

Нейронные сети как инструмент анализа данных в промышленном интернете вещей

О. Н. Петров, Н. Вайнт

Санкт-Петербургский государственный морской технический университет

petr_oleg@mail.ru

Аннотация. Обсуждается проблема предварительной обработки и нейросетевого анализа данных, полученных с устройств промышленного интернета вещей. Рассматриваются различные подходы к подготовке данных, в том числе к процедурам нормализации и масштабирования. Предлагается концепция применения нейросетевых методов обнаружения закономерностей в данных от корреляции к регрессии и авторегрессии. Оптимизация нейронных сетей осуществляется с помощью их представления в формате Open Neural Network Exchange (ONNX) с выполнением средствами ONNX Runtime. Приводятся примеры внедрения и апробации.

Ключевые слова: нейронные сети; данные IoT; обнаружение скрытых знаний; оптимизация ONNX

I. ВВЕДЕНИЕ

A. Актуальность

Обработка больших объемов данных, полученных с промышленного оборудования, представляет собой основной инструмент выявления способов повышения эффективности, надежности и долговечности сложных технических устройств от морских ледостойких стационарных платформ [1] до стационарных установок прямого лазерного выращивания [2].

Совокупность теоретических исследований и практических наработок различных подразделений Санкт-Петербургского государственного морского технического университета (СПбГМТУ) в области обработки больших объемов данных позволяет дать оценку различным подходам к предварительной подготовке данных и сформулировать концепцию интеллектуализации анализа данных на базе современных цифровых технологий.

B. Обзор литературы

Преимущества, достижимые в результате применения интеллектуальных технологий в анализе данных мониторинга сложных технических устройств, подтверждаются многочисленными исследованиями.

Так, в [3] рассматриваются проблемы и особенности разработки интеллектуальных систем промышленного интернета вещей (Industrial Internet of Things (IIoT)). Обсуждаются ограничения осуществления обработки на устройствах и методы аналитики больших данных (Big Data Analytics (BDA)) и отмечается необходимость модифицировать процессы BDA при их применении к массивам данных с устройств IIoT.

В [4] авторы отмечают сложности с использованием традиционных систем управления базами данных при хранении и обработке разнородных данных большого объема, полученных с устройств IIoT. В работе обсуждаются особенности хранения данных для

приложений интернета вещей и концепции управления данными с учетом специфики IIoT.

Авторы [5] оценивают пригодность традиционных методов извлечения скрытых знаний при обработке данных с устройств интернета вещей в режиме реального времени. Обсуждаются ограничения в вычислительной мощности устройств, в возможностях своевременной передачи информации, а также ограничения, обусловленные большим объемом данных и большой скоростью их накопления.

В [6] рассматриваются преимущества, полученные в результате внедрения технологий IIoT в промышленной добыче. При обзоре методов анализа разнородных данных ставится акцент на учет имеющихся знаний об этих данных из предметной области.

В [7] обсуждаются архитектура аналитики данных и основные проблемы при обработке больших данных: значительный объем данных из различных источников, большая скорость генерации новых данных, а также влияние указанных факторов на предобработку, анализ, визуализацию и безопасность данных. Рассматриваются типы и техники аналитики применительно к данным с устройств интернета вещей.

Авторы [8] проводят обзор проблем, возникающих при обработке больших данных, полученных с устройств интернета вещей, а также ключевых методов и архитектур извлечения скрытых знаний из данных.

В [9] авторы делают акцент на необходимость интеграции интернета вещей, облачных и граничных вычислений, искусственного интеллекта для достижения поддержки принятия решений и прогнозирования в реальном времени. В работе обсуждаются проблемы оптимизации моделей нейронных сетей с целью ускорения выполнения с использованием формата Open Neural Network Exchange (ONNX).

