

ХОМ'ЯК АНДРІЙ

Тернопільський національний технічний університет ім. І. Пулюя

<https://orcid.org/0000-0002-1763-868X>E-mail: andrii.khomiak@outlook.com

АНАЛІЗ МЕГ СИГНАЛІВ РУХУ КИСТІ ШЛЯХОМ ПОБУДОВИ РЕГРЕСІЙНОЇ МОДЕЛІ НА ОСНОВІ LSTM

В роботі наведено методологічну основу, застосовану для розробки та оцінки регресора магнітоенцефалографічного (МЕГ) сигналу на основі мережі довгої короткострокової пам'яті (LSTM). Детально описано вибір архітектури, протоколи попередньої обробки даних, режим навчання та стратегію оптимізації. Продуктивність моделі кількісно оцінено за допомогою метрик MSE та R^2 , а також розглянуто перспективи використання вивчених представлень для подальших завдань класифікації.

Ключові слова: магнітоенцефалографія, LSTM, регресія, часові ряди, нейронні мережі, глибоке навчання.

KHOMIAK ANDRII

Ternopil Ivan Puluj National Technical University

ANALYSIS OF MEG SIGNALS OF WRIST MOVEMENT WITH LSTM-BASED REGRESSION MODEL

This article describes the methodological framework for the development and evaluation of a magnetoencephalographic (MEG) signal regressor based on a Long Short-Term Memory (LSTM) network, applied to the analysis of time-series data from wrist movement tasks. The study provides a detailed account of the model's architecture, which is constructed from sequence input, LSTM, and dense layers designed to process multi-channel temporal data. It outlines the data preprocessing pipeline, which includes the standardization of features for each MEG channel using a StandardScaler and the specific formulation of input-target sequences necessary for time-series forecasting. The training and optimization protocols are specified in detail, including the use of the Adam optimizer, a Mean Squared Error (MSE) loss function, and a ReduceLROnPlateau learning rate scheduler to dynamically adjust learning based on validation performance. To ensure numerical stability, particularly when training on long sequences, gradient clipping was implemented. The model's predictive performance was quantitatively assessed using MSE to measure prediction error and the Coefficient of Determination (R^2) to determine the proportion of explained variance. These metrics were evaluated across distinct training, validation, and test datasets to assess the model's learning capacity, its ability to generalize to unseen data, and to monitor for potential overfitting. Finally, the paper considers the application of the latent representations learned by the LSTM regressor. These internal representations, which encapsulate critical temporal features of the MEG signals, are proposed as a source of data-driven features for subsequent, more complex classification tasks, such as distinguishing between different brain states or experimental conditions.

Keywords: magnetoencephalography, LSTM, regression, time series, neural networks, deep learning.

Стаття надійшла до редакції / Received 25.06.2025

Прийнята до друку / Accepted 15.08.2025

Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями

У цій статті представлено методологічну основу та оцінку продуктивності регресора магнітоенцефалографічного (МЕГ) сигналу. Магнітоенцефалографія є неінвазивним методом нейровізуалізації та нейродіагностики, що вимірює слабкі магнітні поля, генеровані синхронною нейронною активністю в мозку. Її виняткова часова роздільна здатність дозволяє фіксувати багату, швидкозмінну динаміку нейронних процесів [1]. Дана робота відображає систематичний підхід, застосований на всіх етапах: від початкового вибору моделі та попередньої обробки даних до тонкощів конфігурації навчання та остаточної оцінки продуктивності моделі, все це з дотриманням суворих наукових принципів. Крім того, розробка цього регресора розглядається як фундаментальний крок. Передбачається, що внутрішні представлення, вивчені моделлю для прогнозування неперервних МЕГ сигналів, фіксують суттєву нейронну динаміку, яка згодом може слугувати потужними ознаками для складніших подальших завдань, таких як класифікація МЕГ сигналів [2].

