



Twitter Sentimental Analysis

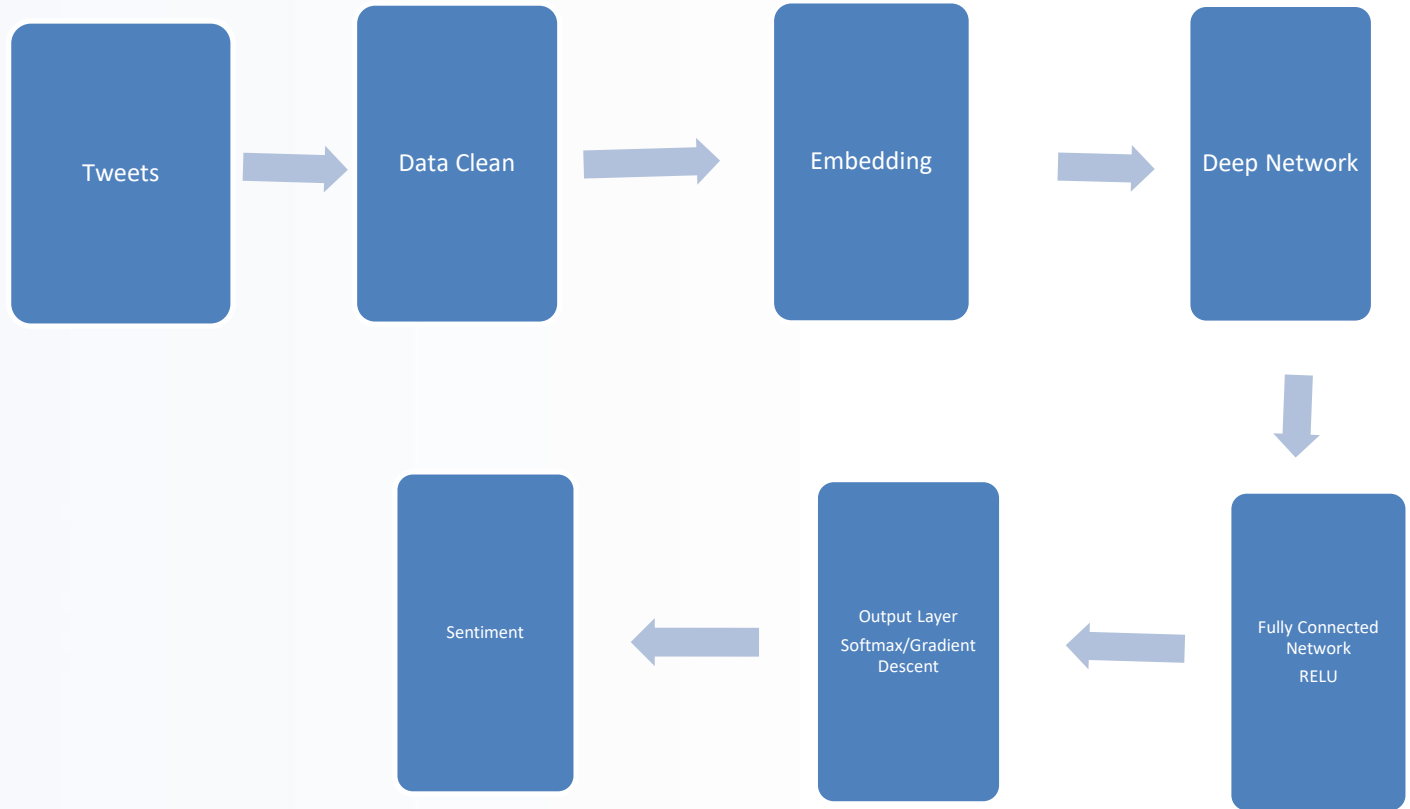
Nisha Pepsi Selvarajan



Sentiment Classification Problem

- *Sentiment classification is the task of looking at a piece of text and telling if someone likes or dislikes the thing they're talking about.*
- The input X is a piece of text and the output Y is the sentiment which we want to predict.

Architecture





Word Embedding

- **Word Embedding** is a representation of text where words that have the same meaning have a similar representation. In other words it represents words in a coordinate system where related words, based on a corpus of relationships, are placed closer together.
- lookup table to map the words represented by numeric indexes to their dense vector representations.





Deep Network/Fully Connected Layer

- Deep network takes the sequence of embedding vectors as input and converts them to a compressed representation. The compressed representation effectively captures all the information in the sequence of words in the text. The deep network part is an RNN.

