ANOMALY DETECTION: ENERGY STORAGE SYSTEMS SAFETY

March 2022

Project Team:

Syrvachev Sergey

1	2	3
INTRODUCTION	RESEARCH	FINALIZE

- Business Problem
- Data Understanding
- Metrics

- Raw Data
- Modeling
- Modeling Results

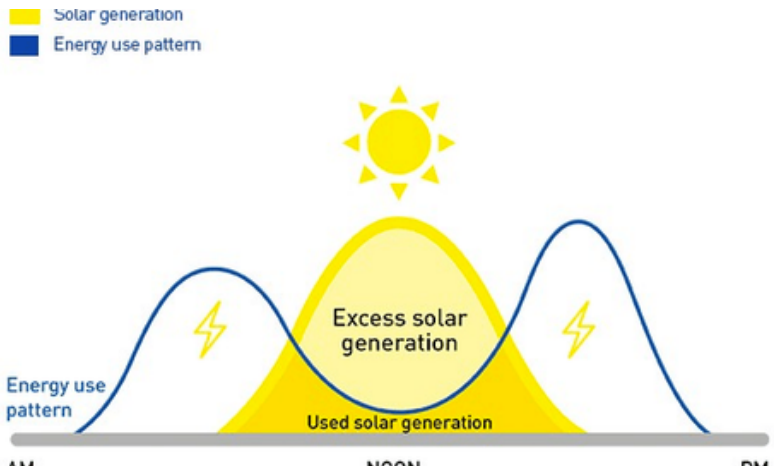
- Business recommendations
- Next Steps

INTRODUCTION

- Business problem
- Key Idea
- Data sources & Methods

BUSINESS PROBLEM

- The rapid growth of renewable sources of energy created **new challenges for energy generation**
- **Unstable** energy production is one of them
- **Energy Storage Systems are the key element** to overcome it

