

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТАВРІЙСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ В.І. ВЕРНАДСЬКОГО**

Журнал заснований у 1918 році

**ВЧЕНІ ЗАПИСКИ
ТАВРІЙСЬКОГО НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ
ІМЕНІ В.І. ВЕРНАДСЬКОГО**

Серія: Технічні науки

Том 33 (72) № 6 2022



Видавничий дім
«Гельветика»
2022

Головний редактор:

Кисельов Володимир Борисович – доктор технічних наук, професор, директор Навчально-наукового інституту муніципального управління та міського господарства Таврійського національного університету імені В.І. Вернадського.

Члени редакційної колегії:

Медведєв Микола Георгійович (відповідальний секретар) – доктор технічних наук, професор, завідувач кафедри загальноінженерних дисциплін та теплоенергетики Таврійського національного університету імені В.І. Вернадського;

Бронін Сергій Вадимович – кандидат технічних наук, доцент, доцент кафедри інформаційних систем та технологій Київського національного університету імені Тараса Шевченка;

Домніч Володимир Іванович – кандидат технічних наук, професор, завідувач кафедри автоматизованого управління технологічними процесами Таврійського національного університету імені В.І. Вернадського;

Дехтяр Анатолій Соломонович – доктор технічних наук, професор, професор кафедри архітектурних конструкцій Національної академії образотворчого мистецтва і архітектури;

Дичко Аліна Олегівна – доктор технічних наук, професор, професор кафедри геоінженерії Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»;

Дубко Валерій Олексійович – доктор фізико-математичних наук, професор, професор кафедри вищої математики Київського національного університету технологій та дизайну;

Єремєєв Ігор Семенович – доктор технічних наук, професор, професор кафедри автоматизованого управління технологічними процесами Таврійського національного університету імені В.І. Вернадського;

Лисенко Олександр Іванович – доктор технічних наук, професор, професор кафедри телекомунікацій Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»;

Огородник Станіслав Станіславович – доктор технічних наук, старший науковий співробітник, професор кафедри загальноінженерних дисциплін та теплоенергетики Таврійського національного університету імені В.І. Вернадського;

Сегай Олександр Михайлович – кандидат технічних наук, доцент, доцент кафедри загальноінженерних дисциплін та теплоенергетики Таврійського національного університету імені В.І. Вернадського;

Чумаченко Сергій Миколайович – доктор технічних наук, старший науковий співробітник, завідувач кафедри інформаційних систем Національного університету харчових технологій;

Цомко Олена – доктор філософії по спеціальності «Безпека і управління інформацією», відділення комп'ютерної інженерії, Інститут Міжнародної освіти, Університет Донгсо, Республіка Корея.

Статті у виданні перевірені на наявність плагіату за допомогою програмного забезпечення StrikePlagiarism.com від польської компанії Plagiat.pl.

**Рекомендовано до друку та поширення через мережу Internet
Вченою радою Таврійського національного університету імені В.І. Вернадського
(протокол № 6 від 29.12.2022 року)**

Науковий журнал «Вчені записки ТНУ імені В.І. Вернадського. Серія: Технічні науки» зареєстровано Міністерством юстиції України (Свідectво про державну реєстрацію друкованого ЗМІ серія KB № 22895-12795P від 11.08.2017 року)

Журнал включено до Переліку наукових фахових видань України (категорія «Б») з технічних наук (спеціальності: 144. Теплоенергетика, 161. Хімічні технології та інженерія, 172. Телекомунікації та радіотехніка) відповідно до Наказу МОН України від 17.03.2020 № 409 (додаток 1), 121. Інженерія програмного забезпечення, 123. Комп'ютерна інженерія, 126. Інформаційні системи та технології, 151. Автоматизація та комп'ютерно-інтегровані технології, 275. Транспортні технології (за видами) відповідно до Наказу МОН України від 02.07.2020 № 886 (додаток 4)

Журнал включено до міжнародної наукометричної бази Index Copernicus International (Республіка Польща)

Сторінка журналу: www.tech.vernadskyjournals.in.ua

ISSN 2663-5941 (Print)

