STUDY MODULE DESCRIPTION FORM

Name of the module/subject **Process Mining**	Code **1010512321010510205**

Field of study **Computing**	Profile of study (general academic, practical) **general academic**	Year /Semester **1 / 2**
Elective path/specialty **Intelligent Decision Support Systems**	Subject offered in: **English, Polish**	Course (compulsory, elective) **elective**

Cycle of study: **Second-cycle studies**	Form of study (full-time,part-time) **full-time**

No. of hours				No. of credits
Lecture: **30**	Classes: **-**	Laboratory: **30**	Project/seminars: **-**	**5**

Status of the course in the study program (Basic, major, other) **major**	(university-wide, from another field) **from field**

Education areas and fields of science and art **technical sciences**	ECTS distribution (number and %) **5 100%**

Responsible for subject / lecturer:

dr inż. Tomasz Pawlak
email: Tomasz.Pawlak@cs.put.poznan.pl
tel. 61 6653022
Faculty of Computing
ul. Piotrowo 3 60-965 Poznań

Prerequisites in terms of knowledge, skills and social competencies:

1	**Knowledge**	Learning outcomes from first-cycle studies defined in the Resolution of the Senate of the PP, especially the effects of K1st_W4, K1st_W5, K1st_W6, K1st_W7, K1st_W8, K1st_W9, K1st_W11 verified in the recruitment process for second-cycle studies. These effects are presented on the website of the faculty www.fc.put.poznan.pl
2	**Skills**	Learning outcomes from first-cycle studies defined in the Resolution of the Senate of the PP, especially the effects of K1st_U1, K1st_U2, K1st_U4, K1st_U5, K1st_U6, K1st_U7, K1st_U9, K1st_U10, K1st_U11, K1st_U12, K1st_U14, K1st_U15, K1st_U19, verified in the recruitment process for the second degree studies. These effects are presented on the website of the faculty www.fc.put.poznan.pl The student should also have the ability to obtain information from the indicated sources.
3	**Social competencies**	Learning outcomes from first-cycle studies defined in the Resolution of the Senate of the PP, especially the effects of K1st_K1, K1st_K2, K1st_K3, K1st_K4, verified in the recruitment process for second-cycle studies. These effects are presented on the website of the faculty www.fc.put.poznan.pl The student should also understand the necessity of expanding their competences and be ready to cooperate within the team. In addition, in terms of social competences, the student must present such attitudes as honesty, responsibility, perseverance, cognitive curiosity, creativity, personal culture, respect for other people.

Assumptions and objectives of the course:

1. Provide students with basic knowledge about the use of modeling tools, execution management and process analysis in distributed systems.
2. Developing students' problem-solving skills related to analysis and design of distributed processes within the technologies used in business.
3. Developing students' skills in analysis and diagnostics of the work of distributed processes using tools used in business.
4. Teaching students the skills of teamwork, analytical thinking, conclusions based on observations and analytical models.

Study outcomes and reference to the educational results for a field of study

Knowledge:

1. The student has ordered, theoretically founded general knowledge in the field of industrial systems architecture in a distributed environment. - [K2st_W1]

2. The student has theoretically developed detailed knowledge related to selected issues in the field of process exploration, such as: design, implementation and analysis of process management systems. - [K2st_W3]

3. The student has knowledge about development trends and the most important new achievements in information technology in the field of enterprise resource planning systems. - [K2st_W4]

4. The student knows advanced methods, techniques and tools used to solve complex tasks and conduct research in the field of process exploration. - [K2st_W6]

Skills:

1. The student is able - when formulating and solving engineering tasks - to integrate knowledge from various areas of computer science (and if necessary also knowledge from other scientific disciplines) and apply a systemic approach, also taking into account non-technical aspects - [K2st_U5]

2. The student can assess the usefulness and the possibility of using new achievements (methods and tools) and new IT products - [K2st_U6]

3. Student is able - using m.in. Conceptually new methods - solve complex IT tasks, including atypical tasks and tasks containing a research component - [K2st_U10]

4. The student is able - in accordance with the given specification, taking into account non-technical aspects - to design a complex device, IT system or process and implement this project - at least in part - using appropriate methods, techniques and tools, including adapting to this purpose existing or developing new tools - [K2st_U11]

5. The student can determine the directions of further learning and implement the process of self-education, including other people - [K2st_U16]

Social competencies:

1. The student understands that in informatics, knowledge and skills quickly become obsolete. - [K2st_K1]

2. The student understands the importance of using the latest knowledge in the field of computer science in solving research and proctological problems. - [K2st_K2]

Assessment methods of study outcomes

The learning outcomes presented above are verified in the following way:

Forming rating:

a) in the field of lectures:

- based on the answers to questions about the material discussed in lectures,

b) in the field of laboratories:

- on the basis of an assessment of the current progress of laboratory tasks,

Summary rating:

a) in the field of lectures, verification of the assumed learning outcomes is accomplished by:

- assessment of knowledge and skills demonstrated in a written colloquium consisting of:

- a set of 8 to 12 closed questions, each of which can be answered with one correct answer from four possible ones. One point is obtained for each correct answer, and 1/3 of the point is taken for incorrect answer.

