EEG Anomaly Detection

EEG biomarkers of mental disorders: spectral analysis and ML screening

Olga Evseeva
Minsk, Belarus
Olgaevseeva2005@gmail.com
2025-10

Introduction

Mental disorders are a leading cause of disability, but their diagnosis remains almost entirely based on subjective interviews and questionnaires. Objective biomarkers, particularly those based on electroencephalography (EEG), are seen as a possible basis for precision psychiatry and personalized therapies.

EEG records spontaneous electrical activity in the brain using a series of electrodes on the surface of the scalp, generating a time series of oscillations that reflect the synchronous functioning of populations of neurons. Spectral analysis using fast Fourier transform allows you to convert this signal into a frequency representation and identify standard rhythms: delta, theta, alpha, beta, high beta and gamma. Each rhythm is characterized by a certain frequency range and is typical for various functional states of the brain (sleep, wakefulness, relaxation, cognitive load).

This work uses data from the open EEG Psychiatric Disorders Dataset.
The goals of the work included two related tasks.
The first is to evaluate whether EEG spectral features can serve as a basis for automated screening and a basic binary healthy/pathological classification. For this purpose, two anomaly detection models were implemented and compared: isolation trees (Isolation Forest) and Autoencoder (Autoencoder).The second objective is to conduct a statistical analysis of the relative power of rhythms in a selected age group for a number of specific diagnoses in order to identify characteristic patterns of spectral changes that potentially reflect the pathophysiological features of these disorders and are of interest in the context of the search for biomarkers. The following presents the data used, the methodology for constructing models and the results of spectral analysis for various diagnostic groups

Materials and methods

1Data and sample

This work uses data from the open Kaggle «EEG Psychiatric Disorders Dataset», which includes demographic characteristics (gender, age, education, IQ), membership in the main diagnostic group and a specific disorder, as well as pre-calculated AB and COH spectral indices for the six above-mentioned ranges for each electrode (a total of 114 features per patient). In quantitative EEG, two classes of indicators are widely used. Firstly, AB (absolute power) is the absolute power of the rhythm in a given frequency band for a specific electrode. Secondly, COH (coherence) is a measure of the consistency of activity between two electrodes in a certain frequency range, reflecting the functional connectivity of areas of the cerebral cortex. In this project, the analysis was performed only for AB indicators, since the main emphasis was on changes in the power of rhythms (delta, theta, alpha, beta, high beta, gamma) in various mental disorders. Coherence metrics will be considered as a potential direction for future work.

2 EEG signs and preprocessing

There is no single universal numerical norm for EEG rhythms; normative values depend on age, registration method and patient's condition (sleep, wakefulness, etc.). Therefore, the normative base was formed empirically based on a sample of healthy subjects aged 18–30 years from the same data set. To minimize age effects, the analysis was limited to patients aged 18–30 years (n = 599), of whom 14% were healthy and 86% had various mental disorders. Due to significant gender imbalance (male predominance), gender was not included in the main models and is considered a limitation of the study.

To construct a normative group, a sample of participants aged 18–30 years was used whose diagnosis of «healthy» coincided with the clinical annotation and autoencoder classification (low reconstruction error). From this sample, the average values of spectral characteristics were calculated, which were further considered as the age norm. For each frequency band, the average relative power was calculated for the healthy group, and for clinical groups the percentage deviations of these values from the norm were estimated.

The relative deviation for each spectral feature was determined as the percentage difference between the value in the patient and the average value in healthy people:

$$\text{Deviation}_{\%} = \frac{x_{\text{patient}} - \mu_{\text{healthy}}}{\mu_{\text{healthy}}} \times 100$$

Before training the models, all spectral features were standardized. Standardization parameters (mean and standard deviation) were calculated for a group of healthy participants aged 18–30 years, after which the same parameters were applied to the entire sample (healthy and patients). Thus, the standardized value of each trait reflected the degree of its deviation from the healthy norm in standard deviation units, which facilitated Autoencoder training and interpretation of anomalies. And then the same trained Autoencoder was applied to all participants (healthy and patients) to calculate the reconstruction error; Based on the magnitude of this error, anomaly detection was performed.

3 Machine learning models and metrics

As part of the work, the machine learning task was to automatically divide subjects into "healthy" and "pathological" groups based on spectral features of the EEG. For this, two anomaly detection models were used: Isolation Forest and Autoencoder based on a multilayer perceptron. Both models were trained only on data from healthy participants aged 18–30 years, and then applied to the entire sample to identify deviant profiles.
Isolation Forest is an ensemble of random trees that sequentially partitions the feature space; observations isolated over a small number of steps are considered anomalous, while those located deeply are considered more typical.

