

РОЗРОБКА ВЕБ-ДОДАТКУ ДЛЯ МОДЕЛЮВАННЯ СТОХАСТИЧНОЇ ЕПІДЕМІОЛОГІЧНОЇ SIR-МОДЕЛІ З НЕЙРОМЕРЕЖЕВОЮ АПРОКСИМАЦІЄЮ

Скутар Ігор Дмитрович

кандидат фіз.-мат. наук,
асистент кафедри прикладної математики та інформаційних технологій,
факультет математики та інформатики,
Чернівецький національний університет імені Юрія Федьковича

Юрченко Олег Ігорович

студент 4-го курсу
кафедри прикладної математики та інформаційних технологій,
факультет математики та інформатики,
Чернівецький національний університет імені Юрія Федьковича

Епідеміологічне моделювання є ключовим інструментом для розуміння та прогнозування динаміки поширення інфекційних хвороб. Однією з найвідоміших математичних моделей у цій галузі є модель SIR (Susceptible-Infectious-Recovered), що була вперше запропонована в 1927 році Вільямом Огілві Кермаком і Ендерсоном Г. Маккендриком у [1]. Ця модель описує динаміку поширення інфекційного захворювання в популяції, розділеній на три класи: сприйнятливі до інфекції (S), інфіковані (I) та ті, що одужали або померли (R).

У наступні десятиліття модель зазнавала численних удосконалень. Детерміновані моделі були розширені для врахування природного приросту населення, різних типів інфекцій, вакцинації тощо. Стохастичні моделі з'явилися як відповідь на потребу відображення невизначеностей у поширенні інфекції, особливо при моделюванні невеликих популяцій або початкових стадій епідемій [2, 3].

Стохастичні модифікації SIR-моделей часто включають вінерові збурення (білий шум), які моделюють непередбачувані флуктуації, наприклад, змінну ймовірність зараження або одужання через зміну поведінки людей, погоду тощо та пуассонові збурення (стрибкоподібні події), які враховують раптові події — наприклад, масові зібрання або локальні спалахи, що різко змінюють кількість інфікованих [4]. Ці типи збурень відображають більш реалістичну картину епідеміологічних процесів, оскільки в реальному світі динаміка захворювання не є гладкою чи передбачуваною. Моделі з диференціальними рівняннями зі стохастичними збуреннями дають змогу врахувати ці аспекти. Для наближеного розв'язання стохастичних диференціальних рівнянь, що містять випадкові збурення, використовується стохастичний аналог методу Ейлера – метод Ейлера-Маруями [5, 6].

Останніми роками методи глибинного навчання, зокрема рекурентні нейронні мережі (RNN), знаходять широке застосування у моделюванні часових рядів, пов'язаних з епідеміологією. Серед них особливе місце займають мережі з довготривалою короткочасною пам'яттю (LSTM — Long Short-Term Memory), які ефективно моделюють залежності на довгих часових інтервалах і здатні працювати з нерівномірними, зашумленими та непередбачуваними даними — саме такими, якими є результати стохастичних епідемічних моделей [4].

Математичні рівняння моделі виглядають так [1–4]:

- для групи сприйнятливих:

$$dS = -\beta \cdot S \cdot I \cdot dt + \sigma \cdot S \cdot dW_S - v_{\text{rate}} \cdot S \cdot dt + \text{reinf}_{\text{rate}} \cdot R \cdot dt,$$

- для групи інфікованих:

$$dI = (\beta \cdot S \cdot I - \gamma \cdot I) \cdot dt + \sigma \cdot I \cdot dW_I + \lambda_{\text{jump}} dN_t + \text{reinf}_{\text{rate}} \cdot R \cdot dt \cdot \text{reinf}_{\text{factor}},$$

- для групи одужавших:

$$dR = \gamma \cdot I \cdot dt + \sigma \cdot I \cdot dW_R - \text{reinf}_{\text{rate}} \cdot R \cdot dt \cdot \text{reinf}_{\text{factor}},$$

