

УДК 004.(4+6): 519.876.2

Т. А. Вакалюк, Я. В. Лук'ячук

Державний університет «Житомирська політехніка»

ТЕОРЕТИЧНІ АСПЕКТИ РОЗРОБКИ МОБІЛЬНОГО ДОДАТКА ДЛЯ ПРОГНОЗУВАННЯ ЙМОВІРНОЇ ВАРТОСТІ ЗЕМЕЛЬНОЇ ДІЛЯНКИ

Розглянуто теоретичні аспекти розробки мобільного додатка для прогнозування ймовірної вартості земельної ділянки. Наведено кількісні і якісні методи прогнозування у практичній діяльності. Детально розглянуто метод аналізу часових рядів. Встановлено, що популярним і широко застосовуваним статистичним методом прогнозування часових рядів є модель ARIMA – авторегресивна інтегрована ковзаюча середня. Для реалізації аналізу даних методом часових рядів та для подальшого формування моделі прогнозування, а також інтеграції її в мобільний додаток, було обрано мову програмування Python. Викладено розробку прогнозування із використанням сучасних засобів на прикладі мобільного додатка із виконанням таких завдань: генерація даних для датасету; дослідження методів та алгоритмів; інтелектуальний аналіз даних; проєктування логіки прогнозування; реалізація прогнозування ймовірної вартості земельної ділянки; проєктування та розробка додатка; тестування додатку. Результатом є мобільний додаток, що розроблений із використанням сучасних засобів прогнозування, який за запитом користувача видає результат прогнозування ймовірної грошової вартості земельної ділянки за потреби користувача.

Ключові слова: *прогнозування, методи прогнозування, метод аналізу часових рядів, розробка мобільного додатка.*

Рассмотрены теоретические аспекты разработки мобильного приложения для прогнозирования возможной стоимости земельного участка. Приведены количественные и качественные методы прогнозирования в практической деятельности. Подробно рассмотрен метод анализа временных рядов. Установлено, что популярным и широко применяемым статистическим методом прогнозирования временных рядов является модель ARIMA – авторегрессивная интегрированная скользящая средняя. Для реализации анализа данных методом временных рядов и для дальнейшего формирования модели прогнозирования и интеграции ее в мобильное приложение был избран язык

© Вакалюк Т. А., Лук'ячук Я. В.; 2020.

программирования Python. Изложены разработка и прогнозирование с использованием современных средств на примере мобильного приложения с выполнением следующих задач: генерация данных для датасета; исследования методов и алгоритмов; интеллектуальный анализ данных; проектирование логики прогнозирования; реализация прогнозирования вероятной стоимости земельного участка; проектирование и разработка приложения; тестирование приложения. Результатом является мобильное приложение, разработанное с использованием современных средств прогнозирования, которое по запросу пользователя выдает результат прогнозирования вероятной денежной стоимости земельного участка по требованию пользователя.

Ключевые слова: *прогнозирование, методы прогнозирования, метод анализа временных рядов, разработка мобильного приложения.*

Forecasting the value of land is time-consuming and costly. Automating or optimizing this process saves time and money. The availability of paper documents is not convenient in today's world, it is more appropriate to use electronic services to obtain appropriate services. The article considers the theoretical aspects of developing a mobile application for forecasting the probable value of the land. Quantitative and qualitative forecasting methods in practice are given. Forecasting using time series analysis requires sufficient information in the past to predict the future state. The method of time series analysis is considered in detail. It is established that the popular and widely used statistical method of time series prediction is the ARIMA model – autoregressive integrated moving average. To implement the analysis of data by the method of time strings and to further form a model of forecasting and its integration into a mobile application, the Python programming language was chosen. After analyzing the data, forming time series, their seasonality, and forecasting, using the ARIMA model, you can proceed to convert the forecasting data model into the Core ML model, for further use (integration) in the mobile application to predict the probable value of the land. The article describes the development of forecasting using modern tools on the example of a mobile application with the following tasks: data generation for the date set; research of methods and algorithms; data mining; designing forecasting logic; implementation of forecasting the probable value of the land plot; application design and development; application testing. The result is a mobile application developed using modern forecasting tools, which at the request of the user gives the result of forecasting the probable monetary value of the land for the needs of the user.

Keywords: *forecasting, forecasting methods, time series analysis method, mobile application development.*

Постановка проблеми. Прогнозування – це метод, в якому

використовуються як накопичений в минулому досвід, так і поточні припущення стосовно майбутнього з метою його визначення. В результаті отримують картину майбутнього, яку можна використовувати як основу в процесі планування. Неможливість знання актуальної грошової вартості свого нерухомого майна – земельної ділянки, не дає впевненості та рішучості у прийнятті рішення. Знання про ймовірну вартість дає розуміння у прийнятті рішення, адже нестабільність ринку нерухомості та землі можуть як підвищити вартість, так і її знизити у дуже короткі терміни часу.

