Philippine Engineering Journal (1999) XX (1): 1- 8

EEE STUDENT MORTALITY-PERFORMANCE ANALYSIS

Gay Marie C. Denoga
Instructor
Department of Electrical & Electronics Engineering
College of Engineering
University of the Philippines
Diliman, Quezon City, Philippines

ABSTRACT

This is a study of student performance/mortality in the EEE curriculum for the UP EEE Classes of 1990 to 1996 based on student grades in the DEEERS database incurred within the time frame from 1991-1997 using Digital Signal Processing techniques, for the purpose of obtaining a measure of, and predicting a student's course completion capability.

I. Introduction

The Department of Electrical & Electronics Engineering (EEE) currently has under its tutelage approximately 1600 students, from freshman to senior level, in its three undergraduate programs alone. Each year, the department accepts 245 students to its undergraduate programs. Not all of these 245 students will make it to their senior year, and even less will graduate.

Most delinquent students fall behind in their studies early on in the course. Prior to dismissal, their performance is assessed by the EEE Department's Readmission Committee for any indication that the student is still capable of completing the course within the bounds of residency that will merit a second chance.

The number of students drops dramatically past the sophomore year, which leads us to presume that performance in major subjects in the first two years of the course is critical. To predict students' course completion capability, grades in particular subjects and the number of times that subject is taken are considered.

Delinquent students have far more retakes and lower grades than those who graduate. The Department of EEE Registration System or DEEERS database uses a numerical grading system with a common range of 1 to 7. '1' is the highest, '3' is pass, '4' is conditional, '5' is fail, '6' is dropped, and '7' is incomplete.

A three-dimensional view of the available data is shown. In Figure 1, the y-axis corresponds to the highest grade obtained by a student in that subject, and corresponds to the frequency that a subject is taken in Figure 2. The bottom left axis for both figures indicates the ith student arranged in ascending student number and batch, and the bottom right axis indicates the jth subject taken in the course. A grade of '0' means that no grade is available.

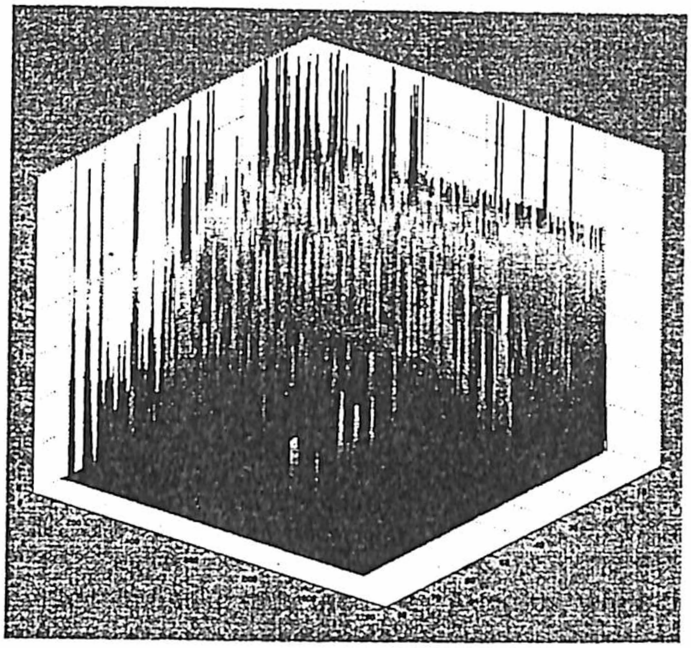

Figure 1. Grades of students in each subject

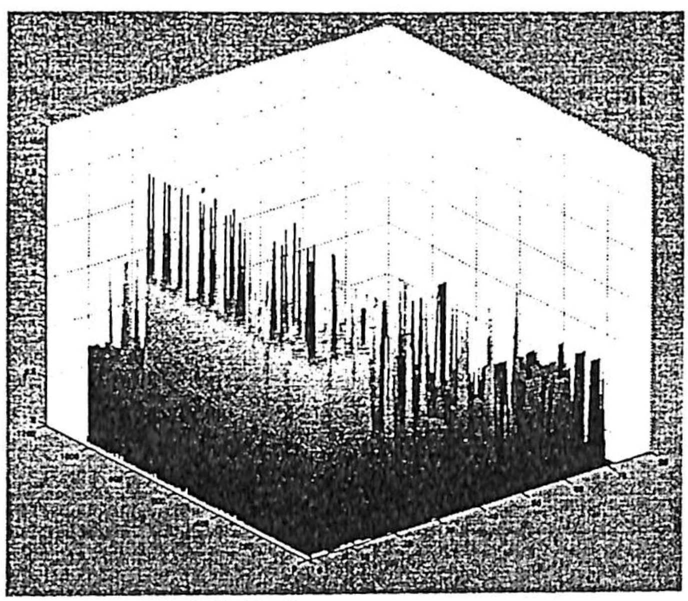

Figure 2. Frequency that a subject is taken

The grades that a student incurs over time can be viewed as a sequence of points that make up an aperiodic digital signal. The frequency content of this time-domain signal can be obtained using the following Fourier transform.

$$X(\omega) = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n} \tag{1}$$

It has a corresponding energy density spectrum:

$$S_{xx} = |X(\omega)|^2 \tag{2}$$

II. Methodology

In order to find a trend, low-pass and high-pass filters with varying cut-off frequencies are applied to the signal.

