

Д. М. Голуб,
*здобувачка другого (магістерського) рівня вищої освіти,
спеціальності 051 «Економіка»,
освітньої програми «Економічна кібернетика»;*
О. В. Ключник,
*здобувачка другого (магістерського) рівня вищої освіти,
спеціальності 051 «Економіка»,
освітньої програми «Економічна кібернетика»;*
науковий керівник
Л. М. Зомчак,
*кандидат економічних наук, доцент,
доцент кафедри економічної кібернетики;
Львівський національний університет імені Івана Франка*

МОДЕЛЮВАННЯ ЙМОВІРНОСТІ ПОКУПКИ В ЕЛЕКТРОННІЙ КОМЕРЦІЇ НА ОСНОВІ ПОВЕДІНКОВИХ ХАРАКТЕРИСТИК КОРИСТУВАЧІВ

В умовах стрімкого розвитку цифрових технологій електронна комерція стає одним із ключових каналів взаємодії між бізнесом і споживачами. Здатність компаній точно ідентифікувати поведінкові патерни користувачів, зокрема прогнозувати ймовірність здійснення покупки, є критично важливою для підвищення ефективності маркетингових стратегій та персоналізації сервісу. Це дослідження присвячене аналізу характеристик споживачів онлайн-магазину з метою побудови моделі, яка дозволяє класифікувати користувачів за фактом здійснення покупки.

У сучасних дослідженнях, присвячених аналізу великих обсягів даних та поведінковому прогнозуванню, широко застосовуються методи багатовимірного аналізу, кластеризації, авторегресивного та регресійного моделювання. Так, у статтях обґрунтовано ефективність використання багатовимірного ранжування [1, 2], кластерного підходу [3], авторегресійних моделей [4] та симульативних моделей [5] для прогнозування економічних змін, логістичної регресії у сфері маркетингових досліджень [6]. Ці методологічні підходи створюють надійне підґрунтя для побудови моделей, спрямованих на прогнозування ймовірності здійснення покупки в електронній комерції на основі аналізу характеристик споживачів.

Метою цього дослідження є ідентифікація споживачів електронної комерції, які переглядали товари на сайті компанії, з подальшим поділом їх на тих, хто здійснив покупку, та тих, хто її не здійснив. Як об'єкт дослідження було обрано підприємство, що реалізує широкий асортимент товарів, зокрема одяг і товари для дому. Для побудови моделі було використано сукупність змінних, що характеризують як соціально-демографічні ознаки споживачів (вік, стать, країна проживання, орієнтовна заробітна плата), так і поведінкові фактори (термін перебування на сайті, наявність кредитної картки, активність користувача,

баланс, кількість переглянутих товарів), а також показник кредитоспроможності (Credit Score). Результуючою змінною виступає факт здійснення покупки (Purchase). Попередній аналіз даних, проведений за допомогою функції str, дозволив з'ясувати, що вибірка складається з 10 000 спостережень і включає числові, категоріальні та бінарні змінні, що створює підґрунтя для застосування методів класифікаційного аналізу. Проаналізувавши категоріальні змінні, можна зробити висновок, що фактор Geography – країни, в яких знаходяться споживачі компанії, містить 3 унікальних значення, це 50,14 % – клієнти з Франції; 25,09 % – клієнти з Німеччини; 24,77 % – клієнти з Іспанії. Фактор Gender – стать споживача містить 2 унікальні значення серед клієнтів, а саме 54,57 % – чоловіки; 45,43 % – жінки.

Частотний аналіз числових змінних відображається графіками. За допомогою цих графіків отримали інформацію про те, які значення приймають змінні, кількість спостережень, які приймають ті чи інші значення. Графік частотного аналізу числових факторів відображає розподіл кожної з досліджуваних числових змінних серед 10 000 спостережень: кредитний рейтинг має пік в районі 650; вік концентрується між 30 та 50 роками; термін перебування клієнта розподілений відносно рівномірно; баланс показує значну кількість нульових значень та два піки в діапазоні 100 000 - 160 000; більшість клієнтів придбали 1 або 2 продукти; значно більше клієнтів мають кредитну картку, ніж не мають; кількість активних та неактивних користувачів майже однакова; заробітна плата розподілена досить рівномірно; і, що важливо, кількість клієнтів, які не здійснили покупку, значно перевищує кількість тих, хто здійснив, вказуючи на дисбаланс класів у цільовій змінній.

Для того, щоб оцінити доцільність побудови моделі використали KS Plot, яка показує, що буде у випадку, використання та не використання нашої моделі: чим вище крива моделі, тим кращою вона є (рис.1).

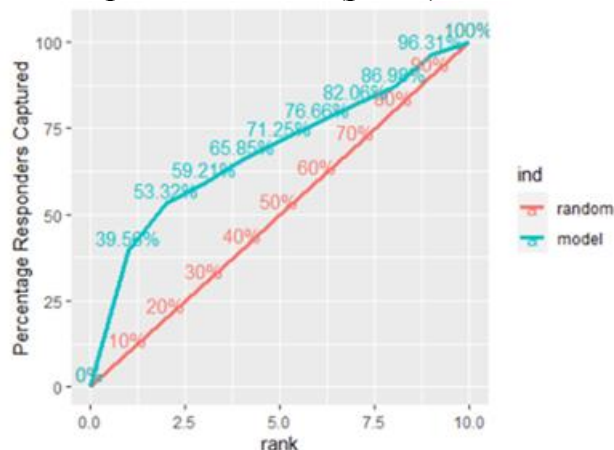


Рис. 1. KS Plot для моделі ідентифікації потенційного споживача

Джерело: побудовано авторами

KS Plot демонструє ефективність моделі у розрізненні покупців та не-покупців, порівнюючи криву моделі (синя) з випадковим вибором (червона). Синя крива, що значно піднімається над червonoю, особливо на початкових

рангах, свідчить про високу здатність моделі ідентифікувати споживачів зі значною ймовірністю покупки. Максимальна вертикальна відстань між кривими відображає розрізнявальну силу моделі, і оскільки синя крива значно вища протягом більшої частини графіка, це підтверджує, що модель градієнтного бустингу ефективно виявляє потенційних покупців порівняно з випадковим підходом.

