

## **Рецензія**

кандидата фізико-математичних наук, доцента,  
доцента кафедри математичного моделювання

Чернівецького національного університету імені Юрія Федьковича

Юрченка Ігоря Валерійовича

на кваліфікаційне дослідження «Оцінки параметрів авторегресійних моделей»

Кнігніцької Тетяни Василівни, яка здобуває науковий ступінь доктора

філософії у галузі знань 11 «Математика та статистика»

за спеціальністю 113 «Прикладна математика»

### **Актуальність дисертаційного дослідження**

У сучасному інформаційному суспільстві, де кількість та різноманітність даних зростають експоненційно, дослідження часових рядів є невід'ємною складовою для прийняття раціональних та обґрунтованих рішень в різних галузях – від економічного прогнозування та медичних досліджень до моніторингу кліматичних змін та розробки інноваційних технологій.

Завдяки зростанню можливостей накопичувальних пристроїв у наш час акумульовано величезні обсяги даних, за допомогою яких науковці отримують багато цінної інформації для покращення життєдіяльності людини. Машинне навчання настільки проникло у повсякденне життя, що люди навіть і не підозрюють того, що користуються ним десятки разів у день. Тому даний напрямок наукових досліджень є надзвичайно актуальним та активно розвивається. Тим не менше досі не знайдено універсального алгоритму машинного навчання, який добре працював би з наборами даних із різних сфер застосування. На практиці, деякий алгоритм машинного навчання може досягати акуратності 99% при прогнозуванні, класифікації чи кластеризації економічних даних, у той же час не перевищувати точності 70% для даних з промислового сектору. Зазвичай вхідні дані є часовими рядами або графами, які є представниками неструктурованих типів даних. Кваліфікаційна наукова праця Кнігніцької Т.В. зосереджена на дослідженні двох основних задач машинного навчання без учителя – визначення міри подібності між часовими рядами та

знаходження оптимальної кількості кластерів для набору даних, який представлено графом.

### **Зв'язок роботи з державними програмами, планами, темами**

Дисертаційне дослідження виконане на кафедрі прикладної математики та інформаційних технологій Чернівецького національного університету імені Юрія Федьковича у рамках наукової тематики «Математичне моделювання і числово-аналітичні методи дослідження динамічних та інформаційних процесів» (номер державної реєстрації: 0102U006591). Тему дисертації затверджено Науково-технічною радою Чернівецького національного університету імені Юрія Федьковича (протокол № 10 від 14 грудня 2016 року).

### **Ступінь обґрунтованості наукових положень, висновків, рекомендацій, сформульованих у дисертації.**

Обґрунтованість наукових положень, результатів та висновків дисертації забезпечені застосуванням добре апробованих методів аналізу часових рядів та теорії стохастичних випадкових матриць графу до розв'язання поставлених завдань. У роботі всі ідеї та обчислення викладено з послідовністю та логічністю. Робота містить багато посилань на наукову літературу. Тому читач при необхідності може легко отримати додаткову інформацію про підходи, які використано у дослідженні. Кваліфікаційне дослідження Книгніцької Т.В. містить приклади застосування та порівняння отриманих результатів з класичними підходами на реальних даних. Результати, отримані у дисертаційному дослідженні, узгоджуються із результатами інших науковців, які використовували аналогічні дані, але інші методи кластеризації. Кваліфікаційне дослідження є добре апробованим на 8 наукових конференціях з публікацією тез чи матеріалів доповідей. Основні результати дисертаційного дослідження опубліковані в 3 наукових статтях, дві з яких індексуються в базах даних Scopus та/або Web of Science. Кількість наукових статей, які додатково відображають результати дисертації – 4.

## Структура дисертації

Дисертація складається з вступу, трьох розділів, висновків до кожного з розділів, загальних висновків, списку використаних джерел, який налічує 122 позиції, та списку публікацій здобувачки за темою дисертації (7 статей та 8 тез конференцій). Загальний обсяг роботи – 150 сторінок. Основні результати дисертації у повній мірі відображені у публікаціях.

## Наукова новизна

Результати і висновки дисертації, які становлять наукову новизну, включають наступні аспекти:

- Описано алгоритм для встановлення подібності між стаціонарними часовими рядами, які описуються моделями  $ARMA(p, q)$ . Отримана метрика використовує відстань між параметрами моделей  $ARMA(p, q)$ , а не між самими вимірюваннями часових рядів. Порівняння запропонованої метрики з класичними підходами (евклідова відстань, DTW, ERP) показали, що описана дисертанткою метрика є більш стійкою до викидів у часових рядах і дає точніші результати для часових рядів з великою кількістю вимірювань. Встановлено, що складність алгоритму обчислення, запропонованого дисертанткою, при розгляді  $N$  часових рядів складає  $O(T * N^2)$ , в той же час аналогічна складність алгоритмів DTW, ERP становить  $O(T^2 N^2)$ , отже, за допомогою отриманої метрики отримуються більш стійкі до шумів кластери.
- Показано, що відносна похибка вимірювань для описаного алгоритму встановлення міри подібності між часовими рядами на основі моделей зростає логарифмічно. Відносна похибка вимірювань для класичних методів (евклідова відстань, DTW, ERP) зростає лінійно. Також при розгляді довгих часових рядів (коли кількість вимірювань часового ряду зростає,  $T > 1000$ ) відносна похибка вимірювань для запропонованого методу спадає, а для класичних методів – залишається сталою величиною.

- Запропоновано новий метод визначення оптимальної кількості кластерів, який базується на проведенні спектрального аналізу стохастичної матриці графу. Власні значення стохастичної матриці графу поділяються на три групи, одна з яких  $i$  є визначальною для оптимальної кількості кластерів.
- За допомогою імітаційного моделювання методом Монте-Карло показано, що запропонований метод дає кращі результати для визначення оптимальної кількості кластерів у порівнянні з низкою класичних методів (марковський алгоритм з двома типами параметрів та метод ліктя). Імітаційне моделювання Монте-Карло використане для утворення багатовимірних даних – графу з фіксованою кількістю сукупностей (кластерів). За цих умов для порівняння запропонованого методу вибору оптимальної кількості кластерів з марковським алгоритмом та методом ліктя дисертантка наперед володіла інформацією про оптимальну кількість кластерів.
- Встановлено, що розроблений алгоритм, який використовує стохастичну матрицю графу для визначення оптимальної кількості кластерів, є менш чутливим до наявності кластерів різного розміру.

### **Практичне значення одержаних результатів**

Результати, отримані у дисертаційному дослідженні Кнігніцької Т.В. можуть бути використані для кластеризації часових рядів та встановлення подібності між часовими рядами. У сучасному світі, де економічні умови постійно змінюються, а глобальні події впливають на ринки, дослідження часових рядів використовується для прогнозування економічних тенденцій. Аналіз часових рядів дозволяє виявляти циклічні зміни, тренди та аномалії, що є ключовим для управління ризиками та прийняття стратегічних рішень у бізнесі та фінансах.

У сфері медицини, особливо у контексті пандемії та глобальних викликів здоров'я, дослідження часових рядів використовується для відстеження поширення захворювань, прогнозування епідемій та аналізу ефективності лікування. Часові ряди можуть містити інформацію про температуру тіла, ритм

серця, рівень глюкози та інші показники, які допомагають у здійсненні ранніх діагнозів та вдосконаленні методів лікування.

У промисловому секторі аналіз часових рядів використовується для моніторингу та підтримки операцій. Відстеження роботи обладнання, виявлення потенційних поломок за допомогою датчиків та передбачення технічного обслуговування – області, де дослідження часових рядів стає визначальним фактором в ефективному управлінні виробництвом. Зміни в кліматі – один із найактуальніших викликів для сучасного світу. Дослідження часових рядів кліматичних даних допомагає відстежувати зміни погодних умов, рівень океанів та інші екологічні параметри. Це необхідно для розробки стратегій адаптації та зменшення впливу кліматичних змін.

### **Зауваження**

- 1) У списку публікацій здобувачки за темою дисертації є роботи, присвячені обробці медичних даних. Тому було б доцільно у дисертаційній роботі навести приклад застосування запропонованих алгоритмів для дослідження даних не тільки ринку акцій, а й медичної статистики.
- 2) У дисертації присутні незначні описки.

Висловлені зауваження не применшують наукові досягнення дисертації Кнігніцької Т.В. та не впливають на висновок про належний рівень дослідження і його загальну позитивну оцінку.

### **Загальний висновок**

Дисертаційна робота Кнігніцької Тетяни Василівни «Оцінки параметрів авторегресійних моделей», подана на здобуття наукового ступеня доктора філософії з галузі знань 11 – Математика та статистика за спеціальністю 113 – Прикладна математика за її актуальністю, науково-теоретичним рівнем, новизною розв'язання поставлених завдань, практичним значенням отриманих результатів, а також їх апробацією цілком відповідає пунктам 6, 7, 8, 9 «Порядку присудження ступеня доктора філософії та скасування рішення

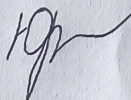


разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії», затвердженого Постановою Кабінету Міністрів України №44 від 12 січня 2022 року (зі змінами, внесеними згідно з Постановою Кабінету Міністрів України №341 від 21 березня 2022 року), а також «Вимогам до оформлення дисертації», затверджених Наказом Міністерства освіти і науки України №40 від 12 січня 2017 року.

Вважаю, що авторка кваліфікаційної наукової роботи Книгніцька Тетяна Василівна заслуговує на присудження їй наукового ступеня доктора філософії з галузі знань 11 – Математика та статистика за спеціальністю 113 – Прикладна математика.

Рецензент

кандидат фізико-математичних наук, доцент,  
доцент кафедри математичного моделювання  
Чернівецького національного університету  
імені Юрія Федьковича



Ігор ЮРЧЕНКО

