Epilepsy & Behavior 127 (2022) 108548

Contents lists available at ScienceDirect

Epilepsy & Behavior

journal homepage: www.elsevier.com/locate/yebeh

Supervised machine learning to predict reduced depression severity in people with epilepsy through epilepsy self-management intervention

Edward J. Camp [a,*], Robert J. Quon [b], Martha Sajatovic [c], Farren Briggs [c], Brittany Brownrigg [c], Mary R. Janevic [d], Stephen Meisenhelter [a], Sarah A. Steimel [b], Markus E. Testorf [a,e], Elaine Kiriakopoulos [a], Morgan T. Mazanec [a], Robert T. Fraser [f], Erica K. Johnson [g], Barbara C. Jobst [a,b]

[a] Department of Neurology, Dartmouth-Hitchcock Medical Center, Lebanon, NH 03756, United States
[b] Geisel School of Medicine, Dartmouth College, Hanover, NH 03755, United States
[c] Case Western Reserve University School of Medicine, Cleveland, OH 44106, United States
[d] Center for Managing Chronic Disease, University of Michigan, Ann Arbor, MI 48109, United States
[e] Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, United States
[f] Department of Rehabilitation Medicine, University of Washington, Seattle, WA 98104, United States
[g] Health Promotion Research Center, University of Washington, Seattle, WA 98105, United States

ARTICLE INFO

Article history:
Received 20 August 2021
Revised 23 December 2021
Accepted 29 December 2021
Available online 15 January 2022

Keywords:
Depression
Epilepsy
Quality of life
Self-management
Machine learning
Support Vector Machine

ABSTRACT

Objective: To develop a classifier that predicts reductions in depression severity in people with epilepsy after participation in an epilepsy self-management intervention.

Methods: Ninety-three people with epilepsy from three epilepsy self-management randomized controlled trials from the Managing Epilepsy Well (MWE) Network integrated research database met the inclusion criteria. Supervised machine learning algorithms were utilized to develop prediction models for changes in self-reported depression symptom severity. Features considered by the machine learning classifiers include age, gender, race, ethnicity, education, study type, baseline quality of life, and baseline depression symptom severity. The models were trained and evaluated on their ability to predict clinically meaningful improvement (i.e., a reduction of greater than three points on the nine-item Patient Health Questionnaire (PHQ-9)) between baseline and follow-up (<=12 weeks) depression scores. Models tested were a Multilayer Perceptron (ML), Random Forest (RF), Support Vector Machine (SVM), Logistic Regression with Stochastic Gradient Descent (SGD), K-nearest Neighbors (KNN), and Gradient Boosting (GB). A separate, outside dataset of 41 people with epilepsy was used in a validation exercise to examine the top-performing model's generalizability and performance with external data.

Results: All six classifiers performed better than our baseline mode classifier. Support Vector Machine had the best overall performance (average area under the curve [AUC] = 0.754, highest subpopulation AUC = 0.963). Our analysis of the SVM features revealed that higher baseline depression symptom severity, study type (i.e., intervention program goals), higher baseline quality of life, and race had the strongest influence on increasing the likelihood that a subject would experience a clinically meaningful improvement in depression scores. From the validation exercise, our top-performing SVM model performed similarly or better than the average SVM model with the outside dataset (average AUC = 0.887).

Significance: We trained an SVM classifier that offers novel insight into subject-specific features that are important for predicting a clinically meaningful improvement in subjective depression scores after enrollment in a self-management program. We provide evidence for machine learning to select subjects that may benefit most from a self-management program and indicate important factors that self-management programs should collect to develop improved digital tools.

* Corresponding author at: Department of Neurology, Dartmouth-Hitchcock Medical Center, One Medical Center Drive, Lebanon, NH 03756, United States.

E-mail addresses: Edward.j.camp.18@dartmouth.edu (E.J. Camp), Robert.j.Quon.GR@Dartmouth.edu (R.J. Quon), Martha.Sajatovic@UHhospitals.org (M. Sajatovic), farren.briggs@case.edu (F. Briggs), Brittany.Brownrigg@UHhospitals.org (B. Brownrigg), mjanevic@umich.edu (M.R. Janevic), stephen.meisenhelter.gr@dartmouth.edu (S. Meisenhelter), sarah.a.steimel.gr@dartmouth.edu (S.A. Steimel), Markus.E.Testorf@Dartmouth.EDU (M.E. Testorf), Elaine.T.kiriakopoulos@dartmouth.edu (E. Kiriakopoulos), Morgan.T.Mazanec@hitchcock.org (M.T. Mazanec), rfraser@uw.edu (R.T. Fraser), ericajohnsonphd@uwalumni.com (E.K. Johnson), Barbara.C.jobst@hitchcock.org (B.C. Jobst).

https://doi.org/10.1016/j.yebeh.2021.108548
1525-5050/

1. Introduction

Epilepsy self-management represents an amalgamation of steps taken to manage the impact that seizures and seizure-related complications have on daily life. These steps are crucial for ensuring patient-centered epilepsy care by transitioning the "ownership" of care from the provider to the patient [1–3]. Epilepsy self-management interventions have become more popular since the establishment of the Prevention Research Centers' Managing Epilepsy Well (MEW) Network in 2007 [4–6]. This network works to develop self-management programs and tools, and disseminates them to persons with epilepsy and their care providers. The efforts of the MEW Network also led to an integrated database (MEW-DB) that pools data from numerous epilepsy self-management studies for secondary analyses [5,7].

