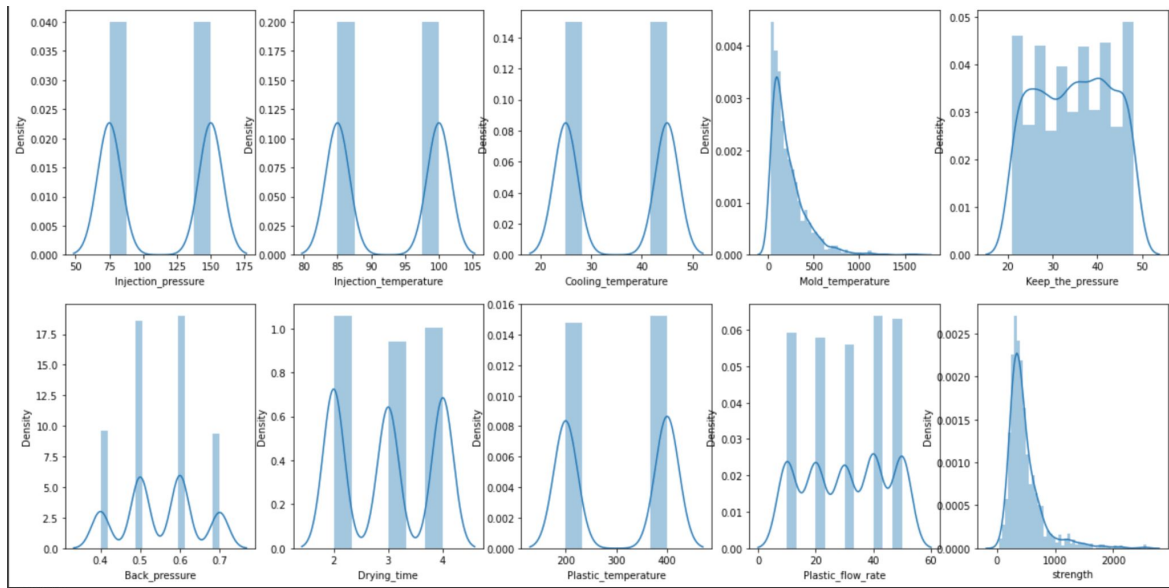


Injection Model with RL

ChrisLo august. 2022

Injection Data Preview



類別變數：

Injection_pressure
Injection_temperature
Cooling_temperature
Keep_the_pressure
Back_pressure
Drying_time
Plastic_temperature
Plastic_flow_rate

連續變數：

Mold_temperature
strength

操作參數分成：類別變數與連續型變數

Injection Data Preprocessing

類別變數：

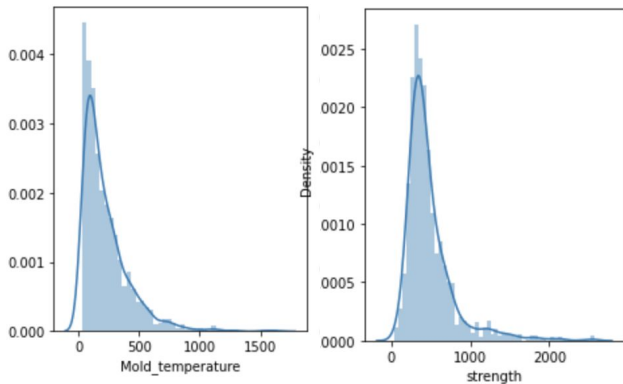
Injection pressure
Injection temperature Cooling_temperature
Keep the pressure
Back pressure
Drying time
Plastic temperature
Plastic_flow_rate

OneHot Encoding

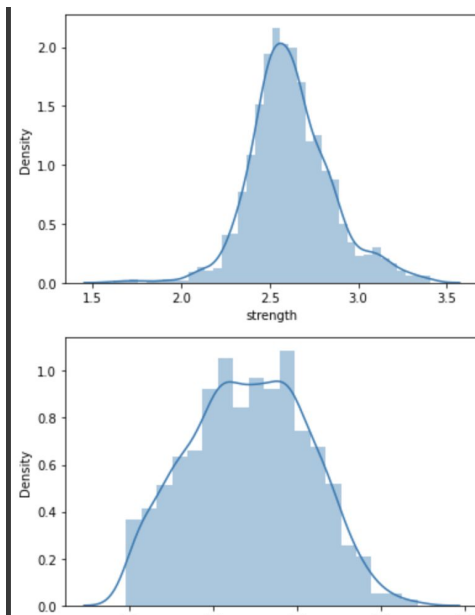
[1,0,0]
[0,1,0]
[0,0,1]