ISSN 2663-595X (Online)

© Таврійський національний університет ім. В.І. Вернадського, 2022

Вакалюк Т.А., Болотіна В.В. НЕОБХІДНІСТЬ РОЗРОБКИ ОНЛАЙН-РЕСУРСУ ДЛЯ НАУКОВОЇ ДІЯЛЬНОСТІ СПІВРОБІТНИКІВ ЗАКЛАДІВ ВИЩОЇ ОСВІТИ.....	71
Haitan O.M., Kolesnyk D.D. 3D COMPUTER GRAPHICS IN THE DEVELOPMENT OF KNOWLEDGE CONTROL SYSTEM IN THE DGBL APPROACH.....	77
Дадиверін В.В., Фещенко І.О., Потапова К.Р., Тарасенко-Клятчєнко О.В., Івасєнко Д.В. СИСТЕМА РОЗПІЗНАВАННЯ ОБЛИЧЧЯ З ВИКОРИСТАННЯМ ПОЄДНАННЯ МЕТОДІВ HAAR ТА SVM	86
Ковалєвський В.В., Вакалюк Т.А. СТАН ДОСЛІДЖЕНЬ У ГАЛУЗІ РОЗРОБКИ ТА ФУНКЦІОНУВАННІ СИСТЕМ ЗАХИСТУ СЕРВІСІВ ЕЛЕКТРОННОЇ КОМЕРЦІЇ.....	92
Лагута В.В., Єгоров О.Й., Доманська Г.А., Тимошенко Л.С. КОМП'ЮТЕРНА СИСТЕМА ДІАГНОСТИКИ ЕЛЕКТРИЧНОЇ ЦЕНТРАЛІЗАЦІЇ З ВИКОРИСТАННЯМ ОБЧИСЛЮВАЛЬНИХ ЗАСОБІВ.....	99
Макарова Л.М., Латанська Л.О., Давлатова Д.Х., Кольцов А.В. ДВОФАКТОРНА НЕЛІНІЙНА РЕГРЕСІЙНА МОДЕЛЬ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ ВЕБ-ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ МОВОЮ PYTHON З ВИКОРИСТАННЯМ DJANGO REST FRAMEWORK.....	104
Марчук Д.К., Левківський В.Л., Марчук Г.В., Голенко М.Ю. СИСТЕМА РОЗПІЗНАВАННЯ ДАКТИЛЬНОЇ МОВИ УКРАЇНСЬКОЇ АБЕТКИ.....	109
Пахомова В.М., Мотилєнко В.А. ДОСЛІДЖЕННЯ МОЖЛИВОСТІ ВИКОРИСТАННЯ RVF ДЛЯ ВИЗНАЧЕННЯ SMURF АТАК НА ОСНОВІ БАЗИ ДАНИХ KDDCUP	115
Рибальченко М.О., Маначин І.О., Потап О.Ю., Шибакінський В.І. РОЗРОБКА КОМПЛЕКСУ ДЛЯ ДОСЛІДЖЕННЯ СИСТЕМ УПРАВЛІННЯ НА БАЗІ SCADA-СИСТЕМИ ТА ПРИКЛАДНОГО ПАКЕТУ ВІЗУАЛЬНОГО МОДЕЛЮВАННЯ.....	122
ЕНЕРГЕТИКА	
Абдулін М.З., Шелешей Т.В., Бєднарська І.С., Будя Ю.С., Юрчук В.С. ПОКРАЩЕННЯ ЕКОЛОГІЧНИХ ПОКАЗНИКІВ КОТЛОАГРЕГАТІВ КВГМ 100 ШЛЯХОМ МОДЕРНІЗАЦІЇ ПАЛЬНИКОВИХ ПРИСТРОЇВ.....	130
Бєлянєвська О.А., Сухий К.М., Пустовой Г.М., Єрьомін О.О., Сухий М.К., Фролова Л.А. КРИТЕРІЇ ПІДБОРУ РОБОЧИХ ПАР ДЛЯ АДСОРБЦІЙНИХ ХОЛОДИЛЬНИХ ПРИСТРОЇВ.....	136
Кочмарський В.З., Костюк О.В., Тимєйчук О.Ю., Куба В.В. АЛГОРИТМ РОЗРАХУНКУ ШВИДКОСТІ ВИДІЛЕННЯ CaSO_3 В КВАЗІСТАЦІОНАРНОМУ РЕЖИМІ РОБОТИ ОБОРОТНИХ СИСТЕМ ОХОЛОДЖЕННЯ.....	141
Сєреда В.В., Соломаха А.С., Притула Н.О., Швець Н.О. ТЕРМОДИНАМІЧНИЙ АНАЛІЗ СИСТЕМИ ТЕРМІЧНОГО ЗНЕСОЛЕННЯ ВОДИ З ВІДКРИТИМ І ЗАКРИТИМ ПОВІТРЯНИМ ЦИКЛОМ.....	146
Черєдніков В.М., Черєднікова О.В. ВИПРОБУВАННЯ З ВИЗНАЧЕННЯ ХАРАКТЕРИСТИК ДВИГУНІВ ВНУТРІШНЬОГО ЗГОРАННЯ ГЕЛІКОПТЕРА НА ЛАБОРАТОРНОМУ СТЕНДІ.....	153
МЕТАЛУРГІЯ	
Тєслєнко О.І., Куц Г.О. СТРУКТУРНІ ТА ТЕХНОЛОГІЧНІ НАПРЯМИ ЗМЕНШЕННЯ ВИКИДІВ ПАРНИКОВИХ ГАЗІВ ПІДПРИЄМСТВАМИ ЧОРНОЇ МЕТАЛУРГІЇ.....	165

