

РОЗРОБЛЕННЯ МОДЕЛІ МАШИННОГО НАВЧАННЯ ДЛЯ ДИФЕРЕНЦІЙНОЇ ДІАГНОСТИКИ ТРАНЗИТОРНИХ ВТРАТ СВІДОМОСТІ СИНКОПАЛЬНОГО ТА НЕСИНКОПАЛЬНОГО ПОХОДЖЕННЯ У ДІТЕЙ

Т. А. Ковальчук, О. Р. Боярчук, С. Є. Богай

*Тернопільський національний медичний університет імені І. Я. Горбачевського
Міністерства охорони здоров'я України*

Перенесений епізод транзиторної втрати свідомості (ТВС) у дитячому віці досі залишається складною діагностичною дилемою для цілого ряду клініцистів. Дослідження спрямовано на створення ефективної моделі машинного навчання для диференційної діагностики ТВС синкопального та несинкопального генезу, що забезпечить лікарям-педіатрам, дитячим кардіоревматологам і неврологам високу точність диференційної діагностики синкопальних і несинкопальних ТВС. Обстежено 140 пацієнтів із синкопе та 58 – із несинкопальними причинами ТВС віком 8-17 років. Для побудови моделі машинного навчання використовували алгоритм XGBoost. Про ефективність запропонованої моделі машинного навчання свідчили метрики точності, влучності, чутливості, f1-міри, специфічності, ROC AUC і PR AUC. Найінформативнішими показниками моделі виявились Modified Calgary Syncope Seizure Score, швидкість ранкового підвищення систолічного артеріального тиску, серцевий індекс, ранкове підвищення діастолічного артеріального тиску, частота епізодів тахікардії упродовж 24-годинного моніторингу за Холтером, нічне зниження діастолічного артеріального тиску, загальний периферичний опір судин, вік дитини, варіабельність добового діастолічного артеріального тиску, рNN50, співвідношення LF/HF, перцентильний розподіл відносно зростання, діастолічного артеріального тиску та індексу маси тіла. Отже, застосування описаної моделі машинного навчання дає змогу диференціювати синкопальні та несинкопальні ТВС у дітей та може використовуватися педіатрами, дитячими кардіоревматологами та неврологами як додатковий інструмент поряд із класичними діагностичними критеріями синкопе (вазовагальних синкопе, синкопе внаслідок ортостатичної гіпотензії та кардіогенних синкопе) і несинкопальних причин розвитку ТВС (епілепсії, первинного неспровокованого епілептичного нападу, психогенних псевдосинкопе, психогенних неепілептичних судом) на першому етапі діагностики.

Ключові слова: синкопе, несинкопальні транзиторні втрати свідомості, диференційна діагностика, модель XGBoost, діти.

DEVELOPMENT OF A MACHINE LEARNING MODEL FOR DIFFERENTIAL DIAGNOSIS OF SYNCOPAL AND NON-SYNCOPAL TRANSIENT LOSS OF CONSCIOUSNESS IN CHILDREN

T. A. Kovalchuk, O. R. Boyarchuk, S. Ye. Bogai

Ivan Horbachevsky Ternopil National Medical University of the Ministry of Health of Ukraine

Background. The occurrence of a transient loss of consciousness (TLOC) episode in childhood presents a complex diagnostic challenge for clinicians. This study aimed to develop an effective machine learning model for the differential diagnosis of syncopal and non-syncopal TLOC, providing pediatricians, pediatric cardiologists, rheumatologists, and neurologists with high diagnostic accuracy.

Materials and methods. A cohort of 140 patients aged 8-17 years with syncope and 58 with non-syncopal TLOC were examined. The XGBoost algorithm was employed to construct the machine learning model.

Results. The proposed machine learning model demonstrated high effectiveness, evidenced by metrics with 95 % confidence interval: accuracy (0.90; CI 0.80-0.99), precision (syncope: 0.92, CI 0.85-1.0; non-syncope: 0.83, CI 0.72-0.95), recall (syncope: 0.92, CI 0.85-1.0; non-syncope: 0.83, CI 0.72-0.95), f1-score (syncope: 0.92, CI 0.85-1.0; non-syncope: 0.83, CI 0.72-0.95), specificity (syncope: 0.83, CI 0.72-0.95; non-syncope:

0.92, CI 0.85-1.0), ROC AUC (0.96, CI 0.90-1.00) and PR AUC (0.86, CI 0.75-0.96). The most informative indicators of the model are the Modified Calgary Syncope Seizure Score, the rate of morning systolic blood pressure increase, cardiac index, rate of morning diastolic blood pressure increase, frequency of tachycardia episodes during 24-hour Holter monitoring, nocturnal diastolic blood pressure decrease, total peripheral vascular resistance, child's age, diastolic blood pressure variability, pNN50, LF/HF ratio, and percentile distribution relative to height, diastolic blood pressure, and body mass index.

Conclusions. The application of this machine learning model enables the differentiation between syncopal and non-syncopal TLOC in children. It can serve as an additional diagnostic tool for pediatricians, pediatric cardiologists, rheumatologists, and neurologists, complementing traditional diagnostic criteria for syncope (vasovagal syncope, syncope due to orthostatic hypotension and cardiac syncope) and non-syncopal TLOC (epilepsy, first unprovoked epileptic seizures, psychogenic pseudo-syncope, and psychogenic nonepileptic seizure) in the initial diagnostic phase.

Keywords: syncope, non-syncopal transient loss of consciousness, differential diagnosis, XGBoost model, children.

Вступ. Європейська асоціація кардіологів розглядає поняття транзиторної втрати свідомості (ТВС) як стан реальної або очевидної втрати свідомості, що характеризується амнезією впродовж періоду без свідомості, відсутністю нормального рухового контролю, втратою чутливості та короткою тривалістю [1]. Усі нетравматичні ТВС прийнято розділяти на синкопальні (рефлекторні синкопе, синкопе внаслідок ортостатичної гіпотензії, кардіогенні синкопе) та несинкопальні ТВС (епілептичні судоми, психогенні ТВС та інші рідкісні причини). Основною відмінністю цих двох груп ТВС є розвиток церебральної гіпоперфузії, що завжди має місце у патогенезі синкопе та відсутня при несинкопальних ТВС [1, 2].

Незважаючи на те, що впродовж останніх двох десятиліть було досягнуто суттєвого успіху в розумінні різних аспектів розвитку синкопе, перенесений епізод ТВС у дитячому віці досі залишається складною діагностичною дилемою для цілого ряду клініцистів [3, 4]. Ідентифікація пароксизмальної події як епілептичної або неепілептичної є першим кроком у діагностичному процесі, що суттєво ускладнюється непередбачуваністю розвитку клінічних симптомів, які розвиваються головним чином поза межами лікувального закладу [5]. Також, не викликає сумніву і той факт, що епілепсія та істинне синкопе можуть потенціювати одне одного на різних етапах розвитку ТВС [6, 7]. Хоча наявність триггеру в розвитку вазовагального синкопе є одним із діагностичних критеріїв, його від-

сутність не завжди обумовлена несинкопальними причинами розвитку ТВС [8]. До того ж інформація від спостерігачів про особливості перебігу пароксизмальної події у дитини не завжди являється інформативною та діагностично релевантною [9]. Варто зауважити і те, що судоми можуть розвиватися у пацієнтів на фоні кардіогенного генезу синкопе [10], а тахікардія або брадикардія часто виникають не лише при кардіогенних синкопе, але і ТВС епілептичного походження [11]. При цьому саме аритмія виступає одним із гетерогенних факторів у патофізіологічному механізмі розвитку раптової неочікуваної смерті при епілепсії [12, 13]. Усі перелічені фактори являються найчастішими причинами труднощів у постановці вірної причини розвитку ТВС у дитячому віці.

