

Method for Predictive Trend Analytics with SNS Information for Marketing

Kohei Arai¹, Ikuya Fujikawa², Yusuke Nakagawa³, Sayuri Ogawa⁴
Information Science Department, Saga University, Saga City, Japan¹
SIC Holdings Co., Ltd, Hakata-ku, Fukuoka City, Fukuoka, Japan^{2,3,4}

Abstract—A method for predictive trend analytics with social media information is proposed for marketing. Through keyword analysis, page view analysis, access analysis, heat map analysis, Google Analytics, real time analysis, company and competitor analysis, trend analysis with the social media data derived from X (former tweeter), Instagram, Facebook, YouTube, TikTok, market trend can be predicted. The proposed method is created in a local server and is extended to AWS cloud. The proposed system, also ensure negative / positive analysis from the acquired social media information. Through some experiments, it is found that by using AI to analyze social data by category, you can visualize the degree of attention for each keyword, model relationships between information, identify trending keywords, and where the keywords are in their lifecycle. It turns out that it's possible to categorize which ones exist and predict which ones will scale up in the next six months. In addition, corporate product development and marketing personnel can identify themes, materials, benefits, etc. that have signs of becoming popular based on insights based on predictive behavioral data obtained from the proposed method and system and utilize them in new business development and new product planning.

Keywords—X (former tweeter); Instagram; Facebook; YouTube; TikTok; market trend; AWS; Google analytics; keyword analysis; page view analysis; access analysis; heat map analysis

I. INTRODUCTION

Trend analysis is the process of predicting the future by analyzing changes in certain elements from the past to the present based on information obtained from SNS and websites. Marketing challenges include "difficult to understand cost-effectiveness," "rapid market changes," and "customer needs are diversifying, making it difficult to predict demand." Predict the market and overcome challenges with trend analysis: For example, if you want to know about changes in the market, you can use trend analysis to quickly catch trends in the field of products and services your company provides. Additionally, if it is used when formulating strategies such as production volume and advertising, it is possible to minimize costs.

In order to plan the right marketing measures, it is necessary to accurately understand and analyze all kinds of information. By conducting trend analysis, it is possible to grasp and analyze "information about competitors and market trends" necessary for marketing. In the process of trend analysis, it becomes easy to understand the differences between the products and services offered by your company and those of your competitors, as well as the differences in marketing. By collecting such information, it will be easier to determine the direction your company should take in future business

development, and it will also be possible to predict market trends that will lead to the development of new products.

Furthermore, impressions and evaluations of products and services differ between company personnel and consumers. By conducting questionnaire surveys when analyzing trends, you can draw out the true feelings of consumers and formulate more effective marketing strategies. By creating personas using the collected data, you can plan measures that are optimized for users. The more detailed the persona settings are, such as age, gender, occupation, hobbies, and information sources, the better you will understand your customers, and the easier it will be to grasp the needs of actual users. By sharing this data within a company, it is possible to share an image between departments and eliminate misperceptions in marketing.

A method for predictive trend analytics with social media information is proposed for marketing. Through keyword analysis, page view analysis, access analysis, heat map analysis, Google Analytics, real time analysis, company and competitor analysis, trend analysis with the social media data derived from X (former tweeter), Instagram, Facebook, YouTube, TikTok, market trend can be predicted. The proposed method is created in a local server and is extended to AWS cloud. The proposed system, also ensure negative / positive analysis from the acquired social media information.

In the following section, some of the related research works are described in Section II, research background in Section III followed by the proposed method in Section IV. Then, some experiments are described in Section V, followed by a conclusion and future research works in Section VI and Section VII respectively.

II. RELATED RESEARCH WORKS

Database marketing has been successfully introduced in [1]. The book "Enterprise One to One: Tools for Competing in the Interactive Age" has also been published [2]. An attempt has been made to apply the concept of CLTV (Customer Lifetime Value) to FMCG (Fast-Moving Consumer Goods) [3]. Furthermore, several studies on CLTV have been introduced and reviewed, with each paper presenting different definitions of customer lifetime value, target industries, business models, and conditions for calculation [4].

