This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

Bipolar disorder prediction with sensor-based semi-supervised learning

D2.6 – Guidelines for selecting features in BD scenarios

Deliverable No.	D2.6	**Due date**	30-JUN-2023
Type	Report	**Dissemination Level**	Public
Version	1.0	**WP**	WP2
Description	Guidelines for selecting features in BD scenarios		

Systems Research Institute Polish Academy of Sciences, Newelska 6, 01-447 Warsaw, Poland

Authors

Name	Email
Katarzyna Kaczmarek-Majer	k.kaczmarek@ibspan.waw.pl
Olga Kamińska	o.kaminska@ibspan.waw.pl
Kamil Kmita	k.kmita@ibspan.waw.pl
Jakub Małecki	jakub.malecki@ibspan.waw.pl

History

Date	Version	Change
01-Jun-2023	0.1	First integrated version of the document
28-Jun-2023	0.2	Version for internal review
30-Jun-2023	1.0	Version ready for submission

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

Executive summary

This deliverable is another output of Task 2.1 and Task 2.2 activities dedicated to feature selection. It further extends the Deliverable D2.3. which was aiming at the feature preprocessing. This task requires that the input data are prepared, so it depends on T.2.1. This assumption was met.

The primary goal for task T2.2. was to verify whether the uncertainty-aware approach for feature selection outperforms the selected baseline approaches in this particular mental health monitoring context.

This report is aimed to summarize the main lessons learned during the feature preprocessing and selection process for the considered context of partially labeled sensor-based data.

First, selected baseline approaches are considered as benchmark, such as e.g., the well-known recursive feature elimination (RFE). The idea of RFE technique is to build a model with all variables and after that the algorithm removes one by one starting from the worst performing one. To find the optimal number of features cross-validation will be used.

Next, analyses were performed to verify the hypothesis that there is one subset of features adequate for all bipolar patients or whether adequate subsets of features shall be further diversified.

Next, we present a cost-constrained approach for feature selection. We are adding a cost factor parameter that controls the trade-off between feature importance and its cost. The experiments were performed on real-life database collected from patients with bipolar disorder during their daily mobile calls. The results indicate that the cost-constrained method allows to achieve better results.

Next, we present an uncertainty-aware approach for feature selection, in which we incorporated the confidence factor developed mainly in WP3.

This deliverable is also closely related to the software component developed under WP4. Features recommended for selection will be used for evaluation in tasks in WP1.

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

Systems Research Institute Polish Academy of Sciences, Newelska 6, 01-447 Warsaw, Poland

Table of Contents

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

List of acronyms

Acronym	Explanation
BIPOLAR	Bipolar disorder prediction with sensor-based semi-supervised learning project
BD	Bipolar disorder
ADP	Acoustic Data Pilot
LDP	Locomotor Data Pilot
SSFC	Semi-supervised fuzzy clustering
SSFCM	Semi-supervised fuzzy c-means algorithm
SSL	semi-supervised learning
SSMC-FCM	Algorithm with Multiple Fuzzification Coefficients
DISSFCM	Dynamic Incremental Semi-Supervised Fuzzy C-Means
S3FCM	Safe Semi-Supervised Fuzzy C-Means
CPR	Confidence Path Regularization
RPCM	Repulsive Possibilistic C-Means
PFCM	Possibilistic Fuzzy C-Means
SSPFCM	Semi-Supervised Possibilistic Fuzzy C-Means
PCM	Possibilistic C-Means algorithm
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses

1.Introduction

1.1.About this document

The aim of this report is to summarized main findings related to feature selection for BD monitoring

1. First, selected baseline approaches are considered as benchmark, such as e.g., the well-known recursive feature elimination (RFE). The idea of RFE technique is to build a model with all variables and after that the algorithm removes one by one starting from the worst performing one. To find the optimal number of features cross-validation will be used.
2. Next, analyses were performed to verify the hypothesis that there is one subset of features adequate for all bipolar patients or whether features shall be further divided, e.g., each patient shall have unique subset of features
3. Next, we present an cost-constrained approach for feature selection. We are adding a cost factor parameter that controls the trade-off between feature importance and its cost. The experiments were performed on a large medical database collected from patients with bipolar disorder during their daily mobile calls. The results indicate that the cost-constrained method allows to achieve better results.
4. Next, we present an uncertainty-aware approach for feature selection, in which we incorporated the confidence factor developed in WP3.
5. Finally, there is the last section, where we list the future extensions of the discussed topics.

2. Related work

Feature selection has been discussed in Deliverable 2.4, which was prepared on the basis of the first work on feature selection, where we distinguished several basic approaches. In this first initial deliverable, we were focusing mainly on the aggregated datasets. Now, we extend this approach to the non-aggregated data.

In the case of aggregated data, we distinguish the following approaches where we differentiate feature selection:

a. due to patient selection;
b. due to the method of data aggregation (mean vs standard deviation);
c. due to the type of label (4 classes vs 2 classes).

However, in the case of non-aggregated data (e.g., all frames from recordings) for a few selected patients [12] were used, the SHAP [17] method was chosen to interpret the acoustic features divided into depression and euthymia. All of the mentioned use cases produced different subsets of features. However, it could be still insufficient and further expert knowledge is needed.

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

In [14], the authors indicate that the most important acoustics features regarding to bipolar disorder episodes classification are following features:

- f0env sma;
- slope0500 sma3;
- pcm fftMag mfcc 1;
- pcm fftMag mfcc 4;
- pcm fftMag mfcc 6.

The authors further show that using recursive feature elimination method leads to the improved classification results.

In [15], the authors individually set the list of important acoustic parameters depending on the aggregation operator and clusters derived from data. The two aforementioned papers take into account novel factors like the method of how to aggregate data, but it has no linkage with expert knowledge that should have higher influence for the considered problem of establishing the best subset of acoustic features for bipolar disorder monitoring.

Both theoretical and practical approaches indicate that specific patterns of vocal modulation are associated with distinct emotional states. Emotions and moods can induce changes in breathing patterns, phonation, and articulation, which are manifested in the speech signal [8]. Initial studies have revealed that even basic features such as pitch, speech rate, loudness, and rhythm have an impact on emotion perception and can be utilized to differentiate between specific affective states [9]. For instance, states like anger and fear are characterized by an increased speech rate, high fundamental frequency (F0) values, and a wide range of intonation.

Drawing from consultations with psychiatric experts, our hypothesis revolves around the notion that there exists a significant correlation between affective states and bipolar disorders. Thus, by leveraging discriminative speech features in emotion recognition, we can potentially differentiate between the aforementioned disorders.

Since we distinguish between other characteristics of both manic and depressive episodes, this distinction has been introduced in the current methods.

