林育志、黃正魁 (2022)，「整合主客視角的新穎協作推薦模型」，*資訊管理學報*，第二十九卷，第二期，頁 103-132。

整合主客視角的新穎協作推薦模型

林育志*
國立中正大學企業管理學系

黃正魁
國立中正大學企業管理學系

摘要

本研究利用項目為基礎的協同過濾想法，提出一種新穎的協作推薦模型。模型根據主觀查詢和客觀規則，透過關聯法則和相似度演算生成推薦結果。並於參與者使用協作推薦系統後，藉由用戶體驗問卷量測使用者對模型的感知有用性、信任度和滿意度。我們以台灣 50 成分股作為實驗標的來收集真實數據集。根據研究結果，新穎協作推薦模型（系統）呈現出更高的感知有用性、信任度和滿意度。

關鍵詞：協作推薦系統、關聯法則、相似度、主觀視角、客觀視角

*本文通訊作者。電子郵件信箱：jokada621@gmail.com

2021/09/06 投稿；2021/11/25 修訂；2021/12/25 接受

Lin, Y.C. & Huang, T.C.K. (2022). Novel Cooperative Recommendation Model from Subjective and Objective Perspectives. *Journal of Information Management*, 29(2), 103-132.

Novel Cooperative Recommendation Model from Subjective and Objective Perspectives

Yu Chih Lin*
Department of Business Administration, National Chung Cheng University

Tony, Cheng Kui Huang
Department of Business Administration, National Chung Cheng University

Abstract

This study utilizes the idea of item-based collaborative filtering to propose a novel cooperative recommendation model. The model adopts the technique of association rule mining and the similarity computation algorithm to generate recommendations from subjective inquiries and objective rules. In addition, a user-experience questionnaire is conducted to measure the perceived usefulness, trust, and satisfaction after participants use the cooperative recommendation system. The experiment adopts the shares from the Taiwan Top50 Exchange Tracker Fund (ETF50) as recommendation items to collect our real-life dataset. According to the result, the novel cooperative recommendation model (system) presents higher perceived usefulness, trust, and satisfaction.

Keywords: Cooperative recommendation system, Association rule mining, Similarity, Subjective perspective, Objective perspective

* Corresponding author. Email: jokada621@gmail.com

2021/09/06 received; 2021/11/25 revised; 2021/12/25 accepted

1. Introduction

People are used to making sensible decisions when facing various problems in work or daily life. As Web 2.0 technology matures and develops, people share and discuss their concerns or interests through online media and platforms, such as social media, blogs, and forums. Such information can be collected and analyzed by a recommender system to help people promptly identify suitable or interesting suggestions. Recommender systems are a helpful alternative since they help users discover items they might not have found otherwise. Currently, recommender systems (Bauer & Nanopoulos 2014; Li, Wu, & Lai 2013; Lu et al. 2015) can be categorized as content-based systems and collaborative filtering systems. The most significant difference between content-based and collaborative filtering systems is that they utilize different user preferences as the calculating basis for recommending items. Balabanovic & Shoham (1997) pointed out that content-based systems recommend similar items according to users' previously browsed or favored items. Collaborative filtering systems recommend clusters of users with similar preferences by forming a neighborhood of purchasing or evaluating (Herlocker et al.1999; Papagelis & Plexousakis 2005; Wang, de Vries & Reinders 2006). The disadvantage of content-based systems is that recommendation is based on users' previous purchasing behavior. If no previous behavioral data is available, a content-based system cannot provide any recommendations for other users. Therefore, scholars have proposed the collaborative filtering approach to solve this problem; that is, the similarity calculations of content-based systems only refer to individual user data.

Collaborative filtering assumes that people are willing to make choices about similar items in the future based on past decisions. The most crucial process of a collaborative filtering model is to determine the similarity predictions for items or users. Thus, the methods of collaborative filtering systems are mainly categorized as user-based collaborative filtering (Konstan et al. 1997) and item-based collaborative filtering

(Lang 1995; Sarwar et al. 2001). The former involves calculating similar preferences or interests of neighbor users; hence, a user-based collaborative filtering system is also called a "neighborhood-based collaborative filtering system". The problem with user-based collaborative filtering is the heavy time-consumption needed to deal with massive amounts of user information (Sarwar et al. 2001). This kind of system searches for neighbors among a vast user population of potential neighbors. Therefore, item-based collaborative filtering explores the relationship between items first instead of users to avoid this trap (Sarwar et al. 2001).

In the Internet era, gathering and analyzing of wider data has now made it possible to find out about the opinions and experiences of human beings. This valuable information can help people make decisions about various requirements, such as service or product quality improvements. A recommendation system is one type of technology that could help satisfy these requirements (Resnick et al. 1994). A recommendation system can reduce pressure on people to make choices when they encounter many uncertain alternative environments.

However, data and the algorithm architecture limit the capability of recommendation systems. Traditional recommendation systems only provide limited results based on user behaviors. Nevertheless, the Internet makes it easier to collect the opinions of crowds and allows the system and users to interact conveniently and quickly. The opinions of the crowd have become very important for recommendation systems. Using the behavior or knowledge of the crowd obtained from the Internet as a source of data can supplement limited recommendation results generated by individual information. For example, when investing in the stock market, if a person only invests in financial stocks, his/her investment payback can only be realized from the portfolio of financial stocks. However, other types of investors are not limited to financial stocks; their investment portfolios also include electronic stocks, transportation stocks, etc. Therefore, as users search for investment recommendations about financial stocks, a new type of new recommendation system could not only give recommendations about the relationships among financial stocks but ones about the relationship between financial stocks and other types of stocks.

Based on the argument above, we propose a novel model that can respond to recommendation items with an objective perspective when a user gives his/her ideas from a subjective perspective (the crowd's opinions/experiences). We call this model the novel cooperative recommendation model (NCRM). NCRM addresses the failure of traditional recommendation systems, which only deal with users' subjective ideas without the reference to objective ones. In this study, we employ the mining association rules technique and develop similarity algorithms to build our NCRM based on the study of Huang, Chen, & Chen (2016).

First of all, the users of NCRM can proffer their subjective search ideas to this model. After that, NCRM responds to optimal similar recommendations. For example, we take five listed stocks, Hon Hai, TSMC, UMC, CDIBH, and FFHC. When the stock price of TSMC is up, it is represented as TSMC↑. On the others, TSMC↓ indicates that the stock price of TSMC is down. When a user provides his/her idea with respect to {FFHC↑, CDIBH↓} in NCRM, there are two rules, {FFHC↑, CDIBH↓} ⇒ {TSMC↑} and {FFHC↑, CDIBH↑} ⇒ {Hon Hai↑}, which are more similar to the user's idea than others. The "⇒" represents that if the left-hand side happens, the right-hand side co-occurs. Therefore, NCRM recommends these two rules to the user.

As mentioned in the introduction above, we know that NCRM can provide more additional recommendations about users' concerns. These extra recommendations are stem from the discovery of the crowd's opinions/experiences. Therefore, the purposes and expected contributions of this study are as follows.

(1) The proposed NCRM is built by using the association rule mining technology and the similarity algorithm to integrate subjective and objective perspectives. We expect that NCRM will contribution to academic research and practical application in the future. For example, follow-up researchers can refer to this work when developing different algorithm models to explore the interactive relationship between subjective and objective data.

(2) In the field of computer science, algorithm performance is often used to evaluate the pros and cons of recommender systems, such as Mean Absolute Error (MAE) (Goldberg et al. 2001), Root Mean Square Error (RMSE) (Cotter & Smyth 2000), and F-measure (Sarwar et al. 2001). However, in the field of information management, user satisfaction with the recommendation system is emphasized (Liang, Lai, & Ku 2006). We are interested in exploring the differences between subjective cognition conditions and objective algorithm results and where users accept the system's recommendations or not. Therefore, we attempt measure users' perceived usefulness, trust, and satisfaction of NCRM by distributing a questionnaire to prove its effectiveness.

This paper is organized as follows. Section 2 reviews the literature on the study topic. Section 3 formally defines the problem. A new model, the novel cooperative recommendation model (NCRM), is developed in Section 4. Section 5 describes experiments that evaluate the effectiveness of our model. Finally, conclusions are presented in Section 6.

2. Literature Review

This section reviews the literature related to recommendation systems, association rules mining, and similarity in Section 2.1, Section 2.2, and Section 2.3, respectively. After that, we review ideas about information platforms in Section 2.4.

