Data mining for association analysis of pes planus and pes cavus

Ji-Yong Jung, Jung-Kyu Choi, Yonggwan Won, and Jung-Ja Kim

Abstract**—**Association rule mining is a useful and widely used method to provide beneficial information in the knowledge discovery area related to diagnosis and prediction of disease. In this study, we analyzed the correlation among symptoms occurred between pes planus and pes cavus by using Apriori algorithm. Database based on clinical diagnosis records about the 1310 patient foot with pes planus or pes cavus were constructed. We extracted disease data, including scoliosis, femoral interversion, anterior knee pain, hallux valgus, intoe, outtoe, plantar fasciitis, pelvic malalignment, lower back pain, heel pain, gastrosoleus tightness, forefoot varus, forefoot valgus, quadriceps tightness, leg length discrepancy, and cerebral palsy, with more than one symptom (pes planus or pes cavus) in total data. In case of only one side of the foot was diagnosed with pes planus or pes cavus, data was excluded. SPSS Clementine 11.1 program was used to analyze the correlation among disease. The results show that both pes planus and pes cavus was correlated with plantar fasciitis from explored association rule. Also, we confirmed other symptoms commonly occurred with pes planus or pes cavus from patients' diagnosis data, and there was an organic associative relation between these diseases. These results could be used as a basic data to make secondary foot disease prevention program and to prevent complication of pes planus or pes cavus. In addition, discovered rule can be used for new application with many data set and applied to useful assistant knowledge on clinical decision making process of a specialist. There are some limitation that we performed analysis for only patients who were diagnosed with pes planus and pes cavus, also some data which are not subdivided into foot deformity was included in total data.

Keywords**—**Association rule mining, Pes planus, Pes cavus, Podiatry diagnosis

I. Introduction

The human foot plays an extremely important role in off loading the body's weight to the ground efficiently, maintaining balance, and facilitating propulsion during standing and walking. Biomechanical structure of the foot has to adapt to a changing pattern of loading as the centre of mass of the body during stance phase [1]. Also, foot function which is connected with lower extremity organically during gait is very important to move for human, and it provides propulsion and progress direction.

Generally, the foot shape is classified into 3 types such as pes planus, normal foot, and pes cavus according to the shape of the arch, and these types of the foot can be identified by visual inspection, footprint analysis, radiographic examination, and resting calcaneal stance position (RCSP) [2-3].

Pes planus is defined as a condition in which the medial arch of the foot is diminished or absent as well as calcaneus is everted, allowing the entire sole to touch the ground. Most primary cause of this deformity were known as excessive flexibility in the joint, neuromuscular disease including cerebral palsy, congential disease such as tarsal coalition, injury including fracture, rheumatoid arthritis, and tibialis posterior dysfunction. Pes planus have been associated with metatarsal stress fractures, plantar fasciitis, achilles tendinitis, tibialis anterior inflammation and patella-femoral joint pain [4].

Pes cavus, which is the opposite of flat foot, is characterized by an abnormally high medial longitudinal arch, inverted hindfoot, and adducted forefoot. In high-arch foot, the rigid arch causes the foot to strike down on its lateral side while walking [5]. This foot deformity is mainly affected by Achilles tendon spasticity or changes in the calf muscle and tendons due to habitual wearing of high-heeled shoes. An increased arch height is linked to high risk of lower extremity injuries such as plantar fasciitis and lateral ankle sprains [6]. In children, neurological cause of high-arch foot is spinal dysraphism, cerebral palsy, cerebellar disease, arthrogryposis, and traumatic pes cavus causes [7].

Data mining is data analysis method for discovering new knowledge by exploring hidden rules or patterns in large volumes of data which may contain useful information [8]. In recent years, various methods of data mining such as decision tree, decision rules, logistic regression, artificial neural network, support vector machines (SVM), naïve Bayesian classifier, Bayesian networks, and k-nearest neighbors have been applied to analyze the data in clinical medicine [9-17]. Medical data mining for provide important information about patients is promising filed of computational intelligence. This field has recently been used to find the rule or correlation between various disease and patient attributes [18-22]. Valuable knowledge induced by data mining increase accuracy of diagnosis and treatment about patients [23].

Association rule mining has high potential to discover

Ji-Yong Jung is with the Division of Healthcare Engineering, Chonbuk National University, 567 Baekje-daero, Deokjin-gu, Jeonju-si, Jeollabuk-do, Republic of Korea (e-mail: cholbun@hanmail.net).

Jung-Kyu Choi is with the Division of Healthcare Engineering, Chonbuk National University, 567 Baekje-daero, Deokjin-gu, Jeonju-si, Jeollabuk-do, Republic of Korea (e-mail: chaosjoozak@nate.com).

Yonggwan Won is with the School of Electronics and Computer Engineering, Chonnam National University, 77 Yongbong-ro, Buk-gu, Gwangju, Republic of Korea (e-mail: ykwon@chonnam.ac.kr).

Jung-Ja Kim is with the Department of Biomedical Engineering, Chonbuk National University, 567 Baekje-daero, Deokjin-gu, Jeonju-si, Jeollabuk-do, Republic of Korea (corresponding author to provide phone: 82-63-270-4102; fax: 82-63-270-2247; e-mail: jungjakim@jbnu.ac.kr).

important information related to disease diagnosis and prediction [24-25]. Especially, association rule mining is the most suitable methodology for exploring pair of attributes which provide useful information in nonrestrictive patterns. The advantage of association rule discovery is that it can analyze the correlation of data at the same time by finding many items to set of rules. Therefore, association rule mining has been applied successfully to biomedical informatics [26-30].

