

Правительство Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»
(НИУ ВШЭ)

Московский институт электроники и математики им. А.Н. Тихонова

ПРАКТИЧЕСКАЯ РАБОТА № 14

ТЕМА РАБОТЫ

«Анализ и обнаружение аномалий в IoT-данных»

Москва, 2025

Цели работы.....	2
Целевая аудитория.....	2
Идея и концепция.....	3
Содержание практической работы.....	3
О наборе данных и задаче работы.....	3
Работа с данными.....	4
Загрузка данных.....	4
Предобработка исходных данных.....	5
Установка и использование Anomaly Detection Extension.....	8
Построение и настройка процесса детекции аномалий.....	9
Визуализация и анализ результатов.....	13
Пример выполненной работы.....	19
Приобретенные навыки.....	19
Обобщенная задача для выполнения индивидуального варианта.....	19
Распределение вариантов.....	21

Цели работы

- Освоить приемы загрузки и предобработки датасета, содержащего телеметрию с IoT-сенсоров (параметры газа, температуры, влажности, и т.д.).
- Научиться использовать расширение Anomaly Detection в RapidMiner для выявления аномальных наблюдений (выбросов).
- Разобраться, как интерпретировать полученные результаты и визуализировать аномалии, чтобы находить потенциальные нештатные ситуации в реальном оборудовании.

Целевая аудитория

Практическая работа ориентирована на студентов, знакомых с базовой работой в RapidMiner и желающих исследовать задачу автоматического обнаружения аномалий на данных IoT.

Идея и концепция

Датасет Environmental Sensor Telemetry Data содержит показания нескольких датчиков с различными параметрами:

column	description	units	
ts	timestamp of event	epoch	Цель – выявить «подозрительные» наблюдения, которые существенно отличаются от остальных и могут указывать на аномальную активность (слишком высокую концентрацию CO, слишком высокую или низкую температуру, нетипичные комбинации признаков и т.д.).
device	unique device name	string	
co	carbon monoxide	ppm (%)	
humidity	humidity	percentage	
light	light detected?	boolean	Для этого предлагается использовать операторы из Anomaly Detection Extension и сравнить, как они обрабатывают данные.
lpg	liquid petroleum gas	ppm (%)	
motion	motion detected?	boolean	Содержание практической работы
smoke	smoke	ppm (%)	О наборе данных и задаче работы
temp	temperature	Fahrenheit	

Набор данных: [Environmental Sensor Telemetry Data](#)

Объем: >130 000 строк. Каждая строка содержит время (ts в формате epoch), устройство, показатели газов (co, lpg, smoke), а также humidity (влажность), temp (температура в °F) и булевы флаги light, motion.

Задача:

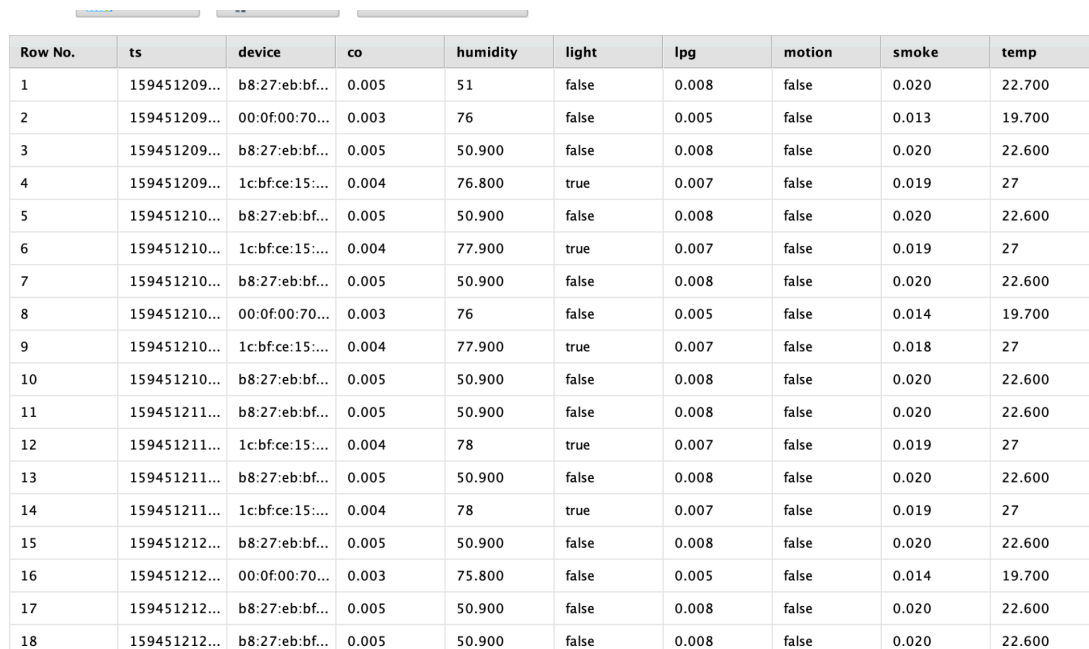
Выявление аномалий — определить, в какие моменты показатели сенсоров могут считаться ненормальными с точки зрения их значений, сочетания признаков и т.д.

