TEACHING PATTERN RECOGNITION CONCEPTS USING WEB BASED TECHNIQUES

Ashraf A. Kassim[1], and Chun-Meng Lau[2]

Abstract *¾ The Internet presents both an opportunity and a challenge to educators: an opportunity to expand the domain of education, and a challenge to develop and implement newer teaching methodologies based upon discovery and cognitive learning. This paper presents our work which involves development of interactive tools that demonstrate certain key concepts of pattern recognition theory for graduate students. The goal is to help students to achieve a good level of understanding of the concepts in a short time and to increase the interest in the subject matter.*

Index Terms ¾ Intelligent Tutoring System, Pattern Recognition System, Web-based Learning.

INTRODUCTION

Highly interactive learning environments that are independent of time and space constraints can now be realized through computer and Internet technology. Distance learning is becoming more and more important for several reasons. More people are interested in higher education [1] and educational institutions have been increasingly turning to the Internet to deliver their services to this group of people [2]. Many of these institutions also supplement regular classroom teaching with additional web based material. Key features of computer and Internet-based learning environments [3] include: interactivity, global accessibility, availability of online resources, learner-controlled pace, convenience, non-discriminatory, cost effective, collaborative learning, online evaluations and etc.

Individual tutoring is the most effective mode of teaching [4] but it poses financial and logistical problems. Also, in engineering education, a major focus is on developing problem solving skills. Although doing exercises given in textbooks enable students to gain insight into how to solve problems, this lacks interaction and feedback. Web-based intelligent learning environments have the potential of bringing the individual tutoring experience to a broader audience.

At the National University of Singapore (NUS) the *Integrated Virtual Learning Environment* or IVLE [5], provides the tools, resources and techniques for group communications, interaction and integrated learning. IVLE provides each course with “websites” for posting course materials, assignments, discussion forums, and many other tools to enable staff and students to interact over the internet.

“Pattern Recognition” [6] is a course offered at the graduate level. Although the lecture materials have been available on IVLE, we wanted to include interactive presentations of certain key pattern recognition concepts with the objective of helping students to better and more accurately understand the concepts, and to increase interest in the subject. Books often provide graphic illustrations that are static but interactive visualizations help to improve the understanding of difficult concepts and ideas.

In this paper, we present our web-based learning environment for introducing key pattern recognition concepts. Our system utilizes information, multimedia and Internet technologies to provide individualized attention and enhance student learning and problem solving skills.

BACKGROUND

In engineering education, the understanding of key concepts that forms the basis of a subject can be strengthened through demonstration of the key concepts through practical experimentation, and interactive simulations.

Enhancing visual elements in a presentation contributes to a greater understanding of the concepts being presented, compared to the sole use of text. Including simulations in courseware enables students to see and understand various concepts visually. This helps students to better, faster and more accurately understand and retain the presented concepts and information. Simulations and interactive demonstrations can certainly be presented easily using the computer technology and be made easily assessable to many using Internet technology.

In recent years, Internet and related technologies have been used in various ways to teach difficult engineering concepts. Java applets have been used to demonstrate the functionality of digital logic circuits [7]-[9], and logic minimization with Karnaugh-maps [10].

PATTERN RECOGNITION

Pattern recognition [6] is an interdisciplinary subject involving research in varied fields such as engineering, computer science, information science, statistics, physics, chemistry, biology and medicine. Pattern recognition concepts have become increasingly recognized as an

[1] A.A. Kassim, Department of Electrical and Computer Engineering, National University of Singapore, ashraf@nus.edu.sg.
[2] C.M. Lau, Department of Electrical and Computer Engineering, National University of Singapore.

important factor in the design of modern information processing and decision-making systems.

Pattern recognition concerns the description, classification or categorization of input data based on its significant features. An example of a pattern recognition system is a character recognition system that receives optical signals as input data and identifies the name of the character. The major approaches to pattern recognition are structural, neural and statistical.

The interrelationships of features sometimes yield important structural information, which facilitates structural description or classification. This is the basis for syntactic or structural pattern recognition, which is used for classification and description. Classification is based on measures of pattern structural similarity. Artificial neural networks (ANNs) emerged from attempts to emulate the computational paradigm of biological systems. ANNs are particularly well suited for pattern association problems. They need to be trained to learn correct response for each of training sample. After training, ANNs are expected to self-organize to enable extrapolation with new yet similar patterns based on “experience“ with training set.

Humans naturally and easily perform acts of recognition, recognizing faces, music, handwriting and etc. Statistical pattern recognition by computers is in some way similar to human recognition. Statistical pattern recognition assumes a statistical basis for classification. A set of measurements or features extracted from input data, form a feature vector ***x*** that are each assigned to one of *c* classes by developing decision or classification strategies, which form classifiers. The design of the classifier involves integrating all available information, such as measurements and a priori probabilities. Features are assumed to have a class conditioned set of probabilities and/or probabilities density functions. The decision rules are formulated in different ways.