Результати оцінок помилок моделі градієнтного бустингу (MSE 0.1315, RMSE 0.3626, RSE 0.8113, MAE 0.1315, MAD 6.575e-05, RAE 0.4056) свідчать про те, що модель в середньому робить відносно невеликі помилки у прогнозуванні ймовірності покупки, а також значно перевершує просту базову модель, яка завжди прогнозує середнє значення. Низькі значення MSE та MAE вказують на близькість прогнозів до фактичних результатів, а значення RSE та RAE менше 1 підтверджують здатність моделі виявляти закономірності в даних. Дуже низьке значення MAD свідчить про стабільність помилок моделі.

У дослідженні розроблено модель класифікації споживачів онлайн-комерції за фактом здійснення покупки з метою оптимізації маркетингових стратегій. На основі аналізу множинних соціально-демографічних та поведінкових ознак 10 000 користувачів застосовано метод градієнтного бустингу. Проведений частотний аналіз категоріальних та числових змінних виявив особливості розподілу досліджуваних характеристик.

Оцінка ефективності моделі за допомогою кривої Колмогорова-Смірнова продемонструвала її статистично значущу здатність диференціювати покупців від не-покупців. Кількісні метрики оцінки помилок (середньоквадратична помилка, середнє абсолютне відхилення тощо) підтвердили адекватність розробленої моделі та її перевагу над базовим прогнозуванням. Отримані результати обґрунтовують можливість ефективної ідентифікації потенційних клієнтів на основі аналізу їхніх цифрових слідів.

Список використаних джерел

1. Zomchak L., Hakava S. Unveiling Disparities and Resilience in Ukrainian Regional Labor Markets: Multidimensional Ranking Approach. *Developments in Information and Knowledge Management Systems for Business Applications*. Cham: Springer Nature Switzerland, 2025. Volume 8. P. 495-516.
2. Зомчак Л.М., Огородник С.П., Бортник Я.В. Місцеві бюджети регіонів України: багатовимірне моделювання динаміки та фінансової стійкості в умовах війни. *Цифрова економіка та економічна безпека*. 2025. Випуск 2(17).
3. Зомчак Л. М., Дида А. О. Регіональна диференціація та диспропорції аграрного сектору України: кластерний підхід. *Цифрова економіка та економічна безпека*. 2024. №6 (15). С. 24-30.
4. Zomchak L., Kukhotska T. Building Food Security Resilience in Ukraine: The Autoregressive Approach to Food Price Forecasting. *Developments in Information and Knowledge Management Systems for Business Applications*. Cham: Springer Nature Switzerland, 2025. Volume 8. P. 403-424.

5. Zomchak L., Miskiv D. Structural model of Ukrainian economic performance: interactions between GDP and industrial output. Smart Economy, Entrepreneurship and Security. 2024. №2(2). P. 7-16.

6. Зомчак Л. М., Вдовин М. Л. Прогнозування успішності банківського маркетингу методами логістичної регресії. Інтелект XXI. 2020. №5. С. 100-104.

І.-М. Р. Гук,
здобувачка другого (магістерського) рівня вищої освіти,
спеціальності 051 «Економіка»,
освітньої програми «Бізнес-економіка»;
науковий керівник
Л. М. Зомчак,
кандидат економічних наук, доцент,
доцент кафедри економічної кібернетики;
Львівський національний університет імені Івана Франка

ЕКОНОМЕТРИЧНЕ ДОСЛІДЖЕННЯ ВЗАЄМОЗВ'ЯЗКУ МІЖ МАШИНОБУДУВАННЯМ І ВВП УКРАЇНИ ІЗ УРАХУВАННЯМ АВТОКОРЕЛЯЦІЇ

Машинобудування є однією з ключових галузей промисловості в Україні, яка забезпечує суттєвий внесок у національну економіку. Вивчення взаємозв'язку між машинобудуванням і ВВП є важливим для розуміння загальних тенденцій економічного розвитку країни, а саме: зростання виробництва у машинобудівній галузі сприяє зростанню ВВП через мультиплікативний ефект та підвищення продуктивності і ефективності у цій галузі може свідчити про загальне економічне оздоровлення.

У контексті дослідження взаємозв'язку між машинобудуванням та ВВП України важливо враховувати ширший економічний ландшафт, зокрема регіональні особливості, структурні диспропорції та інші аспекти макроекономічного розвитку, що акцентуються в сучасних наукових працях, а також на методах, які застосовують дослідники. Так, в статті [1] застосовано багатовимірний ранжувальний підхід для оцінки регіональних ринків праці, що демонструє можливості комплексного аналізу для виявлення макроекономічних диспропорцій. Стаття [2] дозволяє глибше зрозуміти вплив кризових явищ на національні фінансові системи, що має важливе значення при аналізі таких ключових макропоказників, як ВВП. У дослідження [3] описане застосування авторегресійного підходу для прогнозування цін на продовольчі товари. Також в макроекономічних дослідженнях широко застосовують кластерний аналіз, наприклад, в [4] для оцінки регіональних диспропорцій аграрного сектору, а в [5] для виявлення спільних закономірностей економічного розвитку країн ЦСЄ.