

ТЕХНИЧЕСКОЕ ЗАДАНИЕ

# ЗАДАЧА 15

Модель раннего обнаружения  
неисправностей промышленного  
оборудования



## 1. Термины и определения:

Предиктивное ТОиР (predictive maintenance)	Предиктивное техническое обслуживание с упором на прогнозирование отказов и принятие мер, в зависимости от состояния оборудования, для предотвращения отказа или ухудшения характеристик.
Неисправности	Неисправности из перечня М1, М3
М1	Отказы с простоями оборудования
М3	Неисправности без простоя

## 2. Актуальность задачи

Объектами для реализации в рамках «хакатона» выбраны эксгаустеры на агрегатах ЧерМК ПАО «Северсталь»:

1. КОМПЛЕКС АГЛОМАШИНА №4
2. КОМПЛЕКС АГЛОМАШИНА №5
3. КОМПЛЕКС АГЛОМАШИНА №6
4. КОМПЛЕКС АГЛОМАШИНА №7
5. КОМПЛЕКС АГЛОМАШИНА №8
6. КОМПЛЕКС АГЛОМАШИНА №9

Эксгаустер является критически важным компонентом комплекса агломашины - его неплановая остановка приводит к остановке агломашины и, как следствие, к потерям из-за недовыпуска продукции.

Основной причиной остановок эксгаустера является износ ротора. Износ зависит от различных факторов и поэтому его срок службы непостоянен. Заблаговременное прогнозирование превышения предельно допустимых параметров позволит провести замену ротора в плановые остановки агломашины на обслуживание, что позволит исключить (или значительно снизить) внеплановые простои.

### **3. Постановка задачи**

Создание прототипа системы предиктивной аналитики (СПА), предназначенной для:

1. Наиболее полного и своевременного обнаружения скрытых дефектов в оборудовании, оценки возможности повышения уровня надежности эксплуатации оборудования за счет внедрения автоматизированной системы, основанной на применении методов предиктивной аналитики.
2. Сокращение времени простоев оборудования вследствие сокращения непредвиденных остановок по причине развивающихся дефектов
3. Увеличение коэффициента технической готовности технологического оборудования вследствие своевременной диагностики и технического обслуживания.
4. Повышение эффективности технического обслуживания и ремонта оборудования вследствие предиктивного прогнозирования сбоев в работе оборудования на основании анализа данных в режиме реального времени.
5. Повышение качества и снижение сроков ремонта оборудования вследствие своевременного заказа запасных частей на основе предиктивного анализа состояния оборудования.

В рамках задачи требуется реализовать действующий прототип системы предиктивной аналитики, способной прогнозировать неисправности промышленного оборудования.

Основные требования к прогнозной способности моделей на историческом тестовом датасете:

1. Определить наличие или отсутствие неисправности М1 на заданных интервалах тестовой выборки (временные ряды сигналов на тесте имеют искусственные пропуски).
2. Определять периоды, когда были любые неисправности М3 (аномальный режим работы техместа) на протяжении всего тестового интервала.
3. Определить время до простоя М1 с максимально возможным горизонтом - для каждой точки временного ряда тестовой выборки и указать вероятное значение времени до наступления отказа оборудования.
4. Краткое описание оборудования

Эксгаустер — это центробежный нагнетатель, основной задачей которого в агломерационном производстве является просос воздуха через слой шихты, лежащей на колосниковой решётке агломерационной машины.

