AN ABSTRACT OF THE THESIS OF

CHIA-HAO CHANG for the MASTER OF SCIENCE
(Name) (Degree)

in INDUSTRIAL AND GENERAL ENGINEERING presented on June 7, 1979
(Major Department) (Date)

TITLE: APPLICATION OF PATTERN RECOGNITION TO A HUMAN BEHAVIOR PROBLEM

Approved: Redacted for Privacy
Dr. Michael S. Inoue

Pattern recognition techniques and their application to a consumer behavior study are presented. The Local Majority Method (LOMAME) utilizes a set of prototypes and corrective factors which undergo a training cycle before being utilized as pattern classifiers. Its advantages over the Minimum-Distance Method and the Fix and Hodges Method are discussed in terms of the unique discriminating function.

A FORTRAN model of the LOMAME is first applied to standard pattern recognition problems of "A-and-R" and "1101", and shown to be an effective adaptive model for traditional pattern recognition applications.

Next, to evaluate its effectiveness in processing behavior pattern data, prototypes and training samples are selected from 200 equipment survey questionnaires returned by members of the Society for Wang Applications and Programs (SWAP). The consumer preference for a FORTRAN-base calculator keyboard over such other

keyboards as the traditional calculator and the newer BASIC-base keyboards is studied. The results are promising but considered relatively expensive, time-consuming, and inaccurate in comparison to the performance with the bench-mark problems.

Though additional research efforts will undoubtedly improve the direct utility of pattern recognition techniques to consumer survey type applications, a more fundamental use for an adaptive pattern recognizer appears promising. LOMAME prototypes appears to undergo a learning experience that may potentially model the behavioral change of a group of human decision-makers.

APPLICATION OF PATTERN RECOGNITION TO A HUMAN BEHAVIOR PROBLEM

by

CHIA-HAO CHANG

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

MASTER OF SCIENCE

June 1975

APPROVED:

Redacted for Privacy

Professor of Industrial and General Engineering
in charge of major

Redacted for Privacy

Head of Department of Industrial and General Engineering

Redacted for Privacy

Dean of Graduate School

Date thesis is presented June 7, 1974

Typed by Mary Syhlman for Chia-hao Chang

ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to my major advisor, Professor Michael S. Inoue for his invaluable time, counsel, patient guidance, and especially all the encouragement throughout the study and in the preparation of this thesis.

Deep gratefulness is also owed to Professor Emilo Gagliardio for his constructive criticisms, and data for the experiments.

Sincere appreciation is also extended to Wang Laboratories. The data for SWAP case studies were obtained from them.

Many thanks to all the people who gave the help and suggestions during the preparation of this thesis. A special thanks is expressed to Mr. John Gross and Miss Grace Fong for their advice on my writing.

It was my father's most enthusiastic expectation that made my ambition for advance study. This thesis is dedicated to him.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

APPLICATION OF PATTERN RECOGNITION TO A HUMAN BEHAVIOR PROBLEM

CHAPTER I

INTRODUCTION

Marketing Revolution

The scientific revolution is epitomized by our discovery that the sun, and not the earth, is at the center of our solar system. The marketing revolution, according to Keith (1960), also evolved around a philosophically profound discovery. The consumer is at the center of today's business universe, and not the companies that revolve around the customer.

During the course of industrial revolution, the attention of businessmen has gradually shifted from their problem of what they want to produce to the problem of producing what customers want them to market. As a result, marketing executives were born.

Marketing Research

To a marketing executive, his competitive edge is often based on the additional marketing information that he can obtain above and beyond what his competitors also know. Knowing how consumers behave is essential to his making the right decision. Information about his competitors, suppliers, environmental effects, and other direct and indirect factors influencing his company's operation are important. But no new product idea is accepted by the management until a careful study is made of consumers' wants and needs, likes and dislikes.

The scientific process of collecting data on consumer's patterns of behaviors and analyzing them to extract information that are needed for production and marketing decisions is called Marketing Research. It is a branch of Management Science that provides a feedback by which consumers and other environmental forces (the society, government, etc.) can communicate their wish to the company.

A more formal definition of Marketing Research is afforded by the American Marketing Association:

> The gathering, recording, and analyzing of all facts about problems relating to the transfer and sales of goods and services from producer to consumer (Report of the Definition Committee, 1948).

The 1968 survey of Marketing Research gives a detailed and then current description of the variety of activities classified under the banner of Marketing Research. Table 1-1 shows how these activities have benefited from increasingly sophisticated techniques, including some, like regression analysis and operations research, which owe their origin in fields outside the Marketing Research. The Table 1-2 (Kotler, 1972) shows the chronological introduction of these new techniques into Marketing Research.

Though Marketing Research has been remarkably successful in assimilating various techniques developed by such varied fields as statistics, mathematics, industrial engineering, and operations research, it is painfully clear that no successful attempt has been made to introduce techniques from the field of Pattern Recognition and Artificial Intelligence.

TABLE 1-1: Marketing Research Activities (Twedt, 1968).

ADVERTISING RESEARCH

a. Motivation research
b. Copy research
c. Media research
d. Studies of ad effectiveness
e. Other

BUSINESS ECONOMICS AND CORPORATE RESEARCH

a. Short-range forecasting (up to 1 year)
b. Long-range forecasting (over 1 year)
c. Studies of business trends
d. Profit and/or value analysis
e. Plant and warehouse, location studies
f. Diversificiation studies
g. Purchase of companies, sales of divisions
h. Export and international studies
i. Linear programming
j. Operations research
k. PERT studies
l. Employees morale studies
m. Other

PRODUCT RESEARCH

a. New product acceptance and potential
b. Competitive product studies
c. Product testing
d. Packaging research design or physical characteristics
e. Other

SALES AND MARKET RESEARCH

a. Development of market potentials
b. Market share analysis
c. Determination of market characteristics
d. Sales analyses
e. Establishment of sales quotas, territories
f. Distribution channels and cost studies
g. Test markets, store audits
h. Consumer panel operations
i. Sales compensation studies
j. Studies of premium, coupons, sampling, deals
k. Other

TABLE 1-2. Chronological Development of Marketing Research Techniques (Kotler, 1972).

Decade	Technique
Prior to 1910	Firsthand observation Elementary surveys
1910-20	Sales analysis Operating-cost analysis
1920-30	Questionnaire construction Survey technique
1930-40	Quota sampling Simple correlation analysis Distribution-cost analysis Store auditing techniques
1940-50	Probability sampling Regression methods Advanced statistical inference Consumer and store panels
1950-60	Motivation research Operations research Multiple regression and correlation Experimental design Attitude-measuring instruments
1960-70	Factor analysis and discriminant analysis Mathematical models Bayesian statistical analysis and decision theory Scaling theory Computer data processing and analysis Marketing simulation Information storage and retrival
1970-	Nonmetric multidimensional scaling Econometric models Comprehensive marketing planning models Test marketing laboratories Cluster analysis

An Evaluation of the Potential for Using Pattern Recognition in Marketing Research

Pattern Recognition, as a decision-making process, has been successfully employed in medical applications, engineering applications, business applications, scientific applications and so forth. But very little attention has ever been drawn to the applications in Marketing Research problems. One of the objectives of this thesis is to investigate whether this is due to an inherent and unreconcilable difficulty, or merely that the field has been overlooked.

Requirements for a Potential Marketing Research Technique

Before evaluating the potential of applying Pattern Recognition techniques to Marketing Research problems, there is a need to specify what features that are essential to a modern marketing research technique.

The first requirement for an effective marketing research technique is that it be multivariate:

> For the purposes of marketing research or any other applied field, most of our tools are, or should be multivariate. One is pushed to a conclusion that unless a marketing problem is treated as a multivariate problem, it is treated superficially (Gatty, 1966, p. 158).

The social aptitude for a more complete market research is manifest today. According to Sheth (1971) there are three major considerations: (1) there exists today an extensive data base from the past three decades of marketing research efforts; (2) the traditional input-output analysis is no longer considered adequate for current business needs; (3) finally, computers have brought new capabilities to process large mass of data.

Therefore, a modern marketing research technique should take advantage of computers and computerized data bases. Sheth believes that the most important factor in the rapid diffusion of multivariate methods in marketing research is the availability of computer programs.

> In fact we can assert that the lack of computer programs has been a major factor in the imbalance between the extensive data banks in existence today and their weak statistical analysis in most marketing research activities (Sheth, 1971).

Potential of Pattern Recognition Technique

Pattern Recognition method fits both of these two requirements. It unites the multivariate approach with the computer advantages. The data used in a pattern recognition are usually picked from an n-dimentional hyperspace, and represent n variables. With a great mass of such data, pattern recognition searches an empirical law or indicators (e.g., prototype patterns) that can form a criterion for decision-making. The process of search and decision-making is performed by computer. The decision-making process of a pattern recognition method is usually no different from any statistical method. But the essence of a pattern recognition method is not its decision-making process but the searching process for a decision criterion. Once the criterion is found, decisions can be made from it. To search such a criterion from a large number of complex data is usually very time-consuming. Using a modern computer, pattern recognition researchers 'train' the computer to search the unknown criterion based on some adaptive and/or selective algorithms. Besides the advantage of gaining further insight, there is another feature of pattern recognition techniques that is of an even greater practical value. When the pattern recognizer fails to perform satisfactorily, it is sent back for retraining. Unlike most statistical techniques, the retraining can be resumed easily without wasting previous training.

A pattern recognition technique incorporating the adaptive feature within its search process is called a "learning machine."

Outlines of This Thesis

The second chapter introduces the basic concept of Pattern Recognition. Some technical terms used in Pattern Recognition are defined and a survey of previous applications of Pattern Recognition in various areas is presented. The third chapter describes the derivation of the "Local Majority Method." Several sets of testing data are used to demonstrate the applicability of the Local Majority Method. The fourth chapter discusses the SWAP Questionnaire and Equipment Survey as a case study. The development of the questionnaire and data collection procedure are described. The analysis of collected data is presented in the fifth chapter. The analysis includes a statistical analysis for the selection of the pattern and the variables for this study, and uses a computer program (LOMAME) based on the Local Majority Method to analyze those data. The results and recommendations are discussed in the last chapter. The potential of using adaptive pattern recognition techniques in behavior science is evaluated, and the selective pattern recognition techniques are recommended for further investigations.

CHAPTER II

INTRODUCTION TO PATTERN RECOGNITION

Two Stages in Pattern Recognition

Pattern Recognition is a decision-making process involving two stages: the learning stage and the recognition stage (Meisel, 1972).

Learning Stage

The learning stage is initiated by feeding the training data and making the classification by a decision criterion. The decision criterion will be adjusted by some correction algorithms whenever the classification was found to be incorrect. The classification is based on the concept of distance between patterns in a n-dimentional hyperspace.

> The concept of the distance between two points in hyperspace is central to the problem of pattern recognition. When two points differ only slightly in their coordinate values (that is, when the distance between them is small) the corresponding patterns ought to be highly similar. Conversely, patterns that are well separated in hyperspace must be dissimilar (Casey and Nagy, 1971).

The objective for this stage is to find a proper decision criterion that can correctly classify patterns in that hyperspace. The correction algorithms play an important role in this stage. They determine whether the recognition stage will perform satisfactorily or unsatisfactorily.

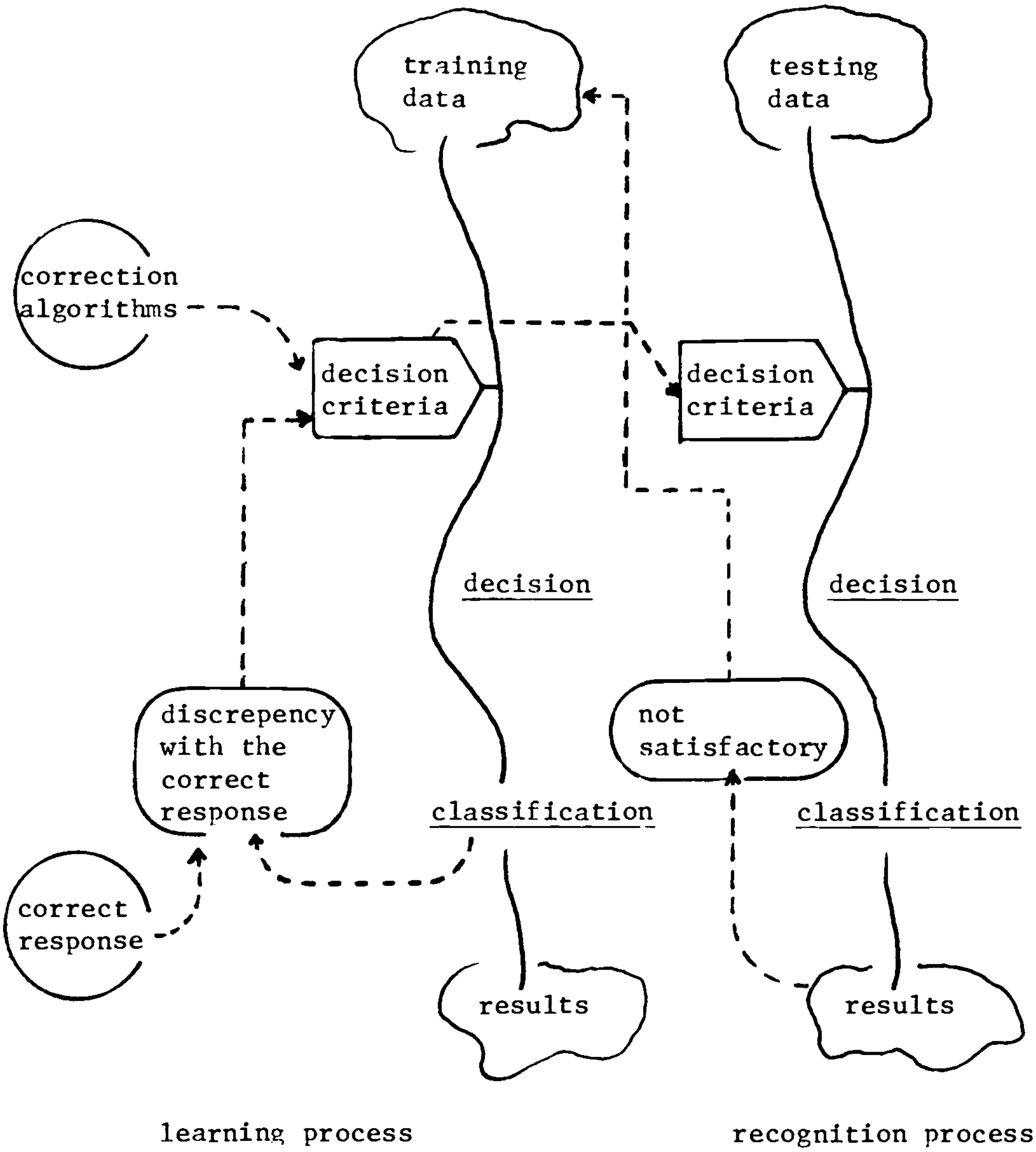

Figure 2-1 Pattern Recognition Process

Recognition Stage

The recognition stage receives a resultant decision criterion from the learning stage and uses it to classify the incoming patterns into the corresponding categories. When those incoming patterns are predetermined, the classification results can be used to indicate how well the decision criterion has worked.

