



УДК 378.016:796.011.3:612.172–057.875 (045):519.8+004.8

ДЕРЕВА РІШЕНЬ ТА ЇХ ЗАСТОСУВАННЯ ДЛЯ КЛАСИФІКАЦІЇ СТУДЕНТІВ РІЗНИХ ГРУП СПОРТИВНО-ПЕДАГОГІЧНОГО ВДОСКОНАЛЕННЯ

Приймак С.Г., к. фіз. вих., доцент,
доцент кафедри біологічних основ фізичного виховання,
здоров'я та спорту

Національний університет «Чернігівський колегіум» імені Т.Г. Шевченка

Заворотинський А.В., к. фіз.-мат. н.,
старший викладач кафедри математики та економіки

Національний університет «Чернігівський колегіум» імені Т.Г. Шевченка

Мета роботи – розроблення моделі дерева рішень залежно від домінування режимів енергозабезпечення реалізації діяльності, яке може бути використане для визначення структури студентів різних груп спортивно-педагогічного вдосконалення залежно від домінування режимів енергозабезпечення реалізації діяльності. Методи дослідження: соматометрія, спектральний аналіз ритму серця, фотоплетизмографія, методи машинного навчання. Студенти диференціюються за групами спортивно-педагогічного вдосконалення відповідно до довжини тіла, відносної потужності I-го навантаження під час виконання проби PWC₁₇₀, сатурації крові киснем у фазу реституції після виконання проби та тривалістю пульсової хвилі в базальних умовах. Виокремлення за групами спортивно-педагогічного вдосконалення дозволяє визначити домінування енергетичних субстратів під час виконання дозованих фізичних навантажень.

Ключові слова: *біопедагогіка, освітній процес, студенти, машинне навчання, дерева.*

Цель работы – разработка модели дерева решений в зависимости от доминирования режимов энергообеспечения реализации деятельности, которое может быть использовано для определения структуры студентов разных групп спортивно-педагогического совершенствования в зависимости от доминирования режимов энергообеспечения реализации деятельности. Методы исследования: соматометрия, спектральный анализ ритма сердца, фотоплетизмография, методы машинного обучения. Студенты дифференцируются по группам спортивно-педагогического усовершенствования в соответствии с длиной тела, относительной мощностью I-й нагрузки при выполнении пробы PWC₁₇₀, сатурации крови кислородом в фазу реституции после выполнения пробы и продолжительностью пульсовой волны в базальных условиях. Разделение по группам спортивно-педагогического усовершенствования позволяет определить доминирование энергетических субстратов при выполнении дозированных физических нагрузок.

Ключевые слова: *биопедагогика, образовательный процесс, студенты, машинное обучение, дерево решений.*

Priymak S.G., Zavorotynskiy A.V. CLASSIFICATION TREE AND THEIR APPLICATION FOR CLASSIFICATION OF STUDENTS OF VARIOUS GROUPS OF SPORTS AND PEDAGOGICAL IMPROVEMENT

The purpose of the work is the development of a decision tree model, depending on the dominance of the energy supply regimes for the realization of activities that can be used to determine the structure of students of different groups of sports and pedagogical improvement, depending on the dominance of the energy supply regimes for the implementation of activities. Methods of research: somatometry, spectral analysis of heart rhythm, photoplethysmography, machine learning methods. These methods of artificial intelligence, quite rightly, have recently been used to provide management of complex cybernetic systems related to the training of specialists in physical education and sports in the process of sports and pedagogical improvement of students. In particular, sophisticated methods of machine learning and intelligent data analysis in physical culture, sports analysis are in place to support decision-making on multidirectional aspects of sporting and pedagogical activity. Students differentiate in groups of SPI according to the body length, the relative strength of the 1st load when performing the PWC₁₇₀ test, the oxygen saturation of the blood into the resuscitation phase after the completion of the test and the duration of the pulse wave in the basal conditions. Separation by the groups of the SPI allows to determine the domination of energy substrates when performing the metered physical.