In our work, the Recursive Feature Elimination (RFE) [13] is considered as a baseline method. Recursive Feature Elimination (RFE) offers several advantages for feature selection. It facilitates the selection of an optimal subset of features by iteratively eliminating the least important ones, ensuring that the chosen features contribute most to the model's performance. RFE is model-agnostic, making it compatible with various machine learning algorithms and enabling its application across different domains. It effectively handles multicollinearity by considering the interdependencies among features. Additionally, RFE enhances interpretability by selecting a reduced set of features, making the model more understandable. However, RFE can be computationally intensive, especially for large datasets or complex models. Its effectiveness relies on the model's ability to accurately assess feature importance, and there is a risk of overfitting if the stopping criterion or feature importance estimates are not properly set. Moreover, RFE assumes that less important features can be eliminated without affecting model performance, which may not always hold true.

RFE can be an effective method of feature selection in certain scenarios, but its performance can vary depending on the dataset and the model being used. In some scenarios in the experiments, we have slightly modified RFE to better fit our case study.

3. Cost-constrained feature selection

In this Section, we present the cost-constrained feature selection method for acoustic data, that was described during XAI workshops with AIME 2023 conference [8].

That method has been implemented to analyze whether the acoustic features are all important and demanding to maintain the high performance of classifying the BD state. We compare it with behavioral features (describing nature of phone usage like the number of incoming/outgoing calls per day or the average length of characters in daily text messages).

That method is based on the greedy forward selection algorithm applicable thanks to information theory.

Let $X1, \ldots, Xp$ be the features, $F = \{1, \ldots, p\}$ be the set of features indexes, Y be a multiclass variable that we consider as a target and $c(k)$ be the cost of the k -th feature. In each step of the algorithm, we add the index of the candidate feature Xk to the set of features already selected in the previous steps $S \leftarrow S \cup \{k*\}$ by maximizing the following equation:

$$k^* = \arg \max_{k \in F \setminus S} [J(X_k, Y | X_S) - \lambda c(k)], \quad (1)$$

The general idea behind Eq. 1. is that it allows us to find the trade-off between the relevance of the feature and its costs and λ controls the balance. The optimal parameter λ value can be calculated by minimizing the loss function with the cross-validation on the training set.

For our bipolar use case, all of features have been divided into 2 groups with different costs as depicted in Table 1.

Table 1. Different types of features

	Feature Type	Cost value
1	BEHAVIORAL FEATURES ABOUT MOBILE CALLS & TEXT MESSAGES e.g.: number of outgoing calls per day, number of incoming calls per day, the average length of characters in daily text messages	1

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

2	ACOUSTIC FEATURES [0-85] e.g.: f0 sma, jitterlocal sma, pcm fftMag mfcc 0 12	5

Now we discuss the main results. Figure 1. presents the AUC of a model trained on features selected. In the first step of both algorithms, the feature number 50 (magnitude of fast Fourier transform coefficients in band 250-650Hz) from the acoustic group is selected. That is obvious because it is the most informative feature. In the second step cost-constrained method chooses feature number 92 (ratio of outgoing calls to all mobile calls per day) from the behavioral group and the traditional method feature number 6 (another magnitude of fast Fourier transform coefficients) from the acoustic group. The cost-constrained method in the third step selects feature number 36 (energy in the specific band), which combined with two previous features results in 0.70 AUC score, at the same time the traditional method results in 0.67 for almost the same budget. As the budget increases, both methods select more features, but the cost-constrained method tends to result in a higher AUC score.

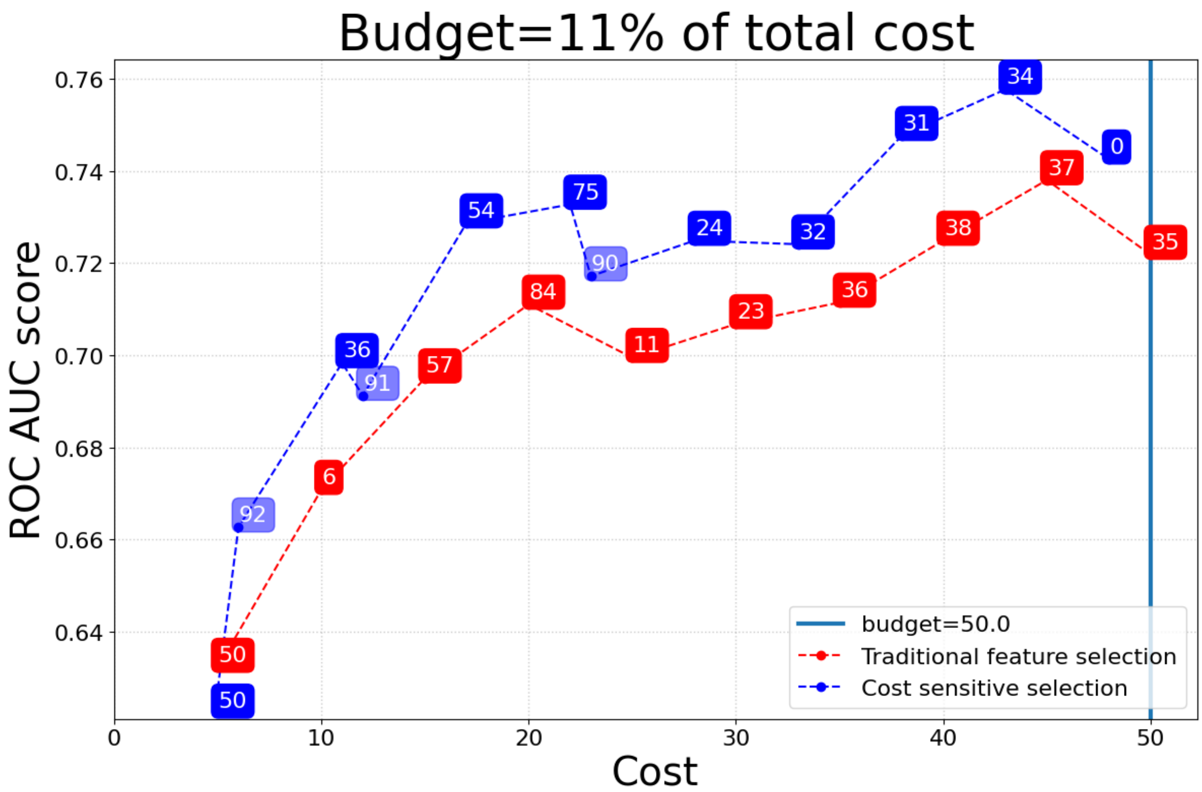

Figure1. ROC AUC score for budget=50.

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

Figure 2. presents the importance of all selected features measured as mutual information with the target variable. Variables in the acoustic group result in greater mutual information with the target class, but at the same time, they are very expensive. On the other hand, we have behavioral features, which are 5 times cheaper, but their mutual information with the target class is lower.

