

ВИКОРИСТАННЯ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ПРОГНОЗУВАННЯ ЦІНОВИХ ХАРАКТЕРИСТИК

Фільов І.А., Коваленко С.В.

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

Зростаюча складність ринкових процесів і велика кількість сторонніх факторів, що можуть вплинути на формування цін, створюють потребу у застосуванні сучасних інтелектуальних методів обробки та аналізу даних. Одним із підходів розв'язання подібних задач є використання штучних нейронних мереж для прогнозування цінових характеристик, зокрема – вартості товарів, послуг або об'єктів нерухомості.

У роботі проведено аналіз існуючих моделей прогнозування на основі нейронних мереж та розроблено власну модель у вигляді API для прогнозування цін залежно від ряду параметрів. Дослідження передбачало обробку історичних даних, що включали технічні характеристики певного типу об'єктів, сезонні фактори, регіональні особливості та інші економічні показники. Для підготовки даних використовувались методи нормалізації, заповнення пропущених значень, пошук та видалення аномалій та перетворення категоріальних змінних у числовий формат. Для рішення описаної задачі запропоновано наступний алгоритм.

На вхід алгоритму подається набір структурованих даних, де кожен рядок містить характеристики об'єктів та відповідне значення цінової ознаки. Алгоритм виконує попередню обробку даних, побудову моделі нейронної мережі, її навчання та подальше використання для прогнозування.

Умовно алгоритм можна розділити на декілька етапів.

1 Здійснюється зчитування вхідного датасету, заповнення пропущених значень, перетворення категоріальних змінних у числові, нормалізація числових ознак.

2 Дані випадковим чином діляться на три частини (навчальна, валідаційна, тестова). Це забезпечує можливість перевірки узагальненої здатності моделі.

3 Визначаються кількість шарів, кількість нейронів на кожному шарі, функції активації та функції втрат.

4 Алгоритм здійснює багаторазове проходження по навчальній вибірці з оновленням ваг за допомогою методу зворотного поширення помилки. У процесі контролюється значення помилки на валідаційній вибірці для запобігання перенавчанню.

5 Після навчання модель тестується на раніше невідомих даних. Результати прогнозування порівнюються з реальними значеннями ціни для обчислення середньої абсолютної та відносної похибки.

6 Навчена модель приймає на вхід нові характеристики об'єктів, обробляє їх і видає прогнозоване значення ціни.

Даний алгоритм може бути адаптований до різних предметних областей, в яких доступні історичні дані та існує потреба у прогнозі вартості товару.