Мета дослідження: створення моделі машинного навчання для диференційної діагностики ТВС синкопального та несинкопального генезу. Розроблена модель має бути ефективною за низкою показників оцінювання та забезпечувати лікарям-педіатрам, дитячим кардіоревматологам і неврологам високу точність диференційної діагностики синкопальних і несинкопальних ТВС.

Матеріал і методи дослідження. Для досягнення мети дослідження нами було обстежено 140 пацієнтів із синкопе (вазовагальні синкопе – 92, синкопе внаслідок ортостатичної гіпотензії – 28, кардіогенні синкопе – 20 дітей) та 58 – із несинкопальними причинами ТВС (епілепсія – 22, первинний неспровокований епілептичний напад

– 18, психогенні псевдосинкопе – 10, психогенні неепілептичні стани – 8 дітей) віком 8-17 років. Для діагностики ТВС застосовували критерії Європейської асоціації кардіологів (2018) [1]. У постановці діагнозу епілепсії використовували рекомендації Міжнародної протиепілептичної ліги (2014) [14]. Пацієнтів із проявами ТВС унаслідок травми голови було вилучено з дослідження.

Усім дітям із ТВС проведено такі клінічні, лабораторні та інструментальні дослідження:

- ретельний збір скарг та анамнезу, включаючи оцінювання за шкалами Calgary Syncope Seizure Score (CSSS) та Modified Calgary Syncope Seizure Score (MCSSS) [15];

- фізикальне обстеження, в т. ч. індекс маси тіла (ІМТ) та тест активного ортостазу [16];

- загальний аналіз крові, глюкоза крові натще;

- інтегральні показники функціонального стану серцево-судинної системи (сistolічний об'єм крові, величина хвилинного об'єму кровотоку, коефіцієнт економічності системи кровообігу, потужність роботи лівого шлуночка серця, об'єм серця, індекс Робінсона, систолічний показник, коефіцієнт функціонального стану системи кровообігу, коефіцієнт функціонального стану кардіореспіраторної системи, коефіцієнт Хільденбранта, величина адаптаційного потенціалу серцево-судинної системи, серцевий індекс, ударний індекс, загальний периферичний опір, питомий периферичний опір) [17];

- електрокардіографія у 12 відведеннях (ритм, електрична вісь серця, тривалість інтервалів R-R, P-Q, Q-T, T-P, S-T, QRS-комплексу, а також вольтаж зубців P, Q, R, S, T);

- ехокардіографія (кінцевий діастолічний та систолічний розміри лівого шлуночка, діаметр правого шлуночка, лівого передсердя, аорти та легеневої артерії, товщина міжшлуночкової перетинки, товщина задньої стінки лівого шлуночка, фракція укорочення лівого шлуночка, кінцевий діастолічний і систолічний об'єми крові, ударний об'єм, хвилинний об'єм крові, фракція викиду лівого шлуночка, відносна товщина лівого шлуночка, маса міокарда лівого шлуночка, індекс маси міокарда лівого шлуночка);

- холтерівське моніторування електричної активності серця впродовж 24 годин (середня ЧСС за день, ніч і добу, циркадний індекс, тривалість епізодів тахікардії, брадикардії, аритмії, екstrasistolії, аналіз атріовентрикулярного проведення, а також тривалості електричної систоли та варіабельності серцевого ритму за часовими (SDANN, RMSSD, pNN50) та частотними (TP, VLF, LF, HF, співвідношення LF/HF) показниками);

- добове моніторування артеріального тиску (середні значення систолічного артеріального тиску (САТ), діастолічного артеріального тиску (ДАТ), середнього артеріального тиску (СрАТ) та пульсового артеріального тиску (ПАТ) за добу, день і ніч, варіабельність артеріального тиску (АТ), нічне зниження АТ, ранкове підвищення АТ (величина та швидкість), показники «навантаження тиском» (індекс часу гіпертензії, індекс площі під графіком) за добу, день і ніч, індекс резистентності судин, частоту серцевих скорочень за добу, день і ніч; ступені нічного зниження АТ (dippers, non-dippers, over-dippers, night-peakers);

- електроенцефалографія;

- консультації спеціалістів – невролога, кардіолога, окуліста.

У роботі дотримані етичні принципи стосовно людей, які виступають суб'єктами дослідження, з урахуванням основних положень Гельсінської декларації Всесвітньої медичної асоціації з біомедичних досліджень, у яких людина виступає їх об'єктом (World Medical Association Declaration of Helsinki, 1964, 2000, 2008), Конвенції Ради Європи про права людини й біомедицину (2007). Діти усіх груп дослідження та їх батьки надали інформовану згоду на проведення дослідження.

Процес розроблення моделі машинного навчання для диференційної діагностики синкопе у дітей складався з п'яти послідовних етапів: (1) вибору типу моделі машинного навчання, (2) підготовки даних, (3) розподілу даних на навчальний і тестовий набори, (4) навчання моделі на навчальному наборі даних, (5) оцінювання ефективності запропонованої моделі на тестовому наборі даних.

Вибір типу моделі машинного навчання.

Відповідно до характеру набору даних та специфіки завдань у цьому дослідженні, було обрано модель XGBoost [18, 19]. Для завдання бінарної класифікації диференційної діагностики ТВС синкопального та несинкопального генезу у дітей було використано функцію втрат «binary:logistic», що використовує логістичну функцію для обчислення ймовірностей приналежності до класу.

Підготовка даних. Спершу всі клінічні та інструментальні показники 140 пацієнтів із синкопе та 58 – із несинкопальними причинами ТВС було завантажено в програмне середовище JupyterLab за допомогою бібліотеки Pandas, перевірено їх на наявність пропущених значень та інші невідповідності, визначено типи ознак. Для кодування категоріальних ознак обрано метод One-Hot Encoding.

Розподіл даних на навчальний і тестовий набори. Розподіл клінічних і лабораторних показників пацієнтів виконано за допомогою функції `train_test_split` бібліотеки `sklearn` із співвідношенням 80 % для навчальних і 20 % для тестових даних. Оскільки набір даних має співвідношення діагнозів «Синкопе» : «Несинкопе» 70 % : 30 %, було використано стратифікований розподіл даних, що дозволяє представити кожну групу таким чином, що результати будуть репрезентативними для всього набору даних.

Навчання моделі на навчальному наборі даних. Для створення моделі диференційної діагностики ТВС використовували пакет XGBoost для Python. Побудову продуктивної моделі здійснили шляхом оптимізації гіперпараметрів за допомогою методу «GridSearchCV» пакету `scikit-learn`. Для запобігання перенавчання моделі застосовували гіперпараметр «`early_stopping_rounds`».

