The Thirty-Second International Florida
Artificial Intelligence Research Society Conference (FLAIRS-32)

A Comparison of Three Recommender Strategies for Facilitating Person-Centered Care in Nursing Homes

Nathan Martindale,[1] Gerald C. Gannod,[1] Katherine M. Abbott,[2] Kimberly Van Haitsma[3]
[1]Department of Computer Science, Tennessee Technological University
[2]Department of Sociology and Gerontology, Miami University-Oxford
[3]College of Nursing, Pennsylvania State University, Polisher Research Institute, Abramson Center for Jewish Life

Abstract

The *Preferences for Everyday Living Inventory* (PELI) is a 72-question instrument used for helping nursing homes assess person-centered care. In particular, the approach allows residents to express their preferences for both care and activities in order to provide direct care workers with insights on how to best provide a high-quality living experience. Among the challenges of using the PELI is its length: 72 questions give rise to issues of survey fatigue while also creating a workflow bottleneck for those providing care. In this paper we explore and evaluate the use of three different recommender strategies that we have applied to the PELI. In particular, we present the use of both *rule-based* and *neighborhood-based* collaborative filtering in order to make recommendations on which preference questions to present to a resident. We illustrate the approaches by providing a domain-specific example, and then compare the approaches across a number of performance and quality metrics.

Introduction

As nursing homes look to improve the quality of care of their residents, there has been a move towards person-centered care that focuses on, among other things, the preferences of individual residents. One such person-centered care approach is to use the Preferences for Everyday Living Inventory for Nursing Home residents (PELI-NH), a 72-question instrument that identifies resident preferences. In past work, we described the Care Preference Assessment of Satisfaction or ComPASS, a system that has been developed to facilitate delivery of the 72-question instrument (Gannod et al. 2018a). However, due to issues related to survey fatigue (as well as literal fatigue) of the residents and the workload constraints of the direct care workers that administer the survey, using the full 72-question instrument can be prohibitive. As such, in its current iteration, ComPASS focuses on a 16 question subset of the PELI-NH that all nursing home providers are required to complete called the Minimum Data Set 3.0 Section F (MDS 3.0). The system guides interviewers and interviewees (i.e., direct care workers and residents) through an interview in order to identify preferences, assess a resident's level of satisfaction on each, and to explore a set of nested questions that are used to learn specifics about how the experience around a given preference can be met or improved for a resident.

As the system moves closer to widespread launch, it is our desire to expand the use of ComPASS to include support for the full 72-question instrument. To this end, we have developed a recommender system that is based on the use of rule-based collaborative filtering in the form of the Apriori algorithm combined with logistic regression (Gannod et al. 2018a). In this approach, we used the 16 core MDS questions to identify frequent itemsets found in a dataset containing 510 interviews from 255 individual respondents. We found in the evaluation of that approach that the algorithm was able to achieve an average precision of 78% and an average recall of 82%. Using these results as a baseline, we have applied and compared two additional recommender variants that use neighborhood-based *item-item* and *user-user* collaborative filtering. In this paper, we present each of these variations and describe an evaluation of the approaches with respect to the same dataset mentioned above.

The remainder of this paper is organized as follows: it discusses background information relevant to the research we are conducting, including a brief discussion of recommender systems, an overview of the Care Preference Assessment of Satisfaction System (ComPASS), and a short survey of related work. In the *approach* section we present the techniques used in ComPASS to provide recommendations for resident preferences as well as the dataset upon which we apply these techniques. The *Example* section illustrates the use of neighborhood-based collaborative filtering on our target dataset. Finally, we provide a comparative analysis of the three techniques used in our work, and close by drawing conclusions and suggesting future investigation.