2.1 Recommendation Systems

Due to information overload, it has become increasingly difficult for users to get suitable data from the internet rapidly. To meet the needs of users, recommendation systems are becoming more important and valuable. Isinkaye, Folajimi, & Ojokoh (2015) collected the recommendation filtering techniques of previous scholars' studies to point out that recommender systems can be classified into content-based filtering, collaborative filtering, and hybrid filtering techniques. Content-based filtering emphasizes the analysis of users' previous items browsed or favored to generate predictions (Balabanovic & Shoham 1997). It can use the techniques of Term Frequency Inverse Document Frequency (TF/IDF) (Salton & Buckley 1988), Naïve Bayes Classifier (Friedman, Geiger, & Goldszmidt 1997), Decision Trees (Breiman et al. 1984; Loh 2011), or Neural Networks (Bishop 2011; Zhang et al. 2018) to generate recommendations. It has a disadvantage that content-based filtering effectiveness is limited by the availability of descriptions of items already defined in users' profiles. Therefore, collaborative filtering recommendation has gradually become the widely used recommendation technology. The idea of collaborative filtering considers similar users to find a product of potential interest to target users. Breese, Heckerman, & Kadie (1998) suggested that collaborative filtering or recommender systems can be categorized as memory-based algorithms and model-based algorithms.

Memory-oriented algorithms can be divided into User-based Collaborative Filtering (UBCF) and Item-based Collaborative Filtering (IBCF) according to differences in filtering data (Resnick et al. 1994; Herlocker et al. 1999; Sarwar et al. 2001; Lemire & Maclachlan 2005). The two most popular similarity measurements of memory-oriented filtering are used to calculate the similarity between items and users. The disadvantage of Pearson Correlation-based and Cosine-based memory-oriented algorithms is that they deal with vast data or calculate the complexity of data, causing the scalability and prediction quality to be poor (Wang et al. 2006).

Schafer et al. (2007) proposed a model based on the User-item Rating Matrix. Model-oriented algorithms create a classification model based on the User-item Rating Matrix. It can effectively solve the problem of the scalability of recommendation systems. Billsus & Pazzani (1998) presented a learning algorithm based on dimensionality reduction through the singular value decomposition (SVD) of an initial matrix of user ratings. In 2006, Simon Funk proposed the Funk-SVD algorithm to resolve two defects of SVD in a recommendation system (Piatetsky 2007). In addition to those mentioned above, model-oriented learning algorithms include association rule (Choa, Kimb, & Kim 2002; Mobasher, Jin, & Zhou 2004; Pan & Li 2010), clustering (McSherry 2004; Kużelewska 2013), decision tree (Caruana & Niculescu-Mizil 2006), artificial neural network (Larose & Larose 2014), link analysis (Cai et al. 2004; Linoff & Berry 2011), regression (Friedman et al. 1997), and matrix completion techniques (Koren, Bell, & Volinsky 2009; Candès & Recht 2009; Keshavan, Montanari, & Sewoong 2010).

Schafer et al. (2007) revealed that collaborative filtering serendipitously generates recommendations even without content in the user's profile. Although collaborative filtering techniques are applied widely in recommendation systems, some potential problems have been discovered, such as cold-start (Park & Chu 2009), data sparsity (Burke 2002; Park et al. 2012), scalability (Park et al. 2012), and synonymy (Landauer, Laham, & Foltz 1998). Scholars have proposed various techniques to solve these potential problems in the application of collaborative filtering (Deerwester et al. 1990;

Billsus & Pazzani 1998; Pazzani 1999; Moshfeghi, Piwowarski, & Jose 2011; Lika, Kolomvatsos, & Hadjiefthymiades 2014).

Hybrid filtering combines different algorithms to gain more accurate and effective recommendations (Schafer et al. 2007). Hybrid filtering approaches can be categorized into weighted hybridization (Claypool et al. 1999), switching hybridization (Billsus & Pazzani 1999), cascade hybridization (Burke 2002), mixed hybridization (Burke, Hammond, & Yound 1997; Ahmad Wasfi 1999; Smyth & Cotter 2000), feature-combination (Basu, Hirsh, & Cohen 1998), feature-augmentation hybrids (Mooney & Roy 1999), and meta-level hybrids (Pazzani 1999).

After studying the recommendation system literature, we refer to the ideas of model-oriented learning algorithms in collaborative filtering to design the mining techniques of NCRM. This study adopts user-item rating as its input; therefore, model-based learning computation to discover interesting rules serves as the foundation of the proposed model.

2.2 Association Rules Mining

Agrawal, Imielinski, & Swami (1993) introduced the association rules algorithm, which has been applied to mine the massive datasets or databases by follow-up scholars (Agrawal & Srikant 1994; Houtsma & Swami 1995). The algorithm results must satisfy at least the predefined thresholds of minimum support and confidence. However, the complex candidate generation process and multiple database scans are two limitations of the Apriori (Agrawal & Srikant 1994) algorithm. When the algorithm mines vast datasets, it is still time-consuming. Therefore, many scholars have attempted to modify or improve the Apriori algorithm to conquer these problems.

Frequent pattern growth mining (called FP-Tree) mines frequent patterns without any candidate generation process and reduces the number of passes over the database, leading to more efficiency and scalability than Apriori (Han & Pei 2000). In addition to FP-Tree, TreeProjection (Agarwal, Aggarwaln, & Prasad 2001), PRICRS (Wang & Tjortjis 2004), and the matrix algorithm (Yuan & Huang 2005) reduce the number of passes over large databases. Sampling technologies are efficient approaches to improving time-consuming mining by identifying an appropriate sample size to replace large datasets (Mannila, Toivonen, & Verkamo 1994; Toivonen 1996; Parthasarathy 2002; Li & Gopalan 2004; Chuang, Chen, & Yang 2005). Some scholars have adopted computing technologies of parallel systems and distributed approaches to improve the efficiency of association rule algorithms, including Fast Distributed Mining of association rules (FDM) (Cheung et al., 1996), Distributed Decision Miner (DDM) (Schuster & Wolff 2001), Fast Parallel Mining (FPM) (Cheung & Xiao 1998) and the Data Allocation Algorithm (DAA) (Manning & Keane 2001). In addition, Parthasarathy et al. (2001) presented a set of placement policies for parallel association mining on shared-memory multiprocessors. Tang & Turkia (2006) proposed a parallelization scheme which can achieve a high speed-up of parallel mining of the frequent itemset algorithm based on frequent pattern trees. The association rules algorithm can overcome constraints to reduce the computational complexity of the mining process and enhance efficiency significantly (Das, Ng, & Woon 2001; Wojciechowski & Zakrzewicz 2002; Do, Hui, & Fong 2003).

If the specified support and confidence thresholds are low, the set of association rules will grow to be unwieldy as the number of transactions increases, and many of these may be redundant (Kotsiantis & Kanellopoulos 2006). Therefore, reducing the redundancy rules could increase the speed of association rules mining (Jaroszewicz & Simovici 2002; Ashrafi, Taniar, & Smith 2004).

Following continuous improvements, the association rule algorithm has efficiently mined rules between vast itemsets. This study uses association rules as the data mining technology of NCRM to efficiently discover the association rules of itemsets.

2.3 Similarity

Similarity is the measurement of the distance between two objects and is the inverse of dissimilarity. Jagadish, Mendelzon, & Milo (1995) developed a framework that comprises three components, including a pattern language, transformation rule language, and query language for posing queries to perform a similarity-based search, rather than equality or inequality searches. Faloutsos et al. (1997) adapted this framework to shrink the data sequences into signatures and searched the signatures instead of the real sequences. Using signatures makes it easier to index and fits many real-life applications of efficient searching. Similarity analysis problems can be addressed by various similarity measures in different application domains, such as text, images, or video datasets, as well as pattern matching, sequence matching, and geometric shape matching. Similarity analysis techniques can be the foundation of pattern recognition, clustering, simplification, or representation.

Similarity analysis techniques are often applied in time series analysis and computational geometry. Researchers have developed similarity analysis techniques for time series analysis and geometric shapes to further address the similarity analysis problems of moving object data mining. Ding et al. (2008) reviewed nine similarity measures to calculate the distance between two datasets of time series. These included lock-step measures (e.g., Lp Norms) (Agrawal, Faloutsos, & Swami 1993; Faloutsos, Ranganathan, & Manolopoulos 1994; Yi & Faloutsos 2000), distance measures (e.g., Dynamic Time Warping (DTW) (Berndt & Clifford 1994; Faloutsos & Lin 1995; Yi, Jagadish, & Faloutsos 1998; Kim, Park, & Chu 2001; Sakurai, Yoshikawa, & Faloutsos 2005; Altiparmak et al. 2006), edit distance (ED) (Levenshtein 1966; Bozkaya, Yazdani,

& Ozsoyoglu 1997; Chen & Ng 2004)), and longest common subsequence (LCSS) (Vlachos, Kollios, & Gunopulos 2002).

