

## ПРИГНІЧЕННЯ КОГНІТИВНИХ УПЕРЕДЖЕНЬ ПІД ЧАС ІДЕНТИФІКАЦІЇ ОСОБИ ТА ЇЇ НАМІРУ В НЕЙРО-КОМП'ЮТЕРНИХ СИСТЕМАХ

Дослідження присвячено проблемі безпеки інформаційних систем у випадку розпізнавання особи та її намірів за допомогою електроенцефалограми. Метод передбачено для неінвазивного нейрокомп'ютерного інтерфейсу, може застосовуватись при багаторівневих дослідженнях (емоційний інтелект, розвиток пам'яті, моніторинг самопочуття, оптимізація прийняття рішень), у чому полягає його перевага. Універсальність нейро-комп'ютерного інтерфейсу дозволяє використовувати його для кількох цілей одночасно, поєднуючи різні методи аналізу одного набору даних (енцефалограми, часу відповіді та адекватності). Запропоновано прикладний комбінований підхід до оцінки якості на основі проектних показників за результатами попередніх досліджень з урахуванням специфіки завдань і ролей користувачів, методів обробки даних, об'єктивності обраних моделей. Експерименти проводилися у змодельованих умовах у два етапи по 3 місяці за результатами тренування нейропластичності та вирішення завдань отримання доступу за наявності прав та в умовах непередбаченого запиту. Враховано індивідуальний стиль виконання завдань, а для прогнозування наміру застосовується стратегія розпізнавання емоцій за моделлю потоку Чиксентмігаї. Результати відповідають сучасним стандартам нейроетики та нейроправа. Когнітивні упередження можуть негативно вплинути на завдання ідентифікації людини та її намірів, для цього пропонуються методи пом'якшення. Зручність використання моделі полягає в простій інтерпретації зібраних даних. Для оцінки використано фронтальна альфа-асиметрія, потенціал пов'язаний з подією, відношення потужності смуги альфа-бета ритмів синхронізації. Результати переконують, що очікуване впровадження може бути ефективним. Робота враховує багатий досвід попередніх досліджень з ідентифікації особи та її намірів на основі аналізу електроенцефалограми. У цьому дослідженні враховано непостійність ЕЕГ образу («відбитку мозку») та акцентовано увагу на важливості його оновлення, проведено експеримент в ході комплексного дослідження. Результатом є обґрунтування вимог до частоти і процедури оновлення «відбитків мозку», запропоновані показники якості розпізнавання, пропозиції до ранжування рівня нейрограмотності на основі тестування типової діяльності. Для вивчення наміру особи запропоновано дослідження фонових когнітивних станів із врахуванням їх тривалості та інтенсивності типових зразків ритмів синхронізації, на основі яких проводилося машинне навчання моделі пригнічення когнітивних упереджень. Дані порівнювалися із моментами втрати контролю, було враховано потребу переосмислення та оновлення мереж синхронізації мозку. Ці результати наразі є базовими і в подальших експериментах вимагають окремої уваги, їх ґрунтовне дослідження буде важливим для покращення якості будь-яких досліджень застосування нейро-комп'ютерних інтерфейсів. Оцінено ризик небажаних когнітивних станів і появи когнітивних упереджень, передбачено вимоги до складання тестів з метою запобігання таким ситуаціям. Враховано вплив досвіду та рівня тренуваності нейрокомп'ютерної грамотності, індивідуальної придатності до роботи в нейро-комп'ютерних середовищах (зокрема потужність нейросигналу та володіння контролем ментальних зусиль, когнітивних властивостей і емоційних фонів).

**Ключові слова:** нейробезпека, неінвазивний нейрокомп'ютерний інтерфейс, ЕЕГ-ідентифікація наміру, пом'якшення когнітивних упереджень.

**Вступ.** Задача ідентифікації особи та її намірів стала типовою для безпеки систем із нейро-комп'ютерним інтерфейсом. Сьогодні зокрема виділяють поняття «відбиток мозку» (типовий ЕЕГ образ особи на основі середніх показників пластичності таких як вольтажі обраних ділянок та інтенсивність утримання цих значень протягом певного часу) та «відбиток розуму» (своєрідний стиль поведінки під час діяльності, визначається за співвідношенням часу реакції та точності відповідей із показниками когнітивних метрик

таких як зосередженість, заангажованість, стрес тощо). Залежно від типу пропонується структурний або функціональний підхід до проектування тестів, а також відповідні алгоритми збирання та обробки даних [1].

Із поширенням пристроїв із неінвазивним нейро-комп'ютерним інтерфейсом, особливо користувацького рівня, впровадженням моніторингу самопочуття та оптимізації прийняття рішень у бізнесі (США) і дослідження освітніх процесів (Китай), легітимізацією таких технологій для слідчої експертизи у досудовій практиці (Франція) стає помітним багатогалузевий вплив нейро-комп'ютерних інтерфейсів та постає логічне запитання про можливі подальші напрямки їхнього розвитку, методи використання і сфери впровадження.

Когнітивні упередження можуть значно погіршити результати розпізнавання особи. Тому пом'якшення когнітивних упереджень (mitigation of cognitive biases) є актуальною задачею в розробці протоколів безпеки інформаційних систем, а системи оснащені нейро-комп'ютерним модулем мають помітну перевагу у вирішенні цієї проблеми та застосування нової технології. Нейро-комп'ютерний інтерфейс, оснащений модулем виявлення небажаних когнітивних шумів, а далі - пригнічення когнітивних упереджень (така технологія може бути реалізована у моделі фільтра енцефалограми), дозволить оптимізувати процес ідентифікації, мінімізувавши небажані втрати часу та кількість помилок.

**Аналіз останніх досліджень.** У дослідженні [2] вчені продемонстрували, що електроенцефалографічні (ЕЕГ) сигнали можуть служити унікальними біометричними ідентифікаторами. Виявлено, що певні компоненти ЕЕГ, які ідентифікують особу, є стабільними в різних задачах та протягом часу. ЕЕГ сигнали дозволяють розрізнити навіть монозиготних близнюків, особливо, якщо використовуються дані з пізніше розвинених областей мозку.

