

УДК 519.7

Клебан Ю. В.,
викладач кафедри економіко-математичного моделювання та інформаційних технологій Національного університету "Острозька академія",

Роман С. В.,
Національний університет "Острозька академія"

ПЛАНУВАННЯ ДІЯЛЬНОСТІ ПІДПРИЄМСТВ НА БАЗІ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ

У статті розроблена модель планування збуту продукції сільськогосподарського підприємства на базі інтелектуального аналізу даних. Розглянуто різні алгоритми до здійснення цього виду аналізу та доведено доцільність їх використання.

Ключові слова: інтелектуальний аналіз даних, збут продукції, сільськогосподарське підприємство, планування діяльності підприємств.

В статье разработана модель планирования сбыта продукции сельскохозяйственного предприятия на базе интеллектуального анализа данных. Рассмотрены различные алгоритмы к осуществлению данного вида анализа и доказана целесообразность их использования.

Ключевые слова: интеллектуальный анализ данных, сбыт продукции, сельскохозяйственное предприятие, планирование деятельности предприятий.

In this paper was developed the model of planning sales of agricultural enterprise based on data mining. We handled different algorithms for implementation of this type of analysis and proved the expediency of the use.

Keywords: data mining, marketing, sales, agricultural enterprise, business planning.

Постановка проблеми. Планування збуту продукції є одним із основних етапів планування діяльності сільськогосподарських підприємств, ефективне функціонування яких залежить від досконалого вивчення ринку, потреб клієнтів, каналів реалізації, якості товарів, їх цін, комунікацій зі споживачами тощо. Основною проблемою є своєчасний збір і обробка великих обсягів інформації, а також побудова аналітичних звітів для керівників підприємств.

Аналіз останніх досліджень і публікацій. Проблеми планування збуту сільськогосподарських підприємств досліджували такі науковці, як: В. Андрійчук, А. Батіг, О. Гудзинський, Й. Завадський, П. Саблук та інші. Зокрема, О. Жук у своїй праці "Особливості планування діяльності сільськогосподарських підприємств" [1] проаналізував їх функціонування на основі таких фінансових показників, як кількість підприємств, площа сільськогосподарських угідь, середній розмір агроформувань, частка виробництва у валовій продукції області, ефективність виробництва, продуктивність праці, рівень рентабельності діяльності господарств. У результаті дослідження було доведено, що сільськогосподарські підприємства потребують розробки ефективного планування, оскільки деякі з вищеперахованих показників мають тенденцію до зниження. Тому у статті було виділено конкретні напрями щодо покращення наявної ситуації, в тому числі й у збутовій політиці, яка потребує інформаційних методів планування.

Проблематикою аналізу даних займалися багато провідних учених, серед яких можна виділити: А. Барсегяна, В. Дюка, М. Купріянова, Д. Макленнена, А. Самойленка, В. Степаненка, Ч. Танга, І. Холода, І. Чубукову тощо. Так, у статті Іванченка Н. "Інтелектуальний аналіз даних для вирішення економічних завдань" [2] було детально проаналізовано його особливості та напрями застосування. Зокрема, автором зазначається визначальна перевага інтелектуальних методів аналізу від інших – це їх здатність досліджувати велику кількість взаємозалежних даних. Також у дослідженні підтверджено доцільність їх використання у галузі сільського господарства, оскільки ці методи показують кращі результати та можуть відшукувати приховані зв'язки та залежності там, де інші методи не в змозі їх визначити. Та незважаючи на численні дослідження у цих сферах, досі не сформовано рекомендацій чи методичних порад про застосування інтелектуального аналізу даних у плануванні збуту сільськогосподарських підприємств.

Мета і завдання дослідження. Метою цієї статті є розробка моделі планування збуту продукції сільськогосподарського підприємства на базі інтелектуального аналізу даних. Конкретними цілями дослідження є: визначення ймовірного кола потенційних покупців і прогнозування обсягів реалізації продукції.

Виклад основного матеріалу. Для успішного просування товарів важливо знати всі актуальні тенденції на ринку, а також, хто є споживачем, і які у нього смаки та вподобання. Відповіді до багатьох питань, які стануть корисними при плануванні діяльності сільськогосподарського підприємства, зокрема, збуту продукції, можуть

дати засоби інтелектуального аналізу даних. Вони отримали назву *Datamining* – це тип аналітичних додатків, які підтримують рішення, розшукуючи за прихованими шаблонами інформацію у базі даних [3, с. 410].