Previous reports have examined factors correlated with health-related quality of life in epilepsy, including clinical seizure features, treatment-related features, and psychiatric comorbidities [8–10]. These studies established associations between commonly collected epilepsy measures to identify significant determinants of quality of life. Depression, measured by the nine-item Patient Health Questionnaire (PHQ-9), consistently demonstrated a strong correlation with many of the unfavorable effects of epilepsy [5,8,11,12], making it a comprehensive metric for gauging the impact of epilepsy self-management interventions.

In this study, we developed machine learning models to predict the efficacy of epilepsy self-management programs for reducing depression in individual persons with epilepsy. We were specifically interested in building a model that could identify a clinically meaningful improvement (i.e., a reduction of greater than three points on the PHQ-9) in depression scores between baseline and follow-up measurements. Using a unique dataset of 93 subjects, we trained several machine learning models, then reported detailed results and analytics from our top-performing model. We hypothesized that age would be a highly important feature, given Bautista et al.'s finding that older age was associated with superior utilization of self-management skills [13]; however, the comparative importance that other sociodemographic and clinical factors hold for predicting intervention-related outcomes previously remained unknown.

2. Materials and methods

The MEW Network provided longitudinal data from five prospective randomized controlled epilepsy self-management intervention trials: HOBSCOTCH (HOme-Based Self-management and COgnitive Training CHanges lives) from Dartmouth-Hitchcock Medical Center, PACES (Program of Active Consumer Engagement in Self-Management) from the University of Washington, FOCUS (Figure out the problem, Observe your routine, Connect your observations and choose a change goal, Undertake a change strategy, and Study the results) from the University of Michigan, and TIME and SMART (Self-Management for People with Epilepsy and a History of Negative Health Events) from Case Western Reserve University. The final dataset consisted of 453 deidentified people with epilepsy. After developing a top-performing model, the MEW Network provided a separate, outside dataset of 41 people with epilepsy from the Community Targeted Self-Management for Epilepsy and Mental Illness (C-TIME) ($n = 21$) and the Community Self-Management for People with Epilepsy and a History of Negative Health Events (C-SMART) ($n = 20$) for a validation exercise. All subjects provided informed consent for participation in the respective studies, and all studies were approved by the IRB of respective testing centers.

2.1. Exclusion criteria

As our goal was to assess depression through a longitudinal analysis, subjects were excluded if they only contributed one visit (e.g., only baseline visit). Further, to control for variable time lengths between baseline and subsequent visits across MEW trials, we only included subjects with post-baseline visits of 12 weeks or less. Subjects were also excluded if they were not actively participating in the treatment-arm of a study (i.e., control group subjects), or if they were missing data for any of the features of interest (described in Section 2.3). After applying these criteria, studies were excluded if they had fewer than five subjects remaining. This resulted in a final dataset consisting of 93 people with epilepsy (Fig. 1).

2.2. Target class

The outcome of interest or target class was a clinically meaningful improvement in PHQ-9 scores. Here, we followed Turkoz et al.'s finding and defined an improvement to be clinically meaningful if a subject's PHQ-9 score decreased by three or more points between the baseline and post-baseline visit [14]. Thus, the model output was a binary classification of success or failure in terms of a clinically meaningful response to the self-management program as assessed by depression.

2.3. Participant features

Features evaluated were age, gender, race, ethnicity, education, study program, baseline subjective quality-of-life score measured by the QOLIE-10, and baseline depression symptom severity measured by the PHQ-9. Baseline PHQ-9 scores were categorized as minimal depression (1–4), mild depression (5–9), moderate depression (10–14), moderately severe depression (15–19), and severe depression (20–27). We Z-scored continuous (e.g., age, QOLIE-10 mean score) and ordinal variables (e.g., education, PHQ-9 category) using Scikit-Learn (sklearn), then one-hot encoded nominal variables.

2.4. Model development

We trained models to predict alterations in depression using baseline features of 93 people with epilepsy from three RCTs (TIME, PACES, and SMART). During model training, 6 classifiers were fit with the preprocessed MEW dataset. Classifiers include the Multilayer Perceptron (ML), Random Forest (RF), Support Vector Machine (SVM), Logistic Regression with Stochastic Gradient Descent (SGD), K-nearest Neighbors (KNN), and Gradient Boosting (GB). To fine-tune model parameters for each classifier, we used sklearn's GridSearchCV with F1-scores as the scoring metric. Nested cross-validation was employed for parameter tuning and to evaluate the general performance of each classifier. The inner cross-validation was handled by sklearn's GridSearchCV class, which performed a stratified 5-fold cross-validation. Outer cross-validation was conducted using repeated ($n = 3$) stratified 5-fold cross-validation (i.e., 15-folds in total). We used sklearn's machine learning Python library for the developing, training, and evaluating the aforementioned classifiers.

