ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р 70984— 2023

СИСТЕМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА НА АВТОМОБИЛЬНОМ ТРАНСПОРТЕ

Системы управления интеллектуальной транспортной инфраструктурой. Требования к испытанию алгоритмов прогнозирования дорожных условий

Издание официальное

Москва
Российский институт стандартизации
2023

gost.e-ecolog.ru

Предисловие

1 РАЗРАБОТАН Обществом с ограниченной ответственностью «А+С Транспроект» (ООО «А+С Транспроект»), Обществом с ограниченной ответственностью «А-Я эксперт» (ООО «А-Я эксперт»)

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 164 «Искусственный интеллект»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 17 октября 2023 г. № 1171-ст

4 ВВЕДЕН ВПЕРВЫЕ

*Правила применения настоящего стандарта установлены в статье 26 Федерального закона от 29 июня 2015 г. № 162-ФЗ «О стандартизации в Российской Федерации». Информация об изменениях к настоящему стандарту публикуется в ежегодном (по состоянию на 1 января текущего года) информационном указателе «Национальные стандарты», а официальный текст изменений и поправок — в ежемесячном информационном указателе «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ближайшем выпуске ежемесячного информационного указателя «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет (*gost.e-ecolog.ru

© Оформление. ФГБУ «Институт стандартизации», 2023

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

Введение

Развитие технологий искусственного интеллекта (ИИ) на прикладном уровне происходит на всех уровнях жизни. На транспорте эти технологии используются в нескольких аспектах — для построения интеллектуальных транспортных систем на различных уровнях управления и для интеллектуализации систем управления интеллектуальной транспортной инфраструктурой (СУИТИ), которые входят в состав интеллектуальных транспортных систем в качестве подсистем. В рамках автомобильной транспортной модальности функционирование СУИТИ основано на применении ряда алгоритмов ИИ для прогнозирования дорожных условий, прогнозирования параметров транспортных потоков.

Вместе с этим СУИТИ проходят широкие испытания. От полноты и достоверности процесса испытаний зависят перспективы этой технологии и сроки начала ее массового использования на автомобильных дорогах общего пользования.

Для решения задач организации и управления дорожным движением одним из алгоритмов является алгоритм прогнозирования дорожных условий. При использовании данных из различных источников СУИТИ могут использовать алгоритм прогнозирования дорожных условий для организации управления дорожным движением на эксплуатируемой улично-дорожной сети населенных пунктов или эксплуатируемых региональных и межмуниципальных автомобильных дорогах, автомобильных дорогах федерального значения.

Для испытания алгоритмов прогнозирования дорожных условий в целях обеспечения доверия к СУИТИ, основанным на использовании методов ИИ, настоящий стандарт устанавливает общие принципы проведения испытаний. В настоящем стандарте приведен перечень весовых коэффициентов для показателей качества алгоритма и тестовых наборов данных с приведением сценариев испытаний описываемого алгоритма. Приведены требования к представительности (полноте и несмещенности) тестовых данных, демонстрационные тестовые наборы данных, правила формирования представительных тестовых наборов данных и принципы расширения (аугментации) тестовых наборов данных.

Настоящий стандарт является частью комплекса стандартов по установлению требований к применению технологий ИИ на автомобильной транспортной модальности для повышения доверия к технологиям ИИ, обеспечения безопасности дорожного движения, жизни и здоровья людей, сохранности их имущества, охраны окружающей среды и эффективности транспортных процессов.

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

СИСТЕМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА НА АВТОМОБИЛЬНОМ ТРАНСПОРТЕ

Системы управления интеллектуальной транспортной инфраструктурой. Требования к испытанию алгоритмов прогнозирования дорожных условий

Artificial intelligence systems in road transport.
Intelligent transport infrastructure management systems. Requirements for testing road condition prediction algorithms

Дата введения — 2024—01—01

1 Область применения

Настоящий стандарт распространяется на СУИТИ и устанавливает требования к испытанию частных алгоритмов, реализованных с использованием методов ИИ, подсистемы интерпретации входных данных о дорожной обстановке в автомобильной транспортной модальности — алгоритмов прогнозирования дорожных условий в СУИТИ.