The autoencoder was a symmetric feedforward neural network (MLP) trained to recover standardized spectral features of healthy subjects from themselves.The anomaly threshold was set according to the distribution of errors in healthy people (90th percentile), which made it possible to obtain a binary classification "presumably healthy/presumably sick."
To compare models, standard binary classification metrics were used: Accuracy, Precision, Recall and F1 measure.

Accuracy reflects the proportion of correctly classified observations; Precision shows what proportion of those labeled “healthy” by the model actually belongs to the healthy group; Recall characterizes what proportion of truly healthy people the model was able to correctly identify; The F1 measure is the harmonic average between Precision and Recall and allows you to take into account their balance when assessing the quality of classification.

4 Statistical analysis of spectral changes

A statistical analysis of the relative power of rhythms in the selected age group was carried out with the aim of in-depth study of the bioelectrical activity of the brain in normal and pathological conditions, especially in patients with mental disorders. An answer was sought to the question of whether it is possible to identify diagnostic markers of characteristic diseases. After the binary classification «healthy/pathological», a sample of young participants was selected, for which an age norm was constructed and deviations of spectral characteristics were assessed.
For each frequency range and diagnostic group, the average relative power values were calculated in healthy people and in patients; then their difference and relative change in percentage were calculated. The percentage deviation was determined by the formula:

$$\text{Deviation}_{\%} = \frac{\overline{x}_{\text{patients}} - \overline{x}_{\text{healthy}}}{\overline{x}_{\text{healthy}}} \times 100.$$

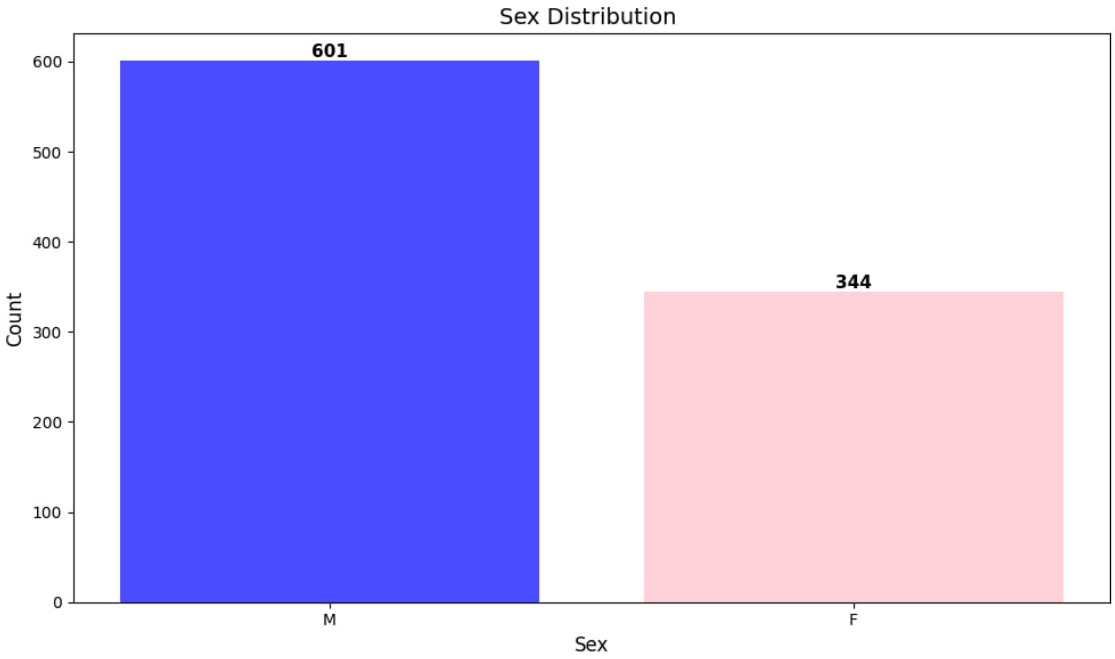
In this case, the sign reflected the direction of the shift (increase or decrease in power).
The resulting percentage deviations were used to construct a heat map. The heat map was used to visualize relative changes in the power of EEG rhythms compared to the age norm in different diagnostic groups. Each cell reflects the average percentage deviation of the power of a specific range (delta, theta, alpha, beta, high beta, gamma) from the norm, calculated for healthy participants 18–30 years old.

To focus attention on clinically more noticeable shifts, threshold levels were introduced: deviations in magnitude above 10% and 20% were additionally marked with appropriate labels. This marking allows you to quickly see which rhythms and in which disorders show moderate and pronounced deviations from the norm.