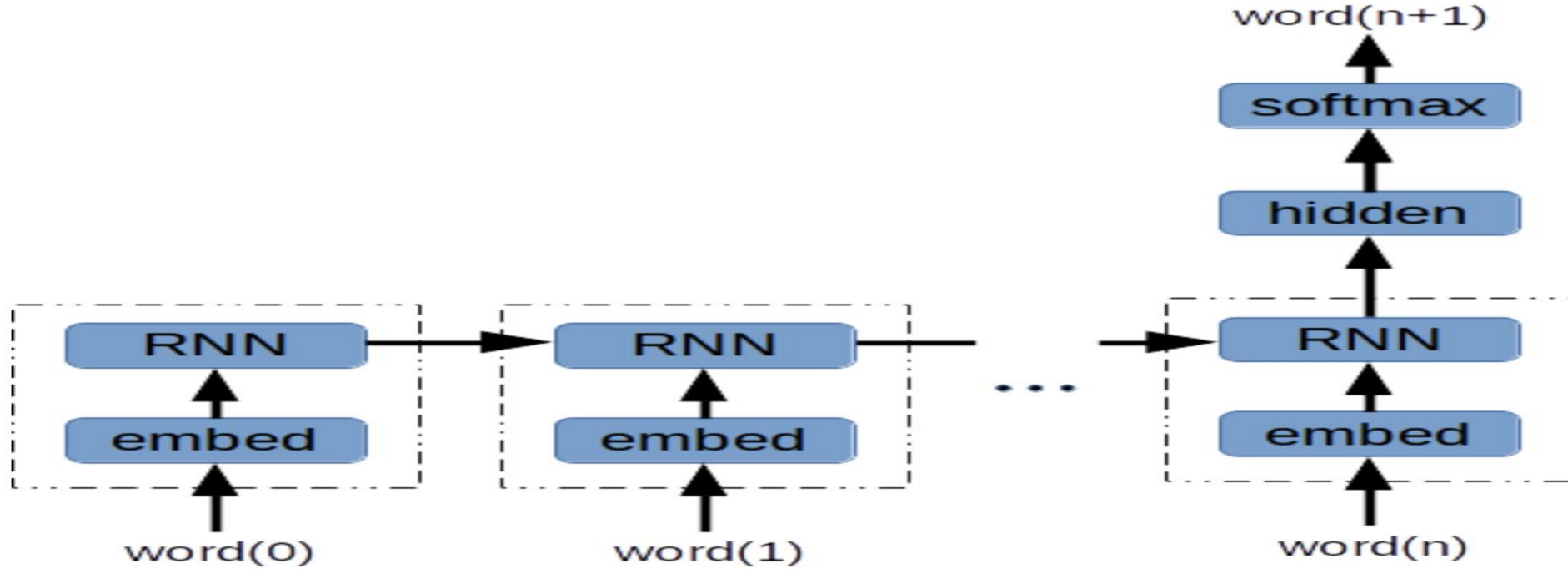


Deep Network/Fully Connected Layer

- Deep network takes the sequence of embedding vectors as input and converts them to a compressed representation. The compressed representation effectively captures all the information in the sequence of words in the text. The deep network part is an RNN.
- Hyper parameters: RNN Hidden Layer size/
batch_size / epochs/learning rate(AdamOptimizer)



