

## 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.**

Однією з найважливіших проблем сучасних футбольних змагань є договірні матчі, результати яких є наперед визначеними, тобто фіксованими, найчастіше відмінними від очікуваних. Задача виявлення договірних матчів є актуальною як у світі, так і в Україні, оскільки угоди за договірними матчами можуть бути пов'язані з іншими формами злочинної діяльності, включаючи відмивання коштів, торгівлю людьми, ухилення від сплати податків, шахрайство, хабарництво.

Для перевірки поточного матчу на фіксований результат використовують математичні методи футбольної аналітики, такі як прогнозування результату матчу на основі аналізу дій учасників матчу протягом всієї гри, що пов'язано з необхідністю використання великої кількості даних, які, зазвичай, не є публічно доступними. Альтернативою є підхід, коли рішення щодо фіксованості матчу приймається з використанням публичної інформації щодо результатів проведених ігор усіх команд, що дозволяє формалізувати пошук матчів, підозрілих на фіксований результат, як задачу виявлення контекстуальних аномалій.

Для вказаної задачі характерно, що вхідні дані характеризуються малою кількістю дискретних числових значень і їх закони розподілу ймовірностей є невідомими. Найбільш адекватними задачі виявлення підозрілих на фіксований результат матчів на основі доступної публичної інформації є статистичні непараметричні гістограмні методи, які не вимагають знання законів розподілу даних. Також перспективним є математичний апарат конформних предикторів та степеневих мартингалів, який не вимагає знання законів розподілу даних, та може бути використаний для вирішення задач виявлення аномалій.

Тому розробка методів виявлення підозрілих на фіксований результат матчів на базі обробки загальнодоступних публічних даних за результатами сезону футбольного турніру з використанням апарату конформних предикторів і степеневих мартингалів є актуальною науковою задачею.

## **Оцінка обґрунтованості наукових результатів дисертації, їх достовірності та новизни.**

Наукова новизна результатів дисертаційного дослідження полягає в наступному:

- 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.

- 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.

- 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.

- Доведені нові властивості степеневого мартингалу, а саме, встановлено, що 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 детектору.

- 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.

- Удосконалено метод кластеризації на основі Гаусівських сумішей в частині регуляризації коваріаційних матриць, що дало змогу зменшити чутливість до початкових умов і отримувати кластери еліпсоподібної форми, які враховують неочевидні зв'язки між точками набору даних.

Достовірність одержаних наукових результатів обумовлена переконливим обґрунтуванням запропонованого підходу на основі коректного використання

сучасних математичних методів конформних предикторів та степеневих мартингалів, ретельно проведеним статистичним моделюванням та перевіркою адекватності розроблених моделей на реальних даних.

Отже, в дисертаційній роботі поставлене наукове завдання 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 «Прикладна математика».

Дисертаційна робота є завершеною науковою працею і свідчить про наявність особистого внеску здобувача у розвиток теорії конформного прогнозування.

Розглянувши звіт подібності за результатами перевірки дисертаційної роботи на текстові співпадиння, можна зробити висновок, що дисертаційна робота Жука Івана Сергійовича є результатом самостійних досліджень здобувача і не містить елементів фальсифікації, компіляції, фабрикації, плагіату та запозичень. Використані ідеї, результати і тексти інших авторів мають належні посилання на відповідне джерело.

### **Мова та стиль викладення результатів**

Дисертаційна робота написана українською мовою.

Дисертація складається з вступу, 4 розділів, висновків, списку літератури та додатків. Загальний обсяг дисертації 187 сторінок.

У вступі обґрунтовано актуальність роботи, сформульовано її мету та основні завдання, наведено відомості про наукову новизну та практичну значущість результатів роботи.

У першому розділі розглядаються проблеми виявлення договірних футбольних матчів за статистичними даними та обґрунтовується актуальність дослідження, спрямованого на пошук потенційно підозрілих щодо фіксованих результатів матчів на основі обробки виключно загальнодоступних публічних даних. Проведено аналіз статистичних методів, а також методів машинного навчання, які використовуються для виявлення аномалій, зокрема методу класифікації на основі конформних предикторів. Сформульовано мету та наукові завдання дисертаційного дослідження.

Другий розділ присвячений розробці імітаційної моделі футбольного сезону з матчами з фіксованим рахунком та алгоритм моделювання договірних

матчів, пов'язаних із заробітком на ставках. Для визначення контекстуального атрибута «сила команди» команди були згруповані за допомогою методів К-середніх і сумішей Гаусса на основі кількості очок і різниці між забитими і пропущеними м'ячами.

Третій розділ присвячений розробці методу виявлення підозрілих за фіксованими результатами футбольних матчів на основі аномального конформного детектора, в якому прийняття рішення відбувається за правилом порогу з використанням степеневого та інтегрального мартингалів.

У четвертому розділі методом імітаційного моделювання розглянуто аналіз особливостей чисельної реалізації розроблених методів та проведено порівняльний аналіз розроблених методів і відомого методу гістограми на основі модельних сезонних даних.

Викладення матеріалу ведеться послідовно і логічно, з використанням загальноприйнятої наукової термінології, стиль викладу відповідає вимогам до сучасної наукової літератури.

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

### **Оприлюднення результатів дисертаційної роботи**

Наукові результати дисертації висвітлені у 8 наукових публікаціях здобувача, серед яких: 3 статті у наукових виданнях, включених на дату опублікування до переліку наукових фахових видань України; 2 статті у періодичних наукових виданнях, проіндексованих у базах даних Web of Science Core Collection та/або Scopus, віднесених до третього квартиля (Q3) відповідно до класифікації SCImago Journal and Country Rank або Journal Citation Reports.

Також результати дисертації були апробовані на 5 наукових фахових конференціях.

Основні результати дисертації опубліковані в авторитетних наукових виданнях із попереднім рецензуванням, що забезпечує високий науковий рівень та дотримання принципів академічної доброчесності. Таким чином, наукові результати описані в дисертаційній роботі повністю висвітлені у наукових публікаціях здобувача.

### **Недоліки та зауваження до дисертаційної роботи.**

1. При проведенні кластеризації команд за силою (п. 2.1) доцільно було б перевірити стійкість встановленої належності тієї чи іншої команди до певної групи за результатами різних сезонів.

2. З тексту роботи незрозуміло, як було обґрунтовано вибір саме закону Пуассона для опису теоретичного закону розподілу ймовірностей кількості забитих голів командами під час їх гри з іншими командами відповідного рангу (п. 2.2) порівняно з альтернативними дискретними законами розподілу.

3. При описі алгоритму імітаційної моделі не вказано, чи використовувався поділ команд по групах, отриманий в результаті кластеризації за реальними даними, чи кластеризація здійснювалася в процесі моделювання за згенерованими даними.

4. При розрахунку міри неконформності матчу стосовно класу матчів у роботі використовувалося відхилення від середньої арифметичної різниці м'ячів зазначеного класу (п. 3.1). Було б доцільно розглянути використання медіанної оцінки середньої, більш стійкої до можливих суттєвих викидів.

5. Принциповим моментом у побудові запропонованого в роботі конформного предиктора є вибір експертно визначеного порога відхилення різниці голів від їх середнього, що прогнозується, причому очевидно, що цей вибір залежить виключно від експертних переваг і може істотно впливати на результат. Виникає питання, чи не можна скористатися результатами попереднього імітаційного моделювання для вибору зазначеного параметра з метою підвищення точності виявлення аномальних ситуацій.

6. При обчисленні ступеневого мартингалу автор використовує термін "поточне спостереження", який зазвичай застосовується при описі рекурентних алгоритмів послідовних спостережень у реальному часі. Залишається незрозумілим, чи використовується запропонована методика побудови мартингального детектора аномалій для прийняття рішення щодо кожного чергового матчу? Як цей метод можна використовувати для обробки всієї сукупності даних по всьому сезону?

7. При описі властивостей ступеневих мартингалів слід чітко виділити результати, отримані автором, і відомі властивості, отримані в роботах V. Vovk, A. Gammerman та ін.

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 математики та статистики.

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  
комп'ютерної математики і аналізу даних  
Національного технічного університету  
«Харківський політехнічний інститут»,  
доктор техн. наук, професор

Леонід ЛЮБЧИК



«18» січня

2024 року



Підпис *проф. Леоніда Любчика*  
ЗАСВІДЧУЮ:  
ВЧЕНИЙ СЕКРЕТАР  
НАЦІОНАЛЬНОГО ТЕХНІЧНОГО УНІВЕРСИТЕТУ  
«ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»  
«18» січня 2024 р.

ЗАПЕЧЕВ. Ю. І.