PROTOTYPE OF STUDENT DEGREE CLASSIFICATION FORECASTING USING FUZZY LOGIC

ROSYUSJULIZA YUSOFF

A report submitted in partial fulfilment of the
requirements for the award of the degree of
Bachelor of Computer Science (Software Engineering)

Faculty of Computer Systems & Software Engineering
University College of Engineering & Technology Malaysia

NOVEMBER, 2006

ABSTRACT

Degree Classification normally classified after the student finished their study. The student degree classification forecasting is important to classify student degree class before the learning process ended. Therefore, the "Prototype of Student Degree Classification Forecasting (POSDC)" is the suggestion prototype that will solve the manual system used currently. The purpose of this prototype is to forecast student degree class by their student id or name. The forecasting result calculated based on their previous CGPA and credit hours countered for each semester. The student will be classified to 4 classes which are First Class Honours, Upper Second-Class Honours, Lower Second-Class Honours and Third Class Honours. There are two users of this prototype which are Academic Advisor and HEA staffs. The technique that used for this prototype is Fuzzy Logic Technique which is fuzzy rules because it is the most suitable technique applied for forecasting and the result shows that fuzzy rules also is a good technique applied for degree classification forecasting.

ABSTRAK

Kelas pengijazahan biasanya ditetapkan selepas seseorang pelajar menamatkan pengajian masing-masing. Ramalan Kelas Pengijazahan pelajar amat penting untuk mengklasifikasikan pelajar mengikut kelas tertentu sebelum proses pembelajaran tamat. Dengan ini, "Prototype of Student Degree Classification Forecasting Using Fuzzy Logic Technique" adalah prototaip yang dicadangkan untuk menyelesaikan masalah semasa yang menggunakan cara manual. Tujuan ramalan kelas pengijazahan ini adalah untuk meramal kelas pelajar berdasarkan pengenalan pelajar atau nama pelajar. Pelajar akan diklasifikasikan kepada empat kelas iaitu Kepujian Kelas Pertama, Kepujian Kelas Kedua Tinggi, Kepujian Kelas Kedua Rendah dan Kepujian Kelas Ketiga. Pengguna prototaip ini terbahagi kepada dua iaitu penasihat akademik dan kakitangan pejabat hal ehwal akademik.Prototaip ini menggunakan teknik "Fuzzy Logic" iaitu "Fuzzy Rules" kerana ia adalah satu teknik yang amat sesuai untuk proses ramalan dan keputusan kajian juga menunjukkan "Fuzzy Rules" juga merupakan satu teknik yang sesuai untuk ramalan kelas pengijazahan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

POSDC	-	Prototype of Student Degree Classification
HEA	-	Hal Ehwal Akademik
CGPA	-	Cumulative Grade Pointer Average
GPA	-	Grade Pointer Average
BCS	-	Bachelor of Computer Science
FSKKP	-	Fakulti Komputer Sains dan Kejuruteraan Perisian
KUKTEM	-	Kolej Universiti Kejuruteraan & Teknologi Malaysia
MAD	-	Mean Absolute Deviation
PC	-	Personal Computer
FL	-	Fuzzy Logic
GA	-	Genetic Algorithm
IIS	-	Internet Information System

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

In this chapter, an introduction to prototype of student degree classification will be presented and followed by the problem statement, the objective and the scope of the project. It will describe the overview of "Prototype of Student Degree Classification Using Fuzzy Logic (POSDC) and the problem statement regarding the system. It will briefly explain about the issues caused by the current system. Besides that, the scopes of the project and objective will detail its reason of development.

1.1.1 Forecasting System

Forecasting is a process of generating information for the future development of process of from the past data and its present development. The forecasting system developed for the future based on previous data or information. Forecasting is important nowadays because people can make planning based on the forecasting's result. There are several types of prediction used today such as weather forecasting, financial forecasting and many more. By using prediction system, people can make planning for the future situations or risk.

1.1.2 Student Degree Classification Forecasting

Student Degree Classification Forecasting is an automatic system to forecast the degree classification whether they can graduate as a First Class Honours, Upper Second-Class Honours, Lower Second-Class Honours or Third Class Honours.

This prototype will forecast the Student's degree class based on their previous CGPA for each semester. The Prediction will be done by individual (student's name). The purpose of this prototype is to help HEA department and Academic Supervisor regarding student progress, so they can take an action to help the students to produce better result. As a benchmark, this prototype will use the previous BCS student's result.

