УДК 681.3

ОБ АНАЛИЗЕ НЕКОТОРЫХ ХАРАКТЕРИСТИК МАШИННОГО ОБУЧЕНИЯ

© 2022 Я. Е. Львович, Ю. П. Преображенский, Е. Ружицкий

Воронежский государственный технический университет
Воронежский институт высоких технологий
Панъевропейский университет (Братислава, Словакия)

В статье обсуждаются некоторые особенности, связанные с машинным обучением.

Ключевые слова: машинное обучение, информация, классификация.

Чтобы получать знания на основе данных, перспективным считается использование систем, базирующихся на машинном обучении. На рисунке 1 приведены области, относящиеся к науке о данных. Данные в ходе анализа могут иметь очень большие размеры.

В них может быть неполнота, элементы разнородности, наличие противоречивости. Необходимо ориентироваться на применение специальных программных продуктов, чтобы вести процесс обработки данных [1, 2].

Иллюстрация методов задач в машинном обучении приведена на рис 2. На основе каких признаков можно использовать методы машинного обучения? Некоторые из них мы видим на рисунке 3.

Ключевые этапы в машинном обучении проиллюстрированы на рисунке 4. Существуют разные типы признаков для объектов.

Мы можем увидеть на рисунке 5, каким образом они продемонстрированы. Способы описания данных зависят от того, какие они будут на входе системы машинного обучения [3, 4].

Их можно увидеть на рисунке 6. Типы задач в машинном обучении даны на рисунке 7.

Обобщенные типы проиллюстрированы на рисунке 8. Для определенного множества данных неизвестная целевая зависимость может быть приближена на основе регрессионных подходов.

Для заданных совокупностей групп распределяются [5, 6] множества объектов в классификационных подходах. Схожие объекты распределяются по непересекающимися группам в подходах, связанных с кластеризацией.

В связи с чем решение задач в ходе классификации является более сложными, если сравнивать с классификацией? По качеству кластеризации трудно говорить об однозначном критерии [7, 8].

Кроме того, не во всех случаях есть данные относительно числа кластеров. Субъективный выбор метрики расстояния будет влиять на то, какие будут пастеризации результаты.

Что мы можем достичь за счёт использования подходов, базирующееся на кластеризации?

Вследствие того, что схожие объекты разбиваются из выборки по группам, возникают возможности по тому, что более лучшим образом будут пониматься данные.

Также данные могут быть сжаты. Если к кластерам мы не можем отнести какой-то из объектов [9, 10], может быть найдена определённая новизна.

Упорядочение исходных данных может осуществляться за счёт применения разных иерархических структур. То есть, реализуется решение таксономических задач.

Действия, которые совершаются в ходе предварительной обработки данных приведены на рисунке 9.

Львович Яков Евсеевич – Воронежский государственный технический университет, доктор техн. наук, профессор, e-mail: office@vivt.ru.
Преображенский Юрий Петрович – Воронежский институт высоких технологий, канд. техн. наук, профессор, e-mail: petrovich@vivt.ru.
Ружицкий Евгений – Панъевропейский университет, канд. техн. наук, доцент, rush_evg_br53@yandex.ru.