Background

Recommender Systems

Recommendation systems are powerful tools in many modern applications, both from the perspective of a user as well as the perspective of a business. Systems such as Amazon and Netflix are both prevalent examples of the usage of powerful recommenders, with Amazon recommending other products one might like and Netflix recommending shows or movies that it predicts one will enjoy (Aggarwal 2016a). These systems are useful in that they improve sales for the

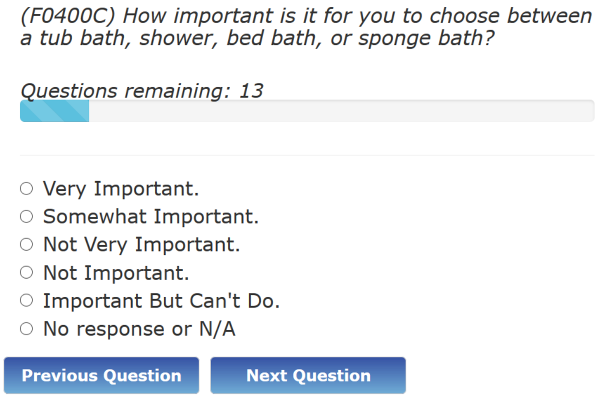

Figure 1: Preference Question Format

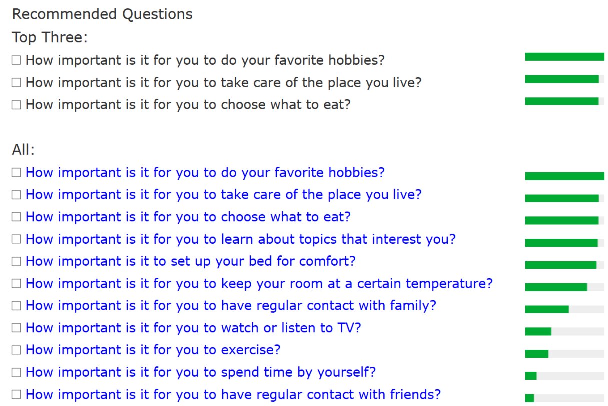

Figure 2: Recommended Preferences

business, and are able to provide novel or useful content to a user, respectively.

An important approach used in many recommender systems is *collaborative filtering*, which bases predictions off of historical user-item interactions. There are several subtopics within collaborative filtering, two of which are relevant to our work: *rule-based* and *neighborhood-based*. Rule-based algorithms use association rules to represent correlated items or things that are frequently bought together. Neighborhood-based algorithms predict based on the observation that similar users behave similarly in rating behavior and similar items are generally rated similarly (Aggarwal 2016a).

These systems are enabled by collecting large amounts of data on what users like, either explicitly through rating systems, or implicitly by keeping track of what items or pages a user clicks on. This data can then be utilized in various approaches to infer other items or pages that a user or other users might similarly take interest in or rate highly (Aggarwal 2016a).

ComPASS

The *Care Preference Assessment of Satisfaction* System (ComPASS) is a web-based implementation of an Excel tool created to support the use of the PELI-NH with quality improvement. ComPASS was developed using the Ruby on Rails framework and seeks to provide a scalable, user-friendly system to facilitate data collection and data analytics of long-term care residents' preferences. ComPASS supports the interview process by providing an interface similar to the one shown in Figure 1. In our initial offering of the system, each interviewer is required to ask the 16 MDS 3.0 subset of questions.

In our work, we have also added additional nested questions that explore specifics regarding the top-level preference question. For instance, in the MDS question regarding preference over type of bathing, nested questions explore topics such as "What type of bathing do you prefer?" or "Would you like a certain room temperature when you bathe?" These additional questions are open ended and provide a way for a care provider to capture details about preferences for incorporation into care plans. After a set of questions has been answered, the interviewer can proceed to the next question by clicking a button at the bottom of the interview. Interview progress is visually marked by a progress bar and is automatically saved at each step, allowing interviewers to pause an interview for any reason and resume it at a later point. ComPASS supports display of resident satisfaction reports in order to provide input and insight into the care of the resident during individual quarterly care conferences.

Our initial investigations have focused on using a *rule-based collaborative filtering* approach, with responses to preference questions used in a manner analogous to selection of products in a market basket (Gannod et al. 2018a). The process of generating a recommendation occurs following an interview. Once a resident has provided responses to the initial 16 MDS questions, the list of recommended questions and categories from the rest of the 72 - 16 = 56 preferences are generated and displayed, and the interviewer can select which questions to add to the rest of the interview.