Аналіз останніх досліджень і публікацій. Прогнозні моделі знаходять своє застосування як у техніці, так і в економіці. Зокрема, можна виділити дослідження В. Волохатої [4], О. Дзєндзелюка [5], О. Любіч [2], В. Мартинюка [6], Є. Реуцького [1], Н. Рилової [3]. Питанням методики застосування прогнозного аналізу приділили увагу багато науковців, серед яких можна виділити внесок Ю. Бажала, Н. Громової, В. Мікловди, Л. Федулової та інших.

Виділення не вирішених раніше частин загальної проблеми. Прогнозування вартості землі займає багато часу і потребує грошових витрат. Автоматизація чи оптимізація цього процесу економить час та кошти громадян. Операції з паперовими документами є не дуже зручними у сучасному світі, більш доцільно використовувати електронні сервіси отримання відповідних послуг.

Метою дослідження є розглянути теоретичні аспекти розробки мобільного додатка для прогнозування ймовірної вартості земельної ділянки.

Виклад основного матеріалу. На сьогодні наявна велика кількість методів прогнозування. Для прогнозування у практичній діяльності застосовуються різні кількісні і якісні методи.

Якісні методи зазвичай базуються на інтуїції, досвіді, експертних оцінках спеціалістів у галузі прийняття рішень, їх використовують за відсутності числових даних або коли отримати їх дуже складно, зокрема дуже витратно. Серед якісних методів виділяються такі: думка «журі», спільна думка працівників збуту, модель очікування споживача ґрунтується на результатах опитувань клієнтів організації, метод експертних оцінок.

Кількісні методи базуються на даних, які впливають з тенденцій зміни певних параметрів, або на статистично достовірних залежностях. Серед кількісних методів прогнозування виділяють такі: аналіз часових рядів, методи Бокса – Дженкінса, каузальні методи прогнозування.

Для прогнозування за допомогою методу аналізу часових рядів потрібно мати достатню кількість даних у минулому. На підставі попередніх значень часових рядів можна спрогнозувати тенденції в економіці та погоді або спланувати пропускну здатність. Стаціонарність, сезонність та автокореляція – деякі властивості часового ряду. Сезонність стосується явища коливань у конкретні часові рамки [7].

Автокореляція означає співвідношення поточного значення з копією попереднього часу (відставання). З огляду на особливості властивостей даних часових рядів для роботи з ними застосовуються спеціалізовані статистичні методи і підходи, які реалізовані та представлені в різних бібліотеках.

Часові ряди – це послідовність спостережень, записаних через однакові часові інтервали. Залежно від частоти спостережень часові ряди, як правило, можуть бути погодинними, щоденними, щотижневими, щомісячними, щоквартальними та річними. Іноді можуть бути також секундні та хвилинні часові ряди, наприклад, кількість відвідувань користувачів щохвилини тощо.

Прогнозування часових рядів має величезне комерційне значення, оскільки такі важливі для бізнесу речі, як попит та продаж, кількість відвідувачів веб-сайту, ціна акцій тощо – це, по суті, дані часових рядів. Аналіз часових рядів передбачає розуміння різних аспектів невід'ємної природи ряду для створення значущих і точних прогнозів [8].

Популярним і широко застосовуваним статистичним методом прогнозування часових рядів є модель ARIMA – авторегресивна інтегрована ковзаюча середня. Це клас моделі, який фіксує набір різних стандартних часових структур у даних часових рядів.

Кожен з цих компонентів явно вказаний у моделі як параметр. Для ARIMA(p, d, q) використовується стандартне позначення, де параметри заміщені цілими значеннями для швидкої вказівки конкретної моделі ARIMA, що використовується. Разом ці три параметри враховують сезонність, тенденцію та шум у наборах даних [5].

Параметри моделі ARIMA визначаються таким чином:

- p – порядок авторегресії (AR), який дозволяє додати попередні значення часового ряду. Цей параметр можна проілюструвати твердженням «завтра, ймовірно, буде тепло, якщо в останні три дні було тепло»;

• d – порядок інтегрування (I ; порядок різниць вихідного часового ряду). Він додає в модель поняття різниці часових рядів (визначає кількість минулих часових точок, які потрібно вирахувати з поточного значення). Цей параметр ілюструє таке твердження: «завтра, ймовірно, буде така сама температура, якщо різниця в температурі за останні три дні була дуже мала» [13];

• q – порядок змінного середнього (МА), який дозволяє встановити похибку моделі як лінійну комбінацію значень помилок, що спостерігалися раніше.

Для відстеження сезонності використовується сезонна модель ARIMA – ARIMA(p, d, q) (P, D, Q) s . Тут (p, d, q) – несезонні параметри, описані вище, а (P, D, Q) слідують тим самим визначенням, але застосовуються до сезонної складової часового ряду. Параметр s визначає періодичність часового ряду (4 – квартальні періоди, 12 – річні періоди і т.д.) [13].

AR(p) – авторегресивна модель по ряду p .

Модель має вигляд:

$$Y(t) = f_0 + f_1 Y(t-1) + f_2 Y(t-2) + \dots + f_p Y(t-p) + E(t), \quad (1)$$

де $Y(t)$ – залежна змінна у момент часу t ; $f_0, f_1, f_2, \dots, f_p$ – оцінювані параметри; $E(t)$ – помилка від впливу змінних, які не враховуються в даній моделі.

