

УДК 681.513.25 + 681.5.015

## ІНДУКТИВНЕ МОДЕЛЮВАННЯ ЯК ЗАСІБ РОЗВ'ЯЗАННЯ ЗАДАЧ ПРОГНОЗНОЇ АНАЛІТИКИ

С.М.Єфіменко

*Міжнародний науково-навчальний центр інформаційних технологій  
та систем (МННЦ ІТС) НАН та МОН України,  
[syefim@ukr.net](mailto:syefim@ukr.net)*

В роботі проаналізовано підходи до прогнозування економічних процесів у системах бізнес-інтелекту. Розглянуто сучасні інструменти прогнозування аналітики, що використовуються для прийняття ефективних бізнес-рішень.

*Ключові слова:* бізнес-інтелект, прогнозна аналітика, індуктивне моделювання, МГУА.

The paper analyzes approaches to prediction of economic processes in business intelligence systems. Contemporary tools of predictive analytics, used for effective business decision making, are considered.

*Keywords:* business intelligence, predictive analytics, inductive modeling, GMDH.

В работе проанализированы подходы к прогнозированию экономических процессов в системах бизнес-интеллекта. Рассмотрены современные инструменты прогнозной аналитики, используемые для принятия эффективных бизнес-решений.

*Ключевые слова:* бизнес-интелект, прогнозная аналитика, индуктивное моделирование, МГУА.

### 1. Вступ

Досягнення успіху та забезпечення конкурентоздатності в сучасних швидкозмінних економічних умовах неможливі без використання достовірної та оперативної інформації. Бізнес-дані стають важливим ресурсом для отримання знань і прийняття важливих управлінських рішень в різних бізнес-областях. Сучасні ефективні рішення потребують істинної та повної інформації, що неможливо зробити засобами традиційних інформаційних систем.

В наш час відбуваються стрімкі перетворення глобального інформаційного простору, що зачіпає суспільство, ринок та бізнес. Спостерігається швидкий розвиток цифрової економіки.

За прогнозом Microsoft у 2020-му році 25% всієї світової економіки стане електронною [1], тоді як у 2005-му році ця цифра складала 15%. Цифровізації економіки сприяють Інтернет речей (Internet of Things, ІОТ) та великі дані (Big Data), мобільні та хмарні технології. Вплив цих технологій на бізнес призведе до того, що безпосереднє володіння фізичними ресурсами стане непотрібним.

Міністерство економічного розвитку України розробило концепцію розвитку цифрової економіки і суспільства України на 2018-2020 роки. Після створення необхідного законодавчого поля більше 80% бізнес-процесів перейдуть в цифровий формат. Завдяки цьому частка високотехнологічного

експорту у випуску промислової продукції буде зростати, з'явиться більше винаходів.

## 2. Бізнес-інтелект і бізнес-аналітика

Бізнес-інтелект охоплює (БІ) стратегії та технології, що використовуються підприємствами для аналізу бізнес-інформації [2]. БІ відноситься до управлінської філософії та інструментарію, який використовується, щоб допомогти організаціям оперувати бізнес-інформацією з метою прийняття ефективних бізнес-рішень. Технології БІ забезпечують історичні, поточні та прогнозні погляди на бізнес-операції.

Класифікацію технологій, які використовуються для проведення бізнес-аналітики, подано в [3]:

Таблиця 1. Сучасні технології БІ

Технологія	Опис
Прогнозне моделювання	Прогнозування значень атрибутів
Описовий інтелектуальний аналіз даних	Розподіл, дисперсія та виключення даних
Аналіз зв'язків	Визначення взаємозв'язків між атрибутами
Класифікація	Визначення класу, до якого належать дані
Кластеризація та аналіз викидів	Розбиття множини даних на класи
Аналіз ситуацій	Тренди та відхилення, послідовні образи, періодичність
OLAP (інтерактивна аналітична обробка даних)	Засоби OLAP дозволяють аналізувати багатовимірні дані
Візуалізація моделей	Використання графіків, даграм, гістограм для полегшення аналізу даних
Розвідувальний аналіз даних (EDA)	Аналіз основних властивостей даних, знаходження загальних закономірностей

Економічний ефект від використання засобів БІ комерційними організаціями може бути різним [4]:

- працівники можуть легко застосувати свої ділові вміння для вирішення багатьох бізнес-питань, зокрема для підвищення швидкості реагування на маркетингові кампанії, що проводяться за допомогою телефону, електронної пошти чи Інтернет;

- фірми можуть ідентифікувати своїх найприбутковіших клієнтів та основні причини їх лояльності, а також визначати потенціальних прибуткових клієнтів;

- можливість аналізувати дані щодо історії відвідування сайтів задля покращення стратегій електронної комерції;

- визначення кримінальної діяльності, пов'язаної із відмиванням коштів нелегального походження;
- аналіз рентабельності потенційних клієнтів та зменшення ризику завдяки більш точній фінансовій оцінці своїх клієнтів
- визначення комбінацій продуктів та сервісів, які покупці найімовірніше придбають;
- аналіз клінічних випробувань експериментальних медичних препаратів;
- визначення найвигідніших ставок для страхових внесків;
- скорочення часу простою обладнання шляхом застосування технічного обслуговування, що ґрунтується на прогнозуванні запасу надійності;
- визначення причин переходу власних клієнтів до конкуруючих організацій.

Організації різного профілю та розмірів можуть зіткнутися з певними проблемами в ефективності використання існуючих даних в своїх системах.

У зв'язку з цим для компаній, що працюють в умовах цифрової економіки, особливого значення набуває якість і швидкість інформаційно-аналітичної підтримки бізнес-управління. Сучасні компанії для задач планування, аналізу і контролю часто використовують аналітичні додатки БІ на основі OLAP-систем.

Проте в нових економічних умовах класичного функціоналу таких систем вже недостатньо для розв'язання нових цифрових задач, оскільки ці системи орієнтовані на ретроспективний аналіз і обробляють інформацію про минулу ситуацію.

Отже, виникає необхідність застосування прогнозної аналітики (predictive analytics), яка доповнює і підсилює можливості БІ в частині прогнозування майбутніх подій.

### **3. Прогнозна аналітика**

Загалом, існує декілька три види аналітики, що співіснують та доповнюють одна одну [5]:

– описова аналітика (descriptive analytics) досліджує минулі факти для того, щоб знайти причини попередніх успіхів або невдач. Вона відповідає на питання «Що сталося?» Значною мірою описова аналітика все ще використовується в наш час. Більшість управлінських звітів щодо продажів, маркетингу, фінансів використовують цей вид бізнес аналітики;

– прогнозна аналітика (predictive analytics) відповідає на питання «Що ймовірно станеться?». Для цього історичні дані поєднуються з правилами, алгоритмами та зовнішніми даними для того, щоб визначити майбутнє значення або ймовірність виникнення події;

– рекомендаційна аналітика (prescriptive analytics) є наступним етапом за прогнозуванням майбутніх подій, і пропонує послідовність дій для отримання

найбільшої користі з прогнозів та показує наслідок кожного прийнятого рішення. Вона відповідає на питання «Що слід робити?»

Прогнозна аналітика визначається в [6] як множина різноманітних статистичних технологій (зокрема, прогнозного моделювання, машинного навчання та інтелектуального аналізу даних), які використовуються для аналізу поточних та минулих фактів з метою прогнозування майбутніх невідомих подій. В процесі аналізу використовуються, як правило, великі масиви даних. Головною ідеєю прогнозової аналітики є завдання визначення одного або декількох параметрів, які впливають на прогнозовану подію.

Процес прогнозного аналізу можна представити у такому вигляді:

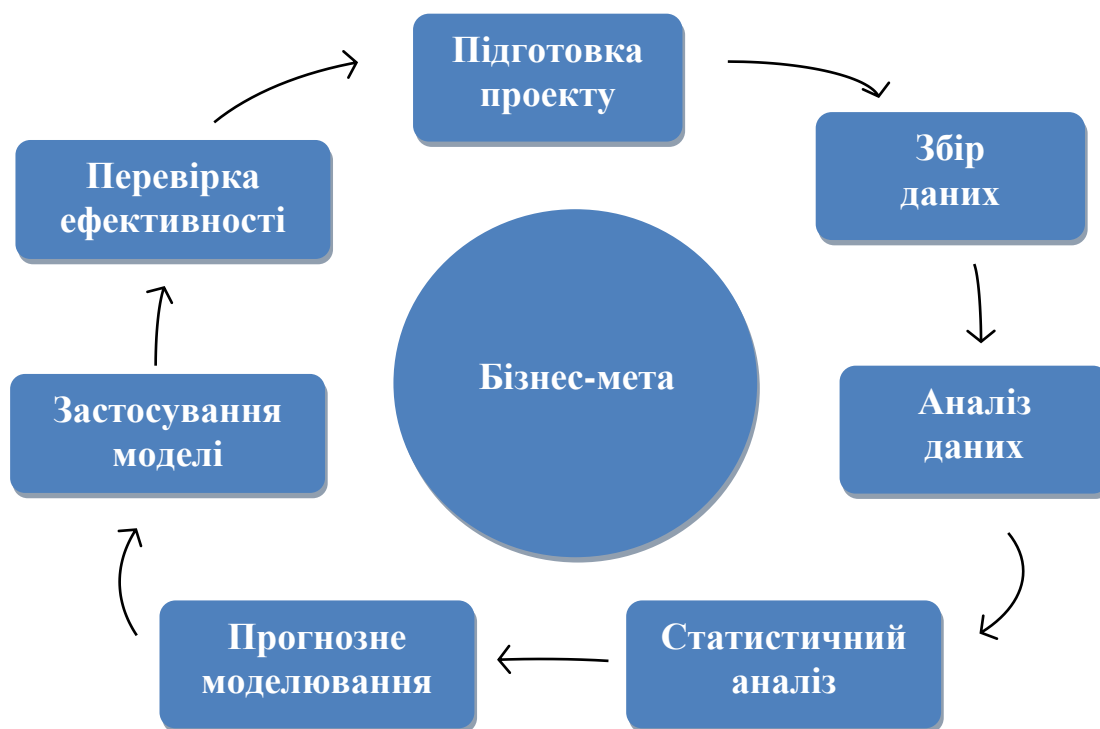


Рисунок 1 – Процес прогнозного аналізу

*Підготовка проекту.* Визначення результатів проекту, складових, масштабів роботи, бізнес-цілі, набору даних, що будуть використані.

*Збір даних.* За допомогою інтелектуального аналізу даних готуються дані з різних джерел. Відбувається це за безпосередньої участі клієнтів.

*Аналіз даних.* Виконання процесу перевірки, очищення та моделювання даних з метою виявлення корисної інформації.

*Статистичний аналіз* дозволяє підтверджувати припущення, гіпотези за допомогою стандартних статистичних моделей.

*Прогнозне моделювання* забезпечує можливість автоматично будувати точні прогнозні моделі.

*Застосування* прогнозної моделі забезпечує можливість використання аналітичних результатів у процесі прийняття рішень для отримання результатів та звітів.

*Перевірка ефективності.* Моделі перевіряються з метою забезпечення очікуваних результатів.

Підсумком застосування прогнозної аналітики є прийняття максимально ефективних для бізнесу рішень. Важливою вимогою до прогнозної моделі є здатність працювати максимально точно та бути статистично значущою.

До прогнозних моделей відносяться [7]:

– моделі класифікації. Вони описують правила або набір правил, відповідно до яких можна віднести новий об'єкт до одного з класів. Такі правила будуються на підставі інформації про існуючі об'єкти шляхом розбиття їх на класи;

– моделі часових рядів. Вони описують функції, що дозволяють прогнозувати зміну неперервних числових параметрів та будуються на підставі інформації про зміну деякого параметра за минулий період часу.

Для підприємств, де необхідна обробка величезної кількості даних і є високі ризики при прийнятті рішень, прогнозна аналітика має особливе значення. Інформація про технологічний процес не завжди використовується ефективно. Оптимізацію можна виконати на виробництві з високим рівнем автоматизації, організованим збором і тривалим зберіганням інформації.

Для цього використовуються інтелектуальні системи які можуть проаналізувати стан технологічного процесу в реальному часі, спрогнозувати подальше протікання процесу, визначити рівень оптимальності і, при необхідності, змінити управляючі параметри, або дати рекомендації диспетчерові. Для розв'язання таких задач за допомогою засобів машинного навчання створюється прогнозна математична модель технологічного процесу. Вона аналізує вхідні параметри, в реальному часі видає прогноз протікання процесу і пропозиції щодо його оптимізації.

