Strategic Decision-making Processes of NPD by Hybrid Classification Model Techniques

You-Shyang Chen[1], Jieh-Ren Chang[2], Chyi-Jia Huang[1]

[1] Department of Information Management, Hwa Hsia University of Technology, Taiwan
[2] Department of Electronic Engineering, National Ilan University, Taiwan
ys_chen@cc.hwh.edu, jrchang@niu.edu.tw, jerry@grand-tek.com

Abstract

Although a successful new product development (NPD) can yield significant profits, it is a high-expenditure and high-risk investment. Thus, the related issues of the NPD success/failure are concerned by practitioners and academicians. With this reason, the research provides a NPD decision-making method to reduce the investment risk when implementing NPD projects for companies. The research develops a hybrid classification model to use a variety of data mining classification techniques, such as Bayes Net, Lazy Learner, Bagging-Bootstrap Aggregating and Decision Trees, to help companies make decision simply and accurately. A total amount of 151 NPD data from 25 companies is collected for further empirical validation to the proposed model. The empirical results show that the critical success factors (CSFs) of influencing NPD are determined as the quantity of products, the product life cycle, the sales territory, project scale and capital. Lastly, the NPD knowledge-based decision-making system is determined based on the execution of Decision Trees. Furthermore, the internet decision-making system has been verified in this research by the use of a series of NPD data from an actual company A as an effective reference for estimating the efficiency of classifying a NPD project in the future.

Keywords: New Product Development (NPD), Decision trees, Critical Success Factor (CSF), Knowledge-Based System (KBS), Project

1 Introduction

Griffin and Hauser [1] indicated that for different industries, New Product Development (NPD) is the common and significant long-term profit making method. NPD has important influence for the firm sustainable operation; once a firm decides to put into NPD, the firm shall invest many manpower and resources, which would be the heavy cost burden for the firm. Thus, for the NPD, the firms with good or bad system would firstly weigh the existing resource status of the firm, and then decide to invest how many resources into the NPD. There is no firm with infinite resources in the world, thus the efficient utilization and investment of resources are the key topic for discussion of any firm. It is known from many literatures and data that the researches on the critical factors for the success of NPD were mainly the researches on Product Development Process Factors, Strategic Factors, Market Environment Factors, Organizational Factors, and Firm's Internal Environment Factors in the product development; besides, the aspects aiming at the sale, marketing, R&D, and manufacture in the product development stage are also the decision factor. If it could construct the knowledge base of product development analysis method, it shall make the decision maker of the firm to make the right decision by putting the key point on the significant success factors.

The research of Chaffee [2] pointed out that if the resource of a firm in that business area has special competitive advantage, the resource shall be the success factor. The research of Nuryakin [3] showed that empirical evidences have the benefits to achieve competitive advantage and product innovation via superior marketing performance for the context of Batik SMEs in Indonesia. Song and Parry [4] summarized the researches of previous scholars and experts and divided the new product success factors into four classes: (1) the firm's competitive environment, (2) the firm's internal condition, (3) NPD process, (4) product competitive advantage. Cleyn et al. [5] found the success key factors of NPD include the product strategy, abundant resource, early stage development capability, product development capability, product marketing capability, product development flow, market related information, high-level manager support, innovative team character, and team establishment. Cooper and Kleinschmidt [6] made a study of high-level managers of industrial business unit aiming at America, Germany, Denmark, and Canada, and found the success factors for the Business of the product development include (1) high-quality new product process, (2) new product strategy defined by business

*Corresponding Author: Jieh-Ren Chang; E-mail: jrchang@niu.edu.tw
DOI: 10.3966/160792642020112106006

unit, (3) sufficient manpower resource and financial resource, (4) NPD expense rate (percentage to the firm turnover), (5) high-quality product development team, (6) high-level manager participating in NPD, (7) innovative culture leading-in team, (8) integrated cross-project team, (9) high-level manager willing to take charge of NPD. Cooper and Kleinschmidt [6] thought the above first, second, third and fourth factors could be the key steps for NPD to the success.

Mohannad [7] had a research to demonstrates a model-based methodology and information technology to engage consumers at large scales to drive new product and manufacturing process development to address two challenges of: (i) the ability to identify rapidly the needs and preferences of different market segments; (ii) the ability to respond quickly and flexibly. An orange beverage has been selected to show that by linking a game-like consumer facing internet web application and a novel computer driven flow manufacturing system, target sensory attributes obtained by consumer groups can be rapidly translated into a new formulation recipe.

V. Khrystoforova and D. Siemeniako build up an Internet-based consumer co-creation model for the new product development processes [8]. Specifically, it is an attempt to determine the level of consumer engagement in an online co-creation process, identifying motives and reasons for the participation in new product development as well as understanding the types of Internet-based co-creation that are mostly preferred by consumers.

This research focuses on the product development proposed knowledge base establishment, to make the firm with limited information and resource, to integrate the research proposed key factors and information of other sales, political and economic environment, product features, main consumer group, and firm capacity analysis, in the expectation of allowing the firm's decision maker to make correct decision with adequate information by the internet decision-making system, and ensure the product could be developed and appear on the market smoothly and successfully, and make profit for the firm.

2 Literature Review

In this section, a brief overview of the aspects of new product development, knowledge base and decision tree C4.5 is presented.

2.1 New Product Development

The new product innovative development is one of the key factors for most firms, only through constant new product innovative development to ensure the firm's sustainable operation [9-10]. Ernst et al. [11] divided the NPD process into three stages: (1) the product development concept phase, (2) the official development stage of the product, and (3) the product development mass production. The data collection for product development is important, and the information from the client side, market side, product development side and purchasing side would quite influence the product success development; the analysis after data collection and final decision making are also very important.

The research of Cooper [12] indicated that, NPD could be divided into the idea stage, preliminary evaluation stage, design concept stage, product development stage, product testing stage, engineering trial stage, and mass production stage. Davidson et al. [13] found that, the key factors influencing the product development include: (1) specific product development target, (2) design team spirit, (3) responsible leader's capability, (4) product development process integration, and (5) flexible R&D process. Horvata et al. [14] devided consumer data into three units: Consumer involvement, food trend, and environmental factor data for NPD and product life cycle (PLC), and its empirical results indicated that over 85% respondents used all the above three data units for the NPD, while they rarely used consumer data for the PLC. Mohrman et al. [15] expressed that NPD is the process producing the valuable product, while the course with value of knowledge. The research of Koskinen [16] showed that the firm must reinforce the knowledge management, which is one of the methods to improve the performance of NPD; Koskinen et al. [17] indicated that the knowledge acquired from the customer side is also the key knowledge source of NPD. To sum up, increasing the knowledge management level could not only improve the performance of NPD, but also shorten the time of NPD, to minimize the product development cost.