Discriminant Functions and Pattern Classifier

The decision criteria can be mathematically expressed by a set of single-valued functions, $G_i(\overline{X})$, $i=1,\ldots,R$ (where R is the number of categories in the n-dimentional hyperspace) of the pattern $\overline{X}$. These functions are called discriminant functions. For any pattern $\overline{Y}$ in the i^{th} category, $G_i(\overline{Y}) > G_j(\overline{Y})$ for $i,j=1,\ldots,R$, $i\neq j$. The device which employs these R discriminant functions to classify patterns is called a discriminant function pattern classifier. In the two-category case, it is known as a pattern dichotomizer (Nilsson, 1965). Since every problem can be simplified into a set of two-category cases, the term 'pattern classifier' often refers to a pattern dichotomizer. The task of a pattern recognition designer is to find a set of appropriate discriminant functions, by which an incoming pattern will be classified into the correct region with a high probability. Such a pattern classifier is said to be a successful predictor of patterns.

According to the physical characteristics and pattern forms, recognition can be classified into visual, symptomatic and auditory.

Visual recognition is to recognize data from pictures or photographs, symptomatic recognition is to recognize patterns from physical measurements, while auditory pattern recognition is to recognize frequency of an utterance as the pattern.

Previous Work of Pattern Recognition

During the last quarter of this century, Pattern Recognition had become widely and successfully applied to many areas. Examples encompass the following areas.

Medical Applications

In the field of medical science, Pattern Recognition has been found useful in visual and symptomatic recognition of diseases. Visual recognition works successfully in x-ray picture analysis (Harlow, et al, 1971), electrocardiographic diagnosis (Nagy, 1968; Meisel, 1972), vectorcardiogram analysis (Specht, 1967), electroencephalogram analysis (Meisel, 1972), microscopic image analysis (Nagy, 1968), cell tissue analysis (Meisel, 1972), and so forth. Based on laboratory tests and/or physical measurements, Pattern Recognition was used to make medical diagnosis; the results were reported to be very encouraging (Kolers and Eden, 1968).

Engineering Applications

Auditory recognition had been used to detect failures of machinery or aircraft by an engine vibration analysis. Visual recognition was used to make radar and sonar signature analysis to identify the types of aircraft by their radar returns (Meisel, 1972). Pattern Recognition was also used for security evaluation in power system operations, and was found to save computer time in some cases where simulation analysis has previously been used (Pang, et al, 1973). Pattern Recognition was also used to develop a dynamic scheduling algorithm for large digital computer systems (Northouse and Fu, 1973). In the bio-medical engineering, Pattern Recognition was used to make decisions in the real time control of an arm aid for physically handicapped operators (Lawrence and Lin, 1972).

Commercial Application

Magnetic-ink character readings and optical-character recognition devices are used by computers to read hand-written and printed data from accounting machines, computer printers, cash registers, and typewriters. Pattern Recognition techniques are used to decipher checks, credit-cards, utility-company bills, sales invoices, insurance-premium notices, etc. (Bohl, 1971; Casey and Nagy, 1971).

Government Application

Pattern Recognition had been used to make investigation on some cases by fingerprint identification (Meisel, 1972) and face-recognition (Fischler and Elschlager, 1973). Auditory pattern recognition was used in speaker identification, and found to be a more accurate method than the traditional fingerprint identification (Sakai and Doshita, 1963; Nagy, 1968).

Public Utilities

An early example of Pattern Recognition was to assist in deciphering hand-sent Morse-codes. The successful results encouraged further works on telephone-message recognition, and even the design of voice-actuated typewriters (Nagy, 1968; Casey and Nagy, 1971). Some Post Offices have installed ZIP code readers to help handle the large volume of mail. Similar devices have been used to read Social Security wage reports (Casey and Nagy, 1971). The aerial photograph analysis was recommended for urban development and city planning (Kawamura, 1971).

Natural Science

In atomic and nuclear physics research, Pattern Recognition was used for bubble and spark chamber photo analysis (Nagy, 1968). In meteorology, Pattern Recognition was used to make weather forecasting (Kolers and Eden, 1968). In seismic signal analysis, auditory Pattern Recognition was applied to earthquake and nuclear explosion detections (Nagy, 1968; Meisel, 1972).

Recreation

On a lighter side, Pattern Recognition was reported being used in palm reading (Oda, et al, 1971) and in musical note recognition (Kolers and Eden, 1968).

Behavior Pattern's Recognition

Most of the works emphasize the recognition of physical patterns. These patterns are quite stationary. Very few works were reported on recognizing non-stationary behavior patterns as we might find in marketing research. This thesis is an attempt to establish the applicability of Pattern Recognition to effectively analyze behavioral patterns.

CHAPTER III

LOCAL MAJORITY METHOD

Introduction

The task of the Pattern Recognition theory is to find a set of proper discriminant functions by which unknown patterns can be correctly classified into a set of N categories. An adaptive method makes a reasonable preliminary guess of those discriminant functions and then adjusts it during the learning stage. The Local Majority Method is derived from both the Minimum-Distance Method and the Fix and Hodges Method. The two major application areas for these methods are (1) linear separable categories and (2) piecewise-linear separable categories.

Minimum-Distance Method

The Minimum-Distance Method is a simple and direct method in Pattern Recognition. The pattern classifier using this method will assign each incoming pattern $\bar{V}$ to the category which contains a pattern nearest from $\bar{V}$ in the hyperspace.

Linear Separable Case

In order to minimize the time and effort needed to measure these distance, each category can be represented by a single prototype point if this category is linear separable from the other categories. Two categories are linear separable if and only if a hyperplane exists which has all patterns of one category on one side and all patterns of another category on the other side (Nilsson, 1965, p. 20). If there are N linear separable categories, there will be N prototype points, $\overline{P}_i$ i=1, 2,...,N. The measure of Euclidean distance, D , from the unknown pattern $\overline{V}$ to $\overline{P}_i$ in a d-dimensional pattern space can be calculated according to the root mean square formula:

$$D_i = ((\overline{V}-\overline{P}_i)\cdot(\overline{V}-\overline{P}_i))^{\frac{1}{2}} = (\sum_{j=1}^{d} (V_j-P_{ij})^2)^{\frac{1}{2}} \qquad (3.1)$$

where V_j and P_{ij} are the j^{th} component of $\overline{V}$ and $\overline{P}_i$ respectively.

The assumption of non-negative distance enables us to consider the square of distance instead of the distance D_i

$$\begin{aligned} D_i^{\,2} &= (\overline{V}-\overline{P}_i)\cdot(\overline{V}-\overline{P}_i) \\ &= \overline{V}\cdot\overline{V}-2\overline{V}\cdot\overline{P}_i + \overline{P}_i\cdot\overline{P}_i \qquad (3.2) \end{aligned}$$

The value $\overline{V}\cdot\overline{V}$ is a constant for all prototype points, and can be dropped. Therefore, the objective to find the minimum $-2\overline{V}\cdot\overline{P}_i + \overline{P}_i\cdot\overline{P}_i$, the discriminant function, can now be defined as:

$$[\text{Max}]\ G_i(\overline{V}) = \overline{V}\cdot\overline{P}_i - \tfrac{1}{2}\,\overline{P}_i\cdot\overline{P}_i \text{ for } i=1, 2,\ldots,N. \qquad (3.3)$$

That is, the pattern classifier is going to find the i which maximize the negative value of $\frac{1}{2}(-2\overline{V}\cdot\overline{P}_i + \overline{P}_i\cdot\overline{P}_i)$

Piecewise Linear Separable Case

In most cases, a single hyperplane is not adequate to linearly separate one category from another. A few more hyperplanes are required. We call these cases piecewise linear separable. Figure 3-1 is an example.

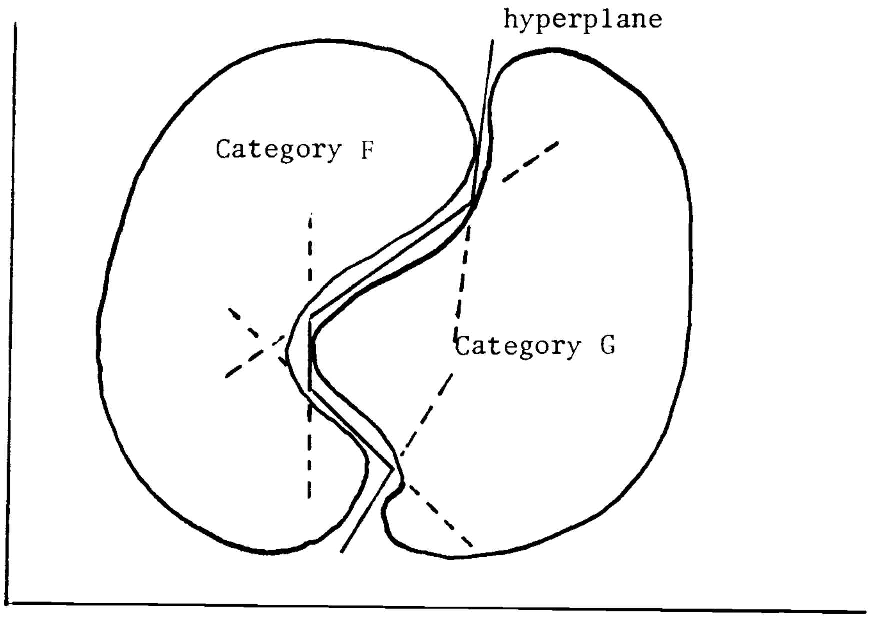

Figure 3.1. The Piecewise Linear Separable Categories.

A single prototype is not enough to make an entire category linearly separated from other categories, there are more prototype points assigned to the i^{th} category, then the consideration will extend to all these prototype points. The procedures can be carried out as follows: First, find the smallest distance, d_i, in each category, where

$$d_i = \underset{j=1,2,\ldots,N_i}{\mathrm{MIN}} [\overline{V}-\overline{P}_i^{(j)}] \text{ for } i=1,2,\ldots,N \tag{3.4}$$

and $\overline{P}_i^{(j)}$ represents the j^{th} prototype point in the i^{th} category.

The pattern classifier will select the smallest d_i, $i=1,\ldots,N$ and the pattern $\overline{V}$ will be assigned to the category associated with that smallest d_i.

Similarly, in a piecewise linear separable case, the discriminant function can be selected as:

$$G_i(\overline{V}) = \underset{j=1,\ldots,N_i}{\mathrm{MAX}} (\overline{P}_i^{(j)}\cdot\overline{V} - \tfrac{1}{2}\,\overline{P}_i^{(j)}\cdot\overline{P}_i^{(j)}) \text{ for } i=1,2,\ldots N \tag{3.5}$$

The incoming pattern $\overline{V}$ will be assigned to a category which has the largest discriminant value (Nilsson, 1965, p. 24).

Fix and Hodges Method

The minimum distance method seeks one pattern that is nearest to the incoming pattern and neglects the influence of the rest of the patterns. A pattern could easily be classified incorrectly since the decision depends on only one nearest sample point. Figure 3.2 is an example. The incoming pattern $\overline{V}$ could be assigned incorrectly to category G by the Minimum-Distance Method.

The Fix and Hodges Method (Nilsson, 1965, p. 120) applies to more general cases. A positive integer k is selected based on

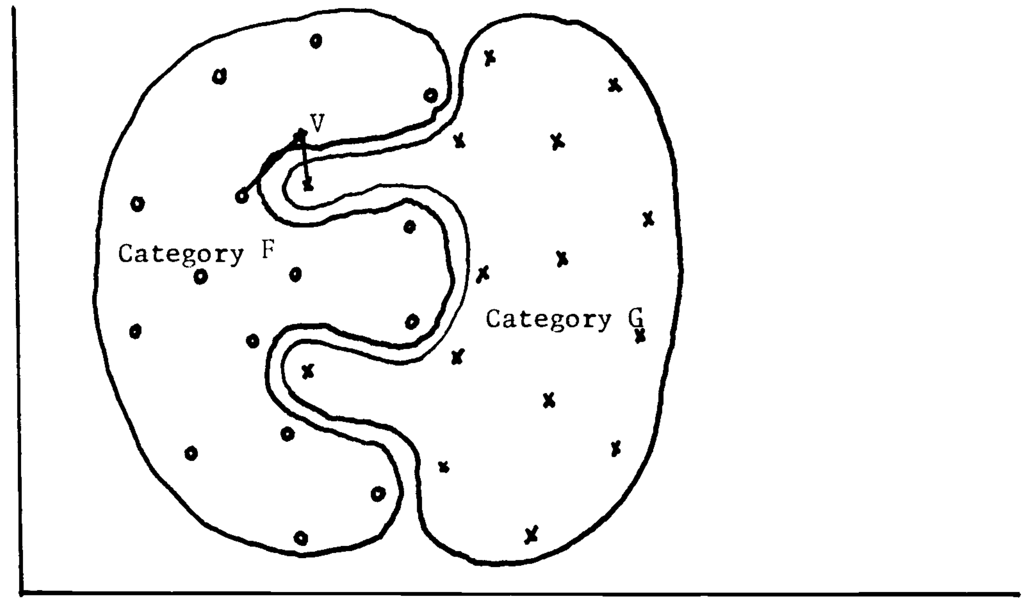

Figure 3-2. Failure of the Minimum Distance Method

the known information and some reasonable guesses. The k should be small compared to the number of the known sample patterns in a different categories. K patterns which are near the incoming pattern $\bar{V}$ are found. If N_i patterns of these k patterns belong to the i^{th} category $\left(\sum_{i=1}^{N} N_i = k\right)$, then N discriminant functions can be selected such that:

$$G_i(\bar{V}) = N_i \quad i=1,2,\ldots,N \tag{3.6}$$

and $\bar{V}$ is assigned to the one with the highest discriminant value.