Розробка регресора для МЕГ-сигналів, а не класифікатора, була мотивована необхідністю прогнозування неперервних значень у складних нейронних часових рядах. Цей підхід спрямований на декодування безперервних поведінкових параметрів, реконструкцію змінних у часі характеристик стимулів або передбачення майбутніх станів самих нейронних сигналів. Вважається, що такий метод виходить за межі простої категоризації мозкових станів чи реакцій, аналізуючи кількісні аспекти функціонування мозку та його взаємозв'язку з неперервними зовнішніми змінними. Таким чином, результати, отримані за допомогою цього регресора, можуть значно поглибити розуміння кодування конкретної безперервної нейронної або поведінкової динаміки в МЕГ-сигналах. Це, у свою чергу, може сприяти розвитку вдосконалених інтерфейсів мозок-комп'ютер (ІМК) [3], які потребують точного безперервного керування, а також застосувань у нейрозворотному зв'язку [4], де першочергове значення мають прогнозування та модуляція поточної мозкової активності в реальному часі.

Аналіз досліджень та публікацій

При розробці надійного регресора МEG сигналу вибір відповідної архітектури моделі був критично важливим першим кроком. Враховуючи часову природу МEG даних, що характеризуються послідовностями вимірювань у часі, було визначено, що моделі, здатні фіксувати часові залежності, є необхідними. Для моделювання часових рядів МEG даних були обрані мережі довгої короткострокової пам'яті (LSTM) через їхню природну придатність для обробки послідовних даних. LSTM [5] є спеціалізованим типом рекурентних нейронних мереж (RNN), розробленим для подолання обмежень традиційних RNN, зокрема проблеми зникаючого градієнта, яка перешкоджає навчанню довгострокових залежностей.

Вибір мереж LSTM замість простіших моделей часових рядів, таких як моделі авторегресії ковзного середнього (ARIMA) (які також використовуються для прогнозування часових рядів) [6], для аналізу МEG сигналів ґрунтувався на апіорному припущенні щодо природи даних. Відомо, що МEG сигнали характеризуються складною, нестаціонарною динамікою, що відображає розподілені та взаємодіючі нейронні процеси. Було висунуто гіпотезу, що часові залежності в цих сигналах часто є нелінійними та можуть охоплювати значні проміжки часу. LSTM, завдяки своїй здатності вивчати такі складні патерни, виявилися краще пристосованими до цієї складності, ніж традиційні статистичні моделі, такі як ARIMA, які переважно фіксують лінійні залежності та мають більш обмежену пам'ять. Цей вибір відображав очікування, що ці складні, довгострокові та нелінійні залежності не тільки присутні, але й є критично важливими для конкретного завдання регресії. Було зрозуміло, що простіші моделі можуть не вловити ці суттєві нюанси, що призведе до неоптимальної продуктивності прогнозування. Хоча цей вибір вимагав більших обчислювальних ресурсів та потенційно більших наборів даних для ефективного навчання, вважалося, що він пропонує потенціал для більш точного та детального моделювання мозкової активності. Успішне застосування такої моделі додатково підтверджує корисність підходів глибокого навчання для складних нейрофізіологічних даних.

Формулювання цілей статті

Метою роботи є розробка та оцінка регресора МEG сигналу на основі архітектури мережі довгої короткострокової пам'яті для аналізу сигналів руху кисті. Це включає:

1. Обґрунтування вибору моделі LSTM для аналізу часових рядів МEG.
2. Детальний опис архітектури моделі та протоколу попередньої обробки даних.
3. Викладення режиму навчання, стратегії оптимізації та методів забезпечення стабільності.
4. Оцінку продуктивності моделі за допомогою кількісних метрик (MSE та R^2) та інтерпретацію результатів.
5. Розгляд потенційного використання вивчених представлень для подальших завдань класифікації.

