

Dataset

Crawled tweets from authors (first 500 from gs scholar sheet) - [dataset link](#)

Findings from dataset

- Users with popular tweets usually have
 - >1000 followers
 - properly written bio with words like 'researcher', 'prof', 'scientist', with twitter accounts of their affiliations tagged or written in the bio; can make a list of popular affiliations
 - website link in bio
- Tweets which are popular
 - Sharing their work and describing it in a thread
 - Asking meaningful questions/ suggestions eg what should be included in a library (huggingface) that almost everyone uses, or what papers should a course's recommended reading have etc.
- Like count correlates positively with
 - Followers count
 - Year in which they are posted
 - Number of urls
 - If it is a quote tweet
 - Total tweets by the author

Can do better

1. Not all tweets are about AI. Should we filter these out?
2. Dataset size is very small
3. User mentions for users which do not exist (deleted or taken down) are not present in entities
4. A better way to input numerical data into the model

Model Improvement

Details

- Dataset building + pre-processing
 - Crawled tweets for first 500 gs scholars (for which twitter ID is known)
 - Considering only English language tweets
 - Consider only tweets posted after 2017
 - Removing RTs as they are a function of both author who re-tweeted and original author and have like_count=0 always
 - [Q] Should we consider adding the original tweet that was re-tweeted into the dataset?
- Tweet pre-processing
 - BertTweet style - replace emojis, replace user mentions with @USER <twitter-name> and urls with HTTPURL, let hashtags be as it is
- Model Training

- 'distilbert-base-cased'
- Class weighting [98, 2] in loss function is used to give more weightage to minority class
- Gradient accumulation enables using bigger batch size: Small batch size used due to memory issues => huge fluctuations in loss => learning is very slow => Accumulated gradients due to memory issue
- Features used: 'text', 'quote_count', 'reply_count', 'retweet_count', 'year', 'month', 'day', 'weekday', 'hour', 'author_id', 'author_loc', 'author_account_age', 'followers_count', 'total_tweets_by_author_count'
- Features are combined with a [SEP] token in between them
- Example sentence
 - How computer vision will help Amazon customers shop online HTTPURL [SEP] 0 quotes [SEP] 0 replies [SEP] 0 retweets [SEP] 2020 year [SEP] 6 month [SEP] 5 day [SEP] 4 weekday [SEP] 21 hour [SEP] author is sondinhtran [SEP] [SEP] 3209 account age [SEP] 2875 followers [SEP] 107 total tweets

Ideas

Dataset Pre-processing:

1. Filter out tweets with less than 10 tokens [not implemented currently]

Tweet Pre-processing strategies:

1. HTTPURL followed by expanded text of the url
 1. The text is very long - think about strategies to reduce this

Model Improvement:

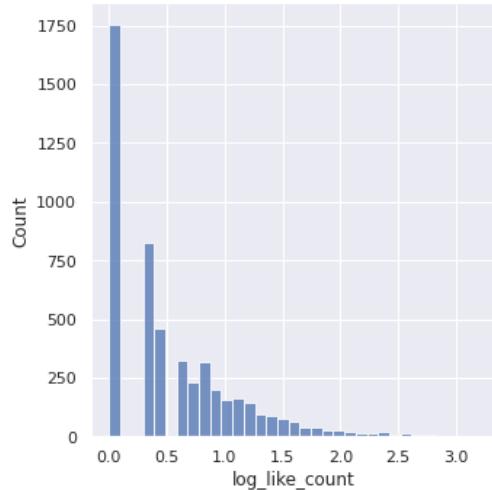
1. Use over-sampling
2. BERT(in general contextual embeddings) does not benefit from class weighting or over-sampling/under-sampling techniques <https://aclanthology.org/D19-5018v2.pdf>
3. Use different loss functions
 1. Sadie loss, KL Divergence; check this link -> <https://github.com/huggingface/transformers/issues/9374>
 2. DICE loss for imbalanced dataset <https://arxiv.org/pdf/1911.02855.pdf> <https://github.com/fursovia/self-adj-dice>
4. tune probability threshold for assigning labels instead of argmax

Analysis

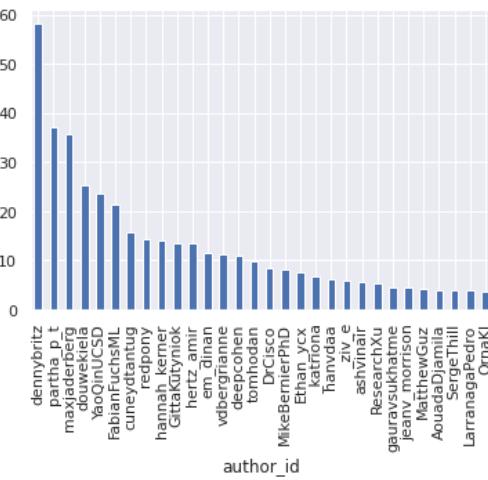
1. Use a better method for interpretability analysis to analyses dependence on input tokens
2. Can classify people into genders based on profile picture - acknowledge that predicting gender based on appearance is problematic and only applies to binary views