When the k is chosen to be 1, the method is reduced to the minimum-distance method (Nilsson, 1965).

Derivation of Local Majority Method

Deficiency of the Fix and Hodges Method

The major shortcoming of the Fix and Hodges method is that each of the k nearest points is considered to possess the same discriminating power in classifying the incoming $\bar{V}$. Unfortunately, it is more reasonable to consider that points nearer to the pattern $\bar{V}$ should have more influence on decision (see Chapter I). In addition, those sample patterns other than these k patterns should also contribute some influence in making the decision of classification. Those sample patterns that are relatively near to the $\bar{V}$ should have a major influence in decision while those which are relatively remote should have a minor influence in decision. Such observations led to the formulation of the so-called local majority rule (Sebestyen, 1962).

The local majority rule suggests a discriminant function described as:

$$G_f(\bar{V}) = \sum_{n=1}^{N_f} (1 + (D_n(\bar{V})/r)^m)^{-1} \tag{3.7}$$

where $D_n(\bar{V})$ represents the distance from the n^{th} pattern in the F category to $\bar{V}$; r is called the radius of the effective neighborhood that influences decision making; m, a positive integer, determines the rate of change of influence with respect to the distance; N_f, the number of patterns in the F category.

In a two-category case, the decision can be made from the sign of the difference of two discriminant functions:

$$G(\bar{V}) = G_f(\bar{V}) - G_g(\bar{V}) = \sum_{n=1}^{N_f} (1+(D_{f_n}(\bar{V})/r)^m)^{-1}$$
$$- \sum_{n=1}^{N_g} (1+(D_{g_n}(\bar{V})/r)^m)^{-1} \qquad (3.8)$$

Using the above form as the discriminant function, a method applying an adaptive technique to search appropriate prototype points was developed.

Local Majority Method's Procedures

The procedure is as follows. A number of sample patterns are used arbitrarily as prototype points for each category. Another set of predetermined patterns are employed to train these prototype points, and these prototype points may be adjusted to more representative positions whenever a wrong judgment is made. The result of this training period may not perfectly discriminate all the patterns in the categories, and a small amount of error must be tolerated.

The characteristics of the local majority method are as follows: (1) the prototype points are used instead of sample patterns; (2) the number of prototype points in each category is chosen to be equal; (3) the radius of the effective neighborhood, r is chosen to be proportional to the mean distance of the prototype points from the sample V, and is reduced by increasing the number of prototype points in each category; (4) the m in the discriminant function is a constant.

Several expressions of r and values of m are possible. For example, an r may be chosen as:

$$r = \frac{20}{20 + N} D$$

where D is the mean distance and N is the number of prototype points each category uses. R will approach D when N is small, and will become a small fraction of D when N is large. Similarly, if m is chosen to be sufficiently large, the discriminant value of the local majority method will correspond to the number of prototype points within the radius as in the Fix and Hodges Method.

Error-Correction Procedures

The essence of an adaptive technique is that it has an error-correction procedure to adjust the mechanism from making more of the same errors.

As mentioned earlier, a discriminant function is selected as a measure, such as Euclidean distance.

$$G_i(\overline{V}) = \overline{V} \cdot \overline{P}_i - \tfrac{1}{2}\, \overline{P}_i \cdot \overline{P}_i \qquad i=1,2,\ldots,N \tag{3.9}$$

In a two-category case, the classification can be based on the difference of two discriminant functions.

$$G_f(\overline{V}) - G_g(\overline{V}) = \overline{V} \cdot (\overline{f}_i - \overline{g}_i) - \tfrac{1}{2}\, (\overline{f}_i \cdot \overline{f}_i + \overline{g}_i \cdot \overline{g}_i) \tag{3.10}$$

where $\overline{f}_i$ and $\overline{g}_i$ represent patterns in category F and category G.

In a d-dimentional hyperspace, this discriminant function can be represented by a dot-product $\overline{Y}\cdot\overline{W}$ where $\overline{Y}$ stands for $(V_1,V_2,\ldots,V_d,1)$ and $\overline{W}$ stands for $[f_{i1}-g_{i1},\ldots,f_{id}-g_{id},\frac{1}{2}(\overline{f}_i\cdot\overline{f}_i+\overline{g}_i\cdot\overline{g}_1)]$. When f_{ij} and g_{ij} respectively indicate the jth component of the d-dimensional vectors $\overline{f}_i$ and $\overline{g}_i$. The incoming V is assigned to category F if $\overline{Y}\cdot\overline{W}$ is positive and category G if $\overline{Y}\cdot\overline{W}$ is negative. The case $\overline{Y}\cdot\overline{W} = 0$ means that $\overline{V}$ is on the hyperplane separating the two categories.

The error-correction rules are the following:

(1) NO CORRECTION: If the training pattern is correctly classified, no change is made. That is no change is made if:

$$\overline{Y}\cdot\overline{W} > 0 \text{ and } \overline{V} \varepsilon F$$
$$\text{or } \overline{Y}\cdot\overline{W} < 0 \text{ and } \overline{V} \varepsilon G \tag{3.11}$$

(2) CORRECTION: Otherwise a change is made

$$\overline{W} = \overline{W} + C\overline{Y} \text{ if } \overline{Y}\cdot\overline{W} < 0 \text{ and } \overline{V} \varepsilon F$$
$$\overline{W} = \overline{W} - C\overline{Y} \text{ if } \overline{Y}\cdot\overline{W} > 0 \text{ and } \overline{V} \varepsilon G \tag{3.12}$$

where the correction increment C is a positive number.

Three types of C were suggested by Nilsson (1965): (1) any fixed number greater than zero; (2) the smallest integer greater than $\overline{W}\cdot\overline{Y}$ / $(\overline{Y}\cdot\overline{Y})$; (3) C = $\overline{W}\cdot\overline{Y}$ / $(\overline{Y}\cdot\overline{Y})$ and $\lambda>1$.

In a piecewise linearly separable case, the correction is usually more complicated. The error-correction procedure for the Local Majority Method is:

(1) NO CORRECTION: No change is made when the training pattern is correctly classified.

(2) TYPE I CORRECTION: If $\overline{V}$ belongs to F and the value of $G(\overline{V})$ in (3.8) ≤ 0, a correctional term is added to each prototype point of F and subtracted from each prototype point of G.

(3) TYPE II CORRECTION: If $\overline{V}$ belongs to G and the value of $G(\overline{V})$ in (3.8) ≥ 0, a correctional term is subtracted from each prototype point of F and added to each prototype point of G.

Correction Term

The effect of the error-correction procedure strongly depends on the correction term employed. Let us consider the following factors.

The Location of the Particular Training Sample

Since the values of discriminant functions are dependent on the distances from the incoming pattern $\overline{V}$ to a prototype point $\overline{f}_i$ (or $\overline{g}_i$), (V_j-f_{ij}), 1,...,d. When $\overline{V}$ is misclassified, it is reasonable to assume that by moving all prototype points the correct category toward $\overline{V}$, or by moving all prototype points of the incorrect category away from $\overline{V}$, more correct decisions can be expected.

The Distance Between Each Prototype Point and Its Training Sample

According to the Local Majority Rule, those which are nearer the training sample have a greater influence than those which are farther from it. Therefore, it is not useful to move every prototype point the same distance toward or away from the sample. The distance moved by the prototype point $\overline{f}_i$ (or $\overline{g}_i$) is inversely related to the distance from the training sample, $DISVF_i$: $(1+(DISVF_i/D))^{-1}$ Prototype points nearer the sample will have a greater move and those farther away from the training sample will have the changes proportionally decreased.

The Amount of Error Made in Misclassification

The method makes decisions by evaluating the sign of $G(\overline{V})$ of (3.8). When the classification is incorrect, the value of $G(\overline{V})$, DIFFG, is indeed a measure of the error; the adjustment should be made proportional to that value. In order to avoid an over-correction, which may make the result worse, the factor was chosen to be

$$\frac{DIFFG}{C + DIFFG}$$

where C is a constant. The factor will have its value range between 0 and 1.

Time Factor

At the beginning, the prototype points are chosen arbitrarily. They may be far away from the appropriate ones which can correctly represent most of the patterns in the category. The training patterns are chosen randomly, and some of them may cause the correction function to impose a bad effect on decisions on other patterns. If this happens at the beginning of the training period, the result will not be a serious mistake, because in the rest of the period a bigger error will only lead to a greater adjustment as discussed on page 26. But, if these cases happen at the end of the training period, the prototype points would have all been changed to their near ideal places, and a wrong change could destroy all the work that had been done. Therefore, the changes should decrease with time, and the factor should be inversely proportional to the cumulative number of previous changes: $\left(\frac{1}{\text{TFCTOR}}\right)$.

The cumulation of all the discussion above leads to the correction term for $G_f(V)$ expressed as

$$(V_j - f_{ij}) \left(\frac{1}{\text{D+DISVF}_i}\right) \left(\frac{\text{DIFFG}}{\text{1+DIFFG}}\right) \left(\frac{1}{\text{TFCTOR}}\right)$$

where V_j and f_{ij} are the j^{th} components of $\overline{V}$ and $\overline{f}_i$,DISVF_i is the distance between the $\overline{V}$ and $\overline{f}_i$;DIFFG is the measure of the error made; and TFCTOR is the time factor which acts as a counter for the frequency of error correction.

Testing Problems

A FORTRAN program for the Local Majority Method was written (Appendix A) and applied to a variety of problems.

"A-and-R" Problem

The "A-and-R" problem is a classical recognition problem for discriminating hand-written characters A and R. Poorly written A and R are difficult to distinguish from each other.

Let us assume each character to be written on a 24 x 24 grid. Figure 3-3 is an example. Let 0 represent a white square and 1 represent a black square. A character can then be represented by a string of 576 variables. Not every variable affects the recognition, and some can be completely ignored. Only a few of them show any significant difference between an A and an R.

Galiardo (1973) found that A and R can be distinguished by observing only 19 variables out of 576. The other variables contribute only very minor influences. The data supplied by Gagliardo to verify the Local Majority Method use 27 variables. For training and testing, two different sets of data were used, A set of experiments were performed by changing: (1) the number of prototype points in each category, (2) the number of training samples, (3) the number of testing samples, and (4) the increase in the time factor. The corresponding results are shown in Table 3.1.

"1101" Problem

"1101" is another classical pattern recognition problem. In this problem, patterns are distinguished by having or not having the chain "1101" in its string of variables. The "1101" is arbitrarily placed in the string. The other variables have the value 1 or 0. Sample data are randomly generated (Figure 3-4).

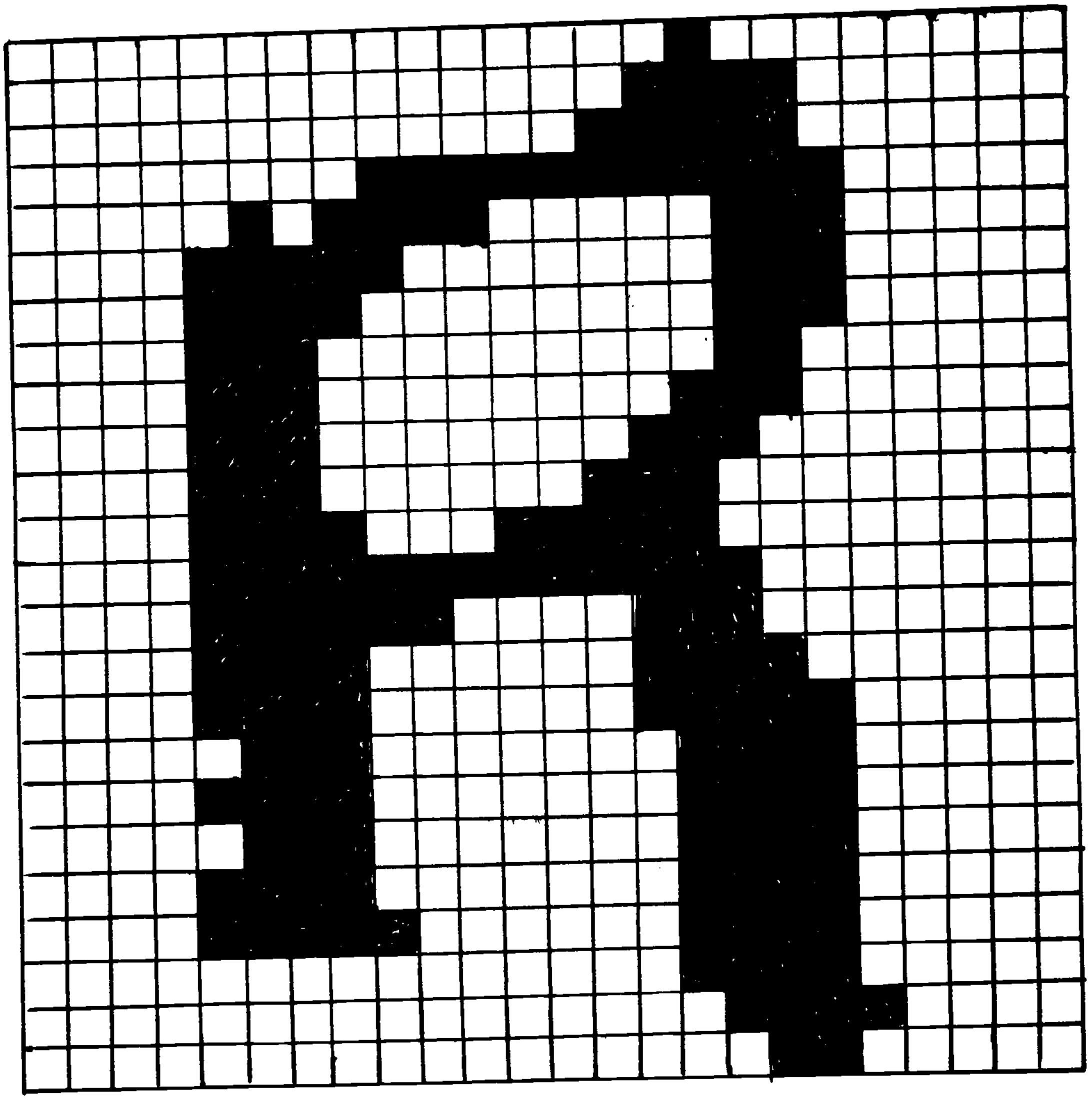

Figure 3-3. "A-and-R" Recognition Problem

TABLE 3-1. Experimental Results for "A-and-R" Problem:

Number of Prototypes Per Category	Number of Training Samples	Number of Testing Samples	Increment in Time Factor	Error Rate
10	70	150	0.1	13/150
10	70	150	0.05	15/150
10	100	120	0.1	11/120
10	100	120	0.05	12/120
10	120	100	0.1	10/100
10	120	100	0.05	11/100
15	60	150	0.1	17/150
15	60	150	0.05	18/150
15	100	110	0.15	7/110
15	100	110	0.1	7/110
15	100	110	0.05	7/110
15	120	90	0.15	6/90
15	120	90	0.1	6/90
15	240	240	0.1	23/240
20	60	140	0.05	19/140
20	70	130	0.1	10/130
20	100	100	0.1	10/100

0011011010 1010011010 0011110100 0011110111	Category that contains '1101'
0111100000 1100010100 1010000011 1011100111	Category that does not contain '1101'

Figure 3-4 "1101" Recognition Problem

It is easy to design a program that deals especially with this "1101" problem. A finite state machine can be built to recognize the particular patterns, and the problem can be solved without any difficulty. However, a pattern recognizer should ideally be a general model. It should be able to solve a variety of problems with different characters. The "A-and-R" problem is more or less related to the value detected at certain fixed variables. By these variables, patterns are easily classified into categories. That is why a satisfactory outcome of the experiments was not unexpected. But the "1101" problem is different. No particular variable or variables are in favor of any category. It looks more difficult than the "A-and-R" problem, and the experiment results attest to the fact. Similar to the "A-and-R" problem, the "1101" problem was used to verify the method by using (1) different number of prototype points in each category, (2) different number of training samples, (3) different number of testing samples and (4) different increments in the time factor. The corresponding results are shown on Table 3.2.

TABLE 3-2. Experimental Results for "1101" Problem.

Number of Prototypes Per Category	Number of Training Samples	Number of Testing Samples	Increment In Time Factor	Error Rate
20	100	100	0.2	29/100
20	100	100	0.175	27/100
20	100	100	0.15	27/100
30	80	100	0.175	16/100
30	80	100	0.15	16/100
30	80	340	0.15	79/340
30	180	240	0.2	68/240
30	180	240	0.175	69/240
30	180	240	0.15	67/240

Buying Intent Problem

Let us consider a marketing situation involving a "semi-luxury" or "luxury" merchandise such as a piece of home furnishing. Let us also assume that the consumer's intent to buy will depend on a set of observed variables:

Residence Area: (1) urban, (2) farm.

Disposible Family Annual Income: (1) less than $10,000, (2) between $10,000 and $15,000, (3) more than $15,000.

Living Status: (1) married, (2) single, separated, or divorced.

Number of Dependents: (1) no dependent, (2) have 1 or 2 dependents, (3) more than 2 dependents.

Transportation Situation: (1) own a car, (2) does not own a car.

Educational Background: (1) have college education, (2) do not have college education.

Occupation and Prospect Future: (1) likely to be good, (2) not good.

By assumption, if a consumer fulfills any of the following conditions he will have the buying intents.

1. The disposable family annual income is more than $15,000 and his business has a good expectation in the near future.

2. The disposable family annual income is more than $15,000 and he is married.

3. The disposable family annual income is more than $15,000, he has received college education and no more than 2 dependents.

4. The disposable family annual income is between $10,000 and $15,000, no dependents, and his business has a good expectation in the near future.

5. The disposable family annual income is between $10,000 and $15,000, no dependents, and own a house in the urban side.

6. The disposable family annual income is between $10,000 and $15,000, married, and own a house.

7. The disposable family annual income is less than $10,000, no dependents, own a house, and his business has a good expectation in the near future.

8. The disposable family annual income is less than $10,000, with no more than 2 dependents, own the house, has a college education, and his business prospect is good.

All other consumers are assumed as having no buying intent.

Ten variables are chosen as follows to describe the condition in the nominal scale values 1 or 0. The consumer is:

Variable 1: 1, living in an urban area

0, living in a farming area

Variable 2: 1, with a disposable annual income larger than $10,000

0, with a disposable annual income less or equal to $10,000

Variable 3: 1, with a disposable annual income larger than $15,000

0, with a disposable annual income less or equal to $15,000

Variable 4: 1, married

0, single, divorced, or separated

Variable 5: 1, with dependent(s)

0, without dependent(s)

Variable 6: 1, with more than 2 dependents

0, with less than or equal to 2 dependents

Variable 7: 1, owning a house

0, renting a house

Variable 8: 1, owning a car

0, not owning a car

Variable 9: 1, college educated

0, not college educated

Variable 10: 1, having a good business prospect

0, not having a good business prospect

Let X_i represent the variable i when its value is 1, and X_i' represent the variable i when its value is 0. The consumer can be classified as follows: when the logical value of the following function equals 1, the consumer is classified as one having the buying intent, and when the logical value of that function equals 0, the consumer is classified as one not having the buying intent.

$$(V_3 \wedge X_{10}) \vee (X_3 \wedge X_4) \vee (X_3 \wedge X_5 \wedge X_6' \wedge X_9)$$
$$\vee(X_2 \wedge X_3' \wedge X_5' \wedge X_{10}) \vee (X_2 \wedge X_3' \wedge X_4 \wedge X_7)$$
$$\vee(X_1 \wedge X_2 \wedge X_3' \wedge X_5' \wedge X_7) \vee (X_2' \wedge X_5' \wedge X_7 \wedge X_{10})$$
$$\vee(X_2' \wedge X_6' \wedge X_7 \wedge X_4 \wedge X_{10}) \quad (3.13)$$

where V represents a logical union
Λ represents a logical intersection

The experiments are performed as follows: the number of prototype points, training samples, testing samples and the incremental value are all fixed. The computer is given data and prints out the error rates in both training and testing stages. If the user decides that the results are not satisfactory, he can order the program back to its training stage, and hopefully, obtain a better result. The same training and the testing samples are used, but the prototype points are modified, and the time factor is increased. In each cycle, the final modified prototypes are also printed out for reference. Some of the results were listed on Table 3.3.

TABLE 3-3. Experiment Results for Buying Intent Problem

Number of Prototype Points Each Category	Number of Training Samples	Incremental Value in Time Factor	Error Rate
20	60	0.05	14%
20	80	0.05	13%
25	50	0.05	9%
25	70	0.05	14%
25	50	0.1	9%
25	70	0.1	14%
25	70	0.5	13%
30	40	0.05	10%
30	60	0.05	11%

The error rates from the training and the testing stages are plotted on Figure 3-5. Both error rates dropped considerably in each re-training cycle. It appears to indicate that data used in the training stage are closely correlated with the data used in the testing stage. In fact, both samples were selected from the same population. When the result in the training stage improved, the result in the testing stage also improved.

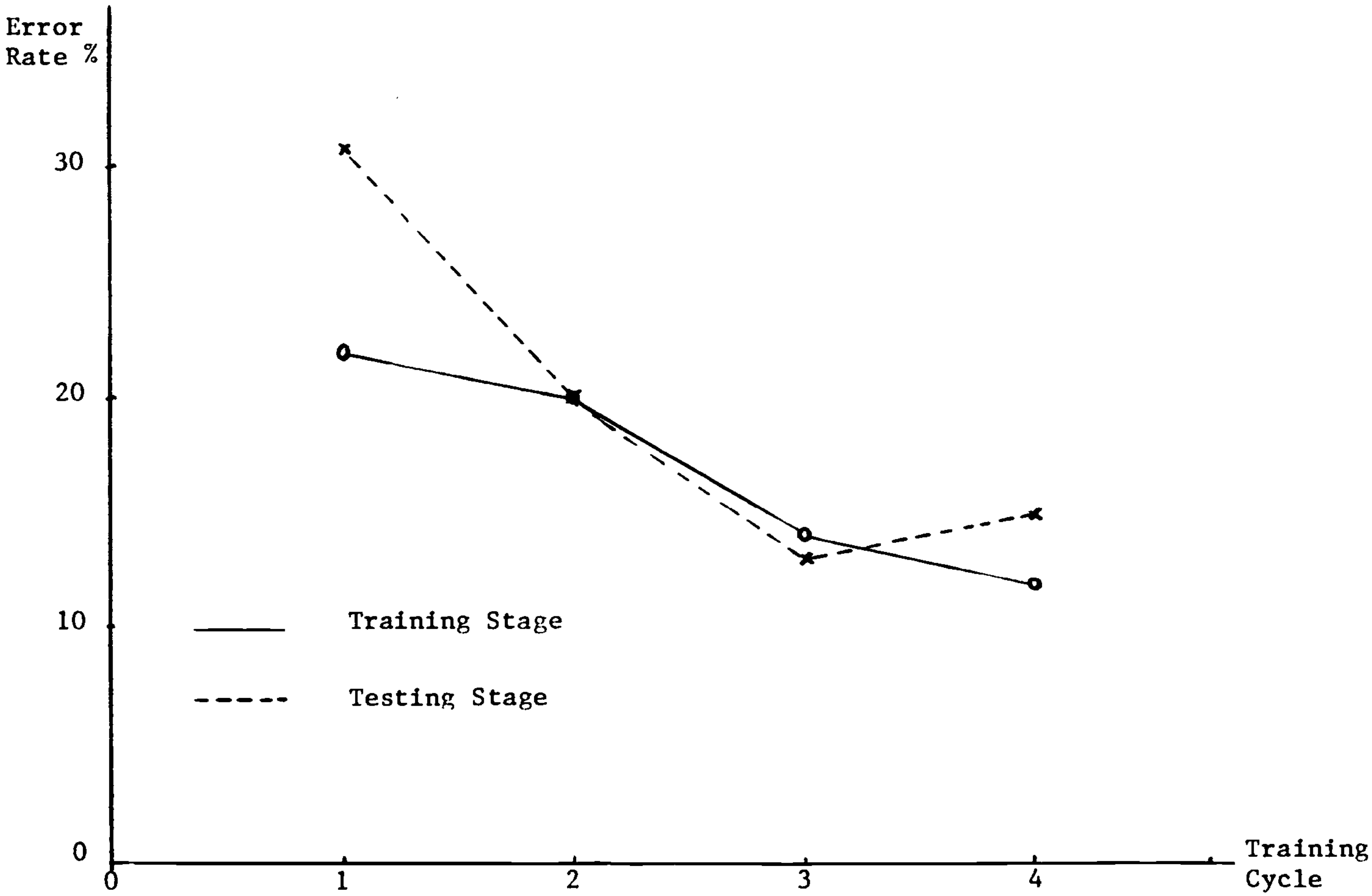

Figure 3-5 Variation of Error Rates in the Repeated Training Cycles: Buying Intent Problem

CHAPTER IV

SWAP EQUIPMENT COMMITTEE QUESTIONNAIRE

SWAP and SWAP Equipment Committee

To test out our hypothesis that the pattern recognition methodology is applicable to the analysis of consumer behaviors, a practical case had to be found. The survey questionnaires sent out by SWAP Equipment Committee provided an opportunity for such an endeavor. SWAP stands for Society for Wang Applications and Programs, and is the Wang Laboratories official user's society. It provides useful information and service to users of Wang Laboratories calculating, word processing, and computing equipments. There are approximately 1500 members of SWAP scattered throughout the world, contributing to and benefiting from this program exchange medium. Several committees are organized by SWAP to carry out its project works. Equipment Committee is one of those committees. The members of Equipment Committee contains people from different positions: a post officer, an attorney, an engineering manager, a consultant, and two university professors. The Committee Chairman is Dr. Michael S. Inoue; Professor of the Department of Industrial and General Engineering, Oregon State University.

Development of SWAP Equipment Committee Questionnaire

The idea of an Equipment Survey Questionnaire was generated shortly after the formation of the Equipment Committee in the spring of 1973. It came out of the spontaneous wish of the Equipment Committee members who simply wanted to let Wang Laboratories know the consensual wish of SWAP members regarding their equipment. The movement accelerated as a member (Kirtland H. Olson) designed the first draft of the SWAP questionnaire, and another member (James H. Cowden) volunteered to defray the cost of printing and mailing the questionnaires.

