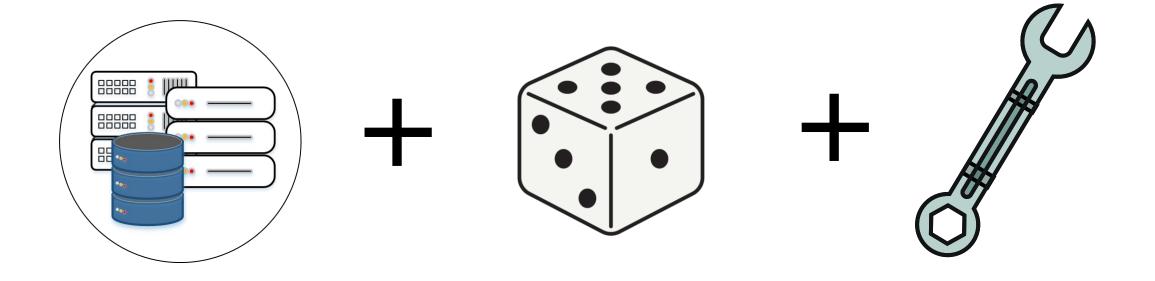


How to model stochastic behavior of failures in telco or IT systems using machine learning?

Sadegh Karimi



Hardware, estimation & maintenance





AirFrame Open Edge Server



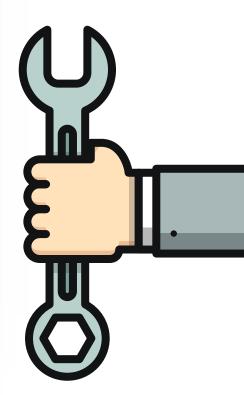
AirScale baseband

Corrective Maintenance (CM)

- Income loss
- Reputation damage
- Less resource planning
- Expensive repairs
- Longer downtime

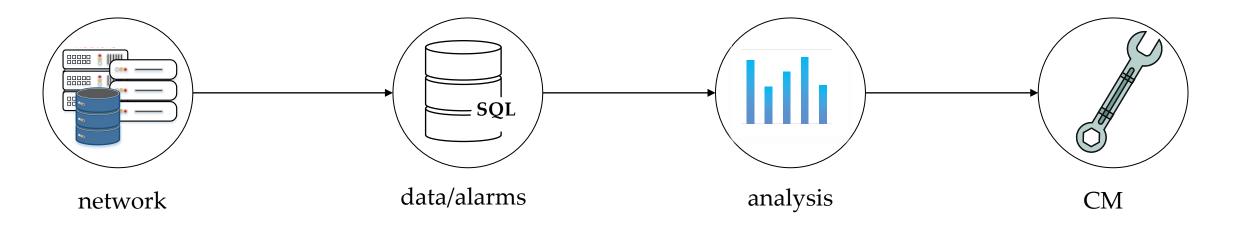
Preventive Maintenance (PM)

- Prevent income loss
- Prevent reputation damage
- More resource planning
- Prevent expensive repairs
- Shorter downtime



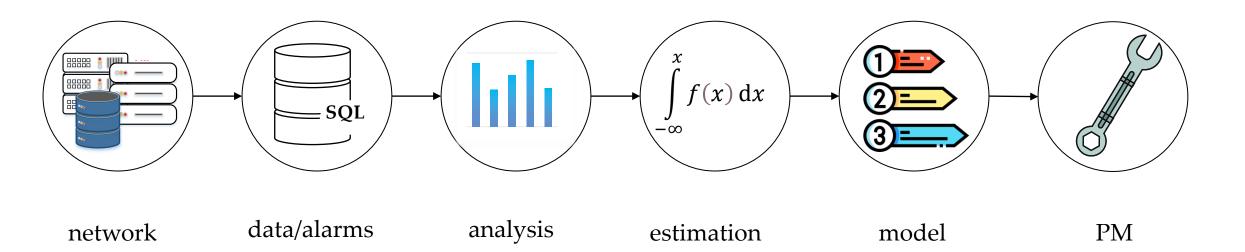
Conventional Method

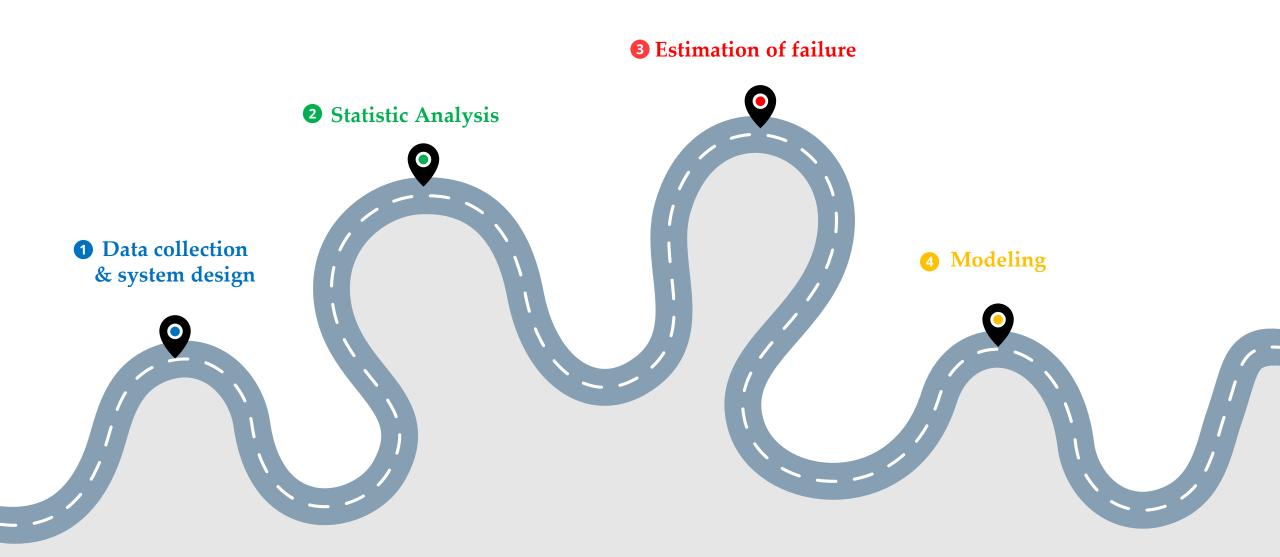
- Failure happens, then perform the fix
- Perform PM on regular basis
- CM is playing active role



This method:

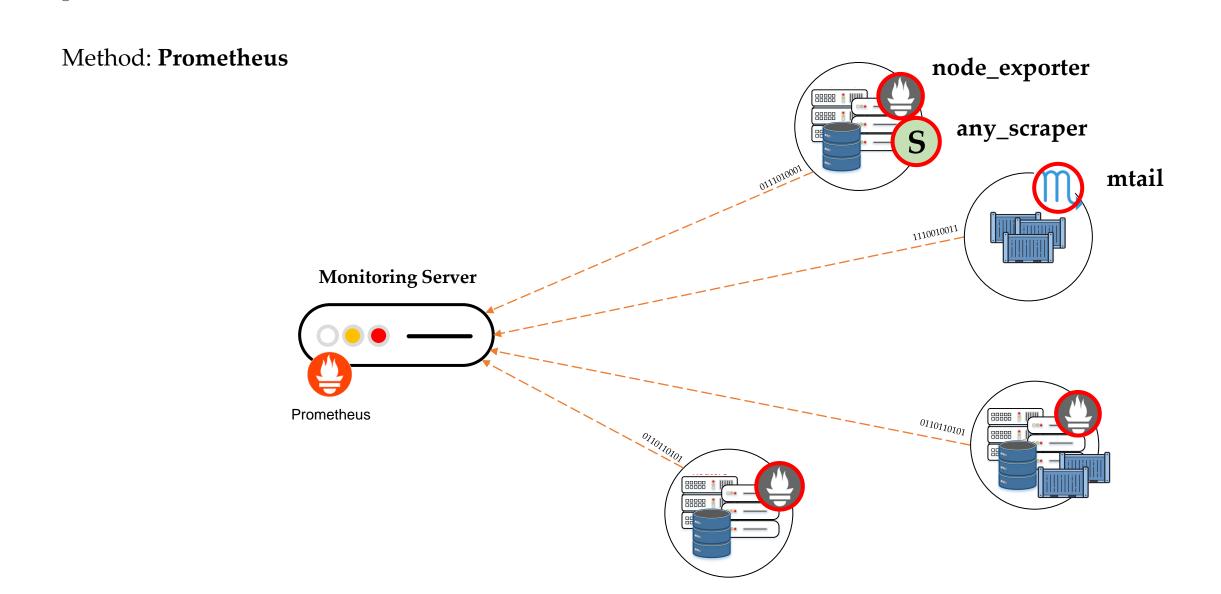
- Anticipate the failure, fix it before happens
- PM is playing active role



