

# 数据分析与可视化

## 1. favorite\_count 与 retweet\_count 比较

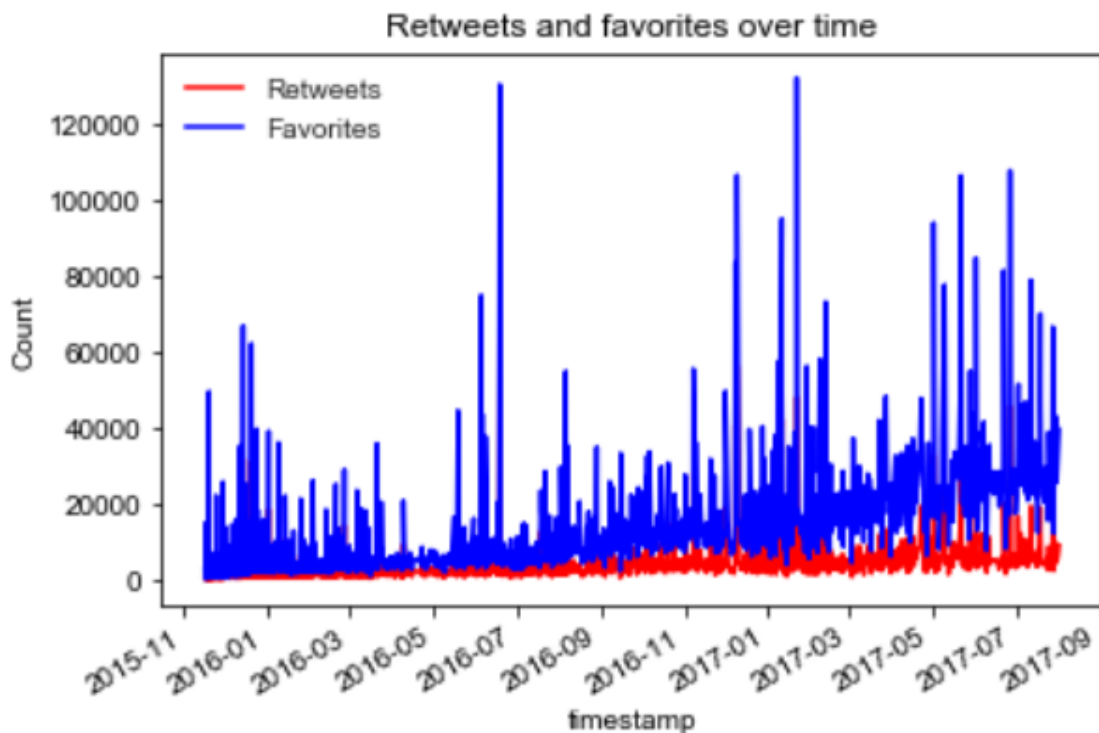
- 分析 favorite\_count 和 retweet\_count。

	favorite_count	retweet_count
mean	8923	2770
std	12400	4715
min	81	15
25%	1972	622
50%	4117	1348
75%	11275	3202
max	132318	79116

结论：分析 favorite\_count 和 retweet\_count 可知 favorite\_count 的平均值为 8923，retweet\_count 的平均值为 2770，可知相较于评论，人们更喜欢点赞。