2.5. Model evaluation

Predictive performance was examined using accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC). This was only evaluated for the best

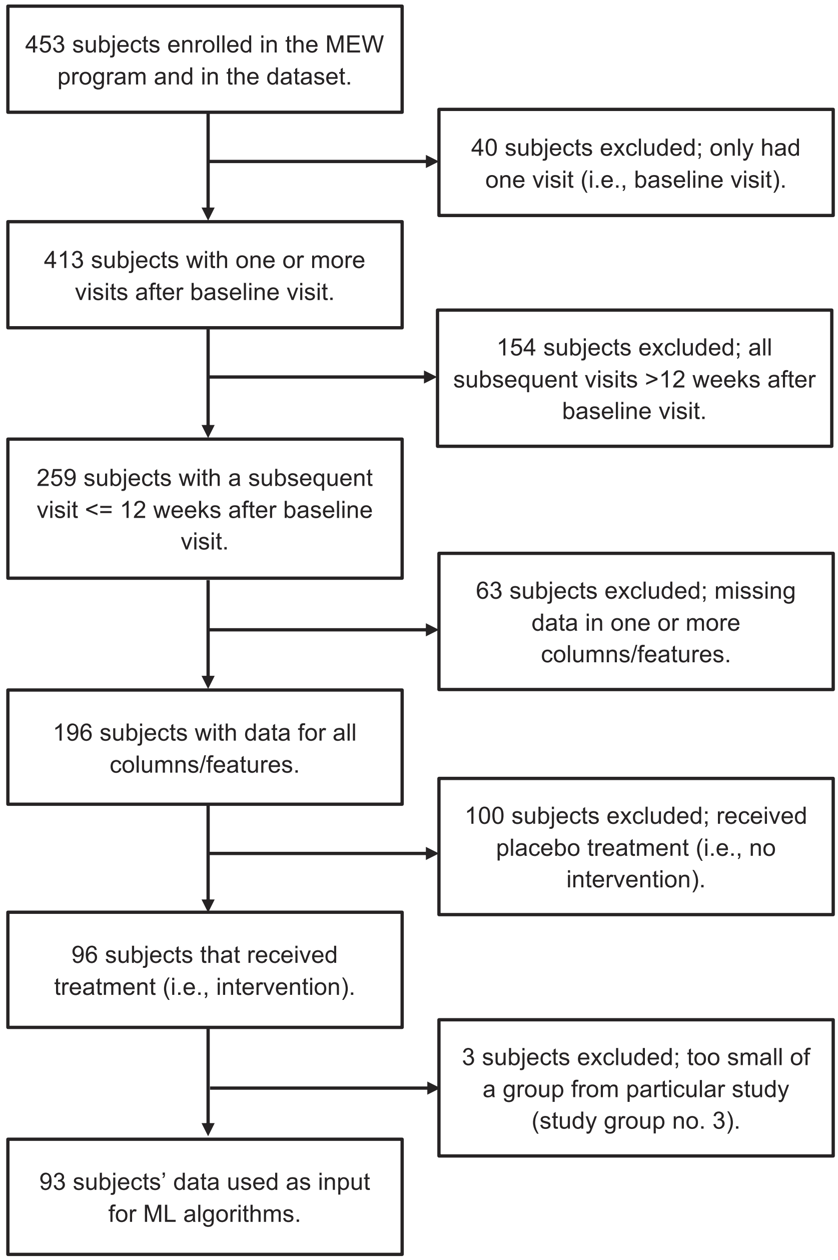

Fig. 1. MEW-DB data exclusion pipeline.

performing model in each outer fold. To compare general classifier performances, we also included a baseline mode classifier as a control; this control model classified all input as the most frequently occurring target class. To demonstrate the best performing classifier's robust performance across all its 15 folds, 1000 iterations of our dataset was bootstrapped and compared to the sample where the outcome labels were randomly shuffled using 95% confidence intervals.

2.6. Analyzing features from the top-performing model

The SHAP Python library was used to measure feature importance and the impact feature values had on a model's classification. SHAP calculates Shapley values, which represent the marginal contribution of each feature after all other features combinations were considered [15]. That is, Shapley values rank the feature's predictive importance and provide feature-specific directionality, explaining how changes in a feature value shifts a model's prediction.

2.7. Model validation exercise

We sought to validate our top-performing model to reduce the potential for a placebo effect and evaluate the generalizability of this classifier. To check for a placebo effect of improved PHQ9 score from natural fluctuations, we used the control group excluded from the MEW dataset ($n = 99$) as a negative control; one patient was excluded because their gender value was not found in the MEW train and test dataset. To examine the model's generalizability, we used outside datasets from independent epilepsy behavioral programs that were separate from the MEW dataset used to train and test our models: (1) C-TIME and (2) C-SMART. C-TIME ($n = 21$) and C-SMART ($n = 20$) were conducted at Case Western Reserve University to assess the implementation feasibility and outcomes of the SMART and TIME self-management interventions in a community setting [16]. With both control group and these new outside datasets, we calculated accuracy, precision, recall, and an F1-score to evaluate our classifier's performance.

3. Results

3.1. Subject characteristics

Table 1 depicts the characteristics of all subjects selected for inclusion in this study ($n = 93$). Subjects' mean age was 43.7 years (SD = 12.6), two-thirds were women, nearly half were African American, and approximately 60% had an education beyond high school. Of the different trials included in the MEW-DB, our filtered sample consisted of SMART (52.7%), PACES (30.1%), and TIMES (17.2%). Average QOLIE-10 scores, baseline PHQ-9 categorizations, and target class percentages are provided (Table 1). Characteristics of subjects from the control group and the outside datasets C-TIME and C-SMART can be found in Supplementary Table 1.