2.2 Knowledge Base

The knowledge management [18-20] is the course of seeing, knowing and acquiring the information in the activity, and also the course transforming the information to the knowledge, including the knowledge acquisition, knowledge expression, and knowledge searching. It also includes the expert opinion. Knowledge is the main basis for decision making and action, especially for the firm's product development, which is the significant issue; if the firm wants to make the precise and correct decision, adequate information would be one of the keys to influence the decision success. Therefore, the knowledge base is the critical factor for a firm. The elemental work of establishing the knowledge base is the big data of data collection, and through the data classification, to transform the data to useful information, and then store the data, so these collected data shall be meaningful.

The research of Car et al. [21] showed that the firm's knowledge assets have the four types of data, information, knowledge, and intelligence (or wisdom) in the context of big data strucutre. After the data collection and knowledge transformation, it shall become the useful information, while the information is not only the firm's knowledge, which may convert to the intelligence of the firm if applied properly. For a firm, the intelligence would make the firm to rapidly, indeed and smoothly manufacture the product and launch in the market.

2.3 Decision Tree C4.5

The most typical algorithm of decision tree [22-24] includes ID3, C4.5 [25-26] and CART, and the decision tree classified mode has the following four characters: (1) simple structure; (2) suitable to train the big data; (3) not needing the knowledge other than training to collect data; (4) with high classification accuracy. C4.5 has an entensive type of algorithms [27], and C4.5 is improved from ID3 algorithm. The decision tree algorithm would put the original data of the training sample in the root of the decision tree, and then divide the original data into two groups: one is the training group needed data; the other is the test group needed data. And the training data shall be applied to establish the decision tree, and it shall evaluate and select the data property to continue being the branch needed basis according to Information Theory in every internal node. Finally, it shall trim the decision tree by testing data, until every branch on the decision tree has one node, to promote the predictive ability and speed.

This research shall establish an "IF-THEN" rule according to the classification standard of every internal node in the process from the tree path to the root node of the decision tree. The decision tree establishment mainly contains three steps, to apply proper algorithm to dispose the training data and establish the decision tree rules, properly trim the decision tree and take the rule out from the decision tree. This research disposes the data based on these three steps, and finds the product development decision making factor according to the finally acquired Rule.

3 Research Methodology

In this section, we present a research framework with its description for solving new product development problem.

3.1 Research Framework

After confirming the research direction, it shall collect the data through expert interview, and summarize the result after the expert interview, for data establishment and data classification; calculate after data classification, then analyze and integrate the result for conclusion as the reference for future decision making of the firm's decision maker. The research framework is respectively the data collection stage, data preprocessing stage, data exploration stage, knowledge base establishment stage and empirical research stage, as shown in Figure 1.

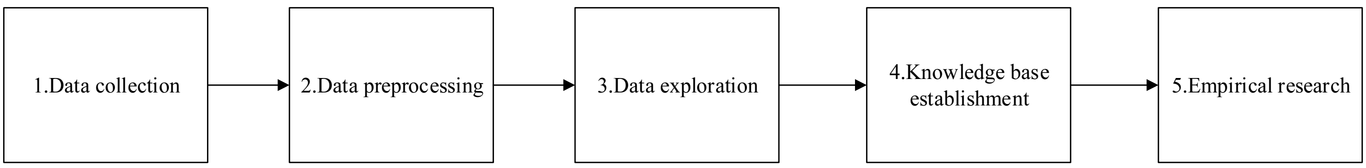

Figure 1. Research framework

3.2 Research Framework Description

Five stages for the research framework are described, as follows:

Data collection stage. This stage shall handle the data collection by expert interview, because the product development data are usually the confidential information of each firm, so the information could be acquired only after the product appears on the market or even after some products' life circle ends, and then this research shall obtain the complete product development data. A lot of data are collected from internet acquisition system during product selling process on the market. The selected experts are those with the qualification and record of service in R&D over ten years and with the leadership experience above manager level, therefore they provided data have high reliability. After the research result comes out, there shall be several interviews, to add the expert related experience and R&D course to analyze and integrate to the research basis.

Data preprocessing stage. After acquiring the expert interview data, it shall classify and process all the data, mainly by the firm, and classify the product according to the expert worked firm type, then classify and establish the document of relevant information as product development course, invested development personnel and product life circle, and finally file all the information to encode for establishment. After finishing the data collection, it shall apply the method of data mining to screen the useful information from these data.

Data exploration stage. The data exploration methods mainly include the following six methods: Association, Classification, Clustering, Time Sequence, Neural Network and Generalization. The purpose of data exploration is to extract the mode and knowledge from mass of data. The data exploration would involve six common tasks: anomaly detection, association rule learning, searching for relation among variables, clustering, classification, regression and Automatic Summarization. Every data exploration method aims at different data property, and up to now, there is no data exploration method able to satisfy all the data property. This research shall find the proper classification method among various data classification methods, and adopt the decision tree as the research basis.

This research shall try to calculate the accuracy by cross validation after finishing the data establishment, and enter the calculation result into the computer as the report for future analysis and research after acquiring the accuracy result via data exploration method, and finally find the key decision making factor of the product development from the result.

Knowledge base establishment stage. If a research only finds the result or reason, it would be a pity for a research. Thus it shall establish the research result to the data base for further study. This research would enter the final result into the data base, for future research.

Empirical research stage. It shall design a system for the final research conclusion, and enter into the new product related information of A firm, and then the system shall judge whether the project is a successful project based on the acquired Rules, to confirm the correctness of result and decision acquired in this research.

4 Research Results

This chapter is divided into six stages for empirical data analysis from real cases according to the research framework: data collection stage, data preprocessing stage, data exploration stage, knowledge base establishment stage, empirical research stage and test result, respectively.

4.1 Data Collection

Expert interview. The expert interview content includes the expert personal data, seniority of product design and development, or leading product project, basic data of designed product, and relevant information. Through the interview, it obtains the information of totally 32 senior project design handlers in design industry, and product related information, including (1) capital sum, (2) product category, (3) product dimension, (4) project invested manpower, (5) R&D time history, (6) sales volume, (7) product life cycle, (8) project design, and (9) production place. The experts had worked in 25 firms, in different scales, from large scale of hundreds of thousands of employees, to small scale of 50 employees; after acquiring the expert interview data, it continues to collect the expert worked firm data, including: the firm's capitalization, registered employee number, operating items, and whether it is the listed company or OTC company. Finally, it shall establish all the data in the computer. The interviewed expert data are as shown in Table 1.

