# Introduction

The purpose of this document is as follows:

* Define the flow of the algorithm
* Record the experiments and done along with their results

## Flow of the algorithm

1. We are taking 5 minutes data at a time, 1 minute of previous run
2. First check if all data is present.
   1. If no data, process terminates
   2. If data is missing (about 1 – 2 minutes): Linear interpolation. Model will characterize it as a no event. ML model will not be run on it because observable event

# Experiments

## Experiment 1

* Data from dates between 3rd and 6th August was used
* The dataset had this distribution
  + No Event: 2886104
  + Refrigerator on: 317
  + Refrigerator off: 414
* Features used were: ['mains\_power','prev\_mean','time\_since\_last\_event','time\_since\_event\_start', 'normalized\_mains\_power']
* We used 100 epochs of batch gradient descent
* Results:
  + - F1 Score: 0.968880516008464
    - Accuracy: 0.9928409947249435
    - Confusion Matrix

[[ 102 0 8]

[ 3 136 6]

[ 0 2 2397]]

* Inference:

For label 0

tp: 102 fp: 3 fn: 8 tn: 2541

For label 1

tp: 136 fp: 2 fn: 9 tn: 2507

For label 2

tp: 2397 fp: 14 fn: 2 tn: 241

## Experiments 2

* Data from dates between 6th and 9th August was used
* The dataset had this distribution (24519 used for training. Most No event instances were removed)
  + No Event: 5000276
  + Geyser off: 360
  + Geyser on: 460
  + Refrigerator off: 604
  + Refrigerator On: 805
* Features used were: ['mains\_power','prev\_mean','time\_since\_last\_event','time\_since\_event\_start', 'normalized\_mains\_power']
* We used 100 epochs of batch gradient descent
* Results:
  + - F1 Score: 0.9405338661823265
    - Accuracy: 0.9917202174987642
* Inference:

For label Geyser off

tp: 108 fp: 3 fn: 6 tn: 7975

For label Geyser on

tp: 114 fp: 1 fn: 32 tn: 7945

For label Refrigerator on

tp: 182 fp: 10 fn: 15 tn: 7885

For label refrigerator 0ff

tp: 247 fp: 20 fn: 14 tn: 7811

For label No event

tp: 7374 fp: 33 fn: 0 tn: 685

## Experiments 3

* Data from dates between 6th and 9th August was used
* The dataset had this distribution (24519 used for training. Most No event instances were removed)
  + No Event: 5000276
  + Geyser off: 360
  + Geyser on: 460
  + Refrigerator off: 604
  + Refrigerator On: 805
* Features used were: ['mains\_power','reactive\_power','time\_since\_last\_event','time\_since\_event\_start', 'normalized\_mains\_power']
* We used 100 epochs of batch gradient descent
* Results:
  + - F1 Score: 0.94071350807468
    - Accuracy: 0. 9917202174987642
* Inference:

For label 0

tp: 109 fp: 1 fn: 5 tn: 7977

For label 1

tp: 115 fp: 3 fn: 31 tn: 7943

For label 2

tp: 182 fp: 15 fn: 15 tn: 7880

For label 3

tp: 247 fp: 19 fn: 14 tn: 7812

For label 4

tp: 7372 fp: 29 fn: 2 tn: 689

## Experiment 4

Same as 3

Features: 'active\_power\_P1', 'active\_power\_P2', 'active\_power\_P3','reactive\_power\_Q1', 'reactive\_power\_Q2', 'reactive\_power\_Q3'