Виклад основного матеріалу

У цьому підході мережа LSTM обробляє вхідні дані шляхом ітерації за часовими кроками та оновлення свого внутрішнього стану, який містить інформацію, накопичену за всі попередні часові кроки. Цей механізм є фундаментальним для розуміння сигналів, де контекст минулого впливає на поточні та майбутні стани. Основна перевага LSTM, і ключова причина їх вибору, полягає в їхній складній архітектурі, яка включає спеціалізовані вентиляльні механізми: вентиль забування, вхідний вентиль та вихідний вентиль. Ці вентиляції спільно регулюють потік інформації через комірку та між часовими кроками. Вентиль забування визначає, яку інформацію з попереднього стану комірки слід відкинути, вхідний вентиль контролює, яка нова інформація з поточного входу зберігається в стані комірки, а вихідний вентиль вирішує, яка інформація зі стану комірки використовується для генерації виходу для поточного часового кроку. Таке вибіркоче збереження та відкидання інформації дозволяє використаним моделям LSTM ефективно вивчати та підтримувати довгострокові залежності, що є критично важливою характеристикою для МEG сигналів, де релевантні нейронні події можуть бути розділені значними часовими лагами. Дійсно, перевага надавалась LSTM, оскільки вони розроблені для утримання інформації протягом тривалих періодів за замовчуванням, здатні долати часові лаги майже в 1000 кроків або більше.

Конкретна архітектура регресора LSTM була ретельно налаштована для обробки багатоканальних часових рядів МEG даних. Модель починається з шару входу послідовностей, призначеного для прийому МEG послідовностей. Критичним обмеженням для цього шару є те, що кожна вхідна послідовність повинна мати однакову кількість каналів, які відповідають ознакам вхідних даних на кожному часовому кроці. Після вхідного шару один або декілька шарів LSTM утворюють ядро мережі. Наприклад, поширена конфігурація, що тестувалася, включала шар LSTM із заданою кількістю прихованих блоків, наприклад, 128 блоків. Кількість прихованих блоків у шарі LSTM є ключовим гіперпараметром, який розглядався, оскільки він визначає репрезентативну здатність шару - по суті, скільки інформації шар може вивчити та зберегти з вхідних послідовностей. Внутрішні обчислення цих блоків LSTM у моделі включають функції активації, такі як гіперболічний тангенс для оновлення стану комірки та модуляції виходу, та сигмоїдні функції для активації вентилів, які контролюють потік інформації. Останнім шаром у регресійній моделі є щільний шар, також відомий як повнозв'язний шар. Для завдання регресії цей вихідний шар був налаштований з кількістю нейронів,

що відповідає розмірності прогнозованої цільової змінної (змінних). У цьому вихідному шарі використовувалася лінійна функція активації, оскільки вона дозволяє моделі виводити значення в усьому діапазоні дійсних чисел, не обмежуючись певним інтервалом.

Навчання нейронних мереж може бути значно ускладнене, якщо вхідні ознаки мають суттєво різні масштаби або одиниці виміру. Для вирішення цієї проблеми в цьому дослідженні стандартизація ознак була стандартним кроком попередньої обробки. Техніка стандартного масштабування трансформує дані шляхом віднімання середнього значення (μ) кожної ознаки та подальшого ділення на її стандартне відхилення (σ). Математичне перетворення, застосоване для кожної ознаки X , задається формулою:

$$X_{\text{масштабоване}} = \frac{X - \mu}{\sigma}$$

Цей процес стандартизував кожну ознаку (наприклад, сигнал кожного МЕГ каналу в часі) в наборі даних так, щоб вона мала середнє значення 0 та стандартне відхилення 1. Обґрунтування використання цього підходу полягає в тому, що багато алгоритмів машинного навчання, включаючи використані нейронні мережі, часто припускають, що вхідні дані розподілені нормально або, принаймні, що ознаки мають порівнянний масштаб. Стандартизація допомогла задовольнити це неявне припущення, центрувала дані навколо нуля та забезпечила більш рівномірний внесок усіх ознак у процес навчання, запобігаючи домінуванню ознак з більшими абсолютними значеннями в оновленнях градієнтів. Було помічено, що це призвело до швидшої збіжності та покращення продуктивності моделі в цих експериментах.