Для демонстрації процесу планування збуту продукції створено та заповнено штучно згенерованими даними базу даних уявного сільськогосподарського підприємства. Реалізація програми для генерації набору даних виконана в середовищі розробки *Microsoft Visual C# Express Edition 2010*. Модельна база даних містить дві таблиці, які далі будуть використовуватися в експерименті (рис. 1):

Customer – таблиця, яка вміщує в собі набір усіх наявних покупців сільськогосподарської продукції та певну інформацію про них;

Demand – таблиця, в якій наявні дані про здійснені замовлення окремого товару за окремими регіонами з обсягами та сумами замовлень.

Для подальших маніпуляцій база даних розміщується на локальному сервері *SQL Server*. Як інструментарій для проведення інтелектуального аналізу використано програмне забезпечення *SQL Server Business Intelligence Development Studio* [5].

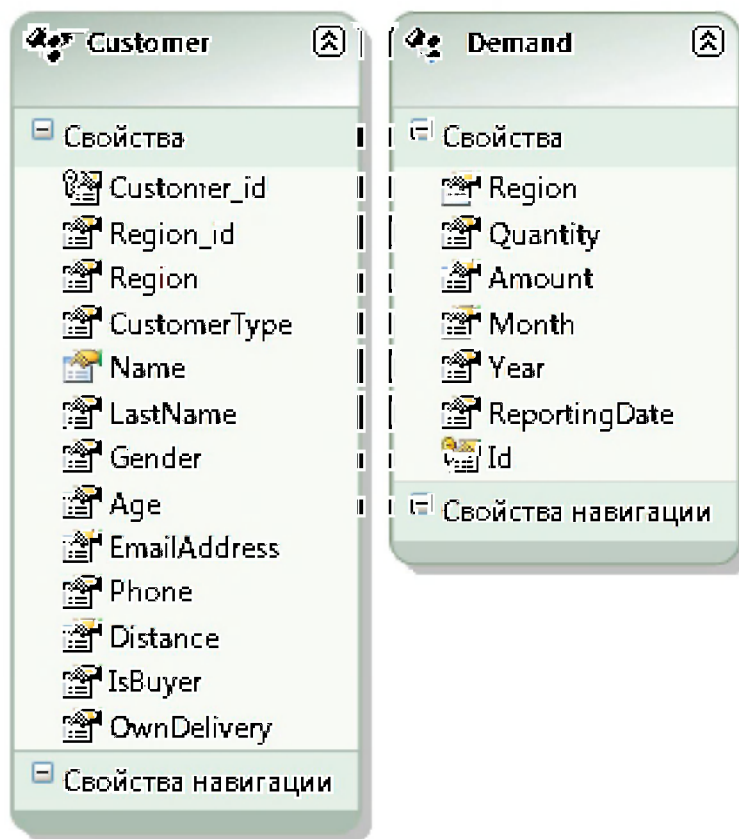


Рис. 1. Об'єктне представлення модельованої бази даних *Agricultural Database*

Розглянемо таку задачу: підприємству необхідно розширити канали збуту за рахунок залучення окремих замовників. Для цього необхідно дослідити їх особливості та розробити закономірності, за якими потенційних клієнтів буде ідентифіковано з набору наявних на ринку покупців. У цьому випадку будемо використовувати першу таблицю нашої бази даних – *Customer*. Як уже зазначалося, вона вміщує в собі набір усіх наявних покупців сільськогосподарської продукції та певну інформацію про них, зокрема: тип покупця (підприємство, підприємець, приватна особа), його вік і стать (якщо це не підприємство), відстань до нього, регіон, а також інформацію про те, чи є у покупця власні засоби доставки та чи є цей покупець клієнтом “нашого” сільськогосподарського підприємства. У цьому випадку використаємо модель кластеризації (*Microsoft Clustering*) [6].

Модель кластеризації групує наявну інформацію про покупців у кластери за певними спільними ознаками, в яких можна аналізувати взаємозв'язки між ними [4]. Використовується два методи створення кластерів та призначення кластерів точкам даних. Першим є алгоритм *K-середніх* (метод жорсткої кластеризації), за якого точка даних може належати лише одному кластеру. Для приналежності кожної точки даних цього кластеру обчислюється одне значення ймовірності. Метод мінімізує різниці між елементами кластеру та максимізує відстані між кластерами. Другий метод – це максимізація очікувань (метод м'якої кластеризації), за якого

точка даних завжди належить до декількох кластерів і для всіх можливих поєднань точок даних з кластерами обчислюються ймовірності. У цьому випадку для встановлення приналежності використовується логарифм правдоподібності даних, що вводяться в модель.