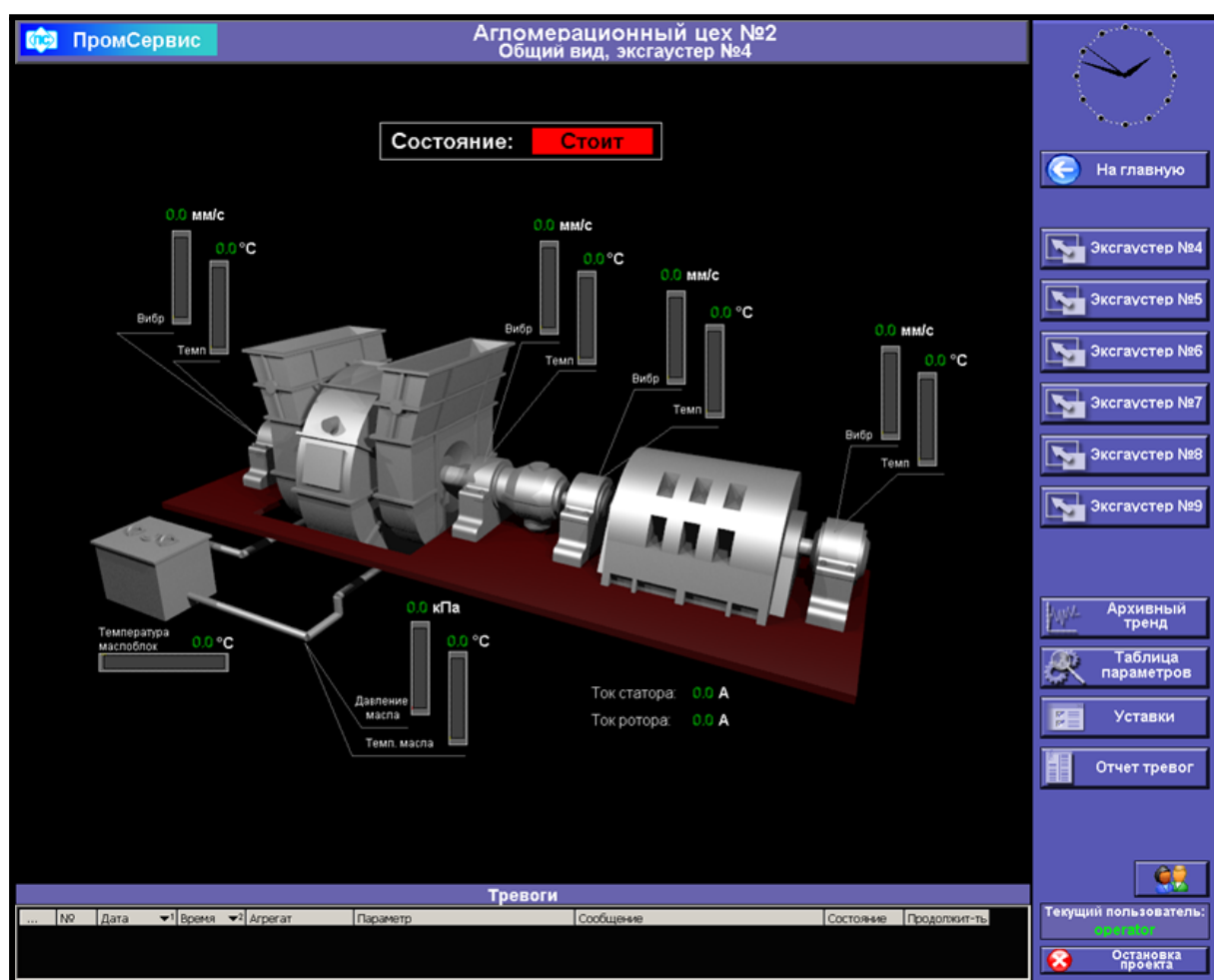
Основные узлы эксгаустера:

- Ротор
- Электродвигатель
- Корпус
- Подшипник опорно-упорный
- Подшипник опорный
- Электроаппаратура
- Кожух муфты

### Характеристики:

- Объемная производительность, отнесенная к условиям всасывания, м<sup>3</sup>/мин- 8000
- Разность между абсолютными статическими конечным и начальным давлениями, мм вод.ст. (кПа)- 1290 (12,6)
- Начальная температура газа, °C- 150
- Начальное давление газа, кгс/см<sup>2</sup> (кПа) - 0,95 (93.1)
- Потребляемая мощность, кВт- 1980
- Мощность электродвигателя, кВт- 2500

### Общий вид эксгаустера





## 5. Требования к решению

**5.1.** В качестве программных средств для реализации модели можно использовать любые методы построения моделей, которые позволят воспроизвести построение модели из исходного кода и результаты прогнозирования (предпочтительно на языке python).

**5.2.** Участники получают обучающую выборку с историческими данными (временные ряды сигналов, событий о неисправностях, перечень наименований сигналов) и тестовую выборку (временные ряды сигналов, наименования сигналов).