Deformity of the foot such as pes planus or pes cavus can cause not only disease came from foot deformity but also complication by biomechanical disorder such as Achilles tendonitis, gastrocnemius spasticity, scoliosis, and plantar fasciitis as it affect negatively to maintain balance among lower extremity, pelvis, trunk, and upper extremity [31-32]. Most patients with foot deformity have more than two deformities, so it is very important to investigate the dynamic relation between diseases. Previous studies have been performed only quantitative analysis based on biomechanical or statistical methods. Also, many studies investigated on cause, characteristic, and treatment of pes planus or pes cavus, while no study so far has conducted about complication in patients with pes planus or pes cavus and correlation of diagnosis based on real clinical data. Accordingly, research is needed to provide various correlations between physical characteristic and disease.

In this study, we analyzed factor affecting disease or dynamic relation between symptoms by applying association rule mining for data processing of patients accompanying with pes planus and pes cavus which occurred with high frequency in foot deformity.

II. Association Rule Mining

Association rule mining is first mentioned in 1993, and numerous methods to generate strong patterns as well as to improve performance of association rules algorithm have been proposed [37]. Main purpose of the association rule mining is to discover the associations or relevant relations among itemsets in large databases [33-38]. Association rule that express transactions occurring concurrently to form of rule means the relation between particular transaction and other transaction appearing simultaneously or progressively when particular transaction occurred. It is defined that condition item as an antecedent of a rule cause consequent item. Given $I = (i_1, i_2,\ldots, i_m)$ as the item's space, which is a set of m distance items (database attributes), let D is a dataset (database) which is a set of transactions and each transaction T_j be defined as a set of items (itemset; subset of I) such that $T_j \subseteq I$ and $T = D = \{T_1, T_2,\ldots, T_3\}$. An association rule can be expressed in the form of $X \rightarrow Y$ where $X \subset I$, $Y \subset I$ and $X \cap Y = \emptyset$.

For example, if the rule such as 'bread → milk, cheese, and ham' (with support = 0.7, confidence = 0.9) was explored using basket analysis with data mining method for supermarket sales database, it means that 'if the bread purchase, milk, cheese, and ham purchase together'. At this time, support and confidence is applied as measure about validity of the rule. Association rule has the two measurements, support and confidence, which is used to determine the significance of an association rule. Support (s) is defined as the frequency that itemsets (X and Y) occur or co-occur in a transaction database D. Confidence (c) represent how "strongly" an itemset X implies another itemset Y [39]. In other words, high support means that the explored rules are occurred frequently in many transactions and these values indicate the importance of that rule in total database. Also, high confidence shows that how to correlate between items which is consist of a prerequisite of the rule in explored rule and other items. In the above example, 0.7 of support indicate probability of bread, milk, cheese, and ham sells together is 70% in customer purchase details of sales database, and 0.9 of confidence indicate that the case of milk, cheese, and ham are also purchased when bread purchased bread is accounts for 0.9. Support and confidence of the rule can be computed by equation (1), and (2).

$$\text{Support}\ (X \rightarrow Y) = \frac{n(X \cup Y)}{|D|} = P\ (X \cup Y) \tag{1}$$

$$\begin{aligned}\text{Confidence}\ (X \rightarrow Y) &= \frac{\text{support}(X \cup Y)}{\text{support}\ (X)} = \frac{n(X \cup Y)}{n(X)} \\ &= \frac{\mathrm{P}\ (X \cup Y)}{\mathrm{P}\ (X)} = \mathrm{P}\ (Y|X)\end{aligned} \tag{2}$$

If it is reinterpreted and defined the point of view for medical diagnosis, support is the percentage of disease A and B occurred together in total disease, and confidence is the percentage of disease B occurred related to disease A. Equations can be expressed by equation (3), and (4).

$$\text{Support (\%)} = \frac{\text{Number of disease A} \cap \text{B}}{\text{Total number of disease}} \tag{3}$$

$$\text{Confidence (\%)} = \frac{\text{Number of disease A} \cap \text{B}}{\text{Total number of disease A}} \tag{4}$$

As following this, the meaning of association rule mining is to interpret qualitative significance to understand easily interconnections between data and quantitative significance as scale of support and confidence about explored rule by indicating many data occurred concurrently to set of explored rules [38,40].

A. Apriori algorithm

Apriori which is efficient algorithm for finding rules have extremely high confidence from transactional databases [41]. It is one of the most widely used algorithm in the data mining literature. Also, it can improve the performance by reducing the number of database scans and the size of the analyzed dataset at each scan [42-44].

Apriori algorithm consists of two steps as follow. In step 1, the Apriori algorithm determines all frequent itemsets by scanning the transaction database. Subset of total set of item I as well as assemblage composed of several item is called itemsets and other else itemsets are called frequent small itemsets.

