

Julho 2021

## Case Predictive Maintenance: Paper CO



# Resumo

01

## Analise Dos Dados

Nos Dados existem colunas com valores constantes. Como não há ganho e eles são redundantes para modelos de aprendizado de máquina, não foi utilizado na análise.

Existe correlação entre alguns sensores. Tag 09 e 14 são quase colineares.

Tag 17 mostra um comportamento de 'stepping', onde todos os passos são sempre inteiros completos.

Os tempos de operação máximo e mínimo através das unidades variam de cerca de 125 ciclos a cerca de 350 ciclos, o que é cerca de 280% de mínimo a máximo.

02

## Modelo de Predição

Foi utilizado o modelo de LSTM tanto para a classificação binária quanto para o modelo de regressão.

Os dados de PM\_Test foram separados randomicamente pelas suas ids para gerar uma amostra de treino e teste.

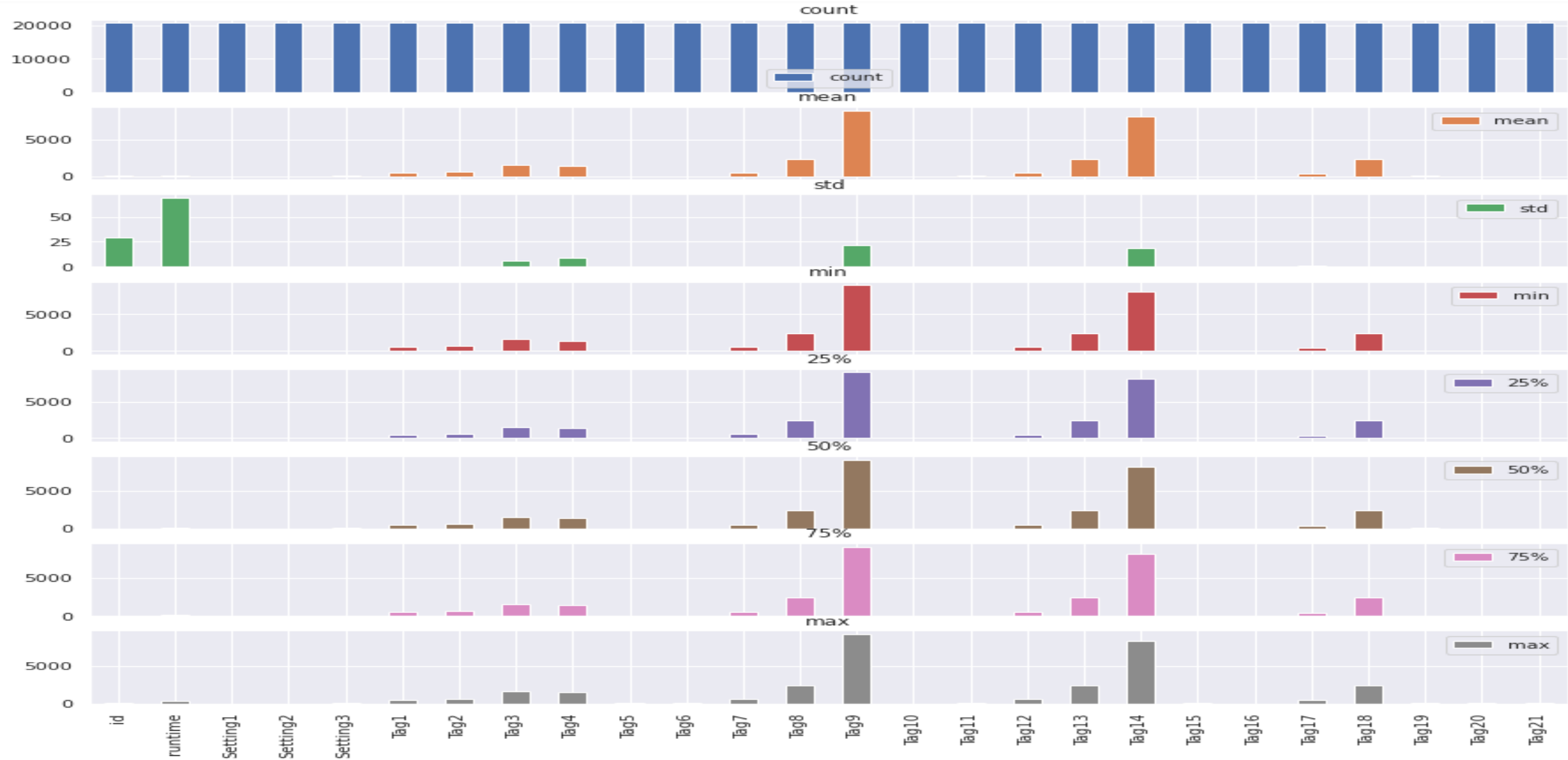
Para o modelo binário, utilizando valores acima de 20 ciclos foi obtido uma precisão de 96% na amostra de treinamento e 91% na amostra de teste

Para o modelo de regressão foi obtido um Mean Absolute Error de 16, 71 e  $R^2$  de 0,71. Na amostra de teste foi obtido o valor de MAE = 13, 99 e  $R^2 = 0, 79$