As shown in Figure 2, top-*3* items are displayed followed by a ranking of questions identified during the logistic regression phase of our approach. The bar to the right of the questions provides a visual representation of the score used during application of logistic regression to rank the quality of a rule. In this context, the score metric is obtained via the use of support and confidence.

Related Work

Different strategies for implementing recommender systems have been suggested, including the use of association rules (Cakir and Aras 2012; Bendakir and Aimeur 2006). In the healthcare domain, the idea of using data mining is not a new one (Patel et al. 2009; Simovici 2012). Applications have included treatment effectiveness and condition identification (Koh and Tan 2005). Hu, for instance, has suggested using the Apriori Algorithm for mining medical data as a means for diagnosis of conditions (Hu 2010). In the area of long-term care, we are unaware of work being done to understand resident preferences using recommendations.

Approach

The length of the Preferences for Everyday Living Inventory (PELI) presents two challenges: it can create literal and figurative survey fatigue in residents, and creates a workflow bottleneck for the staff that administer the instrument. As such, our work is focused on providing a mechanism for identifying recommended preferences for a given resident, providing a filter for focusing PELI survey interviews. In other words, we merely identify preference questions to pose but do not go the extra step of dictating what a resident's preference is. In this section, we describe three collaborative filtering (CF) approaches used to provide recommendations on preferences: a *rule-based* method that combines the use of the *Apriori* algorithm paired with logistic regression (Gannod et al. 2018a), and two *neighborhood-based* methods, including *user-user* and *item-item* collaborative filtering.

Dataset

The Preferences for Everyday Living Inventory for Nursing Home residents (PELI-NH) is a survey instrument containing 72 questions across five domains. The PELI-NH was administered by research assistants using face-to-face interviews with 342 individuals from 28 locations in the suburbs of a major metropolitan East Coast area of the United States. Of the 342 original participants in the baseline (T1), 255 completed a second interview (T2) for an attrition rate of appromimately 25%. This attrition rate was due to a number of factors including death, transfer, change in cognitive ability, withdrawal, or change in the medical stablity of the resident over the 3 months between conducting T1 and T2. These participants were deemed to have a Mini-Mental State Examination (MMSE) $\geq$ 13 (Folstein, Folstein, and McHugh 1975), with an average MMSE score of 24.6 (3.9 SD). Informed consent for participation was established in-person in both the initial interview and follow-up. Demographically, the residents were mostly widowed (44%) white (77%) females (67.8%) with a mean age of 81 years (11.21 SD) and a high school education (54%).

Each entry in the dataset consists of 72 responses to importance questions ranging from a score of 1 - 4, where 1 represents "very important", and 4 represents "not important at all."

Rule-Based Collaborative Filtering

The *Apriori* algorithm (Agrawal and Srikant 1994) is a method for generating association rule sets from co-occuring values in a dataset. Association rules follow the form $\{A\} \rightarrow \{B\}$, where presence of the members in the antecedent $\{A\}$ imply that the member(s) of the consequent $\{B\}$ are also likely present. These rules can be used for generating recommendations and were experimented with in previous work (Gannod et al. 2018a), (Gannod et al. 2018b).

Rules are generated by finding frequent itemsets, or items that are correlated or have a high probability of appearing together in a particular itemset. Several metrics are used to filter out or determine the quality of a particular rule, namely support, confidence, and lift/interest (Tan, Kumar, and Srivastava 2004).

In the application to ComPASS, a set of association rules was generated from the dataset, where antecedents were filtered to include only preferences on the 16 MDS questions, and consequents filtered to include only non-MDS preferences. To determine recommendations, a target user's responses to the 16 MDS questions were compared against the association rules, and the rules with corresponding matches in the antecedents were ranked by a combination of highest support, confidence, and lift. On average, in our experiments this approach achieved a precision of 78% and a recall of 82%. (See results section below for more detail.)