За даними Transparency Market Research [8], до 2019 року ринок прогнозної аналітики досягне 6,5 мільярдів доларів, тоді як за підсумками 2015 року він складав усього 3,6 мільярда. Світовий ринок систем для прогнозного аналізу буде рости в середньому на 17,8% щорічно. Як показує практика, в складні економічні періоди виживають ті компанії, які продовжують інвестувати в технології і інновації. І прогнозна аналітика, безумовно, є однією з них.

#### **4. Програмні засоби прогнозної аналітики**

У 2013-му році компанія Forrester Research опублікувала звіт «Big Data Predictive Analytics Solutions, Q1 2013», у якому назвала лідерів на ринку засобів прогнозної аналітики [9]. Згідно зі звітом найміцнішими позиціями на ринку і найкращими стратегіями серед найбільших розробників засобів

прогнозної аналітики володіють компанії SAS та IBM SPSS. Оцінка проводилася за 51 параметром – від повноти функціональності основної аналітичної системи до розміру клієнтської бази і переваг архітектури пропонувані розробниками рішень.

**SAS (Statistical Analysis System) Enterprise Miner** [10] лідирує в сегменті засобів поглибленої аналітики, займаючи близько третини ринку. Він дозволяє користувачам досліджувати і аналізувати великі обсяги даних, знаходити стійкі закономірності і, ґрунтуючись на фактах та отриманих висновках, приймати виважені рішення. Галузі ефективного застосування: банківський сектор, охорона здоров'я, нафтогазовий сектор, страхові компанії, телекомунікації, транспорт, енергетика.

До головних переваг SAS Enterprise Miner можна віднести:

- передові методи прогнозного та описового моделювання;
- зручний і зрозумілий інтерфейс, що дозволяє користувачам самостійно створювати прогнозні моделі;
- автоматизований процес регламентного застосування моделей;
- можливість пакетної обробки складних процесів;
- швидкий збір і підготовка даних, їх агрегація і дослідження;
- простота масштабування та настроювання рішення;
- висока продуктивність системи навіть при роботі з великим масивом розрізнених даних.

**IBM SPSS (Statistical Package for the Social Sciences)** [11] – поширений інтелектуальний інструмент для прогнозної аналітики. Прогнозна аналітика SPSS допомагає аналізувати закономірності в історичних і поточних даних про операції або результати опитувань для прогнозу потенційних майбутніх результатів і попереджувального реагування на основі отриманих знань з метою досягнення кращих бізнес-результатів.

Ключовим компонентом рішення є **SPSS Modeler** – програмне середовище аналізу даних і текстової аналітики, яка дозволяє створювати інтелектуальні прогнозні рішення, розкриваючи закономірності і взаємозв'язки в даних. **SPSS Modeler Server** підтримує інтеграцію з інструментами для аналізу даних і моделювання, що поставляються розробниками СУБД, зокрема IBM Pure Data System for Analytics. Використовуючи один лише додаток SPSS Modeler, можна створювати, розраховувати і зберігати моделі всередині бази даних. Можна поєднувати аналітичні можливості і простоту використання SPSS Modeler з потужністю і продуктивністю СУБД, використовуючи вбудовані в бази даних алгоритми, що поставляються їх виробниками. Моделі будуються усередині баз даних та доступні до використання за допомогою зручного інтерфейсу користувача SPSS Modeler.

Платформа поглибленого аналізу даних **Dell Statistica** (з 2017-го року – Tibco Software) [12]. Платформа орієнтована на фахівців з роботи з даними і організації, які потребують обробки даних з великої кількості пристроїв ІОТ та різномірних джерел. Функціональність інструментарію допоможе підготувати

структуровані і неструктуровані дані, розгортати аналітичні інструменти на пристрої незалежно від їх місцеположення і користуватися функціями внутрішнього аналізу на платформах MySQL, Oracle і Teradata.

Завдяки Dell Statistica компанії зможуть впоратися з браком традиційних аналітиків даних та зі складністю сучасних середовищ ІОТ, а також враховувати нові джерела та типи даних.