Работа с данными

Загрузка данных

Read CSV:

- В RapidMiner в панели Operators найдите Read CSV.
- Перетащите его в на панель процесса.
- В параметрах укажите путь к файл формата .csv.
- Нажмите Guess Types, чтобы RapidMiner определил типы столбцов.



Row No.	ts	device	co	humidity	light	lpg	motion	smoke	temp
1	159451209...	b8:27:eb:bf...	0.005	51	false	0.008	false	0.020	22.700
2	159451209...	00:0f:00:70...	0.003	76	false	0.005	false	0.013	19.700
3	159451209...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600
4	159451209...	1c:bf:ce:15:...	0.004	76.800	true	0.007	false	0.019	27
5	159451210...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600
6	159451210...	1c:bf:ce:15:...	0.004	77.900	true	0.007	false	0.019	27
7	159451210...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600
8	159451210...	00:0f:00:70...	0.003	76	false	0.005	false	0.014	19.700
9	159451210...	1c:bf:ce:15:...	0.004	77.900	true	0.007	false	0.018	27
10	159451210...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600
11	159451211...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600
12	159451211...	1c:bf:ce:15:...	0.004	78	true	0.007	false	0.019	27
13	159451211...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600
14	159451211...	1c:bf:ce:15:...	0.004	78	true	0.007	false	0.019	27
15	159451212...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600
16	159451212...	00:0f:00:70...	0.003	75.800	false	0.005	false	0.014	19.700
17	159451212...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600
18	159451212...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600

ExampleSet (405,184 examples,0 special attributes,9 regular attributes)

рис. 1: Загрузка датасета

Проверка типов:

- co, humidity, lpg, smoke, temp должны быть Real.
- light и motion (True/False) должны распознаваться как Nominal.
- device — Nominal (строка).
- ts — Real (epoch).

	Name	Type	Missing	Statistics			Filter (9 / 9 attributes): <input type="text" value="Search for Attribute"/>
Data	ts	Real	0	Min 1594512094.386	Max 1595203417.264	Average 1594858017.297	
Statistics	device	Nominal	0	Least 1c:bf:ce [...] (105918)	Most b8:27:eb [...] (187451)	Values b8:27:eb:bf:9d:51 (187451), 00:0f:00:70:91:0a (111815), ...[1 mo]	
Visualizations	co	Real	0	Min 0.001	Max 0.014	Average 0.005	
Annotations	humidity	Real	0	Min 1.100	Max 99.900	Average 60.512	
	light	Nominal	0	Least true (112527)	Most false (292657)	Values false (292657), true (112527)	
	lpg	Real	0	Min 0.003	Max 0.017	Average 0.007	
	motion	Nominal	0	Least true (482)	Most false (404702)	Values false (404702), true (482)	
	smoke	Real	0	Min 0.007	Max 0.047	Average 0.019	
	temp	Real	0	Min 0	Max 30.600	Average 22.454	

рис.2: Проверка типов переменных

Просмотр данных:

- Нажмите правой кнопкой мыши на выходной порт (out) оператора Read CSV и выберите Show Data, чтобы убедиться, что данные загрузились корректно (количество строк/столбцов).

Предобработка исходных данных

В данном датасете нет пропущенных значений, но обычно этот шаг – один из первостепенных в предобработке данных (оператор – Replace Missing Values).

Исключение столбцов:

- Если необходимо исключить столбцы, не нужные для поиска аномалий (например, device), это можно сделать с помощью оператора Select Attributes.
- Откройте расширенные параметры оператора, выберите exclude attributes, а в типе фильтрации – a subset.
- Во вкладке Select attributes выберите необходимые параметры, которые не будут использоваться в дальнейшем.
- Также в этой работе можно оставить все, чтобы посмотреть роль устройства и boolean-флагов (используется этот вариант).

The screenshot shows the 'Parameters' window for the 'Select Attributes' operator. The window has a title bar with 'Process' and 'Parameters' tabs. The 'Parameters' tab is selected. Below the title bar, there is a section titled 'Select Attributes' with a grid icon. The parameters are as follows:

Parameter	Value
type	exclude attributes
attribute filter type	a subset
select subset	

At the bottom, there is a checkbox labeled 'also apply to special attributes (id, label..)' which is unchecked.

рис.3: Расширенные параметры оператора *Select attributes*


Преобразование времени:

- В данной работе ts содержит данные за период: 07/12/2020 00:00:00 UTC – 07/19/2020 23:59:59 UTC.
- В датасете изначально данные представлены в формате 1.5952034172643125E9. Это секунды с эпохи Unix. Приведем их к понятному временному виду.
- Воспользуйтесь функцией *Generate Attributes*, чтобы перевести секунду в миллисекунды (так как оператор *Numerical to Date* в RapidMiner ожидает данные в миллисекундах).
- Создайте новый атрибут с миллисекундами(умножение на 1000):
 $ts_millis = ts * 1000$.
- Затем в операторе *Numerical to Date* используйте именно этот новый атрибут.