Architecture



Layer sizes: embedding layer, number RNN cells, hidden layer

Input Tweets

```
training.1600000.processed.noemoticon.csv
"4","2193526461","Tue Jun 16 08:34:40 PDT 2009","NO_QUERY","xLisaaLollipop","http://xlisaatown.mymincity.com/ &lt;3 Now I'm tha bomb "
"4","2193526487","Tue Jun 16 08:34:40 PDT 2009","NO_QUERY","thenewspies","@shelovesshoes hehehe don't keep up with new books, just seee &amp; buy soo yes thank you darlin
"4","2193526499","Tue Jun 16 08:34:40 PDT 2009","NO_QUERY","mattamj","currently going to 'the cave' then Steps? Work, then C-bus. Fun ahead!"
"4","2193526573","Tue Jun 16 08:34:40 PDT 2009","NO_QUERY","gavin714","packing and getting ready to go to wichita falls! pretty pumped for some koozies, beer, and pool ti
"4","2193526670","Tue Jun 16 08:34:41 PDT 2009","NO_QUERY","Mrs_Nick07","Going to listen to LVA TT to calm myself.. "
"4","2193526672","Tue Jun 16 08:34:41 PDT 2009","NO_QUERY","sytheveenje","#computerarts Nice, people who don't know can just copy it ;) Great contest! Pick me!"
"4","2193526755","Tue Jun 16 08:34:41 PDT 2009","NO_QUERY","faceman101","@damoeldiablo PM myself and Doozie your email address dude. "
"4","2193526868","Tue Jun 16 08:34:42 PDT 2009","NO_QUERY","ctycty","bye twitter . restin' time ! nite "
"4","2193526869","Tue Jun 16 08:34:42 PDT 2009","NO_QUERY","ERR0rR","screw you NZ!! especially middle finger credit to daniel vitory!....lol Sri lanka thru to the semis..
"4","2193526890","Tue Jun 16 08:34:42 PDT 2009","NO_QUERY","MSB7","@NatiKuhnNews Happy bday! "
"4","2193526924","Tue Jun 16 08:34:42 PDT 2009","NO_QUERY","edemayne yes! the 3rd. want a copy? "
"4","2193526931","Tue Jun 16 08:34:42 PDT 2009","NO_QUERY","Killer_Burrito","Okay so the plan is to go to hot topic and buy all the harry potter stuff I can! "
"4","2193526951","Tue Jun 16 08:34:42 PDT 2009","NO_QUERY","tigertuft","#digitalbritain on page 11 and bored already....mind you, nothing can hold my attention these da
"4","2193526999","Tue Jun 16 08:34:42 PDT 2009","NO_QUERY","Foolishoverlord","@shwood loved the cameo, made me chuckle! along with your SXSW talk over the past couple of
"4","2193527055","Tue Jun 16 08:34:42 PDT 2009","NO_QUERY","stephgramling","Headed 2 Islands of Adventure...it is SO hot!!! Can't wait 2 ride Spiderman I love that ride!
"4","2193527300","Tue Jun 16 08:34:44 PDT 2009","NO_QUERY","miss_extreme","WE GET TO KEEP OUR HOUSE! "
"4","2193527364","Tue Jun 16 08:34:44 PDT 2009","NO_QUERY","lovelymskk","@LyssaBrooke really? i think i'm getting the hang"
"4","2193527366","Tue Jun 16 08:34:44 PDT 2009","NO_QUERY","brimil","@jerseyshorejen haha I know that's why I said I like Comcast so much, because of their Twitter crew "
"4","2193527430","Tue Jun 16 08:34:44 PDT 2009","NO_QUERY","katiehogan","good morning. biking to hanna's for breakfast "
"4","2193527454","Tue Jun 16 08:34:44 PDT 2009","NO_QUERY","Aksland","Haugesund airport, waiting for the Oslo plane. Ndc09 next "
"4","2193527491","Tue Jun 16 08:34:45 PDT 2009","NO_QUERY","TheStarr","Why Would I Go Hard When This Sh*t Is So Easy?!! - Nicki Minaj "
"4","2193527664","Tue Jun 16 08:34:46 PDT 2009","NO_QUERY","shynquietsshadow","enjoyinglake tahoe w/my redhead "
"4","2193527670","Tue Jun 16 08:34:46 PDT 2009","NO_QUERY","camdana cole","To Whoever Invented Philly Cheese Steaks, I Thank You, And Right Now So Does My Tummy. "
"4","2193527674","Tue Jun 16 08:34:46 PDT 2009","NO_QUERY","Tweetie_bbe","GUNA LISTEN 2 MUCH BETTER NEXT C WHAT JOE HAS 2 HIDE "
"4","2193527685","Tue Jun 16 08:34:46 PDT 2009","NO_QUERY","COOLIE0B0JULIE0","is going tanning hopfully get a lil more darker... love yas, c ya xoxo"
"4","2193527751","Tue Jun 16 08:34:46 PDT 2009","NO_QUERY","jacodenn","@Deeders_ I hope it all goes well for you at the Dr. today "
"4","2193527841","Tue Jun 16 08:34:46 PDT 2009","NO_QUERY","astroross","@DailyStatsmobi it was an amazing race and the more I think about it the more I want to try it aga
"4","2193527854","Tue Jun 16 08:34:46 PDT 2009","NO_QUERY","lLisAnne","@AmyLynne18 I love you.. can't wait for this weekend "
"4","2193527860","Tue Jun 16 08:34:46 PDT 2009","NO_QUERY","captainboo","@evienyc What are truancy laws like in the US? I do hope those four biches get in trouble "
"4","2193527876","Tue Jun 16 08:34:47 PDT 2009","NO_QUERY","MKeithHarris","@edenniswords good times indeed "
"4","2193527949","Tue Jun 16 08:34:47 PDT 2009","NO_QUERY","Diana54x","I had a good work out day today "
"4","2193527998","Tue Jun 16 08:34:47 PDT 2009","NO_QUERY","MsMaria22","just got off work! i'm exhausted.... worked real hard today. JK LMAO. i only worked 3 hrs. "
```