- A set of 3 to 9 open questions, for which you can get from 2 to 4 points.

b) in the field of laboratories, verification of the assumed learning outcomes is carried out by:

- evaluation of the implementation of project tasks including modeling, implementation and analysis of processes occurring in distributed systems, partly carried out during classes, and partly as part of homework,

- timeliness of tasks implementation;

Obtaining additional points for activity during classes, especially for:

- implementation of additional tasks,

- discussion of additional aspects of the issue,

- effectiveness of using the acquired knowledge while solving a given problem,

- ability to cooperate in a team practically performing a detailed task in a laboratory,

- comments related to the improvement of didactic materials,

- presentations of the self-presentation related to the subject of the classes.

Course description

The lecture program includes the following topics:

1. Introduction to distributed business processes, process learning, process exploration, business process management, management methodology: lean management, six sigma.

2. Methodologies of business process modeling: transition systems, Petri networks, workflow systems, YAWL, BPMN, EPC, causal nets, process trees. Verification and analysis methods of processes based on models.

3. Introduction to data mining: basics of statistical inference, machine learning, supervised and unsupervised learning, methods of assessing knowledge models, discovering local patterns: association rules, sequences and episodes.

4. Preparation of data for analysis: ETL process, typical data formats, preparation challenges, data quality assessment.

5. Basic algorithms of discovering models of distributed processes: algorithm? and its variations
6. Advanced algorithms for the discovery of distributed process models: Heuristic Miner, evolutionary algorithms, state-based algorithms, region-based and formal languages ??algorithms, Inductive Miner.
7. Models of mathematical programming: modeling using expert knowledge, discovering models from data.
8. Verification of compliance of the business process implementation with the model and / or domain knowledge, methods of diagnosing reasons for deviations from the model.
9. Alternative perspectives of business processes: the perspective of resources, costs, and efficiency.

The laboratory program includes the following issues:

1. Workflow systems: Process Maker.
2. ProM tool: analysis of event logs, discovery of business process models, analysis of business process models and process improvement.
3. Disco tool: analysis of event logs, discovery of business process models, analysis of business process models and process improvement.
4. Mathematical programming: modeling of processes and systems using the mathematical programming paradigm.

Teaching methods:

1. lecture: multimedia presentation.
2. laboratory exercises: solving tasks, practical exercises, team work, demonstration.

Basic bibliography:

1. Wil van der Aalst, Process Mining: Data Science in Action, Second Edition, Springer, 2016, http://link.springer.com/978-3-662-49851-4 (available online from computers in university network).

Additional bibliography:

1. H. Paul Williams, Model Building in Mathematical Programming, Fifth Edition, Wiley, 2013.
2. Gopal K. Kanji, 100 Statistical tests, Third Edition, SAGE Publications, 2006.
3. Peter Flach, Machine Learning: The Art. Of Science of Algorithms that Make Sense of Data, Cambridge University Press, 2012.
4. Tomasz P. Pawlak, Krzysztof Krawiec, Automatic synthesis of constraints from examples using mixed integer linear programming, European Journal of Operational Research 261:1141-1157, 2017.
5. Tomasz P. Pawlak, Krzysztof Krawiec, Synthesis of Constraints for Mathematical Programming with One-Class Genetic Programming, IEEE Transactions on Evolutionary Computation 23(1):117-129, IEEE Press, 2019.
6. Patryk Kudła, Tomasz P. Pawlak, One-class synthesis of constraints for Mixed-Integer Linear Programming with C4.5 decision trees, Applied Soft Computing 68:1-12, 2018.
7. Tomasz P. Pawlak, Synthesis of Mathematical Programming models with one-class evolutionary strategies, Swarm and Evolutionary Computation 44:335-348, Elsevier, 2019.
8. Daniel Sroka, Tomasz P. Pawlak, One-Class Constraint Acquisition with Local Search, GECCO '18, pp. 363-370, ACM, 2018.
9. Dokumentacja UML: http://www.omg.org
10. Dokumentacja systemu YAWL: http://www.yawlfoundation.org
11. Dokumentacja systemu Process Maker: https://www.processmaker.com/resources
12. Dokumentacja systemu ProM: http://www.promtools.org

Result of average student's workload

Activity	Time (working hours)
1. Participation in laboratory classes	30
2. Preparation for laboratory classes	15
3. Implementation (as part of your own work) of project tasks	15
4. Participation in consultations related to the implementation of the education process, in particular laboratory classes / project (partly can be carried out electronically)	4
5. Participation in lectures	30
6. Getting to know the indicated literature / didactic materials	10
7. Preparation for the colloquium	17
8. Participation in the colloquium	2
9. Discussing the results of the colloquium	2

Student's workload

Source of workload	hours	ECTS
Total workload	125	5

Contact hours	62	2
Practical activities	60	2