Table 1. Interviewed expert data

Company	Interviewee	Job title	Age	Years of experience
A	Hung	Design Manager	55	> 40
A	Jerry	Senior Engineer	46	> 20
A	Peng	Senior Engineer	48	> 35
A	KY	Senior Engineer	48	> 35
B	Clark	Design Manager	58	> 40
B	David	Design Manager	50	> 25
B	John	Design Manager	57	> 40
C	Vincent	Design Manager	50	> 25
C	Frank	Design Manager	43	> 25
D	Chau	Project Manager	40	> 20
E	Martin	Project Manager	40	> 15
F	David	Design Manager	52	> 30
G	Mark	Design Manager	44	> 20
G	Rock	Design Director	36	> 15
G	Eric	Design Supervisor	32	>12
G	YW	Design Manager	38	>14
H	Jason	Senior Engineer	42	> 15
I	Joe	Owner	36	> 12
J	Kao	Design Manager	48	> 25
J	Jou	Design Manager	47	>25
K	Jason	Owner	40	>15
L	Eric Lee	Owner	42	>15
M, N, O, P, Q	Sam	Design Manager	46	>20
R	Jacky	Design Assistant Manager	41	>16
S	Jeff	Project Manager	38	>12

Table 1. Interviewed expert data (continue)

Company	Interviewee	Job title	Age	Years of experience
T	YWJ	Design Assistant Manager	46	>18
V	Jerry	Design Manager	37	>20
V	Sam Lin	Design Manager	36	>16
W	Johnson	Design Vice President	56	>30
W	Davy	Design Manager	44	>25
X	Jacky	Design Manager	40	>25
Y	Vincent	Design Manager	50	>25

Data classification. It shall firstly remove partial invalid data after arrangement (such as the firm domicile, phone number, person in charge, product feature, product introduction and product appearance introduction, which would not influence the product development).

Property selection. After the screening according to decision tree, the accuracy is 70.5882%, and there are nine results of totally 12 Rules; through the once cross validation and different proportional difference of testing/training, and 10-fold cross validation and different proportional difference of testing/training, it shall make the seriation according to the calculated accuracy, to find the classification method most suitable for this research. The viewing data is screened for unimportant fields of A6, A10, A11, A18 and A19, which respectively represents the product name, product responsible person, product budget, hardware invested manpower, and R&D time history. The left important fields include A0-A5, A7-A9, A12-A17, A20-A23, as shown in Figure 2. Table 2 is the field description of Final Data Sheet, including the filed code, domain name and field attribute.

A0	A1	A2	A3	A4	A5	A7	A8	A9	A12	A13	A14	A15	A16	A17	A20	A21	A22	A23
1	A	2.7E+08	2000	1	1	1	540	4	30000	1000000	1170000	T	T	3	3	3	M	Y
2	A	2.7E+08	2000	1	1	1	600	4	50000	1000000	1200000	T	T	3	3	3	M	Y
3	A	2.7E+08	2000	1	1	1	1500	4	5000	2000000	2220000	T	T	9	2	3	S	N
4	A	2.7E+08	2000	1	1	1	800	5	30000	1200000	2100000	T	T	6	2	3	S	N
5	A	2.7E+08	2000	1	1	1	3000	4	1000	1000000	1470000	T	T	12	1	3	S	N
6	A	2.7E+08	2000	1	1	1	520	4	50000	1000000	1220000	T	T	3	3	3	M	Y
7	A	2.7E+08	2000	1	1	1	580	4	10000	1000000	1220000	T	T	3	3	3	M	N
8	A	2.7E+08	2000	1	1	1	100	3	1000	750000	840000	T	T	6	1	1	S	N
9	B	8.2E+08	2000	2	1	2	350	7	6000	2000000	2010000	T	T	9	1	3	S	N
10	B	8.2E+08	2000	2	1	2	320	8	6000	2000000	2190000	T	T	6	1	3	S	N
11	B	8.2E+08	2000	2	1	2	320	12	700000	6000000	3360000	T	T	18	3	5	L	Y
12	B	8.2E+08	2000	2	1	2	320	6	100000	4000000	2280000	T	T	6	2	5	L	Y
140	Y	1.56E+11	200000	1	1	4	350	8	1300000	2200000	2500000	C	C	3	5	15	L	Y
141	Y	1.56E+11	200000	1	1	4	300	9	2000000	1750000	2000000	C	C	2	5	20	L	N
142	Y	1.56E+11	200000	1	1	4	350	12	750000	1250000	1200000	C	C	3	3	12	M	N
143	Y	1.56E+11	200000	1	1	4	300	8	500000	750000	775000	C	C	2	2	4	S	N
144	Y	1.56E+11	200000	1	1	4	320	6	550000	770000	750000	C	C	2	1	6	S	N
145	Y	1.56E+11	200000	1	1	4	420	4	700000	850000	835000	C	C	3	2	7	S	N
146	Y	1.56E+11	200000	1	1	4	380	7	1100000	1200000	1150000	C	C	3	2	9	M	Y
147	Y	1.56E+11	200000	1	1	4	365	3	1350000	1150000	1250000	C	C	3	3	9	L	Y
148	Y	1.56E+11	200000	1	1	4	320	4	1600000	1650000	1750000	C	C	4	4	10	L	Y
149	Y	1.56E+11	200000	1	1	4	300	6	2200000	1800000	1850000	C	C	3	3	15	L	Y
150	Y	1.56E+11	200000	1	1	4	400	6	650000	650000	700000	C	C	4	3	6	S	Y
151	Y	1.56E+11	200000	1	1	4	420	5	750000	670000	730000	C	C	6	3	5	S	Y
152	Y	1.56E+11	200000	1	1	4	600	6	750000	700000	680000	C	C	3	2	5	S	N
153	Y	1.56E+11	200000	1	1	4	380	6	550000	520000	500000	C	C	4	2	6	S	Y

Figure 2. Final data sheet

Table 2. Individual field description

Filed code	Domain name	Field attribute
A0	Item	Text
A1	Company	Text
A2	Capital	Number
A3	Enterprise scale	Number
A4	Industrial classification	Text
A5	Listed company / Over-The-Counter (OTC) company	Text
A7	Product classification	Text

Table 2. Individual field description (continue)

Filed code	Domain name	Field attribute
A8	Product volume	Number
A9	Human resource	Number
A12	Selling volume	Number
A13	Budget	Number
A14	Research and development expenses	Number
A15	Location of design	Text
A16	Manufacturing location	Text
A17	Development time	Number
A20	Product life cycle	Number
A21	Sales territory region	Number
A22	Project scale	Text
A23	Project success or Project failure	Text

Some important fields are described as follow. A2 denotes company capital from NT$ 5 million to 156 billion. A3 is Enterprise scale which is represented by number of employee. A4 is Industrial classification which includes Computer / Computer Peripheral, Consumer Electronics Manufacturing, Communication Network Industry, Communication equipment and tranditional Manufacturing. A5 denotes Listed company, Over-The-Counter (OTC) company or niether Listed nor OTC. A7 is Product classification with 7 classes.

4.2 Data Preprocessing

The expert interview collected data shall be stored after classification, extraction and encoding, so this stage is the core stage of this research.

Data classification for the second stage. Mainly classify according to the firm and classify the product and the interviewee to the firm he/she worked, and then classify and establish the document of relevant information as product development course, invested development personnel and product life circle, and finally file all the information to encode for establishment.

Data storage. After the data are classified, the data shall be stored in computer as the operational basis.