Оцінювання ефективності запропонованої моделі на тестовому наборі даних. Для оцінювання ефективності моделі використовували такі метрики:

1. Точність (Accuracy) – загальна правильність прогнозів моделі. Показник розраховується як співвідношення правильно передбачених діагнозів до загальної кількості результатів у наборі даних;

2. Влучність (Precision) – правильність прогнозів моделі стосовно кожного окремо взятого класу захворювання. Показник розраховується як співвідношення істинно позитивних результатів до суми істинно позитивних і хибно позитивних результатів;

3. Чутливість (Recall) – ймовірність того, що результат буде позитивним у випадку, коли наявне захворювання. Показник розраховується як співвідношення істинно позитивних результатів до суми істинно позитивних і хибно негативних результатів;

4. F1-міра – гармонійне середнє значення влучності та чутливості, що врівноважує обидва показники;

5. Специфічність (Specificity) – ймовірність того, що результат буде негативним, коли хвороба відсутня. Показник розраховується як співвідношення істинно негативних результатів до суми істинно негативних та хибно позитивних результатів;

6. ROC AUC (площа робочої характеристики приймача під кривою) – графік істинно позитивного показника (true positive rate, TPR) проти хибно позитивного показника (false positive rate, FPR) при різних порогових значеннях класифікації;

7. PR AUC (площа точності запам'ятовування під кривою) – графік співвідношення точності (позитивного прогностичного значення) проти повноти (справжньої позитивної частоти) за різних класифікаційних порогів. Дана метрика є показником компромісу між точністю та чутливістю, особливо репрезентативною у ситуаціях, коли захворювання є рідкісним або співвідношення хибно позитивних і хибно негативних результатів різне.

Модель на основі XGBoost також дала змогу визначити показник важливості для кожної ознаки, що використовувалася у побудові дерева рішень моделі.

XGBoost це алгоритм посилення градієнта, що створює ансамбль дерев рішень для прогнозування цільової змінної. Алгоритм працює шляхом повторного додавання дерев рішень до ансамблю, при цьому кожне нове дерево намагається випра-

вити помилки попередніх дерев. Під час навчання алгоритм оптимізує функцію втрат, що вимірює різницю між прогнозованими та фактичними значеннями цільової змінної. У нашому випадку були створені чотири дерева рішень, що в сумі дали змогу диференціювати синкопальні та несинкопальні ТВС у дітей.

Листові значення (leaf) моделі XGBoost відповідали балам, призначеним кожному кінцевому вузлу в деревах рішень. Ці бали було використано для розрахунку кінцевого прогнозованого значення для кожної вхідної вибірки. Для того, щоб отримати ймовірність приналежності до класу для кожного окремого випадку, було підсумовано листові значення всіх дерев, після чого дану суму використано в сигмоїдній функції:

$$\sigma(x) = 1 / (1 + e^{-x}), \quad (1)$$

де x – сума листових значень.

Результати та їх обговорення. Станом на сьогодні не залишає сумніву той факт, що не існує чітких діагностичних критеріїв, які би допомогли клініцистам розрізняти причини ТВС у закладах первинної або невідкладної медичної допомоги [3, 4]. Ефективність діагностики ТВС можна покращити шляхом поєднання симптомів кожного окремо взятого захворювання, але жодна із таких комбінацій досі не була підтверджена проспективно як діагностичний критерій золотого стандарту. Ряд неепілептичних пароксизмальних явищ, таких як синкопе та психогенні неепілептичні стани, можуть імітувати епілептичні припадки та спричиняти діагностичні труднощі [20, 21]. Саме тому пошук нових методів диференціації ТВС синкопального та несинкопального походження є одним із актуальних і пріоритетних напрямів дослідження.

Раніше нами було створено алгоритм машинного навчання XGBoost для диференційної діагностики синкопе у дитячому віці [27]. Зазначена модель машинного навчання володіє достатньою ефективністю й може використовуватися педіатрами та спеціалістами з дитячої кардіоревматології для диференційної діагностики видів синкопе – вазовагальних синкопе, синкопе внаслідок ортостатичної гіпотензії та кардіогенних синкопе у дитячому віці.

Однак, вона не дає змогу диференціювати ТВС синкопального та несинкопального походження (епілепсія, первинний неспровокований епілептичний напад, психогенні псевдосинкопе, психогенні неепілептичні стани). Саме тому нами було поставлено за мету створити на основі алгоритму XGBoost додаткову модель, що забезпечить високу точність диференційної діагностики педіатричних синкопальних і несинкопальних ТВС.

Модель машинного навчання на основі алгоритму XGBoost неодноразово довела свою ефективність для застосування у побудові алгоритмів диференційної діагностики та прогнозування ризику розвитку захворювань [22-24]. XGBoost має ряд переваг перед іншими моделями класифікації, такими як логістична регресія, дерево рішень, градієнтний бустинг і нейронні мережі. По-перше, XGBoost надає інтерпретовані відомості про важливість функцій, що можна використовувати для визначення того, які змінні найбільше впливають на прогнози та рішення моделі. Ця функція дозволяє легко зрозуміти як модель робить свої прогнози і може бути корисною для визначення потенційних причин або кореляцій, пов'язаних із синкопальними або несинкопальними ТВС. По-друге, XGBoost не чутливий до масштабування функцій, а тому модель може ефективно працювати без необхідності стандартизувати чи нормалізувати функції. По-третє, XGBoost здатний фіксувати нелінійні зв'язки між змінними, що робить його ефективним інструментом у роботі зі складними наборами даних, де лінійні моделі, такі як логістична регресія, можуть бути недостатньо репрезентативними. По-четверте, XGBoost – це метод ансамблю, що поєднує кілька слабких моделей для створення сильної моделі.

Використання гіперпараметру «early_stopping_rounds» у побудові моделі диференційної діагностики синкопальних і несинкопальних ТВС дозволило визначити кількість раундів підсилення, що потрібно виконати після того, як показник перевірки припинив покращуватися. У даній моделі цей параметр був встановлений на показнику 10. Як видно з табл. 1, на п'ятому раунді підсилення мо-

дель не покращується, а точність 1.00000 свідчить про перенавчання моделі. Тобто слід зупинитися на четвертому раунді підсилення.

Точність моделі диференційної діагностики ТВС синкопального та несинкопального генезу на тестовому наборі даних склала 0,90 (95 % ДІ 0,80-0,99).

Переважає більшість метрик оцінок базується на так званій матриці плутанини – спеціальному макеті таблиці, що дозволяє візуалізувати про-

дуктивність алгоритму. Матриця плутанини для нашої моделі представлена на рис. 1. Тестовий набір даних містив 28 пацієнтів, у яких були синкопальні причини розвитку ТВС, та 12 пацієнтів із ТВС несинкопального походження. Дана модель правильно знаходила 26 із 28 випадків синкопе та 10 із 12 не синкопе. У 2 пацієнтів із діагнозом синкопе вона діагностувала несинкопальну ТВС, а у 2 дітей із несинкопальною ТВС – синкопе.