СУИТИ, в состав которых входят алгоритмы прогнозирования дорожных условий с использованием методов ИИ, требования к испытаниям которых установлены в настоящем стандарте, могут быть реализованы в составе интеллектуальных транспортных систем, управляющих дорожной деятельностью в населенных пунктах, городских агломерациях, на сети региональных и межмуниципальных дорог, федеральных автомобильных дорог.

Настоящий стандарт предназначен для применения при проведении всех типов испытаний алгоритмов прогнозирования дорожных условий в СУИТИ.

2 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты:

ГОСТ Р 70249 Системы искусственного интеллекта на автомобильном транспорте. Высокоавтоматизированные транспортные средства. Термины и определения

ГОСТ Р 70250—2022 Системы искусственного интеллекта на автомобильном транспорте. Варианты использования и состав функциональных подсистем искусственного интеллекта

Примечание — При пользовании настоящим стандартом целесообразно проверить действие ссылочных стандартов в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет или по ежегодному информационному указателю «Национальные стандарты», который опубликован по состоянию на 1 января текущего года, и по выпускам ежемесячного информационного указателя «Национальные стандарты» за текущий год. Если заменен ссылочный стандарт, на который дана недатированная ссылка, то рекомендуется использовать действующую версию этого стандарта с учетом всех внесенных в данную версию изменений. Если заменен ссылочный стандарт, на который дана датированная ссылка, то рекомендуется использовать версию этого стандарта с указанным выше годом утверждения (принятия). Если после утверждения настоящего стандарта в ссылочный стандарт, на который дана датированная ссылка, внесено изменение, затрагивающее положение, на которое дана ссылка, то это положение рекомендуется применять без учета данного изменения. Если ссылочный стандарт отменен без замены, то положение, в котором дана ссылка на него, рекомендуется применять в части, не затрагивающей эту ссылку.

Издание официальное

3 Термины и определения

В настоящем стандарте применены термины по ГОСТ Р 70249, а также следующие термины с соответствующими определениями:

3.1 **автомобильная транспортная модальность:** Транспортная модальность, использующая колесные транспортные средства для перемещения по автомобильным дорогам общего пользования.

3.2

автономность (autonomy): Характеристика системы искусственного интеллекта, связанная с ее способностью самостоятельно (без участия человека) выполнять возложенные на нее функции в течение заданного времени и с заданными показателями качества, надежности, безопасности.

[Адаптировано из ГОСТ Р 59277—2020, пункт 3.3]

3.3

адаптируемость: Характеристика системы искусственного интеллекта, связанная с ее способностью в процессе функционирования подстраиваться под изменяющиеся условия эксплуатации и (или) функционирования без необходимости дорабатывать систему на уровне изменения исходного кода.

[ГОСТ Р 70980—2023, пункт 3.2]

3.4

аугментация данных: Способ увеличения количества обучающих данных путем модификации имеющихся данных.

[ГОСТ Р 70983—2023, пункт 3.4]

3.5 **вейвлет-преобразование:** Интегральное преобразование, которое представляет собой свертку вейвлет-функции с сигналом.

Примечание — Вейвлет-преобразование переводит сигнал из временного представления в частотно-временное.

3.6

доверенность: Показатель качества алгоритма искусственного интеллекта, отражающий степень проявления свойства доверия к системе искусственного интеллекта.

[ГОСТ Р 70983—2023, пункт 3.5]

3.7

доверие к системе искусственного интеллекта: Уверенность потребителя и, при необходимости, организаций, ответственных за регулирование вопросов создания и применения систем искусственного интеллекта, и иных заинтересованных сторон в том, что система способна выполнять возложенные на нее задачи с требуемым качеством.

[ГОСТ Р 59276—2020, пункт 3.3]

3.8

интеллектуализация: Повышение степени интеллектуальности системы искусственного интеллекта, то есть увеличение количества интеллектуальных функций в системе искусственного интеллекта.