Можливості Dell Statistica, які спрощують прогнозу аналітику:

- панелі управління з поглибленою візуалізацією, що дозволяють користувачам легко бачити результати аналізу на будь-якому етапі і прикріплювати візуалізацію, пов'язану з конкретними процесами, до верхнього рівня панелей;

- сучасний веб-інтерфейс дозволяє користувачам ділитися звітами, які можна відкрити в будь-якому браузері;

- ефективний контроль над даними, введеними ручним способом.

Крім представлених (далеко не всіх) розробників засобів прогнозу аналітики існує також велика кількість спеціалізованих фірм, які надають послуги у сфері бізнес-аналітики. Однією з найвідоміших є Elder Research [13]. Вона має великий досвід застосування багатьох програмних інструментів (включаючи всі відображені вище), що використовуються для розробки аналітичних рішень, програмування та персоналізованої візуалізації даних.

## 5. Індуктивні засоби прогнозу аналітики

Серед різноманітних засобів прогнозу аналітики можна виокремити декілька, спільною особливістю яких є застосування одного з найефективніших методів індуктивного моделювання – методу групового урахування аргументів (МГУА) [14].

Програмний засіб **Insights** [15] німецької компанії **KnowledgeMiner Software**, створеної ще 1993-му році, окрім МГУА, використовує також для моделювання і прогнозування метод комплексування аналогів та засоби нечіткої логіки. В процесі застосування програмного засобу є можливість побудови лінійних та нелінійних, статичних та динамічних моделей часових рядів, моделей з багатьма входами і одним виходом та багатьма входами і багатьма виходами. Виходи моделі можуть бути представлені як в аналітичній формі (у вигляді рівнянь з оціненими коефіцієнтами) так і графічно, за допомогою системного графу, що відображає взаємозв'язки структури системи.

Insights реалізує векторну обробку, багатоядерну та багатопроекторну підтримку для високопродуктивних обчислень. Вона масштабується до апаратного забезпечення комп'ютерів Macintosh фірми Apple. Незалежно від того, який процесор використовується: двоядерний чи два шестиядерних, засіб автоматично використовує усі можливості обладнання.