Instances of COCA (Cost of Customer Acquisition) have been described, which refers to the cost of acquiring customers [5]. CLTV models and applications for marketing have been proposed and their applicability discussed [6]. Marketing study guides have been published and well-reviewed [7].

The book "Customer Profitability and Lifetime Value" has been published and extensively discussed [8]. Managing customers profitably has also been investigated and discussed [9]. Additionally, the analysis and discussion of "Performance management, which includes integrating strategy execution, methodologies, risk, and analytics," has taken place [10].

The paper "RFM (Recent Frequency Monetary) and CLTV: Using iso-value curves for customer base analysis" has been published, proposing and validating a method for marketing research [11]. Similarly, the paper "Autonomous CRM control via CLTV approximation with deep reinforcement learning in discrete and continuous action space" has been published, attempting to use CLTV approximation for CRM control [12].

On the other hand, CLTV has been well defined and discussed [13]. The paper "EDA of predictive modeling with "R" (a software tool for statistics) for risk management using machine learning" has been published, proposing and validating the use of EDA for predictive modeling [14]. Meanwhile, it is widely acknowledged that EDA is an important and useful technique in data science for analyzing and understanding data better. EDA involves exploring and visualizing the data to identify patterns, relationships, and anomalies.

EDA helps identify missing values, outliers, and other inconsistencies in the data, which can then be addressed before building predictive models. By visualizing the data, EDA also facilitates communicating insights to stakeholders and guiding further analysis. Furthermore, EDA is increasingly recognized as a critical step in any data analysis project as it enables a better understanding of the data, identification of potential issues, and provides insights for further analysis and decision-making. The concept of EDA has also been proposed and discussed [15]. Data analysis and regression have been well proposed for EDA analysis [16].

The paper "Suitability of random forest analysis for epidemiological research: Exploring sociodemographic and lifestyle-related risk factors of overweight in a cross-sectional design" has been published, studying and reporting on the suitability of random forest analysis for epidemiological research [17]. Additionally, EDA has been well defined, described, and investigated for its usefulness [18].

The paper "Customer Profiling Method with Big Data based on BDT and Clustering for Sales Prediction" has been published, proposing and validating a method for sales prediction using big data [19]. Meanwhile, the paper "Modified Prophet+Optuna Prediction Method for Sales Estimations" has been published, also proposing and validating a prediction method for sales using actual sales data [20].

III. RESEARCH BACKGROUND

Google Trends is a free tool provided by Google. According to the Ministry of Information and Communications White Paper 2020 Edition, Google will have an 85.6% share of search engines in the world in 2022. By using Google Trends, it can be checked the trend of popular words searched on Google, which has a high share rate in the world, and the rapidly increasing words. For example, it is not only possible to predict demand based on rapidly increasing keywords, but

also to understand whether the market is expanding or contracting by checking the increase or decrease in the number of searches for a specified word.

This free tool provided by Yahoo! JAPAN is unique in that it allows you to check X trends in real time. Yahoo! Real-time Search has a feature that allows you to analyze emotions by displaying the "emotion percentage" in a pie chart, so it can be seen whether the posted content is negative or positive. If opinions about a company's advertising marketing tend toward negative posts, it can be taken improvement actions based on those comments. If it is searched for your company's name or the name of the product or service it offers using X, it can be performed a so-called "ego search" and check the true intentions of a company in real time without any presumptions.

When using SNS for trend analysis, it is important to take advantage of the characteristics of each SNS. In the case of Instagram, a feature is that it analyzes the situations in which trending products and services are used. In the case of X, a feature is that it analyzes what kind of impressions is being expressed about trending products and services. Additionally, in the case of YouTube, a feature is that it analyzes what kind of marketing measures are being taken by other companies.