Key words: *biopedagogics, educational process, students, machine learning, classification tree.*

Постановка проблеми. Штучний інтелект вже використовується в багатьох сферах нашого повсякденного життя. Він передбачає машинне навчання – комплекс технічних методів і інструментів, завдяки

яким комп'ютер може «думати», створювати математичні алгоритми на базі накопичених даних. Застосування складних інструментів аналізу дозволяє виявити особливості, які не можуть бути визначені описовими ста-



тистичними методами інтерпретації даних, має прикладне значення в багатьох галузях науки і техніки для реалізації проблем асоціації, класифікації, сегментації, діагностики і прогнозування [2, с. 2].

Аналіз останніх досліджень і публікацій. Останнім часом сучасні методи штучного інтелекту (машинне навчання) й інтелектуальний аналіз даних застосовуються в забезпеченні управління складними кібернетичними системами, пов'язаними з підготовкою фахівців із фізичного виховання і спорту в процесі спортивно-педагогічного вдосконалення (далі – СПВ) студентів. Зазначимо, що ці методи у фізичній культурі, спортивній аналітиці застосовуються для підтримки ухвалення рішень із різноспрямованих аспектів спортивно-педагогічної діяльності [2, с. 2]. У даних роботах не розглядається структура та моделювання фізичного стану організму студентів різних груп СПВ, які відрізняються метою, біомеханічними параметрами рухів, характером м'язових скорочень, потужністю і тривалістю роботи, механізмами енергозабезпечення, окремими методами штучного інтелекту [8, с. 116]. Обмежена увага науковців до проблеми застосування методів машинного навчання (дерева ухвалення рішень) для класифікації студентів різних груп СПВ в Україні диктує потребу аналізу можливості використання таких методів у зазначеній сфері.

Постановка завдання. Метою дослідження є розроблення моделі дерева рішень залежно від домінування режимів енергозабезпечення реалізації діяльності, яке може бути використане для визначення структури студентів різних груп СПВ.

Матеріали та методи дослідження. Дослідження проведені упродовж грудня 2010 р. – березня 2013 р. на базі лабораторії психофізіології м'язової діяльності Національного університету «Чернігівський колегіум» імені Т.Г. Шевченка. У дослідженнях брали участь студенти, які спеціалізуються в біатлоні ($n = 27$), боксі ($n = 30$) та волейболі ($n = 27$). Всього обстежено 85 спортсменів чоловічої статі, з яких: 38 – спортсмени масових розрядів (I–III розряди), 46 – кандидати в майстри спорту України і майстри спорту України, 5 – заслужені майстри спорту України, майстри спорту міжнародного класу України. Групи досліджуємих сформовані зі студентів, що відвідують відповідну секцію спортивно-педагогічного вдосконалення за видом спорту, які діють на базі факультету фізичного виховання Національного університету «Чернігівський колегіум» імені Т.Г. Шевченка.