3.2. Model performance

Table 2 shows the mean and standard deviation of performance scores for all classifiers trained on the filtered sample with nested cross-validation. In general, all classifiers were more accurate than the baseline mode classifier. Support Vector Machine had the highest average scores across all performance metrics apart from recall (Table 2). The accuracies and F1-scores listed in Table 2 are visualized in SFig. 1. Comparing the best scores obtained from models for all classifiers on all 15 outer folds (i.e., subpopulations) demonstrated that SVM had the highest performance for all metrics except precision (Table 3). ROC curves generated from the best subpopulations, along with the ROC curve of the control group as a negative control comparison, showed that SVM dominated other models (Fig. 2A). Thus, SVM maintained the highest overall performance and was selected as our top-performing model (average AUC = 0.754, highest subpopulation AUC = 0.963). We also provided the average ROC curve for SVM with a 95% confidence interval (Fig. 2B). In addition, we calculated 95% confidence intervals for the accuracies of 1000 iterations of outcome shuffled and bootstrap samples on all 15 outer folds for SVM. All but one iteration of SVM did the bootstrap sample perform statistically significantly better than the outcome shuffled sample, further supporting SVM's robust performance (SFig. 2).

Table 1
Subject characteristics and baseline assessment scores on the MEW-DB data.

Variables	Total Sample, $N = 93$
Age, mean (SD) [range]	43.7 (12.6) [19–70]
Gender, *n* (%)	
Male	34 (36.6%)
Female	59 (63.4%)
Race, *n* (%)	
White	42 (45.2%)
Black/African-American	44 (47.3%)
Other	7 (7.5%)
Ethnicity, *n* (%)	
Not Hispanic or Latino	84 (90.3%)
Hispanic or Latino	9 (9.7%)
Education, *n* (%)	
High school or less	38 (40.9%)
At least some college	55 (59.1%)
Study, *n* (%)	
TIME	16 (17.2%)
PACES	28 (30.1%)
SMART	49 (52.7%)
QOLIE-10, mean (SD)	2.1 (0.9)
PHQ-9, mean (SD)	7.4 (5.6)
Minimal depression, *n* (%)	35 (37.6%)
Mild depression, *n* (%)	27 (29.0%)
Moderate depression, *n* (%)	20 (21.5%)
Moderately severe depression, *n* (%)	6 (6.5%)
Severe depression, *n* (%)	5 (5.4%)
Outcome/Target class, *n* (%)	
No clinically meaningful decrease in PHQ-9	54 (58.1%)
Clinically meaningful decrease in PHQ-9	39 (41.9%)

3.3. Feature importance and directionality

Using the top performing SVM model, we ranked the relative importance of each feature in terms of its contribution to the classifier's predictive performance. A permutation importance test revealed that baseline PHQ-9 category, study, and race were the most predictive features, respectively (Fig. 3A). We performed an additional feature importance analysis using SHAPLEY values, which represent the marginal contribution of each feature value toward the model's decision. Again, the baseline PHQ-9 category and study were the top two most predictive features, followed by the mean baseline QOLIE-10 score, and race (Fig. 3B).

In addition to feature importance, SHAPLEY values revealed the directional relationship between feature values and the model's prediction. We observed that a higher baseline PHQ-9 category (i.e., being more depressed at baseline) increased the likelihood of a positive response to a self-management intervention (Fig. 3B). Accordingly, a lower baseline QOLIE-10 score (i.e., subjects with higher perceived qualities of life at baseline) predicted an increased likelihood for a positive response to a self-management intervention (Fig. 3B). This suggests that subjects with more severe depression but better subjective qualities of life at baseline were more likely to benefit from a self-management intervention. Regarding the study, PACES was predicted to have the highest increase in likelihood for a positive response to self-management intervention. In terms of race, our model predicted that white subjects were more likely to have a positive response to a self-management intervention. Other feature-specific trends can be evaluated using Supplementary Table 2, which translates feature values to their true labels and associated colors.

3.4. Validation exercise performance

Using the control group, our SVM model had 56.57% accuracy (F1-score = 0.53, precision = 0.56, recall = 0.50, AUC = 0.53), which ranks below all classifiers including the mode classifier (accuracy = 57.89%). Using the outside datasets, our SVM model had 71.43% accuracy (F1-score = 0.67, precision = 0.86, recall = 0.55, AUC = 0.84) with the C-TIME dataset and 80% accuracy (F1-score = 0.67, precision = 1.0, recall = 0.50, AUC = 0.94) with the C-SMART dataset (Supplementary Table 3) (SFig. 3). The relatively high precision suggests that subjects with positive predictions have a high probability of improved depression after participation in an epilepsy self-management program. However, the low recall indicates nearly half of all subjects who benefitted from a self-management intervention (i.e., positive outcome) were identified as such by our classifier.

4. Discussion

Our goal was to utilize the rich aggregate of data collected from various epilepsy self-management programs (MEW-DB) to train machine learning models to predict depression outcomes and develop a computerized decision support tool [17,18]. While the SVM model had the best overall performance, we found that all classifiers generally performed better than our baseline mode classifier. An analysis of SVM features revealed that baseline depres-

Table 2
Average performance scores for all classifiers on the MEW-DB data.

