

Аналіз показників гри жанру «пошук предметів» та оптимізація ігрових характеристик для підвищення платежів

П'ятоволенко Олександр Олександрович, магістрант
Науковий керівник: Хом'як Тетяна Валеріївна, к.ф.-м.н., доцент
НТУ «Дніпровська політехніка», Україна

У сучасному світі ігрова індустрія виявляє неабиякий вплив на економіку та культуру суспільства. З кожним роком вона розширюється та вдосконалюється, завойовуючи нові сегменти ринку та привертаючи мільйони гравців із різних куточків світу. Ігри жанру «Пошук Предметів» (Hidden Object Games) визначаються як один із найпопулярніших жанрів у галузі відеоігор. Цей жанр привертає гравців своєю захоплюючою геймплейною механікою, загадковим сюжетом та можливістю розвивати інтелектуальні навички.

Зростання популярності цього жанру призвело до появи нових викликів для розробників ігор. Оптимізація геймплейних характеристик та аналіз показників стали важливою задачею для студій, які прагнуть не лише задовольнити потреби гравців, а й максимізувати свій прибуток через оптимізацію монетизації гри. Справжня розгадка полягає в збалансуванні між цікавістю гравців та фінансовою стабільністю розробників, яка може бути досягнута за допомогою детального аналізу і оптимізації ігрових характеристик.

Ця наукова робота присвячена вивченню різних аспектів жанру «пошук предметів» з метою визначення оптимальних стратегій оптимізації геймплейних характеристик для підвищення рівня платежів гравців. Також на меті стоїть спроба вирішити задачу класифікації для даного набору даних. Шляхом аналізу попиту гравців, їх взаємодії з ігровим середовищем та реакції на різні геймплейні елементи, стоїть мета розкрити закономірності, які впливають на їх вартість для гравців та рентабельність для розробників.

Ця робота відзначається важливістю своєї теми в контексті розвитку ігрової індустрії та її впливу на економіку. Отже, результати цього дослідження можуть послужити важливим джерелом інформації для розробників ігор, маркетологів та дослідників, що працюють у сфері відеоігор, а також для академічної спільноти, яка вивчає вплив ігор на суспільство та культуру.

Сенс гри полягає в тому, що є пейзаж, на якому розташовані предмети. Приклад такого рівня зображений на рис. 1, де на панелі ліворуч зображені предмети, які потрібно знайти гравцю. Рівень обмежений часом та має допоміжні бонуси, такі як: підказка, бомба, заморожування часу, ліхтар та додатковий час.



Рисунок 1 – Вікно рівня гри Hidden Objects

Гіпотези, які треба перевірити під час аналізу.

Було зроблено наступні припущення, щодо гравців, які використовують покупки та які треба перевірити під час аналізу:

1. *Час першої ігрової сесії.*

Було зроблено припущення, що гравці, які проходять більше рівнів за найпершу ігрову сесію, мають більше шансів зробити покупку в грі.

2. *Win-rate.*

Win-rate - це доля вигравів до всіх стартів рівнів. Було зроблено припущення, що гравці, які робили хоча б одну покупку, мають менший win-rate перед своєю першою покупкою, ніж ті, хто не сплачує.

3. *Кількість пройдених рівнів в день.*

Було зроблено припущення, що гравці які сплачують, мають більше пройдених рівнів перед своєю першою покупкою, ніж ті, хто не сплачує.

4. *Середня кількість монет, які гравець витрачає або заробляє за один пройдений рівень.*

Було зроблено припущення, що гравці які сплачують, витрачають ігрових ресурсів більше, ніж заробляють перед своєю першою покупкою, ніж ті, хто не сплачує.

5. *Кількість бонусів, які гравець витрачає на рівні.*

Було зроблено припущення, що гравці які сплачують, витрачають більше бонусів перед своєю першою покупкою, ніж ті, хто не сплачує. Це відбувається через те, що гра для таких гравців більш складна.

Після всіх трансформацій даних та перед спробою вирішити задачу класифікації була побудована зведена таблиця на основі датасету, яка має вигляд як на рис. 2, де 1 – відповідає за гравців, які сплачували, а 0 – за гравців, які не сплачували. `booster_used` – це скільки в середньому на один рівень витрачається бонусів, `money_ratio` – коефіцієнт, який відображає баланс витрат та заробітку. Якщо `money_ratio < 1`, то гравець витрачає більше ніж заробляє. Якщо `money_ratio > 1`, то гравець отримує більше ніж витрачає. `win_rate` – доля перемог. `level_time` – скільки в середньому використовують гравці часу для проходження рівня. `day_passed` – скільки в середньому проходить рівнів в день. Важливим уточненням є те, що дані для платіжних гравців враховані до моменту першої покупки, а для неплатіжних на дистанції всієї гри. Таблиця на рис. 2 підсумовує усю секцію, яка присвячена секції «Розвідувальний аналіз даних» та по ній можна зробити висновок, що люди, які не сплачують, краще грають ніж платіжні. Це можуть бути фактори різні, такі як: реакція, зір, вік та інше. Видно, що неплатіжні проходять в середньому в день стільки, скільки платіжні, при цьому вони набагато менше витрачають бонусів, у них більше він-рейт та вони більше заробляють ніж витрачають, просто тому що вони мають кращі навички гри. За даних умов гри їм і не потрібно робити покупки, тому що вони отримують в середньому в день стільки контенту, скільки отримують платіжні гравці, застосовуючи більше зусиль та витрачаючи купу бонусів. Виходом із даної ситуації може бути ускладнення гри конкретно для неплатіжних гравців, а для платіжних гравців залишити все як є.

