

# Applied Memology

Дмитрий Сергеев  
ZeptoLab

# Предыстория вопроса

- Небольшой исследовательский проект

# Предыстория вопроса

- Небольшой исследовательский проект
- Предсказание популярности мема по его описанию

# Предыстория вопроса

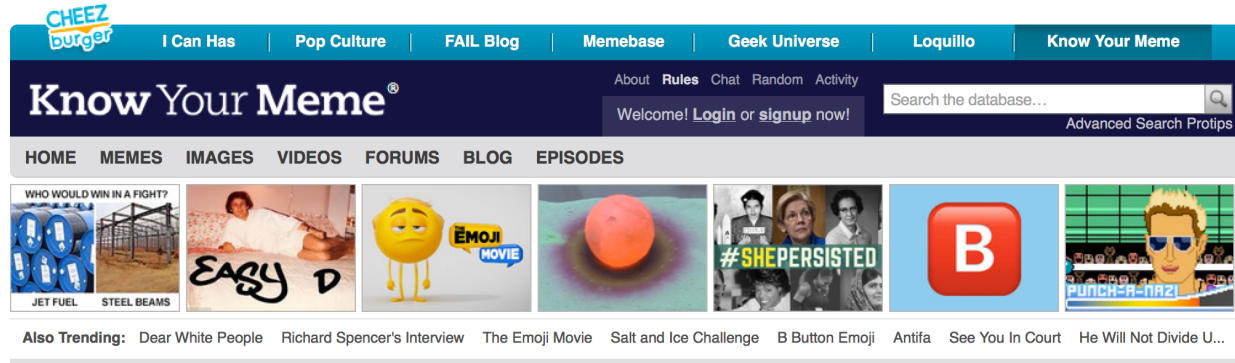
- Небольшой исследовательский проект
- Предсказание популярности мема по его описанию
- Зачем - «for research purposes»

# Предыстория вопроса

- Небольшой исследовательский проект
- Предсказание популярности мема по его описанию
- Зачем - «for research purposes»
- Надо собирать датасет



# Applied Memology



## All Entries

Browsing all 15,266 meme entries.

[Like us on Facebook!](#)

[+ Add an Entry](#)

### Entry Categories

Events 1,031

Memes 11,693

People 1,039

Sites 364

Subcultures 1,136

Confirmed

Submissions

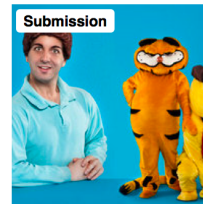
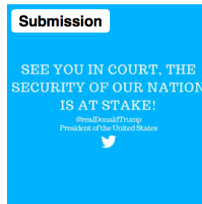
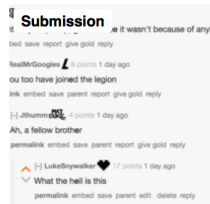
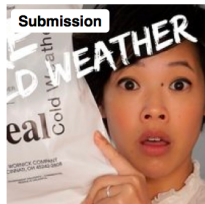
Deadpool

All Entries

Researching

Most Popular

1 2 3 4 5 6 7 8 9 ... 381 382 < >



# Applied Memology



## Doge

*Part of a series on Interior Monologue Captioning. [View Related Entries]  
[View Related Sub-entries]*

Updated 9 months ago by Asdfghjkl.

Added 3 years ago by NovaXP.

★ 281

📺 11,731,983

📺 54

📺 1,633

💬 895

[f Share this Entry](#)

27,394

[G+1](#)

573

[Like us on Facebook!](#)

[f Like](#)

834K

**PROTIP:** Press 'i' to view the image gallery, 'v' to view the video gallery, or 'r' to view a random entry.

## About

**Doge** is a slang term for “dog” that is primarily associated with pictures of Shiba Inus (nicknamed “Shibe”) and internal monologue captions on [Tumblr](#). These photos may be photoshopped to change the dog’s face or captioned with [interior monologues](#) in [Comic Sans](#) font.

## Origin

The use of the misspelled word “doge” to refer to a dog dates back to June 24th, 2005, when it was mentioned in an episode of [Homestar Runner](#)’s puppet show. In the episode titled “Biz Cas Fri 1”<sup>[2]</sup>, Homestar calls Strong Bad his “d-o-g-e” while trying to distract him from his work.