У результаті застосування цієї моделі для таблиці Customer отримали такі результати: з наявної бази усіх можливих клієнтів 51,5% не є покупцями нашого підприємства; жінок-покупців на 2,5% більше, ніж чоловіків; найчастіше зустрічаються покупці віком 26, 33 і 29 років; 60,6% потенційних покупців мають власні засоби доставки товару; на ринку 33,9% підприємств, 33,6% підприємців і 32,5% приватних осіб.

Вибірка була поділена на 10 кластерів, елементи яких спільно об'єднані завдяки ознакам, які їм притаманні. Так, наприклад, у першому, четвертому та шостому кластерах знаходяться лише підприємства. Проте перший складається виключно з підприємств, що знаходяться на півночі, півдні та за кордоном, четвертий – у центрі, на заході, півдні та сході, шостий – у всіх регіонах, з перевагою у західному. Також ці кластери різняться тим, що шостий майже повністю (98,7%) складається з підприємств, які не мають власних засобів доставки та на 77,6 % з тих, хто не є покупцями нашого сільськогосподарського підприємства. Кластери відрізняються регіоном, наявністю власних засобів доставки, статусом покупця, а також, крім того, й віком.

Більш детальну інформацію може надати порівняння окремих кластерів за спільною ознакою. Для прикладу, з наявних кластерів оберемо ті, в яких міститься найбільша частка закордонних покупців – це кластер 8 з 38% та кластер 1 з 31% (рис. 2). На діаграмі кластерів світлішим кольором відображаються ті, в яких частка покупців з-за кордону менша, відповідно темнішим – де більша.

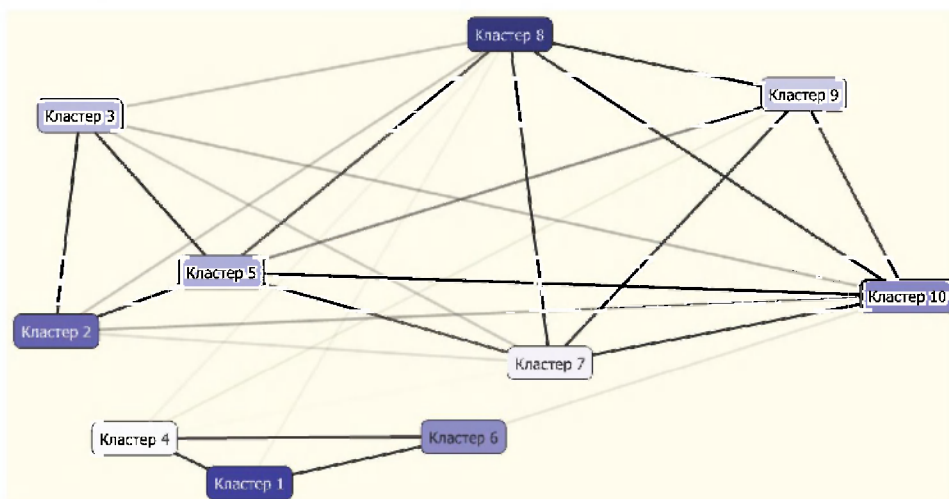


Рис. 2. Діаграма кластерів у методі кластеризації

Кластер 1 містить у собі лише підприємства, у той час як кластер 8 містить лише осіб чоловічої статі, з яких 61,46% є підприємцями, а 38,54% – приватними особами. У кластері 8 лише 26,14% на цей час є покупцями нашого підприємства, тоді як у кластері 1 точно той же відсоток не є покупцями зараз. Отже, з вищезазначених відношень підтверджується те, що за кордоном наше сільськогосподарське підприємство більше співпрацює з підприємствами, ніж з підприємцями чи приватними особами.

Отже, за допомогою моделі кластеризації ми визначили, що наше підприємство більше співпрацює з підприємствами, дещо менше з жінками-підприємцями чи приватними особами; відстань від підприємства до покупця не має значного впливу на визначення потенційного покупця; наявні та можливі покупці розташовані відносно рівномірно залежно від регіону. Тому, підсумовуючи вищезазначене, доходимо висновку, що потенційним покупцем сільськогосподарського підприємства з наявної бази даних будуть підприємці чи приватні особи чоловічої статі з власними засобами доставки товару.