- для групи вакцинованих:

$$dV = v_{\text{rate}} \cdot S \cdot dt \quad (\text{починаючи з } \text{vaccine}_{\text{start}}).$$

Опис змінних та параметрів:

S – кількість осіб, які є сприйнятливими до інфекції;

I – кількість інфікованих осіб;

R – кількість осіб, які одужали від інфекції;

V – кількість вакцинованих осіб;

β – коефіцієнт передачі інфекції (шанс передачі інфекції від інфікованого до сприйнятливого);

γ – коефіцієнт одужання (швидкість, з якою інфіковані одужують);

σ – параметр стохастичних збурень для кожної групи (генерація випадкових змінних);

λ_{jump} – середнє число випадкових "стрибків" (пуассонове збурення) для інфікованих;

v_{rate} – швидкість вакцинування;

$\text{reinf}_{\text{rate}}$ – швидкість реінфекції (перехід від R до I);

$\text{reinf}_{\text{factor}}$ – фактор, який визначає вплив реінфекції на популяцію;

t_{max} – максимальний час моделювання;

dt – крок часу для числового розв'язку;

$\text{vaccine}_{\text{start}}$ – час початку вакцинування.

Наближене обчислення розв'язку для вищенаведеної моделі здійснюється за допомогою метода Ейлера-Маруяма [6].

У досліджуваній моделі поширення захворювання з елементами стохастичності (вінерові та пуассонові збурення) використовується нейромережа LSTM для прогнозування майбутньої динаміки захворювання на основі попередніх даних моделі.

До функцій LSTM належать:

- прогноз кількості інфікованих, вилікуваних та вакцинованих (LSTM тренується на результатах стохастичного моделювання (часові ряди S, I, R); мережа навчається ловити тренди та варіації, які класичні моделі (ARIMA, ковзне середнє) не завжди враховують);
- урахування сезонності та нелінійних ефектів (епідемічний процес часто має нелінійний характер (наприклад, повторні хвилі), LSTM може вловити ці складні взаємозв'язки краще за прості регресійні моделі);
- порівняння якості прогнозів (LSTM порівнюється з результатами використання інших моделей (ARIMA, регресія) для побудови графіків з порівняльною точністю, це дозволяє вибрати найкращий підхід для середньо- та довгострокового прогнозування).

Розроблено веб-застосунок, який реалізує епідеміологічну SIR-модель з використанням стохастичних диференціальних рівнянь для моделювання поширення інфекційної хвороби.

У веб-застосунку використовується удосконалена модель стохастичних диференціальних рівнянь для моделювання чотирьох груп осіб: сприйнятливі (S), інфіковані (I), одужавші (R), вакциновані (V). Моделювання включає ймовірнісні збурення, такі як випадкові стрибки для інфікованих осіб (по типу пуассонових збурень), а також стохастичні збурення для кожної групи (за вінеровим процесом). Враховується також вакцинування, що зменшує кількість сприйнятливих осіб і збільшує кількість вакцинованих (V). Відповідно до запропонованої моделі, є враховані випадки реінфекції, що дозволяє моделювати ситуації, коли людина може знову захворіти після одужання.

Застосунок використовує фреймворк Flask [7], технології HTML/CSS/JavaScript, реалізує метод Ейлера-Маруяма [6] наближеного розв'язання стохастичних диференціальних рівнянь, надає користувачам можливість візуалізувати результати моделювання (Matplotlib, Plotly, Dash), порівняти їх з класичними методами прогнозування (моделі ковзного середнього, ARIMA, лінійної регресії) та експортувати дані у форматах CSV і

PDF. Для покращення прогнозів використовується LSTM-нейромережа (Keras, Scikit-learn) [8, 9], що навчається на результатах стохастичного моделювання.