Після індивідуальної стандартизації МЕГ каналів наступним важливим кроком було структурування цих даних у послідовності, придатні для навчання мережі LSTM. LSTM розроблені для навчання на послідовностях, де порядок точок даних має значення. Для завдання регресії часових рядів, яке мало на меті прогнозування майбутніх значень або одночасних безперервних змінних на основі минулої історії сигналу, підхід полягав у визначенні цільових змінних (відгуків) як самих навчальних послідовностей, але зі зміщеними на один або декілька часових кроків значеннями. У цій методології на кожному часовому кроці t вхідної послідовності мережа LSTM навчалася прогнозувати значення відповідної цільової змінної на наступному часовому кроці, зазвичай $t+1$ для прогнозування на один крок вперед. Отже, предикторами (вхідними послідовностями для LSTM) були навчальні послідовності без останнього часового кроку (або кроків, залежно від горизонту прогнозування), які відповідали зміщеним цільовим значенням. Наприклад, якщо вхідна послідовність у даних складалася з МЕГ даних від часу t_1 до t_k , відповідною цільовою послідовністю були МЕГ дані (або похідна безперервна змінна) від часу t_2 до t_{k+1} . Таке формулювання "послідовність-до-послідовності"[7] дозволило LSTM вивчити часову динаміку, необхідну для прогнозування на основі спостережуваної історії сигналу.

Для оновлення вагових коефіцієнтів мережі під час навчання був обраний алгоритм оптимізації Adam [8]. Adam є алгоритмом оптимізації з адаптивною швидкістю навчання, відомим своєю ефективністю та результативністю, що було визнано корисним у широкому діапазоні завдань глибокого навчання. Він поєднує переваги двох інших популярних методів оптимізації: Momentum та Root Mean Square Propagation (RMSprop)[9]. Adam обчислює індивідуальні адаптивні швидкості навчання для різних параметрів на основі оцінок першого та другого моментів градієнтів. Перший момент (середнє значення градієнтів) допомагає прискорити навчання в узгоджених напрямках, тоді як другий момент (нецентрована дисперсія градієнтів) допомагає масштабувати швидкість навчання для кожного параметра.

Для завдання регресії як функція втрат була обрана середньоквадратична помилка (MSE). MSE є стандартною та широко використовуваною метрикою для оцінки продуктивності регресійних моделей, яка була визнана придатною для кількісної оцінки помилок у прогнозах. Вона кількісно визначає середнє значення квадратів помилок, де помилка - це різниця між прогнозованим та фактичним значенням. Використана формула для MSE:

$$MSE = \frac{1}{n} \sum_{i=1}^n \frac{1}{k} \sum_{j=1}^k (y_{ij} - \hat{y}_{ij})^2$$

де n - кількість зразків, k - кількість каналів в сигналі (10), $y_i \in \mathbb{R}^k$ - k -розмірний вектор дійсного значення дигналів для i -го зразка, $\hat{y}_i \in \mathbb{R}^k$ - прогнозоване значення для i -го зразка. Підносячи помилки до квадрату, MSE сильніше штрафувє більші помилки, що спонукало модель уникати значних відхилень. Це диференційовна функція, що робить її придатною для методів оптимізації на основі градієнта, таких як Adam.

Декілька ключових параметрів керували процесом навчання: модель навчалася протягом 100 епох; навчання проводилося з використанням міні-батчів (для обробки послідовностей різної довжини, вони доповнювалися, щоб забезпечити однакову довжину); навчальні дані перемішувалися на початку кожної епохи; для динамічного коригування швидкості навчання під час тренування була реалізована стратегія планування ReduceLROnPlateau [10] - це адаптивний підхід, при якому швидкість навчання зменшувалася, коли контрольована метрика валідації переставала покращуватися; для пом'якшення проблеми вибухаючих градієнтів та забезпечення чисельної стабільності застосовувалося обрізання

градієнта [11], що обмежує величину градієнтів перед їх використанням для оновлення вагових коефіцієнтів.

