

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА
ЕКОНОМІЧНИЙ ФАКУЛЬТЕТ

FORUM EFBM 2.0

28 вересня - 1 жовтня 2021

Зміни Адаптація Нова Економіка

МАТЕРІАЛИ МІЖНАРОДНОГО ФОРУМУ

Міжнародна науково-практична конференція

**Data Science та інформаційно-аналітичні системи:
застосування в економіці та фінансах**

ЕФ КНУ ЕФ КНУ ЕФ КНУ ЕФ КНУ ЕФ КНУ ЕФ
КНУ ЕФ КНУ ЕФ КНУ

FORUM EFBM 2.0

Зміни. Адаптація. Нова економіка

Матеріали Міжнародного форуму

Міжнародна науково-практична конференція

**Data Science та інформаційно-аналітичні системи:
застосування в економіці та фінансах**

Київ 2021

УДК 330, 331, 336

JEL classification: D00, E00, G00, M10, M20

Рецензенти:

Л.Буяк, доктор економічних наук, професор

Р.Кулієв, доктор економічних наук, професор

Редакційна колегія:

д-р екон. наук, проф. А.І. Ігнатюк (голова); канд. фіз.-мат. наук, доц. Л.А. Анісімова; д-р екон. наук, проф. І.О. Лютий; д-р екон. наук, проф. А.О. Старостіна; д-р екон. наук, проф. О.І. Ляшенко; д-р екон. наук, проф. Н.І. Гражевська; д-р екон. наук, проф. Г.І. Купалова; д-р екон. наук, проф. Г.М. Филюк; д-р екон. наук, доц. М.В. Ситницький; д-р екон. наук, доц. Н.В. Приказюк; д-р екон. наук, проф. Н.В. Ковтун; д-р екон. наук, доц. Б.А. Засадний; к.філол.н., доц. М.В. Петровський; д-р екон. наук, проф. Н.В. Бутенко; д-р екон. наук, доц. О.В. Баженова; д-р екон. наук, проф. Т.В. Гайдай; д-р екон. наук, доц. В.В. Вірченко; д-р екон. наук, проф. А.Б. Камінський; канд. екон. наук, доц. П.В. Кухта; канд. екон. наук, ас. О.Л. Михальська; канд. екон. наук, доц. О.Ю. Мірошніченко; канд. екон. наук, доц. Т.О. Мурована; канд. екон. наук, доц. З.О. Пальян; канд. екон. наук, ас. Н.А. Плешакова; канд. екон. наук, доц. О.Ю. Пригара; канд. екон. наук А.О. Полторацька; канд. екон. наук, доц. І.Я. Софіщенко; канд. філол. наук, доц. А.С. Сизенко; канд. філол. наук, доц. Святюк Ю.В.; д-р істор. наук, проф. О.В. Чеберяко; д-р екон. наук, доц. А.С. Шолойко.

Е Економіка. Фінанси. Бізнес. Управління, матеріали ІІ Міжнародного форуму / За заг. ред. проф. А.І. Ігнатюк: – К., 2021. – Data Science та інформаційно-аналітичні системи: застосування в економіці та фінансах – 55 с.

Збірник містить матеріали ІІ Міжнародного форуму «Економіка. Фінанси. Бізнес. Управління. *Зміни. Адаптація. Нова економіка*», що проходив 28 вересня – 1 жовтня 2021 року на економічному факультеті Київського національного університету імені Тараса Шевченка. Збірник охоплює наукові публікації за основними напрямками економічної науки.

Для науковців, державних службовців, викладачів, аспірантів і студентів економічних спеціальностей закладів вищої освіти.

Автори опублікованих матеріалів несуть повну відповідальність за підбір, точність наведених фактів, цитат, економіко-статистичних даних, галузевої термінології, імен власних та інших відомостей.

ISBN 978-617-696-866-5

ISBN 978-617-696-867-2

© Київський національний університет імені Тараса Шевченка, 2021

FORUM EFBM 2.0

Changes. Adaptation. New economy

Proceedings of the International Forum

International Scientific and Practical Conference

**Data Science, Intelligence Systems and their Applications in
Economy and Finance**

Kyiv 2021

УДК 330, 331, 336

JEL classification: D00, E00, G00, M10, M20

Reviewers:

Buiak Lesia, Doctor of Economics, Professor

Rufat Kuliev, Doctor of Economics, Professor

Editorial Board:

Anzhela Ignatyuk, Doctor of Economics, Professor (Head of Editorial Board); Lyudmyla Anisimova, PhD in Physical and Mathematical Sciences, Associate Professor; Ihor Liutyii, Doctor of Economics, Professor; Alla Starostina, Doctor of Economics, Professor; Olena Liashenko, Doctor of Economics, Professor; Nadiia Hrazhevska, Doctor of Economics, Professor; Halyna Kupalova, Doctor of Economics, Professor; Halyna Filiuk, Doctor of Economics, Professor; Maksym W. Sitnicki, Doctor of Economics, Associate Professor; Nataliia Prukaziuk, Doctor of Economics, Professor; Natalia Kovtun, Doctor of Economics, Professor; Bohdan Zasadnyi, Doctor of Economics, Associate Professor; Mykola Petrovskyi, PhD in Philology; Associate Professor; Nataliia Butenko, Doctor of Economics, Professor; Olena Bazhenova, Doctor of Economics, Associate Professor; Tetiana Gaidai, Doctor of Economics, Professor; Volodymyr Virchenko, Doctor of Economics, Associate Professor; Andrii Kaminskyi, Doctor of Economics, Professor; Pavlo Kukhta, PhD in Economics, Associate Professor; Olena Mykhalska, PhD in Economics, Assistant; Olga Miroshnychenko, PhD in Economics, Associate Professor; Tetiana Murovana, PhD in Economics, Associate Professor; Zinaida Palian, PhD in Economics, Associate Professor; Nataliia Pleshakova, PhD in Economics, Assistant; Olha Pryhara, PhD in Economics, Associate Professor; Alona Poltoratska, PhD in Economics; Iryna Sofishchenko, PhD in Economics, Associate Professor; Anastasiia Syzenko, PhD in Philology; Associate Professor; Yulia Sviatiuk, PhD in Philology, Associate Professor; Oksana Cheberyako, Doctor of Habil. in Historical Sciences, PhD in Economics, Professor; Antonina Sholoiko, Doctor of Economics, Associate Professor.

E Economics. Finance. Business. Management, Proceedings of the II International Forum / Edited by Prof. Anzhela Ignatyuk : Kyiv, 2021. Data Science, Intelligence Systems and their Applications in Economy and Finance – 55 p.

The proceedings contain the papers presented at the II International Forum «Economics. Finance. Business. Management. *Changes. Adaptation. New economy*» held on September 28 – October 1, 2021 at the Faculty of Economics of Taras Shevchenko National University of Kyiv. The proceedings cover research papers in key areas of economics.

For scholars, civil servants, teachers, postgraduate and graduate students of economics in higher educational institutions.

The authors of published papers are solely responsible for the selection and accuracy of presented facts, quotations, economic and statistical data, professional terminology, proper names, and other data.

ISBN 978-617-696-866-5

ISBN 978-617-696-867-2

© Taras Shevchenko National University of Kyiv, 2021

DATA SCIENCE ТА ІНФОРМАЦІЙНО-АНАЛІТИЧНІ СИСТЕМИ: ЗАСТОСУВАННЯ В ЕКОНОМІЦІ ТА ФІНАНСАХ

Денис Бутило КЛАСТЕРИЗАЦІЯ ІНДЕКСНИХ ФОНДІВ АЛЬТЕРНАТИВНИХ ІНВЕСТИЦІЙ НА ОСНОВІ ПОКАЗНИКІВ РИЗИКУ	8
Ірина Вороненко, Наталія Клименко ВПЛИВ АСИМЕТРІЇ ІНФОРМАЦІЇ НА СУСПІЛЬНИЙ ДОБРОБУТ	11
Юлія Горященко УПРАВЛІННЯ ІННОВАЦІЯМИ ТА СУЧАСНІ МОДЕЛІ ЗАБЕЗПЕЧЕННЯ КОНКУРЕНТОСПРОМОЖНОСТІ ГОСПОДАРЮЮЧИХ СУБ'ЄКТІВ	14
Тетяна Затонацька, Олена Максимчук МОДЕЛІ ВИКОРИСТАННЯ ІНСТРУМЕНТІВ DATA SCIENCE В ЕЛЕКТРОННІЙ КОМЕРЦІЇ	16
Лариса Зомчак, Лілія Коваль ПРОСТОРОВЕ МОДЕЛЮВАННЯ СТАЛОГО ТЕРИТОРІАЛЬНОГО РОЗВИТКУ УКРАЇНИ	19
Андрій Камінський АЛЬТЕРНАТИВНІ ДАНІ В КРЕДИТНОМУ РИЗИК-МЕНЕДЖМЕНТІ: НОВИЙ ПІДХІД ДО ЗАСТОСУВАННЯ	21
Катерина Кононова ПРАКТИКА МАШИННОГО НАВЧАННЯ: ОЧІКУВАННЯ ТА РЕАЛЬНІСТЬ	24
Олена Ляшенко, Тетяна Кравець, Анастасія Філогіна ЗАСТОСУВАННЯ СКЛАДНИХ НЕЙРОННИХ МЕРЕЖ ПРИ МОДЕЛЮВАННІ ВОЛАТИЛЬНОСТІ ФОНДОВИХ ІНДЕКСІВ	26
Марина Негрей DATA SCIENCE В FOODTECH: ВИГОДИ ТА ОСНОВНІ ТЕНДЕНЦІЇ	29
Володимир Соловйов, Андрій Матвійчук, Андрій Белінський ЕКОНОФІЗИЧНІ ПІДХОДИ В МОДЕЛЮВАННІ ФІНАНСОВИХ РИНКІВ	32
Олександр Подскребко ПРОГНОЗУВАННЯ ДИНАМІКИ ЗМІНИ ЦІН НА КРИПТОВАЛЮТУ З ВИКОРИСТАННЯМ АЛГОРИТМІВ ГРАДІЄНТНОГО БУСТІНГУ	35
Ірина Федоренко, Вероніка Сачко МОДЕЛЮВАННЯ ВПЛИВУ ТЕХНОЛОГІЧНОГО РОЗВИТКУ НА ГЛОБАЛЬНИЙ СТАЛИЙ ЕКОНОМІЧНИЙ РОЗВИТОК	37
Галина Черноус, Яна Фаренюк МОДЕЛЮВАННЯ МАРКЕТИНГОВОЇ ДІЯЛЬНОСТІ: ПІДХОДИ НА ОСНОВІ ТЕХНОЛОГІЙ DATA SCIENCE	40
Віктор Шпирко, Олександр Бартков'як ЗАСТОСУВАННЯ СТАТИСТИЧНОГО АРБІТРАЖУ ПРИ ТОРГІВЛІ АКЦІЯМИ ТА ДЕРИВАТИВАМИ	43
Віктор Шпирко, Андрій Винокуров МОДЕЛЮВАННЯ ТА ПРОГНОЗУВАННЯ МІГРАЦІЙНИХ ПРОЦЕСІВ В УКРАЇНІ ЗА ДОПОМОГОЮ ГРАВІТАЦІЙНОГО ПІДХОДУ	46
Богдан Якимчук ЗАСТОСУВАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ ІНВЕСТИЦІЙНИХ РІШЕНЬ	49
Валерія Яценко DATA MINING В ОЦІНЦІ ВАЛЮТНИХ РИЗИКІВ В УМОВАХ НЕПОВНОЇ ІНФОРМАЦІЇ	52

DATA SCIENCE, INTELLIGENCE SYSTEMS AND THEIR APPLICATIONS IN ECONOMY AND FINANCE

Denys Butylo CLUSTERING OF ALTERNATIVE INVESTMENT'S ETFs BASED ON RISK INDICATORS	8
Iryna Voronenko, Nataliia Klymenko IMPACT OF INFORMATION ASYMMETRY ON SOCIAL WELFARE	11
Yuliia Horiashchenko INNOVATION MANAGEMENT AND MODERN MODELS OF ENSURING COMPETITIVENESS OF BUSINESS ENTITIES	14
Tetiana Zatonatska, Olena Maksymchuk THE MODELS OF USING DATA SCIENCE TOOLS IN E-COMMERCE	16
Larysa Zomchak, Lilia Koval SPATIAL MODELLING OF UKRAINE SUSTAINABLE TERRITORIAL DEVELOPMENT	19
Andrii Kaminskyi ALTERNATIVE DATA IN CREDIT RISK-MANAGEMENT: NEW APPROACH FOR APPLICATION	21
Kateryna Kononova MACHINE LEARNING PRACTICE: EXPECTATIONS VS REALITY	24
Olena Liashenko, Tetyana Kravets, Anastasiya Filogina APPLICATION OF COMPLEX NEURAL NETWORKS IN STOCK INDICES VOLATILITY MODELING	26
Maryna Nehrey DATA SCIENCE IN FOODTECH: BENEFITS AND THE MAIN TRENDS	29
Vladimir Soloviev, Andriy Matviychuk, Andrii Bielinskyi ECONOPHYSICS APPROACHES IN FINANCIAL MARKET MODELING	32
Oleksandr Podskrebko FORECASTING THE DYNAMICS OF CHANGES IN CRYPTOCURRENCY PRICES USING GRADIENT BOOSTING ALGORITHMS	35
Iryna Fedorenko, Veronika Sachko MODELING THE INFLUENCE OF TECHNOLOGICAL DEVELOPMENT ON GLOBAL SUSTAINABLE ECONOMIC DEVELOPMENT	37
Galyna Chornous, Yana Fareniuk MODELING OF MARKETING ACTIVITY: APPROACHES BASED ON DATA SCIENCE TECHNOLOGIES	40
Viktor Shpyrko, Oleksandr Bartkoviak STATISTICAL ARBITRAGE APPLICATION IN EQUITY AND DERIVATIVES TRADING	43
Viktor Shpyrko, Andrii Vynokurov MODELLING AND FORECASTING OF MIGRATION PROCESSES IN UKRAINE USING GRAVITATIONAL APPROACH	46
Bohdan Yakymchuk APPLICATION OF MACHINE LEARNING METHODS FOR INVESTMENT DECISIONS SUPPORT	49
Valeria Yatsenko DATA MINING IN CURRENCY RISK ASSESSMENT IN THE CONDITIONS OF INCOMPLETE INFORMATION	52

Денис Бутило,

PhD студент,

Київський національний університет імені Тараса Шевченка

КЛАСТЕРИЗАЦІЯ ІНДЕКСНИХ ФОНДІВ АЛЬТЕРНАТИВНИХ ІНВЕСТИЦІЙ НА ОСНОВІ ПОКАЗНИКІВ РИЗИКУ

Ключові слова: оцінка ризику, альтернативні інвестиції, традиційні інвестиції, біржові індексні фонди, кластеризації, портфельний менеджмент

Denys Butylo,

PhD Student,

Kyiv National Taras Shevchenko University

CLUSTERING OF ALTERNATIVE INVESTMENT'S ETFs BASED ON RISK INDICATORS

Keywords: risk assessments, alternative investments, traditional investments, ETF (Exchange Trade Funds), clustering, portfolio management

Сучасні фінансові інвестиції можуть бути поділені на дві широкі категорії: традиційні та альтернативні активи. Традиційні активи включають акції, облігації та депозитні сертифікати в банках. В той же час визначення альтернативних активів є на сьогодні поняттям досить нечітким. Існує два основних підходи до його визначення: “за виключенням” та “за включенням”. Перший підхід інтерпретує такі активи як всі, що не належать до традиційних. Даний метод визначення є досить широким і включає багато видів інвестицій, від досить поширених інвестицій в нерухомість до предметів мистецтва та винних колекцій. Другий підхід орієнтований на чіткому перерахуванні альтернативних класів активів, які відносять до цієї категорії ті чи інші автори [1]. Найчастіше виокремлюють наступні чотири категорії альтернативних активів: хедж-фонди, приватний капітал, реальні активи і структуровані продукти.

На сьогодні альтернативні інвестиції складають значну частину інвестиційного ринку, яка має зростаючий тренд. Зокрема, станом на 2020 рік, загальна капіталізація альтернативних інвестиційних активів на світовому ринку оцінюється в 18,1 трильйона доларів США. Основний напрямок їх інвестиційного використання полягає у розширенні спектру інвестицій, зокрема, диверсифікація ризику шляхом портфельного підходу.

З точки зору портфельного підходу, використання альтернативних інвестицій має ряд переваг і недоліків.

Переваги використання альтернативних інвестицій:

- Дохідність альтернативних інвестицій зазвичай демонструє відсутність кореляції з доходністю традиційних інвестицій. Цей факт сприяє зниженню ризику портфеля за рахунок комбінації обох типів вкладення в портфель.

- Альтернативні інвестиції мають потенціал для відносно більш високих дохідностей. З огляду на те, що очікувана прибутковість портфеля є лінійною функцією, то поєднання альтернативних активів з традиційними сприяє більш високому очікуваному значенню дохідності.

Недоліки:

- Потенційно волатильність дохідності вище, ніж аналогічна волатильність традиційних активів.
- Відносно низька ліквідність;
- Більш високі інвестиційні витрати, пов'язані із порогом входу в інвестиції;

Задача, яка розглядається в нашому дослідженні, полягала в тому, щоб здійснити кластеризацію альтернативних інвестицій на основі біржових індексних фондів. Основою для кластеризації виступила значна статистична база яка вже наявна по альтернативним інвестиціям. Зокрема, ми використовували інформаційні ресурси <https://www.investing.com/> та <https://etfdb.com>. На основі цих інформаційних ресурсів нами сформована статистична база низки економічних показників (дохідність, ліквідність тощо).

Щоб краще зрозуміти природу ризику альтернативних інвестицій, ми розглянули різні підходи до його оцінки для даного типу активів. Здійснивши оцінку різними методами, ми застосували підходи кластеризації, щоб виділити кластери з однорідними компонентами. Такий підхід є цікавими з точки зору відношення «ризик і дохідності» альтернативних активів та виявлення «схритих» закономірностей.

Основою нашої методології є використання біржових індексних фондів (ETF - Exchange Trade Fund) альтернативних інвестицій. Другий методологічний підхід включає побудову репрезентативної вибірки ETF, яка представляє альтернативні інвестиції на ринку США. Крім того, вибірка передбачає охоплення всіх основних типів альтернативних інвестицій. Для реалізації цих підходів ми використовували базу даних ETF (<https://etfdb.com>). Процес Data Mining включав формування вибірки та формування 14 показників ризику. Вибірка була сформована з 83 ETF із загальною капіталізацією в 138,6 млрд дол. Аналіз дохідності був заснований на щотижневих даних за період з 1 січня. 2015 по 31 грудня 2019 р, що включає 260 значень для кожного ETF. Показники ризику включали [2] : K-ratio, міра Фішера, сподівана дохідність, стандартне відхилення дохідності, відношення верхньої семіваріації до нижньої, CVaR, ексцес, бета коефіцієнти то індексів SPY та AGG, середня кореляція, коефіцієнт Херста, автокореляція, середній обсяг торгів та коефіцієнт його ризику

Застосування методів кластеризації дозволило сформувати 5 окремих кластерів альтернативних інвестиційних активів [3].

Кластер 1 включає ETF, які характеризуються негативним K-ratio і прямуючим до нуля середнім значенням дохідності. Значення трьох із критеріїв оцінки ризику для даного кластеру демонструють найбільші значення, а саме: міра ризику Фішера, стандартне відхилення і CVaR. При цьому для даного кластеру не характерні суттєві асиметрія і ексцес. Позитивна характеристика цього кластера - порівняно високий рівень ліквідності. 70% ETF з цього кластера відносяться до групи альтернативних активів, пов'язаних з нафтою та

газом. Це дозволяє зробити висновок, що кластер 1 має високий ризик з негативними К-коефіцієнтом і низьким середнім значенням дохідності.

Кластер 2 складається з ETF, які демонструють позитивний K-ratio, помірні значення міри ризику Фішера, стандартного відхилення, CVaR і ексцесу. Ризик асиметрії не характерний. Перевага цього кластеру це високі обсяги торгів. Це в свою чергу означає низький ризик ліквідності. Ключовими компонентами цього кластера є ETF класів «дорогоцінні метали» (33%) та «нерухомість» (48%).

Кластер 3 характеризується ETF з позитивним K-ratio, помірними рівнями міри ризику Фішера, стандартного відхилення та CVaR. Також їм характерне значне значення ексцесу і ризик асиметрії і ризик чутливості, пов'язаний з доходами SPY. Цей кластер характеризується високим ризиком ліквідності. Даний кластер складається із суміші ETF різних видів альтернативних інвестиційних активів.

Кластер 4 включає ETF з найвищим рівнем K-ratio, найнижчим рівнем стандартного відхилення і CVaR. У той же час ризик представлений ексцесом і високою асиметрією. Прибутковість ETF з цього кластера чутлива до SPY. Кластер демонструє високий ризик ліквідності. Він сформований передусім ETFами 2 типів: хедж-фонди і фонди з довго-короткими позиціями.

Кластер 5 складається з ETF з негативним K-ratio, значним рівнем дохідності, а також помірними значеннями міри Фішера, стандартного відхилення, ексцесу та CVaR. Даний кластер характеризується низьким ризиком ліквідності, через значний обсяг торгів. Кластер складається з різних типів ETF, проте 65% складають сировинні товари та сировинні товари сільськогосподарського типу.

Підхід кластеризації дозволив виділити «неявні» залежності між показниками ризику та сформувані базу до кластерного підходу у портфельному менеджменті альтернативних інвестицій. Підхід сфокусований на використанні наївнодиверсифікованих портфелів за кластерами альтернативних інвестицій.

Список використаних джерел

1. Sokołowska E. (2014). *Alternative Investments in Wealth Management. A Comprehensive Study of the Central and East European Market*. Springer.
2. Andrii Kaminskyi, Maryna Nehrey, Denys Butylo / Integrated approach for risk assessment of alternative investments // *International Journal of Risk Assessment and Management (IJRAM)* (forthcoming).
3. Andrii Kaminskyi, Maryna Nehrey, Denys Butylo / Risk modelling of alternative investments / Системный анализ и моделирование процессов управления / Под ред. докт. экон. наук, проф. В.С. Пономаренко, докт. экон. наук, проф. Т.С. Клебановой, докт. экон. наук, проф. Л.С. Гурьяновой - Братислава-Харьков, ВШЭМ - ХНЭУ им. С. Кузнеця, 2020. – с. 229-245.

Ірина Вороненко,

д.е.н., с.н.с.,

Національний університет біоресурсів
і природокористування України**Наталія Клименко,**

к.е.н., доцент,

Національний університет біоресурсів
і природокористування України

ВПЛИВ АСИМЕТРІЇ ІНФОРМАЦІЇ НА СУСПІЛЬНИЙ ДОБРОБУТ

Ключові слова: Асиметрія, Інформація, Інформаційний простір, Суспільний добробут, Цифровізація

Iryna Voronenko,

Doctor of Economic Sciences, Senior Researcher,

National University of Life and Environmental Sciences of Ukraine

Nataliia Klymenko,

Ph.D in Economics, Associate Professor,

National University of Life and Environmental Sciences of Ukraine

IMPACT OF INFORMATION ASYMMETRY ON SOCIAL WELFARE

Key words: Asymmetry, Information, information space, social welfare, digitization

Інформація займає одне із найважливіших місць в системі економічних відносин, особливо в умовах цифровізації, у силу чого може вважатися найціннішим економічним ресурсом. А вплив інформаційної асиметрії на суспільний добробут та механізм ринкового саморегулювання вимагає наукового дослідження та обґрунтування. В роботі відповідно до запропонованого концептуального підходу здійснено оцінку ефективності послуг у сфері інформаційно-комунікаційних технологій від користування адміністративними послугами онлайн для конкретної особи внаслідок економії часу та коштів на переміщення. Відзначено нерівномірність відкритих даних, що надаються як центральними, так і місцевими органами державної влади. Розглянуто функцію корисності споживача інформаційно-комунікаційних послуг.

Проблему інформаційної асиметрії та її безпосереднього впливу на функціонування ринків і економіки в цілому вперше піднято у науковій літературі близько 70 років тому, проте вона і нині продовжує активно досліджуватись [1 – 3]. Незважаючи на активні дискусії, які точилися серед відомих вчених-економістів протягом останніх десятиліть стосовно ролі, значення та ефектів інформаційної асиметрії, досі залишаються недостатньо дослідженими питання щодо напрямів та інструментів впливу інформаційної асиметрії на систему конкурентних відносин, методів ефективної протидії поширенню інформаційної асиметрії на ринку та частково питання деформації споживчого попиту за умов існування інформаційної асиметрії, при тому, що стабільність попиту на споживчі товари, а зрештою, місткість та стійкість

ринку споживчих товарів і послуг є однією з основних складових, що здатні найбільш точно характеризувати конкурентоспроможність національної економіки.

Пропонуємо розглянути зміну добробуту, пов'язану з інформацією, на прикладі зміни поведінки окремого споживача. Припустимо, що конкретний споживач, обмежений певними фінансовими ресурсами, опинився перед вибором придбати товар серед інших товарів за умови, що він має достовірну інформації щодо його якості. Водночас корисність цього товару він зможе оцінити лише за результатами його експлуатації, аналогічна ситуація і щодо достовірності інформації про вплив цього товару на його здоров'я, яку він також зможе оцінити лише через певний час експлуатації цього товару.

Припустимо, що рівень зацікавленості споживача у придбанні цього товару (z) також не є сталим і характеризується змінною θ , що залежить від оцінки корисності цього товару для нього у конкретний момент часу. Відповідно, конкретні значення θ відповідають конкретним значенням набору інформації у певний момент часу, що припустимо, насамперед, залежать від інформації про якість цього товару (q). У такому разі корисність товару для споживача може бути представлена як:

$$U(q, z), \text{ де } q = (q_1, \dots, q_N). \quad (1)$$

Тоді задача максимізації очікуваної корисності для споживача має вигляд:

$$\max_q E_\theta[U(q, z) | p_q = m, q \geq 0], \quad (2)$$

де $p = (p_1, \dots, p_N)$ – вектор цін, що відповідає q , $p_q = \sum_{i=1}^N p_i q_i = m$, m – споживчий дохід, а θ – набір інформації або передбачуваний розподіл вірогідності споживання z у момент прийняття рішення щодо споживання. Інакше кажучи, E_θ – це очікування щодо споживання у момент прийняття рішення. Очевидно, що споживач буде збільшувати споживання у разі поліпшення очікуваної якості товару та зменшувати споживання у разі збільшення невизначеності щодо якості товару. Отже, вплив інформації на добробут визначається у двох вимірах: очікуваннях споживача, а також ступені невизначеності для нього інформації.

Як кількісний показник виміру асиметрії доступу до інформації запропоновано використовувати коефіцієнт Джині, який, відповідно до проведених розрахунків склав 35,2 %, що навіть перевищує в 1,4 раза аналогічний коефіцієнт доходів Джині для України.

Подолання наявної асиметрії доступу до мережі інтернет в Україні сприятиме підвищенню добробуту як конкретної особи, так і нашої держави загалом, а також підвищенню її конкурентоспроможності, що позитивно вплине на економічний, освітній, політичний, громадський, соціальний простір українців, що є особливо в умовах докорінних змін, що відбулися в інформаційному просторі України.

Водночас окремо зазначимо, що оцінюючи підвищення суспільного добробуту за рахунок використання інформаційно-комунікаційних технологій та послуг у сфері інформаційного простору необхідно врахувати їх вплив, зокрема, за рахунок підвищення суспільного добробуту й юридичних осіб.

Ефективність комунікаційних засобів для мешканця дорівнює відношенню переваг від їх використання в грошовому еквіваленті до витрат за їх користування:

$$E_k = \frac{\sum_{i=1}^n wT_i + \sum_{i=1}^n vd_i}{\sum_{i=1}^n \sum_{j=1}^m \tau_j t_{ij}}, \quad (3)$$

де T_i – кількість часу на отримання i -ї адміністративної послуги традиційним шляхом; w – мінімальна погодинна оплата праці; n – загальна кількість послуг за рік; t_{ij} – час втрачений на отримання i -ї послуги j -ми засобами інформаційно-комунікаційних технологій; d_i – відстань до адміністративного центру, де можна отримати i -ту послугу; v – вартість переміщення на одиницю відстані; τ_j – тариф за користування j -ю послугою у сфері інформаційно-комунікаційних технологій за одиницю часу; m – кількість засобів у сфері інформаційно-комунікаційних технологій.

Звичайно для жителів віддалених населених пунктів втрати часу на переміщення та його вартість роблять показник ефективності суттєво більшим. Водночас існує багато інших позитивних сторін для жителів сільських територій від використання інформаційно-комунікаційних послуг. Окремо зазначимо, що оцінюючи підвищення суспільного добробуту за рахунок використання інформаційно-комунікаційних технологій та послуг у сфері інформаційного простору необхідно врахувати їх вплив, зокрема, за рахунок підвищення суспільного добробуту й юридичних осіб.

Враховуючи, що подолання наявної асиметрії доступу до мережі інтернет в Україні сприятиме підвищенню добробуту як конкретної особи, так і держави загалом, основні цілі та принципи державного регулювання у сфері інформаційного простору України в умовах цифровізації, мають бути спрямовані на подолання цифрового розриву шляхом розвитку цифрових інфраструктур, громадської безпеки, освіти та розвитку цифрових компетенцій, впровадження концепції цифрових робочих місць, цифровізації реальних секторів економіки, реалізації проектів цифрових трансформацій.

Подальші дослідження у даному науковому напрямі, на наш погляд, мають бути спрямовані на удосконаленні моделі моніторингу цифрової трансформації економіки й формуванні багаторівневої системи соціально-економічної та організаційної оцінки цифрової трансформацій, як на національному рівні у розрізі окремих регіонів, типів населених пунктів та галузей економіки, так й на світовому рівні у розрізі розподілу країн за рівнем доходу, регіонами світу тощо.

Список використаних джерел

1. Richard E. Just, Darrell L. Hueth, Andrew Schmitz. The Welfare Economics of Public Policy: A Practical Approach to Project And Policy Evaluation, Edward Elgar Pub. 2004. 712 p.
2. Nehrey Maryna, Hnot Taras. Data Science Tools Application for Business Processes Modelling in Aviation. Cases on Modern Computer Systems in Aviation. IGI Global, 2019. p. 176-190.
3. Вороненко І. В. Взаємозв'язок між добробутом суспільства та інформацією. Бізнесінформ. 2019. № 4. С. 366–371.
4. Державна служба статистики України: сайт. URL: <http://www.ukrstat.gov.ua>

Юлія Горященко,
канд.екон.наук, доцент,
Університет митної справи та фінансів

УПРАВЛІННЯ ІННОВАЦІЯМИ ТА СУЧАСНІ МОДЕЛІ ЗАБЕЗПЕЧЕННЯ КОНКУРЕНТОСПРОМОЖНОСТІ ГОСПОДАРЮЮЧИХ СУБ'ЄКТІВ

Ключові слова: інновації, конкурентоспроможність, господарюючі суб'єкти, моделі, управління

Yuliia Horiashchenko,
PhD (Economics), Associate Professor,
University of Customs and Finance

INNOVATION MANAGEMENT AND MODERN MODELS OF ENSURING COMPETITIVENESS OF BUSINESS ENTITIES

Key words: innovations, competitiveness, economic entities, models, management

Today, the driving forces of economic growth are changing all over the world, henceforth it is a constant generation by economic needs of the ideal image of an innovative product and its creation in this form. The main comparative characteristic of economic entities is their competitiveness. The development of a competitive environment is embodied in the main result of innovation.

Naturally, a competitive economy based on knowledge, innovation and production potential.

Among the modern models of competitiveness it is worth noting:

- The UK Competitiveness Index (UKCI) characterizes the region's competitiveness based on its development, socio-economic well-being and business sustainability. UKCI experts define the competitiveness of a region as the ability of the economy of any locality (country, region, city), first, to attract and retain firms that have a stable or growing market share; second, to ensure the rising living standards of individuals; “Balancing different types of benefits that may be available in one place relative to another” [1, p. 123-124]. The calculation of this index is based on the following factors: “input factors” (R&D costs, number of newly created business units per 1,000 people, level of economic activity in the region, level of academic success of secondary school graduates (number of qualified school graduates) , the share of knowledge-intensive business); “Output factors” (volume of exports and imports per capita, gross value added of the region per capita, labor productivity of the region, level of employment); “Outcome factors” (weekly wages and unemployment rate);

- “Regional Competitiveness Hat Model” (developed with the assistance of the European Commission). The most important source of regional competitiveness is human resources;

- “Pyramid of competitiveness of the region” (B. Gardiner, R. Martin, P. Tyler; 2004) - a key indicator of the level of competitiveness of the regions determines labor productivity;

- theory of competitive advantage (M. Porter): the basis of competitive advantage of the region is labor productivity (driving force);

- the concept of territorial competitiveness (R. Kamagni): “the region has an absolute competitive advantage if it has the best technological, social, infrastructural assets”;

– The theory of industrial competitiveness of the territory (T. Todorov) connects competitiveness with the potential of the country (human resources, capital and physical assets);

– socio-cultural concept of regional competitiveness (A. Anderson). It is based on the criterion of “C - regions” (C - factors: competence, culture, communications, creativity). “With the development of new C - regions, which are concentrated around higher education centers and in cooperation with highly competitive business, they are transformed into regional innovation and production complexes, and the economic importance of the old industrial regions decreases” [1, p. 149-161].

Thus, each of the models of regional competitiveness emphasizes the importance and inevitability of mutual influence and links of socio-cultural potential with the economic dynamics of the region, in some - even attaches special importance to educational or cultural factors of competitive status of regions (model “C - regions”). It is important to note that not all of the above models, given the historical period of their creation, focus on sustainable development; sustainable development is associated primarily with social welfare (more precisely, is its direct consequence) (the model of overall competitiveness of the region). However, as today's practice shows, such a direct transition from prosperity to sustainability is impossible without a special element - information support. Thus, well-being determines the level of quality of life, and information - the level of information support for human development. Only in the combination of well-being (as a basis of socio-economic potential) and information (information culture) (as a basis of socio-cultural potential) can sustainable development be ensured. It should also be noted that the vast majority of models are created by British or American experts, therefore, not adapted to the complex Ukrainian realities.

Scientist I. Verbytskyi [2 p. 141-142] identified indicators of socio-economic development, consisting of six groups of indicators of economic development (production and economic, structural, investment, institutional transformation and business development, foreign economic, scientific and innovative) and four groups of indicators of social development (subsistence level, average wages, per capita income), the level of provision of the population with basic goods in housing, education, household, demographic indicators (natural population growth, birth and death rates, migration), the state of labor resources).

The Ministry of Labor and Social Policy of Ukraine (in monitoring the indicators of social development) identifies the following main indicators of socio-economic development: GDP growth, production of consumer goods, retail trade, agricultural output, housing, income and expenditure, the average monthly salary of one employee, wage arrears, the number of registered unemployed, the consumer price index.

Modern methods of measuring UN human development, in addition to demographic, operational and socio-economic, include indicators that meet sustainable development and the Millennium Development Goals: refugee status and victims of military conflict, victim status, energy consumption and pollution, global health risks, inequality in maternal and child health, gender inequality, etc.

References

1. Antoniuk L. L., Satsyk V. I. Production competitiveness of the regions of Ukraine. Current economic problems. 2011. №5 (119), 149-161.
2. Verbytskyi I. Methodical bases of development of system of indicators of strategic planning of social and economic development of regions. Entrepreneurship, economy, law. 2008. № 1, 139–143.

Тетяна Затонацька,
доктор. екон. наук, професор,
Київський національний університет імені Тараса Шевченка
Олена Максимчук,
експерт з маркетингу

МОДЕЛІ ВИКОРИСТАННЯ ІНСТРУМЕНТІВ DATA SCIENCE В ЕЛЕКТРОННІЙ КОМЕРЦІЇ

Ключові слова: електронна комерція, COVID-19, наука про дані, моделі обробки та прогнозування, регресійний аналіз, модель штучної нейронної мережі.

Tetiana Zatonatska,
Professor, Doctor of Economics,
Taras Shevchenko National University of Kyiv
Olena Maksymchuk,
Marketing expert

THE MODELS OF USING DATA SCIENCE TOOLS IN E-COMMERCE

Key words: e-commerce, COVID-19, data science, big data, forecasting models, machine learning, data mining, regression analysis, artificial neural networks.

Nowadays one of the most important areas of economic development is e-commerce which is rapidly penetrating all spheres of human activity that undoubtedly due to the pandemic and the active transition of traditional business to electronic platforms. Within the framework of such realities the economic side of society's life is not only changing, but also the need for modernization is realized. Despite this trends business is focused on using of the latest production technologies and the development of various scientific ideas. The introduction of social distancing, quarantine and other restrictive measures in response to the COVID-19 pandemic has led to an increase in the use of electronic tools and a rapid growth of the e-retail market, so using of Data Science tools in data analysis is particularly relevant.

Therefore, the purpose of study is to determine the current state, prospects and opportunities of implementation data Science tools in the analysis of e-commerce data in Ukraine.

During the research of development and current state of existing Data Science tools were highlighted the features of models and methods of processing large data sets. Number of methods and models of big data processing increases over time and requires innovative technologies, as well as improvement and development of new means of accumulation of data.

One of the innovative areas of implementation Data Science methods is marketing activities. The spread of new marketing technologies and communication management tools, their availability and clarity are leading to rapid growth of competition in the field of e-commerce. All this requires the creation of new competitive advantages, individualization of Internet marketing communications and marketing tools that focused on activities on Internet.

Social networks are an innovative platform for advertising. According to a study by Nielsen [1] about 53% of consumers aged 35-49 years were confident in advertising on social networks during the 3rd quarter of 2020.

Also, according to a study by Nielsen [2], household goods are the third most popular category of consumer goods purchased online. 51% of consumers buy products such as washing powder, paper towels and household cleaning products through online stores. Most products in this category are well suited for sales through the digital channel, as their purchase is common and delivery is usually inexpensive. In turn, this trend has created new opportunities and challenges for suppliers of household chemicals, as it is necessary to improve sales through the e-commerce channel and relevant logistics infrastructure. Participants of household chemical market actively use Data Science tools to increase e-sales. For example, Henkel implements predictive analytics, marketing analysis, production forecasting, and even e-order supply chain monitoring [3]. Despite these trends, using of technologies and Data science tools is particularly relevant for the market of household chemicals and can be implement in different business processes of companies.

During the study, the following methods and models of data processing were considered: Data mining (cluster analysis, regression analysis, classification), machine learning, artificial neural networks, spatial data analysis and visualization, etc. These methods are implemented by companies by different tools such as Hadoop, NoSQL, R, MapReduce, etc.

For the enterprises of the household chemicals market were built mathematical models, which made it possible to improve the marketing activity on the Ukrainian market. Using of a regression model in the e-metrics analysis allowed to determine that such marketing tool as publishing information about promotional activity on social media effectively affects the growth of e-sales, and kind of publications has no direct impact on the number of sales, so excessive spending of financial resources on

a variety of publications is not effective. In course of research of household enterprise's marketing activity the model of an artificial neural network allowed to investigate influence of each marketing factor on electronic sales by means of weight. This model is effective in planning promotional activities, which will allow company to calculate in advance the required value of each of the factors and optimize financial and marketing resources.

Conclusions. For many companies e-commerce is a vital alternative or complementary product sales strategy that keeps the company running smoothly despite a variety of pandemic-related measures.

Existing research shows that using Data science tools can enable e-commerce companies to reap a number of benefits, such as: improving pricing strategies for goods and services; creation of targeted advertising; improving the link between research and development of new products; improving customer service, etc.

References

1. The Nielsen Total Audience Report Hub. Nielsen. 2021. URL: <https://www.nielsen.com/us/en/insights/article/2020/the-nielsen-total-audience-report-hub/>
2. Nielsen Annual Marketing Report: Era Of Adaptation. Nielsen. 2021. URL: <https://www.nielsen.com/us/en/insights/report/2021/nielsen-annual-marketing-report-era-of-adaptation/>
3. Big data has never been as exciting as it is today. Henkel. 2018. URL: <https://www.henkel.com/spotlight/features/industry-4-0/interview-thomas-zeutschler-769078>

Лариса Зомчак,
канд. екон. наук, доцент,
Львівський національний університет імені Івана Франка
Лілія Коваль,
здобувачка другого (магістерського) рівня вищої освіти,
Львівський національний університет імені Івана Франка

ПРОСТОРОВЕ МОДЕЛЮВАННЯ СТАЛОГО ТЕРИТОРІАЛЬНОГО РОЗВИТКУ УКРАЇНИ

Ключові слова: просторове моделювання, сталий розвиток, панельні дані, регіон.

Larysa Zomchak,
PhD, Associate Professor,
Ivan Franko National University of Lviv
Lilia Koval,
Master's degree student,
Ivan Franko National University of Lviv

SPATIAL MODELLING OF UKRAINE SUSTAINABLE TERRITORIAL DEVELOPMENT

Key words: spatial modelling, sustainable development, panel data, region.

Поняття сталого розвитку за останні роки набуло значного поширення і уже мало хто береться заперечувати критичну необхідність саме такого підходу до подальшого світового розвитку. Важливим є цілісний та науковий підхід до моделювання регіональних процесів. Дослідження спрямоване на виявлення основних причин забруднення регіонів та формування рекомендацій щодо подальшої політики та реформ в Україні на регіональному рівні.

Для економіко-математичного моделювання територіального розвитку на рівні регіонів дослідники найчастіше застосовують методи симульвативного моделювання [1], кластерного аналізу [2, 3], панельного моделювання [4], методи машинного навчання [5] тощо.

Для побудови моделей використано статистичну інформацію Державної служби статистики України про забруднення навколишнього середовища, капітальні інвестиції, ВРП, кількості населення та витрат на збереження навколишнього середовища за 2004 – 2019 роки в розрізі областей України [6].

Оскільки вхідна статистика має панельну структуру, то доцільно застосувати методи панельного моделювання, а для врахування особливостей кожної із областей ще й просторове моделювання. Відтак, у дослідженні використано метод просторово-панельного моделювання. Зміст просторового моделювання полягає у визначенні просторових ефектів, тобто як саме процеси в сусідніх регіонах впливають на процеси та показники досліджуваного регіону.

Одним з головних компонентів просторового моделювання є просторова вагова матриця. У дослідженні матрицю збудовано на основі суміжності, тобто матриця визначає, які регіони межують між собою.

За допомогою тесту Хаусмана обрано просторову модель із фіксованими ефектами ($p_{value}=2.2 \cdot 10^{-16} < 0,05$). Для виявлення просторового лагу використано тест множників Лагранжа, який підтвердив наявність просторового лагу ($p_{value}=3.85 \cdot 10^{-12} < 0,05$).

Тести проведено для трьох найбільш поширених просторових моделей: SAR (просторова авторегресійна модель), SEM (просторова модель помилок) та SARAR (просторова авторегресійна комбінована модель). Для усіх проведених тестів $p_{value} < 0,05$, тому, незалежно від вибору моделі на наступних кроках, можуть застосовуватись моделі з фіксованими ефектами та просторовими лагами.

Побудовано SAR та SEM моделі залежності забруднення в регіоні від забруднення в інших регіонах, валового регіонального продукту, капітальних інвестицій, витрат на збереження навколишнього середовища, кількості населення та фіксованого ефекту регіону. Виявлено досить сильний вплив просторових ефектів на рівень забруднення регіонів України, про що свідчать коефіцієнти просторової авторегресії (наприклад, $\rho = 0.3261$ для SAR-моделі). При чому, коефіцієнти обидвох моделей статично значущі із $p_{value} < 0,001$. Причиною таких впливів можуть бути просторові дифузійні ефекти.

На основі реалізованих моделей було виділено основні напрямки політики, серед яких зосередження інвестицій в «зелених» інноваціях, зростання зацікавлення молоді до природничих та технічних наук, справедливості в розподілі коштів між регіонами, посилення контролю за екологічним слідом регіонів, подолання корупції та створення умов для поширення технологічних та інноваційних знань між регіонами, що пришвидшить темпи розвитку регіонів тощо.

Список використаних джерел

1. Зомчак Л. М., Волошин І. Б. Симультаивна модель інноваційного розвитку регіону. Глобальні та національні проблеми економіки. 2016. Випуск 2. С. 854–858. URL: <http://global-national.in.ua/archive/9-2016/174.pdf>.
2. Babenko, V., Zomchak, L., & Nehrey, M. (2021). Ecological and economic aspects of sustainable development of Ukrainian regions. In *E3S Web of Conferences* (Vol. 280). EDP Sciences.
3. Вдовин М, Зомчак Л, Боднар О. (2021). Кластеризація країн ЦСЄ за показниками ЗЕД. Економіка та суспільство. 2021. Випуск 26. URL: <https://www.economyandsociety.in.ua/index.php/journal/article/download/387/374>
4. Khalid, K., Usman, M., & Mehdi, M. A. (2021). The determinants of environmental quality in the SAARC region: a spatial heterogeneous panel data approach. *Environmental Science and Pollution Research*, 28(6), 6422-6436.
5. Babenko, V., Panchyshyn, A., Zomchak, L., Artym-Drohomyretska, Z., Lahotskyi, T., & Nehrey, M. (2021). Classical machine learning methods in economics research: Macro and micro level examples. *WSEAS Transactions on Business and Economics*, 18, 209-217.
6. Державна служба статистики України: веб-сайт. URL: <http://www.ukrstat.gov.ua> (дата звернення: 01.08.2021)

Андрій Камінський,
доктор економічних наук, професор,
Київський національний університет імені Тараса Шевченка

АЛЬТЕРНАТИВНІ ДАНІ В КРЕДИТНОМУ РИЗИК-МЕНЕДЖМЕНТІ: НОВИЙ ПІДХІД ДО ЗАСТОСУВАННЯ

Ключові слова: кредитний ризик-менеджмент, обсяг та якість даних, алгоритми, data mining, скоринг доходності

Andrii Kaminskyi,
Dr. Sc., Professor,
Taras Shevchenko National University of Kyiv

ALTERNATIVE DATA IN CREDIT RISK-MANAGEMENT: NEW APPROACH FOR APPLICATION

Key words: credit risk-management, data volume and quality, algorithm, data mining, profitability scoring

Кредитний ризик-менеджмент отримав значний поштовх при впровадженні онлайн технологій споживчого кредитування. Даний тип кредитування, з одного боку, потребує адаптації класичних моделей обробки даних про позичальника та їх використання у ризик-менеджменті. Так, якщо банки регулятивно повинні мати візуальний перший контакт з позичальником, то фінансові компанії, які здійснюють онлайн кредитування, це робити не зобов'язані. З іншого боку, онлайн технології кредитування та стрімкий розвиток технологій дозволяє залучити до оцінки позичальника багато нових даних. Питання, яке аналізується в нашому дослідженні, полягає в оцінці ефективності напрямків, в яких використовуються альтернативні дані. Обґрунтовується підхід використання цих даних для побудови скорингів доходності онлайн позичальників, які доповнюють класичний підхід ризик-менеджменту.

Традиційні (класичні) дані, які застосовуються в кредитному ризик менеджменті включають в себе соціально-демографічні, професіонально-кваліфікаційні характеристики, показники добробуту та поведінкові характеристики. Останні виражені, перш за все, інформацією, яка міститься в бюро кредитних історій.

Альтернативними даними виступають всі інші дані, такі як інформація від операторів мобільного зв'язку, поведінка на сайті кредитора, психометричні скоринги, інформація з соціальних мереж тощо.

Наше дослідження стосувалося онлайн-кредитування. Даний сегмент активно розвивається в Україні. За даними НБУ у 2020 році було укладено онлайн близько 10 млн. договорів. За всіма оцінками даний ринок буде продовжувати активно розвиватися. Основним сегментом в структурі онлайн-кредитування є сегмент коротких кредитів. Даний сегмент включає кредити на

термін до 30 днів та розмір в межах 1000-5000 грн. Хоча, надійним позичальникам може бути виданий кредит більшого розміру.

Сегмент коротких онлайн кредитів характеризується високим рівнем ризику та високою доходністю. Така доходність, за нашими дослідженнями породжується двома класами позичальників. Перший клас позичальників складають «повторники». Це позичальники, які дуже часто беруть кредити повторно (5-20 разів на рік). Другий клас позичальників, який генерує високу доходність, включає тих, хто виходить постійно у прострочення, у пролонгацію, та потім сплачує всі пені, штрафи та комісії. Разом з тим, у цього другого класу традиційні характеристики практично нічим не відрізняються від характеристик позичальників, які виходять в прострочення, але не погашають повністю заборгованість. У кредитній історії це відбивається однаковими показниками – наявністю прострочень. Тому нами було взята за основу кластеризація позичальників на основі доходності. Даний підхід є розвитком ідеї, представленої в роботі [1]. Візуальним відображенням такої кластеризації виступають аналоги кривих Стобахова (які також мають назву Whale curves). Крива дозволяє структурувати позичальників на 5 кластерів: A1, A2, B, C, D. Вигляд та опис кластерів подані на Рис.1. Розгорнутий аналіз даного аспекту розглядається в [2].



Рис.1. Кластеризація позичальників на основі доходності

Задача оцінки до якого кластеру належить позичальник є вельми значимою, тому що позичальники відрізняються доходністю. Тому, відповідно, має відрізнятися кредитне рішення, яке буде поєднанням доходності та ризику. Для цього, побудовані скорингові рішення для всіх кластерів, тобто визначаються ймовірності приналежності позичальника до кожного з кластерів. І саме використання альтернативних даних для побудови цих скорингів виявилось ефективним інструментом. Традиційні дані (особливо інформація з бюро кредитних історій) добре прогнозують ризик прострочення. Доходність позичальника в нашому дослідженні краще моделюється альтернативними

даними. Таким чином, оцінка ризику позичальника «класичним» скорингом доповнюється 5 скорингами, які характеризують ймовірність приналежності до відповідного кластеру.

На Рис.2 показана схема розширення класичного скорингового підходу прийняття кредитних рішень. Оцінка ризику доповнюється оцінками доходності. Точка відтинання замінюється на матричне рішення. На Рис.2 відмічено сірим кольором сегменти, в яких можливе позитивне рішення щодо видачі короткого онлайн-кредиту

Скоринги доходності, побудовані на альтернативних даних. Ймовірності приналежності кластерам.	P(A1)							
	P(A2)							
	P(B)							
	P(C)							
	P(D)							
		R1	R2	...	...	...	...	RK
	Класи скорингу кредитного ризику, побудованого на традиційних даних							

Рис.2. Розширення системи кредитного ризик-менеджменту скорингами доходності позичальників

Таким чином, проведене дослідження показує, що альтернативні дані можуть бути ефективно застосовані для побудови скорингів доходності, що доповнюють класичний скоринг ризику, який будується на традиційних даних. Для класичний скорингу ризику ефективними є традиційні дані, особливо дані з кредитної історії позичальника.

Список використаних джерел

1. Van Raaij, E. M., Vernooij, M. J. A., Vernooij, M. J. A., & van Triest, S. P. (2003). The implementation of customer profitability analysis: a case study. *Industrial marketing management*, 32(7), 573-583. ([https://doi.org/10.1016/S0019-8501\(03\)00006-3](https://doi.org/10.1016/S0019-8501(03)00006-3))
2. Kaminskyi A., Nehrey M. Clustering approach to analysis of the credit risk and profitability for nonbank lender. In: *Machine Learning Methods and Models, Predictive Analytics and Applications//Proceedings of the Workshop on the XIII International Scientific Practical Conference Modern problems of social and economic systems modelling (MPSESM-W 2021)*. Kharkiv, Ukraine, April 9, 2021., pp.125-136. (<http://ceur-ws.org/Vol-2927/paper10.pdf>)

Катерина Кононова,
доктор економічних наук, професор,
Харківський національний університет ім. В.Н. Каразіна

ПРАКТИКА МАШИННОГО НАВЧАННЯ: ОЧІКУВАННЯ ТА РЕАЛЬНІСТЬ

Ключові слова: машинне навчання, обсяг та якість даних, обчислювальні потужності, алгоритми, дата-саєнтист

Kateryna Kononova,
D.Sc., professor,
Karazin Kharkiv National University

MACHINE LEARNING PRACTICE: EXPECTATIONS VS REALITY

Keywords: machine learning, volume and quality of data, computing resources, algorithms, data scientist

Зростання обсягів даних, що зберігаються, а також суттєве здешевлення обчислювальних потужностей призвели до сплеску академічних досліджень, а також масового впровадження алгоритмів штучного інтелекту (зокрема, машинного навчання) у виробничі та бізнес-процеси. Безліч навчальних платформ пропонують швидко опанувати базові алгоритми машинного навчання (ML), а сотні компаній – оптимізувати технологічні та економічні процеси з використанням цих методів. Як наслідок, виникає ілюзія легкого входу до ринку та завищені очікування. На практиці виявляється, що такі проекти тривають надзвичайно довго, коштують дорого і далеко не завжди ефективні (за статистикою, лише близько 7 % проектів ML вважають успішними). Чому так відбувається?

1. **Команда.** Подібні проекти передбачають наявність команди, у складі якої працюють висококваліфіковані фахівці (принаймні аналітик, дата-інженер і дата-саєнтист). Вони мають володіти не тільки професійними навичками, але й занурюватися в предметну галузь, що також є досить затратним (і в часовому, і у грошовому еквіваленті).

2. **Дані.** Моделі ML дуже чутливі до обсягів та якості даних. Саме тут найбільш наочно спрацьовує правило «сміття на вході – сміття на виході».

а. Збір даних – процес не тільки тривалий, але й дуже витратний. Зберігання та підтримання актуальності даних потребує правильного проектування та надійних сховищ.

б. Забезпечення якості даних – ще більш нетривіальне завдання. Крім питань попередньої підготовки даних (зазвичай найтриваліший етап у всьому проекті), тут може виникнути чимало додаткових проблем – від розмітки до етики даних.

3. **Обчислювальні потужності.** Сучасні алгоритми потребують значних обчислювальних потужностей, не маючи яких експериментування щодо вибору оптимальної моделі триватиме невиправдано довго.

4. **Алгоритми.** Вибір алгоритмів варто обумовлювати не їхньою складністю або новизною, а специфікою завдання. Під час розв'язання багатьох прикладних завдань методи статистичного моделювання зазвичай

спрацьовують ефективніше за алгоритми ML. Крім того, більшість алгоритмів використовує принцип чорного ящика, що значно ускладнює інтерпретацію результатів. Тому останнім часом дедалі більшої популярності набувають так звані whiteboxing методи [1].

Незважаючи на всі перераховані труднощі, штучний інтелект – один із тих прикладних наукових напрямів, що розвиваються якнайшвидше. Саме тут виникають найсміливіші ідеї та з'являються нетривіальні рішення (рис. 1). Generative AI, Adaptive ML, Generative Adversarial networks, Self-supervised learning обіцяють новий етап у розвитку галузі.

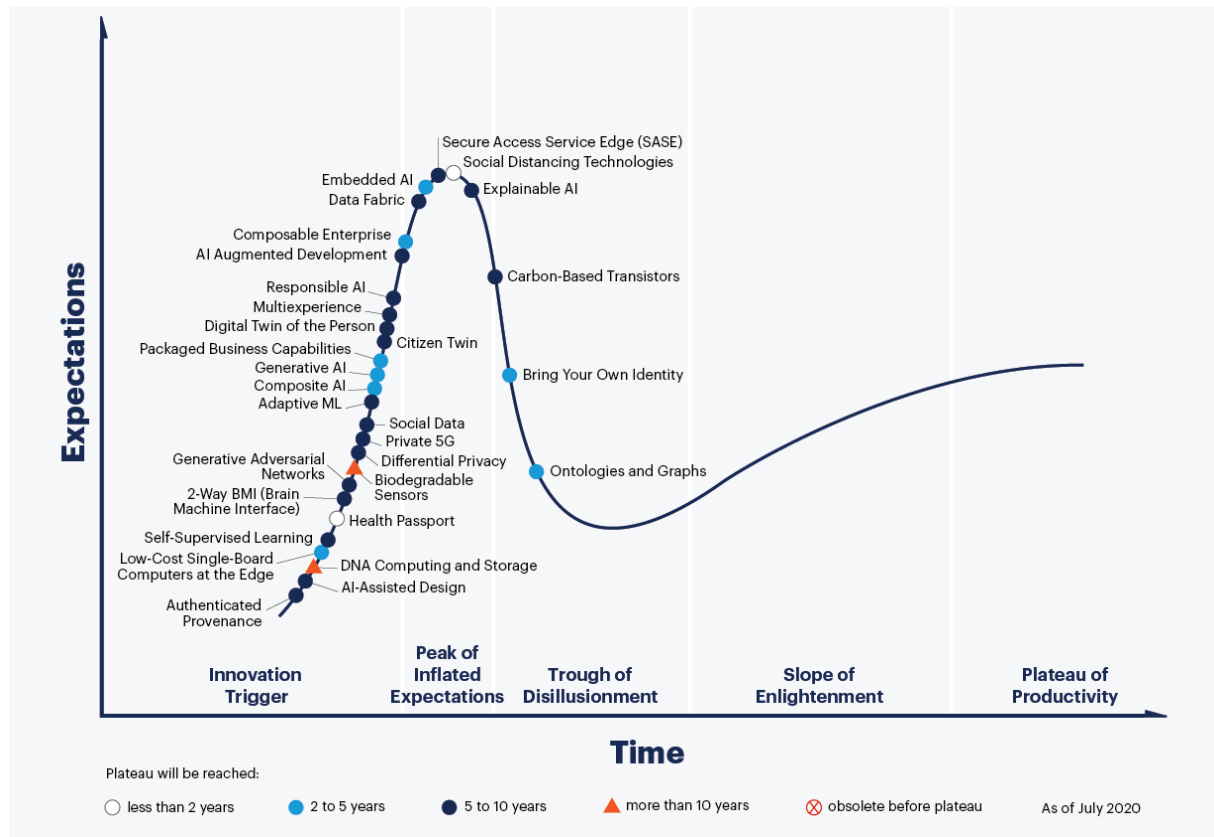


Рис. 1. Життєвий цикл новітніх технологій [2]

У цілому, використання методів машинного навчання виправдане за таких умов: глибоке розуміння процесів, чітка постановка завдання, наявність команди фахівців, даних, обчислювальних потужностей та бюджету. В іншому випадку покладатися на штучний інтелект не варто.

Список використаних джерел

1. Molnar, Ch. Interpretable machine learning. A Guide for Making Black Box Models Explainable, 2019. URL: <https://christophm.github.io/interpretable-ml-book/>
2. The Gartner Hype Cycle for Emerging Technologies, Gartner : website. URL: <https://www.gartner.com/smarterwithgartner/5-trends-drive-the-gartner-hype-cycle-for-emerging-technologies-2020/> (Last accessed: 15.06.2021)

Олена Ляшенко,
доктор економічних наук, професор,
Тетяна Кравець,
к.ф.-м.н., доцент,
Анастасія Філогіна,
студентка 2 курсу магістратури,
Київський національний університет імені Тараса Шевченка

ЗАСТОСУВАННЯ СКЛАДНИХ НЕЙРОННИХ МЕРЕЖ ПРИ МОДЕЛЮВАННІ ВОЛАТИЛЬНОСТІ ФОНДОВИХ ІНДЕКСІВ

Ключові слова: волатильність фондових індексів, моделювання волатильності, складні мережі

Olena Liashenko,
D.Sc., professor,
Tetyana Kravets,
PhD, Associate Professor,
Anastasiya Filogina,
Student,
Taras Shevchenko National University of Kyiv

APPLICATION OF COMPLEX NEURAL NETWORKS IN STOCK INDICES VOLATILITY MODELING

Keywords: stock indices volatility, volatility modeling, complex networks

Дослідження часових рядів, які є однією з найпоширеніших форм даних, широко використовується в різних сферах. Обробка і аналіз великих обсягів даних залишається складним завданням. Стиснення багатовимірної інформації та часових рядів для опису їх мінливості та вивчення еволюційного динамічного механізму моделей волатильності привертає значну увагу дослідників. У нашому дослідженні увага приділяється застосуванню складних мереж для вирішення зазначеної проблеми при розгляді фінансових ринків. Волатильність фінансових систем відображає їх невизначеність, але також може і використовуватися як інструмент для моделювання і прогнозування. За допомогою отриманих мереж виявляються особливі вузли, що можуть слугувати індикатором виходу ринку у нетиповий, нестабільний стан.

Аналіз функціонування фондових ринків є одним із найбільш важливих питань оцінки функціонування економічної системи. Фондові індекси ж належать до основних інструментів аналізу поточної ситуації на ринку та в країні в цілому. Вони використовуються в моделюванні як вихідні дані, а також їх динаміка є об'єктом зацікавленості інвесторів для побудови інвестиційних портфелів тощо.

Моделювання фінансових ринків, в тому числі і фондових, за допомогою теорії складних мереж являє собою новий напрям досліджень. У прикладних

дослідженнях зазвичай застосовують такі типи для мережевого аналізу характеристики, як розмір мережі, мережева щільність, ступінь центральності тощо [1].

Дослідження базується на даних з 2006 по 2021 рік, при цьому використовуються методи як фундаментального, так і технічного аналізу. За даними біржових котирувань були отримані такі показники, що відображають волатильність: довжину реального тіла, колір реального тіла, а також довжину верхньої та нижньої тіні. Далі були створені елементи тіла та масштабу, що були поділені на інтервали, згідно з роботою [2]. Значення тіла були поділені на 4 інтервали: (0, 1%), [1%, 3%], (3%, 7%), [7%, 1), а значення масштабу - на 3 групи: (0, 30%) [30%, 70%] (70%, 1) відповідно до розподілу реального тіла. Відповідно до даного розподілу, різним станам були надані відповідні маркування, що відображають характер волатильності індексу за останні п'ять днів.

У дослідженні були створені орієнтовані зважені мережі, використовуючи методологію перетворення часових рядів у мережу шляхом дослідження еволюційних взаємозв'язків між моделями волатильності [3], для індексів Shanghai Composite, S&P500, DAX30, CAC40, Nikkei225, FTSE100 та IBEX35 за даними порталу Investing [4]. В мережах відображається однакова типова для всіх індексів структура: ядро та петлі. Дана структура відображає той факт, що фондовий індекс загалом розвивається навколо певної групи станів, що і складають собою ядро, і інколи трапляються аномалії, що і мають вигляд петель у мережах.

Для визначення вузлів, що входять до ядра, були розглянуті статистичні характеристики зваженого ступеня вузлів, що дозволило виявити ключові шаблони волатильності на ринку. Близько 65-80% інформації про волатильність індексу міститься у 10% ключових вузлів, що відображає собою принцип Парето. Саме ці вузли і визначаються як ядро мережі.

Відповідно до розподілу зважених ступенів вузлів були виділені також вузли, що входять до петель, і розраховані довжини цих петель. Було виявлено, що близько 80% всіх петель мають довжину 1-7 днів, 10-15% - завершуються протягом 8-15 днів, до 5% - за 16-30 днів. Це говорить про те, що існує 80% ймовірність, що у разі виходу в петлю ринок повернеться до ядра за тиждень, та 95% ймовірність – що за 15 днів. Петлі, з яких ринок не виходить за місяць, а зокрема, за 50 і більше днів, були окремо розглянуті як довгі петлі. На графіку нижче відображені найдовші петлі для індексу S&P. Як видно, нестабільність початку епідемії COVID-19 сформувала останню довгу петлю на ринку.

Кумулятивний розподіл посередництва вузлів у всій мережі використовується для визначення ключових медіа-вузлів. Ці вузли є лише проміжними між петлями та ядром. Близько 30% всіх вузлів мають кумулятивне посередництво близько 70%, тому можна говорити, що є певні вузли, чия медіальність може значною мірою представляти здатність всієї мережі. Також є такі вузли, значення посередництва яких в ядрі та у всій мережі загалом можуть істотно відрізнятись, і деякі вузли, центральність яких в ядровій структурі є нульовою, однак у всій мережі вони є ключовими медіа-вузлами і мають відносно високе посередництво. Ці вузли є «кінцевими

точками» еволюції в структурі ядра, а також «точками початку» петель. Тому у всій мережі вони є важливими засобами для поєднання ядра і петель. На відміну від інших медіа-вузлів, ці вузли діють лише як мости між ядром і петлями. Таким чином можна говорити, що поява цих вузлів є небезпечною, оскільки показує, що на ринку відбудуться певні відхилення від ядрових, звичних станів.



Рис. 1. Періоди найдовших петель для індексу S&P

Джерело: складено авторами на основі даних з порталу Investing [4]

В дослідженні технічні інструменти поєднувалися з фундаментальними, що допомагає з'ясовувати причини знайдених найбільших збурень на ринках. Найдовші петлі припадають на кризовий період осені 2008 року та період кризи, пов'язаної з COVID-19. Загальна тенденція появи довгих петель саме на спадах, що легко пояснюється психологічними факторами, зокрема – панікою трейдерів.

Застосування наведеного методу до фондового ринку може дати багато цінної інформації для інвесторів для прийняття рішень та попередження ризиків, включаючи механізми трансформації ринкової структури та тривалість нестабільності.

Список використаних джерел

1. L. da F. Costa, F. A. Rodrigues, G. Travieso, P. R. Villas Boas. Characterization of complex networks: A survey of measurements. — *Advances in Physics*, Vol. 56, Is. 1, 2007. — p. 167–242. Available at: <https://doi.org/10.1080/00018730601170527>
2. J. H. Fock, C. Klein, B. Zwergel. Performance of Candlestick Analysis on Intraday Futures Data — *The Journal of Derivatives*, Vol. 13, Is. 1, 2005. — p. 28–40. Available at: <https://doi.org/10.3905/jod.2005.580514>
3. S. Liu et al. Modeling the Complex Network of Multidimensional Information Time Series to Characterize the Volatility Pattern Evolution, *IEEE Access*, vol. 6, pp. 29088-29097, 2018, doi: 10.1109/ACCESS.2018.2842069.
4. Офіційний сайт котирування цін на світових біржах Investing [Електронний ресурс]. — Режим доступу: <https://investing.com>

Марина Негрей,
канд. екон. наук, доцент,
Національний університет біоресурсів
і природокористування України

DATA SCIENCE B FOODTECH: ВИГОДИ ТА ОСНОВНІ ТЕНДЕНЦІЇ

Ключові слова: продовольча безпека, data mining, FoodTech, диджиталізація, глобальні стартапи

Maryna Nehrey,
PhD, Associate Professor,
National University of Life and Environment Science of Ukraine

DATA SCIENCE IN FOODTECH: BENEFITS AND THE MAIN TRENDS

Keywords: food security, data mining, FoodTech, digitalization, global startup

Питання продовольчої безпеки, розвиток сталих систем агровиробництва та переробки харчових продуктів є актуальними для світової та вітчизняної наукової та політичної спільноти. Особливо гостро проблема продовольчої безпеки розкрилася в період пандемії COVID-19. Дослідженню даної проблеми присвячена значна кількість сучасних наукових досліджень [1-11].

У світовій практиці формування нових виробничих, логістичних, збутових рішень у сфері продовольства вирішується засобами поступової імплементації екосистеми FoodTech, тобто інтеграції цифрових технологій; автоматизації та роботизації технологічних процесів; застосуванні мережевих ринкових моделей взаємодії по всьому харчовому ланцюгу створення цінності.

Термін FoodTech часто позначає послуги доставки продуктів через Інтернет з магазинів та ресторанів. Однак поняття FoodTech набагато ширше – це інтеграція цифрових та біотехнологій у весь харчовий ланцюг, від ферм та виробництва продуктів харчування до упаковки, зберігання, приготування та утилізації продуктів харчування.

Застосування інструментарію Data Science у сфері FoodTech дає додаткові вигоди, зокрема:

- збільшення ефективності: прогностичних погодних звітів, прискорення on-time доставки, інформації про транспорт, моніторингу цін;
- контроль якості: контроль повного ланцюга поставок, сканування якості вхідних матеріалів;
- удосконалення інформації: прогностичної аналітики, аналіз відгуків клієнтів, аналіз настроїв клієнтів.

Серед основних трендів FoodTech можна виділити наступне:

- Стартапи, що змінюють сільське господарство. Щоб керувати фермами, вони використовують безпілотники, датчики та програмне забезпечення для розробки рішень для покращення сільськогосподарської продукції та якості. AgriTech також застосовується до нової сільськогосподарської продукції, ферм нового покоління та міського сільського господарства.
- Точне землеробство: Стартапи, що підвищують рентабельність фермерських господарств на основі прогнозного аналізу даних сільськогосподарських територій. Вони дають можливість оптимізувати врожайність, зменшуючи при цьому кількість використовуваних продуктів.
- Безпілотники та роботи: Стартапи, які забезпечують фермерів роботами та безпілотниками. Ці інструменти використовуються для збору даних або безпосередньо для заміни людської праці.
- Альтернативні білки: Стартапи, які відповідають виклику глобального попиту на білок у їжу або корм, виготовлений з використанням комах, водоростей або бобових.
- Сільськогосподарські ринки: Стартапи, що розвивають електронну комерцію сільськогосподарського обладнання або економіку співпраці між фермерами.
- Програмне забезпечення для управління фермами: Стартапи, які допомагають фермерам в управлінні, організації та оптимізації всіх завдань у своєму господарстві.
- Ферми майбутнього та міські ферми: Стартапи, які розвивають міські ферми, щоб зменшити відстань між виробництвом та споживанням, або розвивають ферми нового покоління, щоб збільшити врожайність, якість та стійкість.
- Мікроферми: Стартапи, які дозволяють кожному вирощувати свіжі продукти вдома в нових контейнерах або розумних пристроях. Ці початківці полегшують домашнє землеробство, створюючи городи та інші інструменти, придатні для вирощування в кімнатних умовах.

Розвиток FoodTech є особливо актуальним для економіки України, виходячи із внутрішніх та глобальних загроз, зумовлених об'єктивними факторами: зростанням попиту на якісні харчові продукти; погіршення екологічних та кліматичних умов, стану ґрунтів; збитковістю традиційного агровиробництва; поширеністю нових викликів захворювань, зокрема COVID-19.

Список використаних джерел

1. Babenko V., Zomchak L., Nehrey M., Salem A.B.M., Nakisko O. (2022) Agritech Startup Ecosystem in Ukraine: Ideas and Realization. In: Magdi D.A., Helmy Y.K., Mamdouh M., Joshi A. (eds) Digital Transformation Technology.

- Lecture Notes in Networks and Systems, vol 224. Springer, Singapore.
https://doi.org/10.1007/978-981-16-2275-5_19
2. Babenko, V., Nehrey, M. (2021). Digital Agriculture Innovation: Trends and Opportunities. Advanced Trends in ICT for Innovative Business Management. Edited By Katarzyna Szymczyk, Ibrahiem M. M. El Emary Taylor & Francis Books. pp.87-102.
 3. Davydenko, N., Boiko, S., Buriak, A., Demianenko I.(2021) Development of rural areas through fiscal decentralization. Economic Science for Rural Development Conference Proceedings, Issue 55, p. 102-114.
 4. Kaminskyi, A. B., Nehrey, M. V., & Zomchak, L. M. (2021). COVID-19: crisis or new opportunities time for the agricultural sector of Ukraine. In IOP Conference Series: Earth and Environmental Science (Vol. 628, No. 1, p. 012031). IOP Publishing. doi:10.1088/1755-1315/628/1/012031 (Scopus). Available at: <https://iopscience.iop.org/article/10.1088/1755-1315/628/1/012031/meta>
 5. Moro-Visconti, R. (2021). FoodTech and AgriTech Startup Valuation. In Startup Valuation (pp. 363-390). Palgrave Macmillan, Cham.
 6. O'Hara, S., & Toussaint, E. C. (2021). Food access in crisis: Food security and COVID-19. Ecological Economics, 180, 106859.
 7. Qaim, M. (2020). Role of new plant breeding technologies for food security and sustainable agricultural development. Applied Economic Perspectives and Policy, 42(2), 129-150.
 8. Skrypnyk, A., Zhemoyda, O., Klymenko, N., Galaieva, L., & Koval, T. (2021). Econometric Analysis of the Impact of Climate Change on the Sustainability of Agricultural Production in Ukraine. Journal of Ecological Engineering, 22(3).
 9. Udmale, P., Pal, I., Szabo, S., Pramanik, M., & Large, A. (2020). Global food security in the context of COVID-19: A scenario-based exploratory analysis. Progress in Disaster Science, 7, 100120.
 10. Voronenko, I., Skrypnyk, A., Klymenko, N., Zherlitsyn, D., & Starychenko, Y. (2020). Food security risk in Ukraine: assessment and forecast. Agricultural and Resource Economics: International Scientific E-Journal, 6(1868-2021-149), 63-75.
 11. Зомчак Л. Економетричне моделювання функціонування та розвитку аграрного сектору України: монографія / Л.М. Зомчак, М. В. Негрей. – Київ : Компринт, 2018. – 256 с.

Володимир Соловйов,
доктор фізико-математичних наук, професор,
Криворізький державний педагогічний університет
Андрій Матвійчук,
доктор економічних наук, професор,
Київський національний університет імені Вадима Гетьмана
Андрій Белінський,
студент фізико-математичного факультету,
Криворізький державний педагогічний університет

ЕКОНОФІЗИЧНІ ПІДХОДИ В МОДЕЛЮВАННІ ФІНАНСОВИХ РИНКІВ

Ключові слова: міждисциплінарність, еконофізика, складність, ентропія, хаос, необоротність, невірноважність

Vladimir Soloviev,
D.Sc., professor,
Kryvyi Rih State Pedagogical University
Andriy Matviychuk,
D.Sc., professor,
Kyiv National Economic University
Andrii Bielinskyi,
Student of the Physics and Mathematics Faculty,
Kryvyi Rih State Pedagogical University

ECONOPHYSICS APPROACHES IN FINANCIAL MARKET MODELING

Key words: interdisciplinarity, econophysics, complexity, entropy, chaos, irreversibility, non-equilibrium

Financial markets have been attracting the attention of many scientists like engineers, mathematicians, physicists, and others for the last two decades. Such vast interest transformed into a branch of statistical mechanics — econophysics [1].

This study is dedicated to the description of the key paradigm in this interdisciplinary research: informational, (multi-) fractal, chaos-dynamic, recurrent, irreversible, based on complex networks, quantum.

Complexity is a multifaceted concept, related to the degree of organization of systems. Patterns of complex organization and behavior are identified in all kinds of systems in nature and technology. Essential for the characterization of complexity is its quantification, the introduction of complexity measures, or descriptors. Also, we may measure the complexity of the system by using entropy indicators. After the fundamental paper of Shannon [2] in the context of information theory, where entropy denoted the average amount of information contained in the message, its notion was significantly redefined. After this, it has been evolved along with different ways and successful enough used for the research of economic systems [3-4].

The economic phenomena that cannot be explained by the traditional efficient market hypothesis can be explained by the fractal theory proposed by Mandelbrot [5]. Multifractals are a type of fractal, but they stand in contrast to the monofractals. The point of a multifractal analysis is to construct the distribution of different fractal dimensions of the system.

The evolution of the system exhibits sensitive dependence on initial conditions. It means that initially close trajectories that evolve may rapidly diverge from each other and have totally different outcomes. Lyapunov exponent is a measure of the exponential rate of nearby trajectories in the phase-space of a dynamical system. In other words, it quantifies how fast converge or diverge trajectories that start close to each other, quantifying the strength of chaos in the system. Using this measure, it is possible to define chaotic region, including a possible transition between unstable and stable periods.

The CLT [6], which offers the fundamental justification for approximate normality, points to the importance of alpha-stable distribution: they are the only limiting laws of normalized sums of independent, identically distributed random variables. Gaussian distributions, the best-known member of the stable family, have long been well understood and widely used in all sorts of problems. However, they do not allow for large fluctuations and are thus inadequate for modeling high variability. Non-Gaussian stable models, on the other hand, do not share such limitations. In general, the upper and lower tails of their distributions decrease like a power function. Consequently, a probability model with a power tail can be suitable for identifying processes with extreme events. It was discovered that alpha-stable distributions fit better than the Gaussian distribution to financial and spot markets. It is still debatable whether Lévy stable distribution is applicable since there is not enough theoretical material, and there is not a universal analyzing method for estimating parameters of Lévy stable distribution.

Another approach deals with recurrence behavior in the system. Recurrence plot (RP) [7] have been introduced to study dynamics and recurrence states of complex systems. When we create RP, at first, from recorded time series we reconstruct phase-space trajectory. The dots within RP, representing the time evolution of the trajectories, exhibit characteristic large-scale and small-scale patterns. For a qualitative description of the system, the graphic representation of the system suits perfectly. However, the main disadvantage of graphical representation is that it forces users to subjectively intuit and interpret patterns and structures presented within the recurrence plot. In this case it is better to use recurrence quantification analysis.

The irreversibility of time is a fundamental property of non-equilibrium dissipative systems, and its loss may indicate the development of destructive processes [8]. The irreversibility of time series indicates the presence of nonlinearities in the dynamics of a system far from equilibrium, including non-Gaussian random processes and dissipative chaos.

Complex networks include electrical, transport, information, social, economic, biological, neural, and other networks [9]. The network paradigm has become dominant in the study of complex systems since it allows you to enter new quantitative measures of complexity not existing for the time series.

In our paper [10], we have suggested a new paradigm of complex systems modeling based on the ideas of quantum as well as relativistic mechanics. It has been revealed that the use of quantum-mechanical analogies (such as the uncertainty principle, the notion of the operator, and quantum measurement interpretation) can be applied to describing socio-economic processes.

References

1. Mantegna R., Stanley H. An Introduction to Econophysics: Correlations and Complexity in Finance. Cambridge University Press, 2000. P. 164.
2. Shannon C. E. A mathematical theory of communication, *The Bell System Technical Journal*. 1948. Vol. 27, No. 3. P. 379-423. DOI: 10.1002/j.1538-7305.1948.tb01338.x.
3. Richman J. S., Moorman J. R. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*. 2000. Vol. 278, No. 6. P. H2039-H2049. DOI: 10.1152/ajpheart.2000.278.6.H2039
4. Tsallis C. Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World. New York: Springer-Verlag New York, 2009. 382 p.
5. Mandelbrot B. B., Wheeler J. A. The Fractal Geometry of Nature. *American Journal of Physics*. 1983. Vol. 51, No. 3. P. 286-287. DOI: 10.1119/1.13295
6. Andersen E. S. On the fluctuations of sums of random variables II. *Mathematica Scandinavica*. 1954. Vol. 2. P. 194-222. DOI: 10.7146/math.scand.a-10407
7. Eckmann J. P., Kamphorst S.O., Ruelle D. Recurrence Plots of Dynamical Systems. *Europhysics Letters (EPL)*. 1987. Vol. 4. No. 9. P. 973-977.
8. Costa M., Goldberger A. L., Peng C. K. Multiscale entropy analysis of biological signals. *Phys. Rev. E*. 2005. Vol. 71. No. 2. P. 18. DOI: 10.1103/PhysRevE.71.021906
9. Bianconi G. Interdisciplinary and physics challenges of network theory. *Europhysics Letters (EPL)*. 2015. Vol. 111. No. 5. P. 56001. DOI: 0295-5075/111/56001
10. Soloviev V., Saptsin V. Heisenberg uncertainty principle and economic analogues of basic physical quantities. 2011. arXiv:1111.5289 [physics.gen-ph]

Олександр Подскребко,

канд. екон. наук, доцент,

Київський національний університет імені Тараса Шевченка

ПРОГНОЗУВАННЯ ДИНАМІКИ ЗМІНИ ЦІН НА КРИПТОВАЛЮТУ З ВИКОРИСТАННЯМ АЛГОРИТМІВ ГРАДІЄНТНОГО БУСТІНГУ

Ключові слова: криптовалюта, цифрова валюта, Bitcoin, стартап, проект, прогнозування, ринок криптовалют

Oleksandr Podskrebko,

PhD, Associate Professor,

Taras Shevchenko National University of Kyiv

FORECASTING THE DYNAMICS OF CHANGES IN CRYPTOCURRENCY PRICES USING GRADIENT BOOSTING ALGORITHMS

Keywords: cryptocurrency, digital currency, Bitcoin, startup, project, forecasting, cryptocurrency market

З технологічним розвитком людства виникають нові можливості розв'язання різноманітних проблем, змінюються завдання та інструменти їх вирішення, які супроводжуються потужною конкуренцією. Сьогодні інноваційна діяльність є ареною боротьби багатьох тисяч компаній по всьому світу – як великих корпорацій, так і невеликих стартапів [1]. Всі вони змагаються за можливість стати першими серед тих, хто зможе вивести новий продукт, нову послугу на ринок, і завоювати лояльність клієнтів. Подібні зміни також привели до появи нових засобів розрахунків та інструментів фінансування різноманітних проектів. Ринок криптовалют, як система інвестиційних проектів, у зв'язку з стрімким розвитком технологій та глобалізації, ще потребує додаткового дослідження, особливо це стосується принципів функціонування і подальшого розвитку.

У теперішніх умовах нестабільності світової економіки, швидко мінливих умов ведення бізнесу, появою нових і важко передбачуваних факторів у тому числі пандемії COVID-19, що призводить до зміни настроїв суспільства, недовіри до фінансових установ та підштовхує до зростання інтересу до цифрової валюти та автономних інноваційних децентралізованих фінансових систем. Одним із різновидом цифрової валюти є криптовалюта, за облік внутрішніх розрахункових одиниць якої відповідає децентралізована платіжна система, що дозволяє виключити зовнішнього регулятора з процесу здійснення електронних платежів.

Криптовалютою називають такі валюти, які не мають фізичного вираження. За криптовалюту можна купувати товари та послуги або торгувати ними для отримання прибутку. Зараз в світі налічується більш ніж 10000 різноманітних криптовалют, які торгуються публічно. Загальна вартість усіх криптовалют становить понад 2 трлн доларів, у той час коли вартість усіх

біткойнів, найпопулярнішої цифрової валюти, встановлена на рівні близько 866 млрд. доларів.

Біткойн і цифрові валюти засновані на механізмі розподіленої довіри під назвою “блокчейн” – засобу відстеження довірених системі транзакцій розподіленим чином [2]. У таблиці 1 наведено 5 найбільших за ринковою капіталізацією криптовалют.

Таблиця 1

П'ять найбільших за ринковою капіталізацією крипто валют [3]

Криптовалюта	Ринкова капіталізація, млрд. дол.
Bitcoin	866,3
Ethereum	403,6
Cardano	83,1
Binance Coin	69,9
Tether	68,3

Ринком криптовалют та динамікою зміни цін цікавляться: дослідники, корпорації, пересічні громадяни та інші суб'єкти економіки. Прогнозування ціни на ту чи іншу криптовалюту є достатньо складним процесом, у спрощеному вигляді це може бути регресійна модель (див. форм. 1) у якості факторів якої можуть виступати кількість транзакцій, складність блоку, доступний обсяг цифрової валюти. При збільшенні кількості факторів та кількості спостережень, рекомендується використовувати просунуті алгоритми градієнтного бустінгу (наприклад, Light Gradient Boosted Machine) в поєднанні з feature engineering та перевіркою гіпотез о доцільності врахування тієї чи іншої ознаки.

$$y = f(x_1, x_2, x_3), \quad (1)$$

де: y – ціна криптовалюти; x_1 – кількість транзакцій; x_2 – складність блоку; x_3 – доступний обсяг цифрової валюти.

Таким чином, на ціноутворення цифрової валюти впливають багато факторів, деякі з них не піддаються кількісній оцінці (наприклад, дозвіл купувати певні товари за криптовалюту (Tesla та ін.), або офіційне визнання Сальвадором Bitcoin у якості платіжного засобу).

Список використаних джерел

1. Подскребко, А.О., Іванченко Н.О., О.С. Сідлецька. Основні проблеми та перспективи розвитку ринку стартапів в Україні / Науковий журнал “Бізнес інформ”. – 2019. – Вип.4. – С. 303-311.
2. Ivanchenko N., Borysenko Y. Competitive advantages of insurance companies using blockchain technology / Proceedings of the International Scientific and Practical Internet Conference "Business Intelligence: Models, Methods And Techniques". March 3-5, 2021. - К.: NAU, 2021. P. 75-78
3. Постачальник даних та аналітики криптовалют CoinMarketCap [Електронний ресурс] – Режим доступу до ресурсу: <https://coinmarketcap.com/>.

Ірина Федоренко,

канд. екон. наук, доцент,

Вероніка Сачко,

студентка

Київський національний університет імені Тараса Шевченка

МОДЕЛЮВАННЯ ВПЛИВУ ТЕХНОЛОГІЧНОГО РОЗВИТКУ НА ГЛОБАЛЬНИЙ СТАЛИЙ ЕКОНОМІЧНИЙ РОЗВИТОК

Ключові слова: сталий розвиток, технологічний розвиток, регресійний аналіз, індекс глобальної конкурентоспроможності.

.Iryna Fedorenko,

PhD, Associate Professor,

Veronika Sachko,

Student

Taras Shevchenko National University of Kyiv

MODELING THE INFLUENCE OF TECHNOLOGICAL DEVELOPMENT ON GLOBAL SUSTAINABLE ECONOMIC DEVELOPMENT

Key words: sustainable development, technological development, regression analysis, global competitiveness index.

Питання сталого розвитку країни набуло найбільшої актуальності саме з початку XXI століття, бо в ході ряду проведених досліджень було виявлено, що його рівень має прямий зв'язок зі станом економіки країни, соціальної забезпеченості, стабільності та конкурентоспроможності держави. Термін та концепція сталого розвитку привернула до себе більше уваги після Конференції ООН з довкілля та розвитку у 1992 р. і була представлена в Порядку денному на 21 століття, що був прийнятий під час конференції.

Сталий економічний розвиток є одним із найвагоміших індикаторів рівня конкурентоспроможності держави, а технологічний розвиток наразі є рушієм достатньо значущих змін у багатьох сферах життя суспільства та країни в цілому. Моделювання їх взаємозалежності та зв'язку є достатньо важливим та актуальним питанням на сьогоднішній день. Практична цінність результатів дослідження полягає у можливості їх використання для розробки стратегій підвищення рівня сталого та технологічного розвитку у країнах світу та для моделювання їх впливу та взаємозв'язку.

Сучасна концепція сталого розвитку системно поєднує головні компоненти сталого розвитку суспільства: економічну, екологічну і соціальну, що є нерозривними одна від одної. Сталий економічний розвиток

забезпечується за рахунок сталого економічного зростання. Для аналізу та визначення ступеня сталого розвитку держави використовують достатньо комплексний індикатор, а саме Індекс глобальної конкурентоспроможності (ІГК). Він дає змогу найкраще охарактеризувати саме економічну компоненту, а також показати взаємозалежність сталого та технологічного розвитку. Це пов'язано з тим, що одним із факторів, що входять до складу ІГК, є технологічний прогрес, який характеризує швидкість зростання інноваційних технологій, науково-технічного потенціалу, розвиток галузі розробок та досліджень, а також впливає на загальний розвиток багатьох галузей всередині країни. Технологічні зміни є основною рушійною силою у формуванні основ трансформації економіки. З цього випливає, що інновації трактуються як засіб для досягнення економічного прогресу та зростання. Саме тому питання технологічного розвитку є достатньо важливим при визначенні розвитку держави і потребує належної уваги до себе при розробці урядом комплексу стратегій та реформ задля покращення рівня сталого розвитку. З цього випливає необхідність знаходження оптимального універсального інструментарію для моделювання впливу технологічного на сталий розвиток держави. Для цього можуть бути використані різноманітні макроекономічні моделі, що обираються та будуються згідно з їх призначенням та типом обраних даних.

Для моделювання впливу технологічного розвитку на сталий економічний розвиток було обрано регресійний аналіз, а саме моделі множинної та панельної регресії. В ході їх побудови в якості факторів технологічного розвитку були використані макроекономічні показники та Індекс глобальної конкурентоспроможності як індикатор сталого розвитку. На основі кореляційного аналізу були відібрані показники технологічного розвитку International Property Rights Index, Patents applications by residents, R&D expenditure (% of GDP) та High-technology exports, що мають достатньо тісний зв'язок з ІГК. Регресійні моделі були побудовані для України та п'яти розвинених у технологічному плані країн світу – США, Німеччини, Японії, Китаю та Нідерландів. Для множинної регресії кожна країна мала власне рівняння, а для панельної – модель була однаковою для всіх країн, відрізнялися лише їх індивідуальні ефекти. Для перевірки адекватності та можливості використовувати на практиці моделі були перевірені за допомогою спеціальних статистичних тестів та їх гіпотез, а також було визначено які коефіцієнти є значущими, а які можуть бути виключені з моделі. На основі моделей множинної регресії та моделі панельних даних з фіксованими ефектами був побудований також прогноз, який показав близькість змодельованих значень до реальних. Найбільш високий рівень сталого розвитку виявився у США, Німеччині та Нідерландах.

Для України був обраний оптимальний метод моделювання ІГК, а саме моделі множинної регресії, прогнозовані значення за якою є найбільш подібними до первинних даних індексу, рис. 1.

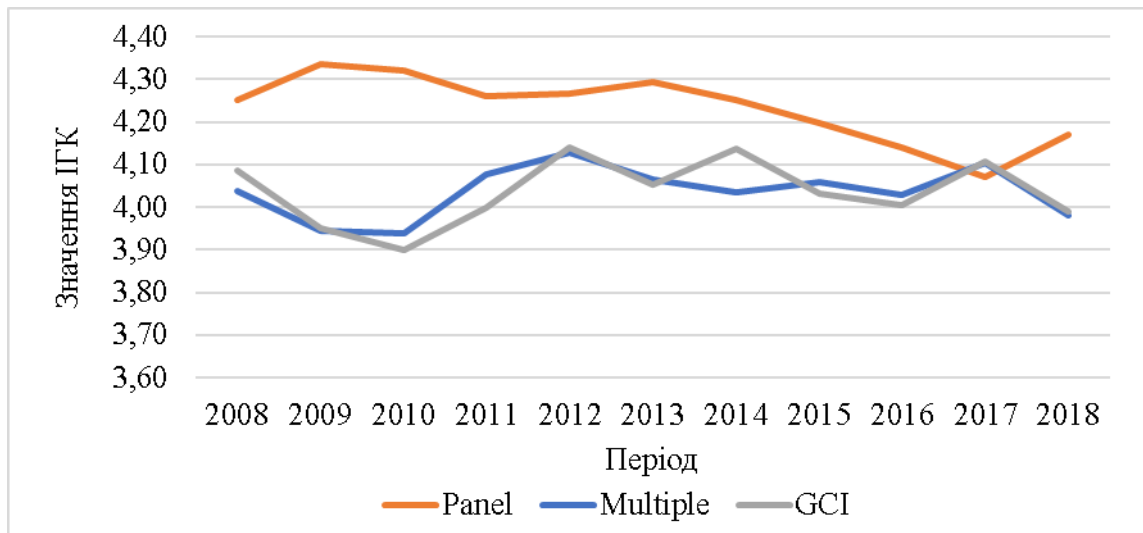


Рис. 1. Початкові та прогнозовані значення ІГК України за 2008-2018 рр.

Практична частина роботи виконувалася за допомогою ПЗ Microsoft Office Excel та Gretl. Для аналізу були використані дані ІГК обраних високорозвинених країн світу та України за 2008-2018 рр. [1-3].

Список використаних джерел

1. World Bank Open Data, електронний доступ: <https://data.worldbank.org/>
2. Economic Data and Statistics on World Economy and Economic Freedom, електронний доступ: <https://www.heritage.org/index/explore>
3. Global Innovation Index, електронний доступ: <https://www.globalinnovationindex.org/home>

Галина Чорноус,
д-р екон. наук, професор,
Яна Фаренюк,
PhD студент,
Київський національний університет імені Тараса Шевченка

МОДЕЛЮВАННЯ МАРКЕТИНГОВОЇ ДІЯЛЬНОСТІ: ПІДХОДИ НА ОСНОВІ ТЕХНОЛОГІЙ DATA SCIENCE

Ключові слова: маркетинг, Data Science, моделювання, маркетинговий комплекс, рекламна активність, ROI, машинне навчання, регресійний аналіз.

Galyna Chornous,
Doctor of Sciences (Economics), Professor
Yana Farenjuk,
PhD student,
Taras Shevchenko National University of Kyiv

MODELING OF MARKETING ACTIVITY: APPROACHES BASED ON DATA SCIENCE TECHNOLOGIES

Key words: marketing, Data Science, modeling, marketing mix, advertising activity, ROI, machine learning, regression analysis.

Базова модель 4P – це широко розповсюджена модель маркетингового комплексу, що складається з 4 основних елементів: product, price, place та promotion. Кожен із них передбачає пошук відповіді на такі запитання: який продукт потребує цільова аудиторія; якою має бути ціна, аби задовольняти виробника і споживача; яка оптимальна модель дистрибуції товару; як має розповсюджуватися інформація та як стимулювати аудиторію до покупки?

Для відповіді на більшість цих запитань на сучасному етапі розвитку маркетингової системи необхідне активне застосування різноманітних Data Science технологій, зокрема методів машинного навчання та економетричного моделювання. Математичне моделювання забезпечує надійний, усесторонній погляд на маркетинг, відображаючи сучасний стан сфери медіа та маркетингу.

Технології Data Science пов'язані з даними та з пошуком у них прихованого знання для формування ефективних бізнес-рішень. Однією з головних переваг є те, що моделювання дає змогу визначити всі драйвери бренду, кількісно оцінити їхні індивідуальні ефекти та врахувати взаємозв'язки між ними для прогнозування майбутньої динаміки. Деякі драйвери при цьому є контрольованими (ціна, дистрибуція чи маркетингова діяльність), а деякі - не контрольованими (погода чи активність конкурента). Технології Data Science створюють можливість дослідити та відобразити, як ці драйвери впливають безпосередньо на бізнес-результати та ефективність компанії (продажі чи інші бізнес-показники).

Моделювання маркетингового міксу (MMM) - це метод розподілу ефекту кожного інструменту маркетингового комплексу та каналу комунікації на цільову метрику та визначення оптимальних показників ефективності для кожного з них. Він дозволяє врахувати фактори в будь-якій комбінації та з високим ступенем точності передбачати, як такі зміни вплинуть на результати бізнесу (реалізувати сценарне прогнозування для бізнес-показників) [1].

MMM передбачає побудову моделі, яка в спрощеному вигляді має наступний вигляд:

$$Sales = Constant + a_1 * Distribution + a_2 * Price + a_3 * Seasonality + a_4 * Adstock(TV) + a_5 * Adstock(Digital) + b_1 * Adstock(TV_Competitors) + b_2 * Adstock(Digital_Competitors) + \varepsilon$$

де *Adstock* - це миттєвий та відкладений ефект від реклами на продажі, тобто розподіл впливу протягом часу: $Adstock(TV)_t = TV_t + a * Adstock(TV)_{t-1}$.

Застосування технологій Data Science відкриває багато можливостей та створює значні переваги для сфери маркетингу та медіа, зокрема:

- детальне розуміння основних рушійних показників ефективності бізнесу та відносного внеску кожного з них [2];
- сценарне прогнозування та бізнес-симуляції;
- оцінка ефективності різних каналів комунікації (розрахунок ROMI – Return on Marketing Investment) та формування оптимального медіа-міксу;
- рекомендації щодо обсягу активності в кожному каналі комунікації та розподілу інвестицій для досягнення бізнес-цілей (доля ринку, продажі тощо);
- оцінка зносу рекламного повідомлення;
- оцінка ефективності креативних матеріалів;
- розуміння Halo-ефекту – як комунікація з одним продуктом компанії впливає на інші (наприклад, вплив реклами молока на продажі масла і навпаки);
- оцінка впливу факторів один на одного, наприклад ціни чи наявності товару на полиці – на віддачу від медіа-активності;
- оптимізація інвестицій протягом року на базі коливань сезонного ROI;
- оцінка ефективності різних архітектур рекламних кампаній (кількість тижнів в ефірі, медійний тиск, перерва між рекламними кампаніями);
- оцінка впливу конкурентів, важливості боротьби за долю в медіа-каналі;
- цінова еластичність та цінова чутливість;
- рекомендації по необхідному рівню дистрибуції;
- портфельна стратегія та регіональне моделювання – оптимізація розподілу інвестицій між продуктами компанії та регіонами.

Підходи аналізу даних та моделювання використовують в сфері маркетингу та реклами, щоб підвищити віддачу від маркетингових інвестицій (їх рентабельність для кожного окремого важеля маркетингу) та відповідно збільшувати продажі та прибуток компанії. В результаті на основі знайдених інсайтів розробляють рекомендації щодо маркетингового міксу, оптимізують портфель брендів і продуктів компанії, визначають необхідний бюджет та медіа-мікс рекламної кампанії, оцінюють вплив конкурентів та прогнозують основні бізнес-показники компанії. Унікальні інструменти та бази даних дозволяються оптимізувати маркетингові бюджети і підвищити ефективність

кожної вкладеної гривні на +5-60%, що підтверджує досвід українських та зарубіжних компаній.

При реалізації проектів моделювання маркетингового міксу за допомогою машинного навчання варто враховувати наступні моменти:

- Доступність даних. Для побудови моделі знадобляться детальні дані по всіх каналах комунікації та факторах впливу. Чим більше точок даних доступно для моделювання, тим більший рівень варіацій буде пояснено та, як наслідок, більш точну модель буде отримано. Якщо немає такої можливості на поточний момент, то варто зосередитися на стратегії роботи з даними.
- Data Science та економетричне моделювання, зокрема - це спеціалізовані сфери, де досвід команди є надзвичайно критичним, адже це ключовий фактор успішної реалізації проекту та пошуку ефективних маркетингових рішень як на стратегічному рівні, так і на рівні тактичних дій.
- Правильне формування припущень. MMM передбачає наявність припущення про забування рекламного повідомлення в показнику «AdStock», тобто припущення, що серед споживачів, які бачать рекламу, спостерігається відставання [3]. Якщо зробити подібне припущення неправильно, можна завищувати чи занижувати вплив певних каналів комунікації зі споживачем.
- Побудова моделі може бути дорогим процесом - як з точки зору прямих витрат, так і з прихованими витратами на збір та підготовку даних, формування припущень тощо. Таким чином, при моделюванні маркетингового міксу необхідно оцінювати фінансові можливості для запуску подібних проектів.

Отже, моделювання маркетингового міксу розглядає історичні зв'язки між маркетинговими витратами та ефективністю бізнесу, щоб допомогти визначити всі бізнес-драйвери і необхідний рівень витрат на маркетингові активності, забезпечуючи при цьому найкращий розподіл на продукти, ринки та маркетингові програми для досягнення бізнес- та комунікаційних цілей компанії. За рахунок реалізації таких проектів існує можливість відповідати на запитання "Що буде, якщо ...?" про ймовірні результати майбутніх змін витрат та на запитання «Як краще?» щодо оптимального способу встановлення та розподілу маркетингових бюджетів. Маркетологи можуть використовувати результати моделювання для коригування маркетингових стратегій, оптимізації маркетингових планів та прогнозування продажів при розгляді різних сценаріїв.

Список використаних джерел

1. Chan D., Perry M. Challenges and Opportunities in Media Mix Modeling / D. Chan, M. Perry // Technical report, Google Inc [Електронний ресурс]. – 2017. – Режим доступу: <https://ai.google/research/pubs/pub45998>.
2. Kucuk S. U. Marketing-Mix Modeling. In: Visualizing Marketing /S. U. Kucuk // Palgrave Macmillan, Cham. – 2017. DOI: https://doi.org/10.1007/978-3-319-48027-5_7.
3. Gijsenberg M. J., van Heerde H. J. Understanding the Role of Adstock in Advertising Decisions / M. J. Gijsenberg, H. J. van Heerde, M. Dekimpe, V. R. Nijs. - 2011. DOI: <http://dx.doi.org/10.2139/ssrn.1905426>.

Віктор Шпирко,

канд. екон. наук, асистент,

Київський національний університет імені Тараса Шевченка

Олександр Бартков'як,

Інвестиційний Аналітик,

ТОВАРИСТВО З ОБМЕЖЕНОЮ ВІДПОВІДАЛЬНІСТЮ "СДКАПІТАЛ"

ЗАСТОСУВАННЯ СТАТИСТИЧНОГО АРБІТРАЖУ ПРИ ТОРГІВЛІ АКЦІЯМИ ТА ДЕРИВАТИВАМИ

Ключові слова: статистичний арбітраж, авторегресія, PCA-аналіз

Viktor Shpyrko

PhD, Assistant of Professor,

Taras Shevchenko National University of Kyiv

Oleksandr Bartkoviak,

Investment Analyst,

LIMITED LIABILITY COMPANY SDCAPITAL

STATISTICAL ARBITRAGE APPLICATION IN EQUITY AND DERIVATIVES TRADING

Key words: Statistical Arbitrage, autoregression, PCA-analysis

Теорія ефективності ринків вказує на те, що ринки є послідовно ефективними та будь-яка подія вже закладена в ціні того чи іншого активу, а отже неправильна оцінка інструменту неможлива. Тим не менше, ринки послідовно демонструють недосконалість, що надає можливість арбітражистам отримувати дохідність, торгуючи на основі перекосів. Зокрема, завдяки популяризації «біржових інвестиційних фондів» (ETF) велика кількість акцій всередині ETF відривається від своєї справедливої оцінки в силу того, що фонд інвестує у весь індекс, створюючи можливості для арбітражистів отримати дохідність на майбутніх, неминучих, корекціях цін окремих компаній [1].

Загалом, останні дослідження вказують на те, що за нинішніх ринкових умов застосування та пошук арбітражів стали досить складними в силу довготривалого росту ринку. Однак, велика кількість можливостей може бути знайдена всередині індексів, зокрема серед компаній з малою капіталізацією, та моментумних акцій, непідкріплених фундаментальними факторами, тощо [2].

Стратегії статистичного арбітражу використовують трейдинговий підхід, що заснований на деяких правилах, які генерують сигнали для входу та виходу з позиції. Загальний підхід до моделювання даних точок отримується з вибору процесу середньої дохідності Орнштейна-Уленбека (OU) [1].

За допомогою моделі авторегресії (AR-1), що є дискретною версією OU-процесу, де $\{a, b, \xi\}$ визначені як у рівнянні (1.2), можна обрахувати регресію для отримання $\{a, b, \text{Var} \{\xi\}\}$ і вирахувати параметри $\{k, m, \sigma^2\}$.

$$X_{n+1} = a + bX_n + \xi_{n+1} \quad (1.1)$$

$$b = \exp(-kdt) \quad a = m(1 - \exp(-kdt)) \quad \text{var}(\xi) = \sigma^2 \frac{1 - \exp(-2kdt)}{2k} \quad (1.2)$$

$$s = -\frac{m\sqrt{2k}}{\sigma} \quad (1.3)$$

де s – нормалізований сигнал, що вказує на кількість одиниць невідповідності з точки зору мінливості процесу OU, а dt - частота генерації сигналу, який може бути виведений з рівності (1.4): вхід в довгу позицію має відбуватися, коли сигнал перетинає точку $s_{long,enter}$ та позиція має утримуватися допоки сигнал не перетне точку $s_{long,exit}$.

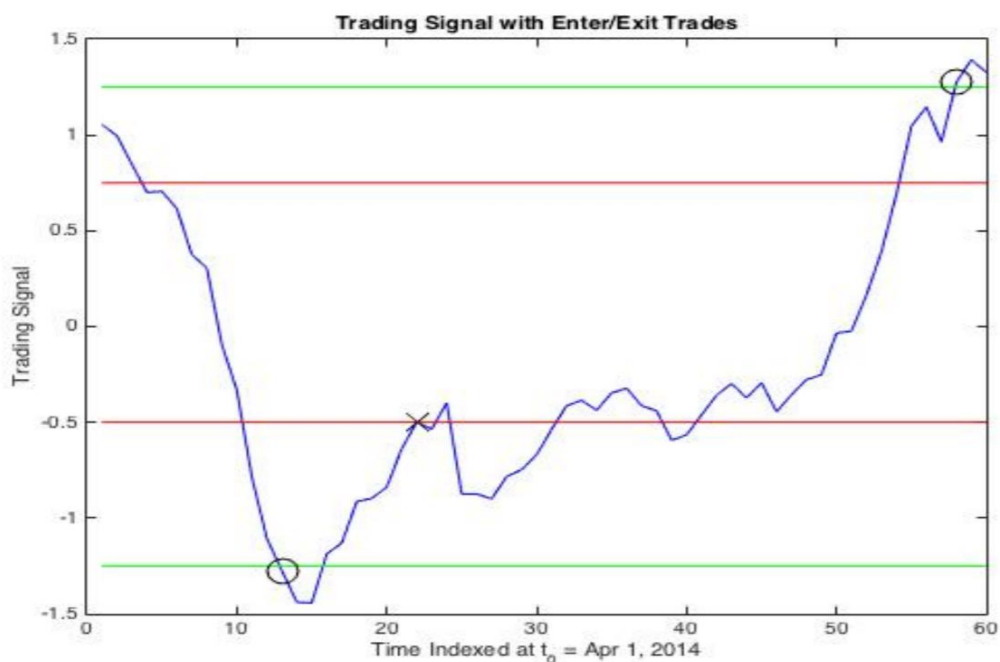


Рис.1. Точки входу та виходу торгових позицій.

Джерело: побудовано автором на основі даних [3].

Відповідний процес має спостерігатися і для коротких позицій:

$$s_{long, enter} = -1.25, s_{long, exit} = -0.5, s_{short, enter} = 1.25, s_{short, exit} = 0.75. \quad (1.4)$$

Одним з найбільш ефективних методів використання даної стратегії є кластерна модель з активами обраними з пулу акцій та деривативів, як найбільш волатильних інструментів, що забезпечує максимізацію дохідності для стратегій з періодом утримання позицій від декількох хвилин до днів.

Кластеризація проводиться на основі комбінації принципового аналізу компонентів (PCA).

Таким чином, моделі статистичного арбітражу надають можливість мінімізувати максимальні збитки, будучи ринково-нейтральними, отримуючи дохідність за будь-яких умов, при ефективному застосуванні з поправками на ринкові умови. Ефективність арбітражу, зокрема, демонструється в періоди падінь: кінець 2016 року – побоювання, щодо вповільнення економіки Китаю, та розпродаж акцій у лютому, в період активного розповсюдження SARS-COV-2 – менеджери, що використовують такі стратегії, отримували вищу дохідність, аніж фонди з довгими стратегіями. Використання моделей також виключає суб'єктивний людський фактор, який є причиною збитків при збільшенні страху на ринках, або надмірної ейфорії.

Список використаних джерел

1. Meucci, A. (2010) Review of Statistical Arbitrage, Cointegration and Multivariate Ornstein-Uhlenbeck
2. Gatev, E., Goetzmann, W.N. and Rouwenhorst, K.G. (2006) Pairs Trading: Performance of a Relative-Value Arbitrage Rule. The Review of Financial Studies, 19, 797-827. <https://doi.org/10.1093/rfs/hhj020>
3. Anran Lu, Atharva Parulekar, Huanzhong Xu (2018) Cluster-Based Statistical Arbitrage Strategy, MS&E 448 Big Data and Algorithmic Trading

Віктор Шпирко,
канд. екон. наук, асистент,
Андрій Винокуров,
студент II курсу магістратури,
Київський національний університет імені Тараса Шевченка

МОДЕЛЮВАННЯ ТА ПРОГНОЗУВАННЯ МІГРАЦІЙНИХ ПРОЦЕСІВ В УКРАЇНІ ЗА ДОПОМОГОЮ ГРАВІТАЦІЙНОГО ПІДХОДУ

Ключові слова: міграційні процеси, країни, часові ряди, прогнозування, регресійна модель, R-квадрат, економічна ситуація.

Viktor Shpyrko,
PhD, Assistant of Professor,
Andrii Vynokurov,
2nd year master's student,
Taras Shevchenko National University of Kyiv

MODELLING AND FORECASTING OF MIGRATION PROCESSES IN UKRAINE USING GRAVITATIONAL APPROACH

Key words: migration processes, countries, data series, forecasting, regression model, R-squared, economic situation.

Міграційні процеси є надзвичайно важливою складовою економічного життя країни. Саме вони показують те, наскільки привабливою та комфортною є країна для життя людей, а отже, наочно демонструють рівень економічного та соціального розвитку держави. Дослідження міграційних потоків між країнами почалося ще дуже давно, а у 40-х роках XX століття Дж. Ціпф [1] запропонував гравітаційну модель міграції, яка виглядає наступним чином:

$$F_{ij} = k \frac{P_i^\alpha P_j^\beta}{D_{ij}^\gamma},$$

де F_{ij} – сила тяжіння між регіонами i та j , P_i – чисельність населення в регіоні вибуття i , P_j – чисельність населення в регіоні прибуття j , D_{ij} – географічна відстань між центрами регіонів i та j , k – коефіцієнт пропорційності, константа. Для того, щоб оцінити цю модель за допомогою багатофакторної лінійної регресії, потрібно прологарифмувати обидві частини цього рівняння, після чого отримаємо наступну формулу:

$$\ln F_{ij} = \ln k + \alpha \ln P_i + \beta \ln P_j - \gamma \ln D_{ij}.$$

На основі цієї формули були побудовані ряд регресійних моделей, дані для яких були взяті з сайту Державної служби статистики України [2]. Для моделі **М1** були використані дані по міграційних потоках України з Польщею, у

моделі **M2** – з Німеччиною, і у моделі **M3** – з Білорусією. Залежною змінною в усіх моделях є міграційний потік між двома країнами, незалежними – населення обох країн та відстань між ними.

Після усіх обрахунків отримуємо наступні результати: найвищий R-квадрат було отримано у моделі **M2** – 0,7; трохи гірший показник можна спостерігати у моделі **M1** – 0,68; а найгірші показники були у моделі **M3** – 0,48.

Також відмітимо, що окрім наведених вище моделей були побудовані й інші, аналогічні за змістом до моделей M1-M3, в яких досліджувалися міграційні потоки України з іншими країнами. Загалом, ці регресії показали середній R-квадрат близько 0,5, що свідчить про помірну, але існуючу залежність між показниками. Усі ці результати дозволяють зробити висновок, що навіть у наші дні можна використовувати гравітаційну модель, розроблену Ціпфом, і аналізувати міграційні процеси саме за її допомогою.

Окрім цього, був побудований прогноз методом експоненційного згладжування на 4 роки вперед для наступних часових рядів: I) міграційний потік між Україною та Польщею; II) кількість людей, прибулих до України. Для обох випадків був обраний коефіцієнт $\alpha = 0,2$. Результати можна побачити у табл. 1.

Таблиця 1

Основні результати проведеного дослідження

Регресійна статистика моделей M1-M3			
Показник	M1	M2	M3
Множинний R	0,75	0,84	0,69
R-квадрат	0,68	0,7	0,48
Нормований R-квадрат	0,54	0,6	0,35
Стандартна похибка	350,64	2617,16	706,98
Кількість спостережень	21	21	21
Прогноз міграційного потоку між Україною та Польщею			
Рік	Міграційний потік (осіб)	Згладжування	
		I	II
2020	993,56	846,59	699,62
2021	1030,30	883,33	736,36
2022	1067,04	920,07	773,10
2023	1103,79	956,82	809,85
Прогноз кількості прибулих до України людей			
Рік	Число прибулих (осіб)	Згладжування	
		I	II
2020	31515,72	35559,66	39603,61
2021	30504,73	34548,68	38592,62
2022	29493,75	33537,69	37581,64
2023	28482,76	32526,71	36570,65

Джерело: побудовано автором на основі даних [2]

Пандемія COVID-19 мала значний вплив на міграційні процеси у всьому світі. Туристичні потоки між Україною та іншими країнами зменшились у декілька разів. Однак, у той самий час, наприклад, Польща та Фінляндія

продовжували приймати наших громадян у якості робочої сили, а люди, які бажали мігрувати до інших країн, продовжують це робити, хоча із певними карантинними обмеженнями.

Підбиваючи підсумки, потрібно сказати, що в Україні дуже гостро стоїть проблема міграції населення. Характерною рисою є саме еміграція, що загалом негативно впливає на економічний стан країни. Пандемія COVID-19 внесла зміни у міграційні процеси, у тому числі ті, що пов'язані з Україною. Але слід зауважити, що епідеміологічна ситуація впливає у першу чергу на туризм, а представлене дослідження вивчає постійну та довготривалу міграцію. А отже, якість прогнозів залишається практично незмінною. Також, після вакцинації від COVID-19 люди зможуть досить вільно подорожувати різними країнами та міграційні процеси поступово відновляться. Актуальними ж проблемами щодо міграції в Україні є підвищення економічної активності населення; недопущення зростання безробіття; забезпечення відповідності між попитом і пропозицією робочої сили. Вирішення цього завдання потребує формування виваженої державної політики з урахуванням регіональної специфіки тенденцій міграції, в тому числі, підвищення конкурентоспроможності робочої сили, забезпечення високого рівня зайнятості, соціального захисту, заробітної плати і доходів населення, їх наближення до аналогічних параметрів економічно розвинутих країн світу.

Список використаних джерел

1. Zipf G.K. Human Behavior and the Principle of Least Effort. Addison-Wesley Press, Reading, Massachusetts, 1949.
2. Офіційний веб-сайт Державної служби статистики України [Електронний ресурс]. – Режим доступу: <http://www.ukrstat.gov.ua>

Богдан Якимчук,
Аспірант,

Київський національний університет імені Тараса Шевченка

ЗАСТОСУВАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ ІНВЕСТИЦІЙНИХ РІШЕНЬ

Ключові слова: машинне навчання, інвестиції, системи підтримки інвестиційних рішень, ритейл, часові ряди

Bohdan Yakymchuk,
PhD student,

Taras Shevchenko National University of Kyiv

APPLICATION OF MACHINE LEARNING METHODS FOR INVESTMENT DECISIONS SUPPORT

Keywords: machine learning, investment, investment decision support systems, retail, time series

Використання методів машинного навчання та штучний інтелект стає одним з ключових пріоритетів для компаній, що хочуть не втрачати темпів розвитку у порівнянні з потенційними конкурентами. Більш глибоке розуміння цільової аудиторії, можливість розпізнавати патерни поведінки клієнтів і створювати персоналізовані пропозиції – короткий перелік сфер застосування методів машинного навчання для стимулювання продаж у будь-якому B2C секторі ритейла, проте компанія будь-якої індустрії може універсально використовувати дані для оптимізації витрат, скорочення відтоку персоналу чи підвищення ефективності логістичної функції.

Близько 50% компаній за даними McKinsey Global Survey 2020 розвивають підхід на основі алгоритмів штучного інтелекту хоча б в одній з функцій бізнесу [1]. Ключові гравці ринку вбачають конкурентну перевагу у використанні машинного навчання для цифрової трансформації та підвищення ефективності інвестицій в R&D. У фінансовому секторі машинне навчання застосовується для високочастотної алгоритмічної торгівлі, пошуку найкращих цільових за параметрами активів, ринкового аналізу та менеджменту інвестиційного портфеля.

Проте у корпоративних фінансових функціях інтеграція методів інтелектуального аналізу даних впроваджується повільніше і рішення базуються на основі фінансових моделей. Такі моделі часто спираються на суб'єктивні оцінки менеджмента, історичні лінійні тренди, а це методологічно не дає можливість покращити оцінку певних метрик чи прогноз їх динаміки на основі нелінійних методів моделювання.

Стандартна фінансова модель, на яку спирається менеджмент при прийнятті рішень спирається на великий ряд припущень. Наприклад, у ритейлі при відкритті нової POS (Point of Sale) менеджмент робить оцінку локації та припускає потенційний трафік та купівельну спроможність населення. Проте часто оцінки на основі історичних бенчмарків чи власного досвіду є достатньо зміщеними, що може привести до ряду похибок у разі низької точності прогнозу:

- некоректний підбір асортименту для цільової аудиторії;

- закупка надлишкового технологічного обладнання для реалізації власної продукції, яка має недостатній рівень попиту;
- неефективна реалізація плану торгового залу і як наслідок зменшення пропускної спроможності POS;
- неоптимальний найм у штат магазину;
- потенційні втрати від надлишку швидкопсувної продукції;
- перевантаженість складу чи високий рівень показника «Out-of-Shelf» (% відсутності товару на полиці торгового залу).

Для покращення точності прогнозування, і як наслідок уникнення вищенаведених наслідків можна запропонувати ряд моделей на основі методів машинного навчання.

1. Модель прогнозування виручки. Часто для прийняття рішення щодо відкриття магазину в якості вхідних даних виступає всього лише інформація про місце знаходження локації та історичні бенчмарки мережі. Такі дані про геолокацію можуть включати в себе:

- кількість житлових комплексів, будинків, квартир та вартість квадратного метра житлової площі в полігоні;
- бінарні метрики знаходження біля інфраструктурних об'єктів, дорожніх розв'язок, ТРЦ чи БЦ;
- площа, кількість та купівельна спроможність населення міста, класифікація району міста та віддаленість від центру.

Такі вхідні дані дозволяють побудувати моделі оцінювання потенціального рівня щоденної виручки. Для вирішення такого кейсу пропонуються методи градієнтного бустінгу, які функціонують як ансамбль дерев рішень.

В основі методу лежить ступінчастий пошук оптимальної моделі. Спочатку ініціалізується первинна модель: $F_0 = \arg \min \sum_{i=1}^n L(y_i, F(x))$. Далі ітераційно розраховуються псевдозалишки $r_{im} = - \left[\frac{\partial L(y_i, F(x))}{\partial F(x)} \right]_{F(x)=F_{m-1}(x)}$ та вирішується одновимірна оптимізаційна задача для пошуку мультиплікатора:

$$\gamma_m = \arg \min_{\gamma} \sum_{i=1}^n L(y_i, F_{m-1}(x) + \gamma h_m(x_i)) \quad (1)$$

Проводиться спуск: $F_m(x) = F_{m-1}(x) + \gamma_m h_m(x_i)$ до моменту задоволення критичного рівня залишків. Фінальна модель $F_m(x)$ – має найкращі властивості і використовується для прогнозування виручки [2]. На основі даних великих продуктових мереж України, застосування алгоритму дозволило скоротити похибку прогнозування до 10% на тестовій вибірці.

2. Модель вибору оптимальної площі на основі прогнозу виручки пропонується реалізувати за наступним алгоритмом (див. рис. 1).

1. Для кожного з магазинів мережі магазинів будуються погодинні криві фізично кількості проданих SKU (Stock Keeping Unit)

2. Усі магазини мережі кластеризуються на основі алгоритму динамічної трансформації часової шкали. Між часовими кривими продаж вимірюється відстань для того, щоб визначити їх подібність і класифікацію, ігноруючи можливі локальні та глобальні зсуви часових рядів.

3. Згодом ідентифікуються сезонні патерни поведінки споживачів за допомоги сезонної декомпозиції на основі рядів Фур'є, що дозволяють виділити внутрішньоденні, тижневі, місячні сезональності розкладаючи часові дані на множину ортогональних тригонометричних функцій [3]:

$$y_t = \beta_0 + \beta_1 t + \sum_{i=1}^s (a_i \cos 2\pi f_i t + b_i \sin 2\pi f_i t) + \epsilon_t \quad (2)$$

де y_t – часовий ряд в момент t , β_0, β_1 – коефіцієнти лінійної регресії, a_i, b_i – коефіцієнти періодизації, а f_i – гармонічна функція періодичності.

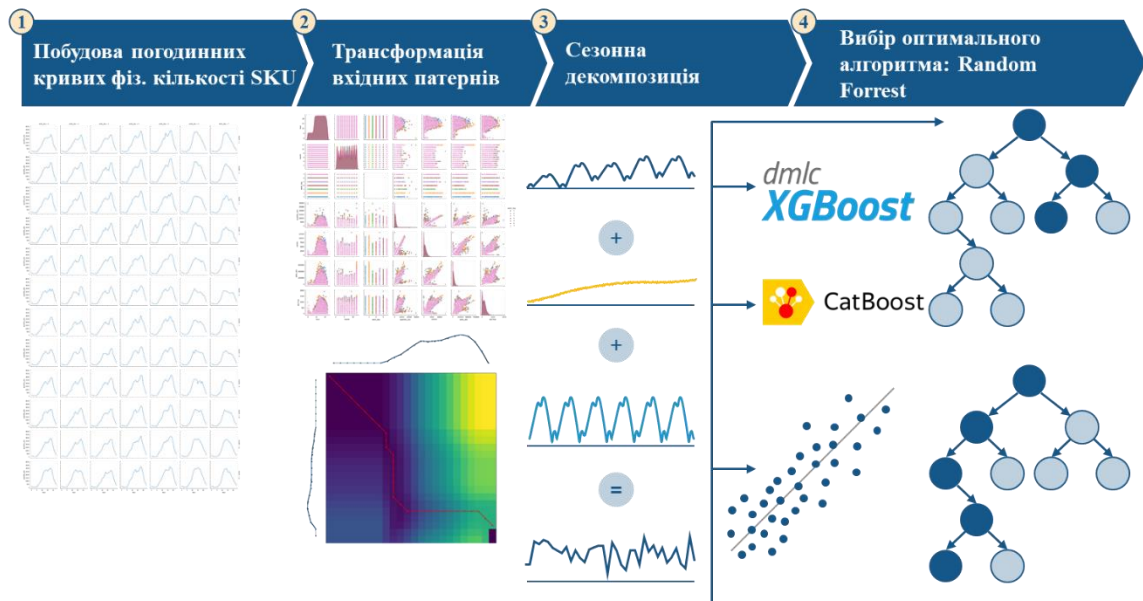


Рис 1. Алгоритм пошуку оптимальної площі POS.

Джерело: розроблено автором.

4. На основі отриманих даних, очищених від сезонного впливу ідентифікуються години максимальної завантаження і відбувається пошук оптимального алгоритму прогнозування виручки на основі кількості кас, площі торгового залу, геоданих магазину та кластерів погодинної динаміки. У нашому випадку таким алгоритмом виявився Random Forrest (з похибкою MAPE – 7%).

Згодом, будуюмо алгоритм вибору оптимальної площі торгового залу для максимізації пропускної спроможності трафіку потенційної точки продаж.

Отримані прогнози виручки дозволяють покращити операційні метрики магазину, обрати найкращу конфігурацію з врахуванням асортименту, а модель вибору оптимальної площі торгового залу – створити найбільш комфортні умови для покупок гостей магазину з врахуванням патернів їх поведінки. Такий перелік моделей дозволяє покращити оцінку проєкту при інтеграції у фінансову модель.

Список використаних джерел

1. Balakrishnan, T., Chui, M., Hall, B., Henke, N., The state of AI in 2020. *McKinsey&Company*, 2020. November 17.
2. Friedman J. H. Greedy function approximation: a gradient boosting machine. *Ann. Stat.* 2001. Vol. 29 (5). Pp. 1189-1232
3. Keogh E. J., Pazzani M. J. Derivative Dynamic Time Warping. First SIAM International Conference on Data Mining. 2002. Vol. 1. P. 11.
4. Iwok I. Seasonal Modelling of Fourier Series with Linear Trend. *International Journal of Statistics and Probability*. 2016. Vol. 5 (6). P. 3.

Валерія Яценко,
Здобувач вищої освіти ступеня доктора філософії
Київський національний університет імені Тараса Шевченка

DATA MINING В ОЦІНЦІ ВАЛЮТНИХ РИЗИКІВ В УМОВАХ НЕПОВНОЇ ІНФОРМАЦІЇ

Ключові слова: Data Mining, обмінний курс, карта Кохонена (самоорганізована карта).

Valeria Yatsenko,
Applicant for the degree of PhD,
Taras Shevchenko National University of Kyiv

DATA MINING IN CURRENCY RISK ASSESSMENT IN THE CONDITIONS OF INCOMPLETE INFORMATION

Keywords: Data Mining, exchange rate, Kohonen map (self-organized map).

Волатильність валютного курсу є одним з найважливіших економічних факторів для країн, що розвиваються та їхніх експортно-орієнтованих галузей. Багато українських вчених підкреслюють виключне значення обмінного курсу для вітчизняного експорту, особливо сільськогосподарського, і стверджують, що девальвація національної валюти є сприятливим для експорту фактором. Однак, зв'язок експортних надходжень та обмінного курсу непрямий та має дуже складну природу. Теоретично, девальвація повинна сприяти збільшенню експортних надходжень за рахунок зниження цін товарів в іноземній валюті і зростанню обсягів гривневої виручки. Проте, таке твердження не справджується повністю: девальвація впливає на експорт в різному ступені залежно від еластичності товарів; збільшення експортних надходжень зміцнює курс національної валюти; зрештою, із девальвацією зростає ціна імпортованих товарів (паливо, комплектуючі, імпортна сировина), що спричинює дорожчання кінцевого товару та носить назву Exchange Rate Pass-Through (ERPT). Крім того, на експортні надходження впливає велика кількість факторів, які здійснюють опосередкований або нелінійний вплив, не піддаються прогнозу або кількісній оцінці. Це засвідчує високий рівень невизначеності глобальної кон'юнктури і необхідність прийняття управлінських рішень в умовах неповної або асиметричної інформації шляхом обробки значних масивів інформації, а саме Data Mining.

Для перевірки зв'язку обсягів експортних надходжень пшениці та валютного курсу було використано нейронну мережу, а саме конкурентну мережу із навчанням без вчителя – самоорганізовану карту Кохонена, який базується на алгоритмі Winner Take All (WTA) [1, с. 147]. Вибірка основних факторів, які потенційно впливають на динаміку вітчизняного експорту пшениці включали 15 макроекономічних індикаторів, поділені на 3 групи: валютні курси гривні та національних валют країн-імпортерів, ціни на пшеницю провідних світових товарних бірж, індекси цін виробників (producer

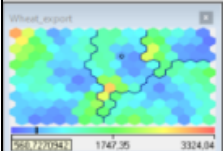
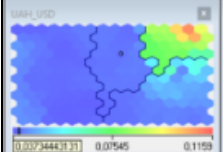
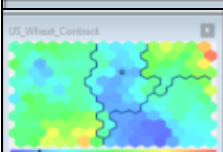
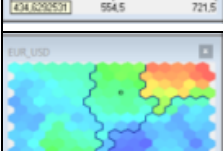
price indice, PPI) країн-імпортерів. Для аналізу використовувались дані країн найбільших імпортерів української пшениці – Німеччини, Австрії та Індонезії.

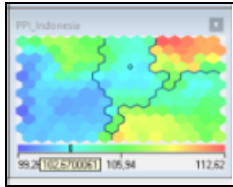
У результаті дослідження зв'язку обсягів експорту та валютного курсу спроектовані та побудовані нейронні мережі типу карти Кохонена, кожна з яких складається із 4 нейронів (за кількістю класифікаційних кластерів) та можливою похибкою навчання 10%.

На основі навчання карти Кохонена можна зробити висновок, що стабільно вагомими факторами на експортні надходження пшениці є ціни контрактів США, як результат взаємодії світових попиту та пропозицій, та індекс цін виробників в Індонезії (табл. 1). Останній фактор можна пояснити тим, що Індонезія є країною, що розвивається, а тому виробнича система суттєво залежить від макроекономічної кон'юнктури.

Таблиця 1

Профілі кластерів експортних надходжень від пшениці

Карти Кохонена	Фактори, вагомість у %	Кластери			
		0	1	2	3
		01.01.2014-01.11.2014	01.04.2017-01.04.2018	01.12.2014-01.03.2017	01.05.2018-01.01.2021
	Обсяг експорту	42.2%	14.9%	5.3%	17.4%
	UAH/USD	100%	89.7%	62.5%	100%
	Ціна ф'ючерсів на Лондонській біржі	42.9%	18.8%	100%	100%
	Ціна контрактів на американській біржі	100%	98.9%	98.7%	95%
	EUR/USD	100%	68.1%	99.9%	91.9%
	PPI Німеччини	87.8%	18.5%	100%	100%
	PPI Австрії	67.2%	25.4%	100%	100%

	PPI Індонезії	100%	91.2%	97.6%	100%
---	---------------	------	-------	-------	------

O

Побудована нейронна мережа була використана для оцінки взаємовпливу динаміки курсу національної валюти та обсягів експорту пшениці. Аналіз показав, що фундаментальні припущення про збільшення експортних надходжень у зв'язку із девальвацією національної валюти в періоди найбільшого знецінення гривні не справджуються - коефіцієнт вагомості обмінного курсу становить 62.5% протягом 12.2014-03.2017. Натомість, високий рівень залежності характерний для показників, які описують макроекономічну ситуацію країн-імпортерів та кон'юнктуру світових ринків.

Таким чином, з метою прийняття ефективних управлінських рішень суб'єкт господарювання повинен аналізувати великі масиви даних різноманітних факторів, які прямо або опосередковано можуть вплинути на обсяг експорту та прибутковість. Такий аналіз слід здійснювати за допомогою методів Data Mining або Machine Learning із врахуванням ймовірності настання факторів в умовах неповної або асиметричної інформації.

Список використаних джерел

1. Стрельченко І. І. Вибір оптимальної топології нейронної мережі в задачах класифікації динамічних економічних систем. Нейро-нечіткі технології моделювання в економіці. 2017. № 6. С. 142-157. URL: http://nbuv.gov.ua/UJRN/Nntm_2017_6_7
2. List of products exported by Ukraine. Retrieved from URL https://www.trademap.org/Product_SelCountry_TS.aspx?nvpm=1%7c804%7c%7c%7c%7cTOTAL%7c%7c%7c2%7c1%7c1%7c2%7c2%7c1%7c1%7c1%7c1%7c1
3. OECD Database Retrieved from URL <https://data.oecd.org/price/producer-price-indices-ppi.htm>
4. Statistical Database Retrieved from URL <https://www.investing.com/>
5. Statistical Database Retrieved from URL <https://tradingeconomics.com/indonesia/producer-prices-change>

Наукове видання

ЕКОНОМІКА. ФІНАНСИ. БІЗНЕС. УПРАВЛІННЯ
Зміни. Адаптація. Нова економіка

EFBM 2021

**МАТЕРІАЛИ МІЖНАРОДНОГО
ФОРУМУ**

Комп'ютерна верстка: Кравець Т.В.

Обкладинка: Коваленко О.Я.

Електронна версія збірника доступна за посиланням:

