Two basic approaches are used to construct Bayesian networks (Kekolahti 2011): data-based and knowledgebased approaches. Data-based method uses conditional independence semantics of Bayes networks to make inferences from data whereas the knowledge-based approach utilizes causal knowledge from domain experts to construct Bayesian network. Before training Bayesian network only data about the domain being modelled are given, i.e. neither the structure, nor the conditional probabilities are known. According to Deventer (2004), when learning the parameters of a distribution in the Bayesian network, it is of advantage, if all nodes are observed. Besides, the usage of hidden nodes, which do not represent an existing value, is sometimes helpful in order to reduce the number of parameters. The data of the modelled processes also can contain continuous variables. Thus, it is necessary to use a training method which is able to deal also with continuous values, ideally with discrete and continuous values at the same time (Deventer 2004). In Bayesian network uncertainty is assessed in point estimates (posterior distribution's statistics of central tendency). It is desirable for a point estimate to be consistent (the larger the sample size, the more accurate the estimate) and unbiased (possessing the smallest variance). Few approaches are applicable for Bayesian network learning (calculation of the estimator): *maximum likelihood estimation, moment’s method, Bayesian approach*. *Expectation maximization* algorithm is frequently used for training with missing data. It is based on maximum likelihood estimation and is able to deal with discrete and continuous data at the same time (Deventer 2004). The EM algorithm enables to derive the machine learning estimates in a mixture model:

* E step: compute the conditional probabilities *,* that arises from the th component for the current value of the mixture parameters;
* M step: update the mixture parameter estimates maximising the  
  expected value of the completed likelihood. It leads to weight  
  the observation for group with the conditional probability .

Using Bayesian networks, any question that can be posed in a probabilistic form can be answered with certain level of confidence. Problems of classification, segmentation, state estimation, fault diagnosis, prediction might be solved using Bayesian model. Bayesian networks might be used to find the probability of independent variable being in a certain state, given certain values of other variables in the network or to find values of a particular set of variables that best explains why set of other variables have certain values. Student’s learning style modelling belongs to the first task. Thus, to predict student’s learning style using BN, it is needed:

* to build and learn generative model from historical data about student’s behavioural activities, influencing respective learning style dimensions
* using Bayes rule, compute the probability that the particular learning style dimension is characteristic to the student, given new evidence of student’s behaviour.