Завдання полягає в тому, щоб визначити $f_0, f_1, f_2, \dots, f_p$. Їх можна оцінити різними способами. Найправильніше шукати їх через систему рівнянь Юла – Уолкера, для складання цієї системи буде потрібен розрахунок значень автокореляційної функції. Можна застосувати простіший спосіб – порахувати їх методом найменших квадратів [9].

Розглянемо авторегресійне ковзне середнє ARMA(p, q). Під позначенням ARMA(p, q) розуміється модель p авторегресійних складових, що містить q ковзаючих середніх. Точніше, модель ARMA(p, q) включає моделі AR(p) і MA(q):

$$X_t = c + e_t + \sum_{i=1}^q \theta_i e_{t-i} + \sum_{i=1}^p \phi_i X_{t-i}. \quad (2)$$

Зазвичай, значення помилки e_t вважають незалежними однаково розподіленими випадковими величинами, узятими з нормального розподілу з нульовим середнім:

$$e_t \sim N(0, \sigma^2), \quad (3)$$

де σ^2 — дисперсія [9].

Припущення можна ослабити, але це може привести до зміни властивостей моделі. Наприклад, якщо не припускати незалежності і 20

однакового розподілу помилок, поведінка моделі суттєво зміниться. Лнійна модель регресії будується, включаючи вказану кількість та тип термінів, і дані готуються за ступенем диференціації для того, щоб зробити її нерухомою, тобто видалити трендові та сезонні структури, що негативно впливають на регресійну модель.

Для параметра можна застосовувати значення 0, яке вказує на те, що не використовувати цей елемент моделі. Таким чином, модель ARIMA може бути налаштована для виконання функції моделі ARMA і навіть простої моделі AR, I або MA. Прийняття моделі ARIMA для часового ряду передбачає, що базовий процес, що породив спостереження, є процесом ARIMA. Це може здатися очевидним, але допомагає мотивувати необхідність підтвердження припущень моделі в необгрунтованих спостереженнях та залишкових помилках прогнозів з моделі.

Для реалізації прогнозування потрібні дані, для генерування яких використовуємо онлайн сервіс <https://extendsclass.com/csv-generator.html>. Результатом роботи даного сервісу є файл із генерованими даними, де знаходиться 1000 записів, які у часовому розрізі являють собою останні п'ять років. Класифікація даних вміщує в собі коефіцієнт земельної ділянки, що є двох видів: коефіцієнт призначення; коефіцієнт розміщення (локації). Кожен із коефіцієнтів має свій діапазон варіантів значення. Так, для коефіцієнта локації він становить від 1 до 10, а для коефіцієнта призначення – від 1 до 27.

Для реалізації аналізу даних методом часових рядів та для подальшого формування моделі прогнозування та інтеграції її в мобільний додаток було обрано мову програмування Python – це високорівнева, мультипарадигмова мова програмування загального призначення, орієнтована на підвищення продуктивності розробників та читабельності коду.

Для аналізу та візуалізації даних у Python були використані такі бібліотеки: NumPy, Matplotlib, SciPy, Pandas. Також за необхідності були використані бібліотеки seaborn, statsmodels, itertools, coremltools, keras. Спочатку за допомогою методів бібліотеки Pandas було завантажено датасет та виведено його (рис. 1). Для подальшого аналізу даних та Time Series візуалізації потрібно попередньо обробити дані.

Для майбутнього порівняння та аналізу дані будемо аналізувати за параметром коефіцієнта призначення, відфільтруємо за ним дані та залишимо ключові поля для time series – дату та значення. Обрано було коефіцієнт – 9, тому що за аналізом датасету це є найпопулярніший коефіцієнт призначення земельної ділянки. Для його визначення було

використано метод count [12].

	id	nameland	dateBuy	goalKof	localKof	vlozuvMoneyMetr	priceOneMetrBuy	area	totalBuyPrice	vlozuvTotal	landCurentPriceTotal
0	100	Ileana	2019-12-04	7	2	954	1797	448	805056	427392	1232448
1	101	Magdalena	2018-04-22	25	10	854	5319	822	4372218	701988	5074206
2	102	Aigneis	2018-01-31	11	5	166	4424	628	2778272	104248	2882520
3	103	Mamie	2018-10-25	9	10	270	7223	471	3402033	127170	3529203
4	104	Leona	2016-03-30	21	7	244	6202	888	5507376	216672	5724048
5	105	Dacia	2016-10-24	3	7	911	2170	438	950460	399018	1349478
6	106	Linzy	2015-08-25	21	7	376	2089	532	1111348	200032	1311380
7	107	Katharina	2018-11-19	19	3	797	9812	644	6318928	513268	6832196
8	108	Priscilla	2017-07-10	19	9	941	9542	458	4370236	430978	4801214
9	109	Cordi	2018-11-07	2	10	886	9681	998	9661638	884228	10545866

Рисунок 1 – Дані датасету

Для коригування даних використано методи бібліотеки Pandas, такі як loc та drop. Для спрощення створення графіка краще мати індекс у форматі data. Тепер можна переглянути графік цього часового ряду (рис. 2).

