Obuda University John von Neumann Faculty of Informatics			Institute of Biomatics	
Name and code: *Database- and Big Data technologies* (*NIXAB1EMNE*) **Credits*: 4*** *Computer Engineering MSc (English language)* *2019/20 year 2. semester*				
Subject lecturer: Rita Fleiner, Péter Piros				
Prerequisites (with code):				
Weekly hours:	Lecture: 2	Seminar.: 0	Lab. hours: 2	Consultation: 0
Way of assessment:	Practice project work & Written exam			
Course description:				
Goal: The aim of the lesson is to familiarize students with advanced database management concepts and procedure.				
Course description: Data and relational data models, relational algebra, SQL deep dive. Logical and physical data model, relations. RDBMS design, dependencies, constraints, normal forms, normalization. Triggers and constrains in SQL. Database fundamentals, instance definition, memory structures in db. Transactions. Index types, hashes. SQL tuning. NoSQL databases: types, concepts, architecture, queries. Translating business problems into data. R language.				

Schedule	
Education week	*Topic*
1.	Introduction. Knowledge assessment. Relational database systems. Advanced SQL exercises.
2.	Database architecture, Database instance. Advanced SQL exercises.
3.	From SQL basics to advanced SQL. Execution plan, database tuning, access paths, indexes, join types, CBO statistics, selectivity, costs, materialization, pipelining. Execution plan analysis.
4.	From SQL basics to advanced SQL. Execution plan, database tuning, access paths, indexes, join types, CBO statistics, selectivity, costs, materialization, pipelining. Execution plan analysis.
5.	NoSQL databases: concepts, types. Key-value stores. Redis: concepts, architecture, queries.
6.	NoSQL databases. Cassandra: concepts, architecture, queries
7.	Basics of Big data. Hadoop framework. Spark.
8.	NoSQL databases. MongoDB: concepts, architecture, queries
9.	Basics of machine learning I. R: Introducing the language and environment, variable types, data preparing
10.	Basics of machine learning II. R: Models, parameter tuning, R packages, Caret, random and grid search
11.	Modelling – Regression, tree-based models, neural networks R: Modelling (Regression, tree-based models, neural networks)
12.	Validation – Model validation, types of cross validation R: Validation, data splitting, cross validation, model selection
13.	Final test
14.	Extra final test

Midterm requirements

Student participation in the classroom is required.
There will be a written test in the 13th week in the lecture and in the lab. It can be rewritten in the 14th week. Students will get optional homework projects to solve and to present continuously during the semester.
The project work is graded with extra points, and the written test is with 50-50 points for the lecture and for the lab. The minimum requirement to complete the course is to achieve at least 26 points at each test (lab and lecture).

Final grade calculation methods

The final grade is formed from the project grade and the exam grade.

Achieved result	Grade
85%-100%	excellent (5)
74%-84<%	good (4)
63%-73<%	average (3)
51%-62<%	satisfactory (2)
0%-50%	failed (1)

Type of requirement

Theoretical test

Type of replacement

In the 14th week for all of the tests.

References

Obligatory: Lecture notes (download form https://elearning.uni-obuda.hu/)

Recommended:
Elmasri, Navathe: Fundamentals of Database Systems

Other materials: -