The Equipment Committee Chairman who is also the SWAP's delegate to the Joint Users Group (JUG) of the Association for Computing Machinery (ACM), attended a JUG meeting on June 4, 1973. Being the only users group representing programmable calculators, the SWAP was promptly charged with the task of investigating the needs and potentials of programmable calculators and their users. Thus, it was decided to make use of fact-finding questionnaires to serve not only the Wang users, but the data processing industry at large.

By July, a copy of the questionnaire was ready, incorporating all information that was thought pertinent to survey the present and potential needs of SWAP members regarding equipment. But the multi-page questionnaire was lengthy and cumbersome to answer. A new criterion was set to limit the questionnaire to no more than one page.

SWAP Membership No. 67-503-19215

SWAP EQUIPMENT COMMITTEE QUESTIONNAIRE

Mr. ☒ Mrs. ☐ Miss ☐ Ms ☐ Godorov (Last name) Edward J. (First name) Bennett Levin & Assoc. (Company) Phila. (City) Pa. (State) 19123 (Zip)

Job title Architectural Eng. Department Mechanical No. of persons in the department 7 Product or service provided Consulting Enging.

	Wang Equipment	Other Calculators	Computers	Time-share
1. (a) Please indicate the information about the computing machineries you have used at this premise:				
The models (how many of each)	(1) 72[illegible]C, 721, 722, 709	(4) Friden 1150		
The date you acquired them	October 1971	Oct. 1968, 69, 70		
The date you plan to dispose (or have disposed) of them	-	-		
Input devices†	K, T	K		
Output devices☆	DD, P	P		
The type of application◊	E	E		
Storage capacity☆	1984(S) 248(R)	5(R)		
Language compilers available•	None	None		
Average hours used per day	5	8		
Average hours of down time per month	1	1/8		
Number of persons using the hardware	2	15		
Time required to train a user	4 Hrs.	1/4 Hrs.		
The type of software you are using◊	E	None		
(b) The improvement you wish to have within 5 years:				
Additional input devices				
Additional output devices				
Additional compilers	X			
(c) What is the source of the software that you use?				
(i) Manufacturer-developed				
(ii) Private software company-developed				
(iii) User developed	X			

2. Information and comments on your calculator equipment:				
The keys you feel may be discarded to save cost	None			
The additional keys you wish to have	Sub Routine#			
The number of addressable registers available	248			
Do you feel the present number is enough?	Yes			
Comments on the present equipment's price	Too Rigid			
Comments on the size of the present equipment	None			
Comments of the weight of the present equipment	None			
How was user trained?	Self & Tewksbury			
How long to train user?	4 Hrs.			

† Keyboard (K), Magnetic tape cassette (M), Marked sense card reader (MC), Card reader (C), Typewriter console (T), Other (O). Describe ____

☆ Digital display (DD), Printer (P), Visual display (V).

◊ Engineering (E), Science (S), Education (ED), Business (B), Medical studies (M), Games and demonstrations (G), Tests and techniques (T).

☆ Bytes (B), Program steps (S), Storage registers (R).

• BASIC (B), FORTRAN (F), COBOL (C), PL/1 (P), RPG (R), APL (A), Other (O) Describe ____

(see other side)

Figure 4-1 SWAP Equipment Committee Questionnaire.

SWAP EQUIPMENT COMMITTEE QUESTIONNAIRE (Continued)

3. Please comment on the following statements concerning programmable calculators (A: agree, S: agree somewhat, D: disagree):
   (a) They will eventually perform a large portion of work now performed by computers __A__; mini-computers __A__; hand-held calculators __D__; time-sharing services __A__; accounting machines __A__; MT/ST __D__.
   (b) They will become compatible with computers __A__; telephone systems __A__; hand-held models __D__; time-sharing systems __A__; all calculators __D__.
   (c) Desk-top models will be replaced by hand-held models __S__.
   (d) Further work is needed to develop calculator keyboard in FORTRAN __D__; in BASIC __D__; IN APL __D__; or in ______.
   (e) Following peripherals are needed: built-in printer __D__; remote devices __A__; others ______.
   (f) Further work on software is required: error diagnostics and debugging __D__; compilers (which?) ______.
   (g) Users should be formally trained by the manufacturer __A__.
   (h) What future do you see for desk-top calculators __Good Future__.
   for hand-held calculators __Good Furture__.

4. General comments on SWAP __I am a new member and I have not evaluated the society.__

TO OUR FELLOW SWAP MEMBERS

The accompanying questionnaire has been prepared by Mr. Chai-hao Chang, an M.S. graduate in Computer Science from Oregon State University and presently a graduate student in Industrial Engineering, under the direction of the Swap Equipment Committee and Mr. Jason Taylor. This is an opportunity for you to express your opinion and to have our voice heard by Wang Laboratories. Please help us to attain this goal by completely filling out this sheet and returning it to Wang Laboratories in the enclosed self-addressed, post-paid envelope.

The returns will be statistically compiled and pertinent results will be presented at the SWAP Symposium. The committee also welcomes other suggestions and volunteers with ideas to work on the Committee. The Committee Chairman is Dr. Michael S. Inoue, Professor, Department of Industrial and General Engineering, Oregon State University, Corvallis, Oregon 97331 (503-754-3745 or -1645). Your cooperation is appreciated.

The SWAP Equipment Committee

Printed in U.S.A.
13-284

The final one-page draft was sent to all Equipment Committee members and to other key members. Their suggestions were taken into account, and the redesigned two-page questionnaire was sent to the Executive Secretary, Jason R. Taylor. Additional questions were added and the questionnaire (Figure 4-1) was finally distributed to all SWAP members in December, 1973 (Inoue and Chang, 1974).

Replies of Questionnaire

By March of 1974, some 200 questionnaires had been returned out of the estimated 1500 SWAP members. The replies came from 41 states in the United States and also from Washington, D. C., and Puerto Rico. Replies also included Canada, West Germany, Finland, Switzerland, New Zealand, England, etc. Questionnaires are returned to Wang Laboratories where they are first reviewed, and then forwarded to Oregon State University for data processing.

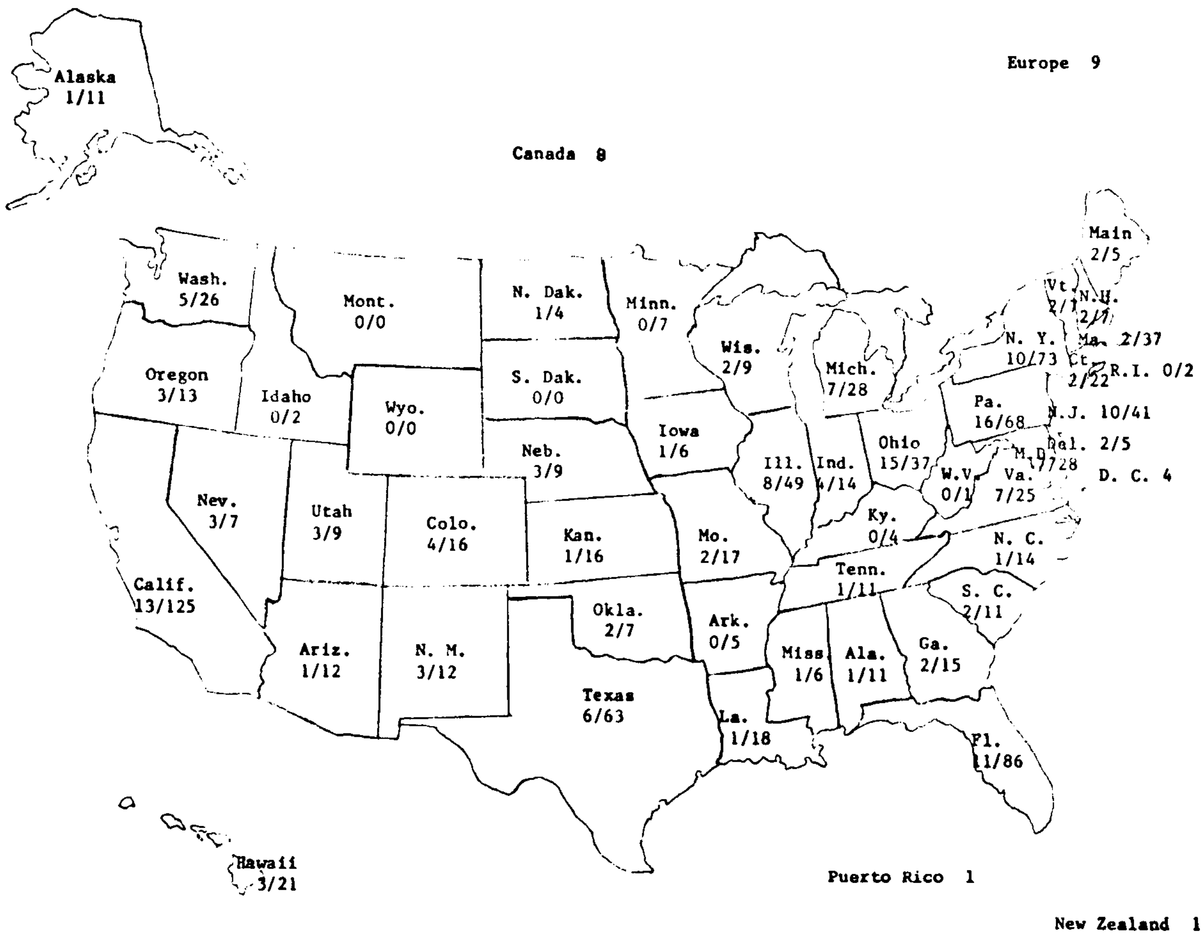

Figure 4-2 Geographical Distribution of Responses/Membership

CHAPTER V

SWAP CASE DATA ANALYSIS

File and Index Procedures

To process and analyze information from some 200 replies posed a major problem. Some questions were controlled opinion type (multiple-choice fillings) while other questions were open-ended. Most of the answers were not quantitative. Many of the replies left blanks to some questions. To edit the information, an alpha-numerical indexing method was adopted using the editing and indexing features of the Oregon State University's time-sharing system. The information on the questionnaire was represented by condensed codes and the data file was alphabetically indexed. The output was used for a statistical analysis.

Validity of Data

The validity of the representativeness of those data was tested by comparing data against information provided by Wang Laboratories.

Equipment Distribution

The distribution of responses on Wang equipment was compared to the SWAP membership equipment distribution. The chi-square value of 2.82 with six degrees of freedom based on 188 replies gave 80% confidence that two populations were homogeneous (Table 5-1).

```
Y(12)    = 2.6440E-02              +1.5919E-01 X( 9)
          +1.2389E-01 X(10)        +5.4172E-01 X(13)
          +2.5835E-01 X(14)
```

ENTERING F VALUE: 1.8150

DEGREES OF FREEDOM: 1 , 85

ANALYSIS OF VARIANCE TABLE

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE
TOTAL	89	2.35155556E 02	2.64219725E 00
REGRESSION	4	9.80373340E 01	2.45093335E 01
RESIDUAL	85	1.37118222E 02	1.61315555E 00

R SQUARED = .41690418

Figure 5-1 Regression Model for the 12^{th} Variable

TABLE 5-1. The Distribution of Responses on Wang Equipments.

Model Series	100/200/300	400	500	600	700	2000	Others
Survey	1.7%	1.7%	4.4%	21.2	63.1%	4%	3.1%
SWAP Membership	2.4%	2.2%	4.1%	23.4%	62.9%	3.36%	1.64%

$$\chi^2_{\nu=6} = 2.82$$

TABLE 5-2. The Distribution of Responses on Application

Fields	Engineering/Science	Business	Education	Miscellaneous
Survey	60.8%	17.5%	15.0%	6.8%
SWAP Membership	58.0%	16.0%	17.0%	9.0%

$$\chi^2_{\nu=3} = 2.06$$

Application Area Distribution

A similar test was given on the application area distribution. The chi-square value from this test could not give a conclusive evidence (Table 5-2).

Works of Local Majority Method

Choice of the Testing Problem

Most of the answers to the open-ended questions were difficult to be quantified. Unfortunately, the Local Majority Method that was chosen to demonstrate the applicability of Pattern Recognition technique to this marketing survey situation was designed to handle quantitative data. The part 3 of the questionnaire contained multiple type questions. Answers were limited to 'agree', 'agree somewhat', and 'disagree'. The relative weights of +2, +1 and -2 were given to these three answers, and weighted averages were used to compile consensus date. Fourteen comments were chosen to form patterns. These fourteen comments are on the following statements:

> Programmable calculators will eventually perform a large portion of work now performed by computer 1 , mini-computer 2 , hand-held calculators 3 , time-sharing services 4 , accounting machines 5 . Programmable calculators will become compatible with computers 6 , telephone systems 7 , hand-held models 8 , time-sharing systems 9 , all calculators 10 . Desk-top models will be replaced by hand-held models 11 .
> Further work is needed to develop calculator keyboard in FORTRAN 12 , in BASIC 13 , in APL 14 (Figure 4-1).