4.3 Data Exploration

After classification, the data shall receive the program operation of various classification methods, to find the optimal classification method, and the applied classification methods are as below: Bayes Net, Naive, IBK, K-Star, Bagging, One R, Zero R, Decision Stump, Hoeffding Tree, REP Tree and C4.5. 10-fold cross validation is to calculate the average accuracy through 10 sampling results, and the final model is the model established by all the training data; this research adopted cross validation of data classification method is as shown in Table 3. It adopts the classifier of cross validation method for accuracy (%), that the operated accuracy shall find the optimal classifier by testing/training allocation proportion. The top three of accuracy in cross validation result are Logistic (74.24%), C4.5, Multilayer Perceptron (70.59%), and LWL (69.28%); the lowest accuracy is Zero R (50.98%).

Table 3. Classifier accuracy (10-fold cross validation)

Methods	Correctly Classified Instances
Logistic	74.24
C4.5	70.59
Multilayer Perceptron	69.28
LWL	69.28
JRIP	68.63
IBK	63.40
Ada Boost M1	63.16
Decision Stump	61.44
Random Tree	60.13
K-STAR	60.13
Bayes Net	59.48
Naive Bayes	59.48
REP Tree	58.82
Bagging	57.52
Hoeffding Tree	57.52
Zero R	50.98

(1) This research adopts the testing/training allocation proportion as below: 95/5, 90/10, 85/15, 80/20, 75/25, 70/30, 65/35, 60/40, 55/45 and 50/50, and after once operation of every classification method, it shall make the cross validation and accuracy (%) of each training/testing proportion, as shown in Table 4. The top three in the accuracy are 95/5 Random Tree (100%), 85/15 Ada Boost M1 (86.96%), and 90/10 Ada Boost M1 (86.67%); the lowest accuracy is 50/50 Decision Stump (67.11%).

Table 4. Accuracy (%) of each classification method in different training/testing proportions

Dataset	Cross validation	95/5	90/10	85/15	80/20	75/25	70/30	65/35	60/40	55/45	50/50
C4.5	70.59	50.00	73.33	73.91	67.74	65.79	63.04	66.67	65.57	66.67	67.11
Decision Stump	61.44	62.50	66.67	73.91	67.74	65.79	63.04	68.52	65.57	65.22	64.47
Hoeffding Tree	57.52	75.00	60.00	65.22	61.29	65.79	67.39	57.41	67.21	65.22	55.26
Random Tree	60.13	100.00	40.00	65.22	45.16	55.26	71.74	66.67	60.66	55.07	57.89
REP Tree	58.82	75.00	60.00	52.17	54.84	63.16	56.52	53.70	45.90	44.93	43.42
Bayes Net	59.48	62.50	80.00	78.26	54.84	60.53	60.87	62.96	63.93	63.77	51.32
Naive Bayes	58.17	75.00	60.00	60.87	61.29	71.05	65.22	55.56	65.57	63.77	57.89
Logistic	74.24	87.50	80.00	73.91	70.97	73.68	76.09	70.37	70.49	72.46	65.79
Multilayer Perceptron	69.28	100.00	73.33	60.87	80.65	78.95	67.39	64.81	67.21	65.22	57.89
IBK	63.40	87.50	66.67	69.57	67.74	71.05	73.91	68.52	63.93	65.22	57.89
K-Star	60.13	62.50	80.00	73.91	70.97	63.16	60.87	61.11	62.30	63.77	64.47
LWL	69.28	75.00	80.00	82.61	77.42	78.95	71.74	66.67	68.85	65.22	63.16
Ada Boost M1	63.16	62.50	86.67	86.96	67.74	63.16	63.04	57.41	67.21	60.87	59.21
Bagging	57.52	87.50	66.67	65.22	67.74	71.05	65.22	59.26	55.74	49.28	46.05
JRIP	68.63	75.00	66.67	47.83	67.74	63.16	63.04	68.52	67.21	56.52	65.79
Zero R	50.98	37.50	60.00	39.13	41.94	44.74	45.65	46.30	45.90	44.93	43.42

(2) The 10-fold accuracy (%) comparison of each classification method in different training/testing proportions is as shown in Table 5. The number in the brackets means the standard deviation, which (also known as mean square deviation) is the index reflecting the dispersion degree of a group of measurement data, and refers the accuracy error's fluctuant range within certain period. Therefore, the higher the accuracy is, the better it is, while the lower the standard deviation is, which takes the accuracy as the most important analysis basis. This research uses the data to acquire the accuracy result through different classification methods of once and ten times of operations and the proportional training/testing method, and finally repeats for ten times of computation to obtain the most precise result. The top three of 10-fold accuracy are 90/10 Logistic 74.39% (9.74), 95/5 Naive Bayes 74.29% (14.51), and 85/15 LWL 72.10% (6.53); the lowest accuracy is 55/45 Multilayer Perceptron 67.39% (3.89).

Table 5. 10-fold accuracy comparison % of each classification method in different training/testing proportions (standard deviation %)

Dataset	Cross validation	95/5	90/10	85/15	80/20	75/25	70/30	65/35	60/40	55/45	50/50
C4.5	68.16 (11.15)	65.83 (14.26)	67.85 (9.73)	68.58 (8.77)	62.47 (7.43)	65.06 (8.38)	61.29 (8.94)	60.72 (5.65)	61.41 (6.60)	62.45 (6.93)	62.47 (7.43)
Hoeff Tree	58.82 (11.41)*	56.85 (10.57)	54.88 (6.97)	56.94 (5.55)	57.14 (4.62)	60.66 (6.63)	59.18 (5.78)	59.45 (4.29)	60.28 (4.41)	59.25 (4.50)	57.14 (4.62)
Random Tree	65.97 (12.36)	66.79 (21.18)	67.39 (10.75)	65.59 (6.62)	62.85 (6.99)	68.71 (7.62)	66.13 (8.06)	65.40 (3.72)	64.54 (7.74)	64.75 (5.59)	62.85 (6.99)
REPT	59.75 (10.13)*	65.95 (17.36)	59.51 (13.02)	56.03 (8.60)	55.43 (7.17)	55.69 (6.29)	55.89 (5.43)	56.65 (6.73)	54.75 (5.18)	54.59 (4.70)	55.43 (7.17)
Bayes	61.52 (11.58)	74.29 (14.51)	64.86 (12.54)	61.16 (7.38)	57.13 (5.31)	60.93 (6.62)	61.54 (5.31)	58.11 (3.94)	60.29 (3.63)	60.12 (4.09)	57.13 (5.31)
Logistic	68.59 (11.56)	71.13 (14.01)	74.39 (9.74)	71.67 (9.00)	64.17 (5.92)	69.52 (5.60)	67.59 (7.64)	66.29 (8.67)	66.97 (6.47)	66.53 (4.78)	64.17 (5.92)
Multi layer Perceptron	68.55 (12.38)	68.93 (15.43)	67.08 (9.82)	68.13 (6.63)	67.59 (5.58)	65.32 (6.19)	66.15 (5.01)	68.58 (6.90)	67.81 (5.04)	67.39 (3.89)	67.59 (5.58)
IBK	62.13 (10.90)	67.20 (15.77)	64.64 (10.10)	65.63 (3.83)	59.73 (4.03)	62.97 (5.85)	60.21 (5.20)	60.00 (5.20)	60.62 (5.04)	61.28 (4.18)	59.73 (4.03)
K-Star	59.21 (7.77)	63.45 (10.27)	60.80 (7.12)	56.51 (5.70)	55.30 (2.43)	56.14 (5.06)	55.46 (4.99)	55.88 (4.25)	56.04 (3.76)	55.02 (3.24)	55.30 (2.43)
LWL	70.21 (10.54)	73.87 (18.70)	73.78 (12.57)	72.10 (6.53)	65.50 (7.58)	69.55 (8.29)	68.74 (5.74)	68.59 (4.62)	67.65 (4.87)	64.20 (6.48)	65.50 (7.58)
Ada Book M1	70.75 (10.81)	70.77 (12.09)	71.21 (12.49)	69.02 (7.41)	64.44 (2.65)	67.47 (8.16)	66.19 (6.32)	66.73 (4.13)	67.31 (2.64)	65.65 (5.54)	64.44 (2.65)
JRIP	68.02 (11.48)	73.04 (12.86)	69.43 (13.26)	66.09 (8.07)	64.07 (6.73)	68.27 (6.62)	66.40 (7.06)	65.22 (6.14)	65.03 (9.56)	64.34 (5.90)	64.07 (6.73)