[ГОСТ Р 70980—2023, пункт 3.6]

3.9

интеллектуальная транспортная инфраструктура: Транспортная инфраструктура, в составе которой имеются системы искусственного интеллекта.

[ГОСТ Р 70980—2023, пункт 3.3]

3.10

интеллектуальная транспортная система: Система управления, интегрирующая современные информационные и телематические технологии и предназначенная для автоматизированного поиска и принятия к реализации максимально эффективных сценариев управления транспортно-дорожным комплексом региона, конкретным транспортным средством или группой транспортных средств с целью обеспечения заданной мобильности населения, максимизации показателей использования дорожной сети, повышения безопасности и эффективности транспортного процесса, комфортности для водителей и пользователей транспорта.

[ГОСТ Р 56829—2015, статья 1]

3.11

интеллектуальность: Наличие у системы искусственного интеллекта свойств автономности и адаптируемости, а также реализация в системе искусственного интеллекта интеллектуальных функций.

[ГОСТ Р 70980—2023, пункт 3.4]

3.12

интеллектуальные функции: Функциональные возможности, которые позволяют технической системе использовать технологии искусственного интеллекта для более эффективной обработки и анализа данных, а также для решения сложных проблем, свойственных когнитивным способностям человека.

[ГОСТ Р 70980—2023, пункт 3.5]

3.13

исходный код: Компьютерная программа в текстовом виде на каком-либо языке программирования.

[ГОСТ Р 54593—2011, пункт 3.4]

3.14

функциональная корректность: Степень обеспечения продуктом или системой необходимой степени точности корректных результатов.

[ГОСТ Р ИСО/МЭК 25010—2015, пункт 4.2.1.2]

3.15

надежность: Свойство объекта сохранять во времени способность выполнять требуемые функции в заданных режимах и условиях применения, технического обслуживания, хранения и транспортирования.

[Адаптировано из ГОСТ Р 27.102—2021, статья 5]

3.16

переобученность: Свойство модели машинного обучения хорошо классифицировать примеры из обучающей выборки, но относительно плохо классифицировать примеры, не участвовавшие в обучении (например, из тестовой выборки).

[ГОСТ Р 70983—2023, пункт 3.13]

3.17 **прогнозирование дорожных условий:** Оценка изменения дорожной обстановки на основе анализа существенных факторов эксплуатации.

3.18

система искусственного интеллекта: Техническая система, в которой используются технологии искусственного интеллекта и обладающая искусственным интеллектом.

[ГОСТ Р 59276—2020, пункт 3.16]

3.19

система управления интеллектуальной транспортной инфраструктурой: Система управления, объектом управления которой является интеллектуальная транспортная инфраструктура.

[ГОСТ Р 70980—2023, пункт 3.10]

3.20

сопровождаемость: Результативность и эффективность, с которыми продукт или система могут быть модифицированы предполагаемыми специалистами по обслуживанию.

[Адаптировано из ГОСТ Р ИСО/МЭК 25010—2015, пункт 4.2.7]

3.21

существенный фактор эксплуатации: Важный входной параметр для алгоритма искусственного интеллекта, который относится к решаемой задаче и изменение значения которого существенно влияет на результат работы алгоритма.

[ГОСТ Р 70983—2023, пункт 3.15]

3.22 **сэмплирование:** Метод корректировки обучающей или тестовой выборки с целью балансировки распределения классов в исходном наборе данных.

3.23

транспортная инфраструктура: Технологический комплекс, предназначенный для перевозки пассажиров и перевалки (перевозки, транспортирования) грузов повышенной опасности в установленном порядке и включающий в себя совокупность объектов (зданий, сооружений, коммуникаций, устройств, оборудования) и транспортных средств.

[ГОСТ Р 56461—2015, пункт 3.2]

3.24

транспортная модальность: Вид транспорта, включая типы транспортных средств и транспортной инфраструктуры, а также среду, в которой функционирует транспорт.