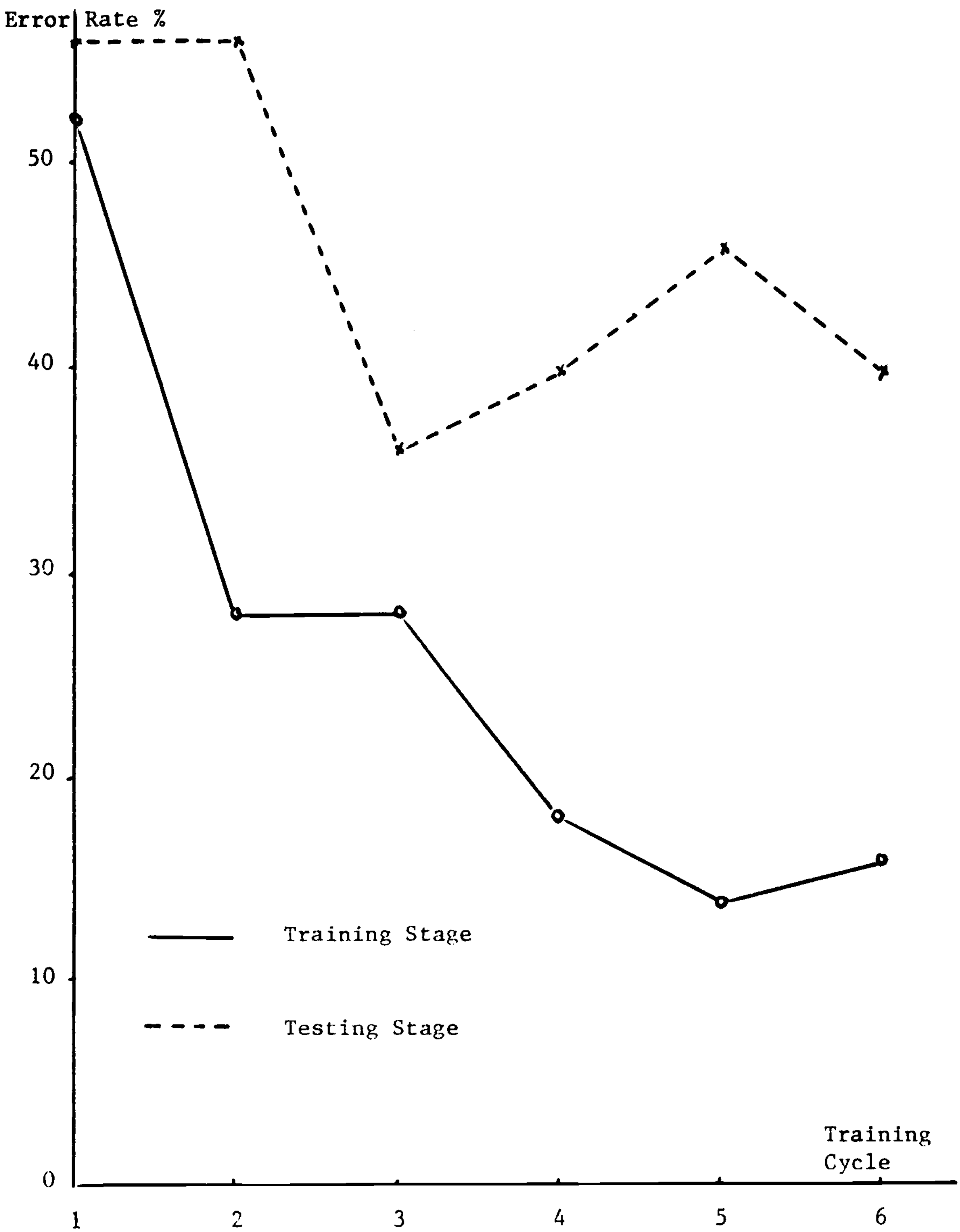

Figure 5-2 Variation of Error Rates in the Repeated Training Cycles: Using Eleven Variables to Classify Patterns

The dependency of each variable on other variables was tested by a regression analysis. They all showed a very low interdependency. The best collective correlation was demonstrated by the 12^{th} variable. For ninety data, when regressed on four best chosen variables, the value of R square test of 12^{th} variable was only .417 while the F test value already dropped to 1.815 (5% point for the distribution of F at 85 degrees of freedom is about 3.95). The best dependent variables for the 12^{th} variable were the 13^{th}, the 14^{th}, the 9^{th}, and the 10^{th} variable in that order. Figure 5-1 contains data for the final regression model.

Findings

Based on the value of the 12^{th} variable, data were classified into two groups, a series of experiments were performed to see the applicability of the Local Majority Method on studying such cases.

Experiment A. Similar to the previous experiments, several sets of data were used as inputs. The program gave satisfactory results as listed in Table 5-3. The average error rate was about 27%.

Experiment B. Although results in experiment A showed that the Local Majority Method could be used to classify these patterns, more details were sought about the performances of the learning process. Three sets of data were used in experiment B, and error rates in both stages were printed out. In each cycle the errors incurred during that training cycle were also printed. Whenever the

```
ORIGINAL PROTOTYPES OF TWO CLASSES

      F CLASS                                  G CLASS

2 2 1 1 1 1 1 1 1 1 1                  1 2-2 1 2 2 1-2 2-2 1
2 2 2 1 0 2 1-2 1 0-2                 -2 2-2 0 0 2-2 2-2 0-2
2 2 2 1 0 2 1-2 1 0-2                 -2 0-2 2 2-2 2-2 2 0-2
1 2 2 1 1 2 2 2 2 2 2                  1 2-2 1 2 2 1-2 2-2 1
-2 2-2 1 1 2 2-2 1 1 2                 2 2-2-2-2 2-2-2 2 1-2
1 2 2 2 2 2 2-2 2-2-2                 -2 2-2 0 0 2-2 2-2 0-2
1 2 2 2 2 2 2-2 2-2-2                  2 2-2 2 1 1 1-2 2-2-2
2 2-2 2 1 1 1-2 1 0-2                 -2 2 1 1-2 1 1-2 1 0 2
1 1 2-2 1 1 0 0 0 0-2                 -2 2 1 1-2 1 1-2 1 0 2
1 1 2-2 1 1 0 0 0 0-2                  1 2-2-2 2-2 2 2 1 2-2
-2-2-2 1-2-2-2 1 1 1 2                 2 2-2 2 1 1 1-2 2-2-2
2 2 1 1-2-2-2-2 2-2-2                  2 2-2 0 2 2 0 0 0 0 1
2 1-2 2-2 1-2-2 1 2-2                  2 2-2 0 2 2 0 0 0 0 1
1 2-2 0 0 2 2 1 2 1 1                  2 2-2 2 0 1 2-2-2-2 1
2 2 1 1-2-2-2-2 2-2-2                  2 2 2 2 2 2 2-2 2-2 1
2 1-2 2 2 2 2-2-2-2 2                  2 1 1 2 1 2 2 2 2 2-2
2 2 1 2 2 2 2-2 2-2 2                  1 2-2-2 2-2 2 2 1 2-2
2 2 1 2 2 2 2-2 2-2 2                  1 2-2 1 2 2 1-2 2-2 1
2 1-2 1-2 2 2 2 1 2-2                  1 2-2 1 2 2 1-2 2-2 1
2 2-2 2 2 2 2 0 2 0 2                  2 2-2 2 2-2-2-2-2-2 1
```

```
MODIFIED PROTOTYPES OF TWO CLASSES

F  1.20-0.03-3.25 1.25 1.61 1.91  .76   .33   .58-0.62-0.01
F   .94  .14  .55 1.25 1.10 2.63  .60-1.05  .70-1.31-1.72
F   .93  .14  .55 1.25 1.10 2.65  .56-1.05  .70-1.31-1.72
F   .69-2.16  .41 1.06 1.62 2.43 1.51  .01 1.16  .07 1.01
F -1.12-2.02-1.60 1.10 1.45 2.54 1.72-1.28  .83-0.28 1.04
F   .53  .18  .81 1.56 2.03 2.70 1.33-1.04 1.20-2.23-1.93
F   .55  .18  .81 1.50 2.05 2.70 1.33-1.04 1.20-2.23-1.83
F  1.00  .17-1.34 1.50 1.55 2.03  .70-0.79  .67-1.04-1.62
F   .52-0.58  .56-0.30 1.75 1.99  .49-0.91  .17-1.72-1.66
F   .55-0.58  .56-0.30 1.75 1.99  .49-0.91  .17-1.72-1.66
F -1.32-2.54-2.04  .42-0.41-0.30-0.66  .20  .69  .10 1.48
F  1.14  .11-0.05 1.39-0.30  .52-0.72-1.00 1.35-2.07-1.72
F  1.14-0.28-1.94 2.54-0.26 1.23-0.90-1.00  .72  .14-1.55
F   .50-0.07-1.06  .74  .91 2.45 1.41  .36 1.19-0.54  .23
F  1.14  .11-0.05 1.39-0.30  .52-0.76-1.00 1.35-2.07-1.72
F  1.03-0.06-1.59 1.61 1.92 2.55 1.63-1.07-0.68-1.85  .77
F   .94  .01  .05 1.64 1.90 2.65 1.50-1.20 1.29-2.34  .62
F   .94  .01  .05 1.64 1.90 2.65 1.50-1.20 1.29-2.34  .62
F  1.05-0.48-1.98 1.45-0.19 2.46 1.20  .43  .55-0.10-1.47
F   .92-0.02-1.75 1.71 1.69 2.48 1.39-0.14 1.70-1.34  .69
*************************************************************
G  1.57 2.73-0.63  .56 1.31  .34  .96-0.92 1.17-0.70  .24
G -0.56 2.75-0.57  .27  .19  .54-0.75 1.06-1.13  .40-1.16
G -0.56 1.49-0.85 1.50 1.35-1.51 1.59-1.10 1.36  .35-1.51
G  1.57 2.79-1.23  .34 1.31  .79  .96-0.92 1.17-0.70  .24
G  1.65 2.72-0.51-1.19-1.00  .79-0.96-1.15 1.39 1.04-1.49
G -0.56 2.75-0.57  .07  .19  .54-0.65 1.06-1.10  .45-1.16
G  2.11 2.77-0.91 1.41  .36-0.15 1.15-0.76 1.25-0.66-1.26
G -0.32 2.74  .79  .40-0.79  .09  .64-0.77  .54  .45  .65
G -0.32 2.74  .79  .40-0.79  .23  .64-0.77  .54  .45  .65
G  1.71 2.72-0.67-1.60 1.20-1.42 1.42 1.10  .62 1.47-1.51
G  2.11 2.77-0.11 1.41  .36-0.15 1.10-0.83 1.25-0.66-1.26
G  1.55 2.79-0.65  .29 1.14  .35  .36  .09  .11  .36  .13
G  1.55 2.79-0.65  .29 1.14  .30  .34  .09  .11  .76  .13
G  2.00 2.77-0.50 1.72  .36-0.10 1.47-1.30-0.96-0.55  .27
G  2.13 2.27 1.62 1.53 1.25  .53 1.05-0.70 1.44-0.75  .19
G  2.01 2.17 1.14 1.00  .21  .77 1.65 1.44 1.63 1.49-1.46
G  1.71 2.12-0.47-0.62 1.00-1.42 1.42 1.12  .52 1.47-1.51
G  1.57 2.73-0.63  .34 1.31  .35  .96-0.92 1.17-0.70  .24
G  1.57 2.73-0.63  .56 1.31  .34  .96-0.92 1.17-0.70  .24
G  1.57 2.76-1.97 1.34 1.45-1.77-0.74-1.34-1.15-1.03  .37
```

FIGURE 5-3. Modified Prototype Points.

testing results did not appear to be satisfactory, the user could order the program to return to the training stage.

Comparing with all other variables, the value of the 13^{th} and the 14^{th} variables of most replies were found to be zero. This meant that very few replies ever gave the comments on those statements. It was decided to drop both the 13^{th} and the 14^{th} variables. Some of the results are listed on Table 5-4. The error rates in training stages (the solid line in Figure 5-2) decreased significantly with the number of cycles, while the error rates in testing stages were not significantly affected (the dashed line in Figure 5-2).

Experiment C. The possibility of using fewer variables to classify patterns was tested in this experiment. The prototype points were modified using both the testing and the training samples from the previous experiments as the new training samples. After only a few modifications, the 2^{nd} and the 6^{th} variables of the prototype points of two categories were clearly distinguished between the F and the G categories (Figure 5-3). For example, the 2^{nd} variable has values that are less than +0.2 for F and greater than +1.4 for G. Since the categories were formed based on the value of the 12^{th} variable, it indicated that the 12^{th} variable has a tendency to depend on the 2^{nd} and the 6^{th} variables. The regression analysis of the 12^{th} variable versus the rest of the variables also provided a similar result. The best dependent variable for the 12^{th} variable is the 6^{th} variable, while the 2^{nd} variable was the second best.

Taking these two variables as the only variables to classify patterns, similar experiments were performed. Figure 5-4 showed the variation of error rates made in both stages. The results were comparatively worse than using 11 variables, but the computer time used for the same number of cycles was reduced to about one-half.

