

laying is established in work. The actual value of the coefficient of accumulating capacity of the lining of cast-iron buckets was determined, which allowed to establish the cause of the increased losses on accumulation and to outline measures to reduce them.

Keywords: *model: storage capacity, thermal work, linings, cast iron bucket, melt*

Рецензент: доцент, канд. техн. наук Кривенко О. В.

Стаття надійшла

УДК 004.651.4

Федосова І. В., Жуков Д. С.

ПРОГНОЗУВАННЯ ПРОСУВАННЯ БАСКЕТБОЛЬНОЇ КОМАНДИ В ТУРНІРІ З ВИКОРИСТАННЯМ ДЕРЕВА РІШЕНЬ

Прогнозування просування баскетбольної команди в турнірі з використанням дерева рішень. У роботі розглядається побудова дерева рішень для прогнозування просування баскетбольної команди в турнірній таблиці. У дослідженні розглядаються переваги і недоліки дерева рішень, а також проблеми навчання і тестування дерев рішень. У статті описується абсолютно новий підхід для вирішення задачі класифікації даних на основі дерев рішень. Використання нових методів дозволить позбутися від нестачі евристичних алгоритмів, які мають пряме відношення до неправильного вибору критерію поділу класифікації при створенні дерева рішень. Такий новий підхід алгоритму побудови дерева рішень продемонстрував в процесі дослідження більш точну класифікацію даних, а також якість прогнозування результатів. У статті був здійснений прогноз можливостей використання в прикладних цілях дерев рішень для прогнозування економічних і фінансових кризових ситуацій. Для запобігання спаду економіки держави або ж відокремленої компанії або підприємства. Однак перед використанням даних необхідно дотримуватися вимог до форми наданих даних. Також важливо зберегти дані в ієрархічності вигляді. У процесі дослідження автори використовують багатовимірні методи, які є корисними для опису нормативних профілів статистики та їх зв'язку з ситуативними змінами. Статистичний підхід багатовимірних методів з непарними даними є найкращим для визначення статистики, пов'язаної з баскетбольними турнірами. Наведено алгоритм CART класифікації за допомогою дерев рішень, математична модель розрахунку основних вузлів дерева рішень. За допомогою мови програмування Python створений і наведено приклад дерева рішень для поставленого завдання.

Ключові слова: *дерева рішень, класифікація, прогнозування, ентропія, баскетбольна команда, CART.*

Вступ.

До появи інтелектуального аналізу даних спортивні організації в основному покладалися на людський досвід, що виходить від тренерів, скаутів, менеджерів, гравців. Вважалось, що ці фахівці перетворюють історичні записи в корисні знання. Але коли обсяг даних, які вони збирали, ставав все більш і більш величезним, спортивні організації шукали

інші методи, щоб використовувати ті дані, які у них вже були. Методи інтелектуального аналізу спортивних даних можуть сприяти підвищенню продуктивності за рахунок використання історичних записів ігор та об'єднання інформації, пов'язаної з іграми, і тому все більше і більше людей присвячують себе цій галузі.

Національна баскетбольна асоціація (НБА) [1] з моменту свого виникнення налічує понад 67 років. За час зростання цієї організації було сформовано 30 команд, розділених на Східну конференцію і Західну конференцію. У регулярному сезоні зазвичай 82 гри для кожної команди і після сезону по схемі серії best-of-seven. По результатам цих ігор ще 8 команд з кожної конференції виходять у плей-офф. У фіналі вони грають до чотирьох перемог однієї команди. Тобто за чемпіонство відбувається ще від 16 до 28 ігор. Таким чином, за найскромнішими підрахунками, було створено не менше 164820 ігор.

Після кожної зіграної гри НБА збирався масив даних, ці існуючі дані можуть дозволити відкрити цінні знання. Коли люди звертають увагу на свою улюблену команду або гравців, їх обов'язково хвилює результат гри. Однак прогнозування результатів спортивних змагань завжди було складною і привабливою роботою. Зацікавленість баскетболом з кожним роком лише росте, зростає й масиви даних, тому прогнозування перемоги в іграх є актуальною практичною задачею. Крім цього побудова моделі для прогнозування результатів гри може бути використана для інших видів спорту, де використовується така сама схема розподілу ігор.

Аналіз останніх досліджень і публікацій.

У статті [2] наводиться аналіз переваг і недоліків дерев рішень, розглядаються питання навчання і тестування дерев рішень. Особлива увага приділяється проблемам збалансованості навчальної вибірки.

Розглядаються також ліси рішень і методи їх навчання. Наводиться короткий огляд методів зменшення взаємозалежності помилок дерев рішень при навчанні лісів рішень. Пропонуються методи подолання недоліків дерев рішень, наводяться результати роботи даних методів.

У роботі [3] розглядається рішення задачі альтернативного вибору на основі нечітких дерев рішень (нечітких позиційних ігор), особливістю яких є використання нечітких якісних оцінок послідовності рішень і станів природи.

У статті [4] автори розробили та програмно реалізували методи індукції дерева рішень для задачі класифікації політрам на основі ряду біохімічних показників. У роботі вивчаються питання обчислювальної складності запропонованого алгоритму.

У роботі [5] пропонується новий підхід до вирішення задачі класифікації даних на основі дерев рішень, що використовує генетичний метод комбінування евристик. Запропонований еволюційний алгоритм побудови дерева рішень показав кращі в порівнянні з іншими моделями результати при короткостроковому прогнозі.

Автор зазначає, що застосування еволюційних методів дозволяє усунути недолік евристичних алгоритмів, пов'язаний з неправильним вибором критерію поділу при побудові дерева рішення, що, у свою чергу, підвищує точність класифікації даних і якість прогнозування в цілому. Даний підхід дозволив підвищити точність класифікації, зберігши всі переваги методу дерев рішень.

У статті [6] автором був зроблений прогноз можливості застосування дерев рішень для прогнозування фінансової кризи в українських страхових компаніях. Використовуючи сформовані набори ознак, побудовано кілька моделей дерев рішень, серед яких шляхом

порівняння класифікаційних характеристик за два періоди обрано найкращу з позиції чутливості та точності модель. На підставі аналізу автором окреслено шляхи вдосконалення практики застосування методу дерев рішень для передбачення кризових явищ у страхових компаніях України.

У роботі [7] описуються основні проблеми, пов'язані з деревом рішень, і поточні точки дослідження. Розглянуто основні алгоритми побудови дерев індукційної класифікації. Автор наводить дослідження і пропонує можливі комбінації зміщення, які можуть зацікавити дослідників при побудові дерев рішень.

Стаття [8] присвячена опису заявки особливості методу дерева рішень у процесі постановки діагнозу. Автором було проаналізовано основні проблемні області предметної галузі, були визначені та розроблені функції, визначено діагностичну систему. Аналіз рішення завдань показав, що найбільш ефективним є метод побудови «дерева рішень», зокрема дерево класифікації, як підвид методу згідно з типом передбачуваного показника. Дерево рішень було побудовано та автором наведено висновки про доцільність його використання в подальшому.

У роботі [9] розглядається задача створення інформаційної системи для предметної області з ієрархічною організацією інформації. Визначається ряд функцій, які необхідно, щоб система, яка розробляється повинна виконувати. Автором формуються вимоги до форми представлення даних, при цьому значну увагу приділено збереженню ієрархічних даних, їх відтворенню, а також відтворенню їх структури.

Постановка проблеми.

Аналізуючи графічний розподіл команд на ігри в турнірі можна побачити схожість з класичним представленням дерева рішень. Дерева рішень використовуються в повсякденних рішеннях, а не тільки в машинному навчанні. Блок-схеми насправді є візуальними уявленнями дерев рішень. Дерево рішень часто є узагальненням досвіду експертів, засобом обміну знаннями про конкретний процес.

Тим самим візуалізація правил представляє деревоподібну структуру даних. Перевага цього алгоритму в тому, що він легко інтерпретується. Багато інших моделей, хоча і більш точні, не володіють цією властивістю і можуть розглядатися як більш «чорний ящик», де важче інтерпретувати, як вхідні дані були перетворені у вихідні дані. Завдяки цій «зрозумілості» і подібності з прийняттям рішень людиною можна легко пояснити свою модель іншій людині, дерева рішень придбали величезну популярність.

Метою даної роботи є побудова дерев рішень для визначення проходження баскетбольної команди до фіналу в турнірі за допомогою методів машинного навчання, а саме алгоритму класифікації CART. Реалізація здійснюється за допомогою мови програмування Python.

Основний матеріал дослідження.

1. Математична модель побудови дерева рішення

При виборі алгоритму для побудови дерев рішень для прогнозування проходження баскетбольної команди до наступного туру було обрано алгоритм CART. Дерево рішень як алгоритм машинного навчання включає в себе потік логічних правил в формі:

ЯКЩО «властивість значення менше x » ТА «значення властивості b менше, ніж y » ТО категорія 1.

Згідно [10], використання багатовимірних методів корисно для опису нормативних профілів ігровій статистики та їх зв'язку з ситуаційними змінними. Отже, цей статистичний

підхід (багатовимірні методи з непараметричними даними) здається більш відповідним для визначення статистики, пов'язаної з грою, яка дозволяє розрізняти команди тих хто виграє і тих, хто програє на великій вибірці, що включає високу продуктивність неоднорідності. Зокрема, багатовимірний метод дерева класифікації і регресії (CART) був запропонований в якості відповідного статистичного інструменту для дослідження і моделювання таких даних [11]. Оскільки він показав, що він є потужним і надійним методом вибору незалежних змінних, які мають найбільший вплив на залежну змінну.

CART - це непараметричний статистичний аналіз, який ідентифікує взаємовиключні і точні підгрупи вибірки, команди яких мають схожі характеристики, які впливають на залежну змінну (наприклад, на результат гри) [11]. Процедура перевіряє всі можливі і / або розділяють змінні і вибирає одну / кілька, які призводять до розбиття на бінарні групи, що найбільш розрізняються щодо залежною змінною. На рисунку 1 представлено схематично відображення дерева рішень згідно алгоритму CART. Кореневий вузол (нерозділені дані) спочатку розгалужується на два дочірніх вузла відповідно до незалежних змінних.

Усередині кожної гілки при розгалуженні дерево продовжує оцінювати незалежні змінні, що залишилися, щоб визначити, яка змінна дає найкращу розбивку. У момент, коли подальше розділення неможливо, встановлюється кінцевий вузол.

Використовуючи статистику, пов'язану з грою, за допомогою алгоритму CART можна виконати пошук по всьому діапазону значень. Та знаходити явного переможця точок відсікання (тобто конкретних значень статистики, пов'язаних з грою) або категорій (наприклад, ситуаційних змінних) відповідно до встановлених відносин між залежними і незалежними змінними. Однак до теперішнього часу цей метод не використовувався в контексті визначення ключової статистики в спорті, пов'язаної з грою, яка розрізняє які команди виграли, а які програли.

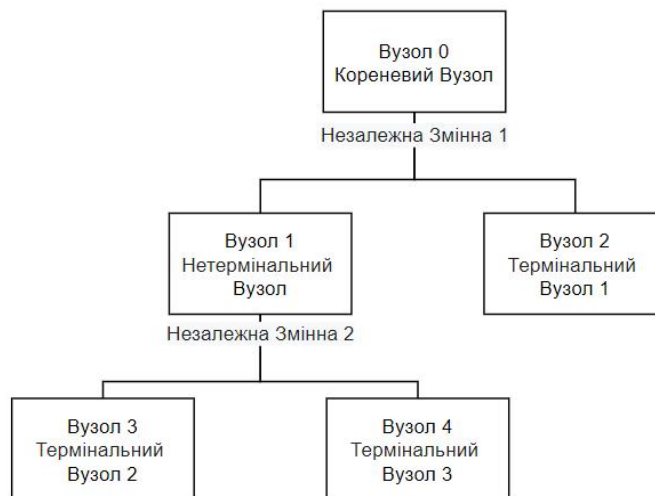


Рисунок 1 – Приклад побудови дерева рішень за допомогою методу CART

Дерева рішень - це спосіб представлення правил в ієрархічній, послідовній структурі, де кожному об'єкту відповідає єдиний вузол, що дає рішення [11]. На ребрах дерева записуються атрибути, від яких залежить цільова функція, в листі - значення цільової функції, а в інших вузлах - атрибути, за якими розрізняються випадки.

За ідею побудови дерев рішень було взято методологію Р. Квінлен [12]. Її інтерпретація полягає в наступному. У разі розподілу даних на тестову та генеральну вибірку, для навчальної вибірки – множини X . Нехай в деякому вузлу дерева сконцентрована множина прикладів \bar{X} , $\bar{X} \in X$. У цьому випадку можливі три ситуації.

1. Множина \bar{X} містить один або більше прикладів, що відносяться до одного класу Y_k . Тоді дерево рішень для \bar{X} – лист, який визначає клас Y_k .

2. Множина \bar{X} не містить жодного прикладу, тобто є порожньою множиною. Тоді це знову лист, і клас, асоційований з листом, вибирається з іншої множини, що асоційована з батьком.

3. Множина \bar{X} містить приклади, які відносяться до різних класів. У такому випадку множину \bar{X} слід розбити на деякі підмножини. Для цього вибирається один з ознак j , що має два і більше відмінних один від одного значень, і \bar{X} розбивається на нові підмножини, кожна з якої містить всі приклади, які мають певний діапазон значень вибраної ознаки. Процедура рекурсивно продовжується до тих пір, доки будь-яка підмножина \bar{X} не буде складатися із прикладів, що відносяться до одного і того ж класу.

Для вибору атрибута розщеплення використовується міра інформативності підпросторів атрибутів, яка ґрунтується на ентропійному методі. Розрахунок міри ентропії [13] дозволяє оцінити розбиття на гілки в дереві рішень. Формула ентропії:

$$S = - \sum_{i=1}^N p_i (\ln p_i) \quad (1)$$

де p_i – ймовірність знаходження системи в i -му стані.

Це дуже важлива концепція, яка використовується у фізиці, теорії інформації та інших областях. Ентропія може бути описана як ступінь хаосу в системі. Чим вище ентропія, тим менше впорядкована система і навпаки. Ентропія допомагає формалізувати «ефективний розподіл даних».

Найпростіша евристика для обробки числових об'єктів у дереві рішень полягає в сортуванні його значень у порядку зростання і перевірки тільки тих порогових значень, в яких змінюється значення цільової змінної.

Крім того, коли в наборі даних є безліч числових ознак, кожне з яких має безліч унікальних значень, вибирається тільки верхній N з порогів, що дає максимальне посилення. Процес полягає в тому, щоб побудувати дерево глибини 1, обчислити ентропію і вибрати кращі пороги для порівняння.

2. Програмна реалізація побудови дерева рішення

Програмна реалізація класифікації команд НБА здійснювалась за допомогою високорівневої мови програмування загального призначення Python. Для програмної реалізації було вирішено побудувати дерево рішень для прогнозування виходу команди в плей-офф. Фрагмент вибірки, та показники для оцінки наведено на рисунку 2.

Усього вибірка нараховує близько 4000 записів, кожний запис характеризується набором показників, розглядаються записи з 1980 року по 2013 рік. Всього у вибірки є 20 показників: `easonEnd` - рік кінця сезону; `Team` - назва клубу; `Playoffs` - чи пройшов клуб в

постсезона (можливість поборотися в турнірі за чемпіонство сезону); W - кількість перемог за сезон; PTS - загальна кількість очок, набраних клубом в іграх сезону; oppPTS - загальна кількість очок, пропущених клубом від опонентів в іграх сезону; FG - загальна кількість реалізованих (забитих) м'ячів; FGA - загальна кількість викинутих м'ячів; 2P - загальна кількість 2-очкових влучань; 2PA - загальна кількість викинутих 2-очкових м'ячів; 3P - загальна кількість 3-очкових влучань; 3PA - загальна кількість викинутих 3-очкових м'ячів; FT - загальна кількість забитих штрафних кидків; FTA - загальна кількість викинутих штрафних кидків; ORB - загальна кількість підборів в нападі (повернення м'яча собі у володіння після промаху команди); DRB - загальна кількість підборів в захисті (право на володіння після промаху команди-суперника); AST - загальна кількість результативних передач (результативна - передача, відразу після якої пішов набір очок команди); STL - загальна кількість перехоплень; BLK - загальна кількість блокшотів; TOV - загальна кількість втрат м'ячів у команди.

```
df = pd.read_csv("C:/NBA_train.csv")
```

```
df.head()
```

	SeasonEnd	Team	Playoffs	W	PTS	oppPTS	FG	FGA	2P	2PA	3P	3PA	FT	FTA	ORB	DRB	AST	STL	BLK	TOV
0	1980	Atlanta Hawks	1	50	8573	8334	3261	7027	3248	6952	13	75	2038	2645	1369	2406	1913	782	539	1495
1	1980	Boston Celtics	1	61	9303	8664	3617	7387	3455	6965	162	422	1907	2449	1227	2457	2198	809	308	1539
2	1980	Chicago Bulls	0	30	8813	9035	3362	6943	3292	6668	70	275	2019	2592	1115	2465	2152	704	392	1684
3	1980	Cleveland Cavaliers	0	37	9360	9332	3811	8041	3775	7854	36	187	1702	2205	1307	2381	2108	764	342	1370
4	1980	Denver Nuggets	0	30	8878	9240	3462	7470	3379	7215	83	255	1871	2539	1311	2524	2079	746	404	1533

Рисунок 2 – Фрагмент вибірки зі статистичними даними по командам НБА

Перед моделюванням будь-яку вибірку даних слід додатково обробити (видалення пустих значень, нормалізація даних). Розбиття всіх даних на тестову та генеральну вибірку виконувалось у співвідношенні 30 % даних для тестової вибірки та 70 % для генеральної вибірки. Розроблена математична модель після її тестування на тестовій вибірці показало точність в 0,729. Спроба покращити показник точності приводить до перенавчання розробленої моделі. При тестуванні основним критерієм розбиття даних було взято міра ентропії. Після навчання тестової вибірки було побудовано дерево рішень для всієї вибірки даних.

При моделюванні генеральної вибірки даних, аналіз відбувався до 2013 року при цьому прогнозування проводилось на наступні роки, для перевірки адекватності моделі. При аналізі отриманого прогнозу можна зробити висновок, що модель адекватна та її прогноз становить 0,642 точності. Дерево було розбито на основні три класи, додаткові значення у разі якщо вони мають не повністю схожу структуру в дереві рішень, відображаються кольором схожим на основний клас. На рисунку 3 наведено фрагмент дерева рішень, а саме розподіл на нульовому рівні, на дві гілки: «правда» та «брехня».

Інформаційні технології

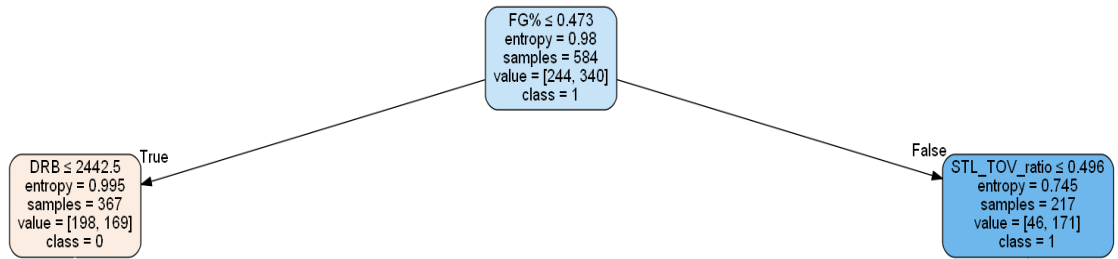


Рисунок 3 – Дерево рішень для розподілу баскетбольної команди в турнірній таблиці, нульовий рівень

На рисунку 4 зображено фрагмент гілки подальшого розбиття гілки «права».

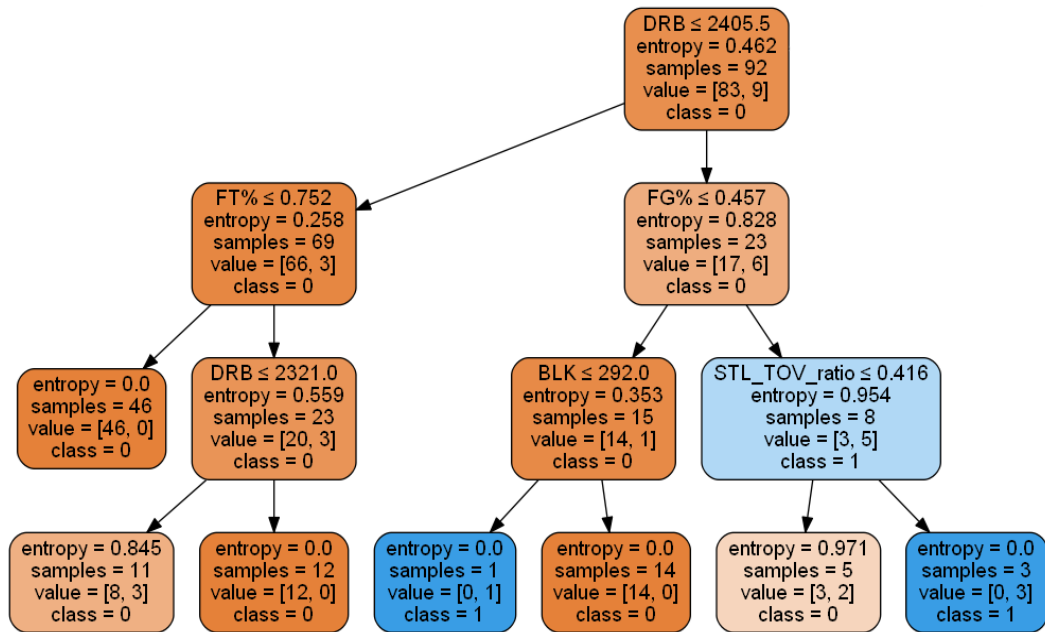


Рисунок 4 – Фрагмент дерева рішень, розгалуження гілки «права»

Варто зазначити, що дерево досить правдоподібно розділяє статистику команд на категорії, порівнюючи в потрібній черговості всі ключові параметри. Тобто, спочатку дерево перевіряє відсоток реалізації кидків з гри. Граничним значенням тут виступає 47,5 %. Адже дійсно, в сучасному баскетболі відсоток реалізації кидків з гри, що дорівнює 50 % вже вважається досить високим показником. Після поділу за кидками, слід перевіряти:

- кількість підбирань команди в захисті (кожен підбір дає право володіння м'ячем команді, чий гравець оформив підбір, що дає можливість команді добрати додаткові очки);
- співвідношення втрат і перехоплень (якщо команда часто втрачає м'яч в нападі, то навіть хороший відсоток кидків і щільна опіка власного щита будуть тільки компенсувати різницю в рахунку, а не давати перевагу);
- кількість блокшотів (їх вплив обумовлено тим, що після блокшота слід можливість оформити підбір у захисті, який, як ми вже знаємо, має вплив на набір очок).

Якщо кожен з цих показників має значення вище середнього, то вже можна робити висновок про те, що команда пройде в турнір постсезона. Це основні показники при аналізі результативності команди при грі.

У разі коли основних даних не достатньо, то слід звертати увагу на другорядні показники:

- кількість результативних передач (збільшення динаміки гри дає більше шансів знайти більш вигідну позицію для набору легких очок в нападі);
- відсоток реалізації штрафних кидків (вони не є основним способом набору очок команди, однак за статистикою 15-25 очок команди в кожній грі припадають на штрафні кидки, що може зіграти свою роль у вирішальному відрізку гри).

ВИСНОВКИ

Проектування системи визначення можливості попадання баскетбольної команди в фінал чи просування її по турнірній таблиці є актуальною задачею. Виконання цієї задачі із використанням класифікації за допомогою дерева рішень, як альтернативного варіанту прогнозування гри, дозволить не лише звичайним вболівальникам робити прогноз та й адміністрації команд своєчасно прокрашувати показники своїх команд та гравців. Спираючись на побудоване дерево рішень було проведено аналіз результатів згідно минулих літ, який показав високу точність збігу з реальними результатами ігор. Але при цьому модель не реалізує та не прогнозує неспортивну поведінку, тобто результати ігор, коли команда спеціально програє супернику.

Список використаних джерел:

1. Tribuna.com [Electronic resource]. – Mode of access: <https://ua.tribuna.com/>
2. Кафтанников, И. Л. Особенности применения деревьев решений в задачах классификации [Электронный ресурс] / И. Л. Кафтанников, А. В. Парасич // Вестник ЮУрГУ. Серия «Компьютерные технологии, управление, радиоэлектроника». – 2015. – Т. 15, № 3. – С. 26–32. – Режим доступа: <https://vestnik.susu.ru/ctcr/article/view/4205>
3. Чернов, В. Г. Нечеткие деревья решений (нечеткие позиционные игры) / В. Г. Чернов // Информационно-управляющие системы. – 2010. – № 5 (48). – С. 8–14.
4. Андрущак, І. Є Програмна реалізація методу індукції дерева рішень для класифікації політрам: питання обчислюваної складності / І. Є. Андрущак, Ю. С. Повстяна // Штучний інтелект. – 2014. – № 2. – С. 100–106.
5. Ржеуцький, А. В. Эволюционный алгоритм построения дерева решений / А. В. Ржеуцький, А. А. Сукончиков // Программные продукты и системы. – 2011. – № 3. – С. 22–24.
6. Литвин А. В. Розробка дерев рішень для прогнозування фінансової кризи в страхових компаніях України / А. В. Литвин // Наукові записки НаУКМА. Економічні науки : зб. наук. пр. – Київ, 2015. – Т. 172. – С. 59–64.
7. Kotsiantis, S. B. Decision trees: a recent overview / S. B. Kotsiantis // Artificial Intelligence Review. – 2013. – Vol. 39, April. – P. 261–283.
8. Басюк, Т. Застосування методу «дерева рішень» для побудови системи встановлення діагнозу / Т. Басюк, І. Дуб'як // Інформація, комунікація, суспільство 2017 : матеріали 6-ої Міжнародної наукової конференції ІКС-2017 (Славське, 18–20 травня

2017 р.). – Львів, 2017. – С. 210–211.

9. *Чирун Л. В.* Інтелектуальний аналіз дерева прийняття рішень в інформаційних системах скринінгового спостереження за імунологічними пацієнтами / *Л. В. Чирун, Є. Я. Лецинський* // Інформаційні системи та мережі. – 2008. – № 610 (1). – С. 233–244.

10. *Moura, F. A.* Analysis of football game-related statistics using multivariate techniques *F. A. Moura, L. E. B. Martins, S. A. Cunha* // Journal of Sports Sciences. – 2014. – Vol. 32, N 20. – P. 1881–1887.

11. *Loh, Wei-Yin* Classification and Regression Trees [Electronic resource] / *Wei-Yin Loh* // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. – 2011. – January. – P. 14–23. – Mode of access: <https://doi.org/10.1002/widm.8>

12. *Quinlan, Ross J.* C4.5 : Programs for Machine learning / *Ross J. Quinlan.* – San Francisco : Morgan Kaufmann Publishers, 1993. – 302 p.

13. *Кочетков, А. В.* О различных смыслах понятия «энтропия» [Электронный ресурс] / *А. В. Кочетков, П. В. Федотов* // Интернет-журнал «НАУКОВЕДЕНИЕ». – 2015. – Т. 7, № 6. – С. 1–13.– Режим доступа:<http://naukovedenie.ru/PDF/72TVN615.pdf>

Федосова И. В., Жуков Д. С.