Database D — Scan D ⇒

Tid	Itemsets
100	1 3 4
200	2 3 5
300	1 2 3 5
400	2 5

C_1

Itemsets	SD
{1}	2
{2}	3
{3}	3
{4}	1
{5}	3

L_1

Itemsets	SD
{1}	2
{2}	3
{3}	3
{5}	3

C_2 — Scan D ⇒

Itemsets
{1 2}
{1 3}
{1 5}
{2 3}
{2 5}
{3 5}

C_2

Itemsets	SD
{1 2}	1
{1 3}	2
{1 5}	1
{2 3}	2
{2 5}	3
{3 5}	2

L_2

Itemsets	SD
{1 3}	2
{2 3}	2
{2 5}	3
{3 5}	2

C_3 — Scan D ⇒

Itemsets
{2 3 5}

C_3

Itemsets	SD
{2 3 5}	2

L_3

Itemsets	SD
{2 3 5}	2

Support = P(X ∩ Y) — Confidence = P(Y | X)

Rules	Support (X Y)	Support (X)	Confidence
{1} → {3}	2	2	100
{2} → {3}	2	3	66.66666667
{2} → {5}	3	3	100
{3} → {5}	2	3	66.66666667
{2} → {3 5}	2	3	66.66666667
{3} → {2 5}	2	3	66.66666667
{5} → {2 3}	2	3	66.66666667
{3} → {1}	2	3	66.66666667
{3} → {2}	2	3	66.66666667
{5} → {2}	3	3	100
{5} → {3}	2	3	66.66666667
{3 5} → {2}	2	2	100
{2 5} → {3}	2	3	66.66666667
{2 3} → {5}	2	2	100

Fig. 1 Apriori algorithm

In step 2, frequent itemstes are used to generate association rule from database. And subsets about all frequent itemsets (not null sets) detected. If percentage of support (*f*) about support (*a*) of subset is over mini minimum confidence (*Cmin*), rules will be generated as follows (5).

$$\frac{\text{Support }(f)}{\text{Support }(a)} \geq Cmin, a \rightarrow (f - a) \tag{5}$$

Discovery process of Apriori repeat the process to find out L_k which satisfying the condition of minimum support by discovering *k*-candidate itemset Ck from (*k-1*)-frequent itemset L_k-1 for search *k*-frequent itemset (L_k) and by calculating support of C_k. Fig. 1 shows the example of execution process for Apriori with assuming 40% of minimum support.

First, calculate support of single item using scanned database *D*, and then determine 1-large itemset L_1 satisfying minimum support in 1-candidate itemset C_1. In Fig. 1, {(1),(2),(3),(5)} were determined to L_1 as their minimum support were bigger than 40%. Large itemsets were discovered such repetition of process.

Characteristic of Apriori algorithm is the process called Apriori-gen which composing of (*k+1*)-candidate itemset C_{k+1} from frequent itemset Lk in any stage as shown in Fig. 1. Process of Apriori-gen and prune is same as following this. For example, we assumed that itemsets which is included in 2-large itemset L_2 are {1 2}, {1 3}, {2 3}, {2 5}, {2 8}. Apriori-gen generate candidates of 3-candidate itemset C_3 by adding item which have last item with more larger number value in itemsets included in L_2. That is, 3-candidate itemset {1 2 3} is generated by connecting each other due to {1 2} and {1 3} is equal in first itme '2' of {1 2}. Also, {2 3 5} and { 2 5 8} is generated as same as above method. However, {2 3 5} and {2 5 8} are not included in L_3 as 2-item subset {3 5} of {2 3 5} and {5 8} of {2 5 8} are not included in L_2, but only {1 2 3} which is included in L_2 is determined to C_3 and then itemset satisfying above conditions is determined to L_3 by calculating support of candidate itemsets [45-46]. The pseudocode about Aprori-gen and prune algorithm was written as shown in Fig. 2.

```
{Apriori-gen}
insert into C_{k+1}
select p.item_1,p.item_2, ..., p.item_k, q.itemset_k
from L_k p, L_k q
where p.item_1 = p.item_1, ..., p.item_{k-1} = q.itemset_{k-1}, p.item_k < q.itemset_k

{Prune step}
forall itemset c ∈ C_{k+1} do
forall k-subsets s of c do
if (s∉L_k) then
delete c from C_{k+1}
end
```

Fig. 2 Apriori algorithm

III. Experimental Method

A. Biomechanical Evaluation Process

Foot structure can be assessed with a variety of clinical and research tools [46-47]. The main objective of this biomechanical evaluation is to investigate the effect of foot deformity such as pes planus or pes cavus on body function. As shown in Fig. 3, clinical measurement are consists of prone, supine, and standing evaluation. And clinical tests are performed using various biomechanical measurement tools, including digital angulometer, digital angle finder, gravity goniometer, and tractography.

(a)

(b)

(c)

(d)

Fig. 3 (a) Prone evaluation, (b) Supine evaluation (c) Standing evaluation, (d) Measurement tools

First, passive inversion and eversion available angle, subtalar neutral joint position of rearfoot/forefoot, and dorsiflexion of ankle were measured in prone evaluation. Second, in supine evaluation, malleolar torsion and leg length differences, midtarsal joint integrity, and great toe range of motion (ROM) into extension were determined. And, finally, standing evaluation is conducted using weight-bearing measurement. The resting calcaneal stance position (RCSP) which is an index of various foot structural and functional abnormalities can provide information on self selected foot postures. Bisected angle of the posterior aspect of the calcaneus is used to this biomechanical evaluation process [48]. Subjects' feet were characterized as pes planus or pes cavus based on RCSP. A neutrally aligned foot had an RCSP between 2° of inversion and 2° of eversion, while pes planus or pes cavus foot had an RCSP of more than or equal to 4° of eversion and inversion, respectively [49-50]. This clinical measurement method is used to the point of view in frontal plane relationship of the forefoot to the rearfoot. The subtalar joint neutral position of each foot was also determined but not used to categorize foot type.