EDA

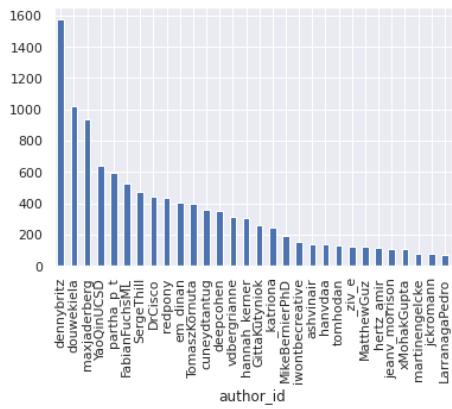
Distribution of log of like counts



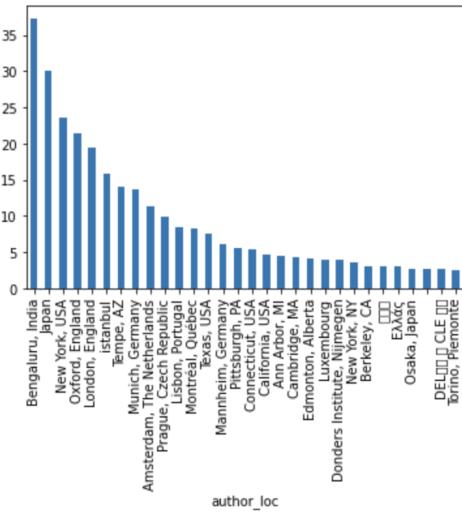
Distribution of mean like count on per author basis



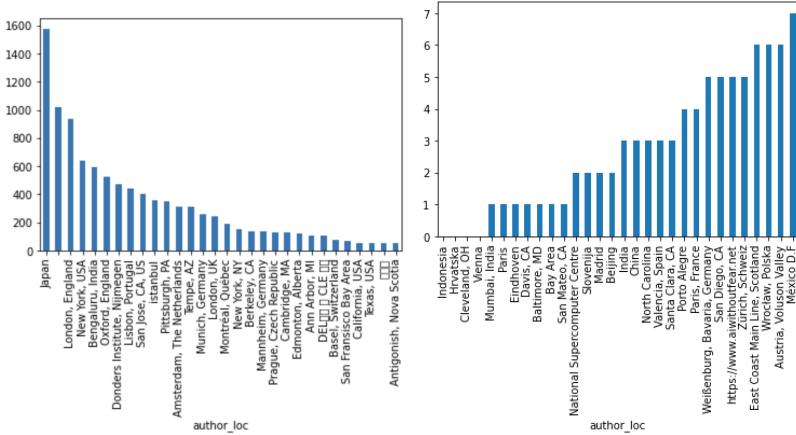
Distribution of max like count on per author basis



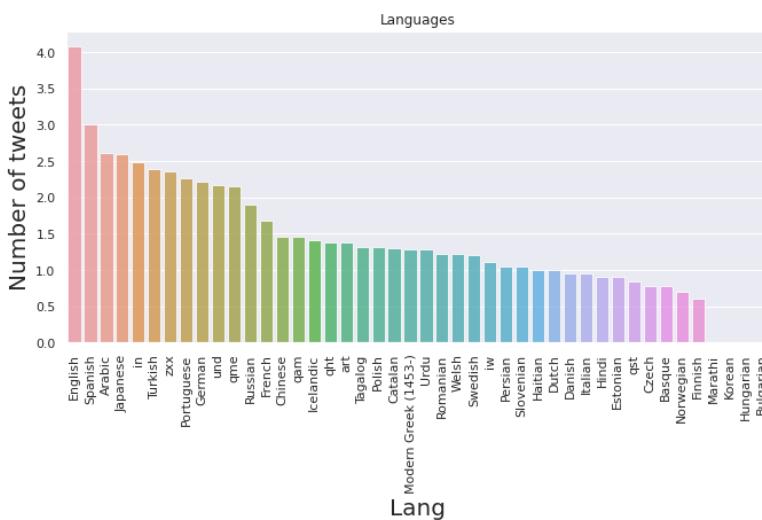
Distribution of mean like count by author location



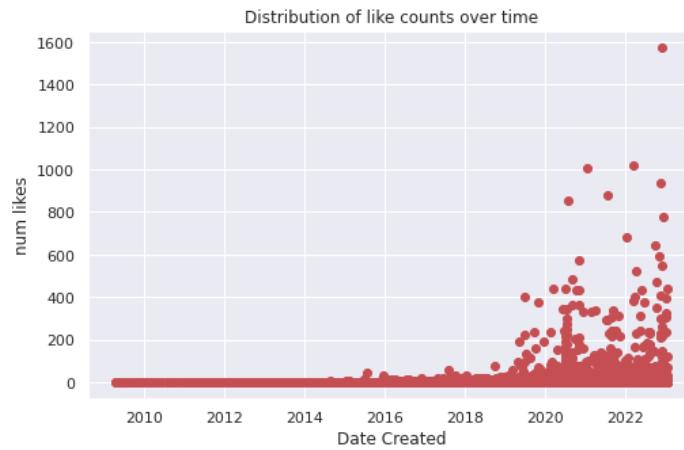
Distribution of max like count by author location



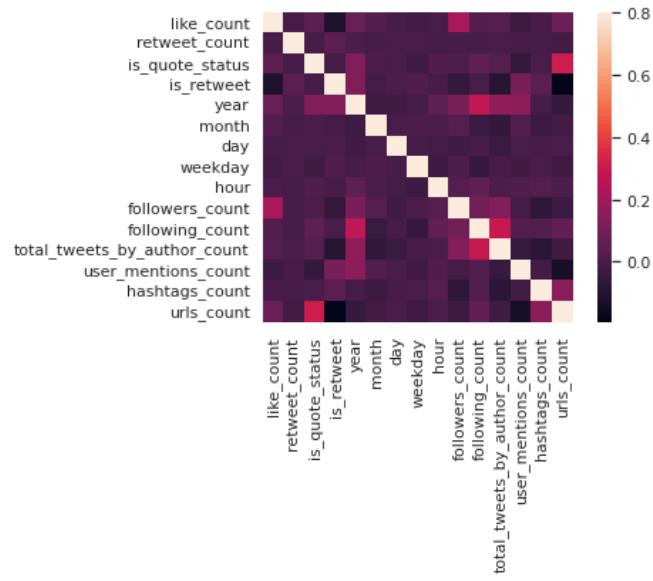
Distribution of number of tweets per language



