Factor Affecting Selecting Web or Facebook Channel for Online Purchasing

Wasun Khan-Am[1], Kritiya Rangsom[2]

[1]Asscoiate Professor: dept. of Information Systems
Faculty of Business Administration, RMUTT
Pathum-thani, Thailand
e-mail: am.siamese@gmail.com
[2]Assistance Professor: dept. of Information Systems
Faculty of Business Administration, RMUTT
Pathum-thani, Thailand
e-mail:gritiya@hotmail.com

Abstract— this article presents a model to predict selecting channel for online purchasing of electronic commerce between web and facebook. There are 400 volunteers who live in Bangkok and nearest area. They all have an experience to purchase a goods from internet and also have willingness to give an answer for the research. By decision tree model that was created, the result is found that all decision rules are produced by occupation, purchase objective, gender, frequency of purchase, age, and income per month. A measure of the tree's predictive accuracy is 71.5%. A model's risk is .285. The cross-validation risk estimate for final tree is .327.

Keywords: Web Purchasing; Facebook Purchasing; Decision Tree; Prediction;

I. INTRODUCTION

A. Background

Nowadays, purchasing a goods from internet is very simple, and anyone is able to do it. The basic purchasing goods from internet is web commerce. Person googles a website for particular shop or store as needed, then shop, paid, and take a goods.

However, a social network such as facebook is a new channel for shopping goods. After login, the member of social network is allowed to participate a social such as post message, visits member's friend, send message and/or image to friend. Furthermore, facebook's member can post goods' image, quote a price in members' page. With this approach, the facebook's member become to a merchant. The merchant's page have an invitation to another facebook's member. Let them go to merchant's page and make a deal.

The research question is "What is a factors using to forecast style of personnel shopping between Web and Facebook". The results will be lent to suggestion for both web and facebook merchant to select goods, select type, and the other thing that is suitable to their client.

B. Objective

The research objective aims to build decision tree model for predict channel of purchasing, between web and facebook.

II. LITERATURE REVIEW

A. Electronic commerce[1-3]

Electronic commerce, E-commerce, is a trading products or services via electronic platform such as internet, mobile phone, EFT, etc. A modern e-commerce is conducted by world wide web.

E-commerce is divided into:

- B2C – Business –to-Consumer
- B2B – Business-to-Business
- C2B – Consumer-to-Business
- C2C – Consumer-to-Consumer
- B2G – Business-to-Government

The B2C is a target of this research

B. Web & Web commerce[1-3]

Web or website is an information system of inter-linked hypertext documents and other digital resources that are accessed via the internet. The hypertext document is a webpage which is created by web author with Hypertext Markup Language.

Web commerce is one form of electronic commerce that do commerce with web. Web owner setup website to display their products or services for customer shopping via web browser. A payment and product delivery may be done on web or outside.

C. Social Network[1-3]

Social Network is one kind of social which its structure made of nodes that may be an individuals or organization. Social network represents and relationships between people, groups, and organization.

The term "social network" was coined in 1954 by Professor J. A. Barnes[4] of the London School of Economics

The popular social network sites (SNS) are: Twitter, Facebook, Likedln, Flickr, google+, and so on.

In this research, the facebook is a target social network site.

D. Decision Tree Learining[5]

A decision tree is a decision support tool that look like a tree in graph theory. Decision tree learning use a decision tree as a predictive model to find out the item's target value. This model approaches used in statistics, data mining and machine learning.

A decision tree use an algorithm for dividing data into groups. By using algorithm, the entire data set are split into some number of pieces. And with a same algorithm the piece will be split into smaller pieces. In general, a piece may be either split of left alone to form a final group. The sample of decision tree shows in figure 1.