Note. * means remarkable.

(3) Via several classification methods of program operation, it is known that the decision tree is the most suitable method for the data of this research. One of the reasons is that C4.5 (decision tree) is easier and faster to make the reader understanding the content and result. In the final node of the decision tree, every node represents one Rule, and the final result of this research's decision tree indicates 12 nodes and means there are 12 Rules. The output result is as shown in Figure 3.

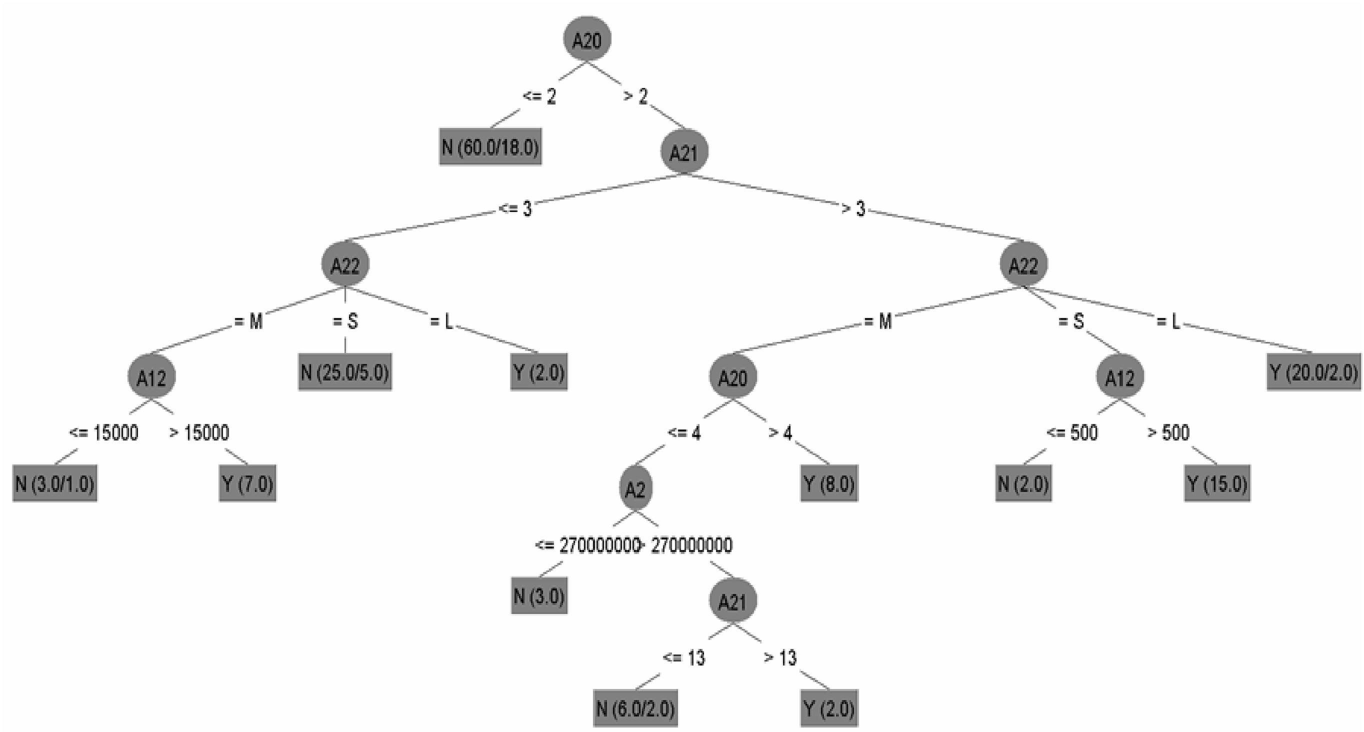

Figure 3. NPD Decision rule by decision trees-C4.5

(4) Rule description: Take 1 Rule from the 12 Rules of this research as example.

(a) Rule 8: IF product life cycle > 2 and product sales country > 3 and project scale = M and product life cycle ≦ 4 and company capitalization > 270,000,000 and sales country number > 13 THEN successful project.

(b) Rule explanation: if the product life cycle is longer than 2 years, the product sales countries are more than 3, the project scale is medium project, the product life cycle is shorter than or equals to 4 years, the company capitalization is greater than NTD 270,000,000, and the sales country number is more than 13, it is a successful project.

(c) Expert illustration: Expert Eric said, when the firm scale is big to certain extent, it shall be more rigorous to the project selection; except for relevant assessment, the product shipment is also a major concern, since there are a great many employees in a middle-large scaled firm, and the daily operation expense would be great. However, adequate manpower, material resource and financial resource shall make it available to undertake some middle-large scaled project, and the product shall be shipped to more countries, which shall not be difficult for the middle-large scaled firm, and the project would be easier to succeed.

4.4 Knowledge Base Establishment

The decision tree results in 12 Rules, which means it also finds the key decision-making factor of product development, including the product life cycle, product sales country number, project scale, project sales volume and company capitalization, in details as below:

(1) Product life cycle: the product life cycle is also called PLC, and every product shall go through the development stage, launching-on-the-market stage, growth stage, maturation stage and product degenerating stage.

Expert opinion. Expert Vincent said, every company hopes their designed product could have long life cycle, but very few products could be, especially products in the information industry; therefore, whether the product development is successful, it is reasonable that the product life cycle is one of the judgement factors.

(2) Product sales country number: country number of product export sales.

Expert opinion. Expert Mark said, the more countries the product could be accepted and used by, the more popular the product is.

(3) Project scale: generally speaking, the small-scale project could not make profit for the firm, while the large-scale project may cause incompletion; and the client may not be interested in the small-scale project, while for the large-scale project, the client may not be

willing to take or have no adequate budget.