Нижче наведено значення гіперпараметрів що були використані під час тренування:

- Архітектура моделі:
 - LSTM нейронна мережа з 3 послідовними шарами та 128 прихованими вузлами на шар
 - Розмірність входу: 10 каналів (дані МEG сенсорів)
 - Розмірність виходу: 10 каналів (прогнозований наступний часовий крок)
 - Загальна кількість параметрів для навчання: 37,600 (розмір моделі 0.151 МБ)
 - Dropout [12] регуляризація: 0.2 ймовірність між шарами
- Конфігурація навчання:
 - Розмір батча: 64 зразки на тренувальний батч
 - Довжина послідовного вікна: 50 часових кроків для темпорального прогнозування
 - Максимальна кількість епох навчання: 200 з раннім зупиненням [13]
 - Швидкість навчання: 0.0001 з адаптивним плануванням
 - Зниження ваг (L2 регуляризація): 0.00001
 - Поріг обрізання градієнтів: 1.0 для запобігання вибуховим градієнтам
- Стратегія оптимізації:
 - Оптимізатор Adam з корекцією зміщення для адаптивних швидкостей навчання
 - Планувальник ReduceLRonPlateau: коефіцієнт зниження 0.5, терпіння 5 епох
 - Раннє зупинення: терпіння 10 епох з мінімальним порогом покращення 0.001
 - Нормалізація даних: StandardScaler застосований для запобігання розбіжності навчання

Основною кількісною метрикою для оцінки точності прогнозування регресора була середньоквадратична помилка (MSE), що узгоджується з її використанням як функції втрат під час навчання. На додаток до MSE, для оцінки якості підгонки моделі використовувався коефіцієнт детермінації, широко відомий як R-квадрат (R^2). R^2 є статистичною мірою, яка представляє частку дисперсії залежної змінної, що прогнозується незалежними змінними в моделі. Значення R^2 розраховувалося за формулою:

$$R^2 = 1 - \frac{\sum_{i=1}^n \sum_{j=1}^k (y_{ij} - \hat{y}_{ij})^2}{\sum_{i=1}^n \sum_{j=1}^k (y_{ij} - \bar{y}_j)^2}$$

де y_i - фактичні значення для i -го зразка, \hat{y}_i - прогнозоване значення з моделі, а \bar{y} - середнє значення фактичних значень.

Отримані значення MSE та R^2 інтерпретувалися в контексті конкретного завдання регресії МEG сигналу. Це включало аналіз здатності моделі фіксувати основну часову динаміку МEG сигналів, її продуктивність генералізації (шляхом порівняння показників на навчальному наборі з показниками на валідаційному та тестовому наборах) та виявлення будь-яких ознак перенавчання або недонавчання. Спільне звітування MSE та R^2 забезпечило більш комплексну оцінку, ніж будь-яка метрика окремо.

Тренування моделі здійснювалося на відеокарті NVIDIA RTX 3060 6GB для апаратного прискорення з використанням інструментів CUDA [8].

На рисунку 1 зображено історію значення функції витрат на тренувальній та валідаційній вибірці (можна побачити характерні для планувальника швидкості навчання криві).