Classifiers	Accuracy	AUC	Precision	Recall	F1-score
Multilayer Perceptron	58.7% (8.2%)	0.734 (0.103)	0.515 (0.082)	0.801 (0.196)	0.613 (0.088)
Random Forest	67.7% (8.4%)	0.700 (0.100)	0.639 (0.112)	0.511 (0.166)	0.559 (0.139)
Support Vector Machine	71.3% (8.6%)	0.755 (0.104)	0.662 (0.123)	0.645 (0.204)	0.640 (0.138)
Logistic Regression with SGD	66.7% (8.7%)	0.692 (0.110)	0.582 (0.195)	0.577 (0.243)	0.562 (0.192)
K Nearest Neighbor	63.4% (7.9%)	0.652 (0.109)	0.570 (0.091)	0.583 (0.135)	0.568 (0.093)
Gradient Boosting	63.8% (5.8%)	0.682 (0.075)	0.571 (0.083)	0.563 (0.147)	0.559 (0.092)
Mode	58.1%	–	–	–	–

Table 3
Best performance scores for all classifiers on the MEW-DB data.

Classifiers	Accuracy	AUC	Precision	Recall	F1-score	Mode
Multilayer Perceptron	79.0%	0.705	0.750	0.750	0.750	57.9%
Random Forest	79.0%	0.830	0.750	0.750	0.750	57.9%
Support Vector Machine	88.9%	0.963	0.800	1.000	0.889	55.6%
Logistic Regression with SGD	84.2%	0.807	0.857	0.750	0.800	57.9%
K Nearest Neighbor	79.0%	0.818	0.750	0.750	0.750	57.9%
Gradient Boosting	73.7%	0.767	0.636	0.875	0.737	57.9%

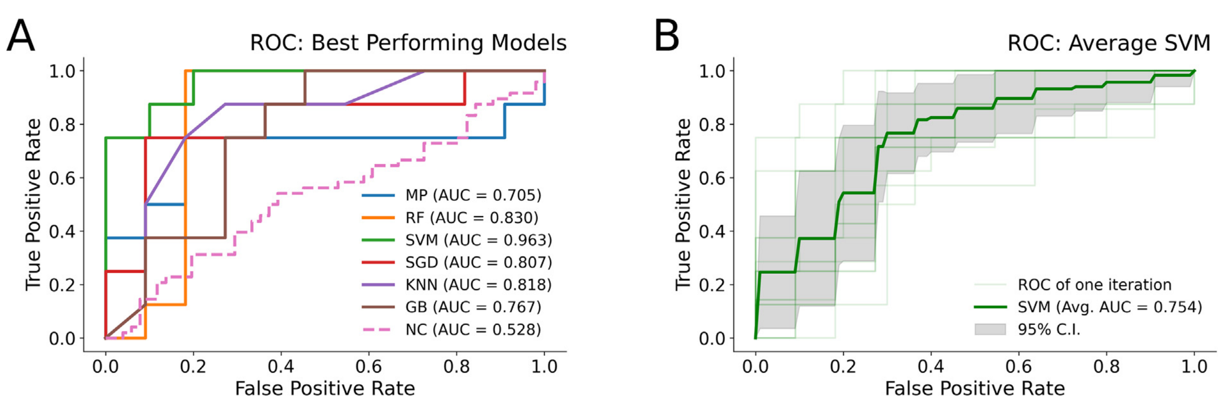

Fig. 2. Performance metrics for the supervised models on the MEW-DB data. (A) ROC curves from each classifier's best performance on their top subpopulation (i.e., outer fold), as well as SVM's performance on the control group (NC). (B) An average ROC curve for SVM, the best performing classifier.

sion (PHQ-9 category), study type (i.e., intervention program goals), baseline quality of life (QOLIE-10), and race influenced the likelihood that a subject would benefit from participation in a self-management intervention.

It was unsurprising that SVM outperformed other selected machine learning classifiers, as SVM has been shown to perform well with smaller datasets that contain a relatively large number of features [19,20]. There is also an abundance of evidence that SVM is superior to other supervised learning classifiers, especially for binary classification [21,22]. Other strengths of SVM are its high generalizability and sound theoretical basis [20,23]. Succinctly, the SVM classifier operates by dividing the data into two separate groups with a linear hyperplane that maximizes the distance between these groups [23,24].

After training and selecting our top-performing classifier, we utilized SHAP to discern underlying feature-level trends in our data [15,25]. The feature of highest importance was baseline PHQ-9 class, where a higher baseline PHQ-9 score (i.e., more severe depression) was associated with a greater predicted positive response to a self-management intervention. This could be explained by the fact that there is greater potential to observe an improvement in depression for subjects starting at a higher baseline level. However, it also indicates that baseline depression, assessed by the PHQ-9, may be a sensitive metric for targeting persons with epilepsy that would benefit most from enrollment in a self-management program.

Our feature analysis revealed that subjects enrolled in PACES had greater relative improvements in depression than subjects in TIME and SMART. This seems less likely due to one program outperforming the other, as all programs aided patients' depression, but may instead reflect differences in the enrolled samples of each study. The PACES program focused on improving medical and psychosocial management and enrolled a general epilepsy sample [26], while SMART focused on reducing negative health events among high-risk individuals who had recent poorly controlled seizures or had crisis care events such as emergency room visits, hospitalizations or self-harm attempts [27]. The TIME intervention focused on improving epilepsy outcomes among people with serious mental illness such as schizophrenia, bipolar disorder, or major depression [28]. Additional research is needed to investigate how diverse sub-groups with additional characteristics beyond those that were a focus of our model-testing respond to epilepsy self-management programs.