У дослідженні [3] висвітлено метод порівняння амплітуд між півкулями мозку (Inter-Hemispheric Amplitude Ratio, IHAR), щоб розрізнити осіб за допомогою ЕЕГ на основі ідентифікації унікальних даних. Завдяки запропонованій функції можна досягти високої точності з малою кількістю сенсорів у простих завданнях. Перевагою підходу є можливість застосування для портативних пристроїв із вбудованими малопотужними мікропроцесорами. Також такий підхід забезпечує більшу зручність при навчанні системи, оскільки не потребує великої кількості навчальних зразків. У запропонованому підході найбільша точність була досягнута для двох фронтальних пар електродів — FC5-FC6 і F7-F8. Було визначено, що для збору даних ЕЕГ найкращим етапом відповідно до результатів була релаксація (точність 90%). Найбільш придатними алгоритмами машинного навчання для класифікації є QDA, SVM, kNN. Крім того, селективне розташування електродів AF3, AF4, F7 і F8 можна використовувати для розробки більш жорсткої апаратної конструкції з високою продуктивністю [1, 2, 4].

Когнітивні упередження, враховуючи специфічність їхнього впливу, можуть суттєво перешкоджати адекватній ідентифікації особи та її намірів, а також призводити до небажаних дій під час прийняття рішень, що в свою чергу може погіршити адекватність протоколу безпеки, робочого порядку і добробуту, самопочуття працівників. З цією метою пропонується використати модуль виявлення ризиків когнітивної упередженості та методи їх пом'якшення у системах з жорсткою умовою ідентифікації (обмежена кількість спроб і час для відповіді). Пропоную виділити три рівні нейрограмотності для цієї задачі: початковий рівень – відсутність досвіду, нетренованість стану контролю, рівень навчання – успішне оволодіння навичкою контролю зосередженості, придатність до використання ментальних команд, та рівень впевненої нейрограмотної особи, що визначається володінням базовим набором ментальних команд і добре розвиненим контролем власного ментального стану. Рівень визначається за типовим тестуванням та підвищується в ході навчання і тренувань [5-7].

Для оцінки ментального стану використовуються типові показники такі як смуги частот, фронтальна асиметрія, відношення потужності різних смуг та потенціал, пов'язаний із подією [2, 6, 8].

Перспективним напрямом розвитку технології є метод врахування когнітивних упереджень, за яким оцінені ризики або реальні прояви когнітивних упереджень фільтруються як небажаний артефакт при задачах ідентифікації особи та намірів.

**Постановка задачі.** Проблема ідентифікації особи за власним нейрообразом передбачає визначення права доступу користувача. Дану задачу можна сформулювати так: «Чи відповідає ЕЕГ метрика особи що здійснює дії нормальним зразкам?»

З цією метою протягом робочого часу збираються зразки із попередньо заданим вибором та інтервалами. Достатньо чотирьох-п'яти контрольних п'ятнадцятихвилинних тестів протягом нормального робочого часу (4-8 годин). Можливі комбінації таких показників як зосередженість, когнітивний стрес, частота кліпання очима, середній час реакції та точність відповідей (причому для подальшого порівняння амплітуд враховуються інтенсивності активації нейронних осередків правої та лівої півкулі мозку) [4, 9, 10].



Рисунок 1 – Схема розташування електродів

Не менш важливою є проблема визначення наміру, яка досліджується в контексті розпізнавання емоцій та ментальних станів. Тут суттєвим є врахування відбитку розуму, який може певною мірою передати індивідуальність користувача.

Ця задача може бути сформульована у такому вигляді: «Чи припустимі наміри особи, яка 1) здійснює дію 2) отримує доступ або права?».

Задача може бути вирішена в межах перегляду емоційного стану особи в часовому околі виконання завдання або у момент здійснення дій.

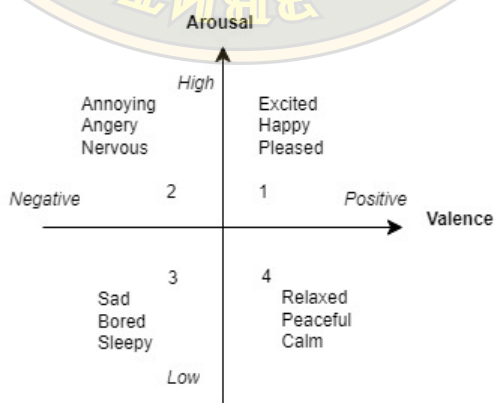


Рисунок 2 – Схема визначення наміру відповідно до емоційного стану

Можливою формою захисту системи є застосування блокування доступу або повторний запит ідентифікації. Технологія ментального паролю може бути запропонована як форма

ключа доступу. Для цього користувачу заздалегідь пропонують обрати варіант ментального завдання (зображення, 3-вимірний об'єкт, графічний ключ, математичний приклад, звукозапис, слово) та фіксують реакцію-відповідь (ментальне зусилля – зосереджена активація нейронних осередків, викликана означеним стимулом із визначеною унікальною відповіддю) (Рис. 3).

**Виклад основного матеріалу.** Дослідження проводилось в імітованих умовах проектної діяльності. До уваги бралися показники ментального стану під час отримання доступу до захищеної інформації. В якості захисту пропонувалися паролі та прості ментальні образи. З метою підвищення якості захисту було використано алгоритм розпізнавання індивідуального нейрообразу особи на основі середніх показників типових артефактів: частота кліпання очима, релаксація, тривалість переосмислення, частота сонливості, а також типових робочих показників з точністю до визначеної похибки: зосередження, середні частоти стану бадьорості та свідомості, тривалість та інтенсивність ментального зусилля, тривалість та точність реакції тощо. Для підвищення якості результатів використано підхід порівняння середніх амплітуд активності правої та лівої півкулі мозку. Кінцеве оцінювання здійснено на основі суб'єктивного тестування із декларацією правдивості даних та експертних оцінок типових показників [3, 5, 10, 11].