0.8744317063988625

0.9778793870489372

For label 0

tp: 91 fp: 7 fn: 23 tn: 7971

For label 1

tp: 125 fp: 7 fn: 21 tn: 7939

For label 2

tp: 110 fp: 17 fn: 87 tn: 7878

For label 3

tp: 240 fp: 6 fn: 21 tn: 7825

For label 4

tp: 7347 fp: 142 fn: 27 tn: 576

# Final Pipeline Results

## Data between 16 – 18 August.

Testing metrics is on completely unseen data

### Training Metrics ###

Without Voting

1.0

1.0

For label AC\_OFF

tp: 60 fp: 0 fn: 0 tn: 3600

For label AC\_ON

tp: 120 fp: 0 fn: 0 tn: 3540

For label Geyser\_OFF

tp: 60 fp: 0 fn: 0 tn: 3600

For label Geyser\_ON

tp: 100 fp: 0 fn: 0 tn: 3560

For label Microwave\_OFF

tp: 100 fp: 0 fn: 0 tn: 3560

For label Microwave\_ON

tp: 100 fp: 0 fn: 0 tn: 3560

For label no\_event

tp: 3120 fp: 0 fn: 0 tn: 540

With Voting

1.0

1.0

For label AC\_OFF

tp: 3 fp: 0 fn: 0 tn: 180

For label AC\_ON

tp: 6 fp: 0 fn: 0 tn: 177

For label Geyser\_OFF

tp: 3 fp: 0 fn: 0 tn: 180

For label Geyser\_ON

tp: 5 fp: 0 fn: 0 tn: 178

For label Microwave\_OFF

tp: 5 fp: 0 fn: 0 tn: 178

For label Microwave\_ON

tp: 5 fp: 0 fn: 0 tn: 178

For label no\_event

tp: 156 fp: 0 fn: 0 tn: 27

### Validation Metrics ###

Without Voting

0.9270201644028547

0.9637096774193549

For label AC\_OFF

tp: 40 fp: 20 fn: 0 tn: 1180

For label AC\_ON

tp: 37 fp: 0 fn: 3 tn: 1200

For label Geyser\_OFF

tp: 40 fp: 0 fn: 0 tn: 1200

For label Geyser\_ON

tp: 20 fp: 0 fn: 0 tn: 1220

For label Microwave\_OFF

tp: 20 fp: 0 fn: 0 tn: 1220

For label Microwave\_ON

tp: 33 fp: 15 fn: 7 tn: 1185

For label no\_event

tp: 1005 fp: 10 fn: 35 tn: 190

With Voting

0.9400560224089636

0.967741935483871

For label AC\_OFF

tp: 2 fp: 1 fn: 0 tn: 59

For label AC\_ON

tp: 2 fp: 0 fn: 0 tn: 60

For label Geyser\_OFF

tp: 2 fp: 0 fn: 0 tn: 60

For label Geyser\_ON

tp: 1 fp: 0 fn: 0 tn: 61

For label Microwave\_OFF

tp: 1 fp: 0 fn: 0 tn: 61

For label Microwave\_ON

tp: 2 fp: 1 fn: 0 tn: 59

For label no\_event

tp: 50 fp: 0 fn: 2 tn: 10

## Data between 8 August to 11 August and 16 Augusts to 18 August

### Training Metrics ###

Without Voting

0.4927548755944336

0.8332310838445808

[[ 32 0 0 0 0 0 20 0 8]

[ 0 19 0 0 0 1 0 14 86]

[ 0 0 239 0 62 0 17 0 102]

[ 0 0 0 362 0 0 0 2 116]

[ 0 0 0 0 26 0 63 0 11]

[ 0 0 0 0 0 0 0 0 100]

[ 27 0 0 0 0 0 184 0 289]

[ 0 0 0 0 0 0 0 361 459]

[ 0 39 22 105 0 17 58 13 6926]]

For label AC\_OFF

tp: 32 fp: 27 fn: 28 tn: 9693

For label AC\_ON

tp: 19 fp: 39 fn: 101 tn: 9621

For label Geyser\_OFF

tp: 239 fp: 22 fn: 181 tn: 9338

For label Geyser\_ON

tp: 362 fp: 105 fn: 118 tn: 9195

For label Microwave\_OFF

tp: 26 fp: 62 fn: 74 tn: 9618

For label Microwave\_ON

tp: 0 fp: 18 fn: 100 tn: 9662

For label Refrigerator\_OFF

tp: 184 fp: 158 fn: 316 tn: 9122

For label Refrigerator\_ON

tp: 361 fp: 29 fn: 459 tn: 8931

For label no\_event

tp: 6926 fp: 1171 fn: 254 tn: 1429

With Voting

0.5283595630707245

0.8404907975460123

[[ 2 0 0 0 0 0 1 0 0]

