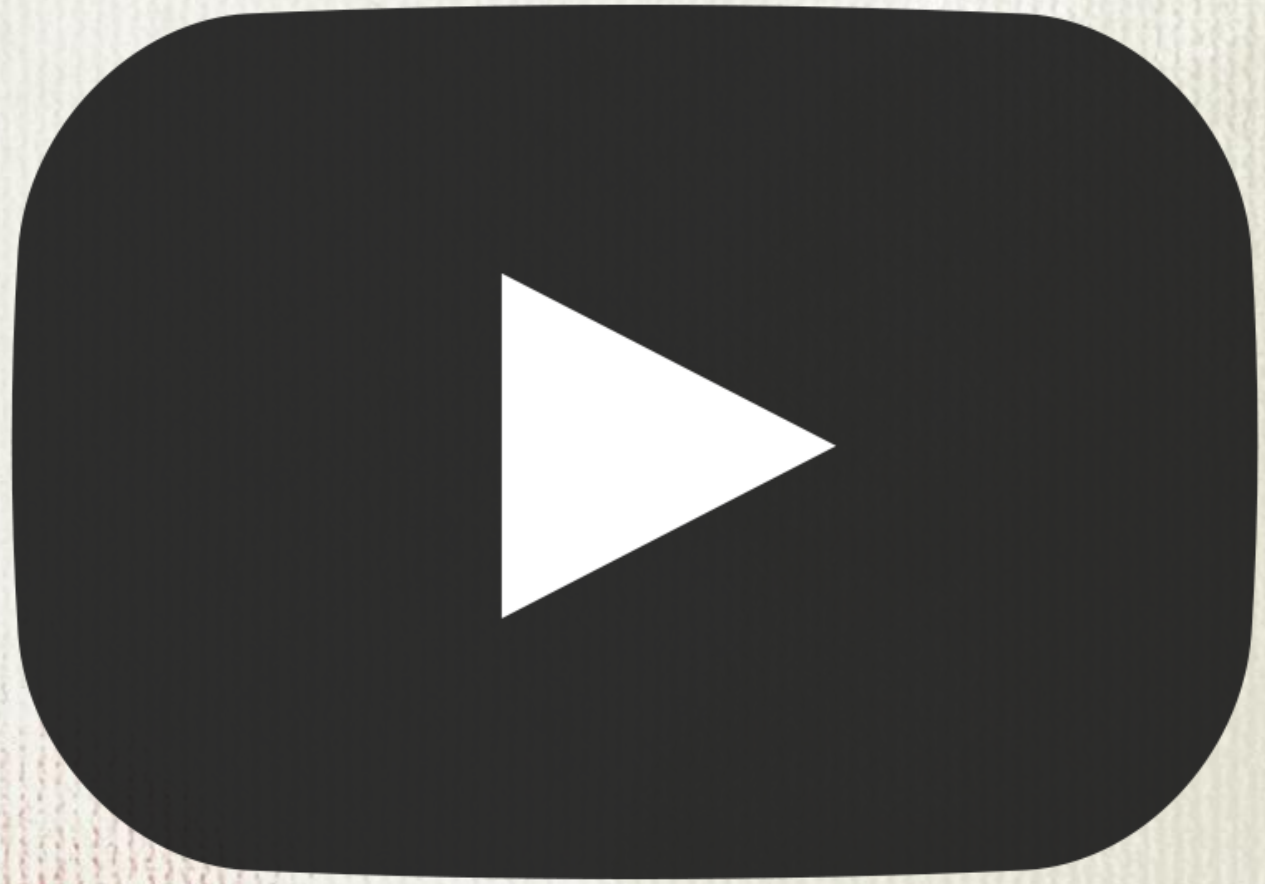
THE INSTAFORTUNE TELLER

ANASTASIYA IVANOVA DS14

LIKE PREDICTION
HASHTAGS RECOMMENDER

Type of Influencer	Number of Followers	Earning
MICRO	5K — 20K	**$88 per post**
MID TIER	20K — 100K	**$200 per post**
MACRO	100K — 1M	**$670 per post**
MEGA	> 1M	**up to 1 m $**

HypeAuditor



WHAT IS BEHIND THE BOT?

Number of post scraped: 241 775

Number of unique profiles: 3716

SCRAPERS

Octoparse

WHAT WAS SCRAPED?

instagram Follow •••

6,390 posts | 355m followers | 55 following

Instagram ← NAME

#ShareBlackStories ← DESCRIPTION

www.facebook.com/covidsupport ← LINK

Followed by gretathunberg, 2dayor4night, greenpeace + 5 more

WHAT'S ON PHOTO ↘

27, 2020. Image may contain: cloud, sky, outdoor and nature" class="FFVAD" decoding="auto"

instagram • Follow

Moscow, Russia ← LOCATION

CAPTION

instagram These cats are bananas. 🍌🐈🍌

Doll artist Anna Nazarenko's (@awesome_oso) menagerie of handmade, hybrid creatures includes broccoli sheep and Milky Way cows, and yes, banana cats. "I love mixing different things and literally anything will become a cute creature if you add a kitty face on it 😊," she says. "I take inspiration from various shapes but it's dull to just sculpt a banana, right?"