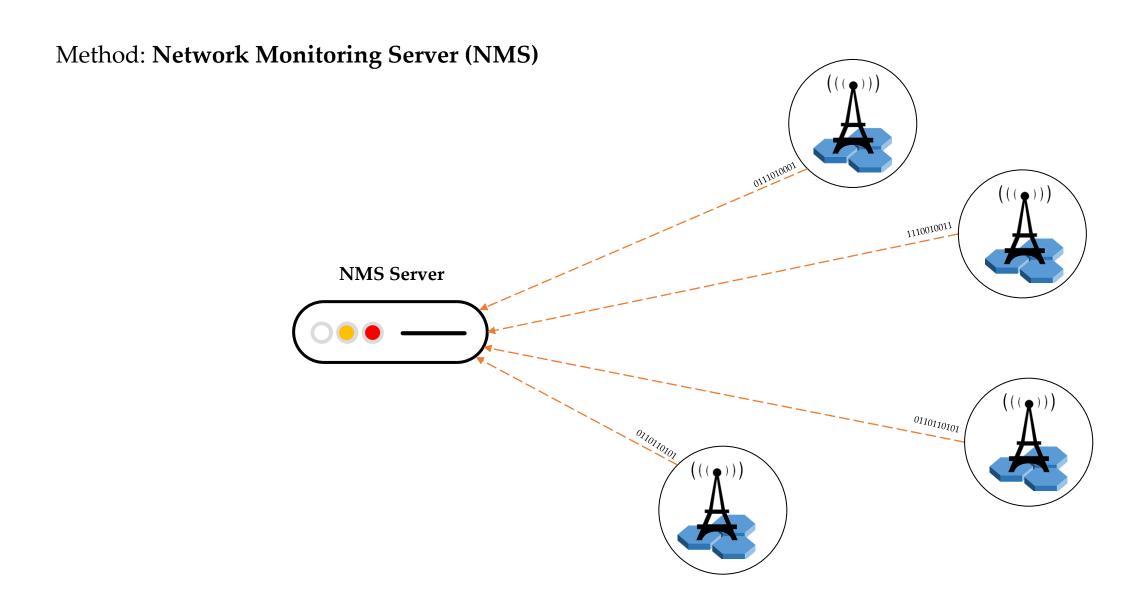


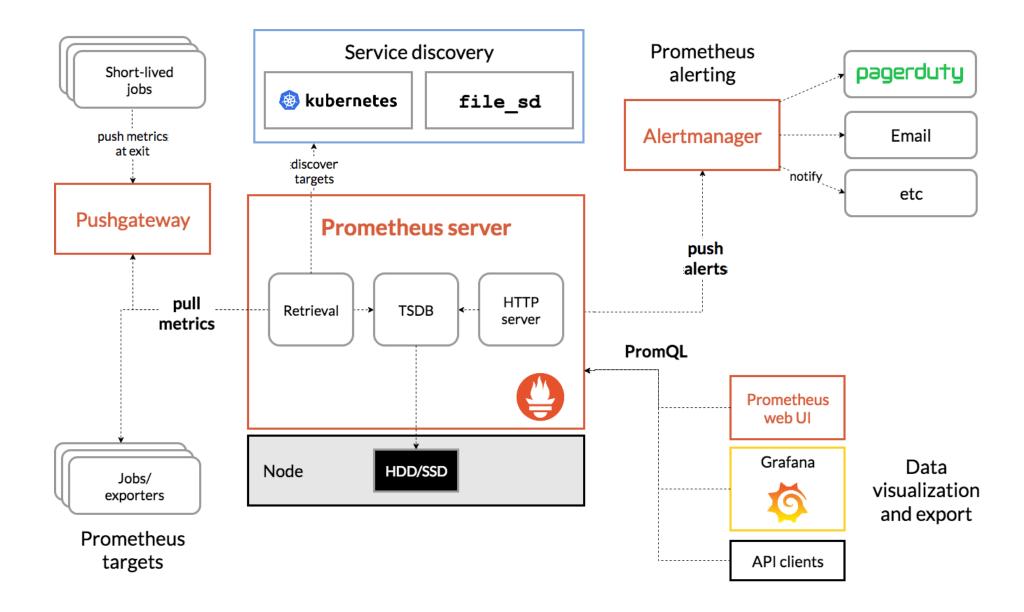
Data collection & system design

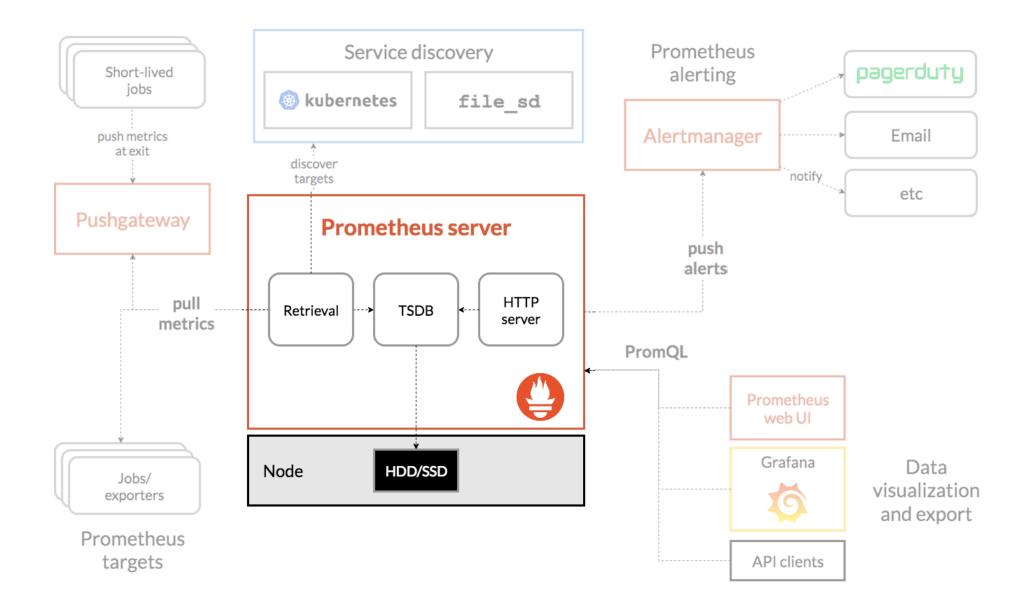
Data Collection

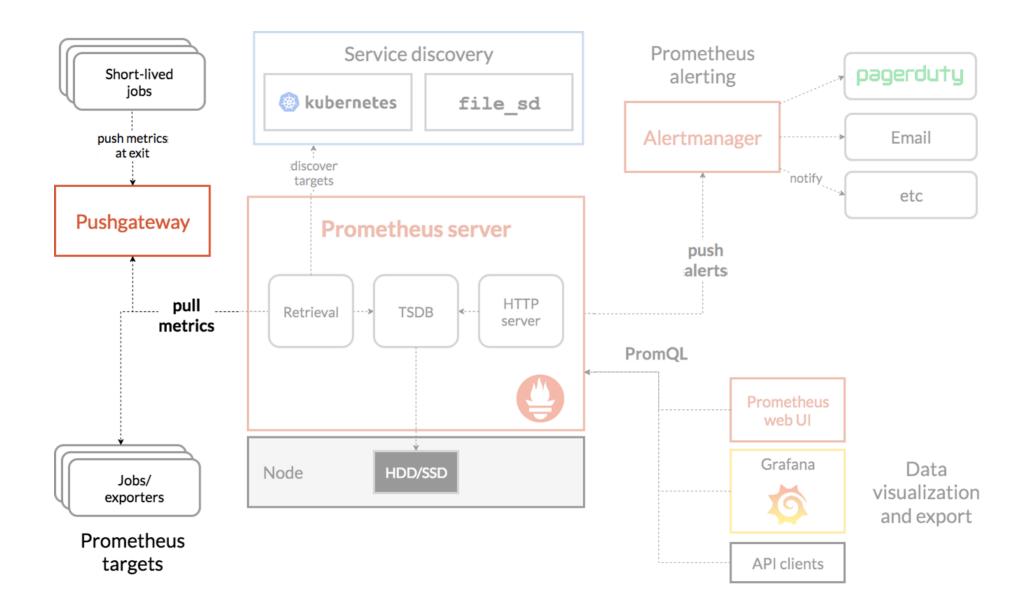


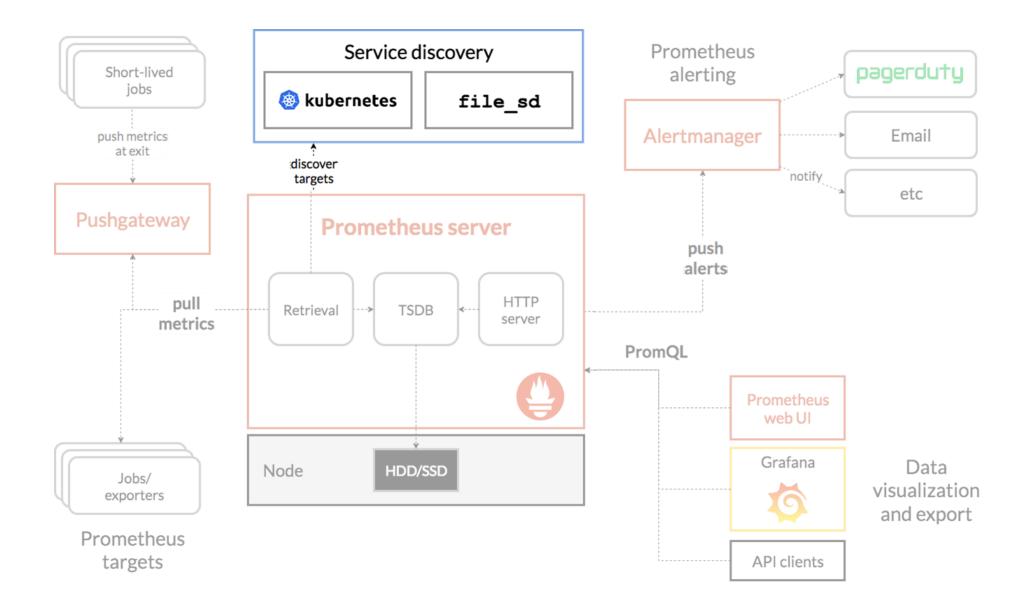
Data Collection

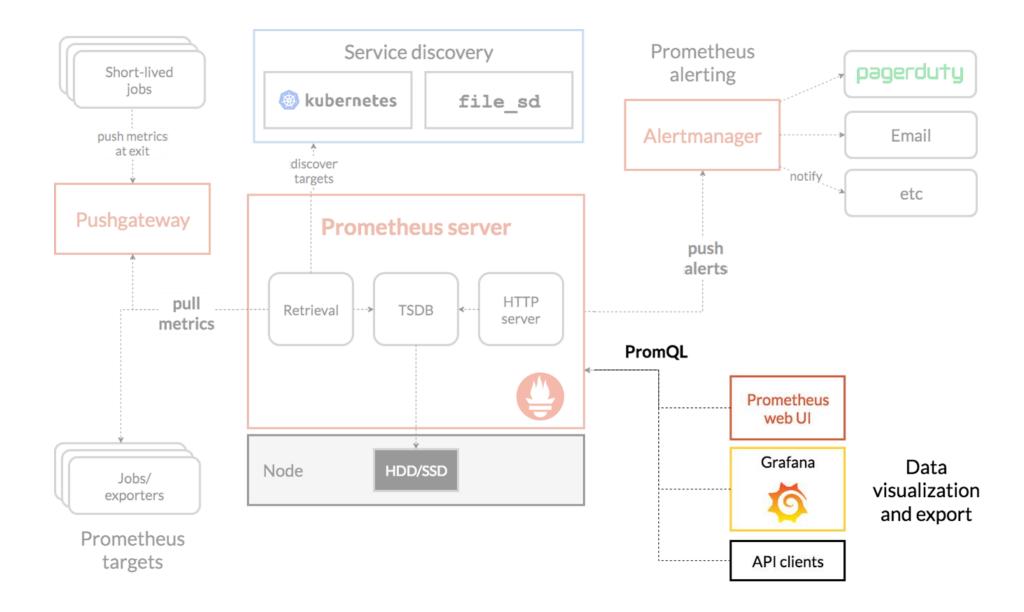


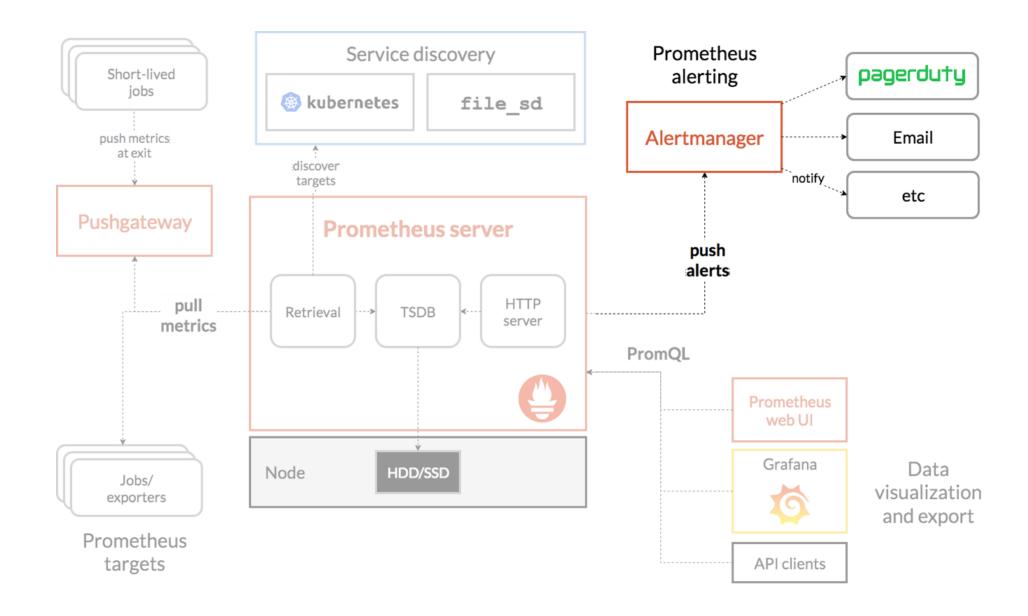


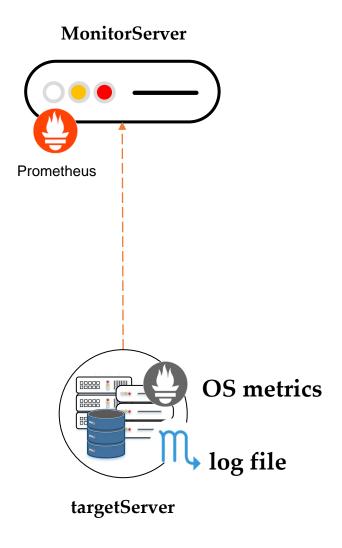






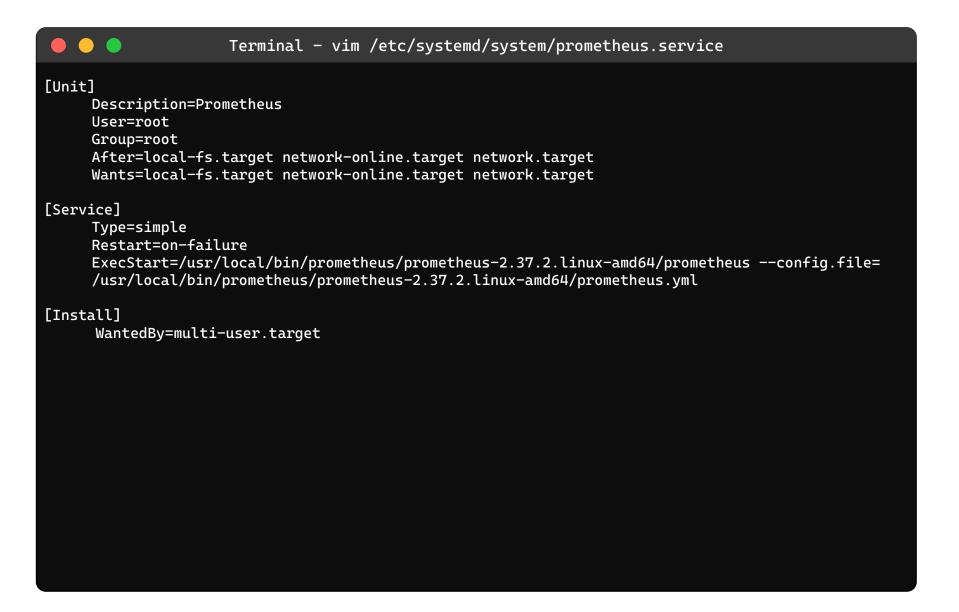




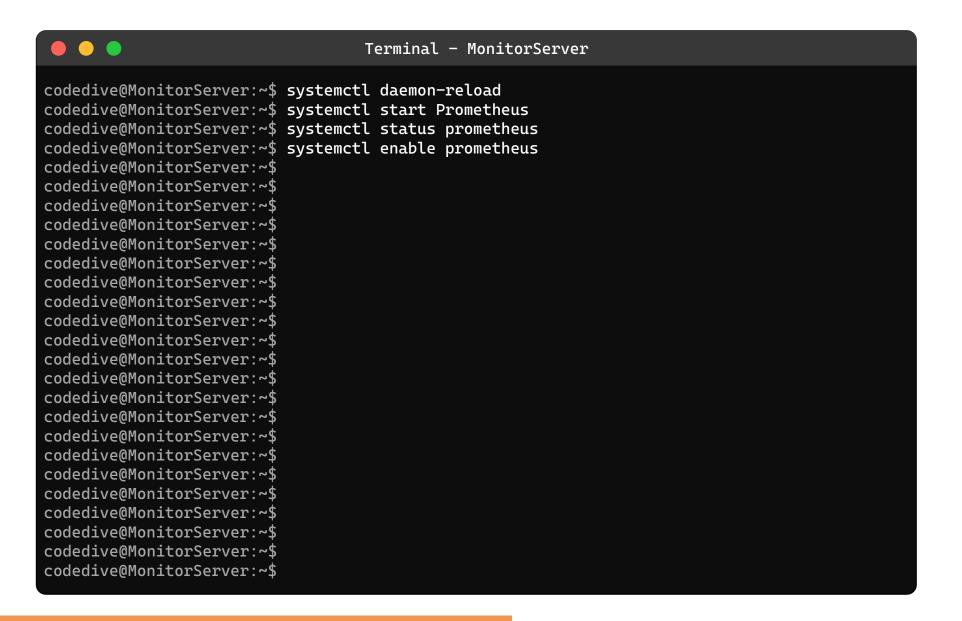


```
Terminal - MonitorServer
codedive@MonitorServer:~$ sudo wget https://github.com/prometheus/prometheus/releases/
download/v2.37.2/prometheus-2.37.2.linux-amd64.tar.gz
codedive@MonitorServer:~$
codedive@MonitorServer:~$ sudo mkdir /usr/local/bin/Prometheus
codedive@MonitorServer:~$
codedive@MonitorServer:~$ mv prometheus-2.37.2.linux-amd64.tar.gz /usr/local/bin/Prometheus
codedive@MonitorServer:~$ cd /usr/local/bin/Prometheus
codedive@MonitorServer:~$
codedive@MonitorServer:~$ tar xvfz prometheus-2.37.2.linux-amd64.tar.gz
codedive@MonitorServer:~$ cd prometheus-2.37.2.linux-amd64
codedive@MonitorServer:~$
codedive@MonitorServer:~$ vim prometheus.yml
codedive@MonitorServer:~$
codedive@MonitorServer:~$
codedive@MonitorServer:~$ sudo ufw allow 9090/tcp
