Collaborative Filtering for Identifying Prescription Omissions in an ICU

Anima Singh and John Guttag
Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, U.S.A.

Keywords: Machine Learning, Collaborative Filtering, Latent Factor Models, Intensive Care Units, Prescription Omissions.

Abstract: Medication errors in critical care are frequent and can lead to adverse consequences. One important category of errors is prescription omission, i.e., failure to prescribe a potentially useful medication. Studies have shown that failure to prescribe a medication can result in adverse consequences leading to patient morbidity or even patient mortality (Aspden et al., 2007; Olsen et al., 2007). In this paper, we present a machine learning based approach to building a system that can be used to provide physicians with an ordered list of possible omissions. We investigated three different collaborative filtering approaches as well as simple prevalence and co-occurrence methods. When evaluated on over 19,000 ICU admissions, each of the CF approaches outperformed both prevalence and co-occurrence based methods. This work highlights the importance of capturing a multi-scale view of the prescription data for the task of identifying omissions. Our results suggest that latent factor models and neighborhood models are better at capturing different kinds of omissions. Latent factor models demonstrated improved performance on identifying omission of rarely prescribed medications while neighborhood models were slightly better at identifying omissions of commonly prescribed medications.

1 INTRODUCTION

Medication-related errors account for 78% of the medical errors in the intensive care unit (ICU) (Rothschild et al., 2005). Patients in critical care are particularly vulnerable to such errors. First, medication errors are more frequent in an ICU because patients are given twice as many medications as patients outside of a critical care setting (Cullen et al., 1997). Second, patients in an ICU are high-risk patients, and therefore any error is more likely to have serious consequences.

The inpatient medication use-process can be broadly categorized into five phases: prescription, transcription, preparation, dispensation and administration. A medication error can occur in any one of the stages in the process. According to the Institute of Medicine (IOM), the prescription stage is most frequently associated with errors (Aspden et al., 2007). One important category of errors is prescription omission, i.e., failure to prescribe a potentially useful medication. Studies have shown that failure to prescribe a medication can result in adverse consequences leading to patient morbidity or even patient mortality (Aspden et al., 2007; Olsen et al., 2007). In addition, these errors translate into longer hospital stays. In this paper, we focus our discussion on ways to detect possible prescription omissions in an ICU.

Medication safety can be improved by incorporating automated systems that can alert caretakers about likely medication errors. Because of the increasing availability of electronic medical records, the prospect of developing automated systems that can use patient information to help guide prescription is promising. Electronic medical records contain information about a patient's demographics, medications, diagnosis and other medical information. While it is challenging for a caretaker to process all the information, computerized systems can process such information, and provide useful insights in detecting errors. In particular, these systems can provide significant benefits in reducing prescription omission errors by providing suggestions about likely omissions.

Although prescription omissions are a serious problem, most admissions do not have omissions. Given a list of medications a patient is currently on, an ideal system should alert a caretaker if there is a high likelihood of an omission. In addition, it should offer suggestions about which medication might be missing.

In this paper, we investigate development of automated systems that can learn to perform the task of identifying possible prescription omissions. One of

Singh A. and Guttag J..
Collaborative Filtering for Identifying Prescription Omissions in an ICU.
DOI: 10.5220/0004233900580064
In *Proceedings of the International Conference on Health Informatics* (HEALTHINF-2013), pages 58-64
ISBN: 978-989-8565-37-2

the challenges in doing this is that there is no reliable data set containing ground truth. There are many data sets that contain information about which drugs are prescribed, but as far as we know there are no significant data sets that contain information about which drugs should have been prescribed but were not. Therefore, rather than trying to directly learn typical errors, we try to learn which drugs are typically prescribed together and how prescriptions relate to diagnoses. The system might learn, for example, that patients with a diagnosis of hypertension who are prescribed Atenolol (a beta blocker) are usually also prescribed Chlorthalidone (a diuretic). We then use the absence of a 'typical' association as an indication that a drug may have been omitted.

The high variation in the ways in which drugs are used, combined with the relatively low incidence of omissions makes it implausible under most circumstance to predict with high certainty that a drug is missing. Therefore we choose to provide a ranked list of the most probable omissions. A physician could look at the first few items on this list as candidate drugs to think about prescribing rather than as an indication that he or she has likely made an error.