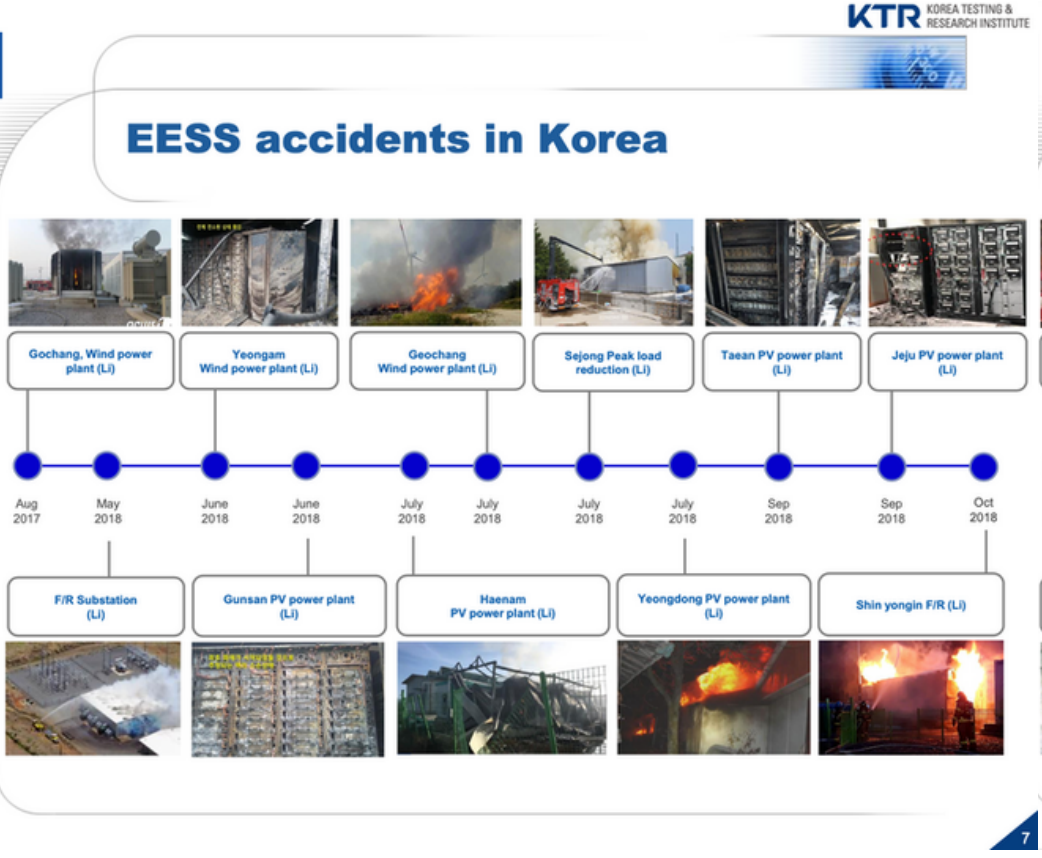


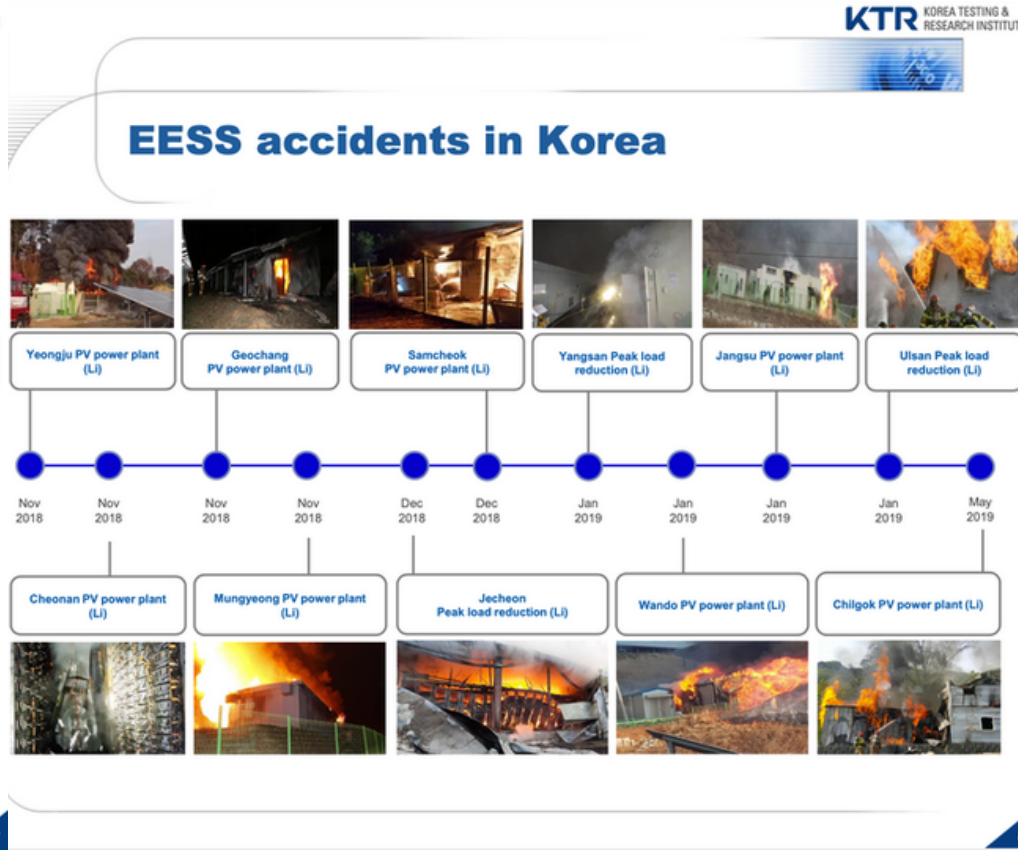
BUSINESS PROBLEM

- The energy storage system consists of dozens of thousands of cells.
- Failure of one cell can lead to catastrophic results and a hazardous situation on site.
- On average, every month, only in Korea two sites is completely destroyed by electrical fire (2018 year)

Source:
[1]Results of Investigation on EESS Fire Accident in Korea
IEC TC 120 WG4 Convenor Misung Kim, 2018

BUSINESS PROBLEM





Source:
[1]Results of Investigation on EESS Fire Accident in Korea
IEC TC 120 WG4 Convenor Misung Kim, 2018