Number of likes v/s time posted per tweet



Correlation matrix



Correlation with like_count

log_like_count	0.574777
followers_count	0.219471
urls_count	0.082099
year	0.071081
is_quote_status	0.042488
total_tweets_by_author_count	0.026494
following_count	0.023791
month	0.019536
hour	-0.000923
retweet_count	-0.004454
hashtags_count	-0.004811
day	-0.007942
weekday	-0.015019
user_mentions_count	-0.031980
is_retweet	-0.115025

Analysis

Error Analysis

Dataset

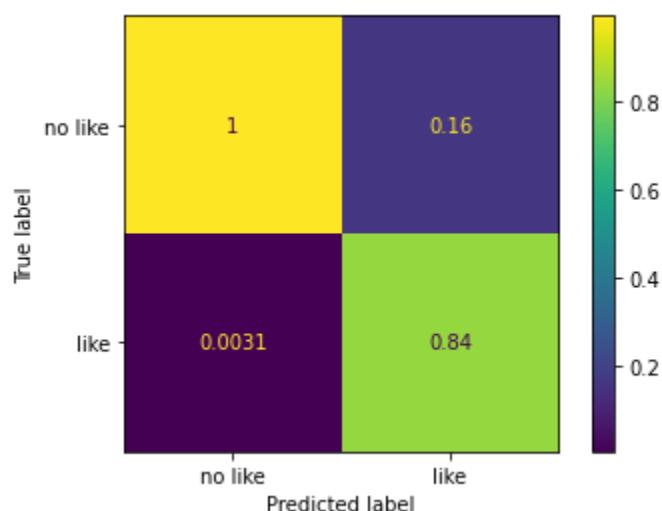
Dataset	Total	Class 0	class
Train	4022	3924	98
Validation	503	491	12
Test	503	491	12

Performance

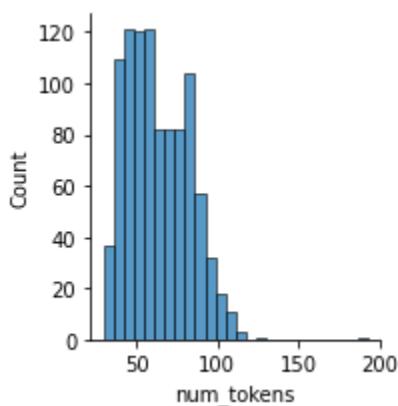
Model	Accuracy	Precision	Recall	F1
MAX_LEN = 128 TRAIN_BATCH_SIZE = 16 VALID_BATCH_SIZE = 2 EPOCHS = 20 LEARNING_RATE = 8e-06 EVAL_STEPS = 100 WARMUP_STEPS = 400 using_distilbert = True lr_sched = True freeze_encoder = False accumulate_gradients = True gradient_accumulation_steps = 32//TRAIN_BATCH_SIZE use_monolingual = True 'distilbert-base-cased' {text with @USER HTTPURL, user mentions appended at end of text}	Training 99.82595723520636 Validation 97.4155069582505	Validation [0.98185484 0.42857143] Test [0.98181818 0.375]	Validation [0.9918533 6 0.25] Test [0.9898167 0.25]	Training [0.9991088 5 0.96296296] Validation [0.9868287 7 0.31578947] Test [0.9858012 2 0.3]

MAX_LEN = 512 TRAIN_BATCH_SIZE = 16 VALID_BATCH_SIZE = 2 EPOCHS = 20 LEARNING_RATE = 1e-05 EVAL_STEPS = 100 WARMUP_STEPS = 400 using_distilbert = True lr_sched = True freeze_encoder = False accumulate_gradients = True gradient_accumulation_steps = 32//TRAIN_BATCH_SIZE use_monolingual = True 'distilbert-base-cased' {text, +text meta + user info}	Validation 99.6023856858847 Test 99.00596421471172	Validation [0.99796334 0.91666667] test [0.99591837 0.76923077]	Validation [0.9979633 4 0.91666667]] Test [0.9938900 2 0.83333333]]	Validation [0.9979633 4 0.91666667]] test [0.9949031 6 0.8]
--	---	--	--	--

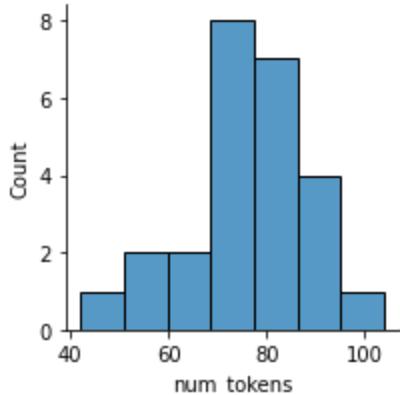
Confusion Matrix



Distribution of number of tokens in the text for non-popular (prediction = 0) tweets



Distribution of number of tokens in the text for popular (prediction = 1) tweets



Interpretability

[1] How does output change on changing information [location, author ID, number of followers, number of re-tweets]?

Location Dependence:

Equivariance is such a beautiful concept and so important for machine learning. @ USER Ed Wagstaff and I spent many hours setting our wits to understanding it from both an intuitive and theoretical points of view. This is the first part of our write - up : fox _ face : HTTPURL HTTPURL 0 quotes 0 replies 14 retweets 2020 year 2 month 25 day 1 weekday 13 hour author is FabianFuchsML Oxford, England 619 account age 2687 followers 152 total tweets

Original

```
882
label: 0, pred:1, text: [CLS] E ##qui ##var ##iance is such a beautiful concept and so important for machine learning . @ US ##ER Ed W ##ag ##sta ##ff and I spent many hours setting our wit ##s to understanding it from both an ##tu ##tive and theoretical points of view . This is the first part of our write - up : fox _ face : H ##TT ##P ##UR ##L H ##TT ##P ##UR ##L 0 quotes 0 replies 14 re ##t ##weet ##s 2020 year 2 month 25 day 1 weekday 13 hour author is Fabian ##F ##uch ##s ##ML Oxford , England 0 account age 26 ##8 ##7 followers 152 total t ##weet ##s
```