Рисунок 1 – Области, относящиеся к науке о данных

Рисунок 2 – Иллюстрация задач в машинном обучении

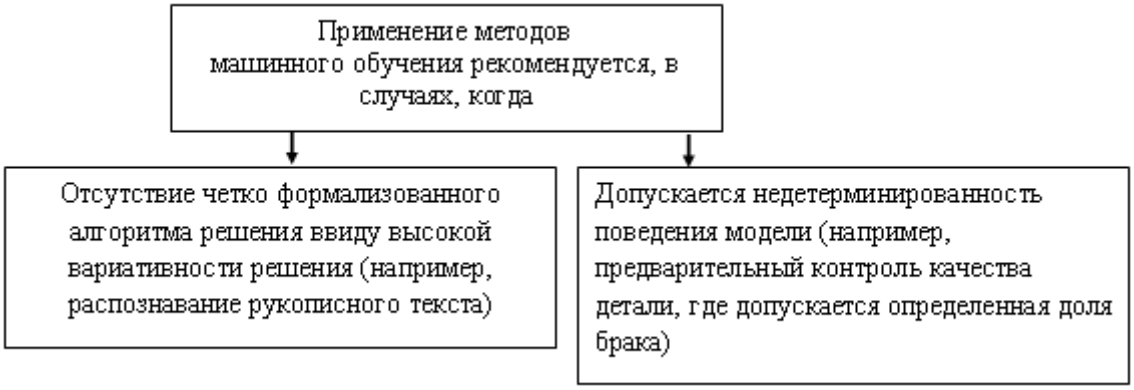

Рисунок 3 – Иллюстрация возможности применения методов в машинном обучении

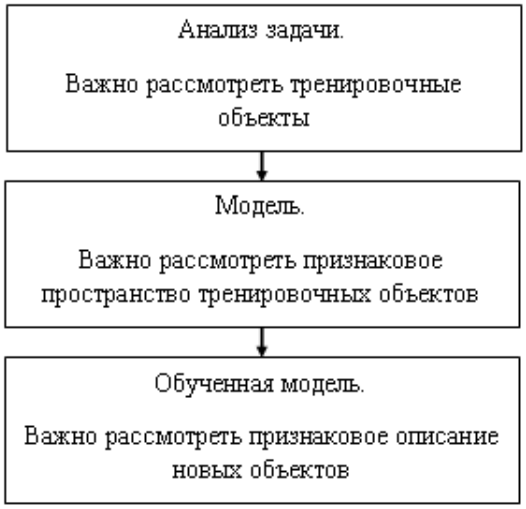

Рисунок 4. Иллюстрация ключевых этапов машинном обучении

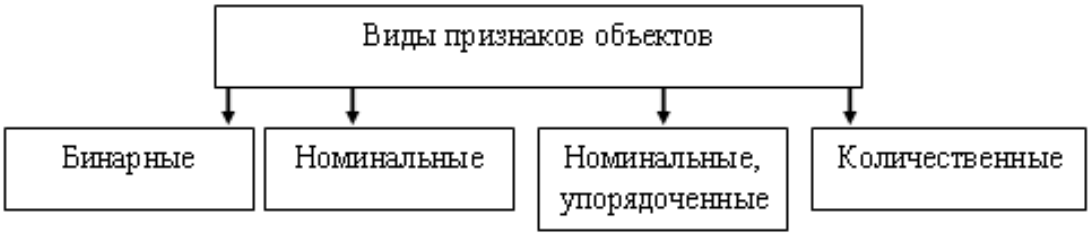

Рисунок 5. Иллюстрация видов признаков объектов

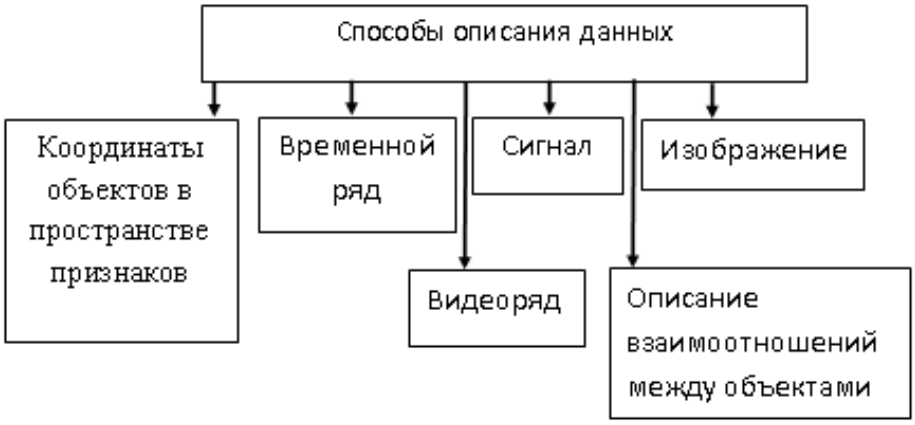

Рисунок 6. Иллюстрация способов описания данных

Рисунок 7. Иллюстрация типов задач в машинном обучении

Рисунок 8. Иллюстрация обобщенных типов задач в машинном обучении

Рисунок 9. Иллюстрация предварительной обработки данных

СПИСОК ИСТОЧНИКОВ

1. Lvovich I. Managing developing internet of things systems based on models and algorithms of multi-alternative aggregation / I. Lvovich, A. Preobrazhenskiy, Y. Preobrazhenskiy, Y. Lvovich, O. Choporov // 2019 International Seminar on Electron Devices Design and Production, SED 2019 – Proceedings. – 2019. – С. 8798413.

