

5. *Набір даних Car Sale Advertisement [Електронний ресурс]. Режим доступу – <https://www.kaggle.com/antfarol/car-sale-advertisements/data#>*

УДК 004.82: 004:85

*Струтовський М. І., студент 2 курсу спеціальності 122 «Комп'ютерні науки»
Нескородєва Т. В., к.т.н., доцент, доцент кафедри інформаційних технологій*

ПЕРЕДБАЧЕННЯ КІЛЬКІСТІ ІНФІКОВАНИХ COVID-19 В УКРАЇНІ

Донецький національний університет імені Василя Стуса, м. Вінниця

Під час пандемії коронавірусної хвороби 2019 (COVID-19) важливо не панікувати і зважливо оцінювати ситуацію. Переглядаючи новини, ми часто чуємо прогнози аналітиків щодо кількості інфікованих, смертей та одужань. Часто ці прогнози є не дуже оптимістичними, найгірший результат передбачення пріоритетний, ніж найкращий.

На 23.04.2020 в Україні офіційно зафіксовано 7170 випадків коронавірусної хвороби COVID-19. Актуальність передбачення кількості інфікованих досить велика. Кожен з нас хоче знати, що чекає українців в найближчому майбутньому. Саме тому я вирішив проаналізувати дані про інфікованих на коронавірус в Україні та зробити деякі висновки.

Для дослідження був використаний набір даних COVID19 Global Forecasting [1] та мова програмування Python. Із набору я обрав дані, які стосуються лише України і почав розвідувальний аналіз (EDA).

Графік залежності кількості інфікованих від кількості днів має вид експоненціальної залежності – на початку періоду швидкість зростання не велика, в наступні періоди швидко зростає. Порівняємо швидкість інфікування з іншими країнами, а саме Італією, Іспанією та Великобританією. На 45 день від першого випадку інфікування для кожної країни кількість інфікованих була така:

- Україна – 4662 особи
- Італія – 27980 осіб
- Іспанія – 11748 осіб
- Великобританія – 1551 особа

Отже, можна відзначити позитивну користь від оперативно введених обмежень в Україні на пересування, навчання, масові скупчення тощо.

Для прогнозування оберемо дві моделі: просту лінійну регресію та логістичну регресію. Ці моделі є у модулі scikit-learn.

Для того, щоб побудувати просту лінійну регресію, потрібно нормалізувати дані за допомогою натурального логарифму. Нормалізувавши, розділяємо дані на навчальну та контрольну вибірки і тренуємо модель. Похибка

даної моделі була досить великою – близько 1500 осіб, тому перейдемо до логістичної регресії.

Спробуємо натренувати логістичну регресію, яка виглядає наступним чином:

$$y = \frac{a}{1 + e^{-c(x-d)}} + b$$

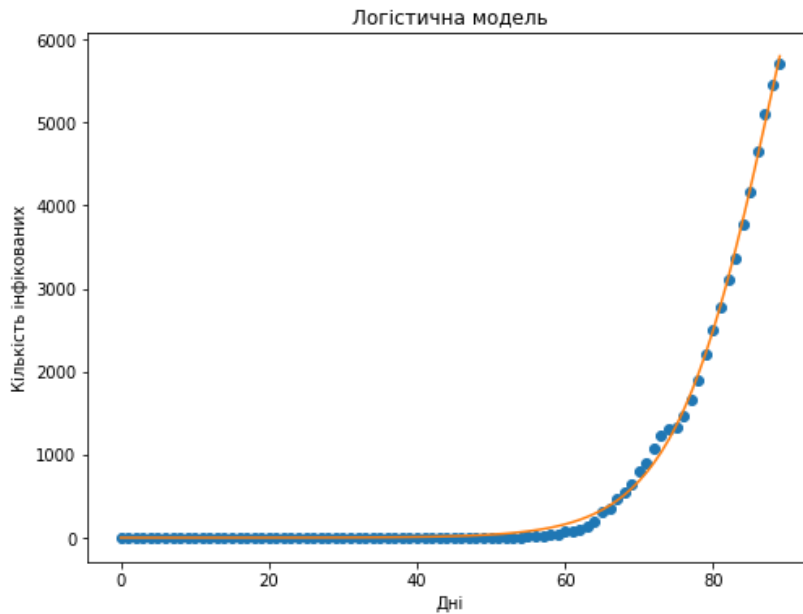


Рисунок 1 – Результат натренованої логістичної регресії

Тепер спробуємо спрогнозувати та отримаємо такий результат:

	День	Передбачено
0	19.04	5394.158774
1	20.04	5798.759985
2	21.04	6199.155890
3	22.04	6591.044223
4	23.04	6970.487207
5	24.04	7334.059117
6	25.04	7678.950773
7	26.04	8003.026225
8	27.04	8304.832847
9	28.04	8583.570818
10	29.04	8839.031177

Рисунок 2 – Передбачення кількості інфікованих

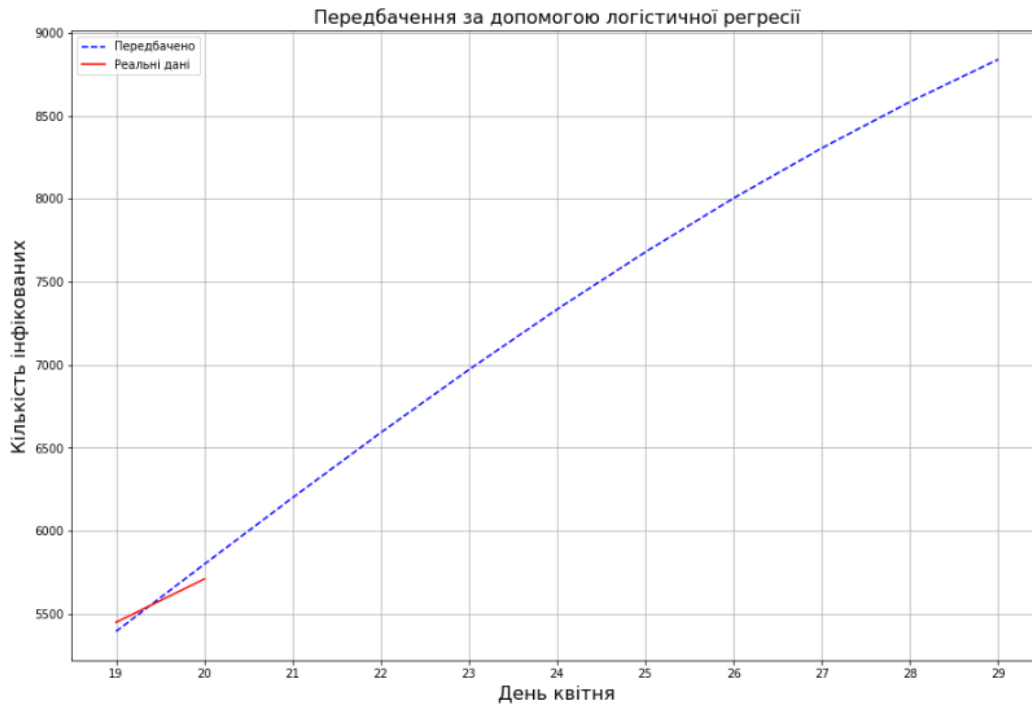


Рисунок 3 – Графік передбачення інфікованих на період 19.04 – 29.04

Отже, можемо бачити, що вже на кінець квітня спрогнозовано близько 9 тис. інфікованих від коронавірусу. Логістична регресія виявилась більш правдоподібною та має мінімальну похибку.

Список літератури

1. COVID19 Global Forecasting [Електронний ресурс]. Режим доступу - <https://www.kaggle.com/c/covid19-global-forecasting-week-4/data>
2. Trevor Hastie, Robert Tibshirani, Jerome Friedman. – “The Elements of Statistical Learning.”, 2009
3. Joel Grus.- “Data Science from Scratch.”, 2018

УДК 004.02

*Стульнев О. Г. студент 5 курсу
спеціальності 124 «Системний аналіз»
Нечволода Л. В. к.т.н., доцент кафедри
інтелектуальних систем прийняття рішень*

ОЦІНКА ЕФЕКТИВНОСТІ РЕКЛАМНОЇ КАМПАНІЇ ДЛЯ ІНТЕРНЕТ-МАГАЗИНУ

Донбаська державна машинобудівна академія, м. Краматорськ

Рекламна робота є невід'ємною складовою частиною комерційної і маркетингової діяльності будь-якого торгового підприємства. Використання Інтернету в якості рекламного каналу – безперечно, об'єктивна необхідність для