В [10] рассматриваются возможные применения интеллектуального анализа данных, полученных с устройств интернета вещей, в том числе реализация предиктивной аналитики для выявления причин отказов сложных технических устройств и осуществления их проактивного обслуживания в целях уменьшения времени простоя и увеличения срока службы.

C. Постановка задачи исследования

По результатам проведения самостоятельных исследований и обращения к опыту научного сообщества можно сформулировать следующие ключевые вопросы:

- Что представляют собой данные, полученные с устройств IIoT?

- Какие знания о получаемых данных оказывают влияние на выбор методов их обработки?
- Какие типы и техники анализа данных наиболее актуальны для рассматриваемых устройств IoT?
- Каков масштаб времени анализа данных?
- Какие интеллектуальные технологии позволяют реализовать выбранные типы и техники анализа данных наиболее эффективно?
- Какие методы и средства могут быть использованы для сокращения времени осуществления интеллектуального анализа?

II. ОСНОВНАЯ ЧАСТЬ

A. Предварительная обработка данных

Данные, получаемые с устройств IoT, могут быть различны как по типу, так и по семантике. С одной стороны, это могут быть числовые (целые или вещественные) данные, текстовые и логические данные, данные о дате и времени, визуальные, аудиальные и т. д. С другой стороны, данные могут отражать физические характеристики устройства или его части (мощность устройства, температуру, скорость перемещения и т. д.), местоположение или ориентацию в пространстве, дату или время снятия других данных, сведения об операторе и режимах использования оборудования, об ошибках и отказах.

Предварительная обработка данных по типам данных предполагает выполнение нижеследующих операций (но не ограничивается ими).

Во-первых, принятие решений о сохранении или преобразовании типов: типы данных, отличные от числовых, могут быть преобразованы к числовым, преобразованы к нечисловым либо оставлены в исходном типе. Например, данные типа даты могут быть преобразованы к числу – количеству дней относительно фиксированной даты, а визуальные данные могут быть преобразованы к нечисловому классу посредством применения методов распознавания образов:

$$(data_i)_{new} = (type_i) data_i.$$

Здесь $(data_i)_{new}$ – преобразованные данные, $data_i$ – исходные данные, i – порядковый номер данных, $type_i$ – новый тип данных.

Во-вторых, принятие решений о наличии или отсутствии нормализации числовых данных, а также о способе нормализации [11]. Независимо от выбранного способа нормализации требуется как применять нормализацию отдельно по каждому из нормализуемых данных [12], так и обращать особое внимание на параметры нормализации:

$$(data_i)_{norm} = f_i(data_i, params_i).$$

Здесь $(data_i)_{norm}$ – нормализованные данные, $data_i$ – исходные данные, i – порядковый номер данных, f_i – функция нормализации, $params_i$ – параметры функции нормализации.

Ключевым фактором в определении параметров нормализации (например, минимума и максимума для минимаксной нормализации или среднего значения и дисперсии для Z-нормализации) является совокупный ответ на два связанных друг с другом вопроса:

- Насколько полны имеющиеся данные?
- Какова семантика данных?

Знания о предметной области (о режимах работы устройства IoT, об ограничениях на условия эксплуатации устройства IoT во внешней среде, о технически возможных диапазонах изменения числовых величин и т. д.) позволяют установить семантически корректные параметры нормализации известных данных. Вследствие использования подобного подхода решается проблема априори предполагаемой неполноты анализируемых данных.

В-третьих, принятие решений о методах оценки близости числовых [13] и нечисловых [14] данных. Применяемый способ и его параметры (например, выбор между L_1 и L_2 расстояниями либо определение p для L_p расстояния) определяются семантикой рассматриваемых данных:

$$D_{X,Y} = d_i(X, Y).$$

Здесь X и Y – примеры данных из числовых или нечисловых данных $data_i$, $D_{X,Y}$ – расстояние между X и Y , d_i – способ вычисления расстояния для данных $data_i$, i – порядковый номер данных.

