New Knowledge Management Model: Enhancing Knowledge Creation with Zack Gap, Brand Equity, and Data Mining in the Sports Business

Fransiska Prihatini Sihotang[1], Ermatita[2]*, Dian Palupi Rini[3], Samsuryadi[4]
Doctoral Program in Engineering Science, Universitas Sriwijaya, Palembang, Indonesia[1]
Faculty of Computer Science, Universitas Sriwijaya, Palembang, Indonesia[2, 3, 4]
Faculty of Computer Science and Engineering, Universitas Multi Data Palembang, Palembang, Indonesia[1]

***Abstract*—This research improves Socialization, Externalization, Combination, and Internalization (SECI) knowledge management model by combining it with Zack's knowledge gap model, brand equity concept, and data mining. Zack's model is incorporated into the SECI model to identify the gap between the knowledge in the organization and the knowledge that the organization should possess. We add the data mining techniques to determine that knowledge gap. The uniqueness of this study lies in the externalization and combination of the SECI model. In the externalization, "what the firm must know" is added; for that, we compile the questionnaire by adopting brand equity and distributing it to the athletes. In the combination, "what the firm knows" is added; we use a database already owned by sports business management. The modifications resulting from both models with data mining in this study were carried out to develop a new knowledge management model in the sports business sector. This new model will be valuable knowledge for sports business management to build strategies and increase their competitiveness in the sports market. In addition, other service business fields besides sports can also apply this new model to improve their knowledge management, which they can then use to improve their marketing strategies.**

Keywords—SECI model; zack model; data mining; brand equity; sport business

I. Introduction

The knowledge creation process is critical to increasing business value and ensuring the success of a business, including sports centers [1], [2]. According to Rot and Sobinska research [3], digital technology can improve knowledge management. Thus, similar to other industries [4], the sports service sector must manage knowledge effectively to boost its competitiveness. Numerous studies have shown that well-organized knowledge significantly enhances a company's innovation [5]. Efficient knowledge management strengthens a company's ability to innovate by ensuring that relevant and accurate information is readily accessible for decision-making. Efficient knowledge management will also create new products, services, and business processes, which give the company a competitive advantage.

The SECI model is the commonly used knowledge management model across various fields. However, it has faced criticism from several researchers. The model, often seen as overly simplistic, addresses the socialization and externalization processes inadequately and overlooks cultural differences among members of an organization. It has been critiqued for its individualistic focus, which limits its ability to represent interactions and interdependencies effectively [6]. Other studies indicate that this model centers only on converting human knowledge, neglecting knowledge stored in databases, external sources, and variations in organizational knowledge. This limitation can lead to a knowledge gap between what the organization currently knows and what it ideally should know [7], [8].

Data mining techniques reveal hidden patterns and relationships within large datasets, supporting more informed and strategic decision-making. Data mining is playing a growing role in data analysis, including applications like predicting stock prices [9], assessing disease risks [10], analyzing customer churn [11], segmenting customers [12], measuring customer satisfaction [13]–[16], forecast customer behavior, streamline supply chains, and improve customer relationship management. [7], [14], [17].

Data mining is also applied in sports [18], such as for predicting injuries in soccer players [19], identifying key attributes and metrics that impact NBA player salaries and performance [20], and detecting tactical patterns through market basket analysis in beach volleyball games [21]. The many studies related to data mining and sports show that sports attract the world's attention. However, many find identifying the optimal data mining method challenging due to the wide variety of processed datasets [17].

This study aims to build on existing research by integrating the SECI model, Zack's knowledge gap model, data mining techniques, and a brand equity questionnaire to create a new knowledge management model specifically for the sports business. To identify the knowledge gaps, we will use two data sources: 1. the database that management already has (what the firm knows), we call it the secondary dataset; 2. a questionnaire containing 20 brand equity questions distributed to athletes/ customers (what the firm must know), we call it the primary dataset.

In addition to the belief that there is a knowledge gap between the two datasets, Rungtrakulchai [22] said that adding a brand equity questionnaire is also important because research has shown that marketing practitioners need to understand the level of brand equity in a particular market area. Other research [23] also said that marketing practitioners can obtain this from customer/consumer responses to the services provided. By merging these models, we aim to provide valuable insights for sports business management, helping organizations to develop effective strategies and improve their competitiveness.

This article consists of several sections. Section II includes a literature review that examines relevant theories and previous research related to this study. Section III provides the research methodology, explaining how we combined various models to create a new knowledge management model. Section IV represents the results and discussion, detailing the outcomes of this combination into the new knowledge management model. Finally, Section V offers the conclusion.

II. Literature Review

A. SECI Model

The SECI model, an adaptable knowledge management framework, encompasses four critical processes: Socialization, Externalization, Combination, and Internalization. Its adaptability is evident in its successful application across various sectors, from cooperatives [24] and universities [25] to software development [26] and even in the design of a knowledge management system for trader education in the marketplace [27]. Miao et al., in their research [28], said that this model's ability to acquire and align knowledge with business strategy has also found a place in the banking sector.

Fig. 1 illustrates the SECI model. The SECI model has been widely adopted and directly applied in shaping knowledge management across various organizations [6] or by first modifying or developing the model [29]. However, some researchers have critiqued the SECI model from different perspectives. The socialization and externalization processes in this model are often ineffective due to limitations on the freedom to exchange ideas at the industrial level. Additionally, some argue that the SECI model is too simplistic, neglects the role of context, and does not adequately account for the cultural differences among organizational members. Another criticism is that the model portrays knowledge transfer as a linear process, even though this may not reflect real-world situations. Some also contend that the model's fundamental structure is individualistic, making it challenging to represent interactions and interdependencies effectively [6].

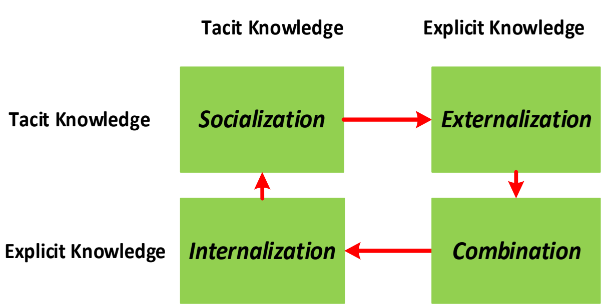

Fig. 1. SECI model [25], [30].