Особливості тотальних розмірів тіла спортсменів вивчали згідно зі стандартизованою методикою: реєстрували показники довжини тіла й окремих сегментів (довжини тулуба, корпусу, нижньої та верхньої кінцівок), маси тіла, обводу грудної клітини (далі – ОГК) у спокої, у фазах вдиху і видиху, життєвої ємності легень (далі – ЖЄЛ), сили м'язів кисті і спини [5, с. 326–328]. Систолічний ($AT_{\text{сист.}}$, мм. рт. ст.) та діастолічний артеріальний тиск ($AT_{\text{діаст.}}$, мм. рт. ст.) визначали за допомогою електромеханічного тонометра AND UA-704 (Японія). Особливості вегетативної регуляції серцевого ритму вивчали на підставі аналізу показників ВРС 5–7 хвилинних фрагментів фотоплетизмограми за допомогою монітора серцевого ритму Polar RS800 (Polar Electro, Finland). Дані аналізувалися за допомогою програмного забезпечення Kubios HRV 2.1 (Kuopio, Finland). Артефакти й екстрасистולי видалялися з електронного запису ручним методом. Аналізувалися такі показники ВРС: RRNN, SDNN, RMSSD, pNN₅₀ [6, с. 186; 9, с. 92]. Серед показників спектрального (частотного) аналізу варіабельності ритму серця (далі – ВРС) та кардіоінтервалографії (далі – КІГ) оцінювалися: загальна потужність спектра (Total Power, TP), потужність високочастотного (High Frequency, HF), низькочастотного (Low Frequency, LF) і наднизькочастотного (Very Low Frequency, VLF) компонентів, внесок зазначених компонентів у загальну потужність спектра, а також співвідношення хвиль LF до HF, розрахованих відповідно до абсолютних (мс^2) одиниць (LF/HF ratio, ум. од); M_0 , AM_0 , ΔX , IH (за Р. Баєвським). Досліджувані показники реєструвалися нами відповідно до рекомендацій спільного засідання Європейського товариства кардіологів та Північно-Американського товариства електростимуляції і електрофізіології щодо єдиних стандартів для аналізу варіабельності ритму серця [11, с. 1048].

Судинний тонус, насичення крові киснем визначали за допомогою фотоплетизмографічної методики із застосуванням пульсоксиметра Ohmeda Biox 3700e Puls-Oximeter (Ohmeda, USA), інтегрованого з комп'ютером для тривалого моніторингу пульсової хвилі з можливістю запису, аналізу й інтерпретації результатів. Нами визначалися: SpO_2 (периферична киснева сатурація), %; $T_{\text{пх}}$ (тривалість пульсової хвилі), с; $T_{\text{дф}}$ (тривалість дикротичної фази пульсової хвилі), с; $T_{\text{АФ}}$ (тривалість анакротичної фази пульсової хвилі), с; $T_{\text{ФН}}$ (тривалість фази наповнення), с; $T_{\text{сист}}$ (тривалість систолічної фази серцевого циклу), с; $T_{\text{діаст}}$ (тривалість діастолічної фази серцево-



го циклу), с; $T_{В_{пх}}$ (час відбиття пульсової хвилі), с; $A_{пх}$ (амплітуда пульсової хвилі), ум. од.; $A_{дх}$ (амплітуда дикротичної хвилі), ум. од.; A_1 (амплітуда інцизури), ум. од. На підставі вищезазначених показників розраховувалися: індекс дикротичної хвилі (далі – ІДХ), ум. од.; індекс відбиття (далі – ІВ), %; індекс жорсткості (далі – ІЖ), $m \cdot s^{-1}$; індекс висхідної хвилі (далі – ІВХ), с. [3, с. 32; 9, с. 99]. Параметри пульсової хвилі реєструвалися за допомогою фотоплетизмографічного датчика на дистальній фаланзі 3 пальця лівої кисті в стані спокою сидячі та через 7 хв. після виконання проби PWC_{170} синхронно з параметрами серцевого ритму. Споживання кисню (VO_2 , мл), частоту дихання (далі – ЧД), дих. циклів $Ч \text{ хв.}^{-1}$), дихальний об'єм (далі – ДО), мл, визначали в стані відносного спокою, під час виконання проби PWC_{170} , у фазі реституції визначали за допомогою спірометалобографа «Метатест-1». На підставі добутку ЧД та ДО нами розраховувався хвилинний об'єм дихання (далі – ХОД), мл. Під час реєстрації вищезазначених показників досліджуваного захищали від впливу аудіовізуальних подразників за допомогою світлоізолюючої тканинної маски чорного кольору та звукопоглинальних навушників, які не створювали дискомфорту.