連續變數：

Mold_temperature
strength



取log10



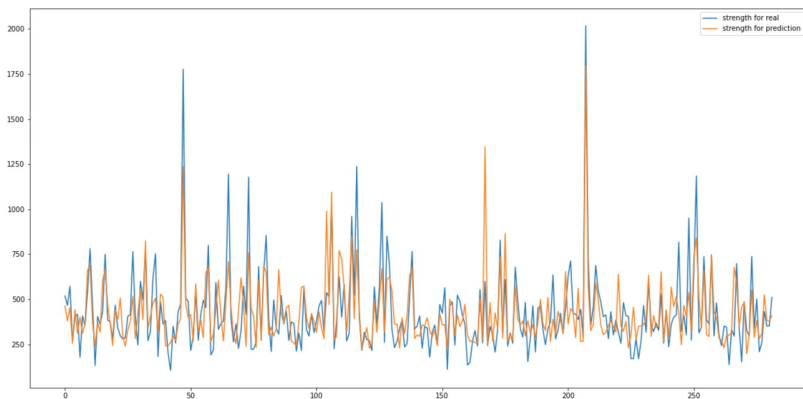
Supervise Learning Method - AutoML

AutoML 提供各種監督學習演算法

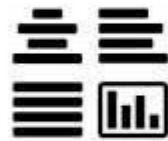
包含:

XGBoost, GLM, DRF, DeepLearning ...

且可以選擇ensemble方法選出最佳組合模型

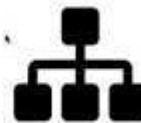


Aspects of Automatic Machine Learning



Data Prep

Model
Generation



Ensembles

RL Environment(State) Definition

State :

t1時間: 參數與 strength
zero padding
zero padding
zero padding
zero padding

t1時間: 參數與 strength
t2時間: 參數與 strength
zero padding
zero padding
zero padding

...

t1時間-參數與 strength
t2時間-參數與 strength
t3時間-參數與 strength
t4時間-參數與 strength
t5時間-參數與 strength

取近5次的數據當作 State

RL Environment(Action) Definition

類別變數可動作範圍

Injection_pressure values : [75 150]

Injection_temperature values : [85 100]

Cooling_temperature values : [25 45]

Keep_the_pressure values : [21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37
38 39 40 41 42 43 44 45 46 47 48]

Back_pressure values : [0.4 0.5 0.6 0.7]

Drying_time values : [2 3 4]

Plastic_temperature values : [200 400]

Plastic_flow_rate values : [10 20 30 40 50]

連續變數 Mold_temperature 可動作範圍

-25%, +25%

RL Environment(Reward) Definition

$$r(s_t, a_t) = -(\alpha \|\Delta_t\|_2 + \beta t^2)$$

$\alpha, \beta = 0.5$

t 為執行次數

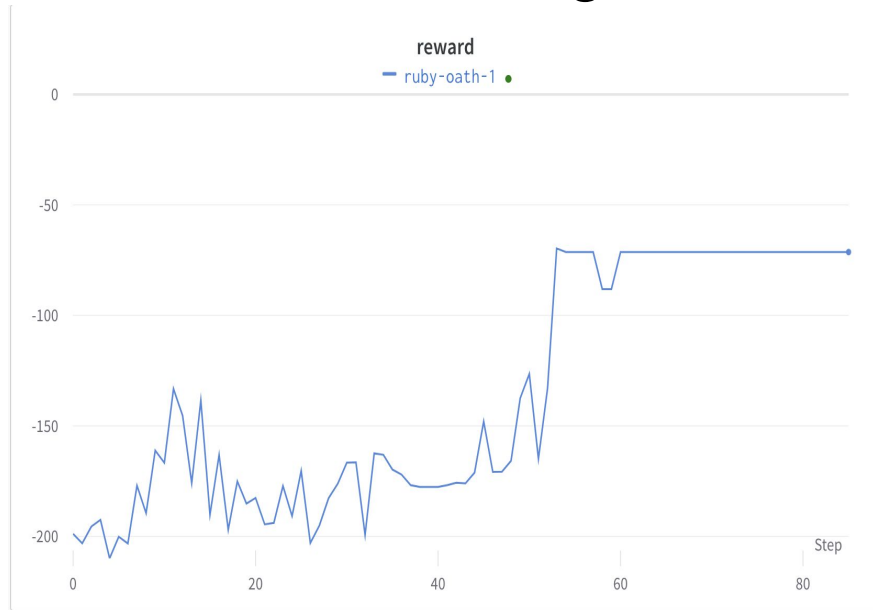
delta_t 為互動後strength 與target strength 的值

