

СПИСОК ЛІТЕРАТУРИ

1. Павленко Т. В. Економічний механізм управління інтелектуальною власністю промислового підприємства. *Актуальні проблеми економіки*. 2009. № 4(94). С. 124–129.
2. Мрочко Л. В., Пирогов А. И. Информационные потребности и интересы личности: связь и соподчинение общего и частного. *Экономические и социально-гуманитарные исследования*. № 3 (11). 2016. С. 125–129.
3. Укрпошта. URL: <https://www.ukrposhta.ua/ua> (дата звернення: 12.10.2020).
4. Рейтинг розвитку поштової служби у світі. URL: <https://ua-news.liga.net/economics/news/ukraina-pidnyalasya-na-sim-pozitsiy-u-globalnomu-reytingu-rozvitku-poshtovih-poslug> (дата: 12.10.2020).

УДК 004.89:316.7

ІНТЕЛЕКТУАЛЬНА СИСТЕМА АНАЛІЗУ І ПРОГНОЗУВАННЯ РЕЗУЛЬТАТІВ СОЦІОЛОГІЧНИХ ДОСЛІДЖЕНЬ

Р. Ю. Широков, М. О. Єнік

Анотація. Роботу присвячено аналізу і прогнозуванню результатів соціологічних досліджень за допомогою розробленої інтелектуальної системи, розбору основних проблем та методів їх вирішення, що виникають у процесі обробки інформації різних форматів соціологічних даних. Виконано визначення поняття «соціологічних досліджень» та проаналізовано важливість коректної адаптації структури заповнення результатів досліджень. Продемонстровано пріоритетні методи обробки «сирих» соціологічних даних для подальшого їх аналізу та прогнозування. Розроблено інтелектуальну систему на мові програмування Python з інтерфейсом аналізу та прогнозування результатів соціологічних досліджень методами аналітичної екстраполяції, а саме ступеневого поліному, та нейронних мереж.

Ключові слова: соціологічне дослідження, інтелектуальна система, екстраполяція, нейронні мережі.

Вступ. У висвітленні основних стадій практики та теорії соціологічних досліджень основним є взаємозв'язок соціального життя та галузей соціуму, які можуть бути як спорідненими між собою, так і кардинально різними за типами та сутністю специфіки вибраної методології соціологічного дослідження.

Проаналізувавши вислови у світоглядах соціологів ХХ століття, таких, як Талкотт Парсон, Сорокін П. О., необхідно орієнтуватися на розподіл соціології, як предмет теоретичного та практичного характеру, в яких теоретична складова вивчає явища взаємодії людей з точки зору *суцього* (аналітику, генетику та механіку). Практична ж складова розрахована на вивчення явища з погляду *належного* (включає основні положення соціології як соціальної політики).

Методологією називають систему принципів наукового дослідження. Саме методологія визначає, якою мірою зібрані факти можуть служити реальною і надійною основою об'єктивного знання. Тому терміном «методологія» прийнято позначати сукупність дослідницьких процедур, технік і методів, включаючи прийоми збору і обробки даних [1, с. 8; 2].

Проблематикою, у загальному вигляді, можна підкреслити пізнання соціологічних досліджень у співвідношенні теоретичних та емпіричних методів і форм. Розглянувши детальніше ці поняття, можна сформулювати їх визначення: теоретичні методи дослідження є сходженням від абстрактного до конкретного за допомогою формалізації, гіпотетичного методу, аксіоматичного методу та створення. Емпіричні методи розгортаються від спостереження, експерименту, вимірювання та обмеження. Розбивши на цикл можна отримати замкнуте коло чинників: спостереження – індукція – дедукція – перевірка – оцінка.

Метою наукової роботи є аналіз вже існуючих інтелектуальних систем аналізу та прогнозування соціологічних досліджень, розбір понять та сутностей досліджень, як

об'єкта проблематики, методів і форм реалізації функціоналу теоретичної частини, що дозволить створити механізм інтелектуальної системи на практичних прикладах.

Основний розділ. У цілому під терміном соціологічне дослідження прийнято розуміти систему поступових методичних, методологічних і організаційних процесів, направлених на пізнання, аналіз та систематизацію соціологічних фактів, розкриття взаємозв'язків або відношення між різними соціальними випадками та процесами з подальшою його обробкою для досягнення мети управління та контролю над об'єктом, що досліджується, прогнозування потенційності його розвитку в майбутньому [3, с. 23].