B. Subjects

In this study, we constructed database using diagnosis data about left and right side of the foot (2620 feet) of 1310 patients (men : 633, women : 677) from C foot clinic hospital. Fig. 4 shows a diagnosis charts of patients. Subjects have more than one disease in left and right side of the foot such as pes planus, pes cavus, scoliosis, femoral interversion (Fem, Inter), anterior knee pain, hallux valgus (HAV), intoe, outtoe, plantar fasciitis, pelvic malalignment, lower back pain (LBP), heel pain, gastrosoleus tightness, forefoot varus, forefoot valgus, quadriceps tightness, leg length discrepancy (LLD) and cerebral palsy (CP).

Fig. 4 Diagnosis chart of patients

C. Methods

Process of analysis in patients with pes planus and pes cavus was carried out as shown in Fig. 5. Diagnosis information of all patients (2620 feet) with more than one disease including pes planus or pes cavus were extracted without identifying before and after being diagnosed based on initial diagnosis in total data. And data warehouse was constructed to present information after data cleaning which is essential procedure for the quality of rules derived from the mining process. Then, Apriori algorithm was used to discover rules in diagnosis data. In case of only one side of the foot was diagnosed with pes planus or pes cavus, it was excluded.

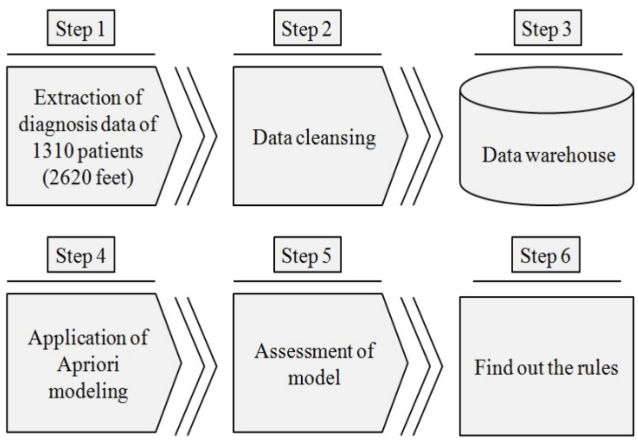

Fig. 5 Association rule mining process

To analyze diagnosis data of patients, Apriori algorithm and web node was applied using Clementine 11.1 (SPSS Inc., Chicage, USA) [49,51]. This software use the node which is connected together to form a stream frame for data processing. And, web node provide a graph by visualizing the relation aspects of items and Apriori algorithm is applicable in case of dichotomous and polytomous [52].

IV. Results

As a result of extracted data from patients with more than one disease including pes planus or pes cavus in left and right foot on all of 1310 patients, 642 patients (436 males and 206 females) records are used to this study.

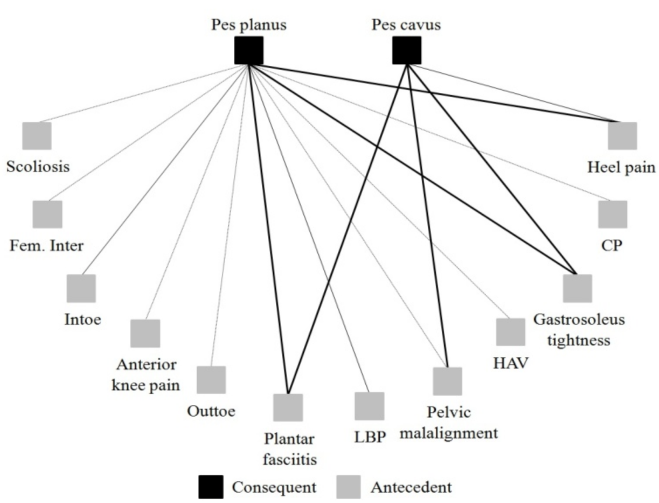

Fig. 6 Result of association rules by Web node

Result of graphic generated using web node about diagnosis record of patients with pes planus or pes cavus is shown as in Fig. 6. Diseases correlated with other symptoms are connected the line respectively and displayed thickly in case of more highly related.

In case of pes planus, it was related to disease like scoliosis, Fem. Inter, anterior knee pain, HAV, intoe, outtoe, plantar fasciitis, pelvic malalignment, LBP, heel pain, gastrosoleus tightness, and CP. Most of all, relation of plantr fasciitis, heel pain, and gastrosoleus tightness was the most highest. And in case of pes cavus, it was related to disease such as plantar fasciitis, pelvic malalignment, heel pain, and gastrosoleus tightness. And the relation of pelvic malalignment was the most highest among others. All of the diseases on pes cavus, especially, are interconnected with pes planus.

Table 1 show rules with the reliability of more than 95% confidence in association rules detected in patients with pes planus and pes cavus by applying Apriori algotirhm. Pes planus was related with gastrosoleus tightness, plantar fasciitis, heel pain, LBP, and HAV while pes cavus was related with gastrosoleus tightness and pelvic malalignment. Support was high (34.36%) in case of pes planus with plantar fasciitis and in more than 20% in case of pes planus with gastrosoleus tightness. In this study, confidence was used to rank the discovered rules in Apriori results [53-56].

