

ТЕХНОЛОГІЯ ПРОГНОЗУВАННЯ ВІДВІДУВАНЬ СУПЕРМАРКЕТІВ

Вінницький національний технічний університет

Анотація

У даній роботі розглянуто актуальність проблеми прогнозування відвідувань супермаркетів на основі попередньої статистики, проведено аналіз сучасних рішень, а також обґрунтовано доцільність використання лінійної регресії з метою прогнозування кількості відвідувань.

Ключові слова: Регресійний аналіз; відвідування супермаркету; технологія прогнозування.

Abstract

In this paper considered the relevance of supermarket visits prediction based on previous statistics, analysis of modern solutions and proved the feasibility of using linear regression to predict the number of visits.

Keywords: Regression analysis; visits to the supermarket; prediction technology.

В умовах розвитку ринкових відносин ефективність роботи підприємства залежить від фінансового планування. При цьому, безпосередня роль приділяється прогнозуванню фінансових результатів підприємства. Прогнозування прибутку як абсолютного кінцевого показника діяльності підприємства дає можливість уникати значних прорахунків та пов'язаних з ними втрат.

Оскільки інформація про відвідуваність супермаркету є слабкоструктурованою, а інформаційні потоки достатньо потужними, то технологія прогнозування відвідувань має базуватися на інтелектуальному аналізі та бути реалізованою у відповідних інтелектуальних програмних засобах.

При цьому вирішуються дві основні задачі:

- прогнозування майбутніх відвідувань клієнта на основі попередньої статистики;
- прогнозування майбутніх витрат клієнта на основі попередньої статистики.

Для визначення доцільності розробки технології прогнозування відвідувань супермаркету буде доречно проаналізувати вже існуючі засоби такого прогнозування.

GeneXproTools – це екстремально гнучка сервісна програма моделювання прогнозів, яка була розроблена для пошуку функцій, класифікацій, прогнозування і логічного синтезу. З GeneXproTools можна створювати аналітичні моделі за декілька хвилин, просто вивантажуючи дані та розпочавши роботу з програмою. Проте для використання даного програмного продукту є необхідним досить потужний обчислювальний пристрій і не кожне підприємство має такий у наявності. А його покупка лише заради прогнозування відвідувань є невиправданою тратою коштів. Крім того користувач для роботи з даним програмним продуктом повинен мати серйозний рівень підготовки та знання в області аналізу даних. Це робить GeneXproTools непридатним для використання у багатьох підприємствах, які не мають у своєму штаті відповідних фахівців.

Програмний засіб «1С: Управління торгівлею» є одним з найбільш широковикористовуваних та багатофункціональних систем керування підприємницькою діяльністю. Він дозволяє здійснювати різноманітні операції пов'язаних з прогнозуванням, менеджментом, маркетинговою діяльністю, управлінням персоналом і т.д. Проте, прогнозування за допомогою даного програмного пакету характеризується низькою точністю отримуваних результатів, а сам пакет має високу вартість, оскільки містить потужний арсенал інструментарію, який не завжди використовується. Отже при розробці технології прогнозування відвідувань супермаркету слід врахувати вимоги до зручності користування, інтуїтивності інтерфейсу, а також відсутності необхідності у спеціальній підготовці для використання. При цьому основною вимогою до створюваного засобу є точність прогнозування прибутку від відвідування супермаркету на визначений термін.

Основні етапи технології прогнозування відвідувань супермаркетів з урахуванням статистичної інформації такі:

1. Визначення набору X параметрів, які будуть необхідними для здійснення прогнозування.

2. Визначення типу прогнозування, а саме для одного клієнта чи одразу для усіх. Якщо обрано прогнозування для одного клієнта, користувачу потрібно вибрати потрібного покупця з бази покупців за допомогою пошуку.
3. Визначення терміну, на який потрібно зробити прогнозування.
4. Імпорт необхідних параметрів із бази покупців та відвідувань в інтелектуальний модуль.
5. Введення величини важливості для кожного з параметрів J_i , числове значення якої визначається в залежності від свіжості даних. Важливість J_i – це показник, який визначається 1 раз і під час аналізу його значення змінюватись не буде.
6. Впорядкування набіру X параметрів за ознакою зменшення їх важливості.
7. Якщо важливість J_i певного параметру більше середнього арифметичного важливості всіх параметрів Sa_J , виключення параметрів Sa_J з набору X характеристик (рис.5.1)

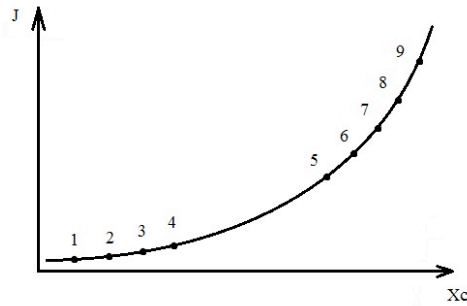


Рисунок 1 – Залежність важливості даних про відвідування J від свіжості параметра X

З рисунка 1 видно, що характеристики 1-4 – мають важливість меншу за середнє арифметичне всіх коефіцієнтів важливості, а отже характеристики 1-4 не включаються до набору X .