Предварительная обработка данных по семантике данных предполагает постановку следующего вопроса: используется ли семантика данных для их анализа и если да, то насколько?

Помимо рассмотренных выше примеров влияния семантики данных на типы данных, интерес представляют также следующие варианты.

Во-первых, исключение данных даты и (или) времени для формирования временных последовательностей (рядов). Такой подход позволяет реализовать предиктивный анализ данных посредством выявления временных закономерностей в данных.

Во-вторых, учет относительной важности данных с целью их масштабирования. Без знаний о предметной области все данные предполагаются имеющими одинаковую важность.

В-третьих, учет влияния или взаимовлияния данных, с целью исключения некоторых данных из анализа и, таким образом, понижения размерности.

Необходимо обратить особое внимание на достоинства и недостатки каждого из рассмотренных выше вариантов.

Первый вариант открывает возможность для дополнительного исследования данных и получения более сложных зависимостей, однако значительно увеличивает время обнаружения этих зависимостей.

Второй вариант не влияет на время анализа данных, однако представление об относительной важности данных, будучи субъективным, подвержено влиянию

человеческого фактора. Использование коэффициентов масштабирования вследствие различной важности данных требует отдельного обоснования.

Третий вариант сокращает время анализа данных, однако, как и во втором варианте, представление о влиянии или взаимовлиянии данных является субъективным и, таким образом, подвержено влиянию человеческого фактора. Сокращение размерности путем удаления из рассмотрения незначимых или коррелирующих с другими данными требует отдельного обоснования.

В. Особенности постановки задачи анализа данных, полученных с устройств IoT

Вопрос реального времени требует особого рассмотрения в случае мониторинга устройств IoT. Для некоторых устройств реальное время исчисляется секундами, для других – минутами, для третьих – сутками. При выборе методов анализа данных и глубины анализа данных необходимо учитывать знания о предметной области: как функционирует устройство, насколько быстро выбор оптимального или неоптимального режима функционирования приводит к увеличению или уменьшению его эффективности, каков риск ошибки в функционировании или отказа устройства в результате отсутствия или задержки в получении результатов анализа данных и т. д.

Принимая во внимание ограничения вычислительных мощностей устройств IoT, а также возможные ограничения вычислительных мощностей электронно-вычислительных устройств оператора, представляется актуальным разделять анализ данных на этап извлечения скрытых знаний из данных и этап реального времени.

Первый этап предназначен для осуществления диагностического анализа данных и установления закономерностей предиктивного и прескриптивного анализа данных. На втором этапе в режиме мониторинга устройства IoT осуществляется предиктивный анализ данных, а в режиме диалога с оператором – прескриптивный анализ данных.

Иной существенной особенностью анализа данных, полученных с устройств IoT, является большое количество категориальных данных: режимов функционирования, кодов или наименований ошибок или отказов, идентификаторов или имен оператора и т. д.

Более того, опыт анализа данных устройств IoT показал, что числовые непрерывные данные (например, мощность, температура, перемещение и т. д.) ведут себя подобно категориальным. А именно, существенным при мониторинге является не точное значение, а отклонение, приводящее к переходу в другую нечеткую величину (например, с нормальной температуры эксплуатации в высокую температуру или очень высокую). Тем не менее, обычно не представляется возможным заранее выделить соответствующие категории числовых данных.

Возможность использования категориальных данных независимо от их типа позволяет принять метод кластеризации в качестве основы для анализа данных.

С. Обнаружение корреляционных зависимостей

Одной из основных задач корреляционного анализа данных является снижение размерности. Вследствие рассмотренных выше особенностей постановки задачи

анализа данных, полученных с устройств IoT, использование обычных статистических методов для выявления корреляционных зависимостей между числовыми или категориальными данными недостаточно эффективно.

Кластеризация данных позволяет выявить количество кластеров и принадлежность данных этим кластерам. Кластеризации подвергаются числовые данные для приведения их к категориальным, категориальные для выявления близости между категориями, а также возможные совокупности как числовых, так и категориальных данных.

Принцип обнаружения корреляционной зависимости между данными может быть сформулирован следующим образом: если при добавлении данных $data_j$ к данным $data_i$ результат их совместной кластеризации C_{ij} близок результату кластеризации C_i , то имеется корреляция данных $data_i$ и $data_j$.

Следует подчеркнуть, что корреляция оценивается между кластерами соответствующих данных. Наличие корреляции данных $data_i$ и $data_j$ внутри кластеров не оценивается.

Для оценки близости результатов кластеризации может быть осуществлен как поэлементный анализ кластеров, так и оценка таблицы сопряженности: если она является перестановочно-эквивалентной матрице с диагональным превосходством, то кластеры близки. Диагональное превосходство может достигаться либо диагональным преобладанием, либо иной выбранной мерой (например, когда диагональный элемент отличается от суммы элементов по строке (столбцу) не более чем на $p\%$, либо когда собственные значения анализируемой матрицы отличны от диагональных элементов не более чем на $p\%$, где p определяется желаемой силой корреляции, т. е. чем p меньше, тем корреляция сильнее).

Результаты кластеризации с различным числом кластеров могут быть признаны близкими, если существует такой способ объединения кластеров для результата с большим количеством кластеров с целью приведения к меньшему количеству кластеров, для которого результаты кластеризации близки по указанному в предыдущем абзаце методу.

В случае выявления корреляционных зависимостей осуществляется снижение размерности, при этом оставляемые и исключаемые данные определяются с учетом семантики данных (при наличии).

Д. Обнаружение регрессионных зависимостей

В том случае, если семантика данных показывает наличие причинно-следственной связи в данных, для которых обнаружена корреляционная зависимость, можно говорить об обнаружении регрессионной зависимости. Вместо исключения таких данных из анализа они используются в предиктивном и прескриптивном анализе.

Если корреляционная зависимость обнаружена для многомерных кластеров (т. е. вместо сравнения C_{ij} с C_i осуществляется сравнение, например, C_{ijkl} с C_{ijk}), может быть сделан вывод об обнаружении регрессионной зависимости (в соответствии с примером ранее, $data_i$ будет зависеть от $data_i, data_j, data_k$).

Е. Обнаружение авторегрессионных зависимостей

При наличии данных, имеющих тип даты и (или) времени, осуществляется выделение прочих данных и формирование наборов данных на основе метода скользящего окна [15].

Параметр задержки d определяет размер нового набора данных:

$$(data_i)_t, \dots, (data_i)_{t-d}, (data_k)_t, \dots, (data_k)_{t-d}$$

Здесь $data_i$ – прогнозируемые данные, $data_k$ – экзогенные данные, отличные от прогнозируемых.

Ф. Нейросетевые методы кластеризации

Корреляционные и регрессионные зависимости категориальных (или приведенных к ним) данных реализуются методами нейросетевой кластеризации.

Наибольший эффект показали нейронные сети соревновательного слоя [16] и самоорганизующиеся карты Кохонена [17]. Для устранения «мертвых» нейронов-кластеров используются штрафы.

Определение количества кластеров осуществляется путем оценки стабильности разделения [18].

Г. Нейросетевые методы прогноза

Модель прогноза формируется путем обучения искусственной нейронной сети NARX (Nonlinear Autoregressive Exogenous) [19]. Параметр d и данные $data_k$ выбираются под конкретную поставленную задачу для максимизации качества обучения с учетом семантики данных.

Н. Концепция применения нейросетевых методов

Нейросетевые методы для проведения анализа данных, полученных с устройств IoT, выбираются исходя из следующего:

- Семантика данных используется для их преобразования и нормализации. Семантика данных не используется для масштабирования, сокращения размерности и априорного установления зависимостей между данными.
- Предпочтительной является работа с категориальными данными. Числовые непрерывные данные преобразуются к категориальным – методами нейросетевой кластеризации.
- Обнаружение корреляционных и регрессионных зависимостей осуществляется нейросетевой кластеризацией.
- Обнаружение автокорреляционных зависимостей осуществляется нейронными сетями прогноза временных рядов.
- Подготовка категориальных данных к использованию в сети прогноза временных рядов (например, в NARX-сети) осуществляется по методу One-Hot Encoding.