Generate Attributes

function descriptions

Edit Parameter List: **function descriptions**
List of functions to generate.

column name	function expressions
ts_millis	ts * 1000 







 Add Entry
 Remove Entry
 Apply
 Cancel

рис.4: Создание новой переменной с помощью оператора *Generate attributes*

Process
Parameters 

 **Numerical to Date**

attribute name

☐ *keep old attribute*

time offset

рис.5: Создание новой переменной типа *datetime* с помощью оператора *Numerical to Date*

Open in Turbo Prep Auto Model Interactive Analysis Filter (405,184 / 405,184 examples): all

Row No.	ts	device	co	humidity	light	lpg	motion	smoke	temp	ts_millis
1	159451209...	b8:27:eb:bf...	0.005	51	false	0.008	false	0.020	22.700	Jul 12, 202...
2	159451209...	00:0f:00:70...	0.003	76	false	0.005	false	0.013	19.700	Jul 12, 202...
3	159451209...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	2020-07-12T03:01:34.735+03:00	
4	159451209...	1c:bf:ce:15:...	0.004	76.800	true	0.007	false	0.019	Press "F3" for focus.	
5	159451210...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020		
6	159451210...	1c:bf:ce:15:...	0.004	77.900	true	0.007	false	0.019	27	Jul 12, 202...
7	159451210...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600	Jul 12, 202...
8	159451210...	00:0f:00:70...	0.003	76	false	0.005	false	0.014	19.700	Jul 12, 202...
9	159451210...	1c:bf:ce:15:...	0.004	77.900	true	0.007	false	0.018	27	Jul 12, 202...
10	159451210...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600	Jul 12, 202...
11	159451211...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600	Jul 12, 202...
12	159451211...	1c:bf:ce:15:...	0.004	78	true	0.007	false	0.019	27	Jul 12, 202...
13	159451211...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600	Jul 12, 202...
14	159451211...	1c:bf:ce:15:...	0.004	78	true	0.007	false	0.019	27	Jul 12, 202...
15	159451212...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600	Jul 12, 202...
16	159451212...	00:0f:00:70...	0.003	75.800	false	0.005	false	0.014	19.700	Jul 12, 202...
17	159451212...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600	Jul 12, 202...
18	159451212...	b8:27:eb:bf...	0.005	50.900	false	0.008	false	0.020	22.600	Jul 12, 202...

ExampleSet (405,184 examples: 0 special attributes, 10 regular attributes)

рис.6: Результирующий вид датасета после преобразований, генерация нового столбца *datetime*

Установка и использование Anomaly Detection Extension

- В RapidMiner откройте меню Extensions → Marketplace.
- Найдите Anomaly Detection.
- Установите и перезапустите RapidMiner.

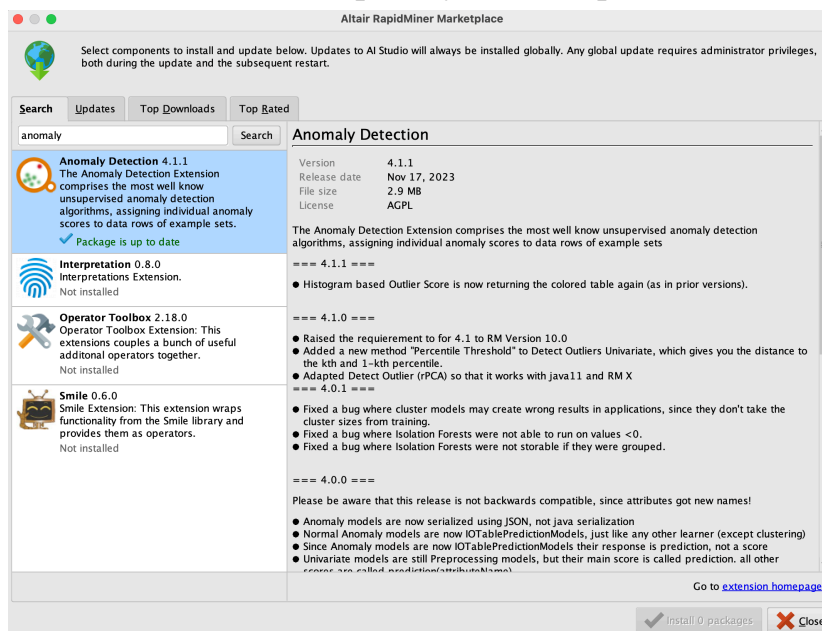


рис.7: Установка расширения из маркетплейса

Построение и настройка процесса детекции аномалий

Detect Outliers (Isolation Forest):

- Перетащите из панели Operators → Anomaly Detection → Detect Outliers (Isolation Forest).
- Подключите предобработанные данные (out) к входу (example set in) этого оператора.