"My creatures must bring joy and smiles. They must be pleasant to look at and they must not break easily." #ThisWeekOnInstagram

Photo by @awesome_oso

891,020 likes

MAY 28 ← DATE

36 features

Profile level features:

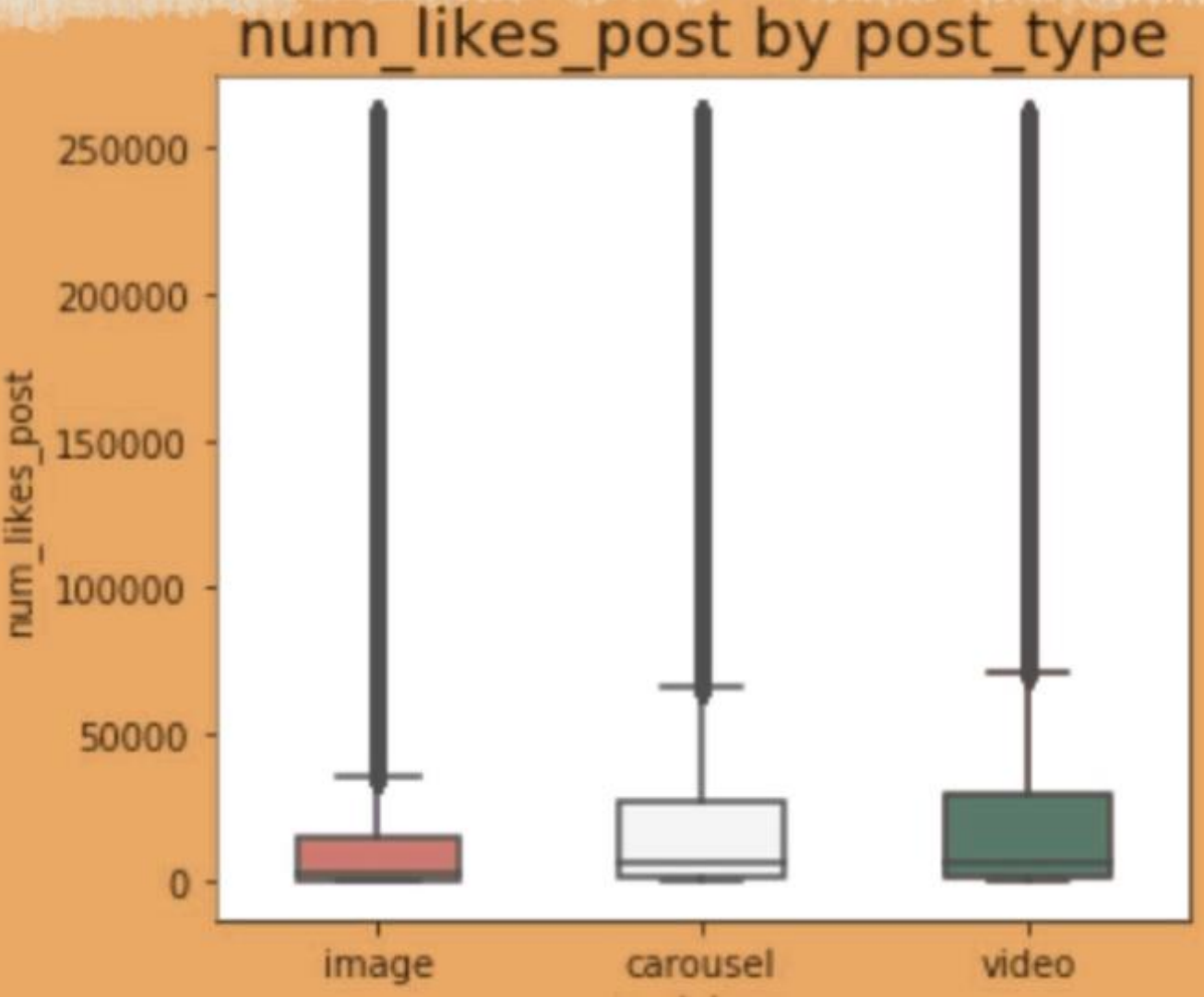
- main_profile_category
- profile_categories
- female followers
- male followers
- country
- profile_name
- num_posts
- full__profile
- profile_namedescription
- personal_link
- num_followers
- influencer_type
- post_frequency

