ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р 71687— 2024

Искусственный интеллект

НАБОРЫ ДАННЫХ ДЛЯ РАЗРАБОТКИ И ВЕРИФИКАЦИИ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ КОСВЕННОГО ИЗМЕРЕНИЯ МЕХАНИЧЕСКИХ СВОЙСТВ ПОЛИМЕРНЫХ КОМПОЗИЦИОННЫХ МАТЕРИАЛОВ

Общие требования

Издание официальное

Москва
Российский институт стандартизации
2024

gost.e-ecolog.ru

Предисловие

1 РАЗРАБОТАН Федеральным государственным бюджетным учреждением «Российский институт стандартизации» (ФГБУ «Институт стандартизации»)

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 11 октября 2024 г. № 1437-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», 2024

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

Содержание

1 Область применения 1
2 Нормативные ссылки 1
3 Термины, определения и сокращения 2
3.1 Термины и определения 2
3.2 Сокращения 3
4 Общие положения 3
4.1 Классификация по методу получения 3
4.2 Классификация по применению данных 3
5 Порядок сбора, подготовки и хранения набора данных 4
5.1 Постановка цели и задач 4
5.2 Сбор данных 4
5.3 Подготовка данных: анализ 4
5.4 Подготовка данных: структурирование 5
5.5 Подготовка данных: постобработка 5
5.6 Хранение и доступ к набору данных 5
6 Общие требования к наборам данных для разработки и верификации моделей машинного обучения . . 6
6.1 Общие требования к описанию наборов данных 6
6.2 Общие требования к наборам данных на этапе разработки моделей машинного обучения 6
6.3 Общие требования к набору данных на этапе верификации моделей машинного обучения 7
Приложение А (рекомендуемое) Пример описания набора данных для косвенного измерения трещиностойкости по другим механическим свойствам полимерных композиционных материалов с использованием моделей машинного обучения 8
Библиография 11

Введение

Одной из основных целей в области проектирования новых композиционных материалов является прогнозирование их надежности и долговечности. Известные способы прогнозирования свойств материалов взаимосвязаны с состоянием внутренней структуры материала, что является ресурсоемкой задачей. Вместе с тем, в отличие от традиционных методов поискового проектирования, растущие требования к высокотехнологичному производству требуют изготовления узлов и элементов конструкций изделий наукоемкой техники с первого раза. Существуют стандартные методы испытаний пултрузионных полимерных композитов по ГОСТ Р 57921 на растяжение, сжатие, изгиб, сдвиг в плоскости армирования и ударную прочность при изгибе. Механические свойства, полученные в результате стандартных испытаний, могут быть использованы для косвенного измерения стойкости и долговечности изделия без дополнительных исследований внутренней структуры материала с использованием машинного обучения.

Установленные в настоящем стандарте общие требования определяют порядок выбора методов сбора, анализа, постобработки набора данных высокотехнологичного производства, а также методы контроля качества и полноты наборов данных для косвенного измерения свойств узлов и элементов конструкций.

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

Искусственный интеллект

НАБОРЫ ДАННЫХ ДЛЯ РАЗРАБОТКИ И ВЕРИФИКАЦИИ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ КОСВЕННОГО ИЗМЕРЕНИЯ МЕХАНИЧЕСКИХ СВОЙСТВ ПОЛИМЕРНЫХ КОМПОЗИЦИОННЫХ МАТЕРИАЛОВ

Общие требования

Artificial intelligence. Datasets for developing and verifying machine learning models for indirect measuring the mechanical properties of polymer composite materials. General requirements

Дата введения — 2025—01—01

1 Область применения

Настоящий стандарт устанавливает требования к сбору и подготовке наборов данных для разработки и верификации систем искусственного интеллекта, используемых для косвенного измерения механических свойств полимерных композиционных материалов с учетом их гетерогенности, анизотропности и зависимости свойств от производственного процесса. Стандарт обеспечивает повышенную доступность и качество данных, что позволяет облегчать разработку и верификацию систем искусственного интеллекта.

В данном стандарте приведена классификация получаемых данных, требования и рекомендации к сбору, подготовке и хранению данных.