The usage rate of SNS used by users differs depending on the age group. Only YouTube has a usage rate that accounts for more than half of all age groups, while the usage rate of other SNSs peaks among people in their teens or 20s and then declines. In particular, regarding TikTok, the usage rate among teenagers is 62.4%, while the usage rate among people in their 40s and above has dropped by around 50 points. The number of SNS users is 330 million on X, 1 billion on Instagram, 1 billion on TikTok, 2.934 billion on Facebook, 573 million on Weibo, 193 million on LINE, Pinterest has 444 million people.

The information that can be obtained with PyTrends¹ is things that are gaining attention, interest by subarea, related information, related keywords, and rapidly increasing rankings by year. It provides search suggestions, obtains category content, and investigates up to five keywords at a time for the above items.

TrendScope is a tool that analyzes all of X's past several years of data and articles published in the media, and extracts data that can be used for management decisions, product development, and marketing. It not only captures the current state of the market from the tweets of consumers, but also analyzes past data, making it possible to grasp signs of the future. We can identify which keywords are currently trending, categorize where they are in their lifecycle, and predict which ones will scale up in the next six months. A company's product development and marketing staff can identify themes, materials, benefits, etc. that have signs of becoming a trend, using insights based on predictive behavior data obtained from TrendScope, and utilize this information in new business development and new product planning.

Trends in the world, for example, McDonald's, after the first trend ended, it spread explosively with the second trend. We use AI to catch up on such trends and develop services and

¹PyTrend: <https://norari-kurari-way.com/python-trend/>

products that are on-trend. Analyzing the keywords that are rapidly increasing on SNS, knowing the trends at that time, and utilizing them for product development. In addition to product development, by analyzing consumer voices posted on SNS, etc., we can improve existing services. We would like to use this knowledge to improve and create new services that people want. It also collects people's tweets on social media, analyzes the tweets using AI to find trends, and visualizes the analysis results on the screen.

Monitor X posts using Yahoo! Real-time search. It can be checked the number of tweets that include a specific keyword and the actual content of the tweets, and it can be also sorted tweets by newest arrival or topic. In the "emotion ratio" section, posts are automatically judged, and the ratio of positive/negative posts can be displayed. Since we are unable to confirm specific tweets that have been judged positive/negative, we can only confirm these figures as reference numbers. There are tools that can help it finds keywords related to a company that are being talked about and prevent the risk of becoming a hot topic.

IV. PROPOSED METHOD

The proposed method allows acquisition of tweet information as an example of SNS information for trend prediction with tweet API in the environment of AWS. The system configuration is shown in Fig. 1.

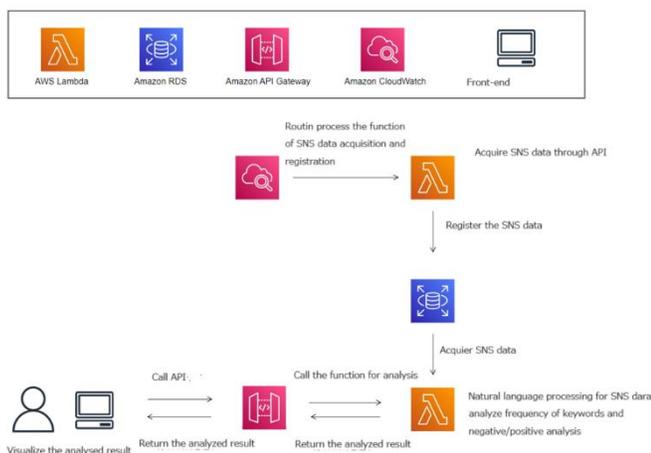


Fig. 1. System configuration of the proposed method for tweet information acquisition relating to trend analysis and prediction based on AWS environment.

