



รายงานการวิจัยฉบับสมบูรณ์

การ Mining ข้อมูลเกี่ยวกับปฏิกิริยาทางบวกของลูกค้าต่อการโฆษณา

ทางสื่อสังคมออนไลน์

DATA MINING FOR CUSTOMER'S POSITIVE REACTION TO ADVERTISING
IN SOCIAL MEDIA

รองศาสตราจารย์ ดร. วีระ บุญจริง

งานวิจัยนี้ได้รับทุนสนับสนุนงานวิจัย

จากงบประมาณเงินรายได้ ประจำปีงบประมาณ พ.ศ. 2559

วิทยาลัยนานาชาติ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

600268027
RE00013

ABSTRACT

The research aims at 1) finding the most important factors influencing positive customer reactions and purchasing merchandises after seeing online social media advertising and 2) identifying characteristics of customer clusters having positive reaction, as well as of purchasing customer clusters, after seeing online social media advertising. Data from 370 respondents are collected by questionnaires using convenience sampling method. Attribute selection and clustering techniques are employed in data analysis to find important factors and identify customer clusters, respectively. It is found that there is a strong correlation between the reason for clicking advertisement on social media and the satisfaction with merchandise, and between purchasing merchandise online and saving information for further consideration. The findings also indicate the characteristics of “Product conscious” and “Price Conscious” clusters for customer’s reaction and purchasing after seeing online social media advertising.

Keywords: Mining, Advertising, Social media, Customer reaction

Table of Contents

	Page
Abstract	i
Table of Contents	ii
List of Tables.....	iii
1 Introduction.....	1
1.1 Motivation.....	1
1.2 Objective	1
1.3 Scope.....	2
1.4 Report Organization.....	2
2 Literature Review.....	3
3 Methodology.....	5
4 Results and Discussion.....	7
5 Conclusion and Recommendation	13
References	15
Appendix	17

List of Tables

Table	Page
3.1 Sample Demographics	6
4.1 Ranked attributes for positive reaction	8
4.2 Ranked attributes for purchasing merchandise	9
4.3 Customers cluster based on positive reaction after seeing social media advertising.....	10
4.4 Customers cluster based on purchasing after seeing social media advertising.....	11



Chapter 1

Introduction

1.1 Motivation

With high popularity of social media, several social media websites have been developed such as Line, Facebook, Twitter, etc. and the usage rate has increased every year. Social media has become a fast and easy way to reach people in almost every group categorized by age, occupation, and education etc. with less cost. The usage of social media is widespread in private and public environments. In the highly competitive business world, social media has become a source of large amounts of data that is extremely useful when data are analyzed properly. The results from analyzing data properly can be useful in variety of disciplines including education, business, politics, social, science, technology, etc.

Marketing campaigns with online advertising are one of the methods that businesses use for increasing purchasing motivation. Finding target customer characteristics and customer reactions to social media advertising helps to reach more customers and is useful information for a marketing campaign. Data mining is one of the known techniques to analyze data to find hidden information from a large amount of available data without having prior hypotheses. Data mining provides a variety of methods such as association, classification, clustering, etc. for analyzing data, but selecting a method to match with the objectives is a challenge.

1.2 Objective

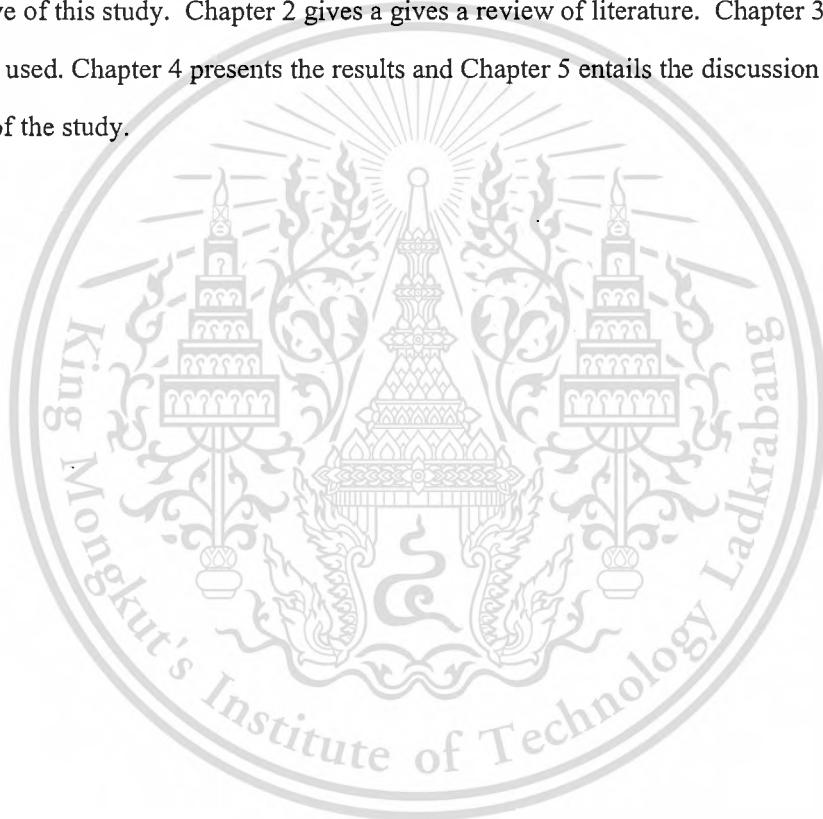
The purposes of this study are to 1) find the most important factors influencing positive customer reactions after seeing online social media advertising, 2) find the most important factors influencing purchasing products advertised online, 3) identify customer clusters characteristics that have positive reaction after seeing online social media advertising, and 4) identify customer clusters characteristics that purchase merchandise after seeing online social media advertising.

1.3 Scope

The study developed a questionnaire based on previous studies. It conducted and distributed a survey to 370 people who used to purchase product or service online through social media. Then, it applied data mining techniques, including attribute ranking and clustering, to answer the research questions.

1.4 Report Organization

This report is divided into five chapters. The first chapter presents the introduction and the main objective of this study. Chapter 2 gives a review of literature. Chapter 3 described methodology used. Chapter 4 presents the results and Chapter 5 entails the discussion and conclusions of the study.



Chapter 2

Literature Review

Data mining has been defined as the process to extract knowledge from large quantities of data in order to discover meaningful patterns and rules [1]. Reference [2] defines data mining as the analysis of (often large) observational data sets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner. Reference [3] defines data mining as a process of knowledge discovery. Reference [4] summarized that data mining contains three key stages: finding patterns, interpreting them in order to check their usefulness, and finally using the patterns to solve business problems. The ultimate goal of data mining is to discover knowledge and it will be useful in several disciplines. In business, data mining is used for strategic benefit such as direct marketing, trend analysis, etc. In direct marketing, data mining is used for targeting people who are most likely to buy certain products and services. For trend analysis, data mining is used to determine trends in the marketplace [5].

Reference [4] explains that there are two main types of DM models as follows: 1) Predictive model: This model is constructed to predict a particular outcome or target variable. Commonly used predictive modeling techniques include multiple regression (for predicting value data), logistic regression (for response prediction) and decision trees (for rule-based value or response models). 2) Descriptive model: This model gives a better understanding of the data, without any single specific target variable. Commonly used descriptive techniques include factor analysis (to extract underlying dimensions from multivariate data), cluster analysis (for grouping a customer database into segments), and association analysis (for discovering relationships between items such as retail products).

A marketing campaign is a specifically defined series of activities used in marketing a new or changed product or service, or in using new marketing channels and methods. Marketing activities are efforts to increase awareness for a particular product or service. Social media is one of the most popular marketing channels due to the ability to reach large numbers of customers with low cost. Social media advertising helps businesses find new potential clients

by using the users' own shared information to identify interest. Rather than re-actively targeting users who search for a certain term, social media advertising proactively targets relevant users before they even begin their searches.

Prior research that relates to the study topics were reviewed as follows: Reference [6] researched the factors that influence the recipients to open direct emails and make an action desired by the company and also studies whether and what elements in the email would influence them to buy the products or services promoted. The results are obtained based on a data mining analysis which includes clustering and classification processes and offer a guide on how organizations should design their email marketing communications in order to have higher response rates. Reference [7] researched on the value of social technologies in organizations based on the 'value focused thinking' approach. The findings highlight innovation of internal processes, creation of organizational identity and new business models, integrated business functions, as well as employee support to be important values of social technology enabled innovation in organizations. Reference [8] researched on the segmenting of consumer reactions to social network marketing. The purpose of the study is to understand how consumers may be segmented with respect to their reactions to social network marketing. The results identified five segments – Passive, Talkers, Hesitant, Active, and Averse –along with significant covariates such as information search, convenience, entertainment, age, and gender that predict membership.

Chapter 3

Methodology

Several aspects related to the factors that influence a respondent to react and purchase a product or service after seeing social media advertising are explored from the literature review. Based on the literature review, the questionnaires are developed and distributed to people who used to purchase product or service through social media such as Facebook, Line, Instagram, etc. The questionnaires are composed of the three parts; the first part is about the demographic data of the respondents, the second part is about the respondent's reaction when seeing social media marketing campaigns, and the third part is open-ended questions about the respondent's opinion. The data has been gathered at the level of samples including 370 respondents aged less than 60 and being employees, freelancers, entrepreneurs, managers, and students as shown in Table 3.1. The data has been collected between January and April 2017 through an online survey by using the convenience sampling method.

The research directions include: determining the most important factors that influence the customers to have a positive reaction after seeing social media advertising; determining the most important factors influencing purchasing products that advertise online in social media; identifying the customers characteristics that have positive reaction after seeing social media advertising; and identifying the customer clusters characteristics that purchase merchandises after seeing social media advertising. The software used to analyze data is WEKA data mining software. The research employs an attribute evaluator called "CorrelationAttributeEval" for determining the most important factors with respect to class attributes. It also uses the "SimpleKMeans" clustering algorithm for grouping customers.

Table 3.1: Sample Demographics

Gender	Frequency	Percentage			
Males	117	31.6			
Females	253	68.4			
Total	370	100			
Age	Frequency	Percentage	Average Income	Frequency	Percentage
<21 years	69	18.6	<=15,000 THB	243	65.7
21-30 years	244	65.9	>15,000 - 20,000 THB	66	17.8
31-40 years	26	7.0	>20,000 - 30,000 THB	27	7.3
>41 years	31	8.4	>30,000 THB	34	9.2
Total	370	100	Total	370	100
Status	Frequency	Percentage	Education	Frequency	Percentage
Single	317	85.7	<Bachelor's degree	54	14.6
Married	38	10.3	Bachelors' degree	283	76.5
Others	15	4.1	>Bachelors' degree	33	8.9
Total	370	100	Total	370	100
Occupation	Frequency	Percentage	Social media used	Frequency	Percentage
Private	78	21.1	Facebook	269	72.7
Government	41	11.1	Line	50	13.5
Freelance	23	6.2	Instagram	31	8.4
Business owner	13	3.5	Twitter	10	2.7
Student	215	58.1	Others	10	2.7
Total	370	100	Total	370	100
Spent/transaction	Frequency	Percentage	Merchandise	Frequency	Percentage
<= 500 THB	117	31.6	Fashion	200	54.1
>500 - 1,000 THB	138	37.3	Accessories	18	4.9
>1,000 - 2,000 THB	51	13.8	Beauty and health	35	9.5
>2,000 - 3,000 THB	22	5.9	IT and software	54	14.6
>3,000 - 4,000 THB	42	11.4	Airplane ticket	12	3.2
			Entertain	33	8.9
			Others	18	4.8
Total	370	100	Total	370	100

Chapter 4

Results and Discussion

The analysis results (as shown in Table 3.1) indicate that most of the respondents are students, females, aged 21-30, single, and educational level of Bachelor's degree, and average income less than 15,000 THB. Highest percentage on the usage of social media is Facebook, followed by Line and Instagram. The most purchasing are in fashion merchandise, followed by IT and software, and beauty and health merchandise. Spending per transaction is between 500 to 1,000 THB.