Location = Indonesia

```
882
label: 0, pred:1, text: [CLS] E ##qui ##var ##iance is such a beautiful concept and so important for machine learning . @ US ##ER Ed W ##ag ##sta ##ff and I spent many hours setting our wit ##s to understanding it from both an ##tu ##tive and theoretical points of view . This is the first part of our write - up : fox _ face : H ##TT ##P ##UR ##L H ##TT ##P ##UR ##L 0 quotes 0 replies 14 re ##t ##weet ##s 2020 year 2 month 25 day 1 weekday 13 hour author is Fabian ##F ##uch ##s ##ML Indonesia 0 account age 26 ##8 ##7 followers 152 total t ##weet ##s
```

- Changing location from Oxford, England to Indonesia changes the prediction from popular to non-popular
- IntegratedLayerGradients does not show location of input as a major contributor
 - Can explore other

Author ID Dependence:

Predictions are not affected by just author id change.

Followers Dependence:

Controversial opinion : Spamming the whole internet with low quality AI - generated content will turn out to be a great thing in the long - run. 0 quotes 0 replies 13 retweets 2022 year 12 month 5 day 0 weekday 13 hour author is dennybritz Japan 5450 account age 50819 followers 6926 total tweets

Followers = 50891 [Original]

```
prediction: tensor([1])
143
label: 1, pred:1, text: [CLS] Con ##tro ##vers ##al opinion : Spa ##mming the whole internet with low quality AI - generated content will turn out to be a great thing in the long - run . 0 quotes 0 replies 13 re ##t ##weet ##s 202 ##2 year 12 month 5 day 0 weekday 13 hour author is den ##ny ##bri ##tz Japan 54 ##50 account age 50 ##8 ##19 followers 69 ##26 total t ##weet ##s
```

Followers = 5891

```

prediction: tensor([1])
143
label: 1, pred:1, text: [CLS] Con ##tro ##vers ##ial opinion : Spa ##mming the whole internet with low quality AI - generated content will turn out to be a great thing in the long - run . 0 quotes 0 replies 13 re ##t ##weet ##s 202
##2 year 12 month 5 day 0 weekday 13 hour author is den ##ny ##bri ##tz Japan 54 ##50 account age 58 ##9 followers 69 ##26 total t ##weet ##

```

Followers = 891

```

prediction: tensor([1])
143
label: 1, pred:1, text: [CLS] Con ##tro ##vers ##ial opinion : Spa ##mming the whole internet with low quality AI - generated content will turn out to be a great thing in the long - run . 0 quotes 0 replies 13 re ##t ##weet ##s 202
##2 year 12 month 5 day 0 weekday 13 hour author is den ##ny ##bri ##tz Japan 54 ##50 account age 89 ##1 followers 69 ##26 total t ##weet ##

```

Followers = 231 [prediction changes]

```

prediction: tensor([0])
143
label: 1, pred:1, text: [CLS] Con ##tro ##vers ##ial opinion : Spa ##mming the whole internet with low quality AI - generated content will turn out to be a great thing in the long - run . 0 quotes 0 replies 13 re ##t ##weet ##s 202
##2 year 12 month 5 day 0 weekday 13 hour author is den ##ny ##bri ##tz Japan 54 ##50 account age 231 followers 69 ##26 total t ##weet ##

```

Followers = 91

```

prediction: tensor([0])
143
label: 1, pred:1, text: [CLS] Con ##tro ##vers ##ial opinion : Spa ##mming the whole internet with low quality AI - generated content will turn out to be a great thing in the long - run . 0 quotes 0 replies 13 re ##t ##weet ##s 202
##2 year 12 month 5 day 0 weekday 13 hour author is den ##ny ##bri ##tz Japan 54 ##50 account age 91 followers 69 ##26 total t ##weet ##

```

Followers = 2

```

prediction: tensor([0])
143
label: 1, pred:1, text: [CLS] Con ##tro ##vers ##ial opinion : Spa ##mming the whole internet with low quality AI - generated content will turn out to be a great thing in the long - run . 0 quotes 0 replies 13 re ##t ##weet ##s 202
##2 year 12 month 5 day 0 weekday 13 hour author is den ##ny ##bri ##tz Japan 54 ##50 account age 2 followers 69 ##26 total t ##weet ##

```

- Reducing the number of followers changes the prediction

re-tweets Dependence:

Re-tweets = 80 [Original]

```

prediction: tensor([1])
190
label: 1, pred:1, text: [CLS] Alpha ##F ##old 2 re ##fines structures it ##erative ##ly & e ##qui ##var ##iant ##ly . What happens when you apply an SE (3) - Trans ##former it ##erative ##ly ? @ US ##ER Ed W ##ag ##sta
##ff and I wrote about the challenges and how they can be addressed . It includes code : laptop _ computer : pretty pictures : sunrise _ over _ mountains & a tech report : me ##mo : H ##TT ##P ##UR ##L H ##TT ##P
##UR ##L 0 quotes 0 replies 80 re ##t ##weet ##s 202 ##1 year 3 month 2 day 1 weekday 11 hour author is Fabian ##F ##uch ##s ##ML Oxford , England 99 ##0 account age 26 ##8 ##7 followers 152 total t ##weet ##

```

Re-tweets = 100

```

prediction: tensor([0])
190
label: 1, pred:1, text: [CLS] Alpha ##F ##old 2 re ##fines structures it ##erative ##ly & e ##qui ##var ##iant ##ly . What happens when you apply an SE (3) - Trans ##former it ##erative ##ly ? @ US ##ER Ed W ##ag ##sta
##ff and I wrote about the challenges and how they can be addressed . It includes code : laptop _ computer : pretty pictures : sunrise _ over _ mountains & a tech report : me ##mo : H ##TT ##P ##UR ##L H ##TT ##P
##UR ##L 0 quotes 0 replies 100 re ##t ##weet ##s 202 ##1 year 3 month 2 day 1 weekday 11 hour author is Fabian ##F ##uch ##s ##ML Oxford , England 99 ##0 account age 26 ##8 ##7 followers 152 total t ##weet ##

```

- Model predicted tweet to be un-popular when re-tweets were increased from 80 to 100

[2] Does the model even understand numbers?

