

Моделі і методи прогнозування рівня урбанізації країн світу

Виконав:

Сенюк Костянтин Костянтинович
студент групи КА-54

Науковий керівник:

д.т.н. проф. Зайченко Олена Юріївна

Вступ

○ **Об'єкт дослідження:**

Об'єктом дослідження було обрано статистичні дані щодо соціально-економічних процесів з метою виявлення темпу росту міського населення

○ **Предмет дослідження:**

Методи прогнозування рівня урбанізації за допомогою АРІКС та нейронної мережі Back Propagation

○ **Мета дослідження:**

Побудова прогностичної моделі для здійснення короткострокового прогнозу рівня урбанізації країн світу

Постановка задачі

- Провести дослідження урбанізаційних процесів, визначитися з даними та методами прогнозування
- Реалізувати обрані алгоритми прогнозування та представити результати їх роботи у вигляді графіків та обраних показників
- Порівняти результати прогнозованих даних різними методами з реальними показниками
- Проаналізувати отримані результати, визначити та описати основні переваги, недоліки та особливості досліджуваних моделей в контексті їх застосування в задачах прогнозування

Актуальність роботи

- ❧ Рішення проблем з переселенням міст
- ❧ Представлення результатів моделювання для стабілізування впливу урбанізаційних процесів на економіку тієї чи іншої країни

Вибірка для проведення дослідження

У якості вхідних параметрів для прогнозуванні було обрано наступні порічні показники починаючи з 1960 року і до 2018:

1. Відсоток людей проживаючих в містах для даної країни
2. Кількість усіх міських жителів для певної країни

Джерелом даних є WDI (World Development Indicators).

Блок-схема процесу



Авторегресійна інтегрована модель ковзного середнього

- Генеруємо нормально розподілений стаціонарний ряд незалежних псевдовипадкових чисел $\varepsilon_1 \dots \varepsilon_t$.
- Записуємо систему рівнянь за формулою:
$$\nabla^d(X_t) = a_1 * \nabla^d(X_{t-1}) + a_2 * \nabla^d(X_{t-2}) + \dots + a_p * \nabla^d(X_{t-p}) + \varepsilon_t - b_1 * \varepsilon_{t-1} - b_2 * \varepsilon_{t-2} - \dots - b_q * \varepsilon_{t-q}$$
Де, $\nabla^d(X_t) = \sum_{i=1}^d (X_{t-i+1} - X_{t-i})$
- Коефіцієнти $b_1 \dots b_q, a_1 \dots a_p$ розраховуємо методом МНК.
- Підставляємо в рівняння що було вище, маємо модель для прогнозу.

Метод зворотного поширення помилки

- Ініціалізація вагових коефіцієнтів
- Вхідний нейрон відправляє сигнал нейронам прихованого шару
- Прихований нейрон сумує зважені вхідні сигнали та застосовує функцію активації. Посилає результат всім елементам вихідного шару

$$z_j = f(z_in_j) = f\left(\sum_i v_{ij} * x_i + v_{0j}\right)$$

- Вихідний нейрон сумує зважені вхідні сигнали та рахує вихідний сигнал

$$y_j = f(y_in_k) = f\left(\sum_j v_{jk} * z_j + w_{0k}\right)$$

Метод зворотного поширення помилки

- Вихідний нейрон рахує помилку та зміну вагових коефіцієнтів

$$\delta_k = (t_k - y_k) * f'(y_{in_k}); \quad \Delta w_{jk} = \eta * \delta_k * z_j$$

- Прихований нейрон сумує вхідні помилки, рахує помилку та зміну вагових коефіцієнтів

$$\delta_{in_j} = \sum_k w_{jk} * \delta_k; \quad \delta_{in_j} * f'(z_{in_j}); \quad \Delta v_{ij} = \eta * \delta_k * x_i$$