Norm is a measurement of mathematical object size. L_p norm is a famous approach for dissimilarity measurement of time series. The p is defined as a series of metrics to calculate the distance between two entities in vector space. The p-value of Manhattan Distance (L_1 norm) equals 1, and the Euclidean Distance (L_2 norm) is called when p equals 2. When p equals ∞, it is known as the maximum norm (L_∞ norm) (Yi & Faloutsos 2000). L_p-norms make it easy to calculate the time series similarity matching, but they cannot handle local time-shifting. Nonmetric distance function techniques such as DTW, LCSS, and ED have been adopted to solve local time-shifting similarity. The Euclidean distance (L_2 norm) can only be used to calculate two sequences of the same length. The time warping distance can be applied to deal with any two sequences of arbitrary lengths. As for the issue of similarity of two sequences, Tsai & Shieh (2009) proposed a three-phase sequential pattern change detection framework to discover sequential pattern changes between two time periods.

After the above evaluation, we refer to the idea of the change detection approach (Tsai & Shieh 2009) to design the similarity algorithms of our model. Therefore, we simplify the change detection approach of dissimilarity measurement to calculate the similarity of the subjective and objective perspectives of NCRM.

2.4 Information Platforms

Ainsbury et al. (2000) pointed out that an information platform automatically collects data, integrates data, performs analysis using multiple content-types, and provides a method for organizing a library of information, thereby providing users insights to make decisions rapidly. An information platform is not only a client/server or n-tiers structure but also can be subdivided into four major sections, including (1) data retrieval, (2) data classification and storage, (3) information browsing, and (4) various formats of desktop integration (Ainsbury et al. 2000). Platform architectures modularize complex systems that steadily maintain operation of specific components (Baldwin & Woodard 2008). Platforms must provide steady architecture that can orchestrate the cross-components of the various services and products being operated on new devices (Boudreau 2010; Mcintyre & Srinivasan 2017). Thereby, platforms assist consumers in product search or selection based on the product's availability and the supplier's geographic region (Najmul Islam, Cenfetelli, & Benbasat 2020). As for consumers, platforms provide opportunities for consumers to meet their needs for services and products. Information platform have been implemented in various application fields, like industrial symbiosis networks, electronic marketplaces, the sharing economy, global labor market, and so on (Fraccascia & Yazan 2018; ITIF October 12, 2018; Standing, Standing, & Love 2010).

According to the literature mentioned above, we adopt information platform concepts in the NCRM system platform. We deploy NCRM on website server to perform the experimental measurements of the study topic. We choose n-tiers structure to build it. Respondents can use internet devices to browse the NCRM website and offer their inquiry conditions. NCRM uses the change detection approach of simplified similarity to calculate the respondents' inquiry conditions and the objective rules. The association rule algorithm generates the objective rules. Finally, NCRM offers suggestions to respondents according to the similarity results.

3. Problem Definition

We address association rules and similarity to query the recommendations between subjective search conditions and objective association rules. Here, we formally define the problem of NCRM algorithms of association rules and similarity.

Let $I = \{i_1, i_2, \ldots, i_n\}$ be a set of n distinct attributes, called items. T is transactions that contain a set of items such that $T \subseteq I$; D is a database with different transaction records T_S. $D = \{t_1, t_2, \ldots, t_m\}$. Each transaction ID is unique in D and each transaction contains a subset of the items in I.

3.1 Association Rules

We use the association rules algorithm to mine the rules from D. An association rule takes the form of $< X \Rightarrow Y >$, where $X, Y \subset I$ are sets of items called itemsets, and $X \cap Y = \emptyset$. Given $X = (i_k \oplus_k i_{k+1})$, $1 \leq k \leq k+1 \leq n$ and $k, n \in Z^+$. $\oplus_k \in$ "~" means that i_k and i_{k+1} appear at the same time. Y is the same as the definition of X. X is called the antecedent or left-hand-side (LHS) while Y is called the consequent or right-hand-side (RHS). This rule means X implies Y. The "⇒" represents that if LHS happens, RHS co-occurs.

For example, assume that we have a database (TABLE 1). The set of items is $I = \{a, b, c, d, e\}$, where, in each entry, the value 1 indicates the item's presence in the transaction, and the value 0 represents absence. Let an association rule be $< (c \sim b) \Rightarrow a >$, meaning that if c and b are bought at the same time, a is also bought by customers.

TABLE 1: A database with 5 transactions and 5 items

Transaction ID	*a*	*b*	*c*	*d*	*e*
1	0	0	1	0	0
2	1	1	0	0	0
3	0	1	0	0	0
4	0	0	0	1	1
5	1	1	1	0	0

Definition 1 Support and confidence are two important thresholds for association rules. Support (Sup(X)) is the percentage of records that contain $X \cup Y$ in the total number

of records in the database. Support value (called *Sp*) corresponds to statistical significance.

$$\mathrm{Sup}(X) = \frac{|\{t \in T \,; X \subseteq t\}|}{|T|} \tag{3-1}$$

Example 1 The itemset $X = (a \sim b \sim c)$ has a support of 1/5 = 0.2.

Definition 2 Confidence refers to how often the rule is valid. Confidence (Conf($X \Rightarrow Y$)) is the percentage of the number of transactions that contain $X \cup Y$ within the total number of records that contain X. Therefore, the Confidence value (called *Cf*) measures the strength of the association rules.

$$\mathrm{Conf}(X \Rightarrow Y) = \frac{\mathrm{Sup}(X \cup Y)}{\mathrm{Sup}(X)} \tag{3-2}$$

Example 2 The rule $< (c \sim b) \Rightarrow a >$ has a confidence of 0.2/0.2 = 1.0 in the database.

Definition 3 Association rules have to satisfy the user-specified minimum support (called mini-*Sp*) and the user-specified minimum confidence (called mini-*Cf*) at the same time. The steps of an association rule mining algorithm are as follows:

- The set of candidates k-itemsets is generated by 1-extension of the large (k -1)-itemsets generated in the previous iteration. The number of items in an itemset is called the length of an itemset. Itemsets of length k are called k-itemsets.
- Passing over the database generates the candidate k-itemsets.
- Itemsets that are less than the mini-*Sp* mean that the rule is not worth consideration, and the others are called large k-itemsets.

The above processes are repeated until no larger itemsets are found.

Example 3 Assume we specify that the mini-*Sp* is 20% and the mini-*Cf* is 100%. There are five rules matched, including $< (d) \Rightarrow e >$, $< (e) \Rightarrow d >$, $< (a) \Rightarrow b >$, $< (a \sim c) \Rightarrow b >$, and $< (b \sim c) \Rightarrow a >$.

We use association rules to mine the rules, which are the objective data of NCRM, from target transaction data sets. The rules are saved in the association database (called *AR*). *AR* includes different association rules: $ar_1, ar_2, \ldots, ar_j, \ldots, ar_r$, where $1 \leq j \leq r$ and $j, r \in Z^+$. Each *ar* is a rule of the association algorithm mined from *D*, and the rule of *ar* is combined by the operators and the different items of *I*.

Definition 4 We define ar_j as $< (i_{j,k} \oplus_k i_{j,k+1}) \Rightarrow i_{j,q} >$, where $1 \leq k \leq k+1 \leq n$, $1 \leq q \leq n$, $(i_{j,k} \oplus_k i_{j,k+1}) \cap i_{j,q} = \emptyset$, and $k, q \in Z^+$. $\oplus_k \in$ "~" represents that $i_{j,k}$ and $i_{j,k+1}$ appear at the same time. $(i_{j,k} \oplus_k i_{j,k+1})$ is the antecedent (called aar_j) and $i_{j,q}$ (called car_j) is the consequent in the database of association rules. We can also express that ar_j is $< aar_j \Rightarrow car_j >$.

3.2 Similarity

We use similarity to measure two pairs of rules (subjective inquiry condition and objective association rules) to discover their optimal matching relationship for recommendation.

Definition 5 The rule of S is the user's inquiry itemsets that contain a set of items such that $S \subseteq I$. Given $S = (i_p \oplus_p i_{p+1})$, $1 \leq p \leq p+1 \leq n$ and $n, p \in Z^+$. $\oplus_p \in$ "~" represents that i_p and i_{p+1} appear at the same time.

Example 4 Assume that the user's inquiry condition S is $(a \sim c)$. Let the database of association rules be $AR = \{< (a \sim b \sim c \sim d) \Rightarrow e >, < (a \sim b \sim c \sim f) \Rightarrow d >, < (a \sim b \sim e \sim f) \Rightarrow c >, < (a \sim b \sim f \sim h) \Rightarrow c >\}$.

Definition 6 (Similarity measurement)**.** Defining the formula of SV_j to measure the similarity between S and aar_j; SV_j is formulated as follows:

$$SV_j = \begin{cases} 1, & \textit{the two rules are the same} \\ 0, & \textit{the two rules are mismatched} \\ Sim(S, aar_j), & \end{cases} \tag{3-3}$$

The items that are included in S and aar_j have no restrictions on the sequence. If the items are the same for the two rules, SV_j will be 1. On the contrary, if all items are mismatched between the two rules, SV_j is equal to 0. Sim (S, aar_j) can be defined as follows:

$$\mathrm{Sim}(S, aar_j) = \frac{NumofComItem(S, aar_j)}{Max(|S|, |aar_j|)} \tag{3-4}$$

$Max(|S|, |aar_j|)$ is the maximal number of items between S and aar_j, $NumofComItem(S, aar_j)$ is the number of common items between S and aar_j.