Процес дослідження зазвичай починається з теоретичної підготовки або формулювання теми дослідження. Це має на меті встановити основи решти частин дослідження. Соціальне дослідження, незалежно від його типу та характеру, передбачає ступінь вимірювання. Це передбачає категоризацію та / або присвоєння цінностей поняттям, що є різноманітними за своїм характером та рівнем функціонування. Це також дуже корисна процедура, оскільки вона служить для забезпечення високої якості соціальних досліджень. Більш за все, вимірювання проводиться з метою сприяння адекватності, однорідності, порівнянню, послідовності, точності опису та оцінки понять [3, с. 89].

Вимірювання може бути як кількісним, так і якісним. Кількісне вимірювання використовує числові значення та атрибути. Якісне вимірювання – ярлики, назви та якості. Дане вимірювання описує атрибути, використовуючи загальноприйняті поняття та символи. Є можливість ввести нові поняття. Загальна процедура передбачає опис категорій та класифікацій.

Змінні є основними елементами вимірювання, що представляють собою емпіричні конструкції, які приймають більше одного значення або інтенсивності. Наприклад, науки (соціологія, математика, хімія), галузі наук (точні науки, природознавство, суспільні науки), вік та освіта є змінними. Константи є протилежністю змінним. Вони можуть приймати лише одне значення або інтенсивність. На початку дослідження визначаються константи та змінні, які відповідають поняттям системи.

Об'єктивність – це принцип дослідження, метою якого є мінімізація особистих упереджень та гарантування того, що соціальна реальність буде представлена такою, якою вона є. Хоча на питання про те, чи повинен соціальний запит бути об'єктивним чи ні, академічна спільнота не отримала однозначної відповіді. Протягом багатьох років академічні погляди на цю проблему розділялися, і дві лінії думок займали дві крайності аргументу [4, с. 82].

У ході дослідження сучасних методів аналізу та прогнозування даних за допомогою інформаційних технологій виявлено такі основні моделі, які широко застосовуються на практиці для конкретного набору даних (середні оцінки студентів):

- 1) аналітична екстраполяція,
- 2) нейронні мережі,
- 3) стохастичне моделювання,
- 4) кореляційний аналіз.

Найбільш привабливими є перші два методи. Щодо останніх двох можна зазначити наступне.

Стохастичне моделювання застосовується у дуже широкому колі аналітичних задач, проте його зазвичай складно привести до формалізованого вигляду, щоб застосувати для обрахування конкретних значень за допомогою обчислювальної техніки. Тому стохастичне моделювання потребує більш розвинутих технологій для трансформації логічних моделей при формалізації.

Кореляційний аналіз застосовується для прогнозування значень за умов, якщо похибки та погрішності мінімальні, а для динамічного ряду, який розглядається в контексті системи, дуже складно визначити конкретні похибки величин, адже вони вкрай випадкові та погано підлягають класичним методам дослідження функцій.

Аналітична екстраполяція дозволяє лінійно передбачити наступне значення, що цілком дає змогу оцінити напрям руху графіку значень (вгору чи вниз), а тому його можна застосувати для первинного прогнозу динаміки динамічного ряду. На відміну від методу нейронної мережі в методі аналітичної екстраполяції існує значна ймовірність похибки, адже функція розглядається лінійно, проте для реального ряду функція може мати нелінійний вигляд. Крім цього слід зазначити, що неможливо зробити прогноз за допомогою методу аналітичної екстраполяції більш ніж на кілька часових відрізків вперед, бо похибка минулих років накопичується, що формує невизначеність наступних значень [6, с. 174]. Слід зазначити, що для даного динамічного ряду не застосовувалася нормалізація в методі аналітичної екстраполяції, бо вхідні значення належать одному діапазону значень.

У даному випадку для даних динамічного ряду був застосований метод функції степеневого поліному (1) [6, с. 176]:

$$y = a_0 + a_1t + a_2t^2 + \dots + a_nt^n \quad (1)$$

По-перше, даний метод можна масштабувати для досягнення максимальної точності (варіювання лише до поліному третього ступеню, бо кореляційне відношення, що слугує мірою щільності для нелінійних моделей, прямує до одиниці при збільшенні ступеню поліному, що не дозволяє визначити тренд динамічного ряду).

