



SOFROS AI/ML. Руководство по установке

Москва
Софрос
2026



ПРАВО ТИРАЖИРОВАНИЯ
КОНФИГУРАЦИИ И ДОКУМЕНТАЦИИ
ПРИНАДЛЕЖИТ КОМПАНИИ SOFROS.

ООО «Софрос», 2026
Москва, 129226, Докукина ул., д.16, стр.3

Телефон: (495) 825-16-15
E-mail: info@sofros.ru
URL: <http://www.sofros.ru>

Номер 1.0.0
редакции:

ОГЛАВЛЕНИЕ

1	Введение	4
1.1	Сокращения и обозначения	4
2	Порядок установки компонентов SOFROS AI/ML	5
2.1	Минимальные технические требования для контейнеров	5
2.2	Схема интеграционной архитектуры	5
2.3	Подготовка к установке	5
2.4	Установка SOFROS AI/ML	6
3	Описание настроек SOFROS AI/ML	8
3.1	Настройки сервиса	8
3.2	Настройки расширения	9

1 ВВЕДЕНИЕ

Настоящий документ содержит описания процессов установки продукта SOFROS AI/ML:

- в **Разделе 2** описан процесс установки прикладных компонентов продукта SOFROS AI/ML, включающий подготовку к установке и пошаговую установку продукта SOFROS AI/ML;
- в **Разделе 3** описаны настройки конфигурационных файлов, расширения и моделей.

1.1 СОКРАЩЕНИЯ И ОБОЗНАЧЕНИЯ

В настоящем документе используются следующие термины и определения:

API	Application Programming Interface — это программный интерфейс приложений, набор правил и инструкций, с помощью которых разные программы взаимодействуют друг с другом и обмениваются данными.
Swagger	Набор инструментов с открытым исходным кодом для разработки, документирования и тестирования API
Backend	Серверная часть веб-приложения или сайта, которая отвечает за обработку данных, управление бизнес-логикой и взаимодействие с базами данных
Docker	Платформа контейнеризации с открытым исходным кодом для разработки, тестирования, доставки и запуска приложений в средах с поддержкой контейнеризации

2 ПОРЯДОК УСТАНОВКИ КОМПОНЕНТОВ SOFROS AI/ML

2.1 МИНИМАЛЬНЫЕ ТЕХНИЧЕСКИЕ ТРЕБОВАНИЯ ДЛЯ КОНТЕЙНЕРОВ

Программные требования:

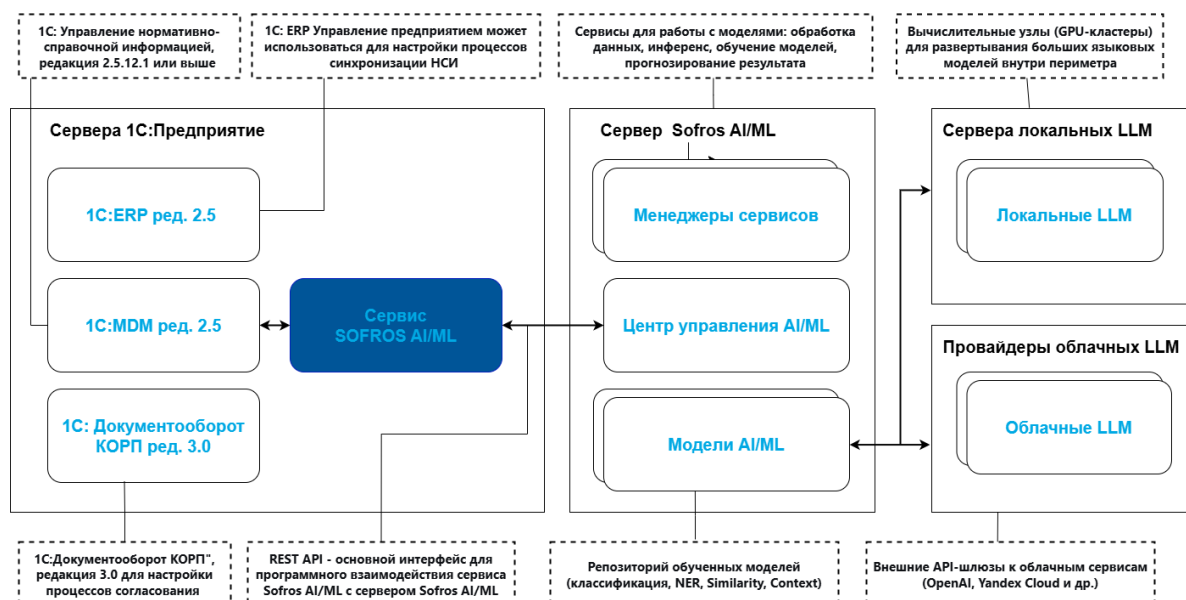
- Docker server версия CE, работающий с Linux-контейнерами

Аппаратные требования:

- Диск: 1000 гб
- Процессор: 12 ядер 3.6 ГГц
- Оперативная память: 32 гб
- Операционная система: Windows или Linux (Windows Server не поддерживается)

2.2 СХЕМА ИНТЕГРАЦИОННОЙ АРХИТЕКТУРЫ

Ниже представлена схема, описывающая состав продукта и предназначение компонентов, системные требования к решению, а также актуальный список доступов к внешним системам для успешной работы сервиса.