Expert opinion. Expert Chau said, the resource of every company for the project depends on the project scale. But not every large-scale project would be successful, however, the result from the data just shows the large-scale project has higher success rate; we could explain that, the large-scale project would gain higher attention of the company for the project condition and cost control, so the success rate would be higher.

(4) Project sales volume: sales volume of a project during the product life cycle.

Expert opinion. Expert YWJ said, the more the product could be sold, the more successful the project is.

(5) Company capitalization: the company capitalization refers to the total capital investment by the shareholders when establishing the company; one company needs funds for operation. When the company starts to expand the business, it shall need capital increment, while the capitalization of the company would expand too.

Expert opinion. Expert Sam said, the greater the capitalization of the company is, the more the resource of the company is, which is the advantage for the project control, cost control and applicable resource comparing to the small and medium size company.

4.5 Empirical Research

The work contents of empirical research stage include the system establishment, new product data input and NPD decision's final result report output. After 12 Rules of the research result is designed to a system, it shall input the project data of A company, and the system shall calculate the decision suggestion, for A company's decision maker as reference, and system correctness validation; the system diagram is as shown in Figure 4.

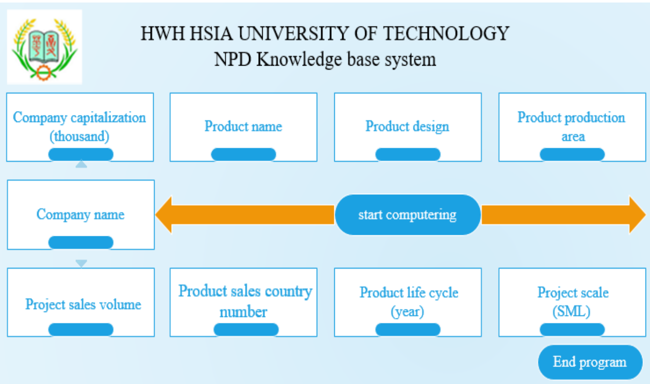

Figure 4. Knowledge base system for NPD

4.6 Test Result

This research mixes the classification model, applies several data exploration classification techniques to solve and reduce the firm development risk, and provides the decision model of NPD. Through different classification methods of training/testing of different proportions of 10-fold accuracy and standard deviation (%), such as 70.75% (10.81%), the analysis result shows the accuracy of 11 classification methods, which is explained as below after clearing up and ordering:

Cross validation. Ada Boost M1 70.75% (10.81%) > LWL 70.21% (10.54%) > Logistic 68.59% (11.56%) > Multilayer Perceptron 68.55% (12.38%) > C4.5 68.16% (11.15%) > JRIP 68.02% (11.48%) > Random Tree 65.97% (12.36%) > IBK 62.13% (10.90%) > Naive Bayes 61.52% (11.58%) > REP Tree 59.75% (10.13%) > K-Star 59.21% (7.77%) > Hoeffding Tree 58.82% (11.41%).

95/5 (training/testing). Naive Bayes 74.29% (14.51%) > LWL 73.87% (18.70%) > JRIP 73.04% (12.86%) > Logistic 71.13% (14.01%) > Ada Boost M1 70.77% (12.09%) > Multilayer Perceptron 68.93% (15.43%) > IBK 67.20% (15.77%) > Random Tree 66.79% (21.18%) > REP Tree 65.95% (17.36%) > C4.5 65.83% (14.26%) > K-Star 63.45% (10.27%) > Hoeffding Tree 56.85% (10.57%).

90/10 (training/testing). Logistic 74.39% (9.74%) > LWL 73.78% (12.57%) > Ada Boost M1 71.21% (12.49%) > JRIP 69.43% (13.26%) > C4.5 67.85% (9.73%) > Random Tree 67.39% (10.75%) > Multilayer Perceptron 67.08% (9.82%) > Naive Bayes 64.86% (12.54%) > IBK 64.64% (10.10%) > K-Star 60.80% (7.12%) > REP Tree 59.51% (13.02%) > Hoeffding Tree 54.88% (6.97%).

85/15 (training/testing). LWL 72.10% (6.53%) > Logistic 71.67% (9.00%) > Ada Boost M1 69.02% (7.41%) > C4.5 68.58% (8.77%) > Multilayer Perceptron 68.13% (6.63%) > JRIP 66.09% (8.07%) > IBK 65.63% (3.83%) > Random Tree 65.59% (6.62%) > Naive Bayes 61.16% (7.38%) > Hoeffding Tree 56.94% (5.55%) > K-Star 56.51% (5.70%) > REP Tree 56.03% (8.60%).

80/20 (training/testing). Multilayer Perceptron 67.59% (5.58%) > LWL 65.50% (7.58%) > Ada Boost M1 64.44% (2.65%) > Logistic 64.17% (5.92%) > JRIP 64.07% (6.73%) > Random Tree 62.85% (6.99%) > C4.5 62.47% (7.43%) > IBK 59.73% (4.03%) > Hoeffding Tree 57.14% (4.62%) > Naive Bayes 57.13% (5.31%) > REP Tree 55.43% (7.17%) > K-Star 55.30% (2.43%).

75/25 (training/testing). LWL 69.55% (8.29%) > Logistic 69.52% (5.60%) > Random Tree 68.71% (7.62%) > JRIP 68.27% (6.62%) > Multilayer Perceptron 65.32% (6.19%) > C4.5 65.06% (8.38%) > IBK 62.97% (5.85%) > Hoeffding Tree 60.66% (6.63%) > K-Star 56.14% (5.06%) > REP Tree 55.69% (6.29%) > Ada Boost M1 67.47% (8.16%) > Naive Bayes 60.93% (6.62%).

70/30 (training/testing). LWL 68.74% (5.74%) > Logistic 67.59% (7.64%) > JRIP 66.40% (7.06%) > Ada Boost M1 66.19% (6.32%) > Multilayer Perceptron 66.15% (5.01%) > Random Tree 66.13% (8.06%) > Naive Bayes 61.54% (5.31%) > C4.5

61.29% (8.94%) > IBK 60.21% (5.20%) > Hoeffding Tree 59.18% (5.78%) > REP Tree 55.89% (5.43%) > K-Star 55.46% (4.99%).

65/35 (training/testing). Multilayer Perceptron 68.58% (6.90%) > Ada Boost M1 66.73% (4.13%) > Logistic 66.29% (8.67%) > Random Tree 65.40% (3.72%) > JRIP 65.22% (6.14%) > C4.5 60.72% (5.65%) > IBK 60.00% (5.20%) > Hoeffding Tree 59.45% (4.29%) > Naive Bayes 58.11% (3.94%) > REP Tree 56.65% (6.73%) > K-Star 55.88% (4.25%) > LWL 68.59% (4.62%).

60/40 (training/testing). Multilayer Perceptron 67.81% (5.04%) > LWL 67.65% (4.87%) > Ada Boost M1 67.31% (2.64%) > Logistic 66.97% (6.47%) > JRIP 65.03% (9.56%) > Random Tree 64.54% (7.74%) > C4.5 61.41% (6.60%) > IBK 60.62% (5.04%) > Naive Bayes 60.29% (3.63%) > Hoeffding Tree 60.28% (4.41%) > K-Star 56.04% (3.76%) > REP Tree 54.75% (5.18%).