Our implementation is based on the statistical approach to pattern recognition. Our main focus is on the classification stage, which involves the ability to classify an object into one of a number of classes using the class conditional probabilities and/or a priori probabilities in the decision process. The Bayes Decision Theory [6], which is important in statistical pattern recognition, involves the use of conditional and a priori probabilities to obtain the decision regions that can be used to assign input features to one of many classes.

IMPLEMENTATION ISSUES

Our implementation was realized in Java which is a platform independent, object oriented language that has all features of C++ with added capabilities such as automatic garbage collection and multi-threading. The platform independence of Java makes it an excellent language for use over the Internet. Java programs can be standalone programs or applets. Standalone programs are interpreted and executed like programs written in other languages while Applets are 'mini-programs' that are downloaded and interpreted by a Java enabled web browsers.

Java allows for fully interactive programs for the Internet. Also, since the Java program runs on the client computer, all actions are quickly processed locally. A major drawback of Java is the slow speed of execution because it needs to be interpreted.

THE IMPLEMENTATION

Interactive simulations of the following pattern recognition concepts were implemented:

- Bayes Decision Theory
- Minimax Strategy
- Sequential Decision Theory

Bayes Decision Theory

In Bayes Decision Theory, decision regions are determined using the class conditional and a priori probabilities of the underlying distributions of the various classes. The Bayes Decision Theory simulation, which assumes that the underlying distributions are Gaussian, allows students to interactively change the parameters and view the resultant decision regions. Fig. 1 shows the simulation for one-dimensional (1D) features for a two-class (ω_1 , ω_2) problem. Students can change the class conditional probabilities by modifying the mean and variance of the underlying distributions of each class through a very user-friendly graphical interface. The graph will then immediately change to reflect the effect of changes to the conditional distributions and the resulting decision regions. The effect of changing the a priori probabilities, $P(\omega_1)$ and $P(\omega_2)$, on the decision regions will also be shown. By clicking the "Working" button, students can view the mathematical steps needed to derive the decision regions as shown in Fig 2. Students can also vary the values of an input feature, *x*, to determine its class.

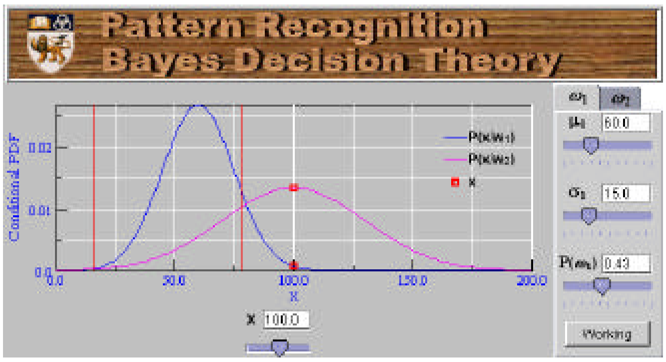

FIGURE. 1
THE BAYES DECISION THEORY (1-D CASE).

A simulation based on the Bayes Decision Theory for two-dimensional (2D) feature vectors (two feature elements) was also implemented as shown in Fig 3. When two features are used, the underlying distributions need to be viewed in 3D and the decision regions are 2D surfaces. It is difficult for students to visualize 3D probability distributions and the resulting 2D decision regions, shown in Fig 3. The simulation helps to explain how the decision regions change when underlying parameters change. The class conditional probabilities can be interactively and dynamically changed by modifying the mean vectors and covariance matrices of the underlying distributions of each class. The "Working" button functions as in the 1D simulation. The user can also see different views of the 3D plot by moving and rotating the graph by moving the mouse pointer in the graph panel while pressing the left mouse button. Furthermore, the plot can be zoomed-in or out to obtain a better view of the plot.