2 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты:

ГОСТ 32656 Композиты полимерные. Методы испытаний. Испытания на растяжение

ГОСТ 33519 Композиты полимерные. Метод испытания на сжатие при нормальной, повышенной и пониженной температурах

ГОСТ 33843 (ISO 15310:1999) Композиты полимерные. Метод определения модуля сдвига в плоскости методом кручения

ГОСТ Р 56805 Композиты полимерные. Методы определения механических характеристик при изгибе

ГОСТ Р 56810 Композиты полимерные. Метод испытания на изгиб плоских образцов

ГОСТ Р 57067 Система внешнего армирования из полимерных композитов. Метод определения межслойной прочности на сдвиг

ГОСТ Р 57715 Композиты полимерные. Определение ударной вязкости по Изоду

ГОСТ Р 57734 Композиты полимерные. Определение энергии удара, затраченной на разрушение образца

ГОСТ Р 57866 Композиты полимерные. Методы определения механических характеристик при изгибе непрерывно-армированных композитов

ГОСТ Р 57921 Композиты полимерные. Методы испытаний. Общие требования

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 ежемесяч-

Издание официальное

ного 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.

3 Термины, определения и сокращения

3.1 Термины и определения

В настоящем стандарте применены следующие термины с соответствующими определениями:

3.1.1

измерять (measure) (глагол): Производить измерение.
[ГОСТ Р ИСО/МЭК 25000—2021, пункт 4.19]

3.1.2 **значение величины:** Число с указанием основы для сравнения, выражающее размер величины.

Примечание — См. [1], статья 1.19.

Пример 1 — Толщина материала: 0,65 мм.
Пример 2 — Сила упругости: 10 Н.
Пример 3 — Коэффициент Пуассона: 0,27.
Пример 4 — Напряжение сдвига: 5,88 МПа.

3.1.3

функция измерения (measurement function): Алгоритм или вычисление, выполняемое для комбинации не менее чем двух элементов показателя качества.
[ГОСТ Р ИСО/МЭК 25021—2014, пункт 4.7]

3.1.4

элемент данных (ЭД) (data element, DE): Единица данных, для которой с помощью набора атрибутов заданы определение, идентификация, представление и допустимые значения.
[ГОСТ Р ИСО/МЭК 11179-1—2010, пункт 3.3.8]

Примечания
1 Поле считается синонимом ЭД.
2 ЭД — это физический объект, «контейнер» значений данных согласно [1], пункт 4.9.

3.1.5

тип данных (datatype): Множество различающихся значений, охарактеризованных свойствами этих значений и операциями над ними.
[ГОСТ Р ИСО/МЭК 11179-1—2010, пункт 3.3.11]

Примечание — Примерами типов данных являются строки символов, тексты, даты, числа, изображения, звуки и т. д. (см. [1], пункт 4.16).

3.1.6 **значение данных:** Содержание элемента данных.

Примечания
1 С «внутренней» точки зрения качество данных относится к самим данным, таким как значения предметной области и возможные ограничения (см. [2], пункт 5.1.1).
2 Номер или категория, присвоенная атрибуту целевого объекта путем проведения измерения (см. [1], статья 4.17).

3.1.7 **запись данных:** Набор связанных элементов данных, рассматриваемых как единое целое.

Примечание — См. [2], пункт 4.15.

3.1.8

машинное обучение (machine learning): Процесс автоматического обучения и совершенствования поведения системы искусственного интеллекта на основе обработки массива обучающих данных без явного программирования.

[ГОСТ Р 59895—2021, пункт 2.1.7]

3.1.9

модель машинного обучения (machine learning model): Математическая конструкция, генерирующая логический вывод или прогноз на основе входных данных и/или информации.

[ГОСТ Р 71476—2024, пункт 3.3.7]

3.2 Сокращения

В настоящем стандарте применены следующие сокращения:

ИИ — искусственный интеллект;

МО — машинное обучение;

ПКМ — полимерные композиционные материалы.

4 Общие положения

Раздел определяет систему классификации наборов данных на основе метода их получения и целей применения, направленную на упорядочивание и систематизацию наборов данных.