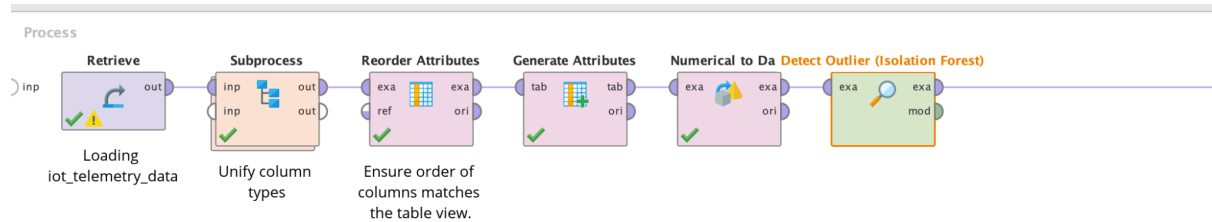


рис.8: Подключение оператора Detect Outliers (Isolation Forest)

Параметры Isolation Forest:

- Number of trees: 100 (Сколько деревьев формируется. Чем больше деревьев, тем надежнее результат (но выше время вычислений)).
- Max leaf size: 1 (Максимальный размер листа (количество наблюдений), при котором ветвление дерева перестаёт продолжаться. Чем меньше размер листа, тем сильнее детализированы деревья и тем более «чувствительным» может быть алгоритм к мелким аномалиям).
- bootstrap ratio: 0.9 (Доля выборки, которую мы используем для построения каждого дерева. Например, 0.9 означает, что каждый раз мы случайным образом берём 90% объектов из исходных данных для обучения дерева).
- use feature heuristic: Если включено (галочка), Isolation Forest при построении деревьев случайно выбирает подмножество признаков на каждом шаге (по аналогии с Random Forest), а не использует все

признаки одновременно. Это может помочь избежать переобучения, если признаков много, и повысить качество выявления аномалий.

- **score calculation:** Определяет, как считается «глубина» для вычисления финального outlier score. `average_path` означает, что итоговый score рассчитывается на основе средней длины пути для каждого дерева. При более коротком пути объект считается более аномальным.


Process	Parameters ×
 Detect Outlier (Isolation Forest)	
<i>number of trees</i>	100
<i>max leaf size</i>	1
<i>bootstrap ratio</i>	0.9
<input checked="" type="checkbox"/> <i>use feature heuristic</i>	
<i>score calculation</i>	average_path

рис.9: Настройка расширенных параметров оператора *Detect Outliers (Isolation Forest)*

Выходные порты:

- Оператор *Detect Outliers (Isolation Forest)* имеет выход `example set (labelled)`: добавляет столбцы `prediction`. Оно содержит величину аномальности (`Outlier Score`).

Оператор *Generate Outlier Flag*:

- Из того же набора (Anomaly Detection), добавьте оператор Generate Outlier Flag после Isolation Forest.
- Подключите out Isolation Forest → example set in Generate Outlier Flag.
- Score column: выберите то поле, которое вернул Isolation Forest (prediction).
- Method: contamination.
- Contamination threshold: 0.05. Это означает, что самый высокий 5% из prediction будет считаться «Yes».
- На выходе появится поле (например, outlier_flag) с двоичными значениями (true/false или Yes/No).

Добавьте перед этим оператор Normalize.

