|  |  |  |
| --- | --- | --- |
| **ФЕДЕРАЛЬНОЕ АГЕНТСТВО**  **ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ** | | |
|  |  |  |
| http://tm.msp.ua:81/TM_IMG/img_RU/610/610659.jpg | **НАЦИОНАЛЬНЫЙ**  **СТАНДАРТ**  **РОССИЙСКОЙ**  **ФЕДЕРАЦИИ** | **ГОСТ Р**  *(проект, окончательная редакция)* |
|  |  |  |

**СИСТЕМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

**Типовые требования к контрольным выборкам исходных данных для испытания систем искусственного интеллекта в образовании**

*Настоящий проект стандарта не подлежит применению до его утверждения*

**Предисловие**

1. РАЗРАБОТАН Федеральным государственным автономным образовательным учреждением высшего образования "Национальный исследовательский университет "Высшая школа экономики"
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 \_\_.\_\_.\_\_\_\_ № \_\_\_-ст
4. ВВЕДЕН ВПЕРВЫЕ

*Правила применения настоящего стандарта установлены в статье 26 Федерального закона "О стандартизации в Российской Федерации". Информация об изменениях к настоящему стандарту публикуется в годовом (по состоянию на 1 января текущего года) информационном указателе "Национальные стандарты", а официальный текст изменений и поправок - в ежемесячно издаваемом 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 органа исполнительной власти в сфере стандартизации в сети Интернет (www.gost.ru)*

|  |
| --- |
| **НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ** |
| **СИСТЕМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**  **Типовые требования к контрольным выборкам исходных данных для испытания систем искусственного интеллекта в образовании**  Artificial intelligence systems. Typical requirements for check samples of initial data for testing artificial intelligence systems in education |
| **Дата введения — \_\_.\_\_.\_\_\_\_** |

1 Область применения

Настоящий стандарт распространяется на технологии искусственного интеллекта в образовании и устанавливает типовые требования к контрольным выборкам исходных данных для испытания систем искусственного интеллекта в образовании, построенных на основе алгоритмов машинного обучения и анализа данных.

2 Терминология

В настоящем стандарте применены термины по ГОСТ Р 1.11.164-1.006.20[[1]](#footnote-1), ГОСТ Р ИСО/МЭК 11179-4-2012[[2]](#footnote-2), а также следующие термины с соответствующими определениями

2.1 **обучающая выборка:** Набор данных, на котором происходит обучение алгоритмов искусственного интеллекта технической образовательной системы.

2.2 **валидационная выборка:** Набор данных, на котором происходит оптимизация качества работы алгоритмов искусственного интеллекта технической образовательной системы.

2.3 **контрольная выборка:** Набор данных, на котором происходит окончательное определение результатов работы алгоритма искусственного интеллекта технической образовательной системы.

2.4 **обучающая** **метрика:** Значение, количественно определяющее качество работы алгоритмов искусственного интеллекта технической образовательной системы, подсчитанное на обучающей выборке

2.5 **валидационная** **метрика:** Значение, количественно определяющее качество работы алгоритмов искусственного интеллекта технической образовательной системы, подсчитанное на валидационной выборке

2.6 **контрольная** **метрика:** Значение, количественно определяющее качество работы алгоритмов искусственного интеллекта технической образовательной системы, подсчитанное на контрольной выборке

 3 Рабочий процесс оценки моделей искусственного интеллекта в области образования

Обобщенный рабочий процесс оценивания моделей искусственного интеллекта представлен на рисунке 1.

