

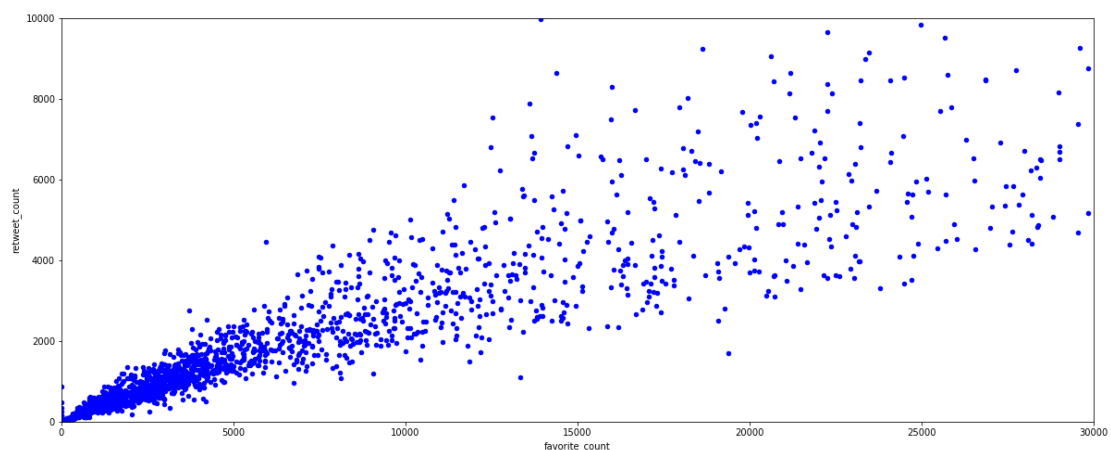
Twitter 上狗狗的信息整理——分析报告

郑浩 2019 年 12 月 3 日 北京

我们从网上得到了一些关于狗狗的信息，这些信息来自推特的内部档案，推特博文和机器学习预测的图片信息。通过整理我们得到了一些有趣的结论。

1. 人们对自己喜欢的狗狗更倾向于将他们转发

通过将喜欢数和转发数制成散点图，我们大体看到，这两者的数据存在一定的正比例的线性关系，相关系数近似于 1。



2. 我们试图找到最受欢迎的狗狗的品种或名字，他应该是金毛犬，可惜没有名字

我们从一些评分数据和喜欢数据中选择了最高的账户，但评分数据由于不规范，参考价值较低，因此我们更倾向于使用喜欢数，可以这个 ID 中没有狗狗的名字，但有他的品种和成长阶段。

3. 这些数据中，金毛犬的数量是最多的，狗狗的成长阶段中幼犬的比例最大

通过统计所有狗狗品种的数量，并将其进行排名，我们发现金毛犬是最多的。

通过对狗的四个阶段数量的统计，并制成饼状图，我们发现幼犬的比例是最大的。