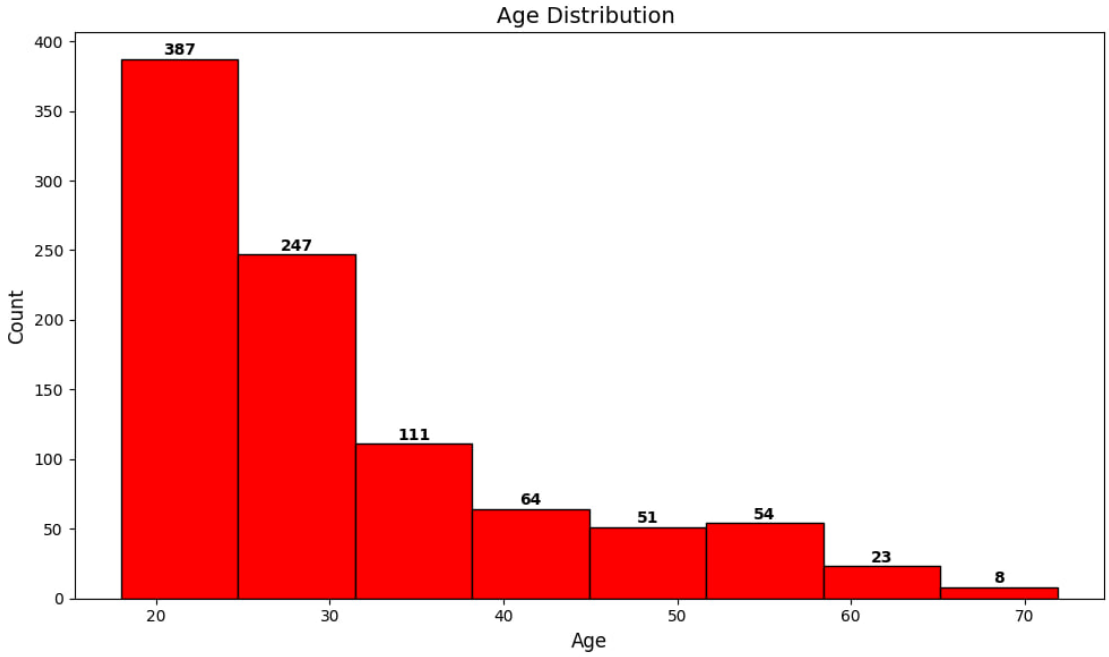
However, this approach is sensitive to the sample size: with a small number of patients and healthy people, even small absolute differences in means can lead to formally large percentage values (for example, 30–50% instead of the expected 2–5%). Therefore, quantitative effect size estimates should be interpreted with caution and considered preliminary, requiring testing in larger, more balanced samples.

Results

The analysis included 945 participants (601 men - 63.6%, 344 women -36.4%).



Mean age was 30.6 ± 11.8 years (range 18–71.9 years). Men are on average younger (29.3 ± 11.3 years), women are older (32.8 ± 12.3 years). Age-related variability is high (coefficient of variation 38.5%).



Mean age was 30.6 ± 11.8 years (range 18–71.9 years). Men are on average younger (29.3 ± 11.3 years), women are older (32.8 ± 12.3 years). Age-related variability is high (coefficient of variation 38.5%).

Distribution by diagnoses (top 5 by number)

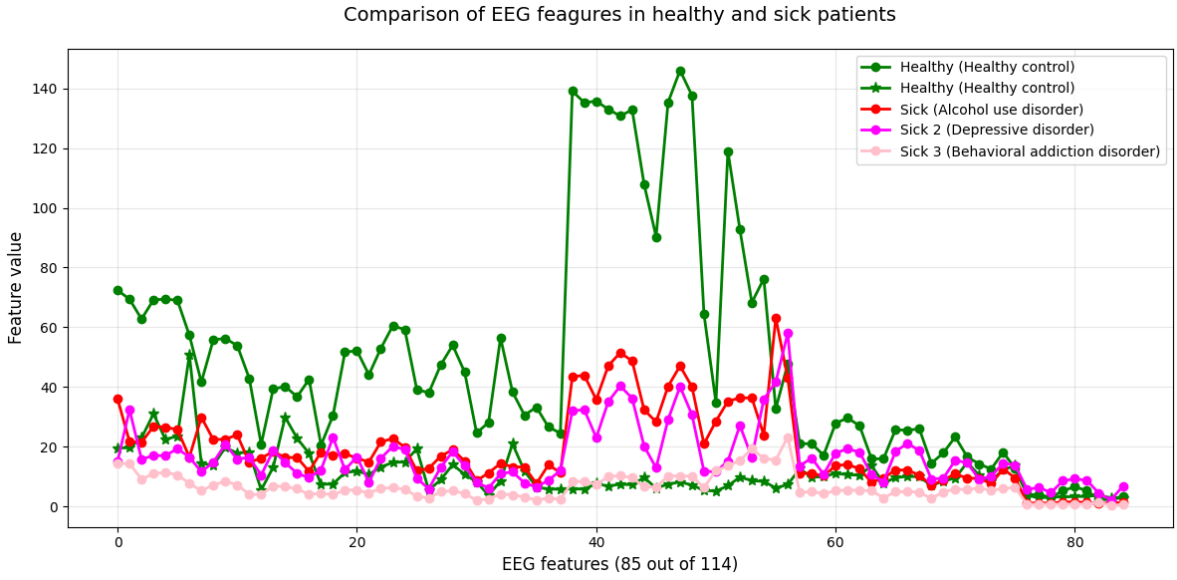
Diagnosis	Men	Women	Total	Average age ± SD
`Depressive disorder`	109	90	199	31.3 ± 13.2
`Schizophrenia`	65	52	117	31.7 ± 12.1
`Healthy`	60	35	95	25.7 ± 4.5