In order to identify the most important factors that influence the customers to have a positive reaction after seeing social media advertising, the class attribute called ReasonForClickingAdvertisementOnSocial Media is selected. The relation between the class attribute and the others is determined using CorrelationAttributeEval. The results present that for the customers (as shown in Table 4.1), there is a strong correlation between the reason for clicking advertisement on social media and the satisfaction with merchandise (0.12727), special rewards (0.08522), not buying (0.0852), and link for searching on more details (0.08507). Other attributes are ranked below 0.07.

Table 4.1.
Ranked attributes for positive reaction

Ranked attributes:	
0.12727	16 Q15_4_Advertise Influence Buying: Satisfaction with merchandises or services
0.08522	22 Q16_3_Marketing Campaign Influence Buying: Special rewards
0.0852	11 Q14_5_Influence of Advertisement on Social Media: Not buying
0.08507	4 Q12_4_Interested in Seeing Advertisement on Social Media: link for searching more details
0.07962	17 Q15_5_Advertise Influence Buying: more refer in social media
0.07643	13 Q15_1_Advertise Influence Buying: interesting price
0.07573	1 Q12_1_Interested in Seeing Advertisement on Social Media: merchandise logo
0.07475	18 Q15_6_Advertise Influence Buying: like marketing campaign
0.07273	9 Q14_3_Influence of Advertisement on Social Media: sent it out for more comments
0.06776	2 12_2_Interested in Seeing Advertisement on Social Media:text, image, clip advertisement
0.06414	24 16_5_Marketing Campaign Influence Buying: others
0.05661	21 16_2_Marketing Campaign Influence Buying: having complimentary
0.03896	5 12_5_Interested in Seeing Advertisement on Social Media: package or merchandise
0.03451	20 16_1_Marketing Campaign Influence Buying: high percentage of discount
0.03409	7 14_1_Influence of Advertisement on Social Media: immediately buy if satisfy
0.03283	10 14_4_Influence of Advertisement on Social Media: irritating and don't like advertisement
0.03194	6 12_6_Interested in Seeing Advertisement on Social Media: others
0.03194	12 14_6_Influence of Advertisement on Social Media: others
0.03086	14 15_2_Advertise Influence Buying: popular brand
0.0304	15 15_3_Advertise Influence Buying: exciting new merchandise
0.02497	3 12_3_Interested in Seeing Advertisement on Social Media: marketing campaign
0.02362	19 15_7_Advertise Influence Buying: others
0.01728	8 14_2_Influence of Advertisement on Social Media: saving information for further consideration
0.0063	23 16_4_Marketing Campaign Influence Buying: having after sales service

The class attribute PurchasingMerchandiseOnline is selected to determine the most important factors influencing purchasing products with social media advertising. The relation between the class attribute and the others is determined using CorrelationAttributeEval. The results present (as shown in Table 4.2) that for the customers, there is a strong correlation between purchasing merchandise online and saving information for further consideration (0.13879), merchandise logo (0.10179), and immediately purchase if satisfied (0.09747). Other attributes are ranked below 0.09.

Table 4.2.
Ranked attributes for purchasing merchandise

Ranked attributes:

0.13879	8	Q14_2	Influence of Advertisement on Social Media: saving information for further consideration
0.10179	1	Q12_1	Interested in Seeing Advertisement on Social Media: merchandise logo
0.09747	7	Q14_1	Influence of Advertisement on Social Media: immediately purchase if satisfy
0.08843	10	Q14_4	Influence of Advertisement on Social Media: irritating and don't like advertisement
0.08675	3	Q12_3	Interested in Seeing Advertisement on Social Media: marketing campaign
0.06656	13	Q15_1	Advertise Influence Buying: interesting price
0.05314	16	Q15_4	Advertise Influence Buying: Satisfaction with merchandises or services
0.04368	5	Q12_5	Interested in Seeing Advertisement on Social Media: package or merchandise
0.04144	2	Q12_2	Interested in Seeing Advertisement on Social Media: text, image, clip advertisement
0.03877	11	Q14_5	Influence of Advertisement on Social Media: Not buying
0.03841	22	Q16_3	Marketing Campaign Influence Buying: Special rewards
0.0357	17	Q15_5	Advertise Influence Buying: more refer in social media
0.03195	4	Q12_4	Interested in Seeing Advertisement on Social Media: link for searching more details
0.03025	18	Q15_6	Advertise Influence Buying: like marketing campaign
0.02892	24	Q16_5	Marketing Campaign Influence Buying: others
0.02329	15	Q15_3	Advertise Influence Buying: exciting new merchandise
0.02039	19	Q15_7	Advertise Influence Buying: others
0.01614	20	Q16_1	Marketing Campaign Influence Buying: high percentage of discount
0.01488	14	Q15_2	Advertise Influence Buying: popular brand
0.0144	12	Q14_6	Influence of Advertisement on Social Media: others
0.0144	6	Q12_6	Interested in Seeing Advertisement on Social Media: others
0.00799	23	Q16_4	Marketing Campaign Influence Buying: having after sales service
0.00383	9	Q14_3	Influence of Advertisement on Social Media: sent it out for more comments
0.0038	21	Q16_2	Marketing Campaign Influence Buying: having complimentary

The EM (Expectation Maximization) and “SimpleKMeans” clustering algorithms are used for grouping similar customers based on positive reaction after seeing social media advertising. The EM algorithm is used to identify the approximated cluster numbers. In this case, the result is 2 clusters. This value is used as a parameter for the “SimpleKMeans” algorithm. The algorithm results are presented in Table 4.3.

Table 4.3. Customers cluster based on positive reaction after seeing social media advertising

No.	Attributes	Cluster 0 (94, 25.4%)	Cluster 1 (276, 74.6%)
1	Interested in Seeing Advertisement on Social Media: merchandise logo	0.2128	0.0725
2	Interested in Seeing Advertisement on Social Media: text, image, and clip	0.266	0.6522
3	Interested in Seeing Advertisement on Social Media: marketing campaign	0.2766	0.1739
4	Interested in Seeing Advertisement on Social Media: link for searching more details	0.2234	0.0725
5	Interested in Seeing Advertisement on Social Media: package or merchandise	0.0213	0.0254
6	Interested in Seeing Advertisement on Social Media: others	0	0.0036
7	Influence of Advertisement on Social Media: immediately purchase if satisfy	0.1596	0.1667
8	Influence of Advertisement on Social Media: saving information for further consideration	0.7447	0.6993
9	Influence of Advertisement on Social Media: sent it out for more comments	0.0426	0.0326
10	Influence of Advertisement on Social Media: irritating and don't like advertisement	0.0532	0.0725
11	Influence of Advertisement on Social Media: not buying	0	0.0254
12	Influence of Advertisement on Social Media: others	0	0.0036
13	Advertise Influence Buying: interesting price	0.3617	0.3659
14	Advertise Influence Buying: popular brand	0.1383	0.1268
15	Advertise Influence Buying: exciting new merchandise	0.0745	0.0507
16	Advertise Influence Buying: Satisfaction with merchandises or Services	0.2872	0.3732
17	Advertise Influence Buying: more refer in social media	0.1064	0.058
18	Advertise Influence Buying: like marketing campaign	0.0213	0.0217
19	Advertise Influence Buying: others	0.0106	0.0036
20	Marketing Campaign Influence Buying: high percentage of discount	0.6383	0.6449
21	Marketing Campaign Influence Buying: having complimentary	0.1277	0.1522
22	Marketing Campaign Influence Buying: special rewards	0.0319	0.0145
23	Marketing Campaign Influence Buying: having after sales service	0.2021	0.1739
24	Marketing Campaign Influence Buying: others	0	0.0145

Cluster 0: Most customers in this group save information for further consideration (74%) and are interested in high percentage of discount (64%) after seeing advertising in the social media. For this cluster, the advertising should emphasize on merchandise quality as well as price to create more positive reaction.

Cluster 1: Most customers in this group save information for further consideration (70%), are interested in high percentage of discount (64%) after seeing advertising in the social media, and are also interested in seeing text, image, and clip advertising (65%). For this cluster, clip advertising may increase the customer's interest and create more positive reaction.

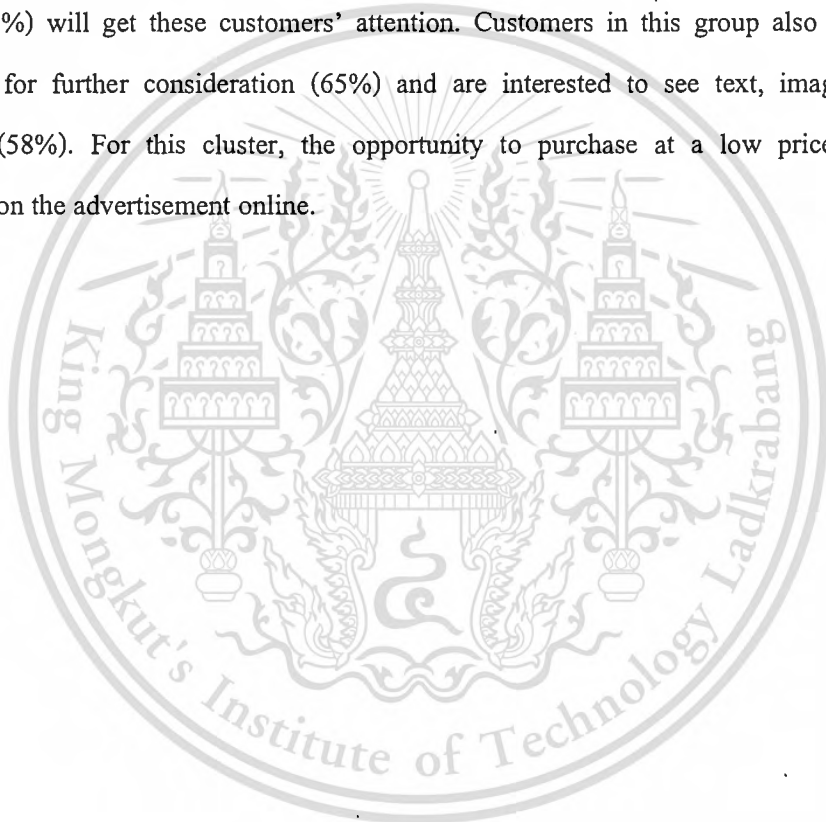
Clustering customers based on purchasing merchandises after seeing social media advertising is similar to clustering based on positive reaction after seeing social media advertising . The results are presented in Table 4.4.