По-друге, вигляд графіку динамічного ряду не схожий на параболічний чи гіперболічний, не схожий на показникову функцію, тому більш доцільно інтерполювати функцію до поліному n -го ступеня, аби він був максимально наближений до вигляду функції динамічного ряду на всьому проміжку.

У розробленій нейронній мережі, що розгорнута у програмі, застосовується 6-ти шарова мережа. Слід окремо зазначити функції активації для кожного шару нейронів: внутрішні заховані шари активуються функцією *ReLU* (2), що дозволяє розрізнити характеристику динамічного ряду більш чітко для даного шару ніж функцію сигмоїди. Функція *ReLU* пояснює модель вибору ймовірності з всіх запропонованих значень даного шару і має вигляд:

$$f(x) = \max(0, x) \quad (2)$$

Сигмоїда (3) застосовується на вихідному шарі для умовного згладжування ймовірностей наприкінці прогнозування.

$$f(x) = \frac{1}{(1 + e^{-x})} \quad (3)$$

Інтелектуальна система аналізу і прогнозування результатів соціологічних досліджень представляє користувачу наступні пункти взаємодій:

- 1) завантаження файлу соціологічних досліджень формату JSON з даними про результати вже відомих досліджень в певному форматі запису;
- 2) перегляд прикладу соціологічних досліджень, які відображаються у головному вікні програми на початку роботи програмного додатку;
- 3) сортування вже обробленої інформації табличного типу
- 4) ознайомлення з результатами аналізу і прогнозування вибраного файлу користувача у вигляді графіків та таблиць, знайденими за методами екстраполяції та нейронних мереж.

Візуальну частину розробленого програмного додатку, а саме інтелектуальної системи, виконує пакет *Tkinter* для мови програмування *TCL*, призначений для роботи з

бібліотекою *Tk*. Сама ж бібліотека *Tk* містить в собі компоненти графічного інтерфейсу користувача, що написані на мові програмування *Tool Command Language*, або *TCL*.

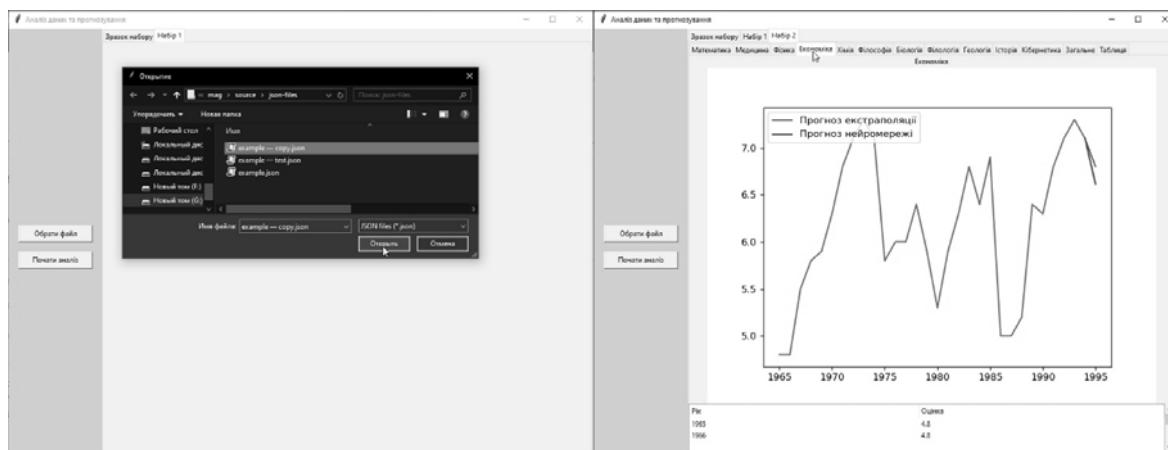


Рис. 1. Візуальна частина ІС

Для машинного навчання було обрано комплексну платформу з відкритим кодом *TensorFlow*. Ця платформа має всеосяжну та гнучку екосистему інструментів, бібліотек, а найголовніше, велику базу клієнтів та користувачів, що дозволяє розробникам використовувати актуальні технології машинного навчання та легко створювати і розгорнути додатки на їх основі.

Структура програмного додатку «Інтелектуальна система аналізу і прогнозування результатів соціологічних досліджень» розділена на файли `app.py` – для візуальної частини програмного додатку, `prediction_models.py` – для реалізації методів аналізу та прогнозування за допомогою нейронних мереж та методу екстраполяції, файлу `TreeViewTable.py` – для коректного відображення інформації типу таблиць.