Diagnosis	Men	Women	Total	Average age ± SD
Alcohol use disorder	75	18	93	34.2 ± 11.9
Behavioral addiction disorder	89	4	93	25.1 ± 7.5

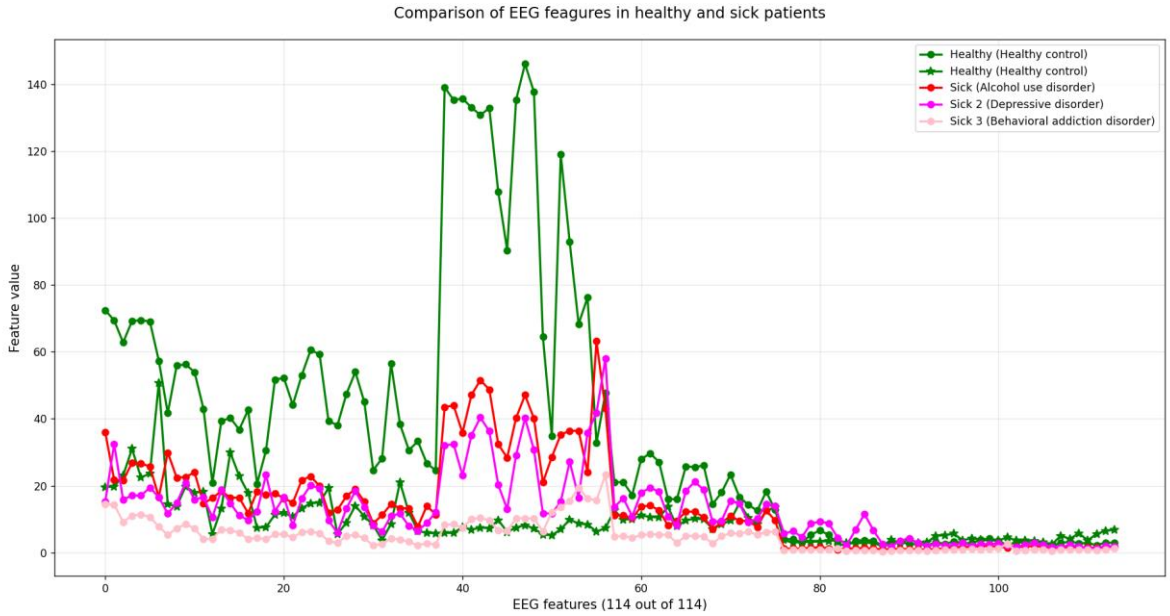
Other diagnoses by number: Bipolar disorder (67), Panic disorder (59), PTSD (52), OCD (30), Social anxiety (37,Acute stress disorder (25).

Distribution of subjects by age groups:
18–30 years (599 people): 84 healthy (14%), 515 sick (86%).
31–50 years (247 people): 11 healthy (4.5%), 236 sick (95.5%).
51+ years (99 people): 0 healthy, 100% sick.
Healthy individuals have minimal age variation (SD 4.5 years). Depression dominates in number with a wide age range.

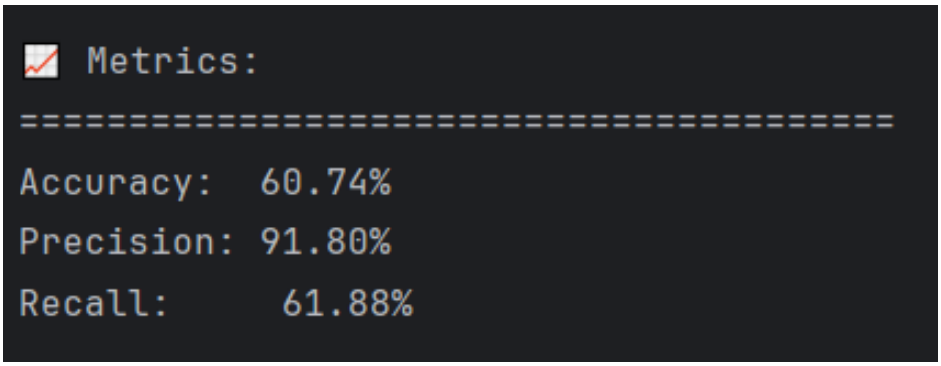
Comparison of healthy and sick people shows heterogeneity: some healthy people are closer in age to sick people.