И. Оптимизация нейронных сетей

Оптимизация нейронных сетей реализована путем их представления в соответствии со стандартом ONNX.

Для выполнения нейронных сетей используется ONNX Runtime (ORT), что позволяет полноценно использовать преимущества аппаратного ускорения.

Л. Пример реализации

Одним из примеров реализации приведенной концепции нейросетевого анализа данных для устройств промышленного интернета вещей является анализ данных со стационарной установки прямого лазерного выращивания.

Пользовательский интерфейс визуализации результатов кластеризации приведен на рис. 1.



Рис. 1. Пользовательский интерфейс визуализации результатов кластеризации

III. ЗАКЛЮЧЕНИЕ

В результате исследования была сформулирована концепция выбора и применения нейронных сетей в задачах анализа данных, полученных с устройств IoT. Установлено, что предварительная обработка полученных данных является ключевым процессом, оказывающим непосредственное влияние на результат. Выявлены особенности, характерные для данных IoT, и даны рекомендации по применению методов предварительной обработки. Предложены нейросетевые методы и конкретные нейронные сети для реализации анализа данных, эффективность которых подтверждена программной реализацией.

В качестве перспектив исследования следует выделить использование моделей генеративного искусственного интеллекта для предварительной обработки данных, а также применение технологии интернета агентов Eclipse LMOSS (Language Model Operating System) с целью распределения подзадач по интеллектуальным агентам (например, при определении количества кластеров). Кроме того, интересным представляется переход от бинарного представления категориальных данных к нечеткому представлению.

СПИСОК ЛИТЕРАТУРЫ

- [1] Применение методов анализа данных к результатам работы системы мониторинга состояния морского инженерного сооружения / В.А. Коршунов, О.Н. Петров, Д.А. Пономарев, А.А. Родионов // Морские интеллектуальные технологии. 2020. № 4-3 (50). С. 73–81.
- [2] Система удалённого мониторинга, сбора и анализа больших данных промышленного оборудования на основе технологий интернета вещей и искусственного интеллекта : свидетельство о гос. регистрации прогр. для ЭВМ № 2025697267 Рос. Федерация / О.Н. Петров, К.И. Андреев, М.А. Вохминцев, Н. Вайнт ; правообладатель ФГБОУ ВО «С.-Петерб. гос. мор. техн. ун-т». № 2025696320 ; заявл. 11.12.2025 ; опубл. 23.12.2025.
- [3] The role of big data analytics in industrial internet of things / I. Yaqoob, I.A.T. Hashem, A. Gani, S. Mokhtar, E. Ahmed, N.B. Anuar, S.T. Khan // arXiv preprint arXiv:1904.05556. 2019. 35 p.
- [4] Data management techniques for Internet of Things / B. Diène, J. J. P. C. Rodrigues, O. Diallo, E.H.M. Ndoye, V.V. Korotaev // Mechanical Systems and Signal Processing. 2020. Vol. 138. P. 106564. — DOI: 10.1016/j.ymssp.2019.106564.
- [5] A Survey on Behavioral Pattern Mining From Sensor Data in Internet of Things / M.M. Rashid, J. Kamruzzaman, M.M. Hassan, S.S.