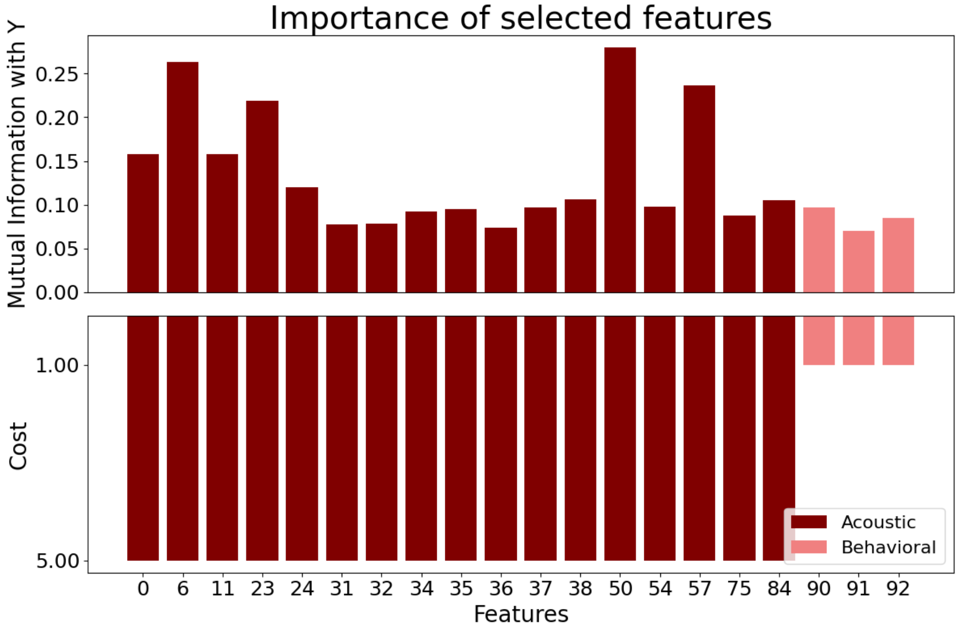

Figure 2. Importance of selected features

The legend for the selected features of Figure 2 is as follows:

Number	Feature name	Feature Type
0	pcm_LOGenergy_sma	Acoustic
6	pcm_fftMag_fband0-650_sma	Acoustic
11	pcm_fftmag_spectralflux_sma	Acoustic
23	pcm_rmsenergy_sma	Acoustic
24	audSpec_Rfilt_sma_compare_0_	Acoustic
31	audSpec_Rfilt_sma_compare_7_	Acoustic
32	audSpec_Rfilt_sma_compare_8_	Acoustic
34	audSpec_Rfilt_sma_compare_10_	Acoustic
35	audSpec_Rfilt_sma_compare_11_	Acoustic
36	audSpec_Rfilt_sma_compare_12_	Acoustic
37	audSpec_Rfilt_sma_compare_13_	Acoustic
38	audSpec_Rfilt_sma_compare_14_	Acoustic
50	pcm_fftMag_fband250-650_sma_compare	Acoustic
54	pcm_fftmag_spectralskewness_sma_compare	Acoustic
57	pcm_fftmag_spectralharmonicity_sma_compare	Acoustic

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

75	pcm_fftMag_mfcc_2_	Acoustic
84	pcm_fftMag_mfcc_11_	Acoustic
91	ratio of incoming calls to all mobile calls per day	Behavioral
92	ratio of outgoing calls to all mobile calls per day	Behavioral

Variables in the acoustic group result in greater mutual information with the target class, but at the same time, they are very expensive whereas behavioral features (5 times cheaper) has lower mutual information with the target class.

4. Uncertainty-aware feature selection

4.1. The proposed Uncertainty-aware Feature Selection UFS approach

In this Section, we discuss the uncertainty aware approach. The primary goal of the ***Uncertainty-aware Feature Selection (UFS)*** approach is to improve the predictive performance due to adequate consideration of imprecision related to the sensor data and psychiatric assessments during the feature engineering and selection stage. In the context of remote monitoring for bipolar disorder (BD), it is common to extend the psychiatrist's evaluation to a specific time frame surrounding the visit, known as the ground truth period.

In the related works (see e.g., [1,2,4,16]) the ground truth period refers to the fixed time window surrounding the visit which is from 7 days before the visit up to 2 days after the visit. This period serves as the basis for extrapolating the label. As a result, all data within this period are categorized uniformly. However, despite the assumption that most phone calls within the ground truth period exhibit certain characteristics of the disease phase, we lack direct labels specifically assigned to these supervised calls. The ground-truth based method has the potential to lead to misleading oversight, which can impact the performance of feature selection algorithms and later, the performance of the predictive model. Therefore, **we address the issue of label uncertainty by assuming that the labels are precise but may be assigned to individual observations with varying levels of confidence.**

Our goal is to model the varying confidence about labeling already at the feature selection stage. The confidence factor is simulated as various scenarios. In practice, these confidence factor calculations can be realized using the *add_confidence* from the *BIPOLAR preprocessing* R package.

Let us consider a small illustrative example when the following data are available:

- patients' visits data - when a patient visited a psychiatrist.
- mobile calls recordings - voice parameters from mobile calls.

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

In most of the state of the art, some ground truth period is assumed at the beginning and then, all observations coming from this period are considered as equally certain, so in the range of -7 to 2 days around the visit day our confidence about the state is the same and equals 1 [Table2].

Table 2. Configuration of labels extension.

BD episode	No. of days before visit	No. of days after visit
euthymia	7	2
depression	7	2
mania	7	2
mixed	7	2

Our goal is to extrapolate the state from a visit date to a wider time window and introduce variability. We'd also like to quantify the confidence of our extrapolation. We're going to model how confident we are when extrapolating the actual patient's state on the time range we defined. For example, we may assume that we are most certain about the state in day 0, i.e. the visit day. Moving away from the visit day our confidence may decrease. This is one of possible scenarios as depicted in Figure 3.

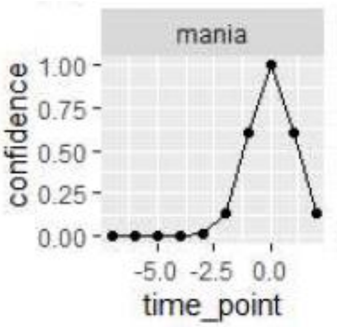

Figure 3. Confidence factor function dependent on time (0 indicates the day of the visit with the psychiatrist) using the gaussian density function.

We can use other time windows and other functions. The important aspect is that confidence factor is always expressed as a number between 0 and 1. For more running examples related to the calculations of the confidence factor, please see the the vignette in the BIPOLAR package[1].

