

Моделювання прогнозу попиту з використання рекомендаційної системи на базі методу LightGBM

Актуальність і мета роботи

- В кожному бізнесі необхідно мати прогнози на майбутнє
- Прогнозування попиту є одним із ключових чинників успішності
- Мета
 - 1. Спрогнозувати помісячний попит на наступний рік
 - 2. Побудувати симуляційні моделі на базі прогнозів з LightGBM

Вхідні дані

- Датасет з продажів товарів у 6 магазинах Греції за 5 років (60 місяців)
- 5 брендів (adult-cola, kinder-cola, orange-power, gazoza, lemon-boost), 3 види упаковок (plastic, glass, can)

- Перші 5 елементів:

	id	date	city	lat	long	pop	shop	brand	container	capacity	price	quantity
0	0.0	31/01/12	Athens	37.97945	23.71622	672130.0	shop_1	kinder-cola	glass	500ml	0.96	13280.0
1	1.0	31/01/12	Athens	37.97945	23.71622	672130.0	shop_1	kinder-cola	plastic	1.5lt	2.86	6727.0
2	2.0	31/01/12	Athens	37.97945	23.71622	672130.0	shop_1	kinder-cola	can	330ml	0.87	9848.0
3	3.0	31/01/12	Athens	37.97945	23.71622	672130.0	shop_1	adult-cola	glass	500ml	1.00	20050.0
4	4.0	31/01/12	Athens	37.97945	23.71622	672130.0	shop_1	adult-cola	can	330ml	0.39	25696.0

- Для групи (shop_1,adult-cola,can):

	date	city	lat	long	pop	shop	brand	container	capacity	price	quantity
	31/01/12	Athens	37.97945	23.71622	672130.0	shop_1	adult-cola	can	330ml	0.39	25696.0
	29/02/12	Athens	37.97945	23.71622	672130.0	shop_1	adult-cola	can	330ml	0.39	29463.0
	31/03/12	Athens	37.97945	23.71622	672130.0	shop_1	adult-cola	can	330ml	0.71	18762.0
	30/04/12	Athens	37.97945	23.71622	672130.0	shop_1	adult-cola	can	330ml	0.75	20679.0
	31/05/12	Athens	37.97945	23.71622	672130.0	shop_1	adult-cola	can	330ml	0.79	21227.0

Обробка датасету

- Перетворюємо категоріальні змінні за допомогою LabelEncoder
- Дату перетворюємо у 2 нових змінних: місяць і рік
- Викидаємо Nan, ['id', 'date']

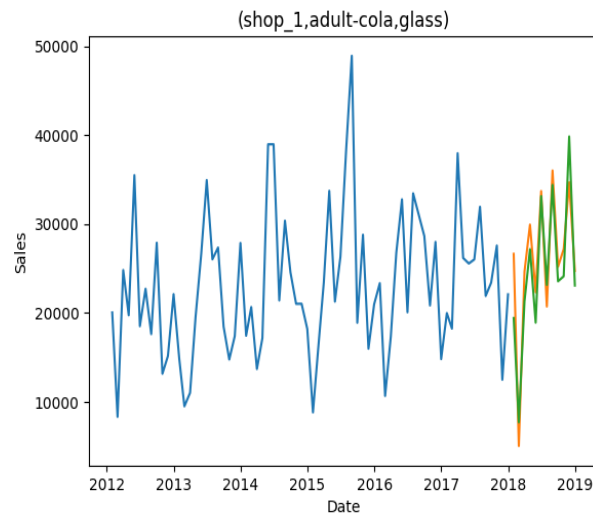
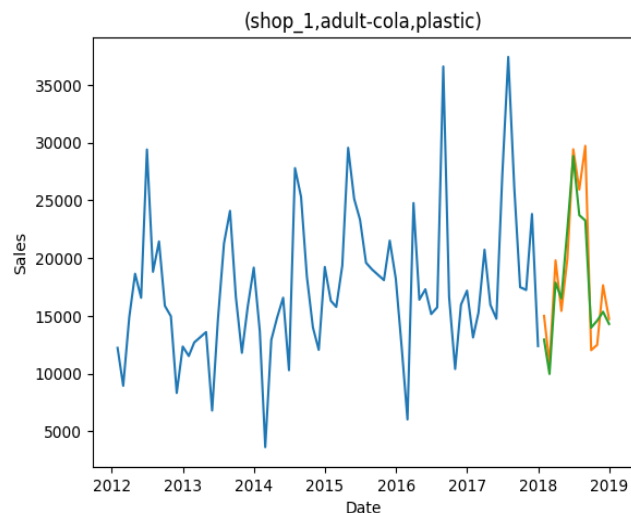
	id	date	city	lat	long	pop	shop	brand	container	capacity	price	quantity
0	0.0	31/01/12	Athens	37.97945	23.71622	672130.0	shop_1	kinder-cola	glass	500ml	0.96	13280.0
1	1.0	31/01/12	Athens	37.97945	23.71622	672130.0	shop_1	kinder-cola	plastic	1.5lt	2.86	6727.0
2	2.0	31/01/12	Athens	37.97945	23.71622	672130.0	shop_1	kinder-cola	can	330ml	0.87	9848.0
3	3.0	31/01/12	Athens	37.97945	23.71622	672130.0	shop_1	adult-cola	glass	500ml	1.00	20050.0
4	4.0	31/01/12	Athens	37.97945	23.71622	672130.0	shop_1	adult-cola	can	330ml	0.39	25696.0