Furthermore, our negative control and external validation exercises serve to strengthen our model's robustness and generalizability for use outside of MEW. Earlier work with PACES reveals that both intervention and control groups experienced eventual improvement in PHQ9 scores after 6 months, which could be a placebo effect from either being enrolled in a self-management study or natural fluctuations [26]. To reduce the potential for placebo contribution, we ran our control group through trained SVM as a negative control validation exercise and found it performed very

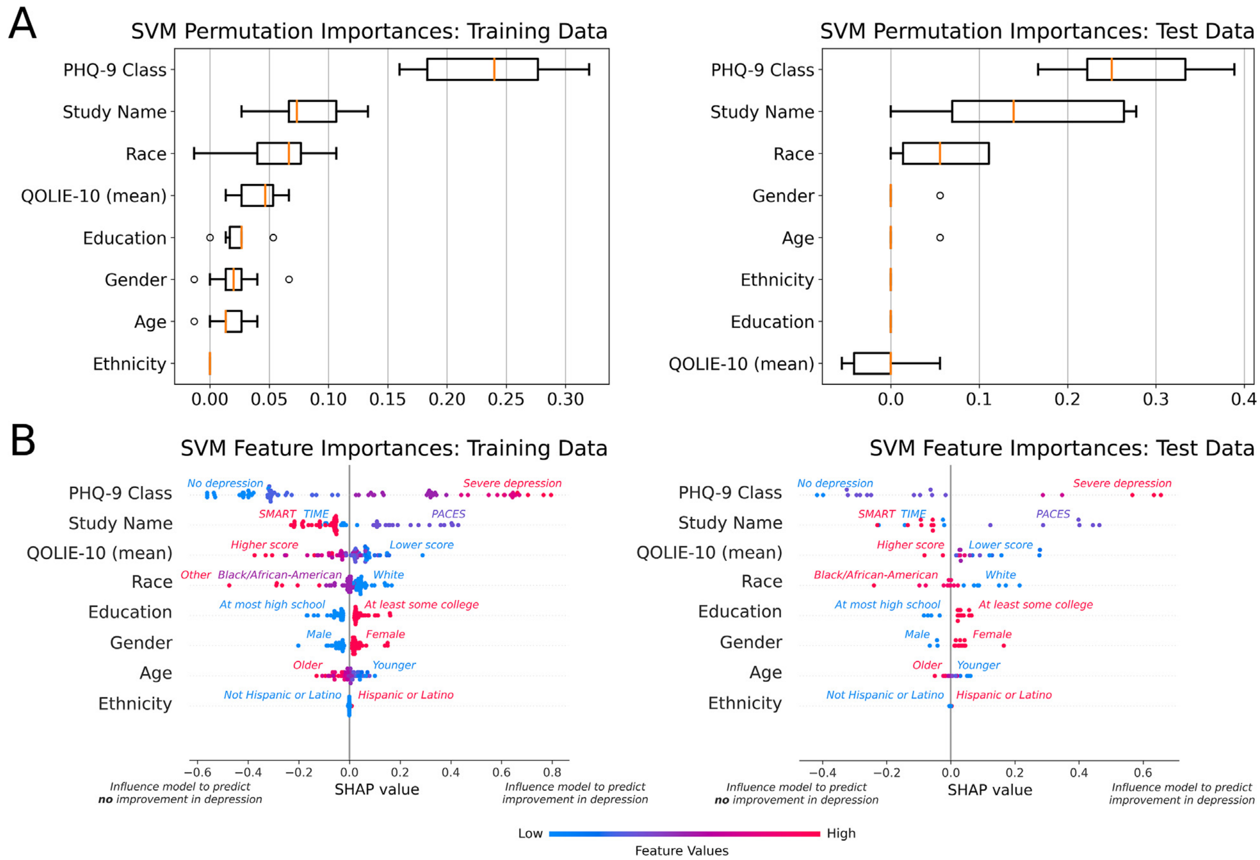

Fig. 3. Feature importance for the SVM classifier on MEW-DB data. (A) A permutation importance test was used to identify the highest ranked features (top–bottom descending rank). (B) An additional feature importance analysis was performed with SHAPLEY values, which provided ranked feature importance (top–bottom descending rank) and the directional relationship between feature values and the model's output.

poor compared to results on the MEW test dataset. Hence, these results help support that the observed effect that our SVM detects is not due to this placebo effect but from the intervention these subjects undergo in the study. Subsequently, an external validation exercise with outside datasets from independent epilepsy behavioral programs yielded results at least on par with the SVM's average performance. This supports that our model is generalizable, making it a good candidate for use as a tool for outside data.

Our study demonstrates the utility of machine learning as a complementary tool for screening people with epilepsy who may benefit from enrollment in a self-management intervention. As machine learning classifiers output (1) a class prediction and (2) a confidence or likelihood related to the prediction, we have designed an online tool that can be accessed at [https://mlmewcalculator.github.io]. Using this tool, providers can input baseline subject features (e.g., age, race, baseline QOLIE-10, baseline PHQ-9) to predict the likelihood that a subject will benefit from enrollment in an epilepsy self-management program. A sample of the user interface and output is provided as a supplementary figure (SFig. 4). This tool may be useful for guiding the clinical selection of patient populations that would benefit most from self-management interventions and providing patients with quantitative approximations of expected improvements after enrolling in an epilepsy self-management program.