Example 5 We match the items in S with the items of all rules in the antecedent of AR. For example, if a user's inquiry is $S_1 = (a \sim b \sim c)$, the antecedent of the rules $(a \sim b \sim c \sim d)$ and $(a \sim b \sim c \sim f)$ are matched. On the other hand, if another user's inquiry is $S_2 = (a \sim b)$, the matched antecedent of the rules includes $(a \sim b \sim c \sim d)$, $(a \sim b \sim c \sim f)$, $(a \sim b \sim e \sim f)$, $(a \sim b \sim f \sim h)$. It is obvious that if the number of items in S is few, the number of rules matched in the antecedent of AR is more.

Definition 7 (Maximal similarity)**.** The largest similarity value between S and all rules in the antecedent of AR, called the maximal similarity MS, is defined as:

$$MS = \max(SV_1, SV_2, \ldots, SV_j) \tag{3-5}$$

Example 6 For instance, if a user's inquiry S = $(a \sim b \sim c)$ is compared with the antecedent of the rules $(a \sim b \sim c \sim d)$ and $(a \sim b \sim e \sim f)$, we get $MS = \max(0.75, 0.5)$. Therefore, $(a \sim b \sim c \sim d)$ is the antecedent rule of maximal similarities.

We use the above algorithm to obtain the maximal similarities as model recommendations based on comparison of two pairs of rules: the subjective inquiry condition and objective association rules.

4. The Proposed Algorithm and System Model

This section proposes the NCRM system model to find recommendations integrating the subjective inquiry condition and objective association rules. The NCRM algorithm is developed by applying the association rules and similarity. The flow chart of the NCRM algorithm is presented in FIGURE 1.

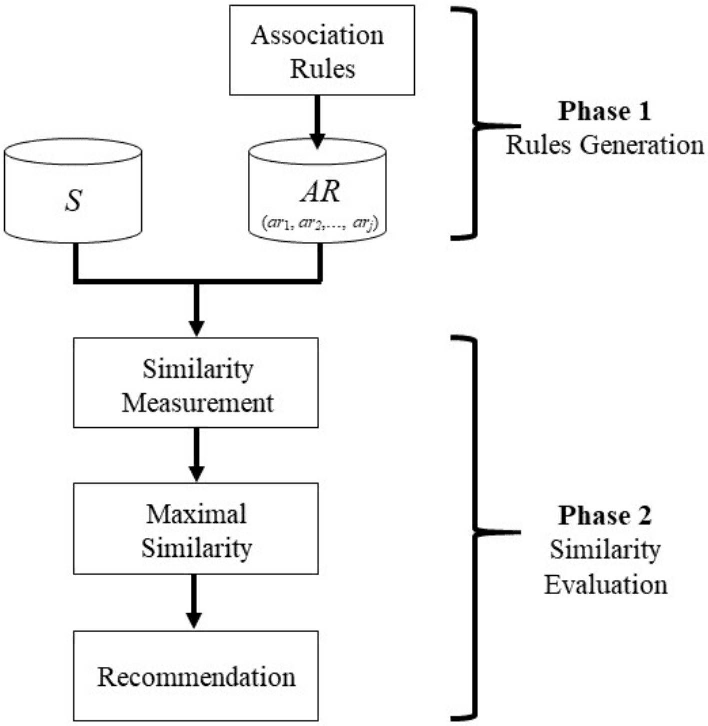

FIGURE 1: Flow chart of the NCRM algorithm

The first phase is rules generation. We use the association rules to mine the extensive target data and store the rules in a database ($AR = \{ar_1, ar_2, \ldots, ar_j\}$).

Example 7 We capture the $AR = \{< (a \sim b \sim c \sim d) \Rightarrow e >, < (a \sim b \sim c \sim f) \Rightarrow d >, < (a \sim b \sim e \sim f) \Rightarrow c >, < (a \sim b \sim f \sim h) \Rightarrow c >\}$ from the database of the association rules.

Then, the user's inquiry condition (S) is gathered from the web interface of NCRM.

Example 8 The user's inquiry condition of S is $(a \sim b \sim c)$.

In the second phase, the pair of S and the AR's LHS can be compared by measuring their similarity. The highest similarity rules of AR's LHS are kept, and the others are discarded. After the similarity evaluation process, NCRM gives the recommendation rules when the AR's association rules with the highest similarity match the user's inquiry.

Example 9 We match the items in S with the items of all rules in the antecedent of AR. For example, if $S = (a \sim b \sim c)$, the antecedent of the rules, $(a \sim b \sim c \sim d)$ and $(a \sim b \sim c \sim f)$, are the highest similarities (MS = max (0.75, 0.75, 0.5, 0.5)), then NCRM gives the following recommendation rules: $< (a \sim b \sim c \sim d) \Rightarrow e >$ and $< (a \sim b \sim c \sim f) \Rightarrow d >$.

System Model

We use the association rules algorithm to mine the rules from raw data. The algorithm can generate the association rules to satisfy the predefined thresholds of minimum support and confidence. The generated association rules are stored in the database. And then, we propose a real-time inquiry and feedback system to make recommendations to the users. NCRM uses the above-mentioned similarity algorithm to mine the recommendations via users' subjective inquiry conditions and objective rules from the database.

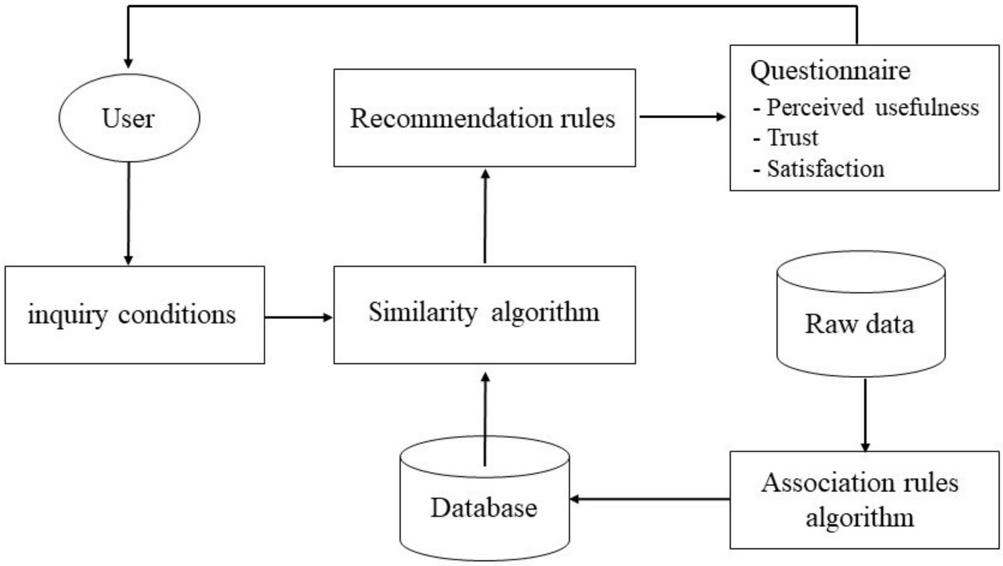

FIGURE 2: System architecture of NCRM

Finally, following Huang (2013), we use demographic distributions to measure user confidence and satisfaction with the novel algorithm recommendation model. After users apply NCRM and obtain recommendations, the website invites them to complete a questionnaire that examines NCRM users' perceived usefulness, trust, and satisfaction. The system architecture of the experiment is shown in FIGURE 2.

5. Experimental Study

In this section, we used the Taiwan Top50 Exchange Tracker Fund (ETF50) stock price fluctuations over a period of six months as the datasets to evaluate the performance of NCRM. First, following Huang & Li (2018), we mined the datasets of association rules saved into the Microsoft SQL Server 2014 Express database as the objective source. Then, the website and similarity algorithm programs were implemented using Sun Java™ language (J2SDK 1.8.0_251). The PC used in this study applied Apache, Tomcat, and Microsoft SQL Server 2014 Express services, with an Intel Core i7-8700 3.2 GHz processor and 32 GB main memory using the Windows 10 professional edition operating system. As follows, we present the results of the experiments using the real dataset in Section 5.1 and the results of the questionnaire to examine NCRM users' perceived usefulness, trust, and satisfaction in Section 5.2.

5.1 Real Dataset

In this section, we examine the actual status of how NCRM works in practice. Therefore, we selected the ETF50 stock price fluctuations between July 15, 2019 and May 4, 2020, as the experimental datasets. All of the experimental procedures are described as follows.

1.First of all, we collected the target data from the Taiwan Economic Journal database system (TEJ). The amount of the data is 10,476, which was gathered according to the listed companies of the ETF50 (TABLE 2) during the 194 trading days from July 15, 2019 to May 4, 2020. Each record's fields include the stock code, listed company name, trading day, opening price, daily high, daily low, and closing price.

