Neighborhood-based Collaborative Filtering for Therapy Decision Support

Felix Gräßer, Hagen Malberg and Sebastian Zaunseder
Institut für Biomedizinische Technik
Technische Universität Dresden
Dresden, Germany
felix.graesser@tu-dresden.de

Stefanie Beckert, Denise Küster and Jochen Schmitt
Zentrum für Evidenzbasierte Gesundheitsversorgung
Universitätsklinikum Dresden
Dresden, Germany
jochen.schmitt@uniklinikum-dresden.de

Susanne Abraham
Klinik und Poliklinik für Dermatologie Universitätsklinikum Dresden
Dresden, Germany
susanne.abraham@uniklinikum-dresden.de

ABSTRACT

Clinical decision support systems (CDSS) providing assistance for diagnosis and treatment decisions are expected to play an increasingly important role in future healthcare. Especially data-driven approaches employing data mining and machine learning techniques to exploit the large volume of daily captured clinical data promise to open up new perspectives. Particularly in e-commerce Recommender Systems (RSs) have evolved considerably over the last years, yielding extremely sophisticated and specialized methods. In healthcare, however, such algorithms have not found wide application although offering wide opportunities. Within this work the idea of RSs, namely neighborhood-based Collaborative Filtering (CF), is transferred to the domain of CDSS aiming at helping to find an optimal personalized therapy for a given patient and time, i.e. consultation under consideration. Particular focus of this work is to adapt neighborhood-based CF methods to exploit high-dimensional clinical data. To leverage trust and reduce risk of the proposed system, an exclusively data-driven approach is extended by a set of evidence-based contraindication rules excluding inappropriate therapies from the recommendation list. The proposed therapy recommendation system is practically evaluated on an exemplary clinical dataset. Its underlying conceptual framework, however, is intended to be transferable to other diseases and medical disciplines.

CCS CONCEPTS

• **Information systems → Expert systems**; • **Applied computing → Health care information systems**; **Health informatics**; • **Human-centered computing →** *Collaborative and social computing theory, concepts and paradigms*;

KEYWORDS

Clinical Decision Support System (CDSS), Health Recommender System, Therapy Recommender System, Collaborative Filtering, Recommender System Evaluation

Second International Workshop on Health Recommender Systems co-located with ACM RecSys 2017, August 2017, Como, Italy

ACM Reference Format:
Felix Gräßer, Stefanie Beckert, Denise Küster, Susanne Abraham, Hagen Malberg, Jochen Schmitt and Sebastian Zaunseder. 2017. Neighborhood-based Collaborative Filtering for Therapy Decision Support. In *Proceedings of the Second International Workshop on Health Recommender Systems co-located with ACM RecSys 2017, Como, Italy, August 2017 (RecSys'17),* 5 pages.

1 INTRODUCTION

Clinical decision support systems (CDSS) aim to assist health professionals with the clinical decision-making tasks. Providing diagnosis or treatment recommendations can foster personalization and contribute to improve quality and efficiency of patient care. CDSS approaches are typically distinguished between data-driven and knowledge-based approaches.

Knowledge-based systems rely on rule-based expert knowledge only. Implementing and updating of such manually encoded evidence-based guidelines, however, puts a challenging bottleneck on providers of CDSS. Furthermore, individual patients' characteristics commonly differ from the strict inclusion criteria on which evidence is based on, i.e. a clinical study was conducted on, which may result in differing therapy responses. Thus, a patient-specific treatment, i.e. an individually optimal therapy option cannot be provided on the basis of guidelines derived from clinical studies only.

Data-driven CDSS, on the other hand, are supposed to automatically extract information from clinical data and facilitate automatic adaptability to evolving databases. In this way, data-driven CDSS are expected to be capable of exploiting the collective clinical experience represented by large-scale databases to improve quality and increase personalization level of automatically generated therapy recommendations. Thus, data-driven CDSS are a promising alternative and are expected to open up new perspectives in medicine. However, clinical data is often characterized by uncertainties and incompleteness (sparsity), high dimensionality and complex interdependencies [23][1], which places high demands on applicable methods. Traditional methods from machine learning, i.e. artificial neural network classifier (ANN) or complex classifier ensembles have proven to be very effective in learning patterns from large-scale databases. However, the stated properties of medical data make the application of such algorithms challenging. Moreover, an essential requirement for acceptance of data-driven CDSS among health professionals are interpretabilty and comprehensibility of the produced results. The blackbox-behaviour of typical machine learning methods restricts insight into the classification process. This can be assumed to be one key factor hindering a wider spread of data-mining and machine learning applications in the context of CDSS to date.

