

Using Machine Learning Algorithms to Predict Patient Portal Use Among Emergency Department Patients With Diabetes Mellitus

Yuan Zhou^a, Thomas K. Swoboda^b, Zehao Ye^a, Michael Barbaro^c, Jake Blalock^d,
Danny Zheng^c, Hao Wang^{c, e} 

Abstract

Background: Different machine learning (ML) technologies have been applied in healthcare systems with diverse applications. We aimed to determine the model feasibility and accuracy of predicting patient portal use among diabetic patients by using six different ML algorithms. In addition, we also compared model performance accuracy with the use of only essential variables.

Methods: This was a single-center retrospective observational study. From March 1, 2019 to February 28, 2020, we included all diabetic patients from the study emergency department (ED). The primary outcome was the status of patient portal use. A total of 18 variables consisting of patient sociodemographic characteristics, ED and clinic information, and patient medical conditions were included to predict patient portal use. Six ML algorithms (logistic regression, random forest (RF), deep forest, decision tree, multilayer perception, and support vector machine) were used for such predictions. During the initial step, ML predictions were performed with all variables. Then, the essential variables were chosen via feature selection. Patient portal use predictions were repeated with only essential variables. The performance accuracies (overall accuracy, sensitivity, specificity, and area under receiver operating characteristic curve (AUC)) of patient portal predictions were compared.

Results: A total of 77,977 unique patients were placed in our final analysis. Among them, 23.4% (18,223) patients were diabetic mellitus (DM). Patient portal use was found in 26.9% of DM patients.

Overall, the accuracy of predicting patient portal use was above 80% among five out of six ML algorithms. The RF outperformed the others when all variables were used for patient portal predictions (accuracy 0.9876, sensitivity 0.9454, specificity 0.9969, and AUC 0.9712). When only eight essential variables were chosen, RF still outperformed the others (accuracy 0.9876, sensitivity 0.9374, specificity 0.9932, and AUC 0.9769).

Conclusion: It is possible to predict patient portal use outcomes when different ML algorithms are used with fair performance accuracy. However, with similar prediction accuracies, the use of feature selection techniques can improve the interpretability of the model by addressing the most relevant features.

Keywords: Machine learning; Diabetic; Patient portal; Feature selection; Prediction performance

Introduction

In recent years, machine learning (ML) technologies have been broadly applied in the field of medicine [1, 2]. With the adoption of these ML technologies, healthcare providers can make more informed decisions based on predicted clinical outcomes such as long-term recovery from stroke, the diagnostic accuracy of various diseases, and the efficiency of treatment [3-5]. Several well-known ML algorithms include random forests (RFs), decision trees (DTs), neural networks, linear or logistic regression (LR), and support vector machines (SVMs). Debnath et al summarized using a variety of ML technologies to augment coronavirus disease 2019 (COVID-19) clinical decisions [6]. Such predictions emerging into healthcare system were also validated with the use of different ML algorithms in different countries [7, 8]. Using the LR and RF algorithms, Safaripour and June Lim reported over 90% of accuracy to predict ED frequent users [9]. Tello et al predicted the hospital bed demand with SVM, one of the ML technologies, which could potentially enhance the efficiency of inpatient bed utilization [10]. In terms of the specific diseases, a previous study reported using ML to help physicians diagnose certain diseases [11]. Many studies focused on patients with diabetic mellitus (DM) and reported diabetic prediction [12], classification [13],

Manuscript submitted January 3, 2023, accepted March 6, 2023
Published online March 28, 2023

^aDepartment of Industrial, Manufacturing, and Systems Engineering, The University of Texas at Arlington, Arlington, TX 76109, USA

^bDepartment of Emergency Medicine, The Valley Health System, Touro University Nevada School of Osteopathic Medicine, Las Vegas, NV 89144, USA

^cDepartment of Emergency Medicine, JPS Health Network, Fort Worth, TX 76104, USA

^dTCU and UNTHSC School of Medicine, Fort Worth, TX 76107, USA

^eCorresponding Author: Hao Wang, Department of Emergency Medicine, John Peter Smith Health Network, Fort Worth, TX 76104, USA.