- Shafin, M. Z. A. Bhuiyan // IEEE Access. 2020. Vol. 8. P. 33318–33341. — DOI: 10.1109/ACCESS.2020.2974035.
- [6] Schmitz, C. Mining 4.0 and Big Data – How Implementing IoT Systems can increase Efficiency underground / C. Schmitz, M. Liebetrau // GeoResources Journal. 2020. № 2. P. 45–49. — URL: <https://www.georesources.net/download/GeoResources-Journal-2-2020.pdf> (дата обращения: 12.04.2026).
- [7] Abughazala M. Modeling Data Analytics Architecture for Data-Driven IoT Applications using DAT / M. Abughazala, H. Muccini // 2023 IEEE 20th International Conference on Software Architecture Companion (ICSA-C). L'Aquila, Italy, 2023. P. 284–291. — DOI: 10.1109/ICSA-C57050.2023.00066.
- [8] Irajy M.S. Big Data Mining Techniques In IoT, Challenges and Architectures / M.S. Irajy, A. Yavari // Transactions on Soft Computing. 2025. Vol. 1, № 1. P. 46–60. — DOI: 10.48314/tsc.vi.38.
- [9] From Sensors to Data Intelligence: Leveraging IoT, Cloud, and Edge Computing with AI / I. Ficili, M. Giacobbe, G. Tricomi, A. Puliafito // Sensors. 2025. Vol. 25, № 6. P. 1763. — DOI: 10.3390/s25061763.
- [10] Amelie J. Integrating Data Mining with Internet of Things (IoT) for Smarter Analytics // Journal of Computer Science & Systems Biology. 2025. Vol. 18, № 2. P. 579. — URL: <https://www.hilarispublisher.com/open-access/integrating-data-mining-with-internet-of-things-iot-for-smarter-analytics-114633.html> (дата обращения: 12.04.2026).
- [11] Singh D. Investigating the impact of data normalization on classification performance / D. Singh, B. Singh // Applied Soft Computing. 2020. Vol. 97, Pt. B. P. 105524. — DOI: 10.1016/j.asoc.2019.105524.
- [12] Singh D. Feature wise normalization: An effective way of normalizing data / D. Singh, B. Singh // Pattern Recognition. 2022. Vol. 122. P. 108307. — DOI: 10.1016/j.patcog.2021.108307.
- [13] Towards robust discriminative projections learning via non-greedy $\ell_{2,1}$ -norm MinMax / F. Nie, Z. Wang, R. Wang, Z. Wang, X. Li // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2021. Vol. 43, № 6. P. 2086–2100. — DOI: 10.1109/TPAMI.2019.2961877.
- [14] Lp- and Ls-Norm Distance Based Robust Linear Discriminant Analysis / Q. Ye, L. Fu, Z. Zhang, H. Zhao, M. Naiem // Neural Networks. 2018. Vol. 105. P. 393–404. — DOI: 10.1016/j.neunet.2018.05.020.
- [15] Dietterich T.G. Machine Learning for Sequential Data: A Review / T. G. Dietterich // Proceedings of the Joint IAPR International Workshop on Structural, Syntactic, and Statistical Pattern Recognition. Berlin; Heidelberg: Springer-Verlag, 2002. P. 15–30.
- [16] Rumelhart D.E. Feature Discovery by Competitive Learning / D. E. Rumelhart, D. Zipser // Cognitive Science. 1985. Vol. 9, № 1. P. 75–112. — DOI: 10.1016/S0364-0213(85)80010-0.
- [17] Kohonen T. Self-organized formation of topologically correct feature maps // Biological Cybernetics. 1982. Vol. 43, № 1. P. 59–69. — DOI: 10.1007/BF00337288.
- [18] Hennig C. Cluster-wise assessment of cluster stability // Computational Statistics & Data Analysis. 2007. Vol. 52, № 1. P. 258–271. — DOI: 10.1016/j.csda.2006.11.025.
- [19] Menezes Jr, J. M. Long-term time series prediction with the NARX network: An empirical evaluation / J. M. Menezes Jr, G. A. Barreto // Neurocomputing. 2008. Vol. 71, № 16-18. P. 3335–3343. — DOI: 10.1016/j.neucom.2008.01.030.