**5.3.** Участники должны разработать веб-сервис, состоящий из модулей:

- Модуль визуализации и выдачи прогнозов о неисправностях и прогнозе даты поломки (отказа) оборудования.
- Модуль расчета и визуализации оценки точности моделей, **находящихся в эксплуатации.**

**5.4.** Используя полученные данные, Участники генерируют модели на тренировочной выборке сигналов, показывают результаты работы данных моделей во времени (ускоренно, на исторических данных как тренировочного, так и тестового набора данных), визуализируют расчет и метрики точности, в соответствии с требованиями настоящего ТЗ и формируют экспорт файла csv с предсказаниями неисправностей на тестовой выборке в соответствующем формате (см. приложение 2).

**5.5.** Для решения задачи №3 для каждого техместа требуется предоставить временной ряд с разметкой, содержащей прогнозируемое время до наступления неисправности (M1).

**5.6.** По окончании выполнения задания Участниками, проводится оценка удобства интерфейса веб-сервиса, полнота исполненного функционала, точность моделей и их фактическая способность выполнять прогнозирование неисправностей и отказов за максимально больший горизонт времени. Методика и критерии оценки изложены в Приложении №1.

### **5.7. Требования к функциям системы**

**5.7.1.** Отслеживание состояния оборудования в реальном времени, при различных режимах функционирования оборудования и его узлов на основе анализа совокупных изменений значений технологических параметров. В рамках хакатона подразумевается симуляция на исторических данных;

**5.7.2.** Осуществление в режиме «реального времени» (на историческом периоде) сравнительного анализа отклонений текущих изменений значений параметров состояния оборудования от значений его нормального и работоспособного состояния;

**5.7.3.** Автоматическое оповещение пользователей о соответствии/несоответствии текущего состояния работы агрегатов штатному режиму функционирования;

**5.7.4.** В случае несоответствия штатному режиму осуществлять классификацию неисправности с указанием проблемного элемента (отдельной части, узла) агрегатов и определение наступления момента отказа оборудования;

**5.7.5.** Формирование необходимой отчетности о текущем состоянии агрегатов с указанием параметров, которые отклоняются от нормальных значений и возможных зарождающихся дефектов агрегатов;

**5.7.6.** Для реализации вышеуказанного функционала СПА должна иметь как минимум следующие возможности:

**5.7.6.1.** Сбор исторических данных для создания моделей предиктивной диагностики. Должна поддерживаться загрузка информации как минимум из локальных источников, предоставленных Участникам;

**5.7.6.2.** Тестирование предиктивных моделей оборудования;

**5.7.6.3.** Хранение данных исторических и текущих данных оборудования, необходимых для построения и выполнения предиктивных моделей;

**5.7.6.4.** Автоматическое уведомления персонала о появлении предупреждения посредством электронной почты или иных способов электронной коммуникации.

**5.7.6.5.** Иметь в своем составе средства логирования результатов работы моделей и оповещения пользователей;

**5.7.6.6.** Возможность модификации количества и структуры предиктивных моделей оборудования (иерархия)

**5.7.6.7.** Иметь архитектуру, допускающую возможность расширения в рамках возможного последующего тиражирования.

**5.8. Требования к тестированию предиктивных моделей на этапе внедрения и эксплуатации:**

СПА должна иметь возможность генерации и проведения тестирования предиктивных моделей, которое заключается в анализе результатов работы предиктивных моделей на исторических эксплуатационных данных с целью подтверждения частоты определения скрытых дефектов в оборудовании, а также в подтверждении возможности заблаговременного определения предиктивными моделями описанных в техническом задании неисправностей.



## **5.9. Требования к функционированию предиктивных моделей**

В СПА должна быть реализована функция выполнения предиктивных моделей оборудования. Результатом выполнения должно являться периодическое последовательное выполнение алгоритма прогнозирования неисправности. Режим работы СПА постоянный, обеспечивающий функционирование моделей и оповещение в круглосуточном режиме (24/7).

## **5.10. Требования к хранению предиктивных моделей**

В СПА должна быть реализована функция хранения предиктивных моделей и полученных с помощью их функционирования результатов - предупреждений о развивающихся неисправностях.

## **5.11. Требования к мониторингу результатов работы системы**

В СПА должна быть реализована возможность наблюдения результатов работы в виде графического интерактивного интерфейса и формирование отчетов в популярных форматах (например pdf).