2. Машков В. Г. Предварительная оценка вероятности принятия правильного решения в автоматизированных системах управления / В. Г. Машков, В. А. Малышев, Ю. В. Никитенко // Моделирование, оптимизация и информационные технологии. – 2021. – Т. 9. – № 3 (34). – С. 12-13.

3. Борзова А. С. Особенности построения системы принятия решений при многовариантной оптимизации структуры цифрового управления логистическим процессом в организационной системе на основе имитационного моделирования / А. С. Борзова, В. В. Муха // Моделирование, оптимизация и информационные технологии. – 2021. – Т. 9. – № 3 (34). – С. 15-16.

4. Диденко С. С. Применение мультиагентных технологий в контекстно-ориентированной среде компонента умного дома / С. С. Диденко // Моделирование, оптимизация и информационные технологии. – 2021. – Т. 9. – № 2 (33). – С. 18-19.

5. Lvovich I. Ya. Modeling of information processing in the internet of things at agricultural enterprises / I. Ya. Lvovich, Ya. E. Lvovich, A. P. Preobrazhenskiy, Yu. P. Preobrazhenskiy, O. N. Choporov // IOP Conference Series: Earth and Environmental Science. Krasnoyarsk Science and Technology City Hall of the Russian Union of Scientific and Engineering Associations. – 2019. – С. 32029.

6. Lvovich I. Optimization of the subsystem for the movement of electronic documents in educational organization / I. Lvovich, A. Preobrazhenskiy, Y. Preobrazhenskiy, Y. Lvovich, O. Choporov // Proceedings - 2021 1st International Conference on Technology Enhanced Learning in Higher Education, TELE 2021. – 1. – 2021. – С. 328-332.

7. Львович Я. Е. Исследование характеристик защищенности мобильных сенсорных сетей / Я. Е. Львович, И. Я. Львович, А. П. Преображенский, Ю. П. Преображенский, О. Н. Чопоров // Радиолокация, навигация, связь. Сборник трудов XXV Международной научно-технической конференции, посвященной 160-летию со дня рождения А. С. Попова. В 6-ти томах. – 2019. – С. 239-244.

8. Новосадов К. С. Анализ спектрально эффективных схем модуляции, применяемых в высокоскоростных системах радиосвязи / К. С. Новосадов // Моделирование, оптимизация и информационные технологии. – 2021. – Т. 9. – № 1 (32). – С. 20-21.

9. Lvovich I. Ya. Modelling of information systems with increased efficiency with application of optimization-expert evaluation / I. Ya. Lvovich, Ya. E. Lvovich, A. P. Preobrazhenskiy, Yu. P. Preobrazhenskiy, O. N. Choporov // Journal of Physics: Conference Series. International Scientific Conference «Conference on Applied Physics, Information Technologies and Engineering – APITECH-2019». Krasnoyarsk Science and Technology City Hall of the Russian Union of Scientific and Engineering Associations; Polytechnical Institute of Siberian Federal University. – 2019. – С. 33079.

10. Печенкин В. В. Моделирование динамики серверной нагрузки стохастическими сетями Петри с приоритетами (на примере системы видеоконференцсвязи) / В. В. Печенкин, А. Т. Х. Аль-Хазраджи, С. С. Гельбух // Моделирование, оптимизация и информационные технологии. – 2021. – Т. 9. – № 1 (32). – С. 10-11.

ABOUT ANALYZING SOME OF THE CHARACTERISTICS OF MACHINE LEARNING

© 2022 Ya. E. Lvovich, Yu. P. Preobrazhensky, E. Ruzhitsky

Voronezh State Technical University (Voronezh, Russia)
Voronezh Institute of High Technologies (Voronezh, Russia)
Pan-European University (Bratislava, Slovakia)

The paper discusses some of the features associated with machine learning.

Keywords: machine learning, information, classification.