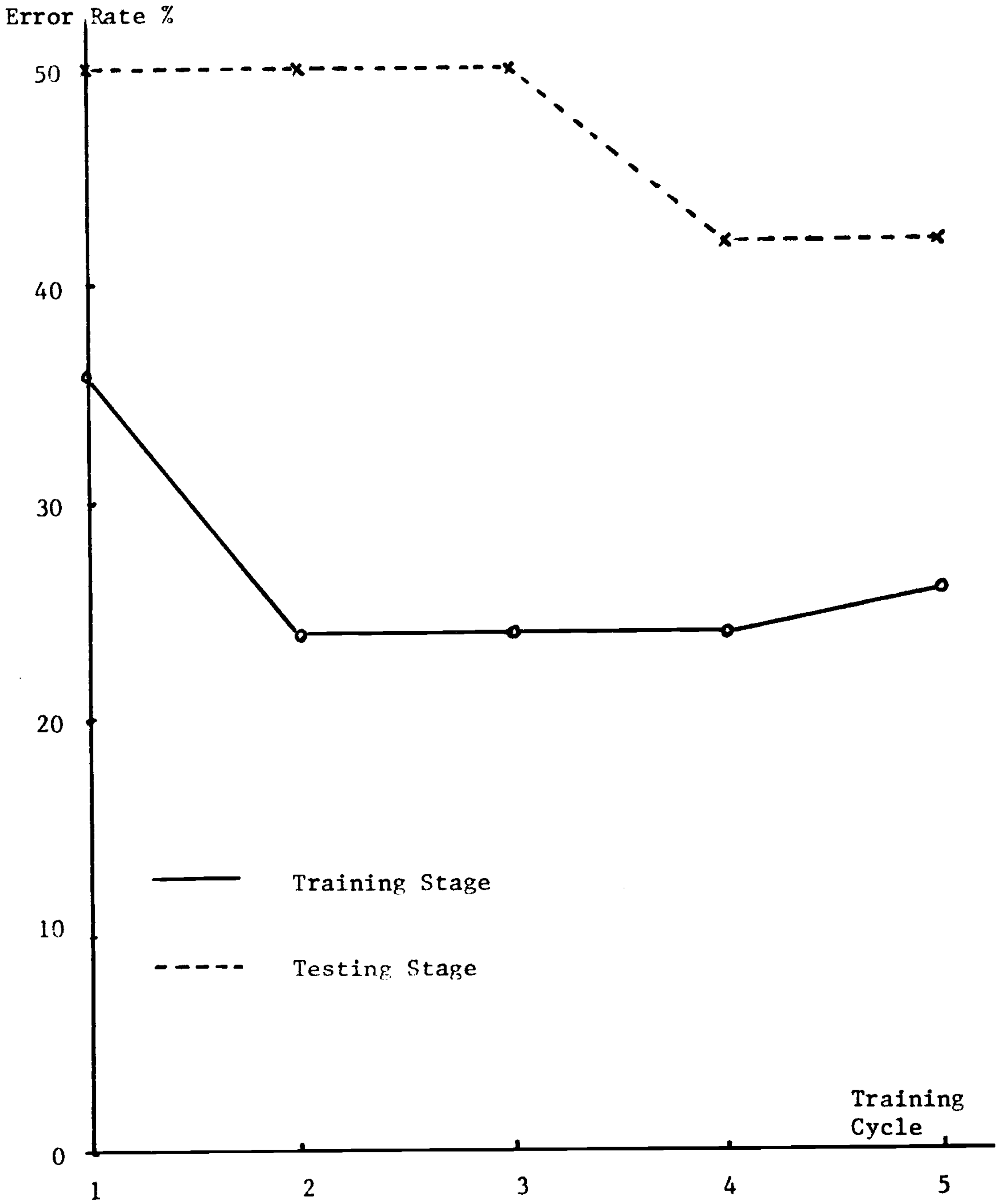

Figure 5-4 Variation of Error Rates in the Repeated Training Cycles: Using the 2nd and the 6th Variables to Classify Patterns

TABLE 5-3. Results From Experiment A

Number of Prototype Points Each Category: 20
Increment Value in Time Factor: 0.2

Set Number	Number of Training Samples	Error Rate in Testing Stage
1	50	11/50
2	50	15/50
3	50	14/50
4	70	11/50
5	70	17/50
6	70	12/50

TABLE 5-4. Results From Experiment B

Number of Prototype Points Each Category: 20
Number of Training Samples: 50
Number of Testing Samples: 50

SET 1: Using 13 variables to make classifications.
Increment Value in Time Factor 0.1.

Training Cycles	Error Rate in Training Stage	Error Rate in Testing Stage
Cycle 1	18/50	14/50
Cycle 2	6/50	12/50
Cycle 3	6/50	16/50

SET 2: Using 12 variables to make Classifications.
Increment Value in Time Factor 0.1.

Training Cycles	Error Rate in Training Stage	Error Rate in Testing Stage
Cycle 1	17/50	16/50
Cycle 2	6/50	9/50
Cycle 3	7/50	16/50

SET 3: Using 11 Variables to Make Classifications
Incremental Value in Time Factor 0.1.

Training Cycles	Error Rate in Training Stage	Error Rate in Testing Stage
Cycle 1	20/50	19/50
Cycle 2	8/50	22/50
Cycle 3	8/50	20/50

SET 4: Using 11 Variables to Make Classifications
Incremental Value in Time Factor: 0.05.

Training Cycles	Error Rate in Training Stage	Error Rate in Testing Stage
Cycle 1	23/50	27/50
Cycle 2	11/50	20/50
Cycle 3	9/50	22/50
Cycle 4	6/50	23/50
Cycle 5	7/50	21/50

SET 5: Using 11 Variables to Make Classifications
Increment Value in Time Factor 0.02.

Training Cycles	Error Rates in Training Stage	Error Rate in Testing Stage
Cycle 1	26/50	28/50
Cycle 2	14/50	28/50
Cycle 3	14/50	18/50
Cycle 4	9/50	20/50
Cycle 5	7/50	23/50
Cycle 6	8/50	20/50

CHAPTER VI

CONCLUSION AND RECOMMENDATION

Evaluation of Experimental Results

The effectiveness of an adaptive pattern recognition technique was demonstrated in Chapters III and V, by subjecting a computer model of the Local Majority Method, "LOMAME", to both physical and behavioral patterns.

The "A-and-R" and the "1101" problems in Chapter III represented typical physical patterns. Table 3-1 and 3-2 show the "LOMAME" to perform satisfactorily as an adaptive pattern recognizer.

The data from SWAP equipment survey (Chapter IV) provided an opportunity to utilize this adaptive pattern recognizer in Chapter V to study behavioral patterns of data processing equipment users. Out of six sets of 50 to 70 samples each, the error rates of 22 percent to 34 percent (Table 5-3) were incurred in classifying patterns of customers opting for a FORTRAN calculator keyboard rather than a BASIC calculator keyboard (Wang, 2200). The error rates are relatively small in comparison to the expected 50 percent.

The increased difficulty of classifying patterns of human business behaviors was foreseen. There are many so-called "unquantifiable" and "intangible" factors entering human decision-making processes that either have not been represented by the data collected or have not been correctly interpreted numerically. Consequently, identifying a decision criterion that can work for all individual characteristics and emotional states seemed to be an impossible task.

Nonetheless, the comparison of the experimental results (Figure 3-5) with the SWAP data (Figure 5-2) illustrated considerable adaptive improvements made by LOMAME through the training period. During testing cycles following repeated training periods, the error rates decreased considerably (from 32 percent to 16 percent) in the buying-intent problem (Figure 5-3). A proper selection of the prototypes greatly influenced the effectiveness of the testing operation. However, even a poor set of prototypes, such as the one producing 56 percent error in the SWAP run (Appendix II) did initially benefit from the repeated application of the training samples (56%, 56%, 36%, 40%, 46%, and 40% error rates in Figure 5-2). No noticeable improvement, and even some deterioration, was observed after the third application of the same set of training data. The stability of data was accomplished in the buying intent case through an assumption that two parent populations had mutually exclusive characteristics. In effect, the patterns were so rigidly categorized as to make them resemble patterns created by a physical process followed some physical principles. Data from SWAP, on the other hand, were answers to questions with uncertain interrelationships. The statistical correlation of FORTRAN keyboard to four other factors (factors 9, 10, 13, and 14 in Figure 4.1 concerning compatibility and other keyboard selection) was $r^2 = .417$ for the regression model (Figure 5-1). The increased exposure to the same training patterns led to better results in classifying the training samples themselves, but not necessarily in classifying new testing samples (Appendix II). More specifically Table 5-4 (set 5) demonstrated that the largest improvement was from 26 to 8 in training errors, but only from 28 to 20 in testing errors.

The Deficiencies of the Local Majority Method

The Local Majority Method was chosen as the adaptive pattern recognition process to be experimented with the classification of behavioral data and was not meant to be the best adaptive technique for such an application. For a particular problem, there may exist a criterion to evaluate which technique is the best, but in general each technique has its own strengths and deficiencies.

The discriminant function based on the local majority rule is more complex than the minimum distance method or the Fix and Hodges method (Chapter III). The inclusion of the correction factor for the decision criterion (the discriminant function) renders the Local Majority Method to be more complicated and requiring additional computing time.

As stated in Chapter II, the correction algorithm plays an important role during the learning stage. The correction term used in the Local Majority Method is based on some intuitive assumptions. In fact, the corrections appeared in the experiment were not satisfactory. The same incorrect classifications were repeated when training cycles were reapplied with the same data set.

More in-depth understanding is needed to choose the time increment factor, the number of training samples, and the number of prototype points. The choice of suitable variables to form a pattern poses another difficulty (Nilsson, 1965). The choice of the eleven variables in the SWAP case study as entirely based on the author's own intuition. Though this technique gave some indications on which variables had significant influences on classification (experiment C), it did not show which variables should be rejected as meaningless indicators of the classification.

Recommendations for Further Studies

Selective and Rejective Techniques

The pattern recognition technique demonstrated in this thesis was an adaptive technique. In classifying categories that are highly complex and not linear separable will have the following deficiencies: (1) they require a large memory space, (2) they are very time-consuming in classifying testing patterns. In order to alleviate these deficiencies, the selective method is recommended. The selective method may spend much more time in searching for a simple criterion of discrimination, but it is more efficient during the recognition stage. The main deficiency of the selective method is that much information will be lost from a strong selection. The experiment C (p. 50) in the SWAP case study showed a similar result: the computing time was saved by using fewer variables, but the error rate did increase.

The rejective method, in contrast to the selective method, starts out with a large pool of prototypes and chooses variables by obtaining the selective set by rejecting the undesired variables. Both the selective and the rejective technique can be used to identify an empirical or a logical expression by analyzing a large number of multivariate data.

Combination of the Techniques

Although an adaptive technique has many deficiencies, it has shown an aptitude for emulating group learning processes. This feature can be useful in studying non-stationary cases.

If a technique can determine a simple decision criterion which is able to adapt to further analysis, it will be an ideal tool for behavior science studies.

Potential Areas for Further Studies

As introduced in Chapter II, a pattern recognition is a combination of two processes. The learning process is to search criteria of discrimination, while the recognition process uses those criteria to classify patterns. The assumption is that if the criteria work well in the learning stage, they should also work well in the recognition stage, and that no further adaptive work is necessary. We have found that this assumption does not work well for behavior patterns as demonstrated by the SWAP case study.

Yet, the adaptive character of the 'learning process' to search and reinforce suitable criteria could provide a potential tool for further studies in most areas of behavior science. The Table 6-1 summarizes the areas and possible fields of applications where an adaptive pattern recognition process may be profitably used.

TABLE 6-1. Potential Areas for Further Studies

Areas	Possible Applications
Marketing Research	Consumers' behavior; motivation analysis; sales forecasting; advertisement study.
Industry	Learning processes; productivity and motivation analysis.
Medical Science	Behavior modification for mentally retarded individuals; recreational therapy.
Psychology	Applied behavior analysis.
Education	Teaching and learning effects.
Politics	Election processes; representative voices in organizations.

This study was undertaken with the objective of identifying the applications of pattern recognizers to behavioral problems. Though more study is needed to comparatively evaluate the advantages of Pattern Recognition techniques over the more traditional statistical approach, what we have found seems to indicate that the pattern recognizers themselves are limited in direct applications. Perhaps the learning process that is associated with the training of prototype points holds greater promises in modeling dynamic behaviors. This could lead to future studies to model human decision-makers with individual sets of decision making characteristics. The metamorphism of a group of individual decision-makers (e.g., political candidates, jurors, etc.) into an integrated decision-making body (e.g., a congressional committee, jury, etc.) could be simulated through an adaptive pattern recognition process undergoing a continuous training stage.

From this new perspective, the contribution that the artificial intelligence studies could make on Industrial Engineering problems appear unlimited. Pattern Recognition may become a new tool to simulate the dynamic behaviors of man-machine systems - the ones in which the human beings never cease their learning process.

BIBLIOGRAPHY

Aaker, David A. 1971. Multivariate Analysis in Marketing Theory and Application. Belmont, Wadsworth. 358 p.

Bongard, Mikhail M. 1970. Pattern Recognition. Spartan. 253 p.

Boyd, Harper W., Jr., and Westfall, Ralph. 1966. Marketing Research. Homewood, Irwin. 791 p.

Casey, Richard G. and Nagy, George. 1971. Advances in Pattern Recognition. Scientific American. Vol. 224, No. 4:56-71.

Fischler, Martin A. and Elschlager, Robert A. 1973. The Representation and Matching of Pictorial Structures. IEEE Trans. on Comp. Vol. C-22, No. 1:678-688.

Fu, King Sun. 1971. Pattern Recognition and Machine Learning. New York, Plenum Press. 343 p.

Gagliardo, Emilo. 1973. ABIOSOFOS 1972 (English Translation). Oregon State University. Technical Report No. 47. 62 p.

Harlow, Charles A. 1971. Feature Extraction in Images. IEEE Systems, Man and Cybernetics Group Annual Symposium Record. pp. 126-131.

Harlow, Charles A. and Eisenbeis, Sharon A. 1973. The Analysis of Radiographic Images. IEEE Trans. on Comp. Vol. C-22, No. 7:678-689.

Kawamura, Joseph G. 1971. Automatic Recognition of Changes in Urban Development from Aerial Photographs. IEEE Trans. on Systems, Man and Cybernetics. Vol. SMC-1, No. 7:230-239.

Keith, Robert J. 1960. The Marketing Revolution. Journal of Marketing. Vol. 24:35-38.

Kolers, Paul A., and Eden, Murray. 1968. Recognizing Patterns. Cambridge, MIT Press. 237 p.

Kotler, Philip. 1972. Marketing Management. Englewood Cliffs, Prentice-Hall. 885 p.

Kulikowski, Casimir A. 1969. Pattern Recognition Approach to Medical Diagnosis. Record of the 1969 IEEE Systems Science and Cybernetics Conference. pp. 198-206.

Lawrence, Peter D. and Lin, Wen-chun. 1972. Statistical Decision Making in the Real-Time Control of an Arm Aid for the Disable. IEEE Trans. on Systems, Man and Cybernetics. Vol. SMC-2, No. 1:35-42.

Lotshaw, Elmer P. 1970. Industrial Marketing: Trends and Challenges. Journal of Marketing. Vol. 34:22-24.

McMillian, Claude and Gonzalez, Richard F. 1968. Systems Analysis. Irwin. 516 p.