- 可视化 favorite\_count 和 retweet\_count 随时间的变化。

将时间分组，可视化 favorite\_count 和 retweet\_count 随时间的变化。



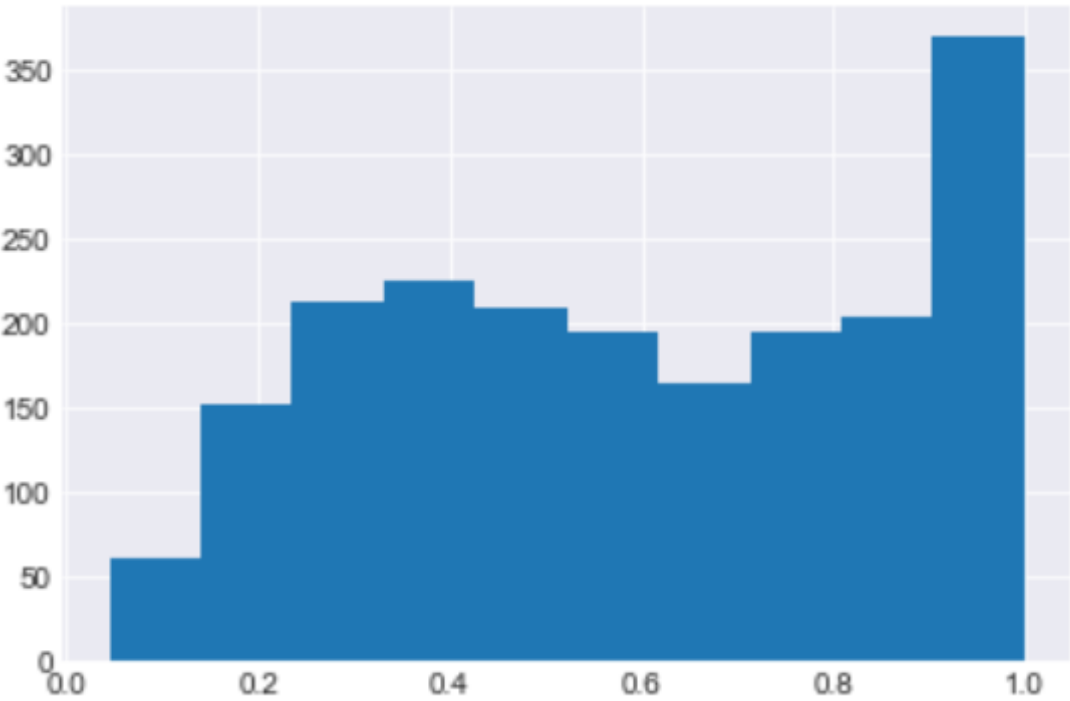
结论：根据 favorite\_count 和 retweet\_count 随着时间的变化的趋势图，可知随着时间的变化，人们点赞和评论都有一定的波动，点赞的波动大于评论的波动，总体来说，随着时间的变化，点赞数高于评论数。