codedive@MonitorServer:~$
codedive@MonitorServer:~$
                          ./prometheus --config.file=prometheus.yml > /dev/null 2>&1 &
codedive@MonitorServer:~$
codedive@MonitorServer:~$
codedive@MonitorServer:~$
codedive@MonitorServer:~$ vim /etc/systemd/system/prometheus.service
```

MonitorServer Prometheus



MonitorServer Prometheus



MonitorServer Prometheus

```
Terminal - TargetServer
codedive@TargetServer:~$
codedive@TargetServer:~$ wget https://github.com/prometheus/node_exporter/releases/download
/v1.4.0/node_exporter-1.4.0.linux-amd64.tar.gz
codedive@TargetServer:~$ tar xvfz node_exporter-1.4.0.linux-amd64.tar.gz
codedive@TargetServer:~$
codedive@TargetServer:~$ cd node_exporter-1.4.0.linux-amd64
codedive@TargetServer:~$ mv node_exporter /usr/local/bin/node_exporter
codedive@TargetServer:~$
codedive@TargetServer:~$ sudo ufw allow 8081/tcp
codedive@TargetServer:~$
codedive@TargetServer:~$ node_exporter --web.listen-address 0.0.0.0:8081 > /dev/null 2>&1 &
codedive@TargetServer:~$
```

MonitorServer Prometheus targetServer

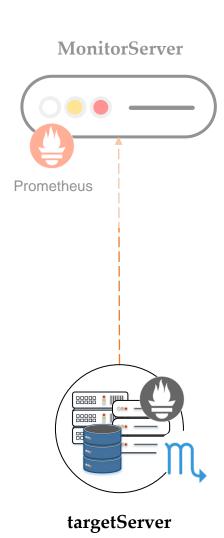
• • •	● ● Terminal - TargetServer									
codedive@TargetServer:~\$ cat anylogfile.log										
	rms, cpu errors				alarm.	fan	degraded	priority	failure	
2022211500	86.245.244.226		0		1	1	3	1	. 6. 2 (3 .2 (