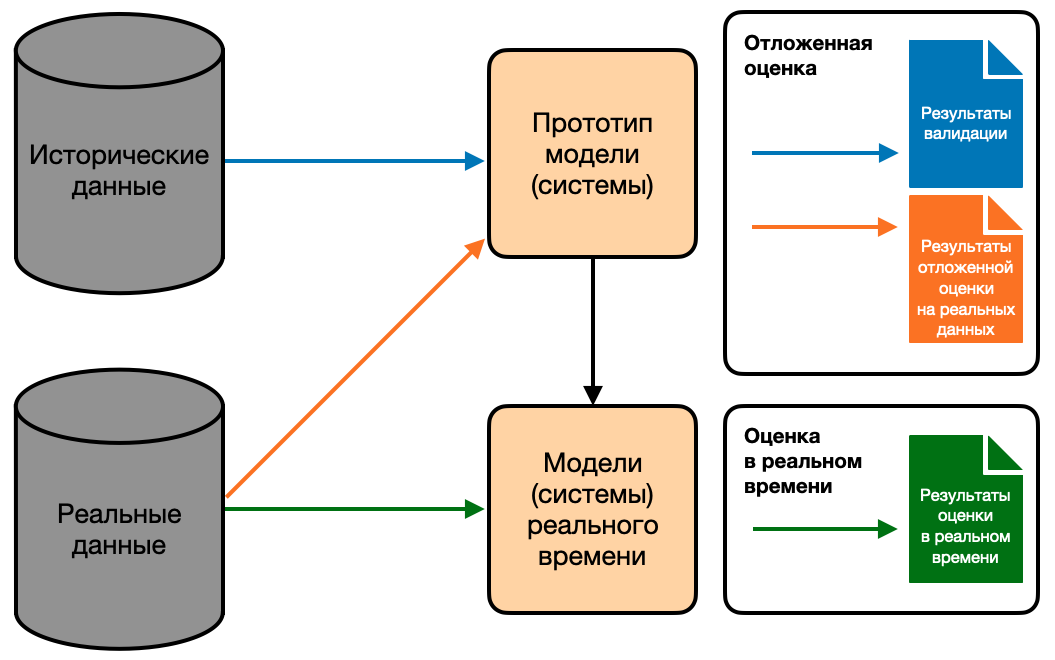


Рисунок 1 – Обобщенный рабочий процесс оценивания моделей искусственного интеллекта

**4 Общие требования к контрольным выборкам исходных данных**

* 1. **Репрезентативность**

Контрольная выборка должна быть репрезентативной: характеристики выборки должны соответствовать характеристикам генеральной совокупности в целом. Параметры генеральной совокупности определяются в соответствии с границами исследуемых данных, и должны быть записаны в метаданных контрольной выборки. Должны быть предусмотрены метрики и меры оценки отсутствия систематических искажений при формировании выборки.

* 1. **Статистическая значимость**

Для каждой системы с использованием ИИ должна быть сформулирована гипотеза и проведена оценка статистической значимости. Должен быть выбран и обоснован уровень значимости.

* 1. **Обучающая, валидационная и контрольная метрики**

При обучении систем искусственного интеллекта на данных валидационные выборки должны использоваться для оценки качества обучения систем и настройки их гиперпараметров, тогда как контрольные выборки должны использоваться для оценки эффективности работы систем и сравнения различных систем между собой. Не допустимо использование контрольных выборок в качестве валидационных в процессе настройки и обучения модели.

* 1. **Период обновления**

Для контрольной выборки должен быть определен период обновления, соответствующий периодам обновлений параметров процессов и объектов, соответствующих собранным данных.

В частности, для данных, описывающих образовательные программы, должен быть назначен период обновления, равный принятому сроку обновления образовательных программ; для данных, описывающих обучающихся, должен быть выбран период обновления, равный средним срокам ротации участников образовательного процесса.

* 1. **Повторное использование контрольной выборки**

Для каждой контрольной выборки должно быть определено максимальное число использований и вестись фактических использований в оценке, с целью недопущения подстройки модели искусственного интеллекта под прохождение конкретной контрольной выборки.

1. **Объекты данных в образовании**

5.1 Данные для контрольных выборок в образовании должны быть собраны в отношении следующих объектов:

- описание и программы образовательных дисциплин, учебно-методические материалы;

- учебные действия и коммуникация участников образовательной деятельности (обучающийся, педагог);

- операции взаимодействия с информационными системами участников образовательной деятельности;

- состояние участников образовательной деятельности;

- продуктивные результаты образовательной деятельности (созданные артефакты);

- оценки действий и артефактов;

- состояние и характеристики окружающей среды.

5.2 Примеры видов наборов данных и требования к контрольным выборкам для каждой категории данных, указанных в 5.1, представлены в Приложении А.

**6 Требования к описанию контрольных выборок данных в образовании**

6.1 Контрольные выборки данных должны иметь определённый набор метаданных. Данный набор можно определить в трёх категориях:

- описательные метаданные, описывают содержание и состояние данных;

- технические метаданные, описывают технические подробности хранения данных;

- операционные метаданные, описывают процессы обработки данных и доступа к ним.