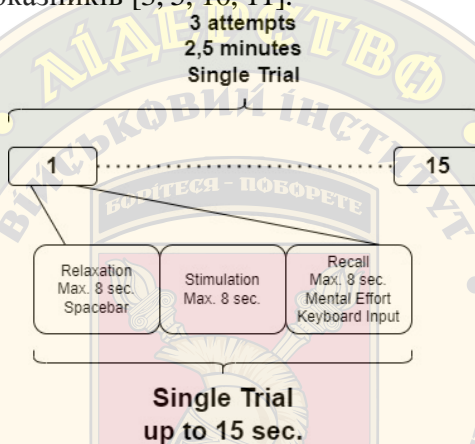


Рисунок 3 – Схема експерименту для збирання зразків

Середній час для одного етапу поліпшення нейропластичності становить 3 місяці. Ця стратегія навчання включає 4 заняття на тиждень по 30 хвилин. Два заняття з навчання грамотності НКІ включають 8 вправ, які базуються на завданнях, що розвивають пам'ять, сприйняття та контроль уваги. Проводяться дві професійні сесії в спеціальному середовищі для закріплення нових навичок у професійному розв'язанні проблем і стані трудової діяльності. Паралельні завдання для вдосконалення свідомого контролю включають розвиток психіки для підвищення емоційного інтелекту та редукції когнітивних упереджень (Табл. 1), [5, 12, 13].

Під час основної частини навчання проводилась якісна оцінка прийняття рішень у 30-хвилинних циклах професійної діяльності на різних етапах проекту, а також у непередбачуваних умовах.

Таблиця 1.

Забезпечення дослідження

Робоче середовище	Microsoft Project, Visual Studio, Jupiter Notebook.
Роб. Станція	Dell Latitude 3570
Дослід. середовище	Emotiv LABS (навчання, тестування), EmotivPRO (збір даних), PsychoPy (додаткові тести).
Методи	ЕЕГ, ІСА, SVM, P300, ERP
Вимоги до суб'єктів	II рівень нейрограмотності, досвід НКІ від 3 міс., хороша нейропластичність, всі темпераменти

При виборі електродів враховувалася оптимальність фіксації сигналів дієвих центрів (Табл. 2). Найвне устаткування дозволяє врахувати описані в методах вимоги до дослідження в межах запропонованих завдань.

Таблиця 2.

Когнітивні функції у відповідності до позицій електродів (система 10-20)	
Позиція	Функції
AF3	Аналіз, логіка, лічба
AF4	Прийняття рішень, увага
T7	Мовлення, слух
T8	Пам'ять
Pz	Сенсорика, просторова свідомість, зосередженість

Проблема когнітивної упередженості має наступний вираз: при підготовці до ідентифікації слід визначити чи високий ризик когнітивної упередженості в найближчий час, а під час ідентифікації – чи міг бути небажаний ментальний стан або рішення викликаний когнітивним упередженням. З метою оптимізації роботи системи пропоную вводити модуль ідентифікації когнітивної упередженості паралельно із модулем ідентифікації.

Додатково систему можна зміцнити інструментом транскарніальної магнито-резонансної томографії, що може пом'якшувати стан упередженості або активувати зосередженість чи необхідний нейтральний стан для осіб нижчого рівня нейрограмотності.

Дослідження EEG здійснюється за методикою вивчення потенціалу (активації) пов'язаного з подією (стимулом, ментальним зусиллям або станом).

Інтенсивність потенціалу активації оцінюється за діапазоном потужності EEG. Спектральний аналіз сигналу здійснюється за алгоритмом швидкого перетворення Фур'є, що дозволяє виділити ритми синхронізації дієвих нейронних осередків (Табл. 3). Таким чином можна отримати образ кожного слова (*когнеми* - за *Нейхгольтом А.*) перформансу (ментального зусилля, реакції, рішення) та оцінити *фон* ментального стану (емоції, зосередженість, релаксація) [14]. Часово-якісне співвідношення цих метрик та реакції відносно прийнятної точності дає уявлення про коректність вирішення задачі розпізнавання.

Таблиця 3.

Відповідність ритмів синхронізації когнітивним станам

EEG Ритм	Когнітивний процес/стан
Альфа (8-13 Гц)	зосередженість, свідомість, увага
Бета (13-30 Гц)	активне мислення, моторика, когнітивні завдання
Тета (4-7 Гц)	сонливість, запам'ятовування, творчість, фантазія
Дельта (0,5-4 Гц)	глибокий сон, відновлювальні процеси, формування спогадів, переосмислення

Вимога повторної ідентифікації висувається у випадку 1) невідповідності образу когнеми доступу нормальному, оновленому протягом останніх 48 годин значенню або 2) невідповідності даних когнітивних метрик типовим, зафіксованим протягом останніх 30-180 хв показникам.

Намір оцінюється як потенційно небезпечний, у випадку 1) відповідності когнітивних метрик небажаний локалізації (рис. 2, квадрант 2) та 2) при здійсненні дій непередбачених типовими інструкціями. Для виявлення цього стану використовуються методи аналізу

незалежних компонент та порівняння середніх амплітуд сенсорів різних півкуль (AF3 із AF4 та T7 із T8).

Результати експериментів та робочих досліджень оцінюються експертами, до уваги беруться також анкети суб'єктивного оцінювання. Оцінювання здійснюється за методом обґрунтування частки успішності дотримання показника відносно ідеального значення у відсотках. Кінцеві результати представлено у формі таблиці кореляції чинників [8, 15, 16].

*Когнітивні метрики перформансу, які порівнюються при ідентифікації наміру особи.* Когнітивний стрес часто є наслідком завдань, які вимагають постійної уваги та зосередженості, що призводить до розумової втоми. Стрес може спровокувати такі емоційні реакції, як тривога, розчарування або дратівливість, що впливає на виконання завдання. Стрес активує вісь гіпоталамус-гіпофіз-наднирники (НРА), що призводить до вивільнення кортизолу, гормону стресу. Стрес відображається на EEG-паттернах, таких як підвищення активності бета-хвиль (вказує на настороженість і тривогу) і зниження активності альфа-хвиль (вказує на розслаблення). Саме когнітивний стрес відповідає непевним намірам особи. Його надмірність може сприйматися як неприпустимий когнітивний стан та сприйматися системою як загроза. В такому випадку варто передбачити заходи захисту для протоколу безпеки такі як додаткова ідентифікація особи, уточнюючі запитання, перевірка робочого місця та звіт.