Followers = 289

```

prediction: tensor([1])
143
label: 1, pred:1, text: [CLS] Con ##tro ##vers ##ial opinion : Spa ##mming the whole internet with low quality AI - generated content will turn out to be a great thing in the long - run . 0 quotes 0 replies 13 re ##t ##weet ##s 202
##2 year 12 month 5 day 0 weekday 13 hour author is den ##ny ##bri ##tz Japan 54 ##50 account age 28 ##9 followers 69 ##26 total t ##weet ##

```

Followers = 290

```

prediction: tensor([0])
143
label: 1, pred:1, text: [CLS] Con ##tro ##vers ##ial opinion : Spa ##mming the whole internet with low quality AI - generated content will turn out to be a great thing in the long - run . 0 quotes 0 replies 13 re ##t ##weet ##s 202
##2 year 12 month 5 day 0 weekday 13 hour author is den ##ny ##bri ##tz Japan 54 ##50 account age 290 followers 69 ##26 total t ##weet ##

```

- The model may not be understanding relationship (increasing or decreasing count) between numbers even though follower count correlates with popularity

[3] What is the dependence on topics? [TODO]

[4] What does the model predictions depend on? Do certain phrases impact model prediction?

Tweets which are predicted to be popular

- Tweets which talk about ‘great/cool opportunity’
- Tweets containing phrases/words
 - ‘PhD’, ‘ICLR’, ‘NeurIPS’, ‘human-like understanding’, ‘learning research’, ‘research’, ‘spent many hours’, ‘Transformer’
- Location dependence Eg: ‘London, England’, ‘Karnataka’, ‘Japan’, ‘Germany’, ‘Oxford’
- Number of followers and total tweets by the author
- Popular authors - "dennybritz", "partha_p_t", "YaoQinUCSD", "douwekiela", "FabianFuchsML", "DrCisco"
- Emojis: smiling_face_with_heart_eyes

```

prediction: tensor([1])
579
label: 1, pred:1, text: [CLS] # N ##eur ##IP ##S : Ex ##cite ##d about presenting the SE ##3 : Trans ##former at N ##eur ##IP ##S ! : ne ##rd _ face : How do you leverage rotation ##al s ##ym ##met ##ries in point clouds and graphs ? Paper + : fire : + an excellent Video : film _ project #for : explanation by @ US ##ER Max Well ##ing here : H ##TT ##P ##UR ##L H ##TT ##P ##UR ##L 0 quotes 0 replies 43 re ##t ##weet ##s 2020 year 9 month 29 day 1 weekday 15 hour author is Fabian ##F ##uch ##s ##ML Oxford , England 83 ##6 account age 26 ##8 ##7 followers 152 total t ##weet ##s

```

Tweets which are predicted to be not popular

- Tweets on QA, KG
- Tweets containing phrases/words
 - ‘Collaboration with’ , ‘Check them out’, ‘BJP’, ‘India’, ‘PhD student’, ‘took more Googling, ‘may be this is the reason’ (showing less confidence in their tweets), ‘spent many hours’

Spurious Features

This tweet is very sensitive to any changes in the words but LayeredIntegratedGradients does not reflect that

```

935
label: 1, pred:1, text: [CLS] I have a math joke , but clearly , it should be obvious ! 0 quotes 0 replies 12 re ##t ##weet ##s 2020 year 7 month 26 day 6 weekday 17 hour author is den ##ny ##br ##z Japan 0 account age 50 ##8 ##19 followers 69 ##26 total t ##weet ##s

```

```

935
label: 1, pred:1, text: [CLS] I have a math joke , but clearly , it should be obvious ! 0 quotes 0 replies 12 re ##t ##weet ##s 2020 year 7 month 26 day 6 weekday 17 hour author is den ##ny ##br ##z Indonesia 0 account age 50 ##8 ##19 followers 69 ##26 total t ##weet ##s

```

```

935
label: 1, pred:1, text: [CLS] I have a math joke , but clearly , it should be obvious ! 0 quotes 0 replies 12 re ##t ##weet ##s 2020 year 7 month 26 day 6 weekday 17 hour author is ant ##oni ##o _ star ##9 ##3 Japan 45 ##8 ##8 account age 50 ##8 ##19 followers 69 ##26 total t ##weet ##s

```

```

prediction: tensor([0])
935
label: 1, pred:1, text: [CLS] I have a math joke , but clearly , it should be obvious ! 0 quotes 0 replies 12 re ##t ##weet ##s 2020 year 7 month 26 day 6 weekday 17 hour author is el ##ias ##z ##wang Japan 45 ##8 ##8 account age 50 ##8 ##19 followers 69 ##26 total t ##weet ##s