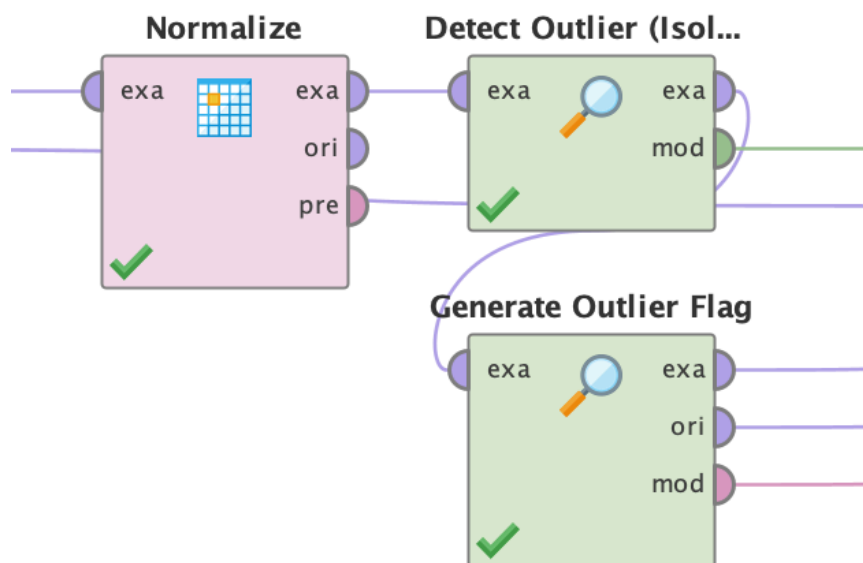


рис.10: Подключение оператора Generate Outlier Flag

Process

Parameters

×

Generate Outlier Flag

method

contamination

☒ define score column

score column

prediction

contamination threshold

0.05

рис.11: Настройка параметров оператора Generate Outlier Flag

Row No.	prediction	outlier_flag	ts	co	humidity	lpg	smoke	temp
1	0.485	No Outlier	-1.734	0.254	-0.837	0.286	0.281	0.091
2	0.572	No Outlier	-1.734	-1.439	1.363	-1.470	-1.466	-1.021
3	0.464	No Outlier	-1.734	0.270	-0.846	0.302	0.296	0.054
4	0.568	No Outlier	-1.734	-0.189	1.433	-0.148	-0.155	1.685
5	0.465	No Outlier	-1.734	0.263	-0.846	0.295	0.290	0.054
6	0.533	No Outlier	-1.734	-0.198	1.530	-0.158	-0.165	1.685
7	0.466	No Outlier	-1.734	0.270	-0.846	0.302	0.297	0.054
8	0.608	Outlier	-1.734	-1.361	1.363	-1.382	-1.379	-1.021
9	0.506	No Outlier	-1.734	-0.235	1.530	-0.194	-0.202	1.685
10	0.466	No Outlier	-1.734	0.265	-0.846	0.298	0.292	0.054
11	0.472	No Outlier	-1.734	0.257	-0.846	0.290	0.284	0.054
12	0.534	No Outlier	-1.734	-0.205	1.539	-0.164	-0.171	1.685
13	0.467	No Outlier	-1.734	0.266	-0.846	0.299	0.293	0.054
14	0.536	No Outlier	-1.734	-0.150	1.539	-0.109	-0.117	1.685
15	0.479	No Outlier	-1.734	0.261	-0.846	0.293	0.288	0.054
16	0.579	No Outlier	-1.734	-1.387	1.345	-1.412	-1.408	-1.021

ExampleSet (405,184 examples,2 special attributes,6 regular attributes)

рис.12: Результат запуска всего процесса и создание новых столбцов prediction и outlier_flag

Визуализация и анализ результатов

Линейный график (Line Plot):

- Выберите X-axis = datetime (ts), Y-axis = co, например.
- Color by = outlier_flag. Теперь точки, где outlier_flag=Outlier, выделятся другим цветом (желтый).
- На графике видно, что пики (очевидные выбросы), действительно определены как аномалии.

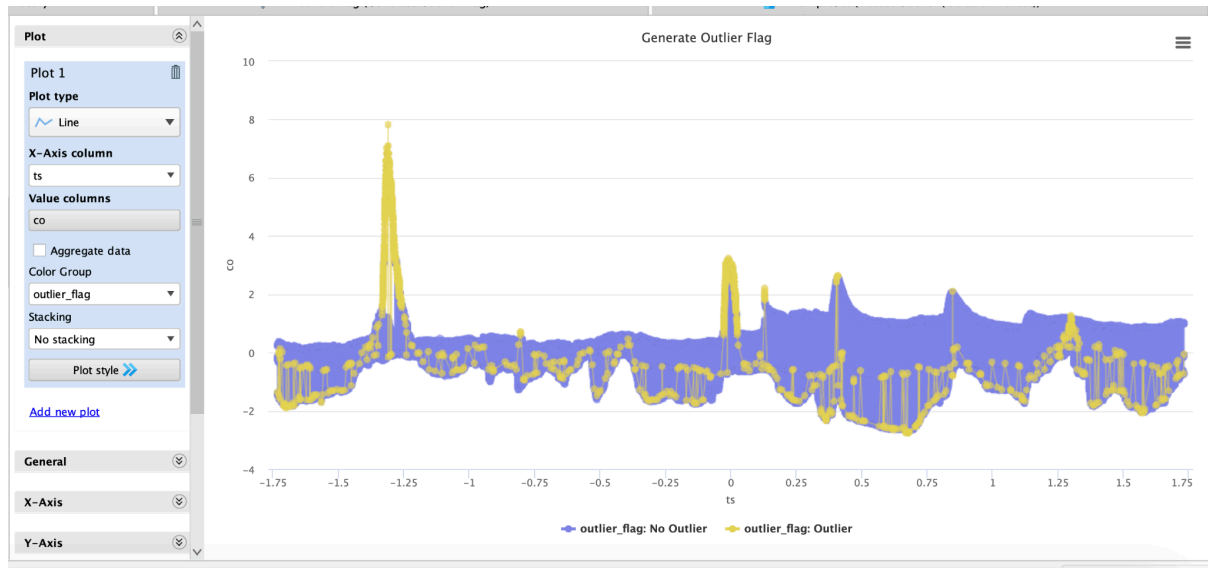


рис.13: Линейный график, раскрашенный на основе outlier_flag

Другой пример:

- X-axis = datetime (ts), Y-axis = temp.
- Color by = outlier_flag. Точки, где outlier_flag=Outlier, выделятся другим цветом (красный).
- На графике видно, что точки, не входящие в основную часть графика, выбивающиеся из распределения (очевидные выбросы), действительно определены как аномалии.