Когнітивний стрес оцінюється за спектральною густиною EEG сигналів, окремо для різних частот. Зростання густини різних частот зазвичай відповідає вищому когнітивному стресу [15].

$$P(f) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n) e^{-j\pi f n / N} \right|^2 \quad (1)$$

де  $P(f)$  – потужність частоти ( $f$ ), ( $N$ ) кількість зразків,  $x(n)$  – значення EEG у мілівольтах. Частота дискретизації сигналу 0.5 Гц.

Потужність частоти сигналу (певного ритму EEG):

$$BP = \sum_{f_1}^{f_2} P(f), \quad (2)$$

де  $f_1, f_2$  – крайні значення частоти смуги в Герцах.

Для оцінки когнітивного стресу використовують бета і гамма-частоти:

$$PSD_{\beta} = \frac{1}{N} \sum_{f=13}^{30} |X(f)|^2, \quad (3)$$

$$PSD_{\gamma} = \frac{1}{N} \sum_{f=30}^{50} |X(f)|^2. \quad (4)$$

Альтернативно можна оцінювати рівень когнітивного стресу за допомогою фронтальної альфа-асиметрії (Frontal Alfa Asymmetry). Цей показник базується на різниці альфа-потужності між лівою та правою лобовими частками. [16] Фронтальна асиметрія альфа-смуги використовується також для оцінки емоційного стану:

$$FFA = \log(\text{Power}_{\alpha, \text{right}}) - \log(\text{Power}_{\alpha, \text{left}}). \quad (5)$$

Потенціали, пов'язані з подіями (Event Related Potential, ERP) – це типові електрофізіологічні відповіді мозку у визначеному відчутті, когнітивному чи емоційному стані або моторного зусилля, що виникли під впливом певної спонуки, та були зафіксовані у визначених часово-потужнісних околах досліджуваної активності. Компонент P300, що виникає приблизно через 300 мс після початку стимулу, зазвичай використовується для оцінки уваги:

$$P300_{ampl.} = \max(ERP_{300ms}) - baseline. \quad (6)$$

Ці показники можна використати для ідентифікації особи за допомогою ментального ключа доступу для оцінки відповіді на стимул. У такому випадку оцінюються ментальні зусилля. Якщо відповідь типова – доступ надається, якщо ні – виконується протокол безпеки. Для ідентифікації особи можна використати дані навчання у типових професійних умовах та задачах, під час перерв та релаксації або отримання доступу [17],

$$\theta/\beta \text{ ratio} = \frac{Power_{\theta}}{Power_{\beta}}. \quad (7)$$

Спектральна густина тета-смуги дозволяє оцінити адекватність контролю. Саме ця частота є індикатором когнітивних упереджень, що можуть виникнути в процесі ідентифікації особи (фреймінг, парейдолія, прив'язка). Активація тета-частоти також може свідчити про розсіяність уваги [18],

$$PSD_{\theta} = \frac{1}{N} \sum_{f=4}^8 |X(f)|^2. \quad (8)$$

Псевдокод прикладу розпізнавання наміру представлено у лістингу 1.

Лістинг 1. Псевдокод прикладу вивчення EEG даних особи в типовому завданні

```
import numpy as np
import mne
from scipy.signal import welch
import matplotlib.pyplot as plt

# Завантажити EEG
# Підтвердження вводу сирих даних
raw = mne.io.read_raw_fif('your_eeg_data.fif', preload=True)

# Задати спектр частот по смугах
theta_band = (4, 8)
beta_band = (13, 30)
alpha_band = (8, 12)

# Розрахувати спектральну густина частот за допомогою функцій Велча
def compute_psd(data, sfreq, band):
    freqs, psd = welch(data, sfreq, nperseg=1024)
    band_power = np.mean(psd[(freqs >= band[0]) & (freqs <= band[1])])
    return band_power

# Розрахувати відношення активації тета-бета частот
def theta_beta_ratio(data, sfreq):
    theta_power = compute_psd(data, sfreq, theta_band)
    beta_power = compute_psd(data, sfreq, beta_band)
    return theta_power / beta_power

# Розрахувати фронтальну асиметрію альфа-частоти
def frontal_alpha_asymmetry(data, sfreq, left_channels, right_channels):
    left_alpha_power = np.mean([compute_psd(data[ch], sfreq, alpha_band) for ch in
left_channels])
    right_alpha_power = np.mean([compute_psd(data[ch], sfreq, alpha_band) for ch in
```

```

right_channels])
    return left_alpha_power - right_alpha_power

# Визначити канали для фронтальної альфа частоти
left_frontal_channels = ['AF3', 'T7']
right_frontal_channels = ['AF4', 'T8']

# Виділити епохи для нормального стану і спонук ідентифікації
events = mne.find_events(raw)
event_id = {'baseline': 1, 'stimuli': 2}
epochs = mne.Epochs(raw, events, event_id, tmin=-0.2, tmax=0.5, baseline=(None, 0),
preload=True)

# Розрахувати потенціал пов'язаний з подією (ідентифікаційною спонуюкою)
erp = epochs['stimuli'].average()

# Дослідити виділені епохи
baseline_data = epochs['baseline'].get_data()
stimuli_data = epochs['stimuli'].get_data()

# Дослідити дані непарного каналу
channel = 'Pz'
sfreq = raw.info['sfreq']

# Розрахувати відношення тета-бета смуг в підготовці та спонуці
baseline_theta_beta = theta_beta_ratio(baseline_data[:, raw.ch_names.index(channel), :], sfreq)
stimuli_theta_beta = theta_beta_ratio(stimuli_data[:, raw.ch_names.index(channel), :], sfreq)

# Розрахувати асиметрію фронтальної альфа частоти в підготовці та спонуці
baseline_alpha_asymmetry = frontal_alpha_asymmetry(baseline_data, sfreq,
left_frontal_channels, right_frontal_channels)
stimuli_alpha_asymmetry = frontal_alpha_asymmetry(stimuli_data, sfreq, left_frontal_channels,
right_frontal_channels)

# Вивести ERP
erp.plot()

print(f"Baseline Theta/Beta Ratio: {baseline_theta_beta}")
print(f"Stimuli Theta/Beta Ratio: {stimuli_theta_beta}")
print(f"Baseline Frontal Alpha Asymmetry: {baseline_alpha_asymmetry}")
print(f"Stimuli Frontal Alpha Asymmetry: {stimuli_alpha_asymmetry}")

