

КЛАСТЕРНИЙ АНАЛІЗ ВХІДНИХ ДАНИХ У ЗАДАЧАХ ПРОГНОЗУВАННЯ ПОПИТУ ЗА ДОПОМОГОЮ НЕЙРОННИХ МЕРЕЖ

Харченко А.О.

*Національний технічний університет
«Харківський політехнічний інститут»,
м. Харків*

Задача прогнозування попиту на продукцію підприємства є досить актуальною для забезпечення ефективного управління роботи корпоративних систем з автоматизованою реєстрацією кількості замовлень. Для підвищення якості прогнозу, визначення його точності та встановлення довірчих інтервалів в умовах дефіциту та нестаціонарності інформації, необхідно проводити попередню обробку вхідних даних.

У даній роботі, продуктивність роботи корпоративної системи та розподілення кадровими ресурсами підприємства у певні періоди доби, обумовлюється точністю прогнозу кількості замовлень з урахуванням добових коливань. При використанні значної кількості вхідних даних, перед етапом прогнозування за допомогою нейронних мереж, доречно застосовувати кластерний аналіз часових рядів. Успішне розв'язання задачі кластеризації визначається відповідністю обраного методу структурі ряду, встановлення типу та вимог до вхідних даних.

Сформуємо часовий ряд $X = \{N_1, N_2, \dots, N_k\}$, де $N(t)$ – значення кількості замовлень за період часу t . Задачею кластерного аналізу даного часового ряду є пошук максимально точного відображення $f : X \rightarrow \{1, 2, \dots, K\}$, яке ставить у відповідність часовому ряду X певний клас K . Результатом пошуку відображення є функція, яка максимізує ефективність кластеризації. При цьому кількість правильно класифікованих об'єктів відповідно до періоду доби визначимо як:

$$E_n = \frac{1}{L} \sum_{i=1}^L [f(X_i) = y_i],$$

де L – загальний розмір вибірки, y_i – номер кластеру (період доби).

Для об'єднання об'єктів у періоди доби існує значна кількість методів: K -means, C -means, EM-алгоритм, нейронні мережі Кохонена та інші. У даній задачі використовувався метод C -means з модифікованим критерієм Гурвіца. Основною перевагою є встановлення α -рівней, які у подальшому використовуватимуться нейронною мережею для прогнозування значення попиту на наступний період часу X_{t+1} та перевірки належності відповідному класу K .

Таким чином, у результаті використання кластерного аналізу вхідних даних у задачі прогнозування попиту на продукцію підприємства з урахуванням добових коливань, ми отримали 4 кластери, що відображають періоди доби, та значення відповідних α -рівней.