```

Visualisations

Mis-classified sentences

```

309
label: 1, pred:0, text: [CLS] Papers from my group @ US ##ER at a ##C ##f ##20 ##20 ##n ##l ##p on O ##A over K ##G , explanations for NL ## better K ##G evaluation , and syntax - guided para ##ph ##ras ##ing . Check them out ! H ##TT ##P ##UR ##L 0 quotes 0 replies 6 re ##t ##weet ##s 2020 year 7 month 5 day 6 weekday 9 hour author is part ##ha _ p _t Bengal ##uru , India 0 account age 254 ##8 followers 74 total t ##weet ##s
498
label: 0, pred:1, text: [CLS] Thank you for this great write - up about A ##F ##2 ! H ##TT ##P ##UR ##L ! The SE ( 3 ) - Trans ##former [ . . . ] is currently too computational ##ly expensive for protein - scale tasks * right arrow . To be honest , I initially thought so as well , but @ US ##ER Min ##ky ##ung Ba ##ek et al made it work ! : smiling _ face _ with _ heart - eyes : H ##TT ##P ##UR ##L H ##TT ##P ##UR ##L 0 quotes 0 replies 17 re ##t ##weet ##s 202 ##1 year 8 month 4 day 2 weekday 9 hour author is Fabian ##F ##uch ##s ##ML Oxford , England 0 account age 26 ##8 ##7 followers 152 total t ##weet ##s
503
label: 1, pred:0, text: [CLS] I am in the UK this week to visit our brand new @ US ##ER London office ! Who 's up for a pint ##in a pub ? Shoot me a message ! : hugging _ face : 0 quotes 0 replies 3 re ##t ##weet ##s 202 ##2 year 6 month 27 day 0 weekday 8 hour author is do ##u ##we ##kie ##la 0 account age 64 ##34 followers 33 ##4 total t ##weet ##s
600
label: 1, pred:0, text: [CLS] @ US ##ER Am ##ged A guy who paid for Twitter calling other people sheep to ##1 0 quotes 0 replies 0 re ##t ##weet ##s 202 ##3 year 1 month 16 day 0 weekday 17 hour author is Dr ##C ##isco Lisbon , Portugal 0 account age 217 ##1 followers 404 ##9 ##2 total t ##weet ##s
612
label: 0, pred:1, text: [CLS] Interest ##red in pursuing a PhD in Re ##lia ##bla # AI in the exciting city of # Munich at the Universities of Excellence @ US ##ER & @ US ##ER ? We now have a call for PhD positions in our Z ##use School re ##t ##A ##I H ##TT ##P ##UR ##L @ US ##ER Research in Bavaria , @ US ##ER Research in Germany , @ US ##ER H ##TT ##P ##UR ##L 0 quotes 0 replies 21 re ##t ##weet ##s 202 ##2 year 11 month 28 day 0 weekday 17 hour author is G ##itt ##fa ##K ##ut ##y ##n ##io ##Munich , Germany 0 account age 1680 followers 34 ##2 total t ##weet ##s
793
label: 0, pred:1, text: [CLS] Stop discrimination of BP ##S faculty members in all Pakistan Universities ! AP ##U ##B ##TA demands H ##EC for BP ##S statutes approval ! @ US ##ER PT ##I @ US ##ER H ##EC Pakistan @ US ##ER @ US ##ER @ US ##ER Official PS ##L @ US ##ER @ US ##ER AR ##Y NE ##WS @ US ##ER 92 News HD Plus 0 quotes 0 replies 21 re ##t ##weet ##s 202 ##2 year 1 month 23 day 6 weekend 7 hour author is u ##omi ##an ##9 Porto , Portugal 0 account age 47 followers 18 total t ##weet ##s
882
label: 0, pred:1, text: [CLS] E ##qui ##var ##iance is such a beautiful concept and so important for machine learning . @ US ##ER Ed W ##ag ##sta ##ff and I spent many hours setting our wit ##s to understand it from both an in ##ftu ##tive and theoretical points of view . This is the first part of our write - up : fox _ face : H ##TT ##P ##UR ##L H ##TT ##P ##UR ##L 0 quotes 0 replies 14 re ##t ##weet ##s 2020 year 2 month 25 day 1 weekday 13 hour author is Fabian ##F ##uch ##s ##ML Oxford , England 0 account age 26 ##8 ##7 followers 152 total t ##weet ##s

```

Correctly classified:

22

[CLS] @ US ##ER A ##ki ##lesh R Yes , using coconut milk [SEP] 0 quotes [SEP] 0 replies [SEP] 0 re ##t ##weet ##s [SEP] 202 ##1 year [SEP] 11 month [SEP] 2 day [SEP] 1 weekday [SEP] 18 hour [SEP] author is part ##ha [SEP] _ p _t Bengal ##uru , India [SEP] 0 account age [SEP] 254 ##8 followers [SEP] 74 total t ##weet ##s [SEP]

574

[CLS] Super excited about 3 projects my group at Google Research has been working on getting announced at # Google ##F ##or ##I ##ndi ##a # 1 V ##aan ##f : collaboration with @ US ##ER II ##S ##c Bangalore to capture India 's speech landscape by collecting and open sour ##cinc speech from 77 ##3 districts H ##TT ##P ##UR ##L H ##TT ##P ##UR ##L [SEP] 0 quotes [SEP] 0 replies [SEP] 28 re ##t ##weet ##s [SEP] 202 ##2 year [SEP] 12 month [SEP] 19 day [SEP] 0 weekday [SEP] 18 hour [SEP] author is part ##ha [SEP] _ p _t Bengal ##uru , India [SEP] 0 account age [SEP] 254 ##8 followers [SEP] 74 total t ##weet ##s [SEP]

