Team 03

鄭皓謙 林宗憲

Reference:

2015, ACM, Predicting Winning Price in Real Time Bidding with Censored Data
2018, ACM, Deep Censored Learning of the Winning Price in the Real Time Bidding

Data preprocessing

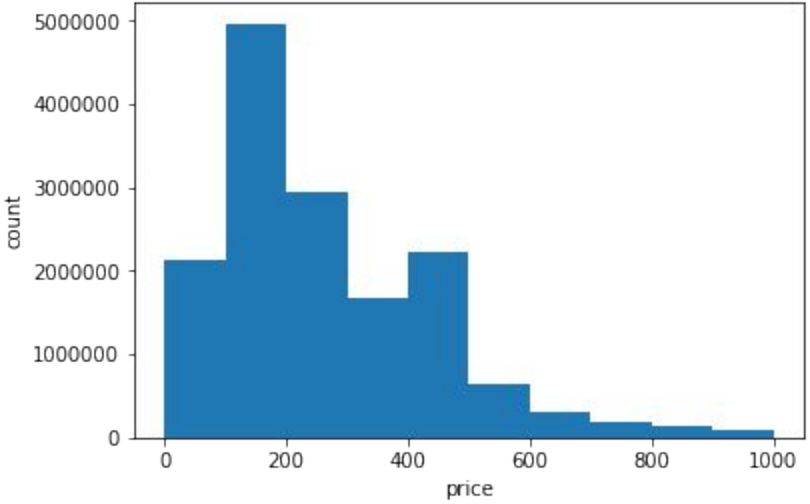
1. Artificial 10% noise.

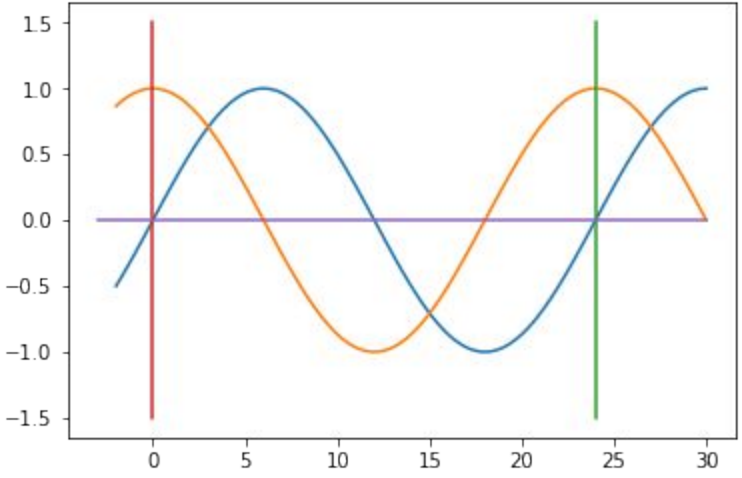
app_type imp_type imp_position device_type

2. impreq_time : weekdays, hours, mins
3. price distribution

z = x/(24*60*2π)

time_1 = sin(x), time_2 = sin(x+π/2)