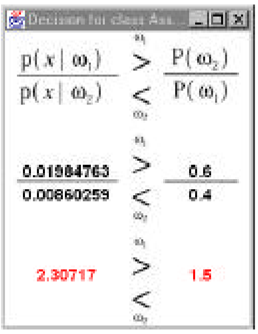

FIGURE. 2
MATHEMATICAL STEPS TO DERIVE THE DECISION REGIONS

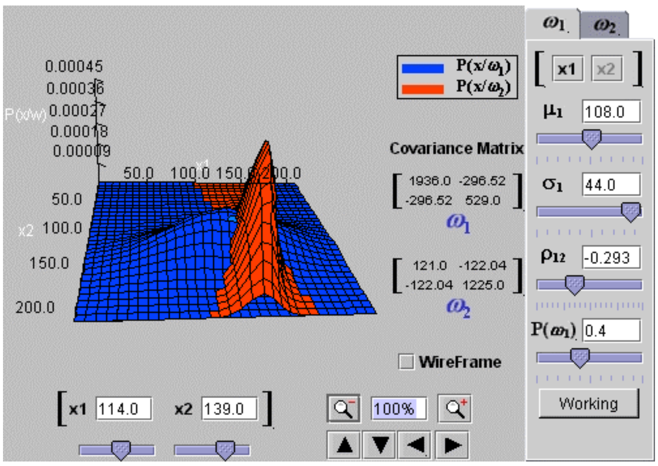

FIGURE. 3
BAYES DECISION THEORY (2-D CASE) SIMULATION

Minimax Strategy

Bayes Decision Theory assumes that the a priori probabilities, P(ω_1) and P(ω_2), are known. There are situations where their actual values are unknown. The Minimax strategy [6] is used in situations where only the class conditional probabilities, P($\boldsymbol{x}|\omega_1$) and P($\boldsymbol{x}|\omega_2$), and loss functions, ψ_{11}, ψ_{12}, ψ_{21} and ψ_{22} are known. The loss function ψ_{ij} is defined as the cost associated with deciding on class i when the actual class is j. The strategy will determine the a priori probability that gives the lowest total cost. The total cost, c is computed as $c = a + b$ P(ω_1) where a and b are functions of the conditional probabilities and loss functions. In the Minimax strategy, the value of P(ω_1) which gives $b\approx 0$ needs to be found.

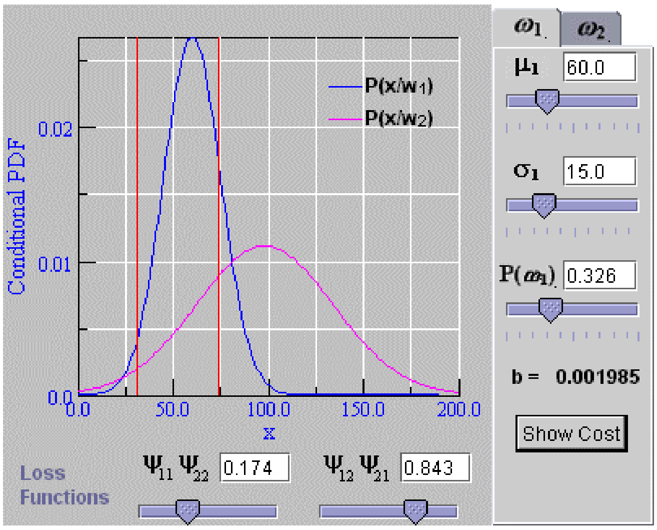

FIGURE. 4
THE MINIMAX SIMULATION

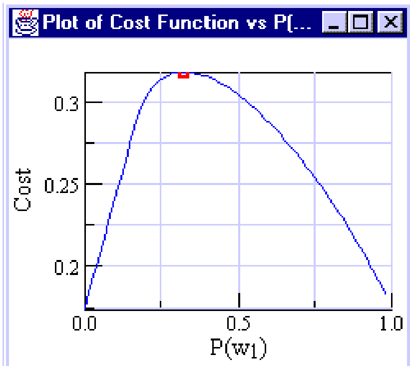

FIGURE. 5
POP-UP WINDOW SHOWING THE COST VERSES P(W1) PLOT

After setting the desired class conditional distributions and the loss functions, students can view the resultant plot of c for different values of P(ω_1) as shown in Fig 5, by clicking the "Show Cost" button in Fig 4. By changing P(ω_1), the user can change the value of b until it becomes zero. As P(ω_1) is varied, the plot of Fig 5 changes dynamically .

Sequential Decision Theory

In the Sequential Decision Theory the number of features used in the classification process is increased until a classification can be made within desired error probabilities. It employs the Sequential Probability Ratio Test (SPRT) to classify objects into classes. Let $\mathbf{x}_m$ represent a vector of m features: x_1, x_2,…., x_m. The classification is made based on the SPRT ratio for an initial value of m=1:

$$\Rightarrow l_m(\mathbf{x}_m)=\frac{p(\mathbf{x}_m|w_1)}{p(\mathbf{x}_m|w_2)} \;\begin{matrix} \overset{w_1}{>} A \\ \underset{w_2}{<} B \end{matrix} \quad \text{where } A > B.$$

If the ratio falls between A and B, another observation needs to be made and the SPRT repeated for m+1. The value for thresholds, A and B, are chosen according to the desired error probabilities (ε_1, ε_2) for the two classes. The simulation first prompts for the required values of ε_1, ε_2, A and B (Fig 6), and then for the conditional probabilities of the first feature, x_1 (Fig 7). The value of feature x_1 can be varied which causes the SPRT value (red dot) to change as shown in Fig 8. If the chosen value of feature x_1 is such that the ratio is greater than A or smaller B, then the decision on the class will be made and shown in a pop-up window. If the ratio falls between A and B, the simulation will request for the next feature, x_2 and its associated probabilities as for x_1. The process continues for x_3, x_4 and so on until the ratio falls outside the value of A and B.