Email: hwang@ies.healthcare

doi: <https://doi.org/10.14740/jocmr4862>

and management benefits from using ML technologies [14]. However, different studies have used different ML algorithms with diverse findings.

With the increasing adoption of ML technologies, the questions about interpretability and liability of their outputs are increasing [15]. Most of the current ML models remain complex black boxes, i.e., their internal logic and inner workings are hidden to the users [16]. As a consequence, the rationale behind the prediction cannot be fully understood, thus making the interpretability of these models a major challenge in current research and practice towards their meaningful use in clinical decision support [17]. Emphasizing the importance of features (i.e., predictors) will significantly improve the ML model interpretability [18]. Some ML models, such as regression and tree ensembles, are capable to assign importance scores to features while fitting the data. Thus, we could rely on feature importance scores for explainability to offer insights into the model behavior [19]. Further, to assess the performance of ML models, the overfitting issue must be taken into account. Overfitting is an undesirable ML behavior that occurs when the ML model gives accurate predictions for training data but not for new data. One of the reasons that can potentially cause overfitting is that the training dataset contains a large amount of irrelevant information (e.g., too many features). To avoid overfitting, feature selection is an accepted pre-processing method which can help mitigate overfitting while optimizing the prediction performance. Studies that consider both ML models' interpretability and overfitting, particularly in the clinical predictions, are sparse.

In this study, we used variables harvested among patients presented at an emergency department (ED) to predict patient portal use. A patient portal is one type of patient personal health record that is tethered to hospital electronic medical records [20]. With the implementation of Health Information Technology for Economic and Clinical Health (HITECH) act, electronic medical records and patient portals have been used broadly in the United States (US) healthcare system [20]. A patient portal has been considered as an effective tool to connect patients and their healthcare providers. However, disparities have occurred among individuals using patient portals, especially among individuals with DM [21-23]. Non-Hispanic White patients tended to utilize patient portals more often than non-Hispanic Black or Hispanic patients [22]. Patients with low socioeconomic status (SES) were also less likely to use patient portals [21]. Therefore, it is important to recognize factors that could promote patient portal use among DM patients. Vulnerable patients tend to be less engaged in healthcare, and many have been uninsured, or underinsured, and utilize the ED as their primary medical home [24]. Accurately predicting patient portal use will help ED healthcare providers identify vulnerable patient populations to further advocate patient portal use. Meanwhile, identifying important factors that promote patient portal use can help implement more efficient interventions. Use of different ML algorithms could possibly yield both findings.

Therefore, our primary goal was to determine the feasibility and accuracy of using different ML algorithms for patient portal use prediction. In addition, our secondary goal was to use a feature selection technique to identify the top "essential"

variables, and further identify the accuracy of patient portal use ML predictions with the use of only "essential" variables among ED DM patients.

Materials and Methods

Study design and setting

This was a single-center retrospective observational study. The study hospital is a level-1 trauma center, regional chest pain, and comprehensive stroke center. The study hospital ED has approximately 125,000 annual patient visits. The entire hospital system uses electronic medical records (Epic) and a patient portal system (MyChart) that is assigned to each patient [25]. Patient portals can be entered by either using a computer via a website or a tablet/smartphone via an app. This study was approved by the regional institutional review board with waived informed consent. This study was conducted in compliance with the ethical standards of the responsible institution on human subjects as well as with the Helsinki Declaration.

Study participants

This study included only adult patients with a history or new onset of DM (including type 1 and type 2) who were presented at the study ED from March 1, 2019 to February 28, 2020. We chose DM patients in this study because different ML technologies have been applied broadly among such a cohort with diverse findings [26, 27]. If patients had multiple ED encounters during the study period, only the last ED visit was included. In addition, we only included DM patients that had patient portal access (i.e., DM patients had at least one ED or clinic visit within the healthcare system prior to the index ED visit). Subjects excluded had one of the following criteria: 1) age < 18; 2) no prior patient healthcare visits at the study healthcare system; 3) prisoners; and 4) no DM patients.