Vakaliuk T.A., Bolotina V.V. THE NEED TO DEVELOP AN ONLINE RESOURCE FOR THE SCIENTIFIC ACTIVITIES OF EMPLOYEES IN HIGHER EDUCATION.....	71
Haitan O.M., Kolesnyk D.D. 3D COMPUTER GRAPHICS IN THE DEVELOPMENT OF KNOWLEDGE CONTROL SYSTEM IN THE DGBL APPROACH.....	77
Dadyverin V.V., Potapova K.R., Feshchenko I.O., Tarasenko-Kliatchenko O.V., Ivasenko D.V. FACE RECOGNITION SYSTEM USING COMBINATION OF HAAR TA SVM METHODS.....	86
Kovalevskiy V.V., Vakaliuk T.A. THE STATE OF RESEARCH IN THE FIELD OF DEVELOPMENT AND FUNCTIONING OF PROTECTION SYSTEMS FOR E-COMMERCE SERVICES.....	92
Laguta V.V., Egorov O.Yo., Domanska G.A., Tymoshenko L.S. COMPUTER SYSTEM OF ELECTRICAL CENTRALIZATION DIAGNOSTICS USING COMPUTER TOOLS.....	99
Makarova L.M., Latanska L.O., Davlatova D.H., Koltsov A.V. A TWO-FACTOR NONLINEAR REGRESSION MODEL FOR ESTIMATING THE SIZE OF WEB APPLICATIONS DEVELOPED IN PYTHON USING DJANGO REST FRAMEWORK.....	107
Marchuk D.K., Levkivskiy V.L., Marchuk G.V., Holenko M.Yu. DACTYLIC LANGUAGE RECOGNITION SYSTEM OF THE UKRAINIAN ALPHABET.....	109
Pakhomova V.M., Motylenko V.A. STUDYING THE POSSIBILITY OF USING RBF FOR DETERMINING SMURF ATTACKS BASED ON THE KDDCUP DATABASE.....	115
Rybalchenko M.O., Manachyn I.O., Potap O.Yu., Shybakinskyi V.I. DEVELOPMENT OF A COMPLEX FOR RESEARCH OF CONTROL SYSTEMS BASED ON SCADA SYSTEM AND APPLIED VISUAL MODELING PACKAGE.....	122
POWER ENGINEERING	
Abdulin M.Z., Sheleshei T.V., Bednarska I.S., Budia Yu.S., Yurchuk V.S. IMPROVEMENT OF ENVIRONMENTAL PERFORMANCE OF KVG 100 BOILER UNITS BY MODERNIZATION OF BURNER DEVICES.....	130
Belyanovskaya E.A., Sukhyi K.M., Pustovoy G.M., Yeromin O.O., Sukhyi M.K., Frolova L.A. CHOICE CRITERIA OF WORKING PAIRS FOR ADSORPTIVE CHILLERS.....	136
Kochmarskii V.Z., Kostyuk O.P., Tymeichuk O.Yu., Kuba V.V. ALGORITHM FOR CALCULATING THE RATE OF CaCO_3 RELEASE IN THE QUASI-STATE MODE OF RCS OPERATION.....	141
Sereda V.V., Solomakha A.S., Prytula N.O., Shvets N.O. THERMODYNAMIC ANALYSIS OF THERMAL DESALINATION SYSTEM WITH OPEN AND CLOSE AIR CYCLE.....	146
Cherednikov V.M., Cherednikova O.V. TESTING FOR DETERMINING THE CHARACTERISTICS OF INTERNAL COMBUSTION ENGINES OF THE HELICOPTER ON THE LABORATORY STAND.....	153
METALLURGY	
Teslenko O.I., Kuts H.O. STRUCTURAL AND TECHNOLOGICAL DIRECTIONS OF REDUCING GREENHOUSE GAS EMISSIONS BY FERROUS METALLURGY ENTERPRISES.....	165