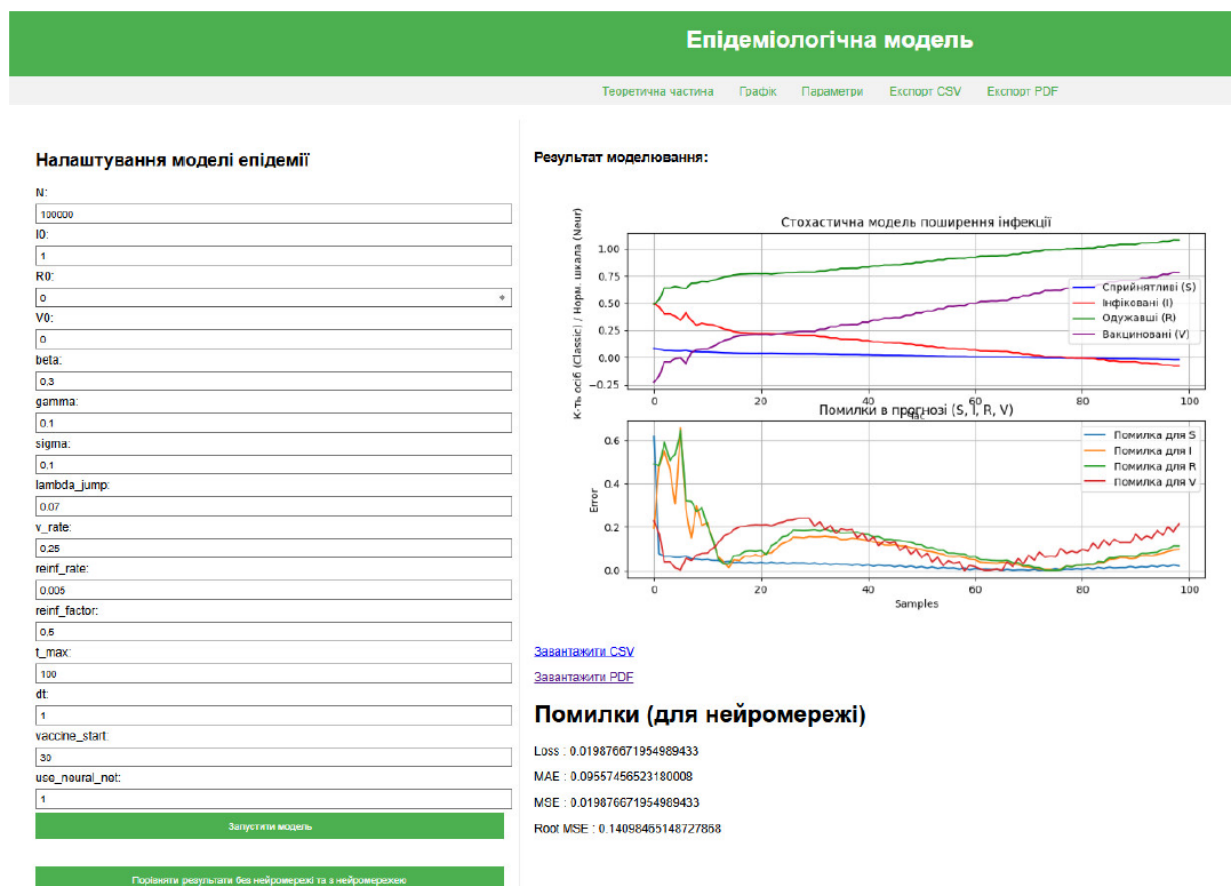


Рис.1. Результати моделювання SIR з вінеровими та пуассоновими збуреннями з використанням нейромережі