DataSet: <http://help.sentiment140.com/>:0 - the polarity of the tweet (0 = negative, 2 , neutral, 4 = positive) 2 million

Word With Dimension

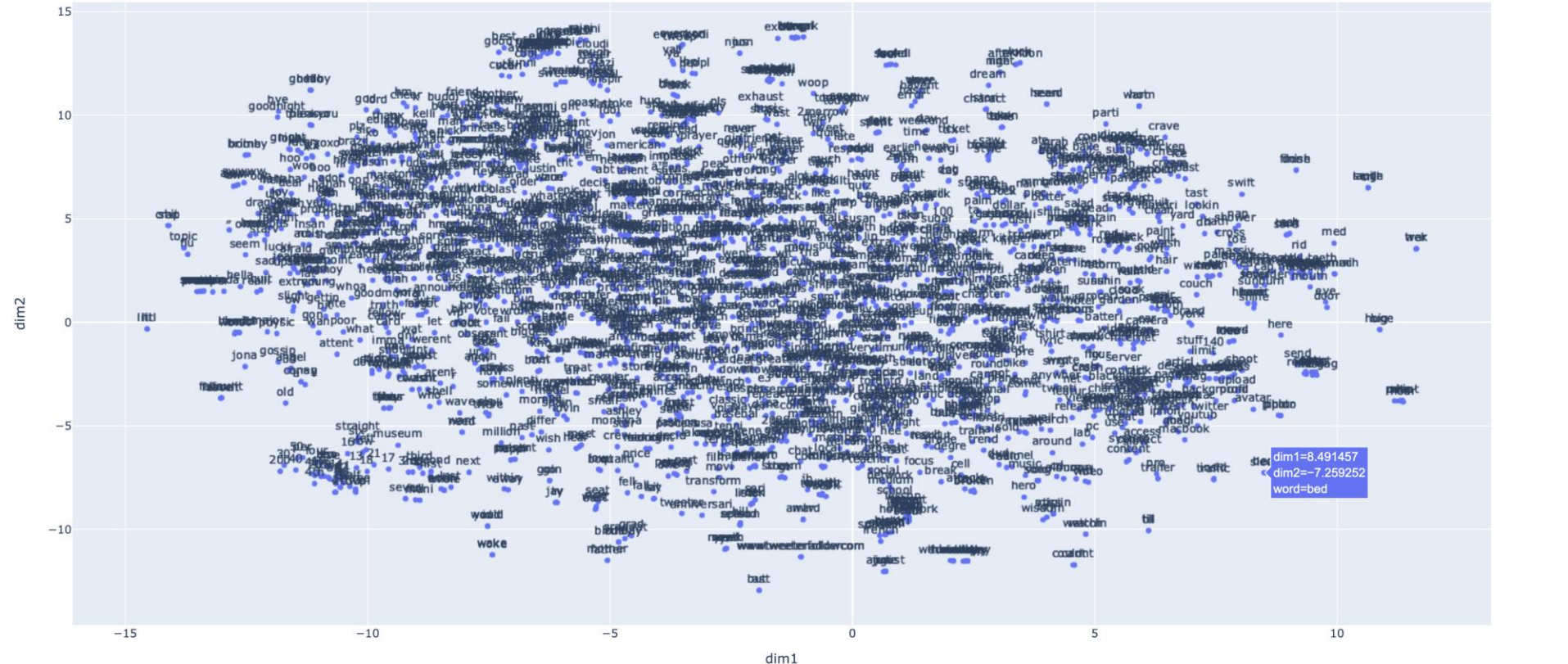
2DEmbedding.csv

```
1 ,word,dim1,dim2
2 0,im,15.472679027293568,-3.7550711550976605
3 1,go,-1.0746235100754926,11.903285157850757
4 2,get,-9.277530556815188,10.20616226905378
5 3,day,-7.991432134169957,-0.10192148713762067
6 4,good,9.303125300473399,-5.8646965607476815
7 5,work,-5.1722989863109134,0.5317016612926622
8 6,like,4.019880703846125,-5.775959525316118
9 7,love,8.164106988126889,-6.7424471181501024
10 8,got,5.657626905180401,7.195427417527587
11 9,u,-9.228081458346205,-6.621883036042413
12 10,dont,-10.144358004614258,8.984030916356124
13 11,today,8.509442209014297,4.343743374628192
14 12,time,-7.860448892742261,-0.03277000898649049
15 13,cant,-11.367901713470607,-3.940236629932148
16 14,thank,-7.43163217542907,-10.147311954787805
17 15,want,-6.883991674063092,-6.1710093752337105
18 16,miss,6.206940195369886,7.842848465270505
19 17,back,-1.3124416898120999,7.682891539386801
20 18,one,8.508185261505394,-8.67765922664008
21 19,lol,10.970133015715728,3.058320417098486
22 20,know,-4.319917811597841,-6.29113856609873
23 21,see,-2.2992120090830257,10.15951096167919
24 22,feel,-8.678200125237412,7.1400266896907825
25 23,think,9.895439985684039,0.01877385309543511
26 24,real,10.957356422178513,-0.08379274766542294
27 25,well,12.223742291885387,3.8693401705826087
28 26,hope,-6.213656491458564,-2.237369559343286
29 27,night,-10.125718642400212,-5.303278340966
30 28,watch,6.864768568747758,7.788471712700728
31 29,need,-6.870252827850719,-6.152261641714214
32 30,still,10.711613844945152,0.09765824969956083
```