2022220912	242.254.113.85		0		3	3	3	1		
2022231054		2	0		0	0	2	2		
2022251123		2	0		2	2	3	2		
2022261254		1	0		1	1	4	1		
2022281333	166.42.237.79	0	0	2	1	0	4	1		
2022291531	240.151.80.141		0	3	0	0	1	2		
2022271723	251.187.13.90	0	0		3	3	2	2		
2022241324	32.40.200.86	0	0	1	0	0	1	Θ		
2022011136	106.129.137.98	Θ	0	2	4	3	2	2		
2022031252	101.25.58.95	1	0	2	3	3	1	2		
2022041521	16.10.150.63	1	2	1	3	3	2	2		
2022102043	168.147.131.18	1	0	2	2	1	3	2		
2022122245	230.12.118.189	1	1	1	4	3	2	3		
2022141247	17.226.234.190	1	0	1	0	0	3	Θ		
2022191539	119.159.65.0	1	0	4	0	0	3	2		
2022201848	129.255.137.20	0	0	2	0	0	1	1		
2022232346	172.91.66.180	1	0	1	3	3	3	2		
2022232339	154.92.117.111	1	0	3	1	1	2	1		
2022251449	12.103.196.116	1	0	1	3	3	2	2		
2022271049	90.205.195.153	0	0	2	0	0	1	1		
2022271943	98.95.46.105	1	0	1	3	3	3	2		
2022281157	158.120.7.77	1	0	3	1	1	2	1		
2022292021	20.28.39.208	1	0	1	3	3	2	2		

MonitorServer Prometheus

```
Terminal - TargetServer
codedive@TargetServer:~$ apt-get update
codedive@TargetServer:~$ apt-get install mtail
codedive@TargetServer:~$ systemctl enable mtail
codedive@TargetServer:~$
codedive@TargetServer:~$ ufw allow 3903/tcp
codedive@TargetServer:~$
codedive@TargetServer:~$ vim /etc/mtail/whateverlogfile.mtail
```