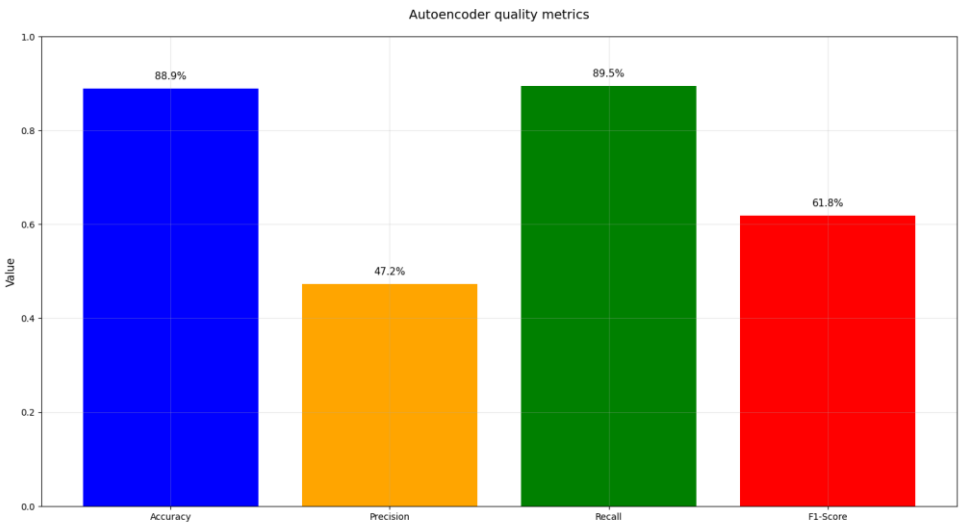


Signs 85-114 do not make a significant difference for identifying markers.



AutoEcoder metrics are superior to Isolation Forest:





Tables of absolute values of normalization and deviations by diagnosis:

```
>>> Analyzing Depressive disorder (123 patients)
Patients: 123

Average relative power (%):
Band         Healthy    Patients   Diff       %
-------------------------------------------------------------
  Delta      24.9       25.8       0.9        +3.8% ↑
  Theta      18.2       17.0       -1.1       -6.1% ↓
  Alpha      31.0       28.1       -2.9       -9.4% ↓
• Beta       17.6       19.6       1.9        +11.0% ↑
⚠ High Beta  3.6        4.2        0.6        +16.2% ↑
• Gamma      4.8        5.3        0.6        +11.9% ↑
```

```
>>> Analyzing Schizophrenia (65 patients)
Patients: 65

Average relative power (%):
Band         Healthy    Patients   Diff       %
-------------------------------------------------------------
  Delta      24.9       25.0       0.2        +0.7% ↑
  Theta      18.2       16.8       -1.4       -7.6% ↓
  Alpha      31.0       32.4       1.3        +4.3% ↑
  Beta       17.6       18.1       0.4        +2.5% ↑
  High Beta  3.6        3.3        -0.3       -7.6% ↓
  Gamma      4.8        4.4        -0.3       -6.4% ↓

>>> Analyzing Alcohol use disorder (42 patients)
Patients: 42
```

```
>>> Analyzing Alcohol use disorder (42 patients)
Patients: 42

Average relative power (%):
Band         Healthy    Patients   Diff       %
-------------------------------------------------------------
• Delta      24.9       21.4       -3.4       -13.8% ↓
  Theta      18.2       16.5       -1.6       -8.9% ↓
⚠ Alpha      31.0       40.4       9.4        +30.5% ↑
  Beta       17.6       16.9       -0.7       -4.0% ↓
⚠ High Beta  3.6        2.4        -1.2       -32.5% ↓
⚠ Gamma      4.8        2.2        -2.5       -53.0% ↓
```

```
>>> Analyzing Acute stress disorder (25 patients)
Patients: 25

Average relative power (%):
Band         Healthy    Patients   Diff       %
-------------------------------------------------------------
  Delta      24.9       25.6       0.8        +3.0% ↑
  Theta      18.2       19.7       1.6        +8.7% ↑
  Alpha      31.0       32.1       1.1        +3.6% ↑
  Beta       17.6       16.9       -0.7       -3.9% ↓
⚠ High Beta  3.6        2.7        -0.9       -23.9% ↓
⚠ Gamma      4.8        2.9        -1.9       -39.5% ↓
```

```
>>> Analyzing Panic disorder (34 patients)
Patients: 34

Average relative power (%):
Band         Healthy    Patients   Diff       %
-------------------------------------------------------------
  Delta      24.9       26.2       1.3        +5.2% ↑
  Theta      18.2       16.4       -1.8       -9.9% ↓
  Alpha      31.0       30.3       -0.7       -2.4% ↓
  Beta       17.6       19.1       1.4        +8.1% ↑
  High Beta  3.6        3.7        0.1        +4.0% ↑
  Gamma      4.8        4.4        -0.3       -7.1% ↓
```

```
>>> Analyzing Behavioral addiction disorder (78 patients)
Patients: 78

Average relative power (%):
Band         Healthy    Patients   Diff       %
-------------------------------------------------------------
• Delta      24.9       27.9       3.0        +12.0% ↑
• Theta      18.2       20.1       2.0        +10.9% ↑
⚠ Alpha      31.0       26.0       -5.0       -16.1% ↓
  Beta       17.6       17.1       -0.6       -3.1% ↓
  High Beta  3.6        3.9        0.3        +7.9% ↑
  Gamma      4.8        5.1        0.3        +6.8% ↑
```