## 2. 图像预测算法的可信度比较

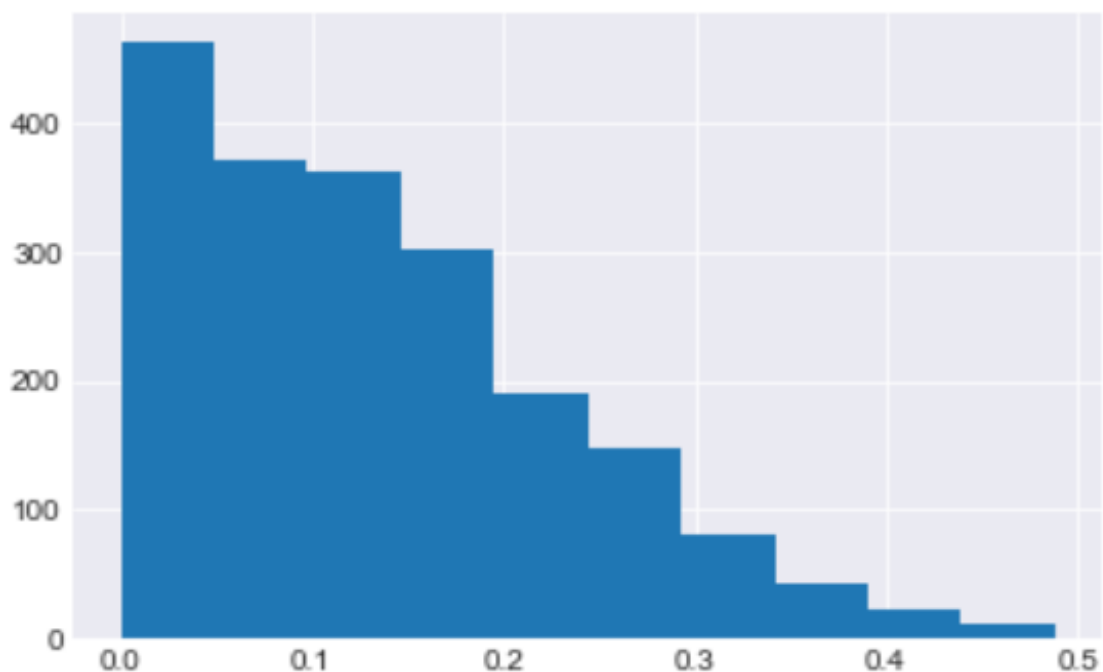
- 分析 p1\_conf, p2\_conf 和 p3\_conf。

	p1_conf	p2_conf	p3_conf
mean	0.593941	1.344195e-01	6.024848e-02
std	0.271954	1.006807e-01	5.089067e-02
min	0.044333	1.011300e-08	1.740170e-10
25%	0.362857	5.393988e-02	1.619283e-02
50%	0.587635	1.174550e-01	4.950530e-02
75%	0.846285	1.951377e-01	9.159438e-02
max	1.000000	4.880140e-01	2.734190e-01

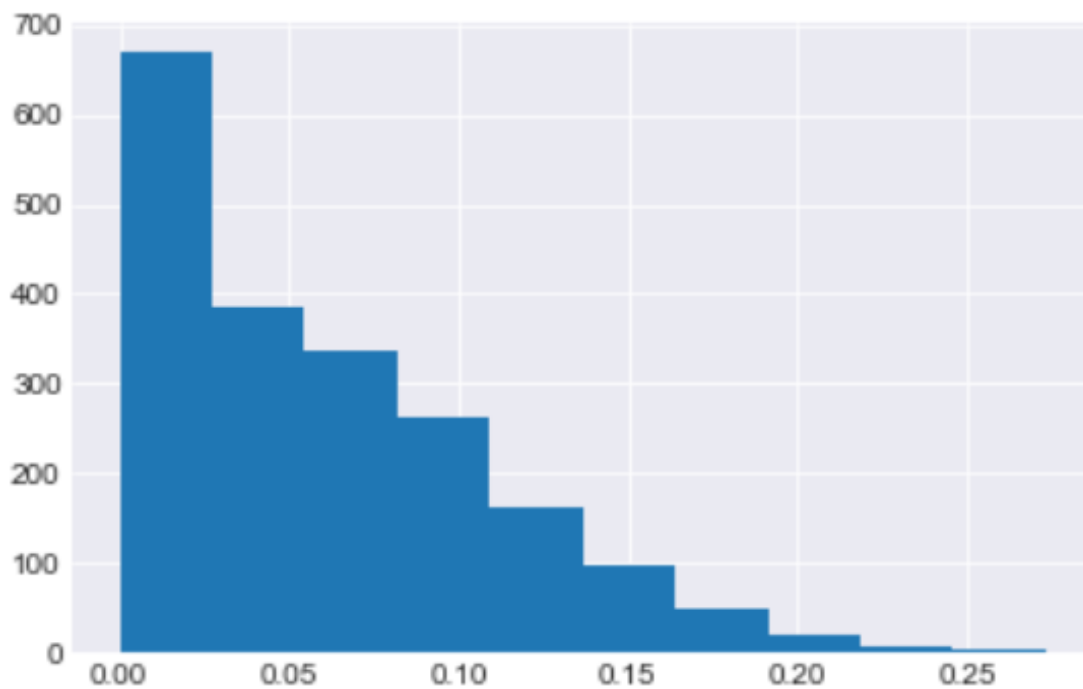
- p1\_conf 直方图：



● P2\_conf 直方图:



● P3\_conf 直方图:

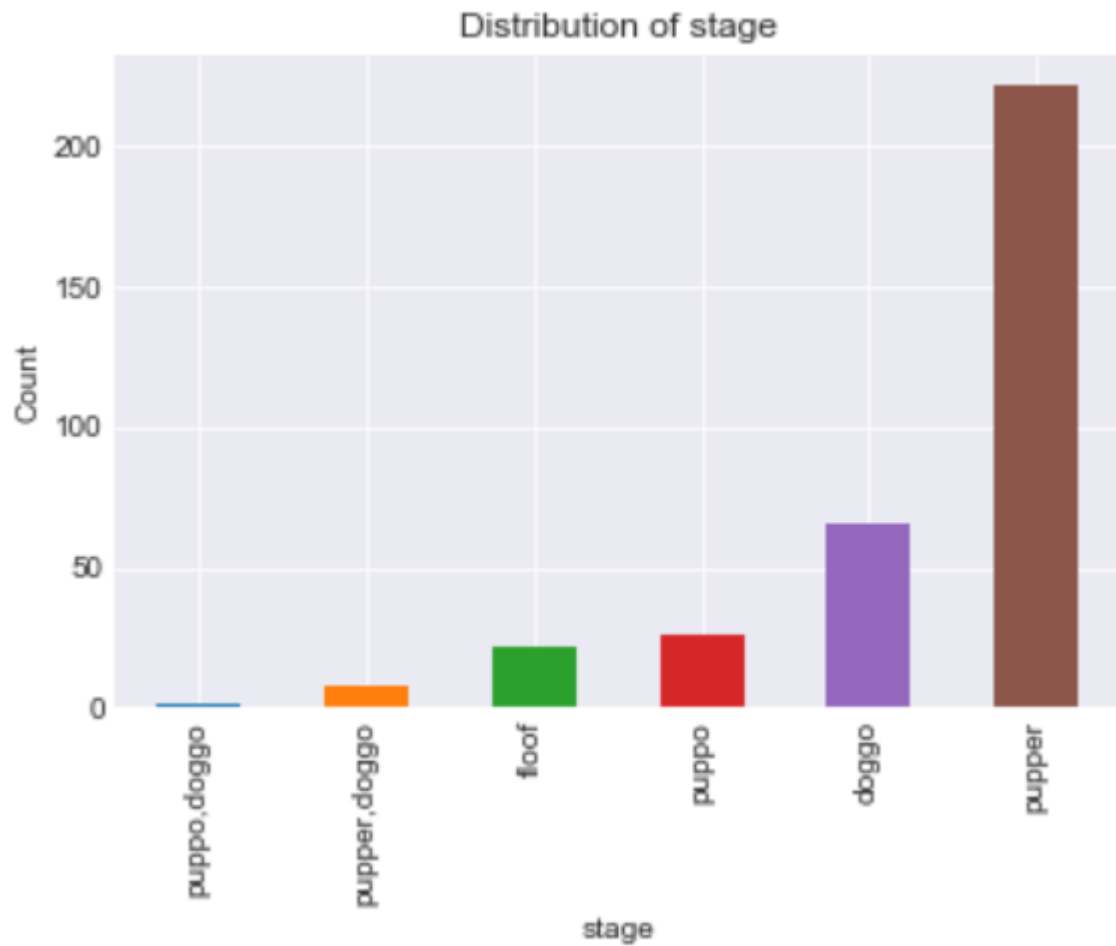


结论: 分析 p1\_conf, p2\_conf, p3\_conf 可知 1 号预测的可信度 > 2 号预测的可信度 > 3 号预测的可信度, 2 号预测和 3 号预测的可信度均小于 0.5, 1 号预测可信度的均值为 0.59, 可信度相对来说还是不高, 因此图形预测算法还有很大的提

升空间。

### 3. 狗狗的评价地位分析

- 可视化狗狗评价地位，按出现的次数顺序绘制柱状图。



结论：由上图可知狗狗的评价地位使用最多的是 pupper,其次是 doggo，再次是 puppo，查阅狗狗字典可知，大多数参与评价的狗狗年纪较轻。