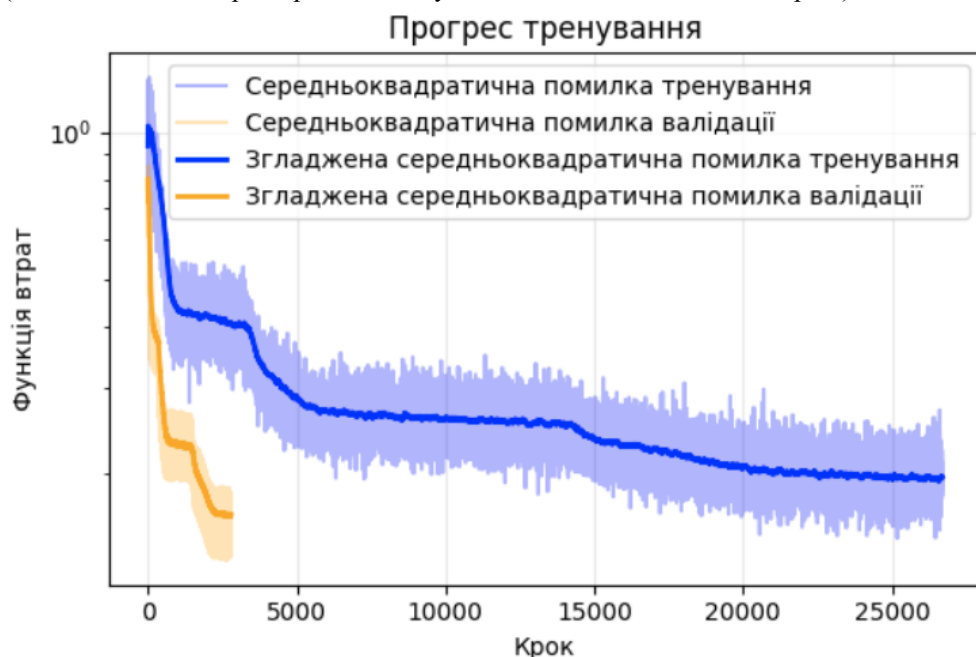


Рис.1. Прогрес тренування моделі

Результати тренування:

- Збіжність навчання: 111 епоха з 200 максимальних
- Фінальна втрата валідації: 0.162 (середньоквадратична похибка)
- Продуктивність на тестовому наборі: $R^2 = 0.827$ (82.7% поясненої дисперсії)
- Тренувальний набір даних: 17,760 зразків у 278 батчах

Висновки з даного дослідження і перспективи подальших досліджень у даному напрямі

У цій роботі детально описано методологічну основу, застосовану для розробки та оцінки регресора МEG сигналу. Вибір архітектури LSTM був зумовлений її природною здатністю моделювати складні часові залежності, присутні в послідовних нейрофізіологічних даних. Попередня обробка даних включала стандартизацію ознак для нормалізації даних МEG каналів, після чого відбувалася відповідна підготовка послідовностей та цільових змінних, адаптована для обробки LSTM. Розроблений режим навчання мав на меті забезпечення надійності та ефективності, включаючи алгоритм оптимізації Adam за його адаптивні можливості швидкості навчання, середньоквадратичну помилку як функцію втрат регресії, та планувальник швидкості навчання ReduceLROnPlateau для динамічного коригування навчання на основі продуктивності валідації. Критично важливим було впровадження обрізання градієнта для забезпечення чисельної стабільності під час навчання глибокої рекурентної архітектури, що пом'якшувало ризик вибухаючих градієнтів. Продуктивність кількісно оцінювалася за допомогою MSE для вимірювання точності прогнозування та коефіцієнта детермінації (R^2) для оцінки якості підгонки, що вказує на частку дисперсії характеристик МEG сигналу, пояснену моделлю. Оцінка на навчальних, валідаційних та тестових наборах надала уявлення про здатність моделі до навчання, її генералізаційну здатність та потенціал для перенавчання або недонавчання. Окрім прямого застосування в прогнозуванні сигналів, ця регресійна структура також була розроблена з перспективою майбутніх досліджень у класифікації МEG сигналів. Передбачається, що латентні представлення, розроблені LSTM під час виконання завдання регресії, інкапсулюють критичні часові та спектральні ознаки МEG сигналів. Ці вивчені ознаки призначені для вилучення та використання як вхідні дані для наступних класифікаційних моделей, потенційно пропонуючи більш керований даними та деталізований набір ознак, ніж традиційні ознаки, створені вручну, для розрізнення між різними станами мозку або експериментальними умовами. Таким чином, продуктивність цього регресора також слугує опосередкованою валідацією якості цих вивчених представлень для таких майбутніх застосувань. Успішна демонстрація можливостей регресора зміцнила б впевненість у корисності LSTM для складних завдань регресії на основі МEG та для вилучення ознак для класифікації. І навпаки, будь-які виявлені обмеження вказали б на критичні сфери, що потребують подальшого дослідження. Майбутня робота може включати дослідження альтернативних архітектур нейронних мереж, більш складних методів попередньої обробки або передових методів регуляризації для подальшого покращення генералізації та надійності моделі для цих складних нейрофізіологічних сигналів, як для регресії, так і для ознак, отриманих для завдань класифікації.