Таблиця 1

Раунди підсилення у ході побудови моделі диференційної діагностики ТВС синкопального та несинкопального генезу

Раунд п/п	ROC AUC	PR AUC
1	validation_0-auc:0.99515	validation_0-aucpr:0.98922
2	validation_0-auc:0.99845	validation_0-aucpr:0.99635
3	validation_0-auc:0.99796	validation_0-aucpr:0.99513
4	validation_0-auc:0.99942	validation_0-aucpr:0.99861
5	validation_0-auc:1.00000	validation_0-aucpr:1.00000

Примітка. Обрана оцінка eval_metric=[«auc», «aucpr»]

Для даної моделі на тестовому наборі даних у групі синкопальних ТВС отримано такі метрики з 95 % рівнем довіри: влучності – 0,92 (ДІ 0,85-1,0), чутливості – 0,92 (ДІ 0,85-1,0), f1-міри – 0,92 (ДІ 0,85-1,0) та специфічності – 0,83 (ДІ 0,72-0,95). Для несинкопальних ТВС зазначені метрики з 95 % рівнем довіри були такими: влучність – 0,83 (ДІ 0,72-0,95), чутливість – 0,83 (ДІ 0,72-0,95), f1-міра – 0,83 (ДІ 0,72-0,95), специфічність – 0,92

(ДІ 0,85-1,0). Оцінки ROC AUC із 95 % рівнем довіри (0,96; ДІ 0,90-1,00) і PR AUC (0,86; ДІ 0,75-0,96) для тестового набору даних представлені на рис. 2 та 3. Отримані результати свідчать про те, що розроблена нами модель машинного навчання на основі алгоритму XGBoost є ефективним інструментом диференційної діагностики ТВС синкопального та несинкопального походження у дітей.

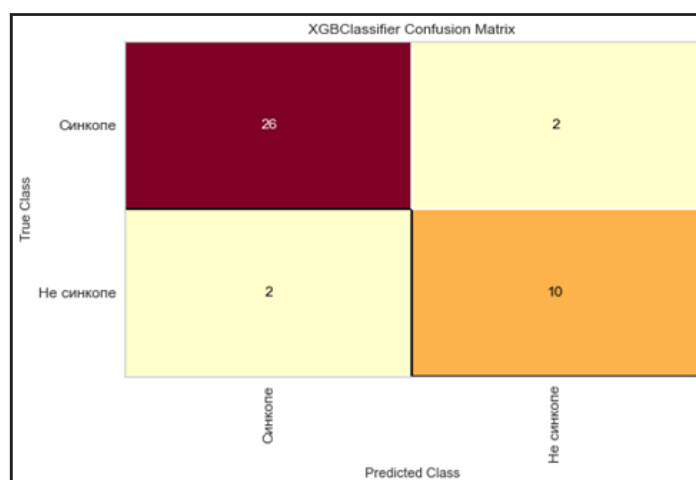


Рис. 1. Матриця плутанини для моделі диференційної діагностики синкопальних і несинкопальних ТВС на тестовому наборі даних

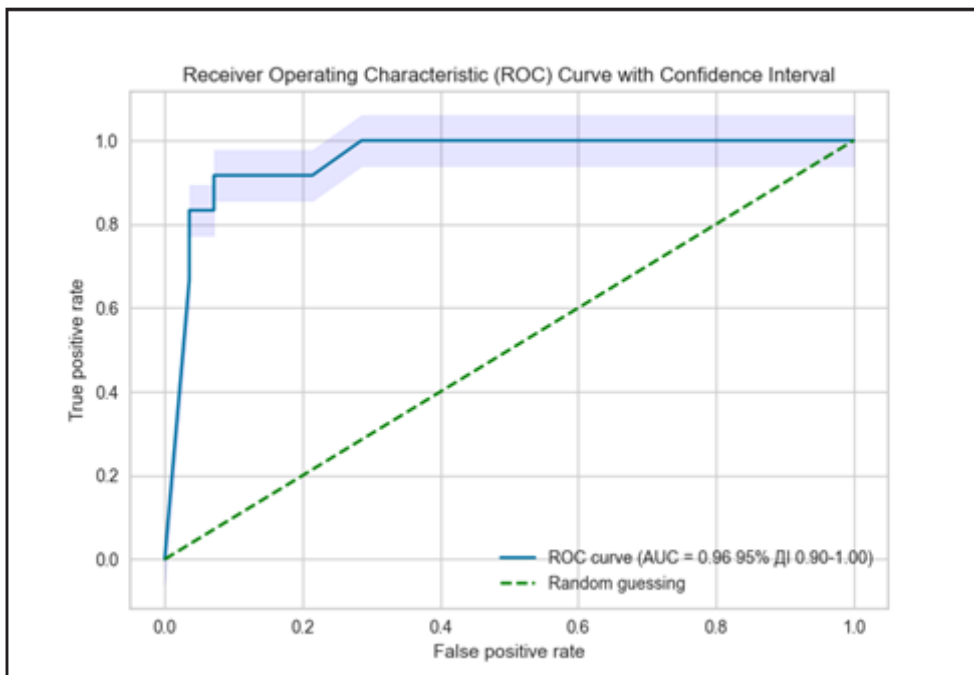


Рис. 2. ROC AUC на тестовому наборі даних моделі диференційної діагностики синкопальних і не-синкопальних ТВС

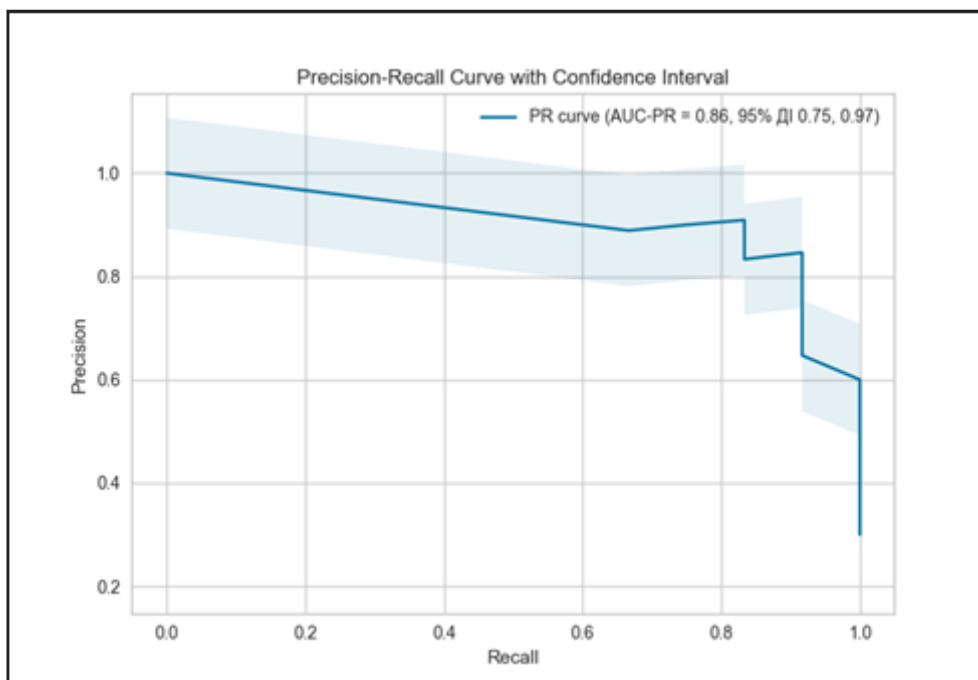


Рис. 3. PR AUC на тестовому наборі даних моделі диференційної діагностики синкопальних і не-синкопальних ТВС