Table 4.4. Customers cluster based on purchasing after seeing social media advertising

No.	Attributes	Cluster 0 (235, 63.5%)	Cluster 1 (135, 36.5%)
1	Interested in Seeing Advertisement on Social Media: merchandise logo	0.0894	0.1407
2	Interested in Seeing Advertisement on Social Media: text, image, and clip advertising	0.5362	0.5852
3	Interested in Seeing Advertisement on Social Media: marketing campaign	0.2383	0.1333
4	Interested in Seeing Advertisement on Social Media: link for searching more details	0.1021	0.1259
5	Interested in Seeing Advertisement on Social Media: package or merchandise	0.0298	0.0148
6	Interested in Seeing Advertisement on Social Media: others	0.0043	0
7	Influence of Advertisement on Social Media: immediately purchase if satisfy	0.1234	0.237
8	Influence of Advertisement on Social Media: saving Information for further consideration	0.7489	0.6444
9	Influence of Advertisement on Social Media: sent it out for more comments	0.0468	0.0148
10	Influence of Advertisement on Social Media: irritating and don't like advertisement	0.0596	0.0815
11	Influence of Advertisement on Social Media: not buying	0.017	0.0222
12	Influence of Advertisement on Social Media: others	0.0043	0
13	Advertise Influence Buying: interesting price	0	1
14	Advertise Influence Buying: popular brand	0.2043	0
15	Advertise Influence Buying: exciting new merchandise	0.0894	0
16	Advertise Influence Buying: Satisfaction with merchandises or services	0.5532	0
17	Advertise Influence Buying: more refer in social media	0.1106	0
18	Advertise Influence Buying: like marketing campaign	0.034	0
19	Advertise Influence Buying: others	0.0085	0
20	Marketing Campaign Influence Buying: high percentage of discount	0.6043	0.7111
21	Marketing Campaign Influence Buying: having complimentary	0.1362	0.163
22	Marketing Campaign Influence Buying: special rewards	0.017	0.0222
23	Marketing Campaign Influence Buying: having after sales service	0.2255	0.1037
24	Marketing Campaign Influence Buying: others	0.017	0

Cluster 0: This group of customers can be called “Product conscious.” Most customers in this group save information for further consideration (74%) after seeing advertising on the social media and purchase when satisfied with merchandise or services (55%). Seeing text, image, and clip advertising is also interesting for this group (53%). For this cluster, the advertising should emphasize on quality and brand to justify the price.

Cluster 1: This group can be called “Price conscious.” Everyone in this group is concerned about price before purchasing merchandise. Interesting price (100%) and high percentage of discount (71%) will get these customers’ attention. Customers in this group also like to save information for further consideration (65%) and are interested to see text, image, and clip advertising (58%). For this cluster, the opportunity to purchase at a low price should be emphasized on the advertisement online.



Chapter 5

Conclusion and Recommendation

Based on the findings from this research, e-commerce business should draw attention to the content that is emphasized in the advertisement online, due to some clusters that are product conscious while others are price conscious. For the product conscious group, price does not matter as much for them and they decide to purchase based on satisfaction with the merchandise after thoroughly considering. Merchandise quality alone can get the product conscious group to purchase or have positive reaction after seeing online advertising in social media. For the price conscious group, they are looking for an interesting price and a high percentage of discount. Presenting interesting price will draw high attention from the price conscious group. However, both groups are interested in the advertisement online, and the advertisement that can draw customer's attention should provide interesting text, image, and clip advertisement. The advertisement should be attention-getting about the pricing so that it would convince the customers to click on the advertisement and continue to purchase the merchandise. The advertisement should also emphasize the high percentage of discount, which can draw attention from customers in both clusters that derive from positive reaction and lead to the purchasing of merchandise after seeing online advertising in social media.

The major limitation of the research consists in the low number of respondents (only 370), so this exploratory research should be followed by a conclusive one to verify the conclusions of the present research. Also, the majority of the respondents were students, aged 21-30, with low income and low spending per transaction, which may have influenced the positive reaction after seeing online advertising in social media.

Future directions of research may include: (1) using association techniques to determine rules related to positive reactions and purchasing, (2) adding some other information about respondents like time spent in front of the computer and time spent in using social network, and analyzing the influence of those factors on positive reactions and the purchasing of merchandise.





References

1. M. J. A. Berry and G. Linoff, *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management* 3rd ed., John Wiley and Sons Ltd., Publication, UK, 2011.
2. D. J. Hand, H. Mannila and P. Smyth, *Principles of Data Mining*, MIT Press, Cambridge, MA, 2001.
3. J. Han and M. Kamber, *Data Mining Concepts and Techniques Second Edition*, Morgan Kaufmann Publishers, United States of America, 2006.
4. B. Leventhal, An introduction to data mining and other techniques for advanced analytics, *Journal of Direct, Data and Digital Marketing Practice*, 12(2), 2010, pp.137–153.
5. M. Negnevitsky, *Artificial Intelligence, A Guide To Intelligent Systems* 3rd ed., Pearson Education Limited, 2005.
6. R. I. Magos and C. A. Acatrinei, *Designing Email Marketing Campaigns - A Data Mining Approach Based on Consumer Preferences*, *Annales Universitatis Apulensis Series Oeconomica*, 17(1), 2015, 15-30.
7. M. Singh and K. Peszynski, “Organisational Value of Social Technologies: An Australian Study”, *The Electronic Journal Information Systems Evaluation*, 17(1), 2014, pp.088-099, available online at www.ejise.com.
8. C. Campbell, C. Ferraro, and S. Sands, “Segmenting consumer reactions to social network marketing”, *European Journal of Marketing*, 48(3/4), 2014, pp.432-452.



This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Annals of Computer Science and Information Systems

Volume 11

Proceedings of the 2017 Federated
Conference on Computer Science and
Information Systems

September 3–6, 2017. Prague, Czech Republic



Maria Ganzha, Leszek Maciaszek, Marcin Paprzycki (eds.)



This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Annals of Computer Science and Information Systems, Volume 11

Series editors:

Maria Ganzha,

Systems Research Institute Polish Academy of Sciences and Warsaw University of Technology, Poland

Leszek Maciaszek,

Wrocław University of Economy, Poland and Macquarie University, Australia

Marcin Paprzycki,

Systems Research Institute Polish Academy of Sciences and Management Academy, Poland

Senior Editorial Board:

Wil van der Aalst,

Department of Mathematics & Computer Science, Technische Universiteit Eindhoven (TU/e), Eindhoven, Netherlands

Frederik Ahlemann,

University of Duisburg-Essen, Germany

Marco Aiello,

Faculty of Mathematics and Natural Sciences, Distributed Systems, University of Groningen, Groningen, Netherlands

Mohammed Atiquzzaman,

School of Computer Science, University of Oklahoma, Norman, USA

Barrett Bryant,

Department of Computer Science and Engineering, University of North Texas, Denton, USA

Ana Fred,

Department of Electrical and Computer Engineering, Instituto Superior Técnico (IST—Technical University of Lisbon), Lisbon, Portugal.

Janusz Górski,

Department of Software Engineering, Gdansk University of Technology, Gdansk, Poland

Mike Hinchey,

Lero—the Irish Software Engineering Research Centre, University of Limerick, Ireland

Janusz Kacprzyk,

Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Irwin King,

The Chinese University of Hong Kong, Hong Kong

Juliusz L. Kulikowski,

Nalecz Institute of Biocybernetics and Biomedical Engineering, Polish Academy of Sciences, Warsaw, Poland

Michael Luck,

Department of Informatics, King's College London, London, United Kingdom

Jan Madey,

Faculty of Mathematics, Informatics and Mechanics at the University of Warsaw, Poland

Stan Matwin,

Dalhousie University, University of Ottawa, Canada and Institute of Computer Science, Polish Academy of Science, Poland

Michael Segal,

Ben-Gurion University of the Negev, Israel

Andrzej Skowron,

Faculty of Mathematics, Informatics and Mechanics at the University of Warsaw, Poland

John F. Sowa,

VivoMind Research, LLC, USA

Editorial Associate:

Katarzyna Wasielewska,

Systems Research Institute Polish Academy of Sciences, Poland

Paweł Sitek,

Kielce University of Technology, Kielce, Poland

T_EXnical editor: Aleksander Denisiuk,

University of Warmia and Mazury in Olsztyn, Poland



Proceedings of the 2017 Federated Conference on Computer Science and Information Systems

Maria Ganzha, Leszek Maciaszek, Marcin Paprzycki
(eds.)



2017, Warszawa,
Polskie Towarzystwo
Informatyczne



2017, New York City,
Institute of Electrical and
Electronics Engineers

This material is reserved for educational use only, not allowed for commercial use.
Forbidden to modify the content, and cite the document when use.

Annals of Computer Science and Information Systems, Volume 11
Proceedings of the 2017 Federated Conference on Computer Science and
Information Systems.

ART: ISBN 978-83-946253-9-9, IEEE Catalog Number CFP1785N-ART
USB: ISBN 978-83-946253-8-2, IEEE Catalog Number CFP1785N-USB
WEB: ISBN 978-83-946253-7-5

ISSN 2300-5963
DOI 10.15439/978-83-946253-7-5

© 2017, Polskie Towarzystwo Informatyczne
Ul. Solec 38/103
00-394 Warsaw, Poland

© 2017, IEEE
10662 Los Vaqueros Circle
Los Alamitos, CA 90720, USA

Contact: secretariat@fedcsis.org
<http://annals-csis.org/>

Cover art:

Jana Waleria Denisiuk,
Elbląg, Poland

Also in this series:

Volume 13: Communication Papers of the 2017 Federated Conference on Computer Science and Information Systems, ISBN WEB: 978-83-922646-2-0, ISBN USB: 978-83-922646-3-7
Volume 12: Position Papers of the 2017 Federated Conference on Computer Science and Information Systems, ISBN WEB: 978-83-922646-0-6, ISBN USB: 978-83-922646-1-3
Volume 10: Proceedings of the Second International Conference on Research in Intelligent and Computing in Engineering, ISBN WEB: 978-83-65750-05-1, ISBN USB: 978-83-65750-06-8
Volume 9: Position Papers of the 2016 Federated Conference on Computer Science and Information Systems, ISBN WEB: 978-83-60810-93-4, ISBN USB: 978-83-60810-94-1
Volume 8: Proceedings of the 2016 Federated Conference on Computer Science and Information Systems, ISBN WEB: 978-83-60810-90-3, ISBN USB: 978-83-60810-91-0, ISBN ART: 978-83-60910-92-7
Volume 7: Proceedings of the LQMR Workshop, ISBN WEB: 978-83-60810-78-1, ISBN USB: 978-83-60810-79-8
Volume 6: Position Papers of the 2015 Federated Conference on Computer Science and Information Systems, ISBN WEB: 978-83-60810-76-7, ISBN USB: 978-83-60810-77-4
Volume 5: Proceedings of the 2015 Federated Conference on Computer Science and Information Systems, ISBN WEB: 978-83-60810-66-8, ISBN USB: 978-83-60810-67-5
Volume 4: Proceedings of the E2LP Workshop, ISBN WEB: 978-83-60810-64-4, ISBN USB: 978-83-60810-63-7
Volume 3: Position Papers of the 2014 Federated Conference on Computer Science and Information Systems, ISBN WEB: 978-83-60810-60-6, ISBN USB: 978-83-60810-59-0
Volume 2: Proceedings of the 2014 Federated Conference on Computer Science and Information Systems, ISBN WEB: 978-83-60810-58-3, ISBN USB: 978-83-60810-57-6, ISBN ART: 978-83-60810-61-3

DEAR Reader, it is our pleasure to present to you Proceedings of the 2017 Federated Conference on Computer Science and Information Systems (FedCSIS), which took place, for the first time outside of Poland, in Prague, Czech Republic, on September 3-6, 2017.

FedCSIS 2017 was Chaired by prof. Pavel Tvrđik, while prof. Jan Janousek acted as the Chair of the Organizing Committee. This year, FedCSIS was organized by the Polish Information Processing Society (Mazovia Chapter), IEEE Poland Section Computer Society Chapter, Systems Research Institute Polish Academy of Sciences, Warsaw University of Technology, Wrocław University of Economics, and Czech Technical University in Prague.

