

Ярослав Медвідь, Михайло Дендюк*

¹ НЛТУ України, Львів, Україна, 23medvid.i@nltu.lviv.ua

² НЛТУ України, Львів, Україна. ORCID: 0000-0002-7631-022X,
dendyuk@nltu.edu.ua

Розроблення інформаційної системи для бронювання номерів в готелі

Анотація. В роботі проведено аналіз даних про бронювання номерів готелів, який містить інформацію про більше як 100 тисяч бронювань з 32 атрибутами. Дослідження виявило ключові аспекти готельного бізнесу, такі як розподіл країн походження гостей, типи кімнат, канали бронювань та частоту повторних гостей. Здійснено порівняльний аналіз між курортними та міськими готелями, визначено різницю в середньому часі очікування перед підтвердженням бронювання та інші ключові моменти.

Ключові слова – бронювання номерів готелів, інформаційна система, машинне навчання, Python, Pandas, Matplotlib, Seaborn.

Аналіз готельних бронювань є важливим у сучасному світі гостювого обслуговування. Зростання конкуренції у готельній галузі та зміни в уподобаннях клієнтів створюють необхідність у глибокому розумінні поведінки та потреб гостей. Дослідження готельних бронювань дає змогу виявити ключові тенденції, що визначають успішність готельного бізнесу. Розуміння ринкових сегментів та особливостей країн походження гостей допомагає готелям адаптувати свої послуги та стратегії маркетингу. Аналіз кореляцій між різними параметрами дає змогу виявити потенційні взаємозв'язки та фактори, які впливають на успішність готельного бізнесу [1]. Порівняльний аналіз між типами готелів допомагає визначити їхні переваги та недоліки у конкретних контекстах. Напрацювання в області аналізу готельних бронювань є важливим кроком у напрямку оптимізації управління готельним бізнесом та підвищення задоволення клієнтів.

Математична модель. Набір даних містить інформацію про бронювання міського готелю та курортного готелю, а також інформацію про дату бронювання, тривалість перебування, кількість дорослих, дітей, також кількість доступних паркувальних місць. Проект має на меті отримати розуміння поведінки

клієнтів під час бронювання готелю [2]. Попит на різні сегменти клієнтів може відрізнятися, прогнозування стає важчим, оскільки може знадобитися інша модель для різних сегментів. Ця інформація може допомогти готелям скорегувати свої стратегії щодо клієнтів і підготуватися до невідомого.

Таблиця 1. Набір даних інформаційної системи

	hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number	arrival_date_day_of_month	stays_i
0	Resort Hotel	0	342	2015	July	27		1
1	Resort Hotel	0	737	2015	July	27		1
2	Resort Hotel	0	7	2015	July	27		1
3	Resort Hotel	0	13	2015	July	27		1
4	Resort Hotel	0	14	2015	July	27		1

Набір даних про бронювання номерів в готелях містить наступні вхідні ознаки. Тип готелю; бронювання скасовано чи ні; кількість днів, які пройшли між датою введення бронювання та датою прибуття; рік, місяць та день прибуття клієнта; кількість діб, на які гість забронював перебування у готелі; кількість дорослих та дітей; країна, з якої приїхав гість; код зарезервованого номера; ID туристичної агенції, яка здійснила бронювання; кількість паркувальних місць, необхідних клієнту [3].

Програмна реалізація. В роботі для проектування інформаційної системи використано бібліотеки NumPy, Pandas, Seaborn та Matplotlib. Для побудови теплової карти, яка показує кореляцію між числовими ознаками датафрейму використано бібліотеку Seaborn. Кореляція вимірює ступінь взаємозв'язку між різними ознаками, де значення близьке до 1 вказує на сильну позитивну кореляцію, близьке до -1 - на сильну негативну кореляцію, а значення близьке до 0 - на відсутність кореляції.

Кореляційні теплові карти використовуються, щоб знайти потенційні зв'язки між змінними та зрозуміти силу цих зв'язків. Lead_time і total_stay позитивно корелюються. Це означає, що якщо клієнти залишаються довше, то час виконання замовлення збільшується [4], гості, які постійно відвідують даний готель, не скасовують свої бронювання.