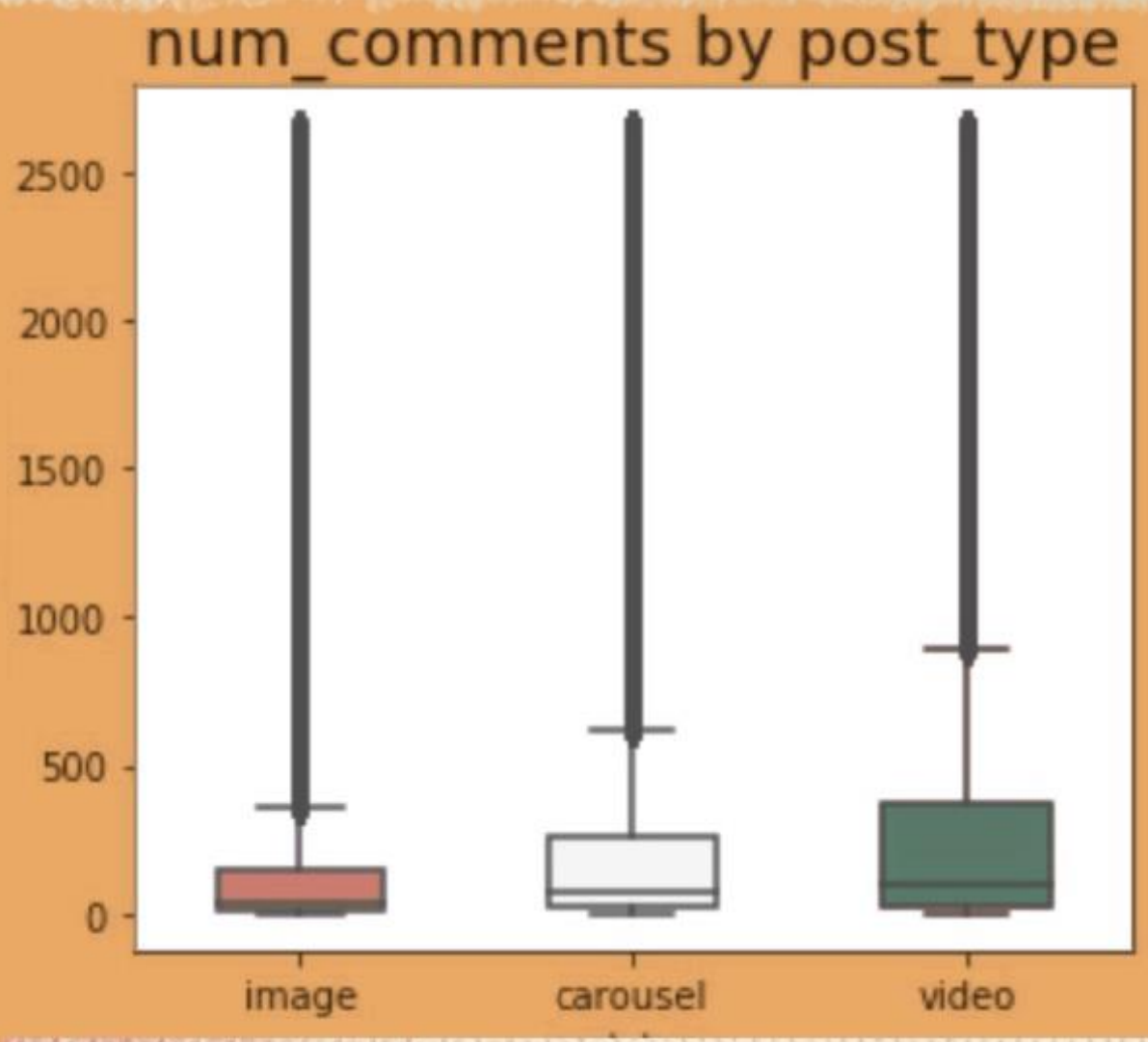
Post level features:

- comments
- post_text
- location
- num_likes
- num_views
- num_comments
- time of publication
- what’ s_on_photo
- post_type (carousel,video,image)
- Hashtags
- emoji_post_text
- emoji_comments
- len_post_text
- num_of_unique_hashtags
- mean_likes
- comments_engagement
- like_engagement
- polarity_post_txt
- subjectivity_post_txt