Часові ряди мають очевидну сезонність, а також загальну тенденцію до зростання. В результаті можемо візуалізувати дані за допомогою методу, який називається seasonal_decompose – розкладанням часових рядів. Як випливає з назви, декомпозиція часових рядів дозволяє розкласти наявний часовий ряд на три різні компоненти: тенденцію, сезонність та шум. На рис. 3 зображено, як виглядає сезонне розкладання часових рядів у Python.

Головне при підборі даних часових рядів у сезонній моделі ARIMA – знайти значення ARIMA(p, d, q) (P, D, Q) s, які оптимізують необхідний показник.

Для кожної комбінації параметрів функція SARIMAX() з модуля statsmodels може підібрати нову сезонну модель ARIMA і оцінити її загальну якість. Оптимальним набором параметрів буде той, в якому потрібні критерії найбільш продуктивні. Для початку потрібно згенерувати різні комбінації параметрів [13].

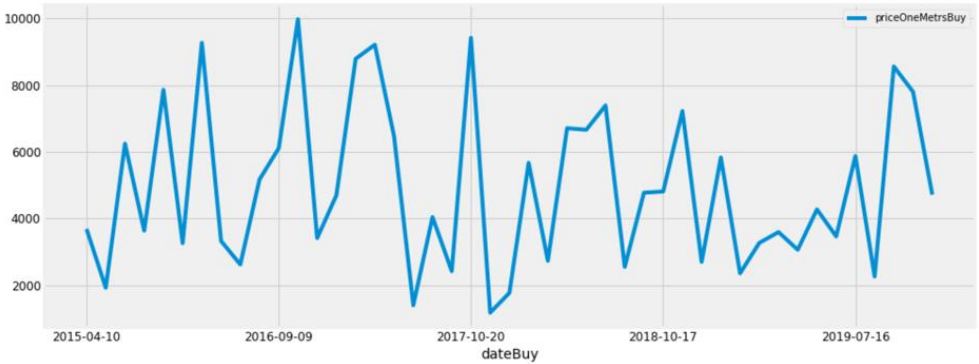


Рисунок 2 – Графік часового ряду

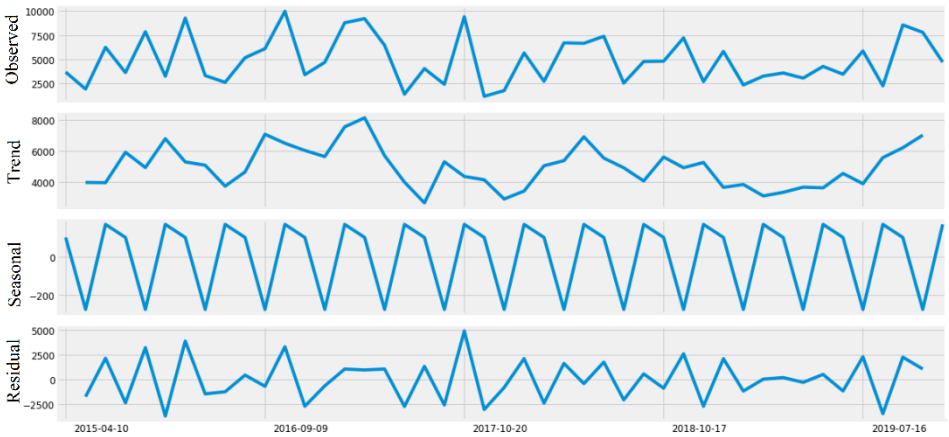


Рисунок 3 – Розкладання часових рядів

Після цього можна використовувати певні триплети параметрів для автоматизації процесу оцінки моделей ARIMA по різних комбінаціях. При оцінці і порівнянні статистичних моделей, що відповідають різним параметрам, враховується, наскільки та чи інша модель відповідає даним і наскільки точно вона здатна прогнозувати майбутні точки даних. При цьому потрібно використовувати значення AIC (Akaike Information Criterion), яке підходить для роботи з моделями ARIMA на основі statsmodels. AIC оцінює, наскільки добре модель

відповідає даним, беручи до уваги загальну складність моделі. Чим менше функцій використовує модель щоб досягти відповідності даним, тим вище її показник AIC. Тому потрібно знайти модель з найменшим значенням AIC [13].