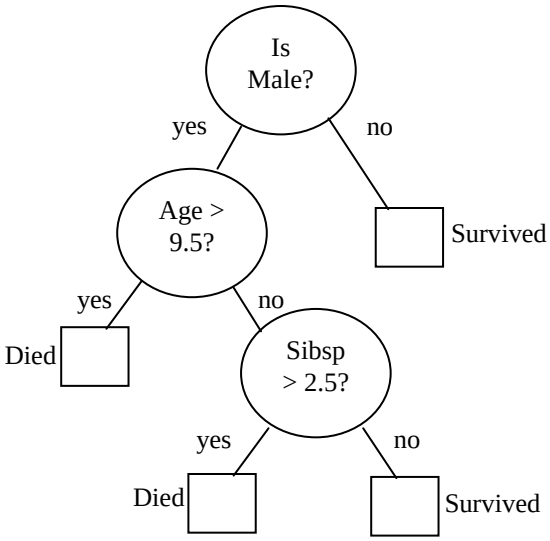

Figure 1. Example of decision tree

E. Marketing Mixed

The marketing mixed or 4Ps Marketing [6] is a business tools used in marketing and marketers. The 4Ps consist of product, price, place, and promotion as table below:

TABLE I. MEANING OF COMPONENT OF MARKETING MIXED

Component	Containing
Product/ Service	a wide variety of product, quality products, brand, packaging design, shape, packaging size, warranty, shape service and so on.
Price	price, a discount price, perceptions of consumers, term of payment.
Place	sales channels, available inventory, transportation, and so on.
Promotion	promotion, advertising, public relations, direct sales, sales promotion through distributors and so on

In this research, 3P including product with wide variety, quality, and shape aspects; price with cheap, reasonable, and bargain aspects; and promotion with discount, point collection aspects were applied.

F. IBM SPSS Decistion Tree

IBM SPSS Decision Tree [7] is a tools for identify groups, discover relationships between them and predict future events. There are four algorithms including:

- CHAID – A fast, statistical, multi-way tree algorithm that explores data quickly and efficiently
- Exhaustive CHIA – A adaptive of CHAID. Examination all possible splits is additional.
- CRT- Classification and regression trees – partitions data and builds accurate subsets
- QUEST- A algorithm for selecting variable without bias and quickly produces accurate tree

In this research, the CHAID algorithms was applied

III. METHODOLOGY

A. Population and Sample

A population of this research are people who live in Bangkok and a nearby province such as Samutprakarn, Nontaburi, Nakron pratom, and Pathum thani. In Department of Provincial Administration announcement [8], there were 5,673,560; 1,223,302; 874,616; and 1,033,837 citizens, including 8,805,315 in five provinces. By using a Taro Yamane formula, the calculated number of sample about 399.84 samples. So that, 400 samples was administered for this research.

B. Sampling Method and Data collection

The purposive sample is employed in this research. The questionnaire is use as a tool. Data collection was conducted by asking question "Have you ever purchasing goods via internet?" before administration, and was made at some of district office, sky train junction(BTS), underground junction (MRT), some of complex building, some of discount store, some of modern trade store, and some of office building. The respondent is a volunteer who willingness to do a questionnaire.

C. Statistics and Analysis

The collected questionnaires are classified in two groups, one for complete questionnaire and another for incomplete questionnaire. The completed questionnaires are processed with a computer program. The statistics were used as instruments to analyze data is decision tree model.

IV. RESULT

After finish a data collection and classification, the 400 complete questionnaires are processed by computer statistics program. There are three parts of result show in next section.

A. Demograhic

The general information about respondent are shown in the followed table.

TABLE II. DEMOGRAHCIC INFORMATION

Demographic	value	Statistics	
		N	*%*
Gender	Male	119	28.8
	Female	281	70.2
Age	< 20	29	7.3
	21-30	256	64.0
	31-40	73	18.2
	> 40	42	10.5
Education	Undergrad.	66	16.5
	Bachalor	280	70.0
	Above Bach.	54	13.5
Occupation	Student	56	14.0
	Public	44	11.0
	Private	300	75.0
Income/month	<= 20,000	73	18.2
	>20,000	327	81.8

The table above shows that the most of respondent is female (70.2%), age between 21 and 30, graduate at bachelor degree level, works in private sector, and earn more than 20,000 baht per month