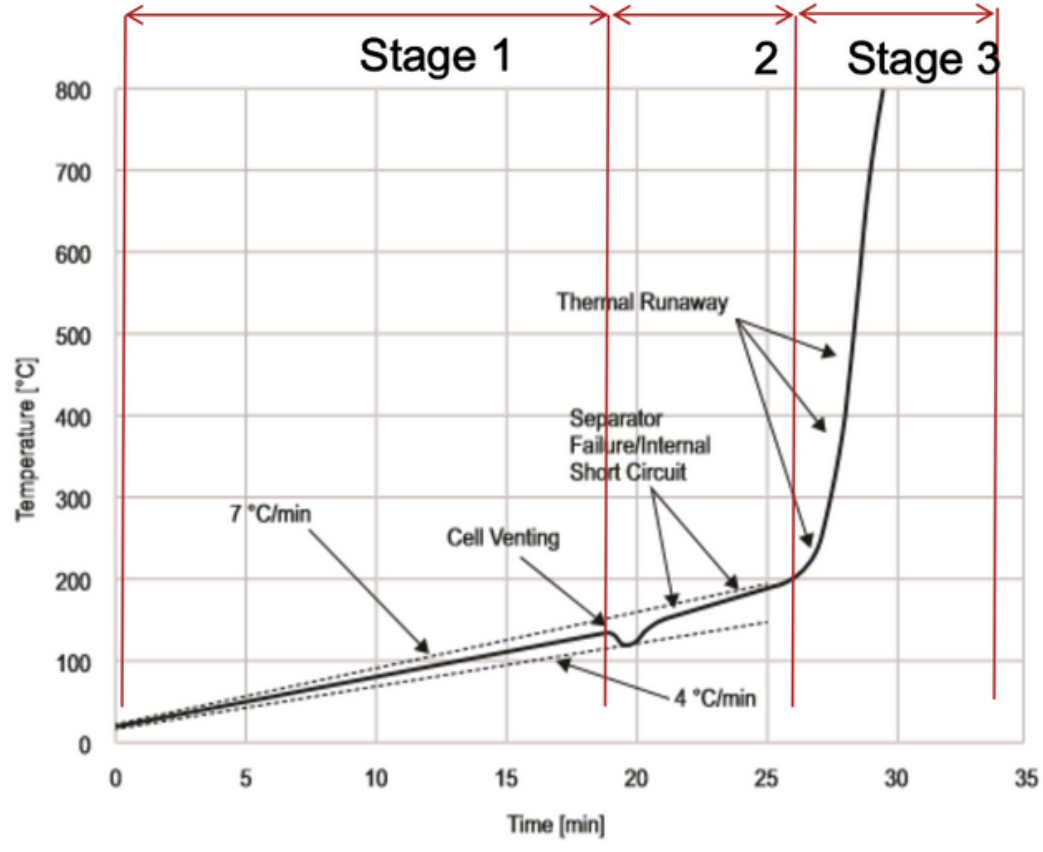
KEY IDEA BEHIND

- Main factors:

- Electrical over-stress
- Thermal over-stress
- Cell internal fault

- Possible to monitor:

- Voltage
- Temperature
- Current



DATA UNDERSTANDING

METHODS:

1. The time series focused on finding cells that behave abnormal compare to other cells.
2. Focus on voltage and temperature behavior.

DATA SOURCES:

1. Data collected from one of the sites during commissioning works and artificially modified for learning purpuses

METRICS

We are focused on cells with abnormal characteristics.

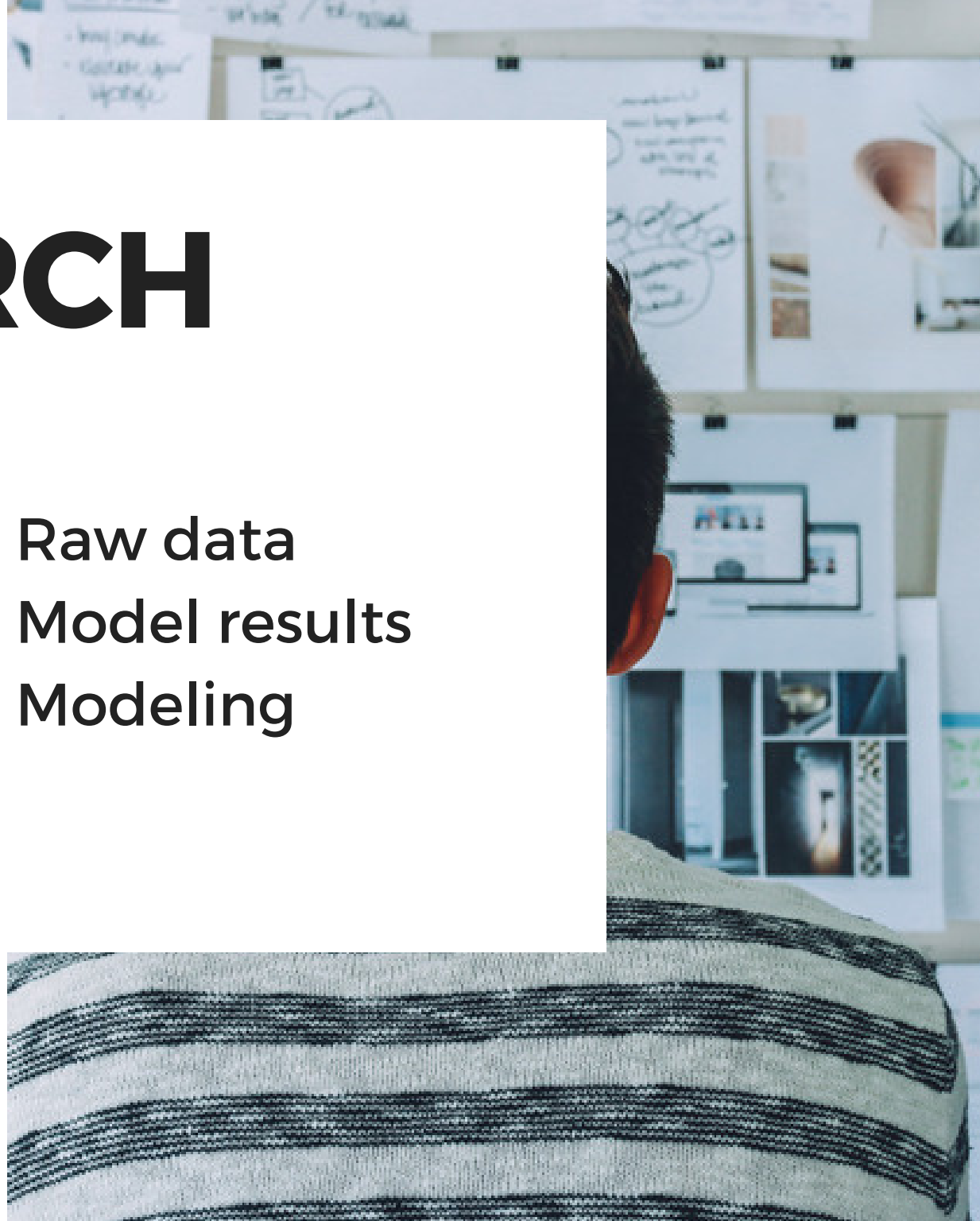
- **Hypothesis:**
  - H0 - Cell is in good condition
  - HA - There is statistically significant proof that the cell has abnormal behavior