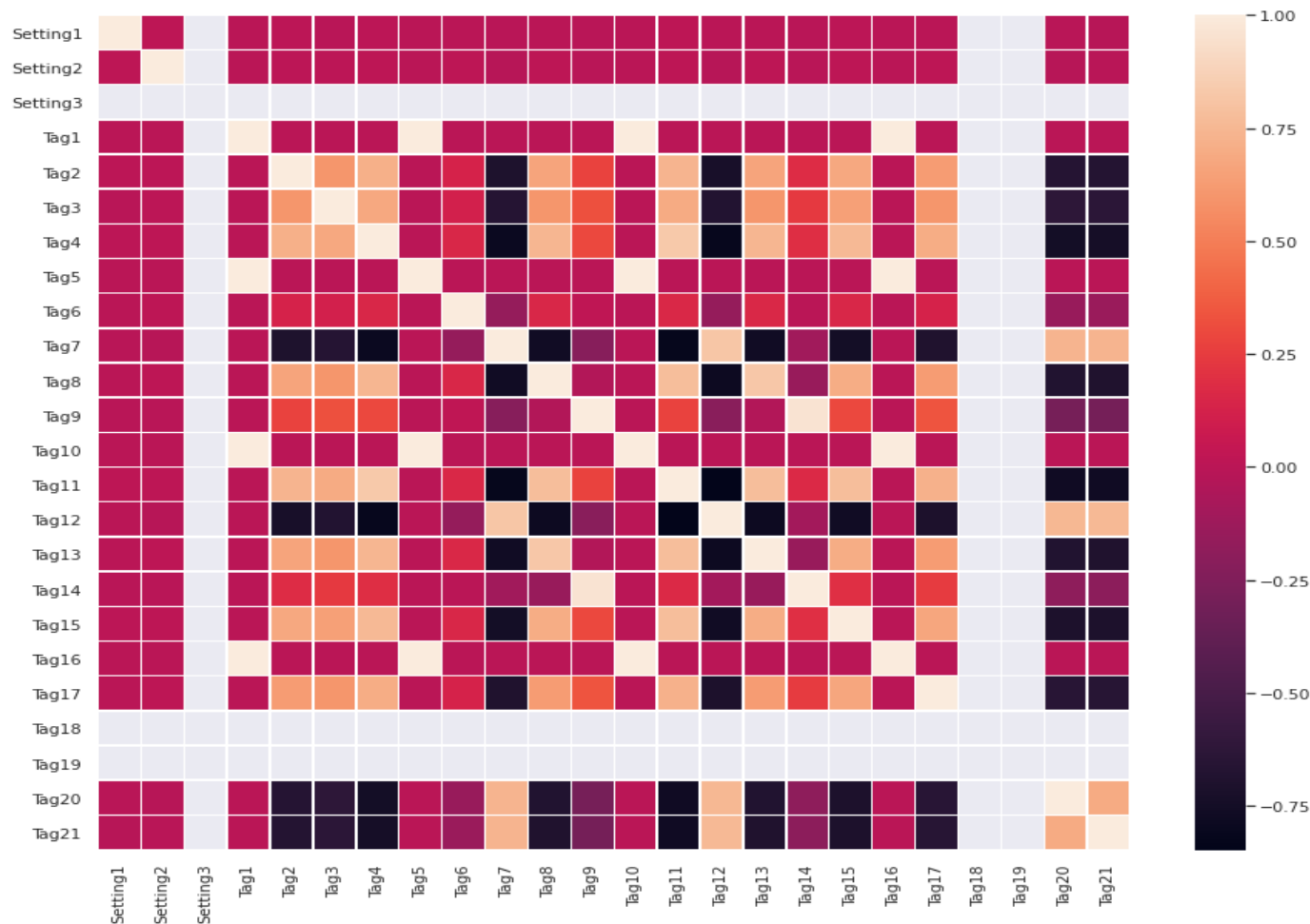
# 01

## Analise dos Dados

Pode ser visto que existem alguns dados bem correlacionados e outros que aparentemente são constantes.



Analizando a Matriz de Correlação é possível ver dados em cinza o que indica dados corrompidos ou constante.

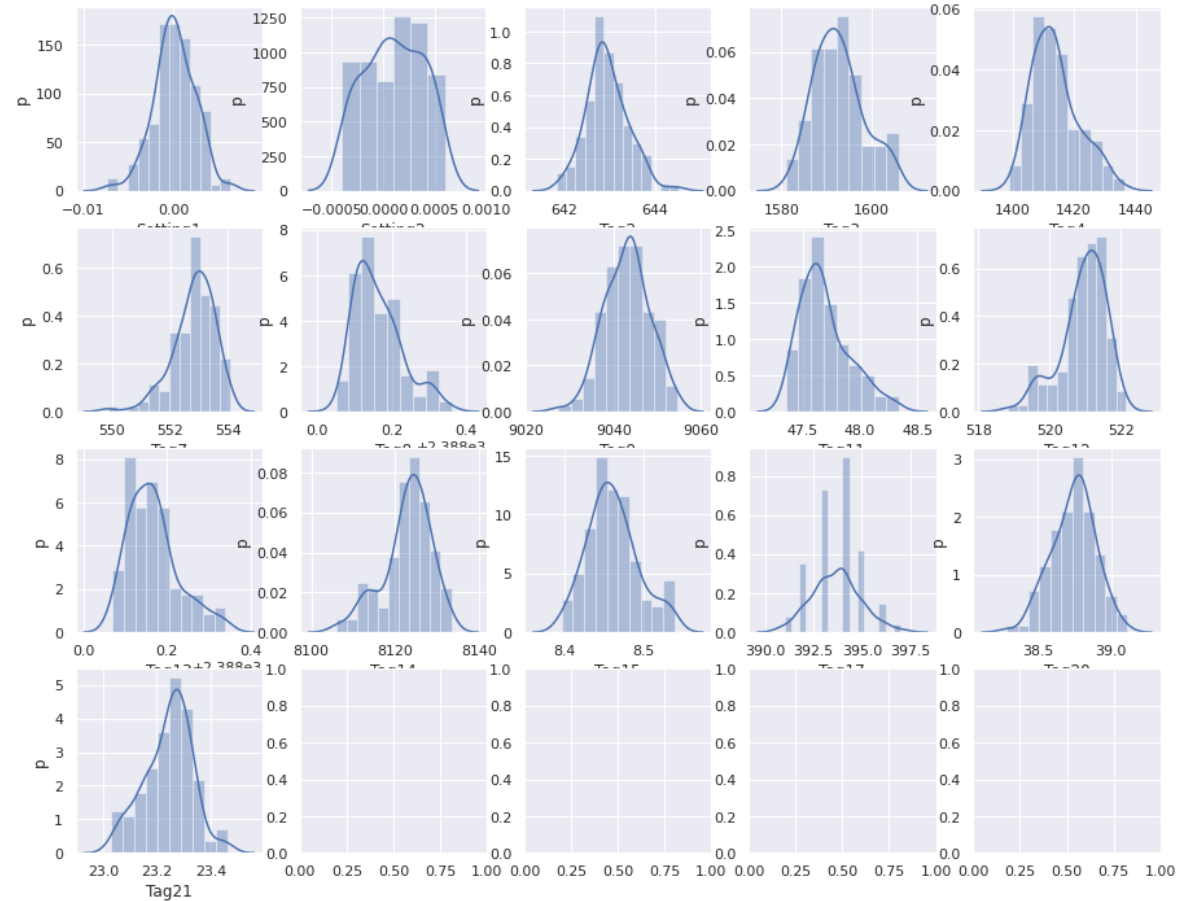
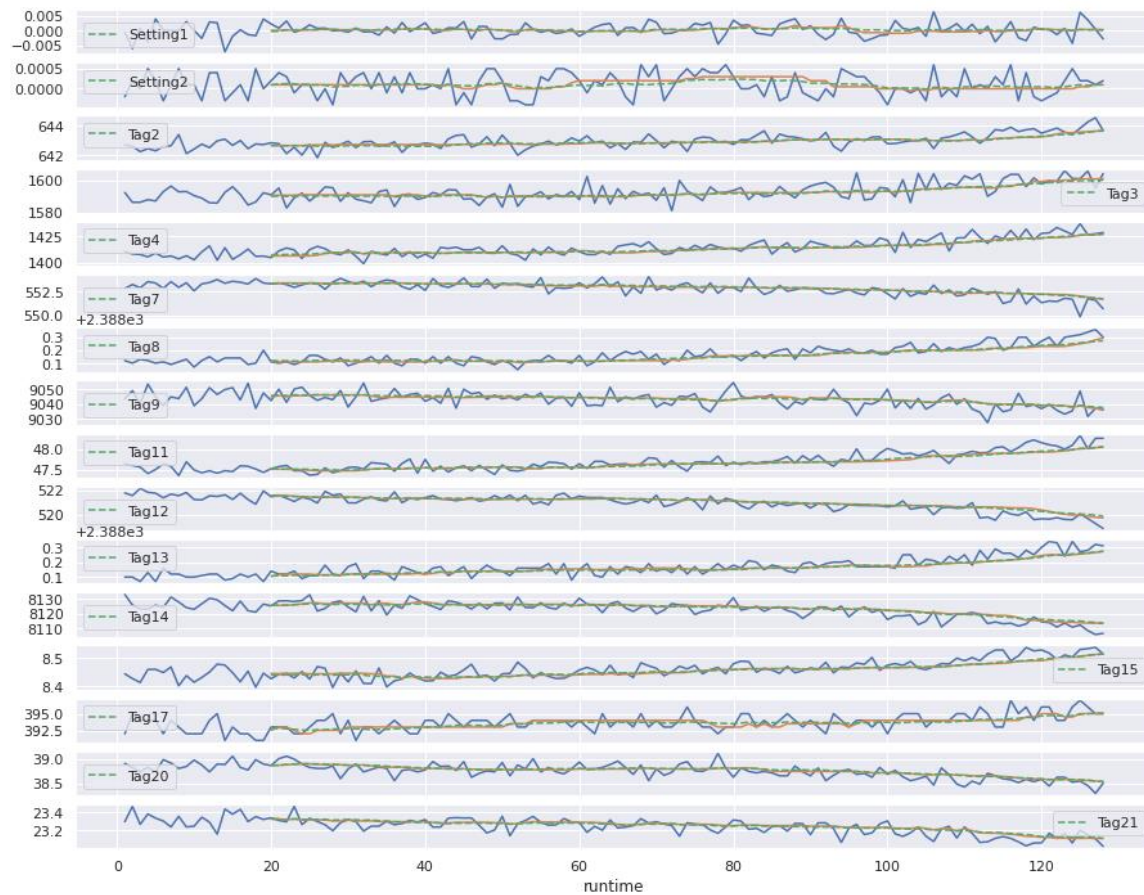