Other researchers have highlighted additional weaknesses of the SECI model, including those pointed out by Almuayqil, who argued that the model focuses solely on converting human knowledge while overlooking knowledge stored in databases and other technological resources. Meanwhile, Yao noted that the SECI model fails to account for external knowledge inputs and the variations in knowledge within the organization [7]. To address these shortcomings, we incorporate the knowledge gap analysis from Zack's model into the externalization and combination phases of the SECI model.

B. Zack gap model

Aligning business strategy with knowledge is crucial for an organization [28]. Learning and knowledge acquisition are necessary to sustain and enhance a competitive advantage. Fig. 2 illustrates the knowledge and strategic gap in Zack's model. Aligning a company's business strategy with its knowledge resources is crucial, as it plays a critical role in driving innovation and improving business performance. This approach has become a primary focus for companies seeking success in today's competitive environment [28].

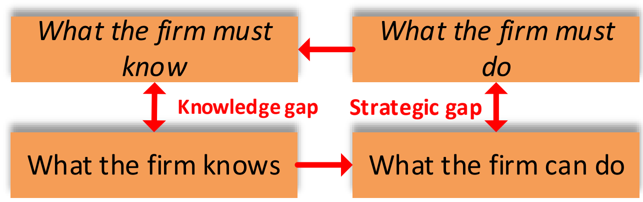

Fig. 2. Zack gap model [8], [28].

Using Zack's model, Gap analysis compares an organization's existing knowledge with its strategic needs. Knowledge and strategy are viewed as opposing poles: what is known or done and what should be known or done. By examining these points, gaps in both organizational knowledge and strategy can be identified. In this study, we focus on the knowledge gap from Zack's model and integrate it with the SECI model. In this research, we will find the knowledge gap by comparing the database that management already has (what the firm knows) with the questionnaire (what the firm must know) distributed to athletes (customers).

C. Data Mining in Knowledge Management

Data mining methods generally consist of classification, association, and clustering. Classification is a technique in data mining to group data based on the data's relationship to sample data [31]. The clustering method is a process of grouping data objects similar to each other into the same cluster and different from objects in other clusters [32]. Based on a literature study conducted by Ngai, the clustering method is mainly used for customer identification, while the classification method is mainly used for customer acquisition and retention [17]. Meanwhile, association in data mining is a technique for obtaining hidden relationship patterns between several items in a dataset [33].

D. Brand Equity

Brand equity is an added value imposed on a particular product or service that reflects consumers' thoughts, feelings, and actions in adapting the product or service and influences customer value perception [22]. It is also related to price, market share, and brand profitability [34]. Marketing practitioners are essential to understand the level of brand equity in a particular market area [22], which is obtained from customer/consumer responses to the services provided [23]. As shown in Fig. 3, brand equity consists of Brand Awareness (BA), Perceived Quality (PQ), Brand Association (BA), Brand Loyalty (BL), and Decision (D) [23].

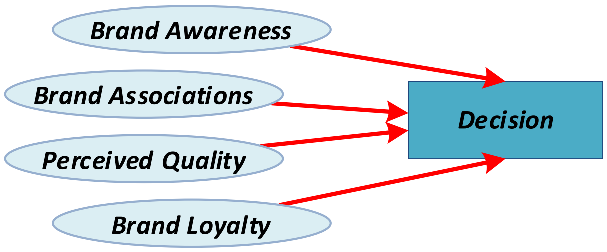

Fig. 3. Brand equity model [22], [35].

III. Research Methodology

This research combines the SECI model, knowledge gap analysis from the Zack model, data mining techniques, and a brand equity. The knowledge gap analysis from the Zack model was added to the SECI model in the externalization and combination section. Externalization of the SECI model is the form of changing knowledge from tacit to explicit, and we match it with "what the firm must know" in the Zack model. The combination of the SECI model is knowledge in an explicit form that already exists, is developed, and disseminated through various media that are more systematic. We match it with " what the firm knows" in the Zack model.

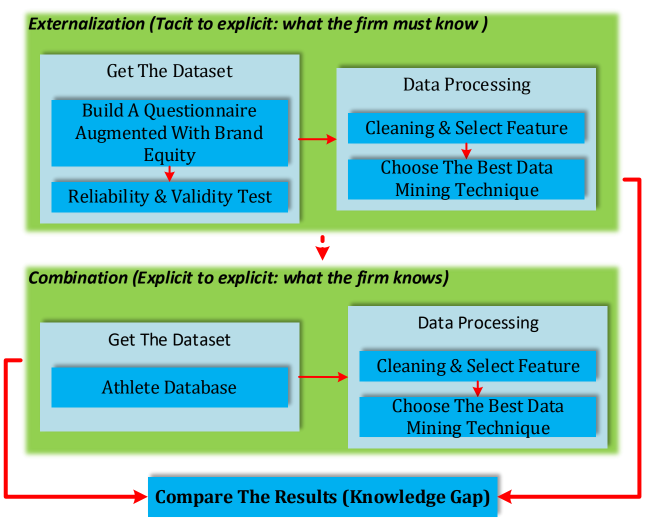

Fig. 4. Combining two models into a new model.

Fig. 4 displays how we match and combine these two models into a new model. To use this new model, we need to acquire a dataset. Since this model will perform knowledge gap analysis, it requires two types of datasets: one that the organization already owns and another that contains knowledge about what the organization should know.

Data cleaning and feature selection need to be performed on both datasets. Next, it is essential to test various data mining techniques and algorithms to identify the one best suited to both datasets. The choice of data mining techniques should be aligned with management needs, depending on the desired outcomes. After selecting the most effective algorithm for both datasets, it is essential to apply this algorithm to both datasets and analyze the results. To finalize this model, we must compare the data mining outcomes of the two datasets to identify any knowledge gaps. This knowledge gap can serve as valuable information for management, helping to enhance decision-making, develop strategies, and improve competitiveness in the sports market.