13
label: 0, pred:0, text: [CLS] @ US ##ER Will Chen @ US ##ER @ US ##ER @ US ##ER Co ##smos - Internet of Block ##cha ##ns Luna story = :full _ moon :: wax ##ing _ g ##ib ##bo ##us _ moon :: first _ quarter _ moon :: wax ##ing _ c ##res ##cent _ moon :: new _ moon :: new _ moon _ face : wa ##ing _ c ##res ##cent _ moon : last _ quarter _ moon :: wa ##ing _ g ##ib ##bo ##us _ moon :: full _ moon :: full _ moon _ face : come back boy , comet : 0 quotes 0 replies 0 re ##t ##weet ##s 202 ##2 year 5 month 14 day 5 weekday 16 hour author is m ##ora ##d ##del ##i Islamic Republic of Iran 0 account age 19 followers 143 total t ##weet ##s 659

label: 0, pred:0, text: [CLS] Un ##u ##s ##lied de ##ten ##ten ##tion and arrest of # Sa ##ad ##H ##iss ##ain ##R ##iv ##f should be end by Pakistani go ##v ##f Sa ##ad Hussain R ##iz ##vi is the hope of this country # Release _T ##LP ##A ##mee ##S ##H ##R 0 quotes 0 replies 2 re ##t ##weet ##s 202 ##1 year 8 month 22 day 6 weekday 4 hour author is q ##ain ##ain ##9 0 account age 60 followers 152 total t ##weet ##s 793

label: 0, pred:1, text: [CLS] Stop discrimination of BP ##S faculty members in all Pakistan Universities ! AP ##U ##B ##TA demands H ##EC for BP ##S statutes approval ! @ US ##ER PT ##I @ US ##ER H ##EC Pakistan @ US ##ER @ US ##ER Official PS ##L @ US ##ER # ##R ##E ##WS @ US ##ER @ US ##ER 32 News HD Plus 0 quotes 0 replies 21 re ##t ##weet ##s 202 ##2 year 1 month 23 day 6 weekday 7 hour author is u ##omi ##fan ##4 ##Porto , Portugal 0 account age 47 followers 18 total t ##weet ##s

524
label: 0, pred:0, text: [CLS] Mad ##ha ##hv (H ##TT ##P ##UR ##L), my first PhD student , successfully defended his thesis on tensor learning today . Con ##gratulations ! Dr , Ni ##mis ##hak ##avi ! One more addition to the academic lineage ##s of # US ##ER Fernando Pereira @ US ##ER Zack Ives NL ##P ##ro ##f ##tensor ##knowledge ##graph H ##TT ##P ##UR ##L 0 quotes 0 replies 0 re ##t ##weet ##s 2019 year 10 month 28 day 0 weekday 6 hour author is part ##ha _ p _t Bengal ##uru , India 0 account age 254 ##8 followers 74 total t ##weet ##s 687

label: 0, pred:0, text: [CLS] Came across this gorgeous on the outskirts of Bangalore today . You can see its body pattern change in a matter of minutes as it moved from the road to shrubs ! Amazing ! H ##TT ##P ##UR ##L 0 quotes 0 replies 2 re ##t ##weet ##s 202 ##1 year 1 month 2 day 5 weekday 16 hour author is part ##ha _ p _t Bengal ##uru , India 0 account age 254 ##8 followers 74 total t ##weet ##s 471

label: 0, pred:0, text: [CLS] It 's really amazing to inter ##n on Google Brain and work with @ US ##ER @ US ##ER Colin Ra ##fe ##f and Nicholas ! I had a wonderful experience for my last internship and am looking forward to my new internship at Brain Toronto ! Also thanks to my advisor Gary for his support of my internship H ##TT ##P ##UR ##L 0 quotes 0 replies 4 re ##t ##weet ##s 2019 year 4 month 22 day 0 weekday 19 hour author is Yac ##Q ##in ##UC ##SD New York , USA 0 account age 248 ##0 followers 56 total t ##weet ##s 43

label: 0, pred:0, text: [CLS] The general space the start ##u is operating in is healthcare data integration | 0 quotes 0 replies 0 re ##t ##weet ##s 2020 year 8 month 18 day 1 weekday 5 hour author is Isa _ Ki ##ko Melbourne , Australia 0 account age 440 followers 253 ##2 total t ##weet ##s

92
label: 0, pred:0, text: [CLS] @ US ##ER Daniel Be ##ag ##le ##ole Thanks for engaging ! But M * depends on f * , which is itself being learned . So the idea of a true M * doesn 't feel like a well defined concept | 0 quotes 0 replies 0 re ##t ##weet ##s 202 ##3 year 1 month 3 day 1 weekday 16 hour author is deep ##co ##hen Pittsburgh , PA 0 account age 228 ##0 followers 600 total t ##weet ##s 27

label: 0, pred:0, text: [CLS] exactly this ! H ##TT ##P ##UR ##L 0 quotes 0 replies 0 re ##t ##weet ##s 2019 year 8 month 9 day 4 weekday 12 hour author is David ##SB ##ati ##sta Berlin , Germany 0 account age 253 followers 53 ##4 total t ##weet ##s

103
label: 0, pred:0, text: [CLS] @ US ##ER Come back ! ##una 0 quotes 0 replies 0 re ##t ##weet ##s 202 ##2 year 5 month 13 day 4 weekday 17 hour author is m ##ora ##d ##del ##i Islamic Republic of Iran 0 account age 19 followers 143 total t ##weet ##s

34
label: 0, pred:0, text: [CLS] Great blog by Ian Good ##ell ##ow and Nicolas Paper ##not on ##c ##y ##bers ##ec ##urity and # M ##L : H ##TT ##P ##UR ##L 0 quotes 0 replies 0 re ##t ##weet ##s 202 ##1 year 7 month 14 day 2 weekday 19 hour author is v ##mo ##ilya ##kov Seattle , WA 0 account age 181 followers 280 total t ##weet ##s

10
label: 0, pred:0, text: [CLS] How expensive are data out ##ages really ? T ##LD ##R , more expensive than you think . H ##TT ##P ##UR ##L 0 quotes 0 replies 0 re ##t ##weet ##s 202 ##1 year 4 month 9 day 4 weekday 1 hour author is v ##ive ##k _ ## Santa Clara , CA 0 account age 61 followers 53 total t ##weet ##s