We use collaborative filtering (CF) as the basis of our approach to learning. This seems a promising choice for at least two reasons: 1) CF has been widely used in recommender systems that infer information about a user's interest in items based on past behavior and preferences, and then provide rankings for other items. 2) CF methods are robust at handling sparse datasets, i.e., datasets in which most of the entries are zeros or missing. This is important in our context, since even the sickest patients receive only a tiny subset of the possible drugs during an admission.

Hasan *et al*. have explored neighborhood based and logistic regression based CF to detect omissions in an outpatient setting (Hasan et al., 2011). Our work focuses on capturing a multi-scale view of the data. In addition to a neighborhood approach, we investigate latent factor models that aim to learn prescription patterns from data. We also explore models that combine both methods.

We evaluated our approach on data from MIMIC II (Saeed et al., 2002), a clinical database of ICU patients. During the learning phase, we assumed that all records accurately reflected the drugs that were prescribed.

Since the database does not include an indication of whether or not a drug was omitted, we simulated omissions by deleting drugs that had been prescribed. Ideally, one would use a deletion model that mirrored the kinds of mistakes actually made in practice. For example, if physicians more often omit rarely prescribed drugs, our deletion model should contain that bias. Unfortunately, there does not seem to be any publicly available information about the actual pattern of omission errors. This led us to use a simple omission model in which we remove one of the prescribed medications uniformly at random.

Using this model, we evaluate different types of CF approaches, and compare them to the performance of prior and co-occurrence based methods. We use each method to rank each medication that was not prescribed based on the likelihood that it is omitted. We then evaluate each method based on the rank of the actually omitted medication in the list. The lower the rank, the better the approach at identifying omissions.

When evaluated on over 19,000 ICU admissions, each of the CF approaches outperformed both prior and co-occurrence based methods. Overall, a latent factor CF approach that captures prescription patterns performed better than a nearest neighbor approach. The latent factor models outperformed the nearest neighbor approach at identifying omitted drugs that are prescribed less often. On the other hand, neighborhood models performed better at identifying omissions of commonly prescribed drugs. This suggests that combining the two approaches can potentially improve performance.

The remainder of the paper is organized as follows. Section 2 discusses problem formulation and notations used in the rest of the paper. Section 3 describes different methods that we investigated. Section 4 summarizes the data used for this study. Our experimental methodology and results are presented in Section 5. Finally, Section 6 presents the summary and future work.

2 PROBLEM FORMULATION

In this section, we present the formal representation of our problem and introduce notations. We consider the first 24 hours of a hospital admission during which a patient was admitted to an ICU. We represent an admission as a list of prescribed medications and the diagnoses of the patient at the time of admission. In the rest of the paper, we will refer to medications and diagnoses as as items.

Let,

- $\{\mathbf{a_1}, \mathbf{a_2}, ..., \mathbf{a_A}\}$ be a set of vectors that represent ICU admissions in the data
- $\{m_1, m_2, ..., m_M\}$ be the set of medications
- $\{d_1, d_2, ..., d_D\}$ be the set of diagnoses
- $I = M + D$ be the number of items

We represent the j^{th} admission $\mathbf{a_j}$ as a binary vector of length I where,

For $i \leq M$,

$$\mathbf{a_j}_i = \begin{cases} 1, & \text{if medication } m_i \text{ is prescribed.} \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

For $i > M$,

$$\mathbf{a_j}_i = \begin{cases} 1, & \text{if the patient has a diagnosis } d_{i-M}. \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

The information about all admissions is represented as a $I \times A$ matrix R, where $R_{ij} \in \{0,1\}$ (Figure 1).

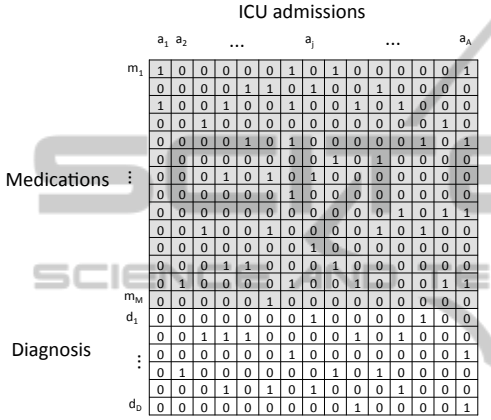

Figure 1: An $I \times A$ binary matrix R where $I = M + D$.