MonitorServer Prometheus targetServer

```
Terminal - vim /etc/mtail/whateverlogfile.mtail
counter hardware_alarm by ip
counter cpu_error by ip
# log_format mtail '$time, $ip, $"hw alarms", $"cpu errors" , $"low volt"'
                  '$"overheat alarm", $"fan degraded", $priority, $failure';
/\s(?P<transfertime>\d+)/ +
# remote IP
/\s(?P<remoteIP>\d+\.\d+\.\d+\.\d+)/ +
# hw alarms
/\s(?P<hw_alarms>\d+)/ +
# cpu errors
/\s(?P<cpu_error>\d+)/ +
  hardware_alarm[$ip]++
  cpu_error[$ip]++
```



```
Terminal - TargetServer
codedive@TargetServer:~$
codedive@TargetServer:~$ mtail --logs /var/log/anyservice/anylogfile.log --progs/etc/
mtail/ > /dev/null 2>&1 &
codedive@TargetServer:~$
```

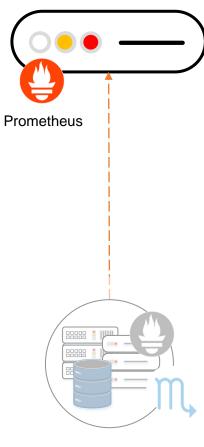
MonitorServer Prometheus targetServer

```
Terminal - MonitorServer
codedive@MonitorServer:~$
codedive@MonitorServer:~$ cd /usr/local/bin/prometheus/prometheus-2.37.2.linux-amd64/
codedive@MonitorServer:~$
codedive@MonitorServer:~$ vim prometheus.yml
```

MonitorServer Prometheus

```
Terminal - vim prometheus.yml
scrape_configs:
. . . . . . . . . .
  - job_name: "nodes"
    # metrics definition
    scrape_interval: 5s
    static_configs:
           - targets: ["63.32.45.224:8081", "63.32.45.224:3903"]
             labels:
                       group: 'nodes'
```

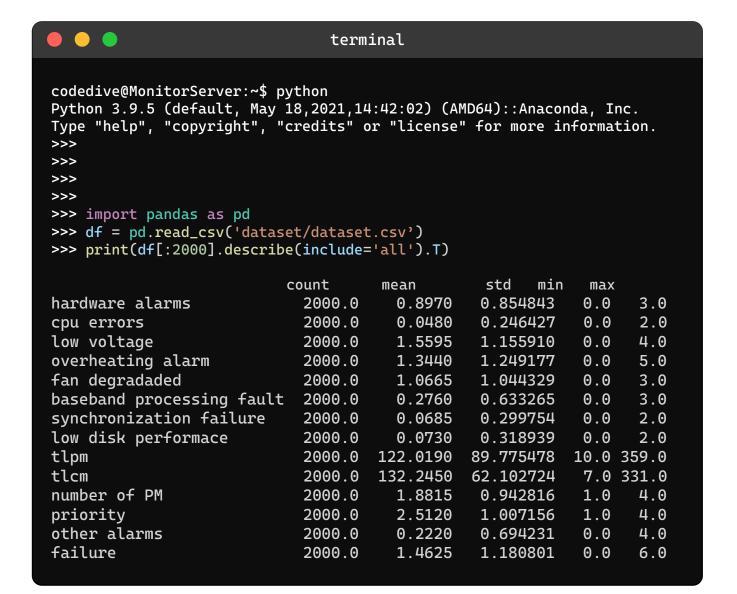
MonitorServer



targetServer

Statistic Analysis

Insight over dataset

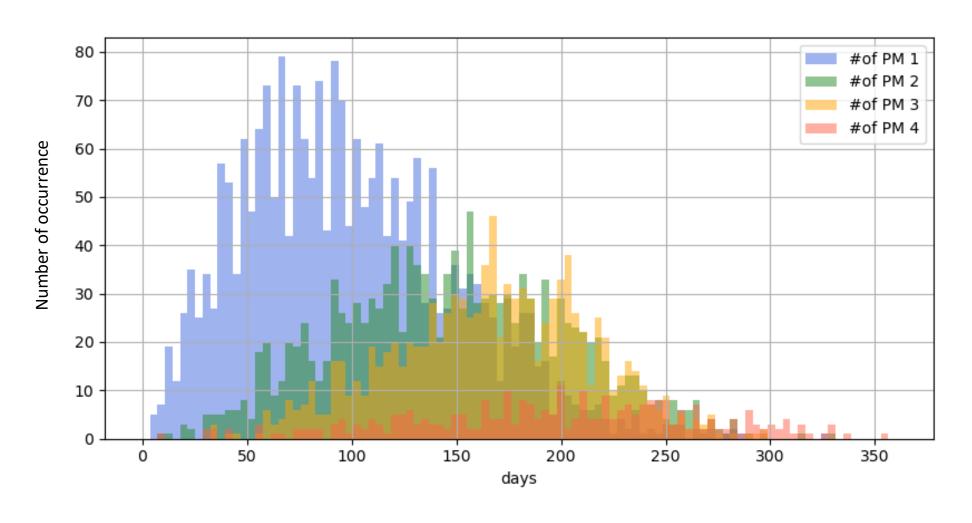


List of features

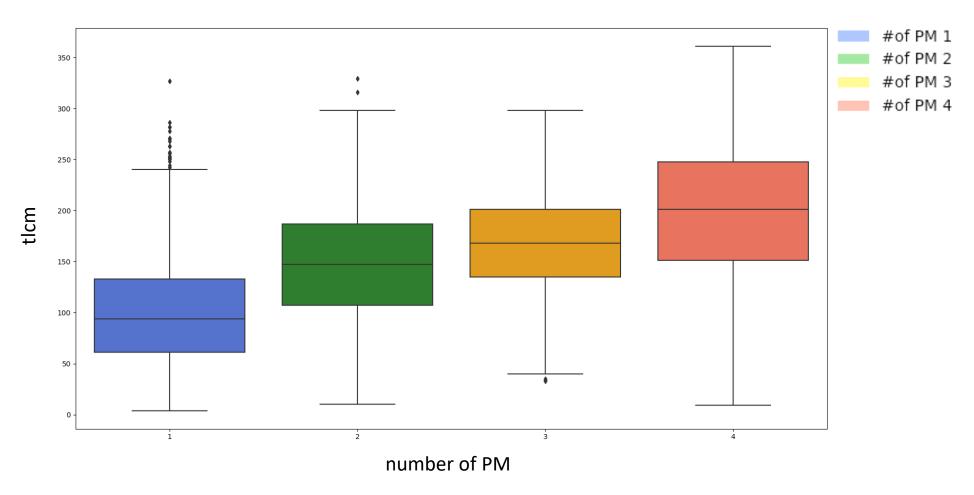
hardware alarms
cpu errors
low voltage
overheating alarm
fan degradaded
baseband processing fault
synchronization failure
low disk performace
tlpm
tlcm
number of PM
priority
other alarms
failure