There are four major functionalities, AWS Lambda, Amazon RDS, Amazon API gateway, and Amazon CloudWatch together with front end under the user interfaces. As a routine process, the function of SNS data acquisition and registration is always working. With the Lambda, SNS data are acquired through API. Then the acquired SNS data are registered to the Amazon RDS. When users call the API and the function for SNS data analysis, natural language processing is activated. The frequency of keywords is analyzed. Also, after the morphological analysis (part of speech analysis and separation of words) as a part of natural language processing, the acquired SNS data of texts are analyzed with negative or positive of emotional impressions through Lambda. Then the

analyzed results are transferred to the users with the front end of visualization terminals.

The database of entities is shown in Table I and Entity Relations are shown in Fig. 2.

TABLE I. DATABASE OF ENTITIES

No	Name	Table	Usage
1	User	Customer	Customer Information Control
2	tweet_data	Acquired tweet data	Acquired tweet data control
3	result_data	Tweet analysis result	Tweet analyzed result control
4	category	Category	Category control
5	industry	Industry society	Industry information control

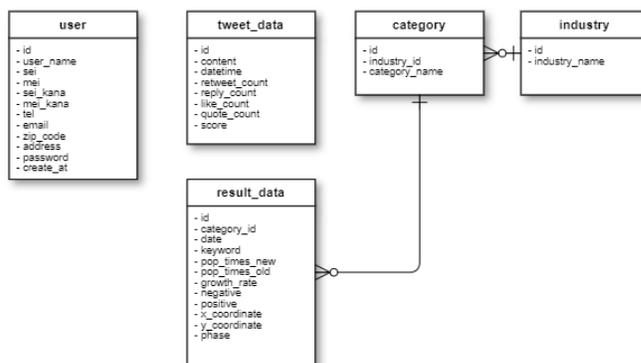


Fig. 2. Entity relations.

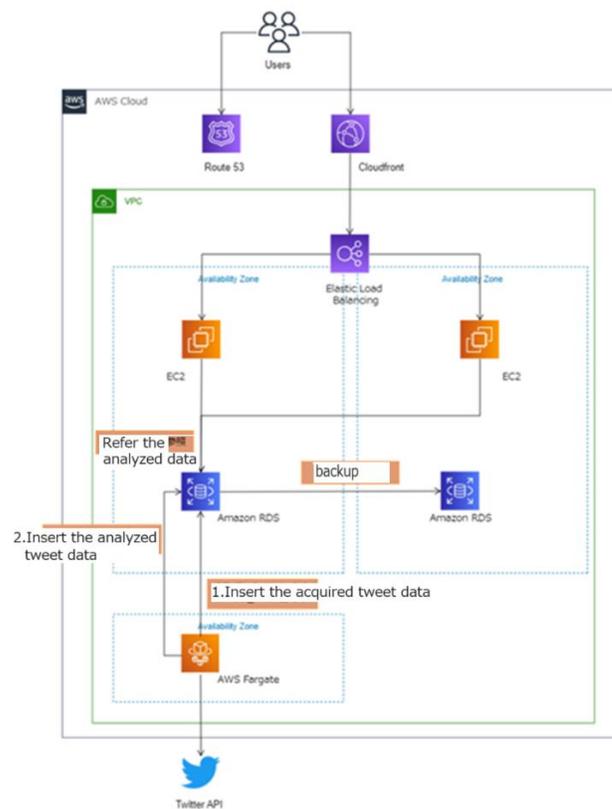


Fig. 3. Detailed configuration of the proposed trend analysis system.

There are five tables in the Amazon RDS, User, Tweet data, Category, Industrial society (Companies in the same industrial society, etc.), and Analyzed result data. The acquired tweet data are categorized based on users, companies. Analyzed results are highly correlated with the categories together with industrial society.

The detailed configuration is shown in Fig. 3. In the figure, backup control scheme and insertion of the acquired SNS data of tweet information as well as referred analyzed results through EC2 are shown. As for the front end, display design of user authentication is shown in Fig. 4. Users have to input their ID and password for authentication and may enter the proposed trend analysis system. After the user authentication, tweet data is collected in accordance with the keywords. Then trend analysis results are appeared together with negative / positive analysis against tweet information of text data. In this stage, “MeaningCloud” of sentiment analysis tool is used as shown in Fig. 5. Thus, the users may realize the impression of tweet information whether or not good or bad. The trend of the users’ impression, then shown in the detailed trend analysis result.