Da Matriz de correção também foi possível ver que:

- sensor 09 e sensor 14 são quase colineares
- sensor 08 e sensor 13 colinear relativamente forte
- sensores 11 e 12 são relativamente fortes colineares inversamente
- sensor 07 correlaciona-se com o par 11 e 12
- o sensor 04 se correlaciona com o par 11 e 12 também

# Plots de tempo e distribuição para Id curtos

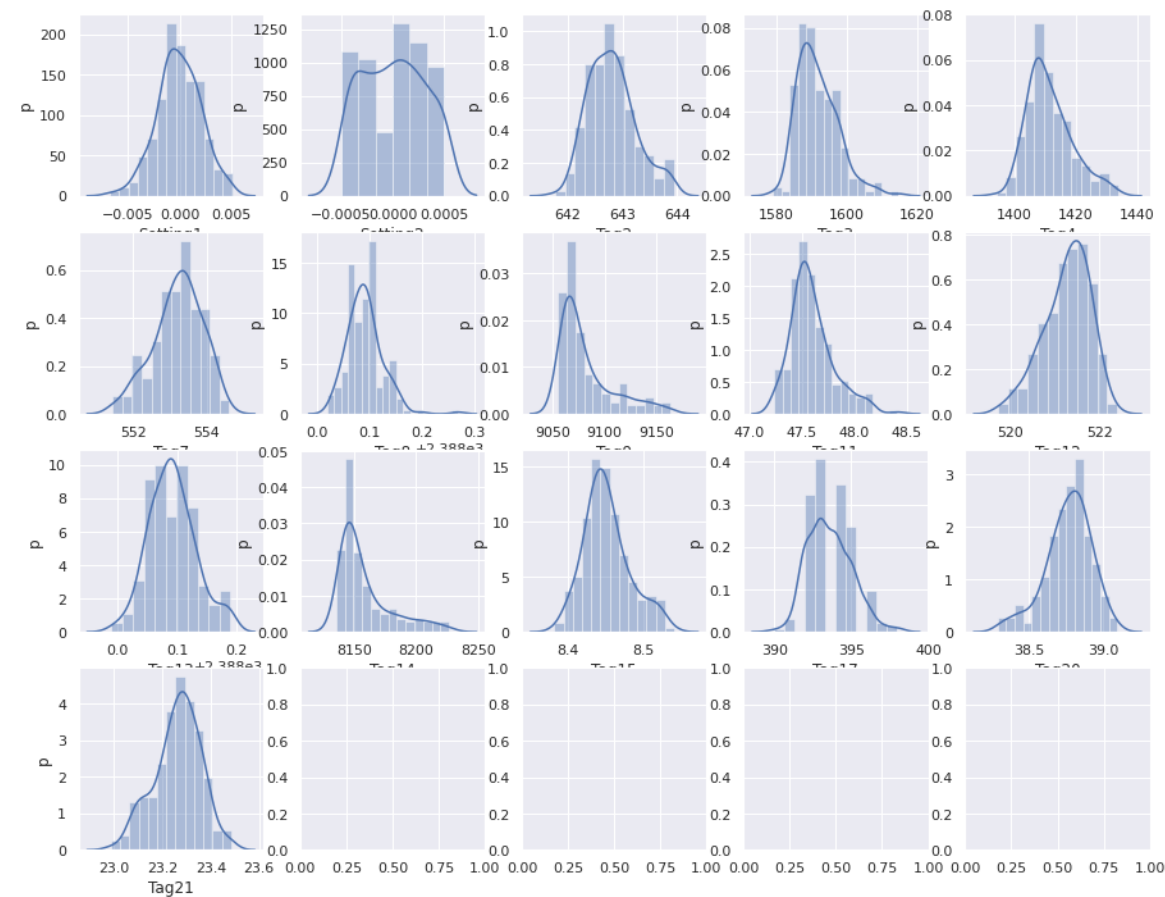
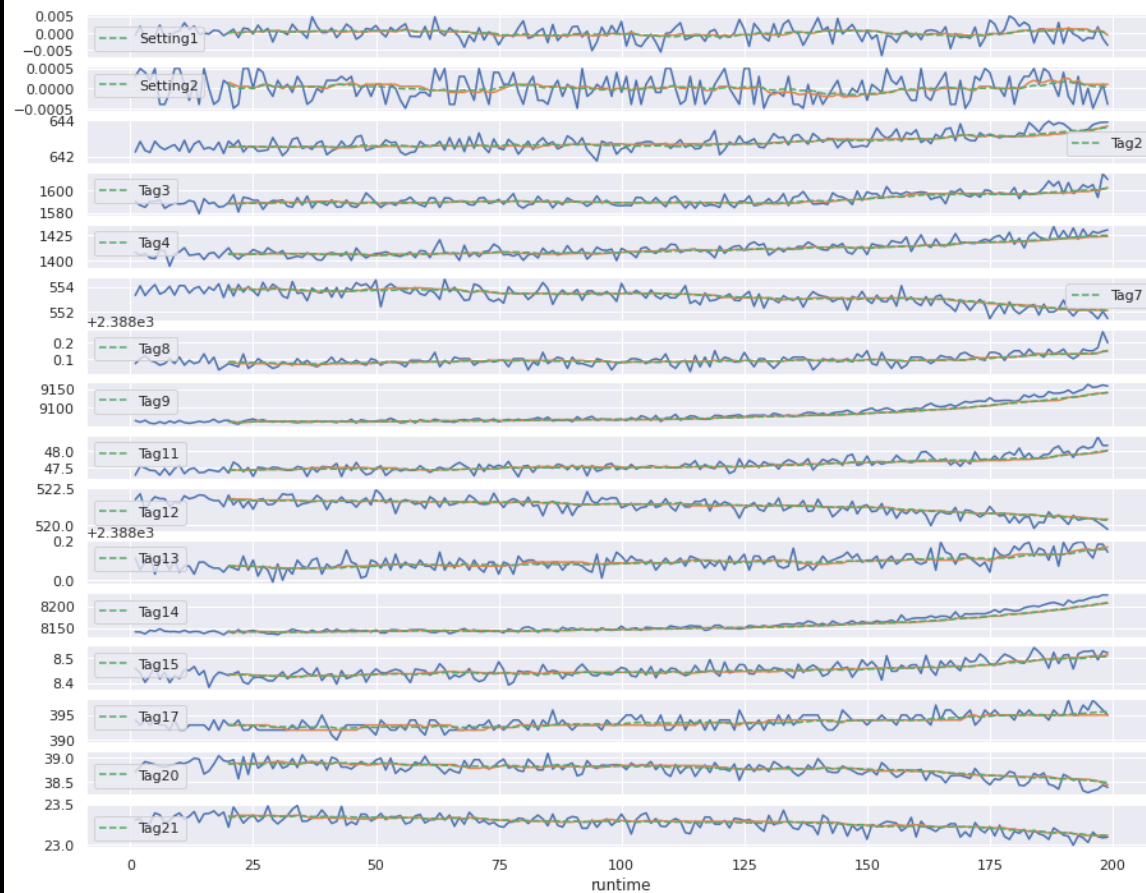
Asset id: 39