Макарова Л.М.

Національний університет кораблебудування імені адмірала Макарова

Латанська Л.О.

Національний університет кораблебудування імені адмірала Макарова

Давлатова Д.Х.

Національний університет кораблебудування імені адмірала Макарова

Кольцов А.В.

Національний університет кораблебудування імені адмірала Макарова

ДВОФАКТОРНА НЕЛІНІЙНА РЕГРЕСІЙНА МОДЕЛЬ ДЛЯ ОЦІНЮВАННЯ РОЗМІРУ ВЕБ-ЗАСТОСУНКІВ, ЩО СТВОРЮЮТЬСЯ МОВОЮ PYTHON З ВИКОРИСТАННЯМ DJANGO REST FRAMEWORK

В умовах нескінченного технологічного розвитку та прагнень до швидкого й зручного отримання інформації є постійна необхідність у створенні нових програмних продуктів, мобільних додатків та веб-застосунків. Кожен розробник ПЗ має за мету якомога швидше та якісніше виконати поставлені задачі, а також створити щось нове раніше за інших. Тобто компанії прагнуть максимально оптимізувати та автоматизувати як бізнес-процеси, так і роботу своїх співробітників. Нині існує багато мов програмування, кожна з яких має своє спрямування, переваги та недоліки, а також розроблена для вирішення конкретних питань та досягнення певних цілей. Для проведення аналізу було обрано мову програмування Python, оскільки вона поширена серед багатьох відомих компаній і є багатофункціональною мовою, що швидко розвивається, легко масштабується та має дуже зручний, логічний синтаксис, широку та всебічну підтримку від ком'юніті, велику базу фреймворків та готових бібліотек для вирішення найрізноманітніших задач. Проаналізувавши існуючі фреймворки Python, було обрано високорівневий та безкоштовний веб-фреймворк Django, що має доступ до вихідного коду, дозволяє легко та швидко створювати веб-сайти, що є безпечними і підтримуваними. У статті розглянуто проблему оцінювання розміру веб-застосунків, що створюються мовою Python з використанням Django Rest Framework. Для досягнення мети роботи, а саме підвищення достовірності оцінювання розміру веб-застосунків, що створюються мовою Python з використанням Django Rest Framework, була побудована двофакторна нелінійна регресійна модель. Було відібрано дані з 71 проекту з відкритим вихідним кодом, проведено аналіз на викиди та використано десятковий логарифм в якості нормалізуючого перетворення для побудови лінійної регресійної моделі. На основі нормалізованих даних та зворотного нормалізуючого перетворення побудовано нелінійну регресійну модель для вихідного негаусівського набору даних (кількості рядків коду, кількості класів та кількості методів). Проведені перевірка якості та порівняння з лінійною двофакторною моделлю показали, що за всіма параметрами якості побудована двофакторна нелінійна регресійна модель є кращою.