USED METRICS:

- **Recall** - Safety is our priority, we will be focused on finding all possible cells that have anomalies
- **Accuracy** - How accurate our results are considering false identified cases.

RESEARCH

- Raw data
- Model results
- Modeling



RAW DATA

THE SYSTEM CONSISTS OF 5 CLUSTERS:

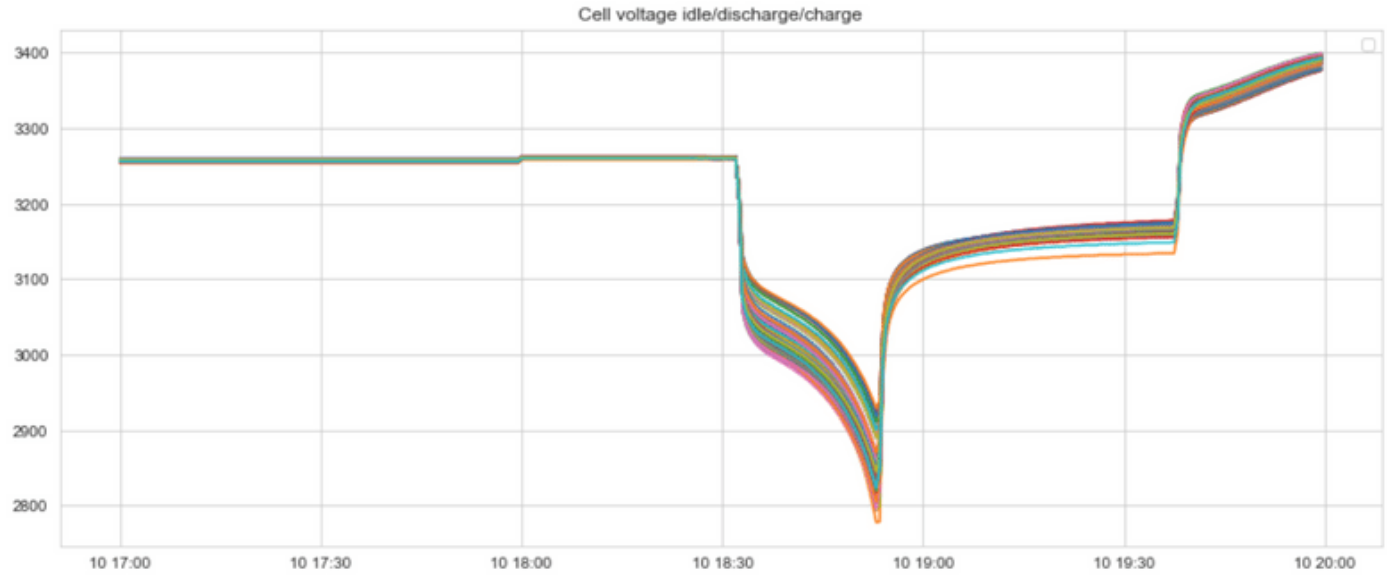
EACH CLUSTER:

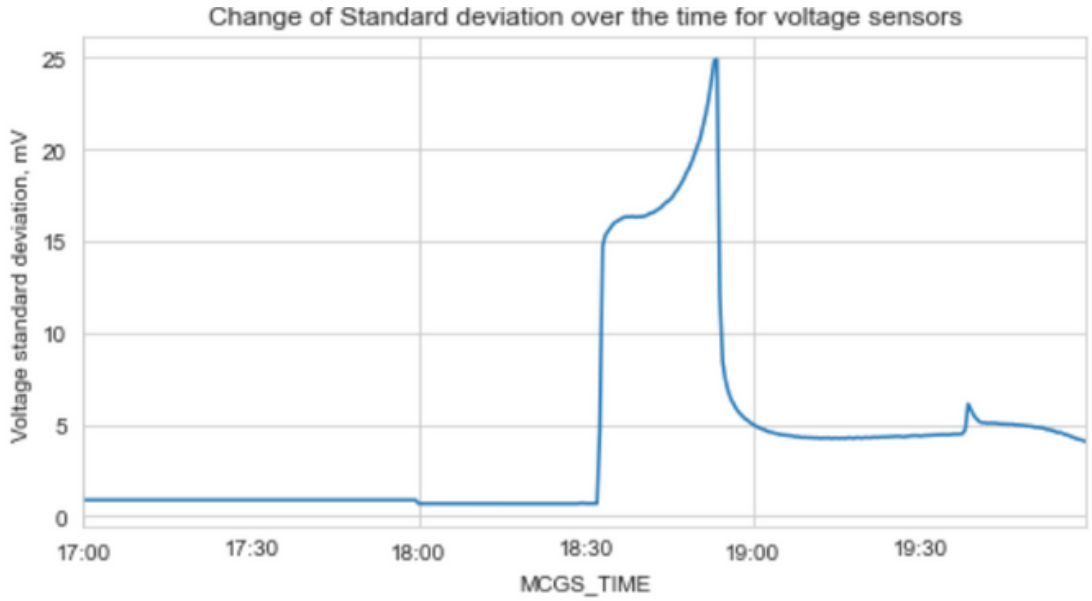
- 240 TEMPERATURE SENSORS
- 240 VOLTAGE SENSORS
- 1 CURRENT SENOR
- DATA READINGS: EVERY 10 SEC