# Plots de tempo e distribuição para Ids médios

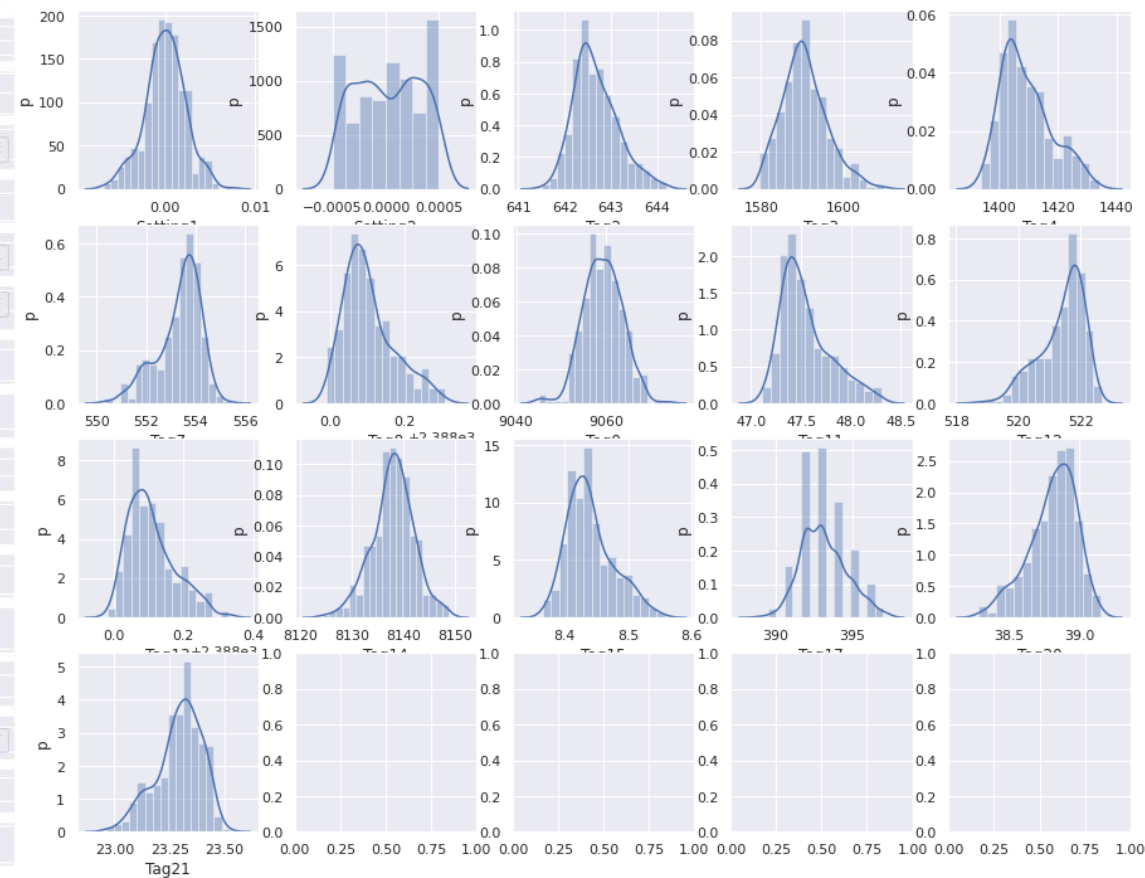
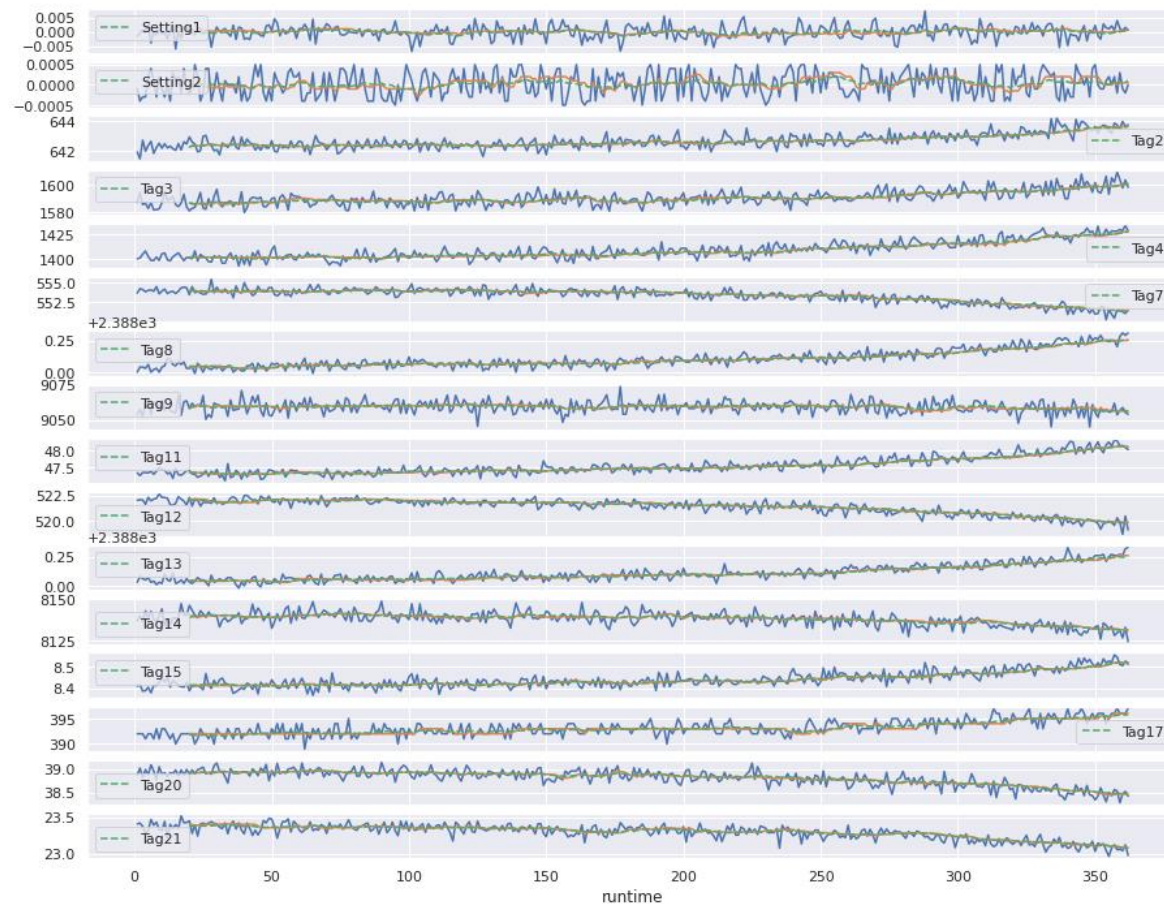
Asset id: 68