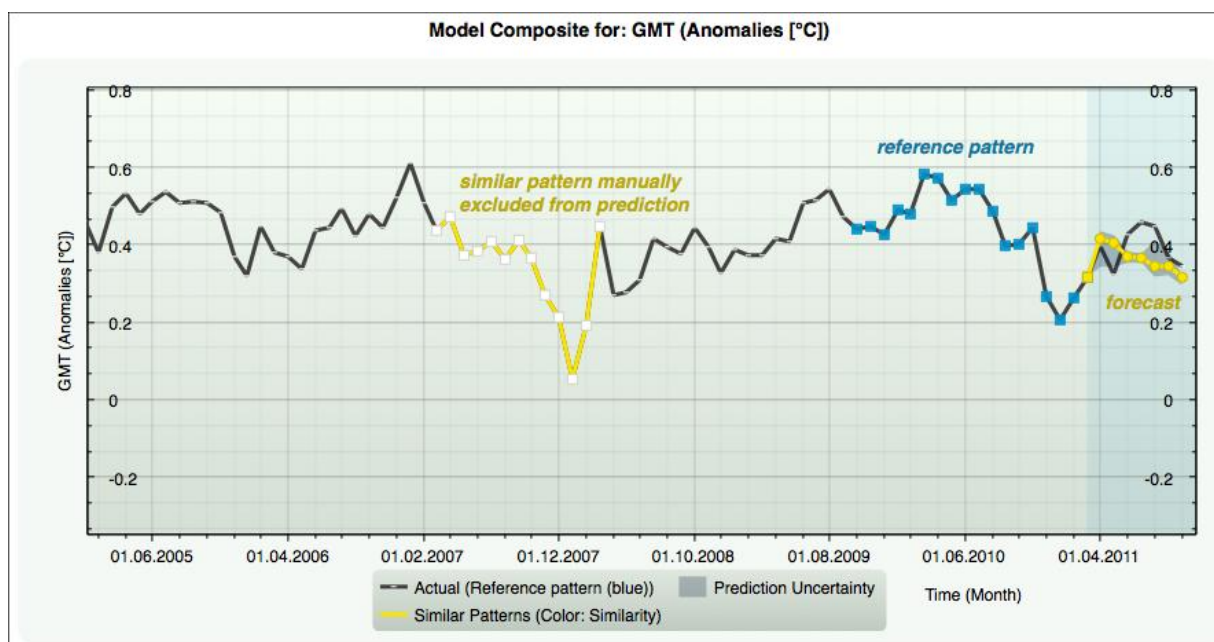


Рис. 2 Приклад використання програмного засобу Insights для прогнозного аналізу

**GMDH Shell** [16] – сучасний програмний засіб прогнозової аналітики на базі класичного алгоритму МГУА, що виконує прогнозування часових рядів, розв'язання задач класифікації та кластеризації. Це потужне рішення для аналізу багатовимірних масивів даних, отриманих з різних напрямків та сфер бізнесу. Засіб пропонує алгоритми інтелектуального аналізу даних – самоорганізовані нейронні мережі та комбінаторну структурну оптимізацію моделей. Також є можливість прискорення обчислень за допомогою Linux-кластера.

Слід зазначити, що GMDH Shell не складає конкуренції Insights у тому сенсі, що він призначений до використання на операційній системі Windows.

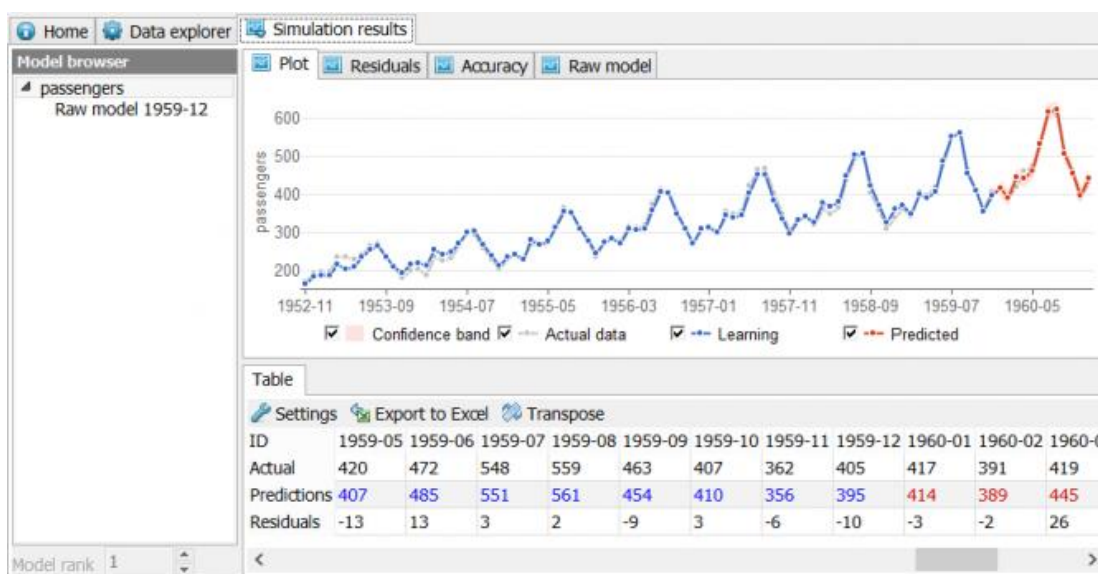


Рис. 3 Приклад використання програмного засобу GMDH Shell для прогнозного аналізу



Програмний засіб для моделювання та прогнозування багатовимірних взаємозв'язаних процесів на основі *оригінальних високопродуктивних рекурентно-паралельних алгоритмів* МГУА у класі дискретних динамічних моделей векторної авторегресії [17].

Засіб реалізовано для використання на багатопроекторних кластерних системах. Однак він може бути вбудований в будь-яку сучасну систему бізнес-інтелекту як новий інструмент.

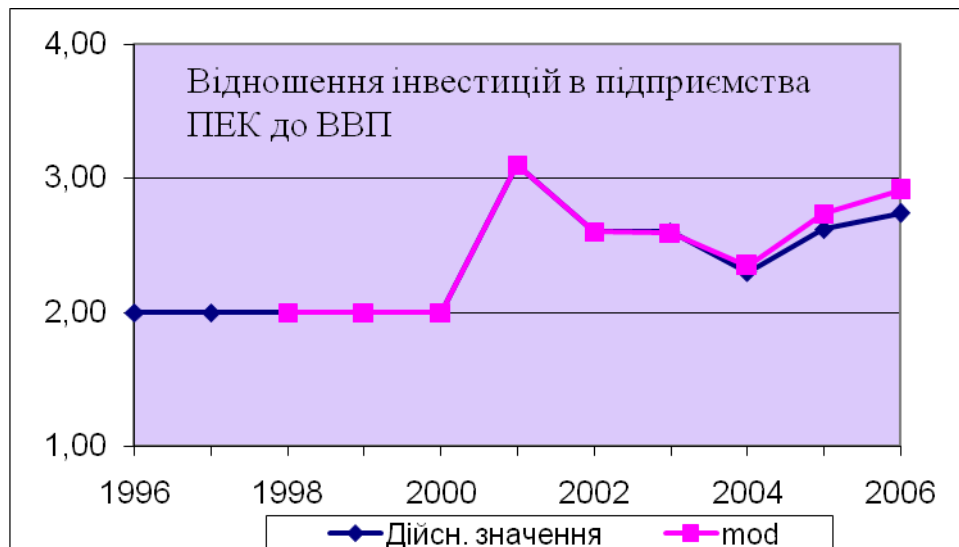


Рис. 4 Приклад використання програмного засобу основі рекурентно-паралельних алгоритмів МГУА для прогнозного аналізу

## Висновки

Сучасні можливості та передові методи прогнозової аналітики стають дієвим інструментом для збільшення продуктивності підприємства. Прогнозна аналітика стала новою тенденцією сучасності, яка відкриває широкі перспективи для подальшого розвитку компаній.

Застосовуючи системи прогнозової аналітики, слід розуміти, що робота таких системи неможлива без достатнього обсягу історичних даних і неефективна без збирання поточних даних. Чим менше даних буде використано, тим менш точними будуть прогнози.

Ефективність використання засобів прогнозової аналітики залежить як від обраних технологій, так і від якості аналітичних інструментів. І перевага тут буде на боці того інструментарію, який надасть передові методи інтелектуального аналізу даних, якими є знання-орієнтовані засоби інтелектуального моделювання на основі МГУА.

## **Література**

1. <http://microsoftblog.com.ua/2016/09/28/25-usiiei-svitovoi-ekonomiki-do-2020-roku-sta/>.
2. [https://en.wikipedia.org/wiki/Business\\_intelligence](https://en.wikipedia.org/wiki/Business_intelligence).
3. Goebel Michael and Le Gruenwald, A survey of data mining and knowledge discovery software tools, Volume 1 , Issue 1 (June 1999) Publisher ACM New York, NY, USA.
4. J Ranjan. Journal of Theoretical and Applied Information Technology 9 (1), 60-70.
5. [https://en.wikipedia.org/wiki/Business\\_analytics](https://en.wikipedia.org/wiki/Business_analytics).
6. [https://en.wikipedia.org/wiki/Predictive\\_analytics](https://en.wikipedia.org/wiki/Predictive_analytics).
7. <http://www.globalcio.ru/workshops/968>.
8. <https://www.transparencymarketresearch.com/pressrelease/predictive-analytics-industry.htm>.
9. <https://www.forrester.com/report/The+Forrester+Wave+Big+Data+Predictive+Analytics+Solutions+Q1+2013/-/E-RES85601>.
10. [https://www.sas.com/ru\\_ua/software/enterprise-miner.html](https://www.sas.com/ru_ua/software/enterprise-miner.html).
11. <https://www.ibm.com/analytics/data-science/predictive-analytics/spss-statistical-software>.
12. <https://www.tibco.com/products/tibco-statistica>.
13. <https://www.elderresearch.com>.
14. Івахненко О.Г. Метод групового урахування аргументів - конкурент методу стохастичної апроксимації // Автоматика. - 1968. - № 3. - С. 58-72.
15. <https://www.knowledgeminer.eu>.
16. <https://gmdhsoftware.com>.
17. Yefimenko S. Building Vector Autoregressive Models Using COMBI GMDH with Recurrent-and-Parallel Computations / In: "Advances in Intelligent Systems and Computing II: Selected Papers from the International Conference on Computer Science and Information Technologies, CSIT 2017, September 5-8, Lviv, Ukraine" / N. Shakhovska, V. Stepashko, Editors. – AISC book series, Volume 689. – Berlin: Springer Verlag, 2017. – 670 p. / – P. 601-613.