Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)

Active Inference for Dynamic Bayesian Networks

Caner Komurlu
Illinois Institute of Technology, Chicago Illinois
ckomurlu@hawk.iit.edu

Abstract

In supervised learning, many techniques focus on optimizing training phase to increase prediction performance. *Active inference*, a relatively novel paradigm, aims to decrease overall prediction error via selective collection of some labels based on relations among instances. In this research, we use dynamic Bayesian networks to model temporal systems and we apply active inference to dynamically choose variables for observation so as to improve prediction on unobserved variables.

1 Introduction

In supervised learning, a mathematical model is trained by tuning its parameters using labeled data in order to automatically predict labels for unseen data. Many studies have focused on training. For example, active learning tries to train a model using fewer labeled data by selecting most informative instances. This helps reducing labeling cost [Settles, 2012].

A relatively new approach, *active inference*, maximizes prediction performance by selective information gathering during prediction [Bilgic and Getoor, 2009]. In this approach, relations between instances are utilized with the intuition that knowing true label of some instances help predicting others.

Dynamic Bayesian network (DBN) is a generative statistical model which asserts probabilities of random variables accounting complex dependencies. Two main properties of DBNs make them powerful: i) factorizing joint probability distributions into conditional probability distributions, ii) dynamically representing random variables in time dimension.

In majority of cases, lack of evidence degrades prediction performance in DBNs over time. As random variables are correlated, observing one contributes to evaluating probabilities on its dependents. In some scenarios, observed variables are specified by definition. Otherwise, selecting variables to observe arises as a problem to tackle. Therefore, active inference can help to detect variables to observe and eventually it can increase prediction performance significantly. This objective revolves around assessment of prediction uncertainty and calculating observation cost. To the best of our knowledge, this is the first time active inference is applied on DBNs.

In the following, Section 2 presents description of these objectives on two practical problems along with results obtained. Section 3 continues with short and long term research plans, followed by a conclusion in Section 4.

2 Preliminary Research and Results

In this section, our proposed method, *active inference for dynamic Bayesian networks*, will be described and evaluated on two practical problems: i) detecting optimal time for observation for tissue engineering, and ii) dynamic detection of optimal observation subset on wireless sensor networks.

2.1 Active Inference for Tissue Engineering

In tissue engineering domain, experts seek conditions for optimal tissue development. One criterion for optimal development is blood vessel network which should develop in tandem with tissue, also named as vascularization. Though many factors affect the performance of vascularization, few are known. Therefore, this phenomenon is partially observable, hence stochastic. Given an initial configuration, e.g initial blood vessel and tissue cell locations, stress level of tissue cells, we try to estimate probabilities of each atomic locations being occupied by blood vessel in a sequence of time stamps.

We modeled the environment as a grid, of which each cell represents an atomic location. We assumed that the direction of blood vessel progress is from bottom to top. We designed a DBN in which each random variable represents a location whose parents are lower neighbors from previous time. Hence occurrence of blood vessel in a location becomes more likely when parents have blood vessel [Komurlu *et al.*, 2014]. Given initial settings, i.e. each location's value at time slice $t = 0$, we compute probabilities at following time slices until the final time slice, $T - 1$. Next, we find most probable complete observation of each time slice. Then for each most probable observation, we compute uncertainty of predictions on the final time slice. The objective is to find the earliest time slice, t^*, on which observation for each location yields an uncertainty at time slice T less than a given threshold σ.

For three different stress levels of tissue cells, we computed uncertainty at each time slice which can be seen in Figure 1. Note that the uncertainty computation is expensive and we cannot merely generate this uncertainty curve for any given initial setting. Therefore, we tried some search methods to find t^* in the search space of uncertainty and we made analytical evaluation. The reference article is hidden as it is under revision of a journal.

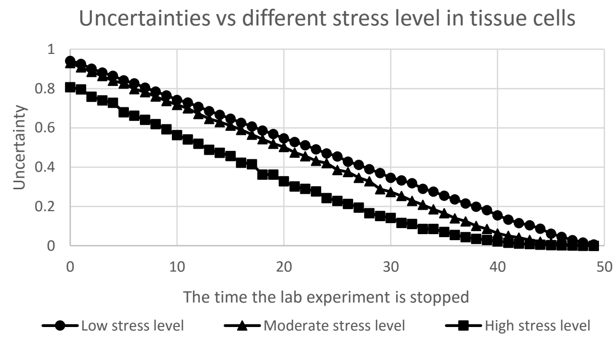

Figure 1: Uncertainty at slice T for each slice as observed.