# Plots de tempo e distribuição para Ids Longos

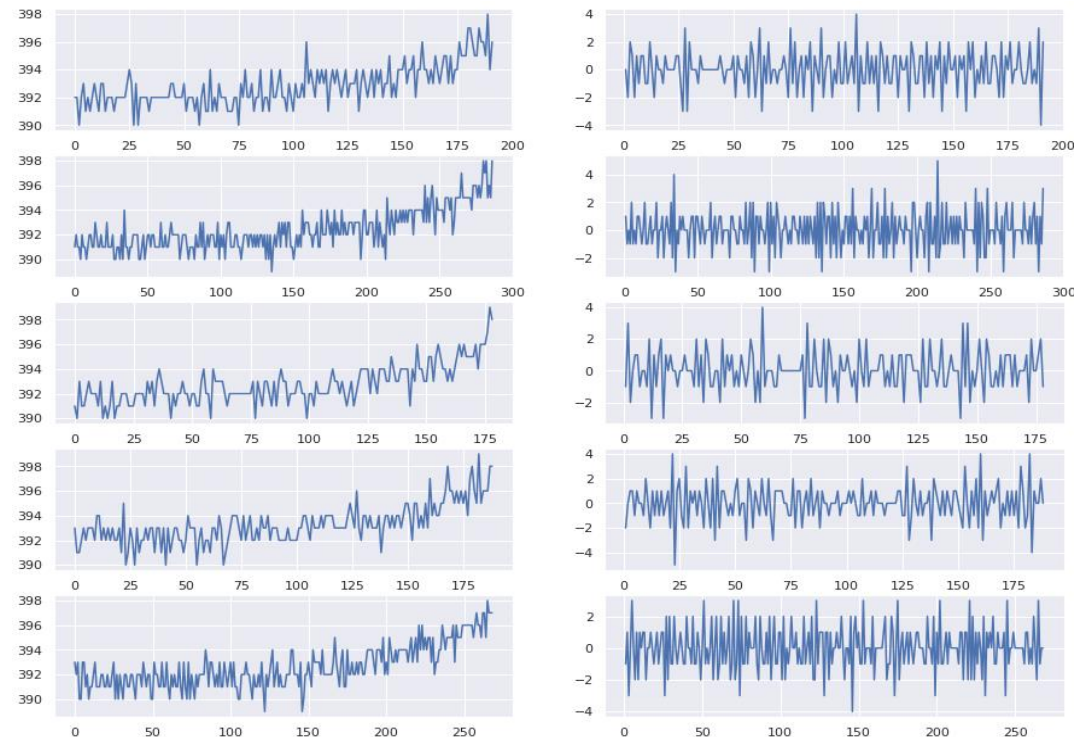
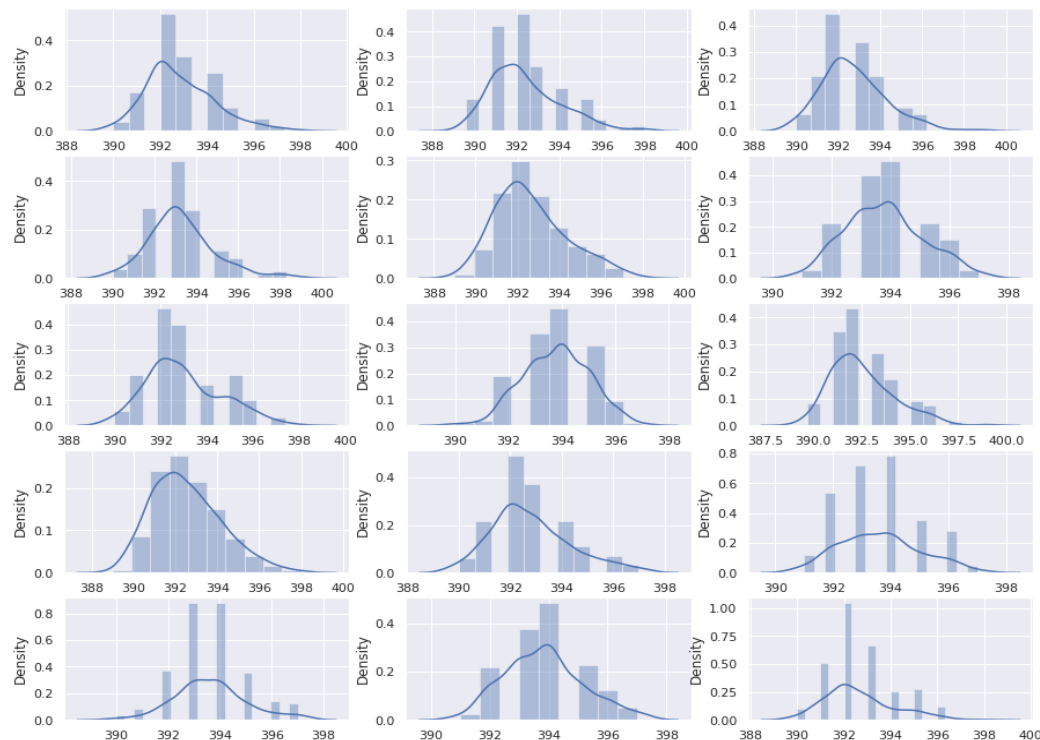
Asset id: 69

Asset id: 69





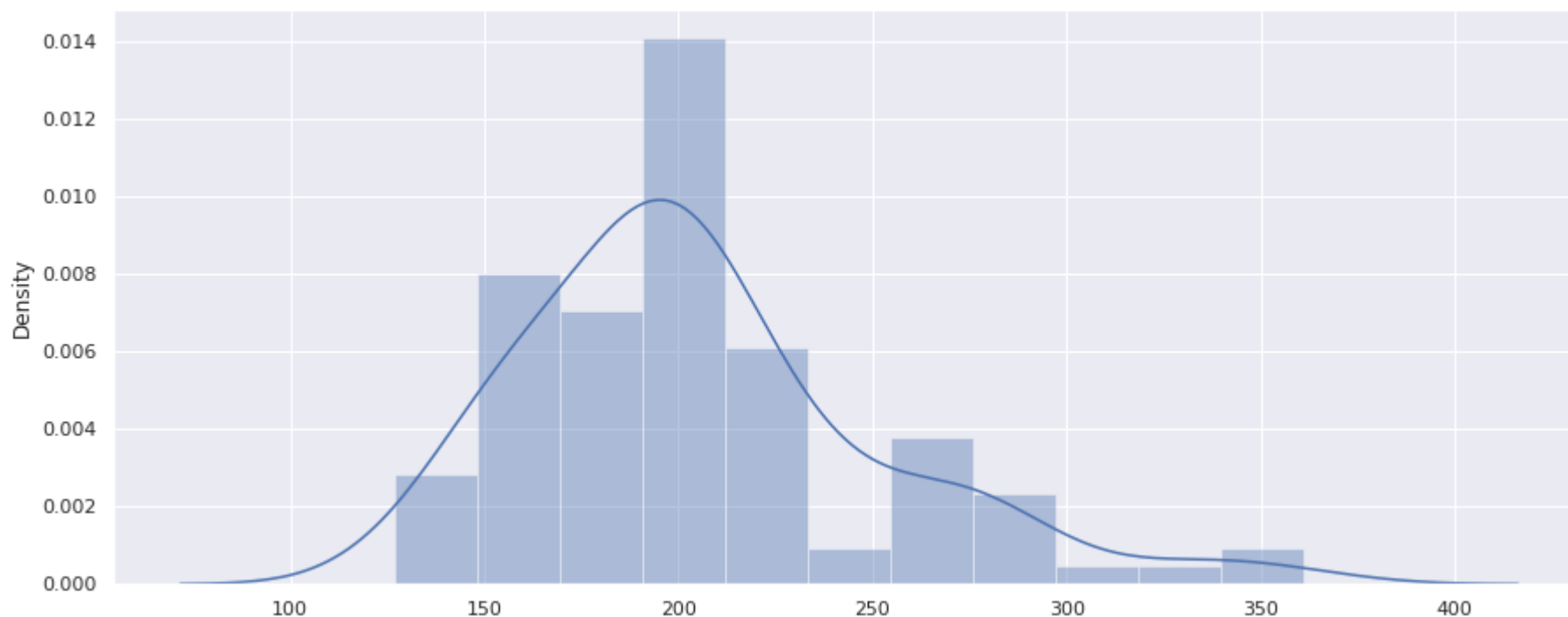
Olhando os gráficos de histograma, É possível ver que Tag17 tem um comportamento diferente dos outros.



Tag 17' mostra um comportamento de 'stepping', onde todos os passos são sempre inteiros completos. Esta propriedade sugere uma quantização de inteiro completo.