The proposed uncertainty-aware feature selection approach is further extension of [3]. In [3], we proposed the Confidence Path Regularization (CPR) able to incorporate this uncertainty into the fuzzy c-means semi-supervised learning.

[1] bipolar/vignettes/adding_confidence.Rmd at main · ITPsychiatry/bipolar · GitHub

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

The developed uncertainty-aware approaches will be demonstrated in two use cases: (1) smartphone sensors for monitoring of depression and mania; (2) locomotor sensors for monitoring of psychomotor disturbances in depression.

Our baseline method used for feature selection (mentioned in paragraph 2) is called RFE (recursive feature elimination) [13]. As a classifier we are using XGboost [11] algorithm with weights where weighting observation means increasing the contribution of an example to cost function. As a weight we are using the confidence which indicates how confident is assessed label. The confidence structure is set in table 2. It means that psychiatrics assessments obtained during visits are extended into 7 days before visit and 2 days after visits and day zero is set as visit day and received highest confidence value. Selected extension is suggested with literature [4] and domain knowledge.

4.2. Acoustic use case

Since we distinguish between other characteristics of both manic and depressive episodes, this distinction has been introduced in the current methods. Used dataset contains 86 acoustic features describing voice, its characteristics and it was collected during unforced telephone mobile calls of patients suffering from bipolar disorder.

First approach we are using all aggregated data from mobile recordings of patients with BD.

Next approaches considered mania and euthymia episodes to distinguish features that differentiate features that best characterize mania. Similar situation Is considered with depression state. Moreover, both mania and depression are divided into male and female cases. In total we are testing 7 use cases.

4.2.1. Recommended features for all BD patients

Figure 5 presents relative importance from the first iteration of UFS methods. On the X-axis we can see relevance of feature, and on Y-axis top 20 features characterized highest importance for classification selected BD state.

The most important feature seems to be f1bandwidth_sma3nz and it's placed into the final list with selected features for that use case presented in Table 3.

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

Systems Research Institute Polish Academy of Sciences, Newelska 6, 01-447 Warsaw, Poland

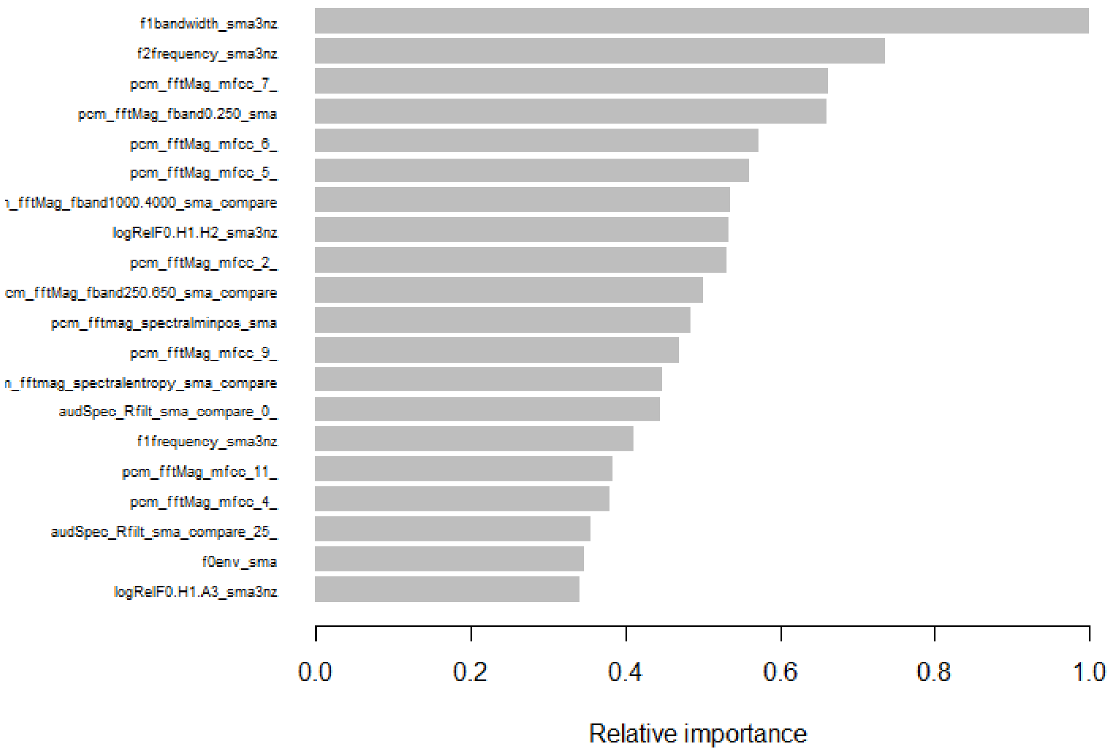

Figure5. Figure shows relative importance for first iteration of UFS method.

Table3. Selected features for all BD states.

f1bandwidth_sma3nz
f2frequency_sma3nz
f1frequency_sma3nz
pcm_fftMag_mfcc_7_
pcm_fftMag_mfcc_9_

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

4.2.2. Recommended features for all patients in the mania state only

In Figure 6, we are presenting relative importance from the first iteration of UFS methods. The most important feature seems to be f1bandwidth_sma3nz and it's placed into the final list with selected features for that use case presented in Table 4.