1.2 Problem Statement

The next sub chapters represent the current system of degree classification in KUKTEM, the process of the current system and the problems faced by the current systems.

1.2.1 The Current System

Prototype of Student Degree Classification Forecasting (POSDC) is not having yet in KUKTEM. KUKTEM used the manual system in order to classify their student's degree class. It means that, the student need to wait until they finished their study to know their degree class.

1.2.2 Process of Current System

The entire student's result for each semester will store in the database and the students need to finish their study in order to know their degree class. The student will classify into 4 classes which are First Class, Upper Second-Class Honours, Lower Second-Class Honours and Third Class.

1.2.3 Problem of Current System

a) The system cannot predict the results earlier and take a long time to know the result

b) The management cannot help the students to improve their results because they did not know the result earlier.

1.3 Objective

The objectives of this prototype are:

a) To forecast the student's degree class whether they will grade as First Class, Upper Second-Class Honours, Lower Second-Class Honours or Third Class Honours.

b) To forecast the student degree class based on their previous result

c) To forecast student's degree class by student id and name..

1.4 Scope

This prototype limited to Academic Supervisor and HEA staffs. All the required data or information for this project gathered from Academic Affair Department (HEA). As the benchmark, this project will use BCS first batch student results because they are the first graduate produced by KUKTEM. The project

developed using ASP.NET for the interface and code. While the Microsoft SQL server 2003 is used for the database.

1.5 Thesis Organization

This thesis consists of 5 chapters and each chapter is devoted to discuss different issues in this project. Below is a summary of the content for each chapter.

a) Chapter 1

Introduction to the project is presented along with the project's problem statement, objectives and scope of the project.

b) Chapter 2

Research and literature review related to the project presented.

c) Chapter 3

Project analysis, design and methodology are presented.

d) Chapter 4

Result from the testing of the system is presented along with statistic and discussion of the result.

e) Chapter 5

Summary of the project is presented.

CHAPTER 2

LITERATURE REVIEW

This chapter presents the literature reviews which discuss the degree classification issues and the fuzzy logic techniques. It also describes about KUKTEM degree classification and the advantages of using fuzzy logic than other technique.

2.1 Student Degree Classification

Nowadays, nearly all students sit their degree for honours. Student degree classification produced to classify the degree's student to four categories which are First Class Honours, Upper Second-Class Honours, Lower Second-Class Honours and Third Class Honours.

Most Universities prohibit such a student from receiving honours and award a class degree based on average mark of the assessed work of candidate has completed. Every country used different scheme in order to classify the student's degree class.

a) First Class Honours(First or 1^{st})
b) Upper Second-Class Honours(2:1)
c) Lower Second-Class Honours(2:2)
d) Third Class Honours (Third or 3^{rd})
e) Pass without Honours (ordinary degree)

f) Fail (no degree awarded)

2.1.1 KUKTEM Degree Classification

Table 2.1 below show the KUKTEM's student possible diploma classification while the Table 2.2 shows the Degree Classification scheme in KUKTEM [1]

Table 2.1: Table Diploma Classification

Diploma	**Conditions**
First Class	CPA > = 3.67
Second Class	2.00 <= CPA < 3.67

Table 2.2: Table Degree Classification

Bachelor	**Conditions**
First Class	CPA > = 3.67
Second Class (Upper)	3.00 <= CPA < 3.67
Second Class (Lower)	2.33 <= CPA < 3.00
Third Class	2.00 <= CPA< 2.33

2.2 Forecasting

Forecasting is the estimation of the value of a variable (or set of variables) at some future point in time. In this note we will consider some methods for forecasting. A forecasting exercise is usually carried out in order to provide an aid to decision-making and in planning the future. Forecasting methods can be classified into several different categories [2].

a) Qualitative methods where there is no formal mathematical model, often because the data available is not thought to be representative of the future (long-term forecasting)
b) Regression Methods is an extension of linear regression where a variable is thought to be linearly related to a number of other independent variables.
c) Multiple equation methods where there are a number of dependent variables that interact with each other through a series of equations (as in economic models).
d) Time series methods where we have a single variable that changes with time and whose future values are related in some way to its past values.

2.2.1 Criteria For Selecting Forecasting Method

Potential Rules for selecting a time series forecasting method. Select the method that gives the smallest bias as measured by cumulative forecast error, gives the Smallest Mean Absolute Deviation(MAD), gives the smallest tracking signal, supports management's belief about the underlying pattern or demand [3].