Змоделюємо таку ситуацію: сільськогосподарському підприємству необхідно здійснити прогноз продажів продукції на наступні періоди. Вихідні дані для побудови моделі: обсяги та суми продажів окремих видів продукції за регіонами протягом трьох років із таблиці Demand.

Для побудови моделі використаємо алгоритм часових рядів (Microsoft TimeSeries), який ґрунтується на алгоритмі ARIMA (Autoregressive Integrated Moving Average). Він видобуває значення з ряду даних і намагається узгодити їх з шаблоном. Якщо ряд даних не є стаціонарним, алгоритм застосовує різницевий порядок. Кожне підвищення різницевого порядку, як правило, робить часовий ряд більш стаціонарним. Наприклад, якщо є часовий ряд (z_1, z_2, \dots, z_n) , над якими виконуються обчислення з використанням одиничного різницевого

порядку, буде отриманий новий ряд (y_1, y_2, \dots, y_{t-1}), де $y_t = z_{t+1} - z_t$. Якщо різницевий порядок дорівнює 2, алгоритм формує новий ряд (x_1, x_2, \dots, x_{t-2}), заснований на ряді, отриманому з рівняння першого порядку. Правильна кількість операцій обчислення різниці залежить від даних. Алгоритм часових рядів визначає оптимальний різницевий порядок ARIMA, використовуючи значення авторегресії.

Таким чином створюються окремі часові ряди для кожного регіону, які містять у собі інформацію про обсяги та суми продажів. Необхідно з'ясувати залежність обсягів продажів кукурудзяної крупи від регіонів. Для цього задаємо в модель необхідні параметри та отримуємо результат (рис. 4), який свідчить про значну відмінність обсягів продажів кукурудзяної крупи у західному районі від інших. Адже, як видно з графіка, крива обсягів продажів на заході майже завжди протягом року залишалася вище від інших. Так, у березні, коли у всіх регіонах простежується зменшення обсягів продажів, у західному регіоні приріст становить 43%. Прогноз на наступні три періоди передбачає збільшення обсягів продажів у березні на заході на 17% та на півночі на 10%, у решти регіонів має місце зменшення – на півдні на 18%, на сході на 29%, у центрі на 12% та за кордоном на 27%.