55/45 (training/testing). Multilayer Perceptron 67.39% (3.89%) > Logistic 66.53% (4.78%) > Ada Boost M1 65.65% (5.54%) > Random Tree 64.75% (5.59%) > JRIP 64.34% (5.90%) > LWL 64.20% (6.48%) > C4.5 62.45% (6.93%) > IBK 61.28% (4.18%) > Naive Bayes 60.12% (4.09%) > Hoeffding Tree 59.25% (4.50%) > K-Star 55.02% (3.24%) > REP Tree 54.59% (4.70%).

50/50 (training/testing). Multilayer Perceptron 67.59% (5.58%) > LWL 65.50% (7.58%) > Ada Boost M1 64.44% (2.65%) > Logistic 64.17% (5.92%) > JRIP 64.07% (6.73%) > Random Tree 62.85% (6.99%) > C4.5 62.47% (7.43%) > IBK 59.73% (4.03%) > Hoeffding Tree 57.14% (4.62%) > Naive Bayes 57.13% (5.31%) > REP Tree 55.43(7.17) > K-Star 55.30% (2.43%).

To sum up, the empirical result of this research is as below: 90/10 Logistic 74.39% (9.74) > 95/5 Naive Bayes 74.29% (14.51) > 85/15 LWL 72.10% (6.53) > Cross validation Ada Boost M1 70.75% (10.81) > 75/25 LWL 69.55% (8.29) > 70/30 LWL 68.74% (5.74) > 65/35 Multilayer Perceptron 68.58% (6.90) > 60/40 Multilayer Perceptron 67.81% (5.04) > 80/20 Multilayer Perceptron 67.59% (5.58) > 50/50 Multilayer Perceptron 67.59% (5.58) > 55/45 Multilayer Perceptron 67.39% (3.89).

Comparing the calculation result of once and ten times, it finds Multilayer Perceptron appears five times at most in 11 operations, and the second is LWL of three times, Naive Bayes, Ada Boost and Logistic of once; the result list acquired in this research is shown in Table 6 Optimal accuracy of two operation results:

Table 6. Optimal accuracy of two operation results

Classification method	Once operation	Ten times of operations
Cross validation	Logistic(74.24%)	Ada Boost M1 70.75% (10.81%)
95/5	Random Tree(100%)	Naive Bayes 74.29% (14.51%)
90/10	Ada Boost M1(86.67%)	Logistic 74.39% (9.74%)
85/15	Ada Boost M1(86.96%)	LWL 72.10% (6.53%)
80/20	Multilayer Perceptron(80.65%)	Multilayer Perceptron 67.59% (5.58%)
75/25	Multilayer Perceptron (78.95%)	LWL 69.55% (8.29%)
70/30	Logistic(76.09%)	LWL 68.74% (5.74%)
65/35	Logistic(70.37%)	Multilayer Perceptron 68.58% (6.90%)
60/40	Logistic(70.49%)	Multilayer Perceptron 67.81% (5.04%)
55/45	Logistic(72.46%)	Multilayer Perceptron 67.39% (3.89%)
50/50	Decision Stump(67.11%)	Multilayer Perceptron 67.59% (5.58%)

From the research result, it is known that different classification method could dispose different types of data; although the overall performance of C4.5 classification method presents above the average, C4.5 presents one method easy and quick to be understood by the users; through the performance of decision tree, it could understand and find the impact factors of product development. This research applies the data over different classification method of once and ten times of operations and proportional training/testing method to obtain the accuracy result as reference for future researchers.

5 Conclusions and Future Direction

This chapter introduces the research finding, management connotation, research limitation, and future research.

5.1 Research Finding

According to the research result finding, the success factors of NPD include:

(1) Key factors of product development decision: the data collected through expert interview shall be processed by data exploration technology, and find the impact factors of product development including the product life cycle, product sales country number, project scale, project sales volume and company capitalization. It shall be the reference for the decision maker to make the decision for NPD with limited information.

(2) Current and previous researches: the previous research found that the system factors influencing the NPD include the organization factor, NPD process,

product business condition forecasting procedure, and the external competitor. But there is not much study on the product research and development part. Thus, it makes this research to find the difference and advantage from the product development angle as the basis, in the expectation of finding the more accurate key factor.

(3) Significance of product development decision to the company: in Taiwan, the small and medium sized firm occupies the majority, and once the firm operator or decision maker makes the wrong decision, it shall make the company investment loss. Thus, how to make the correct decision with limited information shall be the most important topic for discussion.

5.2 Management Connotation

The investors of small and medium sized firm in Taiwan are usually the company operator too, which is different to the foreign firm. The foreign successful middle-large size company has the latent rule, that the investor would not operate the company, while the operator would not invest for the company; the reason is to reduce the mutual interference between the investor and operator in the company operation process. In Taiwan, even some decisions to the company product of non-investor professional manager are made through the decision model from top to bottom; thus cause many decision making chaos. On the contrary, the foreign operator usually adopts the pool of wisdom and efforts of everyone to operate the firm, through multipartite professional combination and multiple decision-making methods, to create the stable decision-making mode with high accuracy. Of course, there are advantages and disadvantages in both ways, the that former has the fast decision-making speed and high efficiency but lower accuracy; however, just as described in this research, since the decision maker is both the investor and company operator, even the decision is wrong, the operator shall assume by own. The latter is more cautious, so the decision-making speed is slow, and the reason is that, if the manager makes wrong decision and brings loss to the company, he/she shall be replaced. Thus, the firm is cautious for the decision to product development, no matter it is decided by the boss or by the group. This research finds the decision making factor of product development through the decision tree, which hopes to extract the correct and efficient decision making information from vast information, and provide to the decision maker via the internet decision-making system to help acquiring the preliminary decision making suggestion from the information.

5.3 Research Limitation

This research collected data are mainly proposed by the R&D supervisor in the view of R&D unit, while the product market is ever-changing, thus it is impossible to include the entire product development decision orientation only by R&D supervisor proposed data. The research limitation is listed as below.

(1) Time factor limitation: the research data source is the product development terminal, which are usually the extremely classified data and unable to be collected the product development data of recent years.

(2) Data source is mainly R&D unit, so it couldn’t completely collect the information of other business units and company’s major operator for the time factor.

5.4 Future Research

The research finds that current companies in Taiwan have no certain rules in product development decision making, but the research could not be completed for the reason of data confidentiality and time limitation, so it hopes the future research could follow up and supplement decision making modes of other classification methods aiming at these parts. It is expected to provide the effective suggestion with high accuracy to the future Taiwan firm decision maker in the aspect of NPD and even more significant commercial decision making.

References

[1] A. Griffin, J. R. Hauser, Integrating R & D and Marketing: A Review and Analysis of the Literature, *Journal of Product Innovation Management*, Vol. 13, No. 3, pp. 191-215, May, 1996.

