

act_report

[TOC]

项目背景

项目名称：
狗狗评分数据清理项目（wraggling adv dog_rage@twitter）

项目内容：
本项目要整理 (以及分析和可视化) 的数据集是推特用户 @dog_rates 的档案, 推特昵称为 WeRateDogs。
WeRateDogs 是一个推特主，他以诙谐幽默的方式对人们的宠物狗评分。这些评分通常以 10 作为分母。但是分子则一般大于 10：11/10、12/10、13/10 等等。

WeRateDogs 的推特档案包括 5000 多条推特的基本信息，但并不包括所有内容。不过档案中有一列包含每个推特的文本，我用这一列数据提取了评分、狗的名字和“地位”（即 doggo、floofer、pupper 和 puppo）——这使数据得以“完善”。在这 5000 多条中，我只筛选出了 2356 条包含评分的推特数据。

text	rating_numerator	rating_denominator	name	doggo	floofer	pupper	puppo
This is Phineas. He's a mystical boy. Only ever appears in the hole of a donut. 13/10 https://t.co/MgUWQ76dJU	13	10	Phineas	None	None	None	None
This is Tilly. She's just checking pup on you. Hopes you're doing ok. If not, she's available for pats, snugs, boops, the whole bit. 13/10	13	10	Tilly	None	None	None	None
This is Archie. He is a rare Norwegian Pouncing Corgo. Lives in the tall grass. You never know when one may strike. 12/10 https://t.co/...	12	10	Archie	None	None	None	None
This is Darla. She commenced a snooze mid meal. 13/10 happens to the best of us https://t.co/tD36da7qLQ	13	10	Darla	None	None	None	None
This is Franklin. He would like you to stop calling him "cute." He is a very fierce shark and should be respected as such. 12/10 #BarkWeek	12	10	Franklin	None	None	None	None
Here we have a majestic great white breaching off South Africa's coast. Absolutely h*ckin breathtaking. 13/10 (IG: tucker_marlo) #BarkWeek	13	10	None	None	None	None	None
Meet Jax. He enjoys ice cream so much he gets nervous around it. 13/10 help Jax enjoy more things by clicking below https://t.co/Zr4hWfAs1H https://t.co/VJBRMnhxl	13	10	Jax	None	None	None	None
When you watch your owner call another dog a good boy but then they turn back to you and say you're a great boy. 13/10 https://t.co/...	13	10	None	None	None	None	None
This is Zoey. She doesn't want to be one of the scary sharks. Just wants to be a snugly pettable boatpet. 13/10 #BarkWeek https://t.co/...	13	10	Zoey	None	None	None	None
This is Cassie. She is a college pup. Studying international doggo communication and stick theory. 14/10 so elegant much sophisticated	14	10	Cassie	doggo	None	None	None
This is Koda. He is a South Australian deckshark. Deceptively deadly. Frighteningly majestic. 13/10 would risk a petting #BarkWeek https://t.co/...	13	10	Koda	None	None	None	None
This is Bruno. He is a service shark. Only gets out of the water to assist you. 13/10 terrifyingly good boy https://t.co/u1XPQMl29g	13	10	Bruno	None	None	None	None
Here's a puppo that seems to be on the fence about something haha no but seriously someone help her. 13/10 https://t.co/BxvuXk0UC	13	10	None	None	None	None	puppo
This is Ted. He does his best. Sometimes that's not enough. But it's ok. 12/10 would assist https://t.co/f8dEDcrKSR	12	10	Ted	None	None	None	None
This is Stuart. He's sporting his favorite fanny pack. Secretly filled with bones only. 13/10 puppered puppo #BarkWeek https://t.co/y70K...	13	10	Stuart	None	None	None	puppo

关于狗地位的评分解释是这样的：

THE DOGTIONARY

doggo
/'dɒɡo/
noun

- A big pupper, usually older. This label does not stop a doggo from behaving like a pupper.
- A pupper that appears to have its life in order. Probably understands taxes and whatnot.

"That's a really good doggo."
"I give my doggo a firm pat every night before bed."

pupper
/'pʌpə/
noun

- A small doggo, usually younger. Can be equally, if not more mature than some doggos.
- A doggo that is inexperienced, unfamiliar, or in any way unprepared for the responsibilities associated with being a doggo.

"H*ck, that's one pettable pupper."
"How many puppers could I fit on my body at once, if I were lying down?"

puppo
/'pʌpə/
noun

- A transitional phase between pupper and doggo. Easily understood as the dog equivalent of a teenager.
- A dog with a mixed bag of both pupper and doggo tendencies.