4.1 Классификация по методу получения

Данные классифицируют в зависимости от способа, которым они могут быть получены:

- данные, которые можно получить прямыми измерениями с помощью средств измерений, характеризующие плотность, прочность на растяжение, модуль упругости, коэффициент теплового расширения и другие характеристики;

- данные, полученные косвенными измерениями, представляют значения физических величин, определяемые с помощью математических манипуляций данными, полученными прямыми измерениями, включая расчетные параметры, такие как коэффициенты анизотропии, распределение напряжений и деформаций, прогнозируемые свойства при изменении условий эксплуатации, а также результаты моделирования поведения материалов под комплексной нагрузкой;

- данные производственного процесса, включающие параметры и условия изготовления ПКМ, такие как температура и давление при формовании, скорость нагрева и охлаждения, время выдержки, свойства исходных компонентов (например, тип полимера и армирующего материала), а также параметры обработки (например, механическая обработка, химическая обработка и постобработка).

Пример — Для однонаправленных композиционных материалов прямыми данными являются плотность, прочность на растяжение, твердость компонентов и т. д., косвенными данными являются трещиностойкость, усталостная прочность и т. д., данными производственного процесса являются температура и давление формирования, время выдержки и т. д.

4.2 Классификация по применению данных

Данные подразделяются по применению в рамках разработки модели МО на следующие категории:

а) для обучения систем ИИ: применяются в моделях МО, которые могут быть включены в системы ИИ для косвенного измерения механических свойств ПКМ; наборы данных для обучения, в свою очередь, возможно разделить:

1) на входные данные для модели МО: набор данных, используемых для подачи информации в модель МО,
2) выходные данные для модели МО: представляют собой результаты косвенных измерений, которые модель МО генерирует после обработки входных данных;

б) для тестирования систем ИИ: применяются для оценки эффективности и точности обученных моделей на этапе их разработки;

в) для внешней верификации: применяются для внешней проверки и тестирования моделей МО с использованием независимых источников данных.

5 Порядок сбора, подготовки и хранения набора данных

Порядок сбора, подготовки и хранения набора данных:
- постановка цели и задач;
- сбор данных;
- подготовка данных: анализ;
- подготовка данных: структурирование;
- подготовка данных: постобработка;
- хранение и доступ к набору данных.

5.1 Постановка цели и задач

Цель сбора данных — получение набора данных, который обеспечивает создание эффективных и точных моделей МО, способных производить косвенное измерение механических свойств ПКМ для различных приложений. Необходимо точно сформулировать, какие механические свойства будут использоваться как входные, а какие данные — как выходные.

Задачи сбора данных:

- формирование строгого списка данных, необходимого для косвенных измерений определенных свойств ПКМ с учетом особенностей материала и условий эксплуатации.

Для косвенного измерения механических свойств ПКМ, армированных однонаправленными волокнами, необходимо учитывать анизотропность материала. Следовательно, наборы данных должны включать параметры, описывающие направленность волокон, такие как угол ориентации волокон, а механические свойства должны быть определены с учетом ориентации волокон в различных направлениях относительно направления волокон;

- определение критериев качества данных для обеспечения надежности и точности моделей МО;
- определение методов сбора данных для обучения и проверки моделей ИИ (см. 5.2);
- формирование необходимого объема данных для выполнения поставленной цели; недостаточное количество данных, их сильная зашумленность (наличие нерелевантных данных) или слишком высокая для имеющихся наборов данных сложность модели могут привести к неэффективному обучению моделей;
- установление контекста использования данных посредством условий и среды сбора данных, в которых МО будет применяться для предсказания механических свойств ПКМ.

Уточнение целей сбора данных и учет аспектов оптимизируют процесс обучения, обеспечивают точность и надежность систем ИИ и повышают их эффективность при косвенном измерении механических свойств ПКМ.

5.2 Сбор данных

Процесс сбора данных может быть осуществлен с помощью следующих источников и способов получения (ранжированы в порядке приоритетного получения):

- лабораторные данные (проспективный сбор) — результат лабораторных испытаний на образцах материалов по ГОСТ Р 57921;
- производственные данные (ретроспективный сбор) — результат испытаний свойств ПКМ, полученных во время производственных процессов и эксплуатации;
- сгенерированные данные, полученные на основе симуляции и моделирования механического поведения ПКМ для получения виртуальных данных, которые могут быть необходимы при недостаточности реальных данных;
- литературные данные, собранные научными группами или организациями при условии, что они соответствуют требованиям к точности и представительности.