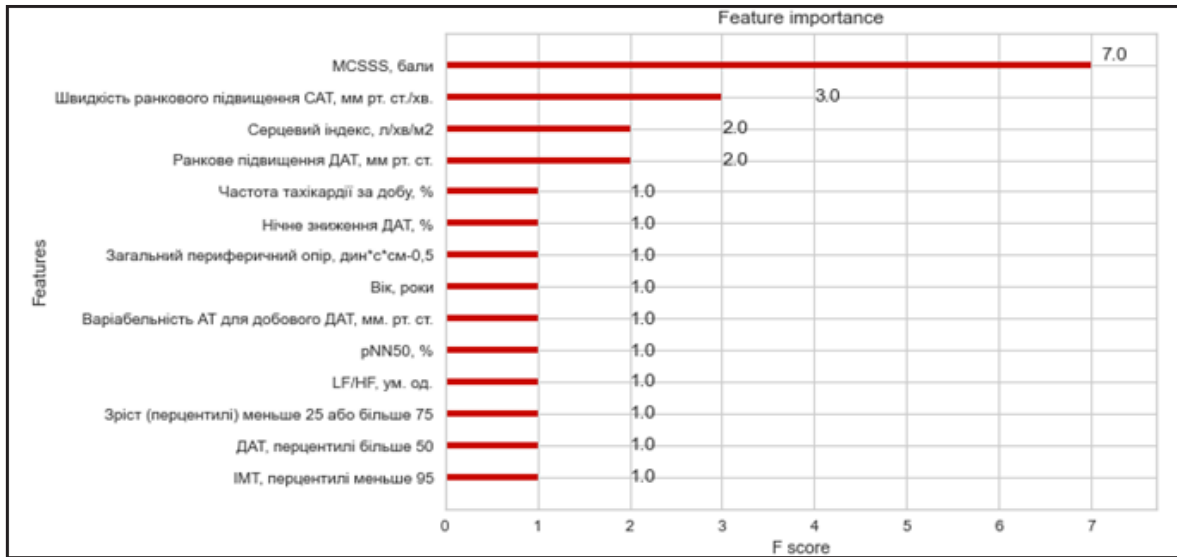


Рис. 4. Важливість ознак у моделі диференційної діагностики синкопальних і несинкопальних ТВС

Публікації про застосування методів машинного навчання у диференційній діагностиці ТВС у дітей досить обмежені. Так, Changjian Li, et al. на основі оцінювання клінічних симптомів розробили модель диференційної діагностики психогенних псевдо-синкопе та вазовагальних синкопе. Чутливість, специфічність і точність цієї моделі підрахунку балів у наборі зовнішньої валідації даних становили 73,7 %, 93,0 % і 87,7 % відповідно [25]. Ouyang C.S., et al. запропонували модель на основі XGBoost для диференціації неепілептичних

пароксизмальних явищ та епілепсії. Результати показали, що точність, площа під кривою, істинно позитивний та істинно від'ємний показник становили 85,17 %, 87,54 %, 89,98 % і 81,81 % відповідно [26].

Важливість ознак у запропонованій нами моделі відображена на рис. 4. Найбільшою інформативністю у диференційній діагностиці синкопальних і несинкопальних ТВС володіли показники MCSSS, швидкість ранкового підвищення САТ, серцевий індекс і ранкове підвищення ДАТ. Дещо менш інформативними, проте важливими показниками

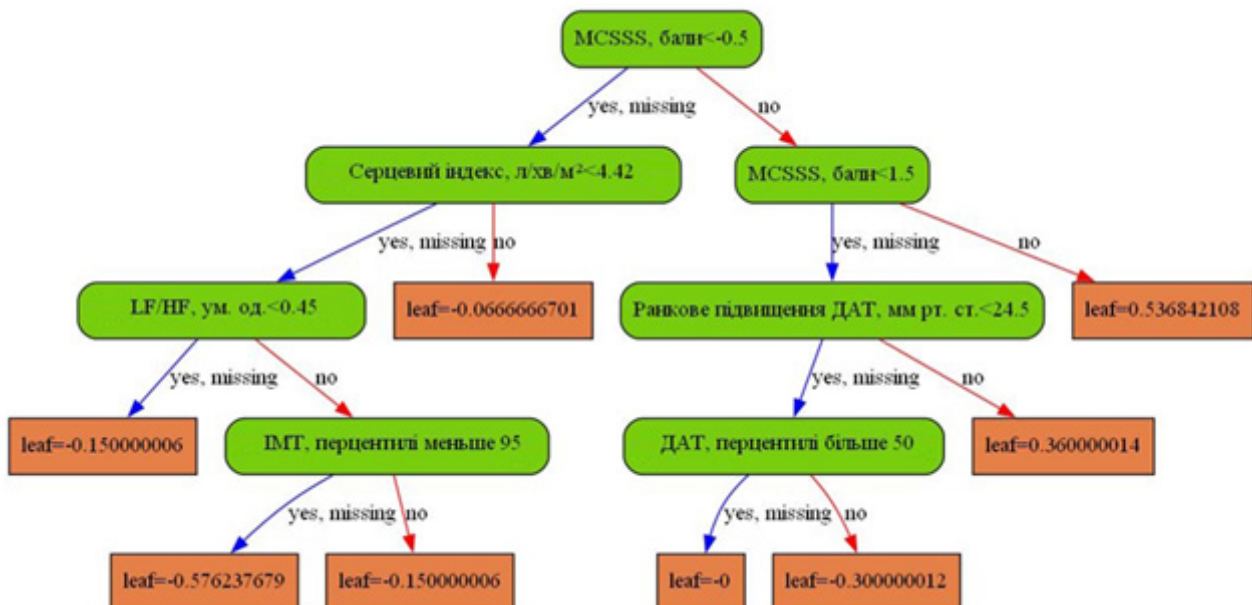


Рис. 5. Перше дерево моделі машинного навчання для диференційної діагностики ТВС синкопального та несинкопального походження



Рис. 6. Друге дерево моделі машинного навчання для диференційної діагностики ТВС синкопального та несинкопального походження

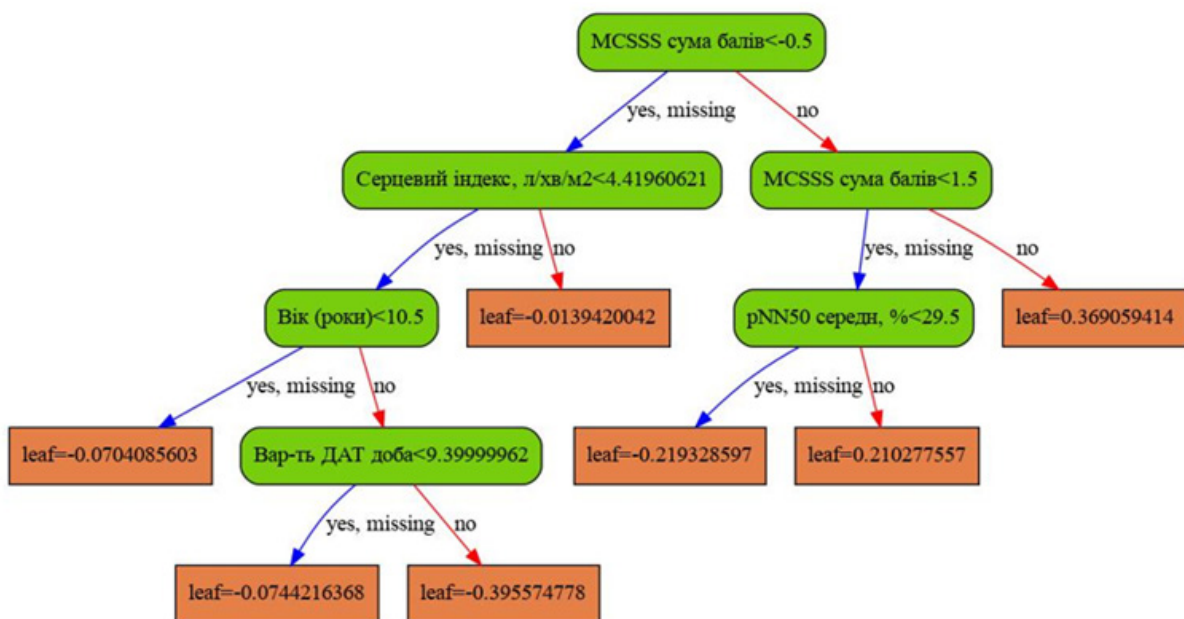


Рис. 7. Третє дерево моделі машинного навчання для диференційної діагностики ТВС синкопального та несинкопального походження