Variables

Response variable

This study conducted classification analyses on a single outcome variable to predict the use of patient portal (i.e., MyChart) by DM patients. This binary outcome variable indicated whether the patients used MyChart in the past 12 months upon their last ED visit during the data collection period (1 = used; 0 = not used).

Predictor variables

After a screening of the potential predictor variables, 18 candidates were selected and used in the analyses. We categorized these predictors into three groups: 1) variables of patient de-

Table 1. Performance Measures of the Six Classifiers Before Feature Selection (All 18 Variables Used in the Model)

Classifier	Accuracy	Sensitivity	Specificity	AUC
Logistic regression	0.8242	0.7542	0.9727	0.8034
Random forest	0.9876	0.9454	0.9969	0.9712
Deep forest	0.9784	0.8966	0.9965	0.9465
Decision tree	0.9820	0.9508	0.9890	0.9699
Multilayer perception	0.8222	0.7814	0.9862	0.9693
Support vector machine	0.6166	0.5265	0.6992	0.5851

AUC: area under the receiver operating characteristic curve.

mographics and social determinants; 2) variables related to patient visit history in the past 12 months (including both clinic and ED visits); and 3) variables related to patient medical conditions. Supplementary Material 1 (www.jocmr.org) shows all the predictor variables and their corresponding data type and values.

Patient portal use prediction using different ML algorithms

Six ML algorithms (i.e., classifiers) were used to predict the outcome variable: 1) LR; 2) RF; 3) deep forest (DF); 4) DT; 5) multilayer perception (MLP); and 6) SVM. Supplementary Material 2 (www.jocmr.org) shows more details on these classification methods.

Training and testing samples for ML

To reduce the selection bias of samples used for training and testing, we employed the 10-fold cross-validation method to train and test each classifier. Cross-validation is a resampling method that uses different portions of the data to train and test a model on different iterations. Specifically, we partitioned the entire dataset into 10 subsets of equal sizes. In each iteration, one subset was used for testing while the remaining nine subsets were used for training. A total of 10 iterations were conducted, where the 10 subsets were rotated such that each subset was used once in an iteration for testing. The overall performance of a classifier was then derived by computing the averaged performance across the 10 iterations.

Analysis

All 18 variables were used to predict patient portal use by the six different ML algorithms. The classifiers' performances were assessed based on four measures: 1) accuracy; 2) sensitivity; 3) specificity; and 4) area under the (receiver operating characteristics (ROC) curve (AUC). Accuracy measures the proportion of the correctly predicted samples (including both positive and negative ones) in all the tested samples. Sensitivity (specificity) measures the proportion of the correctly predicted positive (negative) samples in all the truly positive (neg-

ative) samples. Further, AUC measures the area underneath the ROC curve (range = 0.5 - 1), where the ROC curve is a graph showing the performance of a classification model based on both sensitivity and specificity measures at all classification thresholds. If the classifier is based on random guesses, the AUC is 0.5. If the classifier is perfect, then AUC is 1.

Further, to improve the prediction performance and reduce the risk of overfitting, feature selection was performed through a recursive feature elimination procedure [28]. Recursive feature elimination selection enables the search of a reliable subset of features while looking at performance improvements and maintaining the computation costs acceptable [28]. In our study, the feature selection procedure was specified as follows. Initially, we started with a set of all the 18 predictor variable candidates to predict patient portal use. Based on the feature importance scores, we eliminated the least important variable and kept the 17 top-ranked predictors to run the same prediction. Then, we evaluated the prediction performance derived from models using 17 variables against that using 18 variables. If the performance was acceptable, i.e., either improved or at least maintained, when a smaller set of predictors was used, then we continued removing the least important variable in the reduced set of predictors (i.e., 17 variables) and compared its prediction performance against that before the predictor was removed (i.e., 16 predictors against 17 predictors). This process was continued until the performance could not be maintained appropriately. Finally, the smallest subset of those predictor candidates that could still produce an acceptable performance was chosen. All six classifiers' performances were again assessed. Python (3.7) was used for the derivation and validation of different ML algorithms.