B. Marketing Mixed

The marketing mixed affecting online purchasing goods are shown in the followed table

TABLE III. MARKETING MIXED

Marketing Mixed	Statistics	
	Mean	*Std.Dev*
Product – Quality	4.43	.641
Product – Picture	3.79	.863
Product – Variety	3.85	.766
Price – Cheap	4.01	.761
Price – Resonable	4.24	.807
Price – bargain	3.33	.943
Promotion – Discount	3.89	.817
Promotion – Point Collection	3,04	1.013
Promotion – Warranty	3.92	.932

The table shows almost decision making by marketing mixed are at high level except two aspects are: price – bargain, and promotion – point collection. Those are at moderate level.

C. Decision tree

After analysis data with decision tree, the result of analysis are:

TABLE IV. MODEL SUMARY

Group	*Description*	*value*
Specification	Growing Method	CHAID
	Dependent Variable	Channel_W_F
	Independent Variable	SEX, t_age, t_edu, t_occupy, t_income, t_prod, t_price, t_promotion, obj_buy, t_use, f_buy
	Validateion	Cross Validation
	Maximun Tree Depth	10
	Minimum Cases in Parent Node	50
	Minimum Cases in Child Node	15
Result	Independent Variable Include	SEX, t_age, t_occupy, t_income, obj_buy, t_use, f_buy
	Number of Nodes	20
	Number of Terminal Nodes	11
	Depth	5
	Risk	.285
	Cross-Validation Risk	.327

The decision tree diagram shows in figure 2 and the table for classification shows below:

TABLE V. CLASSIFICATION

Observed value	Predicted		
	Web	*Facebook*	*Percent Correct*
Web	200	52	79.4
Facebook	62	86	58.1
Ovarall Percentage	65.5	34.5	71.5

The table above shows correct percent of model prediction. The first row show: in the first column, 200 respondents select web and model prediction result is web; in the second column, 52 respondents select facebook but model prediction result is web. The second row show: in the first column, 62 respondents select web but model prediction result is facebook; and in the second column, 86 respondents select facebook and the model prediction result is facebook also.

D. RULE

The generated rule from this decision three model are:

```
/* Node 10 */
IF (t_occupy != "student")  AND  (Obj_buy != "co
nvinien"  AND  Obj_buy != "other"  AND  Obj_buy
!= "fashion")  AND  (SEX != "Male")  AND  (f_buy
 != "1 or more pw")
THEN
 Node = 10
 Prediction = 1
 Probability = 0.576271

/* Node 16 */
IF (t_occupy != "student")  AND  (Obj_buy != "co
nvinien"  AND  Obj_buy != "other"  AND  Obj_buy
!= "fashion")  AND  (SEX != "Male")  AND  (f_buy
 = "1 or more pw")  AND  (s_buy = "night")
THEN
 Node = 16
 Prediction = 1
 Probability = 0.923077

/* Node 17 */
IF (t_occupy != "student")  AND  (Obj_buy != "co
nvinien"  AND  Obj_buy != "other"  AND  Obj_buy
!= "fashion")  AND  (SEX != "Male")  AND  (f_buy
 = "1 or more pw")  AND  (s_buy != "night")
THEN
 Node = 17
 Prediction = 1
 Probability = 0.612903

/* Node 12 */
IF (t_occupy != "student")  AND  (Obj_buy != "co
nvinien"  AND  Obj_buy != "other"  AND  Obj_buy
!= "fashion")  AND  (SEX = "Male")  AND  (t_age
!= "work")
THEN
 Node = 12
 Prediction = 1
 Probability = 0.703704

/* Node 13 */
IF (t_occupy != "student")  AND  (Obj_buy != "co
nvinien"  AND  Obj_buy != "other"  AND  Obj_buy
!= "fashion")  AND  (SEX = "Male")  AND  (t_age
= "work")
THEN
 Node = 13
 Prediction = 1
 Probability = 1.000000

/* Node 18 */
IF (t_occupy != "student")  AND  (Obj_buy = "con
vinien")  AND  (t_income != "10001-
20000")  AND  (t_age != "work")  AND  (SEX != "M
ale")
THEN
 Node = 18
 Prediction = 2
 Probability = 0.575758

/* Node 19 */
IF (t_occupy != "student")  AND  (Obj_buy = "con
vinien")  AND  (t_income != "10001-
20000")  AND  (t_age != "work")  AND  (SEX = "Ma
le")
THEN
 Node = 19
 Prediction = 1
 Probability = 0.708333

/* Node 15 */
IF (t_occupy != "student")  AND  (Obj_buy = "con
vinien")  AND  (t_income != "10001-
20000")  AND  (t_age = "work")
THEN
 Node = 15
 Prediction = 1
 Probability = 0.810811

/* Node 9 */
IF (t_occupy != "student")  AND  (Obj_buy = "con
vinien")  AND  (t_income = "10001-20000")
THEN
 Node = 9
 Prediction = 2
 Probability = 0.875000

/* Node 5 */
IF (t_occupy != "student")  AND  (Obj_buy = "oth
er" OR Obj_buy = "fashion")
THEN
 Node = 5
 Prediction = 1
 Probability = 0.969697

/* Node 2 */
IF (t_occupy = "student")
THEN
 Node = 2
 Prediction = 2

 Probability = 0.607143
```