FedCSIS 2017 was technically co-sponsored by: IEEE Region 8, IEEE Czechoslovakia Section, IEEE Poland Section, IEEE Computer Society, IEEE Computer Society Technical Committee on Intelligent Informatics, IEEE Czechoslovakia Section Computer Society Chapter, IEEE Poland Section Gdańsk Computer Society Chapter Poland, SMC Technical Committee on Computational Collective Intelligence, IEEE Poland Section Systems, Man, and Cybernetics Society Chapter, IEEE Poland Section Control System Society Chapter, IEEE Poland Section Computational Intelligence Society Chapter, ACM Special Interest Group on Applied Computing, Łódź ACM Chapter, International Federation for Information Processing, Committee of Computer Science of the Polish Academy of Sciences, Polish Operational and Systems Research Society, Mazovia Cluster ICT Poland, Polski Klaster Badań i Rozwoju Internetu Rzeczy, and Eastern Cluster ICT Poland. FedCSIS 2017 was sponsored by Intel, Profinit and Abra.

FedCSIS 2017 consisted of the following events (conferences, symposia, workshops, special sessions). These events were grouped into FedCSIS conference areas, of various degree of integration. Specifically, those listed without indication of the year 2017 signify "abstract areas" with no direct paper submissions to them (but with submissions to their enclosed events).

- **AAIA'17 – 12th International Symposium Advances in Artificial Intelligence and Applications**
 - AIMA'17 – 7th International Workshop on Artificial Intelligence in Medical Applications
 - AIRIM'17 – 2nd International Workshop on AI aspects of Reasoning, Information, and Memory
 - ASIR'17 – 7th International Workshop on Advances in Semantic Information Retrieval
 - JAWS'17 – 11th Joint Agent-oriented Workshops in Synergy
 - LTA'17 – 2st International Workshop on Language Technologies and Applications
 - WCO'17 – 10th International Workshop on Computational Optimization
- **CSS - Computer Science & Systems**
 - CANA'17 – 10th Computer Aspects of Numerical Algorithms
 - C&SS'17 – 4th International Conference on Cryptography and Security Systems

- CPORA'17 – 2nd Workshop on Constraint Programming and Operation Research Applications
- MMAP'17 – 10th International Symposium on Multimedia Applications and Processing
- WAPL'17 – 6th Workshop on Advances in Programming Languages
- WSC'17 – 9th Workshop on Scalable Computing
- **iNetSApp – International Conference on Innovative Network Systems and Applications**
 - INSERT'17 – 1st International Conference on Security, Privacy, and Trust
 - IoT-ECAW'17 – 1st Workshop on Internet of Things – Enablers, Challenges and Applications
 - WSN'17 – 6th International Conference on Wireless Sensor Networks
- **IT4MBS – Information Technology for Management, Business & Society**
 - AITM'17 – 15th Conference on Advanced Information Technologies for Management
 - ISM'17 – 12th Conference on Information Systems Management
 - IT4L'17 – 5th Workshop on Information Technologies for Logistics
 - KAM'17 – 23rd Conference on Knowledge Acquisition and Management
- **SSD&A – Software Systems Development & Applications**
 - IWCPs'17 – 4th International Workshop on Cyber-Physical Systems
 - LASD'17 – 1st International Conference on Lean and Agile Software Development
 - MIDI'17 – 4th Conference on Multimedia, Interaction, Design and Innovation
 - SEW-37 – The 37th IEEE Software Engineering Workshop
- **DS-RAIT'17 – 4th Doctoral Symposium on Recent Advances in Information Technology**

The 2017 edition of an AAIA'17 Data Mining Challenge, focused on "Helping AI to Play Hearthstone". Its results constitute a separate section in these proceedings. Awards for the winners of the contest were sponsored by: Silver Bullet Solutions and the Mazovia Chapter of the Polish Information Processing Society. Papers resulting from the competition constitute a separate section of these Proceedings.

Each paper, found in this volume, was refereed by at least two referees and the acceptance rate of regular full papers was ~19.3% (96 papers out of 497 general submissions).

The program of FedCSIS required a dedicated effort of many people. Each event constituting FedCSIS had its own Organizing and Program Committee. We would like to express our warmest gratitude to all Committee members for

their hard work in attracting and later refereeing 497 submissions (regular and data mining).

We thank the authors of papers for their great contribution to research and practice in computing and information systems. We thank the invited speakers for sharing their knowledge and wisdom with the participants. Finally, we thank all those responsible for staging the conference in Prague. Organizing a conference of this scope and level could only be achieved by the collaborative effort of a highly capable team taking charge of such matters as conference registration system, finances, the venue, social events, catering, handling all sorts of individual requests from the authors, preparing the conference rooms, etc.

We hope you had an inspiring conference and an unforgettable stay in the beautiful city of Prague. We hope to meet you again for FedCSIS 2018 in Poznań, Poland.

Co-Chairs of the FedCSIS Conference Series

Maria Ganzha, Warsaw University of Technology, Poland and Systems Research Institute Polish Academy of Sciences, Warsaw, Poland

Leszek Maciaszek, Wrocław University of Economics, Wrocław, Poland and Macquarie University, Sydney, Australia

Marcin Paprzycki, Systems Research Institute Polish Academy of Sciences, Warsaw Poland and Management Academy, Warsaw, Poland



Annals of Computer Science and Information Systems
Volume 11

Proceedings of the Federated
Conference on Computer Science and
Information Systems

September 3–6, 2017. Prague, Czech Republic

TABLE OF CONTENTS

CONFERENCE KEYNOTE PAPERS

- Formal Definition of a General Ontology Pattern Language using a Graph Grammar** 1
Eduardo Zambon, Giancarlo Guizzardi
- Application of mean-variance mapping optimization for parameter identification in real-time digital simulation** 11
Abdulrasaq Gbadamosi, José L. Rueda, Da Wang, Peter Palensky
-

12TH INTERNATIONAL SYMPOSIUM ADVANCES IN ARTIFICIAL INTELLIGENCE AND APPLICATIONS

- Call For Papers** 17
- Top k Recommendations using Contextual Conditional Preferences Model** 19
Aleksandra Karpus, Tommaso di Noia, Krzysztof Goczyła
- On the Use of Nature Inspired Metaheuristic in Computer Game** 29
Piotr Andrzej Kowalski, Szymon Łukasik, Małgorzata Charytanowicz, Piotr Kulczycki
- Determining the significance of features with the use of Sobol' method in probabilistic neural network classification tasks** 39
Piotr Andrzej Kowalski, Maciej Kuśy
- Detection and Dimension of Moving Objects Using Single Camera Applied to the Round Timber Measurement** 49
Artem Kruglov, Yuriy Chiryshv, Anastasia Atamanova, Svetlana Zavada
- Bindier Operators in Type-Theory of Algorithms for Algorithmic Binding of Functional Neuro-Receptors** 57
Roussanka Loukanova
- Towards Real-time Motion Estimation in High-Definition Video Based on Points of Interest** 67
Petr Pulc, Martin Holeňa
- Data Clustering with Grasshopper Optimization Algorithm** 71
Szymon Łukasik, Piotr Andrzej Kowalski, Małgorzata Charytanowicz, Piotr Kulczycki
- Co-Evolutionary Algorithm solving Multi-Skill Resource-Constrained Project Scheduling Problem** 75
Paweł B. Myszkowski, Maciej Laszczyk, Dawid Kalinowski
- Efficient selection operators in NSGA-II for Solving Bi-Objective Multi-Skill Resource-Constrained Project Scheduling Problem** 83
Paweł B. Myszkowski, Maciej Laszczyk, Joanna Lichodij

Hybrid Multievolutionary System to Solve Function Optimization Problems	87
<i>Krzysztof Pytel</i>	
Utilizing Multimedia Ontologies in Video Scene Interpretation via Information Fusion and Automated Reasoning	91
<i>Leslie F. Sikos</i>	
Using Classification for Cost Reduction of Applying Mutation Testing	99
<i>Joanna Strug, Barbara Strug</i>	
Evolving KERAS Architectures for Sensor Data Analysis	109
<i>Petra Vidnerová, Roman Neruda</i>	
Measurement of the appropriateness in career selection of the high school students by using data mining algorithms: A case study	113
<i>Ahmet Firat Yelkuvan, Hidayet Takci, Kali Gurkahraman</i>	
<hr/>	
AAIA DATA MINING CHALLENGE	
<hr/>	
Call For Papers	119
Helping AI to Play Hearthstone: AAIA'17 Data Mining Challenge	121
<i>Andrzej Janusz, Tomasz Tajmajer, Maciej Świechowski</i>	
Predicting Unpredictable Building Models Handling Non-IID Data Hearthstone Case Study	127
<i>Dominik Deja</i>	
Helping AI to Play Hearthstone using Neural Networks	131
<i>Lukasz Grad</i>	
Evaluation of Hearthstone Game States With Neural Networks and Sparse Autoencoding	135
<i>Jan Jakubik</i>	
Multi-model approach for predicting the value function in the game of Heathstone: Heroes of Warcraft.	139
<i>Alexander Morgun</i>	
Use of Domain Knowledge and Feature Engineering in Helping AI to Play Hearthstone	143
<i>Przemysław Przybyszewski, Szymon Dziewiątkowski, Sebastian Jaszczur, Mateusz Śmiech, Marcin Szczuka</i>	
An ensemble model with hierarchical decomposition and aggregation for highly scalable and robust classification	149
<i>Quang Hieu Vu, Dymitr Ruta, Ling Cen</i>	
<hr/>	
7TH INTERNATIONAL WORKSHOP ON ARTIFICIAL INTELLIGENCE IN MEDICAL APPLICATIONS	
<hr/>	
Call For Papers	153
Predictive and Descriptive Analysis for Heart Disease Diagnosis	155
<i>František Babič, Jaroslav Olejár, Zuzana Vantová, Ján Paralič</i>	
Subvocal Speech Recognition via Close-Talk Microphone and Surface Electromyogram Using Deep Learning	165
<i>Mohamed S. Elmahdy, Ahmed Morsy</i>	
Developing an SVM classifier for extended ES protein structure prediction	169
<i>Piotr Fabian, Katarzyna Stapor</i>	
Towards a Keyword Extraction in Medical and Healthcare Education	173
<i>Martin Komenda, Matěj Karolyi, Roman Vyškovský, Kateřina Ježová, Jakub Ščavnický</i>	
Medical biophysics as a combination of the classic educational method and e-learning	177
<i>David Kordek, Martin Kopeček, Petr Voda</i>	
A Method For Data Classification In Slovak Medical Records	181
<i>Erik Kučera, Oto Haffner, Erich Stark</i>	
Integration of Virtual Patients in Education of Veterinary Medicine	185
<i>Jaroslav Majernik, Marián Mad'ar, Jana Mojžišová</i>	