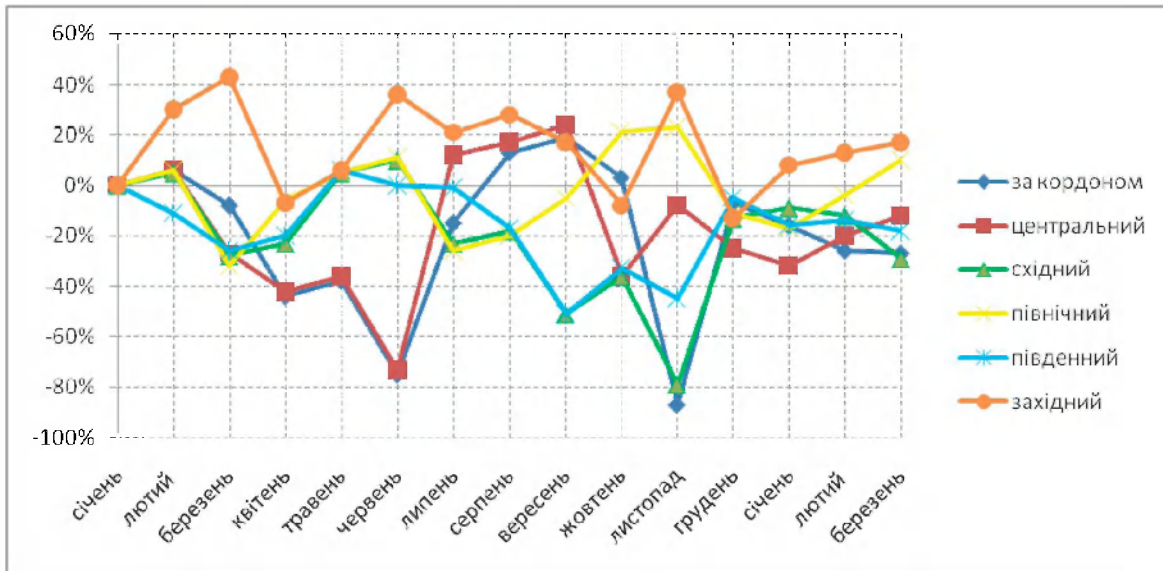


Рис. 3. Обсяги продажів кукурудзяної крупи залежно від регіону

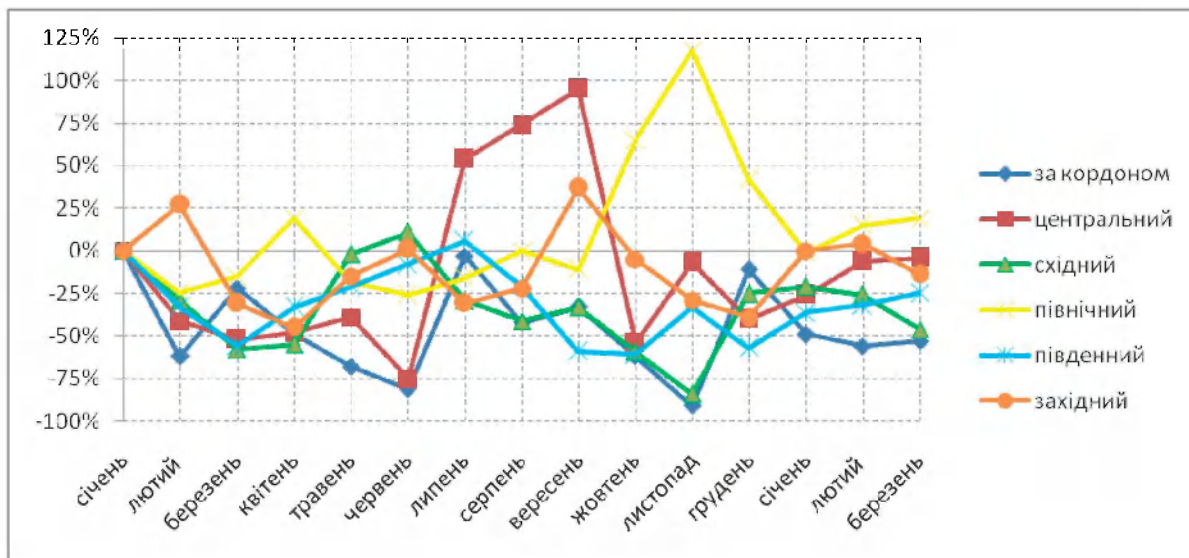


Рис. 4. Суми замовлень кукурудзяної крупи залежно від регіону

Якщо ж простежити суми продажів залежно від регіонів, то має місце така тенденція (рис. 5): перших два квартали на сході, за кордоном та у центрі суми продажів зменшуються та коливаються в межах від -22% до -81%; тоді як у червні-липні відбувається підйом приблизно до 0% у цих регіонах; у вересні відбувається значний ріст сум продажів у центрі (на 95%) та на заході (37%); суттєвий приріст відбувається на півночі на 117% у листопаді та на 42% у грудні; проте за кордоном і на сході навпаки наявний від'ємний приріст у розмірі 95% і 84% відповідно. Прогноз на перших три місяці наступного року показує, що додатний приріст сум продажів

буде присутній лише у північному регіоні (до 19% у березні), у центрі та на заході суми продажів будуть наближені до 0%, тоді як у решти регіонів буде негативний приріст до -53% за кордоном.

Між обсягами та сумами продажів існує прямопропорційна залежність – при збільшенні однієї величини на відсоток інша також збільшиться на відсоток і навпаки. Тому, відповідно до цієї закономірності, можна оцінити динаміку цін на продукцію залежно від регіонів. Для прикладу, динаміка обсягів продажів кукурудзяної крупи у західному регіоні протягом року майже завжди була позитивною, в той час як суми продажів більше мали негативну динаміку (рис. 6). Починаючи з лютого місяця обсяги продажів починають перевищувати суми на 3%, що свідчить про зменшення ціни в цьому регіоні. Найсуттєвіше зменшення ціни на кукурудзяну крупу у західному регіоні відбулося у листопаді аж на 66%, тоді ж найсуттєвіше збільшення було у вересні на 20%, зважаючи на те, що збільшення ціни відбувалося лише у вересні та жовтні. Прогноз вказує на те, що впродовж наступних трьох місяців ціна буде зменшуватися. Ідентичні спостереження за допомогою алгоритму часових рядів можна проводити щодо будь-якого продукту у будь-якому регіоні.