EXPLORATORY DATA ANALYSIS

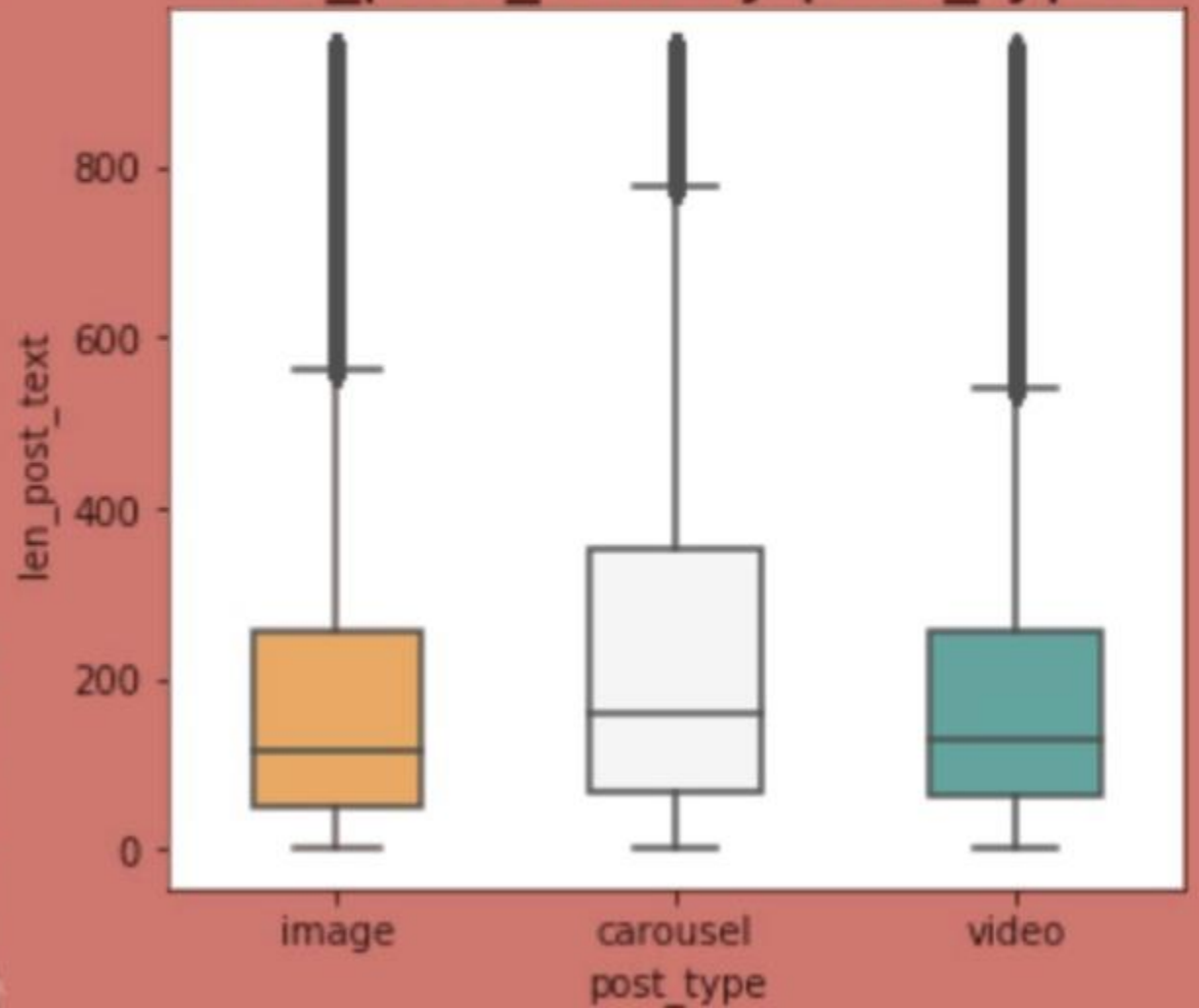
POST TYPES

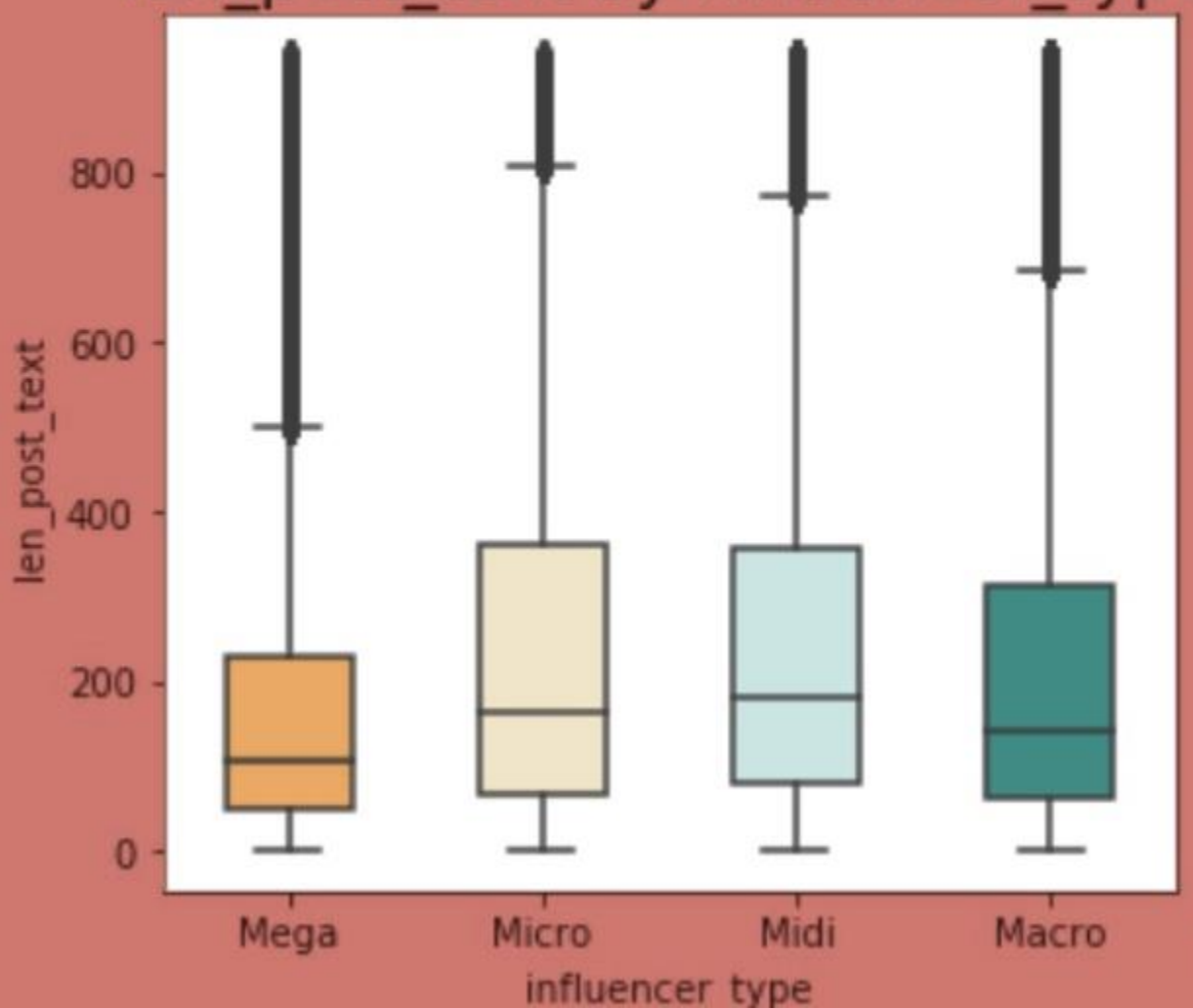

num_likes_post by post_type
num_likes_post
250000
200000
150000
100000
50000
0
image
carousel
video


num_comments by post_type
2500
2000
1500
1000
500
0
image
carousel
video

LENGTH OF TEXT

In number of characters





Most frequent words in posts text

live
work
start
find
love
home
link
bio
give
show
feel
let
well
photo
thing
little
great
moment
weekend
enjoy
remember
wear
watch
part
world
help
create
ready
put
life
perfect
set
happy birthday
challenge
favourite
beautiful
place
check
leave
learn
together
spend
video
face
dream
team
sure
piece
morning
month
change
call
turn
season
share
mean
hope
play
stay safe
inspire
amazing
bring
baby
walk
tonight
step
boy
ad
lover
support
click link
course
include
enough
lockdown
next
family
dress
tomorrow
long
talk
shoot
hour
style
believe
colour
happy
already
guy
yesterday
yet
bit
decide
different
pick
crazy
join
light
small
old
actually
end
someline
head
hard
mum
isolation
miss
drop
travel
couple
story
hair
add
kid
top
hand
eye
real
lovely
product
gorgeous
beauty
never
buy
pretty
picture
stay home
chance win
finish
catch
food
idea
girl
book
fun
experience
quick
fit
break
hear
close
late
child
order
finally
incredible
move
follow
meet
send
happen
build
strip
friend
heart
sunkle
hand
link profile
kind
enough
include
challenge
open
continue
worth
post
goal
bad
ago
before
swipe
special
whole
plan
pic
skin
allow
video
body
dream
hard
mum
piece
morning
month
feature
isolation
change
call
turn
season
share
mean
hope
play
stay safe
inspire
amazing
bring
baby
walk
tonight
step
boy
ad
lover
support
click link
course
include
enough
challenge
favourite
perfect
set
happy birthday
life
help
create
ready
put
world
part
watch
wear
enjoy
remember