"My puppo is still learning what it takes to be a trustworthy doggo."
"I would hug that puppo so passionately."

blep
/'blep/
verb

- An extremely subtle act that occurs without the knowledge of the one who slips. The act includes one's tongue protruding ever so slightly from the mouth, usually just noticeable enough that it attracts the attention it deserves. Can last between three seconds and four days.

"My doggo did a h*ck of a blep the other day."
"Get a load of this blep I captured."

snoot
/'snu:t/
noun

- The nose of a dog. Usually found in places the dog may not fit. The location of the snoot may hint at where the dog's interest lies.

"That is a beautiful snoot."
"I've been trying to boop my neighbor's dog's snoot for six years."

floof
/'flʊf, 'flʊf/
noun

- Any dog really. However, this label is commonly given to dogs with seemingly excess fur. Comical amounts of fur on a dog will certainly earn the dog this generic name.
- Dog fur. The term holds true whether the fur is still on the dog, or if it has been shed off.

"Check out that majestic floof!"
"The floof on my dog has gotten out of control but I don't see anybody complaining any time soon."

- 貌似规则是这样的从年龄上分从年轻到老：pupper < puppo < doggo
- 剩下的3类，blep 为听话、sonnt 为嗅觉好、floof 为毛多

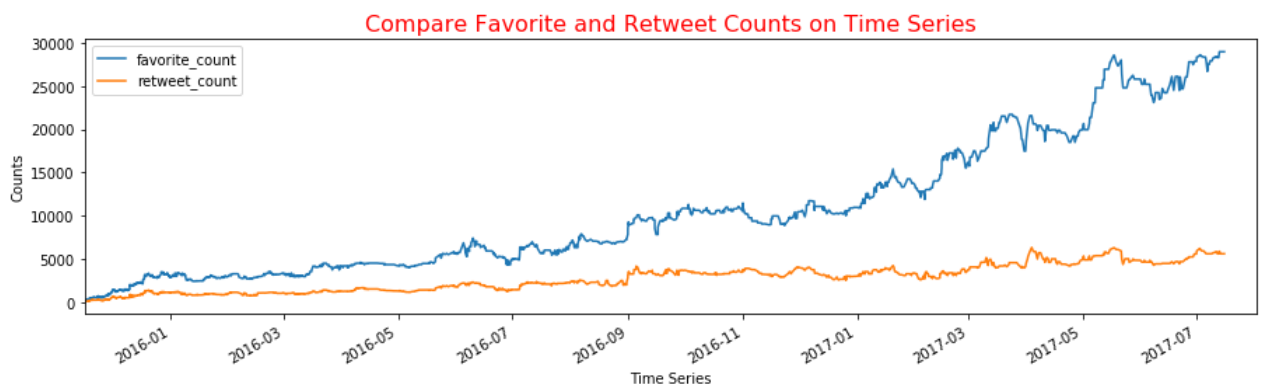
数据文件

- twitt_json.txt (读入为 df_api) 后收集来的数据，用于修正原始数据（提供了链接）
- twitter-archive-enhanced.txt (读入为 df_arc)原始数据
- image-predictions.tsv (读入为 df_img)图像预测结果（要求使用 request 库下载，[/LINK/](#)）
- 最终主数据结构：['tweet_id', 'rating_numerator', 'rating_denominator', 'name', 'retweet_count', 'favorite_count', 'display_text_range', 'p1', 'p1_conf', 'p1_dog', 'dogclass']

结论

/ favorite 数据在逐渐上涨的同时，retweet 数据已经区域饱和

- 2016年上半年之前, favorite 数量大概是 retweet 的两倍
- 但再这之后, favorite 数量大量上涨, retweet 数量上涨十分缓慢(两者之比达到6倍)
- 推测相关因素如下:
 - 可以看出 twitter 增长非常迅速(可惜缺少用户量相关的数据)
 - 但是人们愿意付出更多一点时间 retweet 的时间在减少, 可能原因是当人接触到更多的 twitter 信息后, 能够 retweet 的注意力已经没有什么增长空间了(注意力处于饱和状态)



/ @dog_rate 评论积极情绪为主导

- 对评论使用 word cloud 进行分析
- 第一幅图像为 twitter 英文字符(小鸟图不太美观)
- 第二幅图为了看清词，没有使用蒙版
- 可以看出积极的词汇和对狗狗描述的词占比很大，这就标识 @dog_rate 在评论狗狗时绝大部分都会非常友好，让狗的主人很舒服，也逐渐带来了人气（包括他家特殊的评分系统）