OVERALL:
> 2400 SENSORS WITH > 10 000 READINGS FOR EACH SENSOR

SOME DATA WAS ARTIFICIALLY MODIFIED TO INCREASE THE NUMBER OF ANOMALY SENSORS

VOLTAGE/TEMPERATURE

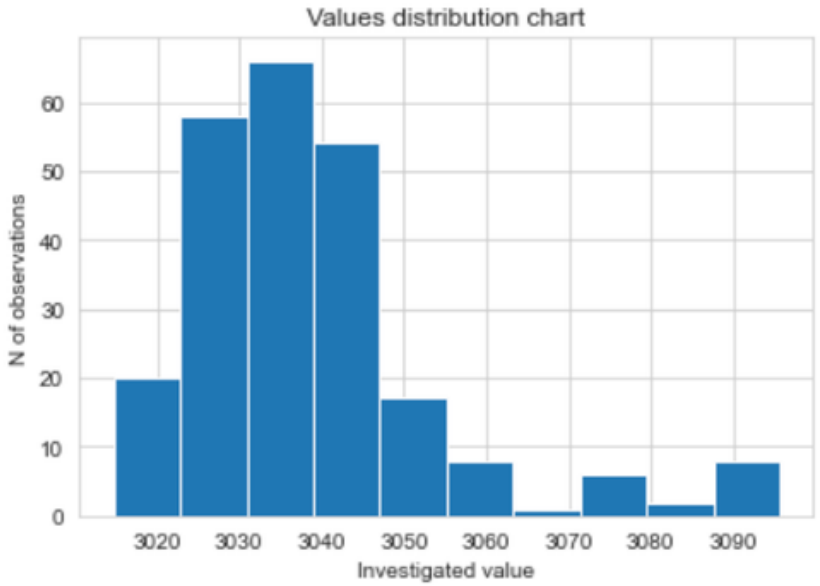




KEY FINDINGS:
DURING CHARGE/DISCHARGE VOLTAGE DIFFERENCE

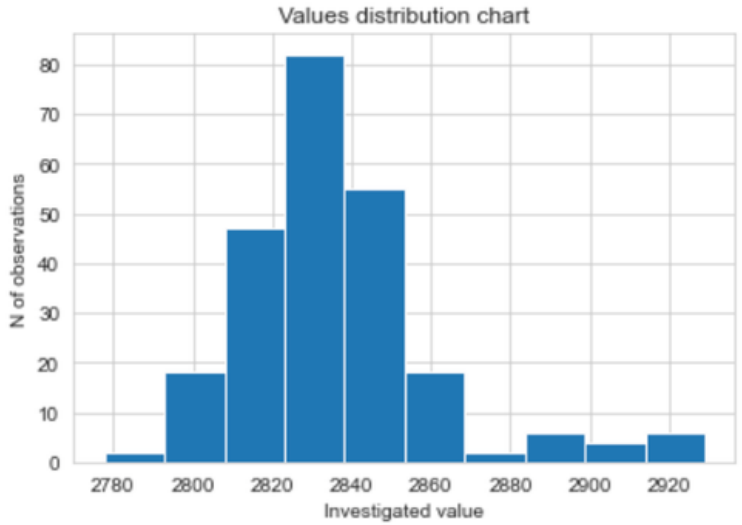
VOLTAGE/TEMPERATURE

Start of discharge

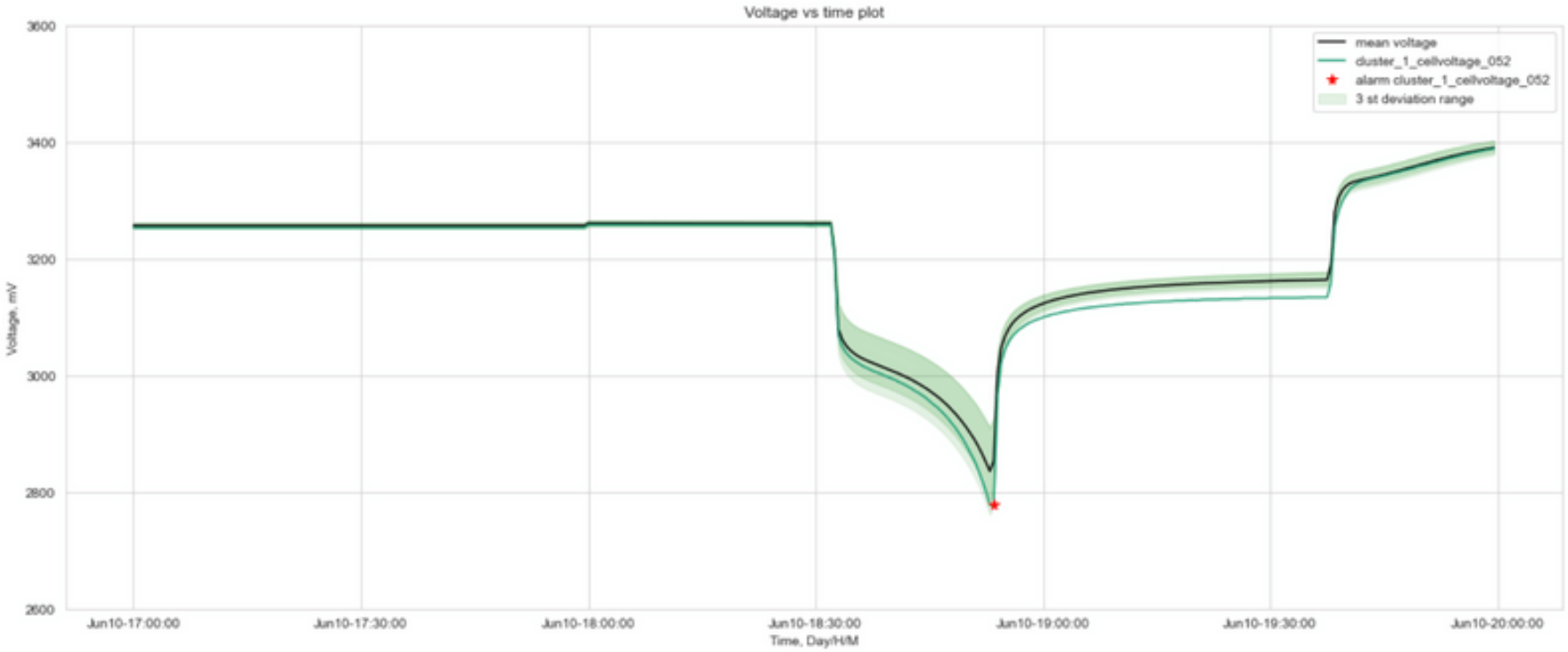


End of discharge

```
Value mean : 2.84 V
Value standard deviation : 0.0249 V
```