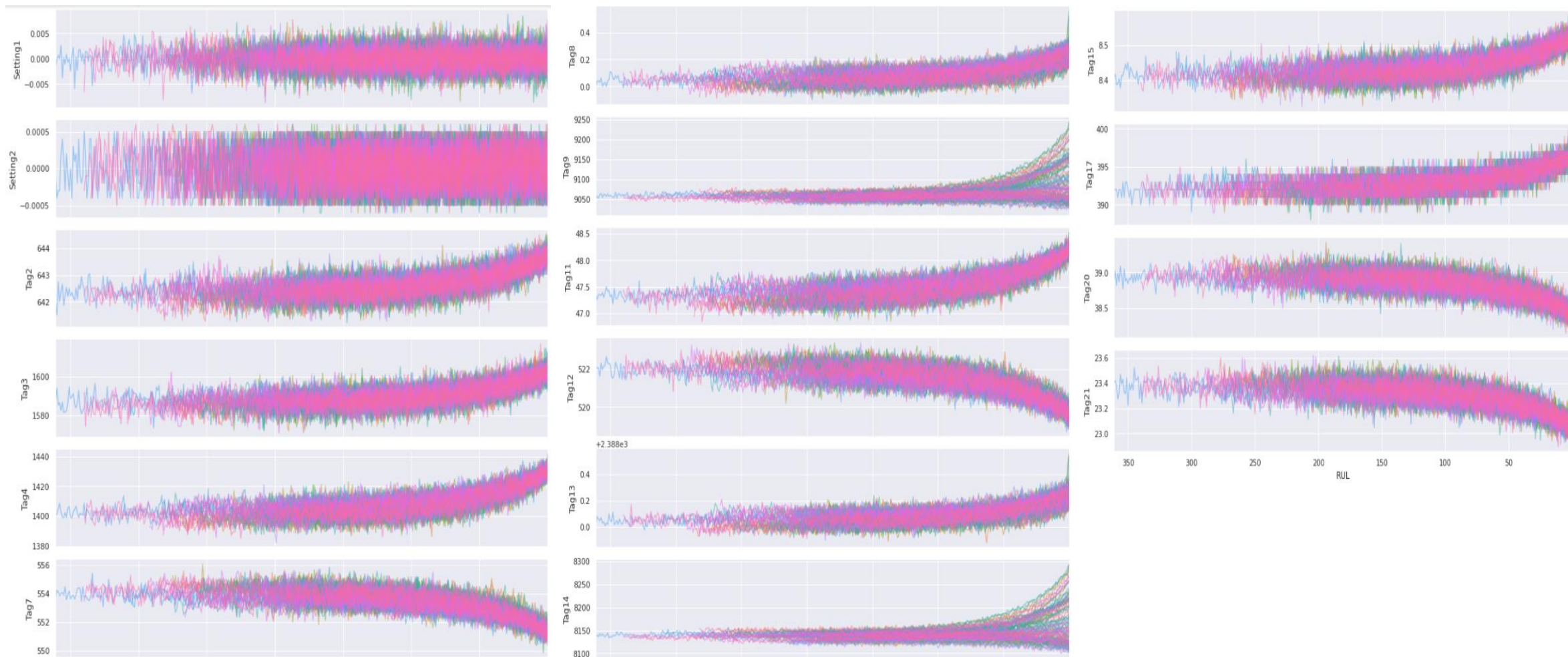
# Visualizando os tempos de execução individuais para cada unidade

É possível ver que os tempos de operação máximo e mínimo através das unidades variam de cerca de 125 ciclos a cerca de 350 ciclos, o que é cerca de 280% do mínimo a máximo. Deve ser seguro dizer que há uma variação significativa entre o RUL de unidades diferentes. No entanto, não se sabe se o estado inicial representa o estado de desgaste no comissionamento, ou algum tipo de estado de desgaste ao longo da vida útil das unidade.



# Visualizando a transição de estado combinada para todas as unidades

parece que todas as unidades têm variações comparáveis ao longo de toda a vida operacional.



# 02

## Modelos de Predição

# Modelo de Predição Binária Utilizando Uma Rede LSTM

A primeira camada é uma camada LSTM com 100 unidades seguida por outra camada LSTM com 50 unidades. A exclusão também é aplicada após cada camada LSTM para controlar o sobreajuste. A camada final é uma camada de saída Densa com unidade única e ativação sigmóide, uma vez que este é um problema de classificação binária.

```
Model: "sequential"
```

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 50, 100)	50400
dropout (Dropout)	(None, 50, 100)	0
lstm_1 (LSTM)	(None, 50)	30200
dropout_1 (Dropout)	(None, 50)	0
dense (Dense)	(None, 1)	51

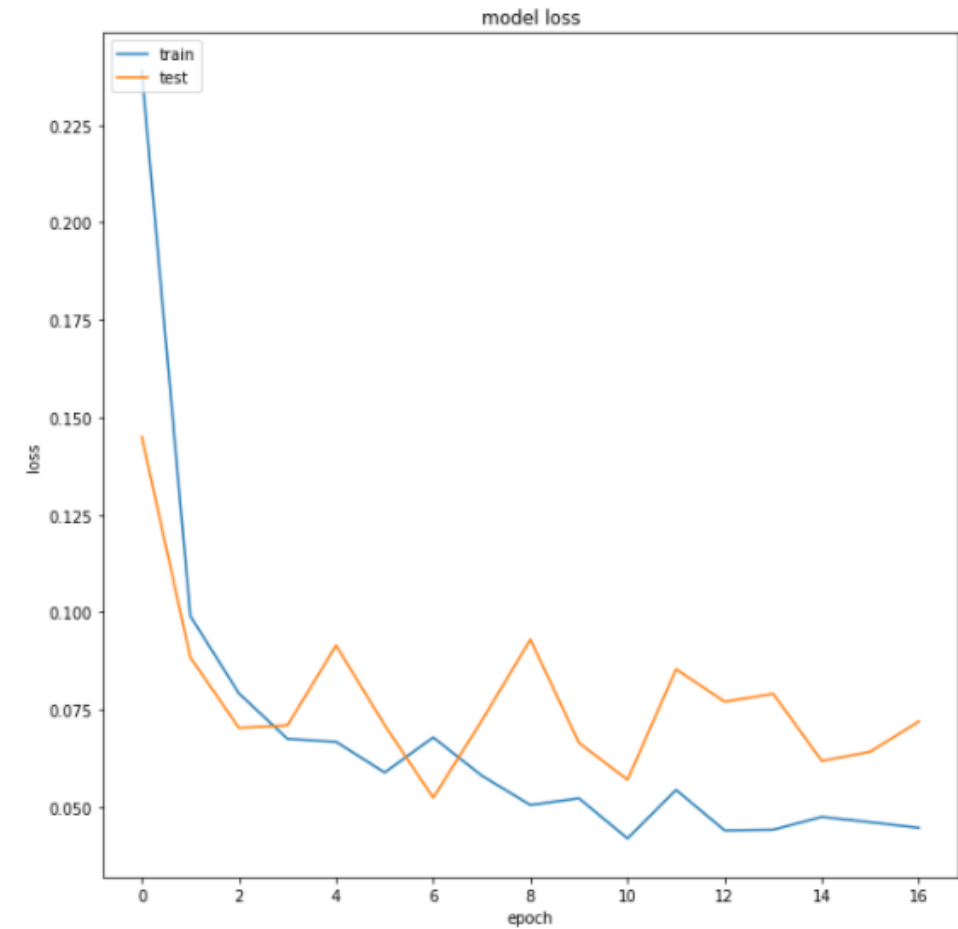
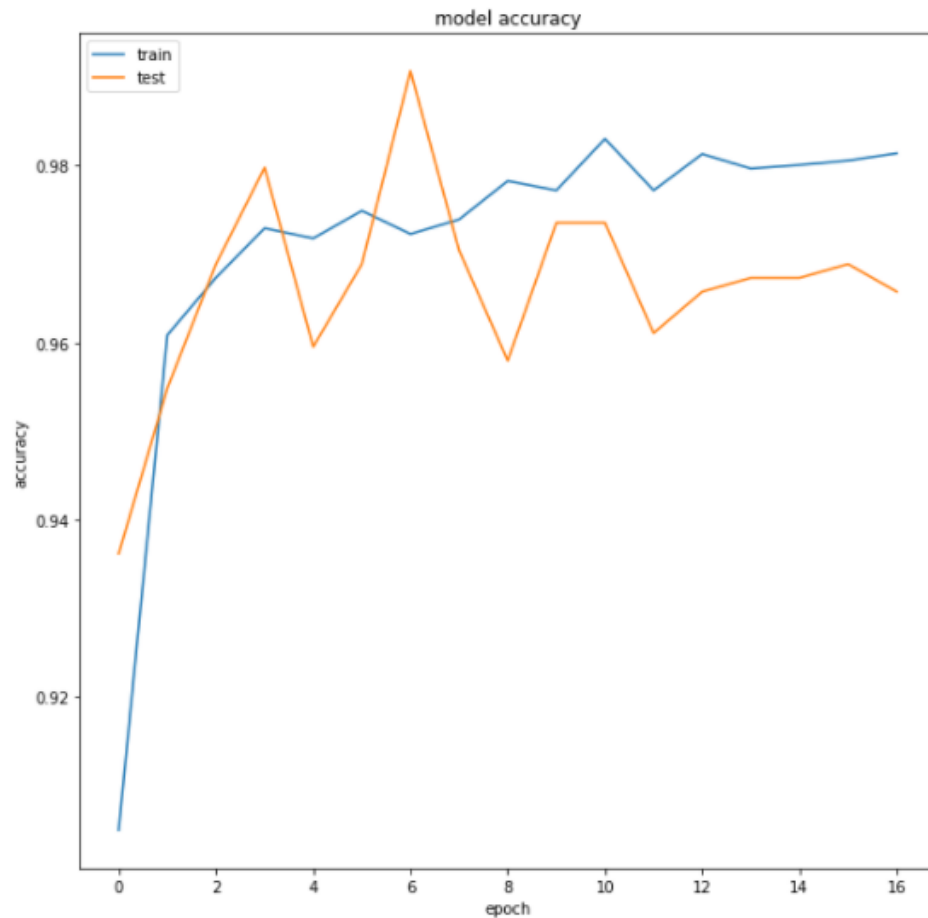
```
Total params: 80,651  
Trainable params: 80,651  
Non-trainable params: 0
```