Література

1. BCI competition IV [Electronic resource] // Wayback Machine. - Mode of access: <https://web.archive.org/web/20240415211536/https://bbci.de/competition/iv/> (date of access: 17.06.2025). - Title from screen.
2. Хом'як А. Методологія класифікації магнітоенцефалографічних сигналів із використанням напівкерованого навчання [Електронний ресурс] / Андрій Хом'як // Herald of Khmelnytskyi National University. Technical sciences. - 2024. - Т. 339, № 4. - С. 110-114. - Режим доступу: <https://doi.org/10.31891/2307-5732-2024-339-4-17>. - Назва з екрана.
3. Schalk G. Brain-computer symbiosis [Electronic resource] / Gerwin Schalk // Journal of neural engineering. - 2008. - Vol. 5, no. 1. - P. P1-P15. - Mode of access: <https://doi.org/10.1088/1741-2560/5/1/p01> (date of access: 17.06.2025). - Title from screen.
4. Marzbani H. Methodological note: neurofeedback: a comprehensive review on system design, methodology and clinical applications [Electronic resource] / H. Marzbani, H. Marateb, M. Mansourian // Basic and clinical neuroscience journal. - 2016. - Vol. 7, no. 2. - Mode of access: <https://doi.org/10.15412/j.bcn.03070208> (date of access: 17.06.2025). - Title from screen.
5. Hochreiter S. Long Short-Term Memory [Electronic resource] / Sepp Hochreiter, Jürgen Schmidhuber // Neural computation. - 1997. - Vol. 9, no. 8. - P. 1735-1780. - Mode of access: <https://doi.org/10.1162/neco.1997.9.8.1735> (date of access: 17.06.2025). - Title from screen.
6. An enhanced ARIMA model for EEG classification [Electronic resource] / Yan Liu [et al.] // WI-IAT '21: IEEE/WIC/ACM international conference on web intelligence, ESSENDON VIC Australia. - New York, NY, USA, 2021. - Mode of access: <https://doi.org/10.1145/3498851.3498955> (date of access: 17.06.2025). - Title from screen.
7. Sutskever I. Sequence to sequence learning with neural networks [Electronic resource] / Pya Sutskever, Oriol Vinyals, Quoc V. Le. - Mode of access: <https://arxiv.org/abs/1409.3215>. - Title from screen.

8. Kingma D. P. Adam: a method for stochastic optimization [Electronic resource] / Diederik P. Kingma, Jimmy Ba. - Mode of access: <https://arxiv.org/abs/1412.6980>. - Title from screen.
9. Teaching machines to read and comprehend [Electronic resource] / Karl Moritz Hermann [et al.]. - Mode of access: <https://arxiv.org/abs/1506.03340>. - Title from screen.
10. Mukherjee K. A simple dynamic learning rate tuning algorithm for automated training of dnns [Electronic resource] / Koyel Mukherjee, Alind Khare, Ashish Verma. - Mode of access: <https://arxiv.org/abs/1910.11605>. - Title from screen.
11. Why gradient clipping accelerates training: a theoretical justification for adaptivity [Electronic resource] / Jingzhao Zhang [et al.]. - Mode of access: <https://arxiv.org/abs/1905.11881>. - Title from screen.
12. Dropout: a simple way to prevent neural networks from overfitting [Electronic resource] / Nitish Srivastava [et al.] // Journal of machine learning research. - 2014. - Vol. 15, no. 56. - P. 1929-1958. - Mode of access: <http://jmlr.org/papers/v15/srivastava14a.html>. - Title from screen.
13. Understanding and improving early stopping for learning with noisy labels [Electronic resource] / Yingbin Bai [et al.]. - Mode of access: <https://arxiv.org/abs/2106.15853>. - Title from screen.
14. Mocanu I. An INTRODUCTION TO CUDA Programming / Irina Mocanu // Journal of information systems & operations management. - 2008. - Vol. 2. - P. 495-506.