TABLE 2: The listed companies of the ETF50

TCC (1101)	ACC (1102)	UNI-PRESIDENT (1216)	FPC (1301)
NPC (1303)	FCFC (1326)	FENC (1402)	CSC (2002)
CST (2105)	HOTAI MOTOR (2207)	YNM (2227)	LTC (2301)
UMC (2303)	DELTA (2308)	HON HAI (2317)	YAGEO (2327)
TSMC (2330)	QISDA (2352)	ASUSTEK (2357)	QCI (2382)
ACL (2395)	NTC (2408)	CHT (2412)	MTK (2454)
CATCHER (2474)	YMTC (2609)	CAL (2610)	THSRC (2633)
CHANG HWA BANK (2801)	CHINA LIFE (2823)	HNFHC (2880)	FUBON FINANCIAL (2881)
CATHAY HOLDINGS (2882)	CDF (2883)	E.S.F.H (2884)	YUANTA GROUP (2885)
MEGA FHC (2886)	TAISHIN HOLDINGS (2887)	SKFH (2888)	SINOPAC HOLDINGS (2890)
CTBC HOLDING (2891)	FFHC (2892)	PCSC (2912)	LARGAN (3008)
TWM (3045)	ASEH (3711)	FAR EASTONE (4904)	PEGATRON (4938)
CHAILEASE (5871)	SCSB (5876)	TCFHC (5880)	FPCC (6505)
PCC (9904)	FT (9910)		

Remark: Listed Company Abbreviation in English (Stock Symbol)

2.Second, trading days of individual stocks were divided into three types: up, down, and flat, according to the formula below. We define the x as today's closing price, y as yesterday's closing price, and k as the sensitivity. When it meets $\frac{(x-y)}{y} \times$

$100\% > k\%$, we mark the status as up (↑). On the contrary, when it satisfies $\frac{(x-y)}{y} \times 100\% < -k\%$, we say it is down (↓). Finally, when it does not satisfy the above two conditions, we mark it as flat (■). We set k as 0 and reorganize the daily transaction data, as shown in TABLE 3 for association rules mining. If k= 0.0, the number of records is 1,089 (x=y), which accounts for 10.40% of the total data. However, if k = 0.2, the number of records is 1,926, which accounts for 18.40%. If k = 0.4, the number of records is 3,708, which accounts for 35.40% of all data.

Example 10 Assume TCC's closing price was 44.05 on July 15, 2019 and was 44.75 on July 16, 2019. When we set the k as 0, we found that the fluctuation of TCC was up (TCC↑) on July 16, 2019.

TABLE 3: The fluctuation of the marked stock sheet

Trading date	TCC	ACC	FPC	…	FT
2019/07/16	TCC↑	ACC■	FPC↓	…	FT↑
2019/07/17	TCC↓	ACC↓	FPC↓	…	FT↑
			…		
2020/05/04	TCC↑	ACC↑	FPC↑	…	FT↑

3. We use the association rules to mine the fluctuation of the marked stock sheets, including the filtering result of the different k-values from the different combinations of mini-*Sp* and mini-*Cf* (TABLE 4, TABLE 5, and TABLE 6). So that NCRM can offer a more efficient similarity response and has a higher confidence and support threshold, our experiment sets the value k as 0.2, with 90% confidence, and 20% support. The number of rules is 1,430, which are saved in the database. For example, we can gather two rules of the highest confidence and support from the database, < (SINOPAC HOLDINGS↑ ~ CSC↑ ~ FFHC↑) ⇒ (HNFHC↑)>, < (SINOPAC HOLDINGS↑ ~ CTBC HOLDING↑ ~ YUANTA GROUP↑) ⇒ (HNFHC↑)>: the parameters for both are k = 0.2, 97.826% confidence, and 23.711% support.

TABLE 4: The association rules quantity matrix of k=0

k=0		mini-*Cf*		
	Rule Qty.	70%	80%	90%
mini-*Sp*	16%	1,384,035	524,631	78,044
	18%	620,214	223,634	22,547
	20%	227,875	68,195	6,590

TABLE 5: The association rules quantity matrix of *k*=0.2

k=0.2		mini-*Cf*		
	Rule Qty.	70%	80%	90%
mini-*Sp*	16%	270,923	92,919	14,236
	18%	122,863	39,753	4,522
	20%	46,682	12,090	**1,430**

TABLE 6: The association rules quantity matrix of *k*=0.4

k=0.4		mini-*Cf*		
	Rule Qty.	70%	80%	90%
mini-*Sp*	16%	5,588	703	29
	18%	2,850	352	9
	20%	1,174	117	1

4. Finally, we code the java programs of NCRM and deploy them on the Apache and Tomcat website (FIGURE 3). Users can fill in the inquiry conditions of subjective responses on the website.

FIGURE 3: The operation of the NCRM website

The similarity algorithm captures the results from the database according to the inquiries. It compares the association rules, and recommends those with higher similarities between inquiries to users.

Example 11 When the user queried (CHT↑ ~ MTK ↑ ~ YAGEO↓) in NCRM, four rules were generated from the similarity comparison of 1,430 association rules, of which the *MS* was 0.333 (TABLE 7). Therefore, NCRM recommended these four rules according to *Cf* value to the user. The rule < (CHT↑ ~ CTBC HOLDING↑) ⇒ (HNFHC↑)> was presented on the first line because it had the highest *Cf* (92.683).

TABLE 7: The results of similarity comparison from example 11

Association Rule	*MS*	*Sp*	*Cf*
< (FUBON FINANCIAL↓ ~ YAGEO↓) ⇒ (QISDA↓)>	0.333	20.619	90
< (CHT↑ ~ CTBC HOLDING↑) ⇒ (HNFHC↑)>	0.333	21.134	92.683
< (CATHAY HOLDINGS↓ ~ YAGEO↓) ⇒ (SKFH↓)>	0.333	20.103	92.308
< (DELTA↓ ~ YNM↓ ~ YAGEO↓) ⇒ (QISDA↓)>	0.333	20.619	90

5.2 Results Analysis

We used an online questionnaire to elicit responses from users who used NCRM to examine their perceived usefulness, trust, and satisfaction. The questionnaire design included three parts. The first part was a nominal scale, mainly to understand the essential information variables of the respondents. Next respondents implemented the NCRM system online. Finally, we used a Likert 7-point scale to measure the users' perceived usefulness, trust, and satisfaction after using NCRM (Appendix A).

TABLE 8: Descriptive statistics of respondents' basic information

Measure	Items	Frequency	Percentage
Gender	Male	66	51.97 %
	Female	61	48.03 %
Age	Under 20	0	0 %
	21-30	24	18.90 %
	31-40	37	29.14 %
	41-50	35	27.56 %
	51-60	28	22.04 %
	Over 60	3	2.36 %
Education	Senior high school or below	8	6.30 %
	College degree	73	57.48 %
	Master degree	45	35.43 %
	Ph.D. degree	1	0.79 %
Industry or Occupation	Military and Police	2	1.58 %
	Government	7	5.51 %
	Education	7	5.51 %
	Financial and Insurance	12	9.45 %
	Manufacturing	24	18.90 %
	Medical	2	1.58 %
	Service	17	13.38 %
	Housekeeping	8	6.30 %
	Student	5	3.94 %
	Retired	2	1.58 %
	Freelance	13	10.23 %
	Others	28	22.04 %
Total investment amount in one month (unit: NTD)	Under 20000	73	57.48 %
	20000-60000	27	21.26 %
	60001-100000	8	6.30 %
	100001-140000	6	4.72 %
	140001-200000	3	2.36 %
	Over 200000	10	7.87 %
Years of investment experience	Under 1 year	25	19.69 %
	1-2 years	40	31.50 %
	3-4 years	12	9.45 %
	5-6 years	7	5.51 %
	7-8 years	11	8.66 %
	Over 8 years	32	25.20 %

The experiment lasted for about six months, from June 26 to December 30, 2020. A total of 149 respondents were included in this study. Twenty-two of them were deleted because they did not fill out the questionnaire completely. Hence, the number of valid samples was 127.

TABLE 8 shows the demographic distributions. As shown in the table, 51.97% of subjects were male and 48.03% were female. The age range was mostly between 31-40 and 41-50 years old, accounting for 56.70%. College graduate (57.48%) and master's degree (35.44%) were the major educational backgrounds of the respondents. The participants were considered to be highly educated. Others (22.04%), manufacturing (18.90%) and service (13.38%) industries were the top three occupational categories of the respondents. The range of investment amount in a month was mostly under 20,000 (57.48%) and from 20,000-60,000 NTD (21.26%). The most significant proportion for years of investment experience was 1-2 years (31.50%), followed by those with over 8 years (25.20%).