REGRESSION

LIKES PREDICTION

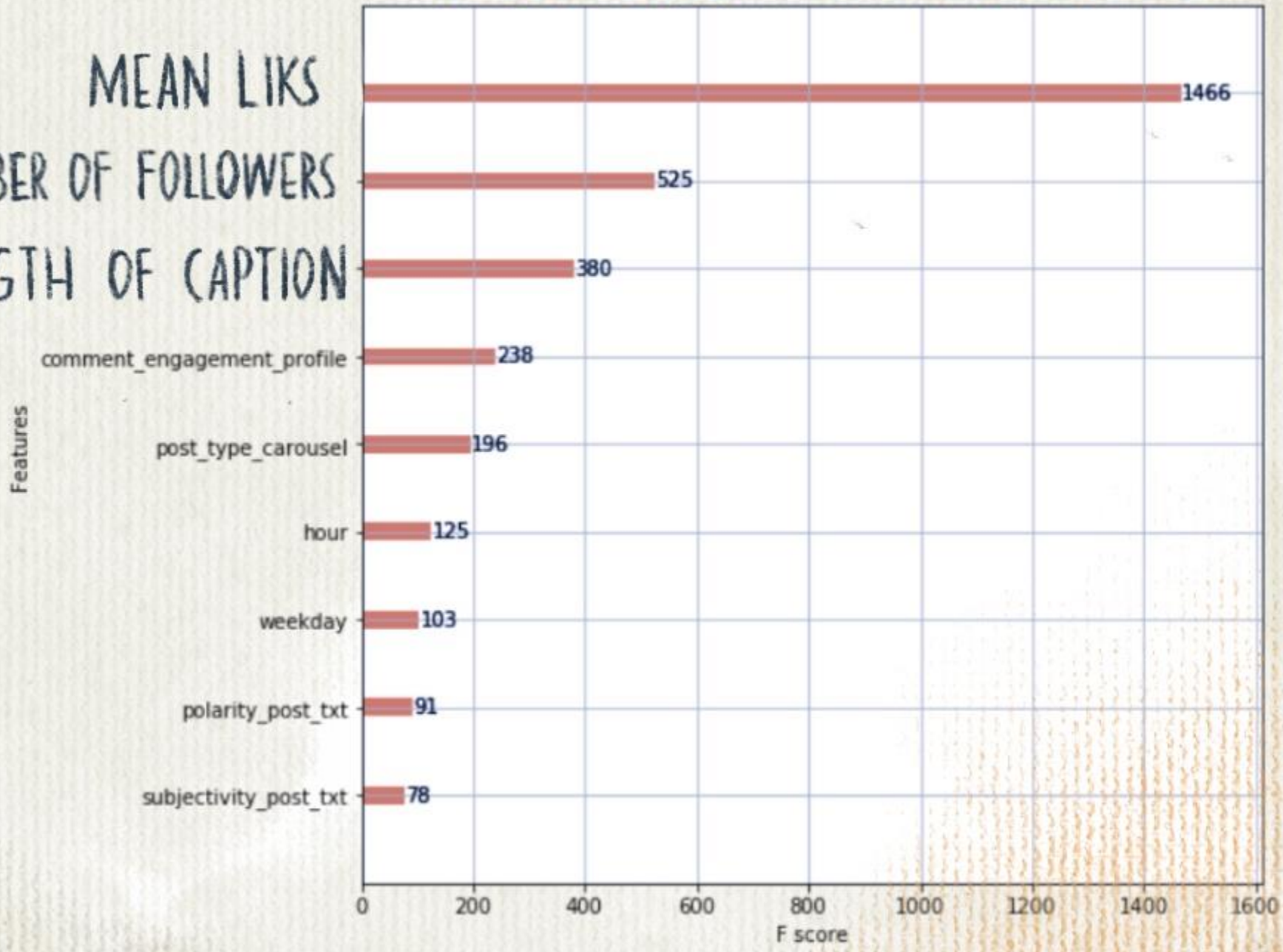
Model	Adjusted R2	Root Mean Square Error
Baseline (mean)	0	264515.4
XGBRegressor	0.77	106187.5
LinearRegression	0.76	112225.9
VotingRegressor	0.74	125721.0
Lasso	0.75	117056.9
ElasticNet	0.67	134077.4
KNeighborsRegressor	0.72	125036.5
DecisionTreeRegressor	0.56	155622.6
GradientBoostingRegressor	0.78	111166.8
Neuronal Network	0.75	131937.6