3 METHODS

In this section, we discuss the methods we investigated for identifying prescription omission.

For an admission $\mathbf{a_j}$, each method considers all medications m_i for which $R_{ij} = 0$. It then assigns a score to each medication. Based on the score, the method generates a ranked list of possible omissions as the output.

3.1 Prior based Approach

The prior based approach assigns a score that is equal to the prior of the medication based on ICU admissions in the training set $\mathfrak{T}$. The prior of medication m_i, $\mathfrak{P}_{m_i}$, is defined as the fraction of ICU admissions in $\mathfrak{T}$ that are prescribed m_i.

$$score(i,\cdot) = \mathfrak{P}_{m_i} \quad (3)$$

A prior based approach will perform well when medications that are prescribed frequently are most likely to be omitted. We use this approach as a baseline to evaluate performance of other models that we investigate.

3.2 Co-occurrence based Approach

Our co-occurrence based approach assigns a score for medication m_i based on how many times it co-occurs with other medications that are observed in admission $\mathbf{a_j}$. The co-occurrence is calculated based on admissions in the training set $\mathfrak{T}$.

$$score(i,j) = \sum_{d \in \{z | R_{z,j}=1, z \leq M\}} \sum_{\mathbf{a_k} \in \mathfrak{T}} R_{i,k} R_{d,k} \quad (4)$$

This approach is the simplest way to capture some information about which sets of medications are prescribed together. However, it only captures pair-wise associations of prescribed medications.

3.3 Collaborative Filtering

Collaborative filtering has been used extensively in recommendation systems to infer user's preferences from historical data. Analogous to the recommendation system application, where there are users and items, our CF formulation uses ICU admissions (as users) and information about the ICU admission (as items).

There are two main categories of CF techniques: neighborhood approaches and latent factor model approaches.

3.3.1 Neighborhood Approach

To make a recommendation for a user u, a neighborhood approach starts by finding users that are similar to u. Next, it takes weighted combinations of their ratings to make recommendations for u. In our work, we treat ICU admissions as users, and consider two ICU admissions to be similar if they share many medications and diagnoses.

The most commonly used neighborhood based-approach is kNN. Given an ICU admission $\mathbf{a_j}$, we find k other ICU admissions that are most similar with respect to a weighted cosine similarity-based distance metric (Xie et al., 2012).

The standard cosine similarity metric calculates distance between two vectors where each component of the vector is weighted equally. However, equal weighting is not appropriate for our application. For example, two ICU admissions that share a less commonly prescribed drug such as Carbamazepine (used to control seizures) should probably be considered more similar than two admissions that both prescribed Acetaminophen. The weighted cosine similarity between two hospital admissions $\mathbf{a_s}$ and $\mathbf{a_t}$ is defined as:

$$S_{\mathbf{a_s},\mathbf{a_t}} = \frac{\sum_{i=1}^{I} w_i \mathbf{a_s}_i \mathbf{a_t}_i}{\sqrt{\sum_{i=1}^{I} (w_i \mathbf{a_s}_i)^2} \sqrt{\sum_{i=1}^{I} (w_i \mathbf{a_t}_i)^2}} \quad (5)$$

where,

- for $i \leq M$, w_i = 1-$\mathfrak{P}_{\mathfrak{m}_i}$
- for $i > M$, w_i = 1- $\mathfrak{P}_{d_{i-M}}$

The predicted score for medication m_i in admission $\mathbf{a_j}$ is obtained by taking a weighted average of the values of m_i in the k nearest neighbors while adjusting for the prior of the medication. The term $(R_{i,t} - \mathfrak{P}_{m_i})$ can be interpreted as information over the prior.

$$score(i,j) = \mathfrak{P}_{\mathfrak{m}_i} + \frac{\sum_{\mathbf{a_t} \in \mathfrak{N}} S_{\mathbf{a_j},\mathbf{a_t}} (R_{i,t} - \mathfrak{P}_{m_i})}{\sum_{\mathbf{a_t} \in \mathfrak{N}} S_{\mathbf{a_j},\mathbf{a_t}}} \quad (6)$$

where, $\mathfrak{N}$ is the set of k nearest neighbors.

This formulation of nearest neighbors is good at capturing localized relationships provided that the distance between the k nearest neighbors and the given ICU admission is small.

3.3.2 Latent Factor Models

Latent factor models aim to uncover latent features that explain the observed data. These models have been successfully used in Netflix movie recommendations (Li et al., 2010). In our application, we attempt to uncover combinations of medications and diagnosis that appear frequently in the dataset. We refer to these as prescription patterns.

Latent factor models are generated using matrix factorization methods. The main idea of matrix factorization is to approximate matrix R using two lower rank matrices X (an $I \times F$ matrix) and Y (an $F \times A$ matrix) such that

$$R_{ij} \approx (XY)_{ij} = \sum_{f=1}^{F} X_{if}.Y_{fj} \quad (7)$$

The F columns of matrix X are the basis vectors. Each basis vector encodes information about prescription patterns learned from the dataset. Each column of matrix Y contains weights for each basis vector. An ICU admission $\mathbf{a_j}$ is approximated as a linear combination of the basis vectors weighted using the j^{th} column of Y.

Again, to adjust for the prior, instead of approximating the matrix R, we approximate the residual matrix $\mathfrak{R}$.

$$\mathfrak{R}_{i,j} = R_{i,j} - \mathfrak{P} \quad (8)$$

where,

$$\mathfrak{P} = \begin{cases} \mathfrak{P}_{m_i}, & i \leq M. \\ \mathfrak{P}_{d_{i-M}}, & i > M. \end{cases} \quad (9)$$

We learn matrices X and Y by solving the following regularized optimization problem using an alternating least squares algorithm (Zamir and Gabriel, 1979):

$$min_{X,Y} \sum_{i,j} (\mathfrak{R}_{ij} - X_{i\cdot}.Y_{\cdot j})^2 + \lambda(\|X_{i\cdot}\|^2 + \|Y_{\cdot j}\|^2) \quad (10)$$

where $X_{i\cdot}$ is the i^{th} row of matrix X and $Y_{\cdot j}$ is the j^{th} column of matrix Y.

Once we estimate matrices X and Y, the predicted score for medication m_i in admission $\mathbf{a_j}$ is obtained by:

$$score(i,j) = \mathfrak{P}_{\mathfrak{m}_i} + X_{i\cdot}.Y_{\cdot j} \quad (11)$$

While neighborhood approach rely on the nearest neighbors to predict omissions, latent factor models have a more holistic approach. They attempt to uncover prescription patterns in the data. A latent factor model with F factors captures the F most prominent basis vectors, which can be viewed as encoding prescription patterns, from the data.

3.3.3 Integrated Approach

In the integrated approach, we combine the latent factor model with the neighborhood model. This integrated model is aimed at capturing both prescription patterns and localized relationships in the data. Here we model the *score* as having three terms:

$$score(i,j) = \mathfrak{P}_{\mathfrak{m}_i} + X_{i\cdot}.Y_{\cdot j} + \mathfrak{H}(i,j) \quad (12)$$

The first two terms are identical to that in Equation 11. The third term attempts to capture any residual information that the baseline and the latent factor model are unable to account for.

After computing matrices X and Y, we compute $\mathfrak{H}(i,j)$ using the neighborhood approach.

$$\mathfrak{H}(i,j) = \frac{\sum_{\mathbf{a_t} \in \mathfrak{N}} S_{\mathbf{a_j},\mathbf{a_t}} (R_{i,t} - \mathfrak{P}_{m_i} - X_{i\cdot}.Y_{\cdot j})}{\sum_{\mathbf{a_t} \in \mathfrak{N}} S_{\mathbf{a_j},\mathbf{a_t}}} \quad (13)$$

4 THE DATA

We obtained our data from the MIMIC II (Multiparameter Intelligent Monitoring in Intensive Care) database. MIMIC II contains comprehensive clinical data of ICU patients in Beth Israel Deaconess Medical Center in Boston, Massachusetts.

MIMIC II contains 24,097 admissions with 928 unique generic medication names. For our study, we removed medications that were prescribed in $< 0.1\%$ of the admissions. Those medications were removed

Table 1: The most commonly prescribed medications in our data.

Medication	N (%)
Sodium chloride electrolyte	9721 (50.2%)
Acetaminophen	9031 (46.6%)
Heparin	8875 (45.8%)
Pantaprazole	8309 (42.9%)
Aspirin	6511 (33.6%)

because it would be difficult to learn any useful patterns from few examples using a data driven approach. This left 316 unique medication names. We also removed admissions in which only one medication was prescribed. This resulted in 19,394 admissions.

The most commonly prescribed medications in the hospital admissions are listed in Table 1.

For each admission, we include information about the 30 diagnoses that form the Elixhauser comorbidity system. The Elixhauser comorbidity system is commonly used in health research. It has been found to be significantly associated with in-hospital and post-discharge mortality. It uses 30 binary variables that represent 30 diagnoses. The most common diagnoses in our data are listed in Table 2.

Table 2: The most common diagnoses in our data.

Diagnoses	N (%)
Hypertension	5959 (30.8%)
Fluid and electrolyte disorders	4233 (21.9%)
Congestive heart failure	3714 (19.2%)
Cardiac arrythmias	3609 (18.6%)
Diabetes (uncomplicated)	3466 (17.9%))

5 EXPERIMENTS AND RESULTS

This section describes the set of experiments used to compare different methods and presents a discussion of the results.

5.1 Experimental Setup

We divided our data into training and test sets. Our training set was composed of 12,994 randomly selected ICU admissions, while our test set consisted of the remaining 6400 admissions. We used the training set for training purposes and for parameter selection. Parameter selection was carried out using 5-fold cross-validation. During parameter selection, we tested k in the range $[10, 100, 500, 1000, 2000]$ for the neighborhood approach. For the latent factor models, we evaluated F in the range $[100, 150, 200, 300]$ and λ in the range $[20, 30, 50, 70, 100]$. The test set was used to evaluate the different methods described in Section 3.

As discussed earlier, MIMIC II does not contain annotations that indicate if there was a prescription omission. Therefore, we used a simulation model to generate them. Since we were unable to find literature with a good model for actual admission errors, we chose a simple model. For each admission $\mathbf{a_j}$ in the test set, we randomly remove one medication m_i that is prescribed.

Our goal is to identify the medication that was omitted during the simulation. For every test admission, each method generates a ranked list of medications that are not prescribed to the patient during the admission. The rankings are based on the scoring systems presented in Section 3.

5.2 Results and Discussion

Table 3 compares the overall performance of the different methods. The comparison is based on the median ranking of the omitted drug over all of the admissions in the test set. A lower rank for the omitted drug means that it appeared near the top of the list and thus corresponds to better performance.

The co-occurrence approach (*Co-occur*) that captures pairwise association of medications significantly outperformed the baseline method (*Prior*). All three CF approaches, the neighborhood approach (*Prior+kNN*), the latent factor model (*Prior+MF*) and the integrated model (*Prior+MF+kNN*) significantly outperformed both *Prior* and *Co-occur*.

Table 3: Median rank of the omitted medication using different methods on the test set. A lower rank means better performance.

Method	Parameters used	Median Rank
Prior	-	17
Co-occur	-	8
Prior+kNN	k=500	5
Prior+MF	F= 100; λ=20	4
Prior+ MF+kNN	F= 100; λ=20; k=1000	4

Overall, *Prior+MF* yielded a small improvement in performance over *Prior+kNN*. However, if one considers only ICU admissions in which the omitted drug has a low prior, the performance improvement is considerably more. Figure 3 shows the performance of *Prior+MF* relative to *Prior+kNN* as a function of the prior of the omitted medication. A positive difference in rank means that *Prior+MF* outperformed *Prior+kNN*. For admissions with omitted medications with a low prior, *Prior+MF* strongly outperforms *Prior+kNN*. For admissions in which the

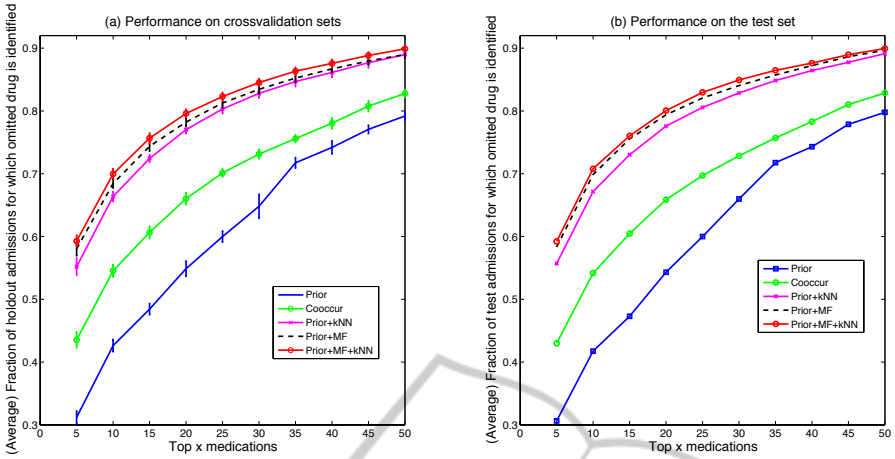

Figure 2: Cumulative performance curves. Each data point shows the fraction of the admissions for which the omitted drug was identified in the top x medications. (a) Cumulative performance curves obtained from 5 fold cross validation. The error bars represent one standard deviation of uncertainty. (b) Cumulative performance curves on the test set.

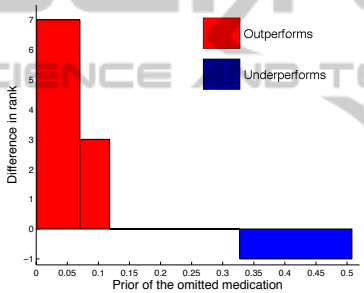

Figure 3: The median difference in rank when *Prior+MF* outperformed and underperformed *Prior+kNN* in the test set. Each bin corresponds to a quartile of the prior of the omitted medications.

omitted medication had a high prior, *Prior+kNN* did slightly better. And for admissions where the omitted medication had a middling prior, the performance of the two methods was almost identical. Keep in mind that our method of simulating drug omissions results in more high prior drugs than low prior drugs being omitted. Thus, when evaluating overall performance, *Prior+MF*'s strong advantage on a few omissions of low prior drugs is largely balanced out by *Prior+kNN*'s small advantage on the omissions of more common drugs.

We believe that the association of performance with prevalence stems from the fact that omitted drugs with high priors are well represented in the k nearest neighbors. Therefore, *Prior+kNN* is able to achieve a better rank for the omitted drug. On the other hand, when the omitted drug has a lower prior, the neighbors are less informative and prescription patterns learned using latent factor models fare better at identifying such omissions.

Figure 2 shows the cumulative performance for each method on both cross validation (CV) and test sets for the top 50 medications. The cumulative performance is measured using the mean fraction of test admissions for which the omitted drug was found in the top x medications. For example, Figure 2(a) shows that for 31.2% of the CV test admissions, *Prior* included the omitted drug within the top 5 medications. $Prior + kNN$ and $Prior + MF$ found the omitted drug (within the top 5 medications) for 55.2% and 58.1% of the CV test admissions respectively. The error bars in Figure 2(a) show one standard deviation of uncertainty.

Consistent with our previous observation, *Prior+MF* outperforms *Prior+kNN*. Notice that the error bars for *Prior+kNN* and *Prior+MF* do not overlap for the top 15 medications. Of course, all of the curves must start to converge as we consider an ever larger number of drugs, since when $x = M$ the curves meet at 1.0.

The results discussed thus far suggest that the latent factor models and the neighborhood approach are better at identifying different kinds of omissions. The proposed integrated model (*Prior+MF+kNN*) combines both methods, and is aimed at capturing multi-scale information - both local relationships and higher level prescription patterns, that exist in the data. Disappointingly, the overall performance of our integrated model was not significantly different from that of the latent factor model in our dataset. However, the cumulative performance curves in Figure 2 show that *Prior+MF+kNN* consistently outperforms both *Prior+MF* and *Prior+kNN* in both CV and test sets. These suggest that a different approach to an integrated model is probably worth

pursuing.

Figure 2(b) shows the cumulative performance on the test set. The trends in the cumulative performance curves are similar. This shows the generalizability of the methods.