A. Get the Dataset

The new knowledge management model for sports business developed in this study requires two data sources, and then knowledge gaps will be sought from these data sources. We call these two data sources secondary and primary datasets. It is essential to ensure that the data sources are credible so that they can produce correct knowledge. In this research, we use a case study in taekwondo sports. So, we collected the secondary datasets from the Taekwondo Indonesia Integrated System (TIIS) application in the South Sumatra Province, which the government uses to process athlete data. Thus, this database is a credible knowledge base for this study. From the TIIS databases, we acquired sixty files in Microsoft Excel format and merged them into a single file for further analysis.

Based on Zack's knowledge gap model, companies' current knowledge (what the firm knows) needs to be compared with the knowledge they do not yet have but should have (what the firm must know). Therefore, it is necessary to discover new knowledge collected using research questionnaires. Therefore, the relationship between the database and the questionnaire is to assess knowledge gaps.

We obtained the primary dataset by distributing questionnaires to users (athletes) of Taekwondo services in South Sumatra Province, Indonesia. The questionnaire was distributed for two months as a Google form since the respondents were spread across many cities. To ensure the accuracy of the research data, parents or guardians filled out the questionnaire for athletes under the age of sixteen. For athletes over sixteen, they filled out the questionnaire directly. However, information related to brand equity was not obtained in the secondary dataset. Therefore, when compiling the questionnaire for the primary dataset, besides asking for the same information as in the secondary dataset, 20 questions related to the Brand Equity of the Club/Dojang where the athletes train were added. The questionnaire consists of Brand Awareness (BA), Perceived Quality (PQ), Brand Association (BA), Brand Loyalty (BL), and Decision (D), as shown in Table I.

The questionnaire was meticulously adapted to the case study, focusing on the sports service business, mainly taekwondo, in South Sumatra Province, Indonesia. Each statement item in the questionnaire was crafted based on interviews with Dojang owners and validated by the Head of the Achievement Division of Taekwondo Indonesia, South Sumatra Province. This rigorous process was undertaken to ensure the credibility of the data source. Each questionnaire item was designed to be answered by respondents by selecting from the options provided: Strongly Agree (SA), Agree (A), Neutral (N), Disagree (D), and Strongly Disagree (SD). Before distributing the questionnaire, a pre-test was conducted to assess its reliability and validity. A small sample of 30 athletes was selected for this pre-test. The results of the pre-test stated that the questionnaire was valid and reliable so that the distribution of the questionnaire could be continued.

TABLE I. QUESTIONNAIRE ITEMS

Code	Questionnaire Statement
BA1	When asked to name a place to practice taekwondo, the first thing I always think of is the club/dojang where I practice.
BA2	Talking about taekwondo reminds me of the Club/Dojang where I train.
BA3	The Club/Dojang where I train is my first choice when I want to start training taekwondo.
PQ1	Sabeum/ trainer at the Club/Dojang where I train is licensed and superior.
PQ2	I feel comfortable in the Club/Dojang where I practice because it is indoors.
PQ3	The supporting facilities at the Club/Dojang where I train are complete and meet standards.
PQ4	The training patterns at the Club/Dojang where I train are varied and not monotonous.
PQ5	The training schedule at the Club/Dojang where I train is relatively routine.
BAs1	My Club/Dojang is the best compared to the others.
BAs2	The training fees at my Club/Dojang are cheap.
BAs3	The training fees at My Club/Dojang are in line with the quality and service.
BAs4	The club/Dojang where I train is easily accessible from my house.
BL1	I will always choose the Club/Dojang where I am currently training as long as I train.
BL2	I will not move to another Club/Dojang.
BL3	I would recommend My Club/Dojang.
D1	I looked for as much information as possible about the Club/Dojang before joining.
D2	I looked for as much information as possible about the Club/Dojang before joining.
D3	I compared the training costs of the Club/Dojang where I train with others before joining.
D4	I consider the distance of the Club/Dojang where I train before joining.
D5	I decided to stay with the Club/Dojang where I train because it meets my expectations.

Following the distribution of the questionnaire, 1468 rows of athlete data were obtained, of which 1348 rows were deemed usable; the rest were damaged. To ensure validity and reliability, we do tests using the SPSS application. The results showed that the calculated r value for each indicator ranged from 0.534 to 0.893, or far above the r table value. Reliability testing was conducted by examining Cronbach's alpha value for each indicator. The research questionnaire is highly reliable if Cronbach's alpha is between 0.9 and 1. Based on the test results, Cronbach's alpha value for each indicator in this research questionnaire was more than 0.9. Based on this value [36], [37], it can be said that this research questionnaire is valid with a confidence level of 99%, and the reliability is very high.

B. Cleaning and Feature Selection

The cleaning process is essential to ensure the dataset is suitable for research. This process includes checking each column for missing data, converting the data types of specific columns to numeric, and removing unnecessary columns. Many studies have highlighted the importance of feature selection as a crucial preprocessing step [17]. Choosing the right features can significantly enhance the performance of machine learning algorithms by reducing data dimensionality, which helps eliminate irrelevant or redundant information. We applied this cleaning procedure to both datasets. However, we excluded the primary dataset's feature selection for 20 brand equity questions because we have to use it all.

We conduct three methods of feature selection in this study: Information Gain (IG), Chi-Square, and Recursive Feature Elimination (RFE). Each of these methods provides a distinct approach to identifying the most relevant features for clustering. The Information Gain measures entropy reduction, the Chi-Square method evaluates the independence of features, and the RFE method recursively removes the least important features. Combining these methods ensures a thorough selection process and improves the data quality used for clustering. The results of the feature selection process in this study are shown in Table II.

TABLE II. FEATURE SELECTION BY VARIOUS METHODS

Feature	IG	Chi-Square	P-Value	RFE
Classification	0.554438	1.424.732.394	4,20E-30	1
ClubName	0.108146	168.885.790	2,12E-31	2
ClubID	0.114757	455.513.151	1,22E-93	3
Age	0.007043	4.471.729	1,07E+05	4
Belt	0.032379	23.657.274	7,29E+00	5
Subdistric	0.049144	36.273.625	1,33E-02	6
Sex	0.027673	44.655.043	2,01E-04	7
BeltID	0.019311	36.009.465	1,52E-02	8
Disctric	0.036292	33.825.733	4,52E-02	9
Distance	0.000000	3.319.825	1,90E+05	10
AgeGroup	0.005404	0.179834	9,14E+05	11
City	0.028890	17.123.501	1,91E+02	12

C. Choose the Best Data Mining Technique

In our case study, taekwondo club management requires grouping its athletes based on their abilities, interests, residence, belts, and other club members' characteristics. So, we use the clustering algorithm. Applying the clustering algorithm in the taekwondo sector helps establish a more organized, efficient, and responsive structure to meet members' needs, ultimately enhancing training programs' quality and sustainability and supporting effective marketing strategy development.