KEY FINDINGS:
NORMAL DISTRIBUTION OF VOLTAGE SENSOR READINGS WITHIN CLUSTER



cluster N 1, discharge anomaly, cellvoltage sensor N 052

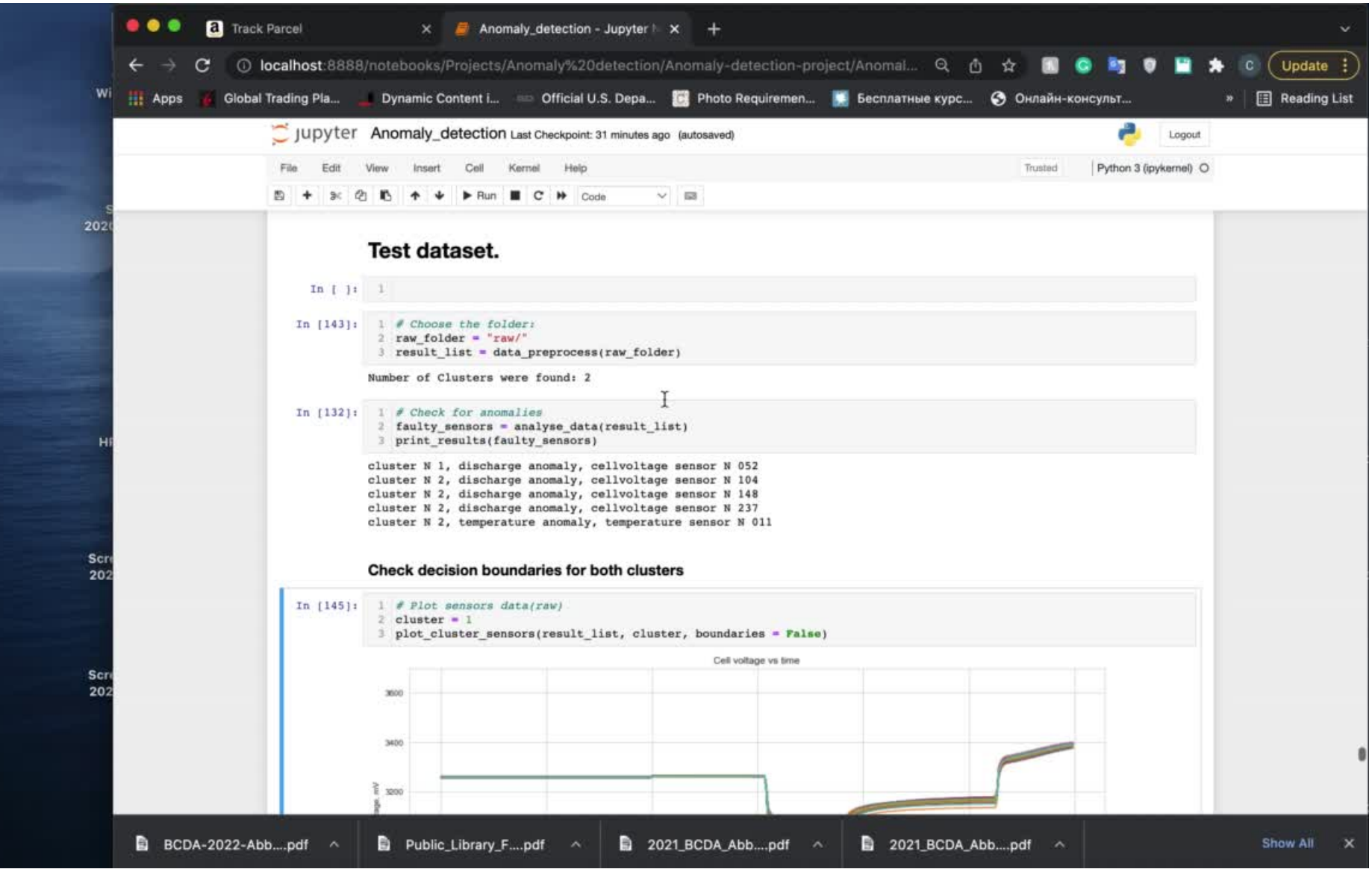
3 WAYS OF RESULT REPRESENTATION:

- THE PLOT OF ANOMALY SENSORS VALUES OVER TIME
- MARKER WHERE THE THRESHOLD WAS BREACHED
- LIST OF ANOMALY SENSORS WITH THE TYPE OF ALARMS

MODEL RESULTS

BEST MODEL AFTER TUNING:

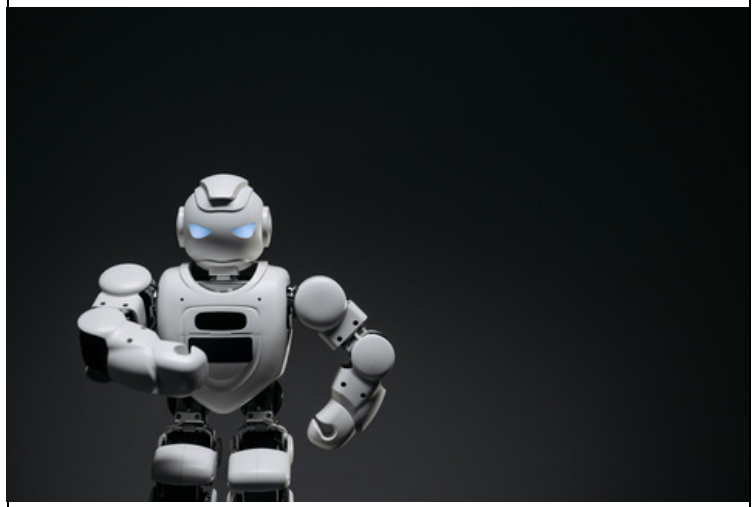
- CUSTOM MODEL BASED ON ASSUMPTIONS
  - 100 % RECALL
  - 99.5 % ACCURACY


Track Parcel
Anomaly_detection - Jupyter
localhost:8888/notebooks/Projects/Anomaly%20detection/Anomaly-detection-project/Anomal...
Update
Apps
Global Trading Pla...
Dynamic Content i...
Official U.S. Depa...
Photo Requiremen...
Бесплатные курс...
Онлайн-консульт...
Reading List
jupyter Anomaly_detection Last Checkpoint: 31 minutes ago (autosaved)
Logout
File Edit View Insert Cell Kernel Help
Trusted
Python 3 (ipykernel)
Run
Code
Test dataset.
In [ ]:
In [143]:
# Choose the folder:
raw_folder = "raw/"
result_list = data_preprocess(raw_folder)
Number of Clusters were found: 2
In [132]:
# Check for anomalies
faulty_sensors = analyse_data(result_list)
print_results(faulty_sensors)
cluster N 1, discharge anomaly, cellvoltage sensor N 052
cluster N 2, discharge anomaly, cellvoltage sensor N 104
cluster N 2, discharge anomaly, cellvoltage sensor N 148
cluster N 2, discharge anomaly, cellvoltage sensor N 237
cluster N 2, temperature anomaly, temperature sensor N 011
Check decision boundaries for both clusters
In [145]:
# Plot sensors data(raw)
cluster = 1
plot_cluster_sensors(result_list, cluster, boundaries = False)
Cell voltage vs time
3600
3400
3200
BCDA-2022-Abb....pdf
Public_Library_F....pdf
2021_BCDA_Abb....pdf
2021_BCDA_Abb....pdf
Show All