Semi-real-time analyses of item characteristics for medical school admission tests	189
<i>Patricia Martinková, Lubomír Štěpánek, Adéla Drabinová, Jakub Houdek, Martin Vejražka, Čestmír Štuka</i>	
Moodle Portal in Virtualized Environment - a Performance Analysis	195
<i>Vladimír Mašín, Martin Kopeček, Josef Hanuš</i>	
Feature Selection Methods Applied to Severe Brain Damages Data	199
<i>Wiesław Paja, Krzysztof Panczerz</i>	
Preprocessing compensation techniques for improved classification of imbalanced medical datasets	203
<i>Agnieszka Wosiak, Sylwia Karbowskiak</i>	
Neuro-Endo-Trainer-Online Assessment System (NET-OAS) for Neuro-Endoscopic Skills Training	213
<i>Vinkle Kumar Srivastav, Britty Baby, Ramandeep Singh, Prem Kalra, Ashish Suri</i>	
<hr/>	
2ND INTERNATIONAL WORKSHOP ON AI ASPECTS OF REASONING, INFORMATION, AND MEMORY	
<hr/>	
Call For Papers	221
Formalization of Pell's Equations in the Mizar System	223
<i>Marcin Acewicz, Karol Pał</i>	
Progress in the Independent Certification of Mizar Mathematical Library in Isabelle	227
<i>Cezary Kaliszzyk, Karol Pał</i>	
Formalization of the Algebra of Nominative Data in Mizar	237
<i>Artur Kornilowicz, Andrii Kryvolap, Mykola Nikitchenko, Ievgen Ivanov</i>	
Introducing Euclidean Relations to Mizar	245
<i>Adam Naumowicz, Artur Kornilowicz</i>	
Is there a computable upper bound on the heights of rational solutions of a Diophantine equation with a finite number of solutions?	249
<i>Krzysztof Molenda, Agnieszka Peszek, Maciej Sporysz, Apoloniusz Tyszk</i>	
Modeling value-based reasoning for autonomous agents	259
<i>Tomasz Zurek, Michail Mokkas</i>	
<hr/>	
7TH INTERNATIONAL WORKSHOP ON ADVANCES IN SEMANTIC INFORMATION RETRIEVAL	
<hr/>	
Call For Papers	263
PitchKeywordExtractor: Prosody-based Automatic Keyword Extraction for Speech Content	265
<i>Yurij Lezhenin, Artyom Zhuikov, Natalia Bogach, Elena Boitsova, Evgeny Pyshkin</i>	
Research on Proposals and Trends in the Architectures of Semantic Search Engines: A Systematic Literature Review	271
<i>Jorge Morales, Andrés Melgar</i>	
Just Walk: Rethinking Use Cases in Mobile Audio Travel Guides	281
<i>Evgeny Pyshkin, Pavel Korobenin</i>	
<hr/>	
11TH JOINT AGENT-ORIENTED WORKSHOPS IN SYNERGY	
<hr/>	
Call For Papers	289
Preface	291
Failure Analysis for Adaptive Autonomous Agents using Petri Nets	293
<i>Mirgita Frasher, Lan Anh Trinh, Baran Cürüklü, Mikael Ekström</i>	
Development of Simulations for Ambient Assisted Living through Pattern Repositories	299
<i>Rubén Fuentes-Fernández, Jorge Gomez Sanz</i>	

Electricity peak demand classification with artificial neural networks	307
<i>Krzysztof Gajowniczek, Rafik Nafkha, Tomasz Zabkowski</i>	
Towards an Agent-based Simulation of Building Stock Development for the City of Hamburg	317
<i>Thomas Preisler, Tim Dethlefs, Wolfgang Renz, Ivan Dochev, Hannes Seller</i>	

2ND INTERNATIONAL WORKSHOP ON LANGUAGE TECHNOLOGIES AND APPLICATIONS

Call For Papers	327
Document Clustering using a Graph Covering with Pseudostable Sets	329
<i>Jens Dörpinghaus, Sebastian Schaaf, Julianè Fluck, Marc Jacobs</i>	
Event Relation Acquisition Using Dependency Patterns and Confidence-Weighted Co-occurrence Statistics	339
<i>Shohei Higashiyama, Kunihiko Sadamasa, Takashi Onishi, Yotaro Watanabe</i>	
A Comparison of Authorship Attribution Approaches Applied on the Lithuanian Language	347
<i>Jurgita Kapočiūtė-Dzikiėnė, Algimantas Venčkauskas, Robertas Damaševičius</i>	
Extraction of specific data from a sound sample by removing additional distortion	353
<i>Dawid Połap</i>	
Deep Learning methods for Subject Text Classification of Articles	357
<i>Piotr Semberecki, Henryk Maciejewski</i>	
A Hierarchical Approach for Sentiment Analysis and Categorization of Turkish Written Customer Relationship Management Data	361
<i>Mehmet Seyfioğlu, Mustafa Demirezen</i>	
Personality Prediction Based on Twitter Information in Bahasa Indonesia	367
<i>Derwin Suhartono, Veronica Ong, Anneke D. S. Rahmanto, Williem, Aryo E. Nugroho, Esther W. Andangsari, Muhamad N. Suprayogi</i>	
Open Class Authorship Attribution of Lithuanian Internet Comments using One-Class Classifier	373
<i>Algimantas Venčkauskas, Arnas Karpavičius, Robertas Damaševičius, Romas Marcinkevičius, Jurgita Kapočiūtė-Dzikiėnė, Christian Napoli</i>	
Unsupervised tool for quantification of progress in L2 English phraseological	383
<i>Krzysztof Wołk, Agnieszka Wołk, Krzysztof Marasek</i>	
Big Data Language Model of Contemporary Polish	389
<i>Krzysztof Wołk, Agnieszka Wołk, Krzysztof Marasek</i>	

10TH INTERNATIONAL WORKSHOP ON COMPUTATIONAL OPTIMIZATION

Call For Papers	397
Optimising SVM to classify imbalanced data using dispersive flies optimisation	399
<i>Haya Alhakbani, Mohammad Majid al-Rifaie</i>	
Correlation clustering: a parallel approach?	403
<i>László Aszalós, Mária Bakó</i>	
An Integer Programming based Ant Colony Optimisation Method for Nurse Rostering	407
<i>Joe Bunton, Andreas Ernst, Mohan Krishnamoorthy</i>	
Ant Colony Optimization Algorithm for Workforce Planning	415
<i>Stefka Fidanova, Gabriel Luque, Olympia Roeva, Marcin Paprzycki, Paweł Gepner</i>	
Comparison of two types of Quantum Oracles based on Grover's Adaptive Search Algorithm for Multiobjective Optimization Problems	421
<i>Gerardo G. Fogel, Benjamín Barán, Marcos Villagra</i>	

Using branching-property preserving Pruefer Code to encode solutions for Particle Swarm Optimization	429
<i>Hanno Hildmann, Dymitr Ruta, Dina Y. Atia, A. F. Isakovic</i>	
Anchored Alignment Distance between Rooted Labeled Unordered Trees	433
<i>Kouichi Hirata, Takuya Yoshino, Yuma Ishizaka</i>	
A Distance-Based Approach for Human Posture Simulations	441
<i>Antonio Mucherino, Douglas S. Gonçalves, Antonin Bernardin, Ludovic Hoyet, Franck Multon</i>	
Solving 0-1 Quadratic Problems with Two-Level Parallelization of the BiqCrunch Solver	445
<i>Camille Coti Etienne Leclercq, Frédéric Roupin, Franck Butelle</i>	
A Fully Fuzzy Linear Programming Model to the Berth Allocation Problem	453
<i>Flabio Gutiérrez Segura, Edwar Luján Segura, Rafael Asmat Uceda, Edmundo Vergara Moreno</i>	
An Electronic Market Model with Mathematical Formulation and Heuristics for Large-Scale Book Trading	459
<i>Ali Haydar Özer</i>	
Distance-2 Collision-Free Broadcast Scheduling in Wireless Networks	469
<i>Valentin Pollet, Vincent Boudet, Jean-Claude König</i>	
Vehicle Oriented Algorithms for the Relocation of Vehicle Sharing Systems	473
<i>Alain Quilliot, Antoine Sarbinowski</i>	
Catching clouds: Simultaneous optimization of the parameters of biological agent plumes using Dirichlet processes to best estimate infection source location	481
<i>James Thompson, Thomas Finnie, Ian Hall, Nina Dobrinkova</i>	
<hr/>	
COMPUTER SCIENCE & SYSTEMS	
Call For Papers	485
<hr/>	
10TH WORKSHOP ON COMPUTER ASPECTS OF NUMERICAL ALGORITHMS	
Call For Papers	487
OpenMP Thread Affinity for Matrix Factorization on Multicore Systems	489
<i>Beata Bylina, Jarosław Bylina</i>	
A Framework for Generating and Evaluating Parallelized Code	493
<i>Jarosław Bylina</i>	
Block Subspace Projection Preconditioned Conjugate Gradient Method for Structural Modal Analysis	497
<i>Sergiy Fialko, Viktor Karpilovskyi</i>	
An algorithm for Gaussian Recursive Filters in a Multicore Architecture	507
<i>Ardelio Galletti, Giulio Giunta, Livia Marcellino, Diego Parlato</i>	
On Memory Footprints of Partitioned Sparse Matrices	513
<i>Daniel Langr, Ivan Šimeček</i>	
Optimizing Numerical Code by means of the Transitive Closure of Dependence Graphs	523
<i>Marek Palkowski, Włodzimierz Bielecki</i>	
A Non-Speculative Parallelization of Reverse Cuthill-McKee Algorithm for Sparse Matrices Reordering	527
<i>Thiago Nascimento Rodrigues, Maria Claudia Silva Boeres, Lucia Catabriga</i>	
Least Square Method Robustness of Computations: What is not usually considered and taught	537
<i>Vaclav Skala</i>	

4TH INTERNATIONAL CONFERENCE ON CRYPTOGRAPHY AND SECURITY SYSTEMS

Call For Papers	543
Enhancing the Imperceptibility of Image Steganography for Information Hiding <i>Mohamed Fouad Abdelmotagally</i>	545
The impact of malware evolution on the analysis methods and infrastructure <i>Krzysztof Cabaj, Piotr Gawkowski, Konrad Grochowski, Alexis Nowikowski, Piotr Żórawski</i>	549
Quantum color image encryption based on multiple discrete chaotic systems <i>Li Li, Bassem Abd-El-Atty, Ahmed Abd El-Latif, Ahmed Ghoneim</i>	555
TARZAN: An Integrated Platform for Security Analysis <i>Marek Rychlý, Ondrej Ryšavý</i>	561
High-Level Malware Behavioural Patterns: Extractability Evaluation <i>Jana Šířastná, Martin Tomášek</i>	569

2ND WORKSHOP ON CONSTRAINT PROGRAMMING AND OPERATION RESEARCH APPLICATIONS

Call For Papers	573
Application of survival function in robust scheduling of production jobs <i>Lukasz Sobaszek, Arkadiusz Gola, Edward Kozłowski</i>	575
A hybrid method for Optimization Scheduling Groups of Jobs <i>Tadeusz Stefański, Jarosław Wikarek</i>	579
Answer Set Programming for Modeling and Reasoning on Modular and Reconfigurable Transportation Systems <i>Walter Terkaj, Marcello Urgo, Daniela Andolfatto</i>	587

10TH INTERNATIONAL SYMPOSIUM ON MULTIMEDIA APPLICATIONS AND PROCESSING

Call For Papers	597
Preface	599
Available Bandwidth Estimation in Smart VPN Bonding Technique based on a NARX Neural Network <i>Giacomo Capizzi, Grazia Lo Sciuto, Francesco Beritelli, Francesco Scaglione, Dawid Potap, Kamil Książek, Marcin Wozniak</i>	601
H.265 Inverse Transform FPGA implementation in Impulse C <i>Stawomir Cichoń, Marek Gorgoń</i>	607
New Content Based-Image Retrieval database structure using Query by Approximate Shapes <i>Stanisław Deniziak, Tomasz Michno</i>	613
Seniors' experiences with online banking <i>Chrysoula Gatsou, Anastasios Politis, Dimitrios Zevgolis</i>	623
Corneal Endothelium Image Segmentation Using Feedforward Neural Network <i>Anna Fabijańska</i>	629
Automatized Generation of Alphabets of Symbols <i>Serhii Hamotskyi, Anis Rojbi, Sergii Stirenko, Yuri Gordienko</i>	639
Soccer Event Recognition Technique based on Pattern Matching <i>Jiwon Lee, Do-won Nam, Sungwon Moon, JungSoo Lee, Wonyoung Yoo</i>	643
Design of Audio Digital Watermarking System Resistant to Removal Attack <i>Guillermo Morales-Luna, Valery Korzhik, Vasily Alekseev</i>	647