Where a variable name in the rule are:

TABLE VI. MEANING OF VARIBLE IN THE RULE

variable	Meaning
SEX	GENDER
t_age	AGE
t_occupy	OCCUPATION
t_income	INCOME PER MONTH
obj_buy	PURCHASING OBJECTIVE
t_use	DURATION FOR PURCHASING
f_buy	FREQUENTLY PURCHASING

V. CONCLUSION

The research result is shown that it is able to predict selecting channel for electronic commerce purchasing. The prediction accurate is 71.5%. The predictor are include: gender, age, occupation, income per month, purchasing objective, using time for purchasing, and frequently purchasing. All marketing mixed are not fits to be a predictor in this research.

VI. LIMITATION AND FUTURE RESEARCH

The limitations of this research are:

1. The number of respondents in this research is only 400. After research has done, it may causes a problem for classification with a lot of predictor. The future research should have more respondent.

2. The number of observe variable in each 4Ps are only three observe variables, it may be a cause of why the 4Ps does not exist in the model. So, the future research may be increase a number of observe variables for each Ps in 4Ps.

3. The future research may use another algorithm for building a decision tree model and compares the model result with this research.

REFERENCES

[1] Jessup, L.M., et al. (2008.) *Information systems today:managing in the digital, 5th ed.*. Prentice-hall Inc.

[2] Laudon, K.C. (2012). *Management Information System:management the digital firm.* Pearson.

[3] Turban, E. and D. King(2011). *Electronic commerce 2012:Managerial and Social Networks perspective*. Pearson Higher ed.

[4] Barnes, John (1954). "Class and Committees in a Norwegian Island Parish." *Human Relations*, (7): 39-58

[5] Han, J. and Kamber, M. (2006). *Data Minig : Concepts and Techiques, 2rd ed.* San Francisco:Morgan Kaufmann publishers.

[6] Kotler, P. (1997). Marketing management analysis, planning, implementation and control (9th ed.). Englewood Cliffs, NJ: Prentice-Hall..

[7] Spss, Data Mining, Statistical Analysis Software, Predictive Analysis, Predictive Analytics, Decision Support Systems.*IBM SPSS Decision Trees: Easily identify groups and predict outcomes.* Retrieved January 26, 2013, from http://spss.co.in/pasw/decisiontrees.aspx

[8] Offical Statistics Registration Systems.(Feb, 2013). *Offical Staistics Registeration Annoucement, The Number of population in Kingdom of Thailan at Bangkok and other provinces. As proof of registration as of December 31, 2012*. Retrieved January 26, 2013, from http://stat.bora.dopa.go.th/stat/y_stat55.html.

