

## STUDENTŲ ĮTRAUKIMO Į MOKSLINĘ VEIKLĄ SKATININAMOJO KONKURSO TEMA

Temos pavadinimas: **Emocijų atpažinimas iš smegenų elektroencefalografijos duomenų taikant klasikinius mašininio mokymosi algoritmus**

Tikslas: pritaikyti kelius klasikinio mašininio mokymosi algoritmus, tokius kaip atraminių vektorių mašinos ir atsitiktiniai miškai, emocijų klasifikavimui iš žmogaus smegenų elektroencefalografijos duomenų

Trumpas temos vykdymo aprašymas (ne daugiau kaip 2000 ženklų):

Šio darbo metu analizuojami atviros prieigos elektroencefalografijos (EEG) duomenys. EEG signalai yra daugiamačiai laiko eilučių duomenys, kuriuos reikia tinkamai apdoroti ir pavirsti savybėmis, kurios tiktų mašininio mokymosi algoritmų apmokymui. Yra suplanuota palyginti klasikinius ir giliojo mokymosi algoritmų klasifikavimo efektyvumą EEG duomenų rinkiniui, išanalizuoti svarbiausias duomenų savybes emocijų klasifikavimui. Studentė(-as), prisijungianti(-s) prie darbo pagal mokslinę veiklą skatinamąjį konkursą, kartu su kitais grupės nariais bus atsakinga(-s) už klasikinių mašininio mokymosi algoritmų taikymą ir atitinkamų savybių analizę. Studentė(-as) darbo metu išmoks pagrindinius EEG duomenų apdorojimo etapus bei mašininio mokymosi pamatus, bei turės galimybę prisijungti prie mokslinio straipsnio rašymo. Studentas turi būti pasiruošęs dirbti su Python programavimo kalba.

Pagrindiniai darbo etapai:

1. Duomenų analizė ir paruošimas: studentė(-as) susipažins su duomenų rinkiniu, atliks pradinį duomenų valymą bei apžvalgą, pritaikys standartinius EEG signalų apdorojimo metodus: filtravimą, artefaktų šalinimą, segmentavimą į laiko langus.
2. Savybių gavimas: studentė(-as) apskaičiuos įvairias savybes iš EEG signalų, tokias kaip laiko srities statistiniai rodikliai (vidurkis, dispersija) ir dažnių srities požymiai (galios spektro tankis įvairiose bangos dažnių ribose). Bus galima išbandyti ir sudėtingesnių savybių išgavimą, tokių kaip koherencija ir entropijos matas.
3. Klasifikavimo modelių kūrimas ir vertinimas: studentė(-as) apmokys ir palygins kelis klasikinio mašininio mokymosi algoritmus, tokius kaip atraminių vektorių mašinos (SVM) ir atsitiktiniai miškai (Random Forest). Studentė(-as) pritaikys kryžminį validavimą, eksperimentuos su modelių parametrais, bei vertins modelių našumą naudojant įvairius rodiklius.

Temą siūlantis mokslininkas/dėstytojas: Taisija Dėmčėnko

**TOPIC OF A COMPETITION PROMOTING STUDENT ENGAGEMENT IN SCIENTIFIC  
ACTIVITIES**

**Topic: Emotion recognition from brain electroencephalography data using classical machine learning algorithms**

**Goal:** adapt classical machine learning algorithms, such as support vector machines and random forests, for emotion classification from human brain electroencephalography data

**Short description (max. 2000 characters):**

This work analyzes an open-access electroencephalography (EEG) dataset. EEG signals are multidimensional time series data that need to be properly processed and converted into features suitable for training machine learning algorithms. The plan is to compare the classification efficiency of classical and deep learning algorithms for the EEG dataset and analyze most important data features for emotion classification. The student joining the work under the competition will be responsible, together with other group members, for the application of classical machine learning algorithms and the analysis of relevant features. During the work, the student will learn about basic stages of EEG data processing and the fundamentals of machine learning, and will have the opportunity to participate in writing a scientific paper. The student must be prepared to work with the Python programming language.

**Main steps:**

1. Data analysis and preparation: the student will familiarize themselves with the dataset, perform initial data cleaning and exploration, apply standard EEG signal processing methods: filtering, artifact removal, segmentation into time windows.
2. Feature extraction: the student will calculate various features from EEG signals, such as statistical indicators in the time domain (mean, dispersion) and characteristics in the frequency domain (power spectrum density in various wave frequency ranges). More complex feature extraction, such as coherence and entropy measures, may also be explored.
3. Classification model development and evaluation: the student will train and compare several classical machine learning algorithms, such as support vector machines (SVM) and random forests. The student will apply cross-validation, experiment with model parameters, and evaluate model performance using various metrics.

**Supervisor researcher/lecturer:** Taisija Dëmčenko