[ГОСТ Р 70983—2023, пункт 3.16]

3.25

удобство применения: Степень, в которой продукт или система могут быть использованы определенными пользователями для достижения конкретных целей с эффективностью, результативностью и удовлетворенностью в заданном контексте использования.

[ГОСТ Р ИСО/МЭК 25010—2015, пункт 4.2.4]

3.26

эффективность: Точность и полнота, с которой пользователи достигают определенных целей.

[Адаптировано из ГОСТ Р ИСО/МЭК 25010—2015, пункт 4.1.1]

4 Общие требования

4.1 Общие требования и методику проведения испытаний алгоритмов прогнозирования дорожных условий определяют по ГОСТ Р 70250.

4.2 При проведении испытаний алгоритмов прогнозирования дорожных условий необходимо провести экспертизу на патентную чистоту алгоритмов.

5 Показатели и критерии качества алгоритма прогнозирования дорожных условий

Организация, осуществляющая тестирование алгоритма прогнозирования дорожных условий, должна применять показатели и критерии для проведения оценки качества этого алгоритма, установленные в ГОСТ Р 70250—2022 (раздел 8).

6 Весовые коэффициенты для оценки алгоритма прогнозирования дорожных условий

6.1 Для конкретизации процедуры оценки качества алгоритма прогнозирования дорожных условий в таблицах 1—5 приведены весовые коэффициенты для критериев и метрик качества. Весовые коэффициенты для критериев представлены в первой графе. Сумма всех коэффициентов первой графы должна быть равна 1. Весовые коэффициенты для всех метрик качества заданного критерия представлены в строке соответствующего критерия. Сумма всех коэффициентов для каждой строки должна быть равна 1. Знак «#» в наименовании метрик является «маской подстановки» для номера критерия, к которому относится соответствующая метрика. Например, для фактора качества «Надежность» метрика «Н#-1» для критерия «Н1» должна читаться как «Н1-1», а для критерия «Н2»—«Н2-1».

6.2 Конкретные весовые коэффициенты для критериев и метрик фактора качества «Надежность» приведены в таблице 1.

Таблица 1 — Весовые коэффициенты фактора качества «Надежность»

Вес критерия	Критерий	Метрика 1 Н#-1	Метрика 2 Н#-2	Метрика 3 Н#-3
0,5	Н1	0,3	0,5	0,2
0,5	Н2	0,25	0,75	—

6.3 Конкретные весовые коэффициенты для критериев и метрик фактора качества «Сопровождаемость» приведены в таблице 2.

Таблица 2 — Весовые коэффициенты фактора качества «Сопровождаемость»

Вес критерия	Критерий	Метрика 1 С#-1	Метрика 2 С#-2	Метрика 3 С#-3	Метрика 4 С#-4
0,6	С2	0,1	0,4	0,4	0,1
0,4	С3	0,75	0,15	0,1	—

6.4 Конкретные весовые коэффициенты для критериев и метрик фактора качества «Удобство применения» приведены в таблице 3.

Таблица 3 — Весовые коэффициенты фактора качества «Удобство применения»

Вес критерия	Критерий	Метрика 1 У#-1	Метрика 2 У#-2	Метрика 3 У#-3	Метрика 4 У#-4	Метрика 5 У#-5
0,3	У1	0,6	0,4	—	—	—
0,2	У2	0,3	0,3	0,2	0,1	0,1
0,5	У3	0,2	0,15	0,5	0,15	—

6.5 Для фактора качества «Эффективность» используются следующие весовые коэффициенты: для критериев Э2 и Э3 — 0,25, для Э4 — 0,5.

6.6 Конкретные весовые коэффициенты для критериев и метрик фактора качества «Корректность» приведены в таблице 4.