	city	lat	long	pop	shop	brand	container	capacity	price	quantity	year	month
0	0	37.97945	23.71622	672130.0	0	2	1	2	0.96	13280.0	2012.0	1.0
1	0	37.97945	23.71622	672130.0	0	2	2	0	2.86	6727.0	2012.0	1.0
2	0	37.97945	23.71622	672130.0	0	2	0	1	0.87	9848.0	2012.0	1.0
3	0	37.97945	23.71622	672130.0	0	0	1	2	1.00	20050.0	2012.0	1.0
4	0	37.97945	23.71622	672130.0	0	0	0	1	0.39	25696.0	2012.0	1.0

Прогнозування “quantity” з LightGBM

- Розділимо датасет на тренувальну і тестову вибірку, виділимо предиктор “quantity” як Y, всі інші змінні як X і застосуємо LightGBM
- Для груп (shop_1, adult-cola, plastic) і (shop_1, adult-cola, glass) отримуємо таку візуалізацію прогнозів:



Синім- тренувальна вибірка
Жовтим- тестова вибірка
Зеленим прогнози

- Загальне значення RMSE- 14609177

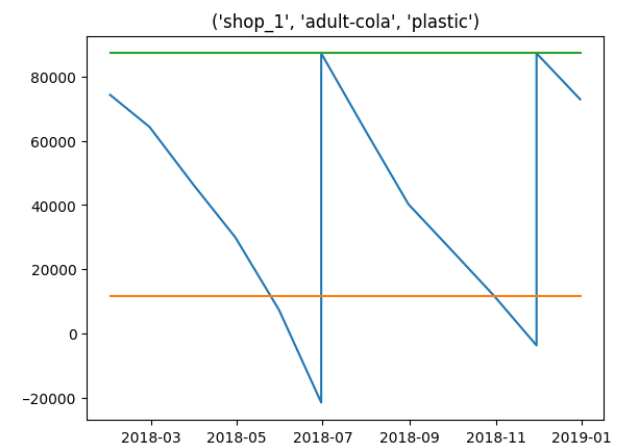
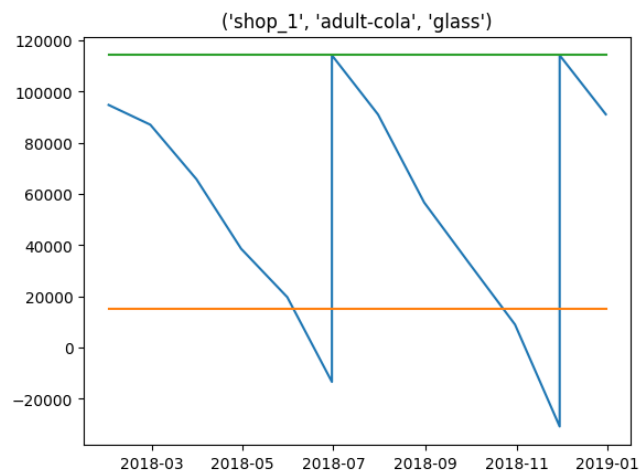
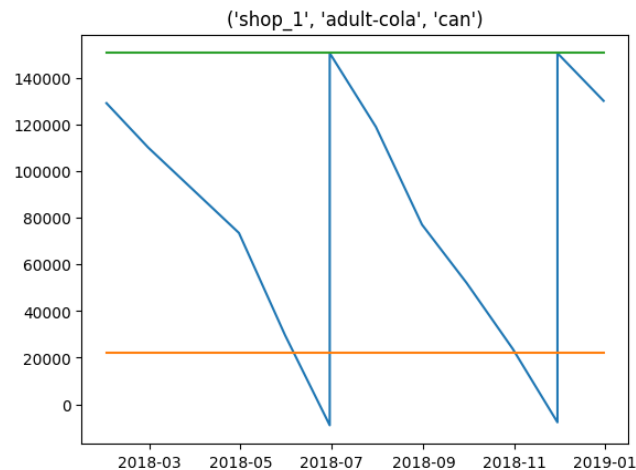
Обчислення значень необхідних і страхових запасів