Meisel, William S. 1972. Computer Oriented Applications to Pattern Recognition. Prentice-Hall. 250 p.

Nagy, George. 1968. State of the Art in Pattern Recognition. Proceeding of the IEEE. Vol. 56, No. 5:836-862.

Nilsson, Nils J. 1965. Learning Machines. McGraw-Hill. 137 p.

Northouse, Richard A. and Fu, King-Sun. 1973. Dynamic Scheduling of Large Digit Computer Systems Using Adaptive Control and Clustering Techniques. IEEE Trans. on Systems, Man and Cybernetics. Vol. SMC-3, No. 3:225-234.

Oda, Moriya, and et al. 1971. A Pattern Recognition Study of Palm Reading. IEEE Trans. on Systems, Man and Cybernetics. Vol. SMC-1, No. 4:171-174.

Pang, Chok K., Koivo, Antti J. and El-Abiad, Ahmed H. 1973. Application of Pattern Recognition to Steady-State Evaluation in a Power System. IEEE Trans. on Systems, Man and Cybernetics. Vol. SMC-3, No. 6:622-631.

Sakai, T. and Doshita, S. 1963. Automatic Speech Recognition System for Conversational Sound. IEEE Trans. on Elec. Comp. Vol. EC-12:835-846.

Schultz, William J. 1972. An Outline of Marketing. Littlefield, Adams. 236 p.

Sebestyen, George S. 1962. Decision-making Process in Pattern Recognition. New York, Macmillan. 162 p.

Sheth, Jadgish N. 1971. The Multivariate Revolution in Marketing Research. Journal of Marketing. Vol. 35:13-19.

Specht, D. F. 1967. Vectorcardiographic Diagnosis Using the Polynomial Discriminant Method of Pattern Recognition. IEEE Trans. on Bio-Medical Engineering. Vol. BME-14:90-95.

Twedt, Dik W. 1968. 1968 Survey of Marketing Research. American Marketing Associations. 70 p.

APPENDICES

APPENDIX A

COMPUTER PROGRAM LOMAME

```
033 FORTRAN  VERSION 3.12            05/17/74  2331

      PROGRAM LONAME
      DIMENSION F(40,50),G(40,50),V(50),DVF(40),DVG(40),RTODVF(40),
     1RTODVG(40),DISVF(40),DISVG(40),NER(40)
C
C THE OBJECT OF THIS PROGRAM IS TO VARIFY THE USAGE OF THE LOCAL
C MAJORITY METHOD IN CLASSIFYING TWO CATEGORIES
C
C NP..NUMBER OF PROTOTYPES EACH CLASS
C ND..NUMBER OF DIMENSION
C NS..NUMBER OF TRAINING SAMPLES
C NTS..NUMBER OF TESTING SAMPLES
C
      NP=TTYIN(5HNP=  )
      ND=TTYIN(5HND=  )
      NS=TTYIN(5HNS=  )
      NTS=TTYIN(6HNTS=  )
      TINCRE=TTYIN(9HICR=      )
      NSA=0
      NSPNTS=NS+NTS
C
C LIST THE ORIGINAL PROTOTYPES
C
      WRITE(3,6)
    6 FORMAT(1H1//,5X,*ORIGINAL PROTOTYPES OF TWO CLASSES*,//,
     110X,*F CLASS*,25X,*G CLASS*,/)
C
C READ IN THE SELECTED ORIGINAL PROTOTYPES OF TWO CLASSES
C
      DO 10 I=1,NP
      READ(60,2) (F(I,J),J=1,ND)
      READ(60,2) (G(I,J),J=1,ND)
    2 FORMAT(11F2.0)
      WRITE(3,7) (F(I,J),J=1,ND),(G(I,J),J=1,ND)
    7 FORMAT(2X,11F2.0,7X,11F2.0)
   10 CONTINUE
C
C MAKE TRAINING SAMPLES A NEW FILE IN LUN 1
C
      DO 80 K=1,NSPNTS
      READ(60,2) (V(J),J=1,ND)
   80 WRITE(1,13) (V(J),J=1,ND)
      REWIND 1
C
C INITIAL VALUES
C
C TFCTOR..TIME FACTOR
C
      TFCTOR=1.
C
C GORF..CATEGORY G OR CATEGORY F
C
 3000 GORF=1.
      IJK=0
      NETR=0
      WRITE(2,4) NP,NS,NTS
    4 FORMAT(1H1,//,5X,*NUMBER OF PROTOTYPES EACH CLASS*,2X,I3,/
     1,5X,*NUMBER OF TRAINING SAMPLES*,2X,I3,/,5X,*NUMBER OF *
     2*TESTING SAMPLES*,2X,I3,//,13X,*DATA*,12X,*CATEGORY*,/)
C
```

```
      OS3 FORTRAN  VERSION 3.12 LUMAME     05/17/74  2331

C NERROR..NUMBER OF ERRORS
C
      NERROR=0
C
C READ IN THE TRAINING SAMPLE V
C
 1000 READ(1,13) (V(J),J=1,ND)
   13 FORMAT(11F4.0)
      IJK=IJK+1
      GORF=-GORF
      VV=0
      DO 30 J=1,ND
      VV=VV+V(J)*V(J)
   30 CONTINUE
      SUMD=0
C
C EVALUATE THE DISTANCES BETWEEN THE TRAINING ONE AND PROTOTYPES
C
      DO 40 I=1,NP
      DVF(I)=VV
      DVG(I)=VV
      DO 50 J=1,ND
      DVF(I)=DVF(I)-2.*V(J)*F(I,J)+F(I,J)*F(I,J)
      DVG(I)=DVG(I)-2.*V(J)*G(I,J)+G(I,J)*G(I,J)
   50 CONTINUE
C
C DISVF..DISTANCE BETWEEN V AND F
C
      DISVF(I)=SQRT(DVF(I))
C
C DISVG..DISTANCE BETWEEN V AND G
C
      DISVG(I)=SQRT(DVG(I))
      SUMD=SUMD+DISVF(I)+DISVG(I)
   40 CONTINUE
      RNP=NP
C
C FIND THE AVERAGE DISTANCE
C
      AVGD=SUMD/(2.*RNP)
C
C VAVGD IS THE MODIFIED AVERAGE DISTANCE
C
      VAVGD=AVGD*20./(20.+RNP)
      FCLASS=0
      GCLASS=0
      DO 60 I=1,NP
      RTODVF(I)=DISVF(I)/VAVGD
      RTODVG(I)=DISVG(I)/VAVGD
      FCLASS=FCLASS+1./(1.+RTODVF(I)**2)
      GCLASS=GCLASS+1./(1.+RTODVG(I)**2)
   60 CONTINUE
C
C TEST IF IT IS THE END OF TRAINING PERIOD
C
      IF(IJK.GT.MS) GO TO 100
      DIFFG=GORF*(FCLASS-GCLASS)
      IF(DIFFG.LE.0.) GO TO 1000
      NETR=NETR+1
```

```
C
C ADJUSTMENTS TO THE PROTOTYPE POINTS
C
      FACTOR=1.5*GORF*DIFFG/(1.+DIFFG)
      FACTOR=FACTOR*AVGD/TFCTOR
      DO 70 I=1,NP
      DO 71 J=1,ND
      F(I,J)=F(I,J)-(V(J)-F(I,J))*FACTOR/(AVGD+DISVF(I))
      G(I,J)=G(I,J)+(V(J)-G(I,J))*FACTOR/(AVGD+DISVG(I))
   71 CONTINUE
   70 CONTINUE
      TFCTOR=TFCTOR+TINCRE
      NER(NETR)=IJK
      GO TO 1000
C
C TO CLASSIFY THE UNKNOWN PATTERN
C
  100 DIFFG=FCLASS-GCLASS
      DG=DIFFG*GORF
      IF(DG.GT.0) NERROR=NERROR+1
      IRSULT=1
      IF(DIFFG.GE.0.) GO TO 200
      IRSULT=2
  200 WRITE(2,300) (V(J),J=1,ND),IRSULT
  300 FORMAT(/,1X,11F2.0,6X,I1)
      IF(IJK.EQ.NSPNTS) GO TO 400
      GO TO 1000
C
C THE NUMBER OF ERRORS IN THIS TEST
C
  400 WRITE(2,5) NERROR,NTS
    5 FORMAT(/,5X,≠THERE ARE ≠,I3,≠ ERRORS IN CLASSIFYING ≠,
     1I3,≠ TESTING SAMPLES≠)
C
C TESTING IF THE RESULTS FROM TRAINING STAGE ARE SATISFACTORY
C
      WRITE(3,8)
    8 FORMAT(//,5X,≠MODIFIED PROTOTYPES OF TWO CLASSES≠,//)
      DO 90 I=1,NP
      WRITE(3,9) (F(I,J),J=1,ND)
    9 FORMAT(2X,≠F  ≠,11F5.2)
   90 CONTINUE
      WRITE(3,10)
   10 FORMAT(2X,≠****************************************≠
     1≠********************≠)
      DO 91 I=1,NP
      WRITE(3,11) (G(I,J),J=1,ND)
   11 FORMAT(2X,≠G  ≠,11F5.2)
   91 CONTINUE
      WRITE(61,12) NS,NETR,NTS,NERROR
   12 FORMAT(/,1X,≠IN CLASSIFYING ≠,I3,≠ TRAINING SAMPLES ≠,1X,
     1I2,≠ ERRORS WERE MADE.≠,/,≠ IN CLASSIFYING ≠,I3,≠ TESTING≠,
     2≠ SAMPLES ≠,I2,≠ ERRORS WERE MADE.≠,/,≠ PLEASE INPUT 01 TO≠
     3≠ NSA IF YOU ARE SATISFIED, ≠,/,≠ OR INPUT 02 TO BACK TO≠
     4≠ TRAINING STAGE.≠,/)
      WRITE(61,15) (NER(K),K=1,NETR)
   15 FORMAT(≠ THE ERRORS IN TRAINING PERIOD WERE AT:≠,/,1X
     1,20I3,/1X,20I3)
      NSA=TTYIN(6HNSA=  )
```

```
        033 FORTRAN  VERSION 3.12 LOMAME      05/17/74  2331

0+178            IF(NSA.EQ.01) GO TO 1001
0+179            REWIND 1
0+180            GO TO 3000
0+181      1001  STOP
0+182            END


 NO ERRORS FOR LOMAME
LENGTH OF SUBPROGRAM        22377
```

APPENDIX B

OUTPUT PRINTS

```
#EQUIP,1=FILE
#EQUIP,2=LP
#EQUIP,3=2
#UNEQUIP,60
#EQUIP,60=*SWAP7
#*LPLABEL,2/SAVE FOR CHANG, CHIA-HAO

#LABEL,2/DATA=*SWAP7, 11 VARIABLES, 140 PATTERNS

#FORTRAN,I=*LOMAME1,R


 NO ERRORS FOR LOMAME
RUN




NP= 20
ND= 11
NS= 50
NTS=50
ICR=0.02

IN CLASSIFYING  50 TRAINING SAMPLES  26 ERRORS WERE MADE.
IN CLASSIFYING  50 TESTING SAMPLES 28 ERRORS WERE MADE.
PLEASE INPUT 01 TO NSA IF YOU ARE SATISFIED,
OR INPUT 02 TO BACK TO TRAINING STAGE.

THE ERRORS IN TRAINING PERIOD WERE AT:
  3  4  5  6  7  8  9 11 12 13 14 15 16 19 20 25 26 27 31 32
 38 39 40 41 44 50
NSA=02

IN CLASSIFYING  50 TRAINING SAMPLES  14 ERRORS WERE MADE.
IN CLASSIFYING  50 TESTING SAMPLES 28 ERRORS WERE MADE.
PLEASE INPUT 01 TO NSA IF YOU ARE SATISFIED,
OR INPUT 02 TO BACK TO TRAINING STAGE.

THE ERRORS IN TRAINING PERIOD WERE AT:
  3  4  5  6 11 16 25 26 31 33 38 40 41 50
NSA=02

IN CLASSIFYING  50 TRAINING SAMPLES  14 ERRORS WERE MADE.
IN CLASSIFYING  50 TESTING SAMPLES 18 ERRORS WERE MADE.
PLEASE INPUT 01 TO NSA IF YOU ARE SATISFIED,
OR INPUT 02 TO BACK TO TRAINING STAGE.

THE ERRORS IN TRAINING PERIOD WERE AT:
  1  5  6 11 16 21 24 25 33 34 38 40 41 50
NSA=02
```

```
IN CLASSIFYING  50 TRAINING SAMPLES   9 ERRORS WERE MADE.
IN CLASSIFYING  50 TESTING SAMPLES 20 ERRORS WERE MADE.
PLEASE INPUT 01 TO NSA IF YOU ARE SATISFIED,
OR INPUT 02 TO BACK TO TRAINING STAGE.

THE ERRORS IN TRAINING PERIOD WERE AT:
   5  6 21 24 33 38 40 41 50
NSA=02

IN CLASSIFYING  50 TRAINING SAMPLES   7 ERRORS WERE MADE.
IN CLASSIFYING  50 TESTING SAMPLES 23 ERRORS WERE MADE.
PLEASE INPUT 01 TO NSA IF YOU ARE SATISFIED,
OR INPUT 02 TO BACK TO TRAINING STAGE.

THE ERRORS IN TRAINING PERIOD WERE AT:
   5  6 19 33 38 40 41
NSA=02

IN CLASSIFYING  50 TRAINING SAMPLES   8 ERRORS WERE MADE.
IN CLASSIFYING  50 TESTING SAMPLES 20 ERRORS WERE MADE.
PLEASE INPUT 01 TO NSA IF YOU ARE SATISFIED,
OR INPUT 02 TO BACK TO TRAINING STAGE.

THE ERRORS IN TRAINING PERIOD WERE AT:
   5  6 21 33 37 38 40 41
NSA=01




END OF FORTRAN EXECUTION
```