Ключові слова: нелінійна регресійна модель, логарифмічне перетворення, веб-застосунки, Python, Django Rest Framework.

Постановка проблеми. Одним з найвагоміших факторів в управлінні процесом розробки ПЗ є розмір програми. Від ефективності оцінювання розміру програм залежить успіх або невдача проекту на ранньому етапі розробки, тому проблема отримання ефективної системи оцінювання кількості строк коду – це важливе завдання, що вимагає удосконалення існуючих методів [1].

Важливим етапом аналізу складної системи є побудова математичної моделі, що здійснюється на підставі змістовної моделі і складається з вихідних, внутрішніх та результуючих параметрів. Однією з умов побудови якісної адекватної моделі є надійність вхідних даних, тому потрібна первинна обробка експериментальних даних із метою виявлення ненадійних вимірів [2].

Виявлення аномальних значень відноситься до проблеми пошуку зразків у даних, які не відповідають очікуваній поведінці.

Під час проведення регресійного аналізу дослідники, зазвичай, віддають перевагу побудові лінійних моделей, які легко оцінюються, і результати яких чітко інтерпретуються. Однак дані рідко носять абсолютно лінійний характер, тому для їх адекватного опису доводиться вдаватися до побудови нелінійних залежностей [3].

TIOBE Software представила рейтинг найпопулярніших мов програмування на серпень 2022 року [4]. Порівняно з минулим роком Python додав до популярності 3,56%, перемістившись з другого на перше місце з показником 15,42%. Це найвищий показник популярності цієї мови програмування за час існування рейтингу. Найнижчий був зафіксований у 2003 році (0,97%), коли Python посідав 13 місце в рейтингу [5].

Таким чином, побудова нелінійних регресійних моделей із використанням нормалізуючих перетворень для підвищення достовірності оцінювання розміру веб-застосунків, що створюються мовою Python з використанням Django Rest Framework, і створення на їх основі інформаційної технології переробки інформації є актуальною та має практичну цінність.

Аналіз останніх досліджень і публікацій. Дослідження щодо розробки ПЗ та оцінювання вартості були численними та різноманітними. Тема все ще дуже актуальна, про що свідчать численні роботи, наявні в літературі [6–8]. Дослідники ретельно досліджували цю тему як з точки зору підходів до оцінювання (регресія, аналогія, експертне судження, функціональні точки, моделювання тощо), так і дослідницького підходу (теорія, опитування, експеримент, тематичне дослідження, моделювання тощо). Дослідження проводяться як у промисловому, так і в академічному контексті. Найбільш часто використовується підхід до оцінювання, заснований на регресії [9].

Оцінювання розміру ПЗ розглядається як фундаментальна діяльність з точки зору завдань управління розробкою ПЗ. Планування робіт і подальше оцінювання часу та зусиль прогнозуються на основі розміру ПЗ. Кількість рядків коду (LOC) і функціональність ПЗ є двома показниками, які часто використовуються для визначення розміру програми. LOC – це прямий показник, який легко підрахувати і яким легко маніпулювати. Існує багато дискусій щодо використання LOC в оцінюванні розміру ПЗ. Критики базуються на аргументах, що кількість рядків коду залежить

від мови програмування та вимагає деталей, які можливо буде складно оцінити до завершення роботи над проектом.

Що стосується валідації методів оцінювання, переважна більшість дослідницьких підходів базується на використанні історичних даних. Крім того, більшість досліджень стосується промислового контексту.

Зібрати інформацію про реальні промислові проекти дуже складно. Більшість компаній не передають дані про свої проекти. Технології змінюються настільки швидко, що стало важко створити стандартну модель для оцінювання. Різні компанії використовують різні підходи, характер і розмір проектів різний.

Звужуючи тему до веб-додатків, одним із перших дослідників, які ввели метрику розміру для вимірювання веб-додатків шляхом статистичного аналізу основних характеристик веб-сайтів, був Тім Брей [10]. Спочатку моделі, що використовувалися для оцінювання веб-додатків були такими ж, як ті, що використовувалися для загальних програмних рішень.