Par o conjunto de treino foi obtido uma acurácia de 98%

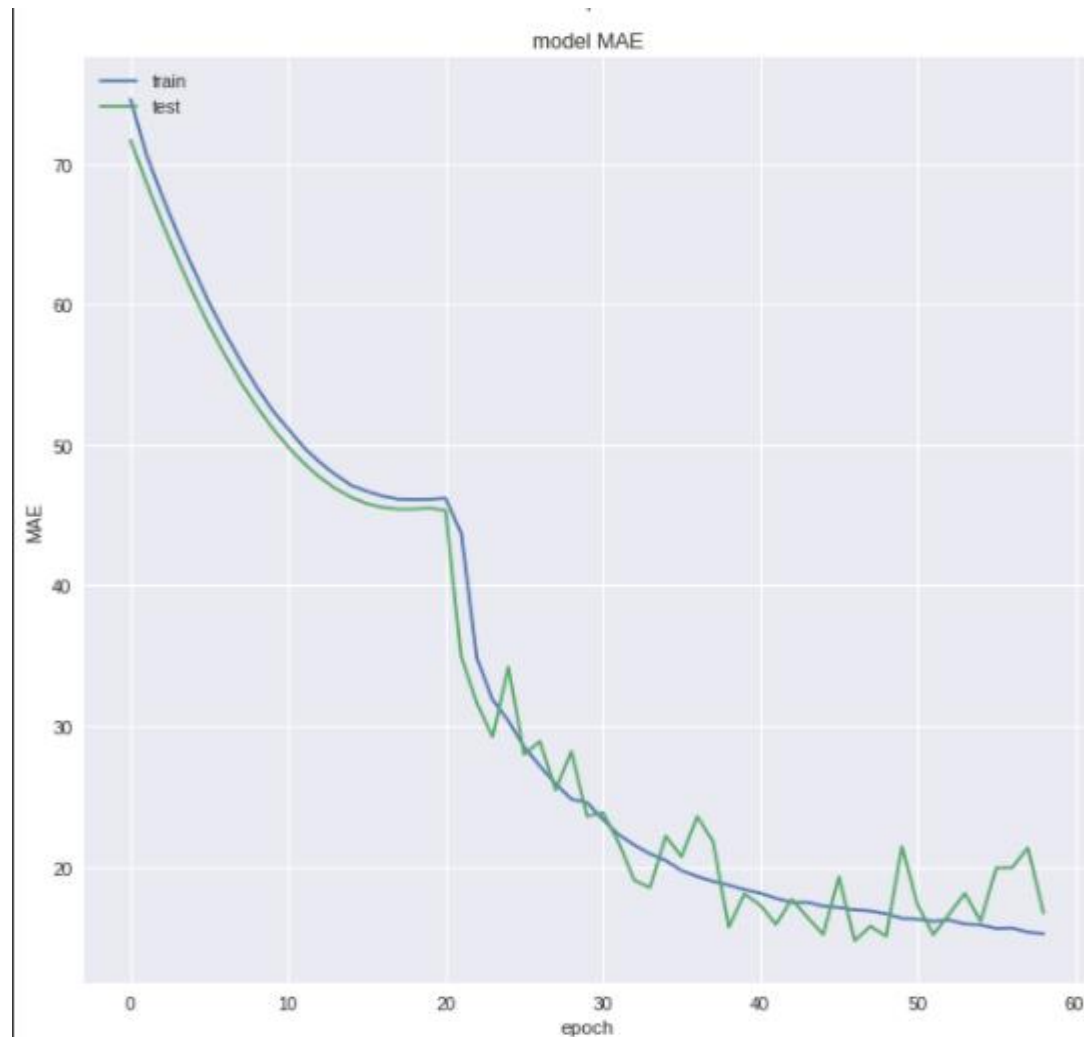
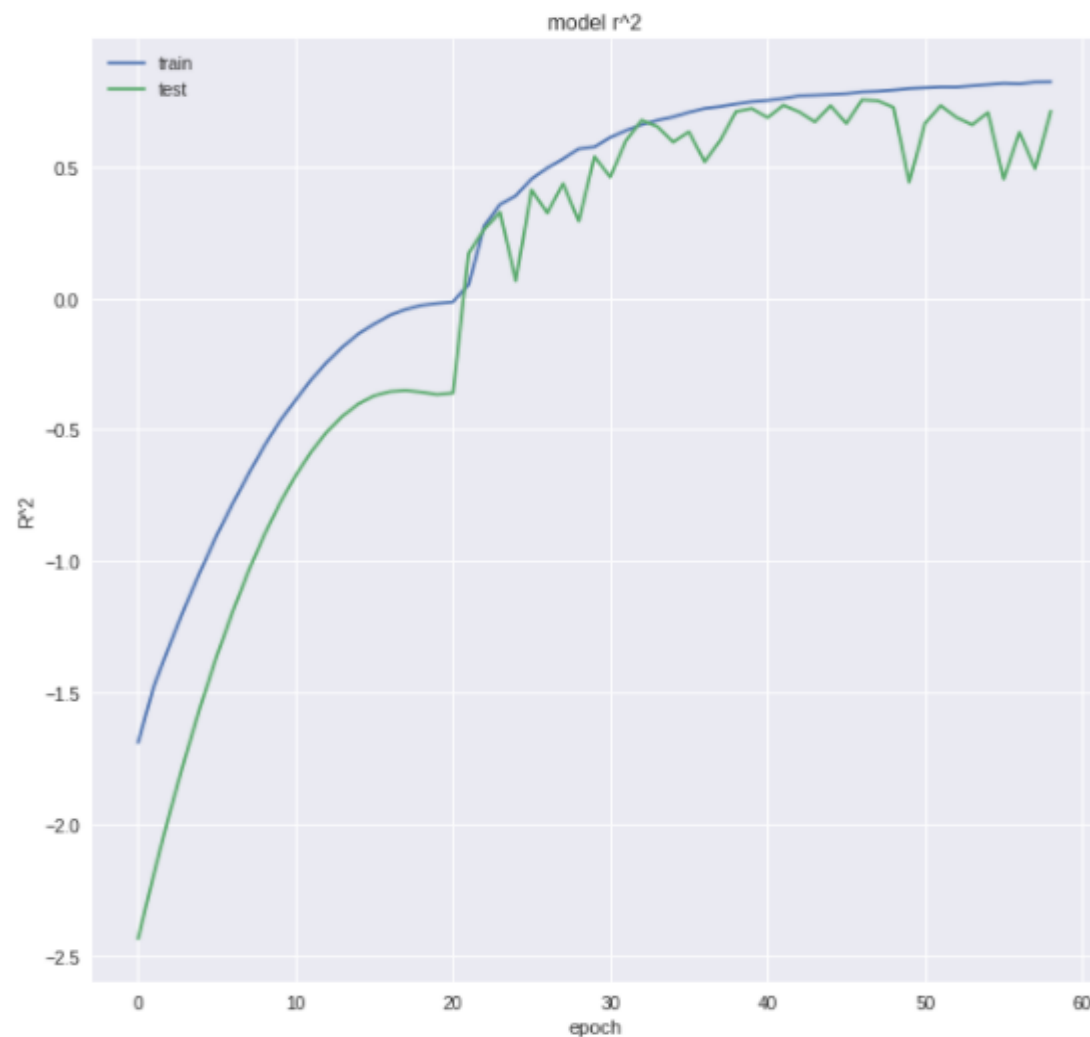