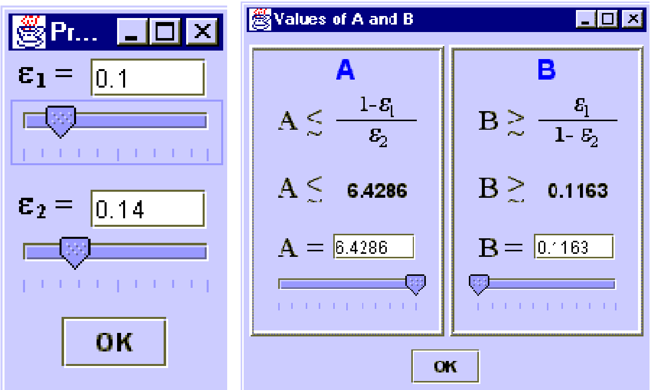

FIGURE. 6
PROMPTS FOR VALUES OF ε_1, ε_2, A AND B.

CONCLUSION

In this paper, we presented details about simulations that have been developed to illustrate certain key concepts in statistical pattern recognition, namely: the Bayes Decision Theory, the Minimax Strategy, and the Sequential Decision Theory. The simulations had been designed and implemented successfully on the web. Students have found the simulations to improve their understanding of the pattern recognition concepts.

REFERENCES

[1] Chaudhury, A, S, and Chew, L, M, "Assessing the Preparedness of Information Institutions for the Knowledge-based Economy", In proc. of 28th Annual Conf. by Canadian Assoc. for Information Science, CAIS2000: Dimension of a Global Information Science, 2000.

[2] Li, H, "Viewpoint: Using Information Technology to Enhance Engineering Education", The International Journal of Engineering Education, Vol. 13, No. 5, 1997, pp 319-324.

[3] Khan, B, H, "Web-based Instruction (WBI): What Is It and Why Is It?", Web-based Instruction. Educational Technology Publications Inc., Englewood Cliffs, NJ, 1997, pp 5-18.

[4] Bloom, B, S, "The 2-Sigma problem: The search for methods of group instruction as effective as one-to-one tutoring", Educational Researcher, Vol. 13, 1984, pp 3-16

[5] About IVLE: http://ivle.nus.edu.sg/public/background.asp.

[6] Duda, R,O, Hart P,E, and Stork, D,G, "Pattern Classification", John Wiley, 2001.

[7] Burks, O, "A Virtual Classroom Approach to Teaching Circuit Analysis", IEEE Transactions on Education, Vol. 39, No. 3, August 1996 pp 287 – 296.

[8] Circuit Builder: http://www.jhu.edu/virtlab/logic/log_cir.htm

[9] CMOS Technology Demonstration – http://tech-www.informatik.uni-hamburg.de/applets/cmos/cmosdemo.html

[10] KV-diagram Applet – http://tech-www.informatik.uni-hamburg.de/applets/kvd/index.html

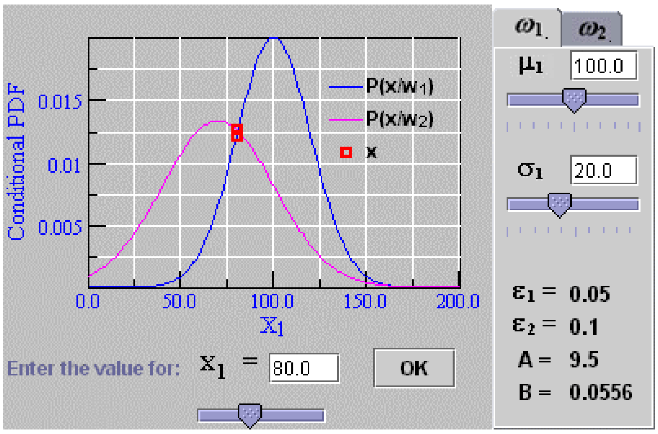

FIGURE. 7
THE SEQUENTIAL DECISION THEORY SIMULATION

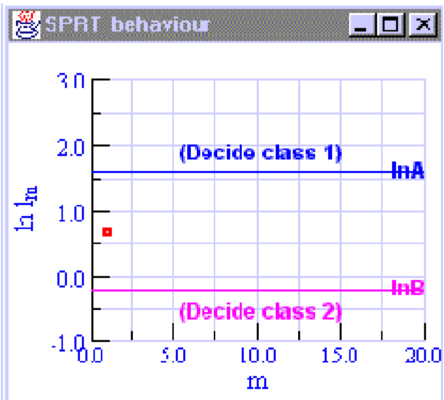

FIGURE. 8
SPRT PLOT