GPU Accelerated 2D and 3D Image Processing	653
<i>Anca Morar, Florica Moldoveanu, Alin Moldoveanu, Oana Balan, Victor Asavei</i>	
Optical Driving for a Computer System with Augmented Reality Features	657
<i>Tomasz Pajys, Krzysztof Murawski, Artur Arciuch, Andrzej Walczak</i>	
The Next Generation of In-home Streaming: Light Fields, 5K, 10 GbE, and Foveated Compression	663
<i>Daniel Pohl, Daniel Jungmann, Bartosz Taudul, Richard Membarth, Harini Hariharan, Thorsten Herfet, Oliver Grau</i>	
Ground plane detection in 3D scenes for an arbitrary camera roll rotation through “V-disparity” representation	669
<i>Piotr Skulimowski, Mateusz Owczarek, Paweł Strumiłło</i>	
The membrane shape mapping of the artificial ventricle in the actual dimensions	675
<i>Wojciech Sulej, Krzysztof Murawski</i>	
Selective Image Authentication Using Shearlet Coefficients Tolerant to JPEG Compression	681
<i>Aleksei Zhuvikin</i>	

6TH WORKSHOP ON ADVANCES IN PROGRAMMING LANGUAGES

Call For Papers	689
Preface	691
Use Case Driven Modularization as a Basis for Test Driven Modularization	693
<i>Michal Bystrický, Valentino Vranić</i>	
Towards Programmable Address Spaces	697
<i>Andrew Gozillon, Paul Keir</i>	
Program analysis for Clustering Programmers’ Profile	701
<i>Daniel José Ferreira Novais, Maria João Varanda Pereira, Pedro Rangel Henriques</i>	
An Approach for Modeling Events in Information Systems	707
<i>Aleksandar Popović, Ivan Luković, Vladimir Dimitrieski, Verislav Đukić</i>	
Properties and Limits of Supercombinator Set Acquired from Context-free Grammar Samples	711
<i>Michal Sičák, Ján Kollár</i>	
Labeling Source Code with Metadata: A Survey and Taxonomy	721
<i>Matúš Sulír, Jaroslav Porubán</i>	

9TH WORKSHOP ON SCALABLE COMPUTING

Call For Papers	731
Techno-economic framework for cloud infrastructure: a cost study of resource disaggregation	733
<i>Mozhgan Mahloo, João Monteiro Soares, Amir Roozbeh</i>	
A Database Performance Polynomial Multiple Regression Model	743
<i>Artur Nowosielski, Piotr Andrzej Kowalski, Piotr Kulczycki</i>	
CloudLightning: a Self-Organized Self-Managed Heterogeneous Cloud	749
<i>Huanhuan Xiong, Dapeng Dong, Christos Filelis-Papadopoulos, Gabriel G. Castañé, Theo Lynn, Dan C. Marinescu, John Morrison</i>	

INTERNATIONAL CONFERENCE ON INNOVATIVE NETWORK SYSTEMS AND APPLICATIONS

Call For Papers	759
------------------------	------------

1ST INTERNATIONAL CONFERENCE ON SECURITY, PRIVACY, AND TRUST

Call For Papers	761
Representation of Attacker Motivation in Software Risk Assessment Using Attack Probability Trees	763
<i>Marko Esche, Federico Grasso Toro, Florian Thiel</i>	
Multimodal Artifact Metrics for Valuable Resin Card	773
<i>Masaki Fujikawa, Kouki Jitsukawa, Shingo Fuchi</i>	
On end-to-end approach for slice isolation in 5G networks. Fundamental challenges	783
<i>Zbigniew Kotulski, Tomasz Nowak, Mariusz Sepczuk, Marcin Tunia, Rafał Artych, Krzysztof Bocianiak, Tomasz Ośko, Jean-Philippe Wary</i>	
Key Exchange Algorithm Based on Homomorphic Encryption	793
<i>Ilya Kuzmin, Sergey Krendelev</i>	
Black Hole Attack Prevention Method Using Dynamic Threshold in Mobile Ad Hoc Networks	797
<i>Taku Noguchi, Takaya Yamamoto</i>	
Dependable Design for Elderly Health Care	803
<i>Kasi Periyasamy, Vangalur Alagar, Kaiyu Wan</i>	
Analysis of DDoS-Capable IoT Malwares	807
<i>Angelo Spognardi, Michele De Donno, Nicola Dragoni, Alberto Giaretta</i>	

1ST WORKSHOP ON INTERNET OF THINGS - ENABLERS, CHALLENGES AND APPLICATIONS

Call For Papers	817
An Unsupervised Evidential Conflict Resolution Method for Data Fusion In IoT	819
<i>Walid Cherifi, Bolesław Szafranski</i>	
An Incremental Evidential Conflict Resolution Method for Data stream Fusion In IoT	825
<i>Walid Cherifi, Bolesław Szafranski</i>	
Combat triage support using the Internet of Military Things	835
<i>Michał Dyk, Mariusz Chmielewski, Andrzej Najgebauer</i>	
An approach to prevention to the DNS Injection attacks on the base of system level comparison method MM	843
<i>Michał Melaniuk</i>	

6TH INTERNATIONAL CONFERENCE ON WIRELESS SENSOR NETWORKS

Call For Papers	847
Analysis of Interferences in Data Transmission for Wireless Communications Implemented in Vehicular Environments	849
<i>Valentin Iordache, Razvan Andrei Gheorghiu, Marius Minea</i>	
Considerations for using ZigBee technology in vehicular non-critical applications	853
<i>Valentin Iordache, Marius Minea, Razvan Andrei Gheorghiu</i>	
Impact of External Phenomena In Compressed Sensing Methods For Wireless Sensor Networks	857
<i>Michal Kochlář, Michal Hodoň</i>	
Adaptation of MANET topology to monitor dynamic phenomena clouds	865
<i>Mateusz Krzysztoń, Ewa Niewiadomska-Szynkiewicz</i>	
Internet connected wireless combustible gas monitoring system for apartment buildings	873
<i>Denis Spirjakin, Alexander Baranov</i>	
Fall Detection using Lifting Wavelet Transform and Support Vector Machine	877
<i>Wipawee Usaha, Hanghan Liang</i>	

Optimizing RTS/CTS to Improve Throughput in Ad Hoc WLANs	885
<i>Emilia Weyulu, Masaki Hanada, Moo Wan Kim</i>	
Modelling and identification of linear discrete systems using least squares method	891
<i>Peter Šarafin, Martin Hudik, Martin Revák, Peter Ševčík, Samuel Žák</i>	
Load balancing of heterogeneous parallel DC-DC converter	895
<i>Samuel Žák, Jaroslav Szabo</i>	

INFORMATION TECHNOLOGY FOR MANAGEMENT, BUSINESS & SOCIETY

Call For Papers	901
------------------------	------------

15TH CONFERENCE ON ADVANCED INFORMATION TECHNOLOGIES FOR MANAGEMENT

Call For Papers	903
Deep Learning for Financial Time Series Forecasting in A-Trader System	905
<i>Jerzy Korczak, Marcin Hernes</i>	
Implementing ERP Systems in Higher Education Institutes: Critical Success Factors Revisited	913
<i>Christian Leyh, Anne Gebhardt, Philipp Berton</i>	
Project Management and Communication Software Selection Using the Weighted Regularized Hasse Method	919
<i>Karolina Muszyńska, Jakub Swacha</i>	
Privacy Preserving BPMS for Collaborative BPaaS	925
<i>Björn Schwarzbach, Michael Glöckner, Sergei Makarov, Bogdan Franczyk, André Ludwig</i>	
Analysis of functions offered by the e-government systems from the perspective of chosen group of users in Poland	935
<i>Oskar Szumski, Witold Chmielarz</i>	
Survey as a source of low quality research data	939
<i>Grzegorz Szyjewski, Luiza Fabisiak</i>	
Data Mining for Customers' Positive Reaction to Advertising in Social Media	945
<i>Veera Boonjing, Daranee Pimchangthong</i>	
Sustainable Decision-Making using the COMET Method: An Empirical Study of the Ammonium Nitrate Transport Management	949
<i>Jarosław Wątróbski, Wojciech Szałabun, Artur Karczmarczyk, Waldemar Wolski</i>	

12TH CONFERENCE ON INFORMATION SYSTEMS MANAGEMENT

Call For Papers	959
IT Governance Program and Improvements in Brazilian Small Business: Viability and Case Study	961
<i>Daniel A. M. Aguillar, Isabel Murakami, Pedro Manso Junior, Plinio Thomaz Aquino Jr.</i>	
Analysis of the Use of Electronic Banking and e-Payments from the Point of View of a Client	965
<i>Witold Chmielarz, Marek Zborowski</i>	
Conceptualization of an Abstract Language to Support Value Co-Creation	971
<i>Christophe Feltus, Erik HA Proper</i>	
Towards Process-Oriented Ontology for Financial Analysis	981
<i>Jerzy Korczak, Helena Dudycz, Bartłomiej Nita, Piotr Oleksyk</i>	
Industry 4.0 and Lean Production – A Matching Relationship? An analysis of selected Industry 4.0 models	989
<i>Christian Leyh, Stefan Martin, Thomas Schäffer</i>	

Adaptation of orchestration graphs in gamification	995
<i>Tomasz Lipczyński, Magdalena Kieruzel, Przemysław Różewski</i>	
Domain-Specific Characteristics of Data Quality	999
<i>Ivo Oditis, Janis Bicevskis, Zane Bicevska</i>	
Process-oriented approach to competency management using ontologies	1005
<i>Ilona Pawełoszek</i>	
Re-Engineering Enterprise Architectures	1015
<i>Murat Paşa Uysal, Ali Halici, A. Erhan Mergen</i>	
Integrated Approach to e-Commerce Websites Evaluation with the Use of Surveys and Eye Tracking Based Experiments	1019
<i>Paweł Ziemba, Jarosław Wątróbski, Artur Karczmarczyk, Jarosław Jankowski, Waldemar Wolski</i>	
The ICT adoption in enterprises in the context of the sustainable information society	1031
<i>Ewa Ziemba</i>	
Example of designing a business process oriented autopoietic knowledge management support system.	1039
<i>Mariusz Żytniewski</i>	
<hr/>	
5TH WORKSHOP ON INFORMATION TECHNOLOGIES FOR LOGISTICS	
<hr/>	
Call For Papers	1049
PhyNetLab: An IoT-Based Warehouse Testbed	1051
<i>Robert Falkenberg, Mojtaba Masoudinejad, Markus Buschhoff, Aswin Karthik Ramachandran Venkatapathy, Daniel Friesel, Michael ten Hompel, Olaf Spinczyk, Christian Wietfeld</i>	
Modeling and Optimization of Multi-echelon Transportation systems - a hybrid approach	1057
<i>Tadeusz Stefański, Paweł Sitek</i>	
Human Machine Synergies in Intra-Logistics: Creating a Hybrid Network for Research and Technologies	1065
<i>Aswin Karthik Ramachandran Venkatapathy, Haci Bayhan, Felix Zeidler, Michael ten Hompel</i>	
Decision Support System for Robust Urban Transport Management	1069
<i>Piotr Wiśniewski, Krzysztof Kluza, Antoni Ligęza</i>	
<hr/>	
23RD CONFERENCE ON KNOWLEDGE ACQUISITION AND MANAGEMENT	
<hr/>	
Call For Papers	1075
Context-aware and pro-active queue management systems in intelligent environments	1077
<i>Radosław Klimek</i>	
DNS as Resolution Infrastructure for Persistent Identifiers	1085
<i>Fatih Berber, Ramin Yahyapour</i>	
Comparison of Selected Modeling Notations for Process, Decision and System Modeling	1095
<i>Krzysztof Kluza, Piotr Wiśniewski, Krystian Jobczyk, Antoni Ligęza, Anna Suchenia (Mroczek)</i>	
The tools and methods of capturing knowledge from customers: empirical investigation	1099
<i>Dmitry Kudryavtsev, Nikita Plyasunov, Liudmila Kokoulina</i>	
Categorizing or Generating Relation Types and Organizing Ontology Design Patterns	1109
<i>Philippe Martin, Jérémy Bénard</i>	
Analysis of Dialogue Stimulated by Science Videos and Reference Materials	1119
<i>Kiyoshi Nosu</i>	
Towards better understanding of context-aware knowledge transformation	1123
<i>Mieczysław Owoc, Paweł Weichbroth, Karol Żuralski</i>	