6.2 Технические и операционные метаданные контрольных выборок должны соответствовать соответствующим метаданным выборок данных, на которых производится обучение модели искусственного интеллекта и метаданным данных, которые обрабатываются в процессе эксплуатации модели.

6.4 Описательные метаданные контрольных выборок данных в образовании должны содержать следующие характеристики:

- контекст сбора данных (название образовательной программы, учебного курса, мероприятия);

- характер данных (описание и план деятельности / фактически осуществленная деятельность / оценка деятельности);

- источник данных (обучающийся, педагог);

- вид данных (данные о характеристиках обучающегося и/или педагога / данные о деятельности);

- уровень данных о деятельности (деятельность / действие / операция);

- вид образования (общее образование, профессиональное образование, дополнительное образование и профессиональное обучение);

- уровень образования (дошкольное образование; начальное общее образование; основное общее образование; среднее общее образование. среднее профессиональное образование; высшее образование - бакалавриат; высшее образование - специалитет, магистратура; высшее образование - подготовка кадров высшей квалификации);

- возраст обучающихся;

- уровень подготовки обучающихся;

- особенности обучающихся.

**Приложение А**

**(обязательное)**

**Примеры наборов данных в образовании и требования к контрольным выборкам**

А.1 Примеры наборов данных в образовании и требования к контрольным выборкам приведены в таблице А.1

Т а б л и ц а А.1 – Примеры наборов данных в образовании и требования к контрольным выборкам

