

АНАЛИЗ НЕКОТОРЫХ ОСОБЕННОСТЕЙ В ХОДЕ РЕАЛИЗАЦИИ МАШИННОГО ОБУЧЕНИЯ

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

Воронежский государственный технический университет (Воронеж, Россия)

Воронежский институт высоких технологий (Воронеж, Россия)

Панъевропейский университет (Братислава, Словакия)

В статье обсуждаются некоторые особенности, которые наблюдаются в ходе реализации подходов, базирующихся на машинном обучении.

Ключевые слова: обработка информации, информационная система, машинное обучение.

В ходе машинного обучения можно наблюдать пропуски по исходным данным. В связи с чем это может быть? Для анкет или аккаунтов бывает так, что нет по всем полям заполнения, которое должны вести клиенты.

Вследствие неверного заполнения данных можно получить выбросы или ошибки. Тогда для интерпретации пропусков следует решать соответствующие проблемы. Могут быть использованы разные подходы.

Например, если есть неполные сведения по анализируемым объектам, то они исключаются. Когда наблюдаются пропуски по столбцам и строкам в таблицах, такие элементы удаляются.

Преимущество такого подхода состоит в том, что матричное представление можно увидеть во многих моделях [1, 2]. Если в строке есть информация по информативным признакам, то она будет потеряна. Это приведёт к тому же, что будет получаться аппроксимация с низким качеством при моделировании.

За счёт осуществления интерполирования пропуски могут быть заменены. По строке, например, можно ориентироваться на медиану или среднее значение.

Относительно соседних значений процесс интерполяции может осуществляться,

если анализируемый признак рассматривается в виде функции от времени [3, 4].

Специальное числовое значение можно применять для того, чтобы наблюдаемые пропуски были закодированы. С точки зрения интерполяции данных могут быть привлечены эксперты [5, 6].

Если исследователи уже сформировали своё видение относительно предметной области, то они смогут осуществить формирование необходимых моделей. Существует риск того, что будет решаться не та задача, которая требуется, а которая была искусственным образом сформирована [7, 8].

Указанные выше подходы на практике могут быть скомбинированы. Специфика задачи влияет на выбор методов. Какие языки программирования могут применяться в ходе разработок для систем машинного обучения?

На данный момент различные пакеты и библиотеки содержат в себе реализованные алгоритмы [9, 10]. Существует вероятность роста сложности в разработках, если отсутствуют специализированные библиотеки.

С другой стороны, когда реализуются полноценные приложения, тогда могут быть сложности в создании моделей на базе специализированных средств.

На рисунке 1 дана иллюстрация возможного построения структуры системы, позволяющей реализовать машинное обучение.

На рисунке 2 показаны основные этапы алгоритма, на основе которого могут формироваться системы для осуществления машинного обучения. В ходе разработок используется язык Python.

Львович Яков Евсеевич – Воронежский государственный технический университет, доктор техн. наук, профессор, e-mail: office@vvt.ru.

Преображенский Юрий Петрович – Воронежский институт высоких технологий, канд. техн. наук, профессор, e-mail: petrovich@vvt.ru.

Ружицкий Евгений – Панъевропейский университет, канд. техн. наук, доцент, rush_evg_br53@yandex.ru.

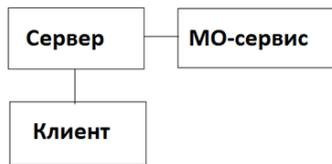


Рисунок 1. Иллюстрация возможной структуры системы машинного обучения

На рисунке 3 дана иллюстрация его ключевых отличий. Деревья решений могут быть использованы в ходе осуществления обработок данных.

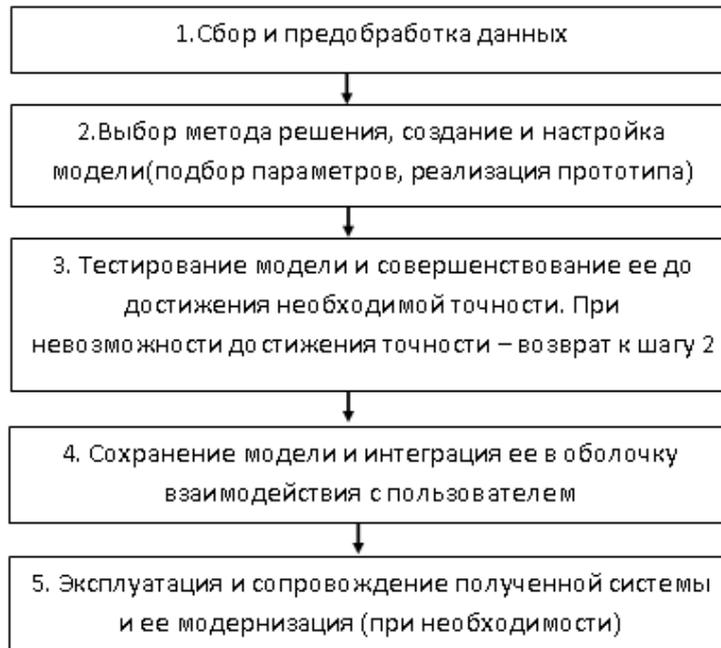


Рисунок 2. Основные этапы алгоритма машинного обучения

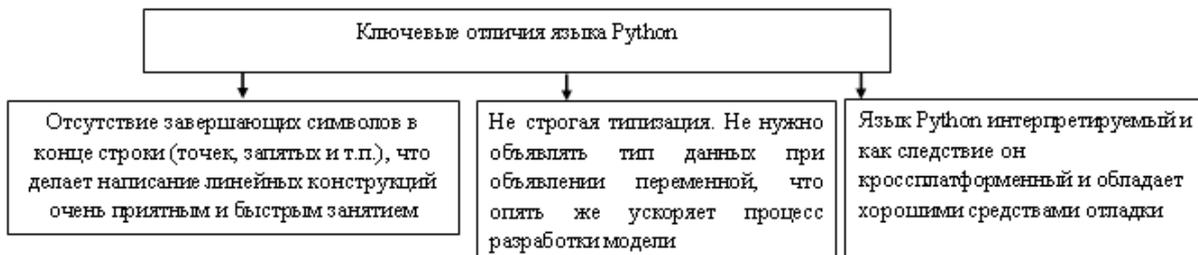


Рисунок 3. Иллюстрация ключевых отличий языка Python

Указанный подход рассматривается в виде логического метода. Почему является интерпретируемым итоговый алгоритм. Это связано с тем, что простые решающие правила объединяются.

Если данные являются разреженными, то в качестве перспективных можно считать методы линейной регрессии. они могут быть полезны и для того, чтобы были по признакам определены характеристики важности.

На практике могут быть реализованы основные шаги для того, чтобы сформиро-

вать полиномиальную регрессию в рамках существующих программных средств.

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.

ANALYSIS OF SOME FEATURES DURING THE IMPLEMENTATION OF MACHINE LEARNING

© 2022 Ya. E. Lvovich, Yu. P. Preobrazhenskiy, E. Ruzhitskiy

*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 that are observed during the implementation of approaches based on machine learning.

Keywords: information processing, information system, machine learning.