Для правильного функционирования продукта SOFROS AI/ML требуется разрешение использования следующих сетевых портов:

- SOFROS AI/ML (http REST, порт 8001 или любой другой доступный)

2.3 ПОДГОТОВКА К УСТАНОВКЕ

Перед установкой сервиса SOFROS AI/ML требуется:

- Убедиться в соблюдении минимальных технических требований SOFROS AI/ML.
- Получить файл лицензии **license.json** от вендора SOFROS AI/ML.
- Настроить расширение для 1С как в общем, так и по моделям.

Ознакомьтесь с содержимым установочного пакета, в котором содержатся следующие файлы и папки:

Имя	Тип	Назначение
data	Папка	Папка с данными для работы моделей, а также с результатами обучений.
images	Папка	Содержит Linux docker образы.
logs	Папка	Папка с логами работы сервиса.
docker-compose.yml	Файл	Файл оркестрации приложения.
install.bat	Файл	Скрипт установки образов из папки images в docker. Для ОС Windows.
install.sh	Файл	Скрипт установки образов из папки images в docker. Для ОС Linux.
models_config.json	Файл	Файл конфигурации моделей
run.bat	Файл	Скрипт запуска сервиса. Для ОС Windows.
run.sh	Файл	Скрипт запуска сервиса. Для ОС Linux.
stop.bat	Файл	Скрипт остановки сервиса. Для ОС Windows.
stop.sh	Файл	Скрипт остановки сервиса. Для ОС Linux.
Соф_SofrosAIML.cfe	Файл	Расширение конфигурации для 1C:MDM Управление нормативно-справочной информацией, редакция 2.5

2.4 УСТАНОВКА SOFROS AI/ML

1. Загрузить установочный пакет сервиса на сервер Docker. Загрузка производится любым доступным способом.
2. Выполнить разархивацию установочного пакета. Разархивация производится любым доступным способом и программным обеспечением. Допускается предварительная разархивация вне сервера Docker.
3. В папку сервиса добавить предоставленный файл лицензии **license.json** и при необходимости настроить файл конфигурации моделей **models_config.json** согласно описанию настроек.
4. При необходимости поменять порт сервиса с 8001 в файле **docker-compose.yml**.
5. Запустить скрипт **install.sh** или **install.bat**. Скрипт произведет локальную загрузку образов в Docker.
6. Запустить скрипт **run.sh** или **run.bat**. Скрипт выполнит команду **compose** и сформирует контейнеры согласно описанию файла **docker-compose.yml**.
7. По умолчанию, если не был изменен порт в файле **docker-compose.yml**, сервис будет доступен в браузере по ссылке http://{ip_сервера_docker}:8001/docs#/.
8. Для необходимости выполнения остановки SOFROS AI/ML выполните запуск скрипта **stop.sh** или **stop.bat**.

9. Добавить расширение в 1С. Расширение добавляется стандартно: «**Администрирование**» → «**Печатные формы, отчеты и обработки**» → «**Расширения**» → «**Добавить из файла**». После добавления расширения необходимо перезапустить клиентский сеанс.
10. Настроить расширение для 1С как в общем, так и по моделям.

3 ОПИСАНИЕ НАСТРОЕК SOFROS AI/ML

3.1 НАСТРОЙКИ СЕРВИСА

Настройки представлены в виде конфигурационных файлов, входящих в пакет приложения. Необходимо произвести настройки конфигурационных файлов для моделей и самого сервиса.

Файл **models_config.json** позволяет настроить модели. Сервис будет загружать только те модели, которые указаны в этом конфигурационном файле. Модели имеют следующие настройки:

- **id** - id модели.
- **name** - название модели.
- **description** - описание модели.
- **data_path** - путь к каталогу с данными модели (включая сохраненные классы, которые будут расположены в подпапке **models** указанного каталога) и ресурсами, необходимыми для работы этой модели.
- **auto_load** - булево значение. Показывает, надо ли модели загружать при старте все свои данные.

