Code-Mixed Sentimental Analysis ID-18

Faculty: Dr.Dipti Misra Sharma
Mentor: Urmi Ghosh
Team:
Archana, Univ of Mumbai
Rakshit, NIT Trichy
Prakriti, Univ of Hyd
Prudhvi, mroads

What is code-mixed sentimental analysis

- Code-mixing is the mixing of two or more languages or language varieties in speech
- We have taken two languages to be Hindi and English
- Example: Iss flying parinday ke feathers gin k batao to manu

Sentimental Analysis: computationally identifying and categorizing opinions expressed in a piece of text

Ex: Modiji this is haalat of your mitron now-a-days

Data collection and Preprocessing

- Extracted tweets with twitter api
- We have use LID [Language Identifier] to identify Hindi and English words in tweet
- Filtered data which are code-mixed
- Annotated the tweets with sentiments
- Tweets taken:
  - Negative : 160
  - Positive : 200
  - Neutral : 158

Annotated Tweets

		Tweet	Neutral	Positive	Negative
0	0	Jaise ek film se koi superstar nahi ban jata waise ek flop se kisika career khatam nahi hota. So chill guys. #sanju #Race3	0	1	0
1	1	And on the other hand...... Bring out your desi swag with #GoodManDiLaaltain! https://t.co/mFTHkUgSxm https://t.co/vr7JAuK4sA	0	1	0
2	2	Ishq di..... baajiyaan..... jeetiyaan.... naa haariyaan Share the love !!!! #Ishqdibaajiyaan #Soorma https://t.co/o4LGNZYZvQ	0	1	0
3	3	RT @StarGoldIndia: College sweetheart jiske Prem ne kiye hearts break hazaron, lekar aa rahi hai entertainment dugna with #Judwaa2OnStarG	0	1	0
4	4	Ab bas theatre mein aane ka intezaar rahega! What an endearing trailer https://t.co/yzJmSyGF9a	0	1	0
5	5	Dussehra ka ENTERTAINMENT hoga kal double! #Judwaa2KiDussehra #Judwaa2 @Varun_dvn @Asli_Jacqueline @foxstarhindi https://t.co/Y1rB1PNgIO	0	1	0
6	6	RT @foxstarhindi: Dugna fun with #Judwaa2KiDussehra! Book now: https://t.co/Mp5OLL7upP Use code JUDWAA2 @Varun_dvn @Asli_Jacqueline @Taapse	0	1	0
7	7	Celebrations hogayi shuru #Judwaa2KiDussehra ke saath! Book now! https://t.co/zeewLc5Ngd Use code JUDWAA2 https://t.co/4ygn8uu4f3	0	1	0
8	8	RT @foxstarhindi: Bappa aur Raja ki jodi aarahi hai kal! #SunoGanpatiBappaMorya song out tom at 10 am @Varun_dvn @Asli_Jacqueline @taapsee	1	0	0
9	9	Ek baar se dil nahi bharta, mud ke dekho humein dobara! Flying high with tan tana tan tan tan tara #Judwaa2 https://t.co/DjBKsICUiv	0	1	0
10	10	@abhilashthapli AAM land mein toh tha mera birthday! AAM girl is happy girl #ThankYou	0	1	0
11	11	@WardaNadiadwala Block toh bust hone hi wala hai! And we both shall be tan tanaing with it! Thankq so much hotness	0	1	0
12	12	@Varun_dvn Fire cracker toh abhi phatne wala hai cake kha kha ke! Chal dubbing slot diya as return gift. Thankq so https://t.co/nEH51TLXnT	0	1	0
13	13	@Riteishd Thankq so much! Koshish jaari hai thanks for always being so encouraging :)	0	1	0
14	14	@rajpalofficial Sir woh toh apna saath mein hoga is saal! Thankq so much :)	0	1	0
15	15	Another feather in @isro cap! Aap toh kamaal karte ho! Don't know about this world but in space we r slaying! #SouthAsiaSatellite	0	1	0
16	16	With this #NaamShabana team bids bye to Delhi. Kuch Dino mein phir milte hai, National Security agenda ke saath :) https://t.co/vqZUOF2pz3	1	0	0
17	17	RT @FFW_Official: @akshaykumar's dailouge Itne dur se aaya hu, kuch toh karne do is @taapsee's favorite dialogue. Which one is your favor	0	1	0
18	18	RT @PlanC_Studios: Teri haathon ki taraf, mere haathon ka safar, Rozana. Feel the love with #Rozana. @taapsee @BajpayeeManoj @PrithviOffic	0	1	0
19	19	My Dilli ka Launda @Saqibsaleem ab toh picture dekh li hai, tell me if that reminded u of a time when u wanted to elope! #RunningShaadi	0	0	1

