

ПРОГНОЗУВАННЯ УСПІШНОСТІ СТАРТАПІВ З ВИКОРИСТАННЯМ МЕТОДІВ DATA MINING

Вінницький національний технічний університет

Анотація

Запропоновано новий підхід до прогнозування успішності стартапів з використання методів Data Mining, метою якого є підвищення точності та швидкості визначення перспективи стартапу.

Ключові слова: *startup, Data Mining, класифікація, пошук асоціативних правил.*

Abstract

A new approach to the prediction of success of the startups using Data Mining methods whose objective is to increase the accuracy and speed of determining the prospects of a startup.

Keywords: *startup, Data Mining, classification, discovery of Association rules.*

Розв'язок задачі прогнозування успішності стартапів базується на математичній моделі, що лежить в основі відповідного процесу. Дослідження математичних моделей, за допомогою яких можна визначити успішність стартапу, є актуальною задачею, оскільки на даний момент переважають моделі, що базуються на економічному прогнозуванні [1], чого недостатньо для точного передбачення успішності. Так як економічне моделювання базується на точних даних, воно не підходить для прогнозування стартапів через невизначеність параметрів оцінки й не може забезпечити високу точність, здатність до аналізу і чуттєвість до змін.

Моделювання процесу [2] прогнозування успішності стартапів можна поділити на такі етапи:

- 1) розробка класів процесу прогнозування стартапів;
- 2) створення параметрів, що характеризують процес прогнозування стартапу;
- 3) визначення числових значень параметрів, що характеризують процес прогнозування стартапу;
- 4) обрахунок показника успішності стартапів.

Використання методів Data Mining [3] при прогнозуванні успішності стартапів дозволить забезпечити високу точність і чуттєвість до змін при їх аналізі. При цьому, доцільним є на етапі розробки класів використовувати алгоритми класифікації Card, на етапі визначення числових значень параметрів алгоритми пошуку асоціативних правил Apriori та алгоритми ранжування параметрів PageRang.

Проведені дослідження показали, що доцільно використання на етапі створення параметрів методу Беркуса, який полягає у визначенні параметрів для розрахунку витрат і визначення сильних сторін проекту [4]. Даний метод має визначенні показники параметрів, що знижує його чуттєвість до змін при ринковій економіці, а також характеризується складністю обрахунків. Значення параметрів обраховуються на основі досвіду аналітика, що значно впливає на об'єктивність отриманих результатів. З метою удосконалення методу доцільно використати підхід, який передбачає присвоєння значень параметрів за допомогою статистичних даних і розрахунок ймовірності успішності стартапу з використанням алгоритмів та методів Data Mining.

Пропонуються такі удосконалення відомого підходу [5] процесу розробки системи прогнозування стартапів, що передбачає такі основні кроки:

- 1) Створення трьох класів системи:
 - успішні стартапи (використовуються для формування статистичних даних),
 - провальні стартапи (використовуються для формування статистичних даних);
 - невизначені стартапи (алгоритм Card).
- 2) Створення однакового набору параметрів для класів.

При цьому, важливо розробити параметр, значення якого показує імовірність успішності стартапу. Для класів «успішні» стартапи та «провальні» даний параметр є сталим і не

потребує обрахунку, так як ці класи створені для формування статистичних даних. Однак в класі «невизначені» стартапи даний параметр доцільно розраховувати за допомогою інших, представлених у класі. Значення цих параметрів будуть отримуватись на основі статистичних даних представлених у класах «успішні» стартапи та «провальні».

3) **Визначення числових значень та ранжування параметрів.**

При цьому, числові значення доцільно визначати за допомогою методу пошуку асоціацій (алгоритм Apriori). Ранжування параметрів та присвоєння їм коефіцієнтів впливу на систему потрібно здійснювати за допомогою проведення запитів впливу значення того чи іншого параметра на загальну успішність стартапу – чим більше співвідношення певного значення параметра до результатів, тим більша важливість цього параметра на загальну оцінку (алгоритм PageRank).

4) **Обрахунок показника успішності стартапів.**

При визначених коефіцієнтах параметрів ймовірність успішності стартапів доцільно обраховувати як суму коефіцієнтів параметрів. Коли стає відомий результат успішності стартапу, запис видаляється з класу «невизначені» стартапи і переходить у клас «успішні» чи «провальні» стартапи.

Наведений вище підхід дозволяє поліпшити точність прогнозування, за допомогою динамічних змін значень параметрів і пошуку асоціативних закономірностей у статистичних даних, які формують ранжування параметрів для загальної системи оцінки.

Висновки

Використання методів Data Mining при прогнозуванні успішності стартапів в інтелектуальних системах, дозволить підвищити точність та швидкість визначення перспективи стартапу.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Как оценить успешность стартапа. Чек лист – дослідження – Режим доступу : <http://www.towave.ru/pub/kak-otsenit-uspeshnost-startapa-chek-list.html>.
2. Романюк О. Н. Організація баз даних і знань. Навчальний посібник / О. Н. Романюк, Т. О. Савчук.. – Видавництво: «УНІВЕРСУМ–Вінниця», 2003. – 32 -35с.,
3. Савчук Т.О.. Концептуалізація моделювання процесу аналізу проблемних ситуацій / Т.О. Савчук, О.В. Смирнова.. – Видавництво : «Вісник Вінницького політехнічного інституту», 2010. – 96-101с.,
4. Методы оценки стартапа [Електронний ресурс]. – Режим доступу : <https://habrahabr.ru/company/findstartup/blog/144521/>.
5. Барсемян А. А. Методи і моделі аналізу даних: OLAP і DATA MINING / А. А. Барсемян, М. С. Купріянов, В. В. Степаненко, І. І. Холод. – Видавництво: «БХВ-Петербург», 2004. – 67 с., ISBN: 5-94157-522-X

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

Давидов Сергій Юрійович — студент кафедри комп'ютерних наук ВНТУ, Вінницький національний технічний університет, м. Вінниця, e-mail: davydov667@gmail.com

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

Sergii U. Davydov — student of the Computer Sciences Chair, Vinnytsia National Technical University, Vinnytsia, e-mail: davydov667@gmail.com