| № | Тип данных | Требования к контрольной выборке |
| --- | --- | --- |
| 1 | Описание курса и учебно-методические материалы |  |
| 1.1 | Описание предмета / курса / тематического модуля в каталоге (содержит требования к участникам, количество контактных часов, зачетных единиц, тип и т.п.) | Выборка должна обеспечивать использование данных для предположения об образовательных результатах, если они не заданы явно |
| 1.2 | Структура активности | Выборка должна обеспечивать использование данных для привязки образовательных результатов к действиям пользователя |
| 1.3 | Образовательные результаты (разметка образовательных результатов) | Выборка должна содержать текстовые описания образовательных результатов. Выборка должна обеспечивать использование данных для соотнесения деятельности обучающегося с результатами обучения. |
| 1.4 | Содержание образования | Выборка должна обеспечивать использование данных для предположения об образовательных результатах, если они не заданы явно |
| 1.5 | Правила оценивания | Выборка должна обеспечивать использование данных для контроля соответствия оценок образовательным результатам |
| 1.6 | Методика оценки / диагностики | Выборка должна обеспечивать использование данных для контроля выводов о достижении образовательных результатов или результатах диагностики. |
| 2 | Данные о продуктивной деятельности обучающегося |  |
| 2.1 | Запись о деятельности и результате деятельности (в привязке к образовательному результату) | Основная единица учета деятельности обучающегося. Выборка должна обеспечивать использование данных для демонстрации навыка или компетенции обучающегося. |
| 2.2 | Созданные обучающимся артефакты (в процессе учебной/проектной деятельности) в привязке к деятельности и образовательному результату (презентации, конспекты, схемы, чертежи, исходные тексты программ) | Выборка должна обеспечивать использование данных для подтверждения демонстрации результатов обучения, навыка или компетенции и её уровня |
| 2.3 | Личный дневник обучающегося, дневник его проектной команды | Выборка должна обеспечивать использование данных для отслеживания прогресса и качества деятельности |
| 2.4 | Записи о прохождении командой значимых вех в проектной деятельности | Выборка должна обеспечивать использование данных для отслеживания прогресса и качества проектной деятельности, проявления командой проекта определенных компетенций, достижения результатов обучения |
| 3 | След оценки деятельности |  |
| 3.1 | Набранные баллы (оценки за задание/урок/тематический раздел/курс) | Выборка должна обеспечивать использование данных для подтверждения демонстрации результатов обучения, навыка или компетенции и её уровня |
| 3.2 | Оценки и отзывы обучающихся | Выборка должна обеспечивать использование данных для подтверждения демонстрации результатов обучения, навыка или компетенции и её уровня |
| 3.3 | Обратная связь (рефлексия) | Выборка должна обеспечивать использование данных для предположения об образовательных результатах, а также для оценки курсов / мероприятий / элементов курсов |
| 3.4 | Экспертная оценка: свидетельства наблюдателей о проявлении навыков / компетенций | Выборка должна обеспечивать использование данных для подтверждения демонстрации результатов обучения, навыка или компетенции и её уровня |
| 3.5 | Полученные сертификаты (в виде изображений/документов) | Выборка должна обеспечивать использование данных для подтверждения демонстрации результатов обучения, навыка или компетенции и её уровня |
| 3.6 | Оценка артефактов, проектов | Выборка должна обеспечивать использование данных для оценки достижения образовательного результата |
| 3.7 | Достижение (получение статуса, бейджа и т.п.) | Выборка должна обеспечивать использование данных для оценки достижения образовательного результата |
| 4 | След взаимодействия |  |
| 4.1 | Отдельные действия пользователя в системе, имеющие смысл с точки зрения обучения (просмотр видео, ответ на конкретный вопрос, запрос подсказки и т.п.) | Выборка должна обеспечивать использование данных для глубокого анализа поведения обучающегося, выявления слабых мест, а также — для оценки учебного курса / мероприятия |
| 4.2 | "Кликстрим": нажатия кнопок, клики, переходы между страницами, скроллинг | Выборка должна обеспечивать использование данных для глубокого анализа поведения обучающегося, выявления слабых мест, а также — для оценки учебного курса / мероприятия |
| 4.3 | Факты и содержание переписки в ИТ-системах (в т.ч. — активность пользователя в открытых системах, социальных сетях и т.п.) | Выборка должна обеспечивать использование данных для глубокого анализа поведения обучающегося, подтверждения некоторых навыков и компетенций, соавторства при совместной работе |
| 4.4 | Факты общения участников образовательной деятельности в очном взаимодействии | Выборка должна обеспечивать использование данных для глубокого анализа поведения обучающегося, подтверждения некоторых навыков и компетенций, соавторства при совместной работе |
| 4.5 | Отдельные события в системах проектной деятельности (создание задачи, изменение статуса задачи и т.п.) | Выборка должна обеспечивать использование данных для глубокого анализа поведения обучающегося или проектной команды |
| 5 | След состояния |  |
| 5.1 | Изменение настроек и профиля в ИТ-системах | Выборка должна обеспечивать использование данных для глубокого анализа изменения поведения обучающегося, для анализа целей |
| 6 | След участия / мероприятия |  |
| 6.1 | Фиксация присутствия в определенном месте | Выборка должна обеспечивать использование данных для анализа поведения обучающегося, подтверждения компетенций; также — для оценки учебного курса / мероприятия |
| 6.2 | Фиксация факта использования содержания образования без подтверждающей продуктивной деятельности | Выборка должна обеспечивать использование данных для анализа поведения обучающегося, подтверждения компетенций; также — для оценки учебного курса / мероприятия |
| 6.3 | Состав и описание команд обучающихся при групповой/командной деятельности | Выборка должна обеспечивать использование данных для подтверждения участия в мероприятии |
| 6.6 | Проявленные роли обучающегося в командном взаимодействии | Выборка должна обеспечивать использование данных для подтверждения способности выполнения роли |
| 6.7 | Описание проекта | Выборка должна обеспечивать использование данных для выявления образовательных результатов, основанных на проектной деятельности |
| 7 | Целеполагание и продвижение по траектории |  |
| 7.1 | Цели пользователя (декларируемые им самим) | Выборка должна обеспечивать использование данных для построения персональной траектории обучения, работы педагогов |
| 7.2 | Рефлексия о достижении целей / продвижении | Выборка должна обеспечивать использование данных для построения персональной траектории обучения, работы педагогов |

|  |  |
| --- | --- |
| УДК 004.896:006.354 | ОКС 35.240.90 |
|  | |
| Ключевые слова: технологии искусственного интеллекта, образование, требования, контрольные выборки, исходные данные, испытание систем искусственного интеллекта в образовании | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Руководитель  разработки | Заведующий лабораторией |  |  |  | И.А. Карлов |
|  | должность |  | личная подпись |  | инициалы, фамилия |

1. ГОСТ Р 1.11.164-1.006.20 «Технологии искусственного интеллекта в образовании. Общие положения и терминология» [↑](#footnote-ref-1)
2. ГОСТ Р ИСО/МЭК 11179-4-2012 «Информационная технология. Регистры метаданных (РМД). часть 4. Формулировка определений данных» [↑](#footnote-ref-2)