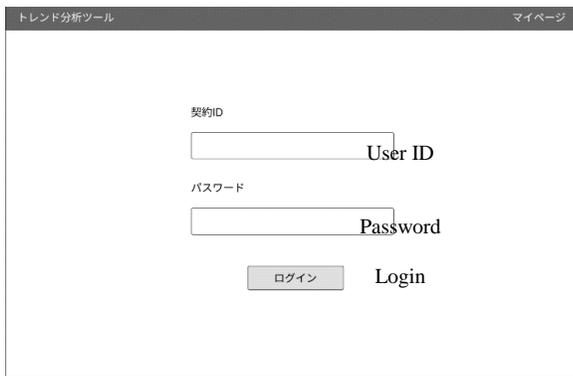


Fig. 4. Login display image.

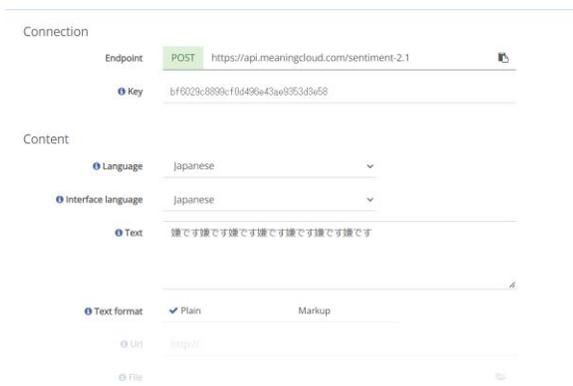


Fig. 5. Screenshot of the “meaningcloud”.

V. EXPERIMENT

The aforementioned functionalities of the proposed trend analysis system are performed and confirmed. After the user authentication, the main menu appears on the screen as shown in Fig. 6. Pul-down menu is then displayed, “Industry name”, “Category”, “Area”, and “Period” appears. In Fig. 6, “Hair salon”, “Shampoo”, “Japan”, and “January 2023” are selected

for each. After all, Analysis of radio button is clicked. Then trend analysis begins.

Fig. 7 shows an example of the detailed trend analysis result with the aforementioned keywords which are shown in Fig. 6.

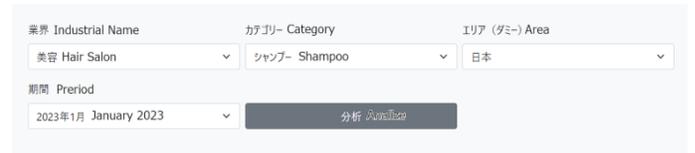


Fig. 6. Main menu of the proposed trend analysis system.

Rank ランク	Keywords キーワード	Frequency 出現数	Grow rate 成長率	Nega/Posi ネガポジ分析
1	ボタニカル Botanical	5394	+40.4%	50 20
2	泥 Mad	5311	+20.3%	50 20
3	美容 Beauty	3455	+11.0%	50 20
4	ラクダ Easy	3222	+60.3%	50 20
5	すごい Great	1600	-18.4%	50 20
6	ママ Mamy	1588	-22.1%	50 20
7	ikuya Ikuya	1200	+40%	50 20
8	福岡 Fukuoka	1023	+40%	50 20
9	サロン Salon	998	+40%	50 20
10	おすすめ Recommend	400	+40%	50 20
11	どろ Mady	322	+40%	50 20
12	ドロアース Droearth	200	+40%	50 20
13	bsi BSI	109	+40%	50 20
14	メーカー Maker	101	+40%	50 20
15	バラ Roses	80	+40%	50 20

Fig. 7. Example of the detailed trend analysis result with the aforementioned keywords.