Particularly in e-commerce applications, Recommender Systems (RSs), namely neighbourhood-based approaches as Collaborative Filtering (CF) algorithms, have gained increasing popularity within the preceding years. In this context, RSs support customers to individually identify most interesting products from a wide range of possible options by predicting a user's preference for products very effectively[20]. Incorporating only a modest number of nearest neighbors into the computation can provide transparency on the process of recommendation generation. Additionally, by applying a suitable similarity metric, such algorithms can cope with heterogeneity and sparsity. Therefore, to address both, challenges related to the data characteristics and comprehensibility issue, we proposed transferring the idea of CF algorithms into the domain of CDSS [9]. The overall objective of the proposed therapy recommender system is to find an optimal personalized therapy for a consultation under consideration by converting estimations of a patient's therapy response into recommendations.

Within this contribution our previous approach is extended by a set of rules derived from evidence-based absolute contraindication criteria in order to increase trust into and reduce risk of the demonstrated overall system. The exemplary therapy decision support system is developed and evaluated targeting therapy recommendations for patients suffering from the autoimmune skin disease psoriasis. Within this clinical application the developed therapy decision support system aims at recommending the potentially most effective systemic therapy for a given patient and consultation. The recommender system's underlying conceptual framework, however, is intended to allow transferring the developed ideas to other diseases and medical disciplines.

The paper at hand is organized as follows. After presenting works related to CDSS for therapy decision support in general and systems making use of CF techniques in particular, the available data, the applied evaluation procedure and the RS algorithm are described, respectively. Finally, we present the results of the proposed method and summarize our findings leading to future works in the field.

2 RELATED WORK

Research on expert systems in clinical context date back to the 70ies. Various approaches were published, deriving therapy decision support from computerized medical guidelines [4, 12]. However, as stated beforehand, knowledge-based approaches suffer from considerable efforts during development and updating of the underlying set of rules and are not always generally applicable. Proposed data-driven approaches on the other hand, typically apply machine learning algorithms to derive therapy recommendations [17] or range from majority voting [15], systems based on association rules [2] to approaches applying case-based reasoning [14]. In spite of gaining increasing popularity in other domains, the use of CF techniques is very limited in the context of CDSS. CF in the medical context was proposed in scientific works [1, 22] but studies on clinical data are rare. There are few works applying CF algorithms for disease risk or mortality prediction [3, 11, 13]. Work loosely related to the idea of using CF for therapy decision support are a nursing care plan recommender [5] and an approach recommending wellness treatment [14].

3 METHODOLOGY AND EVALUATION

Within this work, we propose and evaluate an exemplary therapy decision support system targeting therapy recommendations for patients suffering from the autoimmune skin disease psoriasis. The developed therapy decision support system aims at recommending the potentially most effective systemic therapy out of $M = 21$ therapy options for a given patient and consultation.

3.1 Data Characteristics

The exemplary data at hand comprises excerpts from health records that were collected in the Clinic and Polyclinic for Dermatology, University Hospital Dresden. The collected database comprises 1111 consultations from 213 patients suffering from various types of psoriasis. For each sample, i.e. each consultation in the collected database, patient related attributes containing demographic data, comorbidities and state of health as well as information on current and previous treatments are contained. All relevant therapies applied to a patient up to the consultation under consideration are summarized under previous treatments, whereas therapies which were applied within the last two weeks preceding the respective consultation are collected under current treatments.

For both, previous and current therapies, up to three different outcome indicators are given, namely a therapy effectiveness indicator (good, medium, bad) representing the subjective assessment, an objective health state improvement indicator (Psoriasis Area and Severity Index [6]) and occurrence of adverse effects (yes, no). Overall therapy effectiveness is modeled using a weighted sum of those three parameters as introduced in [8, 9] ranging from 0 (bad response) to 1 (good response). Thus, ground truth, i.e. actually applied therapy along with outcome for a given consultation, is derived from the succeeding consultations therapy response.

Table 1 and 2 summarize patient attributes and therapy information, respectively. All attributes are supplied with scale of measurement, range of values and availability relative to all consultations. In case of comorbidities and therapies the availability is related to all applied comorbidities or therapies, respectively.