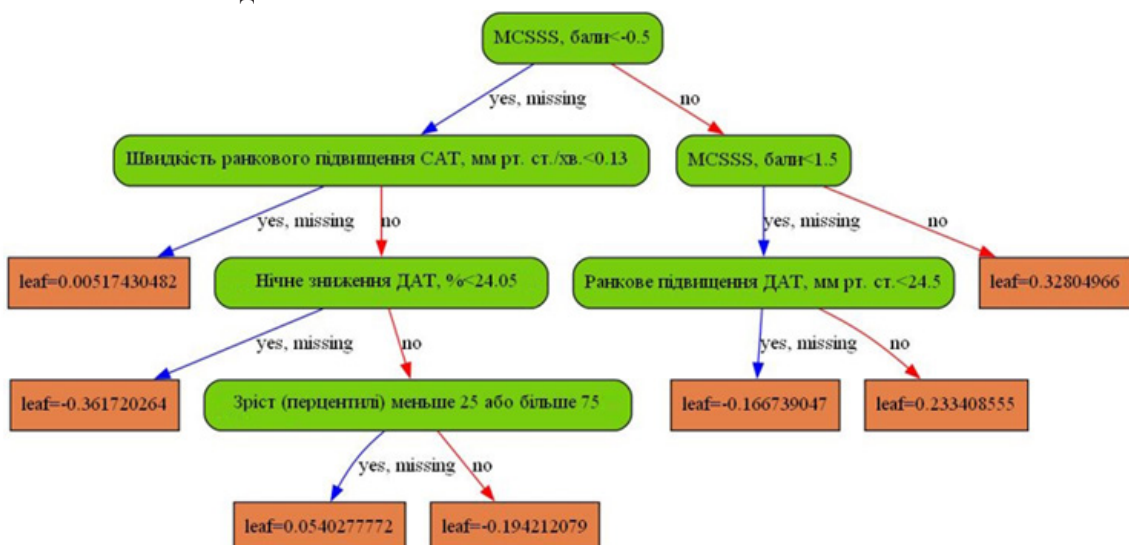


Рис. 8. Четверте дерево моделі машинного навчання для диференційної діагностики ТВС синкопального та несинкопального походження

диференціації ТВС є частота епізодів тахікардії впродовж 24-годинного моніторингу за Холтером, нічне зниження ДАТ, загальний периферичний опір судин, вік дитини, варіабельність добового ДАТ, рNN50, співвідношення LF/HF, перцентильний розподіл стосовно зростання, ДАТ та ІМТ.

У результаті використання пакету XGBoost для Python було створено модель машинного навчання для диференційної діагностики синкопальних і несинкопальних ТВС, що складається з чотирьох дерев (рис. 5-8).

Наведемо приклади практичного використання вищеприписаної моделі для диференційної діагностики синкопальних і несинкопальних ТВС. Для цього застосуємо нашу модель до клінічних і лабораторно-інструментальних показників (табл. 2)

однієї дитини з діагнозом синкопальної ТВС (приклад 1) та однієї – з несинкопальною ТВС (приклад 2).

Розрахунок першого прикладу для дитини з діагнозом синкопе.

1. Leaf першого дерева становить -0.576237679;
2. Leaf другого дерева становить -0.455511838;
3. Leaf третього дерева становить -0.3695574778;
4. Leaf четвертого дерева становить 0.0540277772;
5. Сума Leaf для чотирьох дерев становить -1.37329648758.

Результат сигмоїдної функції: запропонована модель дає прогноз із вірогідністю 80 %, що у дитини має місце синкопе, та ймовірністю 20 %, що у дитини ТВС несинкопального генезу.

Таблиця 2

Вихідні дані пацієнта із синкопальною ТВС і пацієнта із несинкопальною ТВС, що використовуються для прикладу обрахунку запропонованої моделі машинного навчання

Показник	Діагноз пацієнта	
	Синкопе (приклад 1)	Не синкопе (приклад 2)
ІМТ, перцентилі	5,0-85,0	5,0-85,0
ДАТ, перцентилі	50,0-90,0	90,0-95,0
Зріст, перцентилі	>95,0	5,0-10,0
LF/HF, у.о.	1,2	1,7
MCSSS, бали	-4,0	7,0
рNN50, %	23,0	32,0
Варіабельність ДАТ впродовж доби, мм рт. ст.	16,4	8,4
Вік, роки	15,0	15,0
Загальний периферичний опір судин, дин×с×см-0,5	980,2	1516,9
Нічне зниження ДАТ, %	32,4	6,9
Ранкове підвищення ДАТ, мм рт. ст.	47,0	17,0
Серцевий індекс, л/хв/м ²	4,0	3,4
Тривалість епізодів тахікардії за добу, %	6,0	13,9
Швидкість ранкового підвищення САТ, мм рт. ст./хв.	0,7	0,7

Розрахунок другого прикладу для пацієнта з діагнозом несинкопальної ТВС.

1. Leaf першого дерева становить 0.536842108;
2. Leaf другого дерева становить 0.409509927;
3. Leaf третього дерева становить 0.369059414;
4. Leaf четвертого дерева становить 0.32804966;
5. Сума Leaf для чотирьох дерев становить 1.643461109.

Результат сигмоїдної функції: ймовірність діагнозу несинкопальної ТВС у пацієнта становить 84 % , синкопе – 16 %.

Отже, застосування описаної моделі машинного навчання дає змогу диференціювати синкопальні та несинкопальні ТВС у дітей і може використовуватися педіатрами, дитячими кардіоревматологами та неврологами як додатковий інструмент поряд із класичними діагностичними критеріями синкопе (вазовагальних синкопе, синкопе внаслідок ортостатичної гіпотензії та кардіогенних синкопе) і несинкопальних причин розвитку ТВС (епілепсії, первинного неспровокованого епілептичного нападу, психогенних псевдосинкопе, психогенних неепілептичних судом) на першому етапі діагностики. У разі діагностики ТВС синкопального походження в подальшому пропонуємо використовувати модель машинного навчання для диференційної діагностики вазовагальних синкопе, синкопе

Література.

1. ESC Guidelines for the diagnosis and management of syncope / Brignole M., Moya A., J de Lange F. et al. // *European Heart Journal*. – 2018. – № 39 (21). – P. 1883-1948.
2. ACC/AHA/HRS Versus ESC Guidelines for the Diagnosis and Management of Syncope: JACC Guideline Comparison / Goldberger Z. D., Petek B. J., Brignole M. et al. // *J Am Coll Cardiol*. – 2019. – № 74 (19). – P. 2410-2423.
3. Diagnostic criteria to aid the differential diagnosis of patients presenting with transient loss of consciousness: A systematic review / Wardrope A., Newberry E., Reuber M. // *Seizure*. – 2018. – № 61. – P. 139-148.
4. Prehospital capillary lactate in children differentiates epileptic seizure from febrile seizure,

внаслідок ортостатичної гіпотензії та кардіогенних синкопе у дітей. Результати розробки цієї моделі машинного навчання та метрики її ефективності будуть опубліковані в подальшому.