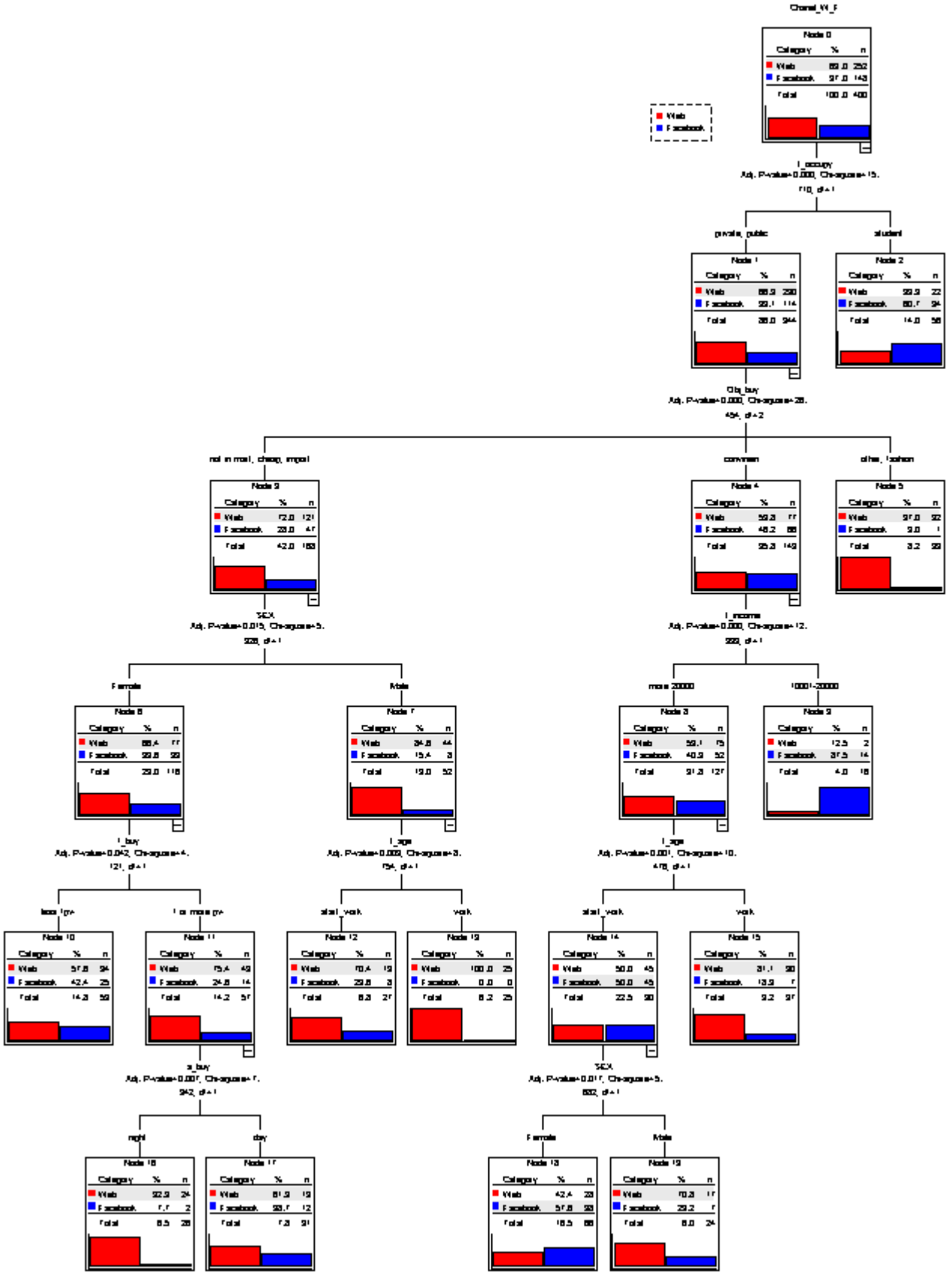

Figure 2. Resulted Decision Tree diagram