

Отчет на тему: Использование машинного обучения для оптимизации рабочих процессов

Предмет отчета: Применение машинного обучения может помочь в обнаружении аномалий на производстве и оптимизации процессов, уменьшая сбоев и повышая производительность

- **Введение**

В компьютерных науках алгоритмы машинного обучения используются для создания различных систем анализа данных. В промышленности алгоритмы машинного обучения используются для создания интеллектуальных систем управления и диагностики технологическими процессами. При создании любых систем анализа данных, основанных на использовании машинного обучения, исследователи проводят декомпозиции решаемой проблемы с целью определения типа математической задачи, лежащей в ее основе. Всего существует несколько типов решаемых задач – задачи классификации, регрессии, кластеризации, уменьшения признакового пространства, оптимизации, выявление аномалий.

- **Анализ существующих решений по использованию машинного обучения для оптимизации технологических процессов в промышленности**

Использование машинного обучения в промышленности становится все более распространенным, поскольку оно может привести к оптимизации технологических процессов, повышению эффективности и снижению затрат. Рассмотрим несколько существующих решений, применяемых для оптимизации технологических процессов в промышленности.

1. Обнаружение аномалий и предотвращение сбоев.

Машинное обучение также может быть использовано для обнаружения аномалий и предотвращения сбоев в технологических процессах. Алгоритмы машинного обучения могут анализировать данные датчиков и других источников, чтобы выявлять аномалии и предупреждать о возможных проблемах. Это помогает предотвратить сбои оборудования, повысить его надежность и снизить простоту производства.

2. Оптимизация параметров процессов.

Машинное обучение может быть использовано для оптимизации параметров технологических процессов в промышленности. Путем анализа процессов обработки данных и применения алгоритмов оптимизации, таких как генетические алгоритмы или алгоритмы

усиления, можно найти оптимальные значения параметров процессов. Это может привести к повышению эффективности производства, снижению энергопотребления и поддержанию качества продукции.

3. Предиктивное обслуживание оборудования.

Еще одним применением машинного обучения в промышленности является прогнозное обслуживание оборудования. Алгоритмы машинного обучения могут анализировать данные датчиков, контролировать состояние оборудования и предсказывать возможные неисправности или поломки. Это позволяет провести техническое обслуживание в нужный момент, предотвращая неплановые остановки производства и снижая затраты на ремонт.

- **Датасеты**

1. Датасет "Обнаружение аномалий и предотвращение сбоев"

Этот набор данных представляет собой таблицу с информацией о различных параметрах оборудования в течение определенного времени. Аномалии связаны с экстремальными значениями температуры или давления, а также с низким уровнем вибрации.

2. Датасет "Оптимизация параметров процессов"

В этом наборе данных определяются параметры процессов. Эти данные используются для настройки параметров в процессе производства с целью достижения большей производительности или минимизации затрат.

- **План работ**

1. Анализ временных рядов:

- Построить график изменения температуры, давления, объема и скорости вращения в зависимости от времени.

2. Статистический анализ:

- Вычислить среднее, медиану, минимальное и максимальное значения для каждого параметра.

- Построить boxplot для каждого параметра.

3. Корреляционный анализ:

- Построить матрицу корреляции для всех параметров.