Table 1. Results of discovered association rule

Consequent	Antecedent	Support (%)	Confidence (%)
Pes planus	Gastrosoleus tightness	25.17	98.29
Pes planus	Plantar fasciitis	34.36	99.24
Pes planus	Heel pain	23.87	98.57
Pes planus	LBP	13.28	95.62
Pes planus	HAV	15.52	96.37
Pes cavus	Gastrosoleus tightness	13.26	98.64
Pes cavus	Pelvic malalignment	15.63	99.25

V. Discussion

Foot has a multi joint mechanism which plays an important role between lower limb and ground during gait. Impairment caused by injury or deformation of the foot affects on muscle, bone, and tendon as well as induces atrophy, instability, and motor disturbance [57]. Deformation of medial longitudinal arch such as pes planus or pes cavus can cause its deformity as well as complications like metatarsal stress fracture, Achilles tendinitis, tibialis anterior inflammation, patella-femoral joint pain by biomechanical disorder in lower extremity. Also, low or high medial arch can cause the chronic plantar fasciitis and influence the severity of heel pain [58]. The plantar fascia is the fibrous connective tissue which runs from the tuberosity of the calcaneus forward to the proximal phalanges. It contributes to support the longitudinal medial arch of the foot and prevent injuries by absorbing the ground shock while walking or running. Plantar fasciitis, that is inflammatory disease to induce pain on plantar fascia, is occurred when excessive weight was overloaded in tiny plantar fascia [59].

In this study, we used Apriori algorithm which is the most suitable methodology for discovering important information related to disease diagnosis to the patients database. And, biomechanical evaluation results on left and right side of the foot through prone, supine, and standing test were included to improve the accuracy and reliability of the association analysis. From the results in this study, not only pes planus which has diminished or absented medial arch but also pes cavus which has abnormally high medial arch of the foot, we confirmed that both pes planus and pes cavus has more higher association with plantar fasciitis than other symptoms such as scoliosis, fem. Int, intoe, outtoe, anterior knee pain, LBP, HAV, CP. In addition, we also determined other correlation between symptoms commonly occurred such as pelvic malalignment, heel pain, and gastrosoleus tightness in pes planus and pes cavus.

Various foot structures are associated with foot function during static or dynamic movement. Most patients with pes planus or pes cavus had more than two symptoms. It means that pes planus or pes cavus could be contributed to change the foot structure and it can cause the serious foot problems. Previous studies reported that foot type categorized as planus and cavus is considered a risk factor in the development of severe foot deformity [60-64]. It is very important to prevent the degeneration as well as complication of disease during patients' treatment process. Appropriate treatment strategy for the individual patient could be considered by understanding the cause of the deformity and relationship between diseases. Rules extracted from apriori algorithm in this study can be utilized to provide the correlation between diseases and assist the clinical diagnosis process related to not only pes planus and pes cavus but also other foot deformities by including diagnosis, screening, prognosis, monitoring, treatment method information.

VI. Conclusion

Association rule mining is efficient and applicable method to discover the association or relevant relations among itemset in large medical databases. In this study, we evaluated the correlation of inter-symptoms which accompanying pes planus and pes cavus using Apriori algorithm which has been successfully employed to extract patterns in the correlation analytic technique of data mining. Diagnosis data of 1310 patients with more than one disease including pes planus or pes cavus were extracted based on initial diagnosis record in total data. And, then apriori algorithm

These results could be used as a basic data to make secondary foot disease prevention program and to protect complication of pes palnus or pes cavus. Finally, discovered rule can be applied effectively to wide range of useful assistant knowledge on decision making process of a specialist.

This study has some limitation. We performed analysis for only patients who were diagnosed with pes planus and pes cavus at a single institution. Also, some data which are not subdivided into pes planus and pes cavus is included in total data. Accordingly, in further research, we will extract a rule which is able to understand and normalized the correlation of pes planus and pes cavus as well as complications occurred complexly by collecting more diagnosis data and analyzing disease pattern according time's flow.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technolgy (2012-3003952).