Далі проводимо ітерацію комбінації параметрів і використовуємо функцію SARIMAX із statsmodels, щоб знайти відповідності із сезонною моделлю ARIMA. Аргумент order визначає параметри (p, d, q), а аргумент seasonal_order вказує сезонний компонент моделі ARIMA (P, D, Q, S). Після перевірки кожної моделі SARIMAX() на екрані з'явиться рейтинг значень AIC.

```
ARIMA(1, 0, 1)x(0, 1, 0, 12)12 - AIC:599.006930031764
ARIMA(1, 0, 1)x(1, 0, 0, 12)12 - AIC:600.9866164364704
ARIMA(1, 0, 1)x(1, 0, 1, 12)12 - AIC:585.8545630100073
ARIMA(1, 0, 1)x(1, 1, 0, 12)12 - AIC:380.9564902449101
ARIMA(1, 1, 0)x(0, 0, 0, 12)12 - AIC:814.4694154586289
ARIMA(1, 1, 0)x(0, 0, 1, 12)12 - AIC:1128.9025224918423
ARIMA(1, 1, 0)x(0, 1, 0, 12)12 - AIC:610.3775539555655
ARIMA(1, 1, 0)x(1, 0, 0, 12)12 - AIC:589.4862497127684
ARIMA(1, 1, 0)x(1, 0, 1, 12)12 - AIC:1175.9845180189166
ARIMA(1, 1, 0)x(1, 1, 0, 12)12 - AIC:365.15019210060865
ARIMA(1, 1, 1)x(0, 0, 0, 12)12 - AIC:784.8459254482028
ARIMA(1, 1, 1)x(0, 0, 1, 12)12 - AIC:2802.1185567014113
ARIMA(1, 1, 1)x(0, 1, 0, 12)12 - AIC:587.4054815626297
ARIMA(1, 1, 1)x(1, 0, 0, 12)12 - AIC:581.2246701184899
ARIMA(1, 1, 1)x(1, 0, 1, 12)12 - AIC:2529.952510226249
ARIMA(1, 1, 1)x(1, 1, 0, 12)12 - AIC:365.1575468939566
```

Рисунок 4 – Ітерування комбінацій параметрів

Відповідно до отриманого висновку, SARIMAX(1,1,1)x(1,1,0,12) отримує найменший показник AIC (365,15). Отже, ці параметри можна вважати оптимальними. Після цього визначаємо модель часових рядів ARIMA. Використовуючи пошук по сітці, визначаємо оптимальний набір параметрів для сезонної моделі даних часового ряду. Цю модель можна проаналізувати більш детально. Додаємо оптимальні параметри в модель SARIMAX.

Для отримання даних про цю модель скористаємось методом summary (рис. 5). Атрибут summary повертає багато даних, але зосередимо увагу на таблиці коефіцієнтів. Столпчик coef визначає важливість кожного параметра і його вплив на часовий ряд. Столпчик P> / z / повідомляє значимість кожного параметра.

При підборі сезонних моделей ARIMA важливо проводити діагностику моделі, щоб переконатися, що жодне з припущень, зроблених моделлю, не було порушено. Об'єкт plot_diagnostics дозволяє швидко провести діагностику моделі і досліджувати будь-яку незвичайну поведінку. Головне завдання – переконатися, що залишки

моделі не корельовані і розподіляються з нульовим середнім значенням.

У даному випадку діагностика показала, що залишки моделі правильно розподіляються.

На верхньому правому графіку червона лінія KDE знаходиться близько до лінії $N(0,1)$ (де $N(0,1)$ є стандартним позначенням нормального розподілу із середнім 0 і стандартним відхиленням 1).

Це ознака того, що залишки нормально розподілені.

```

=====
                        Statespace Model Results
=====
Dep. Variable:                priceOneMetrBuy      No. Observations:      45
Model:                        SARIMAX(1, 1, 1)x(1, 1, 0, 12)  Log Likelihood         -178.579
Date:                          Sat, 16 May 2020      AIC                    365.158
Time:                          21:26:56          BIC                    368.935
Sample:                          0              HQIC                   365.797
                                - 45
Covariance Type:                opg
=====
                                coef      std err      z      P>|z|      [0.025      0.975]
-----
ar.L1                          -0.3497    0.351      -0.996    0.319      -1.038      0.339
ma.L1                          -0.8756    0.127      -6.882    0.000      -1.125     -0.626
ar.S.L12                       -0.2709    0.226      -1.199    0.230      -0.714      0.172
sigma2                          8.384e+06  5.84e-09   1.44e+15  0.000      8.38e+06   8.38e+06
=====
Ljung-Box (Q):                  nan      Jarque-Bera (JB):      0.99
Prob(Q):                        nan      Prob(JB):              0.61
Heteroskedasticity (H):        1.38    Skew:                  0.01
Prob(H) (two-sided):           0.71    Kurtosis:              1.88
=====

```

Рисунок 5 – Результат методу summary

На рис. 6 графік q-q в лівому нижньому кутку показує, що впорядований розподіл залишків (сині точки) мають лінійний тренд вибірок, взятих зі стандартного розподілу $N(0, 1)$. Знову ж таки, це ознака того, що залишки нормально розподілені.

Залишки з плином часу (верхній лівий графік) не показують явну сезонність і здаються білими шумами. Це підтверджується графіком автокореляції (внизу праворуч), який показує, що залишки часових рядів мають низьку кореляцію із запізнілими даними.

Ці графіки дозволяють зробити висновок про те, що обрана модель підходить для аналізу і прогнозування даних часових рядів. Тепер є модель часових рядів, за допомогою якої можна спрогнозувати дані. Для початку потрібно порівняти прогнозовані значення з реальними значеннями часового ряду, що допоможе зрозуміти точність прогнозів. Атрибути `get_prediction()` і `conf_int()` дозволяють отримувати значення та інтервали для прогнозів часових рядів.

Прогнозування почнеться із 2015 року. Аргумент `dynamic = False` включає покрокове прогнозування, а це означає, що прогнози в кожній точці генеруються з використанням повної історії аж до цієї точки.