# **6. Возможный пользовательский путь**

## **6.1. Модуль визуализации и выдачи прогнозов о неисправностях**

**6.1.1.** Пользователь – оператора или диспетчер системы мониторинга оборудования или любой специалист цеха в производстве.

**6.1.2.** У пользователя есть визуальная информация о текущем состоянии контролируемого оборудования. Требования к визуальной части не устанавливаются, за исключением того, что пользовательский интерфейс должен быть удобным для восприятия информации.

**6.1.3.** Визуальная часть позволяет отслеживать состояние интересующего оператора оборудования и перемещает фокус оператора на проблемный объект контроля в случае ухудшения прогноза его «здоровья»

**6.1.4.** Элемент интерфейса, в котором происходит динамическая визуализация состояния оборудования имеет возможность оценить время до отказа оборудования и возможный период его простоя.

**6.1.5.** Функционал модуля позволяет автоматизированную отчетность о прогнозе деградации оборудования всем подписантам на данный сервис. Предполагается, что, как и в интерфейсе модуля, так и в сообщениях (по email), при возникновении проблемы со «здоровьем» оборудования будет визуализироваться (отправляться) диагностика, имеющая, как минимум поля:

- Наименование оборудования (локализация оборудования на котором есть проблема);
- Пользовательское описание предупреждения (описание вероятного дефекта);
- Прогнозная дата отказа;

## **6.2. Модуль оценки точности моделей**

**6.2.1.** Модуль оценки точности функционально в виде виджета встроен в элементы интерфейсов для оценки точности моделей.

## **7. Требования к презентации**

### **7.1. Формат презентации pdf/pptx**

### **7.2. Обязательные разделы:**

- краткое обоснование выбора метода решения,
- указание используемых технологий,
- достигнутые результаты за время конкурса

## 8. Требования к сопроводительной документации

### 8.1. Формат pdf/docx

### 8.2. Обязательные разделы:

- инструкция по запуску/разворачиванию сервиса;
- технологический стек решения;
- описание работы и взаимодействия с моделью;
- возможный потенциал решения в будущем;
- ссылки на доп. материал.

## 9. Источники данных

Участник должен самостоятельно выбрать наиболее релевантные технологические сигналы из доступного перечня сигналов АСУТП оборудования для построения предиктивной модели и обеспечения наиболее качественного прогноза неисправностей.

Перечень предоставляемых файлов:

- X\_train.parquet – тренировочная выборка;
- y\_train.parquet – таргет тренировочной выборки;
- X\_test.parquet – тестовая выборка;
- test\_intervals.xlsx – список интервалов с пропущенными значениями (в рамках задачи №1 из п.3);
- messages.xlsx – дополнительная информация по неисправностям в тренировочной выборке (описание, тип простоя, длительность простоя и т.д.).

Подробнее структура данных представлена в Приложении 1.

## **10. Требования к сдаче решения**

**10.1.** Предоставить результат достигнутой точности модели согласно Приложения №2.

**10.2.** Предоставить ссылку на MVP (прототип) системы для демонстрации работы

**10.3.** Ссылку на общедоступный репозиторий с исходным кодом.

**10.4.** Ссылку на сопроводительную документацию.

**10.5.** Ссылку на презентацию решения.

## **11. Критерии оценки (промежуточные и финальные)**

### **11.1. Подход коллектива к решению задачи.**

Полнота реализованных требований, изложенных в п.5 «Требования к решению» настоящего ТЗ

### **11.2. Техническая реализация.**

Технические показатели разработанных моделей применяются к результатам ретроспективного анализа на тестовой выборке. Методика оценки качества моделей представлена в Приложении 2

### **11.3. Соответствие решения поставленной задаче.**

- соответствие функционала пункту 5.7 настоящего ТЗ;
- фактическая способность выполнять прогнозирование неисправностей и отказов за максимально больший горизонт времени.

### **11.4. Эффективность решения в рамках поставленной задачи**

- максимальное количество корректно предсказанных неисправностей на тестовой выборке;
- скорость работы сервиса;

- удобства интерфейса веб-сервиса.

#### 11.5. Выступление на питч-сессии:

- убедительность и информативность;
- лаконичные и аргументированные ответы;
- соответствие регламенту выступления.