```

*Алгоритм пригнічення когнітивних упереджень при EEG ідентифікації особи та її наміру.*

1. Завантажити дані EEG за розслабленого стану (baseline 30 с) на даний момент. Порівняти зі збереженими в базі даних за допомогою оцінки відношення смуг частот та фронтальної асиметрії.

2. Перевірити стан когнітивної упередженості особи. Якщо ризик вищий за допустимий – застосувати пом'якшення когнітивних упереджень.

3. Ввести умову ідентифікації (stimulus 15-150 с). Завантажити дані EEG та відповіді. Порівняти із типовими індивідуальними даними.

4. Якщо актуальні дані відповідають передбаченим показникам – продовжити. Дані оновити, ідентифікацію підтвердити. Якщо ні – перевірити ще раз. У випадку неспівпадінь – застосувати протокол безпеки.

5. (Додатково, для вищих рівнів захисту) Запропонувати елементарне когнітивне завдання на основі анкети професійної компетентності. Виміряти рівень володіння контролем уваги та ментального стану, оцінити нейрограмотність.

6. Якщо завдання виконані коректно – надати доступ. Якщо ні – запитати повторну ідентифікацію.

7. У разі повторного непроходження ідентифікації або недопустимості даних відповіді – застосувати протокол безпеки.

Псевдокод виконання алгоритму наведено у лістингу 2.

Лістинг 2. Псевдокод пригнічення когнітивних упереджень під час EEG ідентифікації особи та її наміру

```
import mne
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import classification_report, accuracy_score

# Завантажити EEG дані
def load_eeg_data(file_path):
    raw = mne.io.read_raw_fif(file_path, preload=True)
    return raw

# Препроцесінг EEG даних
def preprocess_eeg_data(raw):
    # Apply bandpass filter (0.1-40 Hz)
    raw.filter(0.1, 40, fir_design='firwin')

    # Apply notch filter to remove powerline noise (50 Hz)
    raw.notch_filter(50, fir_design='firwin')

    # Resample data to 200 Hz
    raw.resample(200, npad="auto")

    return raw

# Виділення епох довкола маркерів подій
def extract_epochs(raw, event_id, tmin=-0.2, tmax=0.5):
    events = mne.find_events(raw, stim_channel='STI 014')
    epochs = mne.Epochs(raw, events, event_id, tmin, tmax, baseline=(None, 0), preload=True)
    return epochs

# Виділення потужностей ритмів (смуги частот)
def extract_features(epochs):
    # Define frequency bands
    bands = {
        'delta': (1, 4),
        'theta': (4, 8),
```

```

'alpha': (8, 12),
'beta': (12, 30)
}

features = {}
for band, (fmin, fmax) in bands.items():
    power = mne.time_frequency.psd_welch(epochs, fmin=fmin, fmax=fmax, n_fft=256)
    features[band] = np.mean(power, axis=1)

return features

# Підготовка даних до навчання
def prepare_ml_data(features, labels):
    X = pd.DataFrame(features)
    y = pd.Series(labels)
    return train_test_split(X, y, test_size=0.2, random_state=42)

# Тренування моделі
def train_evaluate_model(X_train, X_test, y_train, y_test):
    model = RandomForestClassifier(n_estimators=100, random_state=42)
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)
    print(classification_report(y_test, y_pred))
    print(f'Accuracy: {accuracy_score(y_test, y_pred)}')

# Головна функція обробки ЕЕГ та визначення когнітивних упереджень
def main(file_path, event_id, labels):
    raw = load_eeg_data(file_path)
    raw = preprocess_eeg_data(raw)
    epochs = extract_epochs(raw, event_id)
    features = extract_features(epochs)
    X_train, X_test, y_train, y_test = prepare_ml_data(features, labels)
    train_evaluate_model(X_train, X_test, y_train, y_test)

# Модуль пригнічення когнітивних упереджень
def preprocess_eeg(eeg_data):
    eeg_data = remove_artifacts(eeg_data)
    eeg_data = apply_bandpass_filter(eeg_data)
    return eeg_data

def extract_features(eeg_data):
    time_features = extract_time_domain_features(eeg_data)
    freq_features = extract_frequency_domain_features(eeg_data)
    spatial_features = extract_spatial_features(eeg_data)
    return concatenate_features(time_features, freq_features, spatial_features)

def mitigate_bias(features):
    normalized_features = normalize(features)
    adapted_features = domain_adaptation(normalized_features)
    return adapted_features

```

```

def train_model(features, labels):
    model = initialize_model()
    model = cross_validate(model, features, labels)
    return model

def identify_intention(eeg_data, model):
    preprocessed_data = preprocess_eeg(eeg_data)
    features = extract_features(preprocessed_data)
    mitigated_features = mitigate_bias(features)
    intention = model.predict(mitigated_features)
    return intention

# Main workflow
eeg_data = collect_eeg_data()
preprocessed_data = preprocess_eeg(eeg_data)
features = extract_features(preprocessed_data)
mitigated_features = mitigate_bias(features)
model = train_model(mitigated_features, labels)
intention = identify_intention(new_eeg_data, model)

# Приклад застосування
if __name__ == "__main__":
    file_path = 'path/to/eeg_data.fif'
    event_id = {'bias_event': 1} # Ідентифікація випадку когнітивного упередження
    labels = [0, 1, 0, 1, 0, 1] # Ознака (0: no bias, 1: bias)
    main(file_path, event_id, labels)

```

**Висновки.** Ефективність розпізнавання було оцінено за показниками точності та вчасності (відповідність образів за оптимальний час). Чинниками якості було визначено «нейрограмотність», «нейропластичність», «кількість учасників», «вік», «стать», «досвід», «складність завдання/стимулу», «темперамент учасника», «метод», «регулярність занять». Вища точність розпізнавання (85%) була досягнута у комбінації: «тип запиту доступу: Зд об’єкт», «тип підтвердження: ментальна команда (3 конфігурації із врахуванням послідовності)», «4 вправи тренування нейропластичності з метою умовної редукції когнітивних упереджень двічі на тиждень», «3 оновлення відбитка розуму на день двічі на тиждень», «метод порівняння середніх амплітуд потенціалів активації нейронних центрів різних півкуль».

