

ПРОЄКТ СИСТЕМИ СИНТЕЗУ ПОЯСНЕНИХ ВЕРБАЛЬНИХ МОДЕЛЕЙ ШТУЧНОГО ІНТЕЛЕКТУ

Фастовський Е. Г.

Науковий керівник – к.т.н., доц. Єльчанінов Д. Б.

Національний технічний університет «Харківський політехнічний
інститут», м. Харків, Україна

Verbal Decision Analysis is an effective tool that can improve Explainable Artificial Intelligence (XAI) by providing verbal explanations of machine learning models and reasoning using words, text, or natural language. Synthesis of explainable verbal AI models involves four main stages: formation of a system of concepts in a particular subject area; formation of an ordinal classification of object/process states; ordering of object/process states from a particular class; determination of the best state of an object/process. By integrating elements of verbal analysis into XAI models, it is possible to improve user interaction, understanding, and perception of AI systems.

Для розробки системи синтезу пояснених моделей штучного інтелекту (ШІ) необхідно врахувати кілька ключових аспектів [1-3]:

- зіставлення вербальної поведінки людини з поведінкою системи ШІ, враховуючи особистісні критерії, щоб генерувати відповідні вербальні та невербальні комбіновані моделі поведінки;

- покращення пояснення систем ШІ в користувацьких інтерфейсах, зосередившись на зрозумілості, персоналізації та візуалізації пояснень, визнанні компромісів, усуненні потенційних помилкових уявлень і прив'язці пояснень до ментальних моделей користувачів задля прозорості та сприйняття користувачами рекомендацій систем ШІ;

- перевірка результатів дослідження в реальних умовах через вивчення персоналізованих пояснень, візуалізацію та забезпечення відображення пояснень у користувацькому інтерфейсі.

Вербальний аналіз рішень [4] – це ефективний інструмент, який може вдосконалити системи пояснюваного ШІ, надаючи словесні пояснення моделей машинного навчання та міркувань за допомогою слів, тексту або природної мови. Наприклад, цей підхід може допомогти людям краще зрозуміти процес прийняття рішень системами ШІ і потенційно зменшити кількість пацієнтів, яким необхідно пройти тривалий і багатоетапний процес діагностики [5].

Синтез пояснювальних вербальних моделей ШІ передбачає чотири основні етапи (рис. 1):

- формування системи понять визначеної предметної галузі;
- формування порядкової класифікації станів об'єкта/процесу;
- впорядкування станів об'єкта/процесу з певного класу;
- визначення найкращого стану об'єкта/процесу.



Рисунок 1 – Діаграма варіантів використання системи

Формування системи понять визначеної предметної галузі складається з таких діяльностей (рис. 2):

- виявлення критеріїв порівняння об’єктів/процесів певної предметної галузі;
- формування порядкових шкал значень критеріїв;
- побудова графа домінування станів об’єктів/процесів;
- порівняння визначених станів та пояснення типу відношення між ними (краще, гірше, непорівнянні).

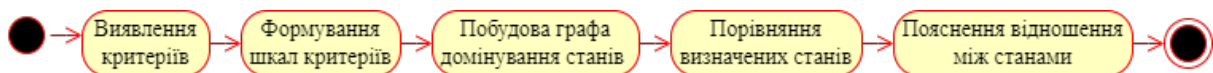


Рисунок 2 – Діаграма діяльностей формування системи понять

Порядкова класифікація гіпотетично можливих станів складається з таких діяльностей (рис. 3):

- визначення назв та лінійного порядку класів;
- визначення найінформативнішого стану;
- пряма класифікація найінформативнішого стану;
- перевірка несуперечливості прямої класифікації;
- непряма класифікація станів;
- класифікація реальних станів;
- визначення та пояснення ефективності певної траєкторії покращення стану.

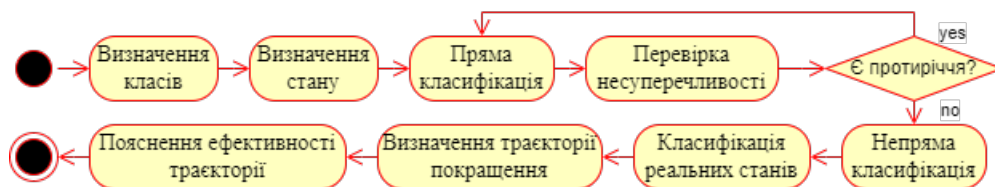


Рисунок 3 – Діаграма діяльностей порядкової класифікації станів

Впорядкування станів об’єкта/процесу з певного класу складається з таких діяльностей (рис. 4):

- побудова єдиної порядкової шкали (ЄПШ) для всіх пар критеріїв;
- побудова ЄПШ для всіх критеріїв;

- перевірка несуперечливості ЄПШ;
- порівняння визначених станів та пояснення типу відношення між ними (краще, гірше, непорівнянні).

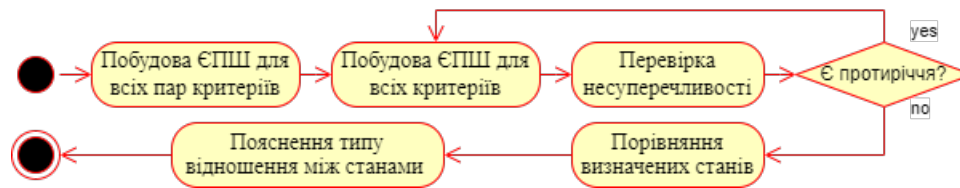


Рисунок 4 – Діаграма діяльностей впорядкування станів

Визначення найкращого стану об'єкта/процесу складається з таких діяльностей:

- вибір пари станів для порівняння;
- визначення відносних недоліків станів за кожним критерієм;
- впорядкування недоліків за важливістю для користувача;
- взаємна компенсація недоліків;
- визначення кращого стану.

Інтегруючи елементи вербального аналізу в пояснені моделі ШІ, можна покращити взаємодію з користувачем, його розуміння і сприйняття систем ШІ.

Список використаних джерел:

1. AKM Bahalul Haque, A.K.M. Najmul Islam, Patrick Mikalef. Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. *Technological Forecasting & Social Change*. January 2023. Vol. 186, Part A. Article 122120.
2. Laato S., Tiainen M., Najmul Islam A.K.M., Mäntymäki M. How to explain AI systems to end users: a systematic literature review and research agenda. *Internet Research*. 2022. Vol. 32, No 7. P. 1–31.
3. Aly A., Tapus A. A model for synthesizing a combined verbal and nonverbal behavior based on personality traits in human-robot interaction. 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI). Mar 2013, Tokyo, Japan. P. 325–332.
4. Moshkovich H., Mechitov A., Olson D. Verbal Decision Analysis. Multiple Criteria Decision Analysis: State of the Art Surveys. *International Series in Operations Research & Management Science*. 2005. Vol. 78. P. 609–633.
5. Evandro Andrade, Samuel Portela, Plácido Rogério Pinheiro, Luciano Comin Nunes, Marum Simão Filho, Wagner Silva Costa, Mirian Caliope Dantas Pinheiro. A Protocol for the Diagnosis of Autism Spectrum Disorder Structured in Machine Learning and Verbal Decision Analysis. *Computational and Mathematical Methods in Medicine*. 2021. Vol. 2021. Article ID 1628959.