Software Product Satisfaction Classification

Somchai Prakancharoen[1], Thavatchai Ngamsantivong[2], Nattapong Luangnaruedom[*3], Suchada Ratanakongnate[4]

[1]Faculty of Science, Rajabhat Chandrakasem University
Bangkok, Thailand
e-mail: somchai.p@chandra.ac.th

[2,3] College of Business Administration, Innovation and Accounting
Dhurakij Pundit University
Bangkok, Thailand
e-mail: [2] Thavatchai.nga@dpu.ac.th, *[3] Natthaphong.Lua@dpu.ac.th

[4]Faculty of Applied Science, KingMongkut's University of Technology North Bangkok
Bangkok, Thailand
e-mail: [4] suchada.rata@gmail.com

Abstract— There are many software product quality attributes which are concerned by both software development companies and customers. This research tried to seek for the most significant attributes then use them on software satisfaction level prediction by classification tree. The samples were collected from thirty complete developed software projects which all of them are coded for customer's satisfaction on that software product. Multivariate linear regression and Discriminant analysis are used to filter the significant attributes. Decision tree with these significant attributes are used to construct the classification rule of software product satisfaction classification. The accuracy of classification was about 60%. The effective attributes on classification are functional suitability, reliability, and portability.

Keywords-software quality; classification

I. INTRODUCTION

Based on requirements about outsource software development, there are composed of functional and nonfunctional requirements. The functional requirement are whole functional activity that customer intend that the application need to accomplish working. While, the nonfunctional requirement are the explanation about how well working of the information system or some function. There are many nonfunctional requirement such as performance, security, reliability, maintainability, etc. There are so costly software development if it must cover many requirements. Therefore, this research objective is to seek for the important attributes that are expected by customer. These attributes will be used to construct for the software product satisfaction level classification.

II. LITERATURE REVIEW

A. *Software product quality [1]*

ISO/IEC 25010 has defined one functional requirement and other seven nonfunctional attributes which all of these use considered as the items that stakeholder must concern. The defined attributes are functional suitability, reliability, operability, performance efficiency, security, compatibility, maintainability and portability. Each attribute has its sub topics. For example, functional suitability can extend to it more detail as functional completeness, functional correctness, and functional appropriateness. Fortunately, many attribute are not clearly explanation. For example, "Reliability: Maturity - Degree to which a system, product or component meets needs for reliability under normal operation. It very difficult to define all the attribute in practical measuring since there are many distingue in vary context. Since there are the problem of interpretation by difference clients even in the same attribute. Hence, this research has defined the concrete explanation of each sub attributes so that the all the answerers will keep understanding in the same definition.

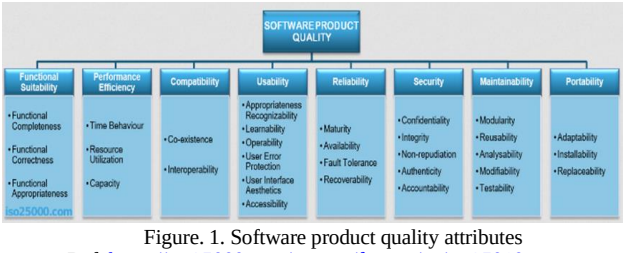

Figure. 1. Software product quality attributes
Ref: https://iso25000.com/images/figures/en/iso25010.png

B. Software quality attributes [2]

The quality attributes are static and dynamic characteristic. Some attributes are used to conduct the software engineering, for example functional suitability. This attribute is the static characteristic. While, performance efficiency attribute is dynamic characteristic since the efficiency of application software could be measured after information system is used for amount of time. After application program processing, the client could make a decision about these dynamic quality attributes.

C. Quality attributes [3]

The software quality attributes are used to explain how to design, develop, test, and validation the developing software project. The output of software development based on the quality attributes is the appropriate software product. The good software product should effectiveness working. Besides, the quality attributes are also used in quality audit, quality control and quality assurance. Therefore, stakeholder should have chosen the significant attribute with their context and limitation.

III. RESEARCH METHODOLOGY

The research methods are explained as in the following activities:

A. Define the interest quality attributes or independent variables

This research has chosen all quality attributes from [1], [2] and [3] as the independent variables.

The explanation about each attribute are totally (all of its sub topics) defined in only one subject. The example for functional suitability preference scale is assigned as shown in table I.