	Нейрогр.	Нейропл.	Кількість	Вік	Стать	Досвід	Складн.	Темперам	Метод	Регулярн.
Нейрогр.	1,000	0,997	-0,943	0,789	0,834	0,999	-0,404	0,756	-0,660	-0,715
Нейропл.	0,997	1,000	-0,967	0,837	0,877	0,993	-0,327	0,807	-0,596	-0,655
Кількість	-0,943	-0,967	1,000	-0,948	-0,970	-0,932	0,077	-0,931	0,373	0,441
Вік	0,789	0,837	-0,948	1,000	0,997	0,768	0,243	0,999	-0,059	-0,134
Стать	0,834	0,877	-0,970	0,997	1,000	0,816	0,167	0,992	-0,136	-0,134
Досвід	0,999	0,993	-0,932	0,768	0,816	1,000	-0,434	0,734	-0,684	-0,737
Складн.	-0,404	-0,327	0,077	0,243	0,167	-0,434	1,000	0,294	0,954	0,929
Темперам	0,756	0,807	-0,931	0,999	0,992	0,734	0,294	1,000	-0,007	-0,082
Метод	-0,660	-0,596	0,373	-0,059	-0,136	-0,684	0,954	-0,007	1,000	0,997
Регулярн.	-0,715	-0,655	0,441	-0,134	-0,210	-0,737	0,929	-0,082	0,997	1,000

Рисунок 4 – Кореляція факторів якості ідентифікації особи та наміру в НКС за результатами експертного оцінювання

Нижча якість розпізнавання (68%) була у комбінації: «тип запиту доступу: слово» «тип підтвердження: акцент увагою (4 слова в довільній конфігурації)», «4 вправи тренування нейропластичності з метою умовної редукції когнітивних упереджень двічі на тиждень», «2 оновлення відбитка розуму на тиждень», «метод аналізу незалежних компонент» (Рис. 5).

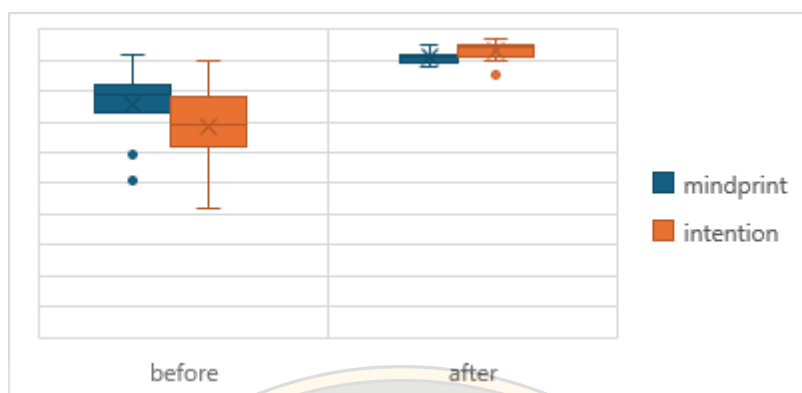


Рисунок 5 – Результуюча успішність ідентифікації за виконання модуля пом'якшення когнітивних упереджень в середньому на 8% вища ніж при звичній ідентифікації

Варто зауважити деякі несподівані результати, зокрема негативну кореляцію до регулярності тренувань. На початкових етапах або при високій частоті оновлення стимулу та умов виникає затримка реакції спровокована упередженістю знаннями та досвідом. У випадку скорочення кількості занять або запитів ідентифікації, нижчій частоті оновлення даних, знижується точність результатів через слабо треновану реакцію. Психологічна специфічність дослідження змушує зважати на високу негативну кореляцію із факторами ефекту спостереження: чим більше різних суб'єктів, тим нижча точність. Таку проблему наразі вирішують на етапі підбору кандидатів, обмежуючи вік від 30 до 50 років. Додатково варто звернути увагу на темперамент, включивши цей показник до обов'язкового тестування. Підбір завдань та розробка тестів, а також навчання суб'єктів мають високу позитивну кореляцію з якістю результатів розпізнавання.

Кількість суб'єктів залежить від кількості даних, які потрібно збирати. Для більшої кількості суб'єктів рекомендовано обмежувати кількість електродів та розробляти складніші завдання, збираючи більше даних. Відкритою проблемою розвитку технології також залишається очевидна непостійність та неоднозначність образів, що вочевидь зумовлено гендерно-віковою різницею учасників тестувань, емоційним контролем (темпераментом), ступенем нейрограмотності та нейропластичності, типовою діяльністю мозку (забування, нейровідновлення, синаптична еволюція та деградація), стресовістю умов та самопочуттям. Вирішення проблеми лежить у полі розробки складних алгоритмів із використанням кількох фільтрів.

Важливо також узгоджувати власні дані з клінічно та правово визначеними стандартами. Сьогодні задачі ідентифікації особи та наміру за допомогою нейро-комп'ютерних інтерфейсів вже застосовуються в експериментальних досудових розслідуваннях (Франція). У 2010 році такі техніки були введені в практику та легітимізовані 2014 [6, 7].

На точність може вплинути збільшення кількості суб'єктів. Крім того, шаблони сигналів мозку можуть суттєво змінюватися з часом, що впливає на продуктивність. Для продовження дослідження цей підхід можна додатково перевірити за допомогою системи ЕЕГ клінічного класу, і його слід дослідити, щоб перевірити, чи можна додатково зменшити кількість електродів.