Одним із перших науковців, які представили метод, спеціально розроблений для веб-додатків, був Рейфер через метрику WO та модель WEBMO [11]. Ряд дослідницьких робіт з оцінювання як розміру, так і трудомісткості веб-додатків також були проведені Мендесом та його співробітниками [12]. Однак наразі не існує єдиної моделі, здатної адекватно оцінити розмір веб-додатку.

Розробка веб-додатків стає дуже популярною на основі концептуальних моделей, і ці моделі можна коригувати та переглядати в будь-який час у процесі розробки додатків. Прогнозування розміру веб-додатків з високою достовірністю все ще є проблемою.

Постановка завдання. Метою роботи є підвищення достовірності оцінювання розміру веб-застосунків, що створюються мовою Python з використанням Django Rest Framework. Щоб досягнути поставленої мети потрібно вирішити такі завдання: виконати аналіз та порівняння моделей оцінювання розміру ПЗ, що вже існують; зібрати проекти з відкритим вихідним кодом зі створення веб-застосунків мовою Python з використанням Django Rest Framework та обрати ті з них, що можуть бути використані для створення регресійної моделі; визначити необхідні метрики з кожного проекту та перевірити отримані дані на викиди; виконати нормалізацію даних, побудувати лінійне рівняння регресії та довірчий інтервал для нього; на основі нормалізованих даних

виконати побудову нелінійного рівняння регресії та довірчого інтервалу.

Виклад основного матеріалу дослідження. Нормальний (або гаусівський) розподіл заклав основу для багатьох статистичних методів, не останнім з яких є класична лінійна регресія.

Практика перетворення даних, щоб вони відповідали нормальному закону за допомогою певного нормалізуючого перетворення, широко використовується через її простоту, часто ця техніка працює добре як швидке та прагматичне рішення. Спотвореність даних впливає на їх оцінювання, тому дані зі значним відхиленням від нормальності можуть бути більш придатними для аналізу після відповідного перетворення.

Тож було обрано десятковий логарифм в якості нормалізуючого перетворення для побудови регресійної моделі для оцінювання кількості рядків коду веб-застосунків, що створюються мовою Python з використанням Django Rest Framework. У логарифмічному перетворенні кожна змінна x замінюється на $\lg(x)$ з основою 10:

$$z = \lg(x), \quad (1)$$

і зворотним до нього є наступне перетворення:

$$x = 10^z. \quad (2)$$

Багатофакторна лінійна регресійна модель в загальному випадку має вигляд:

$$Z_Y = \hat{Z}_Y + \varepsilon = b_0 + b_1 Z_{X_1} + b_2 Z_{X_2} + \dots + b_k Z_{X_k} + \varepsilon, \quad (3)$$

де ε – це випадкова похибка, розподіл якої підпорядковується нормальному закону та має математичне очікування рівне нулю.

Нелінійна регресійна модель будується за зворотним перетворенням (2) та лінійною регресійною моделлю (3):

$$Y = 10^{\varepsilon + b_0} X_1^{b_1} X_2^{b_2} \cdot \dots \cdot X_k^{b_k}. \quad (4)$$

Для побудови двофакторної нелінійної регресійної моделі для оцінювання розміру веб-застосунків, що створюються мовою Python з використанням Django Rest Framework з репозиторію GitHub [13], було зібрано дані 71 проекту з відкритим кодом.

Процес збору метрик проектів, написаних мовою Python, є досить не простою задачею, оскільки фактично немає такого ПЗ, що надає повну інформацію про проект. Завдяки розширенню до PyCharm Statistic [14] було зібрано такі метрики як: кількість файлів з розширенням .py; кількість рядків коду без коментарів та порожніх рядків; кількість порожніх рядків; кількість рядків з коментарями та загальна кількість рядків. Також мовою Python було додатково написано

скрипт, що підраховує кількість класів та методів в кожному з обраних проектів. З цих даних було прийнято рішення обрати наступні фактори:

– загальна кількість рядків коду проекту в тисячах (kilolines of code, KLOC), в якості значення залежної змінної Y ;

– кількість класів проекту в якості першої незалежної змінної X_1 ;

– кількість методів проекту в якості другої незалежної змінної X_2 .

