

Программа для ЭВМ
MD Forecast

**РУКОВОДСТВО АДМИНИСТРАТОРА
ПРИЛОЖЕНИЙ
(Руководство по установке)**

Москва, 2025

Содержание

Инструкция по развертыванию и запуску сервиса прогнозирования MD Forecast	3
1. Аннотация	3
2. Назначение и границы инструкции	3
3. Общая схема работы сервиса	3
4. Исходные данные и доступы	4
5. Требования к инфраструктуре и программному обеспечению	4
5.1. Инфраструктура	4
5.2. Программные компоненты	4
6. Подготовка данных из 1С	5
6.1. Выгрузка данных	5
6.2. Структура каталогов на SFTP	5
6.3. Требования к CSV-файлам	5
7. Развертывание и настройка Airflow	5
8. Настройка и выполнение пайплайнов	6
9. Запуск прогнозных расчетов	6
10. Проверка корректности развертывания	6
11. Контактная информация	7

Инструкция по развертыванию и запуску сервиса прогнозирования MD Forecast

1. Аннотация

Настоящий документ содержит описание порядка развертывания и запуска сервиса прогнозирования MD Forecast (далее — ПО). ПО представляет собой облачный сервис прогнозирования, развертываемый в инфраструктуре облачного провайдера. Локальная установка программного обеспечения на стороне конечных пользователей не требуется. Документ носит **описательный характер** и не является детальной пошаговой инструкцией по установке. Документ предназначен для специалистов, осуществляющих развертывание, настройку и запуск экземпляра ПО.

2. Назначение и границы инструкции

Данная инструкция описывает:

- состав и назначение компонентов сервиса;
- общий порядок развертывания и запуска;
- входные данные и доступы, необходимые для начала работ;
- критерии успешного развертывания.

Инструкция **не описывает**:

- внутреннюю реализацию ML-моделей;
- доработку алгоритмов;
- пользовательскую работу в системе 1С;
- детальные команды и конфигурационные файлы.

3. Общая схема работы сервиса

Сервис функционирует в автоматическом режиме обработки данных:

1. Данные выгружаются из системы 1С в формате CSV.
2. Данные загружаются в облачное хранилище и обрабатываются.
3. Выполняется прогнозирование с использованием ML-моделей.
4. Результаты прогнозов возвращаются в систему 1С и используются в отчетах.

Прямое пользовательское взаимодействие с сервисом отсутствует.

4. Исходные данные и доступы

Для начала развертывания и запуска сервиса специалисту необходимо получить от вендора следующие материалы и доступы:

- доступ к облачной среде, в которой выполняется развертывание (целевой сценарий — Яндекс.Облако);
- доступ к развернутому экземпляру **Apache Airflow**;
- преднастроенные DAG-пайплайны обработки данных;
- параметры подключения к **SFTP-серверу**;
- параметры доступа к **S3-хранилищу**;
- описание форматов входных CSV-файлов;
- сведения о формате файла-триггера, сигнализирующего о завершении выгрузки данных.

Детальные технические параметры и конфигурации предоставляются вендором отдельно.

5. Требования к инфраструктуре и программному обеспечению

5.1. Инфраструктура

Для развертывания сервиса используется облачная инфраструктура.

Рекомендуемая среда развертывания — Яндекс.Облако.

Используемые ресурсы:

- объектное хранилище (S3);
- виртуальные машины для обработки данных и запуска моделей;
- сетевые сервисы для взаимодействия компонентов.

5.2. Программные компоненты

В составе сервиса используются:

- **Apache Airflow** — для оркестрации пайплайнов;
- **Python и R** — для реализации обработки данных и ML-моделей;
- SFTP-сервер — для приема данных из 1С;
- средства доступа к S3-хранилищу.

Локальная установка ПО в виде приложения не предусмотрена.

6. Подготовка данных из 1С

6.1. Выгрузка данных

Необходимо настроить автоматическую выгрузку данных из системы **1С** в формате CSV на SFTP-сервер.

6.2. Структура каталогов на SFTP

На SFTP-сервере используется следующая структура каталогов:

Каталог	Назначение
1С	Основные данные для прогнозирования
bi	Дополнительные данные для аналитики и отчетности
FS	Сценарные корректировки прогноза

Каталог **FS** используется для загрузки CSV-файлов с коэффициентами корректировок, влияющих на итоговый прогноз без изменения ML-моделей.

6.3. Требования к CSV-файлам

CSV-файлы должны соответствовать следующим требованиям:

- кодировка — UTF-8;
- разделитель — точка с запятой (;);
- структура файлов соответствует согласованному формату.

7. Развертывание и настройка Airflow

Apache Airflow используется для управления всеми этапами обработки данных и запуска прогнозных расчетов.

Возможные варианты развертывания:

- в Docker (по согласованию);
- непосредственно в облачной инфраструктуре (предпочтительный вариант).

В Airflow настраиваются:

- подключения к SFTP-серверу;
- подключения к S3-хранилищу;
- доступ к вычислительным ресурсам.

8. Настройка и выполнение пайплайнов

В Airflow должны быть доступны три преднастроенных пайплайна:

1. Пайплайн обработки данных из каталога 1С.
2. Пайплайн обработки данных из каталога bi.
3. Пайплайн обработки данных из каталога FS.

Каждый пайплайн:

- ожидает файл-триггер завершения выгрузки;
- загружает CSV-файлы с SFTP;
- преобразует данные в формат Parquet;
- загружает данные в S3-хранилище;
- инициирует запуск виртуальных машин для обогащения данных.

9. Запуск прогнозных расчетов

После завершения подготовки данных автоматически запускается pipeline прогнозирования.

Прогнозные расчеты:

- выполняются на виртуальных машинах в облаке;
- реализованы с использованием Python- и R-скриптов;
- не требуют ручного запуска со стороны пользователя.

10. Проверка корректности развертывания

Развертывание и запуск считаются успешными при выполнении следующих условий:

- Apache Airflow доступен и функционирует;
- все преднастроенные DAG-пайплайны отображаются в интерфейсе Airflow;
- данные успешно загружаются в S3-хранилище;
- отсутствуют ошибки выполнения пайплайнов;

- прогнозные результаты формируются и передаются в систему 1С.

11. Контактная информация

По вопросам развертывания и запуска ПО обращаться:

ФИО: Старостин Егор Александрович

Электронная почта: starostin_e@serviceplanet.ru

Телефон: +7 926 724-30-85