問題的陳述從主題領域本體的構建及其在邏輯語言框架內形式化的三個階段開始，提出的算法及其軟件實現，以及這三個層次之間關係的分析 形式化。
語義方法的數學理論基於編程語言 Prolog
的思想、可容許集理論、經典的可計算性理論，包括與 oracles
的相對可計算性和證明其合理性的邏輯理論。
在概念上統一的數學理論的基礎上，已經開發出一種方法來在各個學科領域實際實施這些想法：遺傳學、醫學、金融系統 (E.E. Vityaev)、大型企業的文檔管理系統 (A. Man c ivoda)、電信、決策支持系統、支付系統和自動化管理此類系統的任務（D.I. Sviridenko、V. Gumirov、D.E. Palchunov、P. Demenkov）。

文學

1. Ershov Y.L., Goncharov S.S., Sviridenko D.I. Semantic programming. Information processing. Proc. IFIP 10th World Comput. Congress, Dublin. Vol. 10, 1986, 1113-1120.

2. Goncharov S.S., Nechesov A.V. Semantic programming for AI and Robotics. IEEE International Multi-Conference on Engineering, Computer and Information Sciences, 2022, pp.810-815. Doi 10.1109/SIBIRCON56155.2022.10017077.

3. Goncharov S.S., Nechesov A.V. Polynomial analogue of Gandy's fixed point theorem. Mathematics. 2021;9(17):2102. Doi 10.3390/math9172102.

DOI 10.18699/sblai2023-03

Объяснимый искусственный интеллект: проблемы, вопросы и приложения

Марченко М.А.
Институт вычислительной математики и математической геофизики СО РАН,
Новосибирск, Россия
marchenko@sscc.ru

Объяснимый искусственный интеллект (Explainable AI) – математическая модель, которая способна объяснять механизмы, лежащие в основе алгоритмов искусственного интеллекта (моделей машинного обучения). Множество решений, применяющих алгоритмы ИИ, представляют собой подобие «черного ящика». Поэтому зачастую не только конечные пользователи, но и сами разработчики не могут точно определить, как именно модель машинного обучения пришла к тем или иным выводам в ходе обработки исходных данных. Понимание работы алгоритмов искусственного интеллекта позволит разработчикам точно оценивать влияние входных признаков на выходной результат модели, выявлять необъективности и недостатки, связанные с работой модели, а также проводить тонкую настройку и оптимизацию такого рода методов. Для пользователей объяснимость результата работы методов искусственного интеллекта важна в части понимания причин выводов, сделанных моделью, а для экспертов – для объяснения тех выводов, которые на первый взгляд не имеют под собой оснований, например при поиске скрытых закономерностей в данных.

В докладе делается обзор алгоритмически обоснованных подходов к разработке методов искусственного интеллекта, описываются методы интерпретации результатов их применения для решения конкретных задач.

Explainable artificial intelligence: problems, questions and applications

Marchenko M.A.
Institute of Computational Mathematics and Mathematical Geophysics of SB RAS,
Novosibirsk, Russia
marchenko@sscc.ru

The explained artificial intelligence (Explainable AI) is a mathematical model that can explain the mechanisms underlying artificial intelligence algorithms (machine learning models). Many solutions that use AI algorithms are like a “black box”. Therefore, often not only end users, but the developers themselves cannot determine exactly how exactly the machine learning model came to one or another conclusions during the processing of the source data. Understanding the work of artificial intelligence algorithms will allow developers to accurately evaluate the influence of the input features on the output result of the model, to identify bias and disadvantages associated with the work of the model, as well as to carry out subtle tuning and optimizing this kind of method. For users, the explanation of the result of the work of artificial intelligence methods is important regarding the understanding of the causes of the conclusions made by the model, and for experts, to explain those conclusions that at first glance have no grounds, for example, when looking for hidden laws in the data.

The report makes an overview of algorithmically sound approaches to the development of artificial intelligence methods, and methods of interpreting the results of their application are described to solve specific applications.

可解釋的人工智能：問題、問題和應用

馬爾琴科・米哈伊爾・亞歷山德羅維奇
Marchenko M.A.
ICM&MG SB RAS 總監
Marchenko@sscc.ru

Explainable AI
是一種數學模型，能夠解釋人工智能算法（機器學習模型）背後的機制。
許多使用人工智能算法的解決方案就像一個“黑匣子”。
因此，不僅最終用戶，而且開發人員本身往往無法準確確定機器學習模型在處理初始數據期間如何得出某些結論。
了解人工智能算法的運行將使開發人員能夠準確評估輸入特徵對模型輸出的影響，識別與模型運行相關的偏差和缺點，以及微調和優化此類方法。
對於用戶來說，人工智能方法工作結果的可解釋性對於理解模型得出結論的原因很重要，對於專家來說，對於解釋那些乍一看沒有根據的結論，例如，在搜索數據中的隱藏模式時。

該報告概述了基於算法的人工智能方法開發方法，描述了解釋其應用結果以解決特定問題的方法。

DOI 10.18699/sblai2023-33

Три определения меры знания

Михайловский Н.
Высшая ИТ-школа ТГУ и ООО «НТР», Томск, Россия
nickm@ntr.ai

Познающие агенты, включая людей, решают проблемы, строя модели. Модель – это объяснение явления, то есть предположение механизма того, как оно может происходить.