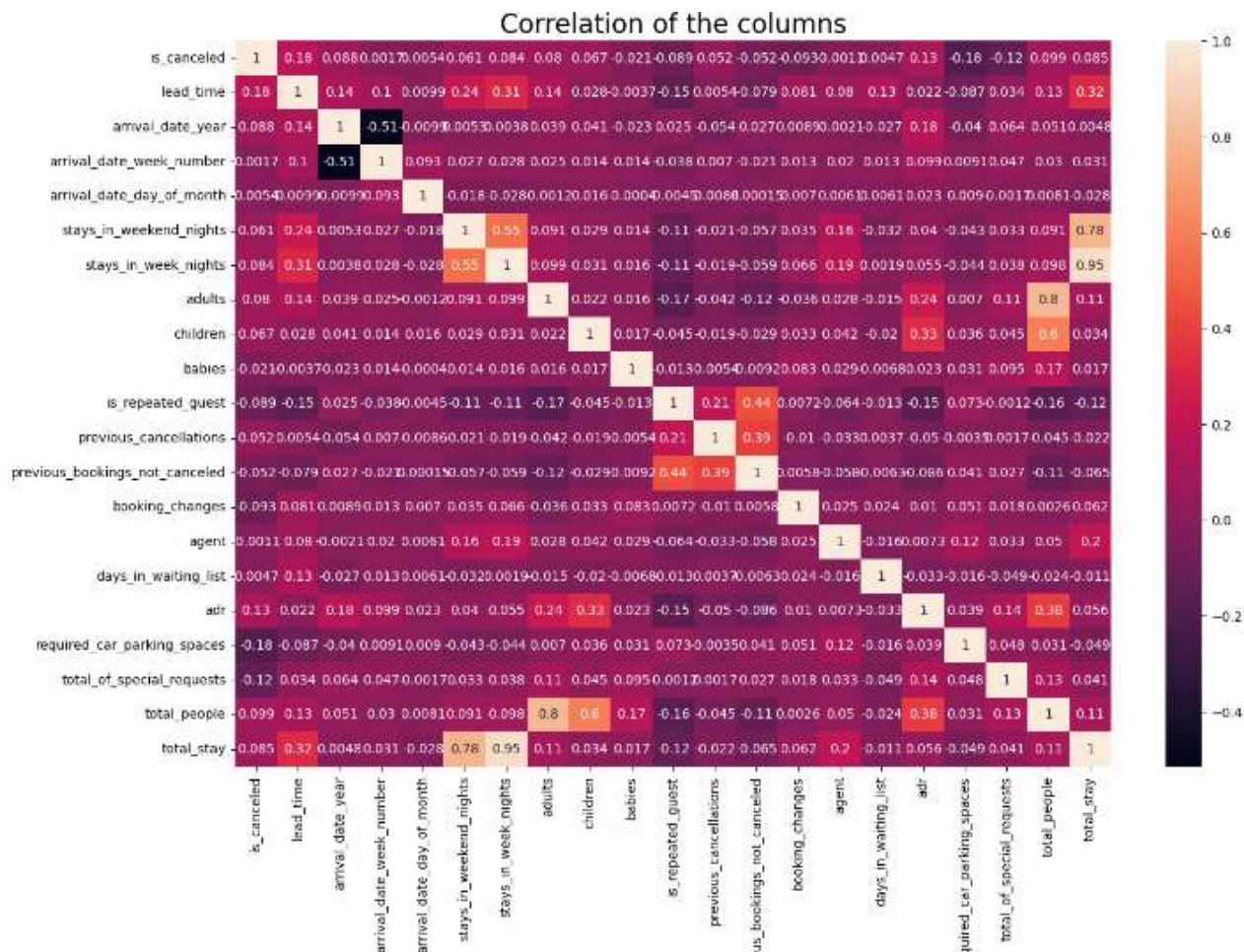


Рисунок 1. Кореляційна теплова карта

Модель машинного навчання, яка використовується в цьому проекті, створена шляхом застосування різних алгоритмів машинного навчання, таких як decision tree, logistic regression, knn, random forest. Далі використовується алгоритм random forest, який забезпечує найвищу точність прогнозу бронювання, спочатку модель проходить навчання за допомогою базового навчання, в якому точність різних алгоритмів наведена в табл. 2.