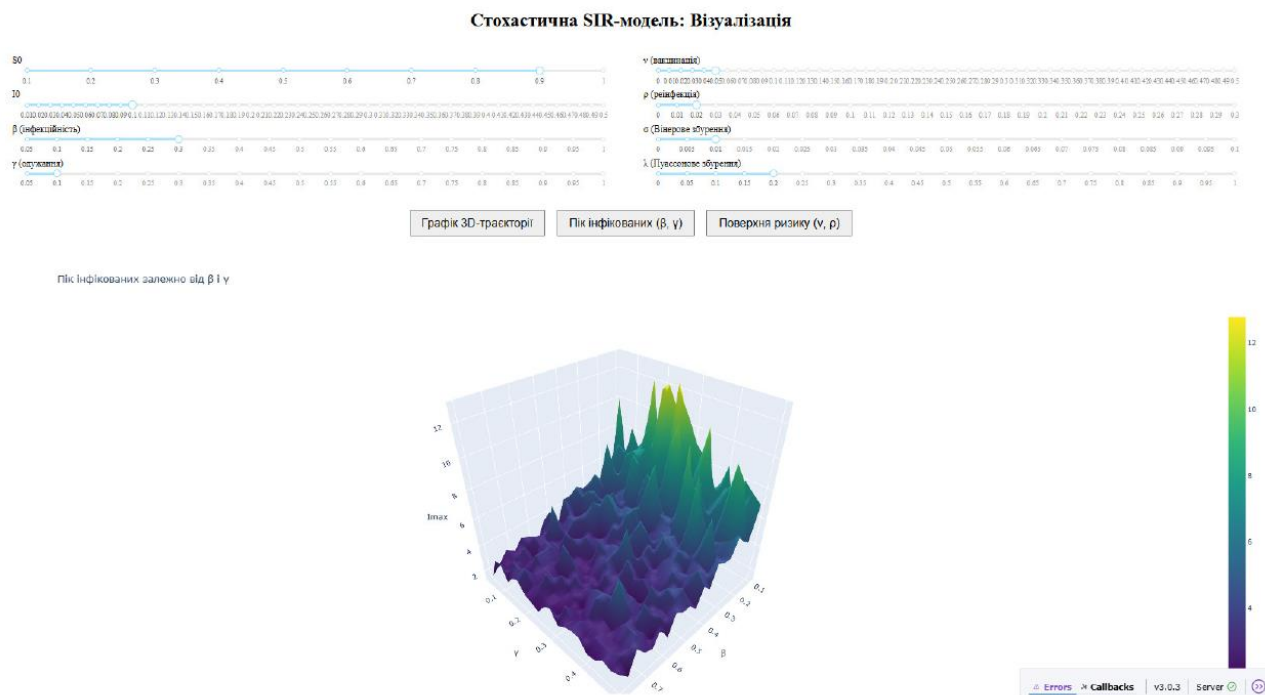


Рис.2. 3D-візуалізація результатів моделювання: пік інфікованих (β, γ, I_{max})

Список літератури:

1. Kermack W.O., McKendrick A.G. A Contribution to the Mathematical Theory of Epidemics // Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character.– 1927.– 115(772).– P.700–721.– <https://doi.org/10.1098/rspa.1927.0118>
2. Allen L. J. S. An Introduction to Stochastic Epidemic Models // In: Brauer F., van den Driessche P., Wu J. (eds) Mathematical Epidemiology. Lecture Notes in Mathematics.– Vol.1945.– Springer, 2008.– https://doi.org/10.1007/978-3-540-78911-6_3
3. Gray A., Greenhalgh D., Hu L., Mao X., Pan J. (2011). A stochastic differential equation SIS epidemic model // SIAM Journal on Applied Mathematics.– 2011.– 71(3).– P.876–902.– <https://epubs.siam.org/doi/10.1137/10081856X>
4. Shiu Kumar, Ronesh Sharma, Tatsuhiko Tsunoda. Forecasting the spread of COVID-19 using LSTM network // BMC Bioinformatics.– 2021.– 22(S6).– DOI:10.1186/s12859-021-04224-2
5. Mao X. Stochastic Differential Equations and Applications.– Woodhead Publishing, 2007.
6. Метод Ейлера-Маруяма [Електронний ресурс].– https://en.wikipedia.org/wiki/Euler%E2%80%93Maruyama_method
7. Фреймворк Flask [Електронний ресурс].– <https://flask.palletsprojects.com/>
8. Бібліотека Keras мови Python [Електронний ресурс].– <https://keras.io/>
9. Бібліотека Scikit-learn мови Python [Електронний ресурс].– <https://scikit-learn.org/stable/>