Ниже представлен пример настроек в файле **models_config.json**:

```
{
  "models": [
    {
      "id": "11111111-0001-3333-4444-555555555555",
      "name": "Модель 1",
      "description": "Модель без автоматической загрузки своих
данных",
      "data_path": "./data/11111111-0001-3333-4444-555555555555",
      "auto_load": false
    },
    {
      "id": "11111111-0002-3333-4444-555555555555",
      "name": "Модель 2",
      "description": "Модель с автоматической загрузкой своих
данных",
      "data_path": "./data/11111111-0002-3333-4444-555555555555",
      "auto_load": true
    }
  ]
}
```

Укажите необходимые для вас настройки, в файле представлено описание параметров со значениями по умолчанию.

Файл **docker-compose.yml** позволяет настроить порт, по которому будет доступен сервис. Ниже представлен пример настроек в файле **docker-compose.yml**:

```
services:
  ai_ml_server:
    container_name: sofros_ai_ml
    image: sofros-ai-ml
    restart: on-failure
```

```
ports:
  - "8001:8000"
volumes:
  - ./data:/app/data
  - ./logs:/app/logs
  - ./models_config.json:/app/models_config.json
  - ./license.json:/app/license.json
```

Чтобы изменить порт, надо поменять левое число, указанное в блоке **ports** сервиса **ai_ml_server**.

3.2 НАСТРОЙКИ РАСШИРЕНИЯ

Настройка расширения производится следующим образом:

1. Установить общие настройки расширения. Настройки находятся на вкладке «**Sofros AI/ML**» - «**Настройки подключения к серверу Sofros AI/ML**».

☆ Настройка подключения к серверу Sofros AI/ML 🔗 ⋮ □ ✕

Записать и закрыть Записать Еще ▾

Общие настройки

☒ Логировать запросы к сервису Sofros AI ML

Размер порции обучающих примеров:

Модель поиска вариантов значений характеристик: 🔗

Модель поиска наименований характеристик: 🔗

Модель классификации

Порог уверенности модели классификации: 🔗

Порог неклассификации модели классификации: 🔗

Порог вероятности модели классификации: 🔗

Не выводить результаты классификации с вероятностью ниже (%): 🔗

Модель экстракции

Тип подготовки при расширении класса: ▾

Модель семантического поиска

Максимальное количество результатов с переранжированием:

☒ Проверять создание семантически-близкой характеристики

☒ Запрещать создание семантически-близкой характеристики

Уверенность для семантического поиска характеристик: 🔗

Интеллектуальные ассистенты

Интеллектуальный ассистент поиска значений характеристик: 🔗

Алгоритм формирования промпта поиска значений характеристик: 🔗

Шаблон промпта поиска значений характеристик: 🔗

Общие настройки:

- «**Логировать запросы к сервису Sofros AI/ML**» - отвечает за включение/отключение логирования запросов к сервису и ответов от него в журнал регистрации.
- «**Размер порции обучающих примеров**» - определяет максимальный размер порции при отправке и удалении обучающих примеров.
- «**Модель поиска вариантов значений характеристик**» - определяет модель машинного обучения, которая будет использоваться при семантическом поиске значений характеристик в обработках «**Поиск значений характеристик**» и «**Разметка характеристик**».

- **«Модель поиска наименований характеристик»** - определяет модель машинного обучения, которая будет использоваться при семантическом поиске характеристик в обработках **«Поиск значений характеристик»** и **«Разметка характеристик»**.

Настройки модели классификации:

- **«Порог уверенности модели классификации»** - результаты классификации со значением уверенности меньше указанного в настройке выводиться не будут.
- **«Порог неклассификации модели классификации»** - результаты классификации со значениями уверенности и вероятности меньше указанного в настройке будут считаться неклассифицированными.
- **«Порог вероятности модели классификации»** - результаты классификации со значением вероятности меньше указанного в настройке выводиться не будут.
- **«Не выводить результаты классификации с вероятностью ниже (%)»** - результаты классификации со значением вероятности ниже указанного в настройке выводиться как предполагаемые значения не будут.

Настройки модели экстракции:

- **«Тип подготовки при расширении класса»** - устанавливает тип предподготовки класса (Simple – простой, Full – полная предподготовка).

Модель семантического поиска:

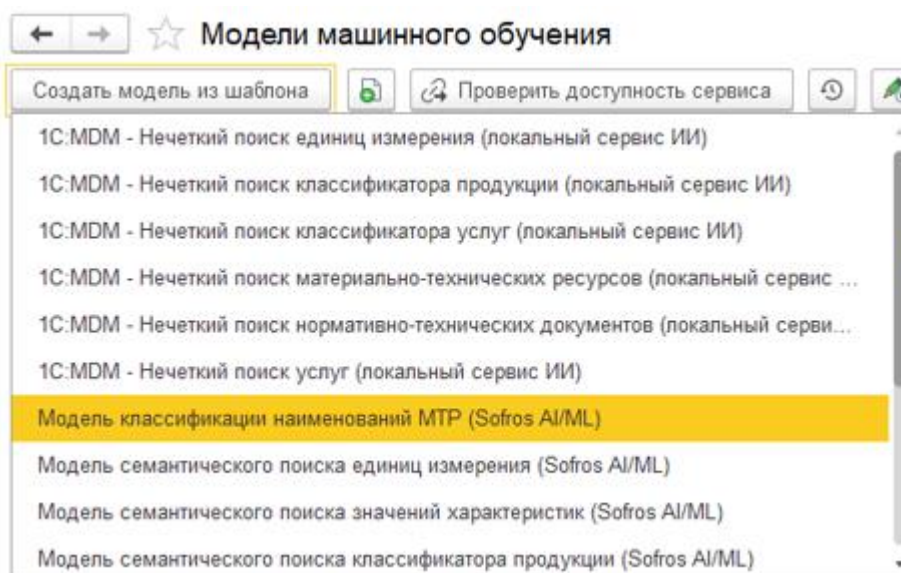
- **«Максимальное количество результатов с переранжированием»** - ограничивает количество результатов для моделей семантического поиска с включенной настройкой **«Использовать реранк»**.
- **«Проверять создание семантически-близкой характеристики»** - при создании новой характеристики происходит семантический поиск созданных ранее характеристик на наличие семантически-близкой.
- **«Запрещать создание семантически-близкой характеристики»** - при успешном нахождении семантически-близкой характеристики запрещает создание новой.
- **«Уверенность для семантического поиска характеристик»** - значения с уверенностью, больше указанной в настройке, считаются семантически-близкими.

Интеллектуальные ассистенты:

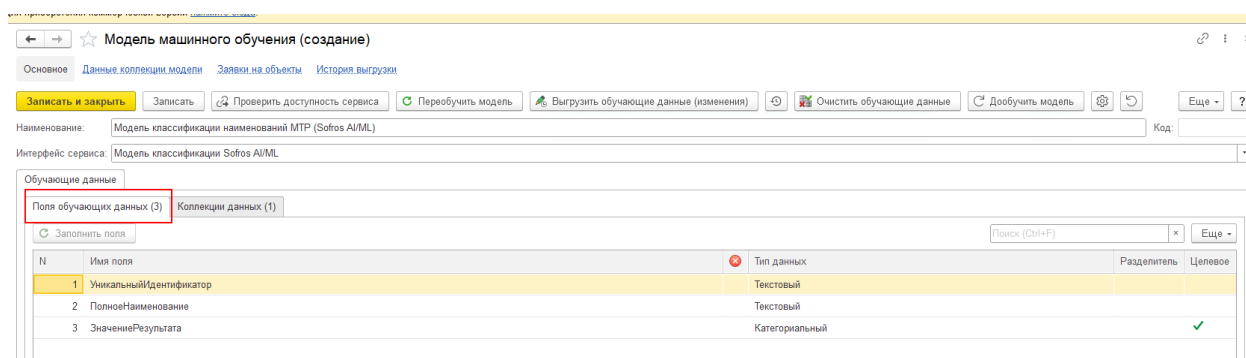
- **«Интеллектуальный ассистент поиска значений характеристик»** - интеллектуальный ассистент, используемый для дообогачивания элементов МТР значениями характеристик и классов новыми характеристиками.
- **«Алгоритм формирования промпта поиска значений характеристик»** - алгоритм, используемый для заполнения параметров шаблона промпта.
- **«Шаблон промпта поиска значений характеристик»** - шаблон промпта, используемый для обогащения списка характеристик МТР новыми значениями.

2. Выполнить настройку моделей машинного обучения, интеллектуальных ассистентов и сервисов LLM.

Модель классификации. Создание происходит по команде **«Создать модель из шаблона»** - **«Модель классификации наименований МТР (Sofros AI/ML)»**.

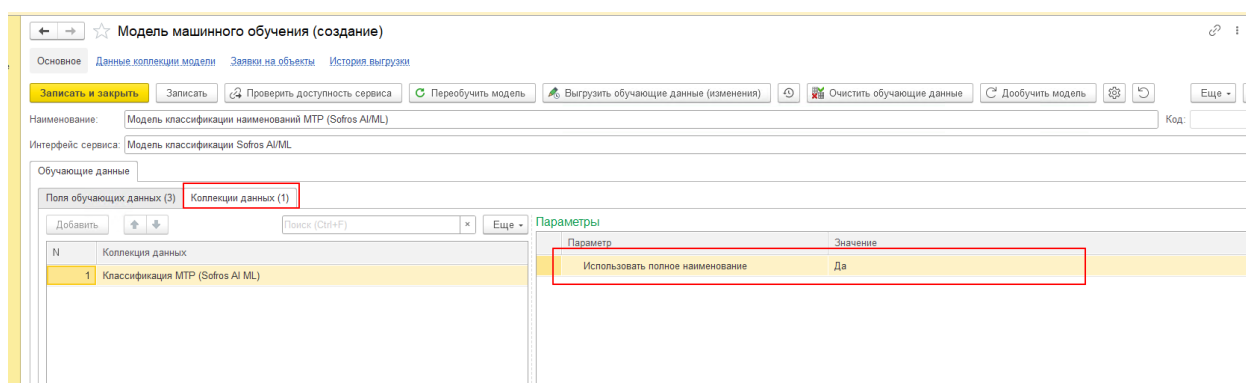


На вкладке «Поля обучающих данных» определен набор реквизитов, получаемых из запроса:

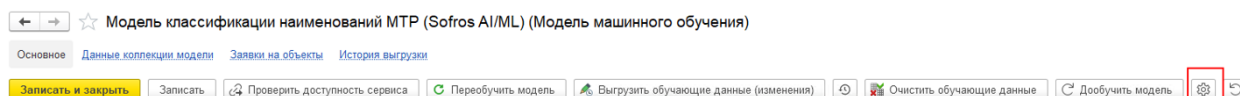


На вкладке «Коллекции данных» определены следующие параметры:

- «Коллекция данных» – запрос, по которому будет производится выборка коллекции обучающих данных.
- «Параметры» – Набор параметров, определенный для выборки коллекции обучающих данных. Для модели классификации наименований MTP определен один параметр «Использовать полное наименование», значение по умолчанию – «да». Если установить значение параметра «нет», то в качестве классифицируемой строки будет отправляться реквизит «Наименование».



Необходимо записать элемент, после чего перейти по команде «**Дополнительные настройки модели**».



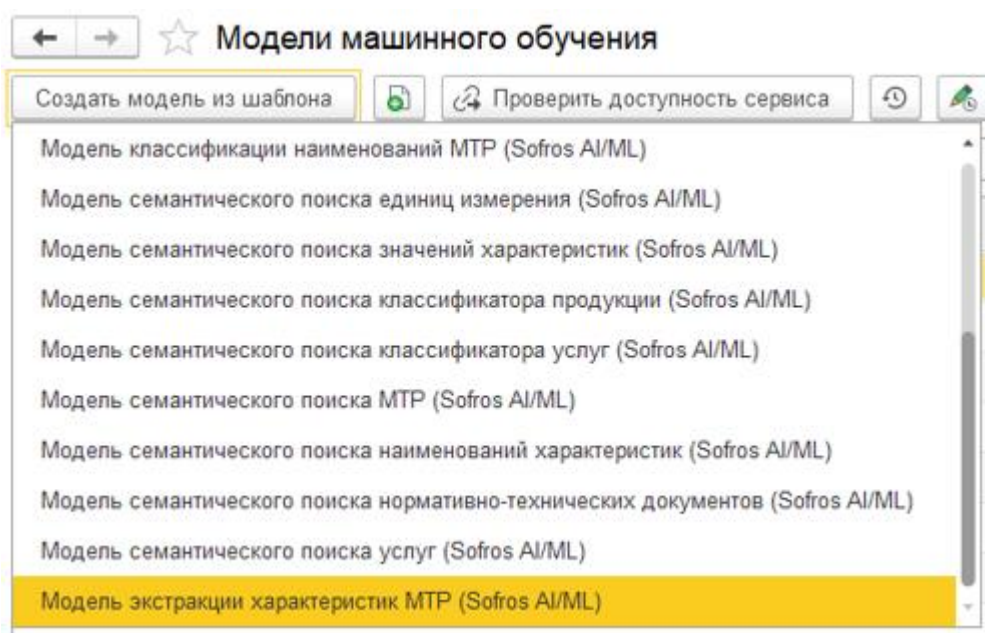
У модели классификации определен следующий набор настроек:

1. «**Сервер**» - адрес сервера, на котором расположен сервис Sofros AI/ML.
2. «**Порт**» - порт сервера, на котором расположен сервиса Sofros AI/ML.
3. «**Идентификатор модели**» - идентификатор модели классификации.
4. «**Вариант наивного Байеса**» - используемый вероятностный алгоритм классификации.
5. «**Игнорировать токены длиной меньше**» - токены в строке, длина которых меньше указанного в настройке значения, будут проигнорированы в процессе классификации.
6. «**Игнорируемые шумовые токены**» - токены, указанные в таблице, будут проигнорированы в процессе классификации.