Загалом прогнози відповідають істинним значенням, демонструючи загальний тренд (рис. 7). Тепер можна використовувати модель ARIMA для прогнозування майбутніх значень. Атрибут `get_forecast()` об'єкта часового ряду може обчислити значення на вказану кількість кроків.

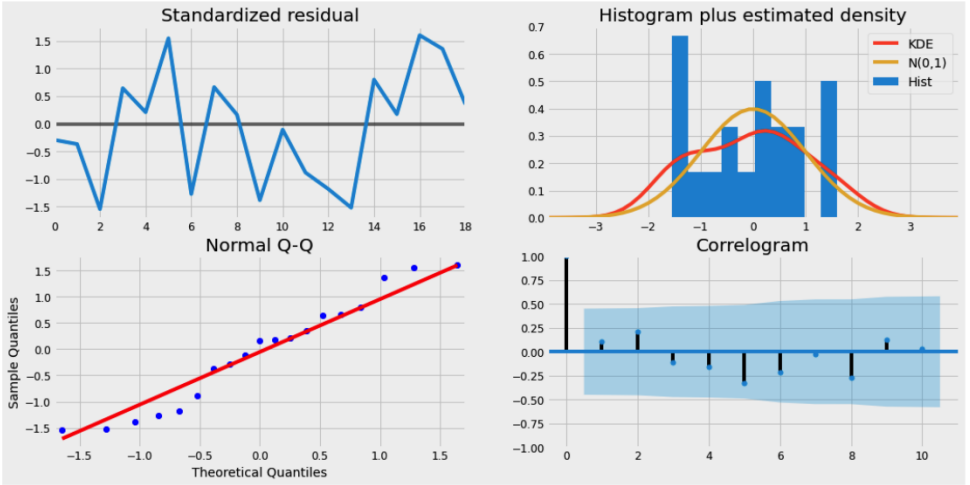


Рисунок 6 – Графік діагностики моделей

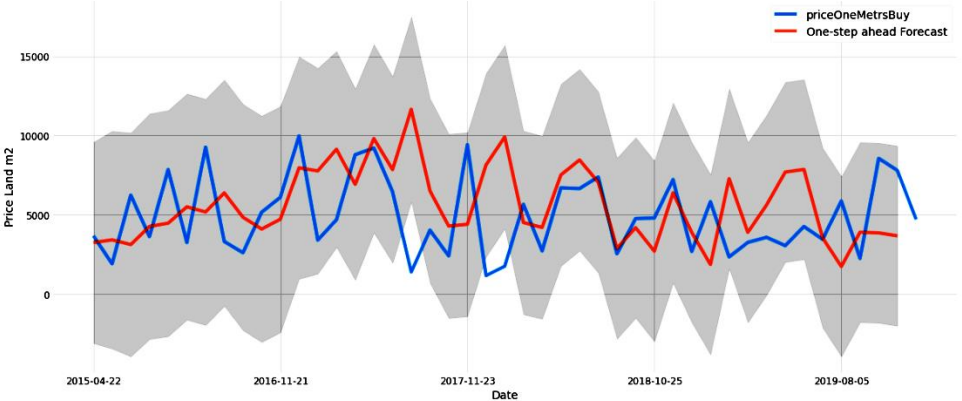


Рисунок 7 – Графік прогнозованих значень у порівнянні з реальними значеннями

Отримані дані (рис. 8) показують, що часові ряди мають свою характерну тенденцію росту та спаду. Це дає розуміння певного шаблону коливання ціни на земельну ділянку та вказує на періодичне продовжуване стабільне зростання, що повторюватиметься час від часу.

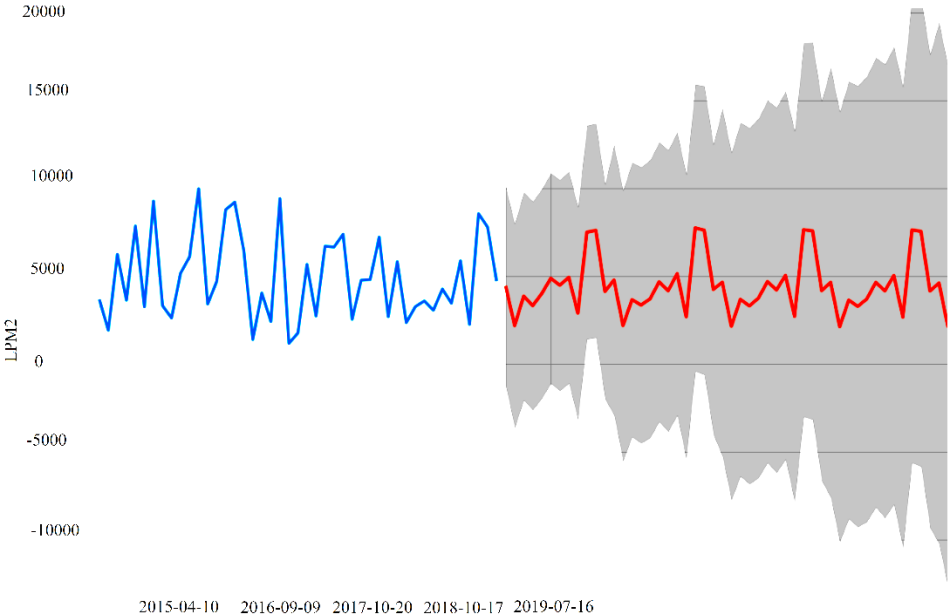


Рисунок 8 – Графік прогнозування

Після виконання аналізу даних, формування часових рядів, їх сезонності та прогнозування, використовуючи модель ARIMA, можна приступити до конвертації моделі даних прогнозування в модель Core ML, для подальшого її використання (інтеграції) в мобільному додатку для прогнозування ймовірної вартості земельної ділянки.