Neighborhood-based Collaborative Filtering

Neighborhood-based collaborative filtering is a popular approach to generating recommendations due to its simplicity and efficiency (Ning, Desrosiers, and Karypis 2015). This method works by taking user-item ratings and determining the "neighborhood" of a prediction target by finding users or items that are most similar, and using ratings of those similar users to fill in or make a prediction about the target's ratings. There are two common methods of implementing this that we tested, *user-user* (i.e., *user-based*) and *item-item* (i.e., *item-based*) CF approaches. In *user-based* CF, a user's neighborhood is constructed by taking the items the target user has rated, and finding other users who have rated those same items most similarly. The *item-based* CF method constructs an item's neighborhood by finding the correlations in how the items are generally rated. The target item is then predicted from the data in its neighborhood.

In our dataset, we represent every resident interview as a row (or "user"), and every preference question as a column (or "item"). The primary difference between *user-based* and *item-based* CF is the direction in which entities are clustered into neighborhoods; in *user-based* CF, a neighborhood is composed of similar rows, and in *item-based* CF it is composed of similar columns. This means that *user-based* CF identifies similar residents, while *item-based* CF identifies similar preferences.

User-Based. In our *user-based* CF implementation, the 16 MDS responses of a given resident are compared to those responses of every other resident. The top N most similar residents are collected based on a similarity function mentioned below, and their responses on the 56 additional questions aggregated and averaged. These averages are classified based on a threshold value, with the resultant binary $\{important, not\ important\}$ values returned as the predicted responses of the target user.

Item-Based. In our *item-based* CF implementation, for each individual preference item, the top N other similar preference items, or items most commonly rated the same throughout the rest of the dataset, are collected, determined by a similarity metric with respect to the rest of the data. The weighted average of the target user's responses on those similar items is calculated and classified in the set $\{important, not\ important\}$ based on a threshold and used as the predicted response for the target preference item.

The idea behind the item-based CF is that if there are preferences that tend to be rated similarly in the dataset, we can use a resident's responses on a few of those similar preferences in order to predict how they will respond to another similar target item. That is, a resident's response to an MDS question is used to provide the rating for non-MDS preferences that have been found to be similar to the given MDS question. For example, if "Phone privacy", "Lock up items", and "Privacy" tend to be rated closely to each other by each resident, then when a new resident gives a response to "Phone privacy" and "Lock up items", we can predict that they will likely respond to "Privacy" similarly.

Calculating Similarity. We used two methods for calculating similarity in our experiments: the Pearson correlation coefficient (Resnick et al. 1994; Adomavicius and Tuzhilin 2005) and cosine similarity (Breese, Heckerman, and Kadie 1998; Adomavicius and Tuzhilin 2005).

The Pearson correlation coefficient returns a result in $[-1, 1]$, and cosine similarity returns a result in $[0, 1]$. In both cases, values closer to 1 are more similar, thus finding the top N most similar entries consists of calculating $sim(target, x)$ for every user/item x in the dataset, and returning the N entries with the highest values.

Aggregating and Predicting. Once the most similar entries have been collected, there are several commonly used methods to aggregate similar rankings to synthesize the prediction. The simplest is to take either an average of the rankings or an average weighted by the degree of similarity (Adomavicius and Tuzhilin 2005). A more sophisticated variant, which was utilized in this paper for user-based recommendations, is a weighted average with *mean-centering* (Aggarwal 2016b). Mean centering takes the weighted average of the similar ratings and adding it to the average of all of the user's ratings - accounting for residents who are more likely in general to consider items important, or vice versa.

Example

To demonstrate how the recommendation process works in the *user-based* CF approach, we have displayed the MDS responses of four residents in Table 1, where T is the target user, or resident whose preferences on the non-MDS items we wish to predict, and S_1, S_2, and S_3 are the three most similar residents based off of their MDS responses, using the Pearson correlation coefficient. Each value represents that user's stated importance level for the given preference (i.e., 1 = very important, 4 = not important at all).