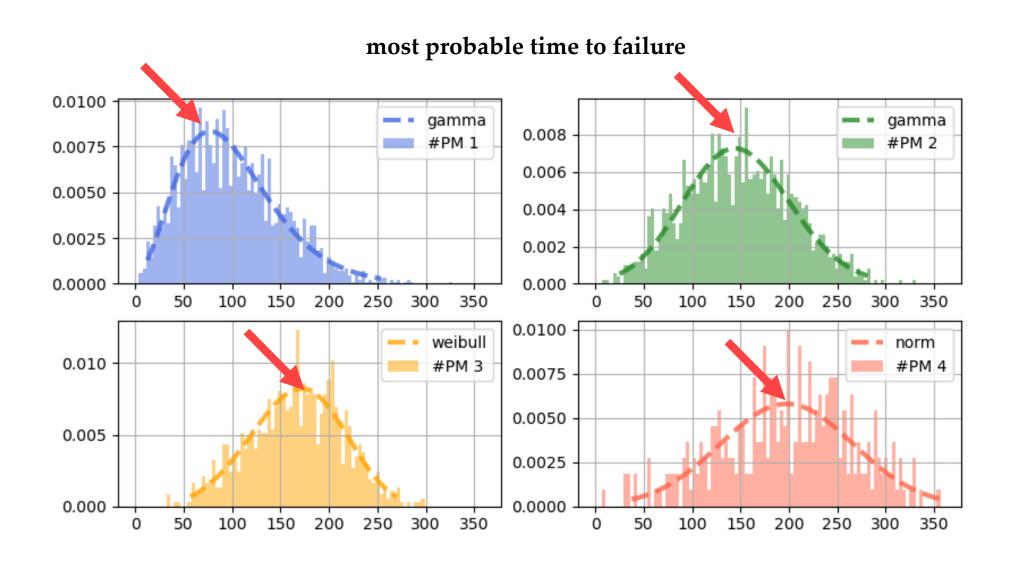
Time to last corrective maintenance



Time to last corrective maintenance

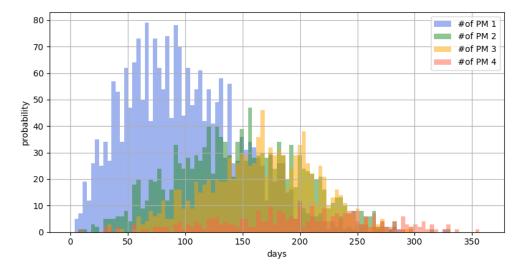


Fitting the Probability function



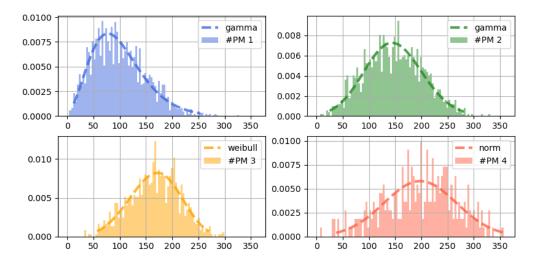
Insight over dataset

```
import numpy as np
   import pandas as pd
   from scipy.stats import gamma, weibull_min, norm
   import matplotlib.pyplot as plt
   from fitter import Fitter
   df = pd.read_csv('dataset/dataset.csv')
   plt.figure(1)
   plt.hist(df.loc[df['number of PM'] == 1,'tlcm'], bins = 100,
                range=(0,360),alpha=0.5, label='#of PM 1', density=False,
12
                facecolor='royalblue')
13
   plt.hist(df.loc[df['number of PM'] == 2,'tlcm'], bins = 100,
                range=(0,360),alpha=0.5, label='#of PM 2', density=False,
                facecolor='forestgreen')
17
   plt.hist(df.loc[df['number of PM'] == 3,'tlcm'], bins = 100,
                range=(0,360),alpha=0.5, label='#of PM 3', density=False,
                facecolor='orange')
21
   plt.hist(df.loc[df['number of PM'] == 4,'tlcm'], bins = 100,
                range=(0,360),alpha=0.5, label='#of PM 4', density=False,
                facecolor='tomato')
27
   plt.xlabel('days')
   plt.ylabel('probability')
   plt.grid()
   plt.legend(loc='best')
```



Insight over dataset

```
33 # plot and fitting the pdf based on different #number of PMs
35 plt.figure(2)
36 plt.subplot(2, 2, 1)
37 dataIn = df.loc[df['number of PM'] == 1,'tlcm']
38 f = Fitter(dataIn, distributions=['gamma', "weibull_min", "norm"])
39 f.fit()
40 print('parameter for #PM=1 are: {}'.format(f.get_best(method = 'sumsquare_error')))
41 bFitPara = f.get_best(method = 'sumsquare_error')
42 a = bFitPara['gamma']['a']
43 l = bFitPara['gamma']['loc']
44 s = bFitPara['gamma']['scale']
46 x = np.linspace(gamma.ppf(0.01, a, loc = 1, scale = s), gamma.ppf(0.99, a, loc = 1,
   scale = s), 100)
   plt.plot(x, gamma.pdf(x, a,loc = 1, scale = s), '--', color='royalblue', lw=3,
                alpha=0.8, label='gamma')
51
    plt.hist(dataIn, bins = 100, range=(0,360), alpha=0.5, label='#PM 1', density=True,
                 facecolor='royalblue')
54
   plt.grid()
   plt.legend(loc='best')
57
59 # plot for the remaining values of PM = {1,2,3,4} are almost same
62
```



Estimation of failures

Goal of estimation

Goal:

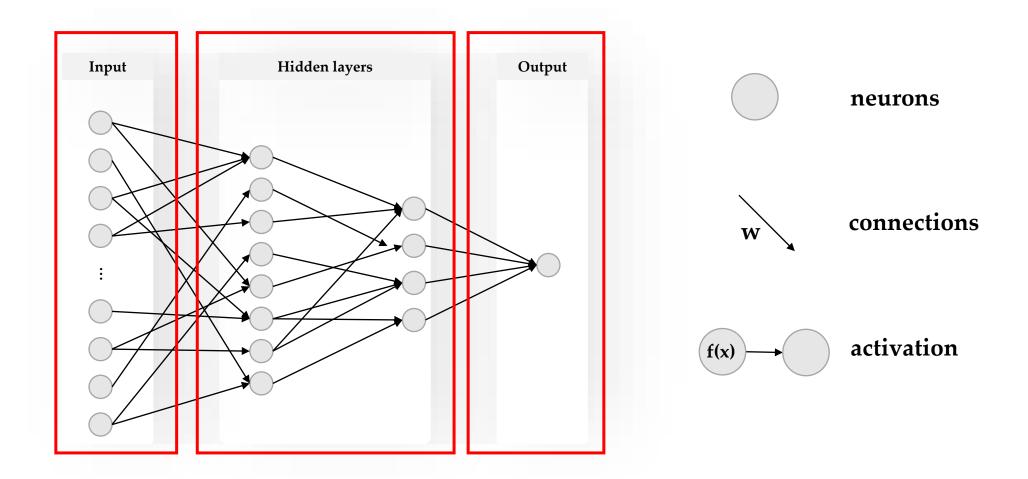
Anticipate the most probable failures

Methods:

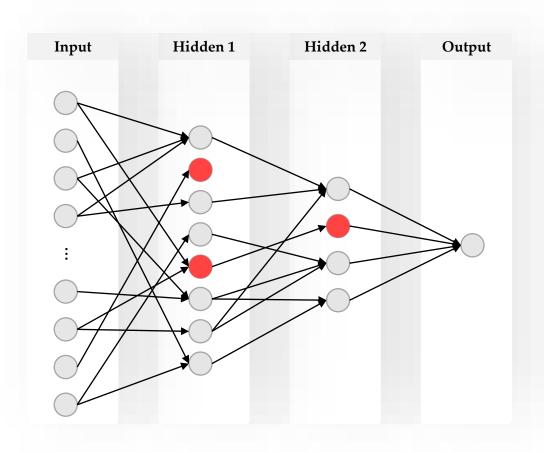
Method 1: fit probability function over data

Method 2: use the neural network for estimation

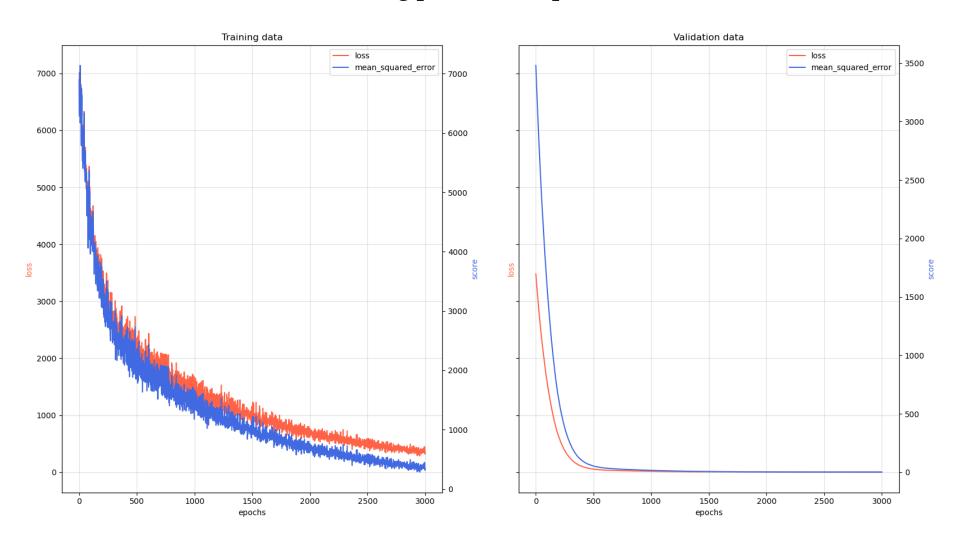
What is neural network?