FINALIZE

- Business recommendations
- Next Steps
- QA

Recommendations



01

MODEL OPTIMIZATION

Model can be adjusted to different BMS manufacturers

02

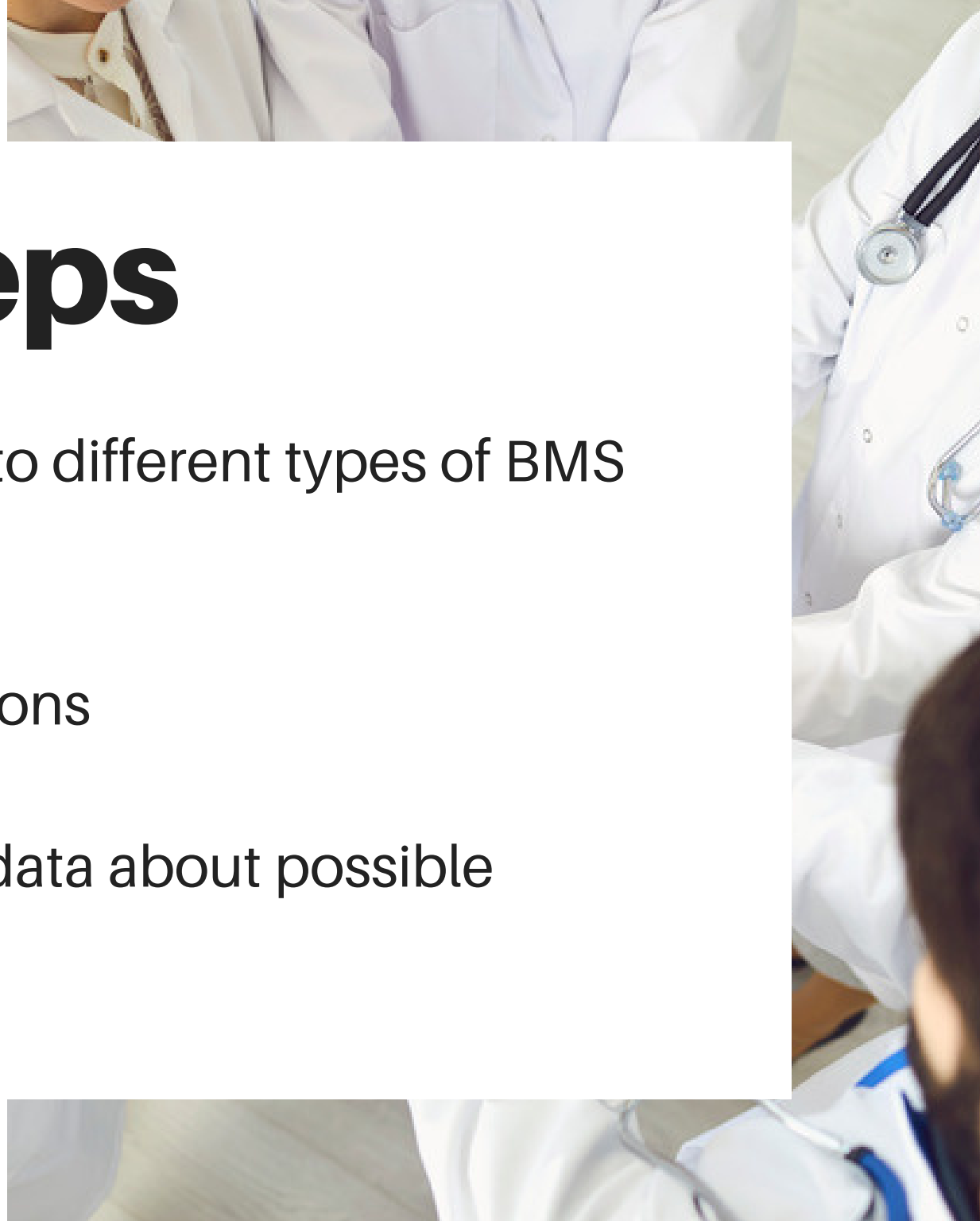
USE OF MODEL

Model should be used in the field for last optimization

03

INFORMATION INPUTS

The model can be extended for real-time monitoring



Next steps

- Extend the model to different types of BMS manufacturers
- Real-time applications
- Gather additional data about possible anomalies

Q & A:

Thank you for joining today's presentation.

SYRVACHEV SERGEY

- DATA SCIENTIST
- DEST.STUDIO@GMAIL.COM
- LINKEDIN: /SSYRVACHEV
- GITHUB: 314KA4Y
- MEDIUM: @SERGEYSYRVACHEV