In January 2023, the top three of the keywords collected from the tweet information with the keyword of “Shampoo” in the hair salon companies in Japan are (1) Botanical, (2) Mad and (3) Beauty. The number of keywords appeared in tweet information for the corresponding top three keywords are 5394, 5311, and 3455. Furthermore, the grow rate of the corresponding keywords is 40.4 %, 20.3 %, and 11.0 %, respectively. Moreover, the sentiment analysis results show

The example of the final trend analysis result in particular for the number of keywords of “Botanical” and the period during January 2022 and October 2022 as well as Negative / Positive analyzed result is shown in Fig. 13.

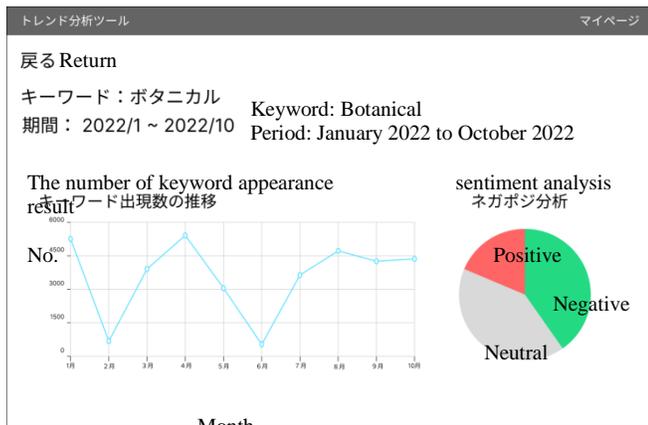


Fig. 13. Example of the number of keywords of “Botanical” and the period during January 2022 and October 2022 as well as Negative / Positive analyzed result.

VI. CONCLUSION

A method for predictive trend analytics with social media information is proposed for marketing. Through keyword analysis, page view analysis, access analysis, heat map analysis, Google Analytics, real time analysis, company and competitor analysis, trend analysis with the social media data derived from X, Instagram, Facebook, YouTube, TikTok, market trend can be predicted. The proposed method is created in a local server and is extended to AWS cloud. The proposed system, also ensure negative / positive analysis from the acquired social media information.

Through some experiments, it is found that by using AI to analyze social data by category, you can visualize the degree of attention for each keyword, model relationships between information, identify trending keywords, and where the keywords are in their lifecycle. It turns out that it's possible to categorize which ones exist and predict which ones will scale up in the next six months. In addition, corporate product development and marketing personnel can identify themes, materials, benefits, etc. that have signs of becoming popular based on insights based on predictive behavioral data obtained from the proposed method and system and utilize them in new business development and new product planning.

VII. FUTURE RESEARCH WORKS

Further investigations are required for not only tweet information derived trend analysis but also Instagram, Facebook, YouTube, TikTok, and so on for market trend has to be predicted in the near future.

ACKNOWLEDGMENT

The authors would like to thank to Professor Dr. Hiroshi Okumura and Professor Dr. Osamu Fukuda for their valuable discussions.