8. Проведення аналогічної процедури для кожного з отриманих значень x_j , досліджуючи важливість різних інтервалів значень, не включаючи у вихідну множину ті з них, значення J_i у яких є меншими за середнє арифметичне важливості всіх значень.

9. Побудова трендової моделі залежності кількості відвідувань та суми покупок клієнта від дня тижня.

10. Здійснення на основі побудованої трендової моделі прогнозування майбутніх відвідувань супермаркету і покупок, враховуючи важливість параметрів X та відкидаючи параметри, що не задовольняють умови описані в пункті 7.

11. Виведення результатів прогнозування у спеціальному текстовому полі інтелектуального модуля. У разі, якщо це прогнозування для одного клієнта, включити до результатів дату кожного візиту та суму витрат під час цього візиту, а також загальну суму витрат протягом терміну, на який здійснювалося прогнозування. Якщо прогнозування здійснювалося для усіх користувачів одразу – вивести лише загальну об'єм прибутку, що являє собою суму витрат усіх покупців.

12. Формулювання файлу звіту прогнозування, який включає всі отримані дані, а також дату і час здійснення прогнозу. Сформований файл в подальшому можна буде відтворити в рамках інтелектуального модуля [1].

Вхідними даними для означеної технології є дані про попередні відвідування та витрати клієнтів супермаркету, які, для зручності імпортується з динамічної бази даних відвідувань та прибутків. Залежність витрат відвідувача супермаркету можна подати у вигляді:

$$Y = \beta_0 + \beta_1 \chi_1 + \dots + \beta_k \chi_k + u, \quad (1)$$

де Y – вектор залежних змінних, що характеризує витрати покупця, $(\chi_1, \chi_2, \dots, \chi_k)$ – вектор незалежних змінних, який характеризує певний день тижня, u – випадкова похибка, розподіл якої в загальному випадку залежить від незалежних змінних, але математичне сподівання якої дорівнює нулю [2].

Згідно з цією моделлю математичне очікування залежної змінної є лінійною функцією незалежних змінних:

$$E(Y) = \beta_0 + \beta_1 \chi_1 + \dots + \beta_k \chi_k + u, \quad (2)$$

Вектор параметрів $(\beta_0, \beta_1 \dots \beta_k)$ є невідомим. Задача лінійної регресії полягає в оцінці вектора параметрів $(\beta_0, \beta_1 \dots \beta_k)$ на основі експериментальних значень Y і $(X_1, X_1 \dots X_k)$. Запропонована математична модель є трендовою моделлю залежності суми витрат клієнта від дня тижня, в який здійснюється відвідування супермаркету [3].

Таким чином, запропоновано технологію прогнозування відвідувань супермаркету, що базується на регресійному аналізі, яка дозволить з достатньою точністю визначити день, коли покупець наступного разу відвідає супермаркет, та кількість коштів, яку він при цьому витратить.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Т. О. Савчук, В. В. Колодний, А. В. Козачук. Система моделювання неперервних випадкових величин з графічно заданими особливостями «GenGraph»: Міжнародний науково-технічний журнал Інформаційні технології та комп'ютерна інженерія – м. Вінниця, 2011 – №2 (21). – С.49-55.

2. Колесницький О. К. Принципи побудови архітектури спайкових нейрокомп'ютерів / О. К. Колесницький // Вісник Вінницького політехнічного інституту. – Вінниця: УНІВЕРСУМ-Вінниця. – 2014. – №4 (115), С.70-78. [Електронний ресурс]. Режим доступу - <http://visnyk.vntu.edu.ua/article/view/3697/5416>

3. Введение в анализ данных [Електронний ресурс]. – Режим доступу: <http://alexanderdyakonov.narod.ru/lpotdyakonov.pdf>

Савчук Тамара Олександрівна — к.т.н., доцент, систем, професор кафедри комп'ютерних наук ВНТУ, Вінницький національний технічний університет, м. Вінниця

Бистрик Максим Васильович — студент групи ІКН-16М, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: max.bystryk@gmail.com;

Tamara O. Savchuk — Cand. Sc. (Eng), Assistant Professor, Professor of the Computer Sciences Chair, Vinnytsia National Technical University, Vinnytsia

Maksym V. Bystryk — Faculty of Information technologies and Computer engineering, Vinnytsia National Technical University, Vinnytsia, email : max.bystryk@gmail.com;