#### **Приложение №1**

- Формат структуры тренировочной выборки (X\_train и y\_train)
- M1 - отказы с простоями оборудования (1 если событие происходит в момент времени)
- M3 - неисправности без простоя (2 если событие происходит в момент времени)
- IDTM\_N – Признак неисправности (M1 и M3) по техническому месту N (таргет)
- IDS\_N – Значение сигнала N

DateTime	IDS_1	IDS_2	...	IDS_N	IDTM_1	...	IDTM_N
01.01.2020 0:00	0.25	0.75	...	0.75	0	...	0
01.01.2020 0:01	0.26	0.76	...	0.76	0	...	0
01.01.2020 0:02	0.27	0.77	...	0.11	0	...	0
01.01.2020 0:03	0.28	0.78	...	0.54	0		0
01.01.2020 0:04	0.29	0.79	...	0.119	0		0
01.01.2020 0:05	0.30	0.80	...	0.184	0		0
01.01.2020 0:06	0.31	0.81	...	0.249	0		0
01.01.2020 0:07	0.32	0.82	...	0.314	0		0
01.01.2020 0:08	0.33	0.83	...	0.379	0		0
01.01.2020 0:09	0.34	0.84	...	0.444	2		2
..	...	...	...	...	0	...	0
01.01.2020 10:02	0.27	0.75	...	0.15	0	...	0
01.01.2020 10:03	0.20	0.78	...	0.16	0	...	0
01.01.2020 10:03	0.20	0.71	...	0.01	0	...	0
01.01.2020 10:02	0.27	0.75	...	0.15	0		0
01.01.2020 10:02	0.20	0.78	...	0.16	0		0
01.01.2020 10:02	0.20	0.71	...	0.01	2		2
01.01.2020 10:02	0.27	0.75	...	0.15	2		2
01.01.2020 10:02	0.20	0.78	...	0.16	2		2
01.01.2020 10:02	0.20	0.71	...	0.01	2		2
01.01.2020 10:02	0.27	0.75	...	0.15	2		2
01.01.2020 10:02	0.20	0.78	...	0.16	1	...	1
01.01.2020 10:02	0.20	0.71	...	0.01	1	...	1
01.01.2020 10:02	0.27	0.75	...	0.15	1	...	1

В наборе сигналов тестовой выборки (X\_test.parquet) присутствуют интервалы с пропущенными значениями. Список интервалов предоставлен в отдельном файле (test\_intervals.xlsx). Для этих интервалов необходимо предсказать наличие неисправности. В случае предсказанной неисправности, определить тип и время начала развития и наступления неисправности.



Формат списка с интервалами (test\_intervals.xlsx):

	start	finish	tm
0	2022-09-10 23:26:41	2022-09-11 01:25:58	NaN
1	2022-01-09 06:57:35	2022-01-09 08:30:54	NaN
2	2022-09-25 14:07:45	2022-09-25 15:35:51	NaN
3	2022-04-03 09:29:12	2022-04-03 10:58:54	NaN
4	2022-09-09 13:23:11	2022-09-09 14:24:22	NaN
...	...	...	...
76	2023-02-12 10:11:52	2023-02-12 11:40:53	NaN
77	2022-06-11 09:44:06	2022-06-11 11:31:58	NaN
78	2022-01-21 17:28:44	2022-01-21 19:16:22	NaN
79	2022-11-20 02:54:35	2022-11-20 04:24:06	NaN
80	2022-07-29 11:06:38	2022-07-29 12:36:52	NaN

Формат структуры тестовой выборки (X\_test.parquet):

DateTime	IDS_1	IDS_2	...	IDS_N
01.01.2022 0:00	0.25	0.75	...	0.75
01.01.2022 0:01	0.26	0.76	...	0.76
01.01.2022 0:02	0.27	0.77	...	0.11
..	...	...	...	...
01.01.2022 10:02	0.27	0.75	...	0.15
01.01.2022 10:03	0.20	0.78	...	0.16
01.01.2022 10:03	0.20	0.71	...	0.01

## Приложение №2

### Методика оценки результата

1. По задаче №1 результаты оцениваются по метрике F1-score по заполненной форме списка интервалов в качестве теста для M1.
2. По задаче №2 результаты оцениваются проверкой совпадения предсказанного периода неисправности M3 для каждого техместа\* с фактическим периодом неисправности M3 по следующей формуле:

$$J = \frac{TP}{TP+FN+FP}, \text{ где}$$

TP – сумма моментов времени, верно определенных как M3

FP – сумма моментов времени, ошибочно определенных как M3

FN – сумма моментов времени, ошибочно определенных как не M3

\* кейсы с малым количеством M3 могут быть исключены из оценки данной метрики

3. В задаче № 3 предсказанное время до наступления М1 оценивается по взвешенной метрике WRMSE для каждого техместа в соответствии с формулой:

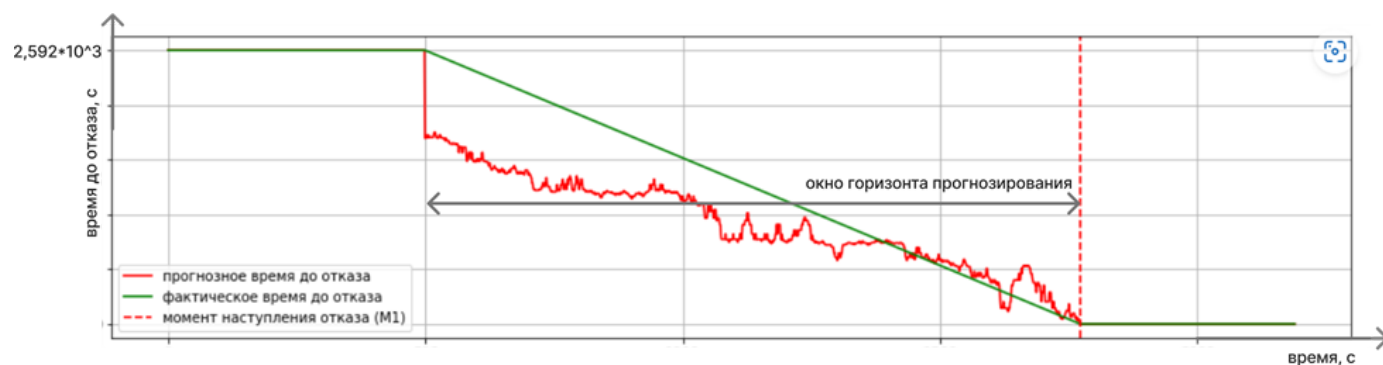
$$WRMSE = \sqrt{\sum_{i=1}^n \frac{w_i (\hat{x}_i - x_i)^2}{n}}$$

$w_i = \frac{1}{\alpha t_i}$ , где  $t_i$  - фактическое время до наступления отказа М1

для момента времени, соответствующего  $i$ ,  $\alpha$  - коэффициент масштабирования штрафа по времени

$x_i$  - фактическое время до отказа для момента времени  $i$ ;

$\hat{x}_i$  - прогнозируемое время до отказа для момента времени  $i$ .



Итоговая метрика WRMSE усредняется по всем техместам.

Метрика WRMSE оценивается на максимальном горизонте прогнозирования в месяц от момента наступления М1.

В процессе оценки целесообразности определенных подходов в прогнозировании, подходы к оценке метрик могут быть скорректированы.

## Формат представления результата для задачи 2 и задачи 3

		Задача для тестовой выборки	
		Задание 2	Задание 3
		Определить периоды, когда были неисправности M3 для каждого техмоста (таких рядов будет N)	Спрогнозировать время до наступления M1 (с)
	DateTime	M3	M1
	01.01.2020 0.00	0	3600
	01.01.2020 0.01	0	3590
	01.01.2020 0.02	0	3580
	01.01.2020 0.03	2	3570
	01.01.2020 0.04	2	3560
	01.01.2020 0.05	2	3550
	01.01.2020 0.06	2	3540
	01.01.2020 0.07	2	3530
	01.01.2020 0.08	2	3520
	01.01.2020 0.09	2	3510
	...	...	...
	01.01.2020 10.02	2	3300
	01.01.2020 10.03	2	3290
	01.01.2020 10.03	2	3280
	01.01.2020 10.02	2	3270
	01.01.2020 10.02	2	3260
	01.01.2020 10.02	2	3250
	01.01.2020 10.02	2	3240
	01.01.2020 10.02	0	3230
	01.01.2020 10.02	0	3220
	01.01.2020 10.02	0	3210
	01.01.2020 10.02	0	3200
	01.01.2020 10.02	0	3190
	01.01.2020 10.02	0	3180