Simulation Driven Development – Validation of requirements in the early design stages of complex systems – the example of the German Toll System 1127
Bernd Pfitzinger, Tommy Baumann, Thomas Jestädt

A view on the methodology of analysis and exploration of marketing data 1135
Maciej Pondel, Jerzy Korczak

SOFTWARE SYSTEMS DEVELOPMENT & APPLICATIONS

Call For Papers 1145

4TH INTERNATIONAL WORKSHOP ON CYBER-PHYSICAL SYSTEMS

Call For Papers 1147

Prediction of Traffic Intensity for Dynamic Street Lighting 1149
Marzena Bielecka, Andrzej Bielecki, Sebastian Ernst, Igor Wojnicki

An efficient real-time architecture for collecting IoT data 1157
Vincenza Carchiolo, Michele Malgeri, Mark Philip Loria, Marco Toja

Implementation of a Simplified State Estimator for Wind Turbine Monitoring on an Embedded System 1167
Theis Bo Rasmussen, Guangya Yang, Arne Hejde Nielsen, Zhao Yang Dong

Visual simulator for MavLink-protocol-based UAV, applied for search and analyze task 1177
Piotr Śmigielski, Mateusz Raczyński, Łukasz Gosek

1ST INTERNATIONAL CONFERENCE ON LEAN AND AGILE SOFTWARE DEVELOPMENT

Call For Papers 1187

Selecting Requirements Documentation Techniques for Software Projects: a Survey Study 1189
Aleksander Jarzębowski, Katarzyna Połocka

Process Mining Methods for Post-Delivery Validation 1199
Paweł Markowski, Michał Przybytek

Application of a process improvement method for improving usability 1203
Stanisław Plebanek

Measuring dimensions of Software Engineering projects' success in Academic context 1207
Aneta Poniszewska-Maranda, Rafał Włodarski

Making agile retrospectives more awesome 1211
Adam Przybytek, Dagmara Kotecka

The lemniscate knowledge flow model 1217
Paweł Weichbroth, Kamil Brodnicki

Enterprise Architecture Approach to SCRUM Processes, Sprint Retrospective Example 1221
Jan Werewka, Anna Spiechowicz

4TH CONFERENCE ON MULTIMEDIA, INTERACTION, DESIGN AND INNOVATION

Call For Papers 1229

Emerging Trends and Novel Approaches in Interaction Design 1231
Krzysztof Marasek, Andrzej Romanowski, Marcin Sikorski

Comparative analysis of multitouch interactive surfaces 1235
Przemysław Kucharski, Dawid Sielski, Krzysztof Grudzień, Wiktor Kozakiewicz, Michał Basiuras, Klaudia Greif, Jakub Santorek, Laurent Babout

Interacting with Digital Memorials in a Cemetery: Insights from an Immersive Practice 1239
Cristiano Maciel, Vinicius Carvalho Pereira, Carla Leitão, Roberto Pereira, José Viterbo

Aesthetic Categories of Interaction: Aesthetic Perceptions on Smartphone and Computer	1249
<i>Mati Mõitus, David Lamas, Liina Kukk</i>	
User-Centered Design Case Study: Ribbon Interface Development for Point of Sale Software	1257
<i>Zdzisław Sroczyński</i>	
Implementation and verification of speech database for unit selection speech synthesis	1263
<i>Krzysztof Szklanny, Sebastian Koszuta</i>	
Creating an Interactive and Storytelling Educational Physics App	1269
<i>Krzysztof Szklanny, Łukasz Homoncik, Marcin Wichrowski, Alicja Wieczorkowska</i>	
What Looks Good with my Sofa: Ensemble Multimodal Search for Interior Design	1275
<i>Ivona Tautkute, Aleksandra Możejko, Wojciech Stokowiec, Tomasz Trzciniński, Łukasz Brocki, Krzysztof Marasek</i>	
Mobile devices' GPUs in cloth dynamics simulation	1283
<i>Marcin Wawrzonowski, Dominik Szajerman, Marcin Daszuta, Piotr Napieralski</i>	
Robust face model based approach to head pose estimation	1291
<i>Adam Wojciechowski, Krzysztof Fornalczyk</i>	
Design and Implementation of Fire Safety Education System on Campus based on Virtual Reality Technology	1297
<i>Kun Zhang, Jintao Suo, Jingying Chen, Xiaodi Liu, Lei Gao</i>	
<hr/>	
THE 37TH IEEE SOFTWARE ENGINEERING WORKSHOP	
<hr/>	
Call For Papers	1301
Fundamentals of a Components Sharing Network to Accelerate JavaScript Software Development	1303
<i>Daniel Souza Makiyama, Plinio Thomaz Aquino Jr.</i>	
Aspect-driven Context-aware Services	1307
<i>Karel Cemus, Filip Klimes, Tomas Cerny</i>	
Evaluation of Mutant Sampling Criteria in Object-Oriented Mutation Testing	1315
<i>Anna Derezinska, Marcin Rudnik</i>	
Documentation Management Environment for Software Product Lines	1325
<i>Stanislaw Jarzabek, Daniel Dan</i>	
Interface-based Semi-automated Testing of Software Components	1335
<i>Tomas Potuzak, Richard Lipka, Premek Brada</i>	
<hr/>	
4TH DOCTORAL SYMPOSIUM ON RECENT ADVANCES IN INFORMATION TECHNOLOGY	
<hr/>	
Call For Papers	1345
A general optimization-based approach for thermal processes modeling	1347
<i>Paweł Drąg, Krystyn Styczeń</i>	
The North Sea Bicycle Race ECG Project: Time-Domain Analysis	1353
<i>Dominika Długosz, Trygve Eftestøl, Stein Ørn, Tomasz Wiktorski, Aleksandra Królak</i>	
A case study on machine learning model for code review expert system in software engineering	1357
<i>Michał Madera, Rafał Tomoń</i>	
The Realisation of Neural Network Structural Optimization Algorithm	1365
<i>Grzegorz Nowakowski, Yaroslav Dorogy, Olena Doroga-Ivaniuk</i>	
Author Index	1373

Data Mining for Customers' Positive Reaction to Advertising in Social Media

Veera Boonjing
International College,
King Mongkut's Institute of Technology Ladkrabang,
Ladkrabang, Bangkok, 10520,
Thailand
Email: veera.bo@kmitl.ac.th

Darane Pimchangthong
Faculty of Business Administration,
Rajamangala University of
Technology Thanyaburi, Klong 6,
Thanyaburi, Pathum Thani, 12110,
Thailand
Email: daranee_p@rmutt.ac.th

Abstract— The paper aims at 1) finding the most important factors influencing positive customer reactions and purchasing merchandises after seeing online social media advertising and 2) identifying characteristics of customer clusters having positive reaction, as well as of purchasing customer clusters, after seeing online social media advertising. Data from 370 respondents are collected by questionnaires using convenience sampling method. Attribute selection and clustering techniques are employed in data analysis to find important factors and identify customer clusters, respectively. It is found that there is a strong correlation between the reason for clicking advertisement on social media and the satisfaction with merchandise, and between purchasing merchandise online and saving information for further consideration. The findings also indicate the characteristics of “Product conscious” and “Price Conscious” clusters for customer's reaction and purchasing after seeing online social media advertising.

I. INTRODUCTION

WITH high popularity of social media, several social media websites have been developed such as Line, Facebook, Twitter, etc. and the usage rate has increased every year. Social media has become a fast and easy way to reach people in almost every group categorized by age, occupation, and education etc. with less cost. The usage of social media is widespread in private and public environments. In the highly competitive business world, social media has become a source of large amounts of data that is extremely useful when data are analyzed properly. The results from analyzing data properly can be useful in variety of disciplines including education, business, politics, social, science, technology, etc.

Marketing campaigns with online advertising are one of the methods that businesses use for increasing purchasing motivation. Finding target customer characteristics and customer reactions to social media advertising helps to reach more customers and is useful information for a marketing campaign. Data mining is one of the known techniques to analyze data to find hidden information from a large amount of available data without having prior hypotheses. Data

mining provides a variety of methods such as association, classification, clustering, etc. for analyzing data, but selecting a method to match with the objectives is a challenge.

The purposes of this study are to 1) find the most important factors influencing positive customer reactions after seeing online social media advertising, 2) find the most important factors influencing purchasing products advertised online, 3) identify customer clusters characteristics that have positive reaction after seeing online social media advertising, and 4) identify customer clusters characteristics that purchase merchandise after seeing online social media advertising.

The rest of paper is organized as follows. Section II gives a review of literature. Section III describes methodology. Results are given in Section IV. Section V concludes the paper and gives discussion.

II. LITERATURE REVIEW

Data mining has been defined as the process to extract knowledge from large quantities of data in order to discover meaningful patterns and rules [1]. Reference [2] defines data mining as the analysis of (often large) observational data sets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner. Reference [3] defines data mining as a process of knowledge discovery. Reference [4] summarized that data mining contains three key stages: finding patterns, interpreting them in order to check their usefulness, and finally using the patterns to solve business problems. The ultimate goal of data mining is to discover knowledge and it will be useful in several disciplines. In business, data mining is used for strategic benefit such as direct marketing, trend analysis, etc. In direct marketing, data mining is used for targeting people who are most likely to buy certain products and services. For trend analysis, data mining is used to determine trends in the marketplace [5].

Reference [4] explains that there are two main types of DM models as follows: 1) Predictive model: This model is constructed to predict a particular outcome or target variable. Commonly used predictive modeling techniques

This work was supported by King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand.

include multiple regression (for predicting value data), logistic regression (for response prediction) and decision trees (for rule-based value or response models). 2) Descriptive model: This model gives a better understanding of the data, without any single specific target variable. Commonly used descriptive techniques include factor analysis (to extract underlying dimensions from multivariate data), cluster analysis (for grouping a customer database into segments), and association analysis (for discovering relationships between items such as retail products).

A marketing campaign is a specifically defined series of activities used in marketing a new or changed product or service, or in using new marketing channels and methods. Marketing activities are efforts to increase awareness for a particular product or service. Social media is one of the most popular marketing channels due to the ability to reach large numbers of customers with low cost. Social media advertising helps businesses find new potential clients by using the users' own shared information to identify interest. Rather than re-actively targeting users who search for a certain term, social media advertising proactively targets relevant users before they even begin their searches.