Перш за все необхідно виконати первинний аналіз даних, а саме: важливо переконатися, що вибірка не містить викидів. Також важливо виконати перевірку предикторів на наявність мультиколінеарності. У регресійній моделі лінійна залежність між двома або більше незалежними змінними демонструє присутність мультиколінеарності, що безперечно є негативним явищем і не дозволяє здійснити оцінювання окремого впливу кожного фактору на залежну змінну, оскільки маємо пов'язаність факторів між собою або високий ступінь кореляції.

Серед майбутніх предикторів в моделі множинної лінійної регресії наявність мультиколінеарності визначимо за коефіцієнтами впливу дисперсії (VIFs). Для лінійної моделі множинної регресії з k -предикторами X_i , $i=1, \dots, k$, VIFs – це діагональні елементи оберненої кореляційної матриці $k \times k$ k -предикторів. Якщо значення VIFs більше за 10 то дані мають проблеми з мультиколінеарністю, а у разі, коли значення VIFs знаходяться у межах від 1 до 5, то мультиколінеарності немає [15]. Виконавши перевірку впевнились, що обрані метрики, а саме: кількість класів X_1 та кількість методів X_2 проекту не мультиколінеарні (відповідні діагональні елементи мають значення 3,1), тому можемо використовувати їх в якості двох незалежних змінних.

Для перевірки даних на викиди було використано квадрат відстані Махаланобіса і для його розрахунку побудовано коваріаційну матрицю згідно з [16, 17]. При $m = 3$, $\alpha = 0,05$ та кількості значень $n = 71$ отримали квантілі розподілів Пірсона та Фішера відповідно $\chi^2 = 7,81$ та $F = 2,74$.

В табл. 1 наведені вихідні дані, нормалізовані дані, а також значення квадрату відстані Махаланобіса, та додатковий TS (Test Statistic) для розрахунку критерію Фішера (фрагмент даних).

Отримані викиди у кількості 9 видаляємо із первинного набору даних і повторно виконуємо перевірку на викиди для набору даних із 62 проектів, в результаті якої викидів виявлено не було.

Таблиця 1

Фрагмент обробки та дослідження вихідних емпіричних даних

№	Y	X ₁	X ₂	Z _y	Z _{x1}	Z _{x2}	d _i ²	TS для d _i ²
...								
10	0,295	2	7	-0,5302	0,3010	0,8451	14,42	4,60
11	0,355	5	15	-0,4498	0,6990	1,1761	8,24	2,63
12	0,573	25	24	-0,2418	1,3979	1,3802	6,43	2,05
13	28,757	473	1987	1,4587	2,6749	3,2982	8,05	2,57
14	13,389	344	465	1,1267	2,5366	2,6675	1,67	0,53
15	4,977	90	170	0,6970	1,9542	2,2304	0,21	0,07
16	0,926	16	96	-0,0334	1,2041	1,9823	1,38	0,44
17	3,393	67	146	0,5306	1,8261	2,1644	0,12	0,04
18	4,503	313	236	0,6535	2,4955	2,3729	1,38	0,44
19	4,037	87	182	0,6061	1,9395	2,2601	0,08	0,02
20	0,529	12	26	-0,2765	1,0792	1,4150	5,08	1,62
21	6,181	131	216	0,7911	2,1173	2,3345	0,34	0,11
22	0,888	4	31	-0,0516	0,6021	1,4914	4,76	1,52
23	27,301	723	592	1,4362	2,8591	2,7723	3,45	1,10
24	1,227	74	68	0,0888	1,8692	1,8325	2,28	0,73
25	306,187	1974	17285	2,4860	3,2953	4,2377	29,23	9,33
...								