- Після того, як ми обрахували помісячні прогнози, можна визначити, що продажі = попит.
- За допомогою формул визначимо необхідні запаси і страхові запаси:
- $\iota = F_{\Gamma}^{-1}(\alpha, k', \theta') + x_{min}$
- $S_s = \iota - \mu_x$, де
- $\mu'_x = \mu_x - x_{min}$
- F_{Γ}^{-1} - Функція квантилів
- (Inverse Cumulative Distribution Function)

shop_name	brand_name	container	Inventory	Safety Stocks
shop_1	adult-cola	can	150571.056898	21961.563567
shop_1	adult-cola	glass	114225.605027	15135.046746
shop_1	adult-cola	plastic	87197.113214	11602.774827
shop_1	gazoza	can	226097.928287	31134.116941
shop_1	gazoza	glass	200884.662984	26799.385314
...	...	...	...	...
shop_6	lemon-boost	glass	169165.639329	21604.171443
shop_6	lemon-boost	plastic	123283.869835	15652.651447
shop_6	orange-power	can	311882.338196	41576.103342
shop_6	orange-power	glass	250910.464141	34191.035666
shop_6	orange-power	plastic	196162.851061	25746.853132

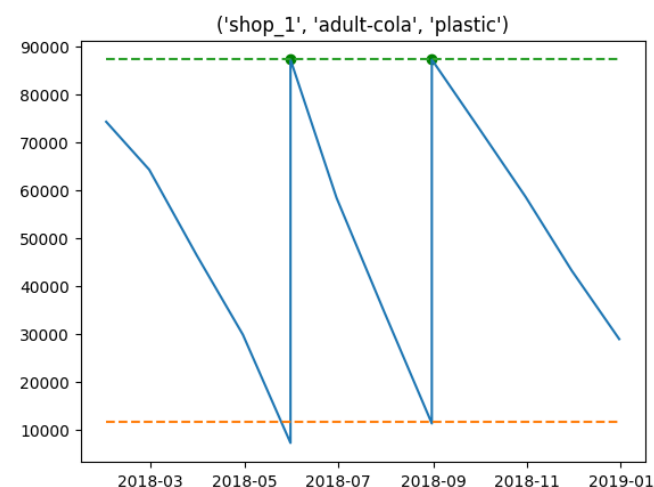
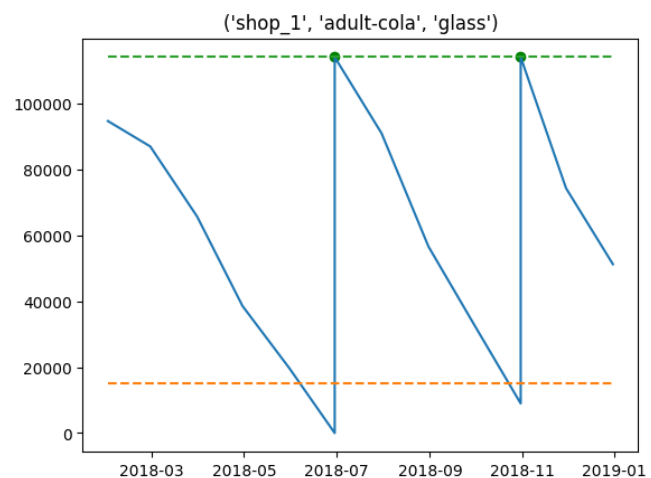
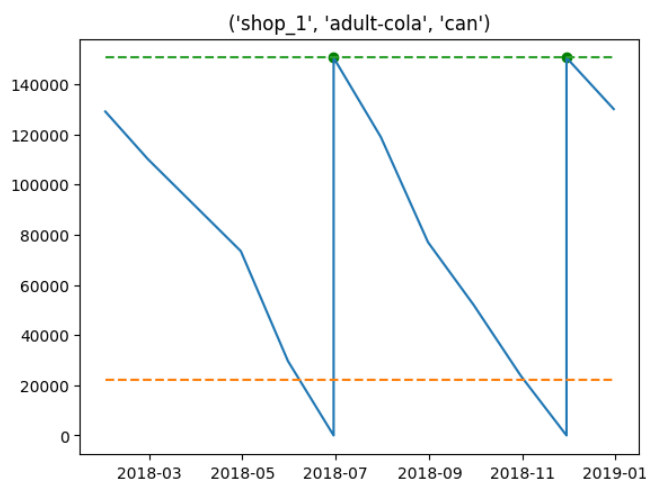
Симуляція запасів