	booster_used			money_ratio			win_rate		
	count	mean	median	count	mean	median	count	mean	median
is_paying									
0	1023	0.971338	0.789474	1023	1.470912	1.069035	1023	0.622926	0.641509
1	447	2.728507	1.333333	447	1.050710	0.745914	447	0.531015	0.500000
<hr/>									
	level_time			day_passed					
	count	mean	median	count	mean	median			
is_paying									
0	1023	69.006354	67.0	1023	6.825024	5.5			
1	447	74.200224	66.0	447	8.866890	5.0			

Рисунок 2 – Загальна зведена таблиця за таргет-класом та характеристиками

Вирішення задачі класифікації за алгоритмом Random Forest.

В результаті модель, яка побудована за алгоритмом Random Forest, має показник AUC 0.8 (рис. 3) та має силу впливу характеристик, як на рис. 4.

В результаті використання алгоритму Random Forest вдалося побудувати математичну модель, яка має AUC score на тестових даних 0.8. Тому можна вважати, що модель буде мати на практиці гарну ефективність. Також варто зауважити, що модель вважає ключовими параметрами money_ratio, booster_used та win_rate, що також підтверджено при розвідувальному аналізі даних. Також в результаті проведеного аналізу підтверджені друга, четверта та п'ята гіпотези. Для опрацювання першої гіпотези потрібно більше даних.

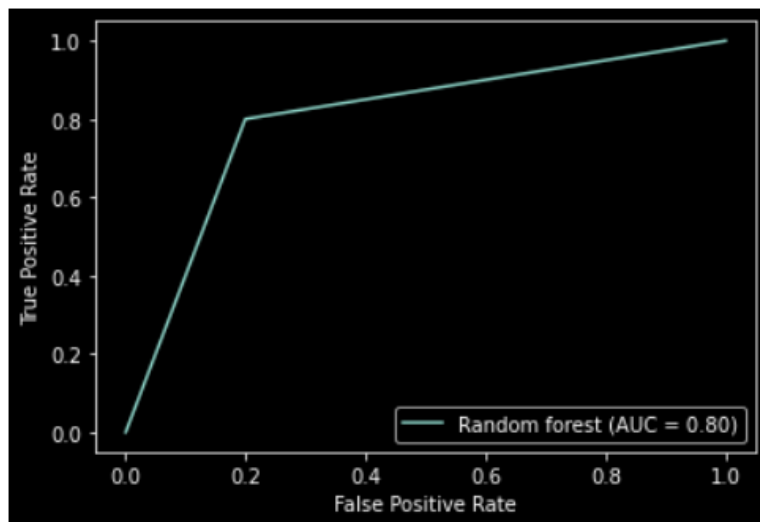


Рисунок 3 – Крива ROC та AUC для тестових даних

	importance
money_ratio	0.495836
booster_used	0.255309
win_rate	0.135400
day_passed	0.057061
level_time	0.056395

Рисунок 4 – Важливість характеристик на визначення класу алгоритму Random Forest

Для підвищення платежів в даній грі потрібно ускладнити рівні для неплатіжних гравців шляхом зменшення часу на рівні. Тоді ці гравці будуть мати більшу потребу використовувати бонуси, відповідно будуть більше витратити монет. Цей хід зменшить показники money_ratio та win_rate і збільшить показник booster_used. Тоді ці гравці за ігровими характеристиками будуть більш схожими на платіжних гравців, що в свою чергу, можливо, буде спонукати зробити першу покупку в грі.

Література:

1. Мармоза А. Т. (2022). Теорія статистики підручник. «Центр учбової літератури». https://fpk.in.ua/images/biblioteka/3fmb_finan/Teoriya-statystyky-Marmoza-A.T.pdf
2. Narkhede S. (2018). Understanding AUC - ROC Curve. Towards Data Science. <https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5>
3. Wes McKinney. (2012). Python for Data Analysis. Cambridge. <https://github.com/wesm/pydata-book>