Calculating T's predicted preferences for each non-MDS item is done by taking a mean centered weighted average of S_1, S_2, and S_4's preferences for that item. Running the recommender with the resident T in Table 1 results in a set of predictions on all 56 items, some of which are shown in Table 2. Classification into $\{$*important*, *not important*$\}$ is done based off of a threshold, or in this case 2.5. Any predicted value in column **Pred** less than 2.5 is a prediction that the resident will consider it important, and any value greater than that indicates a prediction that it not important. Actual responses are classified similarly, where a response of 1 or 2 is considered important, and 3 and 4 is not.

The table demonstrates that when similar residents rate something as important, the prediction is closer to 1, and more likely that the target resident will also consider it important. "Wake up time" and "Tobacco" are both good examples of this - given the preference "Wake up time" ("How important is it to you to choose when to get up in the morning?"), all three similar residents indicated that it was very important to them. The prediction was 0.848, or that it would be important, and when compared to the actual dataset, the target resident had indeed responded that it was very important. On the other hand, all three similar residents stated that tobacco was not important to them, and the prediction and actual response of the target resident agreed.

Also shown in Table 2 are some examples where the system did not make the correct prediction - in particular "Bathroom needs" ("how important is it that your daily caregiver knows your needs when going to the bathroom?"). All three similar residents indicated this was very important to them and so the recommender predicted the target resident would respond similarly, yet in the resident's actual response they stated that it was not important to them.

The particular run containing the set of predictions shown for the target resident T in Table 2 based on their MDS responses in Table 1 resulted in an accuracy of 85.71%, a precision of 88.0%, and a recall of 95.65%.

Question	T	S_1	S_2	S_3
Bath type	1	1	1	1
Clothes	2	1	1	1
Personal belongings	1	1	1	1
Bedtime	1	1	1	1
Care discussions	1	1	1	1
Phone privacy	1	1	1	2
Lock up items	2	1	1	1
Snacks	1	2	1	3
Groups of people	1	1	1	1
Religion	4	4	2	3
Fresh air	1	1	1	1
Animals	2	1	1	2
News	1	1	1	2
Reading materials	1	1	1	1
Music	1	1	1	1
Favorite activities	1	1	1	1

Table 1: Residents similar to target resident *T* using MDS

Evaluation

Neighborhood-based collaborative filtering experiments were run, varying the type of neighborhood collection (user-based versus item-based), similarity method (Pearson correlation coefficient versus cosine similarity), N, and the classification threshold. The results from these experiments were compared against those listed in our previous work (Gannod et al. 2018a) utilizing the Apriori algorithm.

A number of initial experiments were run to determine the best neighborhood size to use for user-based and item-based respectively, which consisted of multiple runs of the recommender utilizing only the Pearson correlation coefficient. These neighborhood sizes ranging from 3 to 12, and

Question	Pred	T	S_1	S_2	S_3	Correct
Name	1.209	1	1	2	1	✓
Wake up time	0.848	1	1	1	1	✓
Hair care	1.569	1	1	3	1	✓
Caregiver gender	3.541	3	4	4	3	✓
Bathroom needs	0.848	4	1	1	1	
Alcohol	2.544	2	1	4	3	
Tobacco	3.848	4	4	4	4	✓
Time with friends	1.463	1	1	1	3	✓
Privacy	0.848	1	1	1	1	✓
Roommate	3.848	4	4	4	4	✓
Take out	1.823	2	1	2	3	✓
Alone time	1.516	1	1	2	2	✓
Gifts	1.488	2	2	1	2	✓
Outdoor tasks	1.820	1	3	1	2	✓
Television	1.155	3	1	1	2	
Movies	2.876	1	2	4	3	
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 2: Subset of Non-MDS Predictions

from the results an N of 5 was selected as the empirical elbow for user-based recommendation, and an N of 6 was selected as the empirical elbow for item-based recommendation. The threshold for most experiments was kept in the center of the target range at 2.5. While lowering this threshold does increase precision, recall quickly plummets, and thus significantly lowers accuracy as can be seen in the last two entries of Table 4.