This research compares the K-means and K-medoid clustering algorithms to find the most suitable one. We compare these two algorithms because there have been many studies that prove the accuracy of both in clustering [38]–[40]. In order to determine the optimal number of clusters of this algorithm, we use the Elbow Method, the Silhouette Coefficient, the Calinski-Harabasz Index, the Davies-Bouldin Index, the Bayesian Information Criterion (BIC), and Akaike Information Criterion

(AIC). Based on various tests that we conducted, we know that for both datasets that we have, the best algorithm to apply to both datasets is K-Medoids, with 2 clusters. Table III displays the results of each method in determining the optimum number of clusters. The tests using various methods show that the most suitable algorithm for both datasets is K-Medoids, with an optimal number of two clusters.

TABLE III. DETERMINING THE OPTIMAL CLUSTERS

Methods	Cluster number with K-Means		Cluster number with K-Medoids	
	Secondary Dataset	Primary Dataset	Secondary Dataset	Primary Dataset
Elbow Method	2	2	2	2
Silhouette Score	2	2	2	2
Calinski-Harabasz Index	2	2	8	8
Davies-Bouldin Index	2	2	2	2
BIC	7	8	6	6
AIC	7	8	6	6

IV. RESULT AND DISCUSSION

A. Brand Equity Result

Based on what has been explained previously, we added 20 questions related to Brand Equity to our research questionnaire. Respondents can answer each question by selecting the options provided: Strongly Agree (SA), Agree (A), Neutral (N), Disagree (D), and Strongly Disagree (SD). We separate the brand equity questionnaire results from other data (which is the same as secondary data) and present them in Table IV. This result is also a knowledge gap that we found in our case study on this research.

TABLE IV. BRAND EQUITY RESULT

Questionnaire Code	SA	A	N	D	SD
BA1	71%	23%	6%	0%	0%
BA2	66%	27%	7%	0%	0%
BA3	64%	28%	8%	0%	0%
PQ1	62%	29%	9%	0%	0%
PQ2	44%	18%	32%	4%	2%
PQ3	55%	32%	13%	0%	0%
PQ4	60%	32%	8%	0%	0%
PQ5	60%	33%	7%	0%	0%
BAs1	55%	27%	18%	0%	0%
BAs2	62%	31%	7%	0%	0%
BAs3	59%	33%	8%	0%	0%
BAs4	62%	31%	7%	0%	0%
BL1	61%	31%	8%	0%	0%
BL2	61%	29%	10%	0%	0%
BL3	60%	33%	7%	0%	0%
D1	53%	33%	14%	0%	0%
D2	54%	32%	13%	1%	0%
D3	45%	27%	23%	5%	0%
D4	54%	32%	12%	2%	0%
D5	63%	29%	8%	0%	0%

These results are precious for sports business management because they can help them create marketing strategies. For example, in question item PQ3, "The supporting facilities at the Club/Dojang where I train are complete and meet standards," the results of the questionnaire distribution show that the supporting facilities at the Club/Dojang have not been fully met and do not meet standards. Thus, sports business management can improve equipment and supporting facilities to provide comfort for its members in training and retain those members.

Then, the BL3 question item "I would recommend My Club/Dojang" shows that most athletes would recommend the Club/Dojang where they train others. Thus, sports business management can approach and offer various promotions to their members to maintain good relationships and retain these members. Then, the BAs3 question item “The training fees at My Club/Dojang align with the quality and service” revealed that not all respondents were satisfied with the quality of service compared to the costs. This is very valuable for management, as it allows them to improve the quality of their services in the future, commensurate with the costs incurred by their members.

B. Data Mining Result

The tests using various methods show that the most suitable algorithm for both datasets in the Taekwondo case is K-Medoids, with an optimal number of two clusters. Therefore, we processed our datasets with the selected algorithm using two clusters for both datasets. Fig. 5 shows the results of the clustering we performed on both datasets, which shows that there are differences between the two (knowledge gap). The differences align with the research objective of identifying knowledge gaps. The proposed algorithm is flexible and not more suitable for one data type than another.

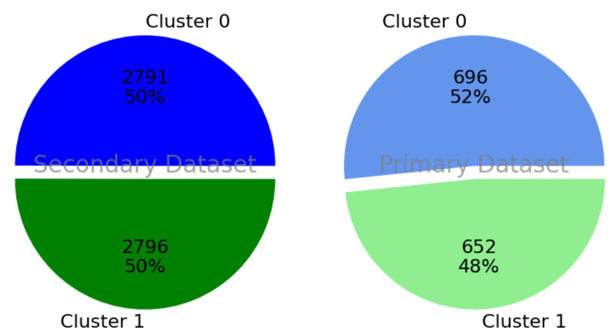

Fig. 5. The clustering result of both dataset.

Based on the results displayed in the clustering process using both datasets, although only a little, there is still a difference (knowledge gap) between the secondary and the primary datasets. The difference lies in the percentage of the secondary dataset, which is exactly 50:50, while for the primary dataset, it is 52:48. Although the difference is only 2%, it is essential for sports business management. Knowing this difference can help them make more careful decisions regarding policies to increase their business's competitiveness.

Based on the differences between the two datasets (knowledge gap), we analyze them based on several categories: athlete gender, athlete age categories, athlete belt, and the distance of the Club/Dojang where the athlete trains from their respective homes. These results are helpful for sports business management in determining their future customer targets.

