	Data Mining and Data Warehouse 3 CREDITS					
	Prerequisites:		Familiarity with the DataBase			
	Learning outcomes:		As a result of discipline studying students should know: • purpose, structure and operations for data warehouse • tasks and techniques of data mining be able: • to use tools to create data warehouse and reports on it • to use data mining tools have an idea about advanced techniques of data mining			
№	**Lecture**	**Hours**	**Laboratory works**			**Ref**
			Content		**Hours**	
1.	**Data Warehousing (DWH)**					
	Decision Support Systems. DWH definition, usage and trends. DWH vs. DBMS. DWH process and architecture. OLAP: data marts, data cubes, multidimentinal data schemas: stars, snowflakes and and fact constellations. OLAP vs. OLTP. MOLAP, ROLAP, HOLAP. ETL (Extract, Transform, and Load) process.	2 h	Tables, relationship tables, fact tables, indexes. Column names and data types, initial index plan, fact table partitioning. SQL query and OLAP query. Creating a OLAP cube in MS SQL Analysis Services (MSAS).		3 h	[4][2][3]
2.	**Data Warehouse implementation, data cubes**					
	Computation of data cubes, modeling OLAP data, OLAP queries manager (Roll-up, drill down, slicing, dicing), data warehouse back end tools, complex aggregation at multiple granularities, tuning and testing of data warehouse.	2 h	Computation of data cubes, modeling OLAP data, OLAP queries manager (Roll-up, drill down, slicing, dicing), data warehouse back end tools, complex aggregation at multiple granularities, tuning and testing of data warehouse.		2 h	[4][2][3]
3.	**Introduction to Data Mining**					
	Definitions and motivations: why Data Mining? Data Mining basics, related concepts and techniques. Data Mining algorithms overview.	2 h	Learning basic concepts using real-world examples. Data exploration with some tool.		2 h	[1] [2]
4.	**Data Mining Algorithms: Classification and Prediction**					
	Classification tasks definition and applications. Classification tasks techniques. Decision trees. Naïve Bayes. Nearest neighbors. Support Vector Machines (SVM). The prediction task. Time series prediction.	3 h	Practice of classification? with various data sets. Classification? with a tool. Create a Mining Model (classification) by using OLAP Data in MSAS.		3 h	[1][3]
5.	**Data Mining Algorithms: Associate Rules and Clustering**					
	Item sets. Generating item sets and rules. Correlation analysis. Apriori algorithm. Basic issues in clustering. Partitioning methods and hierarchical methods of clustering.	3 h	Practice of clustering and association rules with various data sets. Clustering and association rules with a tool. Create a Mining Model (clustering) by using OLAP Data in MSAS.		3 h	[1][2] [3]
6.	**Data Mining Cycle. Applications and Trends in Data Mining. Advanced techniques**					
	Cycle of gathering, preprocessing, analysing data, interpreting and using the results. Data Mining Applications. Business Intelligence Applications. Data Mining Software. Major Challenges in Data Mining. Text mining, web mining, spatial mining, temporal mining.	3 h	Case study involving data gathering, preprocessing, and applying data mining techniques		2 h	[1]
	TOTAL	**15 h**			**15 h**	

References

[1]   Ian H. Witten, Eibe Frank, **Data Mining: Practical Machine Learning Tools and Techniques** (Second Edition),

[2]   Jiawei Han and Micheline Kamber, **Data Mining: Concepts and Techniques**,

[3]   Jamie MacLennan, ZhaoHui Tang, Bogdan Crivat, **Data Mining with Microsoft SQL Server 2008**

[4]   Ralf Kimball, Margy Ross. **The Data Warehouse Toolkit**. John Willey and Sons, 2002.