Results

A total of 77,977 unique patients were placed in our final analysis. Among them, 23.4% (18,223) patients were DM. The average age of these patients was 44 year old with standard deviation of 16 years. Forty-nine percent (49%) of patients were male and 51% were female patients. Patient portal use was found in 26.9% of DM patients. To predict the patient portal use among ED DM patients, the six different ML algorithms were used, and their performance accuracies were compared. Table 1 summarizes the performance measures of the six classifiers. Overall, the RF and DT classifiers outperform the

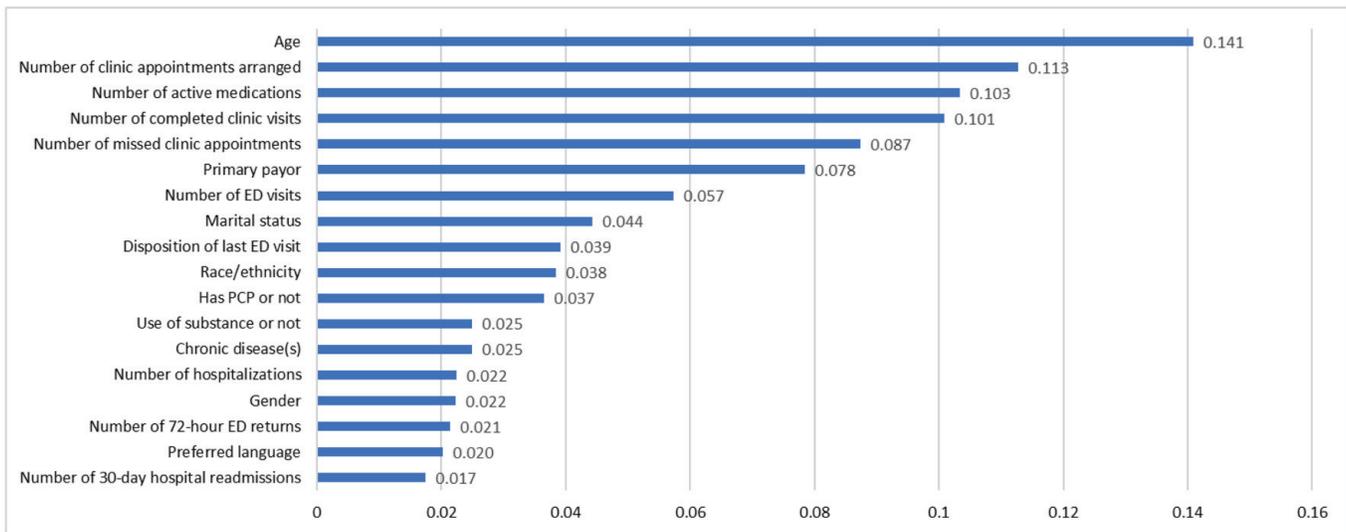


Figure 1. Feature importance scores.

other classifiers by providing an accuracy greater than 98%. The performances of these two classifiers are comparative. RF is slightly better than DT in specificity (0.997 vs. 0.989) and AUC (0.971 vs. 0.970), whereas DT is marginally better than RF in sensitivity (0.951 vs. 0.945).

Figure 1 ranks the individual variables based on their effects on the prediction model, with a higher score indicating a larger effect. It is noted that, among the top eight most important variables, three are demographical variables/social determinants (age, primary payor, and marital status), four are patient healthcare engagement-related variables (number of appointments arranged in clinics, number of completed and missed clinic appointments, and number of ED visits in the past 12 months), and one is individual medical condition variable (number of active medications). By following the recursive elimination feature selection procedure mentioned above, the top eight most important variables were selected in the final models. Table 2 shows the performance measures derived from the same six ML prediction models using those eight top-ranked variables.