Fig. 6 displays the distribution of athletes' gender on both datasets. The graph shows a difference in the percentages of the two datasets. In the secondary dataset, in cluster 0, the ratio of female and male athletes is 41:59, while in cluster 1, it is 37:63. In the primary dataset, the ratio is 36:64 for cluster 0 and 42:58 for cluster 1. Ensuring the gender distribution of Club/Dojang members is very important as it will influence promotion policies and the approach towards members/ potential members. These results align with previous research that discussed the influence of gender on purchasing decisions [41].

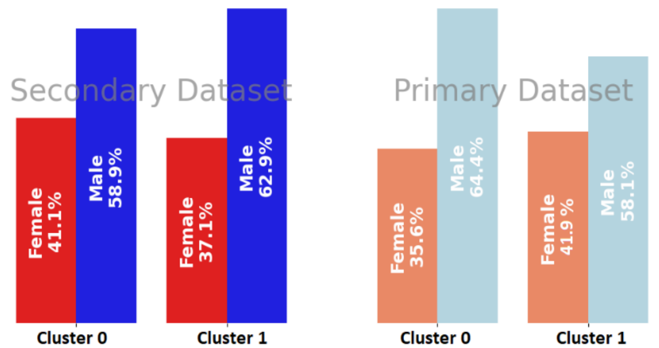

Fig. 6. Athletes gender distribution.

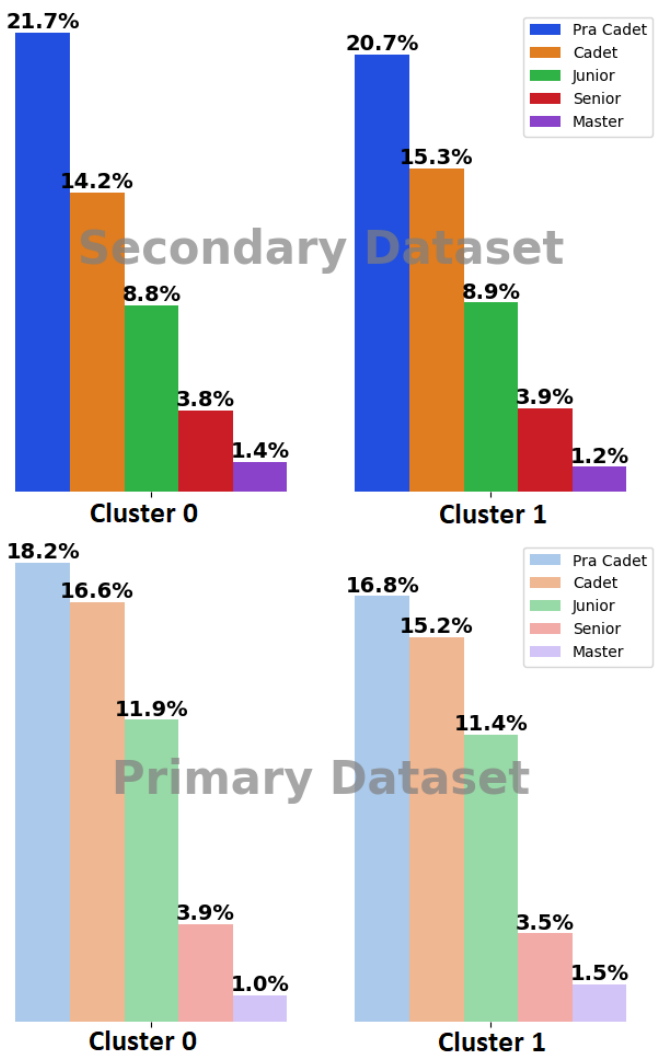

Fig. 7. Athletes age categories distribution.

For competition purposes, participants are classified by their birth year to determine their eligibility within these age groups, ensuring a fair matchup among competitors. Fig. 7 shows the differences between the two datasets. The most striking difference is in the Cadet class, which is intended for athletes aged 12 to 14. This difference is important for sports business management to know and to ensure the accuracy of their data, which will help them create marketing strategies. These findings are essential for sports business management, as understanding the specific age groupings within different clusters can inform training programs, marketing strategies, and other decision-making processes. Similarly, if an organization aims to expand or diversify its athlete base, knowing where certain age groups are underrepresented or overrepresented could help make strategic adjustments.

By ensuring accurate data analysis and a clear understanding of the athlete demographic across clusters, sports organizations can better align their offerings with the needs and preferences of their target audience, ultimately enhancing the effectiveness of training programs and marketing strategies.

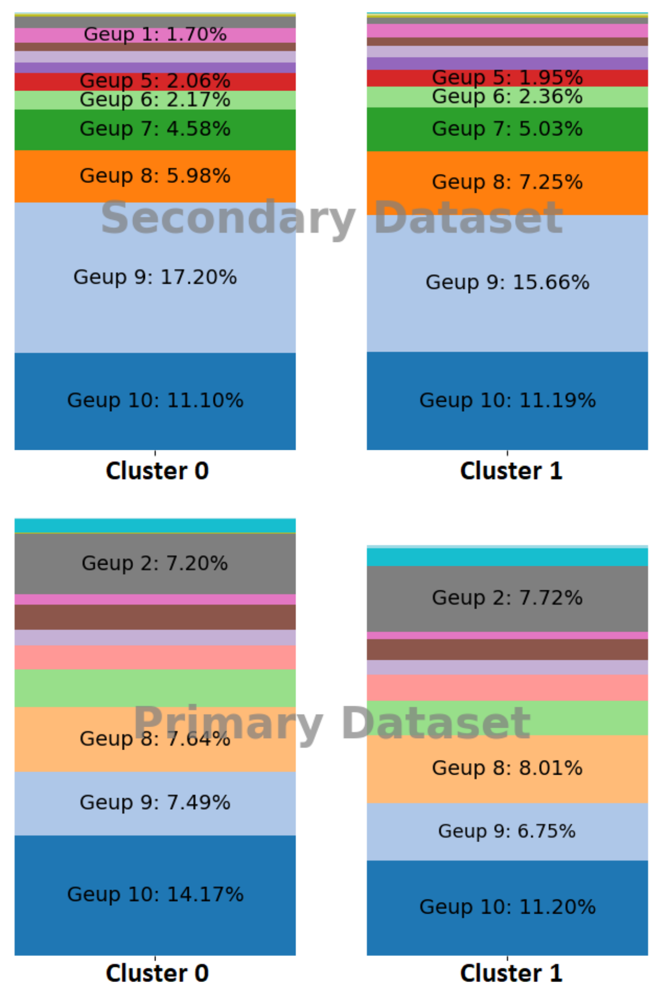

Fig. 8. Distribution of athlete belts.