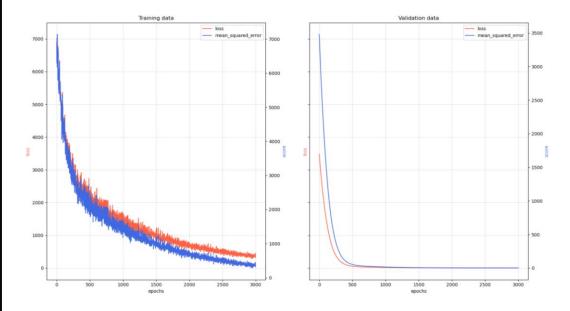
```
NN Model
   import pandas as pd
 2 from matplotlib import pyplot as plt
   from tensorflow.keras import models, layers, metrics
    import shap
   import tensorflow as tf
   tf.random.set_seed(5)
   df = pd.read_csv('dataset/dataset.csv', index_col=None)
   # create neural network model
   model = models.Sequential()
   model.add(layers.Dense(name="Hidden 1", input dim=13, units=8, activation='linear'))
   model.add(layers.Dropout(name="drop1", rate=0.15))
15 model.add(layers.Dense(name="Hidden 2", units=4, activation='linear'))
16 model.add(layers.Dropout(name="drop2", rate=0.15))
   model.add(layers.Dense(name="Output", units=1, activation='linear'))
   model.compile(optimizer='Adadelta', loss='mse',
                 metrics=[tf.keras.metrics.MeanSquaredError()])
21
   dataOut = df['failure']
    dataIn = df.drop(['failure'], axis=1)
24
25 x training = dataIn[0:1000].values
26 y_training = dataOut[0:1000].values
    training = model.fit(x=x_training, y=y_training, batch_size=8, epochs=3000,
                 shuffle=True, verbose=1, validation split=0.3)
31
32
```



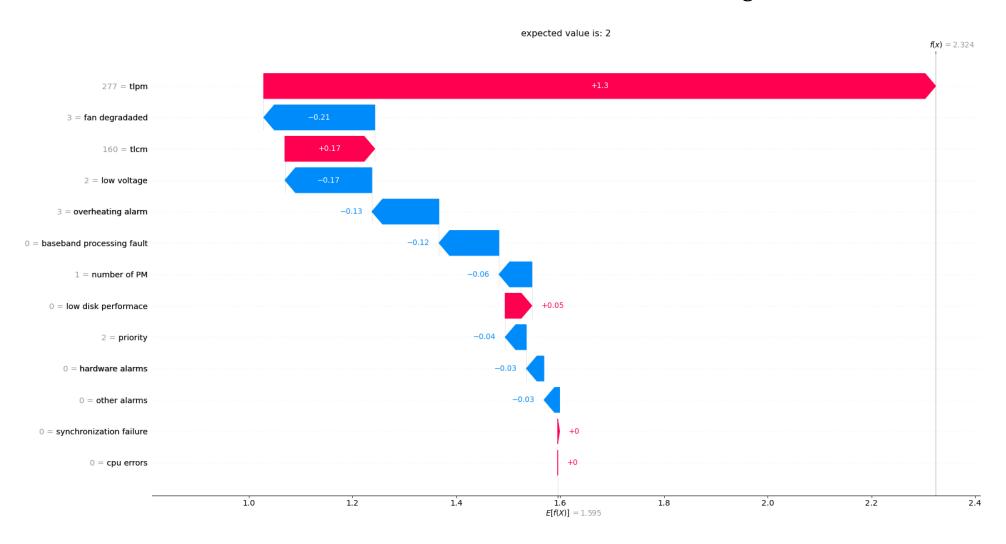
learning process vs. epochs



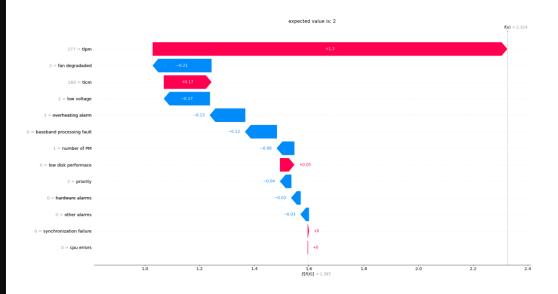
```
plots
33 # plots for training and validation on dataset
   fig, ax = plt.subplots(nrows=1, ncols=2, sharey=True)
36 plt.figure(1)
37 # training plot
38 ax[0].set(title="Training data")
39 ax11 = ax[0].twinx()
40 lns01 = ax[0].plot(training.history['loss'], color='tomato', label='loss')
41 ax[0].set_xlabel('epochs')
42 ax[0].set ylabel('loss', color='tomato')
43 lns02 = ax11.plot(training.history[metric], label=metric, color='royalblue')
44 ax11.set ylabel("score", color='royalblue')
46 lns = lns01+lns02
47 labs = [l.get_label() for l in lns]
48 ax[0].legend(lns, labs, loc='upper right')
   ax[0].grid(alpha=0.4)
51 # validation plot
52 ax[1].set(title="Validation data")
53 ax22 = ax[1].twinx()
54 lns11 = ax[1].plot(training.history['val_loss'], color='tomato', label='loss')
55 ax[1].set_xlabel('epochs')
56 ax[1].set_ylabel('loss', color='tomato')
57 lns12 = ax22.plot(training.history['val '+metric], label=metric, color='royalblue')
58 ax22.set_ylabel("score", color="royalblue")
60 lns = lns11 + lns12
61 labs = [l.get_label() for l in lns]
62 ax[1].legend(lns, labs, loc='upper right')
63 ax[1].grid(alpha=0.4)
```

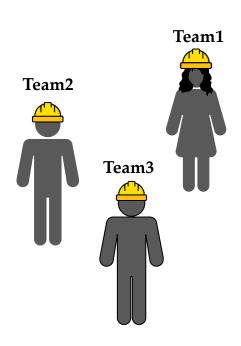


effect of each feature on estimated target

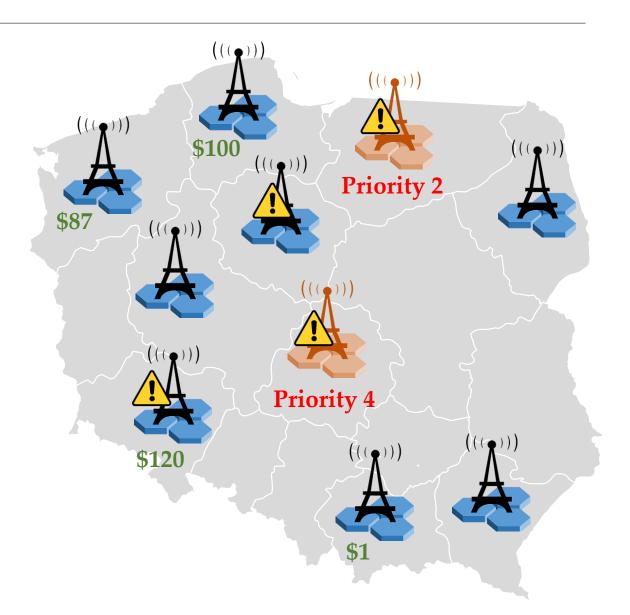


```
shap
65 # impact of feature to the value of target
   dataInExplainer = dataIn[-1:].values
   dataOutExpected = dataOut[-1:]
   explainer = shap.Explainer(model, x_training)
   shap_values = explainer(dataInExplainer)
   values = shap values.values[0]
   base values = shap values.base values[0][0]
   datas = df[-1:].values[0]
   columns = list(df.columns)[:-1] # exclude failure column
77
   plt.figure(2, figsize=(10, 8), dpi=80)
   exp = shap.Explanation(values, base_values, data=dataInExplainer[0],
                           feature names=columns)
   shap.plots.waterfall(exp, max_display=13, show=False)
   plt.rcParams["font.size"] = "12"
   plt.title('expected value is: {}'.format(dataOutExpected.values[0]))
   plt.grid()
   plt.show()
87
90
92
94
```