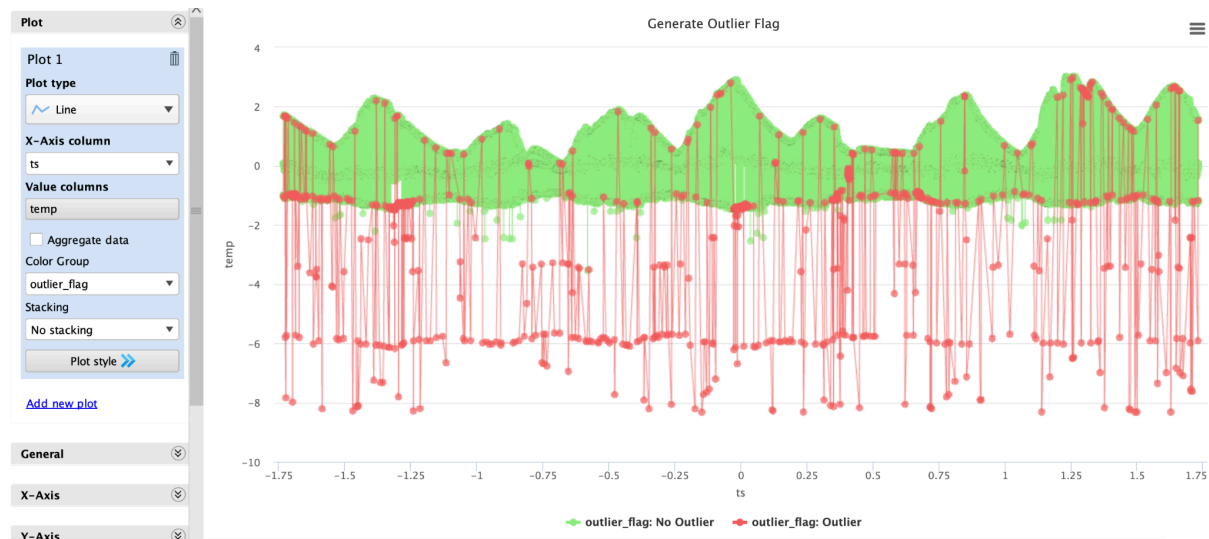


рис.14: Линейный график, раскрашенный на основе outlier_flag
 Диаграмма рассеяния (Scatter Plot):

- X-axis: temp, Y-axis: humidity, Color Group: prediction (anomaly).
- Посмотрите, какие комбинации temp + humidity (температура + влажность) становятся аномальными.

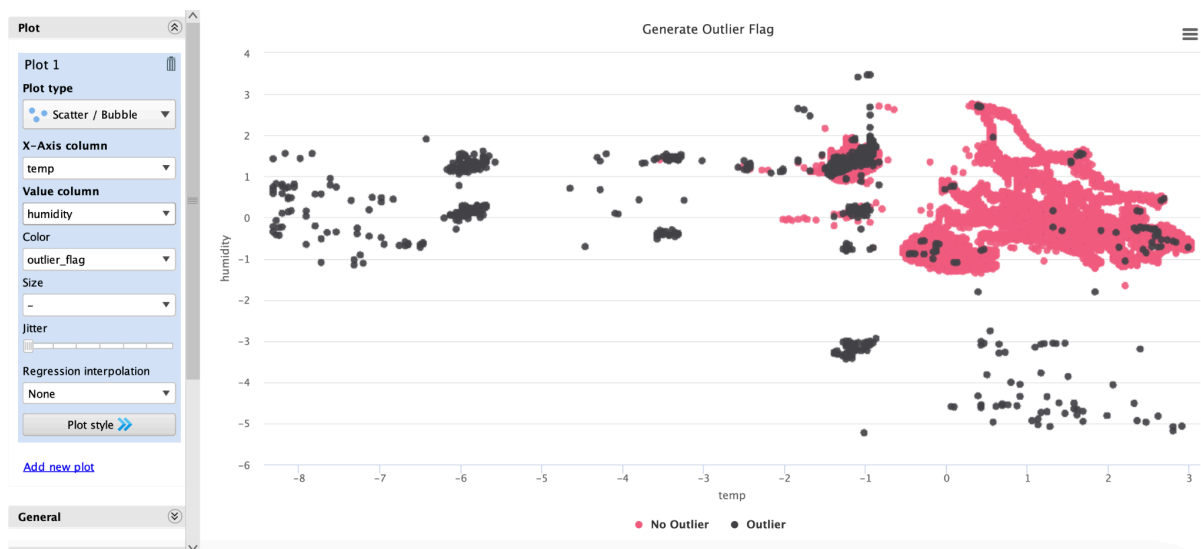


рис.15: Диаграмма рассеяния, раскрашенная на основе outlier_flag

Другой вариант:

- Пары признаков: «co и temp»
- X-axis: co.
- Y-axis: temp.
- Color by: outlier_flag.



рис.16: Диаграмма рассеяния, раскрашенная на основе outlier_flag

Вывод по аномалиям и интерпретация:

Видно, что большая часть данных (синие точки, "No Outlier") группируется в двух основных областях:

- Одна большая горизонтальная область наверху (где значения temp высокие, около 0, а co около 0 или немного отрицательный). Это основное «нормальное» состояние.
- Вторая, менее плотная область снизу, где temp сильно отрицательная, а co преимущественно небольшой и отрицательный.
- Некоторые аномалии явно находятся на краю облака основных значений (например, справа, где очень большие значения co). Такие аномалии легко объяснить: это экстремально высокие значения газа CO, что логично считать аномалией.
- Другие зелёные точки расположены внутри основных областей, это значит, что хотя по двум этим признакам (co, temp) точки выглядят нормально, они могут быть аномальными по другим параметрам (humidity, smoke, lpg, или булевым признакам).
- Горизонтальные линии точек (например, temp около 0) означают, что значения температуры в этот период были стабильны, но изменялось содержание CO. И наоборот, вертикальные «столбцы» точек означают стабильное CO, но колебания температуры. Возможно, аномалией была именно нестандартная стабильность (датчик мог "зависнуть").