- Вихідні та приховані нейрони змінюють вагові коефіцієнти
 $w_{jk}(new) = w_{jk}(old) + \Delta w_{jk}; \quad v_{ij}(new) = v_{ij}(old) + \Delta v_{ij}$
- Перевірка умови закінчення роботи алгоритму

Метод зворотного поширення помилки

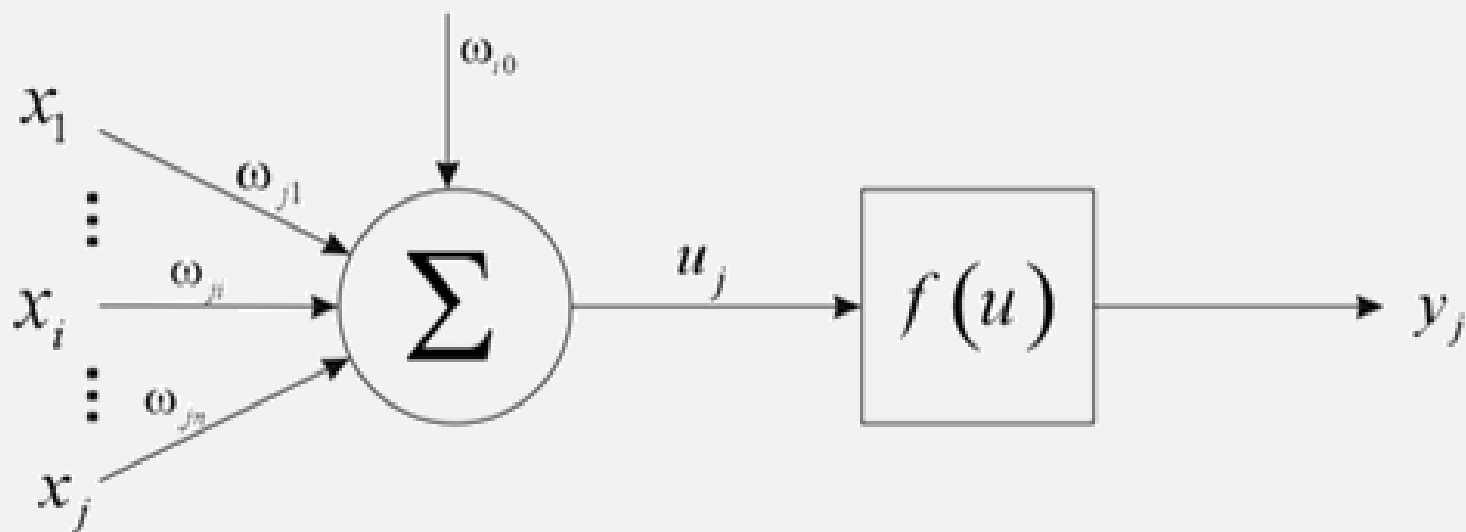


Рисунок 1 – Модель нейрона

Метрики

- Коефіцієнт детермінації

$$R^2 = \frac{D(\hat{y})}{D(y)} = \frac{\sum_{t=1}^n (\hat{y}_t - \bar{y})^2}{\sum_{t=1}^n (y_t - \bar{y})^2}; \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

- Середньоквадратична похибка

$$MSE = \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n}$$

- Середня абсолютна похибка

$$MAE = \frac{\sum_{t=1}^n |y_t - \hat{y}_t|}{n}$$

Аналіз результатів

- Для прикладу були взяті світові відсотки міського населення
- Отримані метрики і вибір кращої моделі:

	Coefficient of determination	Mean squared error	Mean absolute error
► Network	0,99890038627705	0,004206711447542...	0,0584172666386607
ARIMA	0,970848792203972	3,30259775919482	1,56075837532822