Core ML – це платформа Apple для інтеграції моделей машинного навчання в певну програму. Пакет core ml tools Python – це основний спосіб конвертувати сторонні моделі у формат Core ML. На рис. 9 зображено схему взаємодії, роботи, інтеграції Core ML [11].



Рисунок 9 – Схема інтеграції Core ML

Для початку було проведено тренування моделі за параметрами датасету, які будуть братися як вхідні дані для прогнозування в мобільному додатку. А саме це параметри: коефіцієнт призначення, локальний коефіцієнт, площа земельної ділянки, вартість оціненої землі.

Засобами бібліотеки `coremltools`, а саме функціями `converters.sklearn.convert`, модель було конвертовано. Потім задано назву, опис вхідних полів даних, опис моделі. В завершенні було збережено в окремий файл із розширенням `*.mlmodel`. Вже збережений файл було додано до проекту розробки мобільного додатка.

У проєкті модель була інтегрована, також був доданий раніше створений датасет, який був проаналізований та розібраний на складові властивості об'єкта: ключ – значення, отримані результати розбору беруться як вхідні дані для прогнозування. Дані були представлені у вигляді View екранів, де можна переглянути всі земельні ділянки та дані про них. У результаті отримано мобільний додаток для отримання даних ймовірної грошової вартості земельної ділянки.

Приклад роботи мобільного додатка прогнозування ймовірної вартості земельної ділянки представлено на рис. 10а. У додатку вже наявні дані про земельну ділянку. Перейшовши за обраною земельною ділянкою, можна переглянути всі відомості про неї, натиснувши на кнопку «Деталі» буде представлено інформацію про розрахунки, які були сформовані під час формування генерації даних датасету (рис. 10б). За допомогою функції «Спрогнозувати» можна отримати інформацію про ймовірну прогнозовану вартість земельної ділянки (див. рис. 10б та 10в).

Результатом роботи моделі, що була інтегрована в мобільний додаток, є ймовірна вартість земельної ділянки за 1 м². На рис. 11а зображено екран мобільного додатка, де відбувається прогнозування. Приклад виконання роботи моделі в додатку представлено на рис. 11б.

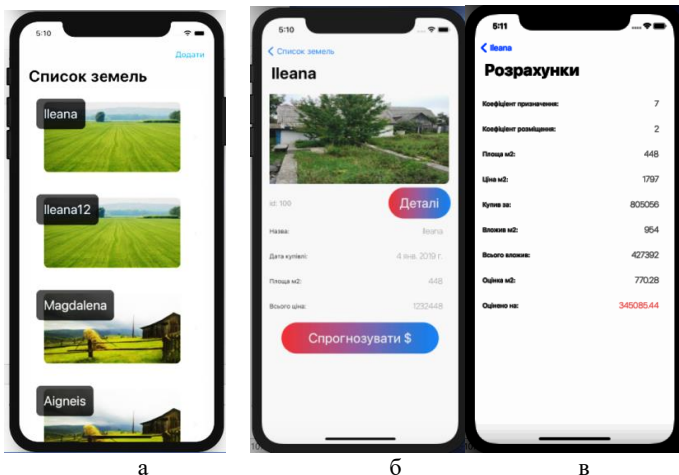


Рисунок 10 – Мобільний додаток

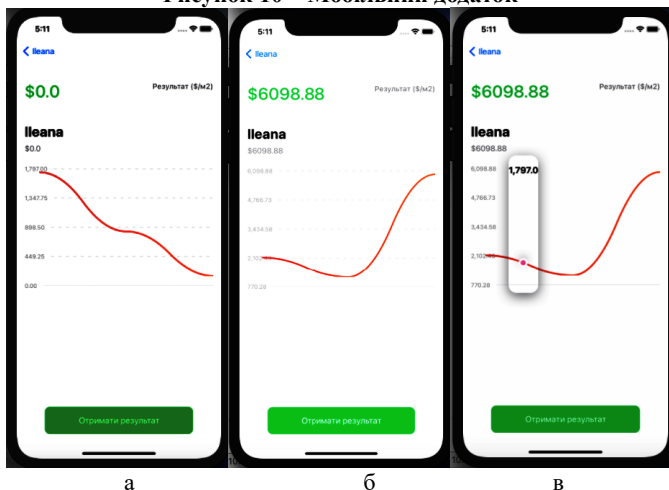


Рисунок 11 – Екран прогнозування

Натиснувши кнопку «Спрогнозувати» відкривається екран прогнозування, на якому потрібно натиснути кнопку «Отримати результат», після чого буде отримано результат імовірної вартості земельної ділянки, який представлено за 1 м² у доларах.