Fig. 8 shows the distribution of athletes based on their belts. The most striking difference between the two datasets is that most athletes are in the Geup 9 belt in the secondary dataset, followed by Geup 10 and Geup 8. While in the primary dataset, most athletes are in Geup 10, followed by Geup 8 and Geup 9.

This difference in belt distribution between the two datasets can provide valuable insights for sports management. For instance, understanding these belt distributions can help identify areas for targeted training programs or recruitment efforts. Additionally, the variation in belt rankings could influence the design of marketing strategies, ensuring that resources are allocated to the appropriate athlete segments, thus optimizing training and business operations.

Meanwhile in Fig. 9, we observe similarities and differences in the distribution of athlete residence distances relative to the Club/Dojang in the two datasets. An apparent similarity is that most athletes live <5 km from their Club/Dojang in primary and secondary datasets.

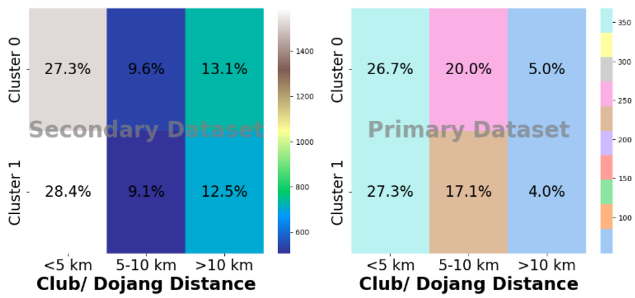

Fig. 9. Club / dojang distance based on secondary dataset.

However, the most notable difference is in the second largest category. In the secondary dataset, the second most common distance is 5-10 km, whereas in the primary dataset, the second largest group lives at a distance of >10 km. This shift in distance distribution between the two datasets highlights potential variations in athlete demographics and accessibility to training facilities, which could be valuable information for optimizing location-based training strategies or marketing efforts to improve participation rates across different distance segments.

C. Discussion

As mentioned before, in this study, the Zack knowledge gap model is integrated into the SECI model to identify gaps between the organization's current knowledge and what it should know. To determine the criteria of "what the firm must know", we adopted brand equity. We use brand equity because marketing practitioners need to understand the level of brand equity in a particular market area [22], which is obtained from customer/consumer responses to the services provided [23]. Fig. 10 shows the new knowledge management model for sports business that we derived from this research.

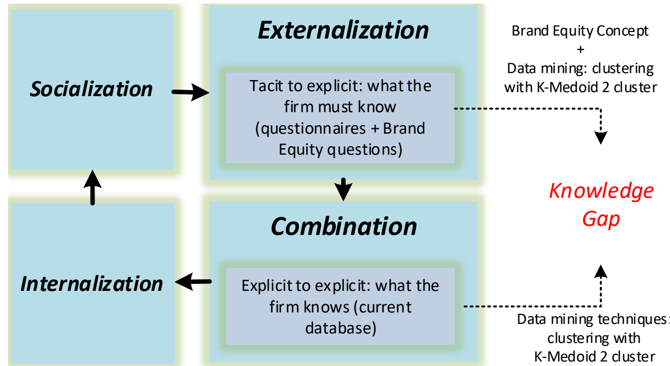

Fig. 10. The new model.

The novelty of this research lies in the externalization and combination stages of the SECI model. We match the externalization in the SECI model with "what the firm must know" in the Zack model and the combination in the SECI model with "what the firm knows" in the Zack model. By merging these models, we aim to provide valuable insights for sports business management, helping organizations to develop effective strategies and improve their competitiveness. Table V shows the additional features of this new model compared to the old SECI model.

TABLE V. COMPARISON OF OLD AND NEW MODELS

Component	Old SECI	New Model
Separating the knowledge between what is already known and what should be known	No	Yes, by using secondary and primary datasets.
Input from the customer's perspective	No	Yes, with the brand equity questionnaire.
Using data mining for data processing	Not mentioned	Yes, with adjustments depending on the needs of the organization.

This new model's advantage is the separation between what is known and what should be known. This model uses customer input to determine what should be known. The results of the separation prove that there is a knowledge gap between the two. This is important to provide awareness for management not to simply rely on what is already known but to update data and get input from customers directly.

We also used clustering data mining techniques to analyze the knowledge gap between both datasets. Based on our research, the K-medoids algorithm with two clusters was optimal on both datasets. For application in other sports businesses besides Taekwondo Club/Dojang and other business sectors besides sports, the management must know the proper data mining techniques and algorithms to apply to their databases to use this new model.

V. Conclusion

This study combines SECI and Zack models, brand equity concepts, and data mining methods to develop a new knowledge management model for sports business services. In this study, we use two datasets to identify knowledge gaps. The novelty of this study is in the externalization and combination of the SECI model. By using this new model, management will gain vital insights that they can use to formulate strategies and improve their competitiveness in the sports services market. To minimize the knowledge gap, sports business management must ensure that their knowledge is the latest data, so they must collect regular data from the athletes. This new model can also be used by other businesses besides sports. In addition to choosing the proper data mining techniques and algorithms, we need to customize the creation of the brand equity questionnaire according to each business's needs. Future work can focus on extending the application of this model to other industries, such as healthcare or retail, to validate its flexibility and effectiveness in different contexts. In addition, further research can explore integrating advanced data mining techniques, such as machine learning algorithms, to improve knowledge gap analysis and increase prediction accuracy.

References

[1] T. Koç, K. Kurt, and A. Akbıyık, "A Brief Summary of Knowledge Management Domain: 10-Year History of the Journal of Knowledge Management," Procedia Comput. Sci., vol. 158, pp. 891–898, 2019, doi: 10.1016/j.procs.2019.09.128.