Таблица 4 — Весовые коэффициенты фактора качества «Корректность»

Вес критерия	Критерий	Метрика 1 К#-1	Метрика 2 К#-2	Метрика 3 К#-3	Метрика 4 К#-4	Метрика 5 К#-5	Метрика 6 К#-6	Метрика 7 К#-7	Метрика 8 К#-8
0,1	К1	0,5	0,5	—	—	—	—	—	—
0,2	К2	0,2	0,05	0,1	0,05	0,1	0,3	0,1	0,1
0,3	К3	0,3	0,4	0,3	—	—	—	—	—
0,4	К4	1,0	—	—	—	—	—	—	—

6.7 Конкретные весовые коэффициенты для критериев и метрик фактора качества «Доверенность» приведены в таблице 5.

Таблица 5 — Весовые коэффициенты фактора качества «Доверенность»

Вес критерия	Критерий	Метрика 1 Д#-1	Метрика 2 Д#-2	Метрика 3 Д#-3	Метрика 4 Д#-4
0,5	Д1	0,4	0,4	0,15	0,05
0,5	Д2	0,05	0,3	0,15	0,5

6.8 Представленные в таблицах 1—5 весовые коэффициенты для критериев и метрик являются рекомендуемыми, однако при испытании частного алгоритма для конкретной задачи можно выбирать специфические коэффициенты для этой задачи.

7 Тестовые наборы данных и сценарии испытания алгоритма прогнозирования дорожных условий

7.1 Общее описание демонстрационных наборов данных

7.1.1 В настоящем разделе описаны тестовые наборы данных и сценарии испытания алгоритма прогнозирования дорожных условий, а именно приведены требования к представительности (полноте и несмещенности) тестовых наборов данных, приведены фрагменты тестовых наборов данных — демонстрационные наборы данных, описаны правила формирования представительных тестовых наборов данных, включая, в случае необходимости, описание представительной совокупности тестовых сценариев, а также разъяснены принципы расширения (аугментации) тестовых наборов данных.

7.1.2 При подготовке тестовых наборов данных для проведения испытаний алгоритмов прогнозирования дорожных условий необходимо обеспечить единство измерений в части обеспечения единообразия средств измерений в соответствии с требованиями к эталонам Государственной системы обеспечения единства измерений.

7.2 Требования к представительности (полноте и несмещенности) тестовых наборов данных

7.2.1 Тестовый набор данных должен быть репрезентативен, содержать целевой набор значений характеристик и ситуаций по отношению к каждому существенному фактору эксплуатации.

7.2.2 Для обеспечения представительности тестового набора данных необходимо:

а) проводить испытания алгоритма прогнозирования дорожных условий на тестовом наборе, подготовленном в соответствии со статистическими закономерностями распределения значений существенных факторов эксплуатации, что позволит обеспечить тестирование алгоритма на соответствие достоверной, объективной, правдивой модели распределения дорожно-транспортных ситуаций;

б) обеспечить повторные испытания со специально смещенной выборкой относительно количества ситуаций независимо от их реального статистического распределения. Для этих целей можно воспользоваться следующей процедурой:

1) выделить ситуации, частота встречаемости которых в реальном окружении периферийного оборудования СУИТИ для сбора данных о дорожных условиях ниже средней частоты встречаемости всех ситуаций более, чем на три стандартных отклонения (редкие ситуации), если закон распределения соответствующей характеристики позволяет использовать критерий трех стандартных отклонений, в противном случае использовать критерии редкости, соответствующие выявленному закону распределения.

Примечание — Уровень определения редких ситуаций является рекомендуемым и может быть изменен,

2) для выделенных ситуаций подготовить для испытаний «смещенную» выборку, которая позволяет лучше проанализировать поведение алгоритма в части редких ситуаций — например, путем сэмплирования с изменением частоты соответствующих ситуаций, которая приближена к среднему по всем вариантам ситуаций, или путем ввода весовых коэффициентов при оценке

ошибки прогнозирования. Конкретный способ должен быть явно определен в программе испытаний алгоритма до начала этих испытаний,

3) провести испытания с подготовленной выборкой, чтобы убедиться, что алгоритм прогнозирует дорожные условия с достаточным качеством (см. ГОСТ Р 70250).

