Пояснительная записка
к первой редакции национального стандарта
ГОСТ Р «Информационные технологии. Искусственный интеллект. Оценка робастности нейронных сетей. Часть 1. Обзор»

1. Основание для разработки проекта стандарта

Национальный стандарт разработан в соответствии с Программой национальной стандартизации на 2021 год, утверждённой Федеральным агентством по техническому регулированию и метрологии.

Шифр задания по ПНС-2021: 1.11.164-1.096.21.

2. Краткая характеристика объекта и аспекта стандартизации

Настоящий стандарт представляет определяет эталонную модель рабочего процесса для определения робастности нейронной сети, на основе которого строятся методы и метрики, применимые к различным типам экспериментов, то есть статистические, формальные и эмпирические методы.

Стандарт сначала рассматривает статистические подходы, которые являются наиболее распространенными подходами, используемыми для оценки робастности. Для них характерен подход к тестированию, определяемый методологией с использованием математических метрик. Затем исследуются подходы к получению формального доказательства, которые все чаще используются для оценки робастности. Наконец, в документе представлены эмпирические подходы, основанные на субъективных наблюдениях, которые дополняют оценку робастности, когда статистические и формальные подходы недостаточны или нецелесообразны

3. Технико-экономическое, социальное или иное обоснование целесообразности разработки стандарта на национальном уровне

Разработка национального стандарта включена в Перспективную программу стандартизации по приоритетному направлению «Искусственный интеллект» на период 2021–2024 годы, утверждённую Министерством экономического развития Российской Федерации и Федеральным агентством по техническому регулированию и метрологии 22.12.2020 г.

Данный стандарт предоставляет информацию о существующих методах оценки робастности нейронных сетей. Хотя существует множество методов валидации систем, не связанных с искусственным интеллектом, они не всегда непосредственно применимы к системам искусственного интеллекта и, в частности, к нейронным сетям. Архитектуры нейронных сетей представляют собой особую проблему, поскольку их трудно объяснить, и иногда они ведут себя неожиданным образом из-за их нелинейной природы. В результате необходимы новые подходы. Отмечается, что характеристика робастности нейронных сетей является открытой областью исследований, и существуют ограничения как для подходов тестирования, так и для верификации.

Методы, представленные стандартом, позволяют ответить на актуальные вопросы по валидации системы искусственного интеллекта. Например, статистические методы позволяют проверить, достигают ли свойства системы желаемого целевого порога. Формальные методы позволяют проверить, доказуемы ли свойства в области использования. Эмпирические методы позволяют оценить степень, в которой свойства системы поддерживаются в тестируемом сценарии.

Кроме того, данный стандарт позволит заинтересованным сторонам использовать единый терминологический аппарат, повысит скорость распространения, единство восприятия информации, повысит стабильность терминологии, создаст предпосылки для взаимного проникновения

отечественных и мировых исследований в области «Информационные технологии – Искусственный интеллект».

4. Сведения о соответствии проекта национального стандарта техническим регламентам Евразийского экономического союза, федеральным законам, техническим регламентам и иным нормативным правовым актам Российской Федерации, которые содержат требования к объекту и/или аспекту стандартизации

Проект национального стандарта соответствует техническим регламентам Евразийского экономического союза и законодательству Российской Федерации.

5. Сведения о соответствии проекта национального стандарта международному стандарту, региональному стандарту, региональному своду правил, стандарту иностранного государства и своду правил иностранного государства, иному документу по стандартизации иностранного государства и о форме применения данного стандарта (документа) как основы для разработки проекта национального стандарта Российской Федерации, а в случае отклонения от международного стандарта, регионального стандарта, регионального свода правил, стандарта иностранного государства и свода правил иностранного государства, иного документа по стандартизации иностранного государства - мотивированное обоснование этого решения и/или иные сведения о научно-техническом уровне проекта национального стандарта

Первая редакция национального стандарта идентична положениям разрабатываемого международного стандарта ISO/IEC TR 24029-1 Information technology — Artificial Intelligence (AI) — Assessment of the robustness of neural networks — Part 1: Overview.

6. Сведения о проведённых научно-исследовательских работах, технических предложениях, опытно-конструкторских, опытно-технологических и проектных работах, а также аналитических работах, послуживших основой для разработки первой редакции проекта национального стандарта

Проект национального стандарта разработан на основе выполненных научно-исследовательских работ в рамках проекта «Мониторинг и стандартизация развития и использования технологий хранения и анализа больших данных в цифровой экономике Российской Федерации», реализуемого научно-учебной лабораторией методов анализа больших данных Федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский университет «Высшая школа экономики» (НИУ ВШЭ).

7. Сведения о наличии в Федеральном информационном фонде стандартов переводов международных, региональных стандартов, стандартов и сводов правил иностранных государств, на которые даны нормативные ссылки в стандарте, использованном в качестве основы для разработки проекта национального стандарта Российской Федерации

Проект национального стандарта использует утверждённые действующие национальные стандарты, идентичные международным и региональным стандартам, либо имеющиеся в Федеральном информационном фонде стандартов переводы соответствующих стандартов.

8. Сведения о взаимосвязи проекта национального стандарта с проектами или действующими в Российской Федерации другими национальными и межгосударственными стандартами, сводами правил, а при необходимости также предложения по их пересмотру, изменению

или отмене (одностороннему прекращению применения на территории Российской Федерации межгосударственных стандартов)

Проект национального стандарта взаимосвязан в части терминологии с национальным стандартом ГОСТ Р ИСО/МЭК 20546–2021 «Информационные технологии. Большие данные. Обзор и словарь», вводимым в действие с 1 ноября 2021 года.

Утверждение национального стандарта не потребует пересмотра, изменения или отмены (одностороннего прекращения применения на территории Российской Федерации) международных, региональных и национальных стандартов.

9. Перечень исходных документов и другие источники информации, использованные при разработке стандарта, в том числе информация об использовании документов, относящихся к объектам патентного или авторского права

Проект национального стандарта не имеет исходных документов и других источников информации, использованных при разработке стандарта, в том числе не использует документов, относящихся к объектам патентного или авторского права. Источником информации является разрабатываемый международный стандарт ISO/IEC TR 24029-1 Information technology — Artificial Intelligence (AI) — Assessment of the robustness of neural networks — Part 1: Overview и его перевод на русский язык, выполненный самостоятельно разработчиком и утверждённый в установленном порядке.

10. Сведения о технических комитетах по стандартизации, в областях деятельности которых возможно пересечение с областью применения разрабатываемого проекта национального.

Технические комитеты по стандартизации в смежной области деятельности отсутствуют.

11. Сведения о разработчиках стандарта

Проект национального стандарта разработан федеральным государственным автономным образовательным учреждением высшего образования «Национальный исследовательский университет «Высшая школа экономики» (НИУ ВШЭ) в лице Научно-учебной лаборатории методов анализа больших данных.

Контактная информация

Электронная почта: austyuzhanin@hse.ru

Номер телефона: +7(495) 772-9590

Руководитель организации-разработчика

заведующий научно-учебной лабораторией методов анализа больших данных НИУ ВШЭ, к.ф.-м.н., доцент _____ А.Е.Устюжанин

Руководитель разработки

эксперт научно-учебной лаборатории методов анализа больших данных, к.ф.-м.н. _____ Ю.И.Пахомов