TABLE 9: The descriptive statistics of respondents' after using NCRM

Measure	Items	Frequency	Percentage
NCRM is useful	Very strongly disagreed	2	1.58 %
	Strongly disagreed	3	2.36 %
	Disagreed	2	1.58 %
	Average	33	25.98 %
	Agreed	34	26.77 %
	Strongly agreed	43	33.86 %
	Very strongly agreed	10	7.87 %
Trust the investment recommendations of NCRM	Very strongly disagreed	3	2.36 %
	Strongly disagreed	3	2.36 %
	Disagreed	5	3.94 %
	Average	30	23.62 %
	Agreed	39	30.71 %
	Strongly agreed	37	29.14 %
	Very strongly agreed	10	7.87 %
Satisfaction at the experience of using NCRM	Extremely unsatisfied	3	2.36 %
	Very unsatisfied	4	3.15 %
	Unsatisfied	3	2.36 %
	Average	30	23.62 %
	Satisfied	35	27.56 %
	Very satisfied	45	35.44 %
	Extremely satisfied	7	5.51 %

After uses implemented NCRM we surveyed their perceived usefulness, trust, and satisfaction with NCRM. The distribution of the measurements is shown in TABLE 9. Davis (1989) pointed out when a person uses a particular system, perceived usefulness is the degree to which a user perceives a positive use-performance relationship. Most subjects agreed that NCRM is useful (26.77% of the samples agreed, 33.86% strongly agreed, and 7.87% are agreed very strongly). Thus, according to our findings, subjects expressed high perceived usefulness after using NCRM.

Traditionally, trust is defined as a person's beliefs based on his or her perceptions about certain attributes. Trust is the confidence in the trustworthiness and integrity of trading partners (Morgan & Hunt 1994). As for trust in the investment recommendations of NCRM in our experiment, 67.72% of samples agreed or strongly (30.71% of the samples agreed, 29.14% strongly agreed, and 7.87% very strongly agreed). Finally, satisfaction with using NCRM was 68.51% (27.56% of the samples were satisfied, 35.44% were very satisfied, and 5.51% were extremely satisfied), while only 7.87% were dissatisfied. Most subjects were satisfied with the stock investment recommendations provided by NCRM.

The questionnaire items of users' perceived usefulness, trust, and satisfaction showed reliability and validity. The results of reliability indicators in all dimensions these complied with the standards, as shown in TABLE 10. The square roots of the AVE of perceived usefulness (0.933), trust (0.935), and satisfaction (0.914) are greater than the correlation coefficients of the two different structures. Therefore, the discriminant validity of the questionnaire is sufficient.

TABLE 10: Reliability of the measurement

Construct	Items	Factor loading	CR	AVE	Cronbach's α
Perceived usefulness	PU1	0.942	0.971	0.870	0.962
	PU2	0939			
	PU3	0.886			
	PU4	0.948			
	PU5	0.947			
Trust	TR1	0.955	0.972	0.874	0.964
	TR2	0.966			
	TR3	0.939			
	TR4	0.873			
	TR5	0.939			
Satisfaction	US1	0.927	0.968	0.835	0.961
	US2	0.904			
	US3	0.913			
	US4	0.912			
	US5	0.911			
	US6	0.916			

To sum up, most respondents gave NCRM a high degree of recognition for its usefulness, trustworthiness, and satisfaction level.

6. Conclusions

Most existing studies of recommendation systems focus on calculating results based on objective data and emphasize the comparison of execution efficiency. Further, we found that few studies have addressed recommendation systems by constructing models from a combination of subjective and objective perspectives. Therefore, this study proposes a novel recommendation model from these two perspectives and attempts to figure out users' perceived usefulness, trust, and satisfaction with the proposed model.

NCRM is a data-driven model for novel cooperative recommendation that explores objective data by subjective inquiry. NCRM is built by combining association rule mining technology and the similarity algorithm to implement a recommendation system of subjective and objective perspectives. NCRM uses the former to discover the rules from objective data. After that, according to subjective inquiry conditions, the execution of the similarity approach of NCRM provides recommendations, which sometimes are able to exceed users' expectations from NCRM.

According to the result of the experimental questionnaire, the novel cooperative recommendation model (system) presents high perceived usefulness, trust, and satisfaction. This is a new and interesting finding.

NCRM represents a new and promising research direction in recommendation systems, integrating the data mining approach and similarity. In the future, the model can be expanded by considering more algorithm applications that integrate various data mining methods and similarities to build more effective and valuable recommendation systems. More real-life data sets can be employed to check the proposed model's value or more complex research methods can be used to examine users' various types of cognitions for systems. For example, a system could accommodate the time lag factors of stock price fluctuations from the investor's point of view to provide recommendations. Future work could also develop theoretical research frameworks to measure user satisfaction with recommendation systems. However, recommendations generated by different algorithms and datasets may affect the users' views, which could be both a research limitation and exciting topic to be explored.

References

Agarwal, R. C., Aggarwaln, C. C. & Prasad, V. V.V. (2001). A tree projection algorithm for the generation of frequent itemsets. *Journal of Parallel and Distributed Computing,* 61(3), 350-371. doi:10.1006/jpdc.2000.1693

Agrawal, R. & Srikant, R. (1994). *Fast algorithms for mining association rules*. In Proceedings of the 20th International Conference Very Large Data Bases, 487-499.

Agrawal, R., Faloutsos, C. & Swami, A. (1993). Efficient similarity search in sequence databases. *Lecture Notes in Computer Science,* 730, 69-84.

Agrawal, R., Imielinski, T. & Swami, A. (1993). *Mining association rules between sets of items in large databases*. In Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, 207-216.

Ahmad Wasfi, A. M. (1999). *Collecting user access patterns for building user profiles and collaborative filtering*. Proceedings of the 4th International Conference on Intelligent User Interfaces - IUI '99, 57-64. doi:10.1145/291080.291091

Ainsbury, R. D., Hussein, H. K. A., Hinnant, M. C., Lahham, M. M., Ludin, S. L., Putterman, D. S., Shotton, F. R. & Tejada, W. M. (2000). Method and apparatus for performing data collection, interpretation, and analysis, in an information platform. *In Google Patents*.

Altiparmak, F., Ferhatosmanoglu, H., Erdal, B. & Trost, D. (2006). Information mining over heterogeneous and high-dimensional time-series data in clinical trials databases. *IEEE Transactions on Information Technology in Biomedicine: a Publication of the IEEE Engineering in Medicine and Biology Society,* 10(2), 254-263.

Ashrafi, M. Z., Taniar, D. & Smith, K. (2004). A new approach of eliminating redundant association rules. *Lecture Notes in Computer Science,* 3180, 465-474.

Balabanovic, M. & Shoham, Y. (1997). Fab: Content-based, collaborative recommendation. *Communications of the ACM,* 40(3), 66-72.

Baldwin, C. Y. & Woodard, C. J. (2008). The architecture of platforms a unified view. *Harvard Business School Finance Working Paper,* 09-034, 1-31.

Basu, C., Hirsh, H. & Cohen, W. (1998). *Recommendation as classification: using social and content-based information in recommendation*. In: Proceedings of the 15th National Conference on Artificial Intelligence, Madison, WI, 714-720.

Bauer, J. & Nanopoulos, A. (2014). Recommender systems based on quantitative implicit customer feedback. *Decision Support Systems,* 68, 77-88.

Berndt, D. J. & Clifford, J. (1994). *Using dynamic time warping to find patterns in time series*. Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining, 359-370.

Billsus, D. & Pazzani, M. J. (1999). A hybrid user model for news story classification. In Kay, J. (Eds.), *Proceedings of the 7th International Conference on User Modeling*, 99-108, Banff, Canada: Springer-Verlag, New York.

Billsus, D. & Pazzani, M.J. (1998). *Learning collaborative information filters*. Proceedings of the 15th International Conference on Machine Learning, 46-54.

Bishop C. M. (2011). *Pattern Recognition and Machine Learning*. Springer.

Boudreau, K. (2010). Open platform strategies and innovation: granting access vs. devolving control. *Management Science, 56*, 1849-1872.

Bozkaya, T., Yazdani, N. & Ozsoyoglu, Z. (1997). *Matching and indexing sequences of* *different lengths*. Proceedings of the 6th International Conference on Information and Knowledge Management, 128-135. doi: 10.1145/266714.266880

Breese, J. S., Heckerman, D. & Kadie, C. (1998). *Empirical analysis of predictive algorithms for collaborative filtering*. Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence, 43-52.

Breiman, L., Friedman, J., Stone, C. J. & Olshen, R.A. (1984). *Classification and regression trees*. Chapman and Hall/CRC.