FEATURE IMPORTANCE

MEAN LIKS

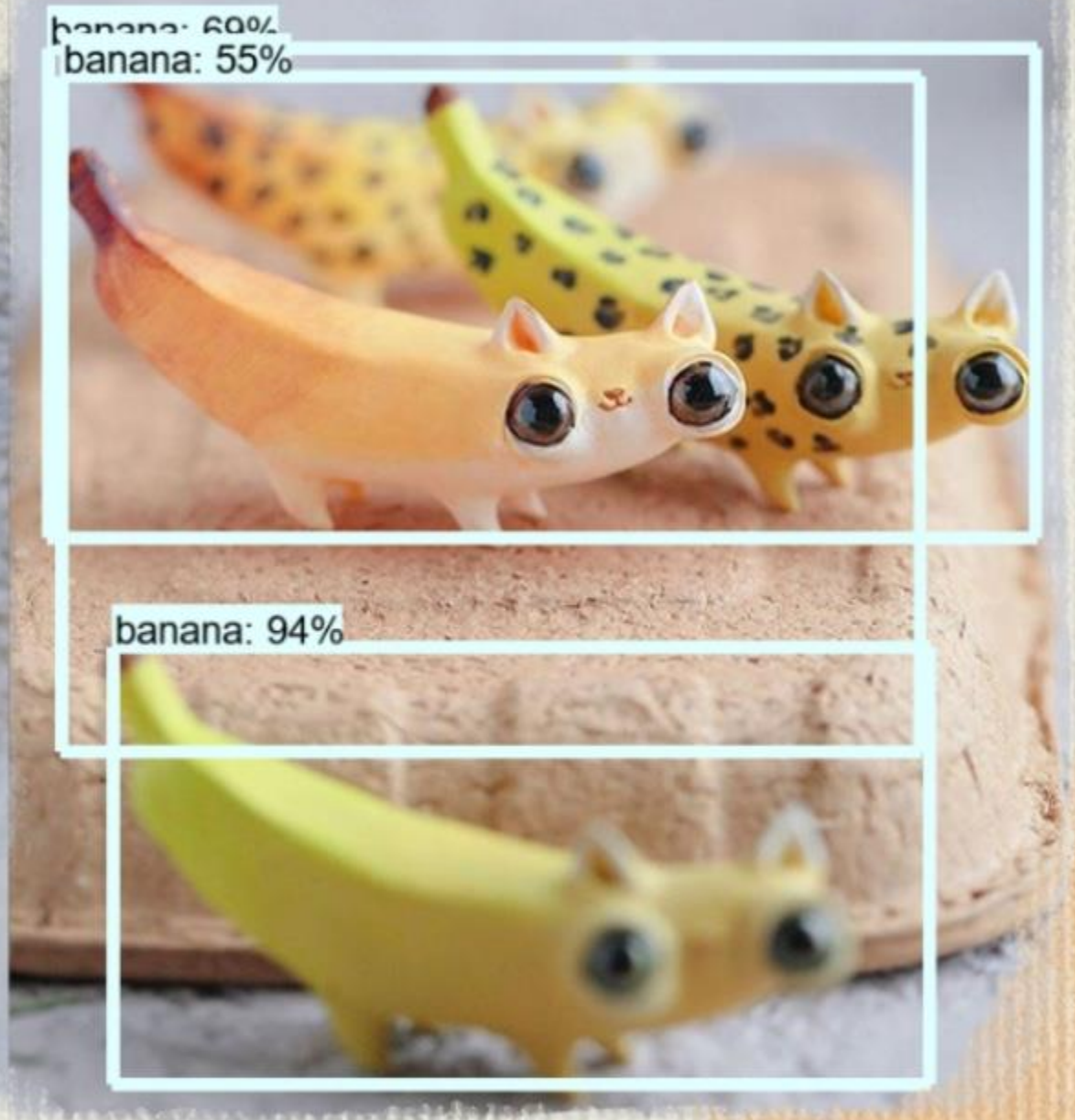
NUMBER OF FOLLOWERS

LENGTH OF CAPTION



OBJECT DETECTION

THRESHOLD - 50% PROBABILITY



#HASHTAG
SEARCHER

PREPROCESSED CAPTION + WHAT'S ON PHOTO

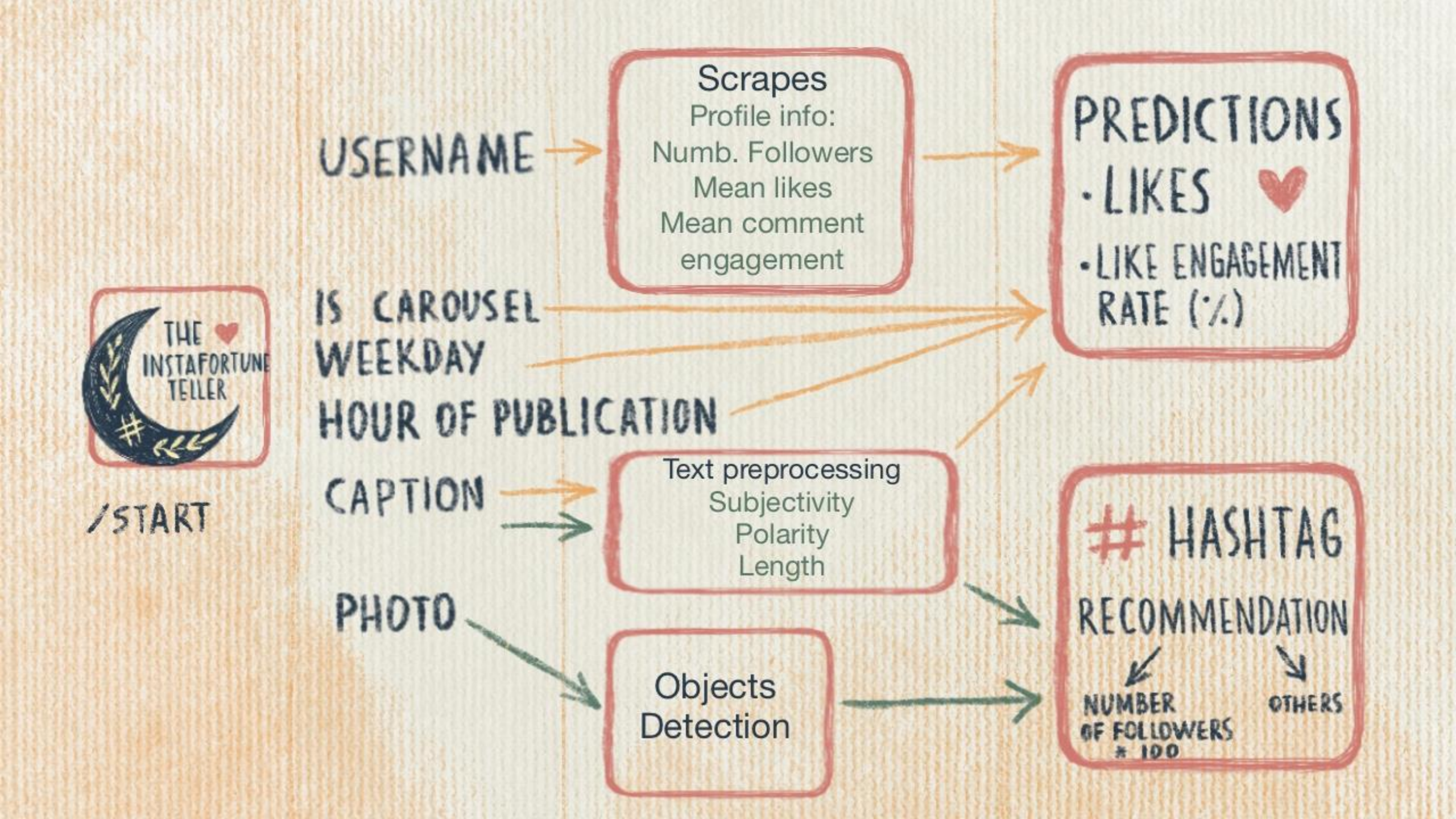
=

WORDS
BASE HASHTAG
SIMILAR

1. FOLLOWERS*100
AND
2. OTHERS

	original_hash	popularity	base_hash	similar	num_posts
0	love	1835000000	love	['loveit', 'lovely', 'loveyourself', 'loveyou'...	['76764148', '66763685', '57749282', '50459517...
1	instagood	1150000000	instagood	['instagoodmyphoto', 'instagoods', 'instagoodn...	['6830190', '3757690', '1799149', '324464', '2...
2	fashion	812700000	fashion	['fashionblogger', 'fashionista', 'fashionstyl...	['119437330', '103271371', '38380855', '341789...
3	photooftheday	797300000	photooftheday	['photoofthedays', 'photoofthedaychallenge', '...	['1820219', '140999', '94889', '57590', '33228...
