

УДК 004.6:004.77

## АНАЛІЗ ДАНИХ У ВЕБ-МАГАЗИНІ ДЛЯ ПРОДАЖУ КОМП'ЮТЕРНИХ ІГОР: МЕТОДИ ПЕРСОНАЛІЗАЦІЇ ТА ПРОГНОЗУВАННЯ

Терещенко О.О.

e-mail: oleksandr.tereshchenko1@nure.ua

Науковий керівник – к.т.н., доц. Тітова О.В.

Харківський національний університет радіоелектроніки, каф. ІНФ  
м. Харків, Україна

This research is focused on creating a data analytics system for an online computer game store that includes personalized recommendations, forecasting sales trends, and detecting suspicious activity. The project aims to improve customer service, prevent fraudulent behavior, and increase sales efficiency using collaborative filtering and the ARIMA model. The system can predict user preferences, forecast demand for different games, and detect unusual patterns, providing both convenience for customers and valuable information for the store. In the future, these methods can be expanded to take into account social media factors and advertising campaigns, making the platform even more efficient.

Робота присвячена розробці системи аналізу даних у вебмагазині комп'ютерних ігор, яка включає алгоритми персоналізованих рекомендацій, прогнозування купівельної активності та виявлення аномальних дій. Використання методів машинного навчання дозволяє покращити користувацький досвід, запобігти шахрайству та підвищити ефективність продажів.

Основні аспекти дослідження

1. Алгоритми персоналізованих рекомендацій.

Для формування персональних рекомендацій використовується метод колаборативної фільтрації, зокрема алгоритм матричної факторизації (SVD). Його основна формула:

$$\hat{r}_{ui} = \mu + b_u + b_i + q_i p_u$$

де:

$\hat{r}_{ui}$  – прогнозована оцінка користувача  $u$  для гри  $i$ ,

$\mu$  – середня оцінка всіх користувачів,

$b_u, b_i$  – зміщення користувача та гри,

$q_i, p_u$  – латентні фактори користувача та гри.

2. Прогнозування купівельної активності.

Для передбачення попиту на ігри використовується модель ARIMA:

$$Y_t = c + \sum_{i=4}^p \phi_i Y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t$$

де:

$Y_t$  – значення продажів у момент часу  $t$ ,

$c$  – константа,

$\phi_i$  – коефіцієнти автокореляції,

$\theta_j$  – коефіцієнти моделі ковзного середнього,

$\varepsilon_t$  – випадкова похибка.

3. Виявлення аномалій та накручення рейтингів.

Для виявлення підозрілих активностей використовується алгоритм Isolation Forest. Аномальність точки оцінюється за глибиною її розділення в дереві:

$$s(x, n) = 2 - \frac{E[h(x)]}{c(n)}$$

де:

$E[h(x)]$  – довжина шляху до ізоляції точки  $x$ ,  $c(n)$  – нормалізаційний фактор для вибірки з  $n$  об'єктів.

Запропонований підхід дозволяє створити ефективний інструмент аналізу даних для вебмагазину ігор, що сприятиме покращенню користувацького досвіду, запобіганню шахрайству та підвищенню продажів. Подальші дослідження можуть включати розширення системи для аналізу соціальних факторів та рекламних кампаній.

Список використаних джерел:

1. Shumway, R. H., & Stoffer, D. S. (2017). *Time Series Analysis and Its Applications*. Springer.
2. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. Springer.
3. Brockwell, P. J., & Davis, R. A. (2016). *Introduction to Time Series and Forecasting*. Springer.
4. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
5. O. Chorna, P. Didyk, S. Titov, O. Titova. Usage of Clustering Algorithms for Automating Route Planning in Transportation Routing Tasks // Системи обробки інформації. № 1 (176) 2024, p. 115-123.