2018 CSE3/5CI Computational Intelligence Assignment

Proposed by A/Prof. Justin D. Wang

This assignment contributes **30%** of your overall marks for students enrolled in CSE3CI, and **20%** of your overall marks for students enrolled in CSE5CI. Please read this sheet carefully before doing your assignment.

Summary: The assignment aims at consolidating your knowledge base and developing practical skills to build a fuzzy system for forecasting the electricity price. The task is formulated as a time-series prediction problem for business application, and the goal is to model the behaviour of underlying dynamics of the electricity market. In principle, the merits of such a fuzzy forecasting system can be evaluated by two aspects:

- **The number of fuzzy rules in the rule-base and the number of variables used in the antecedent part of the fuzzy rules (the smaller the better);**
- **System performance in terms of the accuracy (the smaller the better), i.e., the average relative error between your fuzzy system outputs and the actual outputs for both the training data set (learning capability) and the test data set (generalization capability).**

This is a **GROUP-based** assignment (the number of the group members: 1~3, no more than 3 please; all group members will receive a same marks) and for both 3rd and 5th year students. You are **NOT** permitted to work as a mixed-group (i.e., all group members must be from the same grade) when completing this assignment. The length of the assignment report is about 1200 words with codes used in Python.

Copying, Plagiarism: Plagiarism is the submission of somebody else's work in a manner that gives the impression that the work is your own. The Department of Computer Science and Information Technology at La Trobe University treats plagiarism very seriously. When it is detected, penalties are strictly imposed.

Date due and late submission policy: May 11, 2018 (Friday)

- All assignments are due at **10:00 am**.
- A penalty of 5% per day will be imposed on all late assignments up to 5 days. An assignment submitted more than five working days after the due date **will NOT be accepted and zero mark will be assigned**.
- Assignment without the signed declaration of authorship attached **will NOT be accepted and zero mark will be assigned.**
- Students will **not** be granted an extension of the assignment deadline. Students are requested to submit an application for special consideration through Student Centre. In addition, students are advised to submit whatever incomplete work they have already done for the assignment.

Where to Submit: Your assignment report (hardcopy) is to be submitted at a labelled box opposite to BG 139 lab.

Problem Description (A Fuzzy System for Forecasting Electricity Price)

Develop a fuzzy forecasting system for data analysis using Python. The system performs a forecasting task for power marketing price. The data used in this assignment is from the real world (Queensland, Australia), and it has been split up into two parts, i.e., a training dataset which will be used to build your fuzzy forecasting system, and a testing dataset which will be used to evaluate your system performance in terms of generalization capability. The data sets can be downloaded from the Assignment directory in LMS.

Related Concept

Outliers: Roughly, an ***outlier*** is an observation that lies an abnormal distance from other values in a random sample from a population. You can read more about this concept via the links below:

http://www.itl.nist.gov/div898/handbook/prc/section1/prc16.htm
http://mathworld.wolfram.com/Outlier.html

Average Relative Error: A metric for measuring forecasting systems performance is defined by:

$$RErr = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\text{SystemOutput}_i - \text{TargetOutput}_i}{\text{TargetOutput}_i}\right|, \qquad (1)$$

where *N* is the cardinality of the data set.

System Inputs and Outputs

Let the *temperature* and *total demand of electricity* at time instant ***t*** be *T(t)* and *D(t),* respectively. The goal of the fuzzy forecasting system is to predict the RRP price by using some historical data as system inputs. In this assignment, the historical data set used for building the fuzzy system at time instant ***t*** is composed of a subset of the set M={*T(t-2), T(t-1), T(t), D(t-2), D(t-1), D(t)*}. The output of your system at time instant ***t*** is a forecasting value of the Recommended Retail Price (RRP) of electricity at the next time instant ***t+1***, denoted by *P(t+1)**.***

Note that you should select a subset of the set M as the system's input variables by using correlation analysis.

Tasks Description (the maximum marks for each item below is 20)

- Remove outliers of the output variable from the datasets (both training and test), and give a list of the outliers; and then rebuild the training and the test datasets;
- Select appropriate values or fuzzy subsets for linguistic variables used in your fuzzy rules;

- List the fuzzy rules that are generated by using statistical analysis (correlation coefficients) with heuristics;
- Implement your fuzzy system in Python, where all membership functions involved in your system should be plotted clearly;
- Report your system performance in terms of the average relative error for both training and testing datasets, and analyze the effects of membership functions and defuzzification methods.

Remarks

- Assessment will be done by looking at the average relative prediction accuracy for both the training data set and the test data set.
- Either Mamdani-type or Sugeno-type fuzzy rules can be applied.
- Your report should provide a full list of Python codes used in your system with some graphical illustrations. It will be appreciated to show some fine-tuning of the system's parameters to produce sensible results. It is encouraged to appropriately use appendices to detail your results.

Assessment Criteria

(100-80 marks) - An excellent, well-written report. You have produced a working system that produces sensible results. The report summarises the approach taken well. You have analysed the performance of the system and presented the results in an interesting and sound way. A thorough and systematic analysis of the effect of different membership functions and different defuzzification techniques is presented.

(79-60 marks) - A well-written report. You have produced a working system that produces good results. You have exhibited some initiative in the approach taken and the results are presented clearly. An analysis of the effect of different membership functions and different defuzzification techniques is presented.

(59-40 marks) - A reasonable report that presents an account of the approach taken and the final system. The system performs reasonably well and the results are presented reasonably clearly. Either different membership functions or different defuzzification techniques have been explored.

(39-20 marks) - A report that presents some results of a working system. Demonstrating some understandings on fuzzy forecasting system design.

(19-0 marks) - Either no report submitted or a report that shows little or no understanding of how to develop a fuzzy system.

- End of Assignment Paper -