Burke R. (2002). Hybrid recommender systems: survey and experiments. *User Modeling and User-Adapted Interaction,* 12(4), 331-370. doi:10.1023/A:1021240730564

Burke, R. D., Hammond, K. J. & Yound, B. C. (1997). The FindMe approach to assisted browsing. *IEEE Expert,* 12(4), 32-40. doi:10.1109/64.608186

Cai, D., He, X., Wen, J.R. & Ma, W.Y. (2004). *Block-level link analysis*. Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 440-447. doi:10.1145/1008992.1009068

Candès, E. J. & Recht, B. (2009). Exact matrix completion via convex optimization. *Found Comput Math,* 9, 717-772. doi:10.1007/s10208-009-9045-5

Caruana, R. & Niculescu-Mizil, A. (2006). *An empirical comparison of supervised learning algorithms*. ICML '06: Proceedings of the 23rd International Conference on Machine Learning, 161-168. doi:10.1145/1143844.1143865

Chen, L. & Ng, R. (2004). *On the marriage of Lp-norms and edit distance*. Proceedings of the 30th VLDB Conference, 792-803.

Cheung, D. W. & Xiao, Y. (1998). Effect of data skewness in parallel mining of association rules. *Lecture Notes in Computer Science,* 1394, 48-60.

Cheung, D. W., Han, J., Ng, V. T., Fu, A. W. & Fu, Y. (1996). *A fast distributed algorithm for mining association rules*. 4th International Conference on Parallel and Distributed Information Systems. Miami Beach, FL, USA, 31-44, doi: 10.1109/PDIS.1996.568665

Choa, Y. H., Kimb, J. K. & Kim S. H. (2002). A personalized recommender system based on web usage mining and decision tree induction. *Expert Systems with Applications,* 23, 329-342.

Chuang, K., Chen, M. & Yang, W. (2005). Progressive sampling for association rules based on sampling error estimation. *Lecture Notes in Computer Science,* 3518, 505-515.

Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D. & Sartin, M. (1999). *Combining content-based and collaborative filters in an online newspaper*. ACM SIGIR Workshop on Recommender Systems - Implementation and Evaluation, 1-8.

Cotter, P. & Smyth, B. (2000). *PTV: Intelligent personalized TV guides*. Proceedings of the 17th National Conference on Artificial Intelligence and 12th Conference on Innovative Applications of Artificial Intelligence, 957-964.

Das, A., Ng, W.-K. & Woon, Y.-K. (2001). Rapid association rule mining. *Proceedings of the 10th International Conference on Information and Knowledge Management* 474-481, ACM Press.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly,* 13(3), 319-340. doi:10.2307/249008

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K. & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science,* 41(6), 391-407. doi:10.1002/(sici)1097-4571

Ding, H., Trajcevski, G., Scheuermann, P. I., Wang X. & Keogh E. J. (2008). *Querying and mining of time series data: experimental comparison of representations and distance measures*. Proceedings of the VLDB Endowment, 1542-1552. doi:10.14778/1454159.1454226

Do, T. D., Hui, S. C. & Fong, A. (2003). Mining frequent itemsets with category-based constraints. *Lecture Notes in Computer Science,* 2843, 76-86.

Faloutsos, C. & Lin, K., (1995). *Fastmap: A fast algorithm for indexing, datamining, and visualization of traditional and multimedia datasets*. In Proceedings of ACM-SIGMOD, 163-174.

Faloutsos, C., Jagadish, H. V., Mendelzon, A. O. & Milo T. (1997). *A signature technique for similarity-based queries*. Proceedings. Compression and Complexity of SEQUENCES 1997 (Cat. No.97TB100171), 2-20. doi:10.1109/SEQUEN.1997.666899

Faloutsos, C., Ranganathan, M. & Manolopoulos, Y. (1994). Fast subsequence matching in time-series databases. *ACM SIGMOD Record,* 23(2), 419-429. doi:10.1145/191843.191925

Fraccascia, L. & Yazan, D. M. (2018). The role of online information-sharing platforms on the performance of industrial symbiosis networks. *Resources, Conservation and Recycling,* 136, 473-485. doi:10.1016/j.resconrec.2018.03.009

Friedman, N., Geiger, D. & Goldszmidt, M. (1997). Bayesian network classifiers. *Machine Learning, 29*(2-3), 131-163. doi:10.1023/A:1007465528199

Goldberg, K., Roeder, T., Gupta, D. & Perkins, C. (2001). Eigentaste: A constant time collaborative filtering algorithm. *Information Retrieval Journal,* 4(2),133-151. doi:10.1023/A:1011419012209

Han, J. & Pei, J. (2000). Mining frequent patterns by pattern-growth: methodology and implications. *ACM SIGKDD Explorations,* 2(2), 30-36.

Herlocker, J. L., Konstan, J. A., Borchers, A. & Riedl, J. (1999). *An algorithmic* *framework for performing collaborative filtering*. Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 230-237.

Houtsma, M. & Swami, A. (1995). *Set-oriented mining of association rules in relational databases*. Proceedings of the 11th International Conference on Data Engineering, Taipei, Taiwan, 25-33. doi:10.1109/ICDE.1995.380413

Huang, C. K. & Li, C. J. (2018). *Substitution association rule mining in the stock market with big data analysis*. Paper Presented at the 24th Cross Strait Conference of Information Management Development and Strategy, The University of Hong Kong, Hong Kong.

Huang, T. C. K. (2013). Recommendations of closed consensus temporal patterns by group decision making. *Knowledge-Based Systems,* 54, 318-328. doi:10.1016/j.knosys.2013.10.003

Huang, T. C. K., Chen, Y. L. & Chen, M. C. (2016). A novel recommendation model with Google similarity. *Decision Support Systems,* 89, 17-27.

Isinkaye F.O., Folajimi Y.O. & Ojokoh B.A. (2015). Recommendation systems: principles, methods, and evaluation, *Egyptian Informatics Journal,* 16, 261-273.

ITIF. (October 12, 2018). ITIF technology explainer: what are digital platforms? Retrieved from https://itif.org/publications/2018/10/12/itif-technology-explainer-what-are-digital-platforms

Jagadish, H. V., Mendelzon, A. O. & Milo, T. (1995). *Similarity-based queries*. Proceedings of the 14th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, 36-45. doi:10.1145/212433.212444

Jaroszewicz, S. & Simovici, D. A. (2002). Pruning redundant association rules using maximum entropy principle. *Lecture Notes in Computer Science,* 2336, 135-142.

Keshavan, R. H., Montanari, A. & Sewoong, O. (2010). Matrix completion from a few entries. *IEEE Transactions on Information Theory,* 56(6), 2980-2998.

Kim, S. W., Park, S. & Chu, W. W. (2001). *An index-based approach for similarity search supporting time warping in large sequence databases*. Proceedings - International Conference on Data Engineering, 607-614. doi:10.1109/ICDE.2001.914875

Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R. & Riedl, J. (1997). GroupLens: applying collaborative filtering to Usenet news. *Communications of the ACM, 40*(3), 77-87. doi:10.1145/245108.245126

Koren, Y., Bell, R. & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *IEEE Computer, 42*(8), 30-37. doi:10.1109/MC.2009.263

Kotsiantis, S. & Kanellopoulos, D. (2006). Association rules mining: a recent overview. *GESTS International Transactions on Computer Science and Engineering,* 32(1), 71-82.

Kużelewska, U. (2013). Advantages of information granulation in clustering algorithms. In: Filipe J., Fred A. (Eds.), *Agents, and Artificial Intelligence. ICAART 2011. Communications in Computer and Information Science,* 271, 131-145, Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-29966-7_9

Landauer, T., Laham, D. & Foltz, P. (1998). Learning human-like knowledge by singular value decomposition: a progress report. Jordan, M. I., Kearns, M. J., and Solla, S. A. (Eds.), *Advances in Neural Information Processing Systems* 10, 45-51, Cambridge: MIT Press.

Lang K. (1995). *NewsWeeder: Learning to filter Netnews*. Proceedings of the 12th International Machine Learning Conference, 5000, 331-339.

Larose, D. T. & Larose, C. D. (2014). *Discovering knowledge in data: an introduction to data mining*. Hoboken: John Wiley & Sons Inc.

Lemire, D. & Maclachlan, A. (2005). *Slope one predictors for online rating-based collaborative filtering*. Proceedings of the 2005 SIAM International Conference on Data Mining, 471-475. doi:10.1137/1.9781611972757.43

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady, 10*(8), 707-710.

Li, Y. & Gopalan, R. P. (2004). Effective sampling for mining association rules. *Lecture Notes in Computer Science, 3339*, 391-401.

Li, Y.-M., Wu, C.-T. & Lai, C.-Y. (2013). A social recommender mechanism for e-commerce: combining similarity, trust, and relationship. *Decision Support Systems, 55*(3), 740-752.

Liang, T. P., Lai, H. J. & Ku, Y. C. (2006). Personalized content recommendation and user satisfaction: theoretical synthesis and empirical findings. *Journal of Management Information Systems,* 23(3), 45-70. doi:10.2753/mis0742-1222230303

Lika, B., Kolomvatsos, K. & Hadjiefthymiades, S. (2014). Facing the cold start problem in recommender systems. *Expert Systems with Applications,* 41(4-2), 2065-2073. doi:10.1016/j.eswa.2013.09.005

Linoff, G. S. & Berry M. J. A. (2011). *Data mining techniques: for marketing, sales, and customer relationship management, 3rd edition*. Wiley Computer, New York.