Алгоритм Random forest має найкращу точність серед усіх алгоритмів (табл. 2). Ця модель дають змогу менеджерам готелів пом'якшити втрати доходів від скасування бронювання та пом'якшити пов'язані з цим ризики. Машинне навчання має значний вплив на процес бронювання в індустрії готельного бізнесу. Алгоритми машинного навчання аналізують історичні дані з бронювання, щоб прогнозувати пікові періоди попиту. Це дає змогу готелям краще планувати ресурси і ціни. Завдяки аналізу даних про попит, конкурентів та сезонність, системи можуть автоматично коригувати ціни на номери в режимі ре-

ального часу, щоб максимізувати прибуток. Використовуючи алгоритми рекомендацій, готелі можуть пропонувати персоналізовані пакети та послуги на основі уподобань і історії попередніх бронювань кожного клієнта. Аналітика даних дає змогу готелям більш точно визначити цільову аудиторію для рекламних кампаній і підвищити їх ефективність. За допомогою оброблення природної мови системи можуть аналізувати відгуки клієнтів, виявляти основні проблеми та покращувати обслуговування. Машинне навчання використовується для створення чат-ботів, які можуть відповідати на запитання клієнтів, допомагати з бронюванням і розв'язувати проблеми в режимі реального часу. Алгоритми допомагають управляти доступними номерами в різні сезони та дні тижня, забезпечуючи ефективне використання ресурсів. Завдяки цим та іншим методам машинне навчання робить процес бронювання більш ефективним, покращує досвід клієнтів і підвищує прибутковість готелів.

Таблиця 2. Матриця оцінки різних алгоритмів

	Logistic regression	KNN	Decision tree	Random forest
Accuracy	0.804	0.8344	0.8313	0.8403
Recall	0.615	0.7446	0.7456	0.7638
Precision	0.811	0.7955	0.7881	0.7971
F1 score	0.698	0.7692	0.7663	0.7801

Висновок. Аналіз отриманих даних готельних бронювань дав можливість зрозуміти деякі ключові аспекти цього бізнесу. Набір даних містить інформацію про тип готелю, кількість скасованих бронювань, тривалість передбаченого перебування, а також деякі інші характеристики клієнтів. Було визначено найпопулярніші типи кімнат, які резервують клієнти та розподіл застосованих типів кімнат. Дослідження ринкових сегментів також відобразило основні канали бронювань. Аналіз кореляцій між числовими ознаками вказує на те, що деякі ознаки можуть бути взаємопов'язаними, що може впливати на прийняття управлінських рішень. Проведений аналіз надає цінну інформацію, яка може бути використана для оптимізації стратегій управління готельним бізнесом, враховуючи маркетинг, бронювання та обслуговування гостей.

Список використаних літературних джерел

- [1] Akhtar, N., Zubair, N., Kumar, A., & Ahmad, T. (2017). Aspect based sentiment oriented summarization of hotel reviews. *Procedia Computer Science*, 115, 563-571.
- [2] Al Shehhi, M., & Karathanasopoulos, A. (2020). Forecasting hotel room prices in selected GCC cities using deep learning. *Journal of Hospitality and Tourism Management*, 42, 40-50.
- [3] Antonio, N., Almeida, A., & Nunes, L. (2019). An automated machine learning based decision support system to predict hotel booking cancellations. *An automated machine learning based decision support system to predict hotel booking cancellations*, 1, 1-20.
- [4] Morales, D. R., & Wang, J. (2010). Forecasting cancellation rates for services booking revenue management using data mining. *European Journal of Operational Research*, 202(2), 554-562.

Development of an information system for hotel room reservations

Yaroslav Medvid', Mykhailo Dendiuk

The paper analyses data on hotel room reservations, which includes information on more than 100,000 reservations from 32 attributes. The study revealed key aspects of the hotel business, such as the distribution of guests' countries of origin, room types, booking channels and frequency of repeat guests. A comparative analysis was made between resort and city hotels, the difference in average waiting time before booking confirmation and other key points were determined.