7.2.3 Такой принцип подбора тестового набора данных позволит обеспечить проверку возможностей алгоритма по прогнозированию всех ситуаций в различных контекстах (вариантах комбинаций значений существенных факторов эксплуатации), поскольку для обеспечения доверия к результатам работы алгоритма необходимо, чтобы точность прогнозирования не зависела от частоты встречаемости конкретной ситуации при эксплуатации алгоритма.

7.2.4 В случае если для каких-либо параметров, используемых в тестовых наборах данных, неизвестен закон распределения, необходимо осуществлять проверку статистических гипотез и подбирать закон посредством статистических методов (например, методом максимального правдоподобия, если критерии его применения позволяют) на опытных данных. В случае если под опытные данные подходят несколько законов распределения, организация, осуществляющая испытания алгоритма, обязана выбрать закон распределения на основе подходящих статистических критериев, оценивающих похожесть закона распределения.

7.2.5 При испытании СУИТИ необходимо руководствоваться следующими принципами проведения испытаний: объективность испытаний, обоснованность применяемых методов (методик) испытаний, обеспечение единства измерений (аттестация методик измерений), воспроизводимость результатов испытаний и др.

7.3 Фрагменты тестовых наборов данных (демонстрационные наборы данных)

7.3.1 Основной демонстрационный набор тестовых данных прогнозирования дорожных условий на примере метеорологических условий и климатических факторов приведен в [1].

7.3.2 Тестовый набор описывает отдельное подмножество существенных факторов эксплуатации — метеорологические условия и климатические факторы. Тестовый набор данных содержит следующие поля данных:

- дата измерения;
- наименование метеостанции;
- влажность;
- атмосферное давление;
- вид осадков;
- интенсивность осадков;
- порыв ветра;
- направление ветра;
- скорость ветра;
- точка росы;
- температура воздуха;
- коэффициент сцепления;
- температура дорожного покрытия;
- температура тела дороги на глубине 4—7 мм;
- облачность;
- концентрация реагентов;
- точка замерзания.

7.3.3 Приведенный классификатор является примером перечня существенных факторов эксплуатации для рассматриваемого алгоритма прогнозирования дорожных условий, при этом сценарии дорожно-транспортных ситуаций должны генерироваться на основе приведенного перечня существенных факторов эксплуатации при испытании алгоритма прогнозирования дорожных условий с учетом тех или иных законов распределения существенных факторов эксплуатации.

7.3.4 При формировании конкретного тестового набора данных для тестирования алгоритма прогнозирования дорожных условий по образцу приложенного к настоящему стандарту демонстрационного набора данных должна быть реализована защита от введения в набор недостоверных данных по значениям параметров дорожных метеорологических условий по третьему уровню.

7.4 Правила формирования представительных тестовых наборов данных (включая, в случае необходимости, описание представительной совокупности тестовых сценариев)

7.4.1 Демонстрационный тестовый набор данных содержит в себе файл «Демонстрационный набор данных.xlsx» — таблицу с временным рядом значений параметров дорожных условий с трех автоматических дорожных метеостанций для тестирования алгоритмов прогнозирования дорожных условий.

7.4.2 Сам тестовый набор данных должен включать в себя исходный набор данных с выборкой подтвержденных соответствующих ситуаций. Формат файла должен быть доступен для применения подходов расширения (аугментации) обучающей выборки, описанных в 7.5.

7.4.3 Организация, которая проводит испытания алгоритмов прогнозирования дорожных условий, может использовать иной масштаб времени по аналогии с представленным в демонстрационном наборе (ежедневные измерения) — например, прогнозирование может осуществляться по часам.

7.4.4 В случае необходимости, организация, проводящая испытания, может дополнять тестовые наборы данных дополнительными полями, предназначенными для расширения тестового набора данных новыми параметрами и существенными факторами эксплуатации. Например, прогнозирование климатических факторов может осуществляться при помощи метеорологического радара, кроме того, в состав существенных факторов эксплуатации могут включаться другие параметры дорожной обстановки.