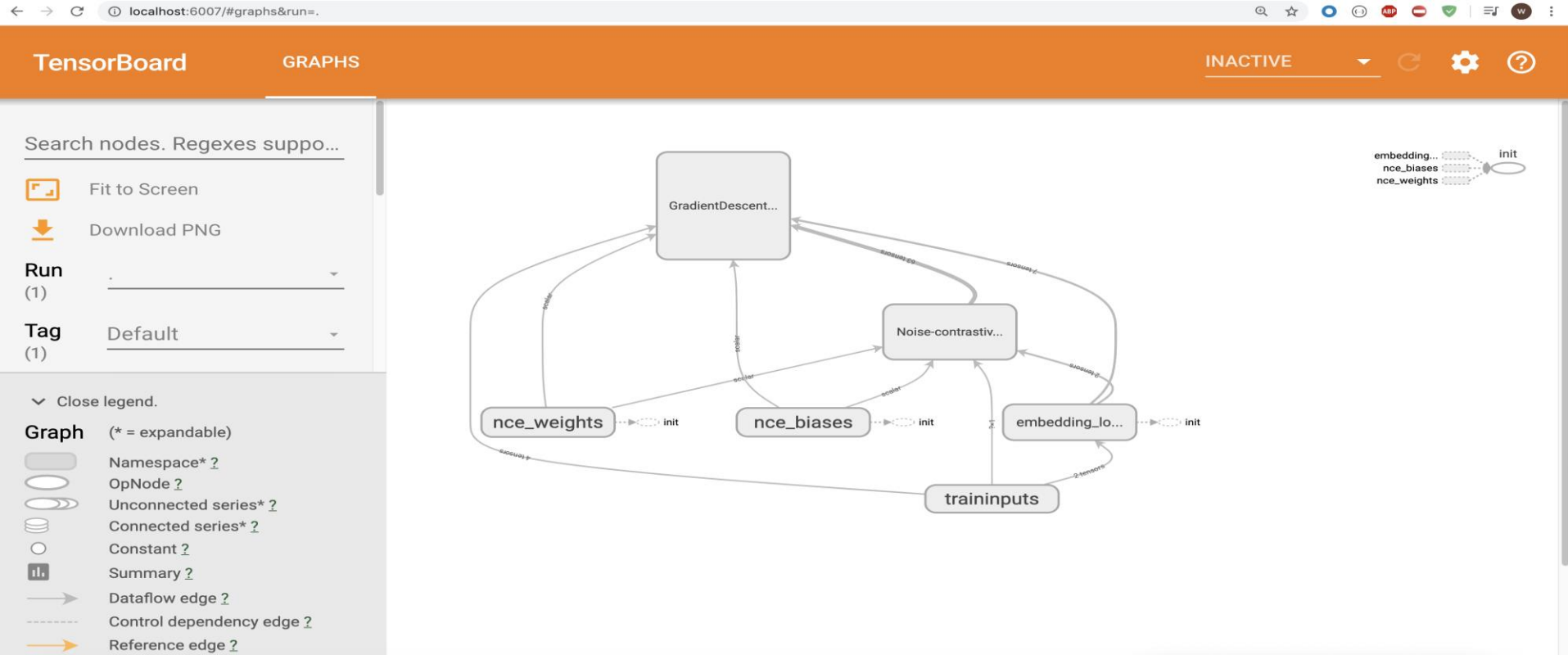


Word Embedding

Embedding Two-D Plot



Word Embedding



Noise-contrastive estimation Loss, Embedding Lookup



localhost:6006/#scalars



TensorBoard

SCALARS

CUSTOM SCALARS

INACTIVE



- ☐ Show data download links
- ☒ Ignore outliers in chart scaling

Tooltip sorting method: default