This study must be viewed in light of a few limitations. Although the dataset was filtered to better harmonize time differences between baseline and post-baseline visits (i.e., different study follow-up time frequencies and durations), each program had different eligibility criteria and settings that could bias our results. Studies included in this MEW-DB sub-sample also did not require subjects to have clinically diagnosed depression, so we relied on subjective measures of depression (PHQ-9 scores). We also acknowledge that a 5-point PHQ-9 change is commonly used for clinical significance; however, this threshold was not used in our study, as it heavily skewed the balance of the target class in favor of no clinically meaningful change. Instead, a 3-point reduction in PHQ-9 was used to maintain a more balanced dataset, and as supported by Turkoz et al.'s finding, this threshold is a clinically meaningful improvement. Nonetheless, future work will use a larger sample population to assess 5-point changes in PHQ-9 scores. Regarding sample size, the generalizability of these findings was further limited by our exclusion criteria, which limited the number of subjects and studies included in model training. Additionally, our filtered dataset only contained data from three different studies, two of which were collected from the same testing center. We acknowledge that this may introduce sampling bias to the demographic results, such as race; however, we consider race an indicator of socioeconomic status and not a biological variable. More details on the socioeconomic status and study methods may be found in the original reports [26–28]. Despite these limitations, some strengths of this endeavor include a good representation of minorities in the dataset, insight into clinical and demographic features important for recommendations to a self-management intervention, and an online, user-friendly tool that clinicians may use to gauge potential outcomes if a person with epilepsy participates in a self-management program.

5. Conclusion

Machine learning may be useful for selecting subjects that may benefit most from enrollment in an epilepsy self-management program. Our model provides novel insights into features that are important for predicting an improvement in subjective depression in people with epilepsy. This proof-of-concept study indicates important factors that self-management programs can collect to develop improved digital tools to aid providers in their treatment recommendations. It may provide pragmatic next steps for epilepsy self-management interventions, highlighting important variables to collect and consider in future iterations. As the MEW-DB continues to grow, our findings encourage future work utilizing machine learning to guide treatment decisions and predict epilepsy-related health outcomes.

Funding sources

This study was funded and supported by The Susan Ellerie Diamond Endowment Research Development Award.

Declaration of interests

M.S. has received research grants from the following entities in the past 3 years: Nuromate, Otsuka, Alkermes, International Society for Bipolar Disorders (ISBD), National Institutes of Health (NIH), Centers for Disease Control and Prevention (CDC), and Patient-Centered Outcomes Research Institute (PCORI). She has consulted for: Alkermes, Otsuka, Janssen, Myriad, Health Analytics, and Frontline Medical Communications. She has royalties with: Springer Press, Johns Hopkins University Press, Oxford Press, and UpToDate. She also has CME activities in: American Physician's Institute, MCM Education, CMEology, Potomac Center for Medical Education, Global Medical Education, Creative Educational Concepts, and Psychopharmacology Institute. B.C.J. has received research funding from the following entities: NIH, National Science Foundation, DARPA, Centers for Disease Control and Prevention, Neuropace, Inc., Harvard-Pilgrim, Inc., and the Diamond Foundation. She serves as Associate Editor for the journal Neurology. She also has received travel support for activities in her roles as Committee Chair and Vice Chair of the Council of Education of the American Epilepsy Society.

All other authors have no conflicts of interest to report. We confirm that we have read the Journal's position on issues involved in ethnical publication and affirm that this report is consistent with those guidelines.

Acknowledgements

We would like to thank the participants in the epilepsy self-management intervention trials that contributed data to the MEW Network and this work. We are also grateful for the staff and study coordinators of each respective trial sites, without whom this study would not have been possible.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.yebeh.2021.108548.

References

[1] Helmers SL, Kobau R, Sajatovic M, Jobst BC, Privitera M, Devinsky O, et al. Self-management in epilepsy: Why and how you should incorporate self-management in your practice. Epilepsy Behav 2017;68:220–4. https://doi.org/10.1016/j.yebeh.2016.11.015.

[2] Chassin MR, Loeb JM. The ongoing quality improvement journey: next stop, high reliability. Health Aff (Millwood) 2011;30:559–68. https://doi.org/10.1377/hlthaff.2011.0076.

[3] Brady TJ, Anderson LA, Kobau R. Chronic disease self-management support: public health perspectives. Front Public Health 2015;2. https://doi.org/10.3389/fpubh.2014.00234.

[4] Sajatovic M, Jobst BC, Shegog R, Bamps YA, Begley CE, Fraser RT, et al. The Managing Epilepsy Well Network. Am J Prev Med 2017;52:S241–5. https://doi.org/10.1016/j.amepre.2016.07.026.

[5] Sajatovic M, Johnson EK, Fraser RT, Cassidy KA, Liu H, Pandey DK, et al. Self-management for adults with epilepsy: Aggregate Managing Epilepsy Well Network findings on depressive symptoms. Epilepsia 2019;60:1921–31. https://doi.org/10.1111/epi.16322.

[6] Dilorio CK, Bamps YA, Edwards AL, Escoffery C, Thompson NJ, Begley CE, et al. The Prevention Research Centers' Managing Epilepsy Well Network. Epilepsy Behav 2010;19(3):218–24. https://doi.org/10.1016/j.yebeh.2010.07.027.

[7] Sahoo SS, Zhang G-Q, Bamps Y, Fraser R, Stoll S, Lhatoo SD, et al. Managing information well: Toward an ontology-driven informatics platform for data sharing and secondary use in epilepsy self-management research centers. Health Informatics J 2016;22(3):548–61. https://doi.org/10.1177/1460458215572924.