Функціонально додаток має можливість при натискуванні на графік результату побачити початкове значення – вартість, за яку була куплена

земельна ділянка, і друге (середнє) значення – вартість оцінювання земельної ділянки. Останнім значенням є ймовірна грошова вартість земельної ділянки, яка прогнозована за допомогою моделі. На графіку можна побачити тенденцію грошової вартості земельної ділянки, що в результаті дає розуміння про продаж земельної ділянки – прийняття відповідного рішення.

Порівнюючи прогнозовані дані моделі ARIMA та COREML, бачимо, що тенденція графіків однакова (рис. 11б та 11в). Це свідчить про те, що реалізація прогнозування часовими рядами та її інтеграція пройшли успішно.

Висновки. Результатом є мобільний додаток, що розроблений із використанням сучасних засобів прогнозування, який за запитом користувача видає результат прогнозування ймовірної грошової вартості земельної ділянки за потреби користувача.

У статті викладено розробку прогнозування із використанням сучасних засобів на прикладі мобільного додатка із виконанням таких завдань: генерація даних для датасету; дослідження методів та алгоритмів; інтелектуальний аналіз даних; проектування логіки прогнозування; реалізація прогнозування ймовірної вартості земельної ділянки; проектування та розробка додатка; тестування додатка.

Даний мобільний додаток був розроблений із використанням сучасних засобів прогнозування. В даній розробці використовувались дані, які вже містять оціночну вартість земельної ділянки. На основі цього показника та низки інших реалізується прогнозування беручи їх як вхідні дані конкретного об'єкта ймовірної грошової вартості земельної ділянки.

Бібліографічні посилання

1. Реуцький Є. Метод прогнозування характеристик точності вимірювальних каналів енергетичних систем. *Енергетика і автоматика*. 2016. № 3. С. 171–181.
2. Любіч О., Домрачев В. Застосування моделі ARIMA для прогнозування доходів зведеного бюджету України. *Екон.-мат. моделювання соц.-екон. систем: зб. наук. пр.* Вип. 16. Київ: МННЦІТС НАН та МОН України, 2011. С. 33–37.
3. Рилова Н., Оксаніч І. Синтез ARIMA-моделей для прогнозування коефіцієнтів виходу кондиційних напівпровідникових матеріалів. *Системи обробки інформації*. 2015. Вип. 5 (130). С. 102–107.
4. Волохата В. Управління залученими ресурсами банку на основі цінового прогнозування з використанням мультиплікативних

- моделей ARIMA. *Інвестиції: практика та досвід*. 2014. № 24. С. 72–78.
5. Дзєндзелюк О., Костів Л., Рабик В. Побудова ARIMA-моделей часових рядів для прогнозування метеоданих на мові програмування R. *Електроніка та інформаційні технології*. 2013. Вип. 3. С. 211–219.
 6. Мартинюк В. Прогнозування надходження податкових платежів до державного бюджету за допомогою використання ARIMA-моделі. *Наукові праці НДФІ*. 2011. № 2 (55). С. 46–55.
 7. Основні методи прогнозування. URL: https://pidru4niki.com/16400221/menedzhment/osnovni_metodi_prognozuvannya_klasifikatsiya_prognoziv (дата звернення: 25.11.2020).
 8. Кількісні методи прогнозування. URL: <https://studfile.net/preview/5014249/page:10/> (дата звернення: 25.11.2020).
 9. Методи прогнозування. URL: https://wiki.tntu.edu.ua/%D0%9C%D0%B5%D1%82%D0%BE%D0%B4%D0%B8_%D0%BF%D1%80%D0%BE%D0%B3%D0%BD%D0%BE%D0%B7%D1%83%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F#.D0.A0.D0.B5.D0.B3.D1.80.D0.B5.D1.81.D1.96.D0.B9.D0.BD.D1.96_.D0.BC.D0.B5.D1.82.D0.BE.D0.B4.D0.B8_.D0.BF.D1.80.D0.BE.D0.B3.D0.BD.D0.BE.D0.B7.D1.83.D0.B2.D0.B0.D0.BD.D0.BD.D1.8F (дата звернення: 25.11.2020).
 10. Time Series Analysis In Python – A Comprehensive Guide With Examples ML+. URL: <https://www.machinelearningplus.com/time-series/time-series-analysis-python/> (дата звернення: 25.11.2020).
 11. Apple Developer Documentation. URL: <https://developer.apple.com/documentation/coreml> (дата звернення: 25.11.2020).
 12. ARIMA ModelPythonExample – TimeSeriesForecasting. URL: <https://towardsdatascience.com/machine-learning-part-19-time-series-and-autoregressive-integrated-moving-average-model-arima-c1005347b0d7> (дата звернення: 25.11.2020).
 13. 8host.com. 2020. *8Host.Com*. URL: <https://www.8host.com/blog/prognozirovanie-vremennyh-ryadov-s-pomoshhyu-arima-v-python-3/> (дата звернення: 25.11.2020).

Надійшла до редколегії 27.11.2020.