1-Fuzzy Logic : Albert Einstein: „If mathematics describes reality, it is not accurate. And if it is accurate, it does not describe the reality“.Trying to meet the reality with its imprecision and uncertainty.The basic element is a fuzzy set.

Fuzzy set is a set that besides full membership or no membership admits partial membership.It belongs to multivalued logics.

Fuzzy logic may be more suitable for many real decision-making tasks than classical logic,because it allows the design of complex systems.

2-Probabilistic “computation”: The most important method: Bayesian network.

Powerful probabilistic model that uses a graphical representation to show probabilistic relationships between events.

3- Difference between Machine Learning and Datamining:

Machine learning focuses on the prediction based on the known properties learned from the training data.

Data mining is focused on the discovery of (previously) unknown properties in the data.

4-Types of learning 1 - learning “with a teacher” (supervised)

✤ inputs and outputs given by a kind of "teacher" in order to create a general function - a rule that maps inputs to the output

✤ the goal is to reduce the output error, ie to approach the desired output

✤ learning (training) sets the model so that the current output from the method is as close as possible to the desired output value

✤ a principle similar to teaching a child to pronounce letters

5-Types of learning 2 - learning without a teacher (unsupervised)

✤ no outputs given (labels of input data)

✤ the structure itself is searched - datamining

✤ or search-learn of suitable inputs (features learning)

✤ the learning algorithm does not receive the required output, it only has data with input information

✤ the algorithm / technique itself decides how to rearrange the data

✤ creates groups (classes, clusters) according to similarity (similar to a human).

6-Machine learning tasks

✤Classification - inputs are divided into two or more classes. The goal is to find a model that assigns

"unknown" inputs to the appropriate class, i.e. identifies a class based on attribute values

✤ usually learning “with a teacher" (supervised)

✤regression - also supervised, but the outputs are more continuous than discrete

the relationship between input and output is “searched” ,estimate the value of an attribute (output) depending on other attributes

7-Naive Bayes :it is not searching the hypothesis space,. but only probabilities based on the occurrence frequency of attribute values (it is possible to apply this approach also on big datasets)

Naive Bayes – A simple classifier that works directly with individual data components. For each component, a score is generated, which with everyone else determines which class we classify it into.

✤ advantage (+):✤ possibility to classify also not complete patterns (missing values)

✤ disadvantage (-):✤ zero posterior probability if missing patterns in training sets

✤ under estimation if low mutual frequency

8-Bayesian nets :✤ Dealing with the fact that features (attributes) are not mutual independent - it would be necessary to consider all possible combinations of outcomes

✤ Bayesian networks describe conditional independence from a subset of data(attributes)

✤ Knowledge burdened by uncertainty (inaccurate, uncertain, vague)

✤ It is used in decision making area

9-Decision trees: are used mainly because of the advantageous interpretability of the model for users

✤ the decision tree consists of a root and nodes ✤ the root represents the whole node,

Types of trees

✤ binary - only two branches grow from one node✤ maximum number of descendants = 2

✤ non-binary - two or more branches grow from one node

✤ the maximum number of is generally n

✤ they are generally called B-tree

Decision trees - supervised method ✤ input variables (features) – predictors ✤ output variable ✤ categorical - classification into classes ✤ continuous – regression

10 -Regression trees : ✤ The principle of regression trees is to estimate the continuous value of the output ✤ similar to example with playing tennis, but the output is not YES / NO, but how many hours of the game were played✤ It is necessary to calculate the average (value in the terminal leaves - there is no category, but the average), standard deviation (for branching) and variance coefficient (stoppingcondition)

11-Ensemble method - Random forest :✤ Random forest is a method that uses multiple trees for the final classification

✤ Since the creation of trees in the form of TDIDT is a de facto deterministic method, it is necessary to work with the dataset itself in order to be able to create different trees

✤ the bootstrap-bagging (bootstrap aggregating) method is used

✤ the aim is to avoid overfitting and to find a robust model for the dataset

✤ in principle, multiple models cooperate in the ensemble method

12-Support vector machines (SVM):is basically a method of linear separation, which works very well for nonlinear separable problems, because it works with transformations of original data.In its basic form, it is a binary classifier that decides which of the two classes a given sample belongs to.

13-Fuzzyfication : is the process of assignment of measured values of input variables into fuzzy sets via membership function .

14-Defuzzification - a process of approximation of noncrisp terms by a crisp value

15-Algorithm for DM :

1-C4.5 – Algorithm for generating decision trees or rule sets. It generates subtrees to. Decision rules do not have tobe binary.

2- k–means – An algorithm that divides a data set into k clusters based on the distance between samples. There are many mutations: Fuzzy k-means, k-means plus plus…

k–means is a clustering algorithm that is a typical example of unsupervised training.

3-kNN – k–Nearest Neighbors – It is a simple classifier that assigns a class to a given sample based on a given number to other nearest samples. The sample is included in the class to which most samples belong to k.

4-DBSCAN - The algorithm differs from the approaches of other algorithms in that clusters are not based only on the distances of points from centroids, but on the density of occurrence in the searched area. The main idea is that there is a higher density of points inside the clusters than outside them, where there is noise.

5-RS - Random Search - This is a random generation of solutions, which are further evaluated and only the best solution is stored until the maximum number of iterations is run out.