[ 0 1 0 0 0 0 0 0 5]

[ 0 0 13 0 3 0 1 0 4]

[ 0 0 0 21 0 0 0 0 3]

[ 0 0 0 0 1 0 4 0 0]

[ 0 0 0 0 0 0 0 0 5]

[ 0 0 0 0 0 0 9 0 16]

[ 0 0 0 0 0 0 0 18 23]

[ 0 2 1 5 0 1 3 1 346]]

For label AC\_OFF

tp: 2 fp: 0 fn: 1 tn: 486

For label AC\_ON

tp: 1 fp: 2 fn: 5 tn: 481

For label Geyser\_OFF

tp: 13 fp: 1 fn: 8 tn: 467

For label Geyser\_ON

tp: 21 fp: 5 fn: 3 tn: 460

For label Microwave\_OFF

tp: 1 fp: 3 fn: 4 tn: 481

For label Microwave\_ON

tp: 0 fp: 1 fn: 5 tn: 483

For label Refrigerator\_OFF

tp: 9 fp: 9 fn: 16 tn: 455

For label Refrigerator\_ON

tp: 18 fp: 1 fn: 23 tn: 447

For label no\_event

tp: 346 fp: 56 fn: 13 tn: 74

### Validation Metrics ###

Without Voting

0.7409838574911662

0.8918181818181818

[[ 40 0 0 0 0 0 0 0 0]

[ 0 30 0 0 0 0 0 3 7]

[ 0 0 39 0 1 0 0 0 0]

[ 0 0 0 40 0 0 0 0 0]

[ 0 0 0 0 20 0 0 0 0]

[ 0 0 0 0 0 5 0 2 33]

[ 3 0 0 0 0 0 62 0 135]

[ 0 0 0 2 0 0 0 413 65]

[ 67 5 3 8 0 0 3 20 2294]]

For label AC\_OFF

tp: 40 fp: 70 fn: 0 tn: 3190

For label AC\_ON

tp: 30 fp: 5 fn: 10 tn: 3255

For label Geyser\_OFF

tp: 39 fp: 3 fn: 1 tn: 3257

For label Geyser\_ON

tp: 40 fp: 10 fn: 0 tn: 3250

For label Microwave\_OFF

tp: 20 fp: 1 fn: 0 tn: 3279

For label Microwave\_ON

tp: 5 fp: 0 fn: 35 tn: 3260

For label Refrigerator\_OFF

tp: 62 fp: 3 fn: 138 tn: 3097

For label Refrigerator\_ON

tp: 413 fp: 25 fn: 67 tn: 2795

For label no\_event

tp: 2294 fp: 240 fn: 106 tn: 660

With Voting

0.7477862085004943

0.9090909090909091

[[ 2 0 0 0 0 0 0 0 0]

[ 0 2 0 0 0 0 0 0 0]

[ 0 0 2 0 0 0 0 0 0]

[ 0 0 0 2 0 0 0 0 0]

[ 0 0 0 0 1 0 0 0 0]

[ 0 0 0 0 0 0 0 0 2]

[ 0 0 0 0 0 0 3 0 7]

[ 0 0 0 0 0 0 0 23 1]

[ 3 0 0 1 0 0 0 1 115]]

For label AC\_OFF

tp: 2 fp: 3 fn: 0 tn: 160

For label AC\_ON

tp: 2 fp: 0 fn: 0 tn: 163

For label Geyser\_OFF

tp: 2 fp: 0 fn: 0 tn: 163

For label Geyser\_ON

tp: 2 fp: 1 fn: 0 tn: 162

For label Microwave\_OFF

tp: 1 fp: 0 fn: 0 tn: 164

For label Microwave\_ON

tp: 0 fp: 0 fn: 2 tn: 163

For label Refrigerator\_OFF

tp: 3 fp: 0 fn: 7 tn: 155

For label Refrigerator\_ON

tp: 23 fp: 1 fn: 1 tn: 140

For label no\_event

tp: 115 fp: 10 fn: 5 tn: 35