2.2 Active Inference for Wireless Sensor Networks

Second problem we focused on is battery optimization on wireless sensor networks (WSNs). In WSNs, sensors rely on their own batteries and they consume most energy when they communicate. So our approach is, observing a subset of sensors and predicting the rest in lieu of observing all. The size of the subset is defined by a budget B. As our first objective, we resort to DBNs as predictive models. We represented each sensor as a Gaussian variable. We learned the topology from training data [Komurlu and Bilgic, 2016]. We selected Gaussian Process (GP) and Kalman filter (KF) as baseline models.

The second objective here is to choose an observation set at each time slice so as to minimize the prediction error. On our DBN, we designed a selection method, net impact-based selection (NBS), based on each variable's impact on its neighbors. We compared our selection method against random selection (RND) and sliding window selection (SW).

We used Intel Lab Data [Deshpande *et al.*, 2004]. We focused on temperature and humidity readings. Our results show that on smaller B, e.g. %10 of all sensors, GP and KF yield smaller error than our DBN as they utilize local attributes. On larger B, our DBN outperforms the baseline models. We are able to compare NBS with baseline methods on only DBN, since GP and KF are not compatible with NBS. Our results show that on all B, NBS is either the best method or ties with the best baseline. Figure 2 shows prediction errors of model-method combinations we tried.

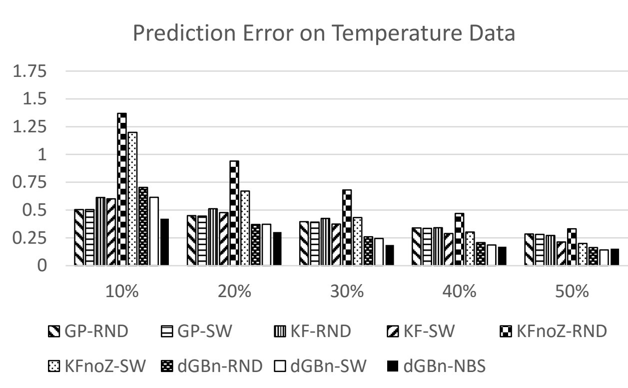

Figure 2: Error of all model-methods on B from 10% to 50%

3 Future Work

In short-term, we will work on information fusion, that is collaborating different domains in one DBN modeling. This collaboration will help exploit correlations across domains. For example, in WSN problem, we modeled temperature and humidity separately. We believe that a DBN that involves both domains can exploit dependencies between two domains and can help active inference with a larger observation space that will eventually reduce prediction error more than it does on DBNs modeling each domain separately, given same budget.

In long-term, we will focus on variance-based active inference. In the first phase, which we call maximum variance selection, we will select variables with maximum variance, at each time slice, since higher variance can result higher prediction error. Yet, computing variance of a variable in a DBN is not trivial as it requires computing all ancestors' variances and incorporating all evidences including ones provided to children. In the second phase, which we call expected variance reduction, we will try to find variables for which evidence will result maximum reduction on other variables' variances. Note that as we do not know the observed value beforehand, we need to compute expectation over all possible values that can be observed.

4 Conclusion

In the first two phases of this research, we formulated active inference for dynamic Bayesian networks (DBNs) in the context of two different problems. We showed that in tissue engineering, active inference helps detecting the optimal time to make an observation. We also showed that in wireless sensor networks (WSNs), active inference helps optimizing battery consumption. In future research, we will work on information fusion to provide more flexibility to active inference. Then we will continue with variance based active inference methods.

Acknowledgements This material is based upon work supported by the National Science Foundation under grant no. IIS-1125412.

References

[Bilgic and Getoor, 2009] Mustafa Bilgic and Lise Getoor. Reflect and correct: A misclassification prediction approach to active inference. *ACM Transactions on Knowledge Discovery from Data*, 3(4):1–32, November 2009.

[Deshpande *et al.*, 2004] Amol Deshpande, Carlos Guestrin, Samuel R Madden, Joseph M Hellerstein, and Wei Hong. Model-driven data acquisition in sensor networks. In *Proceedings of the Thirtieth international conference on Very large data bases-Volume 30*, pages 588–599. VLDB Endowment, 2004.

[Komurlu and Bilgic, 2016] Caner Komurlu and Mustafa Bilgic. Active inference and dynamic gaussian bayesian networks for battery optimization in wireless sensor networks. In *Proceedings of AAAI Workshop on Artificial Intelligence for Smart Grids and Smart Buildings*, 2016.

[Komurlu *et al.*, 2014] Caner Komurlu, Jinjian Shao, and Mustafa Bilgic. Dynamic bayesian network modeling of vascularization in engineered tissues. In *Proceedings of the Eleventh UAI Bayesian Modeling Applications Workshop*, 2014.

[Settles, 2012] B. Settles. *Active Learning*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool, 2012.