

ВІДГУК

офіційного опонента на дисертаційну роботу

Жука Івана Сергійовича

на тему «Математичні моделі і методи виявлення за публічно доступними даними підозрілих на фіксований результат футбольних матчів»,
представлену на здобуття ступеня доктора філософії

в галузі знань 11 Математика та Статистика
за спеціальністю 113 Прикладна математика

Актуальність теми дисертації.

1. Матчі з фіксованим результатом є реальною проблемою, яка загрожує основоположним принципам та авторитету футболу у багатьох країнах світу, у тому числі й в Україні. Результати договірних матчах, пов'язаних з виграшем за ставками, можна розглядати як нетипові, аномальні, що дозволяє формалізувати пошук матчів з фіксованим результатом. Застосування математичних методів футбольної аналітики, таких як прогнозування результату матчу, аналізу ставок або дій учасників матчу протягом всієї гри щодо виявлення підозрілих на фіксований результат матчів вимагає великої кількості даних, які не завжди доступні для аналізу.

При використанні загально доступних результатів матчів усього сезону, задача пошуку підозрілих щодо фіксованості результату футбольних матчів може розглядатись як задача виявлення контекстуальних аномалій, що вирішується в галузі інтелектуального аналізу даних. Її особливістю є: дискретнозначність даних, закони розподілу яких невідомі; відсутність розмічених даних для навчання класифікаторів; наявність вибірок даних малого об'єму. Недоліком статистичних непараметричних гістограмних методів пошуку аномалій в дискретнозначних даних є низька ефективність при малій вибірці. В роботі запропоновано для виявлення підозрілих щодо фіксованості результату футбольних матчів використати математичний апарат конформних предикторів та степеневих мартингалів, який не вимагає знання законів розподілу і навчаючої вибірки даних і використовує інформацію про міру неконформності даних.

Тому тема дисертаційної роботи Жука І.С., що пов'язана з теоретичним обґрунтуванням й розробкою методів математичних моделей і методів виявлення за публічно доступними даними підозрілих на фіксований результат футбольних матчів є актуальною.

Оцінка обґрунтованості наукових результатів дисертації, їх достовірності та новизни.

Наукова новизна результатів дисертаційного дослідження полягає в 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 апіорної 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 $\frac{1}{\eta^{1-\eta}}$, де η — параметр чутливості степеневого мартингалу $M_k^{(\eta)}$.

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.

Наукові положення, що висунуті в дисертації, а також висновки і рекомендації автором достатньо обґрунтовані на підставі використання положень і методів математичного аналізу, теорії ймовірностей і математичної статистики, машинного навчання, теорії конформних предикторів, імітаційного моделювання.

Достовірність одержаних результатів підтверджена коректністю поставленої задачі, застосуванням добре апробованих методів проведення дослідження, математичною суворістю проведених теоретичних досліджень, перевіркою розроблених методів шляхом імітаційного моделювання, а також на реальних даних.

Отже, поставлене в дисертаційній роботі наукове завдання розробки методів виявлення підозрілих на фіксований результат матчів з використанням апарату конформних предикторів і степеневих мартингалів на базі обробки виключно загальнодоступних публічних даних виконано повністю, здобувач повною мірою оволодів методологією наукової діяльності.

Оцінка змісту дисертації, її завершеність та дотримання принципів академічної доброчесності.

За своїм змістом дисертаційна робота здобувача Жука І. С. повністю відповідає Стандарту вищої освіти зі спеціальності 113 Прикладна математика та напрямкам досліджень відповідно до освітньої програми 46343 Прикладна математика.

Дисертаційна робота є завершеною науковою працею, а розроблені у роботі методи виявлення підозрілих щодо фіксованості результату футбольних матчів на основі конформного аномального детектора, степеневого мартингалу і інтегрального мартингалу є внеском у розвиток теорії конформного прогнозування.

Розглянувши звіт подібності за результатами перевірки дисертаційної роботи на текстові співпадіння, можна зробити висновок, що дисертаційна робота Жука І. С. не містить елементів фальсифікації, компіляції, фабрикації, плагіату та запозичень. Використані ідеї, результати і тексти інших авторів мають належні посилання на відповідне джерело. Сумнівів щодо самостійного виконання роботи здобувачем немає.

Мова та стиль викладення результатів

Дисертаційна робота написана українською мовою. Вона викладена на 188 сторінках машинописного тексту, складається із вступу, 4 розділів, загальних висновків, списку використаних джерел та 1 додатку. Обсяг основного тексту дисертації складає 165 сторінок. Роботу проілюстровано 21

таблицею та 42 рисунками. Список використаних джерел містить 166 найменувань.

Дисертаційна робота оформлена відповідно до вимог наказу МОН України від 12 січня 2017 р. № 40 «Про затвердження вимог до оформлення дисертації».

У вступі обґрунтовано актуальність теми дослідження, визначено наукову задачу, мету, об'єкт і предмет досліджень, які відповідають темі дисертації, сформульовано задачі і методи дослідження, наукову новизну й практичне значення отриманих результатів.

В першому розділі розкрито стан і зміст задачі виявлення за публічно доступними даними підозрілих на фіксований результат футбольних матчів. Показано, що вона відноситься до класу задач виявлення контекстних аномалій, які вирішуються в області інтелектуального аналізу даних. Проведено аналіз статистичних методів, а також методів машинного навчання, які використовуються для виявлення аномалій. Для вирішення задачі виявлення за публічно доступними даними підозрілих на фіксований результат футбольних матчів запропоновано використати математичний апарат конформних предикторів, степеневих та інтегральних мартингалів.

У наступних розділах дисертації в повному обсязі викладено зміст власних досліджень здобувача. В другому розділі для визначення контекстуального атрибуту «сила команди» проведено групування команд методами K -середніх та Гаусівських сумішей за ознаками кількості очок та різницею між забитими і пропущеними м'ячами в одновимірному та двовимірному просторах. Розроблено імітаційну модель футбольного сезону з матчами з фіксованим результатом та проведено її аналіз шляхом статистичного моделювання. В третьому розділі розроблено методи виявлення підозрілих щодо фіксованості результату футбольних матчів за наявності даних про весь сезон з використанням конформного аномального детектора, степеневого та інтегрального мартингалів. В четвертому розділі шляхом імітаційного моделювання за даними модельного сезону проведено аналіз ефективності розроблених методів та їх порівняння з відомим гістограмним методом. Також розроблені методи використано для виявлення матчів, які вважаються договірними, в сезоні 2014–2015 рр. Серії *B* Італії.

У загальних висновках підсумовано отримані найбільш важливі наукові та практичні результати дисертації, наголошено на якісних і кількісних показниках здобутих результатів, викладено рекомендації щодо їх використання.

Список використаних джерел характеризує високий ступінь вивченості здобувачем проблематики виявлення аномалій, яка вирішується в області інтелектуального аналізу даних.

Дисертацію написано стилістично грамотно на високому науковому рівні. Вона має внутрішню єдність. Застосована в роботі наукова термінологія є загальноновизнаною, стиль викладення теоретичних і практичних досліджень, нових наукових положень, висновків і рекомендацій забезпечує доступність їх сприйняття та використання.

Оприлюднення результатів дисертаційної роботи

Наукові результати дисертації висвітлені у 8 наукових публікаціях в тому числі 3 статті у наукових фахових виданнях України, зокрема 2 статті опубліковано в Східно-Європейському журналі передових технологій (Eastern-European Journal of Enterprise Technologies), який включено до списку міжнародної наукометричної бази Scopus з квантилем Q3.

Результати дисертації Жука І.С. пройшли всебічну апробацію на 5 міжнародних та вітчизняних науково-практичних конференціях

Недоліки та зауваження до дисертаційної роботи.

В якості зауважень щодо змісту дисертації можна визначити наступне:

1. Відсутня формалізована загальна схема етапів дослідження результату матчу на фіксованість, що ускладнює розуміння подальшого застосування отриманих у дисертаційній роботі моделей та методів.
2. При одновимірній кластеризації команд на групи використано лише метод К- середніх, хоча метод Гаусівських сум в двовимірному випадку показав кращу ефективність.
3. В якості основи для імітаційної моделі було обрано розподіл Пуассона, хоча в теорії ймовірності існують і інші розподіли для дискретних випадкових величин, що можуть бути застосовані для даної задачі.
4. В роботі порівняння розроблених методів виявлення підозрілих на фіксований результат матчів було проведено лише з відомим гістограмним методом. Було б доцільно провести порівняння з існуючими методами пошуку аномалій інтелектуального аналізу даних.
5. В роботі було б бажано надати порівняння отриманих результатів із відомими методами та більш конкретні рекомендації щодо застосування запропонованих методів для виявлення підозрілих на фіксований результат матчів в інших видах спорту.
6. На деяких рисунках відсутні відомості про величини, що відображені на осі абсцис чи осі ординат, наприклад, на рис. 1.2, 1.6, 2.1, 2.3, 2.4.

Вважаю, що висловлені зауваження не є визначальними і не зменшують загальну наукову новизну та практичну значимість результатів та не впливають на позитивну оцінку дисертаційної роботи.

Висновок про дисертаційну роботу

Вважаю, що дисертаційна робота здобувача ступеня доктора філософії Жука Івана Сергійовича на тему «Математичні моделі і методи виявлення за публічно доступними даними підозрілих на фіксований результат футбольних матчів» виконана на високому науковому рівні, не порушує принципів академічної доброчесності та є закінченим науковим дослідженням, сукупність теоретичних та практичних результатів якого розв'язує наукове завдання, що має істотне значення для галузі знань 11 Математика та Статистика. Дисертаційна робота за актуальністю, практичною цінністю та науковою новизною повністю відповідає вимогам чинного законодавства України, що передбачені в п.6 – 9 «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії», затвердженого Постановою Кабінету Міністрів України від 12 січня 2022 р. № 44.

Здобувач Жук Іван Сергійович заслуговує на присудження ступеня доктора філософії в галузі знань 11 Математика та Статистика за спеціальністю 113 Прикладна математика.

Офіційний опонент:

професор кафедри систем
штучного інтелекту
Національного університету
«Львівська Політехніка»,
д.т.н., професор



М.П.

«14» січня 2024 року