Висновки. Запропонована нами модель машинного навчання на основі алгоритму XGBoost являється ефективним інструментом диференційної діагностики ТВС синкопального та несинкопального походження у дітей із точністю 0,90, влучністю, чутливістю та f1-мірою 0,92 для синкопе і 0,83 для не синкопе, специфічністю 0,83 для синкопе та 0,92 для не синкопе, ROC AUC 0,96 і PR AUC 0,86. Найінформативнішими показниками моделі є MCSSS, швидкість ранкового підвищення САТ, серцевий індекс, ранкове підвищення ДАТ, частота епізодів тахікардії впродовж 24-годинного моніторингу по Холтеру, нічне зниження ДАТ, загальний периферичний опір судин, вік дитини, варіабельність добового ДАТ, рNN50, співвідношення LF/HF, перцентильний розподіл стосовно зростання, ДАТ та ІМТ, що в сумі, поряд із діагностичними критеріями захворювань, дозволяють покращити диференційну діагностику синкопальних і несинкопальних ТВС у практиці педіатра, дитячого невролога та кардіоревматолога.

syncope, and psychogenic nonepileptic seizure / Brody E. I., Genuini M., Auvin S. et al. // *Epilepsy Behav*. – 2022. – № 127. – P. 108551.

5. How to distinguish seizures from nonepileptic manifestations / Leibetseder A., Eisermann M., LaFrance W.C. Jr. et al. // *Epileptic Disord*. – 2020. – № 22 (6). – P. 716-738.

6. Loss of Consciousness in the Young Child / Villafane J., Miller J. R., Glickstein J. et al. // *Pediatr Cardiol*. – 2021. – № 42 (2). – P. 234-254.

7. Differential Diagnosis of Seizure and Syncope by the Means of Biochemical Markers in Emergency Department Patients / Masoumi B., Mozafari S., Golshani K. // *Int J Prev Med*. – 2022. – № 8 (13). – P. 58.

8. Vasovagal syncope or psychogenic pseudosyncope: a major issue in the differential diagnosis of apparent transient loss of consciousness in children / Liao Y., Du J., Benditt D. G., Jin H. // *Sci Bull (Beijing)*. – 2022. – № 67 (16). – P. 1618-1620.
9. Value of witness observations in the differential diagnosis of transient loss of consciousness / Chen M., Jamnadas-Khoda J., Broadhurst M., et al. // *Neurology*. – 2019. – № 92 (9). – P. e895-e904.
10. Cardiac Syncope with Anoxic Seizure Activity / Talibi S., Douglas C., Pope B. // *Case Rep Emerg Med*. – 2020. – № 8 (2020). – P. 6749382.
11. Cardiac and neuronal HCN channelopathies / Rivolta I., Binda A., Masi A., DiFrancesco J. C. // *Pflugers Arch*. – 2020. – № 472 (7). – P. 931-951.
12. Gene mutations in comorbidity of epilepsy and arrhythmia / Yu C., Deng X. J., Xu D. // *J Neurol*. – 2023. – № 270 (3). – P. 1229-1248.
13. The brain-heart interaction in epilepsy: implications for diagnosis, therapy, and SUDEP prevention / Costagliola G., Orsini A., Coll M. et al. // *Ann Clin Transl Neurol*. – 2021. – № 8 (7). – P. 1557-1568.
14. ILAE official report: a practical clinical definition of epilepsy / Fisher R. S., Acevedo C., Arzimanoglou A. et al. // *Epilepsia*. – 2014. – № 55 (4). – P. 475-82.
15. Calgary score and modified Calgary score in the differential diagnosis between neurally mediated syncope and epilepsy in children / Zou R., Wang S., Zhu L. et al. // *Neurological sciences*. – 2017. – № 38 (1). – P. 143-149.
16. Task Force of Clinical Guidelines for Child Orthostatic Dysregulation, Japanese Society of Psychosomatic Pediatrics. Japanese clinical guidelines for juvenile orthostatic dysregulation version 1 / Tanaka H., Fujita Y., Takenaka Y. et al. // *Pediatrics international*. – 2009. – № 51 (1). – P. 169-179.
17. Рівень функціонування адаптаційних механізмів серцево-судинної системи в дітей із синкопе різного генезу / Ковальчук Т. А., Лучишин Н. Ю. // *Сучасна педіатрія. Україна*. – 2022. – № 3 (123). – P. 16-26.
18. XGBoost Model for Chronic Kidney Disease Diagnosis / Ogunleye A., Wang Q. G. // *IEEE/ACM Trans Comput Biol Bioinform*. – 2020. – № 17 (6). – P. 2131-2140.
19. Detection of the chronic kidney disease using XGBoost classifier and explaining the influence of the attributes on the model using SHAP / Raihan M. J., Khan M. A., Kee S. H., Nahid A. A. // *Sci Rep*. – 2023. – № 13 (1). – P. 6263.
20. Ictal autonomic activity recorded via wearable-sensors plus machine learning can discriminate epileptic and psychogenic nonepileptic seizures / Zsom A., Tsekhan S., Hamid T. et al. // *Annu Int Conf IEEE Eng Med Biol Soc*. – 2019. – № 2019. – P. 3502-3506.
21. Psychogenic Non-epileptic Seizures and Pseudo-Refractory Epilepsy, a Management Challenge / Anzellotti F., Dono F., Evangelista G. et al. // *Front Neurol*. – 2020. – № 11. – P. 461.
22. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost / Hou N., Li M., He L. et al. // *J Transl Med*. – 2020. – № 18 (1). – P. 462.
23. Machine learning for the prediction of acute kidney injury in patients with sepsis / Yue S., Li S., Huang X. et al. // *J Transl Med*. – 2022. – № 20 (1). – P. 215.
24. Ultrasound-based radiomics XGBoost model to assess the risk of central cervical lymph node metastasis in patients with papillary thyroid carcinoma: Individual application of SHAP / Shi Y., Zou Y., Liu J. et al. // *Front Oncol*. – 2022. – № 12.
25. Differential Diagnosis Between Psychogenic Pseudosyncope and Vasovagal Syncope in Children: A Quantitative Scoring Model Based on Clinical Manifestations / Li C., Zhang Y., Liao Y. et al. // *Front Cardiovasc Med*. – 2022. – № 9.
26. EEG autoregressive modeling analysis: A diagnostic tool for patients with epilepsy without epileptiform discharges / Ouyang C. S., Yang R. C., Chiang C. T. et al. // *Clin Neurophysiol*. – 2020. – № 131 (8). – P. 1902-1908.
27. Алгоритм машинного навчання XGBoost для диференційної діагностики синкопе у дитячому віці / Ковальчук Т. А., Боярчук О. Р., Богай С. Є. // *Вісник Харківського національного університету імені В. Н. Каразіна. Серія: Медицина*. – 2023. – Вип. 47. – С. 33–46.

References.