Smoothing

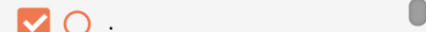


Horizontal Axis

STEP RELATIVE WALL

Runs

Write a regex to filter runs

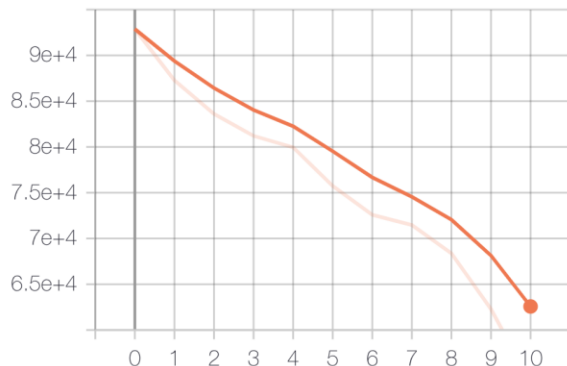


TOGGLE ALL RUNS

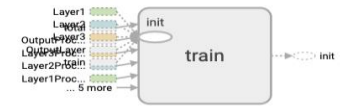
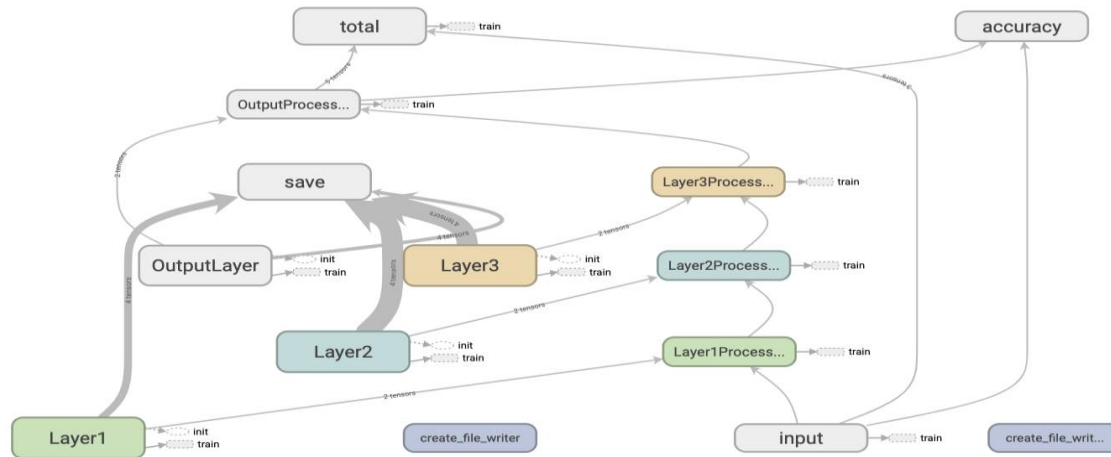
Filter tags (regular expressions supported)

