Jiasheng-Li_SYS 6014 20sp_Final Proposal

March 25, 2020

1 Overview

1.1 Background

There are many approaches to analyze time series data, including time domain, frequency domain and deep learning techniques. Once the patterns in the time series are found, the forecasting and prediction of future time series in a given timestep can be achieved based on the patterns found. However, in many researches, authors emphasized on the precision and other criteria that judge the performance of the learning model. And little attention has been paid on the application, including which the decision analysis, of the predicted sequence. Therefore, this project aims to apply decsion analysis based on a Long-short Term Memory(LSTM) network to solve decision making problems given specific data set.

2 Problem Formulation

Based on the data set and specific problem, the decision making can be categorized as two type, each of which can be described using similar mathematiocal representations. This kind of decision making is called Naturalistic Decision Making, which assumes that there is an optimal solution. The two types are: maximize profit, or minimize the potential risk. In this project, **minimize potential risk** will be used as an example. And when data sets are found and determined, this notebook will be updated. ## Context Considering a machine time series that describes the operation state using many attributes indexed by date. And assume that some attributes are related to the operation state: Malfunction or Operational, and that the patterns of such time series can be learned by LSTM. Therefore the expected output of the prediction model will be a predicted series that indicates the future operation state of the machine. Noticed by the potential risk, the decision makers need to decide actions to minimize the loss based on cost, payoffs, prior beliefs and updated information. ## Action Set

$$\mathbb{A} = \{A_s : stopoperation, A_c : continueoperation\}$$

State Space If the oepration state attribute can be treated as a random variable, then state space can be described as:

$$\mathbb{X} = \{X_n : Operational, X_f : malfunction\}$$

Data generation process The actual data distribution will be updated based on data set. The predicted time series will be generated by LSTM. ## Parameter space

$$\Theta = \mathbb{R} \times [argmin\{attributes\}, argmax\{attributes\})$$

for each of attributes of data set. ## Prior beliefs In this project, the prior belief is the congestion of potential malfuntion happending date based on previous data set. For application use, the precision might use 'day' as unit. ## Information Update This procecss will be achieved using the predicted time series of LSTM. ## Payoffs It can be described as: the expected payoff can be described as the economic based on the decsions made by the given potential risk information. For example: if the decision maker choose to continue operating the machine, because he think the probability of the actual malfuntion is small, then there will be a loss.

$$\mathbf{E}(A_s|X_n) < 0, \mathbf{E}(A_s|X_f) = 0, \mathbf{E}(A_c|X_n) = 0, \mathbf{E}(A_c|X_f) < 0$$

Utility Funtion

$$\mathbf{U} = \sum \mathbf{E}(A|X) \times P(A|X)$$

[ ]: