

МАТЕМАТИЧНІ МЕТОДИ, МОДЕЛІ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В ЕКОНОМІЦІ

УДК 338.27::[519.816+303.094+004.658+004.031.4+003.63]

DOI: <https://doi.org/10.32840/1814-1161/2020-6-29>

Твердохліб І.П.

кандидат економічних наук, доцент,
доцент кафедри інформаційних систем у менеджменті
Львівського національного університету імені Івана Франка

Тимчишин М.А.

студент економічного факультету
Львівського національного університету імені Івана Франка

Tverdokhlib Ivan

Candidate of Economic Sciences, Associate Professor,
Associate Professor of Information Systems in Management Department
Ivan Franko National University of Lviv

Tymchyshyn M.A.

Student of the Faculty of Economics
Ivan Franko National University of Lviv

ВИМІРНЕ МОДЕЛЮВАННЯ СХОВИЩА ДАНИХ ОЦІНОК ТОЧНОСТІ ПРОГНОЗІВ ІНФОРМАЦІЙНОГО WEB-РЕСУРСУ ПІДТРИМКИ РІШЕНЬ У СФЕРІ ПРОГНОЗУВАННЯ

DIMENSIONAL MODELING OF THE DATA WAREHOUSE FOR ESTIMATING THE ACCURACY OF FORECASTS OF THE INFORMATION WEB-RESOURCE OF SUPPORT OF DECISIONS IN THE SCOPE OF FORECASTING

Розглядається проблема інформаційно-аналітичної підтримки рішень у сфері прогнозування. З метою накопичення наявної в інформаційних джерелах інформації про якість соціально-економічних прогнозів пропонується ідея створення спеціалізованого сховища даних з оцінками їх точності як ключової компоненти архітектури web-ресурсу підтримки рішень у сфері прогнозування. Оцінювання якості прогнозів здійснюється на основі відомої в прогнозуванні концепції класів точності. Узагальнено наявні моделі консолідації оцінок точності прогнозів та сформовано перелік типових інформаційно-аналітичних запитів до сховища даних web-ресурсу підтримки рішень у сфері прогнозування. Отримано вимірну модель сховища даних оцінок точності прогнозів як сукупність трьох схем типу «зірка»/«сніжинка», які сукупно уможливають реалізацію усіх типових запитів прогнозиста. Отримані схеми можуть бути використані для створення відповідних вітрин такого сховища даних, які забезпечують базові режими використання web-ресурсу підтримки рішень у сфері прогнозування.

Ключові слова: прогнозування, прогноз, точність прогнозів, підтримка рішень, сховище даних, вітрина даних, вимірне моделювання, схема «зірка», схема «сніжинка».

Рассматривается проблема информационно-аналитической поддержки решений в области прогнозирования. С целью накопления имеющейся в информационных источниках информации о качестве социально-экономических прогнозов предлагается идея создания специализированного хранилища данных с оценками их точности как ключевой компоненты архитектуры web-ресурса поддержки решений в области прогнозирования. Оценка качества прогнозов осуществляется на основе известной в прогнозировании

концепции классов точности. Обобщены имеющиеся модели консолидации оценок точности прогнозов и сформирован список типовых информационно-аналитических запросов к хранилищу данных web-ресурса поддержки решений в области прогнозирования. Получена многомерная модель хранилища данных оценок точности прогнозов как совокупность трех схем типа «звезда»/«снежинка», которые совместно обеспечивают реализацию всех типичных запросов прогнозиста. Полученные схемы могут быть использованы для создания соответствующих витрин такого хранилища данных, обеспечивающих базисные режимы использования web-ресурса поддержки решений в области прогнозирования.

Ключевые слова: прогнозирование, прогноз, точность прогнозов, поддержка решений, хранилище данных, витрина данных, размерное моделирование, схема «звезда», схема «снежинка».

The article considers the actual problem of information-analytical decisions support in the scope of forecasting. In order to accumulate the available information in information sources about quality of socio-economic forecasts, the idea of creating a specialized data warehouse with estimates of their accuracy as a key component of the decisions support web-resource architecture in the scope of forecasting is offered. Such a data warehouse will allow the use of information scattered in space and time about quality of committed forecasts in the process of new socio-economic forecasts reasoning. Estimation of forecasts quality is carried out on the basis of the known in forecasting concept of accuracy classes. The existing models of consolidation accuracy estimates of forecast are generalized and the list of standard information-analytical inquiries to the data warehouse of the decisions support web-resource in the scope of forecasting is formed. Based on a set of typical queries with using the dimensional modeling technique, the data space structure of the forecast quality assessment warehouse is outlined. The temporal character of the relations between the designed data warehouse indicators is clarified, the approach for time modeling in the accumulation process of forecasts accuracy estimations in such warehouse is offered. A dimensional model of the estimating the accuracy of forecasts data warehouse is obtained as a set of three schemes of the "star" / "snowflake" type, which together enabled the implementation of all typical queries of the forecaster. The obtained schemes can be used to create appropriate showcases of such data warehouse, which provide the basic modes of decisions support web-resource in the scope of forecasting use. The results of this scientific and practical study form the basis for the implementation of industrial OLAP-servers specialized warehouse of accuracy estimates of socio-economic forecasts and analysis the accumulated array of such estimates by Data Mining methods to obtain empirical relationships between the circumstances of forecasts realization. This kind of empirical dependencies will be the recommendations that will be provided to the forecaster by informational decisions support web-resource in the scope of forecasting.

Keywords: forecasting, forecast, accuracy of forecasts, decision support, data warehouse, data mart, dimensional modeling, star scheme, snowflake scheme.

Постановка проблеми. В усі часи людина прагнула забезпечити стабільність свого існування через передбачення майбутнього [1]. Прогнози [2] як засіб врахування майбутніх вірогідних тенденцій в управлінських рішеннях стали нормою у сфері управління [3; 4; 5]. Проте не завжди прогнози збігалися з дійсним перебігом процесів або розвитком подій. Тому зусилля значної плеяди фахівців зосереджувалися на аналізі причин недостовірності передбачень у різних сферах діяльності людини та обґрунтуванні рекомендацій щодо покращення ефективності прогнозування [3; 6–8]. Очевидно, що такого плану емпіричні рекомендації залежать від обсягу аналізованих масивів оцінок ефективності прогнозів та їх достовірності. До останнього часу науковці стараються удосконалити окремі аспекти методології прогнозування [9–12], що хоча і покращує вірогідність прогнозів в окремих випадках, проте кардинально не вирішує проблеми якості передбачень. У 70-і роки ХХ ст. була запропонована концепція класів точності прогнозів як практичної шкали оцінювання їх якості [3, с. 19] під час обґрунтування управлінських рішень. Аналіз стану ефективності соціально-економічних прогнозів у динаміці на підставі концепції класів точності з консолидацією оцінок та верифікацією причин неякісних передбачень в економіці і підприємстві може сформувати науково-обґрунтовану основу для прийняття раціональних рішень у сфері прогнозування. Неодмінною умовою створення такої наукової основи є потреба у спеці-

алізованих інформаційних web-ресурсах, призначенням яких має бути накопичення та консолидація оцінок якості прогнозів.

Аналіз останніх досліджень і публікацій. Ідея консолидації даних з оцінками якості соціально-економічних прогнозів на основі концепції класів точності [3, с. 19] була обґрунтована у працях [6; 13–15]. Засобом реалізації цієї ідеї у [16] визначено спеціальний інформаційний Web-ресурс підтримки рішень у сфері прогнозування, який трактується як «спеціалізоване сховище консолидованої інформації з наявними на даний момент часу оцінками якості передбачень, яке функціонує у мережі Інтернет і призначене для підтримки рішень прогнозистів» [17, с. 66]. Також у [16; 17] окреслено головні завдання такого ресурсу. Зокрема, це динамічне накопичення та консолидація оцінок точності прогнозів, пошук тенденцій і залежностей між обставинами здійснення прогнозів, візуалізація консолидованої інформації щодо точності прогнозів, тиражування актуальних тенденцій точності передбачення серед прогнозистів тощо. Загалом цей інформаційний Web-ресурс має інтегрувати усі наявні на даний час оцінки точності соціально-економічних прогнозів у пам'яті комп'ютерної системи та забезпечувати окреслені у моделях [6; 13–15] як алгоритми консолидації оцінок, так і форми подання результатів.

Аналіз праць [6; 13–17] засвідчує наявність багаторівневої формалізації предметної області проблеми ефективності прогнозування, а саме:

– концептуалізація процесу використання Web-ресурсу для підтримки рішень у сфері прогнозування у формі агентної інформаційної моделі [16; 17];

– формалізація процедур консолідації інформації щодо ефективності прогнозів у формі статистичних економіко-математичних моделей на основі концепції класів точності [13, 6];

– структуризація вхідної та результатної інформації для Web-ресурсу (зразки див. [6; 14; 15]). Оприлюднені у них емпіричні залежності ефективності прогнозів від обставин їх розроблення та здійснення, які були отримані на основі консолідації і аналізу масиву із близько 2000 зібраних з наукових джерел оцінок їхньої точності, засвідчують корисність такого роду нових знань у сфері прогнозування.

Виділення не вирішених раніше частин загальної проблеми. У праці [17] обґрунтовано формалізоване подання інформаційного web-ресурсу підтримки рішень у сфері прогнозування як динамічної системи виду $R(t) = \langle B(t), S(t), I(t), A(t) \rangle$, де t – момент часу (період), а $B(t), S(t), I(t), A(t)$ позначають компоненти web-ресурсу $R(t)$ у період t відповідно розподілену БД описів прогнозів, його сервіс-орієнтовану архітектуру, консолідовану на цей період інформацію щодо точності передбачень та множини агентів, які взаємодіють з інформаційним ресурсом. Окрім загальних вимог до структури вхідної і консолідованої інформації щодо якості прогнозів таке подання залишає невирішеними низку питань щодо ефективності реалізації такого web-ресурсу. Зокрема, існують сучасні ефективні технології опрацювання даних, що поєднують накопичення та агрегацію інформації у пам'яті комп'ютерних систем. З огляду на значну кількість розроблених фахівцями прогнозів та їх розпорошеність у часі і просторі, доцільність використання табличних БД для реалізації компоненти $B(t)$ виглядає сумнівною навіть з урахуванням існуючих апробацій [6; 14–16]. Натомість технологія сховищ даних [18] уможливить поєднання компонент $B(t)$ та $I(t)$ в одному спеціалізованому сховищі оцінок якості здійснених прогнозів і забезпечить ефективне накопичення розподіленої інформації про їх якість з подальшою консолідацією таких оцінок у потрібному прогнозістичному контексті. Але структура простору даних такого спеціалізованого сховища нині невідома і потребує прискіпливого обґрунтування.

Формулювання цілей статті (постановка завдання). Перед тим як сформулювати мету статті, розглянемо на основі [6; 15] типову математичну модель консолідації оцінок якості соціально-економічних прогнозів з використанням концепції класів точності [3; 6]. Аналогічно як у [6; 13] позначимо через E, S, P, M, R множини відповідно типів економічних систем, соціально-економічних процесів, економічних показників оцінювання стану процесів, методів прогнозування, моніторингових результатів прогнозування. Нехай для прогнозу $r \in R$ відомими є $x_{\text{espmrt}}, y_{\text{espmrt}}$, величини яких визначають відповідно прогнозне та фактичне значення p -го показника s -го процесу для e -ї економічної системи за період t при прогнозуванні m -м методом. Тоді можна обчислити відносну похибку прогнозу δ_{espmrt} (у %) за формулою [13]

$$\delta_{\text{espmrt}} = \left| \frac{y_{\text{espmrt}} - x_{\text{espmrt}}}{y_{\text{espmrt}}} \right| \times 100. \quad (1)$$

Якщо позначити через α_i, β_i відповідно ліву та праву межі точності (у %) для i -го класу точності [3, с. 19], то можна буде визначити кількість k_{iem} прогнозів i -го класу, ($i = \overline{1,5}$), для m -го методу в умовах e -ї економічної системи як

$$k_{\text{iem}} = \sum_{s \in S} \sum_{p \in P} \sum_{r \in R} \sum_{t \in T} B(i, \delta_{\text{espmrt}}), \quad (2)$$

де T – множина часових періодів з прогнозами, а булева функція $B(i, z)$ приймає такі значення для $i = \overline{1,5}$:

$$B(i, z) = \begin{cases} 1 & \text{при } z \in [\alpha_i; \beta_i]; \\ 0 & \text{інакше} \end{cases} \quad (3)$$

На підставі (2) можна розрахувати емпіричні вірогідності γ_{iev} (у%) попадання прогнозів для e -ї економічної системи з використанням m -го методу прогнозування у i -й клас точності та усереднений розподіл ξ_{ei} ймовірності точності прогнозування за класами в цілому для даного типу економічної системи за формулами

$$\gamma_{\text{iem}} = \frac{k_{\text{iem}}}{\sum_{i=1}^5 k_{\text{iem}}} \times 100; \quad \xi_{\text{ei}} = \sum_{m \in M} \omega_{\text{em}} \times \gamma_{\text{iem}},$$

$$\omega_{\text{em}} = \frac{\sum_{i=1}^5 k_{\text{iem}}}{\sum_{m \in M} \sum_{i=1}^5 k_{\text{iem}}}. \quad (4)$$

Аналогічно можна проаналізувати якість прогнозування у розрізі економічних процесів. Для цього використаємо такі співвідношення:

$$k_{\text{iesm}} = \sum_{p \in P} \sum_{r \in R} \sum_{t \in T} B(i, \delta_{\text{espmrt}}),$$

$$\gamma_{\text{iesm}} = \left[k_{\text{iesm}} I \left(\sum_{i=1}^5 k_{\text{iesm}} \right) \right] \times 100 (i = \overline{1,5}), \quad (5)$$

де $k_{\text{iesm}}, \gamma_{\text{iesm}}$ – відповідно кількість прогнозів та ймовірність попадання їх у i -й клас точності для m -го методу прогнозування, що застосовувався для передбачення розвитку s -го процесу в умовах e -ї економічної системи.

Отже, **метою статті** є висвітлення результатів дослідження структури простору даних спеціалізованого сховища оцінок точності прогнозів як ключового компонента інформаційного web-ресурсу підтримки рішень у сфері прогнозування [17]. Обов'язковими вимогами до простору даних такого сховища, які необхідно враховувати під час дослідження, має бути уможливлення реалізації існуючих моделей консолідації оцінок якості прогнозів із [13–15] на кшталт моделі (1)–(5), форм подання користувачам web-ресурсу результатної інформації із [6; 14; 15] та підтримка режимів функціонування web-ресурсу згідно агентної моделі [17, с. 68].

Виклад основного матеріалу. Для досягнення мети дослідження потрібно було виконати три завдання, а саме:

– узагальнити моделі консолідації оцінок точності соціально-економічних прогнозів та форм подання результуючої інформації із [6; 13–15];

– обґрунтувати вибір методології проектування сховища оцінок точності прогнозів з урахуванням його специфіки;

– розробити за допомогою вибраної методології модель простору даних такого сховища з урахуванням режимів використання web-ресурсу підтримки рішень у сфері прогнозування.

Нижче стисло описуються отримані результати виконання останніх 2-х завдань.

1. Вимірне моделювання як методика проектування СД оцінок точності прогнозів. На практиці використовують декілька підходів проектування схем СД. Найчастіше застосовують методологію Інмона (підхід на основі нормалізованих відношень), підхід Кімбала (методологія вимірного моделювання СД) та гібридний підхід (методологія Data Vault) [19, с. 227–252; 20]. Кожна з цих методологій має свої переваги та недоліки.

У процесі дослідження структури простору даних web-ресурсу підтримки рішень у сфері прогнозування було застосовано методологію вимірного моделювання [21; 22, с. 179–191; 23] для розроблення схеми СД оцінок точності прогнозів. Мотиви вибору підходу Кімбала [23] як основи методики проектування схеми нашого СД такі: 1) у цій методології відштовхуються від множини запитів, які має підтримувати проектуване СД; 2) результатом проектування буде схема СД у формі “зірки” або “сніжинки” [19, с. 472; 20, с. 28; 22, с. 179]; 3) так спроектоване СД забезпечує найкращу продуктивність виконання найважливіших для користувачів запитів за рахунок часткової денормалізації даних у ньому; 4) природня реалізація консолідації даних, що зберігаються у СД.

Методологія вимірного моделювання СД досить повно описана у багатьох джерелах, зокрема і у [22, с. 179–191]. Загалом процес проектування вимірної моделі СД оцінок точності прогнозів для інформаційного Web-ресурсу підтримки рішень у сфері прогнозування включає 10 операцій, згрупованих нами у 8 етапів. Схематично методика вимірного проектування схеми СД зображена на рис. 1 у форматі блок-схеми. Кожен етап методики на схемі показаний як ручна технологічна операція, що опрацьовує певну порцію вхідної проектної інформації та продукує нові знання у формі нової порції проектних даних. Графічно технологічні операції зображені символом ручної операції відомої графічної мови блок-схем. Зміст операції розкрито в її назві, причому назви операцій співпадають з назвами процедур описаної у [22, с. 179–191] методики вимірного моделювання СД. Напрямок передачі проектної інформації вказуються стрілками на лініях.

Структура порцій проектної інформації зазначено як коментарі у фігурних дужках, які окреслюють необхідні для операції змістовні компоненти. Отже, як випливає з рис. 1, методика на основі концептуалізації проблеми підтримки рішень у сфері прогнозування із [6; 13–15] забезпечить отримання вимірної моделі СД у формі «зірка» або «сніжинка» через поетапне уточнення та деталізацію вимог до нашого СД.

2. Проектування вимірної моделі СД оцінок точності прогнозів. З огляду на обмежений обсяг статті неможливо навіть стисло описати результати усіх проектних процедур. Тому зосередимося на ключових, які окреслюють процес розроблення вимірної моделі простору даних для web-ресурсу підтримки рішень у сфері прогнозування.

2.1. Формування переліку бізнес-запитів. Першою з таких важливих є 2-га технологічна операція схеми рис. 1, метою якої є визначення та опис типових запитів до СД оцінок точності прогнозів. Побудована множина таких запитів уможливілює формування у СД багатовимірного простору даних оцінок точності соціально-економічних прогнозів та напрямки (виміри) консолідації інформації. Сформований нами перелік типових запитів до СД оцінок точності прогнозів, згрупованих згідно рекомендацій [22, с. 183] із зазначенням формату подання результатної інформації, описано у табл. 1. Контент типових запитів до СД оцінок точності прогнозів окреслено текстом умови у 3-й графі цієї таблиці, у 2-й зазначається код запиту (його номер у групі), а в останній – очікуваний формат результату виконання запиту. Номер і назва групи запитів з урахуванням як рекомендацій [22, с. 183], так і специфіки проблеми вписано у 1-у колонку табл. 1.

Загалом було виділено 21 типовий запит, які згрупували у 4 групи і які сукупно охоплюють найважливіші аспекти використання проектного СД як компоненти web-ресурсу підтримки рішень у сфері. Перша група запитів призначена для контролю стану поповнення сховища даних інформацією про здійснені прогнози. Запити порівняльного аналізу уможливілюють оцінювання у динаміці розміру багатовимірного простору наявних у СД прогнозів у розрізі ознак (вимірів) консолідації. Третя група – запити аналізу тенденцій – призначена для візуального подання тенденцій росту обсягів інформації про здійснені прогнози у різних зрізах.

Найбільшою за кількістю та оцінюваними аспектами процесу прогнозування є четверта група типових запитів до СД. Їх ще додатково об'єднано у 3-и підгрупи. Підгрупа запитів кластеризації прогнозів забезпечує отримання прогнозистом певних агрегованих характеристик щодо множини наявних у СД прогнозів у розрізі класів точності. Підгрупа запитів аналізу відношень між оцінками точності прогнозів є однією із найвагоміших і реалізує процедури вибірки з СД консолідованої інформації щодо ефективності прогнозування у найзатребуваніших вимірах. Власне результати цих запитів формують основу для обґрунтування рішень у сфері прогнозування. Остання підгрупа четвертої групи типових запитів до проектного СД – запити ранжування авторів і джерел оприлюднення прогнозів – призначена для оцінювання якості діяльності прогнозистів.

2.2. Формування кінцевого набору фактів для СД. Цей етап дослідження поєднує операції № 3–№ 6 блок-схеми рис. 1. Згідно методології вимірного моделювання потрібно для отриманої на попередньому кроці множини типових запитів ідентифікувати необхідні для їх реалізації виміри і показники простору даних та оптимізувати їх структуру. Аналіз умов 21-го типового запиту до проектного СД (див. табл. 1) призвів до окреслення принаймні 15-и вимірного простору даних, у якому можлива їх реалізація. Саме такий набір вимірів забезпечить у СД потрібну консолідацію оцінок точності прогнозів згідно вимог моделей [13–15] та, зокрема, (1)–(5). Загальна ха-

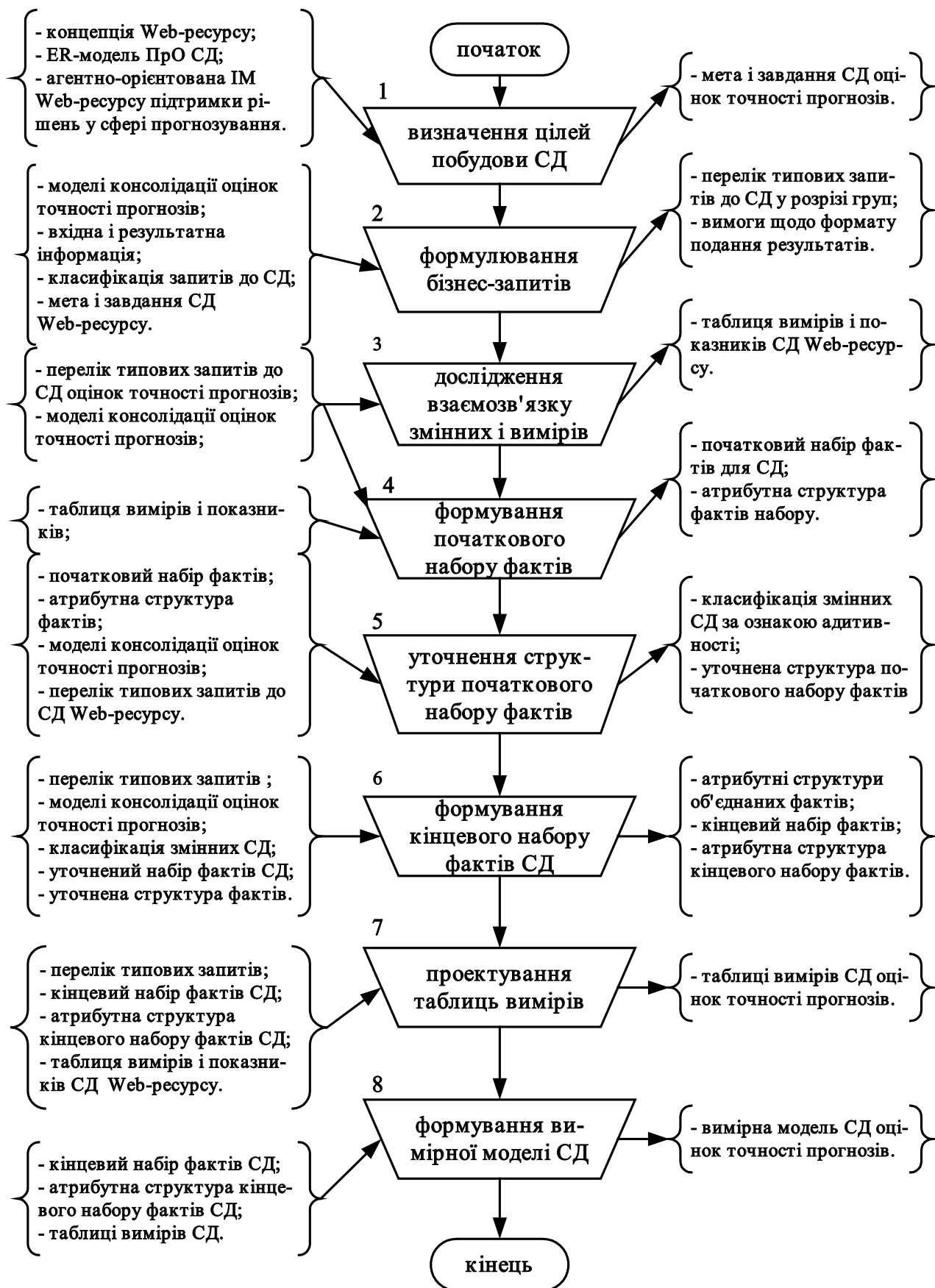


Рис. 1. Блок-схема методики вимірного проектування схеми СД оцінок точності соціально-економічних прогнозів

Джерело: розроблено на основі [22, с. 179–191]

Типові запити до СД оцінок точності прогнозів (фрагмент)

Група запитів	Характеристики запитів у розрізі груп		
	№ п/п	Умова запиту	Подання результату
1. Запити перевірки	1.1	Підтвердити наявність у СД вказаного прогнозу.	Текст
	1.4	Належність прогнозу до вказаного класу точності.	Текст
2. Запити порівняльного аналізу	2.2	Отримати приріст кількості наявних у СД прогнозів у розрізі базових ознак консолідації оцінок точності прогнозів.	Таблиця
	2.3	Візуально співставити кількості {наявних у СД прогнозів; статистичного забезпечення прогнозів; оцінок точності прогнозів} у поточному та базовому періодах.	Графік (стовпчикова діаграма)
4. Запити аналізу відношень, ранжування та кластеризації			
4.1. Запити кластеризації прогнозів.	4.1.4	Отримати емпіричну оцінку середньозваженої відносної похибки наявних у СД прогнозів для вказаного класу точності.	Текст
4.2. Запити аналізу відношень між оцінками точності прогнозів.	4.2.2	Візуалізувати розподіли ймовірностей попадання прогнозів у класи точності до зазначеного періоду для вказаної ознаки консолідації.	Діаграма [15]
	4.2.3	Отримати сукупність if-then-правил емпіричних залежностей між обставинами здійснення прогнозів зазначеного рівня достовірності для вказаного класу точності і типу економіки.	Таблиця [14]
	4.2.4	Візуалізувати значущість факторів прогнозування для зазначеного типу економіки.	Діаграма [15]
4.3. Запити ранжування авторів і джерел оприлюднення прогнозів.	4.3.3	Отримати емпіричну оцінку середньозваженої відносної похибки наявних у СД прогнозів вказаного автора/джерела.	Таблиця
	4.3.4	Ранжувати авторів/джерела прогнозів за величиною середньозваженої відносної похибки прогнозу.	Таблиця

Джерело: розроблено авторами

характеристика необхідних вимірів СД оцінок точності прогнозів наведена у табл. 2.

Необхідність введення двох осей часу (дійсного та модельованого) як вимірів простору даних проєктованого сховища зумовлена об'єктивно існуючим часовим розривом між датами оприлюднення прогнозів та моментами часу їх введення у СД. Вісь дійсного часу має відображати часові мітки здійснення прогнозів у фізичному світі. Натомість вісь модельованого часу характеризує часовий аспект появи і поширення інформації щодо прогнозів у пам'яті СД. Тобто має містити часові мітки реального часу про події внесення прогнозу та його характеристик у СД. Схема траєкторії опису прогнозу у СД у часі зображена на рис. 2. Горизонтальна вісь на рис. 2 позначає часові мітки (періоди) дійсного часу t_a і призначена для відображення моментів реального (фізичного) часу нашого світу. Вертикальна вісь t_m містить часові мітки модельованого часу, які є часовими моментами дійсного часу, коли були виконані певні процедури внесення даних щодо соціально-економічного прогнозу.

Час здійснення r -го прогнозу позначено на осі t_a поділкою $t_a(r)$ як момент фізичного часу, у який було здійснено прогнозування (збігається з часовим періодом показника прогнозування у r -му прогнозі). Часто автор прогнозу оприлюднює одною датою результати передбачення розвитку процесу на декілька періодів як вперед, так і назад. Такі часові мітки на осі дійсного часу позначено як $t(r,1), t(r,2), \dots$ для наступних періодів та як $t_a(r,-1), t_a(r,-2), \dots$ для ретроспективного прогнозування, а на осі модельо-

ваного часу відповідно як $t_m(r+1), t_m(r+2), \dots$ та як $t_m(r-1), t_m(r-2), \dots$. На осі модельованого часу відзначено моменти завершення процедур введення у СД базового опису r -го прогнозу (мітка $t_m(r,x)$), внесення фактичного значення показника для r -го прогнозу (мітка $t_m(r,y)$) та обчислення точності r -го прогнозу (мітка $t_m(r,\delta)$), де x, y, δ відповідно прогнозне і фактичне значення показника та відносна похибка прогнозу. Власне ці точки на вертикальній лінії до дійсної осі часу і формують траєкторію опису прогнозу у СД у часовому вимірі. Такі траєкторії створюються і для усіх інших передбачень, які мають бути збережені у СД оцінок точності прогнозів.

Щодо показників запитів, то було ідентифіковано 21-у змінну простору даних проєктованого сховища web-ресурсу підтримки рішень у сфері прогнозування. Для ілюстрації фрагмент табличного опису множини показників СД оцінок точності прогнозів із збереженням їх нумерації наведено у табл. 3. Звернемо увагу на глибшу деталізацію показників моделі (1)–(5) та описаних у [13–15].

Після ідентифікації вимірів і змінних проєктованого СД на основі аналізу взаємозв'язку змінних з вимірами сформували множину початкових фактів. Для цього змінні, що мали спільні набори вимірів, об'єднували в один факт. У результаті отримали 16 початкових фактів СД, які уможливають реалізацію усіх типових запитів прогнозиста.

Для уможливлення автоматичної консолідації оцінок точності прогнозів у СД потрібно було змінні фактів початкового набору дослідити на адитивність, а самі факти – на рівень їх деталізації та узгодженості за ча-

Характеристики вимірів інформаційного простору СД оцінок точності прогнозів web-ресурсу підтримки рішень у сфері прогнозування

№ п/п	Назва виміру	Позначення	Призначення виміру у СД	Примітка
1	Прогноз	r	Ідентифікує прогноз у СД	Номер/код прогнозу
2	Тип економіки	e	Вказує тип економіки, для процесу якої здійснений прогноз	{ринкова; перехідна; командно-адміністративна; ...}
3	Процес	s	Ідентифікує соціально-економічний процес	Номер/код процесу
4	Показник	p	Ідентифікує у СД показник виміру стану процесу	Номер/код показника
5	Метод прогнозування	m	Вказує на методи прогнозування, що використовувалися у передбаченнях	Номер/код методу
6	Напрямок прогнозування	h	Визначає тип прогнозування	{ретроспектива; перспектива}
7	Авторство	a	Окреслює у СД множину авторів прогнозів	Прізвище чи код автора
8	Джерело оприлюднення	b	Окреслює у СД сукупність джерел оприлюднення прогнозів	Назва/код джерела
9	Клас точності	i	Множина класів точності прогнозів	Номер класу {1–5}
10	Регіон процесу	j	Ідентифікує у СД регіони, для яких розробляли прогнози	Назва/код регіону
11	Регіон здійснення прогнозу	l	Визначає множину регіонів, у яких розробляли прогнози	Назва/код регіону
12	Статистичне джерело	g	Окреслює у СД множину статистичних джерел для верифікації прогнозів	Назва/код статджерела
13	Чинник процесу прогнозування	μ	Визначає сукупність базових обставин процесу прогнозування	Назва чинника
14	Час дійсний	t_{δ}	Вказує часові мітки дійсного часу для прогнозів	Тип даних часу
15	Час модельований	t_m	Вказує часові мітки модельованого часу для прогнозів	Тип даних часу

Джерело: розроблено авторами

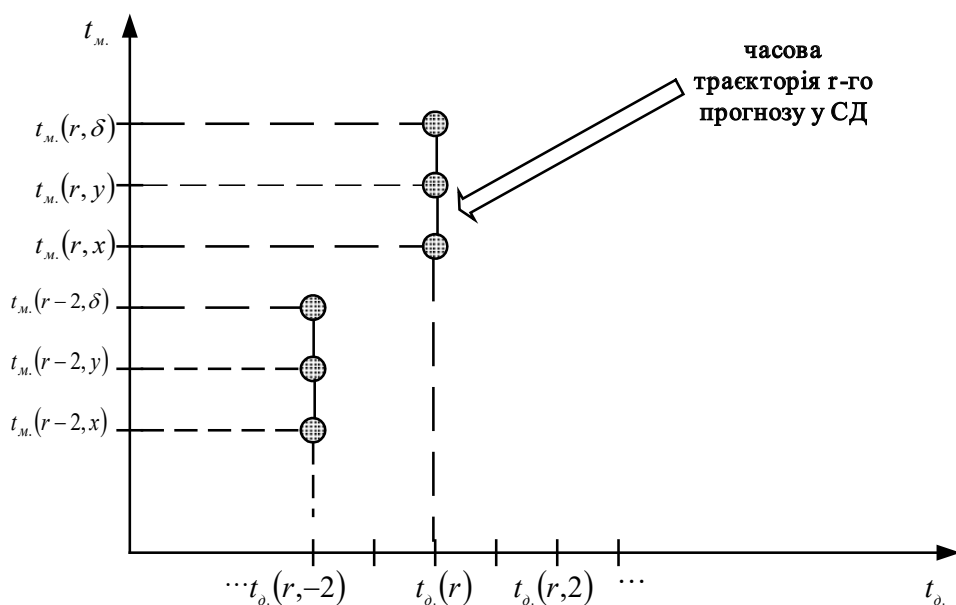


Рис. 2. Принципова схема формування опису прогнозу у СД у часовому вимірі

Джерело: розробка авторів

Характеристики змінних інформаційного простору СД оцінок точності прогнозів Web-ресурсу підтримки рішень у сфері прогнозування (фрагмент)

№ п/п	Змінна СД		Призначення змінної у СД
	назва	позначення	
3	Відносна похибка прогнозу	$\delta_{espmhabijlt_s,rt_m,g}$	Відносна похибка у % на основі g -го статистичного джерела прогнозу p -го показника s -го процесу j -го регіону для e -ї економічної системи при прогнозуванні m -м методом у r -му прогнозі, отриманого a -м автором для періоду t_s та внесеного у СД у період t_m .
4	Ознака належності до класу точності	B_{ir}	Ознака належності r -го прогнозу до i -го класу точності.
14	Дата прогнозу	D_r^1	Період часу дійсної часової осі, для якого був виконаний r -й прогноз.
18	Дата оцінки точності прогнозу	D_r^5	Період часу модельованої часової осі, коли була здійснена оцінка точності r -го прогнозу у СД.
20	Ймовірність попадання прогнозів у класи точності	$\gamma_{iespmat_m}$	Ймовірність у % попадання прогнозів методом m показника p оцінювання s -го процесу e -ї економіки, здійснених a -м автором, на основі множини прогнозів СД до періоду t_m .
21	Значущість чинників прогнозування	$\theta_{\mu t_m}$	Вага у % μ -го чинника процесу прогнозування у if-then-правилах на період t_m .

Джерело: розроблено авторами

совими вимірами. Зазвичай фахівці класифікують показники вимірної моделі даних у контексті можливості їх підсумовування на адитивні, напіваадитивні та неадитивні [19, с. 474; 22, с. 180]. Загалом змінні фактів початкового набору формують дві групи за ознакою адитивності (напіваадитивних не виявлено), а саме:

– адитивні, куди входять такі показники СД, як *кількість наявних прогнозів, кількість прогнозів з фактами, кількість прогнозів з оцінками, сумарна кількість прогнозів, кількість прогнозів з оцінками, відносна похибка прогнозу*;

– неадитивні, множину яких складають такі змінні, як *прогнозне значення, дата прогнозу, дата здійснення прогнозу, дата внесення прогнозу, фактичне значення, відносна похибка прогнозу, дата внесення статданих, дата оцінки прогнозу, ознака належності класу, ліва межа класу, права межа класу, ознака належності класу, частка прогнозів, ймовірність прогнозування класу точності, ймовірність попадання у клас точності, вага чинника прогнозування, середня відносна похибка, середньозважена відносна похибка, середньозважена відносна похибка*.

Оскільки більшість неадитивних змінних початкового набору фактів може бути обчислена на основі значень інших уже адитивних показників (див. форми (4), (5)), то їх замінюємо у фактах на адитивні. Окрім того, для усунення неузгодженості початкових

фактів за рівнем деталізації приймаємо рішення, що найменшою часовою міткою для агрегації оцінок точності прогнозів у часовому вимірі буде день зі стандартною часовою ієрархією «день – місяць – рік». З метою уможливлення підтримки типових запитів з емпіричною оцінкою значущості чинників процесу програмування у Факт_12 початкового набору потрібно додатково увести такі виміри як «множина правил», «правило», «клас точності».

Завершує цей етап процедура формування кінцевого набору фактів для проектного СД. З цією метою уточнені факти початкового набору інтегруються. У результаті отримали такий набір фактів для СД оцінок точності прогнозів (див. рис. 3, підготовлено за допомогою MS Visio).

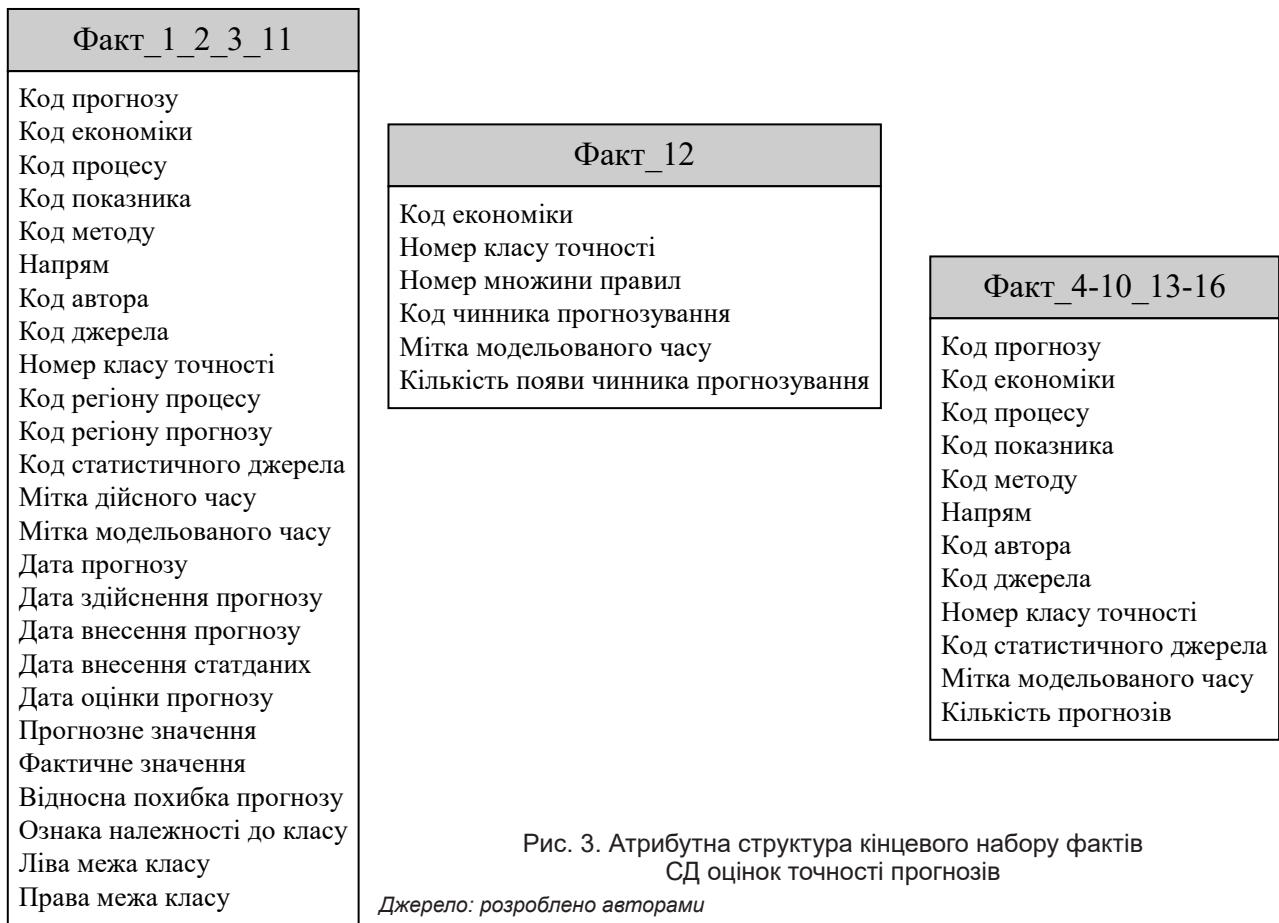
2.3. *Формування вимірної моделі СД оцінок точності прогнозів*. Включає останні дві операції блок-схеми рис. 1. Метою є побудова схеми СД на основі отриманого кінцевого набору фактів (див. рис. 3). Для цього спочатку розробляємо консольні таблиці вимірів простору даних, а потім поєднуємо їх з фактами на рис. 3. Кожне таке поєднання окреслює відповідний куб даних сховища оцінок точності прогнозів. Загалом було отримано три куби даних, які сукупно забезпечують реалізацію усіх типових запитів до web-ресурсу підтримки рішень у сфері прогнозування і які охарактеризовані у табл. 4.

Таблиця 4

Загальна характеристика вимірних моделей кубів СД оцінок точності прогнозів

Куб даних		Кількість вимірів	Ім'я факту набору	Тип вимірної моделі куба
ім'я	призначення куба			
Stat_progn	Містить оцінки точності здійснених прогнозів	14	Факт_1_2_3_11	“зірка”
Kons_progn	Містить агреговані дані щодо оцінок точності прогнозів	12	Факт_4-10_13-16	“сніжинка”
Scen_progn	Містить ідентифіковані типові сценарії прогнозування у формі асоціативних правил	6	Факт_12	“сніжинка” (див. рис. 4)

Джерело: розроблено авторами



Висновки з цього дослідження і перспективи подальших розвідок у цьому напрямі. Здійснене авторами дослідження структури простору даних для web-ресурсу підтримки рішень у сфері прогнозування засвідчує можливість його реалізації за допомогою технології сховищ даних. Вимірне моделювання предметної області СД оцінок точності прогнозів показало, що забезпечити базові режими використання такого web-ресурсу можливо сховищем з трьома вітринами даних, кожна з яких формує свій куб даних. Отримані вимірні моделі кубів даних створюють основу для реалізації СД оцінок точності прогнозів у середовищі OLAP-сервера.

Встановлено наявність темпоральних відношень між показниками простору даних web-ресурсу підтримки рішень у сфері прогнозування, що потрібно враховувати під час реалізації кубів даних засобами OLAP-сервера. Отримані схеми кубів даних мають бути ще доповнені процедурами агрегації оцінок точності прогнозів вздовж вимірів під час реалізації у середовищі OLAP-сервера. Зазначені завдання визначають напрям подальшого розвитку цього дослідження.

Бібліографічний список:

1. Д. Эндриус, Д. Франклин. «Мир в 2050 году». Москва : Эксмо, 2012. 219 с.
2. Прогноз. *Википедия: свободная энциклопедия*. URL: <https://ru.wikipedia.org/wiki/Прогноз> (дата звернення: 20.12.2020)
3. Пятакова Ф.С. Региональное прогнозирование потребления товаров и услуг. Киев : Техника, 1991. 215 с.
4. Кологривов Я.І. Побудова сценаріїв розвитку світової економіки до 2030 р. у контексті великих економічних циклів Кондратьєва. *Системні дослідження та інформаційні технології*. 2012. № 2. С. 125–137.
5. Щербіна В.В. Модель підвищення ефективності макроекономічного прогнозування та планування доходів державного бюджету України. *Академічний огляд*. 2017. № 1. С. 44–53.
6. Твердохліб І., Петрик Г. Ефективність економіко-математичного моделювання процесів трансформаційної економіки: прагматичний аспект. *Вісник Львівського університету. Серія економічна*. 2005. Вип. 34. С. 227–241.
7. Ейнес Ю. Рекомендаційні техніки для аналізу криміналістичних даних. *Електроніка та інформаційні технології*. 2018. Вип. 10. С. 64–70.
8. Белз О. Усунення нестационарності часових рядів під час статистичного моделювання соціально-економічних процесів. *Актуальні проблеми економіки*. 2013. № 11. С. 163–171.
9. Майер-Шенбергер В., Кукьер К. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим / пер. с англ. И. Гайдюк. Москва : Манн, Иванов и Фербер, 2014. 240 с.
10. Pavlyshenko V.M. Forecasting of events by tweets data mining. *Electronics and information technologies*. 2018. Issue 10. P. 71–85.
11. Boldyreva A., Alexandrov M., Koshulko O., Sobolevskiy O. Queries to Internet as a tool for analysis of regional police work and forecast of crimes in regions. Proc. of 15th Mexican Intern. Conf. on Artificial Intelligence, Springer, LNAI. Vol. 10061. 2016. Pp. 291–302.
12. Trokhimchuck P.P. Some Problems of Polymetric Modeling in the Econometrics. *International Journal of Engineering Research And Management (IJERM)*. 2020. Vol. 07. Issue 06. P. 43–52. URL: <http://www.ijerm.com/> (дата звернення: 20.12.2020)
13. Твердохліб І. Статистична модель узагальнення оцінок точності прогнозів соціально-економічних процесів. *Вісник Львівського університету. Серія економічна*. 2006. Вип. 35. С. 547–554.
14. Твердохліб І.П. Технологія data mining як інструментальний засіб удосконалення методології прогнозування соціально-економічних процесів. *Актуальні проблеми економіки*. 2008. № 12(90). С. 247–258.
15. Твердохліб І.П. Економетрія точності прогнозів соціально-економічних процесів трансформаційної економіки. *Zastowanie metod ilosciowych w ekonomii i zarzadzaniu dowiadczania polskie i ukrainskie*. Pod redakcją naukowa Bohdana Kopytko. Katowice, 2009. С. 79–96.
16. Твердохліб І. Концепція інформаційного Web-ресурсу підтримки рішень у сфері прогнозування. *Інформація, комунікація, суспільство 2014* : матеріали 3-ї міжнар. наук. конф. ICS-2014. Львів : Видавництво Львівської політехніки, 2014. С. 28–29.
17. Твердохліб І.П. Сервіс-орієнтована архітектура інформаційного web-ресурсу підтримки рішень у сфері прогнозування. «Сучасні проблеми інформатики в управлінні, економіці та освіті» : [матеріали XIII міжнар. наук. семінару] / за наук. ред. д.е.н., проф. М.М. Єрмошенка. Київ : Національна академія управління, 2014. С. 65–70.
18. Сховище даних. *Вікіпедія: вільна енциклопедія*. URL: [https://uk.wikipedia.org/wiki/Сховище_даних#:~:text=Сховище_даних_\(дата_звернення:_20.12.2020\)](https://uk.wikipedia.org/wiki/Сховище_даних#:~:text=Сховище_даних_(дата_звернення:_20.12.2020))
19. Пасічник В.В., Шаховська Н.Б. Сховища даних : Навчальний посібник. Львів : «Магнолія 2006», 2008. 496 с.
20. Джабраїлов Ш.В., Орлова Ю.А., Камбарова Ш.Ш. Сравнительный анализ методологий проектирования хранилищ данных. *Молодой ученый. Международный научный журнал*. 2020. № 20(310). С. 27–30.
21. Яцишин А.Ю. Побудова мультибазових сховищ даних на основі структурованості даних та запитів. *Восточно-Европейский журнал передовых технологий*. 2015. № 1/2 (73). С. 11–17.
22. Постіл С.Д. CASE-технології. Міждисциплінарне інформаційне моделювання : навчальний посібник. Ірпінь : Університет ДФС України, 2018. 304 с.
23. Черняк Л. Взгляд Ральфа Кимбалла на хранилища данных. *Открытые системы. СУБД*. 2007. № 5. URL: <https://www.osp.ru/os/2007/05/4265198/> (дата звернення: 20.12.2020)

References:

1. D. Endryus, D. Franklin (2012). Megachange: The World in 2050. Moscow: Eksmo. (in Russian)
2. Prognost. *Wikipedia: the free encyclopedia*. Available at: <https://ru.wikipedia.org/wiki/Прогноз> (accessed 20 December 2020).
3. Pyatkova F.S. (1991) *Regional'noe prognozirovanie potrebleniya tovarov i uslug* [Regional forecasting of consumption of goods and services]. Kiev: Tekhnika. (in Russian)
4. Kologhryvov Ja.I. (2012) Pobudova scenarijiv rozvytku svitovoi ekonomiky do 2030 r. u konteksti velykykh ekonomichnykh cyklyv Kondratjjeva [Construction of scenarios for the development of the world economy until 2030 in the context of large economic cycles Kondratiev]. *System Research & Information Technologies*, no. 2, pp. 125–137. (in Ukrainian)
5. Shherbina V.V. (2017) Modelj pidvyshhennja efektyvnosti makroekonomichnogo proghnozuvannja ta planuvannja dokhodiv derzhavnogo bjudzhetu Ukrajinj [Model of increase of efficiency of macroeconomic forecasting and planning of incomes of the state budget of Ukraine]. *Academic review*, no. 1, pp. 44–53. (in Ukrainian)
6. Tverdokhlib I., Petryk Gh. (2005) Efektyvnistj ekonomiko-matematichnogo modeljuvannja procesiv transformacijnoji ekonomiky: pragmatychnyj aspekt [The

- effectiveness of economic and mathematical modeling of transformational economics: a pragmatic aspect]. *Visnyk of the Lviv University. Series Economics*, issue 34, pp. 227–241. (in Ukrainian)
7. Ejnes Ju. (2018) Rekomendacijni tehniky dlja analizu kriminalistychnykh danykh [Recommended techniques for the analysis of forensic data]. *Electronics and information technologies*, issue 10, pp. 64–70. (in Ukrainian)
 8. Belz O. (2013) Usunennja nestacionarnosti chasovykh rjadiv pid chas statystychnogho modeljuvannja socialjno-ekonomichnykh procesiv [Elimination of nonstationarity of time series during statistical modeling of socio-economic processes]. *Actual Problems of Economics*, no. 11, pp. 163–171. (in Ukrainian)
 9. Viktor Mayer-Schönberger, Kenneth Cukier (2014) BIG DATA: A Revolution That Will Transform How We Live, Work, and Think. Moscow: Mann, Ivanov i Ferber. (in Russian)
 10. Pavlyshenko B.M. (2018) Forecasting of events by tweets data mining. *Electronics and information technologies*, issue 10, pp. 71–85.
 11. Boldyreva A., Alexandrov M., Koshulko O., Sobolevskiy O. (2016) Queries to Internet as a tool for analysis of regional police work and forecast of crimes in regions. Proc. of *15th Mexican Intern. Conf. on Artificial Intelligence*, Springer, LNAI. Vol. 10061, pp. 291–302.
 12. Trokhimchuck P.P. (2020) Some Problems of Polymetric Modeling in the Econometrics. *International Journal of Engineering Research And Management (IJERM)*, vol. 07, issue 06, pp. 43–52. Available at: <http://www.ijerm.com/> (accessed 20 December 2020).
 13. Tverdokhlib I. (2006) Statystychna modelj uzagalnennja ocinok tochnosti proghnoziv socialjno-ekonomichnykh procesiv [Statistical model of generalization of estimates of accuracy of forecasts of social and economic processes]. *Visnyk of the Lviv University. Series Economics*, issue 35, pp. 547–554. (in Ukrainian)
 14. Tverdokhlib I.P. (2008) Tekhnologhija data mining jak instrumentalnyj zasib udoskonalennja metodologhiji proghnozuvannja socialjno-ekonomichnykh procesiv [Data mining technology as a tool for improving the methodology of forecasting socio-economic processes]. *Actual Problems of Economics*, no. 12(90), pp. 247–258. (in Ukrainian)
 15. Tverdokhlib I.P. (2009) Ekonometrija tochnosti proghnoziv socialjno-ekonomichnykh procesiv transformacijnoji ekonomiky [Econometrics of accuracy of forecasts of social and economic processes of transformational economy]. *Zastosowanie metod ilosciowych w ekonomii i zarzadzaniu dowiadczania polskie i ukrajskie (Poland, Katowice, 2009)* (eds. B. Kopytko), Katowice, pp. 79–96. (in Ukrainian)
 16. Tverdokhlib I. (2014) Koncepcija informacijnogho Web-resursu pidtrymky rishenj u sferi proghnozuvannja [The concept of information Web-resource to support solutions in the field of forecasting]. Proceedings of the *3rd International Academic Conference «Information, Communication, Society 2014» (Ukraine, Lviv – Slavske, May 21-24, 2014)*, Lviv: Lviv Polytechnic National University, pp. 28–29. (in Ukrainian)
 17. Tverdokhlib I.P. (2014) Servis-orijentovana arkhitektura informatsiinogho web-resursu pidtrymky rishenj u sferi proghnozuvannja [Service-oriented architecture of information web-resource to support solutions in the scope of forecasting]. Proceedings of the *XIII International Scientific Seminar «Modern problems of computer science in management, economics and education» (Ukraine, Kyiv – Svietyaz, June 30 – July 4, 2014)* (eds. prof. M.M. Yermoshenko), Kyiv: National Academy of Management, pp. 65–70. (in Ukrainian)
 18. Skhovyshche danykh [Data Warehouse]. *Vikipediia: vilna entsyklopediia*. Available at: https://uk.wikipedia.org/wiki/Сховище_даних#:~:text=Сховище_даних (accessed 20 December 2020).
 19. Pasichnyk V.V., Shakhovska N.B. (2008) Skhovyshcha danykh: Navchalnyi posibnyk [Data Warehouses: Tutorial]. Lviv: "Mahnoliia 2006". (in Ukrainian)
 20. Dzhabrailov Sh.V., Orlova Yu.A., Kambarova Sh.Sh. (2020) Sravnitel'nyy analiz metodologiy proektirovaniya khranilishch danykh [Comparative analysis of data warehouse design methodologies]. *Young scientist. International scientific journal*, no. 20 (310), pp. 27–30. (in Russian)
 21. Yatsyshyn A.Yu. (2015) Pobudova multybazovykh skhovyshch danykh na osnovi strukturovanosti danykh ta zapytiv [Construction of multibase data warehouses based on structured data and queries]. *Eastern European Journal of Advanced Technologies*, no. 1/2 (73), pp. 11–17. (in Ukrainian)
 22. Postil S.D. (2018) CASE-tekhnologhii. Mizhdystsyplinarne informatsiine modeliuvannja: navchalnyi posibnyk [CASE-technologies. Interdisciplinary information modeling: a textbook]. Irpin: Universytet DFS Ukrainy. (in Ukrainian)
 23. Chernyak L. (2007) Vzgljad Ral'fa Kimballa na khranilishcha danykh [Ralph Kimball's perspective on data warehouses]. *Open systems. DBMS*, no. 5. Available at: <https://www.osp.ru/os/2007/05/4265198/> (accessed 20 December 2020).