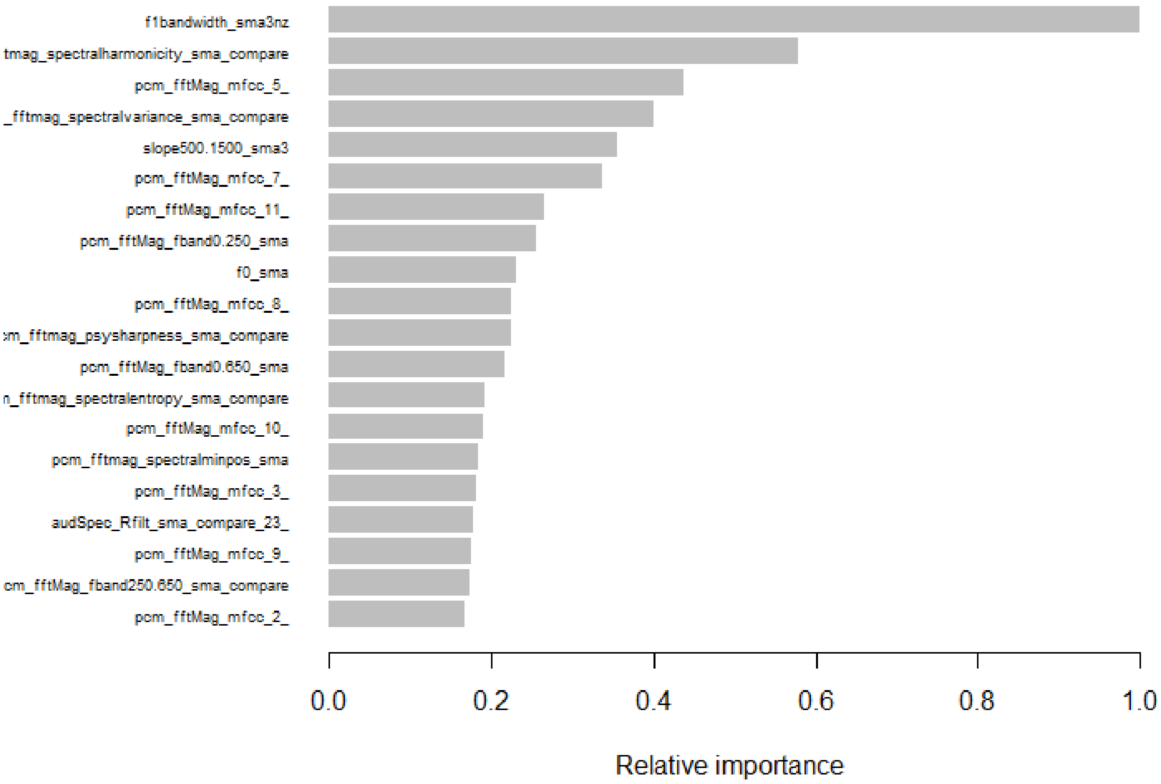

Figure 6. Figure shows relative importance for first iteration of UFS method for mania state.

Table4. Selected features for mania state.

f1bandwidth_sma3nz
pcm_fftmag_spectralvariance_sma_compare
pcm_fftMag_mfcc_5_
slope500.1500_sma3
pcm_fftMag_mfcc_4_

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

4.2.3. Recommended features for all patients in the depressive state only

In Figure 7, we are presenting relative importance from the first iteration of UFS methods. The most important feature seems to be f1frequency_sma3nz and it's placed into the final list with selected features for that use case presented in Table 5.

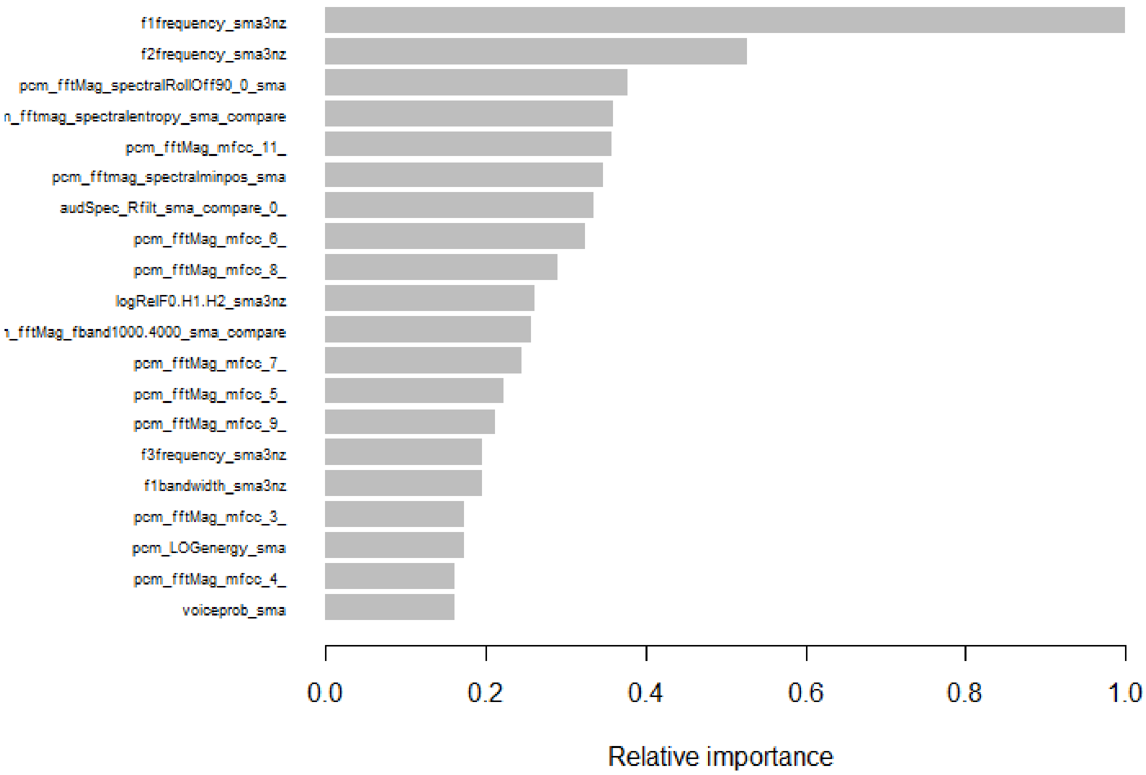

Figure 7. Figure shows relative importance for first iteration of UFS method for depression state.

Table5. Selected features for depression state.

f1frequency_sma3nz
f2frequency_sma3nz
f3frequency_sma3nz
slope0.500_sma3
pcm_fftMag_fband1000.4000_sma_compare

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

4.2.4. Recommended features for all patients in the mania state for female patients

In Figure 8, we are presenting relative importance from the first iteration of UFS methods. The most important feature seems to be slope500-1500_sma3 and it's placed into the final list with selected features for that use case presented in Table 6.

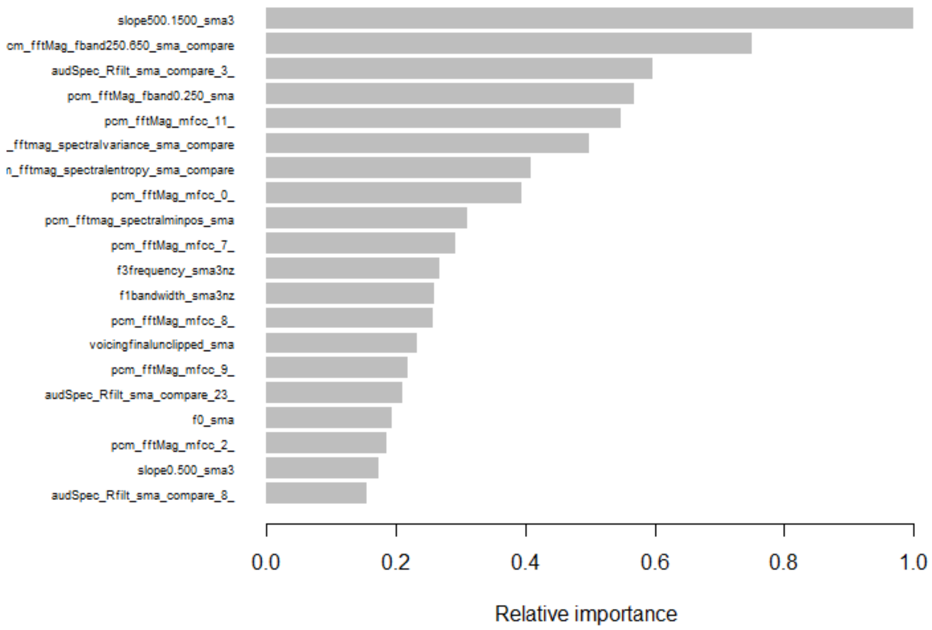

Figure 8. Figure shows relative importance for first iteration of UFS method for mania state with female users.

Table 6. Selected features for mania states with female users.

Slope500-1500_sma3
pcm_fftMag_fband250.650_sma_compare
pcm_fftMag_fband0.650_sma
pcm_fftMag_fband0.250_sma
pcm_fftmag_spectralentropy_sma_compare

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

4.2.5. Recommended features for all patients in the mania state for male patients

In Figure 9, we are presenting relative importance from the first iteration of UFS methods. The most important feature seems to be pcm_fftMag_mfcc_6_ and it's placed into the final list with selected features for that use case presented in Table 7.

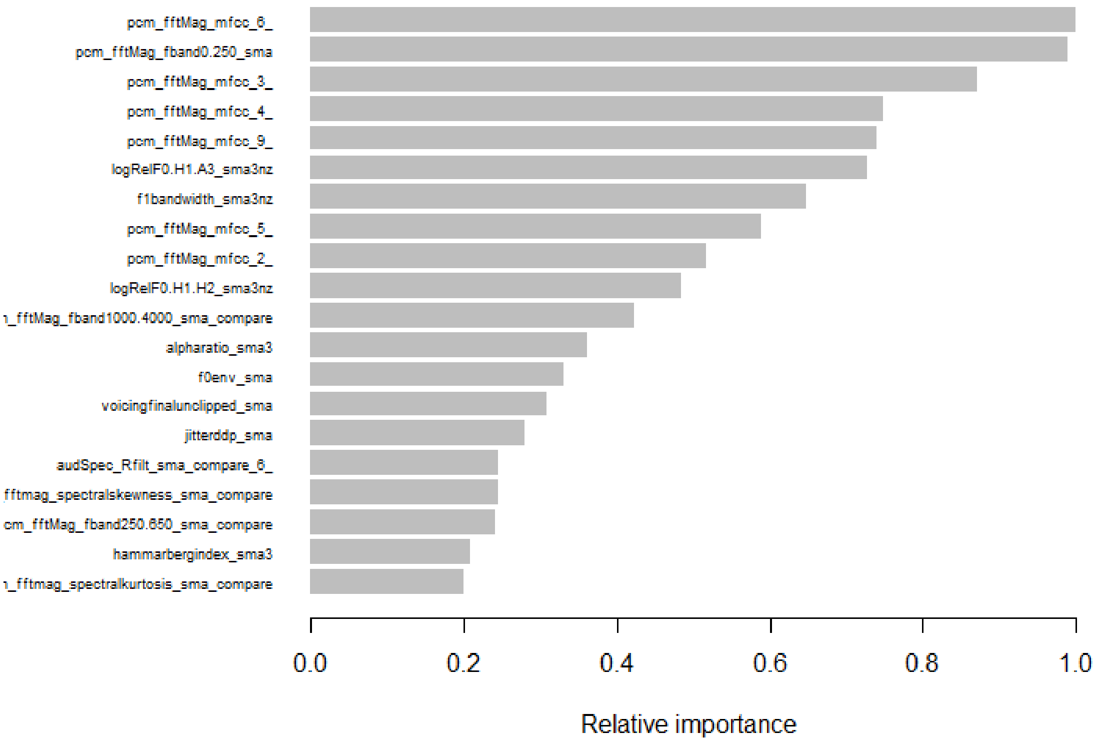

Figure 9. Figure shows relative importance for first iteration of UFS method for mania state with male users.

Table7. Selected features for all BD states.

pcm_fftMag_mfcc_6_
pcm_fftMag_mfcc_3_
pcm_fftMag_mfcc_4_
pcm_fftMag_mfcc_9_
pcm_fftMag_fband0.250_sma

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

4.2.6. Recommended features for all patients in the depression state for female patients

In Figure 10, we are presenting relative importance from the first iteration of UFS methods. The most important feature seems to be pcm_LOGenergy_sma and it's placed into the final list with selected features for that use case presented in Table 8.

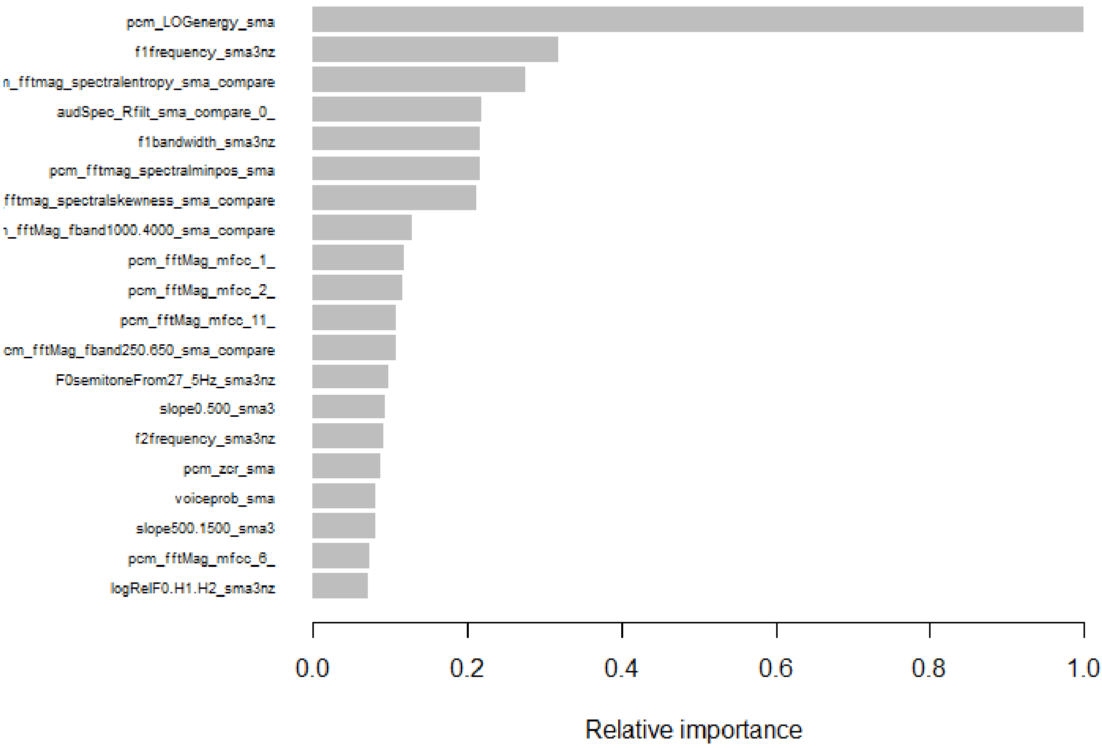

Figure 10. Figure shows relative importance for first iteration of UFS method for depression state with female users.