Probable financial loss:

```
for each node:
   p = find probability of failure
   c = find income per day
   Ci = p * c
```

Probable non-financial loss:

```
for each node:
    p = find probability of failure
    m = rate the node based on its importance 0-100
    Pi = p * m
```

Sets and Indices

I: Set of all available task, $i \in I$

M: Set of all available Workforces, $m \in M$

Parameters

 c_i : Probable financial loss due to task i

 p_i : Probable non-financial loss due to task i

 T_m : Available time in hours for workforce m

 a_{im} : time consumed by workforce m to complete task i

Variables

 y_{im} : is equal to 1 if workforce m is assigned to task i

 W_i : is equal to 1 if task i is selected to be performed

Cost function

$$Max\{\alpha_1 \times F_1 + \alpha_2 \times F_2\}$$

Where:

$$F_1 = \sum_{i \in I} c_i w_i$$
 Cost function of financial loss

$$F_2 = \sum_{i \in I} p_i w_i$$
 Cost function of non-financial loss

Constraints

$$w_i = \{0,1\} \quad \forall i \in I$$

$$y_{im} = \{0,1\} \quad \forall i \in I, m \in M$$

$$\sum_{m \in M} y_{im} = w_i \quad \forall i$$

assign only a task to a workforce

$$\max \sum_{i \in I} a_{im} y_{im} \le T_m \quad \forall m \in M$$

workforce time limitation

```
import pyomo.opt as po
   import pyomo.environ as pe
      = ['BTS01', 'BTS03', 'BTS04', 'BTS05', 'BTS08', 'BTS10']
       = ['Team01', 'Team02', 'Team03']
6 Ci = {'BTS01': 109, 'BTS03': 89, 'BTS04': 23, 'BTS05': 54, 'BTS08': 86, 'BTS10':
7 68}
8 Pi = {'BTS01': 17, 'BTS03': 2, 'BTS04': 100, 'BTS05': 21, 'BTS08': 71, 'BTS10': 34}
9 Tm = {'Team01': 2, 'Team02': 2, 'Team03': 3}
10
   Aim = {
11
12
        ('BTS01', 'Team01'): 1, ('BTS01', 'Team02'): 3, ('BTS01', 'Team03'): 1.5,
13
        ('BTS03', 'Team01'): 2, ('BTS03', 'Team02'): 5, ('BTS03', 'Team03'): 0.5,
        ('BTS04', 'Team01'): 3, ('BTS04', 'Team02'): 4, ('BTS04', 'Team03'): 4,
        ('BTS05', 'Team01'): 1, ('BTS05', 'Team02'): 1, ('BTS05', 'Team03'): 5,
15
        ('BTS08', 'Team01'): 3, ('BTS08', 'Team02'): 9, ('BTS08', 'Team03'): 3,
16
17
        ('BTS10', 'Team01'): 4, ('BTS10', 'Team02'): 5, ('BTS10', 'Team03'): 1
18 }
19
20 # set solver and create model
21 solver = po.SolverFactory('glpk')
   model = pe.ConcreteModel()
23
24 # Set & Indices definition
25 model.i = pe.Set(initialize=I)
   model.m = pe.Set(initialize=M)
27
28 # Parameter Declaration
   model.ci = pe.Param(model.i, initialize=Ci)
30 model.pi = pe.Param(model.i, initialize=Pi)
31 model.tm = pe.Param(model.m, initialize=Tm)
32 model.aim = pe.Param(model.i, model.m, initialize=Aim)
```

Sets and Indices

I: Set of all available skills, $i \in I$

M: Set of all available Workforces, $m \in M$

Parameters

 c_i : Probable financial loss due to task i

 p_i : Probable non-financial loss due to task i

 T_m : Available time in hours for workforce m

 a_{im} : time consumed by workforce m to complete task i

```
33 # Variable declaration
   model.yim = pe.Var(model.i, model.m, domain=pe.Binary)
35 model.wi = pe.Var(model.i, domain=pe.Binary)
37 # cost function
38 def costfunc(model):
       alphaC = 0.7; alphaP = 0.3
       fC = sum(model.ci[i] * model.wi[i] for i in model.i)
       fP = sum(model.pi[i] * model.wi[i] for i in model.i)
       return alphaC * fC + alphaP * fP
42
   model.obj = pe.Objective(rule = costfunc, sense = pe.maximize)
   # contstraint definition
   def ruleC1(model, i):
       return sum(model.yim[i,m] for m in model.m) == model.wi[i]
   model.C1 = pe.Constraint(model.i, rule=ruleC1)
51
   def ruleC2(model, m):
       return sum(model.aim[i,m] * model.yim[i,m] for i in model.i) <= model.tm[m]
54
   model.C2 = pe.Constraint(model.m, rule=ruleC2)
   # solve the model
   solver.solve(model)
59 # print outout of the model
   model.pprint()
61
62
63
```

Variables

 y_{im} : is equal to 1 if workforce m is assigned to task i

 W_i : is equal to 1 if task i is selected to be performed

Cost function

$$Max\{\alpha_1 \times F_1 + \alpha_2 \times F_2\}$$

Where:

$$F_1 = \sum_{i \in I} c_i w_i \qquad \text{Cost fur}$$

Cost function of financial loss

$$F_2 = \sum_{i \in I} p_i w_i$$

Cost function of non-financial loss

Constraints

$$w_i = \{0,1\} \quad \forall i \in I$$

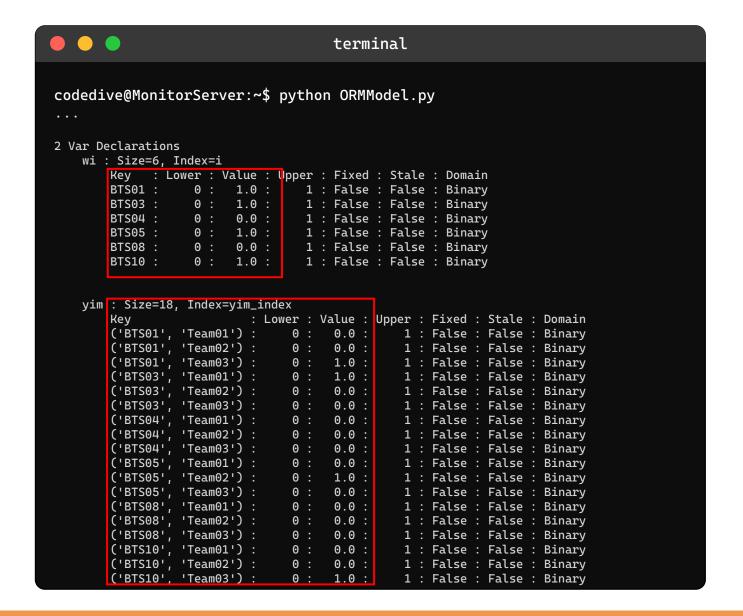
$$y_{im} = \{0,1\} \quad \forall i \in I, m \in M$$

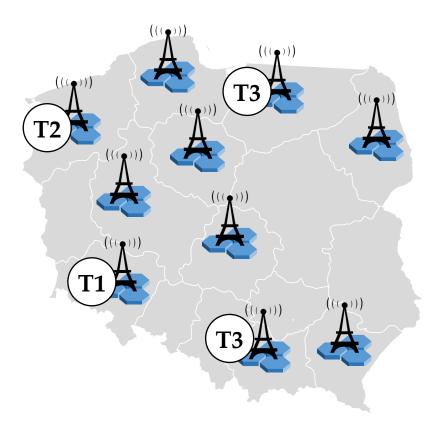
$$\sum_{m \in M} y_{im} = w_i \quad \forall i$$

assign only a task to a workforce

$$\max \sum_{i=1}^{n} a_{im} y_{im} \le T_m \quad \forall m \in M$$

workforce time limitation





Thank you