References

[1] R. J. Abboud, "(i) Relevant foot biomechanics,"*Current Orthopaedics,* vol. 16, no. 3, pp. 165-179, Jun. 2002.
[2] W. H. Lee and S. W. Lee, "A study of the relationship between normal adults resting calcaneal stance position and postural sway," *Journal of Korean Physical Therapy Association,* vol. 11, no. 2, pp. 5-17, Mar. 2004.
[3] J. Burns, J. Crosbie, A. Hunt, and R. Ouvrier, "The effect of pes cavus on foot pain and plantar pressure," *Clinical Biomechanics,* vol. 20, no. 9, pp. 877-882, Nov. 2005.
[4] A. E. Hunt and R. M. Smith, "Mechanics and control of the flat versus normal foot during the stance phase of walking," *Clinical Biomechanics,* vol. 19, no. 4, pp. 391-397, May. 2004.
[5] M. Eslami, C. Tanaka, S. Hinse, M. Anbarian, and P. Allard, "Acute effect of orthoses on foot orientation and perceived comfort in individuals with pes cavus during standing," *Foot,* vol. 19, no. 1, pp. 1-6, Jul. 2009.
[6] D. W. Carson, G. D. Myer, T. E. Hewett, R. S. Heidt, and K. R. Ford, "Increased plantar force and impulse in American football players with high arch compared to normal arch," *Foot,* vol. 22, no. 4, pp. 310-314, Dec. 2012.
[7] J. H. Kim, I. S. Park, B. O. Kim, Y. S. Yang, Y. G. Won, and J. J. Kim, "Podiatric clinical diagnosis using decision tree data mining, *Journal of Korean Institute of Electrical Engineers,* vol. 48, no. 2, pp. 28-37, Mar. 2011.
[8] G. Richards, V. J. Rayward-Smith, P. H. Sonksen, S. Carey, and C. Weng, " Data mining for indicators of early mortality in a database of clinical records," *Artificial Intelligence in Medicine,* vol. 22, no. 3, pp. 215-231, Jun. 2001.
[9] G. Vannozzi, A. Cereatti, C. Mazza, F. Benvenuti, and U. Della Croce, "Extraction of information on elder motor ability from clinical and biomechanical data through data mining, *Computer Methods and Programs in Biomedicine,* vol. 88, no. 1, pp. 85-94, Oct. 2007.
[10] T. Imamura, S. Matsumoto, Y. Kanaqawa, B. Tajima, S. Matsuya, M. Furue, and H. Oyama, "A technique for identifying three diagnostic findings using association analysis," *Medical and Biological Engineering and Computing,* vol. 45, no. 1, pp. 51-59, Jan. 2007.
[11] R. Bellazzi and B. Zupan, "Predictive data mining in clinical medicine: Current issues and guidelines," *International Journal of Medical Informatics,* vol. 77, no. 2, pp. 81-97, Feb. 2008.
[12] F. E. Harrell, "*Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis,* Springer, New York, 2001.
[13] D. W. Hosmer and S. Lemeshow, "*Applied Logistic Regression, 2nd edition*", Wiley, New York, 2000.
[14] G. Schwarzer, W. Vach, and M. Schumacher, "On the misuses of artificial neural networks for prognostic and diagnostic classification in oncology," *Statistics in Medicine*, vol. 19, no. 4, pp. 541-561, Feb. 2000.
[15] N. Crisianini and J. Shawe-Taylor, "*An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*," Cambridge University Press, Cambridge, UK, New York, 2000.
[16] I. Kononenko, "Machine learning for medical diagnosis: history, state of the art and perspective," *Artificial Intelligence in Medicine*, vol. 23, no. 1, pp. 89-109, Aug. 2001.
[17] E. H. Herskovits and J. P. Gerring, "Application of a data-mining method based on Bayesian networks to lesion-deficit analysis," *Neuroimage*, vol. 19, no. 4, pp. 1664-1673, Aug. 2003.

[18] Z. Huang, J. Li, H. Su, G. S. Watts, and H. Chen, "Large-scale regulatory network analysis from microarray data: Modified Bayesian network learning and association rule mining," *Decision Support System*, vol. 43, no. 4, pp. 1207-1225, Aug. 2007.

[19] H. Ishibuchi, I. Kuwajima, and Y. Nojima, "Prescreening of candidate rules using association rule mining and pareto-optimality in genetic rule selection," *in Proceeding of KES '07 Knowledge-Based Intelligent Information and Engineeirng Systems,* Santiago, Chile, 2007, pp. 509-516, 2007.

[20] M. Karabatak and M. C. Ince, "An expert system for detection of breast cancer based on association rules and neural network," *Expert Systems with Application,* vol. 36, no. 2, pp. 3465-3469, Mar. 2009.

[21] A. M. Shin, I. H. Lee, G. H. Lee, H. J. Park, H. S. Park, K. I. Yoon, J. J. Lee, and Y. N. Kim, "Diagnostic analysis of patients with essential hypertension using association rule mining," *Healthcare Informatics Research,* vol. 16, no. 2, pp. 77-81, Jun. 2010.

[22] Z. Wang and W. Hoy, "Is the Framingham coronary heart disease absolute risk function applicable to Aboriginal people," *Medical Journal of Australia,* vol. 182, no. 11, pp. 597-598, Jun. 2005.

[23] Z. S. Zubi and R. A. Saad, "Using some data mining techniques for early diagnosis of lung cancer," *in Proceeding of the 10th WSEAS Internal Conference on Artificial intelligence, knowledge engineering and data bases,* Cambridge, UK, 2011, pp. 32-37.

[24] W. T. Hwang and D. S. Kim, "Improved association rule mining by modified trimming," *Journal of Korean Institute of Electrical Engineers,* vol. 45, no. 3, pp. 15-21, Jun. 2008.

[25] C. Ordonez, "Association rule discovery with the train and test approach for heart disease prediction," *IEEE Transaction Information Technology in Biomedicine,* vol. 10, no. 2, pp. 334-343, Apr. 2006.

[26] H. Fu, "Cluster analysis and association analysis for the same data," *in Proceeding of 7th WSEAS Internation Conference of Artificial intelligence, knowledge engineering and data bases,* Cambridge, UK, 2008, pp. 576-581.

[27] L. Lei, M. Cui, and Z. Mi, "Study on association rule mining in multi-component Chinese medicine research and development for treatment of lung cancer," *China Journal of Chinese Materia Medica,* vol. 35, no. 16, pp. 2192-2195, Aug. 2010.

[28] C. Ordonez, E. Omiecinski, L. de Braal, C. A. Santana, N. Ezquerra, J. A. Taboada, D. Cooke, E. Krawczynska, and E. V. Garcia, "Mining constrained association rules to predict heart disease," *IEEE International Conference on Data Mining,* USA, 2001, pp. 433-440.