Even with the use of the top eight contributors, the overall model performance accuracies among all six ML algorithms were either better or not changed (Table 2). The RF and DT classifiers still outperform the other classifiers by providing an accuracy of approximately 98%. The performances of these

two classifiers are comparative. Similar sensitivity (0.9374 vs. 0.9493) specificity (0.9932 vs. 0.9901), and AUC (0.9769 vs. 0.9710) are reported between RF and DT.

Discussion

In this study, different ML algorithms were used to predict patient portal use among ED DM patients. We found overall good predictive accuracies from five out of six different ML models, which indicates the applicability of using ML technology for patient portal prediction among individuals with DM. However, special attention should be paid when using ML prediction since different ML models might result in different performance accuracies [29]. In our models, though using the same dataset, SVM prediction yielded the poorest performance when compared with the other prediction models. In addition, some ML predictions provide results via a “black-box” analysis [16] which might have less value for physicians/administrators when determining the weight of the variables. However, the step of ML analysis using identified variables in our study overcomes the difficulties of “black-box” analysis. Using different ML models to determine the prediction performance accuracy, not only did we compare the accuracy of each ML model and

Table 2. A Comparison of Performance Measures of Six Different ML Algorithm Feature Selections (Top Eight Ranked Variables)

Classifier	Accuracy	Sensitivity	Specificity	AUC
Logistic regression	0.8351	0.7566	0.9839	0.8124
Random forest	0.9876	0.9374	0.9932	0.9769
Deep forest	0.9773	0.8962	0.9875	0.9510
Decision tree	0.9818	0.9493	0.9901	0.9710
Multilayer perception	0.8263	0.7798	0.9861	0.9634
Support vector machine	0.7052	0.6891	0.7134	0.6971

ML: machine learning; AUC: area under the receiver operating characteristic curve.

choose the best one for the predictions, but also identified important variables for outcome predictions. Determining these essential variables for outcome prediction can help develop effective interventions which yield improved outcomes. Moreover, when these essential variables were identified, subsequent ML predictions provide further validation about the value of these essential variables for future interventions.

This study does not emphasize findings of which ML models best predict patient portal use among ED DM patients. Instead, the dataset is a resource for different ML prediction models and validates the value of stepwise ML prediction. This study explores the usefulness of our “method” instead of being a healthcare prediction project. More importantly, this study shows the importance of choosing variables for ML predictions. More variables included in the ML prediction models seem to be less important than the “essential” variables included in the model. The primary ML prediction models used 18 variables, whereas the secondary ML prediction models only used eight variables. However, both yielded similar performance accuracies. This result is consistent with the “Garbage in, garbage out” phenomenon, frequently noted when ML is used in the field of medicine [30]. With the development of electronic medical records, numerous variables can be easily retrieved. However, most variables might not play an important role in the ML model predictions [18]. We suggest that screening, expert opinion, and group discussions to carefully choose essential variables should be the critical steps before initiating ML predictions.

Our study used a relatively large sample to predict patient portal use outcomes with the use of six different ML algorithms. We performed performance accuracy comparisons, one with 18 variables and the other, with only eight variables. We believe that these choices increase the validity of our results.

Our study also has its limitations. First, we only studied six ML prediction models, excluding other ML models from study and comparisons. Second, our dataset was limited to a single hospital ED, so uncertainty remains as to whether our stepwise prediction model can be generalized to different datasets. Third, since the choice of variables for the initial ML predictions largely depends on expert opinions and discussion, our study does not have data to validate this determination. Therefore, future studies are warranted to validate the value of using ML prediction models to predict clinical outcomes and further apply models for real-time predictions in medicine.

