Облікова картка дисертації

I. Загальні відомості

Державний обліковий номер: 0824U000367

Особливі позначки: відкрита

Дата реєстрації: 15-01-2024

Статус: Наказ про видачу диплома

Реквізити наказу МОН / наказу закладу: № НСВС/21/24 від 01.03.2024

II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Жук Іван Сергійович
2. Ivan S. Zhuk

Кваліфікація:

Ідентифікатор ORCID ID: Не застосовується

Вид дисертації: доктор філософії

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 113

Назва наукової спеціальності: Прикладна математика

Галузь / галузі знань: математика та статистика

Освітньо-наукова програма зі спеціальності: Прикладна математика

Дата захисту: 13-02-2024

Спеціальність за освітою: Прикладна математика

Місце роботи здобувача: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Берестейський, буд. 37, Київ, 03056, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

III. Відомості про організацію, де відбувся захист

Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради): ДФ 26.002.77; ID 3827

Повне найменування юридичної особи: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Берестейський, буд. 37, Київ, 03056, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

IV. Відомості про підприємство, установу, організацію, в якій було виконано дисертацію

Повне найменування юридичної особи: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Берестейський, буд. 37, Київ, 03056, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

V. Відомості про дисертацію

Мова дисертації: Українська

Коди тематичних рубрик: 28.23.25

Тема дисертації:

1. Математичні моделі і методи виявлення за публічно доступними даними підозрілих на фіксований результат футбольних матчів

2. Mathematical models and methods for detecting football match-fixing using publicly available data

Реферат/анотація:

1. Метою роботи є підвищення ефективності виявлення підозрілих на фіксований результат футбольних матчів на базі обробки виключно загальнодоступних публічних даних за результатами сезону футбольного турніру. Однією з найважливіших проблем футболу, що зіставна з проблемою допінгу, є договірні матчі. Результати таких матчів або певний перебіг подій в них є наперед визначеними, тобто фіксованими. У договірних матчах, пов'язаних з виграшом за ставками, завданням є отримання результату, відмінного від очікуваного. Тому такі результати можна розглядати як нетипові, аномальні. Для перевірки поточного матчу на фіксований результат використовують математичні методи футбольної аналітики, такі як: прогнозування результату матчу, аналіз ставок або дій учасників матчу протягом всієї гри. Їх перевагою є оперативність

прийняття рішень, а недоліком □ необхідність використання дуже великої кількості даних, які, зазвичай, не є публічно доступними. Альтернативним може розглядатись підхід, коли рішення щодо фіксованості матчу приймається за результатами усього сезону. При цьому загально доступною є публічна інформація щодо результатів проведених ігор усіх команд, що дозволяє формалізувати пошук матчів, підозрілих на фіксований результат як задачу виявлення контекстуальних аномалій. Найбільш адекватними розглянутій задачі виявлення підозрілих на фіксований результат матчів на основі доступної публічної інформації є статистичні непараметричні гістограмні методи. Це обумовлено тим, що вхідні дані характеризуються малою кількістю дискретних числових значень і їх закони розподілу ймовірностей є невідомими. Водночас, ефективність використання цих методів залежить від об'єму вибірки. Разом з цим, перспективним є математичний апарат конформних предикторів та степеневих мартингалів, який не вимагає знання законів розподілу даних, використовує інформацію про міру неконформності даних та може бути використаний для вирішення задач класифікації даних. Тому актуальною науковою задачею є розробка методів виявлення підозрілих на фіксований результат матчів з використанням апарату конформних предикторів і степеневих мартингалів на базі обробки виключно загальнодоступних публічних даних за результатами сезону футбольного турніру. Наукова новизна одержаних результатів полягає у наступному: 1. Розроблено новий метод виявлення підозрілих щодо фіксованості результату футбольних матчів, який відрізняється від відомих застосуванням конформного аномального детектора із запропонованою мірою неконформності поточного матчу, що забезпечує можливість визначення порогу прийняття рішення у відповідності до заданого значення апріорної ймовірності появи аномальних даних. 2. Розроблено новий метод виявлення підозрілих щодо фіксованості результату футбольних матчів, який відрізняється від відомих застосуванням степеневого мартингалу і правилом прийняття рішення на основі порівняння поточного значення степеневого мартингалу з попереднім, що дозволяє за рахунок зміни параметра чутливості налаштовувати степеневий мартингал на виявлення аномалій відповідного рівня і знаходити їх. 3. Розроблено новий метод виявлення підозрілих щодо фіксованості результату футбольних матчів, який відрізняється від відомих застосуванням інтегрального мартингалу і правилом прийняття рішення на основі порівняння поточного значення інтегрального мартингалу з попереднім, що дає змогу виявляти аномальні матчі без налаштування параметрів. 4. Доведені нові властивості степеневого мартингалу: - за яких завгодно малих значень ступеня конформності (p-value) поточного спостереження значення степеневого мартингала для поточного спостереження є більшим за значення цього ж мартингала для попереднього спостереження; - збільшення значення степеневого мартингала для поточного спостереження по відношенню до попереднього еквівалентно виконанню правила конформного аномального детектора зі значенням рівня аномальності, який дорівнює □^(1/(1-□)), де □ □ параметр чутливості степеневого мартингалу M_k^((□)). 5. Отримала подальший розвиток імітаційна модель футбольного сезону, яка на відміну від існуючих враховує розбиття матчів на класи за контекстуальними атрибутами «сила команди» і «тип гри» □ домашня або виїзна, що забезпечує моделювання договірних матчів з фіксованим результатом, які мають аномальний характер. 6. Удосконалено метод кластеризації на основі Гаусівських сумішей в частині регуляризації недіагональних елементів коваріаційних матриць, що дало змогу зменшити чутливість до початкових умов і отримувати кластери еліпсоподібної форми, які враховують неочевидні зв'язки між точками набору даних.

2. The aim of the thesis is to increase the efficiency of identifying suspects for fixed results of football matches based on the processing of exclusively publicly available public data based on the results of the football tournament season. One of the most important problems of football, comparable to the problem of doping, is fixed matches. The results of such matches or a certain sequence of events in them are predetermined, that is, fixed. In fixed matches related to winning bets, the task is to obtain a result that is different from the expected one. Therefore, such results can be considered atypical, or abnormal. Mathematical methods of football analytics are used to check the current match for a fixed result, such as prediction of the match result, and analysis of bets or actions of the match participants throughout the game. Their advantage is the promptness in decision-making, and their disadvantage is the need to use a very large amount of data, which, as a rule, is not publicly available. As alternative there can be considered an approach where the decision regarding the match fixedness is made based

on the results of the whole season. At the same time, public information about the results of the games played by all teams is publicly available, which allows formalizing the search for matches suspicious of a fixed result as a task of detecting contextual anomalies. Statistical non-parametric histogram methods are the most adequate for the considered task of identifying matches suspicious for a fixed result based on available public information. This is due to the fact that the input data are characterized by a small number of discrete numerical values and their probability distribution laws are unknown. At the same time, the effectiveness of using these methods depends on the size of the sample. Along with this, the mathematical apparatus of conformal predictors and power martingales is promising, which does not require knowledge of data distribution laws, uses information about the degree of non-conformity of data and can be used to solve data classification problems. Therefore, an urgent scientific task is the development of methods for detecting suspicious matches with a fixed result using the apparatus of conformal predictors and power martingales based on the processing of exclusively publicly available public data based on the results of the football tournament season. The scientific novelty of the obtained results is as follows: A new method of detecting football matches suspicious for fixed results has been developed, which differs from the known ones by the use of a conformal anomaly detector with a proposed measure of non-conformity of the current match, which provides the possibility of determining the decision threshold in accordance with the given value of the a priori probability of the appearance of anomalous data. A new method of detecting football matches suspicious for fixed results has been developed, which differs from the known ones by using a power martingale and a decision-making rule based on a comparison of the current value of the power martingale with the previous one, which allows, by changing the sensitivity parameter, to adjust the power martingale to detect anomalies of the appropriate level. A new method of detecting football matches suspicious for fixed results has been developed, which differs from the known ones by the use of an integral martingale and a decision rule based on the comparison of the current value of the integral martingale with the previous one, which makes it possible to detect anomalous matches without adjusting the parameters. New properties of the power martingale are proved: for any small values of the p-value of the current observation, the value of the power martingale for the current observation is greater than the value of the same martingale for the previous observation; increasing the value of the power martingale for the current observation in relation to the previous one is equivalent to fulfilling the rule of the conformal anomaly detector with the value of anomaly threshold equals to □^(1/(1-□)), where □ is the sensitivity parameter of the power martingale M_k^((□)) . The simulation model of the football season got further development, which, unlike the existing ones, takes into account the grouping of matches into classes according to the contextual attributes "team strength" and "type of game" - home or away, which provides simulation of contractual matches with a fixed result, which are anomalous. The method of clustering based on Gaussian mixtures has been improved in terms of regularization of off-diagonal elements of covariance matrices, which made it possible to reduce the sensitivity to initial conditions and to obtain clusters of an elliptical shape that take into account non-obvious connections between the points of the data set.

Державний реєстраційний номер ДіР:

Пріоритетний напрям розвитку науки і техніки: Інформаційні та комунікаційні технології

Стратегічний пріоритетний напрям інноваційної діяльності: Розвиток сучасних інформаційних, комунікаційних технологій, робототехніки

Підсумки дослідження: Нове вирішення актуального наукового завдання

Публікації:

- Zhuk, I., & Chertov, O. (2023). Framework based on conformal predictors and power martingales for detection of fixed football matches. Eastern-European Journal of Enterprise Technologies, 2(4 (122), 6–15; https://doi.org/10.15587/1729-4061.2023.276977
- Chertov, O., & Zhuk, I. (2023). Detection of fixed football matches based on the theory of conformal predictors using the modified Stepanets indicator function. Eastern-European Journal of Enterprise Technologies, 3(4

(123), 22–32; https://doi.org/10.15587/1729-4061.2023.282645

- Чертов О. Р., Жук І. С., “Імітаційна модель футбольного сезону з матчами з фіксованим результатом”, Наукові вісті КПІ, №1–2 , с. 11-35, 2022. doi: https://doi.org/10.20535/kpisn.2022.1-2.287916
- Chertov O.R., Zhuk I.S. Clusterization of soccer season teams using K-means and GMM methods Intelligent Solutions-S: Proceedings of the International Symposium, September 28, 2023, Kyiv-Uzhorod, Ukraine / Ministry of Education and Science of Ukraine, Taras Shevchenko National University of Kyiv and [etc]; Vitaliy Ye. Snytyuk (Editor). 2023, pp. 14-15.
- Жук І. Застосування конформних предикторів і степеневих мартингалів для виявлення підозрілих матчів футбольних турнірів. Прикладна математика та комп'ютинг. ПМК, 2022 : п’ятнадцята наук. конф. магістрантів та аспірантів, Київ, 16—18 лист. 2022 р. : зб. тез доп. / [редкол.: Дичка І. А. та ін.]. — К. : Просвіта, 2022. — 368 с. ISBN 978-617-7010-23-3. С.24 – 29.
- Chertov, O., Zhuk, I., & Serdyuk, A. "Search of the Deviation from the Natural Process Using Stepanets Approach for Classification of Functions," 2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), 2021, pp. 720-724, doi: 10.1109/IDAACS53288.2021.9660997
- Жук І. Виявлення зовнішнього впливу в інформаційних потоках мережі Internet як проблема ідентифікації образу. Філософія і науково-технічна творчість у хронотопі технічного університету: Матеріали ІІІ Міжнародної науково-практичної конференції / Укладачі: Новіков Б. В., Покулита І. К., Колотило М. О., Гавва О. В. – К. : Ліра-К, 2020. – 410 с.; С.144 – 147.
- Жук І. С., Чертов О. Р. Використання математичного апарату наближень Степанця для виявлення штучних втручань у сигналах різної природи. Інтегровані інтелектуальні робототехнічні комплекси (ІІРТК-2020). Тринадцята міжнародна науково-практична конференція 19-20 травня 2020 р., Київ, Україна. – К. : НАУ, 2020. – 305 с.; С.276 – 278.

Наукова (науково-технічна) продукція: методи, теорії, гіпотези; аналітичні матеріали

Соціально-економічна спрямованість: забезпечення промисловості чи населення новим видом інформаційно-комунікаційних послуг

Охоронні документи на ОПІВ:

Впровадження результатів дисертації: Впроваджено

Зв'язок з науковими темами: 0120U102216

VI. Відомості про наукового керівника/керівників (консультанта)

Власне Прізвище Ім'я По-батькові:

1. Чертов Олег Романович

2. Oleh R. Chertov

Кваліфікація: д. т. н., професор, 05.13.06

Ідентифікатор ORCID ID: 0000-0003-0087-1028

Додаткова інформація:

Повне найменування юридичної особи: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Берестейський, буд. 37, Київ, 03056, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

VII. Відомості про офіційних опонентів та рецензентів

Офіційні опоненти

Власне Прізвище Ім'я По-батькові:

1. Виклюк Ярослав Ігорович

2. Vyklyuk Yaroslav I.

Кваліфікація: д. т. н., професор, 01.05.02

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи: Національний університет "Львівська політехніка"

Код за ЄДРПОУ: 02071010

Місцезнаходження: вул. Степана Бандери, буд. 12, Львів, 79013, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Любчик Леонід Михайлович

2. Leonid M. Lyubchyk

Кваліфікація: д. т. н., професор, 05.13.07

Ідентифікатор ORCID ID: 0000-0003-0237-8915

Додаткова інформація:

Повне найменування юридичної особи: Національний технічний університет "Харківський політехнічний інститут"

Код за ЄДРПОУ: 02071180

Місцезнаходження: вул. Кирпичова, буд. 2, Харків, Харківський р-н., 61002, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Рецензенти

Власне Прізвище Ім'я По-батькові:

1. Стіренко Сергій Григорович

2. Sergii G. Stirenko

Кваліфікація: д. т. н., професор, 05.13.05

Ідентифікатор ORCID ID: 0000-0001-5478-0450

Додаткова інформація:

Повне найменування юридичної особи: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Берестейський, буд. 37, Київ, 03056, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Ланде Дмитро Володимирович

2. Dmytro Lande

Кваліфікація: д. т. н., професор, 05.13.06

Ідентифікатор ORCID ID: 0000-0003-3945-1178

Додаткова інформація:
;https://scholar.google.com/citations?user=6VY4hNQAAAAJ&hl;https://www.scopus.com/authid/detail.uri?authorId=19934022600;https://www.webofscience.com/wos/author/record/943541

Повне найменування юридичної особи: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Берестейський, буд. 37, Київ, 03056, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

VIII. Заключні відомості

Власне Прізвище Ім'я По-батькові голови ради Шелестов Андрій Юрійович

Власне Прізвище Ім'я По-батькові головуючого на засіданні Шелестов Андрій Юрійович

Відповідальний за підготовку облікових документів Жук Іван Сергійович

Реєстратор

УкрІНТЕІ

Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності

Юрченко Тетяна Анатоліївна