Preface

Early classification of time series is valuable in many real-world applications where data is generated over time. The aim of early classification is to predict the class label of incoming time series as early as possible before observing its complete sequence. In general, whenever early prediction time improves, the prediction accuracy decreases. In other words, one can achieve better accuracy by waiting for more data points in the series, but it will delay the response time. In time-sensitive applications, it is worth sacrificing some classification accuracy in favour of early predictions, preferably early enough for taking actionable decisions. Thus, there exists a trade-off between earliness and accuracy. However, existing approaches do not consider trade-off optimization well in their decision criteria.

Time Series Classification (TSC) is one of the major research areas that developed over the past few years, mainly due to its practical applicability in various domains such as agriculture, healthcare, medicine, finance, and industries. The main objective of TSC is to maximize prediction accuracy. In contrast, an early classification of time series has two conflicting objectives, i.e., accuracy and earliness. Nowadays, the early classification of time series attracts researchers more due to its useful applications in various domains such as early disease prediction, early gas leakage prediction, drought prediction, etc.

This thesis focuses on the problem of early classification of time series by learning optimal decision criteria. The problem of early classification has been identified as the composition of two sub-problems. The first one is to design the early classifier that can label the incomplete time series. The second is to define the decision criteria that can estimate the right time for making an online decision. Initially, we propose an early classification model for Univariate Time Series (UTS), which relies on two factors (i) a set of probabilistic classifier and (ii) a confidence threshold. The confidence threshold ensures the reliability of class prediction defined by measuring the uncertainty in predicted output. In this method, decision policy is more inclined toward accuracy and does not take trade-off optimization into consideration. In this regard, a further optimization-based approach has been adapted for early classification and defines the early stopping rules for optimal decision making, which have been learned through optimization between accuracy and earliness simultaneously.

Furthermore, this optimization-based approach has been extended for Multivariate Time Series (MTS), which is more challenging than UTS because of the multiple variables involved in decision making. An ensemble-based system has been designed to label the incomplete MTS, and collective output from all the variables has been utilized for decision making. These proposed methods are highly effective for small training data sets, but feature transformation is required for training the classifiers. Finally, a deep learning-based hybrid classifier has been proposed that can capture the temporal information from the raw sensory data effectively to perform the classification task. Moreover, the optimal confidence threshold has been defined by balancing the trade-off between accuracy and earliness. The proposed approaches have been evaluated on publicly available datasets and they demonstrated effective solutions for early classification on time series.