3.2 Evaluation Procedure

The quality of RSs is typically evaluated concerning accuracy metrics for preference prediction performance as Root Mean Square Error (*RMSE*) and decision support metrics for ranked lists of items derived from information retrieval research as *precision* and *recall* [10]. Generally, quality is evaluated offline and retrospectively based on a test dataset comprising ratings on previously consumed items. In the context of therapy RSs this implies that evaluation metrics are computed on the actually applied therapy associated to a consultation for which outcome is known. However, the focus of a clinical recommender system should not only be to meet the therapy decision of the attending physician but finding therapies with possibly good outcome and rejecting bad ones. Therefore, an additional output-driven precision metric $precision_o@N$ is used in this study as introduced in [16] and previously demonstrated in

[9]. $Precision_o@N$ is computed for each evaluated consultation on the top-N recommendations matching the actually applied therapy only. That means, precision is defined as the ratio of recommendations having good outcome (true-positive TP) and all cases matching the actually applied therapy and having good or bad outcome (true-positive TP and false-positive FP). Employing an effectiveness indicator threshold, data was divided into instances showing good outcome (effectiveness > 0.5) and the remaining ones.

Consequently, $precision_o@N$ can be improved by increasing the number of good outcome recommendations and rejecting bad outcome recommendations from the top-N recommendation list.
To make most out of the already limited amount of data, the proposed algorithm is evaluated using a leave one out cross validation (LOOCV) on the entire dataset. Multiple consultations from the patient for which outcome prediction and recommendations are evaluated are excluded from the training dataset during evaluation.

3.3 Demographic-enhanced Collaborative Filtering Recommender

The CF algorithm applied in this contribution and initially presented in [8, 9] uses both, information on therapy history, i.e. previously applied therapies and associated therapy response, along with all information on a patient's type of disease, comorbidities and demographic data to represent consultations in the database. Furthermore, the attribute vector was extended to incorporate additional information on disease progression and associated therapies. To that end, attributes from the temporal sequence of state of health and applied therapies, i.e. lag features from preceding consultations, are added. The underlying assumption is that therapy history together with the stated patient and disease progress describing data carries sufficient information to reliably compare consultations. Additionally, by not relying on therapy history solely, the cold-start limitation can be overcome in cases where only limited or no information on therapy history is available. The overall objective is to make prediction on patients having similar therapy history and characteristics. Therefore, similarity is computed between a vector representation of the consultation under consideration and representations of all other consultations in the database. Attributes representing consultations are (i) of various level of measurement and differ in range and are (ii) only intermittently available. A similarity metric capable of coping with both, missing values and varying levels of measurement, is the Gower Similarity Coefficient [7], which is applied in this work. Here, the level of measurement of the individual attributes is respected for each attribute comparison using a data type-specific similarity coefficients. Therapy outcome predictions are estimated based on a weighted sum of the k nearest consultations to a consultation under consideration. In a subsequent recommendation step a list of therapy options, ordered by outcome prediction, is available for further processing.

3.4 Exclusion rules

Particularly in the context of CDSS, trust plays a crucial role to leverage acceptance and applicability of such systems. However, in contrast to other domains of RS applications, e.g. e-commerce applications, particular in the area of health and medicine failures in recommendations can accompany high risk. Therefore, to increase confidence in automatically generated therapy recommendations the risk of inaccurate or even health endangering recommendations must be minimized.
For this purpose we implemented a set of exclusion rules based on European S3-Guidelines on the systemic treatment of psoriasis [18] as summarized in table 3. Therapy options which are included in the recommendations provided by the CF algorithms and are affected by an exclusion criterion are removed from the list. Additionally, therapies showing good outcome in a patient's previous consultation were moved to the top of the list. Finally, from the modified recommendation list the top-3 entries are presented to the user.

Table 1: Patient describing attributes

Attribute	Scale	Range	Availability %
Patient Data			
Year of Birth	interval	1931 - 1998	100
Gender	nominal	1,2	100
Weight	interval	50 - 165	50
Size	interval	99 - 204	36
Planned Child	nominal	1,2,3	100
Year of First Diagnosis	interval	1950 - 2014	90
Family Anamnesis	ordinal	1,2,3	50
Type of Psoriasis	nominal	1,2,3,4,5,6	100
Comorbidities			
Comorbidity	nominal	1,2,3,...,34	-
Status	ordinal	1,2,3	100
Under Treatment	dichotomous	0,1	45
State of Health			
PASI Score	interval	0 - 43	69

