

1. Statystyka dla studentów kierunków technicznych i przyrodniczych, J. Koronacki, J. Mielniczuk, WNT, Warszawa, 2006 (I wydanie 2004)
2. Statystyka w zarządzaniu. A. Aczel, PWN, Warszawa, 2002

Additional

1. Przystępny kurs statystyki. Tom 1, A. Stanisz, Statsoft, Kraków, 1997
2. Przewodnik po pakiecie R. P. Biecek, Oficyna Wyd. GiS. Wrocław 2008
3. Statystyka praktyczna, W. Starzyńska, PWN, Warszawa, 2006

Breakdown of average student's workload

	Hours	ECTS
Total workload	100	4
Classes requiring direct contact with the teacher	60	2.0
Student's own work (literature studies, preparation for laboratory classes/tutorials, preparation for tests/exams, project preparation)[1]	40	12.0

[1] niepotrzebne skreślić lub dopisać inne czynności