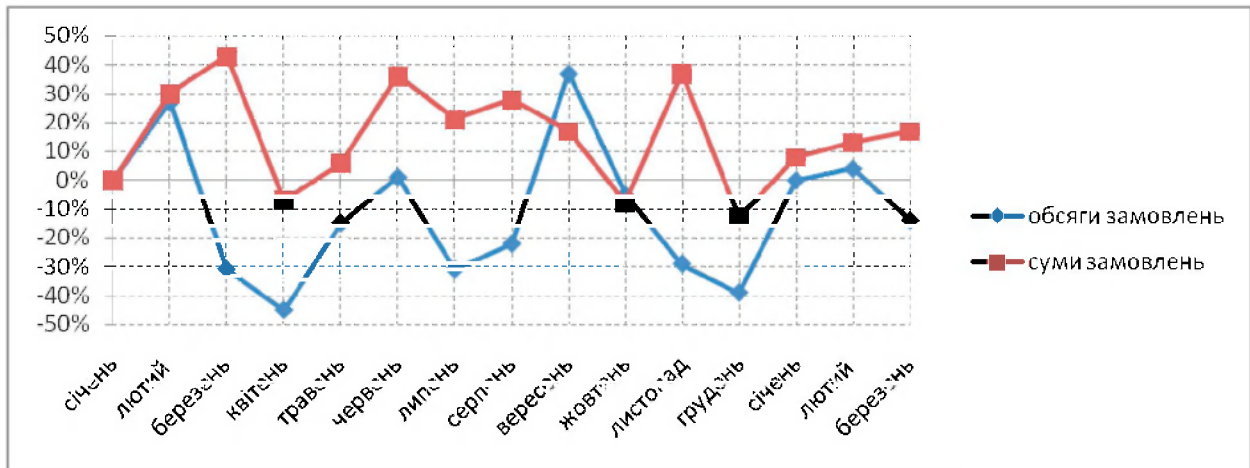


Рис. 5. Обсяги та суми замовлень кукурудзяної крупи на заході

Висновки. Отже, інтелектуальний аналіз даних є надзвичайно корисним інструментом для планування збуту продукції сільськогосподарського підприємства. Його результати є простими для інтерпретації та корисними як для поточного, так і довгострокового планування діяльності підприємства. За допомогою Data Mining легко порівнювати однорідні дані та знаходити залежності між показниками там, де здавалося б їх не існує взагалі. Також цей тип аналітичних додатків можна використовувати для вирішення будь-яких інших завдань підприємства, про які наявна інформація у базі даних.

Тому для успішного планування збуту сільськогосподарських підприємств ми рекомендуємо використовувати засоби інтелектуального аналізу. Зокрема, модель кластеризації чудово підходить для вивчення ринку, на якому функціонує підприємство, а також особливостей та вподобань покупців на ньому. Використовуючи базу даних із інформацією про тих, хто вже купує та міг би купувати продукцію, ця модель дає змогу вирішити за рахунок залучення яких замовників можна розширити канали збуту. Своєю чергою, алгоритм часових рядів є корисним інструментом для прогнозування кількісних показників функціонування сільськогосподарського підприємства та пошуку їх взаємозв'язків з якісними. З його допомогою, наприклад, можна розрахувати сподіваний обсяг продажів на такі періоди та визначити різноманітні залежності ціни від попиту, часу, місця реалізації тощо.

Література:

1. Жук О., Особливості планування діяльності сільськогосподарських підприємств // Всеукраїнський науково-виробничий журнал "Інновація економіка". – С. 38–40.
2. Іванченко Н. Інтелектуальний аналіз даних для вирішення економічних завдань // Вісник КНУТД. – № 2. – 2009. – С. 148–150.
3. Ситник В. Ф. Системи підтримки прийняття рішень: Навч. посіб. – К.: КНЕУ, 2004. – 614 с.
4. Tarhini A. Practical Data Mining: Advanced Tutorial with Visual Studio & SQL Server 2008 R2. – American university of Beirut, 2011.
5. Сайт корпорації Microsoft [Електронний ресурс]. – Режим доступу: <http://www.microsoft.com/sqlserver/en/us/solutions-technologies/business-intelligence.aspx>.
6. Сайт Microsoft Developer Network (MSDN) [Електронний ресурс]. – Режим доступу: <http://msdn.microsoft.com/ru-ru/library/ms175595.aspx>.