Table 2: Therapy describing attributes

Attribute	Scale	Range	Availability %
Systemic Therapy	nominal	1,2,3,...,15	-
Effectiveness	ordinal	1,2,3	98
ΔPASI	interval	-37 - 25	42
Adverse Effect	dichotomous	0,1	100

4 EVALUATION RESULTS

Both $precision@3$ and outcome-driven $precision_o@3$ of the CF-based algorithm are highly dependent on the neighborhood size k incorporated into the outcome prediction computation (see figure 2). Outcome prediction accuracy and recommendation precision show extrema in the neighborhood of around $k = 10 \dots 20$. Integrating too many neighbors, i.e. increasing k, provokes a performance decline due to noise influencing prediction accuracy and thus recommendation precision. In contrast, outcome-driven $precision_o@3$

Table 3: Contraindications and exclusion rules.

Contraindication	Excluded Therapy
arterial hypertonia	Ciclosporin
renal disease	Ciclosporin
hepatic disease	Methotrexat
cancerous disease	any Biologics
planned child	Methotrexat, Acitretin
psoriasis arthritis	Fumaderm, any UV therapies
	Acitrecin, Ciclosporin

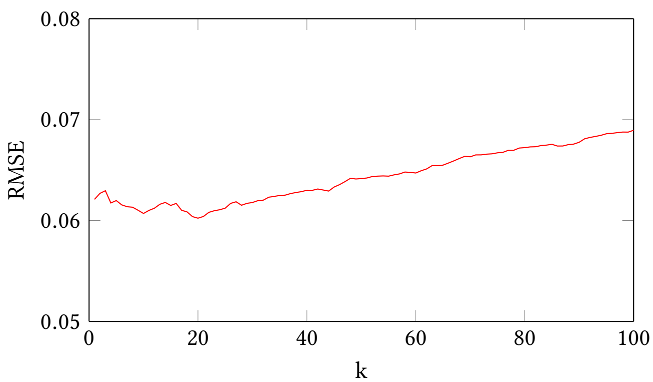

Figure 1: RMSE (——) computed between effectiveness estimated by the CF recommender and effectiveness of actually applied therapies.

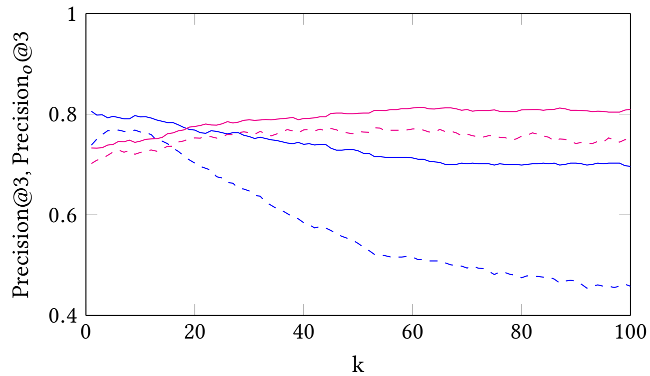

Figure 2: Overall ***precision*****@3 of therapy recommendations and outcome-driven** ***precision$_o$*****@3.** ***precision*****@3 and** ***precision$_o$*****@3 for the CF output (- - -, - - -) and with additional exclusion rules applied (——,——) are shown.**

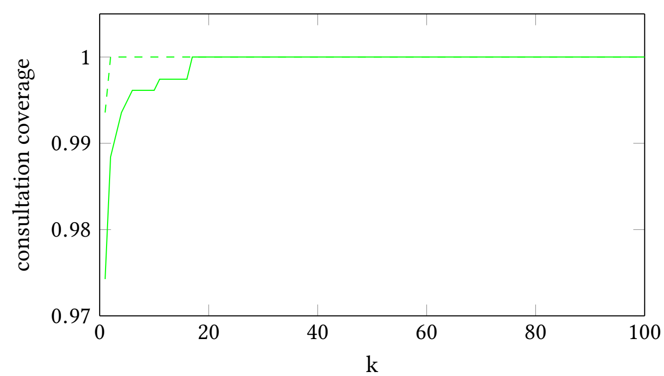

Figure 3: Consultation coverage, i.e. ratio of overall consultations for which recommendations could be provided. Consultation coverage for the CF output (- - -) and with additional exclusion rules applied (——) are shown.

seems to benefit from a somewhat larger neighborhood. However, when comparing $precision_o$@3 for different k it must be kept in mind that $precision_o$@N highly depends $precision$@N.