Table 8. Selected features for all BD states.

pcm_LOGenergy_sma
loghnr_sma
pcm_fftMag_mfcc_12_
f1frequency_sma3nz
f2frequency_sma3nz

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

4.2.7. Recommended features for all patients in the depression state for male patients

In Figure 11 we are presenting relative importance from the first iteration of UFS methods. The most important feature seems to be pcm_fftMag_mfcc_7_ and f1bandwidth_sma3nz.
Both features are placed into the final list with selected features for that use case presented in Table 9.

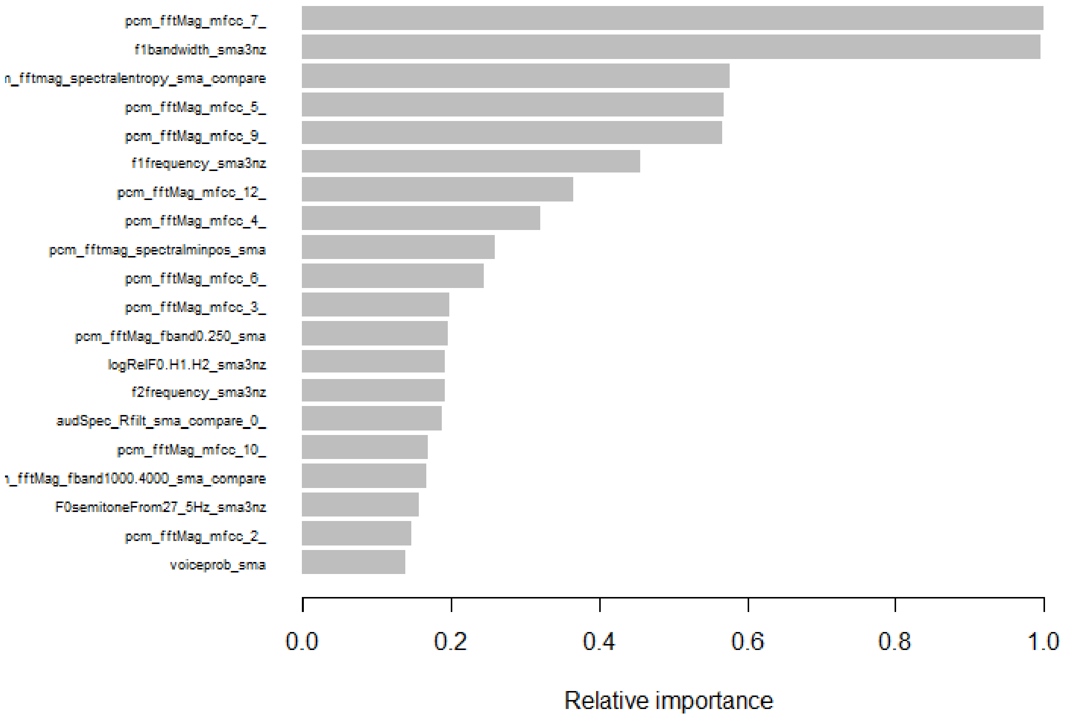

Figure 11. Figure shows relative importance for first iteration of UFS method for depression state with male users.

Table9. Selected features for depression state with male users

pcm_fftMag_mfcc_7_
f1bandwidth_sma3nz
f1frequency_sma3nz
audSpec_Rfilt_sma_compare_25_
pcm_fftMag_mfcc_5_

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

We can observe that all of the use cases received a different list. However, some parameters like 'f1bandwidth_sma3nz',' f1frequency_sma3nz' or 'pcm_fftMag_mfcc_**n**' occuring in most of use cases.

Particularly in relation to literature, we observe mostly divisions based on the type of BD episodes instead of distinction on specific BD phase focus on selected gender. General selection of important features sometime could be insufficient. Previous divisions (included in Report 2.4) did not provide much domain knowledge (division by data aggregation method) or were too detailed (division per each patient).

For the reasons given above and trying to keep the balance between excessive details on use cases and medical recommendation, we recommend the division for selecting important features for mania and depression separately regardless of gender - included in the Table 10 and 11.

4.3. Locomotor use case

Next results are coming from dataset received from Bulgarian hospital [6]. We were using a similar method for selecting the most informative features, however in that case we could assign confidence factor due to the fact that all of the observations were equally certain. Figure 12 presents an order of 10 the most important feature with its variable importance. As we can see – the feature describing age is the most informative. From those features we are recommending the first 5 features as the most informative and results are presented in Table 12.

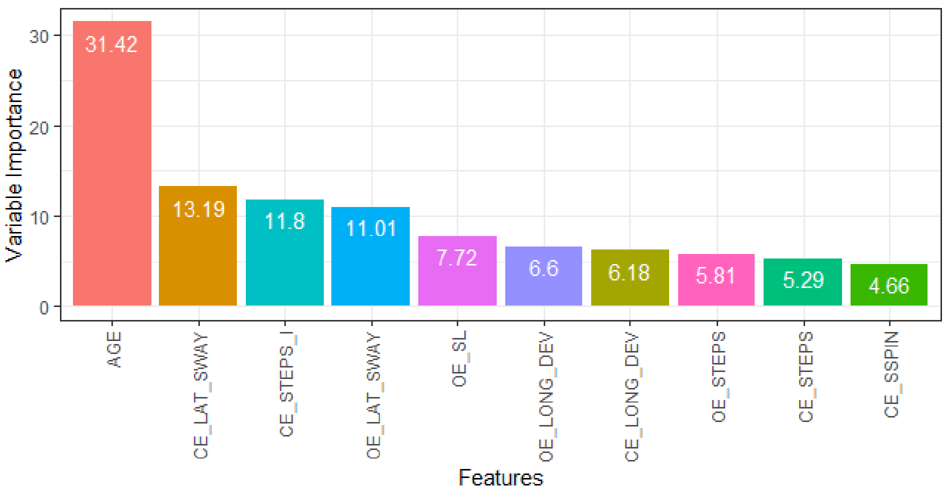

Figure12. Variable importance for locomotr data

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

5. Recommended features

In this Section, we summarize the main findings about the recommended approach to derive the most relevant features describing patients in bipolar disorder.