Аналіз результатів

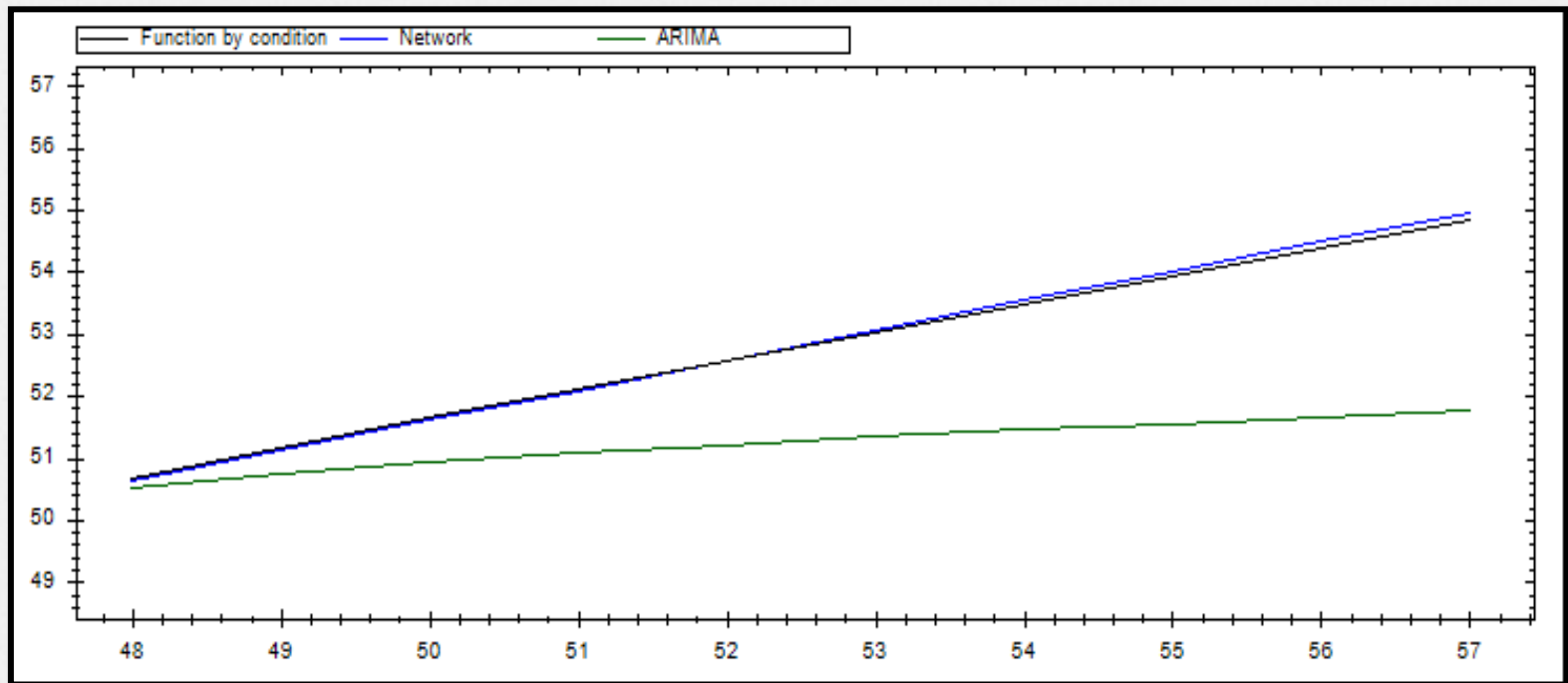


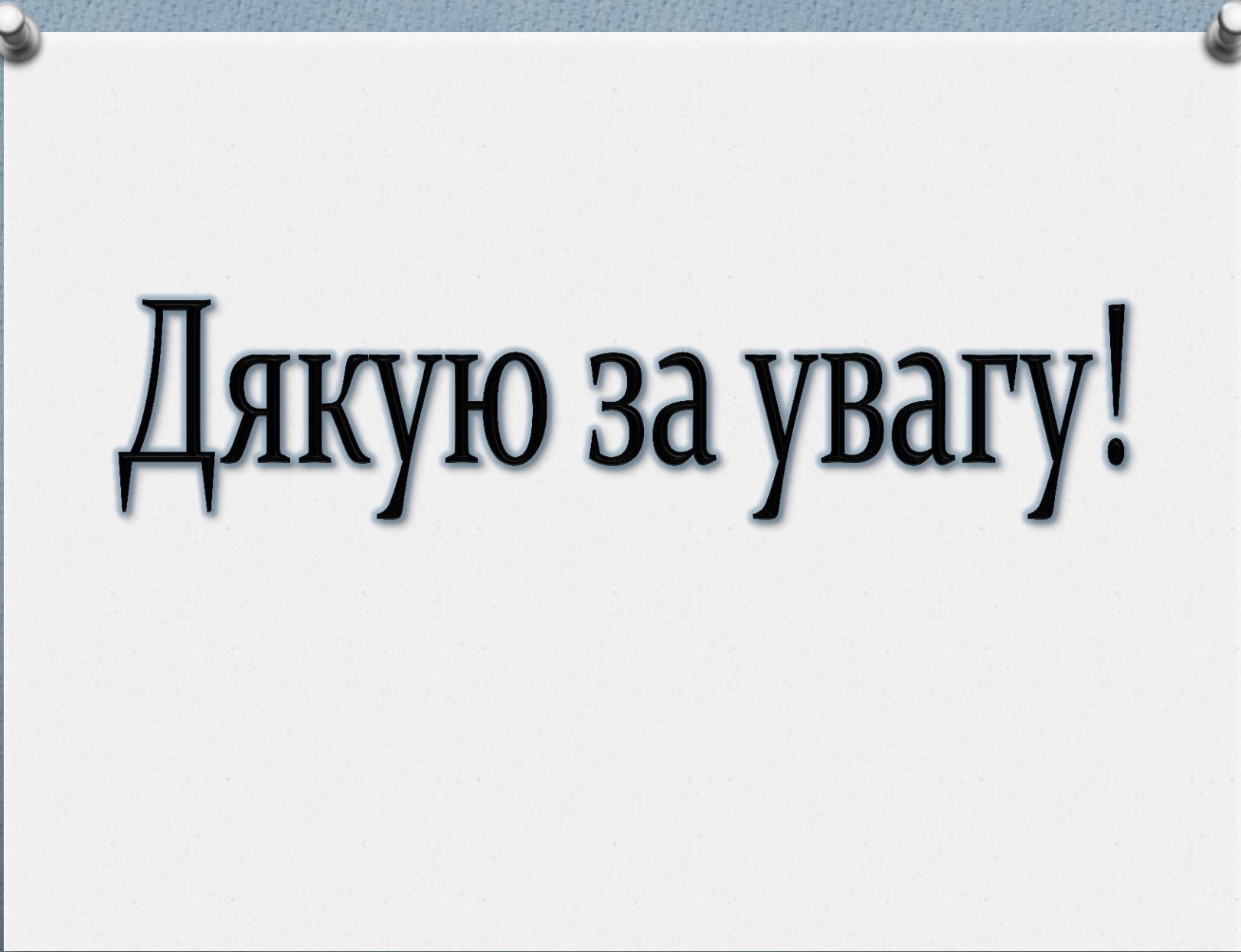
Рисунок 2 – Графіки прогнозованих даних на тестовій вибірці

Висновки

- Сформовано вибірку, на основі якої побудовано прогноз за допомогою АРІКС та нейронної мережі Back Propagation
- На основі проведених моделювань було проведено порівняльний аналіз результатів АРІКС і нейронної мережі, який показав, що мережа в більшості варіантів дає кращі результати при прогнозуванні долі міського населення, в порівнянні з АРІКС
- Розроблено програмний продукт для прогнозування рівня урбанізації

Подальше дослідження

- Вдосконалення моделей для досягнення більшої точності
- Використання інших методів прогнозування та їх порівняльна характеристика
- Включити до аналізу більшу кількість похідних процесів, що впливають на рівень урбанізації



Дякую за увагу!