Висновки. Дослідивши сутність створення та обробки соціологічних досліджень можна зробити висновок, що створення таких інтелектуальних систем можливо за наявності виконання основних вимог теоретичних та емпіричних методів і форм. Розглянуто основоположні твердження та вибрано напрямок розвитку в одному із методів збору «сирих» соціологічних даних для подальшого дослідження: аналізу й прогнозування вже існуючих (підготовлених) соціологічних результатів, а даному випадку, для визначення тенденцій розуміння або звернення уваги на певні критерії в сфері освіти.

Щодо доцільності обраних методів аналізу та прогнозування результатів соціологічних досліджень використано аналітичну екстраполяцію, як один із методів, який можна застосувати в програмуванні універсальним методом прогнозування, точність якого не критично залежить від вигляду вхідних даних, а це в свою чергу можна застосувати для динамічного ряду, який обрано як логічну модель для створеної програми та метод нейронних мереж для аналізу та прогнозування значень для даного динамічного ряду, тому, що метод задовольняє потребам аналізу і прогнозування. Наведено етапи розробки ІС та її основну архітектуру. Розроблена інтелектуальна система використовує названі методи для аналізу та прогнозування результатів соціологічних досліджень створюючи графіки кривих, що визначають напрямок певної категорії даних (зростання, або спад) та проводять аналіз вже існуючих соціологічних даних. По-перше, виконується первинний аналіз динаміки руху графіку динамічного ряду, що дає змогу оцінити зростання або спадання графіку для наступної точки, а також порівняти з результатами метода аналітичної екстраполяції. По-друге, метод обраховує значення виходячи з «тренувальної» моделі, яка є прикладом поведінки динамічного ряду, та яку можна брати в основу для порівняння поточного динамічного ряду з «тренувальним» за допомогою алгоритму багатопарової структури, де кожен нейрон умовно відповідає за конкретну характеристику динамічного ряду в заданій точці.

Аннотация. Работа посвящена анализу и прогнозированию результатов социологических исследований с помощью разработанной интеллектуальной системы, разбор основных проблем и методов их решения, возникающих в процессе обработки информации различных форматов социологических данных. Дано определение понятия «социологических исследований» и проанализированы важность корректной адаптации структуры заполнения результатов исследований. Продемонстрировано приоритетные обработки «сырых» социологических данных для дальнейшего их анализа и прогнозирования. Разработана интеллектуальная система на языке программирования *Python* с интерфейсом результатов анализа и прогнозов результатов социологических исследований методами аналитической экстраполяции, а именно ступенчатого полинома, и нейронных сетей.

Ключевые слова: социологическое исследование, интеллектуальная система, экстраполяция, нейронные сети.

Abstract. The work is devoted to the analysis and forecasting of sociological researches results by means of the intellectual system, analysis of the basic problems and methods of their decision arising in the course of information processing of various formats of sociological data. The definition of the concept of «sociological research» is performed and the importance of correct adaptation of the structure of filling the research result us analyzed. The priority methods of raw sociological data processing for their further analysis and forecasting are demonstrated. The intelligent system in the Python programming language with an interface for analyzing and predicting the results of sociological research using analytical extrapolation, namely a power polynomial and neural networks, has been developed.

Key words: sociological research, intelligent system, extrapolation, neural networks.

СПИСОК ЛІТЕРАТУРИ

1. Ядов В. А. Социологическое исследование: методология, программа, методы: учеб. пособие. Москва : Наука, 2007. 328 с.
2. Кириленко В. І. Методологія сучасної економічної науки: генеза, суперечливості та напрями розвитку. Київ, 2015. URL: <http://securities.usmdi.org/?p=22&n=91&s=936> (дата звернення: 19.03.2019).
3. Sotirios Sarantakos, Social Research: Macmillan International Higher Education, 2012. 536 p.
4. Вербець В. В. Методологія та методика соціологічних досліджень: навч. посіб. Рівне : РДГУ, Інститут соціальних досліджень, 2006. 167 с.
5. Сурмин Ю. П., Туленков Н. В. Методология и методы социологических исследований: учеб. пособие. К.: МАУП, 2000. 304 с.
6. Donald Khosrowi. Extrapolation of causal effects – hopes, assumptions, and the extrapolator’s circle: Published online: 07 Jan 2019. – 45–58 p. – URL: <https://www.tandfonline.com/doi/full/10.1080/1350178X.2018.1561078>