Варто додати, що деякі пристрої вже дозволяють збирати інформацію впродовж тривалого часу із високою якістю. Наприклад, ЕЕГ-гарнітури як Emotiv MN8 Earbuds, що дозволяють фіксувати рівень уваги та стресу протягом користування. 2023 компанія Apple оголосила, що запатентувала нову гарнітуру Air Pods Sensor System із датчиками ЕЕГ. Система сенсорів

g.Pangolin дозволяє знімати ЕЕГ сигнал як мозку так і периферії. Перевагою є зручність використання протягом тривалого часу [8].

#### ЛІТЕРАТУРА:

1. Brunner, Clemens, et al. "BNCI Horizon 2020: Towards a Roadmap for the BCI Community." *Brain-Computer Interfaces*, vol. 2, no. 1, 2015
2. Yang, YY., Hwang, A.H.C., Wu, C.T. et al. Person-identifying brainprints are stably embedded in EEG mindprints. *Sci Rep* 12, 17031 (2022). <https://doi.org/10.1038/s41598-022-21384-0>
3. Jayarathne, I., Cohen, M., & Amarakeerthi, S. "Person identification from EEG using various machine learning techniques with inter-hemispheric amplitude ratio." *PLoS ONE*, vol. 15, no. 9, 2020, e0238872. DOI: 10.1371/journal.pone.0238872
4. López-Silva, Pablo, and Luca Valera, editors. *Protecting the Mind: Challenges in Law, Neuroprotection, and Neurorights*. Springer International Publishing, 2023.
5. Nam, Chang S., Anton Nijholt, and Fabien Lotte, editors. *Brain-Computer Interfaces Handbook: Technological and Theoretical Advances*. 1st ed., CRC Press, 2018.
6. Mykhalchuk V., Druzhynin V. Targeting mind in challenging conditions of biased nature of performance (preliminary results on individual experience). 1st international scientific and practical conference «Information Systems and Technology: Results and Prospects» (IST 2024)", Kyiv, 2024.
7. *Imagerie Cérébrale: Enjeux Épistémologiques, Éthiques Et Politiques, Espace Ethique d'Ile de France*, 2017. 97p.
8. Park Sang-Eon et al. Digital Assessment of Cognitive-Affective Biases Related to Mental Health. *PLOS Digital Health*, vol. 1, no. 1, 2024, article e0000595, <https://doi.org/10.1371/journal.pdig.0000595>.
9. Schreiner, Leonhard, et al. "Mapping of the Central Sulcus Using Non-Invasive Ultra-High-Density Brain Recordings." *Scientific Reports*, vol. 14, no. 6527, 19 Mar. 2024, doi:10.1038/s41598-024-57167-y.
10. Korteling, J. E., Gerritsma, J. Y. J., & Toet, A. (2021)1. Retention and transfer of cognitive bias mitigation interventions: A systematic literature study23. *Frontiers in Psychology*, 12, 629354. <https://doi.org/10.3389/fpsyg.2021.629354>
11. Mykhalchuk, V. "Cognitive Design of Educational Brain-Computer Interfaces." *CEUR Workshop Proceedings*, vol. 3309, 2022, 1. DOI: 10.1007/978-3-319-64274-1\_8.
12. Siritzky M. Emma. Standard Experimental Paradigm Designs and Data Exclusion Practices in Cognitive Bias Research. *Cognitive Research: Principles and Implications*, vol. 8, no. 1, 2023, article 520, <https://doi.org/10.1186/s41235-023-00520-y>.
13. Tmienova N., Mykhalchuk V. Brain-Computer Interface as Tool of Cognitive Optimization (Case of Biases Reducing in Decision-Making and Control Improvement). *CEUR Workshop Proceedings*, vol. 3309, 2022, 1. DOI: 10.1007/978-3-319-64274-1\_8.
14. A Review on Evaluating Mental Stress by Deep Learning Using EEG Signals. *Proceedings of the 1st International Scientific and Practical Conference 'Information Systems and Technology: Results and Prospects' (IST 2024)*, 2024, pp. 45-56.
15. An Evaluation of Mental Workload with Frontal EEG. *PLOS ONE*, vol. 12, no. 4, 2017, article e0174949, <https://doi.org/10.1371/journal.pone.0174949>.
17. *Mental Workload Monitoring: New Perspectives from Neuroscience*. *Advances in Neuroergonomics and Cognitive Engineering*, edited by Hasan Ayaz and Umer Asgher, Springer, 2020, pp. 1-12, [https://doi.org/10.1007/978-3-030-32423-0\\_1](https://doi.org/10.1007/978-3-030-32423-0_1).
18. Real-Time Mental Workload Estimation Using EEG. *Advances in Neuroergonomics and Cognitive Engineering*, edited by Hasan Ayaz and Umer Asgher, Springer, 2020, pp. 13-24, [https://doi.org/10.1007/978-3-030-32423-0\\_2](https://doi.org/10.1007/978-3-030-32423-0_2).

#### REFERENCES:

1. Brunner, Clemens, et al. "BNCI Horizon 2020: Towards a Roadmap for the BCI Community." *Brain-Computer Interfaces*, vol. 2, no. 1, 2015
2. Yang, YY., Hwang, A.H.C., Wu, C.T. et al. Person-identifying brainprints are stably embedded in EEG mindprints. *Sci Rep* 12, 17031 (2022). <https://doi.org/10.1038/s41598-022-21384-0>