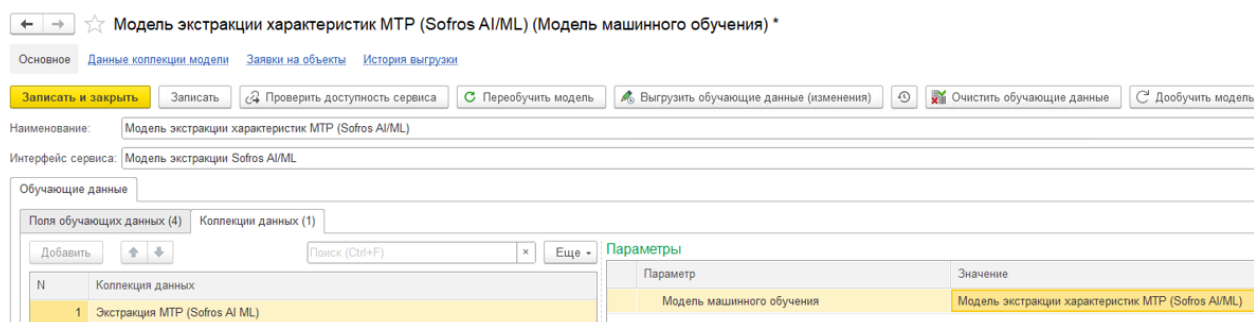
Для использования модели классификации в форме элемента справочника «Материально-технические ресурсы» следует выполнить ряд действий:

1. В форме списка справочника «**Материально-технические ресурсы**» перейти в настройки классификации по команде «Еще» - «**Параметры классификации реквизитов**».
2. В открывшейся форме есть ряд настроек:
 - а. «**Модель МО**» - модель машинного обучения, которая будет использоваться для классификации элементов справочника «**Материально-технические ресурсы**».
 - б. «**Использовать**» - признак использования функционала классификации.
 - с. В таблице «**Параметры реквизитов**» установить набор реквизитов, используемых для классификации.

Модель экстракции. Создание происходит по команде «Создать модель из шаблона» - «Модель экстракции характеристик MTP (Sofros AI/ML)».



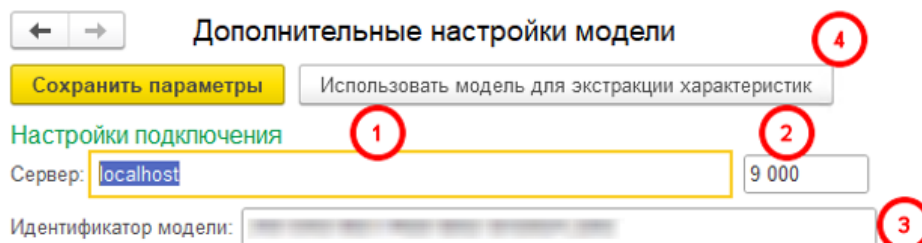
На вкладке «**Коллекции данных**» - «**Параметры**» определен один параметр «Модель машинного обучения», как параметр нужно задать созданную модель.



Далее необходимо перейти по вкладке «**Дополнительные настройки модели**».



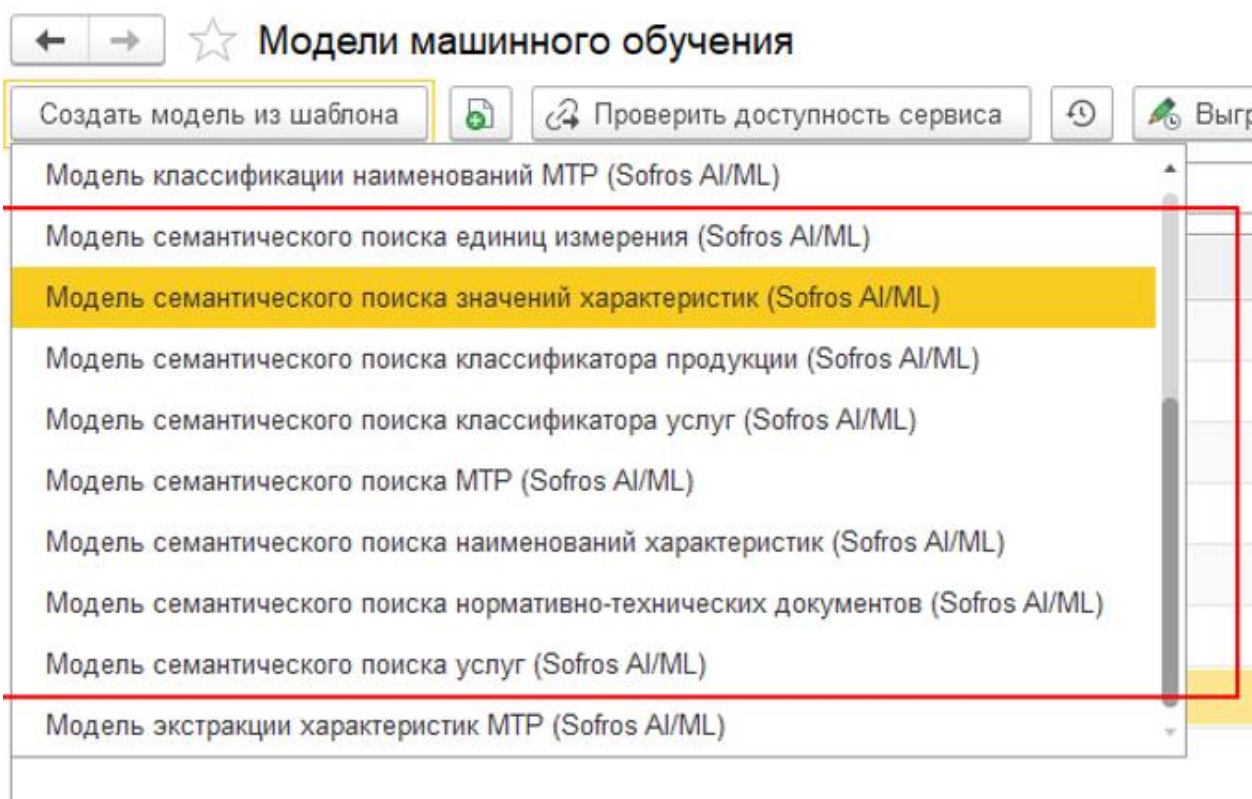
У модели экстракции определен следующий набор настроек:



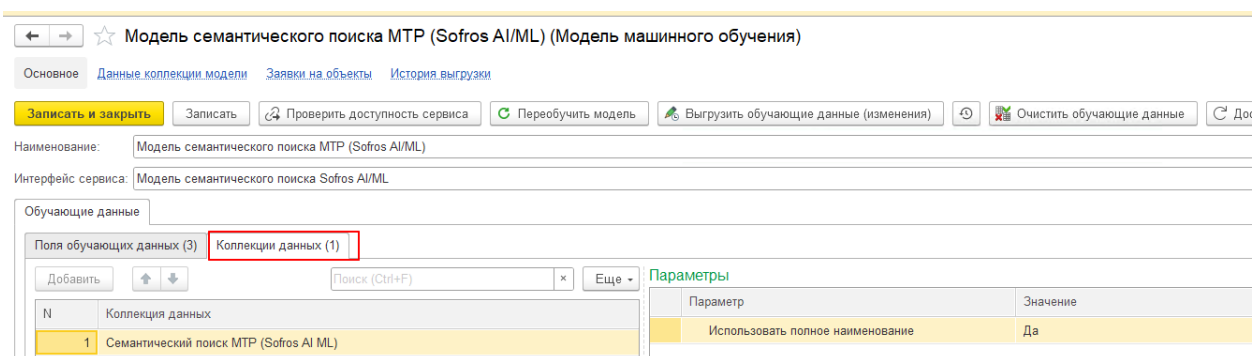
1. «**Сервер**» - адрес сервера, на котором расположен сервис Sofros AI/ML.
2. «**Порт**» - порт сервера, на котором расположен сервиса Sofros AI/ML.
3. «**Идентификатор модели**» - идентификатор модели экстракции.

4. «Использовать модель для экстракции характеристик» - устанавливает данную модель как основную для экстракции характеристик.

Модель семантического поиска. В расширении Sofros AI/ML определено несколько моделей семантического поиска, настройка для всех аналогична, в качестве примера будет представлена настройка модели семантического поиска материально-технических ресурсов – «**Модель семантического поиска MTP (Sofros AI/ML)**».



На вкладке «**Коллекции данных**» - «**Параметры**» определен один параметр «**Использовать полное наименование**», значение по умолчанию – «да». Если установить значение параметра «нет», то поиск будет производиться по реквизиту «**Наименование**».



У модели семантического поиска определен следующий набор настроек:

← →

Дополнительные настройки модели

Сохранить параметры

Настройки подключения

Сервер: localhost

Идентификатор модели: [blurred text]

Дополнительные настройки

Ограничение по количеству результатов:

Минимальная вероятность результата поиска:

Использовать реранк: ☐

Настройки реранка

Использовать старый реранкер: ☐ Нижний порог вероятности реранкера:

1. **«Сервер»** - адрес сервера, на котором расположен сервис Sofros AI/ML.
2. **«Порт»** - порт сервера, на котором расположен сервиса Sofros AI/ML.
3. **«Идентификатор модели»** - идентификатор модели семантического поиска.
4. **«Ограничение по количеству результатов»** - устанавливает максимальное количество результатов поиска по модели. Если включено использование реранка, значение в настройке не может быть выше значения в настройке «Максимальное количество результатов с переранжированием» («Sofros AI/ML» - «Настройки подключения к серверу Sofros AI/ML»).
5. **«Минимальная вероятность результатов поиска»** - результаты поиска с вероятностью ниже значения, указанного в настройке, не будут выводиться.
6. **«Использовать реранк»** - включает использование переранжирования при поиске значения. Увеличивает точность предсказания, но снижает скорость выполнения поиска.
7. **«Использовать старый реранкер»** - включает использование старого алгоритма переранжирования.
8. **«Нижний порог вероятности реранкера»** - результаты переранжирования с вероятностью ниже значения, указанного в настройке, не будут выводиться.