## Meme

### Status

Confirmed

### Type:

Character

### Year

2013

### Origin

[Tumblr](#)

### Tags

animal, dog, shiba inu, shibe, such doge, super shibe, japanese, super, tumblr, much, very, many, comic sans,

# Data

	name	added	views	comments	status	year	tags	about	origin	spread
0	Last Time The Cubs Won the World Series	Nov 03, 2016 at 11:22AM EDT	17157	60	Submission	Sports	Chicago	Last Time the Cubs Won the World Series refers...	NaN	Jokes and fascination with the Cubs' had small...
1	PenorCut's Fiverr Videos	Jul 20, 2015 at 07:54PM EDT	36629	39	Submission	Researching	YouTube	NaN	NaN	NaN
2	AmourShipping	Nov 02, 2016 at 03:01PM EDT	32293	81	Submission	Pairing	Unknown	Amourshipping refers to the romantic pairing o...	On September 12th, 2013, a discussion thread f...	NaN



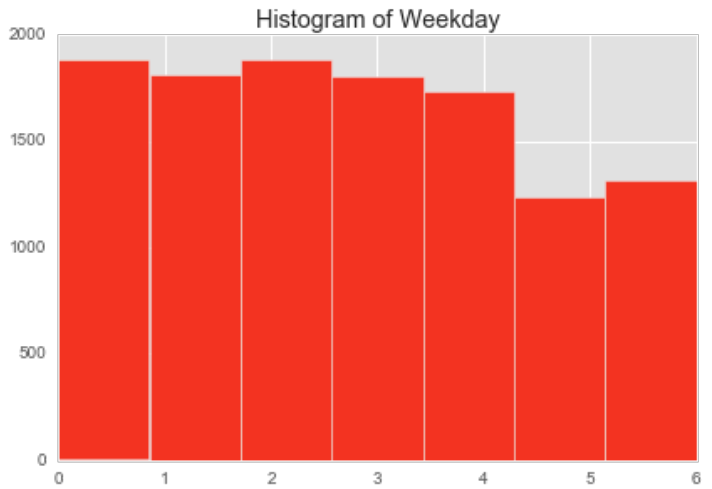
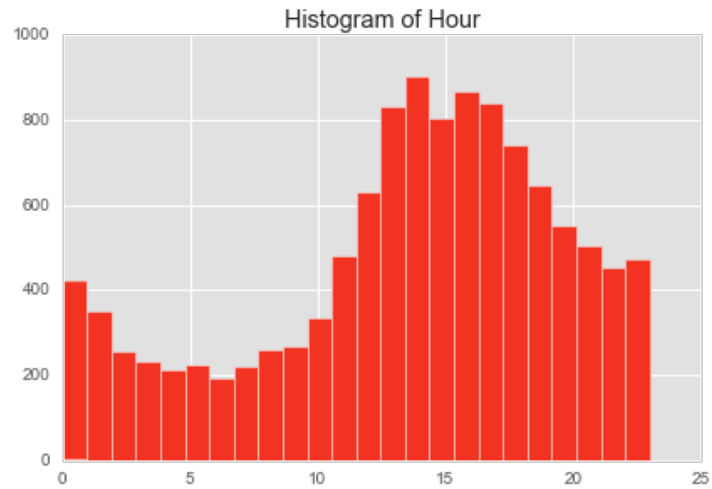
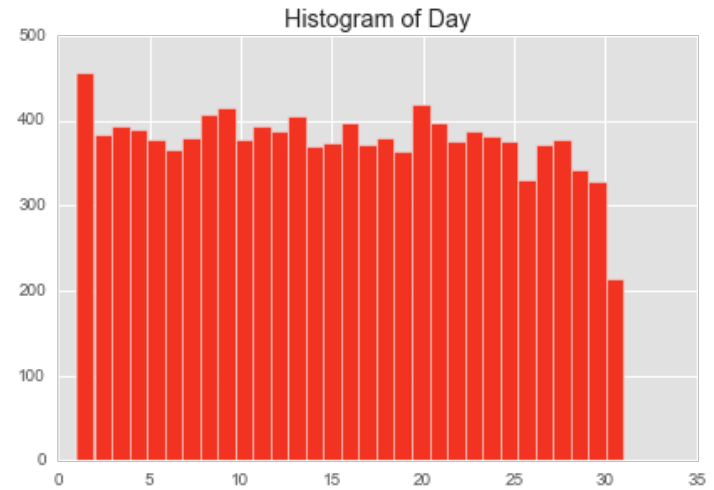
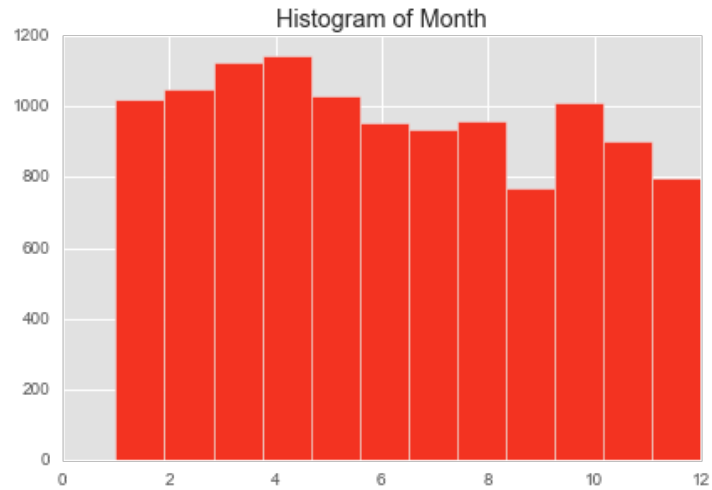
# Data