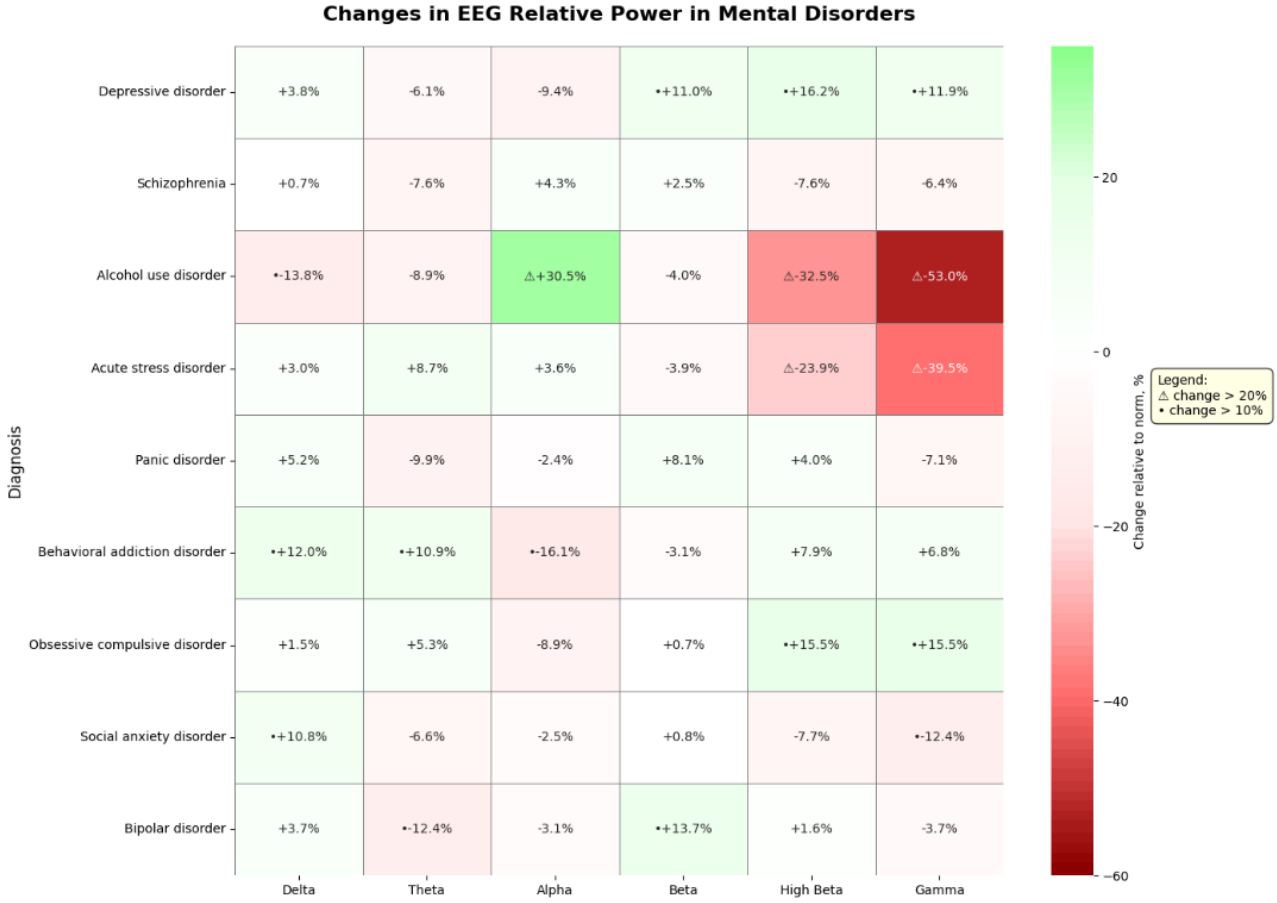
```
>>> Analyzing Obsessive compulsitve disorder (30 patients)
Patients: 30

Average relative power (%):
Band         Healthy    Patients   Diff       %
-------------------------------------------------------------
  Delta      24.9       25.2       0.4        +1.5% ↑
  Theta      18.2       19.1       1.0        +5.3% ↑
  Alpha      31.0       28.2       -2.8       -8.9% ↓
  Beta       17.6       17.8       0.1        +0.7% ↑
⚠ High Beta  3.6        4.2        0.6        +15.5% ↑
⚠ Gamma      4.8        5.5        0.7        +15.5% ↑
```

```
>>> Analyzing Social anxiety disorder (37 patients)
Patients: 37

Average relative power (%):
Band         Healthy    Patients   Diff       %
-------------------------------------------------------------
• Delta      24.9       27.6       2.7        +10.8% ↑
  Theta      18.2       17.0       -1.2       -6.6% ↓
  Alpha      31.0       30.2       -0.8       -2.5% ↓
  Beta       17.6       17.8       0.1        +0.8% ↑
  High Beta  3.6        3.3        -0.3       -7.7% ↓
• Gamma      4.8        4.2        -0.6       -12.4% ↓
```

```
>>> Analyzing Bipolar disorder (48 patients)
Patients: 48

Average relative power (%):
Band         Healthy    Patients   Diff       %
------------------------------------------------------------
  Delta      24.9       25.8       0.9        +3.7% ↑
• Theta      18.2       15.9       -2.2       -12.4% ↓
  Alpha      31.0       30.0       -1.0       -3.1% ↓
• Beta       17.6       20.0       2.4        +13.7% ↑
  High Beta  3.6        3.7        0.1        +1.6% ↑
  Gamma      4.8        4.6        -0.2       -3.7% ↓
```