To implement the low-pass filter, we sample the signal in the frequency domain and take only the first *n* samples, while the last *n* samples are considered for the high-pass filter. Filter cut-off frequency varies as *n* varies, while the total number of samples taken is N=512. The corresponding total normalized energy contained in this set of samples is then computed using the following equation.

$$\text{Total normalized energy for the } n \text{ samples } \{X_1, X_2, \ldots X_n\} = \sum_{i=1}^{n} |X_i|^2 \Big/ \sum_{k=1}^{512} |X_k|^2 \tag{3}$$

A comparison of the total normalized energy obtained from the grades of a student who graduated versus that of a student who did not graduate shows that the former is generally higher than the latter when a low-pass filter is used. This is because a student with consistently good grades will have a corresponding signal with a range of 3 and below . The occurrence of failures result in low-high-low transitions, increasing the high frequency content of the signal and decreasing the total normalized energy found in the passband of the low-pass filter.

This allows us to identify potentially successful students from delinquent ones provided the delineating or threshold value of total normalized energy generally applies to the entire EEE student population.

This threshold value is more apparent when the original time-domain signal is encoded prior to obtaining the total normalized energy.

The cipher of 12-bit sequences in Table I is used. The sequence contains few or no transitions for good grades, and an increasing number of transitions for lower grades

Table 1
Cipher

111111111111	1
111111000000	1<x<2
111000111000	2≤x<3
110011001100	=3
101010101010	>3

The plots in Figures 3 and 4 show the frequency response and the power spectrum of each grade code.

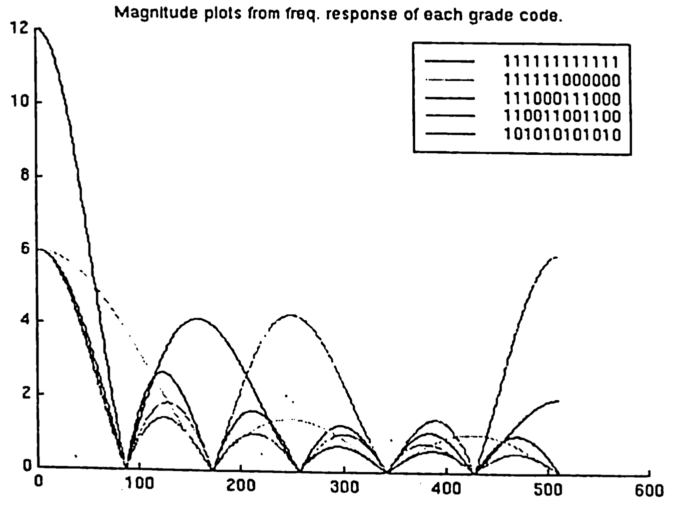

Figure 3. Magnitude plots from frequency response of each grade code.

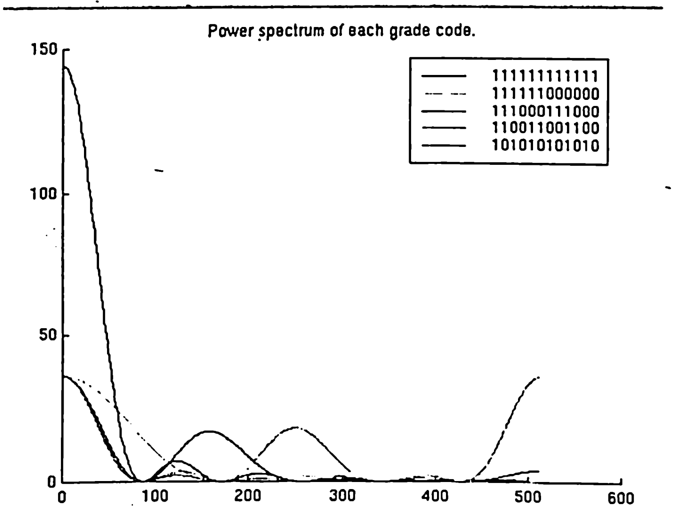

Figure 4. Power spectrum of each grade code.

III. Results

Encoding results in an increase in the dc and low frequency components of signals pertaining to students with good grades and the opposite for students with poor grades.