REFERENCES

- [1] Stone, Merlin and Shaw, R, “Database marketing”. Aldershot, Gower. 1988.
- [2] Peppers, D., and M. Rogers, "Enterprise One to One: Tools for Competing in the Interactive Age." New York: Currency Doubleday, 1997.
- [3] Hanssens, D., and D. Parcheta (forthcoming). “Application of Customer Lifetime Value (CLV) to Fast-Moving Consumer Goods.”, 2011.
- [4] Nakamura and Higa, A Review of Past Research on Customer Lifetime Value Measurement, Japan Society for Management Information National Research Presentation Conference Abstracts, DOI <https://doi.org/10.11497/jasmin.2011s.0.530.0>, 2011.
- [5] <https://swifterm.com/how-to-calculate-cost-of-customer-acquisition-cac-or-coca/> accessed on 11 May 11, 2023.
- [6] Berger, P. D.; Nasr, N. I., “Customer lifetime value: Marketing models and applications”. *Journal of Interactive Marketing* 12 (1): 17–30. doi:10.1002/(SICI)1520-6653(199824)12:1<17::AID-DIR3>3.0.CO;2-K 1988..
- [7] Fripp, G, "Marketing Study Guide" Marketing Study Guide, 2014.
- [8] Adapted from "Customer Profitability and Lifetime Value," HBS Note 503-019, 2014..
- [9] Ryals, L. *Managing Customers Profitably*. ISBN 978-0-470- 06063-6. p.85, 2008.
- [10] Gary Cokins, *Performance Management: Integrating Strategy Execution, Methodologies, Risk and Analytics*. ISBN 978-0-470-44998-1. p. 177, 2009.
- [11] Fader, Peter S and Hardie, Bruce GS and Lee, Ka Lok, “RFM and CLV: Using iso-value curves for customer base analysis”. *Journal of marketing research* (SAGE Publications Sage CA: Los Angeles, CA) 42 (4): 415-430. doi:10.1509%2Fjmk.2005.42.4.415, 2005.
- [12] Tkachenko, Yegor, “Autonomous CRM control via CLV approximation with deep reinforcement learning in discrete and continuous action space”. arXiv preprint arXiv:1504.01840. doi:10.48550/arXiv.1504.01840, 2015.
- [13] V. Kumar, *Customer Lifetime Value*. ISBN 978-1-60198-156-1. p.6, 2008.
- [14] Hirokazu Iwasawa, Yuji Hiramatsu, "EDA (Exploratory Data Analysis) Predictive Modeling with R: For Risk Management Using Machine Learning Tokyo Tosho pp.46-62, 2019.
- [15] Yasuhito Mizoe, "Concept of Exploratory Data Analysis," *Estrela*, No.65, August 1999, pp.2-8, 1999.
- [16] Mosteller, F. and J.W. Tukey, “Data Analysis and Regression”, Addison- Wesley, 1977.
- [17] Noora Kanerva, Jukka Kontto, Maijaliisa Erkkola, Jaakko Nevalainen, Satu Männistö, “Suitability of random forest analysis for epidemiological research: Exploring sociodemographic and lifestyle-related risk factors of overweight in a cross-sectional design.” *Scandinavian Journal of Public Health*, Vol 46(5) pp.557-564, 2018.
- [18] Tukey, J.W., “Exploratory Data Analysis”, Addison-Wesley, 1977.
- [19] Kohei Arai, Zhang Ming Ming, Ikuya Fujikawa, Yusuke Nakagawa, Ryoya Momozaki, Sayuri Ogawa, Customer Profiling Method with Big Data based on BDT and Clustering for Sales Prediction, *International Journal of Advanced Computer Science and Applications*, 13, 7, 22-28, 2022.
- [20] Kohei Arai, Ikuya Fujikawa, Yusuke Nakagawa, Ryoya Momozaki, Sayuri Ogawa, Modified Prophet+Optuna Prediction Method for Sales Estimations, *International Journal of Advanced Computer Science and Applications*, 13, 8, 58-63, 2022.

AUTHOR'S PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was

a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is a Science Council of Japan Special Member since 2012. He is an Adjunct Professor of University of Arizona, USA for 1998 to 2020 and is Adjunct Professor of Nishi-Kyushu University as well as Kurume Institute of Technology (Applied AI Laboratory) up to now. He also is Vice Chairman of the Science Commission "A" of ICSU/COSPAR since 2008 then he is now award committee member of ICSU/COSPAR. He wrote 77 books and published 710 journal papers as well as 550 conference papers. He received 66 of awards including ICSU/COSPAR Vikram Sarabhai Medal in 2016, and Science award of Ministry of Education of Japan in 2015. He is now Editor-in-Chief of IJACSA and IJISA. <http://teagis.ip.is.saga-u.ac.jp/index.html>