Виконання проби PWC_{170} здійснювалося на велоергометрі ВЭ-02 відповідно до стандартів її виконання [1, с. 165]. У стані спокою, після I та II навантажень, у фазах реституції (через 3 хв. після I та 7 хв. після II навантажень) визначалися вищезазначені показники. Всього отримано 148 показників для кожного з 85 студентів.

Виклад основного матеріалу дослідження. Для побудови моделей класифікації студентів різних груп СПВ використано дерева рішень. Дерева рішень (або дерева класифікацій, регресійні дерева) – тип моделей, що використовуються в машинному навчанні й аналізі даних для прогнозування та класифікації. Дерево складається з набору листя та гілок. На ребрах (гілках) дерева рішень зазначено атрибути, від яких залежить функція втрат, на листі – значення даної функції, а в інших вузлах – атрибути, за якими відрізняються екземпляри. Щоб класифікувати новий екземпляр, потрібно спуститися по дереву, відповідно до його характеристик, до листка і видати значення. Кожен листок являє собою значення цільової змінної, яке залежить від попередніх вузлів за руху від коріння. Дерево може бути навченим через поділ вихідних наборів змінних на підмножини, що базуються на зміні значень вхідних змінних. Такий процес буде повторюватися всере-

дині кожної з отриманих підмножин. Такі рекурсивні дії закінчуються тоді, коли підмножина має ті ж значення цільової змінної, а отже, не додає цінності для прогнозу чи класифікації. Процедура «згори донизу» є прикладом жадібного алгоритму, це найбільш поширена сьогодні стратегія побудови дерев рішень, але є й інші [4].

Основними перевагами дерев рішень є швидкий процес навчання; генерація правил в областях, де експертові важко формалізувати власні знання; формулювання правил природною мовою; інтуїтивно зрозуміла класифікаційна модель; висока точність прогнозування порівняно з іншими методами (статистика, нейронні мережі) тощо [7, с. 99–100].

Нині є багато алгоритмів, що реалізують дерева рішень, проте найбільш поширеними є такі два: CART (Classification and Regression Tree) – алгоритм побудови бінарного дерева рішень – дихотомічної класифікаційної моделі. Кожний вузол дерева під час розбивки має тільки двох нащадків; C4.5 – алгоритм побудови дерева рішень, в якому кількість нащадків вузла необмежена.

Побудова моделі здійснювалася з використанням мови Python (v. 3.6.3 Anaconda custom). Для побудови дерева рішень використовувався алгоритм CART (Classification & Regression Trees), реалізований у пакеті scikit-learn (v. 0.19.1) [7, с. 86].

У результаті побудови моделі визначилася впливовість окремих ознак на рівень успішності спортивно-педагогічної діяльності. Виявлено, що із загального об'єму 148 показників відокремлюються 5 найвпливовіших ознак, які з високою вірогідністю диференціюють студентів за групами СПВ (рис. 1), зокрема: довжина тіла (см); відносна потужність I-го навантаження

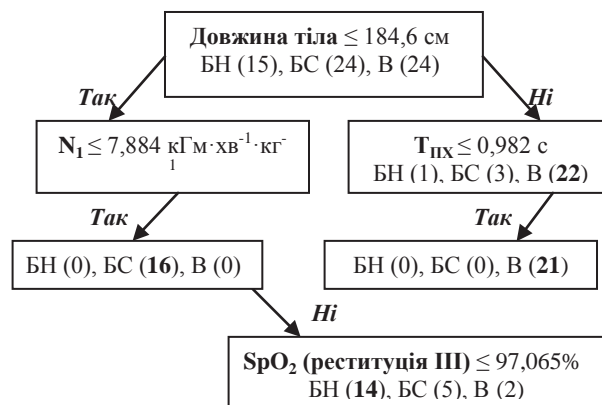


Рис. 1. Дерево розв'язків диференціації студентів за групами СПВ

Перелік умовних позначок: БН – біатлон; БС – бокс; В – волейбол



під час виконання проби PWC_{170} , розрахованої на 1 кг маси тіла досліджуваного (N_1 , $\text{кг}\cdot\text{м}\cdot\text{хв}^{-1}\cdot\text{кг}^{-1}$); сатурація крові киснем (SpO_2 , %) у фазу реституції після виконання проби PWC_{170} ; тривалість пульсової хвилі ($T_{\text{пх}}$, с) у базальних умовах [10, с. 33].