Predictions on the course completion capability of each student are based on the total normalized energy contained in each set of *n* samples with respect to a threshold. If the value is higher than the threshold, the student is predicted to graduate.

Such predictions are compared with actual graduation data. In Figure 5, a '1' on the y-axis indicates that the student graduated, while a '0' indicates that the student did not graduate. The x-axis indicates the ith student in the batch.

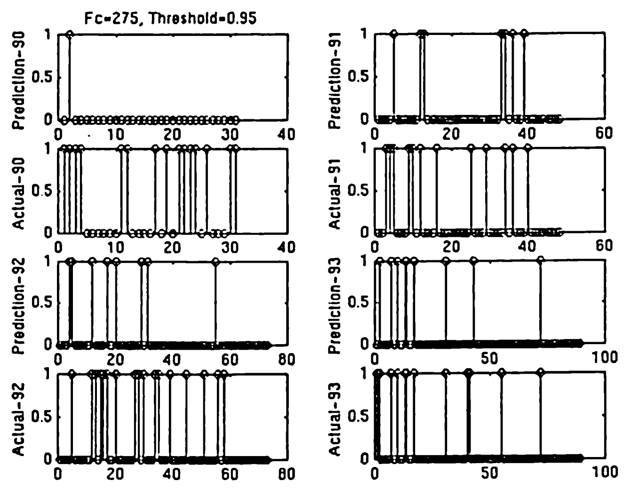

Figure 5. Predictions vs. Actual Data

Accuracy is computed by counting the number of correctly predicted outcomes (*i.e.* the student predicted to graduate did graduate, and the student predicted to fail did not graduate), then dividing by the total number of outcomes.

A sweep of all possible *n* and threshold levels for each batch of students yields the optimal *n* and threshold level that provides an acceptable degree of prediction accuracy that is consistent for all batches.

Using $n = 275$ and total normalized energy in those *n* samples ≥ 0.95, a prediction accuracy of 70-to-94% is achieved for Batch 1991 to 1993, and only 55% for Batch 1990. In Figure 6, graphs at x = 1,2,3 and 4 correspond to results for Batch 1990, 1991, 1992, and 1993 respectively.

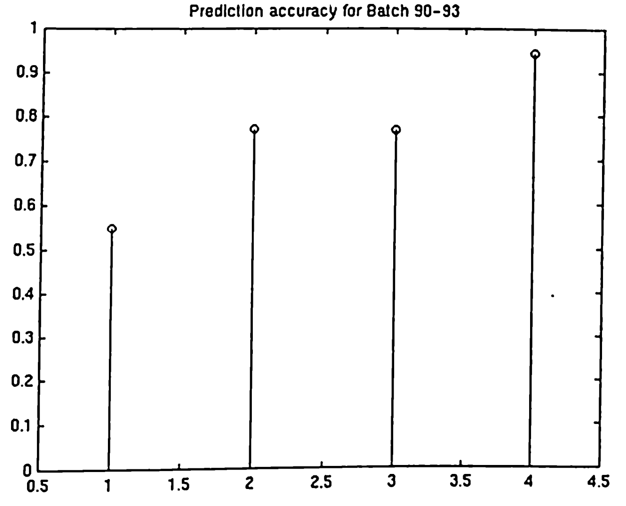

Figure 6. Prediction accuracy for Batch 90-93

If we use the method to predict only those students who will graduate on time, prediction accuracy for Batch 1990 increases to 100%.

IV. Conclusions

As the number of subjects involved in the prediction process increases, the accuracy of the predictions generally increases. The periods with the most significant improvements in accuracy support the premise that student performance in subjects for the first two years are critical in determining whether the student can handle the course or not.

Using a filter that considers only the subjects taken during the first two years of the course, the best cut-off is found to be at the 275th sample point. The optimal threshold level is 0.95, or at least 95% of the total power of the signal obtained from the student's grades must be within the passband of this filter. Prediction accuracy of this filter is above 50% for all four batches, and generally increases as the threshold level increases. This attests to the reliability of the filter.

Results of this study may improve with the use of a larger data set in the search for an optimal cut-off frequency and threshold level. I also recommend that alternative encoding schemes be found.

References

1. Ingle, Vinay F. & Proakis, John G. (1997). *Digital Signal Processing Using Matlab*, Vol.4. PWS Publishing Company, Boston.
2. Lunneborg, Clifford E. (1994). *Modeling Experimental and Observational Data*, Wadsworth Publishing Company, California.
3. Proakis, John G. & Manolakis, Dimitris G. (1996). *Digital Signal Processing: Principles, Algorithms, and Applications,* 3rd edition, Prentice-Hall, Inc., New Jersey.
4. Taylor, John Keenan (1990). *Statistical Techniques for Data Analysis,* Lewis Publishers, Inc., Florida.
5. Thompson, James R. (1989). *Empirical Model Building,* John Wiley & Sons, Inc., New York.