Conclusion

It is possible to predict patient portal use outcomes when different ML algorithms are used with fair performance accuracy. However, with similar prediction accuracies, the use of feature selection techniques can improve the interpretability of the model by addressing the most relevant features.

Supplementary Material

Suppl 1. List of predictor variables, data type, and values.

Suppl 2. Classifiers.

Acknowledgments

None to declare.

Financial Disclosure

None to declare.

Conflict of Interest

None to declare.

Informed Consent

This study was approved by institutional review board with waived informed consent.

Author Contributions

Conceptualization: YZ, TKS, and HW. Methodology: YZ, ZY, and HW. Validation: MB, JB, DZ, and HW. Formal analysis: YZ and ZY. Resources: HW, MB, and DZ. Writing-original draft preparation: YZ, ZY, and HW. Writing-review and editing: YZ, TKS, ZY, MB, JB, DZ, and HW. Supervision: HW.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

1. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol.* 2019;20(5):e262-e273. [doi](#)
2. Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. eDoctor: machine learning and the future of medicine. *J Intern Med.* 2018;284(6):603-619. [doi](#)
3. Jiang T, Gradus JL, Rosellini AJ. Supervised machine learning: a brief primer. *Behav Ther.* 2020;51(5):675-687. [doi](#) [pubmed](#) [pmc](#)
4. Heo J, Yoon JG, Park H, Kim YD, Nam HS, Heo JH. Machine Learning-Based Model for Prediction of Outcomes in Acute Stroke. *Stroke.* 2019;50(5):1263-1265. [doi](#)
5. Lalehzarian SP, Gowd AK, Liu JN. Machine learning in orthopaedic surgery. *World J Orthop.* 2021;12(9):685-699. [doi](#) [pubmed](#) [pmc](#)
6. Debnath S, Barnaby DP, Coppa K, Makhnevich A, Kim EJ, Chatterjee S, Toth V, et al. Machine learning to assist clinical decision-making during the COVID-19 pandemic. *Bioelectron Med.* 2020;6:14. [doi](#) [pubmed](#) [pmc](#)