4	beautiful	661000000	beautiful	['beautifuldestinations', 'beautifulday', 'bea...	['43766926', '22586677', '9268518', '8490783',...

TELEGRAM BOT

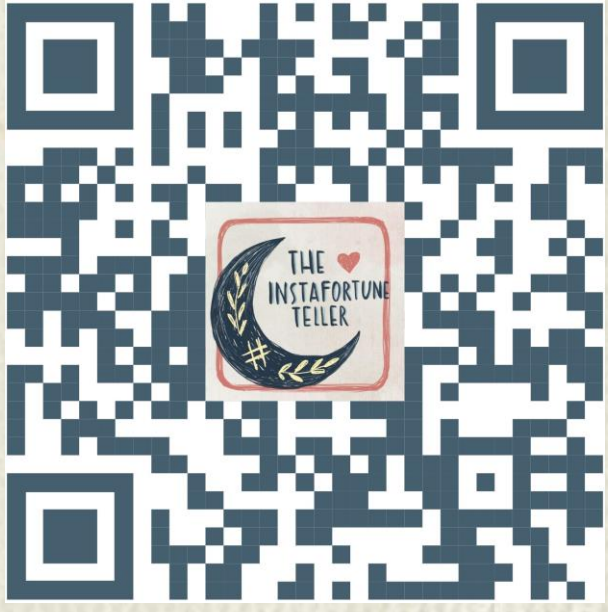

THE INSTAFORTUNE TELLER
/START
USERNAME
Scrapes
Profile info:
Numb. Followers
Mean likes
Mean comment engagement
IS CAROUSEL
WEEKDAY
HOUR OF PUBLICATION
CAPTION
Text preprocessing
Subjectivity
Polarity
Length
PHOTO
Objects Detection
PREDICTIONS
• LIKES
• LIKE ENGAGEMENT RATE (%)
HASHTAG RECOMMENDATION
NUMBER OF FOLLOWERS * 100
OTHERS

FURTHER STEPS:

Bot limitation:

- Only public accounts.
- For profiles with more than 500 followers.
- Database of hashtag does not update.
- Works from my local machine – performance is not very fast.

What can be done next?

- Search for similar posts from Macro and Mega influencers to show them as inspiration.
- Generate caption text base on photo and key words.

THANK YOU!