# Modelo de Regressão Utilizando Uma Rede LSTM

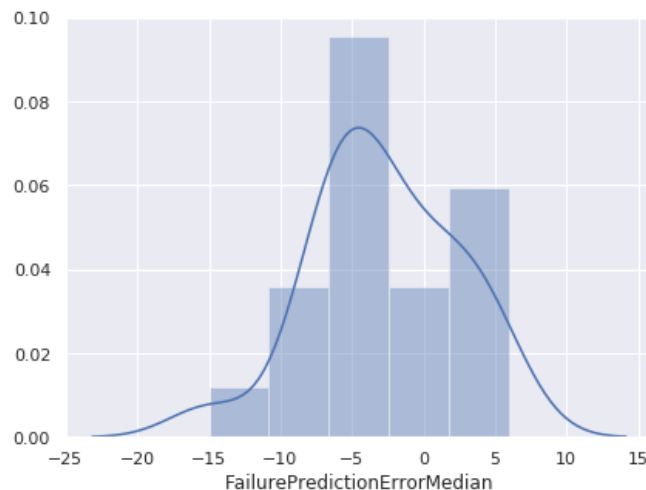
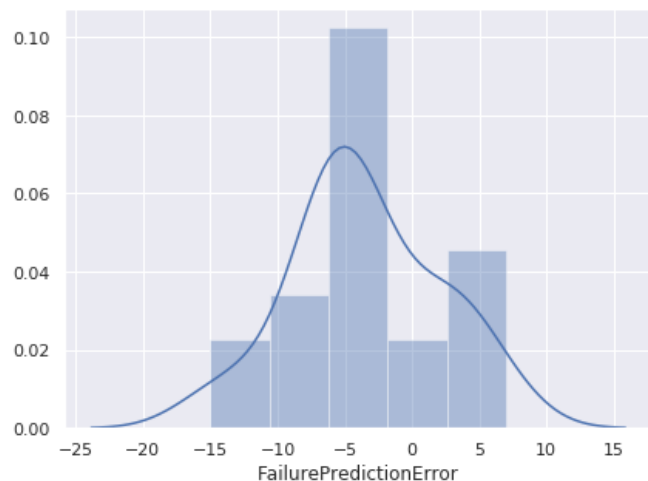
Pra a amostra de teste foi obtido uma precisão de 91%



Para a amostra de teste foi obtido um MAE de 13.99 e  $R^2$  de 0.80

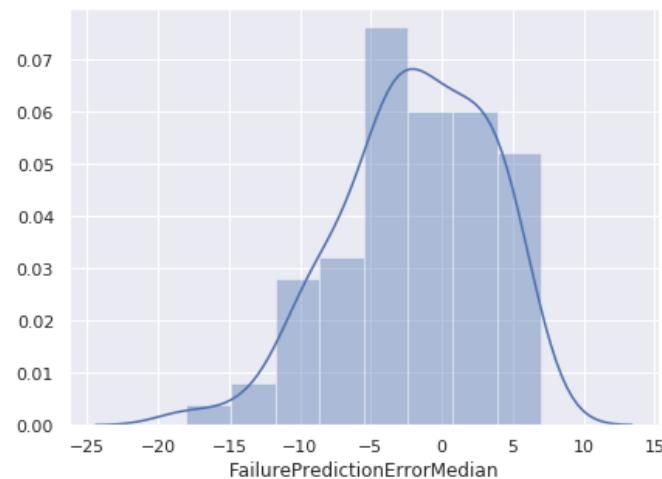
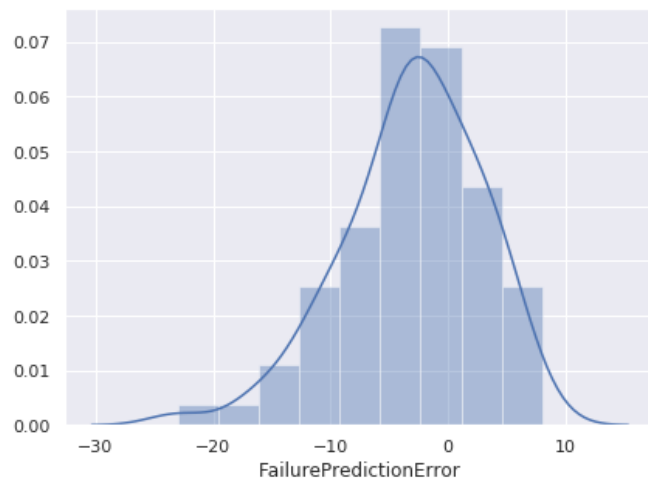


# Analizando se as falhas da unidade são previstas corretamente para o conjunto de dados



Números  $> 0$  predizem um colapso mais tarde do que se pensava, enquanto Números  $< 0$  estão predizendo um colapso antes do que se pensava. Os valores que preveem o colapso muito cedo são menos problemáticos do que os valores que preveem um colapso muito tarde.

Geralmente, parece que mais de 90% das falhas são detectadas pelo menos 1 ciclo ou mais antes do tempo. Este é, do ponto de vista do ML, apenas um resultado ok, o que não é surpreendente considerando o fato de que foi usado SLTM para regressão sem tratamento.



Podemos fazer melhor, mas deve-se notar que prever a tempo uma falha para mais de 90% das unidades já é uma coisa muito prática quando se trata de construir sistemas de monitoramento de condição real. Especialmente considerando o fato de que o modelo SLTM fornecido pode ser implementado com muita facilidade.