Loh, Wei-Yin. (2011). Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* 1, 14-23. doi:10.1002/widm.8

Lu, J., Wu, D., Mao, M., Wang, W. & Zhang, G. (2015). Recommender system application developments: a survey. *Decision Support Systems,* 74, 12-32.

Mannila, H., Toivonen, H. & Verkamo, A. I. (1994). *Efficient algorithms for discovering association rules*. In AAAI-94 Workshop on Knowledge Discovery in Databases, 181-192.

Manning, A. M. & Keane, J. A. (2001). Data allocation algorithm for parallel association rule discovery. *Lecture Notes in Computer Science,* 2035, 413-420.

Mcintyre, D. P. & Srinivasan, A. (2017). Networks, platforms, and strategy: emerging views and next steps. *Strategic Management Journal,* 38, 141-160.

McSherry D. (2004). Explaining the pros and cons of conclusions in CBR. In: Funk P., González Calero P.A. (Eds.), *Advances in Case-Based Reasoning. ECCBR 2004. Lecture Notes in Computer Science,* 3155, 317-330, Springer, Berlin, Heidelberg. doi:10.1007/978-3-540-28631-8_24

Mobasher, B., Jin, X. & Zhou, Y. (2004). Semantically enhanced collaborative filtering on the web. *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science),* 3209, 1-20. doi:10.1007/978-3-540-30123-3_4

Mooney, R. J. & Roy, L. (1999). *Content-based book recommending using learning for text categorization*. Proceedings of the 5th ACM Conference on Digital Libraries, 195-204.

Morgan, R. M. & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *Journal of Marketing,* 58(3), 20-38.

Moshfeghi, Y., Piwowarski, B. & Jose, J. M. (2011). *Handling data sparsity in collaborative filtering using emotion and semantic-based features*. Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '11, 625-634. doi:10.1145/2009916.2010001

Najmul Islam, A. K. M., Cenfetelli, R. T. & Benbasat, I. (2020). Organizational buyers' assimilation of B2B platforms: effects of IT-enabled service functionality. *The Journal of Strategic Information Systems,* 29(1), 1-28. doi:10.1016/j.jsis.2020.101597

Pan, C. & Li, W. (2010). *Research paper recommendation with topic analysis*. International Conference on Computer Design and Applications (ICCDA 2010), 4, 264-268. doi:10.1109/ICCDA.2010.5541170

Papagelis, M. & Plexousakis, D. (2005). Qualitative analysis of user-based and item-based prediction algorithms for recommendation agents. *Engineering Applications of Artificial Intelligence,* 18(7), 781-789.

Park, D. H., Kim, H. K., Choi, I. Y. & Kim, J. K. (2012). A literature review and classification of recommender systems research. *Expert Systems with Applications, 39*(11), 10059-10072. doi:10.1016/j.eswa.2012.02.038

Park, S. T. & Chu, W. (2009). *Pairwise preference regression for cold-start recommendation*. In Proceedings of the 3rd ACM Conference on Recommender Systems - RecSys ’09, 21-28. doi:10.1145/1639714.1639720

Parthasarathy, S. (2002). *Efficient progressive sampling for association rules*. 2002 IEEE International Conference on Data Mining, 2002. Proceedings, 354-361. doi:10.1109/ICDM.2002.1183923

Parthasarathy, S., Zaki, M. J., Ogihara, M. & Li, W. (2001). Parallel data mining for association rules on shared-memory systems. *Knowledge and Information Systems, 3*, 1-29.

Pazzani, M. J. (1999). A framework for collaborative, content-based, and demographic filtering. *Artificial intelligence review,* 5(6), 393-408.

Piatetsky, G. (2007). Interview with Simon Funk. *ACM SIGKDD Explorations Newsletter,* 9(1), 38-40. doi:10.1145/1294301.1294311

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P. & Riedl, J. (1994). *GroupLens: An open architecture for collaborative filtering of Netnews*. Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work, 175-186. doi:10.1145/192844.192905

Sakurai, Y., Yoshikawa, M. & Faloutsos, C. (2005). *FTW: Fast similarity search under the time warping distance*. Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, 326-337. doi:10.1145/1065167.1065210

Salton, G. & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing and Management,* 24(5), 513-523. doi:10.1016/0306-4573(88)90021-0

Sarwar, B., Karypis, G., Konstan, J. & Reidl, J. (2001). *Item-based collaborative filtering recommendation algorithms*. Proceedings of the 10th International Conference on World Wide Web, 285-295. doi:10.1145/371920.372071

Schafer, J. B., Frankowski, D., Herlocker, J. & Sen, S. (2007). Collaborative filtering recommender systems. In: Brusilovsky, P., Kobsa, A., and Nejdl, W., (Eds.), *The Adaptive Web, LNCS 4321* (pp. 291-324). Berlin Heidelberg (Germany): Springer. doi:10.1007/978-3-540-72079-9_9

Schuster, A. & Wolff, R. (2001). *Communication-Efficient distributed mining of association rules*. Proceedings of the 2001 ACM SIGMOD Int’l. Conference on Management of Data. Santa Barbara, California, 473-484.

Smyth, B. & Cotter, P. (2000). A personalised TV listings service for the digital TV age. *Knowledge-Based Systems,* 13(2-3), 53-59. doi:10.1016/s0950-7051(00)00046-0

Standing, S., Standing, C. & Love, P. E. D. (2010). A review of research on e-marketplaces 1997-2008. *Decision Support Systems,* 49, 41-51. doi:10.1016/j.dss.2009.12.008

Tang, P. & Turkia, M. P. (2006). Parallelizing frequent itemset mining with FP-trees. *Computers and Their Applications*, 1-6.

Toivonen, H. (1996). *Sampling large databases for association rules*. Proceedings of the 22nd VLDB Conference Mumbai (Bombay). India, 134-145.

Tsai, C. Y. & Shieh, Y. C. (2009). A change detection method for sequential patterns. *Decision Support Systems,* 46, 501-511.

Vlachos, M., Kollios, G. & Gunopulos, D. (2002). *Discovering similar multidimensional trajectories*. Proceedings 18th International Conference on Data Engineering, 673-684. doi:10.1109/ICDE.2002.994784

Wang, C. & Tjortjis, C. (2004). PRICES: An efficient algorithm for mining association rules, *Lecture Notes in Computer Science,* 3177, 352-358.

Wang, J., de Vries, A. P. & Reinders, M. J. T. (2006). *Unifying user-based and item-based collaborative filtering approaches by similarity fusion*. Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '06, 501-508.

Wojciechowski, M. & Zakrzewicz, M. (2002). Dataset filtering techniques in constraint-based frequent pattern mining. *Lecture Notes in Computer Science,* 2447, 77-83.

Yi, B. K. & Faloutsos, C. (2000). *Fast time sequence indexing for arbitrary Lp norms*. Proceedings of the 26th International Conference on Very Large Data Bases, 385-394.

Yi, B. K., Jagadish, H. V. & Faloutsos, C. (1998). *Efficient retrieval of similar time sequences under time warping*. Proceedings 14th International Conference on Data Engineering, 201-208. doi:10.1109/ICDE.1998.655778

Yuan, Y. & Huang, T. (2005). A matrix algorithm for mining association rules, *Lecture Notes in Computer Science,* 3644, 370-379.

Zhang, S., Yao, L., Sun, A. & Tay, Y. (2018). Deep learning based recommender system: a survey and new perspectives. *ACM Computing Surveys,* 1(1), 1-35.

Appendix

Appendix A. Part II Questionnaire items

Perceived Usefulness (PU)

PU1. NCRM can help me earn money by investing in stocks.

PU2. Using NCRM can improve the profitability of my stock investment.

PU3. The benefits of NCRM far outweigh its disadvantages.

PU4. On the whole, there are advantages to using NCRM to invest in stocks.

PU5. Overall, NCRM is useful.

Trust (TR)

TR1. I agree NCRM is reliable.

TR2. I agree NCRM is trustworthy.

TR3. I believe that using NCRM is worthy of a guarantee.

TR4. When NCRM continues to update information and expand applicable stocks, it will help my stock investment profit.

TR5. When I make the stock investment decisions, I want to refer to NCRM's recommendation results.

User satisfaction (US)

US1. I am satisfied with the results of NCRM's recommendation.

US2. I am satisfied with the system functions of NCRM.

US3. I am satisfied with the recommended quality of NCRM.

US4. Using NCRM can improve the efficiency of my stock investment decision-making.

US5. Using NCRM can reinforce my decision to integrate other stock analysis information.

US6. Overall, I am satisfied with the experience of using NCRM.