[8] Johnson EK, Jones JE, Seidenberg M, Hermann BP. The relative impact of anxiety, depression, and clinical seizure features on health-related quality of life in epilepsy. Epilepsia 2004;45(5):544–50. https://doi.org/10.1111/j.0013-9580.2004.47003.x.

[9] Chen Y-y, Huang S, Wu W-Y, Liu C-R, Yang X-Y, Zhao H-T, et al. Associated and predictive factors of quality of life in patients with temporal lobe epilepsy. Epilepsy Behav 2018;86:85–90. https://doi.org/10.1016/j.yebeh.2018.06.025.

[10] Luoni C, Bisulli F, Canevini MP, De Sarro G, Fattore C, Galimberti CA, et al. Determinants of health-related quality of life in pharmacoresistant epilepsy: Results from a large multicenter study of consecutively enrolled patients using validated quantitative assessments: *Quality of Life in Pharmacoresistant Epilepsy*. Epilepsia 2011;52(12):2181–91. https://doi.org/10.1111/j.1528-1167.2011.03325.x.

[11] Ettinger AB, Good MB, Manjunath R, Edward Faught R, Bancroft T. The relationship of depression to antiepileptic drug adherence and quality of life in epilepsy. Epilepsy Behav 2014;36:138–43. https://doi.org/10.1016/j.yebeh.2014.05.011.

[12] Boylan LS, Flint LA, Labovitz DL, Jackson SC, Starner K, Devinsky O. Depression but not seizure frequency predicts quality of life in treatment-resistant epilepsy. Neurology 2004;62(2):258–61. https://doi.org/10.1212/01.WNL.0000103282.62353.85.

[13] Bautista RED, Shoraka AR, Shapovalov D. Factors associated with superior self-management skills among individuals with epilepsy. Epilepsy Behav 2014;41:221–6. https://doi.org/10.1016/j.yebeh.2014.10.012.

[14] Turkoz I, Alphs L, Singh J, Jamieson C, Daly E, Shawi M, et al. Clinically meaningful changes on depressive symptom measures and patient-reported outcomes in patients with treatment-resistant depression. Acta Psychiatr Scand 2021;143(3):253–63. https://doi.org/10.1111/acps.13260.

[15] Sundararajan M, Najmi A. The many Shapley Values for model explanation n.d.:11.

[16] Sajatovic M, Needham K, Colón-Zimmermann K, Richter N, Liu H, Garrity J, et al. The Community-targeted Self-management of Epilepsy and Mental Illness (C-TIME) initiative: A research, community, and healthcare administration partnership to reduce epilepsy burden. Epilepsy Behav 2018;89:175–80. https://doi.org/10.1016/j.yebeh.2018.10.004.

[17] Contreras I, Vehi J. Artificial intelligence for diabetes management and decision support: literature review. J Med Internet Res 2018;20(5):e10775. https://doi.org/10.2196/10775.

[18] Triantafyllidis AK, Tsanas A. Applications of machine learning in real-life digital health interventions: review of the literature. J Med Internet Res 2019;21:e12286. https://doi.org/10.2196/12286.

[19] Cervantes J, Garcia-Lamont F, Rodríguez-Mazahua L, Lopez A. A comprehensive survey on support vector machine classification: Applications, challenges and trends. Neurocomputing 2020;408:189–215. https://doi.org/10.1016/j.neucom.2019.10.118.

[20] Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of Support Vector Machine (SVM) learning in cancer genomics. Cancer Genomics Proteomics 2018;15. https://doi.org/10.21873/cgp.20063.

[21] Statnikov A, Wang L, Aliferis CF. A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. BMC Bioinf 2008;9:319. https://doi.org/10.1186/1471-2105-9-319.

[22] Liang X, Zhu L, Huang D-S. Multi-task ranking SVM for image cosegmentation. Neurocomputing 2017;247:126–36. https://doi.org/10.1016/j.neucom.2017.03.060.

[23] Priya A, Garg S, Tigga NP. Predicting anxiety, depression and stress in modern life using machine learning algorithms. Procedia Comput Sci 2020;167:1258–67. https://doi.org/10.1016/j.procs.2020.03.442.

[24] Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20:273–97. https://doi.org/10.1007/BF00994018.

[25] Lundberg SM, Lee S-I. A unified approach to interpreting model predictions n.d.:11.

[26] Fraser RT, Johnson EK, Lashley S, Barber J, Chaytor N, Miller JW, et al. PACES in epilepsy: results of a self-management randomized controlled trial. Epilepsia 2015;56:1264–74. https://doi.org/10.1111/epi.13052.

[27] Sajatovic M, Colon-Zimmermann K, Kahriman M, Fuentes-Casiano E, Liu H, Tatsuoka C, et al. A 6-month prospective randomized controlled trial of remotely delivered group format epilepsy self-management versus waitlist control for high-risk people with epilepsy. Epilepsia 2018;59:1684–95. https://doi.org/10.1111/epi.14527.

[28] Sajatovic M, Tatsuoka C, Welter E, Perzynski AT, Colon-Zimmermann K, Van Doren JR, et al. Targeted Self-Management of Epilepsy and Mental Illness for individuals with epilepsy and psychiatric comorbidity. Epilepsy Behav 2016;64:152–9. https://doi.org/10.1016/j.yebeh.2016.08.012.