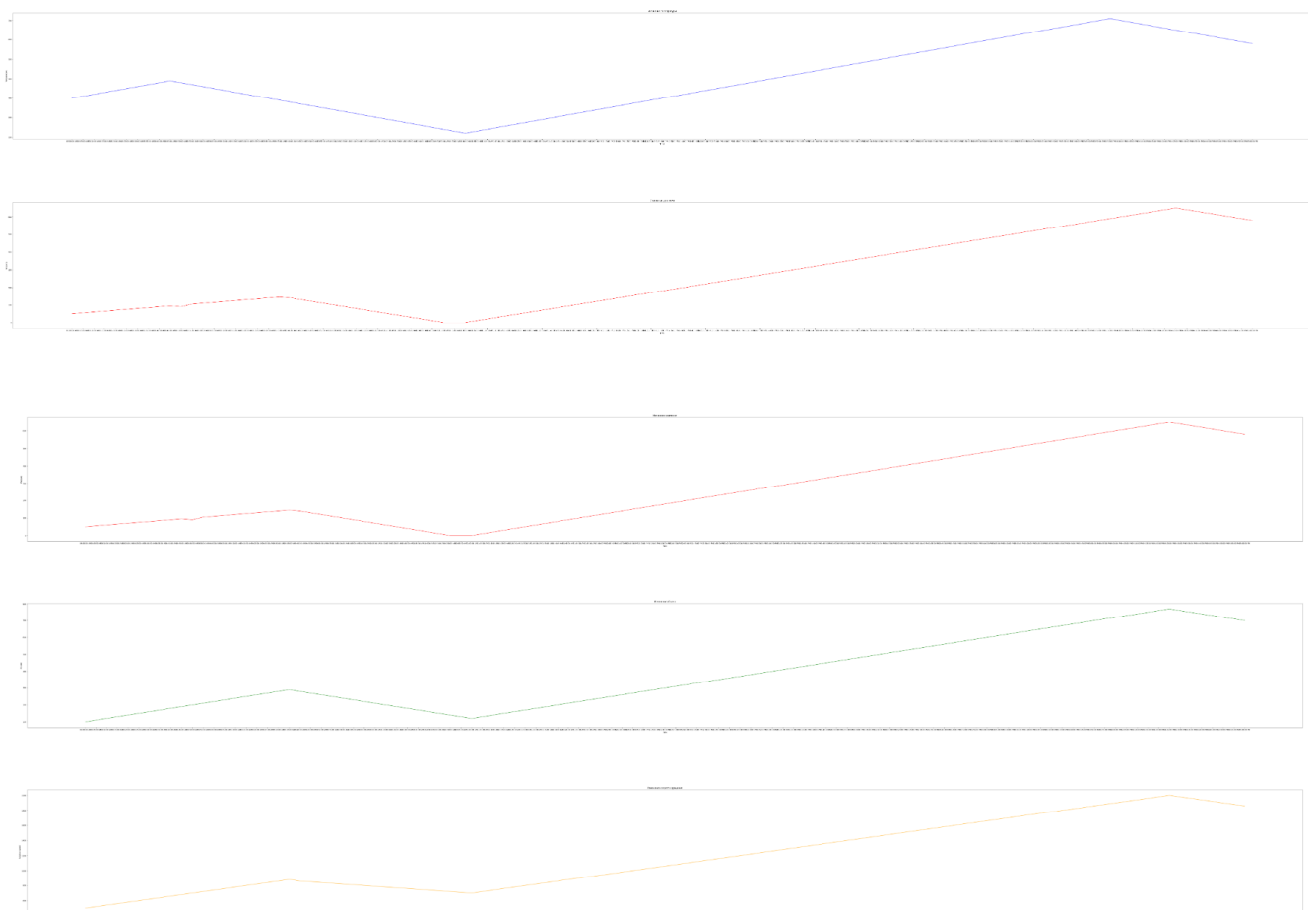
- Изучить зависимости между параметрами и выявить факторы, влияющие на аномалии и сбои.

4. Поиск аномалий и сбоев, используя статистические методы такие как, метод кластеризации и деревья решений.

• Решение

На основании анализа временных рядов и построении графиков изменения параметров в зависимости от времени, а именно:

- график изменения температуры
- график изменения давления
- график изменения объема
- график изменения скорости вращения