ПРОГНОЗИРОВАНИЕ ПРОДВИЖЕНИЯ БАСКЕТБОЛЬНОЙ КОМАНДЫ В ТУРНИРЕ С ИСПОЛЬЗОВАНИЕ ДЕРЕВА РЕШЕНИЙ

Прогнозирование продвижения баскетбольной команды в турнире с использованием дерева решений. В работе рассматривается построение дерева решений для прогнозирования продвижения баскетбольной команды в турнирной таблице. В исследовании рассматриваются преимущества и недостатки дерева решений, а также проблемы обучения и тестирования деревьев решений. В статье описывается совершенно новый подход для решения задачи классификации данных на основе деревьев решений. Использование новых методов позволит избавиться от недостатка эвристических алгоритмов, которые имеют прямое отношение к неправильным выбором критерию разделения классификации при создании дерева решений. Такой новый подход алгоритма построения дерева решений продемонстрировал в процессе исследования более точную классификацию данных, а также качество прогнозирования результатов. В статье был осуществлён прогноз возможностей использования в прикладных целях деревьев решений для прогнозирования экономических и финансовых кризисных ситуаций. Для предотвращения спада экономики государства или же обособленной компании или предприятия. Однако перед использованием данных необходимо придерживаться требований к форме предоставленных данных. Также важно сохранить данные в иерархичном виде. В процессе исследования авторы используют многомерные методы, которые являются полезными для описания нормативных профилей статистики и их связи с ситуативными изменениями. Статистический подход многомерных методов с непарными данными является самым подходящим для определения статистики, связанной с баскетбольными турнирами. Приведен алгоритм CART классификации с помощью деревьев решений, математическая модель расчета основных узлов дерева решений. С помощью языка программирования Python создан и приведен пример дерева решений для поставленной задачи.

Ключевые слова: *деревья решений, классификация, прогнозирование, энтропия, баскетбольная команда, CART.*

Fedosova I. V., Zhukov D. S.

CONSTRUCTION OF DECISION TREES FOR DETERMINING THE PASSAGE TO THE FINAL OF A BASKETBALL TEAM USING MACHINE LEARNING

Construction of decision trees for determining the passage to the final of a basketball team using machine learning. The paper considers the construction of a decision tree to predict the progress of a basketball team in the standings. The study examines the advantages and disadvantages of a decision tree, as well as the problems of training and testing decision trees. This article describes a completely new approach to solving the problem of data classification based on decision trees. The use of new methods will allow to get rid of the lack of heuristic algorithms that are directly related to the wrong choice of the classification separation criterion when creating a decision tree. This new approach of the decision tree construction algorithm demonstrated in the process of research a more accurate classification of data, as well as the quality of forecasting results. In the article, a forecast was made of the possibilities of using decision trees for applied purposes for predicting economic and financial crisis situations. To prevent a recession in the economy of the state or a separate company or enterprise. However, before using the data, you must adhere to the requirements for the form of the data provided. It is also important to keep the data in a hierarchical manner. During the research, the authors use multidimensional methods that are useful for describing the normative profiles of statistics and their relationship to situational changes. The statistical approach of multivariate methods with unpaired data is most suitable for determining statistics related to basketball tournaments. An algorithm for CART classification using decision trees, a mathematical model for calculating the main nodes of a decision tree are presented. Using the Python programming language, an example of a decision tree for the task at hand has been created and presented.

Keywords: *decision trees, classification, forecasting, entropy, basketball team, CART.*

Рецензент: **канд. техн. наук**, доц. Проніна О. І.

Стаття надійшла 23.11.2020 р.

УДК 004.93

Левицька Т. О., Король М. Д.

СИСТЕМА ПІДТРИМКИ РІШЕННЯ ПРИ УПРАВЛІННІ ЗАПАСАМИ ДРІБНООПТОВИХ ТОВАРІВ МЕТОДАМИ НЕЧІТКОЇ ЛОГІКИ

Дана стаття присвячена аналізу проблем управління запасами дрібнооптових товарів. Системи нечіткого виводу широко використовуються для вирішення завдань управління різними об'єктами і процесами. При цьому створення нечіткої моделі визначається на основі формальних параметрів досліджуваної системи в термінах лінгвістичних змінних. Однією з проблем управління запасами товарів є існуюча