1. Brignole, M., Moya, A., de Lange, F. J. et al. (2018). ESC Guidelines for the diagnosis and management of syncope. *Eur Heart J.*, 39(21), 1883-948. doi: 10.1093/eurheartj/ehy037.
2. Goldberger, Z. D., Petek, B. J., Brignole, M. et al. (2019). ACC/AHA/HRS Versus ESC Guidelines for the Diagnosis and Management of Syncope: JACC Guideline Comparison. *J Am Coll Cardiol.*, 74 (19), 2410-2423. doi: 10.1016/j.jacc.2019.09.012.
3. Wardrope, A., Newberry, E., Reuber, M. (2018). Diagnostic criteria to aid the differential diagnosis of patients presenting with transient loss of consciousness: A systematic review. *Seizure*, 61, 139-148. doi: 10.1016/j.seizure.2018.08.012.
4. Brody, E. I., Genuini, M., Auvin, S., Lodé, N., Brunet, S. R. (2022). Prehospital capillary lactate in children differentiates epileptic seizure from febrile seizure, syncope, and psychogenic nonepileptic seizure. *Epilepsy Behav.*, 127. doi: 10.1016/j.yebeh.2021.108551.
5. Leibetseder, A., Eisermann, M., LaFrance, W. C. Jr., Nobili, L., von Oertzen, T. J. (2020). How to distinguish seizures from non-epileptic manifestations. *Epileptic Disord.*, 22 (6), 716-738. doi: 10.1684/epd.2020.1234.
6. Villafane, J., Miller, J. R., Glickstein, J. et al. (2021). Loss of Consciousness in the Young Child. *Pediatr Cardiol.*, 42 (2), 234-254. doi: 10.1007/s00246-020-02498-6.
7. Masoumi, B., Mozafari, S., Golshani, K., Heydari, F., Nasr-Esfahani, M. (2022). Differential Diagnosis of Seizure and Syncope by the Means of Biochemical Markers in Emergency Department Patients. *Int J Prev Med.*, 13, 58. doi: 10.4103/ijpvm.IJPVM_129_20.
8. Liao, Y., Du, J., Benditt, D. G., Jin, H. (2022). Vasovagal syncope or psychogenic pseudosyncope: a major issue in the differential diagnosis of apparent transient loss of consciousness in children. *Sci Bull (Beijing)*, 67 (16), 1618-1620. doi: 10.1016/j.scib.2022.07.024.
9. Chen, M., Jamnadas-Khoda, J., Broadhurst, M. et al. (2019). Value of witness observations in the differential diagnosis of transient loss of consciousness. *Neurology*, 92 (9), e895-e904. doi: 10.1212/WNL.0000000000007017.
10. Talibi, S., Douglas, C., Pope, B. (2020). Cardiac Syncope with Anoxic Seizure Activity. *Case Rep Emerg Med.*, 8. doi: 10.1155/2020/6749382.
11. Rivolta, I., Binda, A., Masi, A., DiFrancesco, J. C. (2020). Cardiac and neuronal HCN channelopathies. *Pflugers Arch.*, 472 (7), 931-951. doi: 10.1007/s00424-020-02384-3.
12. Yu, C., Deng, X. J., Xu, D. (2023). Gene mutations in comorbidity of epilepsy and arrhythmia. *J Neurol.*, 270 (3), 1229-1248. doi: 10.1007/s00415-022-11430-2.
13. Costagliola, G., Orsini, A., Coll, M., Brugada, R., Parisi, P., Striano, P. (). The brain-heart interaction in epilepsy: implications for diagnosis, therapy, and SUDEP prevention. *Ann Clin Transl Neurol.*, 8 (7), 1557-1568. doi: 10.1002/acn3.51382.
14. Fisher, R. S., Acevedo, C., Arzimanoglou, A. et al. (2014). ILAE official report: a practical clinical definition of epilepsy. *Epilepsia*, 55 (4), 475-82. doi: 10.1111/epi.12550.
15. Zou, R., Wang, S., Zhu, L. et al. (2017). Calgary score and modified Calgary score in the differential diagnosis between neurally mediated syncope and epilepsy in children. *Neurol Sci.*, 38 (1), 143-149. doi: 10.1007/s10072-016-2740-5.
16. Tanaka, H., Fujita, Y., Takenaka, Y. et al. (2009). Japanese clinical guidelines for juvenile orthostatic dysregulation version 1. *Pediatr Int.*, 51 (1), 169-79. doi: 10.1111/j.1442-200X.2008.02783.x.
17. Kovalchuk, T. A., Luchyshyn, N. Yu. (2022). The level of functioning of adaptive mechanisms of the cardiovascular system in children with syncope of various genesis. *Modern pediatrics. Ukraine*, 3 (123), 16-26. doi 10.15574/SP.2022.123.16.
18. Ogunleye, A., Wang, Q. G. (2020). XGBoost Model for Chronic Kidney Disease Diagnosis. *IEEE/ACM Trans Comput Biol Bioinform.*, 17 (6), 2131-2140. doi: 10.1109/TCBB.2019.2911071.
19. Raihan, M. J., Khan, M. A., Kee, S. H., Nahid, A. A. (2023). Detection of the chronic kidney disease

using XGBoost classifier and explaining the influence of the attributes on the model using SHAP. *Sci Rep.*, 13 (1), 6263. doi: 10.1038/s41598-023-33525-0.

20. Zsom, A., Tsekhan, S., Hamid, T. et al. (2019). Ictal autonomic activity recorded via wearable-sensors plus machine learning can discriminate epileptic and psychogenic nonepileptic seizures. *Annu Int Conf IEEE Eng Med Biol Soc.*, 3502-3506. doi: 10.1109/EMBC.2019.8857552.

21. Anzellotti, F., Dono, F., Evangelista, G. et al. (2020). Psychogenic Non-epileptic Seizures and Pseudo-Refractory Epilepsy, a Management Challenge. *Front Neurol.*, 11, 461. doi: 10.3389/fneur.2020.00461.

22. Hou, N., Li, M., He, L. et al. (2020). Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost. *J Transl Med.*, 18 (1), 462. doi: 10.1186/s12967-020-02620-5.

23. Yue, S., Li, S., Huang, X. et al. (2022). Machine learning for the prediction of acute kidney injury in patients with sepsis. *J Transl Med.*, 20 (1), 215. doi: 10.1186/s12967-022-03364-0.

24. Shi, Y., Zou, Y., Liu, J. et al. (2022). Ultrasound-based radiomics XGBoost model to assess the risk of central cervical lymph node metastasis in patients with papillary thyroid carcinoma: Individual application of SHAP. *Front Oncol.*, 12. doi: 10.3389/fonc.2022.897596.

25. Li, C., Zhang, Y., Liao, Y. et al. (2022). Differential Diagnosis Between Psychogenic Pseudosyncope and Vasovagal Syncope in Children: A Quantitative Scoring Model Based on Clinical Manifestations. *Front Cardiovasc Med.*, 9. doi: 10.3389/fcvm.2022.839183.

26. Ouyang, C. S., Yang, R. C., Chiang, C. T., Wu, R. C., Lin, L. C. (2020). EEG autoregressive modeling analysis: A diagnostic tool for patients with epilepsy without epileptiform discharges. *Clin Neurophysiol.*, 131 (8), 1902-1908. doi: 10.1016/j.clinph.2020.04.172.

27. Kovalchuk, T. A., Boyarchuk, O. R., Bogai, S. Ye. (2023). XGBoost machine learning algorithm for differential diagnosis of pediatric syncope. *The Journal of V.N. Karazin Kharkiv National University. Series «Medicine»*, 47, 33-46. doi: 10.26565/2313-6693-2023-47-04. [In Ukrainian].

ORCID:

Tetiana A. Kovalchuk: 0000-0003-2455-3278

Oksana R. Boyarchuk: 0000-0002-1234-0040