Модель контекстного поиска. Для настройки модели контекстного поиска необходимо создать настройку подключения к сервису Ollama, после чего создать интеллектуального ассистента. Для создания настройки подключения нужно перейти в раздел «Администрирование» - «Настройки работы с ассистентами ИИ» - «Настройки подключения сервисов LLM» и создать новый сервис LLM со следующими реквизитами:

- «Наименование» - произвольное значение.
- «Интерфейс сервиса» - «Sofros AI/ML API».
- «Адрес сервиса» - адрес сервера с Ollama формата **http://{ip_сервера_c_ollama}:11434**. Если Ollama слушает порт, отличный от 11434, его надо заменить.

← → ☆ Подключение сервиса LLM (создание) *

Основное Заявки на объекты

Записать и закрыть Записать ⚙

Наименование:

Интерфейс сервиса: Sofros AI/ML API

Адрес сервиса: http://localhost:11434

После выполнения настроек и записи перейти в раздел «Дополнительные настройки» и заполнить следующие поля:

← → ☆ Sofros AI/ML (Подключение сервиса LLM)

Основное Заявки на объекты

Записать и закрыть Записать ⚙

← → Форма дополнительных настроек

Сохранить и закрыть Сохранить Отмена

Настройки модели

Сервер: 1

Порт: 2

Идентификатор модели: 3

Настройки поиска в интернет

Ключ API: 4

1. «Сервер» - адрес подключения к сервису Sofros AI/ML.
2. «Порт» - порт подключения к сервису Sofros AI/ML.
3. «Идентификатор модели» - идентификатор модели контекстного поиска.
4. «Ключ API» - ключ API сервиса Ollama, необходимый для выполнения web-поиска.

Далее переходим в раздел «Администрирование» - «Настройки работы с ассистентами ИИ» - «Интеллектуальные ассистенты» и создаем нового ассистента со следующими настройками:

← → ☆ Модель контекстного поиска Sofros AI/ML (Интеллектуальный ассистент)

Основное Взаимодействие с ассистентом Заявки на объекты Шаблоны промптов

Записать и закрыть Записать Проверить доступность ассистента

Наименование: Модель контекстного поиска Sofros AI/ML Код: 000000003

Родитель:

Настройки ассистента

Сервис LLM: Sofros AI/ML Модель: llm431b1c10u

Системный промт: Ассистент по характеристикам товаров. Сначала ответь из своих знаний. Если уверен — дай JSON без вызова web_search. Если не хватает данных — вызови web_search только для недостающего. Не вызывай поиск без нужды. Ответ — строго JSON объект, содержащий только наименование характеристики и соответствующее ей значение. Если для какого-то ключа нет данных, значение должно быть null. Числа записывай как числа (без кавычек), true/false как булевы значения. Не обрезай данные.

Дополнительный контент (Нет) Поиск в ИБ (Нет) Поиск в интернет (Нет)

- «Сервис LLM» - выбираем ранее созданный сервис LLM «Sofros AI/ML».
- «Модель» - выбираем модель из списка, предоставленного сервисом Ollama. Инициализация списка может занять некоторое время.

- **«Системный промпт»** - промпт, передаваемый модели до начала диалога с пользователем. Заполняется опционально, если поле не заполнено, подставляется значения по умолчанию из сервиса Sofros AI/ML.

Для настройки web-поиска необходимо перейти на страницу «Поиск в интернет» и установить значение «Истина» для настройки **«Использовать поиск в интернет»**. Для ограничения web-поиска набором определенных ресурсов необходимо указать ресурсы в табличной части и напротив ресурса установить галочку **«Использовать»**.

Дополнительный контент (Нет) Поиск в ИБ (Нет) Поиск в интернет (Нет)

☐ Использовать поиск в интернет

Настройки поиска в интернет

Результатов: 3 Максимальное количество результатов поиска (ссылок).

Дополнять ответ результатом поиска: ☐ Добавлять в ответ информацию об использованных для ответа источниках информации найденных в интернете

Добавить ↑ ↓

Использовать	Ссылка на ресурс
--------------	------------------

Для проверки корректности настроек и доступности сервиса Sofros AI/ML можно использовать команду **«Проверить доступность ассистента»**.

← → ☆ **Модель контекстного поиска Sofros AI/ML (Интеллектуальный ассистент)**

Основное [Взаимодействие с ассистентом](#) [Заявки на объекты](#) [Шаблоны промптов](#)

Записать и закрыть Записать Проверить доступность ассистента