Adding additional rules clearly increases recommendation precision and outcome-driven precision. Continuing successful therapies has favorable effect on recommendation precision. Furthermore, therapy options which were successfully applied at neighboring patients, i.e. located in the top-3 list, can be contraindicated therapies for a patient and consultation under consideration. Eliminating them from the recommendation list moves non-contraindicated actually applied therapies to the top instead.

For small neighborhoods, the coverage of possible therapy options can be too low to facilitate recommendations for a consultation under consideration leading to low overall consultation coverage. Removing recommendations from the recommendation list affects the coverage additionally as shown in 3. Consequently, adding additional exclusion rules demands for an increased neighborhood k to facilitate satisfactory consultation coverage.

5 CONCLUSION AND FUTURE WORKS

Our analyses show that the proposed prototype combining a data-driven CF approach with evidence-based knowledge can provide reliable personalized therapy recommendations. However, the selection and extraction of appropriate attributes and applying a suitable similarity metric heavily affects the obtained results. Therefore, future efforts will concentrate on those aspects, namely feature selection methods [19] and metric learning algorithms [21, 24], to further improve the proposed CF performance and contribute to create a basis for applicability and acceptance of suchlike CDSS. Besides the used features, it is shown that the neighborhood size k plays a vital role in terms of outcome prediction accuracy and recommendation quality. For the exemplary application an appropriate neighborhood size was determined by cross validation. However, this size can neither be expected to provide the best results in case of an extended dataset nor it can be readily transferred to other problems. Aspects related to the neighborhood size will be further investigated in future studies, particularly incorporating more data. In fact, one major limitation of this work is the rather small database our studies are based on. Therefore, future work will address applying the proposed methods to more comprehensive datasets which we assume will improve the recommendation quality significantly. Particularly, data provided by additional dermatologists needs to be incorporated into the database to prevent learning a limited number of experts recommendations and improve generalization capability. As a consequence, another aspect which needs to be investigated in future works to both, improve recommendation quality and cope with scalability issues when applied to large-scale data, is identifying clusters in the database. This is intended to be done offline prior to actual recommendation generation.

ACKNOWLEDGMENTS

This work is part of the project *Therapieempfehlungssystem* which is funded by the *Roland Ernst Stiftung für Gesundheitswesen*.

REFERENCES

[1] André Calero Valdez, Martina Ziefle, Katrien Verbert, Alexander Felfernig, and Andreas Holzinger. 2016. *Recommender Systems for Health Informatics: State-of-the-Art and Future Perspectives*. Springer International Publishing, Cham, 391–414.

[2] Jonathan H. Chen and Russ B. Altman. 2014. Automated Physician Order Recommendations and Outcome Predictions by Data-Mining Electronic Medical Records. *AMIA Summits on Translational Science proceedings* (2014), 206–210.

[3] Darcy Davis, Nitesh V. Chawla, Nicholas Christakis, and Albert László Barabási. 2010. Time to CARE: A collaborative engine for practical disease prediction. *Data Mining and Knowledge Discovery* 20, 3 (2010), 388–415.

[4] Paul A. De Clercq, Johannes A. Blom, Hendrikus H. M. Korsten, and Arie Hasman. 2004. Approaches for creating computer-interpretable guidelines that facilitate decision support. *Artificial Intelligence in Medicine* 31, 1 (2004), 1–27.

[5] Lian Duan, W. Nick. Street, and E. Xu. 2011. Healthcare information systems: data mining methods in the creation of a clinical recommender system. *Enterprise Information Systems* 5, 2 (2011), 169–181.

[6] T. Fredriksson and U. Pettersson. 1978. Severe Psoriasis - Oral Therapy with a New Retinoid. *Dermatologica* 157, 4 (1978), 238–244.

[7] J. C. Gower. 1971. A General Coefficient of Similarity and Some of Its Properties. *Biometrics* 27, 4 (1971), 857–871.

[8] Felix Gräßer, Stefanie Beckert, Denise Küster, Jochen Schmitt, Susanne Abraham, Hagen Malberg, and Sebastian Zaunseder. 2016. Application of recommender system methods for therapy decision support. *2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)* (2016), 1–6.

[9] Felix Gräßer, Stefanie Beckert, Denise Küster, Jochen Schmitt, Susanne Abraham, Hagen Malberg, and Sebastian Zaunseder. 2017. Therapy Decision Support Based on Recommender System Methods. *Journal of Healthcare Engineering* 2017 (2017).