Анализ boolean-флагов:

- light=FALSE и motion=TRUE может быть нетипичным состоянием ночью, если обычно нет движения без включенного света.
- Обратите внимание, появляются ли такие комбинации в аномальных строках.
- Для облегчения задачи можно создать новый столбец через Generate Attributes, объединяющий комбинации light_num и motion_num.

Expression

```
1 if(light_num==1 && motion_num==1, "L=1,M=1",  
2     if(light_num==1 && motion_num==0, "L=1,M=0",  
3     if(light_num==0 && motion_num==1, "L=0,M=1", "L=0,M=0"))
```

рис.16: Создание функции внутри оператора Generate Attributes для генерации столбца, объединяющего бинарные столбцы

- Создайте Bar Chart с группировкой по light/motion и outlier_flag.

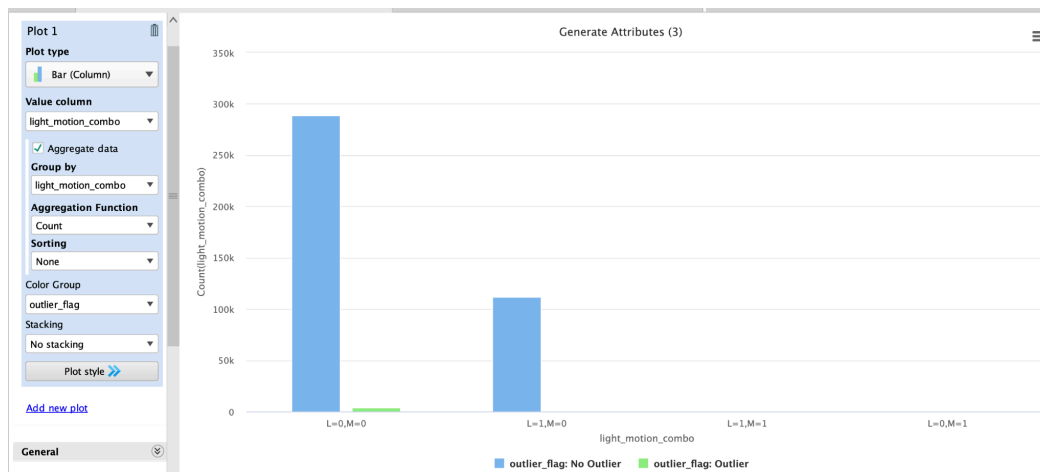


рис.16: Столбчатая диаграмма для совмещенного из бинарных переменных столбца, раскрашенная с помощью outlier_flag

График представляет распределение количества примеров по четырем комбинациям признаков:

- L=0,M=0 (свет выключен, движения нет)

- $L=1, M=0$ (свет включен, движения нет)
- $L=0, M=1$ (свет выключен, движение есть)
- $L=1, M=1$ (свет включен, движение есть)
- Большинство измерений относятся к категории $L=0, M=0$ (выключенный свет, нет движения), это основное состояние.
- Значительное количество наблюдений относится к категории $L=1, M=0$ (свет включен, нет движения).
- Комбинации, связанные с движением ($M=1$), почти не встречаются, либо вообще не представлены.
- Это говорит о том, что датчики движения срабатывают крайне редко (или вообще не активировались за всё время наблюдений).

Вывод по аномалиям:

- Аномалии (зеленый цвет) в основном сосредоточены в категории $L=0, M=0$ (свет выключен, движения нет). Это значит, что большинство аномалий возникают, когда датчики находятся в состоянии покоя (ночью или когда никого нет).

Распределения признаков в аномальных и нормальных:

- Необходимо построить диаграмму (Histogram).
- Value column: со (например).
- Color: outlier_flag.
- График покажет, в каких диапазонах со (или другого сенсора) сконцентрировано большинство аномалий.