Here are the main findings formulated as guidelines in steps:

1. The key point for any analyses is to prepare an appropriate dataset according to established assumptions. During the in-depth analysis of the methods in this project, we wanted to keep the best trade-off between the level of detail of the target group and the universality of the method. Therefore, in the tested example, a feature selection method was prepared according to the division of leading states (mania-euthymia vs. depression-euthymia) with an additional division for each gender: women-mania, men-mania, women-depression, men-depression. We need to define classes and rules for extrapolation of labels.
2. Having prepared the final dataset, we need to consider the different levels of confidence assigned to the labels coming from psychiatrics visits.
3. In the next step, we perform uncertainty-aware feature selection (e.g., using method called 'weighted_FS' available in bipolar-preprocessing package whereas a function parameter we are passing dataset with acoustic features, labels and confidence).
4. Implemented method returns best subset of features, that best defines the declared classes. Steps 1-4 can be repeated with different assumptions (e.g., about the number of classes).
5. We highly recommend comparing obtained results with the state of the art.

We now provide lists of the recommended acoustic features for the considered use cases in Tables 10-12.

Table 10. Recommended features indicating maniac overactivity

MANIAC approach
f1bandwidth_sma3nz
pcm_fftmag_spectralvariance_sma_compare
pcm_fftMag_mfcc_5_
slope500.1500_sma3
pcm_fftMag_mfcc_4_

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

Table 11. Recommended features indicating depressive symptoms

DEPPRESSION approach
f1frequency_sma3nz
f2frequency_sma3nz
f3frequency_sma3nz
slope0.500_sma3
pcm_fftMag_fband1000.4000_sma_compare

Table 12. Recommended features recommended during using locomotor sensors

LOCOMOTR features	Description
AGE	patients age
CE_LAT_SWAY:	closed-eyes lateral sway
CE_STEPS_I	closed-eyes number of steps during test
OE_LAT_SWAY	open-eyes lateral sway
OE_SL	open-eyes step length

.

For future work, we consider to test another classifiers or algorithm for feature selection, however, it's should be appropriate to presented use case. Data sharing could largely improve the validity of the results for this particular context of mental health monitoring. However, this is not easily possible due to the sensitive nature of speech collected from personal devices. In the long term, the proposed uncertainty-aware approach could be applied for overcoming these problems by, e.g., sharing parameters estimated for the models.

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

References

[1] G. Casalino, M. Dominiak, F. Galetta, and K. Kaczmarek-Majer, “Incremental Semi-Supervised Fuzzy C-Means for Bipolar Disorder Episode Prediction,” in 2020 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS). Bari, Italy: IEEE, May 2020, pp. 1–8.

[2] G. Casalino, G. Castellano, F. Galetta, and K. Kaczmarek-Majer, “Dynamic Incremental Semi-supervised Fuzzy Clustering for Bipolar Disorder Episode Prediction,” in Discovery Science, A. Appice, G. Tsoumakas, Y. Manolopoulos, and S. Matwin, Eds. Cham: Springer International Publishing, 2020, vol. 12323, pp. 79–93.

[3] K. Kmita, G. Casalino, G. Castellano, O. Hryniewicz, and K. Kaczmarek-Majer, “Confidence path regularization for handling label uncertainty in semi-supervised learning: Use case in bipolar disorder monitoring,” p. 8.

[4] M. Faurholt-Jepsen, J. Busk, M. Frost, M. Vinberg, E. M. Christensen, O. Winther, J. E. Bardram, and L. V. Kessing, “Voice analysis as an objective state marker in bipolar disorder,” Translational Psychiatry, vol. 6, no. 7, pp. e856–e856, Jul. 2016.

[5] M. Faurholt-Jepsen, J. Busk, M. Frost, J. E. Bardram, M. Vinberg, and L. V. Kessing, “Objective smartphone data as a potential diagnostic marker of bipolar disorder,” Australian & New Zealand Journal of Psychiatry, vol. 53, no. 2, pp. 119–128, 2019, pMID: 30387368.

[6] Haralanov S, Haralanova E, Milushev E, Shkodrova D. Locomotor movement-pattern analysis as an individualized objective and quantitative approach in psychiatry and psychopharmacology: clinical and theoretical implications . Psychiatry and Neurosciences Vol. III, Chapter32. Springer Nature Switzerland AG. 2019. pp.387-416. https://doi.org/10.1007/978-3-319-95360-1_32

[7] Kaminska O., Klonecki T., Kaczmarek-Majer K. ‘Feature selection in bipolar disorder episode classification using cost-constrained methods’, XAI workshops , AIME 2023, Portoroz, Slovenia.

[8] R. W. Picard, Affective computing. MIT press, 2000.

[9] C. E. Williams and K. N. Stevens, “Emotions and speech: Some acoustical correlates,” The journal of the acoustical society of America, vol. 52, no. 4B, pp. 1238–1250, 1972

[10] R. Nakatsu, J. Nicholson, and N. Tosa, “Emotion recognition and its application to computer agents with spontaneous interactive capabilities,” in Proceedings of the seventh ACM international conference on Multimedia (Part 1), 1999, pp. 343–351

[11] Chen, Tianqi & Guestrin, Carlos. (2016). XGBoost: A Scalable Tree Boosting System. 785-794. 10.1145/2939672.2939785.

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.

[12] K. Kaczmarek-Majer, G. Casalino, G. Castellano, M. Dominiak, O. Hryniewicz, O. Kamińska, G. Vessio, N. Díaz-Rodríguez, "PLENARY: Explaining black-box models in natural language through fuzzy linguistic summaries", Information Sciences, Volume 614, 2022, Pages 374-399,

[13] Guyon, I., Weston, J., Barnhill, S., & Vapnik, V., "Gene selection for cancer classification using support vector machines", Mach. Learn., 46(1-3), 389--422, 2002.

[14] Kaminska O., K.Kaczmarek-Majer, O.Hryniewicz, "Acoustic feature selection with fuzzy clustering, self organizing maps and psychiatric assessments", (2020) in Proceedings of the IPMU 2020: Information Processing and Management of Uncertainty in Knowledge-Based Systems, pp 342-355

[15] Kaminska O. Kaczmarek-Majer K. and Hryniewicz O. Impact of clustering unlabeled data on classification: case study in bipolar disorder. FedCSIS, Sofia 2022 https://doi.org/10.15439/2022F210

[16] A. Gr¨unerbl, A. Muaremi, and V. Osmani, "Smartphone-based recognition of states and state changes in bipolar disorder patients," IEEE Journal of Biomedical and Health Informatics, vol. 19(1), 2015

[17] Lundberg, Scott M and Lee, Su-In, "A Unified Approach to Interpreting Model Predictions", Advances in Neural Information Processing Systems, vol 30, 2017

This project has received funding from SMALL GRANT SCHEME Call under grant agreement NOR/SGS/BIPOLAR/0239/2020-00.