2.2.2 Focus Forecasting

Focus forecasting" refers to an approach to forecasting that develops forecasts by various techniques, and then picks the forecast that was produced by the

"best" of these techniques, where "best" is determined by some measure of forecast error [4].

2.3 Techniques in Artificial Intelligence

The next sub chapters discussed the several techniques applied in artificial intelligence nowadays.

2.3.1 Fuzzy Logic

The concept of Fuzzy Logic was conceived by Lotfi Zadeh, a professor at the University of California at Berkley and presented not as a control methodology, but as a way of processing data by allowing partial set membership rather than crisp set membership or non-membership

Fuzzy logic is the most popular new technique that had been used nowadays. Fuzzy logic has rapidly become one of the most successful of today's technologies for developing sophisticated control system [4].

People do not require precise, numerical information input, and yet they are capable of highly adaptive control. If feedback controllers could be programmed to accept noisy, imprecise input, they would be much more effective and perhaps easier to implement [5].

Fuzzy Logic is a problem-solving control system methodology that lends itself to implementation in systems ranging from simple, small, embedded micro-controllers to large, networked, multi-channel PC or workstation-based data acquisition and control systems. It can be implemented in hardware, software, or a combination of both. FL provides a simple way to arrive at a definite conclusion

based upon vague, ambiguous, imprecise, noisy, or missing input information. Fuzzy Logic's approach to control problems mimics how a person would make decisions, only much faster

Fuzzy Logic Notation and operations are based on classical logic and the propositional calculus, the modern form of notation for classical logic [6].

Table 2.3: .Truth Table for AND and OR Logical Operators

P	Q	P AND Q	P OR Q
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	1

The Fuzzy Set Theory has been used in several areas of human knowledge as the link between the imprecise (subjective) models of the real world and their mathematical representations [4]. Each fuzzy number represents an importance degree and uses a linguistic term as a meaning.

2.3.2 Genetic Algorithm

Genetic Algorithms is a specific type of optimization on techniques that are based on a set of heuristics and involve a non-trivial and careful setting of a number of parameters. GA's are also known to help solve complex, non-linear problems that often lead to cases where the search space shows a curvy, noisy "landscape" with numerous local minima. A category of GA's is designed to address so-called routing or scheduling problems. [4].

2.3.3 Expert Systems

Expert systems are systems capable of offering solutions to specific problems in a way and level comparable to experts. Expert systems also the programs which emulate human expertise in well-defined problem domains. Expert systems have been applied to solve problems in various domains. Expert systems are the problem solving programs that mimic the way human expert reason. Expert systems are a revolutionary approach to integrating computers into the management decision making process

Expert system maybe forwarded or back warded chaining. In forward chaining systems reasoned from antecedent truth to consequent truth while back warded chaining reversed this.

Modularization of expert system programs is essential for both their construction and their debugging, but is even more important conceptually. At the lowest level modularization is accomplished by assigning rules to block [6].

2.4 Selected technique for POSDC

From four techniques that mentioned above, the fuzzy logic technique is the good technique that can apply to the POSDC. This is because the fuzzy logic technique addresses such applications perfectly as it resembles human decision making with an ability to generate precise solutions from certain or approximate information [4].So, fuzzy logic is the suitable technique to apply in POSDC.

2.4.1 Fuzzy Logic Concept

Fuzzy system represents two distinct methodologies that deal with uncertainty. In Fuzzy systems, the reverse situation prevails. The input and output

variables are encoded in "fuzzy" representations, while their interrelationships take the form of well-defined if/then rules. Fuzzy logic systems address the imprecision of the input and output variables directly by defining 'then' with fuzzy numbers that can be expressed in linguistic term as cold, warm and hot. Furthermore they allow far greater flexibility in formulating system descriptions at the appropriate level of detail. Fuzziness has a lot to do with the parsimony and hence the accuracy and efficiency of a description. This means that complex process behavior can be described in general terms without precisely defining the complex phenomena involved [7].

Fuzziness is a feature of their representation in a milieu of symbols and is generally a property of models, computational procedures, and language.
In fuzzy approaches, relations possess the computational potency and significance that functions possess in conventional approaches. Fuzzy if/then rules and their aggregations, known as fuzzy algorithms, both of central importance in engineering applications, are fuzzy relations in linguistic disguise. They offer and alternative and often complementary language to conventional (analytic) approaches to modeling systems (involving differential or different equations). They encode knowledge about a system in statements of the form as if (a set of conditions) and then (a set of consequences can be inferred).