3. Jayarathne, I., Cohen, M., & Amarakeerthi, S. "Person identification from EEG using various machine learning techniques with inter-hemispheric amplitude ratio." *PLoS ONE*, vol. 15, no. 9, 2020, e0238872. DOI: 10.1371/journal.pone.023888
4. López-Silva, Pablo, and Luca Valera, editors. *Protecting the Mind: Challenges in Law, Neuroprotection, and Neurorights*. Springer International Publishing, 2023.
5. Nam, Chang S., Anton Nijholt, and Fabien Lotte, editors. *Brain-Computer Interfaces Handbook: Technological and Theoretical Advances*. 1st ed., CRC Press, 2018.
6. Mykhalchuk V., Druzhynin V. Targeting mind in challengable conditions of biased nature of performance (preliminary results on individual experience). 1st international scientific and practical conference «Information Systems and Technology: Results and Prospects» (IST 2024)", Kyiv, 2024.
7. *Imagerie Cérébrale: Enjeux Épistémologiques, Éthiques Et Politiques*, Espace Ethique d'Ile de France, 2017. 97p.
8. Park Sang-Eon et al. Digital Assessment of Cognitive-Affective Biases Related to Mental Health. *PLOS Digital Health*, vol. 1, no. 1, 2024, article e0000595, <https://doi.org/10.1371/journal.pdig.0000595>.
9. Schreiner, Leonhard, et al. "Mapping of the Central Sulcus Using Non-Invasive Ultra-High-Density Brain Recordings." *Scientific Reports*, vol. 14, no. 6527, 19 Mar. 2024, doi:10.1038/s41598-024-57167-y.
10. Korteling, J. E., Gerritsma, J. Y. J., & Toet, A. (2021)1. Retention and transfer of cognitive bias mitigation interventions: A systematic literature study23. *Frontiers in Psychology*, 12, 629354. <https://doi.org/10.3389/fpsyg.2021.629354>
11. Mykhalchuk, V. "Cognitive Design of Educational Brain-Computer Interfaces." *CEUR Workshop Proceedings*, vol. 3309, 2022, 1. DOI: 10.1007/978-3-319-64274-1\_8.
12. Siritzky M. Emma. Standard Experimental Paradigm Designs and Data Exclusion Practices in Cognitive Bias Research. *Cognitive Research: Principles and Implications*, vol. 8, no. 1, 2023, article 520, <https://doi.org/10.1186/s41235-023-00520-y>.
13. Tmienova N., Mykhalchuk V. Brain-Computer Interface as Tool of Cognitive Optimization (Case of Biases Reducing in Decision-Making and Control Improvement). *CEUR Workshop Proceedings*, vol. 3309, 2022, 1. DOI: 10.1007/978-3-319-64274-1\_8.
14. A Review on Evaluating Mental Stress by Deep Learning Using EEG Signals. *Proceedings of the 1st International Scientific and Practical Conference 'Information Systems and Technology: Results and Prospects' (IST 2024)*, 2024, pp. 45-56.
15. An Evaluation of Mental Workload with Frontal EEG. *PLOS ONE*, vol. 12, no. 4, 2017, article e0174949, <https://doi.org/10.1371/journal.pone.0174949>.
17. *Mental Workload Monitoring: New Perspectives from Neuroscience*. *Advances in Neuroergonomics and Cognitive Engineering*, edited by Hasan Ayaz and Umer Asgher, Springer, 2020, pp. 1-12, [https://doi.org/10.1007/978-3-030-32423-0\\_1](https://doi.org/10.1007/978-3-030-32423-0_1).
18. *Real-Time Mental Workload Estimation Using EEG*. *Advances in Neuroergonomics and Cognitive Engineering*, edited by Hasan Ayaz and Umer Asgher, Springer, 2020, pp. 13-24, [https://doi.org/10.1007/978-3-030-32423-0\\_2](https://doi.org/10.1007/978-3-030-32423-0_2).

## SUPPRESSION OF COGNITIVE BIASES IN OF A PERSON AND HIS INTENTION IDENTIFICATION TASK FOR NEURO-COMPUTER SYSTEMS

*The study is devoted to the problem of security of information systems in the case of recognition of a person and his intentions using an electroencephalogram. The method is designed for a non-invasive neurocomputer interface, can be used in multi-level research (emotional intelligence, memory development, monitoring well-being, optimization of decision-making), which is its advantage. The versatility of the neuro-computer interface allows you to use it for several purposes at the same time, combining different methods of analysis of one data set (encephalogram, response time and adequacy). An applied combined approach to quality assessment based on design indicators based on the results of previous studies, taking into account the specifics of user tasks and roles, data processing methods, and the objectivity of the selected models, is proposed. The experiments were carried out in simulated conditions in two stages of 3 months each based on the results of neuroplasticity training and solving tasks of obtaining access in the presence of rights and in the conditions of an unexpected request. The individual style of performing tasks is taken into account, and the emotion recognition strategy based on the Chikszentmihalyi's model is used to predict the intention. The results meet modern standards of neuroethics and neurolaw. Cognitive biases can negatively affect the task of identifying a person and his intentions, for this, mitigation methods are proposed. The ease of use of the model lies in the simple interpretation of the collected data. Frontal alpha asymmetry, event-related potential, and alpha-beta band power ratio of synchronization rhythms were used for assessment. The results are convincing that the expected implementation can be effective. The work takes into account the rich experience of previous research on the identification of a person and his intentions based on the analysis of the electroencephalogram. In this study, the impermanence of the EEG image ("brain imprint") is taken into account and attention is focused on the importance of its renewal, an experiment was conducted in the course of a comprehensive study. The result is a substantiation of requirements for the frequency and procedure of updating "brain prints", proposed indicators of the quality of recognition, proposals for ranking the level of neuroliteracy based on testing typical activities. In order to study the intention of a person, it is proposed to study the background cognitive states, taking into account their duration and intensity of typical patterns of synchronization rhythms, on the basis of which machine learning of the cognitive bias suppression model was carried out. Data were compared with moments of loss of control, the need to rethink and update brain synchronization networks was taken into account. These results are currently basic and require special attention in further experiments, their thorough investigation will be important for improving the quality of any research on the use of neuro-computer interfaces. The risk of unwanted cognitive states and the appearance of cognitive biases has been assessed, and requirements for taking tests are provided to prevent such situations. The influence of experience and training level of neurocomputer literacy, individual suitability to work in neuro-computer environments (in particular, neurosignal power and control of mental efforts, cognitive properties and emotional backgrounds) are taken into account.*

*Key words: neurosecurity, non-invasive neurocomputer interface, EEG-identification of intention, mitigation of cognitive biases.*