

платформах та великого порогу входження нового користувача. Ці недоліки були усунуті в розробленому програмному продукті.

Для розробки клієнтської частини веб додатку, був використаний фреймворк React та Redux, а для серверної частини була обрана бібліотека Django. Це дуже популярні та дуже функціональні бібліотеки.

Спроектвана база даних, містить в собі всі дані за співробітника, його посади, департамент, навички, списки контактів та критерії для роботи алгоритму лінійної згортки.

Для прийняття рішення, керівнику необхідно обрати критерії прийняття рішення, обрати необхідний департамент та визначитись з переліком співробітників. Потім, за допомогою алгоритму лінійної згортки, вираховуються коефіцієнти на підставі яких система рекомендує співробітника на керівну посаду.

### **Список літератури**

1. Фаулер М. *Архитектура корпоративных программных приложений.*: Пер. с англ. — М.: Издательский дом «Вильямс», 2006. — 544 с.
2. В.Д. Ногин *Принятие решений в многокритериальной среде: количественный подход (2-е изд.)* – М.: ФИЗМАТЛИТ, 2005, 176 с.
3. Alter S. L. *Decision support systems: current practice and continuing challenges.* Reading, Mass.: Addison-Wesley Pub., 1980.

**УДК 004.82: 004:85**

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

## **АНАЛІЗ СВІТОВИХ ДАНИХ ПРО ЗАХВОРЮВАНІСТЬ НА COVID-19**

*Донецький національний університет імені Василя Стуса, м. Вінниця*

Пандемія COVID-19 – пандемія коронавірусної інфекції COVID-19, викликана коронавірусом SARS-CoV-2. Почалася з виявлення в кінці грудня 2019 року в місті Ухань провінції Хубей центрального Китаю перших випадків пневмонії невідомого походження у місцевих жителів, пов'язаних з місцевим ринком тварин і морепродуктів Хуанань. 31 грудня 2019 року влада Китаю поінформували про спалах невідомої пневмонії Всесвітню організацію охорони здоров'я (ВООЗ). За даними на 29 квітня (08:10 за GMT), підтверджено 3 146 651 випадок зараження в 210 країнах і територіях, 218 178 смертей 961 833 чоловік видужала, летальність від усіх зареєстрованих завершилися випадків зараження становить 18% [1]. Регіони, найбільш піддані пандемії за кількістю хворих, - США, Європа (Італія, Іспанія, Німеччина, Франція, Великобританія, Росія), Китай, Іран, Бразилія. Регіони, які найбільше постраждали від пандемії за

кількістю загиблих, - США, Європа (Італія, Іспанія, Франція, Великобританія, Бельгія, Німеччина), Іран. Регіони, в яких летальність найбільш висока - Африка (Бурунді, Судан, Мавританія, Алжир), Багамські острови, Європа (Бельгія, Франція, Італія, Великобританія). Тому на сьогодні актуальним є дослідження факторів (предикторів  $X$ ), які впливають на захворюваність (відгук  $Y_1$ ), смертність (відгук  $Y_2$ ) і одужання (відгук  $Y_3$ ).

Кількість захворілих (значення відгуку  $Y_1$ ), може бути сильно занижена - в першу чергу, за рахунок тих, у кого хвороба протікає в легкій формі. Деякі країни (наприклад, Італія і Швейцарія) проводять політику відмови від тестування осіб, що мають лише незначні симптоми захворювання.

Аналіз проведемо за допомогою спеціалізованого статистичного пакету R [2]. Для дослідження був використаний набір даних Novel Corona Virus 2019 Dataset [3], що містить 19 928 спостережень по 8 показниках:

- SNo – ID, унікальний порядковий номер;
- ObservationDate – дата спостереження;
- Province.State – штат, область, провінція;
- Country.Region – країна, регіон;
- Last.Update – дата остатнього оновлення;
- Confirmed (відгук  $Y_1$ ) – кількість підтверджених людей заражених COVID-19;

COVID-19;

- Deaths (відгук  $Y_2$ ) – кількість померлих людей від COVID-19;
- Recovered (відгук  $Y_3$ ) – кількість хворих, які одужали від COVID-19;

Переглянемо загальну інформацію з кожного показника:

SNo	ObservationDate	Province.State	Country.Region
Min. : 1	04/10/2020: 321	:10182	US : 3830
1st Qu.: 4983	04/11/2020: 321	Diamond Princess cruise ship: 135	Mainland China: 3067
Median : 9964	04/27/2020: 321	Gansu : 101	Canada : 799
Mean : 9964	04/28/2020: 321	Hebei : 101	Australia : 628
3rd Qu.:14946	04/29/2020: 321	Anhui : 99	France : 532
Max. :19928	04/06/2020: 320	Beijing : 99	UK : 486
	(other) :18003	(other) : 9211	(other) :10586
	Last.Update	Confirmed	Deaths
3/8/20 5:31	: 1232	Min. : 0	Min. : 0.0
2020-04-10 23:01:43:	321	1st Qu.: 11	1st Qu.: 0.0
2020-04-11 22:52:46:	321	Median : 118	Median : 1.0
2020-04-28 02:31:20:	321	Mean : 3550	Mean : 219.6
2020-04-29 02:32:29:	321	3rd Qu.: 787	3rd Qu.: 10.0
2020-04-30 02:32:27:	321	Max. :299691	Max. :27682.0
(other)	:17091		Max. :132929.00

Рисунок 1 – Загальна інформація

Після цього побудуємо матрицю діаграм розсіювання:

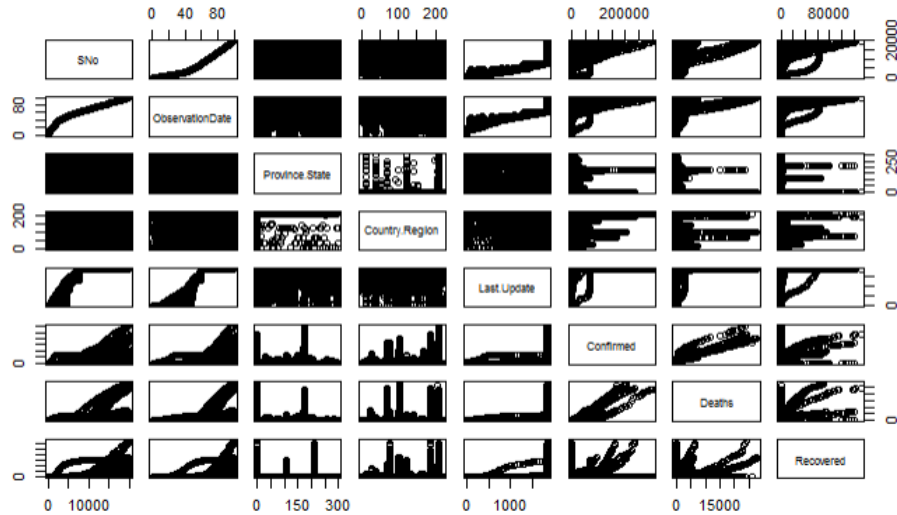


Рисунок 2 – Матриця розсіювання

Виділемо залежності між кількістю заражених і кількістю летальних випадків:

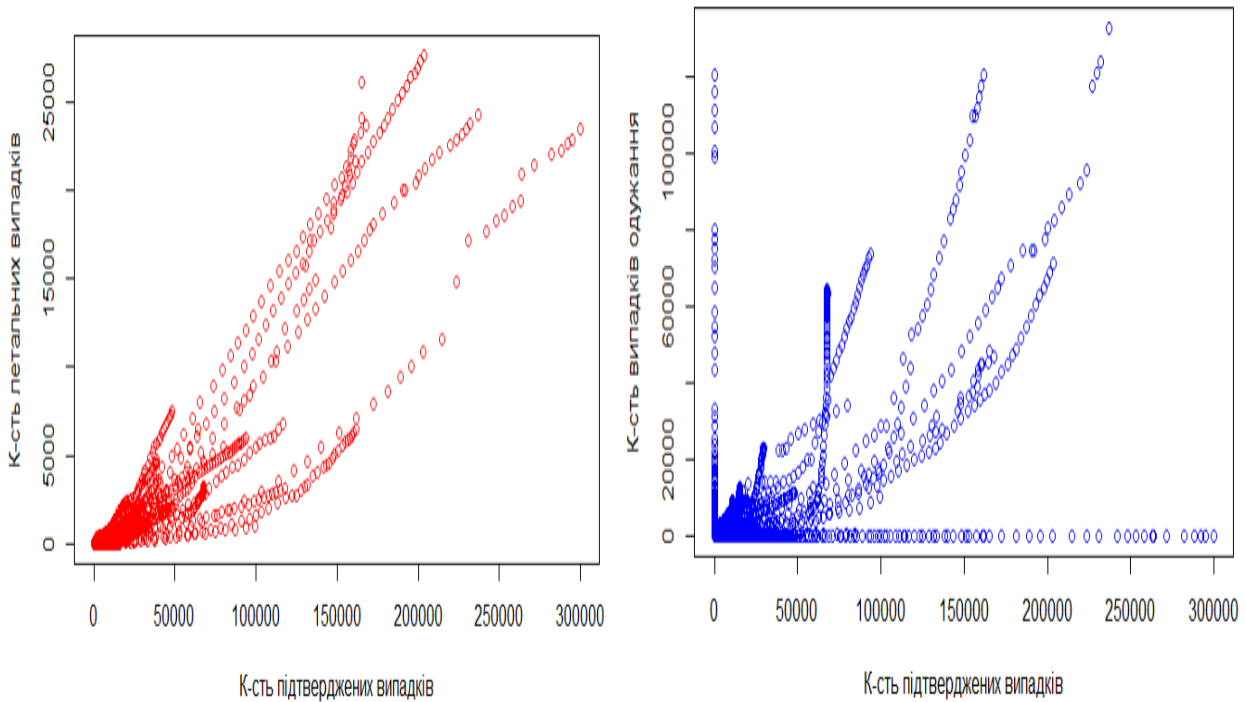


Рисунок 3 – Кореляційні поля між відгуком і предикторами

Графічний аналіз (рис.3) дозволяє зробити висновки, що зі зростанням кількості підтверджених випадків захворюваності зростає кількість летальних випадків і кількість одужавших. Також можемо помітити, що кореляційне поле має вигляд пучок зростаючих кривих, які відповідають різним країнам. Для чисельної оцінки зазначених залежностей потрібно побудувати регресійні моделі окремо по країнах та перевірити їх якість і адекватність.

### Список літератури

1. Ситуація з розповсюдження вірусу в реальному часі (дані університету Джонса Хопкінса) (англ.). Дата звернення 19 квітня 2020. (англ.)
2. Набір даних Novel Corona Virus 2019 Dataset [Електронний ресурс]. Режим доступу – <https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset#novel-corona-virus-2019-dataset>
3. NV.UA [Електронний ресурс]. Режим доступу – <https://nv.ua/ukr/world/countries/koronavirus-v-bilorusi-skilki-zarazhenih-i-tempi-poshirennya-ostanni-novini-50085697.html>
4. Humdata.org [Електронний ресурс]. Режим доступу – <https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases>

**УДК 004.82: 004:85**

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

## **АНАЛІЗ ДАНИХ ПРО РЕЙТИНГ ФІЛЬМІВ НА ПЛАТФОРМІ ОЦІНЮВАННЯ IMDB ЗАСОБАМИ МОВИ R**

*Донецький національний університет імені Василя Стуса, м. Вінниця*

У сучасному світі жодна людина не уявляє свого життя без перегляду фільмів, особливо зараз це заняття стало популярним у зв'язку із ситуацією у світі. При виборі фільму для перегляду люди часто користуються основним критерієм – рейтингом. Рейтинг у фільмів буває двох видів: від глядачів та від критиків, людей, які професійно аналізують та розбираються в даній сфері. Але як залежить рейтинг фільму глядачів від рейтингу критиків, чи є між ними схожість? Чи впливає рейтинг на популярність та успіх фільму? Дати відповідь на це питання ми зможемо провівши дослідження на основі даного набору даних.

Його актуальність обумовлена наступними факторами: зростання ринку фільмів, необхідність визначення факторів, що суттєво і несуттєво впливають на популярність фільмів. Аналіз проведемо за допомогою пакету R. Для дослідження був використаний набір даних IMDB Movie Data, що містить 1000 спостережень по 10 показниках:

1. Rank – порядок ранжування фільму.
2. Title – Назва фільму.
3. Genre – Жанр фільму.
4. Description – короткий опис фільму.
5. Director – ім'я режисера фільму.
6. Actors – Список основних зірок фільму, розділений комою.
7. Year – рік виходу фільму.
8. Runtime – Тривалість фільму в хвилинах.
9. Rating – Рейтинг фільму 0-10 (від глядачів).
10. Votes – Кількість голосів.