Research Statement

of Abrar Zahin (ECEE PhD Candidate at Arizona State University)

Machine learning is capable of extracting information from data, allowing machines to make high-quality predictions, and being widely used in numerous real-world applications, such as image processing, human genetics, text mining, brain networking, and recommender systems. However, most advanced machine learning approaches suffer from huge time and computational costs when operating on large-scale data. Therefore, my research is focused on developing energy and query-efficient algorithms for large-scale network systems.

Active Model Selection of Non-Tree Structured Graphical Models: In several applications, it might be costly to obtain many samples across all the variables in the underlying system. For instance, in a sensor network, obtaining one sample across all the sensors is equivalent to obtaining a synchronized measurement from all the sensors. Similarly, in neuroscience and proteomics, it might be easier to obtain data from a small subset of sensors rather than from all sensors. In all these situations, a machine learning algorithm must select some informative sensors to obtain the measurements to achieve the desired task. This kind of machine learning algorithm is known as an active learning algorithm or interactive learning algorithm. Earlier researchers have developed active learning algorithms which can handle either one of the following two challenges: (a) the existence of unobserved measurements from a fraction of sensors, and the underlying network is tree-structured, and (b) non tree-structured network, but measurements from all sensors must be observed. Our developed algorithm is the first one that can tackle both of these challenges simultaneously.

A Semi-supervised Learning Algorithm for Fall Action Detection: We developed a semi-supervised learning algorithm for fall action detection in this work. In an Internet of Things (IoT) framework, abundant data make a supervised learning algorithm inefficient, as labeling enough data is expensive and time-consuming. Furthermore, supervised learning algorithms seldom take advantage of understanding data generation processes which can be achieved only with a considerable amount of data. A semi-supervised learning algorithm uses both labeled and unlabeled data to develop an efficient learning algorithm. The main advantage of a semi-supervised algorithm lies in better understanding the data from a fundamental level which provides a more efficient and robust algorithm to a practitioner. This advantage was also prominent in our developed algorithm, which could perform better than a traditional supervised learning algorithm.

A Machine-Learning based Smart Healthcare Monitoring System: We proposed a novel and energy-efficient framework for a smart healthcare system to detect fall action. Typically, a surveillance network consists of wireless camera nodes, which generates an extensive number of images. These images are then transmitted to the processing center for detecting any possible anomalies. Due to the limited resource and computational capability, it becomes unaffordable for a camera node to maintain the stream of images for a very long time. Thus, reducing the amount of transmitted data (in an image) can be beneficial for saving energy and combating overwhelming data bandwidth. Furthermore, it is crucial to design an efficient image compression and transmission scheme to develop prolonged camera sensor networks. In order to tackle these challenges, we developed a new algorithm for the smart healthcare system by using deep learning and various techniques from signal processing. Our developed algorithm is simple to implement and of low complexity. Our algorithm if implemented will be energy efficient in various IoT environments without compromising the signal quality.

Noisy Model Selection of Non-Tree Structured Graphical Models: In this work we consider the problem of learning the structure of a networked system when underlying generated samples are noisy. Earlier researchers have considered networks which have only tree structures, and showed even though the true structure recovery is not possible, one can deduce useful information of the underlying network. Motivated by this work we considered this problem for non tree-structured networks prevalent in real applications unlike trees, and developed an algorithm for deducing useful clustering information. We showed that it is not possible for any algorithm to obtain more useful information about the network from noisy data than ours. Furthermore, we developed a novel test for identifying the latent ancestors in a non-tree structured graph. The algorithm developed in this work can be modified to develop different other important algorithms as well.

Sequential Edge Change Detection of Graphical Models: In several large network systems, it is computationally expensive to store data and process it. Hence, one is required to develop algorithms which can learn the underlying structure on the fly. That is, the underlying structure of the network is learned in a sequential manner. Furthermore, the network can go through several changes, and tracking this change in the underlying network is equally important. Motivated by these factors, we proposed a learning algorithm for recovering the structure and detecting the changes in the underlying network.

Future Research Direction: Tech giants today, like Google, Netflix, Facebook, Twitter, Apple, and IBM use various recommendation system algorithms extensively for their products. The main objective of recommender systems is to turn data on users and their preferences into predictions of users' possible future likes and interests. A recommendation system achieves this goal by understanding the underlying network of similarity among different people: which suggests how one person's interest affects others. My current Ph.D. research is focused on developing efficient algorithms for these recommendation systems. Hence, in the future, my research experience will allow me to develop algorithms that can tackle the following important challenges in any large-scale recommendation systems: learning the underlying structure when (a) only noisy samples are generated, and (b) when there is a budget constraint on the number of queries that can be obtained. There is another aspect of the recommendation system where my expertise can be helpful: the Sequential Recommendation System (SRS). SRSs treat the user-item interactions as a dynamic sequence and take the sequential dependencies into account to capture a user's current and recent preference for more accurate recommendations. Current graph-based recommendation system learning algorithm generally models user preferences in a static manner, while SRS requires capturing users' evolved and dynamic preferences. My recent work on online structure learning of underlying recommendation network structure will be an excellent fit for this scenario.