TABLE I. SOME QUALITY ATTRIBUTE AND SATISFACTION LEVEL DEFINITION

Attribute	Detail	Satisfaction scale
Functional attribute	71-100 of functional complete, correctness and appropriate	3
	51-70% of functional complete, correctness and appropriate	2
	0-50% of functional complete, correctness and appropriate	1
Relaibility	>14 days: Mean Time Between Failure	3
	>7=>14 days: Mean Time Between Failure	2
	>1=>7 days: Mean Time Between Failure	1

B. Dependent variable

The software product satisfaction is the variable that could be predicted by the research chosen attributes. Customers have to consider and judge what are the mark scale of the quality attributes. Moreover, all sample customers will consider on the overall satisfaction on the software product which are deployed and using at least six months.

TABLE II. SOFTWARE PRODUCT SATISFACTION LEVEL

Attribute	Detail	Satisfaction scale
Satisfaction	81-100 % overall satisfaction	3
	51-80 % overall satisfaction	2
	0-50 % overall satisfaction	1

C. Data collection

Thirty delivered complete software products from five software development companies are mark giving by the owner of the software products. There are thirty software products used as the research samples. The application of whole samples are covered about accounting field based on web based application. The sample data were collected during 2019-2021 from Bangkok and perimeter provinces of Thailand.

D. Quality attributes filtering

Multivariate linear regression and Discriminant analysis [4, 5] are used in order to seek for significant quality attributes.

E. Classification tree generation

Filtered quality attributes are used to construct decision tree [6]. The class variable is satisfaction level.

F. Classification accuracy measurement

The model will be measure its accuracy by cross validation, root mean square, recall, and f-measure.

IV. RESEARCH RESULT

The result of research are presented as following.

A. Define the interest quality attributes

The software quality attributes (variable name), label, data type, and data range are shown in table III.

TABLE III. VARIABLE DETAIL

Variable	Label	Data type	Data range	Type-variable
fs	functional suitability	ratio	1,2,3	Independent
pe	performance efficiency	ratio	1,2,3	Independent
cp	compatibility	ratio	1,2,3	Independent
us	usability	ratio	1,2,3	Independent
rl	reliability	ratio	1,2,3	Independent
sc	security	ratio	1,2,3	Independent
mt	maintainability	ratio	1,2,3	Independent
pt	portability	ratio	1,2,3	Independent
sf	software satisfaction	category	1,2,3	Dependent

B. *Data collection*

The thirty developed-delivered-implemented software application are data preparation and partial observation are shown in table IV.

TABLE IV. PARTIAL OBSERVATION

	fs	pe	cp	us	rl	sc	mt	pt	sf
1	1.00	2.00	1.00	1.00	1.00	1.00	2.00	1.00	1.00
2	1.00	2.00	1.00	3.00	1.00	2.00	2.00	2.00	1.00
3	1.00	2.00	1.00	2.00	1.00	2.00	2.00	1.00	1.00
4	1.00	2.00	1.00	3.00	1.00	2.00	3.00	3.00	1.00
5	1.00	2.00	1.00	2.00	2.00	2.00	2.00	2.00	1.00
6	1.00	2.00	1.00	3.00	1.00	1.00	2.00	1.00	1.00
7	1.00	2.00	2.00	2.00	1.00	1.00	2.00	2.00	1.00
8	1.00	2.00	2.00	1.00	1.00	1.00	2.00	1.00	1.00
9	1.00	3.00	1.00	1.00	1.00	1.00	3.00	3.00	1.00
10	1.00	3.00	1.00	2.00	2.00	1.00	3.00	2.00	1.00

C. *Attributes filtering*

The attribute filtering in this research use two techniques (multivariate linear regression and discriminant analysis) to seek for the effective quality attributes. The results of two technique are considered which attributes will be selected as the significant quality attributes.

- Multivariate linear regression

The standardized regression equation as shown in (1).

$$sf = 0.493\,fs + 0.114\,pe - 0.008cp + 0.067us + \\ +0.397rl + 0.096sc - 0.043mt + 0.191\,pt \quad (1)$$

$$R^2 adjust = 0.748, sig\,0.000$$

TABLE V. REGRESSION STANDARDIZED COEFFICIENT

Independent Variable	Standardized coefficient	p-value
fs	0.49	0.004
pe	0.11	0.39
cp	-0.01	0.95
us	0.07	0.52
rl	0.40	0.016
sc	0.10	0.35
mt	-0.04	0.66
pt	0.19	0.1
R^2	0.748	

There are only two significant attributes, “fs” and “rl”, with 0.49 and 0.40 standardized coefficient respectively.