Выбор способа зависит от цели разработки и верификации моделей МО, доступных ресурсов и технических возможностей. Важно обеспечить достоверность, точность и представительность источников данных и методов получения.

5.3 Подготовка данных: анализ

Необходимо использовать различные подходы и статистические методы для идентификации пропусков, выбросов, дубликатов и других аномалий.

Способы анализа данных:

- визуализация данных (гистограммы и диаграммы рассеяния для выявления аномалий и зависимостей);
- проверка на выбросы (идентификация и обработка выбросов);
- корреляционный анализ (оценка взаимосвязей между переменными);
- обработка пропусков (выбор метода заполнения пропущенных значений);
- кластерный анализ (группировка данных для выявления особенностей различных групп).

Выбор способа анализа данных зависит от характера данных, задач, доступных ресурсов и требований к точности. Этот этап подготовки набора данных для обучения и верификации моделей МО позволяет обнаружить и устранить аномалии и некорректные данные.

5.4 Подготовка данных: структурирование

Для структурирования собранных данных необходимо использовать следующие подходы:

- выбор признаков: отбор релевантных признаков с важной информацией помогает сократить размерность данных и улучшить производительность моделей;
- нормализация и стандартизация: приведение числовых признаков к одному масштабу обеспечивает эффективное обучение моделей и предотвращает доминирование одних признаков над другими;
- работа с текстовыми данными: применение методов разбиения текстового документа на более мелкие единицы, векторизации и снижения размерности для работы с текстовыми данными.

Выбор способа структурирования данных зависит от типа, объема и сложности данных, а также требований к точности и надежности моделей ИИ.

5.5 Подготовка данных: постобработка

На этапе определения критериев постобработки данных для обеспечения контроля качества можно применять разнообразные подходы для обнаружения, предотвращения и устранения проблем, связанных с качеством наборов данных:

- обработка и удаление значений, выходящих за пределы разумных или физически обоснованных значений, которые могут исказить результаты моделей;
- выбор и применение методов заполнения пропущенных значений или удаления соответствующих записей, чтобы избежать искажения общей структуры данных;
- оценка качества данных и отсечение ненадежных или некачественных записей или источников данных;
- исключение данных, не соответствующих контексту и условиям сбора, которые могут быть непрезентативными или неприменимыми для поставленных задач;
- обнаружение и исправление ошибок или несогласованности в данных, чтобы обеспечить их корректность и надежность;
- применение методов обработки шума или выбросов, которые могут возникать при сборе данных;
- установление процесса контроля качества данных и верификации моделей МО для обнаружения и устранения потенциальных проблем с данными на регулярной основе.

Выбор способа постобработки данных зависит от типа данных, характера исследования и требований к точности и надежности моделей ИИ.

5.6 Хранение и доступ к набору данных

Хранение и управление набором данных, а также обеспечение безопасности и защиты данных от несанкционированного доступа или потери осуществляется при помощи следующих действий:

- определения базы данных, которая обеспечит достаточную емкость и производительность для сохранения больших объемов данных;
- разработки системы регулярного резервного копирования данных, чтобы предотвратить потерю информации в случае сбоев или ошибок;
- внедрения мер безопасности для защиты данных от несанкционированного доступа и утечек информации, предварительной обработки ЭД;
- создания и управления метаданными, которые содержат информацию о происхождении данных, параметрах сбора и другую контекстную информацию, необходимую для правильной интерпретации и использования данных;

- обеспечения удобного доступа к данным для их использования в различных приложениях и моделях МО;

- планирования системы хранения данных с учетом возможного увеличения объемов данных в будущем;

- ведения документации о структуре данных и методах доступа, а также разработки метрик для оценки производительности и качества системы хранения данных;

- разработки системы аудита и мониторинга для отслеживания активности пользователей, обнаружения необычных событий и быстрой реакции на возможные проблемы.

6 Общие требования к наборам данных для разработки и верификации моделей машинного обучения

6.1 Общие требования к описанию наборов данных

Общие требования к наборам данных для разработки и верификации моделей МО должны обеспечивать надежность, точность и статистическую значимость данных, чтобы обучение и оценка моделей были корректными и достоверными (см. приложение А).