2.4.2 Fuzzy Rules

A set of fuzzy rules can be learned from data that represent either a no chaotic or a chaotic process. The rules can than be used to model this process and control it. One of the advantages of using fuzzy rules is that they express in a linguistic form the process which may lead to time series with the use of fuzzy logic techniques can be achieved by articulating fuzzy rules by experts after observing past data and using their experience, using the rules in a fuzzy inference machine; and by extracting fuzzy rules from data and their consecutive use in a fuzzy inference machine [4].

Both approaches may include refining the rules and choosing an appropriate inference method for better processing. The approaches are illustrated in case examples below.

Example 1

Predicting water flow to a sewage plant by using rules articulated by experts. Through analyzing smaller portions of the data about the problem and analyzing the "fuzzy graph" of more data points. The fuzzy labels and their membership functions were defined in advance. These rules are now used in a fuzzy inference machine to predict the water flow on test data.

Figure 2.1 show the fuzzy rules for the water flow to sewage plant prediction problem that articulated by experts.

```
IF      TIME VERY EARLY
        PREVIOUS VERY LOW
THEN NEXT VERY LOW

IF      TIME EARLY
        PREVIOUS LOW
THEN NEXT MEDIUM

IF      TIME MIDDAY
        PREVIOUS MEDIUM
THEN NEXT HIGH

IF      TIME LATE
        PREVIOUS MEDIUM
THEN NEXT LOW

IF      TIME VERY LATE
        PREVIOUS LOW
THEN NEXT LOW
```

Figure 2.1 Fuzzy Rules for the water flow to sewage plant prediction problem articulated by experts.

Near optimal values for the number of fuzzy labels and their membership functions are found experimentally. The following step described the experiment:

a) Define the number l of previous values (lags) for the time-series data to be used for predicting k future values; here $l = 9$ and $k = 1$ is used.

b) Define the number of fuzzy discretization intervals and their membership functions; here 29 fuzzy labels are used for each of the 10 input and output values.

c) Extract fuzzy rules; for this purpose 700 out of 1000 values are used and the second method from subsection. [8]

2.4.3 Time-Series Prediction

Prediction of time-series events is called time-series prediction. When a prediction is done on the basis of only one independent variable it is called a multivariate prediction. A general form of a heuristic rule for time-series prediction is:

IF (previous time-moment values for the features from the feature space are d), THEN (next time-moment values for the predicted variable(s) will be s)

2.4.4 Advantages of using Fuzzy Logic

Fuzzy Logic offers several unique features that make it a particularly good choice for many control problems [5].

a) It is inherently robust since it does not require precise, noise-free inputs and can be programmed to fail safely if a feedback sensor quits or is destroyed. The output control is a smooth control function despite a wide range of input variations.

b) Since the Fuzzy Logic controller processes user-defined rules governing the target control system, it can be modified and tweaked easily to improve or drastically alter system performance. New sensors can easily be incorporated into the system simply by generating appropriate governing rules

c) Fuzzy Logic is not limited to a few feedback inputs and one or two control outputs, nor is it necessary to measure or compute rate-of-change parameters in order for it to be implemented. Any sensor data that provides some indication of a system's actions and reactions is sufficient. This allows the sensors to be inexpensive and imprecise thus keeping the overall system cost and complexity low.

d) Because of the rule-based operation, any reasonable number of inputs can be processed (1-8 or more) and numerous outputs (1-4 or more) generated, although defining the rule base quickly becomes complex if too many inputs and outputs are chosen for a single implementation since rules defining their interrelations must also be defined. It would be better to break the control system into smaller chunks and use several smaller Fuzzy Logic controllers distributed on the system, each with more limited responsibilities.

e) Fuzzy Logic can control nonlinear systems that would be difficult or impossible to model mathematically. This opens doors for control systems that would normally be deemed unfeasible for automation

2.4.5 Difference between Fuzzy Logic and Conventional Control Method

Fuzzy Logic incorporates a simple, rule-based IF X AND Y THEN Z approach to a solving control problem rather than attempting to model a system mathematically. The FL model is empirically-based, relying on an operator's experience rather than their technical understanding of the system. For example, rather than dealing with temperature control in terms such as "SP =500F", "T <1000F", or "210C <TEMP <220C", terms like "IF (process is too cool) AND (process is getting colder) THEN (add heat to the process)" or "IF (process is too hot) AND (process is heating rapidly) THEN (cool the process quickly)" are used [5].