

РОЗРОБКА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ ПРОГНОЗУВАННЯ ДАНИХ

Ярош К. Є.¹, Коваленко С. В.¹, Коваленко С. М.²

¹Національний технічний університет

«Харківський політехнічний інститут»,

²Харківський національний технічний університет сільського господарства ім. П. Василенка, м. Харків

В роботі розглянуті питання прогнозування статистичних даних. Основою для прогнозування служить інформація, що зберігається в базі даних, та може бути представлена у вигляді часових рядів. На базі цього побудовано напрям, що має назву Time-Series DataMining (DataMining часових рядів).

За оцінками зарубіжних і вітчизняних фахівців налічується понад ста методів та моделей прогнозування, в зв'язку з чим виникає задача вибору методів, які давали б адекватні прогнози для досліджуваних процесів [1]. Виникає проблема вибору моделі. Тому більш універсальною представляється ідея спільного застосування декількох алгоритмів прогнозування.

Припустимо, що для прогнозування часового ряду $y(t)$ використовуються p алгоритмів, $b_i(t)$ – прогноз i -го алгоритму на момент часу t . Знайдемо точний прогноз $a(t)$ як лінійну комбінацію b_p [2]:

$$a(t) = \sum_{i=1}^p w_{it} b_i(t), \quad \sum_{i=1}^p w_{it} = 1, \quad t = 1, \dots, T. \quad (1)$$

Ваги алгоритмів w_{it} , що використовуються для прогнозування у момент часу t , обчислюються за даними передісторії $[1, \dots, t - 1]$.

Адаптація вагових коефіцієнтів робиться за формулою:

$$w_t = \operatorname{argmin} \sum_{\tau=1}^{t-1} \theta^{t-\tau-1} \left(\sum_{i=1}^p w_i b_i(\tau) - y(\tau) \right)^2 + \lambda \sum_{i=1}^p (w_i - w_{i,t-1})^2,$$

де мінімум береться по вектору ваг $w = (w_1, \dots, w_p)$, які відповідають умові нормування (1). Параметр θ – це коефіцієнт «забування» передісторії, λ – коефіцієнт регуляції (згладжування). Другий доданок є штрафом за відхилення вектору ваг w_t від вектору ваг w_{t-1} у попередній момент часу.

Метою роботи було створення програмного забезпечення для аналізу статистичної інформації. У ході було розглянуто часові ряди, їх місце в новій області – DataMining. Був проведений аналіз методів та моделей для прогнозування часових рядів. Розглянута проблема вибору адекватної моделі. Запропоновано алгоритм для динамічної адаптації композиції моделей прогнозування. Створено програмне забезпечення, що робить композиції на основі базових алгоритмів та за допомогою вагових коефіцієнтів, що динамічно адаптуються, та знаходить найбільш точний прогноз.

Література:

1. Тихонов Е. С. Методы прогнозирования в условиях рынка / Е. С. Тихонов // Невинномисск., 2006. – С. 6.
2. Воронцов К. В. Динамически адаптируемые композиции алгоритмов прогнозирования / К. В. Воронцов, Е. В. Егорова // М., 2006. – С. 2.