За допомогою методу найменших квадратів було знайдено коефіцієнти лінійного рівняння регресії: $b_0 = -1,6796$, $b_1 = 0,0534$, $b_2 = 0,9571$ та з використанням формули (3) побудовано лінійну регресійну модель для нормалізованих даних:

$$Z_Y = -1,6796 + 0,0534Z_{X_1} + 0,9571Z_{X_2} + \varepsilon.$$

Перевіряючи залишки ε на відповідність нормальному закону розподілу за допомогою критерію згоди χ^2 Пірсона, підтвердили нормальний закон розподілу ($\chi^2_{кр} = 9,49$, $\chi^2 = 0,22$), тобто лінійна регресійна модель була побудована обґрунтовано. Далі за рівнянням (4) можемо перейти до нелінійної регресійної моделі:

$$Y = 10^{\varepsilon - 1,6796} X_1^{0,0534} X_2^{0,9571}.$$

Для нелінійної регресійної моделі були отримані значення коефіцієнту детермінації $R^2 = 0,8936$, середньої величини відносної похибки $MMRE = 0,2342$ і рівню прогнозування $PRED(0,25) = 0,6290$.

Було проведено порівняння з лінійною двофакторною моделлю ($R^2 = 0,8555$, $MMRE = 0,3352$, $PRED(0,25) = 0,4776$), яке показало, що за всіма

параметрами якості двофакторна нелінійна регресійна модель є кращою: за параметром R^2 на 4,45%, за параметром $MMRE$ на 30,13%, за параметром $PRED(0,25)$ на 31,70%.

Висновки. У статті було розглянуто побудову двофакторної нелінійної регресійної моделі для оцінювання розміру веб-застосунків, що створюються мовою Python з використанням Django Rest Framework на основі логарифмічного перетворення, завдяки чому було підвищено достовірність оцінювання розміру веб-застосунків. Отриману модель, що розроблена на основі незалежних змінних, а саме значення кількості класів X_1 та кількості методів X_2 проекту, можна використовувати для виконання прогнозу значень результативного показника Y .

В подальшому планується побудова довірчого інтервалу та інтервалу прогнозування нелінійної регресії, а також розробка програмного забезпечення для прогнозування розміру веб-застосунків, що створюються мовою Python з використанням Django Rest Framework.

Список літератури:

1. Разработка программного обеспечения – Краткое руководство. URL: <https://coderlessons.com/tutorials/akademicheskii/programmnaia-inzheneriia/razrabotka-programmnogo-obespecheniia-kratkoe-rukovodstvo> (дата звернення: 21.09.2022).
2. Павленко П.М. Основы математического моделирования систем и процессов : навч. посіб. Київ : Книжкове вид-во НАУ, 2013. 201 с.
3. Adam Hayes. Nonlinearity. URL: <https://www.investopedia.com/terms/n/nonlinearity.asp> (дата звернення: 21.09.2022).
4. The software quality company TIOBE. TIOBE Index for September 2022. URL: <https://www.tiobe.com/tiobe-index/> (дата звернення: 21.09.2022).

Науковий журнал

**ВЧЕНІ ЗАПИСКИ
ТАВРІЙСЬКОГО НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ
ІМЕНІ В.І. ВЕРНАДСЬКОГО**

Серія: Технічні науки

Том 33 (72) № 6 2022

Коректура • *Н. Славогородська*

Комп'ютерна верстка • *В. Гайдабрус*

Адреса редакції:

Таврійський національний університет імені В.І. Вернадського

м. Київ, вул. Джона Маккейна, 33

Електронна пошта: editor@tech.vernadskyjournals.in.ua

Сторінка журналу: www.tech.vernadskyjournals.in.ua

Формат 60×84/8. Гарнітура Times New Roman.

Папір офсетний. Цифровий друк. Обл.-вид. арк. 27,17. Ум. друк. арк. 34,86. Зам. № 1222/530

Підписано до друку 30.12.2022. Наклад 150 прим.

Видавництво і друкарня – Видавничий дім «Гельветика»

65101, м. Одеса, вул. Інглезі, 6/1

Телефони: +38 (095) 934 48 28, +38 (097) 723 06 08

E-mail: mailbox@helvetica.ua

Свідоцтво суб'єкта видавничої справи

ДК № 7623 від 22.06.2022 р.