loss

loss
tag: loss/loss/scalar_summary



Sentimental Classification



Sentimental Classification

TensorBoard

SCALARS

CUSTOM SCALARS

INACTIVE

- ☐ Show data download links
- ☒ Ignore outliers in chart scaling

Tooltip sorting method: default

Smoothing

0.6

Horizontal Axis

STEP

RELATIVE

WALL

Runs

Write a regex to filter runs



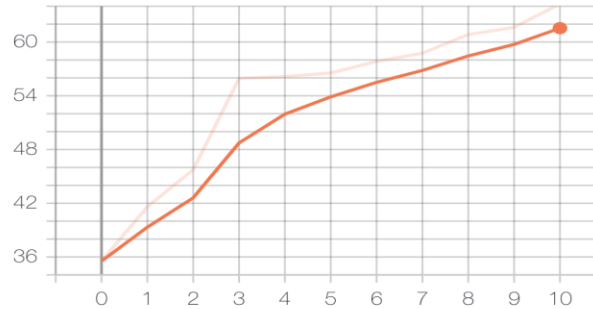
TOGGLE ALL RUNS

/tmp/dsads

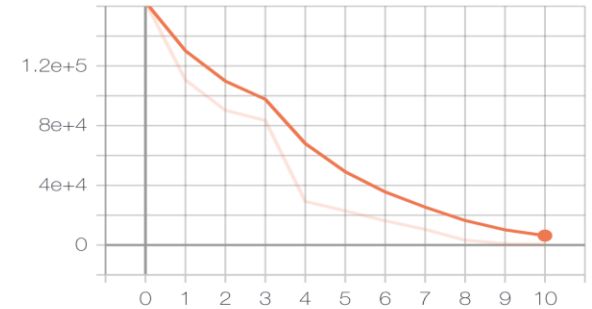
Filter tags (regular expressions supported)

loss

accuracy
tag: loss/accuracy/scalar_summary



loss
tag: loss/loss/scalar_summary





Sentimental Classification

```
/Users/nselvarajan/untitledyu/bin/python "/Users/nselvarajan/Downloads/sentiment-analysis-tensorflow-master 3/Test1.py"
```

```
WARNING:tensorflow:From /Users/nselvarajan/untitledyu/lib/python3.6/site-packages/tensorflow_core/python/ops/resource_variable_ops.py:1635: calling BaseResourceVariable.__init__ (from tensorflow.python.ops.resource_v
```

```
Instructions for updating:
```

```
If using Keras pass *_constraint arguments to layers.
```

```
2020-03-12 02:02:20.810593: I tensorflow/core/platform/cpu_feature_guard.cc:142] Your CPU supports instructions that this TensorFlow binary was not compiled to use: AVX2 FMA
```

```
2020-03-12 02:02:20.823012: I tensorflow/compiler/xla/service/service.cc:168] XLA service 0x7fbeb30d590 initialized for platform Host (this does not guarantee that XLA will be used). Devices:
```

```
2020-03-12 02:02:20.823026: I tensorflow/compiler/xla/service/service.cc:176] StreamExecutor device (0): Host, Default Version
```

```
Epoch 1 completed. Total loss: 220976.748046875
```

```
Epoch 2 completed. Total loss: 126231.708984375
```

```
Epoch 3 completed. Total loss: 88405.412109375
```

```
Epoch 4 completed. Total loss: 70276.42138671875
```

```
Epoch 5 completed. Total loss: 57585.16748046875
```

```
Epoch 6 completed. Total loss: 46172.174072265625
```

```
Epoch 7 completed. Total loss: 39570.009765625
```

```
Epoch 8 completed. Total loss: 31314.65625
```

```
Epoch 9 completed. Total loss: 28257.273193359375
```

```
Epoch 10 completed. Total loss: 24124.6376953125
```

```
Accuracy using test dataset: 0.6787148714065552
```

```
Process finished with exit code 0
```




Predict Sentiment

- Given a text, load the word dict model from training:
- Count the number of positive words in the text.
- Count the number of negative words in the text.
- If the number of positive word appearances is greater than the number of negative word appearances return a positive sentiment, conversely, return a negative sentiment.