6 SUMMARY & FUTURE WORK

In this paper, we investigated a collaborative filtering framework for identifying prescription omission in ICU admissions. We investigated three different types of CF methods: a nearest neighbor approach, a latent factor model and an integrated model. Using prescription medication data and diagnosis information from 19,000 hospital admissions, we tackled the problem of identifying which medications are likely to be omitted. The novelty of our work lies in using CF to investigate multi-scale information, i.e., localized relationships as well as overall prescription patterns, contained in prescription data for the task of identifying prescription omissions.

All of the CF approaches ranked omitted medication higher than a prior based algorithm and a simple co-occurrence approach. Overall, latent factor models outperformed neighborhood models. Further analysis showed that latent factor models are much better at identifying omissions of rarely prescribed medications. Neighborhood models, on the other hand, were able to yield a slightly better ranking for commonly prescribed medications. This highlights the relevance of learning both strong local relationships and overall patterns from the data for the application of identifying omissions.

For our methods, we envision a use model where a computerized systems will suggest a list of medications that are likely to be omitted given the information about the patient. A physician then scans through the top x medications to look for potential omissions.

We conclude with a brief discussion of some future work. First, our current integrated approach optimizes latent factor models independently of the neighborhood approach. Since our results indicate that the neighborhood approach and latent factor model perform better for different kinds of omissions, an improved approach that intelligently uses localized relationships and/or prescription patterns will likely result in better performance. We plan to investigate advanced models that optimizes latent factor parameters, taking into account the information about the neighbors.

Our work explores CF methods using medication and diagnosis data for identifying omissions. Clearly, incorporating other information in electronic medical records, e.g., a patient's allergies should yield improved results. Furthermore, we investigated only data-driven approaches. Hybrid computerized systems that augment data driven methods with expert knowledge are also worth pursuing.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial assistance provided by Quanta Computers Inc.

REFERENCES

Aspden, P., Wolcott, J., Bootman, J. L., and Cronenwett, L. R. (2007). *Preventing Medication Errors: Quality Chasm Series*. The National Academies Press.

Cullen, D., Sweitzer, B., Bates, D., Burdick, E., Edmondson, A., and Leape., L. (1997). Preventable adverse drug events in hospitalized patients: a comparative study of intensive care and general care units. *Critical Care Medicine*, 25(8):1289–1297.

Hasan, S., Duncan, G. T., Neill, D. B., and Padman, R. (2011). Automatic detection of omissions in medication lists. *Journal of the American Medical Informatics Association : JAMIA*, 18(4):449–58.

Li, Y., Hu, J., Zhai, C., and Chen, Y. (2010). Improving one-class collaborative filtering by incorporating rich user information. In *Proceedings of the 19th ACM international conference on Information and knowledge management*, CIKM '10, pages 959–968, New York, NY, USA. ACM.

Olsen, S., Neale, G., Schwab, K., Psaila, B., Patel, T., Chapman, E. J., and Vincent, C. (2007). Hospital staff should use more than one method to detect adverse events and potential adverse events: incident reporting, pharmacist surveillance and local real-time record review may all have a place. *Quality & safety in health care*, 16(1):40–4.

Rothschild, J., Landrigan, C., Cronin, J., Kaushal, R., Lockley, S., Burdick, E., Stone, P., Lilly, C., Katz, J., Czeisler, C., and Bates, D. (2005). The critical care safety study: The incidence and nature of adverse events and serious medical errors in intensive care. *Critical Care Medicine*, 33(8):1694–1700.

Saeed, M., Lieu, C., Raber, G., and Mark, R. (2002). Mimic ii: A massive temporal icu patient database to support research in intelligent patient monitoring. In *Computers in Cardiology, 2002*, pages 641 – 644.

Xie, J., Leishman, S., Tian, L., Lisuk, D., Koo, S., and Blume, M. (2012). Feature Engineering in User's Music Preference Prediction. *Journal of Machine Learning Research: Workshop and Conference Proceedings*, 18:183–197.

Zamir, S. and Gabriel, K. R. (1979). Lower Rank Approximation of Matrices by Least Squares With Any Choice of Weights. *Technometrics*, 21(4).