- Для кожної групи проведемо симуляцію запасів, щоб побачити візуально динаміку змін продажів протягом року
- Правило- якщо кількість запасів перетинає значення страхових запасів, відбувається негайне поповнення до значень необхідних
- Оскільки в деяких групах кількість інвентаря може сягати значень < 0 , що нелогічно, то мінімальне значення стане 0, а різницю 0 і від'ємного значення позначимо як прогнозовані втрачені продажі:



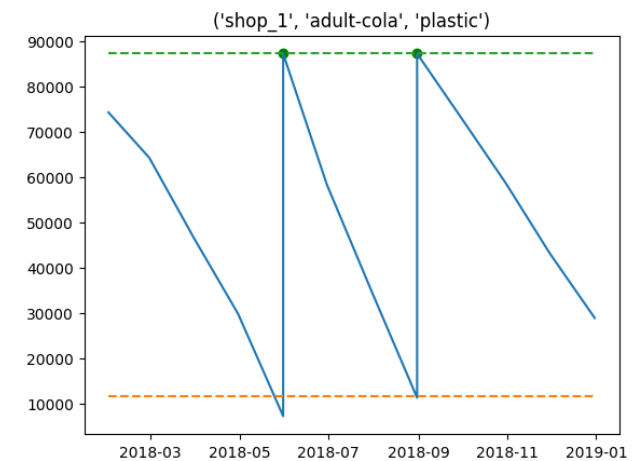
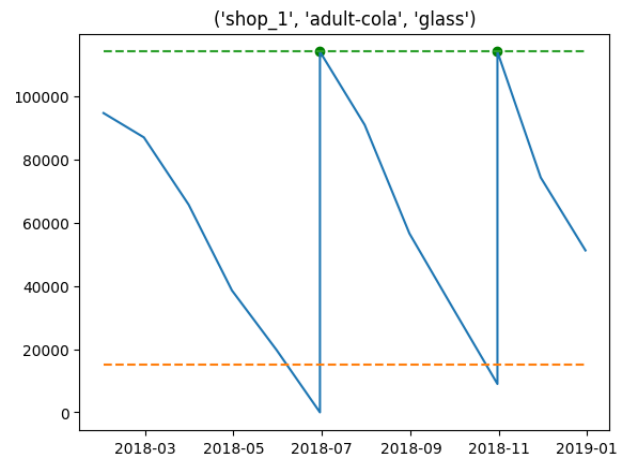
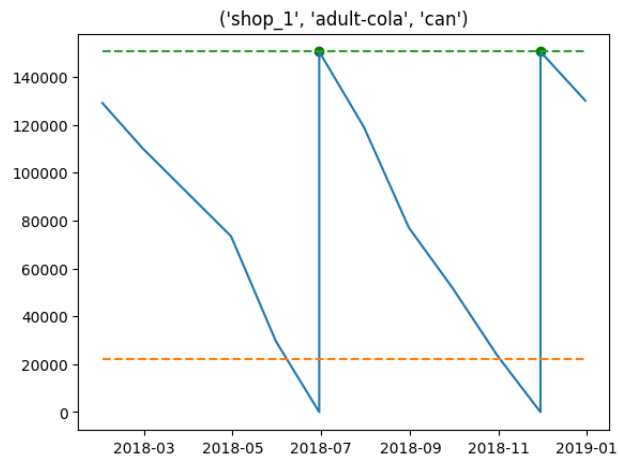
Симуляція запасів 2

- Оскільки в деяких групах кількість інвентаря може сягати значень < 0 , що нелогічно, то мінімальне значення стане 0, а різницю 0 і від'ємного значення позначимо як прогнозовані втрачені продажі:



- Причина того, що результати є менше за 0 – це обмеження у наявності даних- немає багатьох змінних, які б впливали на якість прогнозів – час доставки, вартість доставки, собівартість товарів, тощо

Результат симуляції



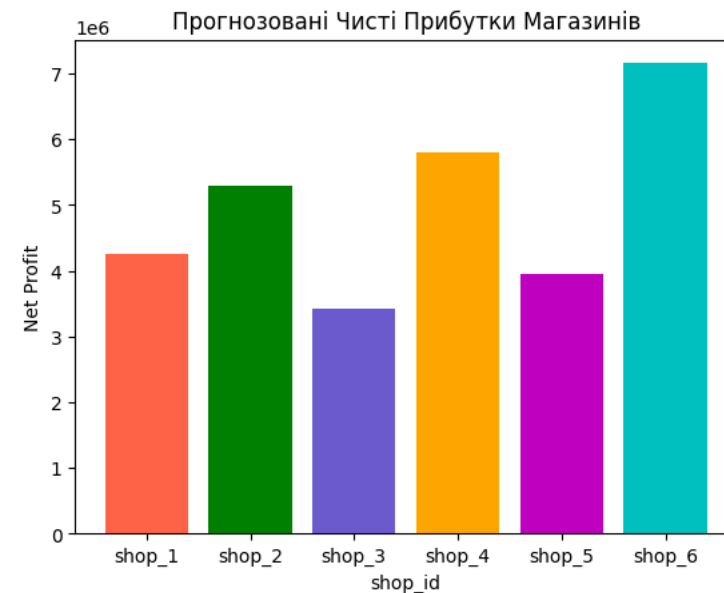
Сенс такої симуляції полягає в тому, що можна дізнатися цікаві факти про майбутні продажі, цикли повторних замовлень, потенційні збитки і доходи

shop_name	brand_name	container	Inventory	Safety Stocks	profits	losses	netto	replenishment periods
shop_1	adult-cola	can	150571.056898	21961.563567	210622.823768	9859.066319	200763.757449	6
shop_1	adult-cola	glass	114225.605027	15135.046746	305842.755275	13818.206891	292024.548384	5

Результат симуляції 2

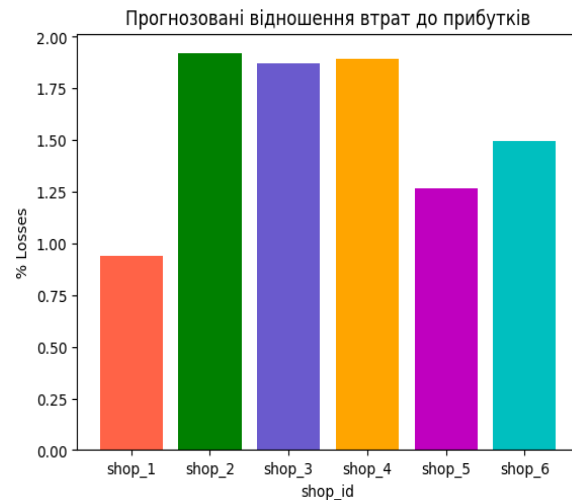
- Порахувавши потенціальні доходи і збитки за 12 місяців, згрупуємо датасет по магазинах і обчислимо їх потенційні чисті прибутки:

	shop_name	brand_name	container	Inventory	Safety Stocks	profits	losses	netto
0	shop_1	adult-cola	can	150571.056898	21961.563567	210622.823768	9859.066319	200763.757449
1	shop_1	adult-cola	glass	114225.605027	15135.046746	305842.755275	13818.206891	292024.548384
2	shop_1	adult-cola	plastic	87197.113214	11602.774827	526267.049118	0.000000	526267.049118
3	shop_1	gazoza	can	226097.928287	31134.116941	172921.889489	3736.865464	169185.024025
4	shop_1	gazoza	glass	200884.662984	26799.385314	269928.523799	0.000000	269928.523799



Валідація

- Для того, щоб визначити, чи валідні запропоновані симуляції, можна порахувати відношення % потенційних збитків до потенційних доходів.
- Якщо цей відсоток % буде незначним, то прогнози, значення необхідних і страхових запасів будуть придатними до втілення.



- Оскільки втрати не перевищують 2%, результати симуляцій можна вважати валідними

Висновки

- Було спрогнозовано попит на наступний рік з LightGBM
- Було пораховано необхідні і страхові запаси для кожної групи (shop_id, brand, container)
- Проведено симуляції запасів для наступного року.