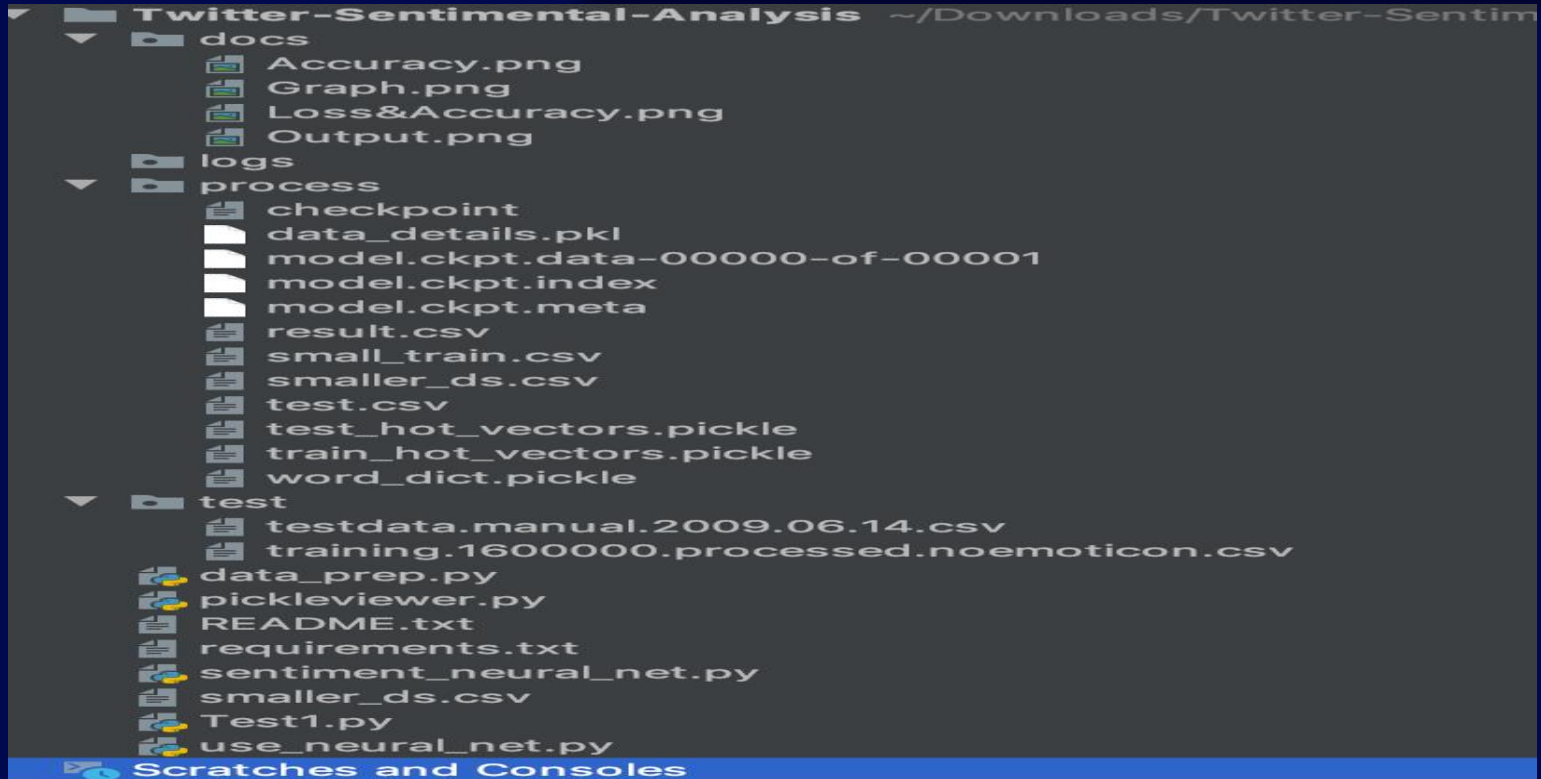


Output

```
2020-03-12 02:32:11.685229: I tensorflow/compiler/xla/service/service.cc
2020-03-12 02:32:11.685241: I tensorflow/compiler/xla/service/service.cc
Epoch 1 completed. Total loss: 605643.794921875
Epoch 2 completed. Total loss: 255532.17333984375
Epoch 3 completed. Total loss: 179960.20336914062
Epoch 4 completed. Total loss: 127754.30297851562
Epoch 5 completed. Total loss: 101332.6337890625
Epoch 6 completed. Total loss: 87579.12744140625
Epoch 7 completed. Total loss: 94325.10192871094
Epoch 8 completed. Total loss: 60958.75518798828
Epoch 9 completed. Total loss: 40296.262145996094
Epoch 10 completed. Total loss: 30400.178588867188
Accuracy using test dataset: 0.5301204919815063
[1]
Positive: This is my very first neural network project, I love it.
[0]
Negative: Corona virus is very scary and distrubing
[0]
Negative: I hate you!
[1]
Positive: Game of thrones is intresting and exciting

Process finished with exit code 0
```

Project Structure





Requirements

requirements.txt

```
1 tensorflow==2.1.0rc2
2 scipy==0.18.1
3 pandas==1.0.1
4 matplotlib==3.1.3
5 plotly==4.5.3
6 numpy==1.17.0
7 nltk==3.4.5
8 scikit_learn==0.18.1
9
```




Thank You!!!