[2] D. Komaludin, "Model Manajemen Pengelolaan Pusat Olahraga dan Rekreasi Melalui Pengembangan Sumber Daya Manusia Berbasis Pengetahuan (Knowledge Based Worker)," J. Fak. Kegur. Ilmu Pendidik., vol. 1, no. 1, pp. 19–33, 2020.

[3] A. Rot and M. Sobinska, "Challenges for Knowledge Management in Digital Business Models," in 2020 10th International Conference on Advanced Computer Information Technologies (ACIT), IEEE, Sep. 2020, pp. 555–558. doi: 10.1109/ACIT49673.2020.9208867.

[4] M. Firdaus Abdullah, A. Yazid Abu Bakar, and U. Kebangsaan Malaysia Mohamad Nizam Nazarudin, "The Significance of Knowledge, Skills, Capabilities and Competencies in Fitness Centres' Management," Acad. Strateg. Manag. J., vol. 20, no. 6, p. 2021, 2021.

[5] E. Ode and R. Ayavoo, "The mediating role of knowledge application in the relationship between knowledge management practices and firm innovation," J. Innov. Knowl., vol. 5, no. 3, pp. 209–217, 2020, doi: 10.1016/j.jik.2019.08.002.

[6] A. O. Adesina and D. N. Ocholla, "The SECI Model in Knowledge Management Practices," Mousaion South African J. Inf. Stud., vol. 37, no. 3, 2019, doi: 10.25159/2663-659x/6557.

[7] F. P. Sihotang, Ermatita, Samsuryadi, and D. Palupi Rini, "Knowledge Management Model Review in Relation to Marketing and Branding," in 2022 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS), IEEE, Nov. 2022, pp. 490–495. doi: 10.1109/ICIMCIS56303.2022.10017560.

[8] C. M. Sarungu, Titan, D. F. Murad, and Sunardi, "Storing, diving and distributing of comprehensive knowledge using knowledge management in the library and knowledge center," Proc. - 2017 2nd Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICITISEE 2017, vol. 2018-Janua, pp. 288–293, 2018, doi: 10.1109/ICITISEE.2017.8285513.

[9] A. Gupta, P. Bhatia, K. Dave, and P. Jain, "Stock market prediction using data mining techniques," in 2nd International Conference on Advances in Science & Technology (ICAST-2019), Maharashtra,India: University of Mumbai, 2019. [Online]. Available: http://ssrn.com/link/2019-ICAST.html

[10] N. Deepa, J. Sathya Priya, and T. Devi, "Towards applying internet of things and machine learning for the risk prediction of COVID-19 in pandemic situation using Naive Bayes classifier for improving accuracy," Mater. Today Proc., vol. 62, pp. 4795–4799, 2022, doi: 10.1016/j.matpr.2022.03.345.

[11] S. M. Shrestha and A. Shakya, "A Customer Churn Prediction Model using XGBoost for the Telecommunication Industry in Nepal," Procedia Comput. Sci., vol. 215, pp. 652–661, 2022, doi: 10.1016/j.procs.2022.12.067.

[12] Y. Sun, H. Liu, and Y. Gao, "Research on customer lifetime value based on machine learning algorithms and customer relationship management analysis model," Heliyon, vol. 9, no. 2, p. e13384, Feb. 2023, doi: 10.1016/j.heliyon.2023.e13384.

[13] Z. A. Diekson, M. R. B. Prakoso, M. S. Q. Putra, M. S. A. F. Syaputra, S. Achmad, and R. Sutoyo, "Sentiment Analysis for Customer Review: Case Study of Traveloka," in 7th International Conference on Computer Science and Computational Intelligence 2022, Bandung, Indonesia: Elsevier B.V., 2023, pp. 682–690. doi: 10.1016/j.procs.2022.12.184.

[14] N. N. Moon, I. M. Talha, and I. Salehin, "An advanced intelligence system in customer online shopping behavior and satisfaction analysis," Curr. Res. Behav. Sci., vol. 2, no. August, p. 100051, Nov. 2021, doi: 10.1016/j.crbeha.2021.100051.

[15] A. Patel, P. Oza, and S. Agrawal, "Sentiment Analysis of Customer Feedback and Reviews for Airline Services using Language Representation Model," Procedia Comput. Sci., vol. 218, pp. 2459–2467, 2023, doi: 10.1016/j.procs.2023.01.221.

[16] P. Savci and B. Das, "Prediction of the customers' interests using sentiment analysis in e-commerce data for comparison of Arabic, English, and Turkish languages," J. King Saud Univ. - Comput. Inf. Sci., vol. 35, no. 3, pp. 227–237, Mar. 2023, doi: 10.1016/j.jksuci.2023.02.017.

[17] M. Z. Abedin, P. Hajek, T. Sharif, M. S. Satu, and M. I. Khan, "Modelling bank customer behaviour using feature engineering and classification techniques," Res. Int. Bus. Financ., vol. 65, no. October 2022, p. 101913, Apr. 2023, doi: 10.1016/j.ribaf.2023.101913.

[18] R. Komitova, D. Raabe, R. Rein, and D. Memmert, "Time Series Data Mining for Sport Data: a Review," Int. J. Comput. Sci. Sport, vol. 21, no. 2, pp. 17–31, 2022, doi: 10.2478/ijcss-2022-0008.

[19] M. Mandorino, A. J. Figueiredo, G. Cima, and A. Tessitore, "A Data Mining Approach to Predict Non-Contact Injuries in Young Soccer Players," Int. J. Comput. Sci. Sport, vol. 20, no. 2, pp. 147–163, 2021, doi: 10.2478/ijcss-2021-0009.

[20] V. Sarlis and C. Tjortjis, "Sports Analytics: Data Mining to Uncover NBA Player Position, Age, and Injury Impact on Performance and Economics," Inf., vol. 15, no. 4, 2024, doi: 10.3390/info15040242.

[21] S. Wenninger, D. Link, and M. Lames, "Data Mining in Elite Beach Volleyball-Detecting Tactical Patterns Using Market Basket Analysis," Int. J. Comput. Sci. Sport, vol. 18, no. 2, pp. 1–19, 2019, doi: 10.2478/ijcss-2019-0010.