[10] Asela Gunawardana and Guy Shani. 2009. A Survey of Accuracy Evaluation Metrics of Recommendation Tasks. *J. Mach. Learn. Res.* 10 (Dec. 2009), 2935–2962.

[11] Shahzaib Hassan and Zeeshan Syed. 2010. From Netflix to Heart Attacks: Collaborative Filtering in Medical Datasets. *Proceedings of the 1st ACM International Health Informatics Symposium* (2010), 128–134.

[12] Ellen Kilsdonk, Linda W. Peute, Rinke J. Riezebos, Leontien C. Kremer, and Monique W M Jaspers. 2013. From an expert-driven paper guideline to a user-centred decision support system: A usability comparison study. *Artificial Intelligence in Medicine* 59, 1 (2013), 5–13.

[13] Maytiyanin Komkhao, Jie Lu, and Lichen Zhang. 2012. Determining Pattern Similarity in a Medical Recommender System. *Data and Knowledge Engineering* 7696 (2012), 103–114.

[14] Thean Pheng Lim and Wahidah Husain. 2013. Recommender System for Personalised Wellness Therapy. *International Journal of Advanced Computer Science and Applications* 4, 9 (2013), 54–60.

[15] Haifeng Liu, Guotong Xie, Jing Mei, Weijia Shen, Wen Sun, and Xiang Li. 2013. An Efficacy Driven Approach for Medication Recommendation in Type 2 Diabetes Treatment Using Data Mining Techniques. *Medinfo 2013: Proceedings of the 14th World Congress on Medical and Health Informatics, Pts 1 and 2* 192 (2013), 1071.

[16] Jing Mei, Haifeng Liu, Xiang Li, Yiqin Yu, and Guotong Xie. 2015. Outcome-driven Evaluation Metrics for Treatment Recommendation Systems. *Studies in health technology and informatics* 210 (2015), 190–194.

[17] Geert Meyfroidt, Fabian GÃČÂijiza, Jan Ramon, and Maurice Bruynooghe. 2009. Machine learning techniques to examine large patient databases. *Best Practice and Research Clinical Anaesthesiology* 23, 1 (2009), 127 – 143.

[18] A. Nast, P. Gisondi, A.D. Ormerod, P. Saiag, C. Smith, P.I. Spuls, P. Arenberger, H. Bachelez, J. Barker, E. Dauden, E.M. de Jong, E. Feist, A. Jacobs, R. Jobling, L. KemÃĺny, M. Maccarone, U. Mrowietz, K.A. Papp, C. Paul, K. Reich, S. Rosumeck, T. Talme, H.B. Thio, P. van de Kerkhof, R.N. Werner, and N. Yawalkar. 2015. European S3-Guidelines on the systemic treatment of psoriasis vulgaris âĂŞ Update 2015 âĂŞ Short version âĂŞ EDF in cooperation with EADV and IPC. *Journal of the European Academy of Dermatology and Venereology* 29, 12 (2015), 2277–2294.

[19] P. Pudil, F. Ferri, J. Novovicova, and J. Kittler. 1994. Floating search method for feature selection with non monotonic criterion functions. *Pattern Recognition* 2 (1994), 279–283.

[20] Xiaoyuan Su and Taghi M. Khoshgoftaar. 2009. A Survey of Collaborative Filtering Techniques. *Advances in Artificial Intelligence* 3, Section 3 (2009), 1–19. arXiv:421425

[21] Kilian Q. Weinberger and Lawrence K. Saul. 2009. Distance metric learning for large margin nearest neighbor classification. *The Journal of Machine Learning Research* 10 (2009), 207–244.

[22] Martin Wiesner and Daniel Pfeifer. 2014. Health recommender systems: Concepts, requirements, technical basics and challenges. *International Journal of Environmental Research and Public Health* 11, 3 (2014), 2580–2607.

[23] David Windridge and Miroslaw Bober. 2014. A Kernel-Based Framework for Medical Big-Data Analytics. *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics: State-of-the-Art and Future Challenges* (2014), 197–208.

[24] Eric P. Xing, Andrew Y. Ng, Michael I. Jordan, and Stuart J. Russell. 2002. Distance Metric Learning with Application to Clustering with Side-Information. *Advances in Neural Information Processing Systems (NIPS)* 15 (2002), 505–512.