

## Topics for the diploma examination in second-cycle studies

### **ADVANCED ANALYTICS – BIG DATA**

1. Supervised and unsupervised learning techniques.
2. Selection of variables and observations for a model.
3. Data classification methods.
4. Decision tree models.
5. Tree-based models.
6. Artificial neural networks.
7. Data grouping (aggregation) methods.
8. Non-parametric models of survival time-to-event analysis.
9. Parametric models of survival time-to-event analysis.
10. Semi-parametric models of survival time-to-event analysis.
11. Bayesian approach in survival analysis.
12. Predictive model development procedure.
13. Examples of machine learning models used in regression and classification tasks.
14. Data quality.
15. Data imputation.
16. Multiple imputation.
17. Longitudinal analyses.
18. Quantile regression.
19. Adaptive regression.
20. Cluster analysis.
21. Models for assessing customer lifetime value.
22. The data analysis process in the context of visualisation.
23. Selection of data visualisation methods.
24. Decision engine in automated credit assessment processes.
25. Selection and preparation of data for scoring model development.
26. Scorecard.
27. Profitability analysis of the credit application acceptance process.
28. The problem of risk estimation for rejected applicants.
29. Big data solutions in the area of data repositories.
30. Specific characteristics of analytical environments used in big data.

**ADVANCED ANALYTICS – BIG DATA**

31. Selected algorithms used in big data analytics.
32. The MapReduce algorithm.
33. Typical challenges related to big data.
34. Distributed processing.
35. Ethical issues related to big data.
36. The 3V and 5V concepts in the context of big data.
37. Structured and unstructured data in the context of big data.
38. Comparison of Lambda architecture and Kappa architecture.
39. Concept and business applications of data warehouses.
40. Differences between batch and stream data processing.
41. Business applications of real-time data analytics.
42. Iteration techniques.
43. The concept of a function in programming: typical elements and objectives.
44. Program execution through interpretation of source code.
45. Library (package) management in the Python language environment.
46. The CIA triad (confidentiality, integrity, availability) in information security.
47. Spear-phishing cyberattacks.
48. Basic cybersecurity recommendations for SMEs.
49. Methods of performing table joins in a relational database.
50. Classification of SQL single row functions.
51. Reasons for creating views in a relational database. Clauses specific to view-creation statements.
52. Set operators – formal syntax and interpretation of results.
53. Types of problems solved by subqueries. Clauses where subqueries occur, operators.