[22] R. Rungtrakulchai, "The effects of brand equity and brand personality on customer value: The case of Uniqlo in Thailand," in Proceedings of 2018 5th International Conference on Business and Industrial Research: Smart Technology for Next Generation of Information, Engineering, Business and Social Science, ICBIR 2018, Bangkok, Thailand: IEEE, 2018, pp. 491–495. doi: 10.1109/ICBIR.2018.8391247.

[23] M. F. Juddi, S. Perbawasari, S. Dida, and A. R. Nugraha, "Branding of Purwakarta Regency Through Gempungan di Buruan Urang Lembur in Health Services," in Proceedings of the Asia-Pacific Research in Social Sciences and Humanities Universitas Indonesia Conference (APRISH 2019), 2021, pp. 154–161. doi: 10.2991/assehr.k.210531.019.

[24] E. Zuraidah, "Knowledge Management System Untuk SDM Menggunakan Seci Model (Studi Kasus: Koperasi Karyawan)," J. Inform., vol. 5, no. 1, pp. 157–168, Apr. 2018, doi: 10.31311/ji.v5i1.2481.

[25] R. Nurcahyo and D. I. Sensuse, "Knowledge Management System Dengan Seci Model Sebagai Media Knowledge Sharing Pada Proses Pengembangan Perangkat Lunak," J. Teknol. Terpadu, vol. 5, no. 2, pp. 63–76, 2019.

[26] F. Ramadhani and M. ER, "A Conceptual Model for the Use of Social Software in Business Process Management and Knowledge Management," Procedia Comput. Sci., vol. 161, pp. 1131–1138, 2019, doi: 10.1016/j.procs.2019.11.225.

[27] M. Ihsan Nugraha and J. S. Suroso, "Designing Knowledge Management System on Seller Education Tokopedia," in 2018 International

Conference on Information Management and Technology (ICIMTech), IEEE, Sep. 2018, pp. 27–32. doi: 10.1109/ICIMTech.2018.8528191.

[28] M. Miao, S. Saide, and D. Muwardi, "Positioning the Knowledge Creation and Business Strategy on Banking Industry in a Developing Country," IEEE Trans. Eng. Manag., pp. 1–9, 2021, doi: 10.1109/TEM.2021.3071640.

[29] A. Tizkar Sadabadi and A. Abdul Manaf, "IKML Approach to Integrating Knowledge Management and Learning for Software Project Management," Knowl. Manag. Res. Pract., vol. 16, no. 3, pp. 343–355, Jul. 2018, doi: 10.1080/14778238.2018.1474165.

[30] N. Norani, H. M. Belal, Q. Hasan, and M. Kosaka, "Knowledge Leadership : Managing Knowledge to Lead Change in PKT Logistics Group," in International Conference on Service Science and Innovation and Serviceology 2018, ICSSI 2018 & ICServ 2018, Taichung, Taiwan, 2019, pp. 156–160. [Online]. Available: https://www.researchgate.net/profile/Kuanlin-Chen/publication/339107708_Embarking_Advanced_Service_on_Servitization_Transformation_Make_Sense_Using_System_Thinking/links/5e3df772a6fdccd96590d854/Embarking-Advanced-Service-on-Servitization-Transformation-M

[31] I. Oktanisa and A. A. Supianto, "Perbandingan Teknik Klasifikasi Dalam Data Mining Untuk Bank Direct Marketing," J. Teknol. Inf. dan Ilmu Komput., vol. 5, no. 5, p. 567, 2018, doi: 10.25126/jtiik.201855958.

[32] D. Triyansyah and D. Fitrianah, "Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing," J. Telekomun. dan Komput., vol. 8, no. 3, pp. 163–182, 2018, doi: 10.22441/incomtech.v8i3.4174.

[33] I. Zulfa, R. Rayuwati, and K. Koko, "Implementasi data mining untuk menentukan strategi penjualan buku bekas dengan pola pembelian konsumen menggunakan metode Apriori (studi kasus: Kota Medan)," Tek. J. Sains dan Teknol., vol. 16, no. 1, pp. 69–82, 2020, doi: 10.36055/tjst.v16i1.7601.

[34] U. Narimawati, D. Munandar, and S. Mauluddin, "The Effectiveness of The Knowledge Management Model for Private Universities' Identity Branding," Cent. Eur. Manag. J., vol. 30, no. 4, pp. 696–704, 2022, doi: 10.57030/23364890.cemj.30.4.64.

[35] R. Sethi and Y. Jangir, "Measurement of Brand Equity: A Significant Aspect of Brand Equity Management," Aayushi Int. Interdiscip. Res. J., vol. 9, no. 1, pp. 9–11, 2022, [Online]. Available: https://www.aiirjournal.com/uploads/Articles/2022/02/5499_04.Dr. Ruhi Sethi & Yasha Jangir.pdf

[36] I. Ghozali, Aplikasi Analisis Multivariate Dengan Program IBM SPSS 26, 10th ed. Semarang: Badan Penerbit Universitas Diponegoro, 2021.

[37] I. Ariyanti, "Uji Validitas dan Reliabilitas Instrumen Angket Kemandirian Belajar Matematik," THETA J. Pendidik. Mat., vol. 1, no. 2, pp. 53–57, 2019.

[38] V. H. Antonius and D. Fitrianah, "Enhancing Customer Segmentation Insights by using RFM + Discount Proportion Model with Clustering Algorithms," Int. J. Adv. Comput. Sci. Appl., vol. 15, no. 3, pp. 902–911, 2024, doi: 10.14569/IJACSA.2024.0150390.

[39] L. Zahrotun, U. Linarti, B. H. T. Suandi As, H. Kurnia, and L. Y. Sabila, "Comparison of K-Medoids Method and Analytical Hierarchy Clustering on Students' Data Grouping," Int. J. Informatics Vis., vol. 7, no. 2, pp. 446–454, 2023, doi: 10.30630/joiv.7.2.1204.

[40] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means algorithm: A comprehensive survey and performance evaluation," Electron., vol. 9, no. 8, pp. 1–12, 2020, doi: 10.3390/electronics9081295.

[41] K. F. Subroto and T. E. Balqiah, "The effects of gendered marketing on brand perception and purchase intention," Contemp. Res. Manag. Bus., pp. 111–114, 2022, doi: 10.1201/9781003295952-29.