Prior research that relates to the study topics were reviewed as follows: Reference [6] researched the factors that influence the recipients to open direct emails and make an action desired by the company and also studies whether and what elements in the email would influence them to buy the products or services promoted. The results are obtained based on a data mining analysis which includes clustering and classification processes and offer a guide on how organizations should design their email marketing communications in order to have higher response rates. Reference [7] researched on the value of social technologies in organizations based on the 'value focused thinking' approach. The findings highlight innovation of internal processes, creation of organizational identity and new business models, integrated business functions, as well as employee support to be important values of social technology enabled innovation in organizations. Reference [8] researched on the segmenting of consumer reactions to social network marketing. The purpose of the study is to understand how consumers may be segmented with respect to their reactions to social network marketing. The results identified five segments – Passive, Talkers, Hesitant, Active, and Averse – along with significant covariates such as information search, convenience, entertainment, age, and gender that predict membership.

III. METHODOLOGY

Several aspects related to the factors that influence a respondent to react and purchase a product or service after seeing social media advertising are explored from the literature review. Based on the literature review, the questionnaires are developed and distributed to people who used to purchase product or service through social media such as Facebook, Line, Instagram, etc. The questionnaires

are composed of the three parts; the first part is about the demographic data of the respondents, the second part is about the respondent's reaction when seeing social media marketing campaigns, and the third part is open-ended questions about the respondent's opinion. The data has been gathered at the level of samples including 370 respondents aged less than 60 and being employees, freelancers, entrepreneurs, managers, and students. The data has been collected between January and April 2017 through an online survey by using the convenience sampling method.

The research directions include: determining the most important factors that influence the customers to have a positive reaction after seeing social media advertising; determining the most important factors influencing purchasing products that advertise online in social media; identifying the customers characteristics that have positive reaction after seeing social media advertising; and identifying the customer clusters characteristics that purchase merchandises after seeing social media advertising. The software used to analyze data is WEKA data mining software. The research employs an attribute evaluator called "CorrelationAttributeEval" for determining the most important factors with respect to class attributes. It also uses the "SimpleKMeans" clustering algorithm for grouping customers.

IV. RESULTS

The analysis results indicate that most of the respondents are students, females, aged 21-30, single, and educational level of Bachelor's degree, and average income less than 15,000 THB. Highest percentage on the usage of social media is Facebook, followed by Line and Instagram. The most purchasing are in fashion merchandise, followed by IT and software, and beauty and health merchandise. Spending per transaction is between 500 to 1,000 THB.

In order to identify the most important factors that influence the customers to have a positive reaction after seeing social media advertising, the class attribute called ReasonForClickingAdvertisementOnSocial Media is selected. The relation between the class attribute and the others is determined using CorrelationAttributeEval. The results present that for the customers, there is a strong correlation between the reason for clicking advertisement on social media and the satisfaction with merchandise (0.12727), special rewards (0.08522), not buying (0.0852), and link for searching on more details (0.08507). Other attributes are ranked below 0.07.

The class attribute PurchasingMerchandiseOnline is selected to determine the most important factors influencing purchasing products with social media advertising. The relation between the class attribute and the others is determined using CorrelationAttributeEval. The results present that for the customers, there is a strong correlation between purchasing merchandise online and saving information for further consideration (0.13879),

merchandise logo (0.10179), and immediately purchase if satisfied (0.09747). Other attributes are ranked below 0.09.

The EM (Expectation Maximization) and "SimpleKMeans" clustering algorithms are used for grouping similar customers based on positive reaction after seeing social media advertising. The EM algorithm is used to identify the approximated cluster numbers. In this case, the result is 2 clusters. This value is used as a parameter for the "SimpleKMeans" algorithm. The algorithm results are presented in Table I.

Cluster 0: Most customers in this group save information for further consideration (74%) and are interested in high percentage of discount (64%) after seeing advertising in the social media. For this cluster, the advertising should emphasize on merchandise quality as well as price to create more positive reaction.

Cluster 1: Most customers in this group save information for further consideration (70%), are interested in high percentage of discount (64%) after seeing advertising in the social media, and are also interested in seeing text, image, and clip advertising (65%). For this cluster, clip advertising may increase the customer's interest and create more positive reaction.

Clustering customers based on purchasing merchandises after seeing social media advertising is similar to clustering based on positive reaction after seeing social media advertising. The results are presented in Table II.

Cluster 0: This group of customers can be called "Product conscious." Most customers in this group save information for further consideration (74%) after seeing advertising on the social media and purchase when satisfied with merchandise or services (55%). Seeing text, image, and clip advertising is also interesting for this group (53%). For this cluster, the advertising should emphasize on quality and brand to justify the price.

Cluster 1: This group can be called "Price conscious." Everyone in this group is concerned about price before purchasing merchandise. Interesting price (100%) and high percentage of discount (71%) will get these customers' attention. Customers in this group also like to save information for further consideration (65%) and are interested to see text, image, and clip advertising (58%). For this cluster, the opportunity to purchase at a low price should be emphasized on the advertisement online.

V. CONCLUSION AND DISCUSSION

Based on the findings from this research, e-commerce business should draw attention to the content that is emphasized in the advertisement online, due to some clusters

Table I. Customers cluster based on positive reaction after seeing social media advertising

No.	Attributes	Cluster 0 (94, 25.4%)	Cluster 1 (276, 74.6%)
1	Interested in Seeing Advertisement on Social Media: merchandise logo	0.2128	0.0725
2	Interested in Seeing Advertisement on Social Media: text, image, and clip	0.266	0.6522
3	Interested in Seeing Advertisement on Social Media: marketing campaign	0.2766	0.1739
4	Interested in Seeing Advertisement on Social Media: link for searching more details	0.2234	0.0725
5	Interested in Seeing Advertisement on Social Media: package or merchandise	0.0213	0.0254
6	Interested in Seeing Advertisement on Social Media: others	0	0.0036
7	Influence of Advertisement on Social Media: immediately purchase if satisfy	0.1596	0.1667
8	Influence of Advertisement on Social Media: saving information for further consideration	0.7417	0.6993
9	Influence of Advertisement on Social Media: sent it out for more comments	0.0426	0.0326
10	Influence of Advertisement on Social Media: irritating and don't like advertisement	0.0532	0.0725
11	Influence of Advertisement on Social Media: not buying	0	0.0254
12	Influence of Advertisement on Social Media: others	0	0.0036
13	Advertise Influence Buying: interesting price	0.3617	0.3659
14	Advertise Influence Buying: popular brand	0.1383	0.1268
15	Advertise Influence Buying: exciting new merchandise	0.0745	0.0507
16	Advertise Influence Buying: Satisfaction with merchandises or Services	0.2872	0.3732
17	Advertise Influence Buying: more refer in social media	0.1064	0.058
18	Advertise Influence Buying: like marketing campaign	0.0213	0.0217
19	Advertise Influence Buying: others	0.0106	0.0036
20	Marketing Campaign Influence Buying: high percentage of discount	0.6383	0.6449
21	Marketing Campaign Influence Buying: having complimentary	0.1277	0.1522
22	Marketing Campaign Influence Buying: special rewards	0.0319	0.0145
23	Marketing Campaign Influence Buying: having after sales service	0.2021	0.1739
24	Marketing Campaign Influence Buying: others	0	0.0145

Table II. Customers cluster based on purchasing merchandises after seeing social media advertising

No.	Attributes	Cluster 0 (235, 63.5%)	Cluster 1 (135, 36.5%)
1	Interested in Seeing Advertisement on Social Media: merchandise logo	0.0894	0.1407
2	Interested in Seeing Advertisement on Social Media: text, image, and clip advertising	0.5362	0.5852
3	Interested in Seeing Advertisement on Social Media: marketing campaign	0.2383	0.1333
4	Interested in Seeing Advertisement on Social Media: link for searching more details	0.1021	0.1259
5	Interested in Seeing Advertisement on Social Media: package or merchandise	0.0298	0.0148
6	Interested in Seeing Advertisement on Social Media: others	0.0043	0
7	Influence of Advertisement on Social Media: immediately purchase if satisfy	0.1234	0.237
8	Influence of Advertisement on Social Media: saving Information for further consideration	0.7489	0.6444
9	Influence of Advertisement on Social Media: sent it out for more comments	0.0468	0.0148
10	Influence of Advertisement on Social Media: irritating and don't like advertisement	0.0596	0.0815
11	Influence of Advertisement on Social Media: not buying	0.017	0.0222
12	Influence of Advertisement on Social Media: others	0.0043	0
13	Advertise Influence Buying: interesting price	0	1
14	Advertise Influence Buying: popular brand	0.2043	0
15	Advertise Influence Buying: exciting new merchandise	0.0894	0
16	Advertise Influence Buying: Satisfaction with merchandises or services	0.5532	0
17	Advertise Influence Buying: more refer in social media	0.1106	0
18	Advertise Influence Buying: like marketing campaign	0.034	0
19	Advertise Influence Buying: others	0.0085	0
20	Marketing Campaign Influence Buying: high percentage of discount	0.6043	0.7111
21	Marketing Campaign Influence Buying: having complimentary	0.1362	0.163
22	Marketing Campaign Influence Buying: special rewards	0.017	0.0222
23	Marketing Campaign Influence Buying: having after sales service	0.2255	0.1037
24	Marketing Campaign Influence Buying: others	0.017	0

that are product conscious while others are price conscious. For the product conscious group, price does not matter as

much for them and they decide to purchase based on satisfaction with the merchandise after thoroughly considering. Merchandise quality alone can get the product conscious group to purchase or have positive reaction after seeing online advertising in social media. For the price conscious group, they are looking for an interesting price and a high percentage of discount. Presenting interesting price will draw high attention from the price conscious group. However, both groups are interested in the advertisement online, and the advertisement that can draw customer's attention should provide interesting text, image, and clip advertisement. The advertisement should be attention-getting about the pricing so that it would convince the customers to click on the advertisement and continue to purchase the merchandise. The advertisement should also emphasize the high percentage of discount, which can draw attention from customers in both clusters that derive from positive reaction and lead to the purchasing of merchandise after seeing online advertising in social media.

The major limitation of the research consists in the low number of respondents (only 370), so this exploratory research should be followed by a conclusive one to verify the conclusions of the present research. Also, the majority of the respondents were students, aged 21-30, with low income and low spending per transaction, which may have influenced the positive reaction after seeing online advertising in social media.

Future directions of research may include: (1) using association techniques to determine rules related to positive reactions and purchasing, (2) adding some other information about respondents like time spent in front of the computer and time spent in using social network, and analyzing the influence of those factors on positive reactions and the purchasing of merchandise.

REFERENCES

- [1] M. J. A. Berry and G. Linoff, *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management* 3rd ed., John Wiley and Sons Ltd., Publication, UK, 2011.
- [2] D. J. Hand, H. Mannila and P. Smyth, *Principles of Data Mining*, MIT Press, Cambridge, MA, 2001.
- [3] J. Han and M. Kamber, *Data Mining Concepts and Techniques* Second Edition, Morgan Kaufmann Publishers, United States of America, 2006.
- [4] B. Leventhal, An introduction to data mining and other techniques for advanced analytics, *Journal of Direct, Data and Digital Marketing Practice*, 12(2), 2010, pp.137-153.
- [5] M. Negnevitsky, *Artificial Intelligence, A Guide To Intelligent Systems* 3rd ed., Pearson Education Limited, 2005.
- [6] R. I. Magos and C. A. Acatrinei, *Designing Email Marketing Campaigns - A Data Mining Approach Based on Consumer Preferences*, *A nales Universitatii Apulensis Series Oeconomica*, 17(1), 2015, 15-30.
- [7] M. Singh and K. Peszynski, "Organisational Value of Social Technologies: An Australian Study, *The Electronic Journal Information Systems Evaluation*, 17(1). 2014, pp.088-099, available online at www.ejise.com.
- [8] C. Campbell, C. Ferraro, and S. Sands, "Segmenting consumer reactions to social network marketing". *European Journal of Marketing*, 48(3/4), 2014, pp.432-452.