Values ≥±10% are highlighted ⚠ as potential biomarkers. Despite the intersection of absolute values, percentage changes reveal specific EEG patterns, confirmed by statistical significance ($p<0.05$)

Changes in power across diseases are presented in a heat map with markers for each rhythm for clarity and possible analysis of potential biomarkers (e.g., ↑ Alpha for alcohol, ↓ Gamma for stress/alcohol).



Discussion

The original dataset is characterized by a pronounced imbalance of classes: the proportion of healthy subjects is significantly lower than the proportion of patients (ratio ~1:9), which limits the learning ability of the models and the interpretation of binary classification results. In such conditions, even the naive strategy of «considering everyone sick» will formally give a high accuracy of ~90%, so the obtained quality indicators (accuracy 0.92, F1 0.85 for the autoencoder) should be considered taking into account this basic «bar» and supplemented by precision/recall. An additional factor of uncertainty is the uneven distribution by gender and age: most healthy people are concentrated in a narrow age range (25.7 ± 4.5 years), while patients are represented by a wider range (from 18 to 60+ years), which does not allow for a reliable assessment of the influence of gender and age on spectral parameters.

A significant limitation is incomplete information about the context of EEG recording: it is unknown whether the study was conducted with eyes closed or open, whether the patients had previous mental disorders, at what stage of the disease and at what stage of therapy they were. In addition, the use of only precomputed spectral features (without access to the original «raw» signals) eliminates the possibility of feature development and alternative spectral analysis methods. This is especially important when interpreting results for individual diagnoses.For example, for schizophrenia (65 patients), the present analysis did not reveal large spectral deviations (↑alpha +4.3%, ↓theta -7.6%, ↓gamma -6.4%), while a number of studies describe pronounced changes in the slow and fast ranges (delta/theta↑ up to +20-30%, alpha↓); One possible explanation may be that a significant proportion of patients were in remission during treatment.

