Before deciding to apply Bayes network for a particular task, Bayes network’s advantages and drawbacks should be considered. Kekolahti (2011) and other authors stated the following advantages of the parametric Bayesian modelling: possibility to model sequential data and causal relationships, to combine prior knowledge and data, to use Bayes’ manage situations where some data is missing, possibility of structural learning of Bayes network, possibility of making inferences about probabilities of different causes given the consequences, possibility to incorporate different types of information. The following general disadvantages of parametric Bayesian modelling were identified (Goštautaitė 2019): BN only encodes directional relationship (not bi-directional) – cyclic relationship is not allowed; there is a lack of support of feedback loops due to acyclic nature of network (Kekolahti 2011); BN reflects impact, but not the cause and effect relationship; data to be modelled should be independent and identically distributed; inferences are made using past historical data only – Bayesian network doesn’t know how to react to unforeseen situations; here so called “cold start” problem should be mentioned; inferences are subjective and probabilistic, they are useful only in cases when prior knowledge is reliable and are not suitable for science where strong causality cases are under consideration; inferences are brute: in reality, there is a low likelihood of independence of variables; joint probability distribution may be biased when it is not modelling complete data, i.e. it may only be suitable under closed-world assumptions (Castillo 2006); typical Bayesian network is finite dimensional model having finite number of parameters, therefore it is not flexible; for some problems nonparametric Bayesian models are more suitable; continuous variables are handled only in a limited manner [20]; there is a high computational cost in models with a large number of parameters; learning a Bayesian network (i. e. selecting a probabilistic model that explains a given set of data) is complex – there is no efficient deterministic way to find the best network, and randomized algorithms like Markov chain Monte Carlo might be needed to find a good network; Bayesian network’s structure learning is NP-complete problem; simple Bayes network models only linear hypothesis and likelihood distributions; to use a probability distribution function alone to represent uncertainty is not enough, because it fails to show the ignorance (called “sensitivity”), or uncertainty about the function itself, i.e. the Bayesian approach has no general way to represent and handle the uncertainty within the background knowledge and the prior probability function (Wang 2004). Wang (2004) emphasizes that in order to apply a Bayesian network to a practical domain, one of the following requirements must be satisfied:

* the implicit condition of the initial probability distribution, that is, the domain knowledge used to determine the distribution initially, can be assumed to be immune from future modifications, or
* all modifications of the implicit condition can be treated as updating, in the sense that when new knowledge conflict with old knowledge, the latter is completely abandoned.

Bayesian network is not able to combine conflicting beliefs that are based on different implicit conditions and to carry out inference when premises are based on different implicit conditions. Confusion between explicit and implicit conditions of probability evaluations (distinction between revision and updating as well) causes underestimation about the limitation of “Bayesianism” (Wang 2004). Naive Bayes modelling makes an assumption that all the attributes (features) are conditionally independent. For real cases when it is necessary to model directly or indirectly dependent of each other attributes, an extension of the Naive Bayes model that also includes some (conditional) dependencies between the random variables might be usable (Friedman *et. al.* 1997). An example is augmented Naive Bayes network which augments network by adding edges between the attributes to include the information of interdependence between the attributes (Friedman *et. al.* 1997). TreeAugmented Naive Bayes Model in which each variable has only one or two parents, except for the root variable, belongs to the category of such models. More accurate than simple Bayes models should learn as the new sample data arrives – adaptive Bayesian networks may be used for that purpose (Goštautaitė 2019). This category of Bayesian models can handle concept drift, i. e. detect when concept change has occurred and adapt to the changes accordingly. Adaptive Bayesian networks also handle the trade-off between cost of updating and the gain in performance, deciding whether it is necessary to initiate updating process.