[29] D. Fu, S. Zhou, and P. Guo, "Research on a distributed network intrusion detection system based on association rule mining," *in Proceeding of the 1st International Conference on Information Science and Engineering (ICISE),* Nanjing, China, 2010, pp. 1816-1818.

[30] C. Chen, S. Mabu, K. Shimada, and K. Hirasawa, "Network intrusion detection using class association rule mining based on genetic network programming," *IEEE Transaction on Electrical and Electronic Engineering*, vol. 5, no. 5, pp. 553-559, Sep. 2010.

[31] Y. C. Chen, S. Z. Lou, C. Y. Huang, and F. C. Su, "Effects of foot orthoses on gait patterns of flat feet patients," *Clinical Biomechanics,* vol. 25, no. 3, pp. 265-270, Mar. 2010.

[32] G. Solis, M. S. Hennessy, and T. S. Saxby, "Pes cavus: a review," *Foot and Ankle Surgery,* vol. 6, no. 3, pp. 145-153, Jul. 2000.

[33] A. Savasere, E. Omiencinsky, and S. Navathe, "An efficient algorithm for mining association rules in large databases," *in Proceeding of the 21st International Conference on Very Large Data Bases (VLDB'95),* Zurich, Switzerland, 1995, pp. 432-434.

[34] R. SriKant and R. Agrawal, "Mining generalized association rules," *in Proceeding of the 21st International Conference on Very Large Data Bases (VLDB'95),* Zurich, Switzerland, 1995, pp. 407-419.

[35] E. H. Han, G. Darypis, and V. Kumar, "Scalable parallel data mining for association rules," *in Proceeding of ACM SIGMOD International Conference on Management of data,* Montreal, Canada, 1997, pp. 277-288.

[36] J. Han and Y. Fu, "Discovery of multiple-level association rules from large databases," *in Proceeding of the 21st International Conference on Very Large Data Bases (VLDB'95),* Zurich, Switzerland, 1995, pp. 420-431.

[37] N. Pasquier, Y. Bastide, R. Taouil, and L. Lakhal, "Efficient mining of association rules using closed itemset lattices," *Information Systems,* vol. 24, no. 1, pp. 25-46, Mar. 1999.

[38] Y. C. Lee, T. P. Hong, and W. Y. Lin, "Mining association rules with multiple minimum supports using maximum constraints," *International Journal of Approximate Reasoning,* vol. 40, no.11-12, pp. 44-54, Nov. 2005.

[39] J. J. Kim, H. T. Ceong, and Y. G. Won, "Weighted Association rule mining for item groups with different properties and risk assessment for networked systems," *IEICE Transaction on Information and Systems,* vol. E92.D, no. 1, pp. 10-15, Jan. 2009.

[40] K. Y. Jung, "Optimal associative neighborhood mining using representative attribute," *Journal of Korean Institute of Electrical Engineers,* vol. 43, no. 4, pp. 50-57, Jul. 2006.

[41] M. N. Moreno, S. Segrera, V. F. Lopez, and M. J. Polo, "A method for mining quantitative association rules," in *Proceeding of the 6th WSEAS International Conference on Simulation, Modeling and Optimization*, Lisbon, Portugal, 2006, pp. 173-178.

[42] B. Tunc and H. DAG, "Generating classification association rules with modified apriori algorithm," in *Proceeding of the 5th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases,* Madrid, Spain, 2006, pp. 384-387.

[43] M. Craus and A. Archip, "A generalized parallel algorithm for frequent itemset mining," *in Proceeding of the 12th WSEAS International Conference on Computers,* Kos Island, Greece, 2008, pp. 520-523.

[44] R. J. Kuo and C. W. Shih, "Association rule mining through the ant colony system for National Health Insurance Research Database in Taiwan," *Computer and Mathematics with Applications,* vol. 54, no. 11, pp. 1303-1318, Dec. 2007.

[45] R. Agrawal, T. Imielinski, and A. Swami, "Mining association rules between sets of items in large database," *Proceeding of ACM SIGMOD International Conference on Management of Data*, Washington, USA, 1993, pp. 207-216.

[46] M. B. Pohl and L. Farr, "A comparison of foot arch measurement reliability using both digital photography and caliper method," *Journal of Foot and Ankle Research,* vol. 14, no. 3, pp. 14, Jul. 2010.

[47] N. Yalcin, E. Esen, U. Kanatli, and H. Yetkin, "Evaluation of the medial longitudinal arch: a comparison between the dynamic plantar pressure measurement system and radiographic analysis," *Acta Orthop Traumatol Turc,* vol. 44, no. 3, pp. 241-245, Jun. 2010.

[48] R. Agrawal and R. Strikant, "Fast algorithm for mining association rules," *the Proceeding of the 20st International Conference on Very Large Data Bases (VLDB'94),* Santiago, Chile, 1994, pp. 487-499.

[49] J. Song, H. J. Hillstrom, D. Secord, and J. Levitt, "Foot type biomechanics. comparisoin of planus and rectus foot types," *Journal of the American Podiatric Medical Association,* vol. 86, no. 1, pp. 16-23, Jan. 1996.

[50] W. R. Ledoux and H. J. Hillstrom, "The distributed plantar vertical force of neutrally aligned and pes planus feet," *Gait Posture,* vol. 15, no. 1, pp. 1-9, Feb. 2002.

