

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ
ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ
Кафедра математичних методів системного аналізу

До захисту допущено:
Завідувач кафедри
Оксана ТИМОЩУК
«__» _____ 2024 р.

ДИПЛОМНА РОБОТА
на здобуття ступеня бакалавра
за освітньо-професійною програмою «Системний аналіз і управління»
спеціальності 124 «Системний аналіз»
на тему: «Дослідження тижневого трекеру ОЕСР як надійного показника
економічної діяльності»

Виконав:

студент ІV курсу, групи КА-03

Сапельніков Вячеслав Леонідович _____

Керівник:

к.ф.-м.н., доцент Барановська Леся Валеріївна _____

Консультант з нормоконтролю:

к.ф.-м.н. Статкевич Віталій Михайлович _____

Консультант з економічного розділу:

доцент, к.е.н., Дученко Марина Михайлівна _____

Рецензент:

доцент, к.т.н., Харченко Костянтин Васильович _____

Засвідчую, що у цій дипломній роботі немає
запозичень з праць інших авторів без
відповідних посилань.

Студент _____

Київ – 2024

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Навчально-науковий інститут прикладного системного аналізу
Кафедра математичних методів системного аналізу

Рівень вищої освіти – перший (бакалаврський)

Спеціальність – 124 «Системний аналіз»

Освітня програма «Системний аналіз і управління»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Оксана ТИМОЩУК

«__» _____ 2024 р.

ЗАВДАННЯ

на дипломну роботу студенту

Сапельнікову Вячеславу Леонідовичу

1. Тема роботи «Дослідження тижневого трекеру ОЕСР як надійного показника економічної діяльності», керівник роботи Барановська Леся Валеріївна, кандидат фізико-технічних наук, доцент кафедри ММСА, затверджені наказом по університету від «__» _____ 2024 р. №
2. Термін подання студентом роботи 11.06.2024.
3. Вихідні дані до роботи: дані індексів економічної активності, показники ВВП за 2004-2022 рр.
4. Зміст роботи: дослідження предметної області; теоретичний опис процесу оцінки тижневого трекеру ОЕСР як надійного показника економічної активності; програмна реалізація та аналіз результатів; функціонально-вартісний аналіз програмного забезпечення.

5. Перелік ілюстративного матеріалу (із зазначенням плакатів, презентацій тощо):
графіки темпу росту економічних показників.

6. Консультанти розділів роботи.

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Економічний	доцент, к.е.н., Рощина Н.В.		

7. Дата видачі завдання _____

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання етапів роботи	Примітка
1.	Вивчення літератури за темою роботи.		Виконано
2.	Підготовка першого розділу.		Виконано
3.	Підготовка другого розділу.		Виконано
4.	Початок розробки програмного продукту.		Виконано
5.	Підготовка третього розділу.		Виконано
6.	Підготовка економічної частини.		Виконано
7.	Оформлення розділів відповідно до нормконтролю.		Виконано
8.	Підготовка презентації доповіді.		Виконано
9.	Оформлення дипломної роботи.		Виконано

Студент

Вячеслав САПЕЛЬНИКОВ

Керівник

Леся БАРАНОВСЬКА

РЕФЕРАТ

Дипломна робота: 90 с., 10 рис., 9 табл., 2 дод., 33 джерел.

ДОСЛІДЖЕННЯ ТИЖНЕВОГО ТРЕКЕРУ ОЕСР ЯК НАДІЙНОГО ПОКАЗНИКА ЕКОНОМІЧНОЇ ДІЯЛЬНОСТІ

Об'єкт дослідження – тижневий трекер економічної активності, розроблений Організацією економічного співробітництва та розвитку (ОЕСР).

Предмет дослідження – показники економічної активності, які відстежуються тижневим трекером ОЕСР.

Мета роботи – проаналізувати предмет дослідження, оцінити надійність та точність тижневого трекера ОЕСР як інструмента для моніторингу економічної діяльності.

Методи дослідження – економетричний аналіз, порівняння даних трекера ОЕСР з іншими показниками економічної активності, такими як ВВП, індикатор економічних настроїв та індекси споживчих настроїв.

Актуальність – дослідження необхідності оперативного моніторингу економічної активності в умовах швидких змін на глобальних ринках.

Результати роботи – проведено аналіз точності тижневого трекера ОЕСР, встановлено його високу кореляцію з традиційними показниками економічної активності.

Шляхи подальшого розвитку предмету дослідження – інтеграція трекера ОЕСР з іншими інструментами аналізу, розширення його застосування на різні галузі економіки. Проведення досліджень щодо впливу сезонних факторів та інших змінних на точність трекера.

Ключові слова: ТРЕКЕР, ОЕСР, ВВП, НАУКАСТИНГ, ПРОГНОЗУВАННЯ, ІНДЕКС, ЕКОНОМЕТРИЧНИЙ АНАЛІЗ

ABSTRACT

Diploma thesis: 90 p., 10 fig., 9 tabl., 2 app., 33 references.

A STUDY OF THE OECD WEEKLY TRACKER AS A RELIABLE INDICATOR OF ECONOMIC ACTIVITY

Object of research – a weekly tracker of economic activity, developed by the organization of economic cooperation and development (OECD).

Purpose of study – to analyze the subject of the study, evaluate the reliability and accuracy of the OECD weekly tracker as a tool for monitoring economic activity.

Research methods – econometric analysis, comparison of OECD tracker data with other indicators of economic activity, such as GDP, indicator of economic moods and consumer mood indices.

Relevance – investigation of the need for operational monitoring of economic activity in the face of rapid changes in global markets.

Results – the accuracy of the OECD weekly tracker was analyzed, its high correlation with traditional indicators of economic activity was established.

Ways to further develop the subject of research – integration of the OECD tracker with other analysis tools, expanding its application into different sectors of the economy. Studies on the influence of seasonal factors and other variables on tracker accuracy.

Key words: TRACKER, OECD, GDP, NOWCASTING, FORECASTING, INDEX, ECONOMETRIC ANALYSIS

ЗМІСТ

ВСТУП.....	9
РОЗДІЛ 1 ДОСЛІДЖЕННЯ ПРЕДМЕТНОЇ ОБЛАСТІ.....	12
1.1 Наукастинг	12
1.2. Наукастинг за допомоги Google Trends	13
1.3. Наукастинг ВВП за допомоги якісних індикаторів	15
1.4. Наукастинг ВВП за допомоги індексу ОЕСР	18
1.5. Опис даних	19
1.6. Постановка задачі.....	21
1.7. Висновки до розділу 1.....	22
РОЗДІЛ 2 ТЕОРЕТИЧНИЙ ОПИС ОЦІНКИ ТИЖНЕВОГО ТРЕКЕРУ ОЕСР ЯК НАДІЙНОГО ПОКАЗНИКА ЕКОНОМІЧНОЇ АКТИВНОСТІ	23
2.1. Основні поняття про моделі прогнозування.....	23
2.2. Основні поняття про методи оцінки точності прогнозів.....	25
2.3. Агрегація часу.....	26
2.4. Еталонні показники	27
2.5. Емпірична модель.....	28
2.5. Висновки до розділу 2.....	30
РОЗДІЛ 3 ПРОГРАМНА РЕАЛІЗАЦІЯ ТА АНАЛІЗ РЕЗУЛЬТАТІВ	32

3.1. Обґрунтування вибору мови програмування.....	32
3.2. Підготовка даних до аналізу.....	34
3.3. Результати кореляції	35
3.3.1. Загальні спостереження	35
3.3.2. Спостереження під час глобальної фінансової кризи	36
3.3.3. Спостереження до та після пандемії COVID	39
3.4. Аналіз результатів	41
3.4.1. Порівняння з вихідним дослідженням	41
3.4.2. Результати наукастингу	42
3.4.3. Специфічні для країн результати.....	45
3.5. Висновки до розділу 3.....	47
РОЗДІЛ 4 ФУНКЦІОНАЛЬНО-ВАРТІСНИЙ АНАЛІЗ ПРОГРАМНОГО	
ПРОДУКТУ	48
4.1 Постановка задачі проектування	49
4.2 Обґрунтування функцій програмного продукту	49
4.3 Обґрунтування системи параметрів програмного продукту.....	52
4.4 Аналіз експертного оцінювання параметрів.....	55
4.5 Аналіз рівня якості варіантів реалізації функцій	59
4.6 Економічний аналіз варіантів розробки ПП	60
4.7 Вибір кращого варіанту ПП техніко-економічного рівня	67
4.8 Висновки до розділу 4.....	68

ВИСНОВКИ	69
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	71
ДОДАТОК А ЛІСТИНГ КОДУ	76
ДОДАТОК Б ПРЕЗЕНТАЦІЯ.....	83

ВСТУП

Валовий внутрішній продукт (ВВП) є ключовим показником активності та здоров'я економік. Хоча це головний показник економіки країни з часів Бреттон-Вудської конференції 1944 року, він має серйозний недолік – швидкість його випуску. На відміну від фінансових даних, ключові макроекономічні змінні, такі як ВВП, доступні лише з низькою частотою та зі значним запізненням: приблизно з місячною затримкою в розвинутих економіках і з квартальною затримкою в країнах з економікою, що розвивається [1]. За даними ОЕСР [2], попередні дані в США та Великобританії публікуються приблизно через місяць після закінчення звітного кварталу, тоді як в ЄС затримка публікації становить 2–3 тижні довше. Раннє прогнозування ВВП може мати вирішальне значення для прийняття рішень у разі раптових потрясінь, як-от глобальна фінансова криза, пандемія чи війна, коли економічні умови швидко змінюються. Це ставить перед сучасними економістами нове завдання: розробити нові методи своєчасного прогнозування ВВП.

Після 2000 року для прогнозування макроекономічних змінних на поточний і наступний квартал використовували якісні індикатори, такі як опитування споживачів, у поєднанні з кількісними індикаторами, такими як роздрібні продажі та промислове виробництво [3]. Останнім часом із розвитком технологій з'явилися альтернативні джерела даних, також відомі як Big data. Big data стосуються величезних обсягів структурованих і неструктурованих даних, які генеруються з таких джерел, як соціальні мережі, онлайн-транзакції та державні дані. Ці дані можуть надати більш повне уявлення про економічну діяльність, включаючи аспекти, які не враховуються традиційними макроекономічними змінними, як-от поведінка споживачів [4]. Незважаючи на свої обмеження, Big data стає дедалі важливішим інструментом економічного аналізу. Він використовується урядами,

підприємствами та дослідниками для кращого розуміння економіки та прийняття більш обґрунтованих рішень [5].

Google Trends збирає статистичні дані про те, як часто шукають певний термін, і тому служить хорошим джерелом великих даних для короткострокового макроекономічного прогнозування. Перш ніж приймати фінансові рішення, наприклад купувати що-небудь, споживачі часто використовують пошукові системи для отримання інформації. Цілі та наміри споживачів та інших агентів, а особливо споживчі витрати, розкриваються даними Google Search [4]. Вибір використовувати пошукові запити Google замість даних з іншої пошукової системи пояснюється домінуванням Google на ринку онлайн-пошуку та доступом до найбільшої частки пошукових даних [6]. Зважаючи на це, дані Пошуку Google є корисним джерелом інформації для прогнозування макрозмінних.

Наукастинг (англ. *nowcasting*) — це процес прогнозування стану економічних показників на теперішній час, найближче майбутнє та зовсім недавнє минуле (термін є скороченням слів «now» і «forecasting»). Термін наукастинг з'явився в економіці в 21 столітті і зараз широко розробляється економістами та економетристами для прогнозування макроекономічних факторів, таких як рівень безробіття, торговий баланс і зростання ВВП, до публікації офіційної оцінки ВВП. Це дозволяє центральним банкам проводити макроекономічну політику та дослідження з потоком випусків даних у реальному часі [5].

У 2020 році був створений новий високочастотний індикатор, здатний прогнозувати зростання ВВП для 47 країн ОЕСР і G20 — тижневий трекер ОЕСР (OECD Weekly Tracker), автором якого є Nicolas Woloszko. За словами автора створений трекер стабільно перевершує модель авторегресії четвертого лагу темпів зростання ВВП. У цьому аналізі було зроблено більше, а саме порівняно показники тижневого трекеру ОЕСР зі звичайними індексами, які допомагають оцінити поточне зростання ВВП: PMI, ESI та їх поєднання. Даним аналізом зроблено внесок у попереднє дослідження, емпірично спостерігаючи за підвищенням точності

індексу ОЕСР порівняно з іншими моделями, а також перевіряючи точність індексу протягом безкризового періоду. Робота пояснює відмінності в ефективності тижневого трекеру ОЕСР у прогнозуванні зростання ВВП для відповідної країни.

Таким чином, наше дослідницьке запитання: «Чи служить тижневий трекер ОЕСР точним показником для прогнозування зростання ВВП?»

РОЗДІЛ 1 ДОСЛІДЖЕННЯ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Наукастинг

Наукастинг є відносно новою галуззю економічного прогнозування, яка має на меті забезпечити прогнозування економічних показників у реальному або майже реальному часі, таких як ВВП, інфляція та зайнятість. Trehan [7] використовував моделі мостових рівнянь для прогнозування вибраних місячних макроекономічних змінних, а потім використав цю інформацію для прогнозування квартального ВВП. Rünstler і Sedillot [8] використовували подібні моделі для прогнозування квартального зростання ВВП для країн Єврозони, використовуючи кілька місячних рядів даних. Evans [9] приніс значний прогрес у галузі наукастингу, зробивши одні із перших досліджень, які оцінили як рівні, так і короткострокове зростання ВВП Сполучених Штатів.

Технічне прогнозування має на меті передбачити економічне зростання за допомогою змінних, які регулярно оновлюються та доступні оперативно. Центральні банки та державні установи широко використовують і підтримують розвиток макроекономічного наукастингу для встановлення монетарної політики. Наукастинг допомагає центральним банкам швидко й ефективно реагувати на зміни в економіці, приймати більш обґрунтовані рішення щодо монетарної політики та підтримувати стійке економічне зростання. Використовуючи передові методи статистики та машинного навчання, моделі наукастингу можуть аналізувати великі обсяги даних у реальному часі, щоб створювати точні та своєчасні прогнози економічної діяльності. Це спричинило швидке розширення економічної літератури з наукастингу з використанням масивів даних великої розмірності, які включають наукастинг економіки до оголошення офіційних цифр.

Це змусило дослідників використовувати як кількісні, так і якісні індикатори у своїх моделях. Кількісні показники базуються на кількісних даних, таких як

офіційна економічна статистика, і зазвичай включають змінні, пов'язані з промисловим виробництвом, робочою силою, продажами та споживанням. Якісні показники базуються на якісних даних, таких як опитування чи суб'єктивні оцінки.

За даними Giannone та ін. [10], дані опитування мали помітний вплив на прогнози ВВП. Пізніше Bańbura та Rünstler [11] підтвердили, що дані опитування можуть дати цінну інформацію про майбутнє зростання ВВП, коли точні дані недоступні. Це призвело до пошуку альтернативних якісних індикаторів, таких як дані Google Trends, які є доступними в режимі реального часу та охоплюють велику кількість населення. Дані Google Trends відображають обсяг пошукових запитів в Інтернеті за певними ключовими словами, які можуть дати розуміння поведінки та вподобань споживачів, але не пов'язані безпосередньо з офіційною економічною статистикою. Незважаючи на такі обмеження, як якість даних і потенційних маніпуляцій, перші дослідження показують, що дані Google Trends є корисним джерелом інформації.

1.2. Наукастинг за допомоги Google Trends

Використання даних веб-пошуку для прогнозування економічної статистики вперше було представлено в опублікованій статті, автором якої є Ettredge та ін. [12], який вивчав рівень безробіття в США. Ця ідея була додатково досліджена та розширена в різних областях.

Ще з 2008 року було помічено зростаючий інтерес до використання Google Trends для наукастингу в багатьох галузях, таких як психологія, фінанси і політологія. Choi і Varian були піонерами в галузі економіки та продемонстрували потенціал даних пошуку Google як інструменту прогнозування різних економічних показників. Відтоді багато авторів використовували Google Trends для

прогнозування макроекономічних змінних, включаючи рівень безробіття, роздрібні продажі, приватне споживання, продажі автомобілів тощо [13]. Крім Choi та Varian, в економіці використання даних пошуку Google досліджували такі вчені, як Askitas та Zimmermann, D'Amuri, Marcucci і Suhoу. Усі ці дослідження передбачають об'єднання кількох пошукових запитів у єдиний складений індекс. Потім цей індекс використовується як предиктор у регресійній моделі часових рядів.

У 2009 році Choi і Varian заявили, що моделі, які включають релевантні змінні Google Trends, як правило, перевершують моделі, які не включають цих предикторів. У деяких випадках підвищення точності завдяки додатковому використанню Google Trends є дуже високим. Наприклад, автори виявили збільшення на 18% точності прогнозу для роздрібних продажів у категорії «Motor vehicle and parts dealers» і збільшення на 12% при прогнозуванні «New Housing Starts» порівняно зі звичайними моделями.

Goel та ін. провели опитування щодо використання даних веб-пошуку та визнали їх корисність для прогнозування, проте вказали, що вони не забезпечують значного підвищення передбачуваності [14]. Choi і Varian не погоджуються з цією точкою зору та припускають, що, незважаючи на обмеження, їхні дослідження показують, що включення даних пошукових систем може призвести до суттєвого покращення ефективності прогнозування. Siliverstovs and Wochner припускають, що в середньому прогнози, зроблені на основі інтенсивності онлайн-пошуку, тісно збігаються з фактичними реальними результатами. Це вказує на те, що прогнози майже неупереджені ефективні та мають високий рівень інформаційної точності. Однак існує значна кількість варіацій у точності прогнозів між різними міжрегіональними порівняннями [16].

D'Amuri та Marcucci у своєму дослідженні використовують Google Index (GI), показник інтенсивності пошуку роботи в Інтернеті, як найточніший провідний показник для прогнозування місячного рівня безробіття в США. Вони виконують

прогноз поза вибіркою з використанням Google Trends і більш традиційних моделей прогнозування безробіття в США [15].

Вони прийшли до висновку, що моделі, які включають GI, перевершують традиційні моделі та дають менші похибки при прогнозуванні рівня безробіття до, під час і після Великої рецесії. Порівняно з Опитуванням професійних прогнозистів і прогнозами на державному рівні, моделі на базі Google виявилися більш точними. D'Amici та Marcucci вважали, що використання Інтернет-даних ставатиме все більш популярним в економічних дослідженнях.

Kohns і Bhattacharjee написали одну з перших статей, у якій досліджували переваги пошукових даних, окрім стандартних макроекономічних даних США під час пандемії Covid-19. Автори стверджують, що на початку певного кварталу, до того, як стануть доступними інші макроекономічні дані, наукастинг можна покращити, використовуючи достатньо широкий діапазон пошукових термінів. Автори підкреслюють, що вони залишили невивченим той факт, що Google Trends може мати змінну в часі важливість по відношенню до досліджуваної макроекономічної змінної. Пошукові терміни можуть передбачати переломні моменти в одні періоди, але не в інші, оскільки вони сильно залежать від контексту. Ці висновки свідчать про те, що дані з Google Trends можуть служити дійсним показником економічної активності [17].

1.3. Наукастинг ВВП за допомоги якісних індикаторів

Спочатку для прогнозування ВВП прогноз економічного розвитку покладався лише на прості показники, такі як рахунки національного доходу, торгова статистика та дослідження ділової активності. Одним із широко використовуваних індексів на основі опитувань для прогнозування економічної активності до

публікації офіційних даних є індекс менеджерів із закупівель (PMI). Це був один із найперших індикаторів економіки США на основі опитувань. PMI базується на опитуванні менеджерів із закупівель у виробничому секторі, яких запитують про їхні очікування щодо майбутнього виробництва, нових замовлень і працевлаштування. PMI містить низку показників ділової та економічної активності. Koenig оглядає та узагальнює ранні дослідження на PMI [18]. Він досліджує, чи PMI прогнозує ВВП, і демонструє сильну кореляцію між PMI і цільовою ставкою федерального фонду, що є ключовим компонентом грошово-кредитної політики, встановленої Федеральним комітетом з відкритих ринків. Це означає, що PMI надає актуальну інформацію про стабільність і рух економіки.

Хоча PMI широко використовується, він не є цілком надійним при прогнозуванні ВВП. Зазвичай він публікується через кілька тижнів після закінчення місяця, що означає, що він може не відобразити найновішу інформацію про економіку. Крім того, він не враховує інші важливі сектори економіки, такі як послуги, роздрібна торгівля та будівництво. Крім того, Woloszko також стверджує, що хоча якісні індикатори надають більш швидку інформацію, вони можуть стати менш проникливими щодо ВВП під час економічних спадів. На такі показники, як PMI та опитування довіри, часто впливає оптимізм або песимізм респондентів, що обмежує їхню здатність точно відобразити масштаби наукастингу [19].

Іншим індексом, заснованим на опитуваннях, є індикатор економічних настроїв (ESI), який щомісяця проводять усі країни-члени Європейського Союзу. ESI складається з п'яти різних галузевих показників довіри з різною вагою. Ці галузеві індикатори включають індикатор довіри в промисловості, індикатор довіри в сфері послуг, індикатор довіри споживачів, індикатор довіри в будівництві та індикатор довіри в роздрібній торгівлі.

Вважається, що ESI наразі є найпоширенішим композитним показником у ЄС, який використовується для моніторингу та прогнозування бізнес-циклів в окремих країнах або в регіоні в цілому. Автори дійшли висновку, що ESI може

запропонувати корисну інформацію про зростання реального ВВП для майже всіх 28 країн ЄС, як у короткостроковій, так і в довгостроковій перспективі, що дозволяє відображати зростання ВВП, коли інформація про кількісні показники не оприлюднюється. ESI може бути корисним для прогнозування майбутніх темпів зростання ВВП. Також зазначають, що ESI є більш корисним, ніж CCI через його сильніший зв'язок із коливаннями ВВП. У країнах ЄС-28 ESI був доречним для прогнозування зростання ВВП, за винятком Польщі. Автор зазначає, що польські настрої виявилися нижчими за середні за всіма показниками порівняно з ЄС-28. Дійсно, якщо взяти лише одну країну ЄС, було виявлено, що ні ESI, ні PMI не можуть точно передбачити економічну динаміку в Польщі.

Зазначають, що порівняння PMI та ESI, які не завжди узгоджуються, показало, що іноді вони показують чіткі відмінності в тенденціях, особливо в поворотні моменти. Отримані дані свідчать про те, що прогнози ВВП у реальному часі, створені на основі моделі, що використовує ESI, працюють краще, ніж прогнози, створені за допомогою моделі, що використовує показники PMI. Тим не менш, стверджують про сильну кореляцію між PMI (як у промисловості, так і в сфері послуг) і довірою споживачів у євроні та Сполучених Штатах.

Реальна діяльність може бути оцінена економістами за допомогою добре відомих агрегованих часових рядів, таких як рівень інфляції, промислове виробництво або зростання ВВП. Однак, коли настає раптовий шок, джерело інформації про стан економіки в реальному часі вкрай необхідно. Щоб подолати цю проблему, Деніел Льюїс, колишній економіст Федерального резервного банку Нью-Йорка, розробив Щотижневий економічний індекс (WEI) для щотижневого вимірювання реальної економічної активності в США та опублікував його в березні 2020 року. У WEI більшість щотижневих серій походять із приватних джерел, як-от комерційні соціологічні компанії або бізнес-групи, які збирають інформацію для своїх членів [20].

WEI використовує показники роздрібних продажів, споживчої довіри, початкових і триваючих звернень за страхуванням на випадок безробіття, індекс чисельності, утримання федерального прибуткового податку, виробництво сталі, продаж палива, виробництво електроенергії та залізничний рух із загальної кількості 10 тижневих серій, що фіксують ключові аспекти реальної діяльності. Більшість сезонних даних також видалено, щоб відобразити зміни у відсотках за 52 тижні.

Існує багато різних типів WEI, кожен з яких базується на різних джерелах даних. Прикладом WEI є тижневий трекер ОЕСР, який використовується для надання інформації про економічну діяльність.

1.4. Наукастинг ВВП за допомоги індексу ОЕСР

Цей підрозділ зосереджений навколо статті Woloszko, оскільки він був єдиним, хто написав про тижневий трекер ОЕСР. Щотижневий трекер ОЕСР, розроблений Woloszko в 2020 році, є інструментом, який використовується для оцінки темпів зростання щотижневого ВВП країн у багатьох країнах із 5-денною затримкою. Weekly Tracker працює, застосовуючи єдиний алгоритм машинного навчання на панелі даних Google Trends для 46 країн ОЕСР і G20. Трекер збирає інформацію, пов'язану зі споживанням, ринками праці, житлом, промисловою діяльністю та економічною невизначеністю, отриману з даних Google Trends. Таким чином, трекер очікується особливо корисний для оцінки змін у діяльності, коли вона змінюється дуже швидко в результаті значного шоку.

Woloszko підкреслює, що ефективність моделі тижневого трекеру ОЕСР оцінювалася шляхом щоквартального запуску моделювання в псевдореальному часі. Ці симуляції показують, що Weekly Tracker має високу ступінь доречності для

прогнозування ключових економічних показників, таких як зростання ВВП, ділові цикли та економічні кризи майже в усіх 46 досліджуваних країнах. Weekly Tracker надзвичайно успішно виявив негативний вплив пандемії COVID-19 на економіку в усіх країнах [19].

Автор зазначає, що квартальна модель, яка використовує Google Trends, має середньоквадратичну помилку, яка в середньому на 17% нижча для країн ОЕСР і G20 порівняно з авторегресійною моделлю, яка використовує лише лаги річного зростання ВВП. Хоча щотижневий трекер ОЕСР є цінним ресурсом для щотижневого аналізу активності в реальному часі, він не перевершує моделі, які використовують більш традиційні змінні, коли вони стають доступними. У зв'язку з цим виникає питання, наскільки добре тижневий трекер ОЕСР прогнозує зростання ВВП?

1.5. Опис даних

У даній роботі використано економічні дані, включаючи якісні та кількісні показники, які є корисними при змогах оцінити темпи зростання ВВП вибраних економік. У аналізі, описаному в наступних розділах, використано квартальні темпи зростання, обчислені з використанням відповідних даних, включаючи темпи зростання номінального ВВП, взяті з ОЕСР [2].

Робота зосереджена на наукастингу за допомогою тижневого індексу економічної активності ОЕСР, який є частотним індикатором, який надається щотижня ОЕСР [2]. У методології, яка описує, як створювався індекс, можна прочитати, що він повністю базується на аналізі Google Trends із використанням алгоритму машинного навчання. Google Trends вимірює кількість пошукових запитів у кожному місці та періоді часу, які потім поділяються на 1200 різних категорій, 215 із яких використовуються для аналізу економіки в моделі ОЕСР [2].

У своїй роботі Woloszko уточнює, що індекс наголошує на споживанні, але також розглядає запити про ринок праці, послуги, будівництво та житло, банкрутства, бізнес-послуги, промислову діяльність (частково) та так звану «економічну тривогу», щоб охопити період кризи. Після об'єднання запитів щоквартальні дані Google Trends узгоджуються з квартальними результатами зростання ВВП за останній квартал, щоб дозволити алгоритму співвіднести кількість різних пошукових запитів у Google і зміни в економічній активності. Після цього процесу обсяги та інтенсивність підбираються з огляду на його чутливість до змін економічного зростання, щоб отримати тижневі нелінійні зв'язки, що призводять до більш частих щотижневих показників економічного зростання, які публікуються з лагом у 5 днів після кожного тижня.

Що стосується індексів, робота зосереджується на тих, які частіше використовуються в літературі та дослідженнях щодо прогнозування та наукастингу різних показників економічної діяльності. Перший — PMI (Purchasing Managers' Index), наданий S&P Global, розроблений IHS Markit — компанії об'єдналися на початку 2022 року [21]). Індекс було отримано від S&P Global для ексклюзивного використання для аналізу, виконаного як частина дипломної роботи. PMI публікується щомісяця та створюється шляхом опитування керівників вищої ланки приватного сектору, наприклад, 400 компаній у США), що охоплює виробництво, послуги, будівництво та всю економіку (ціни, зайнятість тощо). Індекс є середньозваженим значенням різних показників, таких як «Нові замовлення (30%), обсяг виробництва (25%), зайнятість (20%), терміни доставки постачальників (15%) і запаси закупівель (10%)».

Другим використовуваним індексом є ESI (Economic Sentiment Indicator). Це чергове опитування, що проведеться Генеральним директором з економічних і фінансових питань Європейської комісії. Опитування враховує промисловість, послуги, споживачів, роздрібну торгівлю та будівництво, вагу відповідно від найважливішого до найменшого (Євростат). ESI також публікується щомісяця,

аналогічно, вимагаючи агрегування до квартальних даних. Основна відмінність між ESI та PMI полягає в тому, що останній опитує лише керівників, тоді як перший також проводить опитування споживачів.

1.6. Постановка задачі

Дана робота полягає в прогнозі квартального зростання ВВП, використовуючи одноетапну модель прогнозування, і оцінці його за допомогою середньоквадратичної помилки. Показати, що наукастинг, зроблений на основі тижневого трекаеру ОЕСР значно перевершує наукастинг, зроблений на основі таких індикаторів як Economic Sentiment Indicator (ESI) та Purchasing Managers' Index (PMI).

Програмний продукт у представлений роботі буде розроблений у вигляді одноетапної моделі прогнозування. Для аналізу та порівняння результатів буде використана кореневе середньо-квадратичне відхилення (RMSE).

Для створення програмного продукту буде використано набір даних, який знаходиться у вільному доступі в мережі Інтернет, а саме: щоквартальний звіт про ріст ВВП, тижневий трекаер економічної діяльності ОЕСР та щомісячний звіт ESI. PMI був наданий S&P Global, розроблений IHS Markit - компанії, об'єднані на початку 2022 року. Індекс було отримано від S&P Global для ексклюзивного використання для аналізу, виконаного в рамках цієї бакалаврської роботи.

1.7. Висновки до розділу 1

Дослідження в галузі наукастингу виявляє значний потенціал в передбаченні економічних показників в реальному часі або наближено в реальному часі. Використання високочастотних індикаторів, таких як Google Trends, дозволяє аналізувати великий обсяг даних в реальному часі та генерувати точні інформаційні прогнози економічної активності. Згідно з результатами попередніх досліджень, моделі, які використовують дані Google Trends, демонструють підвищену точність передбачень економічних показників, таких як обсяги роздрібною торгівлі та показники ринку житла.

Подальші дослідження також підтверджують корисність використання Google Trends для наукастингу економічних змін. Зокрема, аналіз показує, що моделі, які використовують дані Google Trends, зазвичай перевершують традиційні моделі, що базуються на лагах зростання ВВП за попередні періоди, знижуючи середньоквадратичну помилку на 17%.

Дослідження також вказує на значення використання різних якісних індикаторів, таких як PMI та ESI, для наукастингу рівня ВВП. Хоча ці індикатори можуть не завжди бути надійними у прогнозуванні ВВП, вони надають цінну інформацію про стан ринку та можуть доповнювати результати, отримані з використанням Google Trends.

Ці інструменти можуть стати корисними для центральних банків та урядових установ у прийнятті ефективних монетарних політичних рішень та підтримці сталого економічного зростання.

РОЗДІЛ 2 ТЕОРЕТИЧНИЙ ОПИС ОЦІНКИ ТИЖНЕВОГО ТРЕКЕРУ ОЕСР ЯК НАДІЙНОГО ПОКАЗНИКА ЕКОНОМІЧНОЇ АКТИВНОСТІ

2.1. Основні поняття про моделі прогнозування

Модель прогнозування - це формалізована математична або статистична структура, яка використовується для прогнозування майбутніх значень змінної на основі доступних даних і вхідних параметрів. Вона використовується для аналізу та передбачення поведінки системи чи процесу в майбутньому.

Модель прогнозування може бути побудована на основі різних методів, таких як економетричні моделі, статистичні методи, моделі машинного навчання та інші. Вона може враховувати різноманітні фактори, що впливають на змінну, такі як історичні дані, експертні оцінки, економічні тенденції тощо.

Моделі прогнозування можна класифікувати за різними критеріями, такими як тип даних, часовий горизонт, методологія та призначення. Вкажемо основні класи моделей [34].

1. Статистичні моделі: базуються на аналізі історичних даних та використанні статистичних методів для прогнозування майбутніх значень. До них відносяться ARIMA, експоненційні згладжування та моделі машинного навчання.
2. Моделі машинного навчання: використовують алгоритми машинного навчання для виявлення складних залежностей у даних та прогнозування майбутніх значень. До них відносяться нейронні мережі, дерева рішень та метод опорних векторів.
3. Базові моделі: використовують прості статистичні методи, такі як середнє значення або тренд, для прогнозування майбутніх значень.

Для цієї роботи було обрано авторегресійну модель другого порядку. Вона може бути вигідною для наших цілей з кількох причин [34].

1. Врахування автокореляції: враховує автокореляцію між показниками трекера ОЕСР на довгих затримках, що може бути характерно для економічних часових рядів. Вона використовує останні два значення для прогнозування майбутніх значень, що може допомогти у захопленні динаміки економічних змін.
2. Простота моделі: не вимагає складних підгонок параметрів або додаткових аналітичних інструментів. Це може полегшити розуміння та інтерпретацію результатів моделі.
3. Рівень довіри: Параметри моделі можуть бути статистично перевірені на значимість, що дозволяє оцінити рівень довіри до прогнозів, отриманих за допомогою цієї моделі.
4. Гнучкість: модель ARIMA дозволяє легко враховувати додаткові фактори або змінювати параметри для вдосконалення прогнозів в залежності від специфіки дослідження.

Отже, авторегресійна модель другого порядку може бути кращим вибором для прогнозування економічної діяльності на основі показників трекера ОЕСР, оскільки вона ефективно враховує автокореляцію між даними та є достатньо простою для реалізації та інтерпретації.

2.2. Основні поняття про методи оцінки точності прогнозів

У випадку прогнозування економічної діяльності можна використовувати різні методи оцінки точності прогнозів для порівняння різних моделей та визначення їхньої ефективності [32].

1. Середньоквадратична помилка (Mean Squared Error, MSE): вимірює середню квадратичну відстань між прогнозованими значеннями та фактичними значеннями. Чим менше значення MSE, тим краще якість прогнозів.
2. Коренева середньоквадратична помилка (Root Mean Squared Error, RMSE) - це показник, який вимірює середнє квадратичне відхилення між прогнозованими значеннями та фактичними значеннями. Він є популярною мірою точності прогнозів, оскільки враховує не лише величину помилок, але й їхню розкид. Чим менше значення RMSE, тим краще якість прогнозів.
3. Середня абсолютна помилка (Mean Absolute Error, MAE): вимірює середню абсолютну відстань між прогнозованими та фактичними значеннями. Він також дозволяє оцінити точність прогнозів, але не залежить від квадратів помилок, тому він менш чутливий до великих відхилень.
4. Коефіцієнт детермінації (Coefficient of Determination, R-squared): вимірює відсоток варіації в залежній змінній, який пояснюється моделлю. Він дозволяє визначити, наскільки добре модель відповідає даним. Чим більше значення R-squared, тим краще модель пояснює варіацію у даних.

Для цього дослідження було обрано кореневу середньоквадратичну помилку в якості методу оцінки точності прогнозу з наступних причин [32].

1. Врахування величини помилок: враховує не лише напрямок помилок, але й їхню величину. Це важливо, оскільки вона відображає реальну ступінь відхилення прогнозів від фактичних даних.

2. Інтерпретованість: вимірюється у тих же одиницях, що й вихідні дані, що робить його легким у розумінні та порівнянні з іншими метриками.
3. Чутливість до великих помилок: оскільки RMSE використовує квадратичну функцію, великі помилки мають більший вплив на цю метрику, що може бути важливо при оцінці точності прогнозів у критичних ситуаціях.

2.3. Агрегація часу

Розбіжності частоти даних викликають велике занепокоєння для дослідження. Враховуючи, що трекер ОЕСР випускається щотижня, PMI та ESI публікуються щомісяця, тоді як зростання ВВП наводиться на кварталній основі з низькою частотою, необхідно агрегувати високі - частотні показники та показники, щоб не тлумачити результати аналізу неправильно.

Це питання вирішується, використовуючи мостові рівняння, використані Schumacher, де автор порівнює різні способи агрегації часу та робить висновок, що і рівняння MIDAS, і мостові рівняння є порівнянними та схожими за результатами для прогнозування низькочастотних даних з використанням інших індикаторів з більшою частотою [22]. Через подання індексів PMI, ESI та ОЕСР у вигляді рівнів, необхідно спочатку перетворити їх у показники темпу зростання для відповідних частот, застосовуючи експоненціальну функцію до першої логарифмічної різниці рівнів. Наступним кроком «об'єднуємо» дані за відповідний квартал, беручи середні значення (зміни в рівнях не потребують вагових коефіцієнтів) [33]:

$$x_t^Q = w(L^{\frac{1}{3}})x_t^w$$

де $w(L^{\frac{1}{3}})$ — функція агрегатора, а x_t^w — тижневий темп зростання для ОЕСР та місячний для PMI та ESI.

Отже, остаточні регресії виглядають наступним чином:

$$y_t = \beta_0 + \beta(L^{\frac{1}{3}})x_t^w + \epsilon_t$$

2.4. Еталонні показники

Щоб мати просту та відому точку порівняння з використаною моделлю прогнозування, в даній роботі вирішено зробити порівняльний показник на основі минулих результатів зростання ВВП. Еталонний тест створюється за допомогою другої авторегресійної моделі з використанням двох минулих лагів тієї самої змінної для прогнозування майбутнього результату. Цей еталонний тест дуже корисний, коли мова йде про довгострокове прогнозування з двох конкретних причини:

- 1) коливання в короткостроковій перспективі не впливають на основну тенденцію, яка в стаціонарному світі повертається до середнього значення;
- 2) через стаціонарність довгострокові економічні показники схильні визначатися простими факторами [33]:

$$GDP_t = \beta_0 + \beta_1 GDP_{t-1} + \beta_2 GDP_{t-2} + \epsilon_t.$$

2.5. Емпірична модель

Основною проблемою з оприлюдненням ВВП і загальноекономічних даних є відставання публікації. Здебільшого країни випускають першу версію ВВП із затримкою в три-чотири тижні, а потім детальну та остаточну версію видають із більш значним відставанням. Наприклад, затримки ВВП Німеччини за перший квартал 2023 року є такими: 28 днів затримки для попереднього (першого) випуску та майже 2 місяці затримки для остаточного (детального) випуску.

Аби перевірити, чи є корисним і точним щотижневий трекер ОЕСР щодо економічної активності, нам потрібно порівняти результати наукастингу з результатами, отриманими за допомогою PMI та ESI, а також різні комбінації індексів, щоб перевірити, чи враховує індекс ОЕСР усі ефекти, представлені в опитуваннях, або його все ще можна покращити при наукастингу з використанням багатовимірних моделей. Крім того, щоб отримати детальнішу інформацію про індекс на основі Google Trends і його ефективність у періоди різної волатильності, проводиться аналіз за загальний період із 2 кварталу 2004 року до другого кварталу 2022 року та аналізуються лише періоди з низькою економічною невизначеністю, за винятком глобальних фінансових криз (3 та 4 квартал 2008 року, 1 квартал 2009 року) та пандемія COVID-19 (2 та 3 квартал 2020 року). Було застосовано метод прогнозування на один крок наперед із періодом у вибірці 20 кварталів (порівняно з 10, які використовував Woloszko) із 73 розглянутих кварталів

Окрім описаного вище тесту, також виконуються такі регресії [36]:

- 1) наукастинг квартального зростання ВВП за темпами зростання PMI, щоб побачити, як PMI можна порівняти з трекером ОЕСР:

$$(1) GDP_t = \beta_0 + \beta_1 PMI_t + \epsilon_t ;$$

- 2) наукастинг квартального зростання ВВП з використанням темпів зростання ESI, щоб побачити, як ESI можна порівняти з трекером ОЕСР:

$$(2) GDP_t = \beta_0 + \beta_1 ESI_t + \epsilon_t;$$

- 3) наукастинг квартальне зростання ВВП, використовуючи показники тижневого трекеру ОЕСР, щоб побачити, як працює індекс:

$$GDP_t = \beta_0 + \beta_1 OECD_t + \epsilon_t;$$

- 4) наукастинг квартального зростання ВВП з використанням темпів зростання як PMI, так і тижневого трекеру ОЕСР, щоб побачити, як результати покращилися порівняно з рівнянням (1):

$$GDP_t = \beta_0 + \beta_1 PMI_t + \beta_2 OECD_t + \epsilon_t;$$

- 5) наукастинг квартального зростання ВВП з використанням темпів зростання ESI та тижневого трекеру ОЕСР, щоб побачити, як покращилися результати порівняно з рівнянням (2):

$$GDP_t = \beta_0 + \beta_1 ESI_t + \beta_2 OECD_t + \epsilon_t;$$

- 6) прогнозування квартального зростання ВВП за допомогою темпів зростання ESI, PMI та тижневого трекеру ОЕСР, щоб побачити, чи допоможуть ESI та PMI отримати ще точніші результати:

$$GDP_t = \beta_0 + \beta_1 ESI_t + \beta_2 PMI_t + \beta_3 OECD_t + \epsilon_t.$$

2.5. Висновки до розділу 2

У цьому розділі розглянуто питання агрегування часу щодо економічних даних із різною частотою, особливо зосередившись на щотижневому трекері ОЕСР, місячних PMI та ESI та квартальному зростанні ВВП. Враховуючи розбіжності в частоті публікації даних, було вкрай важливо правильно агрегувати високочастотні індекси та показники, щоб уникнути неправильного тлумачення результатів аналізу. Щоб вирішити цю проблему, було застосовано мостові рівняння, представлені Schumacher, які показали порівняльні та схожі результати з MIDAS для прогнозування низькочастотних даних з використанням високочастотних показників. Після перетворення рівнів PMI, ESI та індекс OECD у темпи зростання та агрегування їх за кварталами було досягнуто ефективної методології аналізу.

Крім того, було встановлено орієнтири на основі минулих результатів зростання ВВП, використовуючи другу авторегресійну модель, яка виявилася корисною для довгострокового прогнозування завдяки своїй простоті та здатності вловлювати основну тенденцію, незважаючи на короткострокові коливання. Запропонована емпірична модель передбачала порівняння результатів наукастингу, отриманих від тижневого трекера ОЕСР, з результатами PMI та ESI, а також різних комбінацій цих індексів. Крім того, було проведено дослідження продуктивності індексу Google Trends і його ефективності у періоди різної нестабільності. Для аналізу та порівняння результатів використано різні регресійні моделі, у тому числі ті, що прогнозують квартальне зростання ВВП з використанням різних комбінацій PMI, ESI та щотижневих темпів зростання OECD. Щоб оцінити точність моделей прогнозування, використано середньоквадратичну помилку (RMSE), метрику, яка зазвичай використовується для оцінки ефективності прогнозування моделей.

Загалом дослідження дає зрозуміти ефективність щотижневого трекера OECD у прогнозуванні економічної активності, пропонуючи надійну методологію для

агрегування даних і порівняння з традиційними показниками, такими як PMI та ESI. Подальші дослідження можуть вивчити додаткові моделі прогнозування та вдосконалити методологію для ще більш точних прогнозів.

РОЗДІЛ 3 ПРОГРАМНА РЕАЛІЗАЦІЯ ТА АНАЛІЗ РЕЗУЛЬТАТІВ

3.1. Обґрунтування вибору мови програмування

Вибір мови програмування є ключовим фактором для успіху будь-якого проекту з розробки програмного забезпечення. Мова програмування впливає на швидкість розробки, доступність і якість підтримки сторонніх бібліотек, а також на можливість масштабування і стійкість кінцевого продукту. Крім того, вибір мови визначає вартість проекту, оскільки різні мови мають різні рівні доступності ресурсів і вимоги до обслуговування.

Мова програмування R - це потужний інструмент для статистичних обчислень і графіки, який використовується багатьма аналітиками, статистиками та дослідниками даних. R був створений Россом Іхакою та Робертом Джентлменом в 1993 році як вільне програмне забезпечення для статистичних обчислень та графічних побудов. R підтримує різні парадигми програмування, включаючи процедурне, об'єктно-орієнтоване та функціональне програмування .

Мова програмування R широко використовується в науці про дані та машинному навчанні завдяки своєму багатому набору пакетів та інструментів для статистичного аналізу та візуалізації даних. Наприклад, пакети `dplyr` та `data.table` забезпечують ефективну обробку даних, `ggplot2` та `lattice` - високоякісну візуалізацію, а `caret` та `randomForest` - алгоритми машинного навчання. Крім того, R пропонує пакети для роботи з часовими рядами (`xts`, `zoo`), текстовими даними (`tm`, `quanteda`) та багатовимірними даними (`vegan`, `ade4`).

Однією з головних переваг R є його здатність легко візуалізувати дані та створювати публікаційного рівня графіки. Завдяки цьому R часто використовується для підготовки звітів та презентацій, де візуалізація даних відіграє ключову роль.

Крім того, R має потужний інструмент інтерактивної звітності R Markdown, що дозволяє інтегрувати код, результати аналізу та текст у єдиному документі, який можна експортувати у різні формати, включаючи HTML, PDF та Word .

Мова програмування R також відома своєю активною та великою спільнотою користувачів і розробників, яка забезпечує постійну підтримку та розвиток нових пакетів та інструментів. Існує велика кількість онлайн-ресурсів, таких як курси, форуми, блогпости та підручники, що робить процес навчання R доступним для новачків та зручним для досвідчених користувачів .

Крім того, R легко інтегрується з іншими системами та інструментами, такими як бази даних, веб-сервіси та хмарні платформи. Це дозволяє використовувати R для збирання, обробки та аналізу великих обсягів даних з різних джерел. Наприклад, пакети RPostgreSQL та RMariaDB дозволяють працювати з базами даних, а httr та rvest - з веб-сервісами та веб-скрапінгом відповідно .

Загалом, мова програмування R є потужним, гнучким та доступним інструментом для аналізу даних, що підходить як для новачків, так і для досвідчених аналітиків. Її багатофункціональність, розширюваність та активна підтримка спільноти роблять її незамінним інструментом для багатьох проектів у сфері науки про дані, статистики та машинного навчання.

З огляду на всі вищезазначені переваги, розробка програмного продукту для аналізу та візуалізації даних буде здійснюватися з використанням мови програмування R.

3.2. Підготовка даних до аналізу

Як було зазначено раніше, різні частоти даних створюють значні проблеми для дослідження. Щоб підготувати дані для аналізу та інтерпретації, було прийнято рішення змінити їх у однаковий формат. Це було досягнуто шляхом агрегування за часом високочастотних індексів (OECD, ESI, PMI) і конвертації їх у квартальні темпи зростання для порівняння з квартальним зростанням ВВП. Для спрощення аналізу та для виконання кореляції візьмемо середнє значення індексів високої частоти для даного кварталу, а потім розрахуємо темпи зростання, щоб принизити результати, взявши першу різницю, таким чином, пожертвувавши першим кварталом 2004 року. У результаті отримані дані є стаціонарними, що дозволяє провести дослідження (див. рис. 3.1).

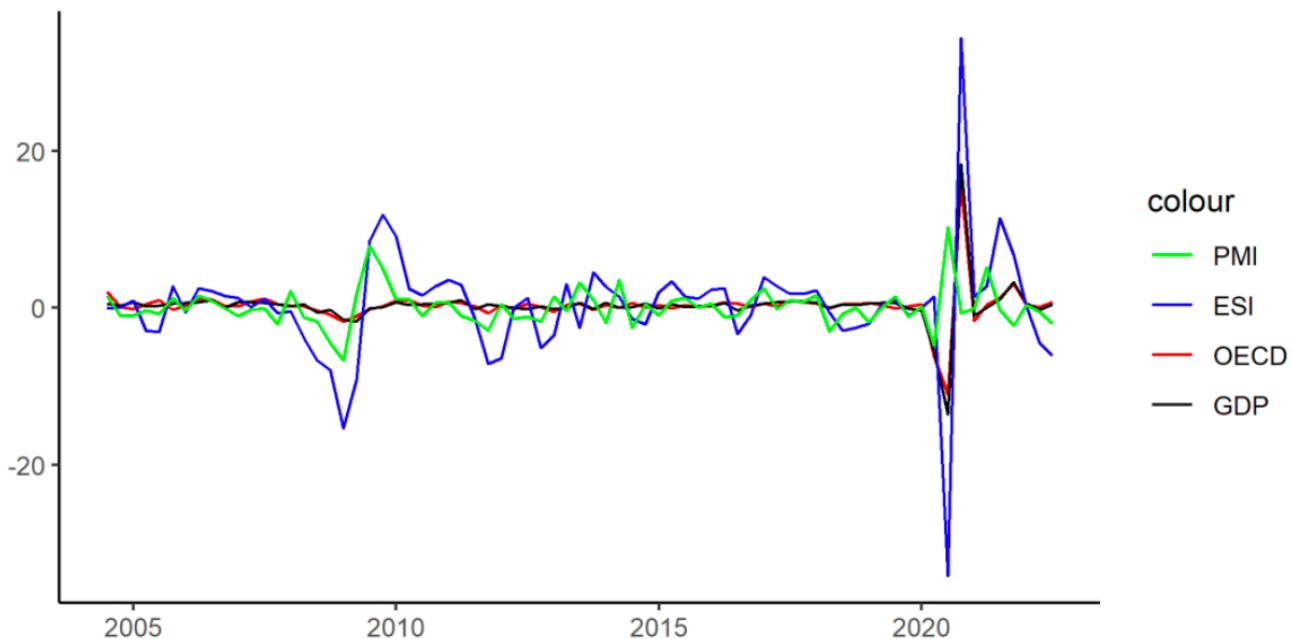


Рисунок 3.1 - Темпи зростання ВВП, індексу ОЕСР, РМІ та ЕСІ для Франції

Наша вибірка даних включає 44 країни, за період з другого кварталу 2004 року до другого кварталу 2022 року. Єдиним винятком є РМІ – для ряду країн дата початку його публікації не є узгодженою. У результаті індекс менеджерів із закупівель завжди закінчується у другому кварталі 2022 року, але дата початку відрізняється залежно від країни. Це дає нам набір даних із 9640 квартальних даних, використаних протягом 73 кварталів.

3.3. Результати кореляції

Для аналізу необхідно зрозуміти, чого можна очікувати від прогнозування і як краще підійти до процесу. Для цього виконуємо кореляцію Пірсона між усіма змінними для відповідних періодів доступності даних і ковзну кореляцію з вікном 20 кварталів кожен. Беручи до уваги розвиток Інтернету, робота також зосереджується на різних періодах часу, включаючи світову фінансову кризу та пандемію 2019 року.

3.3.1. Загальні спостереження

Результати, наведені в таблиці 3.1, ілюструють кореляції між розглянутими змінними в часі. Для десяти країн у таблиці є викид, Ірландія, з яким жоден індекс не зміг співвіднесться більше ніж 0,5765 (РМІ). Щотижневий індекс економічної активності ОЕСР найбільше корелює з темпами зростання ВВП усіх країн протягом значного періоду часу. Можна висунути гіпотезу, що індекс економічної активності

ОЕСР є найбільш точним у прогнозуванні зростання ВВП серед представлених індексів.

Варто зазначити, що кореляції між індексом ОЕСР та ВВП зменшуються в період після глобальної фінансової кризи і стають сильнішими лише після 2014 року. Кореляції до періоду глобальної фінансової кризи становлять у середньому 0,62 для всіх країн, 0,35 після глобальної фінансової кризи до 2014 р., а найвищий результат у середньому 0,90 з 2014 р. до 2 кв. 2022 р. (див. табл. 3.1).

Таблиця 3.1 – Кореляція між відповідними змінними протягом максимального доступного періоду

Країна	ОЕСР ВВП	PMI ВВП	ESI ВВП	PMI ОЕСР	ESI ОЕСР
Австрія	0.9198	0.7178	0.7762	0.6303	0.6843
Чехія	0.8499	0.5704	0.8079	0.5464	0.7257
Данія	0.7149	0.407	0.5934	0.2658	0.5753
Франція	0.9825	0.8199	0.8084	0.817	0.7882
Німеччина	0.9178	0.7476	0.8132	0.6332	0.757
Греція	0.7505	0.7131	0.614	0.8382	0.5573
Ірландія	0.4137	0.5765	0.3401	0.7882	0.6205
Італія	0.9556	0.7429	0.8234	0.735	0.7533
Нідерланди	0.9144	0.6867	0.8029	0.6604	0.8043
Польща	0.8038	0.7457	0.7356	0.7958	0.8272

3.3.2. Спостереження під час глобальної фінансової кризи

Наступним кроком було проаналізовано кореляції під час глобальної фінансової кризи для ОЕСР та ESI. Причиною виключення PMI було те, що дані неповні, тому результати не є репрезентативними, оскільки охоплюють необхідний період.

Тут бачимо щось подібне до попередньої картини, де щотижневий індекс економічної активності ОЕСР у більшості випадків виявляється найточнішим. Однак використання та проникнення Інтернету в країні на той час були не такими високими, якщо порівнювати з періодом пандемії 2019 року, тому результати мали відрізнятись, а кореляція для змінних була нижчою.

Є 10 країн із кореляцією нижчою за 0,5 (ОЕСР і ВВП), а найнижчою є Португалія (0,2081) і Латвія (0,2184) (див. табл. 3.2, рис. 3.2 та рис. 3.3). Причиною цього може бути більш серйозний вплив глобальної кризи на економіку та таким чином більше відхилень у темпах зростання ВВП.

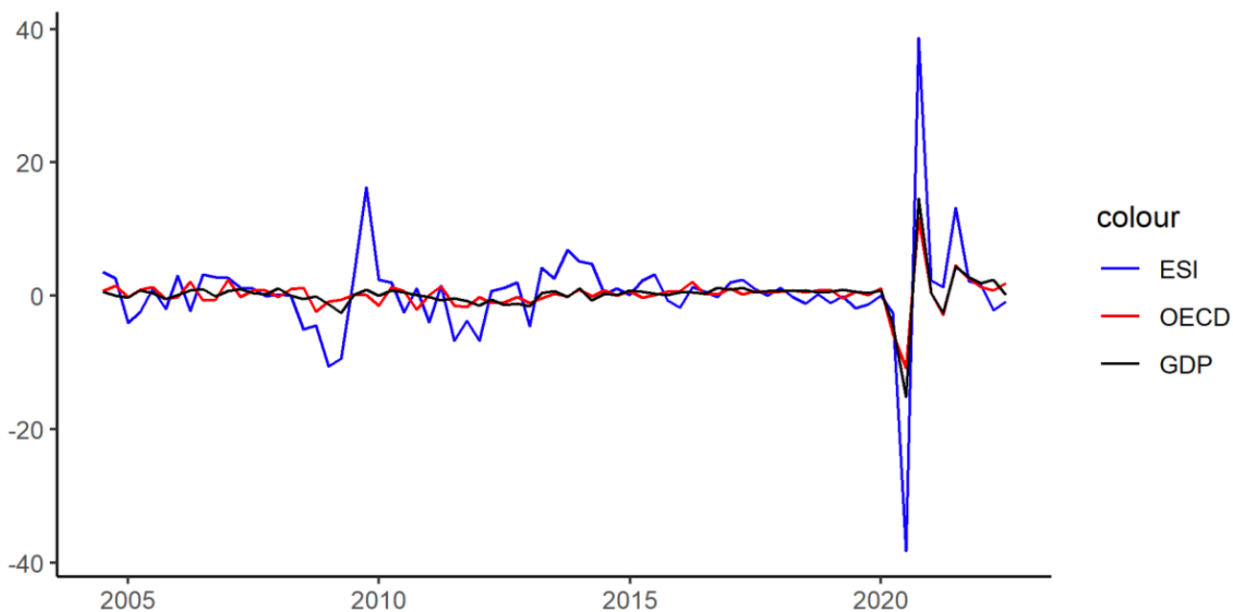


Рисунок 3.2 - Темпи зростання ВВП, індексу ОЕСР та ESI для Португалії

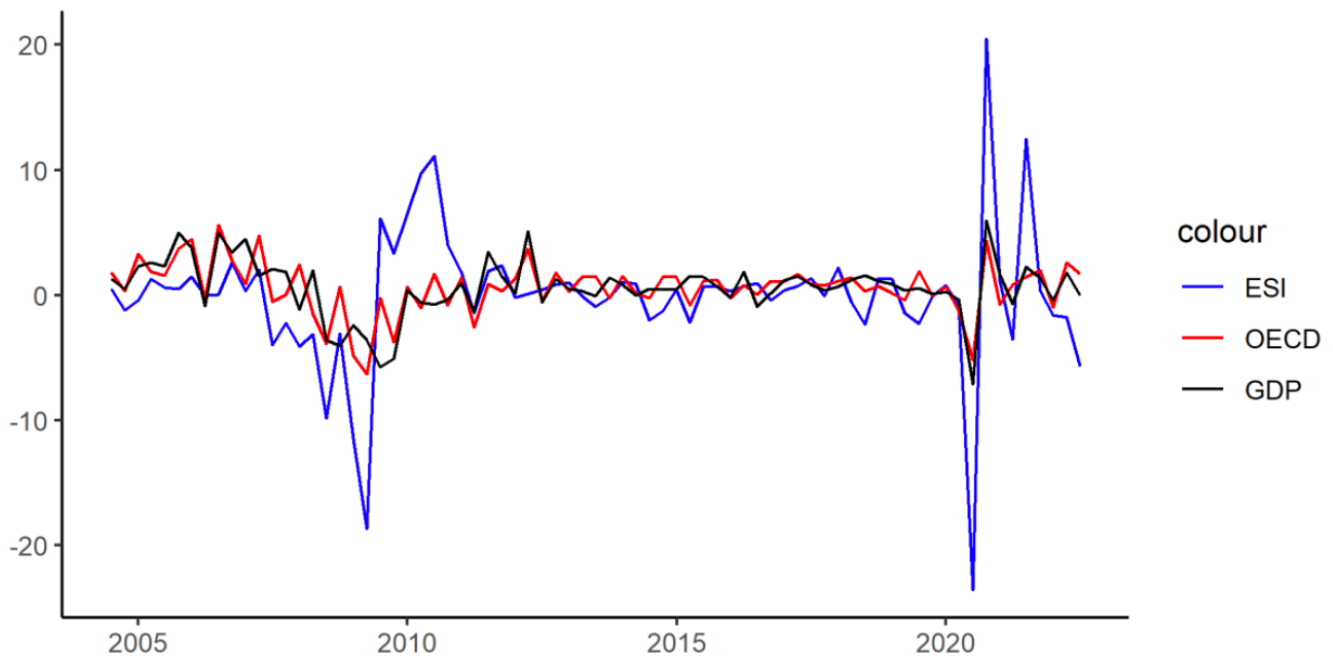


Рисунок 3.3 - Темпи зростання ВВП, індексу ОЕСР та ESI для Латвії

Таблиця 3.2 – Кореляція між відповідними змінними протягом глобальної фінансової кризи (Q2 2007 – Q2 2009)

Країна	ОЕСР ВВП	ESI ВВП	ESI ОЕСР
Австрія	0.7391	0.5857	0.5285
Чехія	0.6905	0.6836	0.3584
Данія	0.6504	-0.0309	0.5521
Франція	0.8848	0.6783	0.7497
Німеччина	0.859	0.5566	0.3417
Греція	0.777	0.8153	0.8333
Ірландія	0.3019	0.6162	0.3123
Італія	0.7821	0.5347	0.69
Нідерланди	0.8908	0.6288	0.8547
Польща	0.4589	-0.1241	0.5064

3.3.3. Спостереження до та після пандемії COVID

Період пандемії та період після неї є більш мінливим для ВВП, ніж попередній час. Тим не менш, у часи аномальної волатильності не можна побачити зниження кореляцій ні для індексів ESI, ні для індексів ОЕСР (див. табл. 3.3). Для деяких країн збільшення коефіцієнтів кореляції перевищує 100% (Данія ОЕСР проти ВВП: до 0,4524 і після 0,9128). Це збільшення могло бути спричинене збільшенням використання Інтернету в період пандемії, збільшенням віддаленої роботи в результаті обмежень, запроваджених у кожній країні.

У середньому можемо побачити з таблиці 3.3, що індекс ОЕСР більше корелює зі зростанням ВВП. Таким чином, можна припустити, що наукастінг з його використанням має бути більш точним, ніж той, що використовує індекс ESI. Іноді кореляції набагато вищі порівняно з ESI, але є випадки, коли коефіцієнт дорівнює або трохи вищий для індексу, заснованого на опитуванні ESI.

Таблиця 3.3 – Кореляція між відповідними змінними до та після пандемії COVID-19

Країна	До COVID			Після COVID		
	ОЕСР ВВП	ESI ВВП	ESI ОЕСР	ОЕСР ВВП	ESI ВВП	ESI ОЕСР
Австрія	0.7391	0.5857	0.5285	0.9626	0.8874	0.8133
Чехія	0.6905	0.6836	0.3584	0.9352	0.9235	0.9036
Данія	0.6504	-0.0309	0.5521	0.9128	0.905	0.8622
Франція	0.8848	0.6783	0.7497	0.9954	0.9236	0.8967
Німеччина	0.859	0.5566	0.3417	0.9482	0.9455	0.8892
Греція	0.777	0.8153	0.8333	0.8032	0.8062	0.6355
Ірландія	0.3019	0.6162	0.3123	0.7116	0.5995	0.8999
Італія	0.7821	0.5347	0.69	0.9714	0.921	0.8445
Нідерланди	0.8908	0.6288	0.8547	0.9704	0.9174	0.9135
Польща	0.4589	-0.1241	0.5064	0.8663	0.8769	0.9223

Це призводить до аналізу, який залучає як індекси ESI, так і індекси ОЕСР у процесі наукастингу. Необхідно висунути гіпотезу, що для деяких країн поєднання обох індексів у цьому процесі повинно дати більш точні результати, ніж будь-який з них окремо.

Якщо придивитися ближче, можна побачимо, що за після-ковідний період є 10 країн (Франція, Канада, Великобританія, Мексика, Туреччина, Південна Африка, Іспанія, Бельгія, Фінляндія та Португалія) з кореляцією між індексом ОЕСР та зростання ВВП понад 0,98 (див. рис. 3.4). Це означає, що прогнози, зроблені за допомогою щотижневого індексу економічної активності ОЕСР у період більшого використання Інтернету, можуть дати більш точні результати.

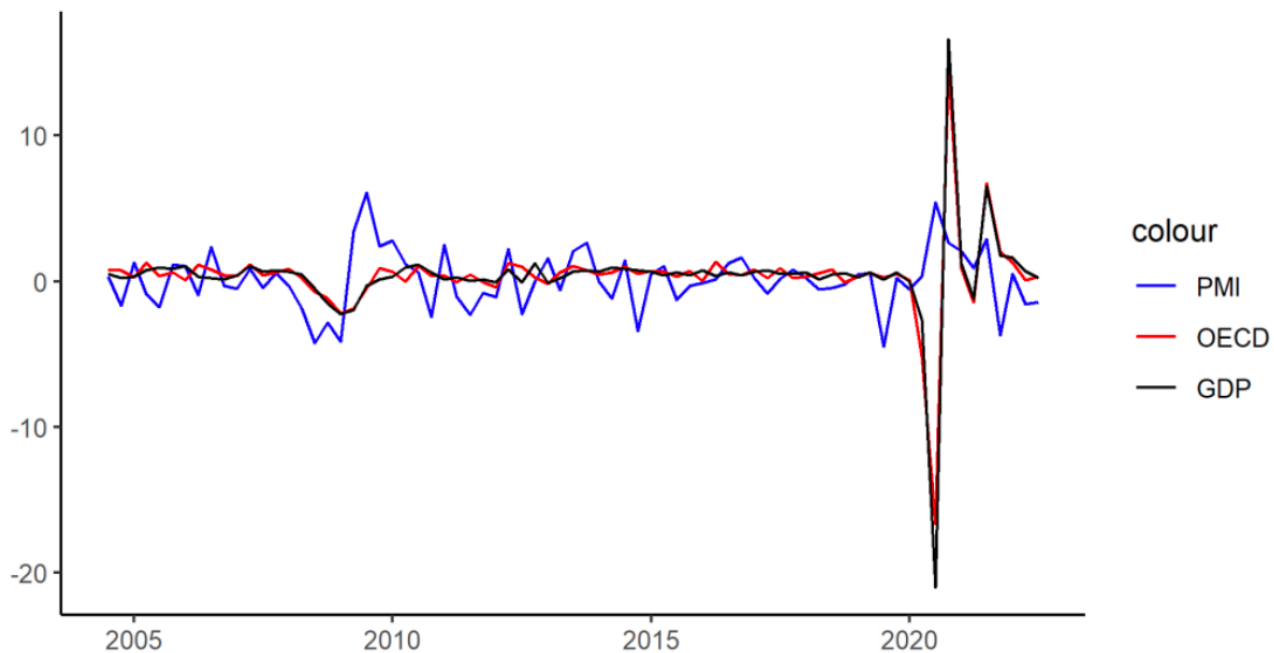


Рисунок 3.4 – Темпи росту ОЕСР, РМІ та ВВП у Великобританії (країна з одним з найвищих показників кореляції між ОЕСР та ВВП в пост-ковідний період)

3.4. Аналіз результатів

3.4.1. Порівняння з вихідним дослідженням

Єдина стаття, яка досліджує ефективність тижневого індексу економічної активності ОЕСР, була представлена у 2020 році самим автором індексу. Woloszko провів декілька аналізів, у тому числі розрахунків RMSE для країн ОЕСР, використовуючи річні темпи зростання [19]. Таким чином, як перший крок, можна порівняти наші результати з результатами, представленими в описаній статті. Важливо підкреслити, що, на відміну від Woloszko, в даній роботі використано квартальні темпи зростання.

Для порівняння можна побачити, що результати RMSE, отримані для 42 країн становлять у середньому 2,33 і 1,31 для Woloszko і нашого аналізу відповідно. Причини вдвічі більш точних результатів прогнозування можуть полягати в різних темпах зростання, а також у різниці у типі використовуваного ВВП (реальний проти номінального). Окрім періоду 2004–2022 рр., також виконано аналіз із обмеженням часу періодами низької волатильності, за винятком глобальної фінансової кризи 2008 р. та пандемії COVID. Результати, які отримано, ще виразніші, середнє значення RMSE становить 0,92, що становить приблизно 30% покращення.

Серед найбільш яскравих відмінностей між нашими роботами можна окреслити всі країни Балтії із середнім покращенням 2,85 у RMSE, а також Ірландію, Словаччину та Туреччину (2,2 покращення у RMSE). Що стосується подібності, ми обидва дійшли висновку, що індекс ОЕСР дає найбільшу похибку в оцінках для Ірландії, Аргентини та Ісландії.

3.4.2. Результати наукастингу

ESI використовується в регресійному аналізі з двох причин. Перш за все, як описано в огляді літератури, цей показник широко використовується для прогнозування змін в економічній діяльності, тому він є дійсним показником порівняння для індексу ОЕСР. Крім того, цей індикатор, як і PMI, може фіксувати різні метрики порівняно із запитами Google.

Результати наукастингу показують, що в середньому ESI працює гірше, а його результати наукастингу мають на 30% вищий RMSE порівняно з індексом на основі Google Trend із таким самим ефектом для загального періоду часу та за винятком періодів високої волатильності. Як індикатор, заснований на опитуванні, ESI має обмежене охоплення приблизно 140 тисяч компаній і споживачів у всьому ЄС [23] у порівнянні з мільярдами людей, які користуються Інтернетом. Таким чином, різниця в точності наукастингу потенційно походить від величезної різниці в охопленні. Крім того, індикатори на основі опитувань фіксують думки, а не лише показники, керовані даними, тоді як індекс Google на основі тенденцій більше зосереджується на обсягах і цифрах.

Крім того, було перевірено, чи може ESI надати більш точні результати в поєднанні з тижневим індексом ОЕСР. Ця багатовимірна регресія дає більш точні результати, ніж просто тижневий індекс ОЕСР у 50% випадків, в інших 50% випадків результати більш точні для однофакторного аналізу. Варто відзначити, що ESI стає незначним майже в 70% випадків. З одного боку, це ставить під сумнів, чи може це допомогти покращити результати, але з іншого деякі дослідження ставлять під сумнів корисність тестів значущості у прогнозуванні, якщо тільки модель не працює добре [24]. Теоретично результати покращуються, оскільки ESI може охоплювати аспекти, не включені в Google Trends – упередження змінної може вплинути на результати або вносить покращення на основі середньої освіти та рівня

досвіду опитаних людей, який у середньому вищий для опитування ESI, а не для Інтернету [25]. У нашому випадку результати суттєво покращуються у випадках, коли ESI є також статистично значущим. Нарешті, коли йдеться про точність моделі, вона в середньому досягає RMSE 1,2 для загального періоду та 0,82 за виключенням криз.

Іншим якісним показником, який широко використовується при прогнозуванні економічної активності, є індекс менеджерів із закупівель (PMI). Проводимо той самий аналіз, що й ESI, і бачимо, що PMI має ще нижчу точність результатів. Він перевершує індекс ОЕСР на 135% за загальний період (RMSE = 2,5) і на 43% без урахування криз (RMSE = 1,3). Що варто відзначити, це випадок Південної Африки – єдиної країни в нашому наборі даних, де PMI показує значно (на 60% нижчий RMSE) кращий прогноз зростання ВВП, ніж Google Trend Index. Можна порівняти ці результати з відповідними результатами для Мексики, де показник PMI є втричі гіршим, ніж індекс ОЕСР. Обидві країни мають однакове використання Інтернету на ~70%, але різну залежність від експорту продукції та різний рівень освіти населення [26]. Частка населення без середньої освіти в Мексиці становить 50%, тоді як у Південній Африці понад 80% отримали цей рівень освіти, що може позитивно вплинути на якість пошуку в Google і відповідей PMI. Що стосується економічних показників, експорт Мексики на 80% залежить від експорту промислової продукції порівняно з 42% у випадку Південної Африки, що має зробити PMI більш точним у випадку Мексики [2]. Проте південноафриканський експорт значно більше базується на сировині та металах, де зміни в цьому показнику навряд чи вловлять запити Google Trends. Цю гіпотезу підтверджує випадок Австралії, де комбінація PMI та ОЕСД призводить до 47% нижчого RMSE порівняно з одним лише індексом Google Trends. Австралійський промисловий експорт сильно залежить від металів – 24%, які також не враховуються індексом ОЕСР [27].

Подібно до ESI, PMI корисний для прогнозування разом із тижневим індексом ОЕСР, коли Google Trends не може охопити аспекти, включені в якісні індикатори економічної активності, і може підвищити точність результатів, коли в країні низьке покриття Інтернетом. Що стосується достовірності результатів, було досягнуто статистичної значущості результатів для кожного випадку вираженого покращення результатів при поєднанні PMI з індексом ОЕСР.

Нарешті, регресійний аналіз також включав наукастинг, виконаний з використанням усіх трьох показників: ESI, PMI та тижневого індексу ОЕСР. У наборі даних є десять країн, для яких усі три показники готують відповідні установи. У середньому модель працює краще, ніж одновимірна модель із тижневим індексом ОЕСР, досягаючи на 5,7% нижчого RMSE. Ця модель також перевершує будь-яку іншу розглянуту модель у 4 з 10 випадків за весь період часу (Італія, Австрія, Данія та Греція), де в решті випадків індекс ОЕСР або його комбінація з ESI має кращі результати. Цікаво, що другий найнижчий RMSE для всіх цих 4 країн забезпечується моделлю, що поєднує ESI та ОЕСР, таким чином, покращення відбувається завдяки факторам, охопленим PMI. Також слід зазначити, що якщо в нашому аналізі немає криз, поєднання всіх трьох індексів дає кращі результати, ніж індекс Google Trends, лише в 3 з 10 випадків (і лише вдвічі краще, ніж будь-яка інша модель). Оскільки цей етап аналізу охоплює 10 країн з розвинутою економікою, модель може бути більш корисною, ніж будь-яка інша, лише у випадку, якщо існують фактори, охоплені PMI і не охоплені ESI, які суттєво впливають на економічний розвиток відповідної країни (або навпаки).

3.4.3. Специфічні для країн результати

Широкий діапазон країн з різними умовами, що входять до вибірки, містить кілька викидів.

Найпомітнішим викидом є Ірландія з надзвичайно високим RMSE при наукастингу за допомогою будь-якої моделі (див. таб. 3.1). Це може бути наслідком зміни системи оподаткування транснаціональних корпорацій у 2015 році, що призвело до спотворення макроекономічної статистики [28].

Ісландія показує високий RMSE як для AR (2), так і для щотижневого трекера ОЕСР, що можна пояснити малою відкритою економікою та населенням. Ісландія має дуже спеціалізовану економіку та вразливість до зовнішніх шоків, що робить модель AR (2) ненадійною [29]. Через невелику популяцію вибірка пошукових запитів у Google є відносно невеликою. Таким чином, тижневий трекер ОЕСР може не зібрати достатньо даних і точно відобразити компоненти ВВП, щоб точно прогнозувати зростання ВВП.

Можемо помітити, що прогностична потужність авторегресійної моделі з двома лагами зростання ВВП є особливо точною для Данії та Норвегії, країн з постійним економічним зростанням. Без урахування кризових періодів для цих країн модель AR (2) перевершує тижневий трекер ОЕСР зниженням RMSE на 6,5% і працює краще, ніж інші прогностичні моделі протягом усього періоду, що розглядається.

Релевантність даних Google Trends, яка залежить від якості пошукових запитів, може сильно залежати від тижневого трекеру ОЕСР. Рівень розвитку країни, використання Інтернету та рівень освіти можуть бути причинами різниці в точності прогнозів тижневого трекеру ОЕСР для різних країн. Порівнюючи середній RMSE для 50% найвищих і найнижчих країн за індексом людського розвитку (за 2021 рік), індексом готовності мережі (за 2022 рік) і відсотком

використання Інтернету (за 2020 рік), взятих з Network Readiness Index [30], and World Bank [26], можна побачити, що країни з вищою освітою, людським розвитком, готовністю користуватися перевагами технологій і проникненням Інтернету в середньому мають нижчий RMSE при накатсингу за допомогою тижневого трекеру ОЕСР (див. рис. 3.5). Ірландія та Ісландія були виключені з цього аналізу як викиди. Оскільки існує позитивна кореляція між рівнем освіти та проникненням Інтернету в країні [31], це може означати, що більш освічені верстви населення, як правило, мають більш релевантні запити Google щодо компонентів ВВП, що призводить до вищої точності прогнозу зростання ВВП разом із тижневим трекером ОЕСР.



Рисунок 3.5 – Індекс ОЕСР та фактори розвитку

3.5. Висновки до розділу 3

Аналіз тижневого індексу економічної активності ОЕСР, що базується на квартальних темпах зростання, здається, є кращим, ніж річні дані, що призвело до більш ніж на 70% нижчого RMSE у середньому. Це особливо ефективно при застосуванні в країнах з високим рівнем використання Інтернету населенням та вищим рівнем освіти. Наукастинг може виграти від додавання якісних індикаторів, таких як PMI або ESI, якщо показники, включені в показники на основі опитувань, які не враховуються тенденціями Google, є важливими для економіки. Тим не менш, для країн зі стабільним економічним зростанням у періоди низької економічної невизначеності авторегресійна модель із двома лагами дає кращі прогнози зростання ВВП, ніж тижневий трекер ОЕСР та будь-яка комбінація індексів. Нарешті, можна зробити висновок, що існують зовнішні економічні чинники, які впливають на зростання ВВП, суттєво впливаючи на точність аналізу за допомогою всіх або частини моделей, подібних до змін в оподаткуванні у випадку Ірландії.

РОЗДІЛ 4 ФУНКЦІОНАЛЬНО-ВАРТІСНИЙ АНАЛІЗ ПРОГРАМНОГО ПРОДУКТУ

В цьому розділі буде проведено оцінку основних характеристик майбутнього програмного продукту, який спеціалізуватиметься на дослідженні демографічного стану. Ця реалізація сприятиме проведенню необхідних досліджень, дозволяючи якісно аналізувати демографічні питання не лише в Україні, але й по всьому світу.

Дослідження також демонструє різні варіанти реалізації для забезпечення найкоректнішої та оптимальної стратегії вибору, яка впливатиме на економічні фактори та сумісність з майбутнім програмним продуктом. Для цього використовується функціонально-вартісний аналіз (ФВА).

Функціонально-вартісний аналіз є технологією, що дозволяє оцінити реальну вартість продукту або послуги незалежно від організаційної структури компанії. ФВА проводиться для виявлення резервів зниження витрат завдяки ефективнішим варіантам виробництва, кращому співвідношенню між споживчою вартістю виробу та витратами на його виготовлення. Для аналізу використовуються економічні, технічні та конструкторські дані.

Алгоритм ФВА включає визначення послідовності етапів розробки продукту, визначення річних витрат та кількості робочих годин, виявлення джерел витрат та кінцевий розрахунок вартості програмного продукту.

4.1 Постановка задачі проектування

У роботі застосовується метод ФВА для проведення техніко-економічного аналізу розробки системи прогнозу стійкості фінансових показників. Оскільки рішення стосовно проектування та реалізації компонентів, що розробляється, впливають на всю систему, кожна окрема підсистема має її задовольняти. Тому фактичний аналіз представляє собою аналіз функцій програмного продукту, призначеного для збору, обробки та проведення аналізу даних по компанії.

Технічні вимоги до програмного продукту є наступні:

- функціонування на персональних комп'ютерах із стандартним набором компонентів;
- зручність та зрозумілість для користувача;
- швидкість обробки даних та доступ до інформації в реальному часі;
- можливість зручного масштабування та обслуговування;
- мінімальні витрати на впровадження програмного продукту.

4.2 Обґрунтування функцій програмного продукту

Головна функція F_0 – розробка можливого програмного продукту, яка дозволяє аналізувати різні характеристики, що безпосередньо впливають на стійкість підприємства. Беручи за основу цю функцію, можна виділити наступні:

F_1 – вибір самої програми.

F_2 – якісний аналіз даних.

F_3 – графічні показники.

Кожна з цих функцій має декілька варіантів реалізації:

Функція F_1 :

а) R.

б) Python.

Функція F_2 :

а) Застосування вбудованих функцій.

б) Створення своїх обчислень значень.

Функція F_3 :

а) Використання шаблонних графіків.

б) Створення своїх.

Варіанти реалізації основних функцій наведені у морфологічній карті системи (див. рис. 4.1).

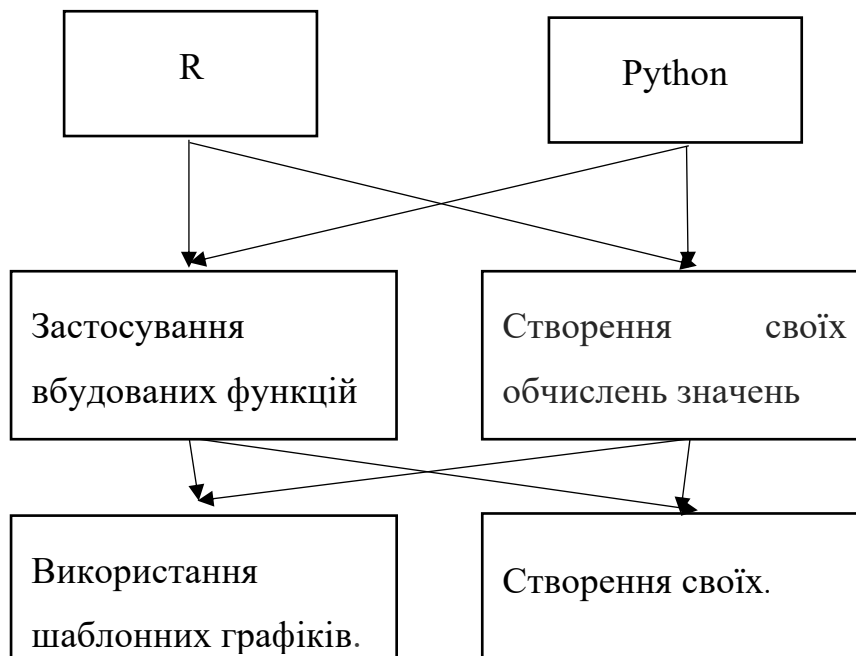


Рисунок 4.1 – Морфологічна карта

Морфологічна карта відображає множину всіх можливих варіантів основних функцій. Позитивно-негативна матриця показана в таблиці 4.1.

Таблиця 4.1 - Позитивно-негативна матриця

Функції	Варіанти реалізації	Переваги	Недоліки
F_1	<i>A</i>	Доступна в реалізації програма для різних обчислень	На написання коду необхідно мати базові навички та вміння
	<i>B</i>	Загальнодоступна програма, доступність багатьох бібліотек	Необхідність повної реалізації алгоритму
F_2	<i>A</i>	Доступність та легкість при написанні	Іноді не відповідає задачі яку треба розв'язати
	<i>B</i>	Ідеально описують усі необхідні характеристики	Достатньо затратно реалізовувати свої алгоритми для подальшої реалізації
F_3	<i>A</i>	Загально прийнята реалізація	Іноді не відповідає очікуваним значенням
	<i>B</i>	При виконанні власних досліджень краще може передавати висновки	Необхідно достатньо багато часу для написання програми для побудови та знаходження всього необхідного в задачі.

На основі аналізу позитивно-негативної матриці робимо висновок, що при розробці програмного продукту деякі варіанти реалізації функцій варто відкинути, тому, що вони не відповідають поставленим перед програмним продуктом задачам. Ці варіанти відзначені у морфологічній карті.

Функція F_1 :

Перевагу даємо реалізації обчислень. Для спрощення роботи по написанню коду варіант Б має бути відкинутий.

Функція F_2 :

Програма допускає обрання обох варіантів. Можливо використати варіанти А чи Б.

Функція F_3 :

Реалізація першого варіанту є сприйнятливою для програми. Це варіант А.

Таким чином, будемо розглядати такий варіанти реалізації ПП:

$$F_1a - F_2a - F_3a$$

$$F_1a - F_2б - F_3a$$

Для оцінювання якості розглянутих функцій обрана система параметрів, описана нижче.

4.3 Обґрунтування системи параметрів програмного продукту

На основі даних, розглянутих вище, визначаються основні параметри вибору, які будуть використані для розрахунку коефіцієнта технічного рівня.

Для того, щоб охарактеризувати програмний продукт, будемо використовувати наступні параметри:

- $X1$ – швидкодія мови програмування;
- $X2$ – об'єм пам'яті для обчислень та збереження даних;
- $X3$ – час навчання даних;
- $X4$ – потенційний об'єм програмного коду.

Гірші, середні і кращі значення параметрів вибираються на основі вимог замовника й умов, що характеризують експлуатацію програмного продукту, як показано у таблиці 4.2.

Таблиця 4.2 - Основні параметри програмного продукту

Назва Параметра	Умовні позначення	Одиниці виміру	Значення параметра		
			гірші	середні	кращі
Швидкодія мови програмування	X1	оп/мс	80	110	120
Об'єм пам'яті	X2	Мб	100	70	50
Час попередньої обробки даних	X3	мс	90	60	50
Потенційний об'єм програмного коду	X4	кількість рядків коду	40	30	20

За даними таблиці 4.3 будуються графічні характеристики параметрів – рис. 4.2 – рис. 4.5.

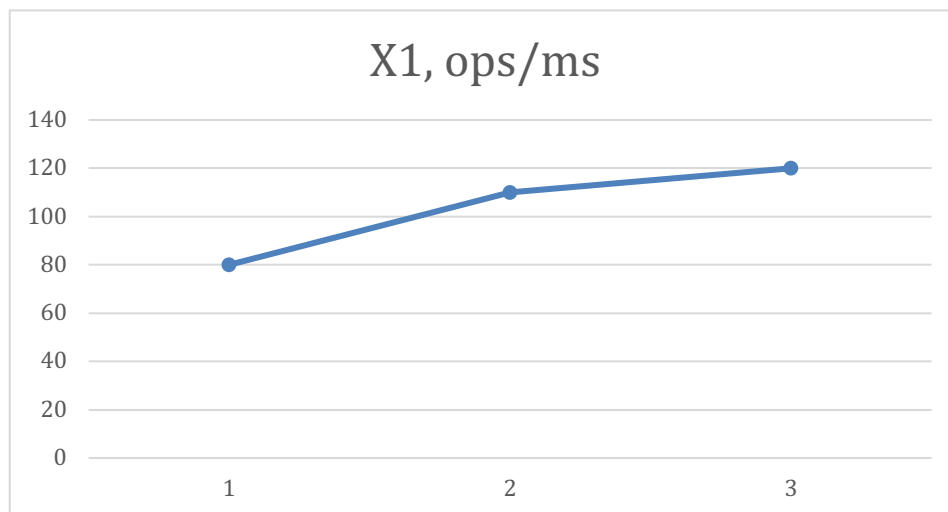


Рисунок 4.2 – X1, швидкодія мови програмування

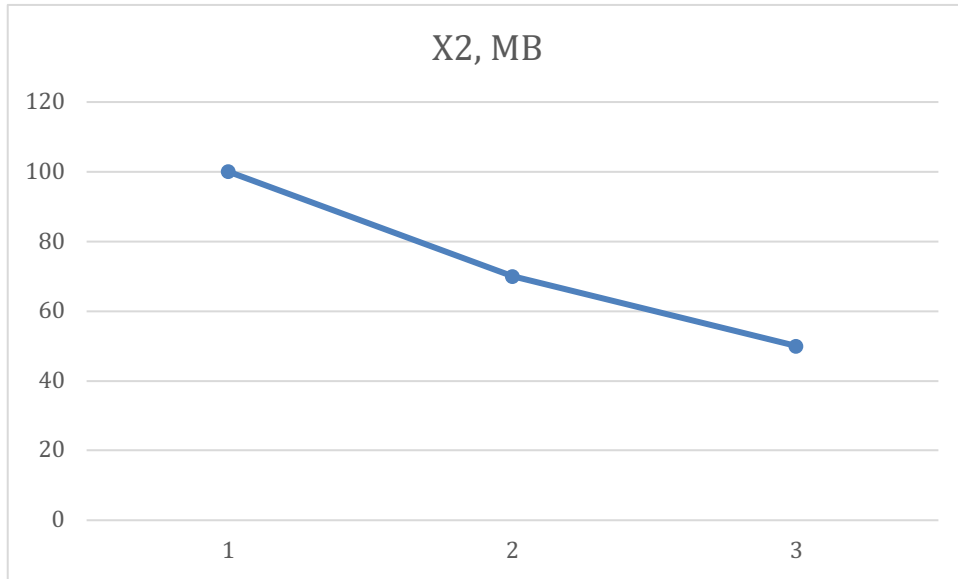


Рисунок 4.3 – X2, об'єм пам'яті

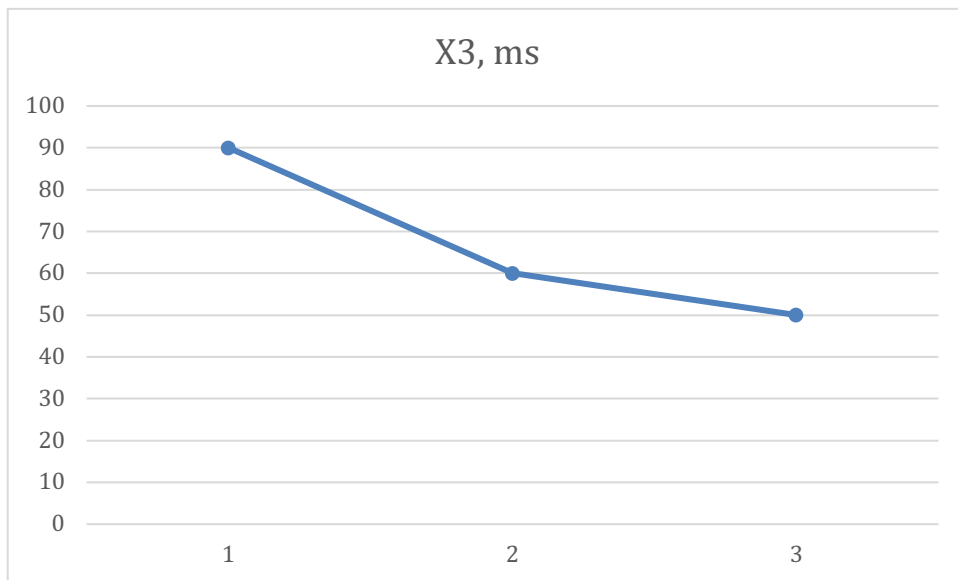


Рисунок 4.4 – X3, час попередньої обробки даних

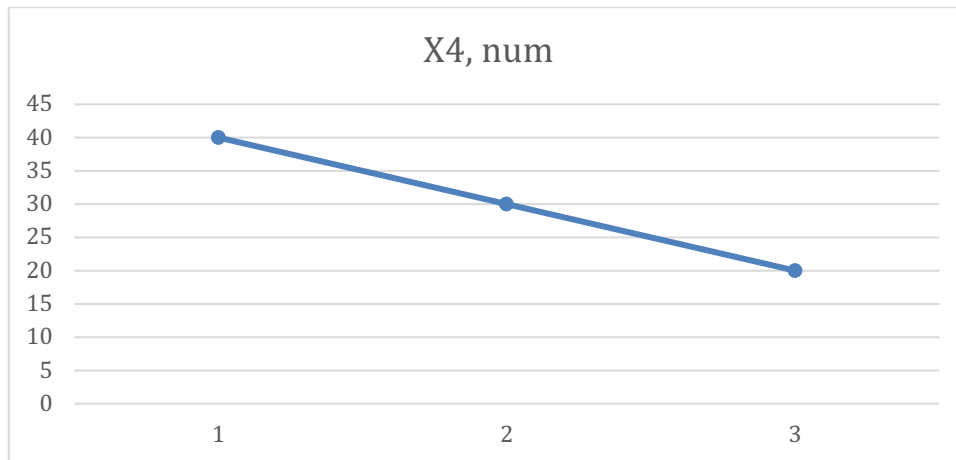


Рисунок 4.5 – X4, потенційний об'єм програмного коду

4.4 Аналіз експертного оцінювання параметрів

Після детального обговорення й аналізу кожний експерт оцінює ступінь важливості кожного параметру для конкретно поставленої цілі – розробка програмного продукту, який дає найбільш точні результати при знаходженні параметрів моделей адаптивного прогнозування і обчислення прогнозних значень.

Значимість кожного параметра визначається методом попарного порівняння. Оцінку проводить експертна комісія із 7 людей. Визначення коефіцієнтів значимості передбачає:

- визначення рівня значимості параметра шляхом присвоєння різних рангів;
- перевірку придатності експертних оцінок для подальшого використання;
- визначення оцінки попарного пріоритету параметрів;
- обробку результатів та визначення коефіцієнту значимості.

Результати експертного ранжування наведені у таблиці 4.3.

Таблиця 4.3 - Результати ранжування параметрів

Позначення параметра	Назва параметра	Одиниці виміру	Ранг параметра за оцінкою експерта							Сума рангів R_i	Відхилення Δ_i	Δ_i^2
			1	2	3	4	5	6	7			
X1	Швидкодія мови програмування	Оп/мс	1	2	2	1	2	1	1	10	-7,5	56,25
X2	Об'єм пам'яті	Мб	4	4	3	4	3	4	3	25	7,5	56,25
X3	Час попередньої обробки даних	мс	2	1	1	2	1	2	2	11	-6,5	42,25
X4	Потенційний об'єм програмного коду	Кількість рядків коду	3	3	4	3	4	3	4	24	6,5	42,25
	Разом		10	10	10	10	10	10	10	70	0	197

Для перевірки степені достовірності експертних оцінок, визначимо наступні параметри:

а) сума рангів кожного з параметрів і загальна сума рангів:

$$R_i = \sum_{j=1}^N r_{ij} R_{ij} = \frac{Nn(n+1)}{2} = 70, \quad (4.1)$$

де N – число експертів,

n – кількість параметрів;

б) середня сума рангів:

$$T = \frac{1}{n} R_{ij} = 17,5 \quad (4.2)$$

Числове значення, що визначає ступінь переваги i -го параметра над j -тим, a_{ij} визначається по формулі:

$$a_{ij} = \begin{cases} 1.5 \text{ при } X_i > X_j \\ 1.0 \text{ при } X_i = X_j \\ 0.5 \text{ при } X_i < X_j \end{cases} \quad (4.6)$$

З отриманих числових оцінок переваги складемо матрицю $A = \|a_{ij}\|$.

Для кожного параметра зробимо розрахунок вагомості K_{vi} за наступними формулами:

$$K_{vi} = \frac{b_i}{\sum_{i=1}^n b_i} \quad (4.7)$$

$$b_i = \sum_{j=1}^N a_{ij} \quad (4.8)$$

Відносні оцінки розраховуються декілька разів доти, поки наступні значення не будуть незначно відрізнятись від попередніх (менше 2%). На другому і наступних кроках відносні оцінки розраховуються за наступними формулами:

$$K_{vi} = \frac{b'_i}{\sum_{i=1}^n b'_i}, \quad (4.9)$$

$$b'_i = \sum_{j=1}^N a_{ij} b_j \quad (4.10)$$

Як видно з таблиці 4.5, різниця значень коефіцієнтів вагомості не перевищує 2%, тому більшої кількості ітерацій не потрібно.

Таблиця 4.5 - Розрахунок вагомості параметрів

Параметри x_i	Параметри x_j				Перша ітер.		Друга ітер.		Третя ітер	
	X1	X2	X3	X4	b_i	K_{Bi}	b_i^1	K_{Bi}^1	b_i^2	K_{Bi}^2
X1	1	0,5	1,5	0,5	3,5	0,22	12,25	0,21	44,875	0,21
X2	1,5	1	1,5	1,5	5,5	0,34	21,25	0,36	77,875	0,36
X3	0,5	0,5	1	0,5	2,5	0,16	9,25	0,16	34,125	0,16
X4	1,5	0,5	1,5	1	4,5	0,28	16,25	0,27	59,125	0,27
Всього:					16	1	59	1	216	1

4.5 Аналіз рівня якості варіантів реалізації функцій

Визначаємо рівень якості кожного варіанту виконання основних функцій окремо.

Абсолютні значення параметрів X2 (Об'єм пам'яті), X3 (час попередньої обробки даних) та X4 (потенційний об'єм програмного коду) відповідають технічним вимогам умов функціонування даного ПП.

Абсолютне значення параметра X1 (швидкість роботи мови програмування) обрано не найгіршим.

Коефіцієнт технічного рівня для кожного варіанта реалізації ПП розраховується так (таблиця 4.6):

$$K_K(j) = \sum_{i=1}^n K_{ei,j} B_{i,j}, \quad (4.11)$$

де n – кількість параметрів;

K_{ei} – коефіцієнт вагомості i -го параметра;

B_i – оцінка i -го параметра в балах.

Таблиця 4.6 - Розрахунок показників рівня якості варіантів реалізації основних функцій ПП

Основні функції	Варіант реалізації функції	Параметри	Абсолютне значення параметра	Бальна оцінка параметра	Коефіцієнт вагомості параметра	Коефіцієнт рівня якості
F1	A	X1	115	25	0,21	5,25
F2	A	X2	75	22	0,36	7,92
	Б	X3	60	25	0,16	4
F3	A	X4	20	30	0,27	8,1

За даними з таблиці 4.6 за формулою:

$$K_K = K_{TY}[F_{1k}] + K_{TY}[F_{2k}] + \dots + K_{TY}[F_{zk}], \quad (4.12)$$

визначаємо рівень якості кожного з варіантів:

$$K_{K1} = 5,25 + 7,92 + 8,1 = 21,27 ;$$

$$K_{K2} = 5,25 + 4 + 8,1 = 17,35 .$$

Як видно з розрахунків, кращим є 1 варіант, для якого коефіцієнт технічного рівня має найбільше значення.

4.6 Економічний аналіз варіантів розробки ПП

Для визначення вартості розробки ПП спочатку проведемо розрахунок трудомісткості.

Всі варіанти включають в себе два окремих завдання:

1. Розробка проекту програмного продукту;
2. Розробка програмної оболонки;

Завдання 1 за ступенем новизни відноситься до групи А, завдання 2 – до групи Б. За складністю алгоритми, які використовуються в завданні 1 належать до групи 1; а в завданні 2 – до групи 3.

Для реалізації завдання 1 використовується довідкова інформація, а завдання 2 використовує інформацію у вигляді даних.

Проведемо розрахунок норм часу на розробку та програмування для кожного з завдань.

Загальна трудомісткість обчислюється як:

$$T_0 = T_P \cdot K_{\Pi} \cdot K_{СК} \cdot K_M \cdot K_{СТ} \cdot K_{СТ.М}, \quad (4.13)$$

де T_P – трудомісткість розробки ПП;

K_{Π} – поправочний коефіцієнт;

$K_{СК}$ – коефіцієнт на складність вхідної інформації;

K_M – коефіцієнт рівня мови програмування;

$K_{СТ}$ – коефіцієнт використання стандартних модулів і прикладних програм;

$K_{СТ.М}$ – коефіцієнт стандартного математичного забезпечення

Для першого завдання, виходячи із норм часу для завдань розрахункового характеру ступеню новизни А та групи складності алгоритму 1, трудомісткість дорівнює: $T_P = 33$ людино-днів. Поправочний коефіцієнт, який враховує вид нормативно-довідкової інформації для першого завдання: $K_{\Pi} = 1.7$. Поправочний коефіцієнт, який враховує складність контролю вхідної та вихідної інформації для всіх семи завдань рівний 1: $K_{СК} = 1$. Оскільки при розробці першого завдання

використовуються стандартні модулі, врахуємо це за допомогою коефіцієнта $K_{CT} = 0,9$. Тоді загальна трудомісткість програмування першого завдання дорівнює:

$$T_1 = 33 \cdot 1,4 \cdot 0,9 = 41,58 \text{ людино-днів.}$$

Проведемо аналогічні розрахунки для подальших завдань.

Для другого завдання (використовується алгоритм третьої групи складності, степінь новизни Б), тобто $T_P = 24$ людино-днів, $K_{П} = 0,8$, $K_{СК} = 1$, $K_{СТ} = 0,7$:

$$T_2 = 24 \cdot 0,8 \cdot 0,7 = 13,44 \text{ людино-днів.}$$

Складаємо трудомісткість відповідних завдань для кожного з обраних варіантів реалізації програми, щоб отримати їх трудомісткість:

$$T_I = 41,58 \cdot 8 = 332,64 \text{ людино-годин.}$$

$$T_{II} = 13,44 \cdot 8 = 107,52 \text{ людино-годин.}$$

Найбільш високу трудомісткість має варіант II.

В розробці беру участь один дата аналітик з окладом 25000 грн.. Визначимо середню зарплату за годину за формулою:

$$C_{ч} = \frac{M}{T_m \cdot t} \text{ грн.,} \quad (4.14)$$

де M – місячний оклад працівників;

T_m – кількість робочих днів тиждень;

t – кількість робочих годин в день.

$$C_{\text{ч}} = \frac{25000}{3 \cdot 21 \cdot 8} = 49,6 \text{ грн.} \quad (4.15)$$

Тоді, розрахуємо заробітну плату за формулою:

$$C_{\text{зп}} = C_{\text{ч}} \cdot T_i \cdot K_{\text{д}}, \quad (4.16)$$

де $C_{\text{ч}}$ – величина погодинної оплати праці програміста;

T_i – трудомісткість відповідного завдання;

$K_{\text{д}}$ – норматив, який враховує додаткову заробітну плату.

Зарплата розробників за варіантами становить:

$$\text{I.} \quad C_{\text{зп}} = 49,6 \cdot 332,64 \cdot 1,2 = 19798,74 \text{ грн.}$$

$$\text{II.} \quad C_{\text{зп}} = 49,6 \cdot 107,52 \cdot 1,2 = 6399,6 \text{ грн.}$$

Відрахування на єдиний соціальний внесок становить 22%:

$$\text{I.} \quad C_{\text{вд}} = C_{\text{зп}} \cdot 0,22 = 19798,74 \cdot 0,22 = 4355,8 \text{ грн.}$$

$$\text{II.} \quad C_{\text{вд}} = C_{\text{зп}} \cdot 0,22 = 6399,67 \cdot 0,22 = 1408 \text{ грн.}$$

Тепер визначимо витрати на оплату однієї машино-години. ($C_{\text{м}}$)

Так як одна ЕОМ обслуговує одного програміста з окладом 25000 грн., з коефіцієнтом зайнятості 0,2 то для однієї машини отримаємо:

$$C_{\text{г}} = 12 \cdot M \cdot K_3 = 12 \cdot 25000 \cdot 0,2 = 60000 \text{ грн.}$$

З урахуванням додаткової заробітної плати:

$$C_{зп} = C_{г} \cdot (1 + K_{з}) = 60000 \cdot (1 + 0.2) = 72000 \text{ грн.}$$

Відрахування на соціальний внесок:

$$C_{від} = C_{зп} \cdot 0.22 = 72000 \cdot 0.22 = 15840 \text{ грн.}$$

Амортизаційні відрахування розраховуємо при амортизації 25% та вартості ЕОМ – 27000 грн.

$$C_A = K_{тм} \cdot K_A \cdot Ц_{пр} = 1,4 \cdot 0,25 \cdot 27000 = 9450 \text{ грн.,}$$

де $K_{тм}$ – коефіцієнт, який враховує витрати на транспортування та монтаж приладу у користувача;

K_A – річна норма амортизації;

$Ц_{пр}$ – договірна ціна приладу.

Витрати на ремонт та профілактику розраховуємо як:

$$C_P = K_{тм} \cdot Ц_{пр} \cdot K_P = 1.4 \cdot 27000 \cdot 0.08 = 3024 \text{ грн.,}$$

де K_P – відсоток витрат на поточні ремонти.

Ефективний годинний фонд часу ПК за рік розраховуємо за формулою:

$$\begin{aligned} T_{эф} &= (D_K - D_B - D_C - D_P) \cdot t_3 \cdot K_B = (365 - 104 - 12 - 16) \cdot 8 \cdot 0.35 = \\ &= 627,2 \text{ години,} \end{aligned}$$

де D_K – календарна кількість днів у році;

D_B, D_C – відповідно кількість вихідних та святкових днів;

D_p – кількість днів планових ремонтів устаткування;

t – кількість робочих годин в день;

K_B – коефіцієнт використання приладу у часі протягом зміни.

Витрати на оплату електроенергії розраховуємо за формулою:

$$C_{\text{ЕЛ}} = T_{\text{ЕФ}} \cdot N_C \cdot K_3 \cdot C_{\text{ЕН}} = 627,2 \cdot 0,2 \cdot 0,3 \cdot 5,22 = 196,44 \text{ грн.},$$

де N_C – середньо-споживча потужність приладу;

K_3 – коефіцієнтом зайнятості приладу;

$C_{\text{ЕН}}$ – тариф за 1 КВт-годин електроенергії.

Накладні витрати розраховуємо за формулою:

$$C_H = C_{\text{ПР}} \cdot 0,67 = 27000 \cdot 0,67 = 18090 \text{ грн.}$$

Тоді, річні експлуатаційні витрати будуть:

$$C_{\text{ЕКС}} = C_{\text{ЗП}} + C_{\text{ВІД}} + C_A + C_P + C_{\text{ЕЛ}} + C_H, \quad (4.17)$$

$$C_{\text{ЕКС}} = 72000 + 15840 + 9450 + 3024 + 196,44 + 18090 = 118600,44 \text{ грн.}$$

Собівартість однієї машино-години ЕОМ дорівнюватиме:

$$C_{\text{М-Г}} = C_{\text{ЕКС}} / T_{\text{ЕФ}} = 118600,44 / 627,2 = 189,09 \text{ грн/год.}$$

Оскільки в даному випадку всі роботи, які пов'язані з розробкою програмного продукту ведуться на ЕОМ, витрати на оплату машинного часу, в залежності від обраного варіанта реалізації, складає:

$$C_M = C_{M-\Gamma} \cdot T, \quad (4.18)$$

$$\text{I. } C_M = 189,09 \cdot 332,64 = 62898,9 \text{ грн.}$$

$$\text{II. } C_M = 189,09 \cdot 107,52 = 20330,96 \text{ грн.}$$

Накладні витрати складають 67% від заробітної плати:

$$C_H = C_{ЗП} \cdot 0,67, \quad (4.19)$$

$$\text{I. } C_H = 19798,74 \cdot 0,67 = 13265,16 \text{ грн.}$$

$$\text{II. } C_H = 6399,6 \cdot 0,67 = 4287,73 \text{ грн.}$$

Отже, вартість розробки ПП за варіантами становить:

$$C_{ПП} = C_{ЗП} + C_{ВІД} + C_M + C_H, \quad (4.20)$$

$$\text{I. } C_{ПП} = 19798,74 + 4355,8 + 62898,9 + 13265,16 = 100318,6 \text{ грн.}$$

$$\text{II. } C_{ПП} = 6399,6 + 1408 + 20330,96 + 4287,73 = 32426,29 \text{ грн.}$$

4.7 Вибір кращого варіанту III техніко-економічного рівня

Розрахуємо коефіцієнт техніко-економічного рівня за формулою:

$$K_{\text{TEP}j} = K_{\text{Kj}} / C_{\text{Фj}}, \quad (4.21)$$

$$K_{\text{TEP1}} = 21,27 / 100318,6 = 2,12 \cdot 10^{-4},$$

$$K_{\text{TEP2}} = 17,35 / 32426,29 = 5,35 \cdot 10^{-4}.$$

Як бачимо, найбільш ефективним є другий варіант реалізації програми з коефіцієнтом техніко-економічного рівня $K_{\text{TEP1}} = 5,35 \cdot 10^{-4}$.

Після виконання функціонально-вартісного аналізу програмного комплексу що розроблюється, можна зробити висновок, що з альтернатив, що залишилися після першого відбору двох варіантів виконання програмного комплексу оптимальним є перший варіант реалізації програмного продукту. У нього виявився найкращий показник техніко-економічного рівня якості $K_{\text{TEP}} = 5,35 \cdot 10^{-4}$.

Цей варіант реалізації програмного продукту має такі параметри:

- Вибір програмного продукту – R;
- Створення своїх обчислень значень;
- Використання стандартного інтерфейсу для побудови значень.

Даний варіант виконання програмного комплексу дає користувачу зручний інтерфейс, швидку реалізацію програми та доступний функціонал для роботи.

4.8 Висновки до розділу 4

В даному розділі було проведено повний функціонально-вартісний аналіз програмного продукту. Також було знайдено оцінку основних функцій програмного продукту.

В результаті виконання функціонально-вартісного аналізу програмного комплексу що розроблюється, було визначено та проведено оцінку основних функцій програмного продукту, а також знайдено параметри, які його характеризують.

На основі аналізу вибрано варіант реалізації програмного продукту.

ВИСНОВКИ

Високочастотні макроекономічні індекси стали повсякденним інструментом для економістів і політиків для реагування на швидкі зміни в економіці. Нещодавно розроблений тижневий трекер ОЕСР отримав мало уваги з боку колег-вчених. Мета цієї роботи — дослідити точність наукастингу індексу та провести аналіз його ефективності відносно PMI, ESI та їх комбінацій. У даній роботі досліджується здатність тижневого трекеру ОЕСР прогнозувати зростання ВВП за загальний доступний період, а також за винятком періодів високої економічної невизначеності (глобальна фінансова криза та пандемія COVID-19). Крім того, порівнюємо точність наукастингу індексу ОЕСР з індексом менеджерів із закупівель (PMI) та індикатором економічних настроїв (ESI) окремо та разом.

Було виявлено, що тижневий трекер ОЕСР забезпечує більш точний наукастинг, виміряні RMSE, ніж будь-який інший показник. Було припущено, що це може бути пов'язано з величезною різницею в охопленні даними. Також виявлено, що темпи зростання ВВП за попередні два квартали особливо ефективні для прогнозування на наступний період для стабільних економік у періоди з низькою волатильністю. Крім того, показано, що незважаючи на відмінності в охопленні, тижневий трекер ОЕСР у поєднанні з одним або двома розглянутими якісними індикаторами може надавати більш точні результати за наявності аспектів, які не враховує Google Trends. Нарешті, аналіз результатів показує, що використання Інтернету серед населення країни та людський розвиток позитивно впливають на точність наукастингу тижневого трекеру економічної активності ОЕСР.

Висновки, представлені в даному дослідженні, актуальні для професіоналів галузі, дослідників і майбутніх досліджень, які потребують високочастотних даних

для оцінки економічних змін у короткостроковій перспективі. Трекер, автором якого є Nicolas Woloszko, дає змогу переглядати економічну ситуацію в реальному часі та в поєднанні з іншими показниками тепер може з високою точністю прогнозувати зміни зростання ВВП, особливо перевершуючи інші показники під час економічної нестабільності.

Дана робота має стимулювати майбутні дослідження цієї теми, особливо під час відстеження та кількісного визначення аспектів, які не охоплюються відповідними категоріями Google Trends, які використовуються для формування трекера. Більше того, можна припустити, що вимірювання та визначення ефектів, що стосуються окремих країн, а не загальних економічних аспектів, підвищать точність наукастингу.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Cepni O., Güney I. E., Swanson N. R. Nowcasting and forecasting GDP in emerging markets using global financial and macroeconomic diffusion indexes. *International Journal of Forecasting*, 2019, 35(2), P. 555–572. URL: https://mycourses.aalto.fi/pluginfile.php/1685887/mod_resource/content/1/Lecture_2_readingNew.pdf
2. OECD, Gross domestic product (GDP) (indicator). Retrieved on January 2nd, 2023. URL: <https://data.oecd.org/gdp/gross-domestic-product-gdp.htm>
3. Turner D. The use of models in producing OECD macroeconomic forecasts, 2016. URL: https://www.researchgate.net/profile/David-Turner-7/publication/309480216_The_use_of_models_in_producing_OECD_macroecomic_forecasts/links/58127c0d08aea2cf64e26083/The-use-of-models-in-producing-OECD-macroeconomic-forecasts.pdf
4. Vosen S., Schmidt T. Forecasting private consumption: survey-based indicators vs. Google trends. *Journal of forecasting*, 2011, 30(6), P. 565-578. URL: <https://www.econstor.eu/bitstream/10419/29900/1/614061253.pdf>
5. Bok B., Caratelli D., Giannone D., Sbordone A. M., Tambalotti A. Macroeconomic nowcasting and forecasting with big data. *Annual Review of Economics*, 2017, 10, P. 615-643. URL: <https://www.econstor.eu/bitstream/10419/189871/1/sr830.pdf>
6. Kennedy A., Hauksson K. M. *Global search engine marketing: Fine-tuning your international search engine results*. Que Publishing, 2012. URL: <https://tinyurl.com/44xsd39y>
7. Trehan B. Forecasting growth in current quarter real GNP. *Economic Review-Federal Reserve Bank of San Francisco*, 1989, (1), 39. URL:

<https://www.proquest.com/openview/b356d10e2bca990e6c93f9e24cbd608e/1?pq-origsite=gscholar&cbl=41365>

8. Rünstler G., Sedillot F. Short-term estimates of euro area real GDP by means of monthly data, 2003, No. 276, ECB Working Paper. URL: <https://www.econstor.eu/bitstream/10419/152710/1/ecbwp0276.pdf>
9. Evans M. Where are we now? real-time estimates of the macro economy, 2005. URL: https://www.nber.org/system/files/working_papers/w11064/w11064.pdf
10. Banbura M., Giannone D., Reichlin L. Nowcasting, 2010. ECB Working Paper No. 1275, URL: <https://ssrn.com/abstract=1717887>
11. Bańbura M., Rünstler G. A look into the factor model black box: Publication lags and the role of hard and soft data in forecasting GDP. *International Journal of Forecasting*, 2011, 27(2), P. 333-346 URL: <https://www.econstor.eu/bitstream/10419/153185/1/ecbwp0751.pdf>
12. Ettredge M., Gerdes J., Karuga G. Using web-based search data to predict macroeconomic statistics. *Communications of the ACM*, 2005, 48(11), P. 87-92. URL: <http://portal.acm.org/citation.cfm?id=1096010>
13. Choi H., Varian H. Predicting the present with Google Trends. *Economic record*, 2012, 88, P. 2-9. URL: <https://www.datascienceassn.org/sites/default/files/Predicting%20the%20Present%20with%20Google%20Trends.pdf>
14. Goel S., Hofman J. M., Lahaie S., Pennock D. M., Watts D. J. Predicting consumer behavior with Web search. *Proceedings of the National academy of sciences*, 2010, 107(41), P. 17486-17490. URL: <http://www.pnas.org/content/early/2010/09/20/1005962107>
15. D'Amuri F., Marcucci J. The predictive power of Google searches in forecasting US unemployment. *International Journal of Forecasting*, 2017, 33(4), P. 801-816. URL:

<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=7dc47a4f68a44c37345573ff35118b9c124a8122>

16. Boriss A. Siliverstovs, Daniel S. Wochner, Google Trends and reality: Do the proportions match?, 2017. URL: <https://doi.org/10.1016/j.jebo.2017.10.011>
17. Kohns D., Bhattacharjee A. Nowcasting Growth using Google Trends Data: A Bayesian Structural Time Series Model. International Journal of Forecasting, 2022. URL: <https://www.sciencedirect.com/science/article/pii/S0169207022000620>
18. Koenig EF. Using the purchasing managers' index to assess the economy's strength and the likely direction of monetary policy, Economic and Financial Policy Review, 2002. URL: <https://ideas.repec.org/a/fip/fedder/y2002nv.1no.6.html>
19. Woloszko N. Tracking activity in real time with Google Trends, 2020. OECD. URL: <https://doi.org/10.1787/6b9c7518-en>
20. Lewis D. J., Mertens K., Stock J. H., Trivedi M. Measuring real activity using a weekly economic index. Journal of Applied Econometrics, 2021, 37(4), P. 667–687. <https://doi.org/10.1002/jae.2873>
21. S&P Global, 2023, PMI by S&P Global. URL: <https://www.spglobal.com/marketintelligence/en/mi/products/pmi.html>
22. Schumacher C. A comparison of MIDAS and bridge equations. International Journal of Forecasting, 2016, 32(2), P. 257-270. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0169207015001144#:~:text=The%20bridge%20equation%20differs%20from,only%20in%20the%20lag%20polynomials.>
23. Eurostat, 2023. Economic Sentiment Indicator, URL: <https://data.europa.eu/data/datasets/c04buuz6wxiqgjkhpwlug?locale=en>
24. Kostenko A. V., Hyndman R. J. Forecasting without significance tests? *manuscript, Monash University, Australia, 2008.* URL:

- <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=f9e981ce476390a452f699e4f36f4acdf1450c16>
25. Jargowsky P. A. Omitted variable bias. *Encyclopedia of social measurement*, 2005, 2, P. 919-924. URL: <https://scholar.archive.org/work/ntxu3hihkb7fakzuvqpaglza/access/wayback/http://www.pub.utdallas.edu:80/~jargo/papers/omitted%20variables.pdf>
 26. World bank. *Individuals using the Internet (% of population)*, 2023. URL: <https://data.worldbank.org/indicator/IT.NET.USER.ZS>
 27. The Treasury of the Australian Government. *Australia's manufactures exports*, 2006. URL: https://treasury.gov.au/sites/default/files/2019-03/05_HoR.pdf
 28. Tedeschi R. The Irish GDP in 2016. After the disaster comes a dilemma. Bank of Italy Occasional Paper, 2018, (471). URL: http://www.bancaditalia.it/pubblicazioni/qef/2018-0471/QEF_471_18.pdf
 29. Edwards S. Monetary policy in Iceland: An evaluation. *Report prepared for the Government of Iceland for a review of Iceland's monetary and exchange rate policy*, 2018. URL: <https://www.government.is/library/01-Ministries/Prime-Ministrers-Office/Endurskodun-a-ramma-peningastefnu/ICELAND%20REPORT%20Edwards%20REVISED%2006%2004%202018.pdf>
 30. Network Readiness Index. *Benchmarking the Future of the Network Economy*, 2023. URL: <https://networkreadinessindex.org/>
 31. Ünver H. Explaining education level and internet penetration by economic reasoning-worldwide analysis from 2000 through 2010. *Int J Inf Secur*, 2014, 7, P. 898-912. URL: <https://infonomics-society.org/wp-content/uploads/iji/published-papers/volume-7-2014/Explaining-Education-Level-and-Internet-Penetration.pdf>
 32. Chauvet M., Potter S. Forecasting Output. *Handbook of Economic Forecasting*, 2013, P. 141–194. URL: <https://doi.org/10.1016/B978-0-444-53683-9.00003-7>

33. Ghysels E., Sinko A., Valkanov R. MIDAS Regressions: Further Results and New Directions. *Econometric Reviews*, 2007, 26(1), P. 53–90. URL: <https://doi.org/10.1080/07474930600972467>

ДОДАТОК А ЛІСТИНГ КОДУ

```
library(dplyr)
library(stringr)
library(PerformanceAnalytics)
library(panelr)
library(zoo)
library(lubridate)
library(midasr)
library(ggplot2)
library(forecast)
library(ppcor)
library(broom)
library(graphics)
library(reshape2)
library(tidyverse)
library(forecast)
library(Metrics)
library(openxlsx)

setwd("C:/Users/sapelnikov/Desktop/Diploma/")

dataframe <- read.csv("france_forecasting.csv")
dataframe$Date <- as.Date(dataframe$Date, format = "%d/%m/%Y")

dataframe$Lag_GDP_1 <- lag(dataframe$GDP)
dataframe$Lag_GDP_2 <- lag(dataframe$GDP, 2)

train <- dataframe[1:20,]
test <- dataframe[21:73,]

train_AR <- dataframe[3:23,]
test_AR <- dataframe[24:73,]

train_ESI <- dataframe[1:20,]
test_ESI <- dataframe[21:73,]

train_ESI_OECD <- dataframe[1:20,]
test_ESI_OECD <- dataframe[21:73,]

### OECD forecasting

predictions <- c()
coefs <- c()
p_values <- c()
t_values <- c()
```

```

se_values <- c()

for (i in 2:length(test[, "GDP"])) {
  train <- rbind(train, test[i-1,])
  model <- lm(GDP ~ OECD, data = train)
  prediction <- predict(model, newdata = test[i,])
  predictions <- c(predictions, prediction)

  coef <- coef(model)
  coefs <- rbind(coefs, coef)

  p_value <- summary(model)$coef[,4]
  p_values <- rbind(p_values, p_value)

  t_value <- summary(model)$coef[,3]
  t_values <- rbind(t_values, t_value)

  se_value <- summary(model)$coef[,2]
  se_values <- rbind(se_values, se_value)
}
rmse <- rmse(predictions, test[2:length(test[, "GDP"]), "GDP"])

results <- data.frame(predictions, coefs, p_values, t_values, se_values)
write.xlsx(results, file = "results_oecd_gdp.xlsx")

### Autoregressive model forecasting

predictions_AR <- c()
coefs_AR <- c()
p_values_AR <- c()
t_values_AR <- c()
se_values_AR <- c()

for (i in 2:length(test_AR[, "GDP"])) {
  train_AR <- rbind(train_AR, test_AR[i-1,])
  model_AR <- lm(GDP ~ Lag_GDP_1 + Lag_GDP_2, data = train_AR)
  prediction_AR <- predict(model_AR, newdata = test_AR[i,])
  predictions_AR <- c(predictions_AR, prediction_AR)

  coef_AR <- coef(model_AR)
  coefs_AR <- rbind(coefs_AR, coef_AR)

  p_value_AR <- summary(model_AR)$coef[,4]
  p_values_AR <- rbind(p_values_AR, p_value_AR)

  t_value_AR <- summary(model_AR)$coef[,3]

```

```

t_values_AR <- rbind(t_values_AR, t_value_AR)

se_value_AR <- summary(model_AR)$coef[,2]
se_values_AR <- rbind(se_values_AR, se_value_AR)
}
rmse_AR <- rmse(predictions_AR, test_AR[2:length(test_AR[, "GDP"]), "GDP"])

results_AR <- data.frame(predictions_AR, coefs_AR, p_values_AR, t_values_AR,
se_values_AR)
write.xlsx(results_AR, file = "results_AR.xlsx")

#### ESI forecasting

predictions_ESI <- c()
coefs_ESI <- c()
p_values_ESI <- c()
t_values_ESI <- c()
se_values_ESI <- c()

for (i in 2:length(test_ESI[, "GDP"])) {
  train_ESI <- rbind(train_ESI, test_ESI[i-1,])
  model_ESI <- lm(GDP ~ ESI, data = train_ESI)
  prediction_ESI <- predict(model_ESI, newdata = test_ESI[i,])
  predictions_ESI <- c(predictions_ESI, prediction_ESI)

  coef_ESI <- coef(model_ESI)
  coefs_ESI <- rbind(coefs_ESI, coef_ESI)

  p_value_ESI <- summary(model_ESI)$coef[,4]
  p_values_ESI <- rbind(p_values_ESI, p_value_ESI)

  t_value_ESI <- summary(model_ESI)$coef[,3]
  t_values_ESI <- rbind(t_values_ESI, t_value_ESI)

  se_value_ESI <- summary(model_ESI)$coef[,2]
  se_values_ESI <- rbind(se_values_ESI, se_value_ESI)
}
rmse_ESI <- rmse(predictions_ESI, test_ESI[2:length(test_ESI[, "GDP"]), "GDP"])

results_ESI <- data.frame(predictions_ESI, coefs_ESI, p_values_ESI, t_values_ESI,
se_values_ESI)
write.xlsx(results_ESI, file = "results_esi.xlsx")

#### ESI + OECD forecasting

```

```

predictions_ESI_OECD <- c()
coefs_ESI_OECD <- c()
p_values_ESI_OECD <- c()
t_values_ESI_OECD <- c()
se_values_ESI_OECD <- c()

for (i in 2:length(test_ESI_OECD[, "GDP"])) {
  train_ESI_OECD <- rbind(train_ESI_OECD, test_ESI_OECD[i-1,])
  model_ESI_OECD <- lm(GDP ~ OECD + ESI, data = train_ESI_OECD)
  prediction_ESI_OECD <- predict(model_ESI_OECD, newdata = test_ESI_OECD[i,])
  predictions_ESI_OECD <- c(predictions_ESI_OECD, prediction_ESI_OECD)

  coef_ESI_OECD <- coef(model_ESI_OECD)
  coefs_ESI_OECD <- rbind(coefs_ESI_OECD, coef_ESI_OECD)

  p_value_ESI_OECD <- summary(model_ESI_OECD)$coef[,4]
  p_values_ESI_OECD <- rbind(p_values_ESI_OECD, p_value_ESI_OECD)

  t_value_ESI_OECD <- summary(model_ESI_OECD)$coef[,3]
  t_values_ESI_OECD <- rbind(t_values_ESI_OECD, t_value_ESI_OECD)

  se_value_ESI_OECD <- summary(model_ESI_OECD)$coef[,2]
  se_values_ESI_OECD <- rbind(se_values_ESI_OECD, se_value_ESI_OECD)
}
rmse_ESI_OECD <- rmse(predictions_ESI_OECD, test_ESI_OECD[2:length(test_ESI_OECD[,
"GDP"]), "GDP"])

results_ESI_OECD <- data.frame(predictions_ESI_OECD, coefs_ESI_OECD,
p_values_ESI_OECD, t_values_ESI_OECD, se_values_ESI_OECD)
write.xlsx(results_ESI_OECD, file = "results_esi_oecd.xlsx")

#### PMI forecasting ####

train_PMI <- dataframe[1:20,]
test_PMI <- dataframe[21:73,]

predictions_PMI <- c()
coefs_PMI <- c()
p_values_PMI <- c()
t_values_PMI <- c()
se_values_PMI <- c()

for (i in 2:length(test_PMI[, "GDP"])) {
  train_PMI <- rbind(train_PMI, test_PMI[i-1,])
  model_PMI <- lm(GDP ~ PMI, data = train_PMI)
  prediction_PMI <- predict(model_PMI, newdata = test_PMI[i,])
}

```

```

predictions_PMI <- c(predictions_PMI, prediction_PMI)

coef_PMI <- coef(model_PMI)
coefs_PMI <- rbind(coefs_PMI, coef_PMI)

p_value_PMI <- summary(model_PMI)$coef[,4]
p_values_PMI <- rbind(p_values_PMI, p_value_PMI)

t_value_PMI <- summary(model_PMI)$coef[,3]
t_values_PMI <- rbind(t_values_PMI, t_value_PMI)

se_value_PMI <- summary(model_PMI)$coef[,2]
se_values_PMI <- rbind(se_values_PMI, se_value_PMI)
}
rmse_PMI <- rmse(predictions_PMI, test_PMI[2:length(test_PMI[, "GDP"]), "GDP"])

rmse_PMI

results_PMI <- data.frame(predictions_PMI, coefs_PMI, p_values_PMI, t_values_PMI,
se_values_PMI)
write.xlsx(results_PMI, file = "results_PMI_france.xlsx")

#### PMI_OECD forecasting ####

train_PMI_OECD <- dataframe[1:20,]
test_PMI_OECD <- dataframe[21:73,]

predictions_PMI_OECD <- c()
coefs_PMI_OECD <- c()
p_values_PMI_OECD <- c()
t_values_PMI_OECD <- c()
se_values_PMI_OECD <- c()

for (i in 2:length(test_PMI_OECD[, "GDP"])) {
  train_PMI_OECD <- rbind(train_PMI_OECD, test_PMI_OECD[i-1,])
  model_PMI_OECD <- lm(GDP ~ OECD + PMI, data = train_PMI_OECD)
  prediction_PMI_OECD <- predict(model_PMI_OECD, newdata = test_PMI_OECD[i,])
  predictions_PMI_OECD <- c(predictions_PMI_OECD, prediction_PMI_OECD)

  coef_PMI_OECD <- coef(model_PMI_OECD)
  coefs_PMI_OECD <- rbind(coefs_PMI_OECD, coef_PMI_OECD)

  p_value_PMI_OECD <- summary(model_PMI_OECD)$coef[,4]
  p_values_PMI_OECD <- rbind(p_values_PMI_OECD, p_value_PMI_OECD)

  t_value_PMI_OECD <- summary(model_PMI_OECD)$coef[,3]
  t_values_PMI_OECD <- rbind(t_values_PMI_OECD, t_value_PMI_OECD)
}

```

```

    se_value_PMI_OECD <- summary(model_PMI_OECD)$coef[,2]
    se_values_PMI_OECD <- rbind(se_values_PMI_OECD, se_value_PMI_OECD)
  }
rmse_PMI_OECD <- rmse(predictions_PMI_OECD, test_PMI_OECD[2:length(test_PMI_OECD[,
"GDP"]), "GDP"])

rmse_PMI_OECD

results_PMI_OECD <- data.frame(predictions_PMI_OECD, coefs_PMI_OECD,
p_values_PMI_OECD, t_values_PMI_OECD, se_values_PMI_OECD)
write.xlsx(results_PMI_OECD, file = "results_PMI_OECD_france.xlsx")

#### PMI_ESI_OECD forecasting ####

train_PMI_ESI_OECD <- dataframe[1:20,]
test_PMI_ESI_OECD <- dataframe[21:73,]

predictions_PMI_ESI_OECD <- c()
coefs_PMI_ESI_OECD <- c()
p_values_PMI_ESI_OECD <- c()
t_values_PMI_ESI_OECD <- c()
se_values_PMI_ESI_OECD <- c()

for (i in 2:length(test_PMI_ESI_OECD[, "GDP"])) {
  train_PMI_ESI_OECD <- rbind(train_PMI_ESI_OECD, test_PMI_ESI_OECD[i-1,])
  model_PMI_ESI_OECD <- lm(GDP ~ OECD + PMI + ESI, data = train_PMI_ESI_OECD)
  prediction_PMI_ESI_OECD <- predict(model_PMI_ESI_OECD, newdata =
test_PMI_ESI_OECD[i,])
  predictions_PMI_ESI_OECD <- c(predictions_PMI_ESI_OECD, prediction_PMI_ESI_OECD)

  coef_PMI_ESI_OECD <- coef(model_PMI_ESI_OECD)
  coefs_PMI_ESI_OECD <- rbind(coefs_PMI_ESI_OECD, coef_PMI_ESI_OECD)

  p_value_PMI_ESI_OECD <- summary(model_PMI_ESI_OECD)$coef[,4]
  p_values_PMI_ESI_OECD <- rbind(p_values_PMI_ESI_OECD, p_value_PMI_ESI_OECD)

  t_value_PMI_ESI_OECD <- summary(model_PMI_ESI_OECD)$coef[,3]
  t_values_PMI_ESI_OECD <- rbind(t_values_PMI_ESI_OECD, t_value_PMI_ESI_OECD)

  se_value_PMI_ESI_OECD <- summary(model_PMI_ESI_OECD)$coef[,2]
  se_values_PMI_ESI_OECD <- rbind(se_values_PMI_ESI_OECD, se_value_PMI_ESI_OECD)
}
rmse_PMI_ESI_OECD <- rmse(predictions_PMI_ESI_OECD,
test_PMI_ESI_OECD[2:length(test_PMI_ESI_OECD[, "GDP"]), "GDP"])

rmse_PMI_ESI_OECD

```

```
results_PMI_ESI_OECD <- data.frame(predictions_PMI_ESI_OECD, coefs_PMI_ESI_OECD,  
p_values_PMI_ESI_OECD, t_values_PMI_ESI_OECD, se_values_PMI_ESI_OECD)  
write.xlsx(results_PMI_ESI_OECD, file = "results_PMI_ESI_OECD_france.xlsx")
```

ДОДАТОК Б ПРЕЗЕНТАЦІЯ

Дослідження тижневого трекеру ОЕСР як надійного показника економічної діяльності

Підготував: студент групи КА-03

Сапельніков Вячеслав Леонідович

Науковий керівник: Барановська Леся Валеріївна

Актуальність

У сучасному глобалізованому світі економічна ситуація може змінюватися дуже швидко, під впливом різних факторів, таких як політичні події, природні катастрофи, пандемії тощо. Традиційні методи оцінки економічної активності, які базуються на кварталних або річних даних, часто не встигають відобразити ці зміни оперативно. Тижневий трекер ОЕСР надає можливість отримувати оновлені дані кожного тижня, що дозволяє швидше реагувати на зміни та приймати обґрунтовані рішення.

Мета роботи

Основна мета роботи полягає у доведенні надійності та корисності тижневого трекеру ОЕСР як інструмента для швидкого та точного аналізу економічної ситуації у сучасних умовах динамічного розвитку глобальної економіки



Постановка задачі

Дана робота полягає в прогнозі квартального зростання ВВП, використовуючи одноетапну модель прогнозування, і оцінці його за допомогою середньоквадратичної помилки. Показати, що наукастинг, зроблений на основі тижневого трекеру ОЕСР значно перевершує наукастинг, зроблений на основі таких індикаторів як Economic Sentiment Indicator (ESI) та Purchasing Managers' Index (PMI).

Традиційні методи прогнозування

- 1) **Purchasing managers Index (PMI)** – індекс, що базується на опитуванні менеджерів із закупівель у виробничому секторі, яких запитують про їхні очікування щодо майбутнього виробництва, нових замовлень і працевлаштування.
- 2) **Economic sentiment Indicator (ESI)** - це комплексний показник, який відображає загальний стан економіки та рівень економічного оптимізму серед підприємств і споживачів. ESI об'єднує дані з різних секторів економіки, таких як промисловість, будівництво, роздрібна торгівля, послуги та споживацькі настрої, що дозволяє отримати цілісне уявлення про економічні настрої в країні або регіоні.

Тижневий трекер ОЕСР



Щотижневий трекер ОЕСР, розроблений Ніколасом Волошко в 2020 році, є інструментом, який використовується для оцінки темпів зростання щотижневого ВВП країн у багатьох країнах із 5-денною затримкою. Weekly Tracker працює, застосовуючи єдиний алгоритм машинного навчання на панелі даних Google Trends для 46 країн ОЕСР і G20

Алгоритм дослідження

Процес оцінки тижневого трекеру ОЕСР включає збір та попередню обробку даних з різних джерел, таких як економічні індикатори, ВВП та індекси споживчих настроїв. Потім проводиться кореляційний аналіз для визначення відповідності та надійності трекеру у порівнянні з традиційними економічними показниками. Нарешті, результати аналізу використовуються для оцінки точності та корисності трекеру ОЕСР як інструмента для оперативного моніторингу економічної активності.

Агрегація часу

Через те, що PMI, ESI та індекс ОЕСР подано як рівні, ми спочатку перетворюємо їх у показники темпу зростання для відповідних частот, застосовуючи експоненціальну функцію до першої логарифмічної різниці рівнів. Наступним кроком ми «об'єднуємо» дані за відповідний квартал, беручи середні значення

$$x_t^Q = w(L^{\frac{1}{3}})x_t^W$$

де $w(L^{\frac{1}{3}})$ — функція агрегатора, а x_t^W — тижневий темп зростання для ОЕСР та місячний для PMI та ESI.

Остаточні регресії виглядають наступним чином:

$$y_t = \beta_0 + \beta(L^{\frac{1}{3}})x_t^W + \epsilon_t$$

Еталонні показники

Для отримання точки порівняння з використаною моделлю прогнозування, було зроблено порівняльний показник на основі минулих результатів зростання ВВП

$$GDP_t = \beta_0 + \beta_1 GDP_{t-1} + \beta_2 GDP_{t-2} + \epsilon_t$$

Емпірична модель

Аби перевірити, чи є корисним і точним щотижневий трекер ОЕСР щодо економічної активності, нам потрібно порівняти результати поточної прогнозу з результатами, отриманими за допомогою PMI та ESI, а також різні комбінації індексів, щоб перевірити, чи враховує індекс ОЕСР усі ефекти, представлені в опитуваннях, або його все ще можна покращити при наукастингу з використанням багатовимірних моделей.

$$GDP_t = \beta_0 + \beta_1 OECD_t + \epsilon_t$$

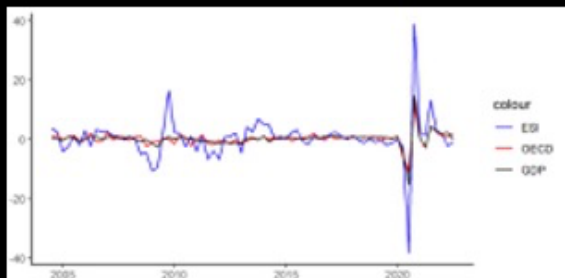
$$GDP_t = \beta_0 + \beta_1 ESI_t + \beta_2 PMI_t + \beta_3 OECD_t + \epsilon_t$$

Результати кореляції

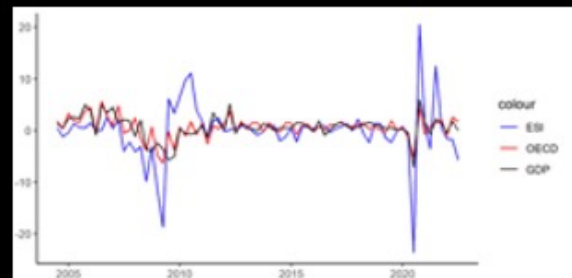
Країна	ОЕСР ВВП	PMI ВВП	ESI ВВП	PMI ОЕСР	ESI ОЕСР
Австрія	0.9198	0.7178	0.7762	0.6303	0.6843
Чехія	0.8499	0.5704	0.8079	0.5464	0.7257
Данія	0.7149	0.407	0.5934	0.2658	0.5753
Франція	0.9825	0.8199	0.8084	0.817	0.7882
Німеччина	0.9178	0.7476	0.8132	0.6332	0.757
Греція	0.7505	0.7131	0.614	0.8382	0.5573
Ірландія	0.4137	0.5765	0.3401	0.7882	0.6205
Італія	0.9556	0.7429	0.8234	0.735	0.7533
Нідерланди	0.9144	0.6867	0.8029	0.6604	0.8043
Польща	0.8038	0.7457	0.7356	0.7958	0.8272

Спостереження під час ГФК

Португалія



Латвія



До та після пандемії COVID19

Країна	До COVID			Після COVID		
	ОЕСР ВВП	ESI ВВП	ESI ОЕСР	ОЕСР ВВП	ESI ВВП	ESI ОЕСР
Австрія	0.7391	0.5857	0.5285	0.9626	0.8874	0.8133
Чехія	0.6905	0.6836	0.3584	0.9352	0.9235	0.9036
Данія	0.6504	-0.0309	0.5521	0.9128	0.905	0.8622
Франція	0.8848	0.6783	0.7497	0.9954	0.9236	0.8967
Німеччина	0.859	0.5566	0.3417	0.9482	0.9455	0.8892
Греція	0.777	0.8153	0.8333	0.8032	0.8062	0.6355
Ірландія	0.3019	0.6162	0.3123	0.7116	0.5995	0.8999
Італія	0.7821	0.5347	0.69	0.9714	0.921	0.8445
Нідерланди	0.8908	0.6288	0.8547	0.9704	0.9174	0.9135
Польща	0.4589	-0.1241	0.5064	0.8663	0.8769	0.9223

Порівняння з вихідним дослідженням

- Результати RMSE, отримані для 42 країн становлять у середньому 2,33 і 1,31 для Волошко і нашого аналізу відповідно
- Серед найбільш яскравих відмінностей між нашими роботами можна окреслити всі країни Балтії із середнім покращенням 2,85 у RMSE, а також Ірландію, Словаччину та Туреччину (2,2 покращення у RMSE)
- Ми обидва дійшли висновку, що індекс ОЕСР дає найбільшу похибку в оцінках для Ірландії, Аргентини та Ісландії

Висновки

- Високочастотні макроекономічні індекси є важливим інструментом для економістів і політиків для швидкого реагування на економічні зміни.
- Дослідження показує здатність трекера прогнозувати зростання ВВП, зокрема у періоди економічної невизначеності, і показує його перевагу в точності порівняно з іншими показниками.
- Використання трекера у поєднанні з іншими індикаторами може покращити точність прогнозів, враховуючи аспекти, не охоплені Google Trends.
- Трекер, розроблений Ніколасом Волошко, дозволяє аналізувати економічну ситуацію в реальному часі та з високою точністю прогнозувати зміни зростання ВВП, перевершуючи інші показники під час економічної нестабільності.

Дякую за увагу