The evaluation metrics for all neighborhood-based collaborative filtering experiments are obtained by averaging the confusion matrix results from each resident prediction. (See Table 3 for an example of the confusion matrices that user-based and item-based output for a particular resident prediction.) The predictions are run using the data from the first interview for all residents in the set except for the target resident. Since the end results of all three approaches are binary classifications, confusion matrices and their corresponding accuracy, precision, specificity, recall, and F1 scores can be directly compared.

		Predicted T	Predicted F
Actual	T	38	0
	F	13	5

(a) user-user

		Predicted T	Predicted F
Actual	T	37	1
	F	17	1

(b) item-item

Table 3: Confusion Matrix Examples

Results and Comparison

Table 4 contains the scores of some of our experiments, where ACC is accuracy, PPV is precision, SPC is specificity, and TPR is recall. The Apriori results were pulled from previous work (Gannod et al. 2018b). The results chosen here for comparison were selected due to having the highest combination of accuracy and F1 score, the primary metrics by which the new algorithms were ranked. The Apriori model that achieved 72.61% ACC was run filtering out rules not meeting a minimum confidence of 80%, a minimum support of 1%, and a minimum lift or interest measure of 1.1%. The second was run filtering a minimum confidence of 75%, a minimum support of 1%, and a minimum lift of 1.1%.

In comparing the Pearson and Cosine similarity methods used in neighborhood-based collaborative filtering, we determined that the Pearson correlation coefficient generally had a higher accuracy and F1 score in both of the user-based and item-based approaches. The correlation coefficient had on average, a .5% higher accuracy, and a 1% higher F1 score.

User-based recommendation experiments outperformed item-based by approximately 4% in accuracy and 5% in precision. The best user-based experiment results, which had the highest accuracy and precision-recall balance (F1), also outperformed our previous best Apriori results by 5% in accuracy, and 2% in F1 score.

In addition to better evaluation scores, the neighborhood-based algorithms have a few advantages over using Apriori in a live system. The Apriori algorithm has a significant training component - before the recommendation system can be used, it must go through the process of analyzing all of the interview data and finding the association rules. This is a resource-intensive process, and if new data is to be continually incorporated into recommendations, it would likely need to be retrained frequently. Neighborhood-based algorithms however, with no pre-processing necessary, can run online. This allows new recommendations to continually take the latest data into account (Rish 2001).

The neighborhood-based algorithms performed better on less data than the Apriori system did. The dataset contains 510 interviews total, two for each of the 255 residents, and in the Apriori experiments the interviews were treated as independent in order to have more data and thus a larger set of association rules. In contrast, all neighborhood-based experiments were run using only the first interview of each resident, effectively utilizing only half of the data needed with the Apriori approach.

Threats to Validity

In our previous work, two major threats have been the small size of the dataset used, as well as the assumption that the two separate interviews for each resident can be considered independent (Gannod et al. 2018a). In the work described in this paper we have eliminated the independence assumption in the neighborhood-based collaborative filtering, but are still working with a relatively small dataset.

Conclusions and Future Investigations

Our goal in this work is to provide a means for streamlining both data collection and reporting in order to effectively impact the lives of nursing home residents. While the use of ComPASS without the recommendation engine is a step towards that goal, we believe that the recommendations will facilitate greater impact by helping the staff identify preferences more rapidly. In this paper, we have demonstrated that recommendations can be supported to a relatively high

ACC	PPV	SPC	TPR	F1	Type	SimMethod	N	Thresh
77.58%	78.98%	35.84%	91.10%	**84.60%**	User	Pearson	5	2.5
76.86%	78.04%	33.69%	90.69%	83.85%	User	Cosine	5	2.5
73.68%	73.70%	7.08%	**94.38%**	82.77%	Item	Pearson	6	2.5
73.09%	72.69%	8.08%	92.45%	81.39%	Item	Cosine	6	2.5
72.61%	76.65%	24.48%	90.14%	82.82%	Apriori	-	-	-
69.79%	79.19%	41.89%	80.21%	79.53%	Apriori	-	-	-
73.54%	81.63%	57.80%	75.15%	78.26%	User	Pearson	5	2.0
68.25%	**82.07%**	**68.40%**	63.68%	71.72%	User	Pearson	5	1.75

Table 4: Comparison of Strategies

degree of success. However, we must still perform direct validation on this hypothesis by engaging residents directly to determine whether the recommendations ultimately do align with actual preferences.

To date, we have experimented primarily upon the use of off-the-shelf collaborative filtering approaches. In future investigations, we are interested in expanding our experiments to include variations in the available dataset so as to study various effects that may be impacted by the population of respondents. This would include determining whether preferences may vary regionally, or by other demographic variables. Finally, we are interested in determining whether other collaborative filtering approaches may provide additional gains in either quality or performance.

Acknowledgements

The authors would like to thank Kaley White for her preliminary work on the neighborhood-based collaborative filtering implementation and initial experiments.

References

Adomavicius, G., and Tuzhilin, A. 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING* 17(6):734–749.

Aggarwal, C. C. 2016a. *Recommender Systems*. Cham: Springer International Publishing.

Aggarwal, C. C. 2016b. *Recommender Systems*. Springer International Publishing, 1 edition.

Agrawal, R., and Srikant, R. 1994. Fast Algorithms for Mining Association Rules in Large Databases. In *Proceedings of the 20th International Conference on Very Large Data Bases*, VLDB '94, 487–499. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Bendakir, N., and Aimeur, E. 2006. Using association rules for course recommendation. In *Proceedings of the AAAI Workshop on Educational Data Mining*, 31–40. AAAI.

Breese, J. S.; Heckerman, D.; and Kadie, C. 1998. Empirical analysis of predictive algorithms for collaborative filtering. *Aaai* 43–52.

Cakir, O., and Aras, M. E. 2012. A recommendation engine by using association rules. *Procedia -Social and Behavioral Sciences* 62:452–456.

Folstein, M. F.; Folstein, S. E.; and McHugh, P. R. 1975. "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* 12(3):189–198.

Gannod, G. C.; Abbott, K. M.; Haitsma, K. V.; Martindale, N.; Jennings, R. A.; and Long, C. N. 2018a. Using Machine Learning to Facilitate the Delivery of Person Centered Care in Nursing Homes.

Gannod, G. C.; Abbott, K. M.; Van Haitsma, K.; Martindale, N.; and Heppner, A. 2018b. A Machine Learning Recommender System to Tailor Preference Assessments to Enhance Person-Centered Care Among Nursing Home Residents. *The Gerontologist*.

Hu, R. 2010. Medical Data Mining Based on Association Rules. *Computer and Information Science* 3(4).

Koh, H. C., and Tan, G. 2005. Data mining applications in healthcare. *J Healthc Inf Manag* 19(2):64–72.

Ning, X.; Desrosiers, C.; and Karypis, G. 2015. A comprehensive survey of neighborhood-based recommendation methods. In *Recommender Systems Handbook, Second Edition*.

Patel, V. L.; Shortliffe, E. H.; Stefanelli, M.; Szolovits, P.; Michael, R.; Bellazzi, R.; Abu-hanna, A.; Larson-meyer, D. E.; Redman, L.; Heilbronn, L. K.; and Martin, C. K. 2009. The coming of age of artificial intellingence in medicine. *Artificial Intelligence* 46(1):152–159.

Resnick, P.; Iacovou, N.; Suchak, M.; Bergstrom, P.; and Riedl, J. 1994. GroupLens: An Open Architecture for Collaborative Filtering of Netnews. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*, CSCW '94, 175–186. New York, NY, USA: ACM.

Rish, I. 2001. An empirical study of the naive Bayes classifier. *IJCAI 2001 workshop on empirical methods in artificial intelligence* 22230:41–46.

Simovici, D. A. 2012. Data Mining of Medical Data: Opportunities and Challenges in Mining Association Rules. In *IALS*.

Tan, P.-N.; Kumar, V.; and Srivastava, J. 2004. Selecting the right objective measure for association analysis. *Information Systems* 29:293–313.