179
label: 0, pred:0, text: [CLS] @ US ##ER mat ##I Yeah !!! H ##TT ##P ##UR ##L 0 quotes 0 replies 0 re ##t ##weet ##s 202 ##3 year 1 month 16 day 0 weekday 15 hour author is Dr ##C ##isco Lisbon , Portugal 0 account age 217 ##1 followers 404 ##9 ##2 total t ##weet ##s

140
label: 0, pred:0, text: [CLS] What did I learn ? ! The OS ##S models are actually reasonably good , but they 're hard to use . I 'm a PhD student in this stuff and it took more Go ##og ##ing than I expected . Maybe this is the reason to use the public API ##s . 0 quotes 0 replies 0 re ##t ##weet ##s 202 ##2 year 10 month 9 day 6 weekday 22 hour author is m ##oi ##na ##de ##em San Fr ##ans ##isco Bay Area 0 account age 135 ##8 followers 160 ##41 total t ##weet ##s

119
label: 0, pred:0, text: [CLS] @ US ##ER Who will make it understand to my dear B ##JP supporters | 0 quotes 0 replies 0 re ##t ##weet ##s 202 ##2 year 4 month 22 day 4 weekday 16 hour author is Mo ##h ##d ##N ##ave ##d India 0 account age 70 ##5 ##4 followers 36 ##37 total t ##weet ##s

536
label: 1, pred:1, text: [CLS] :party _ pop ##per : ; party _ pop ##per : ; party _ pop ##per : Happy to officially announce the @ US ##ER : hugging _ face : AI Research Re ##side ##ncy Program , a great opportunity to launch or advance your career in machine learning research : rocket : Read more and / or apply here : H ##TT ##P ##UR ##L H ##TT ##P ##UR ##L 0 quotes 0 replies 226 re ##t ##weet ##s 202 ##2 year 3 month 22 day 1 weekday 16 hour author is do ##fu ##we ##kie ##la 0 account age 64 ##34 followers 33 ##4 total t ##weet ##s 430

label: 1, pred:1, text: [CLS] Cha ##GP ##T is the most impressive tech demo I 've ever seen . With the previous models I was still skeptical about the broader impact , but now that 's impossible ! Hu ##mble ##d by how close this get to human - like understanding . We 're not so special after all . 0 quotes 0 replies 45 re ##t ##weet ##s 202 ##2 year 12 month 4 day 6 weekday 6 hour author is den ##ny ##bri ##tz Japan 0 account age 50 ##8 ##19 followers 69 ##26 total t ##weet ##s

336
label: 1, pred:1, text: [CLS] The past ~ 8 month s I 've been working on a new company called Run ##loop together with @ US ##ER C ##rist ##ian St ##rat . We 're building simulation environments and R ##L agents for De ## ##F (Dec ##ent ##f ##ized Finance) . Really cool opportunity because : 0 quotes 0 replies 27 re ##t ##weet ##s 202 ##1 year 8 month 31 day 1 weekday 18 hour author is den ##ny ##bri ##tz Japan 0 account age 50 ##8 ##19 followers 69 ##26 total t ##weet ##s

935
label: 1, pred:1, text: [CLS] I have a math joke , but clearly , it should be obvious . 0 quotes 0 replies 12 re ##t ##weet ##s 2020 year 7 month 26 day 6 weekday 17 hour author is den ##ny ##bri ##tz Japan 0 account age 50 ##8 ##19 followers 69 ##26 total t ##weet ##s

754
label: 1, pred:1, text: [CLS] Can neural networks discover the vocabulary of a language from un ##se ##gment ##ed character sequences without supervision ? Work with Ka ##zu ##ya Ka ##wa ##ka ##mi and Phil Blu ##ns ##fm at @ US ##ER Deep ##M ##ind ##A ##I shows they can , and doing so improve ##s the language model H ##TT ##P ##UR ##L 1 / 20 quotes 0 replies 67 re ##t ##weet ##s 2019 year 6 month 25 day 1 weekday 9 hour author is red ##po ##ny London , England 0 account age 108 ##0 ##4 followers 42 ##6 total t ##weet ##s

351
label: 1, pred:1, text: [CLS] Late ##st surprising ##CL ##R 21 submission to stumble upon " Sell + Organ ##izing Intel ##lent Matter : A blue ##print for an AI generating algorithm " ! Me ##sm ##er ##sing emerge ##nt dynamics | each c ##xel contains an individual neural network . H ##TT ##P ##UR ##L H ##TT ##P ##UR ##L 0 quotes 0 replies 104 re ##t ##weet ##s 2020 year 10 month 8 day 3 weekday 14 hour author is ma ##x ##ja ##der ##berg London , England 0 account age 89 ##44 followers 280 total t ##weet ##s

164
label: 1, pred:1, text: [CLS] Karnataka is showing signs of the wrong kind of turn ##around . Know thy con ##found ##ers ! Daily tests are down by a s ##tag ##gering - 33 % . Wen ##from 1 , 8 la ##c to 1 , 2 la ##c daily tests in last two w ##ks . This then leads to R ##t and daily cases showing a false turn ##around . Stay safe ! ka ##ma ##ta ##ka , H ##TT ##P ##UR ##L H ##TT ##P ##UR ##L 0 quotes 0 replies 40 re ##t ##weet ##s 202 ##1 year 5 month 15 day 5 weekday 8 hour author is x ##M ##roh ##ak ##G ##up ##ta C ##LE 0 account age 140 ##7 ##7 followers 61 ##8 total t ##weet ##s

143
label: 1, pred:1, text: [CLS] Con ##tro ##vers ##ial opinion : Spa ##mming the whole internet with low quality AI - generated content will turn out to be a great thing in the long - run . 0 quotes 0 replies 13 re ##t ##weet ##s 202 ##2 year 12 month 5 day 0 weekday 13 hour author is den ##ny ##bri ##tz Japan 0 account age 50 ##8 ##19 followers 69 ##26 total t ##weet ##s

More images [here](#)