К общим требованиям к наборам данных относятся:

- полнота и представительность: набор данных должен обладать достаточным объемом и разнообразием примеров, чтобы правильно отразить характеристики исследуемой области;

- актуальность: данные должны быть актуальными и соответствовать текущим условиям и требованиям задач ИИ;

- аннотации и разметка: набор данных должен содержать соответствующие аннотации и разметку, которые описывают каждый пример данных и его связь с целевыми переменными или метками классов;

- доступность и открытость: рекомендуется обеспечивать доступность набора данных для использования и верификации другими исследователями и специалистами. Открытость данных способствует повторяемости результатов и обмену знаниями;

- документация и метаданные: набор данных должен быть подробно задокументирован, содержать информацию о происхождении данных, исходных источниках, процессах сбора, обработки и структуризации;

- отчетность: при предоставлении набора данных должны быть подготовлены дополнительные материалы, такие как описание задачи, к которой применяется набор данных, статистика данных, итоги анализа, другая информация, которая может помочь пользователям лучше понять и использовать данные.

6.2 Общие требования к наборам данных на этапе разработки моделей машинного обучения

Общие требования к наборам данных на этапе разработки МО:

- сбалансированность в терминах классов или категорий, во избежание предвзятости или искажения результатов модели в пользу определенных классов;

- содержание эталонных выходных данных или значений, которые являются точными или признанными стандартными значениями, для проведения оценки моделей во время разработки;

- масштабируемость и содержание объема данных, достаточных для формирования количества выборок, необходимых для эффективного обучения.

Для успешной разработки моделей МО набор данных должен быть доступным для использования.

Набор данных, полученный средствами измерений, должен быть совместимым с используемыми форматами и инструментами для обучения и верификации моделей.

Требования к данным в контексте разработки и верификации моделей МО включают следующие аспекты:

- надежность и недвусмысленность данных во избежание противоречий и неоднозначности при их интерпретации;

- корректность и актуальность данных для рассматриваемой задачи; устаревшие или неточные данные могут привести к неверным результатам модели;

- получение данных в соответствии с установленными стандартами для обеспечения сопоставимости результатов измерений и обработки данных.

При необходимости обеспечения высокого уровня доверия к результатам измерений прямые данные должны быть получены с использованием калиброванных и сертифицированных приборов.

Требования к данным, полученным косвенными измерениями в контексте разработки и верификации моделей МО включают следующие аспекты:

- косвенные данные основаны на надежной и точной модели, которая связывает их с прямыми измеряемыми величинами; необходимо обеспечить высокую точность этой модели, чтобы минимизировать ошибки при получении косвенных данных;
- модель, используемую для вычисления косвенных данных, проверяют и верифицируют на соответствие исходным физическим законам и предположениям;
- модель должна быть надежной и стабильной, чтобы обеспечить согласованность результатов и избежать значительных изменений в косвенных данных при незначительных изменениях в прямых данных;
- корректность и надежность прямых данных, используемых для определения косвенных величин, критически важна; несоответствие или ошибки в прямых данных могут привести к неточности и искажениям в косвенных данных;
- метод вычисления косвенных данных должен быть объективным и не должен зависеть от субъективных оценок или предположений.

При использовании косвенных данных необходимо учитывать возможную неопределенность, связанную с моделью и вычислениями, и предоставить оценку этой неопределенности.

Методика вычисления косвенных данных должна быть разумной и реализуемой с учетом доступных данных и вычислительных ресурсов.

6.3 Общие требования к набору данных на этапе верификации моделей машинного обучения

Выделяют следующие требования к набору данных на этапе верификации моделей МО:

- репрезентативность относительно реального применения моделей МО: набор данных должен содержать разнообразные сценарии и условия, чтобы верифицировать работу модели на различных ситуациях;
- объективность и независимость от обучающего набора данных помогают предотвратить переобучение модели и дает более объективную оценку производительности;
- тестирование модели может включать различные метрики, такие как точность, полнота, F1-мера и другие, которые оценивают производительность модели в разных аспектах;
- верификационный набор данных должен включать критические случаи и редкие сценарии, чтобы проверить, как модель справляется с экстремальными условиями;
- при сравнении производительности различных моделей или алгоритмов необходимо учитывать статистическую значимость различий;
- если возможно, процесс верификации должен быть автоматизирован, чтобы обеспечить повторяемость и эффективность тестирования моделей.