[51] C. Payne and M. Richardson, "Changes in the measurement of neutral and relaxed calcaneal stance positions with experience," *Foot,* vol. 10, no. 2, pp. 81-83, Jun. 2000.

[52] J. Hu, "Data mining using Clementine," Koryo Information Industry,

[53] S. H. Liao and C. H. Wen, "Artificial neural networks classification and clustering of methodologies and applications – literature analysis from 1995 to 2005," *Expert Systems with Applications,* vol. 32, no. 1, pp. 1-11, Jan. 2007.

[54] S. Mutter, M. Hall, and E. Frank, "Using classification to evaluate the output of confidence-based association rule mining," *in Proceeding of the 17th Australian Joint Conference on Advances in Artificial Intelligence,* Cairns, Australia, 2004, pp. 538-549.

[55] W. Taihua and G. Fan, "Associating IDS alerts by an improved apriori algorithm," *in Proceeding of the 3rd Internation Symposium on Intelligent Information Technology and Security Informatics,* Jinggangshan, China, 2010, pp. 478-482.

[56] S. Y. Kim and H. K. Choi, "Biomechanical analysis of the human foot by using passive elastic characteristics of joints," *Journal of the Korean Society of Precision Engineering,* vol. 21, no. 1, pp. 197-204, Jan. 2004.

[57] J. Nahar, T. Imam, K. S. Tickle, and Y. P. Chen, "Association rule mining to detect factors which contribute to heart disease in males and females," *Expert Systems with Applications,* vol. 40, no. 4, pp. 1086-1093, Mar. 2013.

[58] R. Puttaswamaiah and P. Chandran, "Degenerative plantar fasciitis: A review of current concepts," *Foot,* vol. 17, no. 1, pp. 3-9, Mar. 2007.

[59] J. E. Taunton, M. B. Ryan, D. B. Clement, D. C. McKenzie, and D. R. Lloyd-Smith, "Plantar fasciitis: a retrospective analysis of 267 cases," *Physical Therapy in Sport,* vol. 3, no. 2, pp. 57-65, May, 2002.

[60] H. J. Hillstrom, J. Song, A. P. Kraszewski, J. F. Hafer, R. Mootanah, A. B. Dufour, B. S. Chow, and J. T. Deland, "Foot type biomechanics part 1:

structure and function of the asymptomatic foot," *Gait Posture*, vol. 37, no. 3, pp. 445-451, Oct. 2012.

[61] H. K. Sugathan and D. A. Sherlock, "A modified Jones procedure for managing clawing of lesser toes in pes cavus: long-term follow-up in 8 patients," *Journal of Foot and Ankle Surgery*, vol. 48, no. 6, pp. 637-641, Sep. 2009.

[62] J. C. Levy, M. S. Mizel, L. S. Wilson, W. Fox, K. McHale, D. C. Taylor, and H. T. Temple "Incidence of foot ankle injuries in West Point cadets with pes planus compared to the general cadet population," *Foot and Ankle Internatioanl*, vol. 27, no. 12, pp. 1060-1064, Dec. 2006.

[63] K. R. Kaufman, S. K. Brodine, R. A. Shaffer, C. W. Johnson, and T. R. Cullison, "The effect of foot structure and range of motion on musculoskeletal overuse injuries," *The American Journal of Sports Medicine*, vol. 27, no. 5, pp. 585-593, Oct. 1999.

[64] W. R. Ledoux, J. B. Shofer, J. H. Ahroni, D. G. Smith, B. J. Sangeorzan, and E. J. Boyko, "Biomechanical differences among pes cavus, neutrally aligned, and pes planus feet in subjects with diabetes," *Foot and Ankle International*, vol. 24, no. 11, pp. 845-850, Nov. 2003.

Ji-Yong Jung He received the B. S., M. S. degree from the Chonbuk National University, Republic of Korea, in 2010 and 2012, respectively. He is currently pursuing a doctorate in biomedical informatics at the university. His major research interest is the medical informatics, podiatry, and biomechanics.

Jung-Kyu Choi He received the B. S., M. S. degree from the Chonbuk National University, Republic of Korea, in 2011 and 2013, respectively. He is currently pursuing a doctorate in foot biomechanics at the university. His major research interest is the podiatry, and biomechanics.

Yonggwon Won He received the B. S. in Electronics Engineering from Hanyang Univesity in 1987, and M. S. and Ph. D. degrees in Electrical and Computer Engineering from University of Missouri-Columbia in 1991 and 1995, respectively. He worked with Electronics, and Telecommunication Research Institute (ETRI) from 1995 to 1996, and Korea Telecomm (KT) from 1996 to 1999. He is currently a professor in Chonnam National University in Korea, and the director of Korea Bio-IT Foundary Center at Gwangju. His major research interest is the computational intelligence for image analysis, pattern recognition, network and communication security, bio and medical data analysis.

Jung-Ja Kim She received the B.S., M.S. degree in 1985, 1988 and Ph.D. degree from 1997 to 2002, in Computer Science from Chonnam National University respectively. She worked with electronic telecommunication Laboratory at Chonnam National University from 2002 to 2004, and Korea Bio-IT Foundry Center at Gwangju from 2004 to 2006. She is currently an assistant professor at Chonbuk National University. Her major research interet is the bio and medical data analysis, database security, and pattern recognition.