Actor Critic Training

- 1.以 injection_process_old.csv 第一筆資料當初始參數
- 2.每批次與Injection Model 互動20次, 存入buffer並訓練
- 3.大約訓練540次後, 只需三步即可接近我們要的Strength(500)

Episode 390	Average reward: -177.58	reward history:	[-110.77234656410863, -102.12056368158562, -99.36490576023121, -111.17878768476362, -124.54424556078132, -123.17878768476362, -138.5442,
Episode 400	Average reward: -177.58	reward history:	[-110.77234656410863, -102.12056368158562, -99.36490576023121, -111.17878768476362, -124.54424556078132, -123.17878768476362, -138.5442,
Episode 410	Average reward: -176.78	reward history:	[-110.77234656410863, -102.12056368158562, -99.36490576023121, -111.17878768476362, -109.36490576023121, -123.17878768476362, -138.5442,
Episode 420	Average reward: -175.66	reward history:	[-110.77234656410863, -96.00737464568815, -99.36490576023121, -111.17878768476362, -109.36490576023121, -123.17878768476362, -138.5442,
Episode 430	Average reward: -175.96	reward history:	[-110.77234656410863, -96.00737464568815, -99.36490576023121, -109.04033919381729, -109.36490576023121, -131.04424556078132, -123.36490,
Episode 440	Average reward: -171.11	reward history:	[-110.77234656410863, -96.00737464568815, -99.36490576023121, -111.82387393737193, -109.36490576023121, -110.39502444970998, -122.75744,
Episode 450	Average reward: -147.95	reward history:	[-110.77234656410863, -102.12056368158562, -99.36490576023121, -108.51012378708046, -109.36490576023121, -114.73738220114274, -113.4124,
Episode 460	Average reward: -170.78	reward history:	[-110.77234656410863, -102.12056368158562, -99.36490576023121, -110.12056368158562, -109.36490576023121, -121.04033919381729, -123.3649,
Episode 470	Average reward: -170.70	reward history:	[-110.77234656410863, -102.12056368158562, -99.36490576023121, -110.12056368158562, -115.57206811355553, -115.64581762745689, -127.5687,
Episode 480	Average reward: -165.76	reward history:	[-110.77234656410863, -102.12056368158562, -99.36490576023121, -98.6664225791072, -105.06212079863866, -113.98743110863356, -118.166422,
Episode 490	Average reward: -137.47	reward history:	[-139.33167768336847, -102.12056368158562, -99.36490576023121, -108.83183120968476, -67.20662782686696, -78.00791883036823, -85.4895009,
Episode 500	Average reward: -126.46	reward history:	[-110.77234656410863, -102.12056368158562, -99.2393574400237, -99.13727773803996, -58.354740792190114, -66.23313214036028, -67.59942965,
Episode 510	Average reward: -164.79	reward history:	[-110.77234656410863, -102.12056368158562, -61.507918830368226, -65.98950092847988, -109.36490576023121, -123.82387393737193, -123.3649,
Episode 520	Average reward: -133.04	reward history:	[-110.77234656410863, -102.12056368158562, -61.507918830368226, -65.98950092847988, -70.90889097372781, -75.44152395387476, -84.9088909,
Episode 530	Average reward: -69.68	reward history:	[-110.77234656410863, -58.371034211349354, -61.507918830368226, -48.0594011616607]
Episode 540	Average reward: -71.33	reward history:	[-110.77234656410863, -58.371034211349354, -44.833622249542316]
Episode 550	Average reward: -71.33	reward history:	[-110.77234656410863, -58.371034211349354, -44.833622249542316]
Episode 560	Average reward: -71.33	reward history:	[-110.77234656410863, -58.371034211349354, -44.833622249542316]
Episode 570	Average reward: -71.33	reward history:	[-110.77234656410863, -58.371034211349354, -44.833622249542316]
Episode 580	Average reward: -88.08	reward history:	[-110.77234656410863, -108.62035489370575, -44.833622249542316]
Episode 590	Average reward: -88.08	reward history:	[-110.77234656410863, -108.62035489370575, -44.833622249542316]

Actor Critic Training Result



note

左圖Step 為每互動20次存取平均
Reward