Input

width

height

app

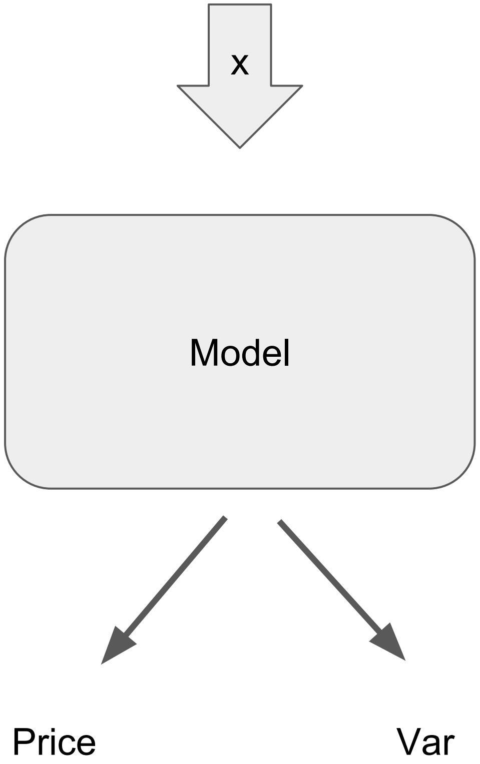
device

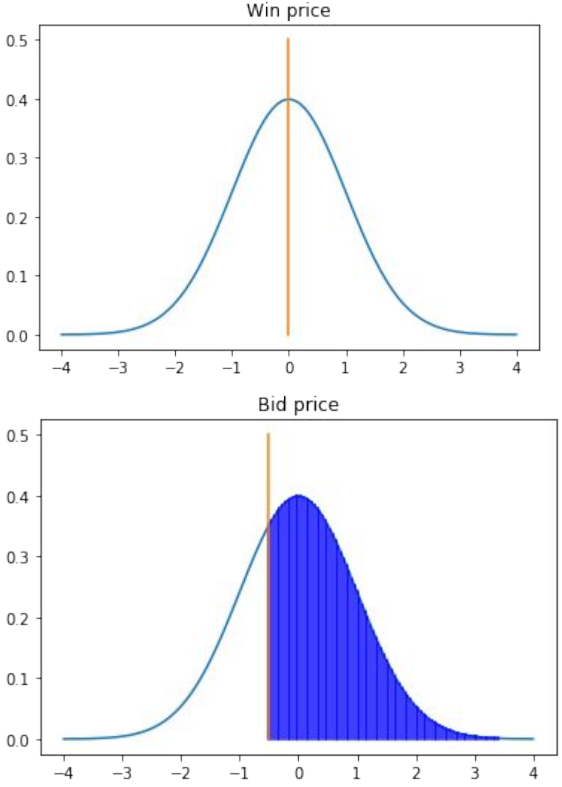
imp_type

imp_position

weekdays

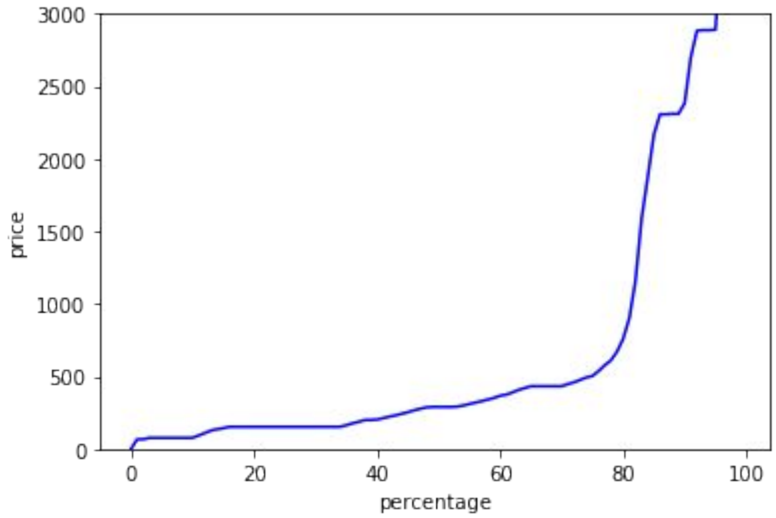
mins





$$\sum_{i\in\mathcal{W}} \log f_{x_i}(v_i - g(x_i|\beta)|\Theta) + \sum_{i\in\mathcal{L}} \log(1 - F_{x_i}(l_i - g(x_i|\beta)|\Theta), \tag{5}$$

Result



	80% win data	80% censored data	off 5% two-tails, win data
Normal log likelihood	875.850517	846.294398	846.158235
Gumbel log likelihood	896.601951	863.916028	/
MSE	877.291040	/	850.321041

Result

1. trim two-tails 5%
2. only win data
3. with normal distribution, neg log likelihood
4. public score : 846.158

$$\sum_{i \in \mathcal{W}} \log f_{x_i}(v_i - g(x_i|\beta)|\Theta)$$

Observations

lower weight on bid data -> higher public score

Using missing data ?

Different threshold between win data bid data.