7.4.5 Кроме того, в случае необходимости прогнозирования дорожных условий не только в точке установки измерительного периферийного оборудования СУИТИ, организация, проводящая испытания, может дополнять тестовые наборы данных дополнительными полями, предназначенными для пространственного прогнозирования. Например, допускается добавлять данные термокартирования.

7.5 Принципы расширения (аугментации) тестовых наборов данных

7.5.1 В целях увеличения репрезентативности обучающей выборки рекомендуется использовать методы из числа следующих:

- сдвиг по времени.

Примечание — Сдвиг по времени предполагает сдвиг данных временного ряда на определенное количество временных шагов для создания новых образцов для обучения. Например, можно сдвинуть данные о транспортном потоке вперед или назад на определенное количество временных шагов, чтобы создать новые образцы для обучения;

- масштабирование.

Примечание — Масштабирование предполагает масштабирование данных временного ряда на определенный коэффициент для создания новых образцов для обучения. Например, можно масштабировать данные транспортного потока на коэффициент 0,5 или 2, чтобы создать новые обучающие выборки;

- добавление шума.

Примечание — Добавление шума — это добавление случайного шума к данным временного ряда для создания новых образцов для обучения. Например, можно добавить гауссовский шум к данным транспортного потока, чтобы создать новые обучающие выборки;

- интерполяция.

Примечание — Подразумевают интерполяцию недостающих данных во временном ряду для создания новых образцов для обучения. Например, можно интерполировать недостающие данные о транспортном потоке для создания новых обучающих выборок;

- сглаживание.

Примечание — Подразумевают сглаживание данных временного ряда для создания новых образцов для обучения. Например, можно использовать фильтр скользящего среднего для сглаживания данных о транспортном потоке, чтобы создать новые обучающие выборки;

- повторная выборка.

Примечание — Означает повторную выборку данных временного ряда с другой частотой для создания новых образцов для обучения. Например, данные о транспортном потоке можно дискретизировать с большей или меньшей частотой, чтобы создать новые обучающие выборки;

- преобразование Фурье.

Примечание — Применяется к данным временного ряда для создания новых образцов для обучения. Например, можно использовать преобразование Фурье для извлечения частотных характеристик из данных транспортного потока, чтобы создать новые обучающие выборки;

- вейвлет-преобразование.

Примечание — Предполагает применение вейвлет-преобразования к данным временного ряда для создания новых образцов для обучения. Например, можно использовать вейвлет-преобразование для извлечения частотно-временных характеристик из данных транспортного потока, чтобы создать новые образцы для обучения.

7.5.2 Этот перечень методов может быть применен к временным рядам с целью генерализации тестового набора данных для более объективного тестирования степени переобученности и устойчивости алгоритмов прогнозирований к шуму, статистическим выбросам и другим дефектам временных рядов, а также настройкам периферийного оборудования, с которого были получены значения характеристик дорожных условий.

Библиография

[1] Демонстрационный тестовый набор данных для алгоритмов прогнозирования дорожных условий/ ООО «А+С Транспроект», ООО «А-Я эксперт». — М., 2023

УДК 62-52:006.354 ОКС 35.240.60

Ключевые слова: искусственный интеллект, автомобильный транспорт, системы управления, интеллектуальная транспортная инфраструктура, алгоритм прогнозирования, дорожные условия

Редактор *М.В. Митрофанова*
Технический редактор *В.Н. Прусакова*
Корректор *И.А. Королева*
Компьютерная верстка *М.В. Малеевой*

Сдано в набор 20.10.2023. Подписано в печать 03.11.2023. Формат 60×84⅛. Гарнитура Ариал.
Усл. печ. л. 1,86. Уч.-изд. л. 1,00.

Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

Создано в единичном исполнении в ФГБУ «Институт стандартизации»
для комплектования Федерального информационного фонда стандартов,
117418 Москва, Нахимовский пр-т, д. 31, к. 2.

gost.e-ecolog.ru gost.e-ecolog.ru

gost.e-ecolog.ru