[2] E. E. Chaffee, Three Models of Strategy, *Academy of Management Review*, Vol. 10, No. 1, pp. 89-98, January, 1985.

[3] Nuryakin, Competitive Advantage and Product Innovation: Key Success of Batik SMEs Marketing Performance in Indonesia, *Academy of Strategic Management Journal*, Vol. 17, No. 2, pp. 1-17, April, 2018.

[4] X. M. Song, M. E. Parry, The Determinants of Japanese New Product Successes, *Journal of Marketing Research*, Vol. 34, No. 1, pp. 64-76, February, 1997.

[5] S. H. D. Cleyn, A. Jacoby, J. Braet, Success Factors in New Product Development: How Do They Apply to Company Characteristics of Academic Spin-offs, *The Journal of Private Equity*, Vol. 13, No. 1, pp. 51-61, Winter, 2009.

[6] R. G. Cooper, E. J. Kleinschmidt, Winning Businesses in Product Development: The Critical Success Factors, *Research-Technology Management*, Vol. 50, No. 3, pp. 52-66, May-June, 2007.

[7] M. Jreissat, S. Isaev, M. Moreno, C. Makatsoris, Consumer Driven New Product Development in Future Re-Distributed Models of Sustainable Production and Consumption, *Procedia CIRP*, Vol. 63, pp. 698-703, 2017.

[8] V. Khrystoforova, D. Siemeniako, Internet-based Consumer Co-creation Experience of the New Product Development Process, *Engineering Management in Production and Services*, Vol. 11, No. 3, pp. 60-68, September, 2019.

[9] K.-C. Wang, Product Design Prediction Using Integrated

Dynamic Kansei Engineering Scheme, *Journal of Internet Technology*, Vol. 15, No. 7, pp. 1217-1225, December, 2014.

[10] X. Lu, J. Yin, G. He, H. Yu, N. N. Xiong, An Architecture-Centric Development Approach for Service-Oriented Product Lines, *Journal of Internet Technology*, Vol. 20, No. 4, pp. 999-1012, July, 2019.

[11] R. Ernst, P. Kueppers, J. Stindt, K. Kuchler, L. Schmitt, Multidrug Efflux Pumps: Substrate Selection in ATP-binding Cassette Multidrug Efflux Pumps - First Come, First Served?, *FEBS Journal*, Vol. 277, No. 3, pp. 540-549, February, 2010.

[12] R. G. Cooper, A Process Model for Industrial New Product Development, *IEEE Transactions on Engineering Management*, Vol. EM-30, No. 1, pp. 2-11, February, 1983.

[13] J. M. Davidson, A. Clamen, R. A. Karol, Learning from the Best New Product Developers, *Research-Technology Management*, Vol. 42, No. 4, pp. 12-18, July-August, 1999.

[14] A. Horvat, G. Granato, V. Fogliano, P. A. Luning, Understanding Consumer Data Use in New Product Development and the Product Life Cycle in European Food Firms-An Empirical Study, *Food Quality and Preference*, Vol. 76, pp. 20-32, September, 2019.

[15] S. A. Mohrman, D. Finegold, A. M. Mohrman Jr., An Empirical Model of the Organization Knowledge System in New Product Development Firms, *Journal of Engineering and Technology Management*, Vol. 20, No. 1-2, pp. 7-38, June, 2003.

[16] K. U. Koskinen, Tacit Knowledge as a Promoter of Project Success, *European Journal of Purchasing & Supply Management*, Vol. 6, No. 1, pp. 41-47, March, 2000.

[17] K. U. Koskinen, P. Pihlanto, H. Vanharanta, Tacit Knowledge Acquisition and Sharing in a Project Work Context, *International Journal of Project Management*, Vol. 21, No. 4, pp. 281-290, May, 2003.

[18] Á. Fidalgo-Blanco, M. L. Sein-Echaluce, F. J. García-Peñalvo, Ontological Flip Teaching: A Flip Teaching Model Based on Knowledge Management, *Universal Access in the Information Society*, Vol. 17, No. 3, pp. 475-489, August, 2018.

[19] G. P. Khiste, D. B. Maske, R. K. Deshmukh, Knowledge Management Output in Scopus During 2007 to 2016, *Asian Journal of Research in Social Sciences and Humanities*, Vol. 8, No. 1, pp. 10-19, January, 2018.

[20] M. Ghaedi, M. Madhoushi, Social Capital, Knowledge Management and Innovation Performance, *International Journal of Entrepreneurship and Small Business*, Vol. 35, No. 4, pp. 579-597, January, 2018.

[21] J. Car, A. Sheikh, P. Wicks, M. S. Williams, Beyond the Hype of Big Data and Artificial Intelligence: Building Foundations for Knowledge and Wisdom, *BMC Medicine*, Vol. 17, Article No. 143, pp. 1-5, July, 2019.

[22] G. Zheng, Y. Wang, Y. Chen, Study of Stress Rules Based on HRV Features, *Journal of Computers*, Vol. 29, No. 5, pp. 41-51, October, 2018.

[23] N. A. Alsabour, On the Role of Dimensionality Reduction, *Journal of Computers*, Vol. 13, No. 5, pp. 571-579, May, 2018.

[24] X. Dong, M. Qian, R. Jiang, Packet Classification Based on the Decision Tree with Information Entropy, *The Journal of Supercomputing*, Vol. 76, No. 6, pp. 4117-4131, June, 2020.

[25] D. Moon, S. B. Pan, I. Kim, Host-based Intrusion Detection System for Secure Human-centric Computing, *The Journal of Supercomputing*, Vol. 72, No. 7, pp. 2520-2536, July, 2016.

[26] M. H. M. Soleimani, M. Mansoorizadeh, M. Nassiri, Real-time Identification of Three Tor Pluggable Transports Using Machine Learning Techniques, *The Journal of Supercomputing*, Vol. 74, No. 10, pp. 4910-4927, October, 2018.

[27] I. I. Sinam, A. Lawan, An Improved C4.5 Model Classification Algorithm Based on Taylor's Series, *Jordanian Journal of Computers and Information Technology*, Vol. 05, No. 01, pp. 34-42, April, 2019.

Biographies

You-Shyang Chen is a professor in Hwa Hsia University of Technology and interested in fields of information management, including internet technique, financial analysis, medical management, operations research, organizational management, soft computing, fuzzy time series, rough set theory, and behavior analysis. He has published over 140 journal and conference papers.

Jieh-Ren Chang is currently an Associate Professor of Electronic Engineering and Dean of Library and Information center of National Ilan University, Taiwan. He is interested in the research fields of Image Processing, knowledge discovery from databases and Internet web, machine deep learning, fuzzy time series, rough set theory, and data mining.

Chyi-Jia Huang is a product manager who received a degree of Department of Information Management in a two-year completion program from the Hwa Hsia University of Technology in 2017. He is a product manager of R&D department in the field of the electronic industry now. His main fields of research are Decision Trees, Data Analysis, and Data Mining.