На навчальному наборі правильність класифікації становила 93,7%, на тестовому наборі – 85,7%. Збільшення або зменшення глибини дерева рішень призводить до погіршення його властивості узагальнення.

Детальний аналіз та зумовленість функціонального стану систем організму студентів різних груп СПВ залежно від домінування режимів енергозабезпечення реалізації діяльності розглянуто в роботі авторів [10, с. 33].

Висновки із проведеного дослідження. Аналіз продемонстрував, що дерева рішень можуть бути застосовані для диференціювання студентів за групами СПВ. Вдосконалення класифікації з використанням дерев рішень може відбуватися в кількох напрямках, серед яких розширення інформаційної бази, вдосконалення алгоритмів вибору ознак, а також формування ансамблів дерев рішень. Подальші розвідки спрямовані на розроблення методичної системи розвитку функціональних можливостей майбутніх учителів фізичної культури в процесі СПВ.

Застосована методика аналізу реалізована для визначення рівня інформативності морфофункціональних особливостей організму студентів, які спеціалізуються у волейболі, боксі та біатлоні, відповідно до успішності реалізації професійної діяльності.

ЛІТЕРАТУРА:

1. Белоцерковский З. Эргометрические и кардиологические критерии физической работоспособности у спортсменов. М.: Советский спорт, 2005. 312 с.

2. Бурева В., Стоянов Е. Применение методов искусственного интеллекта в спорте. Актуальные вопросы технических наук: материалы III Междунар. науч. конф. (г. Пермь, апрель 2015 г.). Пермь: Зебра, 2015. С. 1–12.

3. Галкин М., Змиевской Г., Ларюшин А., Новиков В. Кардиодиагностика на основе анализа фотоплетизмограмм с помощью двухканального плетизмографа. М.: Фотоника, 2008. С. 30–35.

4. Дерево прийняття рішень. URL: https://uk.wikipedia.org/wiki/Дерево_прийняття_рішень (дата доступу: 20.06.2018).

5. Иваницкий М. Анатомия человека / под ред. Б. Никитюка, А. Гладышевой, Ф. Судзиловского. 7 изд. М.: Олимпия, 2008. 624 с.

6. Михайлов В. Вариабельность ритма сердца: опыт практического применения метода. Иваново: Ивановская государственная медицинская академия, 2002. 290 с.

7. Мюллер А., Гвидо С. Введение в машинное обучение с помощью Python: руководство для специалистов по работе с данными Москва; Санкт-Петербург, 2017. 472 с.

8. Носко М., Данілов О., Маслов В. Особливості проведення тренувального процесу при заняттях зі студентами у групах спортивного удосконалення: спортивні ігри. Фізичне виховання і спорт у вищих навчальних закладах при організації кредитно-модульної технології: підручн. для каф. фіз. вихов. та спорту ВНЗ. К., 2011. С. 115–134.

9. Приймак С. Спортивно-педагогічне удосконалення студентів: морфофункціональне забезпечення діяльності: монографія. Чернігів: ПАТ «ПВК «Десна»», 2018. 292 с.

10. Приймак С., Заворотинський А. Моделювання фізичного стану організму студентів різних груп спортивно-педагогічного удосконалення. Наукові записки. Серія «Педагогічні науки». Кропивницький: РВВ Центральноукраїнського державного педагогічного університету імені Володимира Винниченка, 2018. С. 32–38.

11. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability. Standards of Measurement. Physiological interpretation and clinical use. Circulation. 1996. V. 93. P. 1043–1065.