Приложение А
(рекомендуемое)

Пример описания набора данных для косвенного измерения трещиностойкости по другим механическим свойствам полимерных композиционных материалов с использованием моделей машинного обучения

Название набора данных: механические свойства пултрузионного профиля композиционного материала, армированного стекловолокном.

Содержание набора данных: набор данных получен как результат испытаний 600 образцов материала и состоит из 900 ЭД. Статистика результатов механических испытаний представлена в таблице А.1. Выходные данные помечены знаком (*), остальные данные являются входными.

Таблица А.1

Механическое свойство	Среднее значение	Среднеквадратичное отклонение	Методика измерения
Напряжение разрушения при изгибе под углом 0°, МПа	559,2	71,3	По ГОСТ Р 56810
Модуль упругости при изгибе под углом 0°, ГПа	30,3	3,2	По ГОСТ Р 56805
Напряжение разрушения при изгибе под углом 90°, МПа	132,5	25,9	По ГОСТ Р 57866
Модуль упругости при изгибе на 90°, ГПа	11,8	1,3	По ГОСТ Р 56810
Предел прочности при сдвиге 0°, МПа	45,6	7,9	По ГОСТ Р 57067
Модуль упругости при сдвиге 0°, ГПа	6,9	1,2	По ГОСТ 33843
Прочность при сдвиге на 90°, МПа	54,4	4,9	По ГОСТ Р 57067
90° модуль сдвига, ГПа	7,8	0,7	По ГОСТ 33843
0° прочность на сжатие, МПа	494,4	58,0	По ГОСТ 33519
Модуль упругости при сжатии 0°, ГПа	47,4	3,9	По ГОСТ 33519
Прочность на сжатие 90°, МПа	91,9	10,1	По ГОСТ 33519
Модуль упругости при сжатии 90°, ГПа	6,7	0,8	ПО ГОСТ 33519
Энергия разрушения при ударе 0°, кДж/м^2	70,0	27,0	По ГОСТ Р 57715
Энергия разрушения при ударе 90°, кДж/м^2	14,5	3,0	По ГОСТ Р 57734
0° прочность на растяжение, МПа	657,5	85,6	По ГОСТ 32656
0° модуль упругости при растяжении, ГПа	45,1	1,5	По ГОСТ 32656
Прочность при растяжении на 90°, МПа	45,6	4,5	По ГОСТ 32656
Модуль упругости при растяжении 90°, ГПа	5,4	0,5	По ГОСТ 32656
0° KIC для 40 мм*	232,4*	29,6*	По методике лаборатории*

Каждый файл результатов механического испытания предварительно анализируют и определяют пороговое значение для создания данных тепловой карты корреляций в полученном наборе данных. Последняя строка (или столбец) представляет корреляцию трещиностойкости с другими свойствами для лучшего представления качества измерения. Все файлы сканирования суммируются и сохраняются в файле структуры со следующими ветвями (например):

- изображение тепловой карты корреляций в полученном наборе данных;
- имя или номер значения косвенного измерения механических свойств в виде строки.

Цель сбора и использования данных:

- набор данных содержит значение косвенного измерения механических свойств ПКМ, которое можно использовать для МО;
- набор данных целесообразно использовать для тестирования результатов МО;

- набор данных и используемые методы могут быть дополнительно интегрированы со многими машинами для технологии изготовления новых изделий.

Формирование данных:

а) источник данных: указание источников данных, откуда были получены косвенные измерения и какими методами.

Данные были получены путем лабораторных испытаний ПКМ, специально изготовленного для механических испытаний (см. таблицу А.1);

б) технические характеристики: описание технических характеристик, используемых для проведения косвенных измерений, включая оборудование, параметры и методики.

Лабораторные испытания проводились на поверенных установках Instron, в сертифицированной лаборатории механических испытаний;

в) модель косвенного измерения: подробное описание математической или статистической модели, используемой для преобразования прямых данных в косвенные значения механических свойств.