References

1. BCI competition IV [Electronic resource] // Wayback Machine. - Mode of access: <https://web.archive.org/web/20240415211536/https://bbci.de/competition/iv/> (date of access: 17.06.2025). - Title from screen.
2. Khomiak A. Methodology of magnetoencephalography signal classification using semi-supervised learning [Electronic resource] / Andrii Khomiak // Herald of Khmelnytskyi National University. Technical sciences. - 2024. - T. 339, № 4. - S. 110-114. - Mode of access: https://doi.org/10.31891/23_07-5732-2024-339-4-17. - Title from screen.
3. Schalk G. Brain-computer symbiosis [Electronic resource] / Gerwin Schalk // Journal of neural engineering. - 2008. - Vol. 5, no. 1. - P. P1-P15. - Mode of access: <https://doi.org/10.1088/1741-2560/5/1/p01> (date of access: 17.06.2025). - Title from screen.
4. Marzbani H. Methodological note: neurofeedback: a comprehensive review on system design, methodology and clinical applications [Electronic resource] / H. Marzbani, H. Marateb, M. Mansourian // Basic and clinical neuroscience journal. - 2016. - Vol. 7, no. 2. - Mode of access: <https://doi.org/10.15412/j.bcn.03070208> (date of access: 17.06.2025). - Title from screen.
5. Hochreiter S. Long Short-Term Memory [Electronic resource] / Sepp Hochreiter, Jürgen Schmidhuber // Neural computation. - 1997. - Vol. 9, no. 8. - P. 1735-1780. - Mode of access: <https://doi.org/10.1162/neco.1997.9.8.1735> (date of access: 17.06.2025). - Title from screen.
6. An enhanced ARIMA model for EEG classification [Electronic resource] / Yan Liu [et al.] // WI-IAT '21: IEEE/WIC/ACM international conference on web intelligence, ESSENDON VIC Australia. - New York, NY, USA, 2021. - Mode of access: <https://doi.org/10.1145/3498851.3498955> (date of access: 17.06.2025). - Title from screen.
7. Sutskever I. Sequence to sequence learning with neural networks [Electronic resource] / Ilya Sutskever, Oriol Vinyals, Quoc V. Le. - Mode of access: <https://arxiv.org/abs/1409.3215>. - Title from screen.
8. Kingma D. P. Adam: a method for stochastic optimization [Electronic resource] / Diederik P. Kingma, Jimmy Ba. - Mode of access: <https://arxiv.org/abs/1412.6980>. - Title from screen.
9. Teaching machines to read and comprehend [Electronic resource] / Karl Moritz Hermann [et al.]. - Mode of access: <https://arxiv.org/abs/1506.03340>. - Title from screen.
10. Mukherjee K. A simple dynamic learning rate tuning algorithm for automated training of dnns [Electronic resource] / Koyel Mukherjee, Alind Khare, Ashish Verma. - Mode of access: <https://arxiv.org/abs/1910.11605>. - Title from screen.
11. Why gradient clipping accelerates training: a theoretical justification for adaptivity [Electronic resource] / Jingzhao Zhang [et al.]. - Mode of access: <https://arxiv.org/abs/1905.11881>. - Title from screen.
12. Dropout: a simple way to prevent neural networks from overfitting [Electronic resource] / Nitish Srivastava [et al.] // Journal of machine learning research. - 2014. - Vol. 15, no. 56. - P. 1929-1958. - Mode of access: <http://jmlr.org/papers/v15/srivastava14a.html>. - Title from screen.
13. Understanding and improving early stopping for learning with noisy labels [Electronic resource] / Yingbin Bai [et al.]. - Mode of access: <https://arxiv.org/abs/2106.15853>. - Title from screen.
14. Mocanu I. An INTRODUCTION TO CUDA Programming / Irina Mocanu // Journal of information systems & operations management. - 2008. - Vol. 2. - P. 495-506.