7. Thandapani S, Mahaboob MI, Iwendi C, Selvaraj D, Dumka A, Rashid M, Mohan S. IoMT with deep CNN: AI-based intelligent support system for pandemic diseases. *Electronics*. 2023;12:424. [doi](#)
8. Iwendi C, Huescas CGY, Chakraborty C, Mohan S. COVID-19 health analysis and prediction using machine learning algorithms for Mexico and Brazil patients. *J Exp & Theoretical Artificial Intelligence*. 2022;1:21. [doi](#)
9. Safaripour R, June Lim HJ. Comparative analysis of machine learning approaches for predicting frequent emergency department visits. *Health Informatics J*. 2022;28(2):14604582221106396. [doi](#)
10. Tello M, Reich ES, Puckey J, Maff R, Garcia-Arce A, Bhattacharya BS, Feijoo F. Machine learning based forecast for the prediction of inpatient bed demand. *BMC Med Inform Decis Mak*. 2022;22(1):55. [doi](#) [pubmed](#) [pmc](#)
11. Ramasamy LK, Khan F, Shah M, Prasad B, Iwendi C, Biamba C. Secure smart wearable computing through artificial intelligence-enabled internet of things and cyber-physical systems for health monitoring. *Sensors (Basel)*. 2022;22(3):1076. [doi](#) [pubmed](#) [pmc](#)
12. Dinh A, Miertschin S, Young A, Mohanty SD. A data-driven approach to predicting diabetes and cardiovascular disease with machine learning. *BMC Med Inform Decis Mak*. 2019;19(1):211. [doi](#) [pubmed](#) [pmc](#)
13. Butt UM, Letchmunan S, Ali M, Hassan FH, Baqir A, Sherazi HHR. Machine learning based diabetes classification and prediction for healthcare applications. *J Healthc Eng*. 2021;2021:9930985. [doi](#) [pubmed](#) [pmc](#)
14. Contreras I, Vehi J. Artificial intelligence for diabetes management and decision support: literature review. *J Med Internet Res*. 2018;20(5):e10775. [doi](#) [pubmed](#) [pmc](#)
15. Rasheed K, Qayyum A, Ghaly M, Al-Fuqaha A, Razi A, Qadir J. Explainable, trustworthy, and ethical machine learning for healthcare: A survey. *Comput Biol Med*. 2022;149:106043. [doi](#)
16. Carvalho DV, Pereira EM, Cardoso JS. Machine learning interpretability: A survey on methods and metrics. *Electronics*. 2019;8(8):832. [doi](#)
17. Ahmad MA, Eckert C, Teredesai A. Interpretable machine learning in healthcare. In: *Proceedings of the 2018 ACM international conference on bioinformatics, computational biology, and health informatics*. 2018:559-560. [doi](#)
18. Stiglic G, Kocbek P, Fijacko N, Zitnik M, Verbert K, Cilar L. Interpretability of machine learning-based prediction models in healthcare. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. 2020;10(5):e1379. [doi](#)
19. Dunn J, Mingardi L, Zhuo YD. Comparing interpretability and explainability for feature selection. *arXiv preprint arXiv*. 2021.
20. Turner K, Hong YR, Yadav S, Huo J, Mainous AG. Patient portal utilization: before and after stage 2 electronic health record meaningful use. *J Am Med Inform Assoc*. 2019;26(10):960-967. [doi](#) [pubmed](#) [pmc](#)
21. Sun R, Korytkowski MT, Sereika SM, Saul MI, Li D, Burke LE. Patient portal use in diabetes management: literature review. *JMIR Diabetes*. 2018;3(4):e11199. [doi](#) [pubmed](#) [pmc](#)
22. Graetz I, Gordon N, Fung V, Hamity C, Reed ME. The digital divide and patient portals: internet access explained differences in patient portal use for secure messaging by age, race, and income. *Med Care*. 2016;54(8):772-779. [doi](#)
23. Sarkar U, Karter AJ, Liu JY, Adler NE, Nguyen R, Lopez A, Schillinger D. Social disparities in internet patient portal use in diabetes: evidence that the digital divide extends beyond access. *J Am Med Inform Assoc*. 2011;18(3):318-321. [doi](#) [pubmed](#) [pmc](#)
24. Bieler G, Paroz S, Faouzi M, Trueb L, Vaucher P, Althaus F, Daeppen JB, et al. Social and medical vulnerability factors of emergency department frequent users in a universal health insurance system. *Acad Emerg Med*. 2012;19(1):63-68. [doi](#)
25. Wang H, Kirby R, Pierce A, Barbaro M, Ho AF, Blalock J, Schrader CD. Patient portal use among diabetic patients with different races and ethnicities. *J Clin Med Res*. 2022;14(10):400-408. [doi](#) [pubmed](#) [pmc](#)
26. Woldaregay AZ, Arsand E, Walderhaug S, Albers D, Mamykina L, Botsis T, Hartvigsen G. Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes. *Artif Intell Med*. 2019;98:109-134. [doi](#)
27. Kavakiotis I, Tsave O, Salifoglou A, Maglaveras N, Vlahavas I, Chouvarda I. Machine learning and data mining methods in diabetes research. *Comput Struct Biotechnol J*. 2017;15:104-116. [doi](#) [pubmed](#) [pmc](#)
28. Recursive feature selection: addition or elimination? A smart way to carry out exhaustive feature selection.
29. Rahman MM, Ghasemi Y, Suley E, Zhou Y, Wang S, Rogers J. Machine learning based computer aided diagnosis of breast cancer utilizing anthropometric and clinical features. *IRBM* 2021;42(4):215-226. [doi](#)
30. Sanders H. Garbage in, garbage out: how purportedly great ml models can be screwed up by bad data? 2017.