Сформированные данные показали, что некоторые механические свойства (таблица. А.1) в большей степени коррелируют с трещиностойкостью материала: упругий изгиб и эластичность при растяжении, вследствие характера поведения матрицы пултрузионного профиля композиционного материала, армированного стекловолокном; модуль продольного сжатия и растяжения вследствие накопления энергии в местах сжатия пултрузионного профиля;

г) прямые данные: список и описание прямых измеряемых величин, которые были использованы для расчета косвенных данных.

К прямым данным относятся:

- напряжение разрушения при изгибе под углом 0°,
- модуль упругости при изгибе под углом 0°,
- напряжение разрушения при изгибе под углом 90°,
- модуль упругости при изгибе на 90°,
- предел прочности при сдвиге 0°,
- модуль упругости при сдвиге 0°,
- прочность при сдвиге на 90°,
- 90° — модуль сдвига,
- 0° — прочность на сжатие,
- модуль упругости при сжатии — 0°,
- прочность на сжатие — 90°,
- модуль упругости при сжатии — 90°,
- энергия разрушения при ударе — 0°,
- энергия разрушения при ударе — 90°,
- 0° — прочность на растяжение,
- 0° — модуль упругости при растяжении,
- прочность при растяжении на 90°,
- модуль упругости при растяжении — 90°;

д) разметка и аннотации: информация о разметке и аннотациях косвенных данных, включая соответствующие метки классов, если применимо.

Маркировка набора данных во многих случаях требует значительных инвестиций и усилий профессионалов, чтобы гарантировать, что данные маркированы правильно, воспроизводимо и объективно;

е) объем данных: общее количество примеров в наборе данных и объем каждой прямой и косвенной величины.

Для обучения моделей был получен значительный набор данных (900 свойств, 50 партий) механических свойств пултрузионного композиционного материала;

ж) связь с целевыми значениями: если возможно, указание связи косвенных данных с соответствующими целевыми значениями или эталонными измерениями.

Рассчитывается корреляция входных параметров к выходному параметру (трещиностойкости):

- ситуация с высокой корреляцией входных данных для прогнозирования механических свойств трещиностойкости,

- ситуация с нормальной корреляцией входных данных для прогнозирования механических свойств трещиностойкости,

- ситуация с недостаточной корреляцией входных данных для прогнозирования механических свойств трещиностойкости;

и) дата сбора данных: информация о датах сбора данных и возможных временных особенностях;

к) авторство и контактная информация: информация об авторах набора данных и контактные данные для обратной связи.

Приводится информация об организации и сотрудниках организации, которые производили сбор, обработку и структуризацию данных;

л) документация верификации: описание процесса верификации данных, включая результаты тестирования и оценки модели;

м) статистика данных: сводные статистические показатели о распределении и характеристиках прямых и косвенных данных.

Среднее значение и среднее квадратичное отклонение приведено в таблице А.1;

н) контекст использования: информация о том, для каких задач и областей применения предназначен набор данных.

Набор данных используется в научных исследованиях ПКМ, произведенных методом пултрузии.

Библиография

[1] Международный словарь по метрологии: основные и общие понятия и соответствующие термины: пер. с англ. и фр. / Всерос. науч.-исслед. ин-т метрологии им. Д.И. Менделеева, Белорус. гос. ин-т метрологии. Изд. 2-е, испр. — СПб.: НПО «Профессионал», 2010. — 82 с.

[2] ИСО/МЭК 25024:2015 Системная и программная инженерия. Требования и оценка качества систем и программного обеспечения (SQuaRE). Измерение качества данных [Systems and software engineering — Systems and software. Quality Requirements and Evaluation (SQuaRE) — Measurement of data quality]

УДК 004.01:006.354

ОКС 35.020
35.240.99

Ключевые слова: искусственный интеллект, средства измерений, наборы данных, разработка, верификация, модели машинного обучения, косвенное измерение, механические свойства, полимерные композиционные материалы, требования

Редактор *М.В. Митрофанова*
Технический редактор *И.Е. Черепкова*
Корректор *М.И. Першина*
Компьютерная верстка *Л.А. Круговой*

Сдано в набор 05.11.2024. Подписано в печать 20.11.2024. Формат 60×84⅛. Гарнитура Ариал.
Усл. печ. л. 1,86. Уч.-изд. л. 1,58.

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