	name	added	views	comments	status	year	tags	about	origin	spread
5408	Doge	Jul 24, 2013 at 04:29PM EDT	11551088	893	Confirmed	Character	Tumblr	Doge is a slang term for “dog” that is primari...	The use of the misspelled word “doge” to refer...	On October 28th, 2010, a photo of Kabosu was s...

И это Доге!



# Data



# Target

- **совсем мертвые** - если просмотров меньше 1 в день
- **непопулярные** - если от 1 до 6 просмотров
- **средние** - от 6 до 24 просмотров
- **популярные** - от 24 до 74
- **супер популярные (виральные)** - выше 74 просмотров в день

Наблюдений в классах поровну, наивный бейзлайн для точности – 0.2

# Pipeline



NLTK snowball Stemmer

Удаляем стоп-слова

Tf-idf из sklearn, уни+биграммы, min\_df = 0.001

Логит-регрессия, случайный лес + композиция

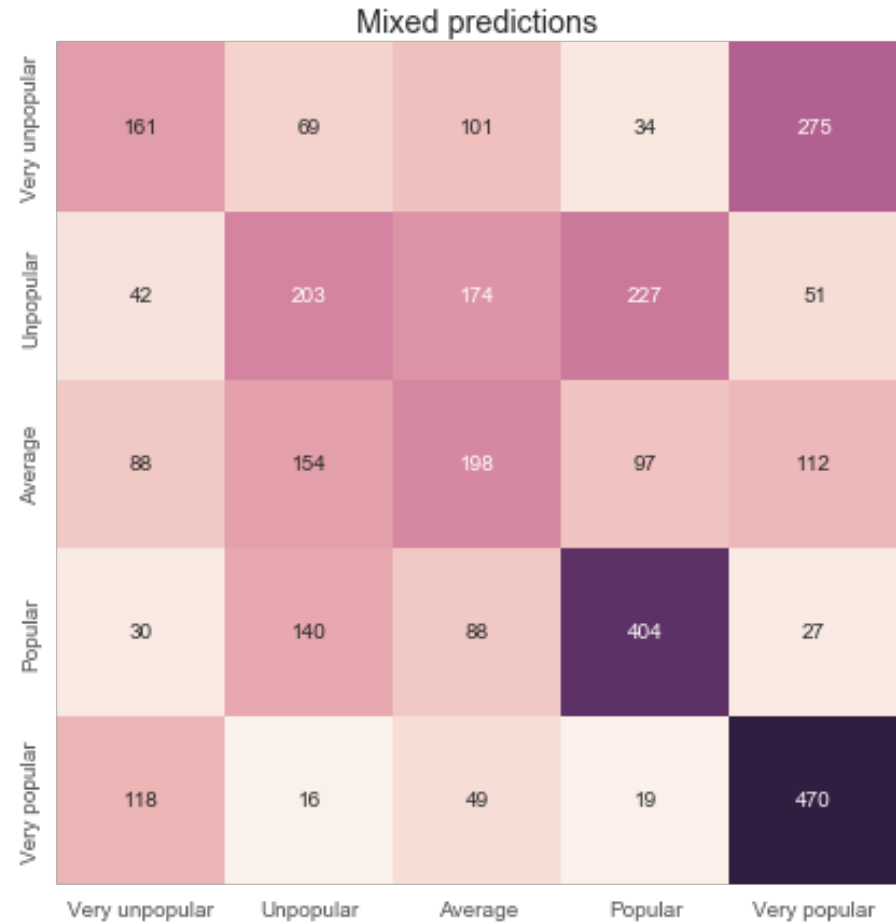
Итоговая точность на тесте – 0.43

# Результат

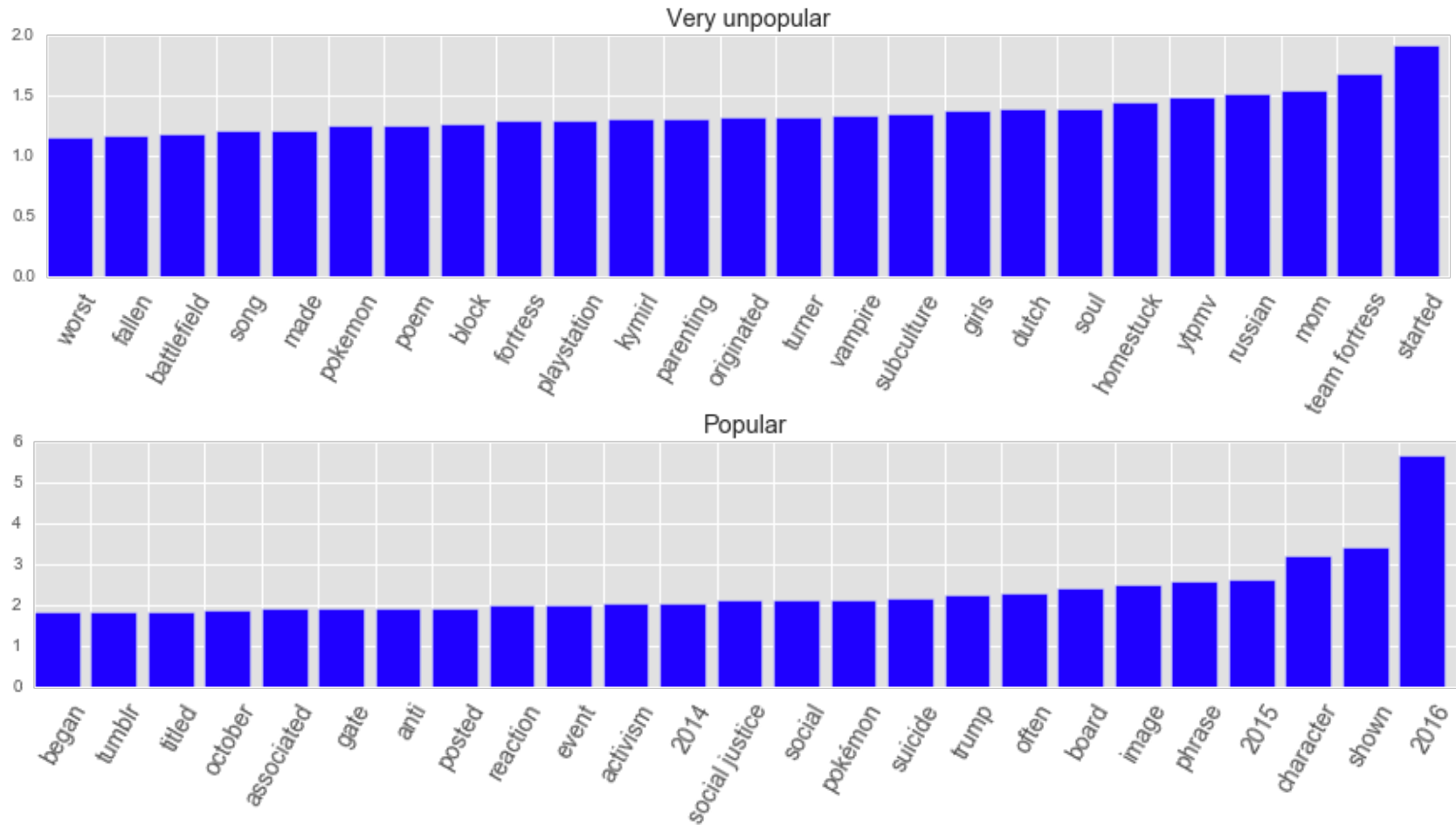
Логит 0.421

Лес 0.408

Смесь 0.429



# Немного картинок



**GO FUN YOURSELF.**

<https://github.com/DmitrySerg/memology>