

ФОРМУВАННЯ КАУЗАЛЬНИХ ПРОЦЕСНИХ ПОЯСНЕНЬ В ІНТЕЛЕКТУАЛЬНИХ СИСТЕМАХ

Лещинський В.О.

Харківський національний університет радіоелектроніки, Харків, Україна

Інтелектуальні інформаційні системи забезпечують можливості підтримки прийняття рішень у виробництві, охороні здоров'я, управлінні бізнес-процесами та фінансовому аудиті, тобто у сферах, де своєчасність рішень безпосередньо визначає результати діяльності організації.

Однак використанню таких систем перешкоджає їхня непрозорість, оскільки складні алгоритми машинного навчання формують рішення без пояснення щодо того, які стани системи, дії та причинно-наслідкові зв'язки між ними зумовили отриманий результат. Концепція пояснювального штучного інтелекту (explainable artificial intelligence, XAI) реалізована як система методів підвищення прозорості моделей машинного навчання, орієнтованих на різні категорії користувачів [1]. Зовнішні кінцеві користувачі потребують пояснень на концептуальному рівні, тоді як внутрішні користувачі, якими є розробники та аналітики систем, потребують детального процесного опису послідовності станів і дій, що сформували рішення, з метою виявлення вузьких місць і вдосконалення відповідних моделей. Існуючі методи XAI, зокрема LIME та SHAP, забезпечують пояснення на рівні важливості окремих ознак, але не відображають причинно-наслідкових зв'язків між послідовними станами процесу формування рішення, що обмежує їх можливості щодо опису процесно-орієнтованої поведінки системи. Каузальне моделювання надає формальний апарат для опису причинно-наслідкових залежностей між подіями та станами складних систем [2], що відкриває можливість побудови пояснень, які відповідають на питання «чому» та «що було причиною» щодо отриманого рішення, а не лише питання «що вплинуло на рішення».

Темпоральні залежності, які представляють собою впорядковані у часі відношення між станами процесу прийняття рішення, у поєднанні з каузальними залежностями дають можливість відобразити логіку рішення як упорядкований у часі ланцюжок дій і відповідних їм змін стану системи.

Інтелектуальна система на практиці може розглядатись як «сіра скринька», оскільки її поведінка лише частково відображена у системних логах (журналах подій), що обмежує використання ймовірнісної оцінки правил формування рішення. Проте теорія можливостей дозволяє оцінювати ступінь правдоподібності та впевненості у причинно-наслідкових правилах в умовах неповноти інформації [3]. У підсумку, існуючі підходи реалізують окремо побудову пояснень на основі впливу вхідних даних, опис процесу прийняття рішення у вигляді комбінації темпоральних та каузальних залежностей, а також оцінку правдоподібності цих залежностей. Зазначене свідчить про актуальність створення інформаційної технології побудови каузальних процесно-орієнтованих пояснень.

Метою роботи є розробка інформаційної технології формування каузальних процесних пояснень в інтелектуальних системах на основі можливісних каузальних залежностей між станами процесу формування рішення.

Запропонована інформаційна технологія включає фази побудови моделі пояснення та формування пояснення на основі отриманої моделі. Фаза побудови моделі пояснення включає аналіз логу процесу, побудову темпоральних залежностей і на їх основі формування й можливісну оцінку каузальних залежностей, а також формування моделі пояснення із множини правил. На даній фазі із журналу подій процесу формування рішення виділяються впорядковані послідовності станів, кожен з яких описується атрибутами найменувань дій, часових міток виконання та залучених ресурсів. Між парами послідовних станів усередині кожної траси журналу формуються темпоральні залежності, які узагальнюються у темпоральні правила, що охоплюють повторювані послідовності дій у кількох реалізаціях процесу. Темпоральні правила перетворюються на каузальні залежності та потенційні каузальні правила, які описують причинно-наслідкові зв'язки між станами. Для кожного правила визначаються оцінки можливості та необхідності [4]. Можливість застосування правила оцінюється за частотою його використання у відомих реалізаціях процесу, а показник необхідності визначається як відносна частота застосування цього правила серед усіх альтернативних правил, що можуть спрацювати в аналогічних ситуаціях. Модель пояснення формується як впорядкована підмножина правил із різним рівнем деталізації, що дозволяє користувачеві самостійно визначати потрібну глибину пояснення. Фаза представлення пояснення включає деталізацію відібраних каузальних правил з урахуванням граничних значень атрибутів станів та побудову пояснення звичайною мовою на основі каузальних правил.

Висновки. Запропонована інформаційна технологія формування каузальних процесних пояснень в інтелектуальних інформаційних системах відрізняється використанням послідовності станів процесу для формування темпоральних і можливісних каузальних залежностей, що створює умови для побудови пояснень в умовах неповноти інформації щодо процесу прийняття рішення в інтелектуальній системі.

Список літератури

1. Arrieta A. B., Díaz-Rodríguez N., Del Ser J. та ін. Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI // *Information Fusion*. 2020. Vol. 58. P. 82–115. DOI: <https://doi.org/10.1016/j.inffus.2019.12.012>
2. Pearl J. *Causality: Models, Reasoning and Inference*. 2nd ed. Cambridge: Cambridge University Press, 2009. 484 p. ISBN: 978-0-521-89560-6.
3. Zadeh L. A. *Fuzzy Sets as a Basis for a Theory of Possibility // Fuzzy Sets and Systems*. 1978. Vol. 1, № 1. P. 3–28. DOI: [https://doi.org/10.1016/S0165-0114\(78\)80002-4](https://doi.org/10.1016/S0165-0114(78)80002-4)
4. Чалий С. Ф., Лещинський В. О. Можливісна оцінка коректності пояснень для кінцевого користувача в системі штучного інтелекту. *Сучасні інформаційні системи*. 2023. Т. 7, № 4. С. 75–79. DOI: <https://doi.org/10.20998/2522-9052.2023.4.10>