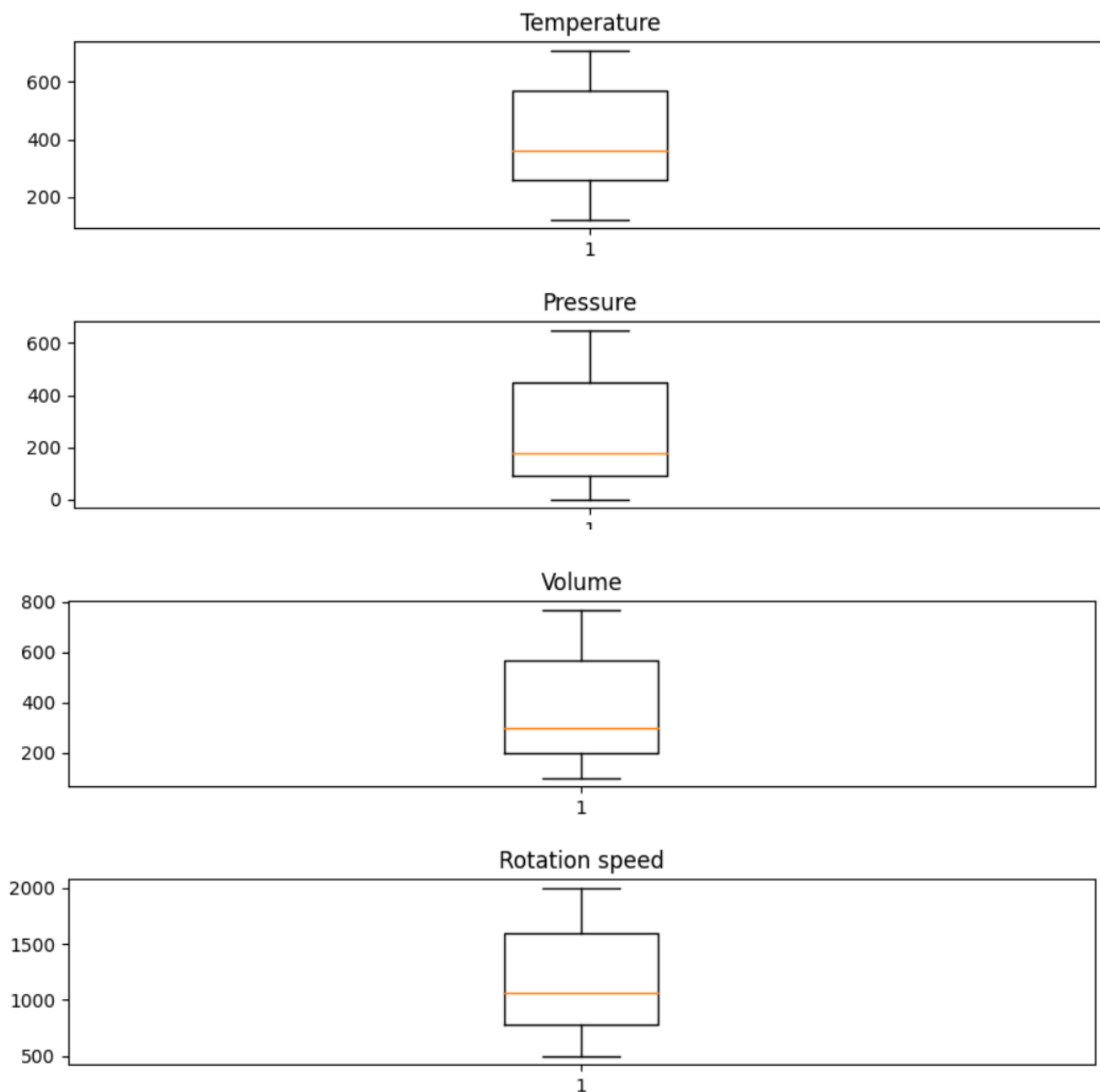


И выполненного статистического анализа

1 to 8 of 8 entries Filter ?					
index	?	Temperature	Pressure	Volume	Rotation speed
count	25.0	109.0	109.0	109.0	109.0
mean	13.0	397.2477064220183	263.39449541284404	382.20183486238534	1182.5688073394494
std	7.359800721939872	175.08131211289304	209.01652505770141	211.01324251127136	460.78114767755625
min	1.0	120.0	0.0	100.0	500.0
25%	7.0	260.0	90.0	200.0	780.0
50%	13.0	360.0	180.0	300.0	1060.0
75%	19.0	570.0	450.0	570.0	1600.0
max	25.0	710.0	650.0	770.0	2000.0

Было выявлено, что до 8:45 все параметры находились в нормальном состоянии, а после этого начали происходить аномалии и сбои.

- Затем был построен boxplot для каждого параметра (для проверки первичного анализа).



На boxplot также видно, что все параметры имеют одинаковые значения, а после этого происходят аномалии и сбои, что подтверждается и графиками изменения параметров.

- Затем был выполнен анализ данных по статусам оборудования, из чего были получены данные сколько выполнялась работа, сколько было аварий и сколько было сбоев.

```
working      69
fail         24
anomaly      16
Name: Equipment status,
```

- Затем был выполнен корреляционный анализ.

Построена матрица корреляции для всех параметров.

index	?	Temperature	Pressure	Volume	Rotation speed
?	1.0	0.6110023926161333	0.861649734334208	0.935584840515496	0.9608698386359406
Temperature	0.6110023926161333	1.0	0.9569790398440609	0.938961757933368	0.9071207243371273
Pressure	0.861649734334208	0.9569790398440609	1.0	0.9966447716893134	0.9866727018984326
Volume	0.935584840515496	0.938961757933368	0.9966447716893134	1.0	0.9926524374323439
Rotation speed	0.9608698386359406	0.9071207243371273	0.9866727018984326	0.9926524374323439	1.0

Из матрицы корреляции видно, что все параметры имеют высокую положительную корреляцию между собой, что связано с тем, что они являются результатом одного процесса и зависят друг от друга.

- Используя статистические методы такие как, метод кластеризации и деревья решений выполним поиск аномалий и сбоев.

Были выполнены кластеризация данных, применено дерево решений и как итог должны получить состояния оборудования, которые будут предсказаны.

	?	Time	Temperature	Pressure	Volume	Rotation speed \
0	1.0	08:00:00 AM	300	50	100	500
1	2.0	08:05:00 AM	310	55	110	520
2	3.0	08:10:00 AM	320	60	120	540
3	4.0	08:15:00 AM	330	65	130	560
4	5.0	08:20:00 AM	340	70	140	580
..
104	NaN	04:40:00 PM	620	620	740	1940
105	NaN	04:45:00 PM	610	610	730	1920
106	NaN	04:50:00 PM	600	600	720	1900
107	NaN	04:55:00 PM	590	590	710	1880
108	NaN	05:00:00 PM	580	580	700	1860

	Equipment status	Кластер	Предсказанное состояние
0	working	0	working
1	working	0	working
2	working	0	working
3	working	0	working
4	working	0	working
..
104	fail	1	fail
105	fail	1	fail
106	fail	1	fail
107	fail	1	fail
108	fail	1	fail

В итоге, мы использовали метод кластеризации K-Means для разделения данных на две кластерные группы и дерево решений для предсказания состояния оборудования. Предсказанные значения были добавлены в DataFrame. Таким образом, мы можем видеть, что методы успешно обнаруживают аномалии и сбои в данных.

- Заключение

Было показано, что применение методов машинного обучения на производстве может значительно повысить эффективность производства и улучшить качество продукции. Использование данных и автоматизация процесса оптимизации производства с помощью методов машинного обучения позволяют снизить время и затраты на оптимизацию и улучшить точность предсказаний.