- Discriminant analysis

The “sf” is set as a grouping variable while “fs”, “pe”, “cp”, “us”, “rl”, “sc”, “mt”, and “pt” are set to be the independent variables. Discriminant analysis is performed with “Enter” method. The standardized coefficient of two discriminant functions are shown in table VI.

TABLE VI. STANDARDIZED COEFFICIENT OF TWO DISCRIMIONANT ANALYSIS FUNCTION

Independent Variable	sf(1)	sf(2)	Decision
fs	0.68	-0.74	chosen
pe	0.28	0.53	chosen
cp	0.01	0.42	chosen
us	0.2	0.35	chosen
rl	0.65	0.44	chosen
sc	0.23	-0.28	-
mt	-0.11	-0.04	-
pt	0.48	-0.18	chosen
Pareto important(yellow)	80.38%	83.50%	
Wilk's lamda	sig0.00		

The discriminant function obtain two discriminant functions with are passed Wilk’s lamda test. The 80% important weighting of attribute for each function are painted in yellow color. For example, total coefficient (ignore the sign) “fs”+ “pe”+ “rl”+ “pt” is 80.38% Pareto important.

In summary, the result of multivariate linear regression and discriminant analysis show that “fs”, “re” are two common attributes which presence in both (1) and Table VI. Likewise, the attribute “pe” is also presence in

both discriminant functions sf(1), and sf(2). Actually, ("cp", "us", and "pt") are all shared the Pareto importance therefore this research choose six attributes ("fs", "re", "pe", "cp", "us", and "pt") for classification tree construction in next session.

D. Classification tree generation

The J.48 classification method give most accuracy in classification on 60.00%, 33% ten folds cross validation, root mean square error value= 0.44, recall=0.60, and f-measure about 0.60. The J.48 decision tree as shown on figure 2.

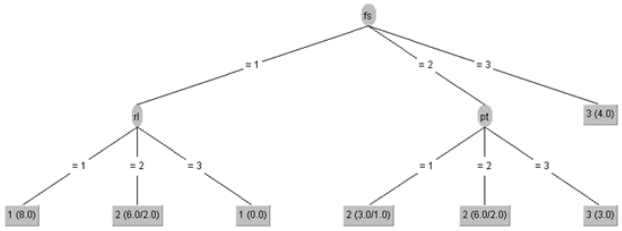

Figure. 2. J.48 decision tree for satisfaction classification rules.

The classification rules are:

1. If fs=1 and rl=1 then sf=1
2. If fs=1 and rl=2 then sf=2
3. If fs=1 and rl=3 then sf=1
4. If fs=2 and pt=1 then sf=2
5. If fs=2 and pt=2 then sf=2
6. If fs=2 and pt=3 then sf=3
7. If fs=3 then sf=3

The rule#3 presents that the satisfaction level value is level 1 satisfaction. Actually, there has no any observations (0.0) thus this rule could be deleted.

V. RESEARCH CONCLUSION

The result of classification show that attribute "fs" (functional suitability), "rl" (reliability), and "pt" (portability) are importance attributes that could increase accurately predict the customer satisfaction level on software product. Therefore, the software house must concern about three attributes. The importance of attribute sorted by descending are "fs", "rl", and "pt" respectively

ACKNOWLEDGMENT

The samples of this research were kindly supported from College of Business Administration, Innovation and Accounting, Dhurakij Pundit University.

REFERENCES

[1] ISO 25000 Portal,Softare quality model,2022. https://iso25000.com/index.php/en/iso-25000-standards/iso-25010

[2] International Association for SoftwareArchitects, Quality Attributes, 2022. Available on https://itabok.iasaglobal.org/quality-attributes/

[3] Syndicode, software architecture quality attributes, San Francisco, CA 94102. Available on https://syndicode.com/blog/12-software-architecture-quality-attributes/

[4] Huberty, C. J. and Olejnik, S. (2006). Applied MANOVA and Discriminant Analysis, Second Edition. Hoboken, New Jersey: John Wiley and Sons, Inc.

[5] Lund Research, Linear Regression Analysis using SPSS Statistics, Laerd Statistics, 2020. Available on https://statistics.laerd.com/spss-tutorials/linear-regression-using-spss-statistics.phpK. Elissa, "Title of paper if known," unpublished.

[6] Prashant Gupta, Decision Trees in Machine Learning, Towards Data Science, 2017. Available on https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052