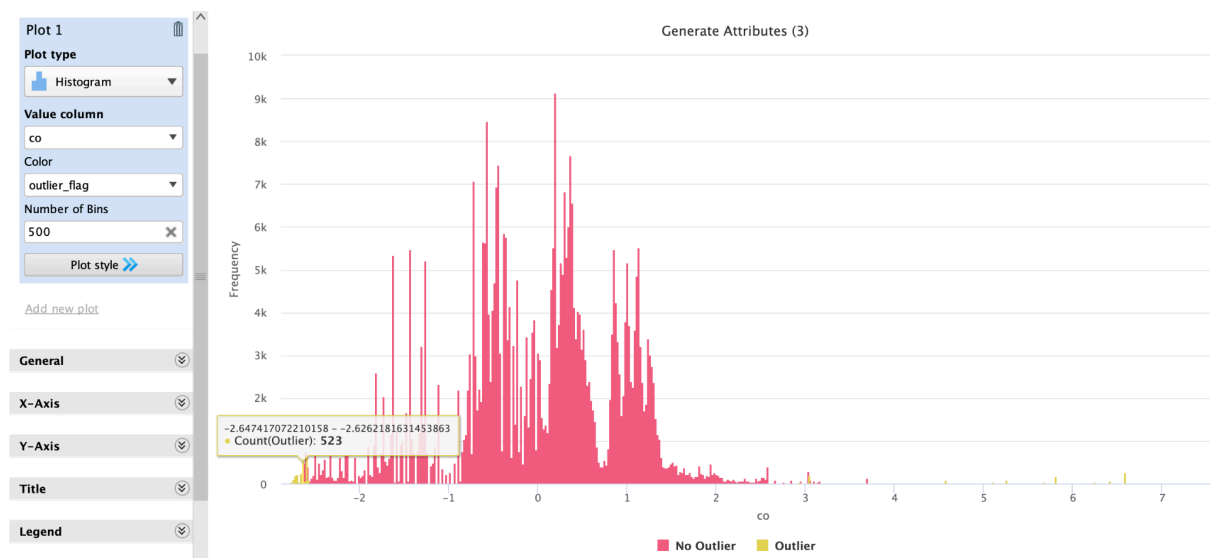


рис.17: Гистограмма для просмотра распределения признаков в аномальных и нормальных случаях

Аномалии (желтые столбцы) в основном сконцентрированы в крайних значениях распределения:

- Положительные экстремумы (значения $CO > 3$, после нормализации) означают значения значительно выше среднего уровня содержания газа в воздухе.
- Отрицательные экстремумы (значения $CO < -2$ после нормализации) означают существенно более низкое содержание CO по сравнению со средним уровнем.
- Высокие значения (> 3) соответствуют значительному превышению типичного уровня содержания газа CO. Это потенциально опасные или нестандартные ситуации, требующие проверки (возможно, короткие резкие всплески газа, связанные с утечками).
- Низкие значения (< -2) указывают на низкий уровень CO относительно средней концентрации. Это может быть нормальной ситуацией, но всё равно заслуживает внимания, поскольку значительно отличается от типичного поведения сенсора. Возможно, в такие моменты в помещении происходила интенсивная вентиляция или другие специфические условия, резко снижающие концентрацию газа.
- Положительные экстремальные аномалии потенциально свидетельствуют об утечках или всплесках уровня газа и требуют оперативного внимания.

- Отрицательные экстремальные аномалии говорят о периодах с необычно низкой концентрацией CO (значительно ниже среднего), что также интересно с точки зрения анализа условий среды (например, проветривания или изменения режима работы помещения).

Пример выполненной работы

- Шаг 1: Оператор Read CSV использован для загрузки датасета формата .csv и определения типов столбцов. Replace Missing Values использован для обработки пропущенных значений (при наличии).
- Шаг 2: Добавлен оператор Detect Outliers (Isolation Forest). На выходе получен столбец prediction (числовой outlier score).
- Шаг 3: Добавлен Generate Outlier Flag → появился outlier_flag.
- Шаг 4: Построен Scatter Plot, построен Bar Chart, построен линейный график с выбранным параметром Color – outlier_flag. Каждый график интерпретирован и сделаны соответствующие выводы, относящиеся к аномалиям.

Приобретенные навыки

- Загрузка и предобработка IoT-данных в RapidMiner (epoch → datetime, конверсия boolean в 0/1).
- Использование Anomaly Detection Extension (Isolation Forest + Generate Outlier Flag).
- Нормализация признаков, настройка contamination и прочих параметров алгоритма.
- Визуализация результатов (Line Plot, Scatter Plot, Histogram, Bar Chart) и интерпретация аномалий.
- Поиск закономерностей (роль boolean-флагов, разброс по датчикам, экстремальные показатели газа и т.д.) в контексте реального IoT-процесса.

Обобщенная задача для выполнения индивидуального варианта

Задача: разработайте универсальный процесс обнаружения аномалий в IoT-данных, применимый к различным наборам данных с телеметрией. В задачу входят следующие этапы:

1. Загрузка и предобработка данных:
 - Импортируйте набор данных, содержащий временные ряды измерений с IoT-сенсоров (например, показатели температуры, влажности, концентрации газов и т.д.).
 - Проведите очистку данных: обработку пропущенных значений, преобразование временных меток из формата epoch в datetime, а также приведение типов признаков к корректным форматам (Real, Nominal).
2. Обнаружение аномалий:
 - Реализуйте алгоритм обнаружения аномалий с помощью Isolation Forest.
 - Настройте параметры алгоритма (количество деревьев, размер листа, долю выборки и т.д.) для достижения оптимальной чувствительности к аномалиям.
 - Создайте дополнительный атрибут, который будет бинарно маркировать наблюдения как нормальные или аномальные на основе вычисленного outlier score.
3. Визуализация и анализ результатов:
 - Постройте различные графики (линейные графики, диаграммы рассеяния, гистограммы, столбчатые диаграммы) для наглядного представления распределения данных и обнаруженных аномалий.
 - Проведите интерпретацию визуализированных результатов, выделив типичные паттерны и потенциально опасные или нештатные ситуации в работе IoT-устройств.
4. Интерпретация и выводы:
 - Проанализируйте, какие показатели и их комбинации приводят к обнаружению аномалий.
 - Обоснуйте возможные причины возникновения аномальных значений и предложите пути их дальнейшей проверки или корректировки системы мониторинга.

Распределение вариантов