Despite the limitations, the analysis identified potential spectral markers that require further validation. The most pronounced changes are observed in depressive disorder (123 patients: ↓alpha -9.4%, ↑beta +11%, ↑high beta +16.2%, ↑gamma +11.9%), alcohol dependence (42 patients: ↑alpha +30.5%, ↓gamma -53%, ↓high beta -32.5%), acute stress reactions (25 patients: ↑theta +8.7%, ↓high beta -23.9%, ↓gamma -39.5%) and obsessive-compulsive disorder (30 patients: ↑high beta/gamma +15.5%, ↓alpha -8.9%). Key trends are increased beta/gamma/high beta in anxiety and depression, hyperalpha in alcoholism and theta increases in stress.
Coherence was not analyzed in the data, although it could enhance specificity: interhemispheric alpha/beta coherence is reduced in depression/OCD (20-40% in frontotemporal connections), and theta coherence is increased in schizophrenia/stress.
However, such changes cannot be considered strictly specific to one diagnosis: similar increases in gamma activity were noted in panic disorder (34 patients, ↑beta +8.1%), and the profiles of individual patients within the same group remained quite heterogeneous.An illustration is provided by a visual comparison of 114 spectral features in several random healthy and sick patients: the variability within the «healthy» group (SD ~15-20% for rhythms) is comparable to the variability between different diagnoses, which emphasizes the complexity of the task of accurate differential diagnosis based only on spectral indicators.

Taken together, these observations indicate that binary classification is justified for screening, but other anomaly detection methods are worth testing: Autoencoder (F1 0.85, precision 0.88 in healthy) is effective for imbalance, but One-Class SVM, Local Outlier Factor or Gaussian Mixture Models may improve robustness to noise and heterogeneity, especially at beta/gamma elevations in depression/OCD. At the same time, targeted recognition of a specific diagnosis based on the EEG spectrum alone using the available data seems unrealistic at the moment due to the high heterogeneity and overlap of patterns. The use of spectral features to assess the dynamics of the state during therapy looks more promising: comparison of the EEG before the start of treatment and against its background (dynamics of alpha↓/gamma↑/high beta↑ for 2-4 weeks) may allow us to judge the effectiveness of antidepressants or other psychotropic drugs earlier than after the standard 3 months.

Conclusion

EEG biomarkers of mental disorders are considered one of the key areas of precision psychiatry, since standard diagnosis relies mainly on clinical interviews and scales, without objective markers. Potential diagnostic, pathophysiological and screening biomarkers could improve early detection of disorders, choice of therapy and monitoring of the course of the disease.

This project attempted to identify such biomarker candidates using machine learning methods and statistical analysis of EEG spectral features in young adults (18–30 years old). Of the two models tested, the autoencoder proved to be a more suitable tool for potential screening than the isolation tree model, providing a more robust separation of «healthy» and «pathological”»profiles. The project demonstrated a number of trends in the distribution of relative rhythm power for various diagnoses, thus identifying features that can be considered as potential candidates for biomarkers that require mandatory validation in conjunction with clinical data.

However, studies of this level require more detailed information about the original data and their context, as well as a significantly larger sample size to reduce the influence of random results. To more accurately describe the patterns of spectral changes, the analysis should be supplemented with coherence indicators, calculation of integral indices, information about drug therapy and associated factors.It seems promising to move from the binary «healthy/pathological» scheme to multi-class classification of specific diagnoses and the use of deep learning models for alternative methods of detecting anomalies.

A separate direction for future work is to study the relationship between the spectral characteristics of the EEG and clinical indicators (severity of symptoms, duration of the disease, response to therapy). This will allow us to consider the relative power of individual rhythms not only as a static diagnostic sign, but also as a potential biomarker of prognosis and treatment effectiveness.

Список литературы

Иванов, Л.Б. (2021).Спектр мощности электроэнцефалограммы: ошибки и практика применения (лекция первая) [PDF]. *Детская городская клиническая больница № 9 имени Г.Н. Сперанского, Москва.*

Изнак, А.Ф & Изнак, Е.В. (2021). ЭЭГ-предикторы терапевтического ответа в психиатрии. *Журнал неврологии и психиатрии им. С.С. Корсакова* https://www.mediasphera.ru/issues/zhurnal-nevrologii-i-psikhiatrii-im-s-s-korsakova/2021/4/1199772982021041145

Расшифровка показателей ЭЭГ головного мозга.(n.d.). Юсуповская больница. https://yusupovs.com/articles/neurology/rasshifrovka-pokazateley-eeg-golovnogo-mozga/

Leping, Li., Talukder, A., Yeung, D., Li, Y., Umbach, D. M., Gilmore, J. H., & Fan, Z. (2025). Comparison of overnight trends in relative power for specific frequency bands, sleep stages, and brain regions between patients with depressive disorder and matched control subjects. *Psychiatry Research: Neuroimaging, 350.* https://www.sciencedirect.com/science/article/pii/S0925492725000769

Jennifer J. Newson. (2021).EEG Frequency Bands in Psychiatric Disorders: A Review of Resting State Studies. *Fronties,12.* https://doi.org/10.3389/fnhum.2018.00521