LID Tagged Output

	0\O
0\O	#\O Bologuruji\Hin=बोलोगुरुजी guruji\Hin=गुरुजी kya\Hin=क्या aap\Hin=आआप् ne\Hin=ने kabhi\Hin=कभी socha\Hin=सोचा tha\Hin=ता ke\Hin=क aap\Hin=आआप् ka\Hin=का channel\Eng India\Eng me\Eng top\Hin=टोप hogan\Hin=होगन .\O
1\O	@\O TechnicalGuruji\Eng Guruji\Eng Wo\Hin=वो Muuch\Hin=मूच (\O Mustache\Hin=मुस्चे)\O Rakh\Hin=रख Kar\Hin=कार Aap\Hin=आआप् Bade\Hin=बड़े Chutiye\Hin=चुटिय Lag\Hin=लाग Rahe\Hin=रह Ho\Hin=ह ?\O Kyun\Hin=क्यूँ ?\O #\
2\O	\O Honor\Eng 7c\Eng or\Eng Mi\Eng 5\O which\Eng is\Eng better\Eng ?\O #\O Bologuruji\Hin=बोलोगुरुजी
3\O	#\O BoloGuruji\Hin=बोलोगुरुजी Hlo\Hin=हो Bhai\Hin=भाई maraa\Hin=मरा saval\Hin=सवाल hai\Hin=हैं ke\Hin=क abhi\Hin=अभी mara\Hin=मरा pass\Eng 4g\Eng phone\Eng hai\Hin=हैं magar\Hin=मगर jab\Eng 5g\Eng aye\Hin=आये tab\Hi
4\O	@\O TechnicalGuruji\Eng Love\Hin=लव you\Hin=यूं guruji\Hin=गुरुजी kanpur\Hin=कानपुर meet\Hin=मीत up\Hin=उप kab\Hin=काब kab\Hin=काब ????\O love\Hin=लव your\Eng new\Eng look\Eng please\Eng up\Eng the\Eng question\Eng ...\O #\O Bologuruji\Hin=बोलोगुरुजी \O
5\O	@\O Bhavu2010\Eng @\O SadhviDeva\Hin=सध्विदेव Jiii\Hin=जी haaa\Hin=हा bapujiii\Hin=बापूजी ko\Hin=को is\Hin=इस shadyantra\Hin=शड्यांत्र me\Hin=मे isliye\Hin=इसलिये hii\Hin=हैं fasayaaa\Hin=फ़साया gaya\Hin=गया takii\Hin=तकी v
6\O	#\O Bologuruji\Hin=बोलोगुरुजी bhai\Hin=भाई hai\Hin=हैं apne\Hin=अपने bh\Hin=भी apple\Hin=एप्ल्अ ki\Hin=की tarah\Hin=तरह notch\Eng rakh\Hin=रख li\Hin=ली choti\Hin=चोटी si\Hin=सी ..\O
7\O	#\O Bologuruji\Hin=बोलोगुरुजी aap\Hin=आआप् apna\Hin=अप्ना number\Hin=नंबर de\Hin=दे do\Hin=दो sir\Hin=सिअर ji\Hin=जी please\Hin=प्लीअज् sir\Hin=सिअर sir\Hin=सिअर
8\O	#\O Bologuruji\Hin=बोलोगुरुजी guruji\Hin=गुरुजी maine\Hin=मैंने halhi\Hin=हली mai\Hin=माई realme\Hin=रिअलमे 1\O ka\Hin=का 4+64\O wala\Hin=वाला variant\Eng buy\Eng kiya\Hin=किया hai\Hin=हैं kya\Hin=क्या yeh\Hin=यह mobile\Hin
9\O	#\O Bologuruji\Eng which\Eng phone\Eng to\Eng buy\Eng Huawei\Eng p20\Eng lite\Eng or\Eng Moto\Eng g6\Eng other\Eng love\Eng your\Eng videos\Eng and\Eng love\Eng you\Hin=यूं bhai\Hin=भाई
10\O	\O Yeh\Hin=यह Kya\Hin=क्या hua\Hin=हुआ gaurav\Hin=गौरव Bhai\Hin=भाई ko\Eng 😳\EMT @\O TechnicalGuruji\Eng @\O emNikhil\Hin=इमनिखिल bro\Hin=ब्रो you\Hin=यूं forgot\Eng to\Eng blur\Eng it\Eng .\O #\O BoloGuruji\Hin=बोलोगुरुजी https\Eng :/\O t\Eng .\O co\Hin=को \O hvoPgay0qT\Hin=ह्वोपगाय0क्त \O
11	गरूजी\O आपने\O दाढ़ी\O निकाल\O दी\O इस\O वजह\O से\O क्या\O आपको\O अजिब\O नहीं\O लग\O रहा\O क्यों\O की\O मुझे\O भी\O न\O चाहते\O हुए\O passport\Eng साइज\O फोटो\O के...\O https\Eng :/\O t\Eng .\O co\Hin=को \O 3lfxPggKPe\Hin
12\O	#\O BoloGuruji\Hin=बोलोगुरुजी Please\Eng start\Eng a\Eng gaming\Eng channel\Eng I\Eng would\Eng love\Eng seeing\Eng some\Eng gameplay\Eng videos\Eng by\Eng you\Eng by\Eng the\Eng way\Eng love\Eng your\Eng vide

Algorithms used and results

Logistic Regression: [0-Neutral, 1-Positive, 2-Negative]

Testing Accuracy: 52%

Classification report:

Class	precision	recall	f1-score	support
0	0.65	0.34	0.45	64
1	0.48	0.89	0.63	70
2	0.50	0.06	0.11	32
avg / total	0.55	0.52	0.46	166

SVM:

Classification report:

Testing Accuracy : 42%

Class	precision	recall	f1-score	support
NEGATIVE	0.63	0.43	0.51	60
NEUTRAL	0.42	0.55	0.47	40
POSITIVE	0.50	0.56	0.53	48
avg / total	0.53	0.51	0.51	148

Naive Bayes:

Classification report:

Testing Accuracy: 54 %

Class	precision	recall	f1-score	support
NEGATIVE	0.63	0.62	0.62	60
NEUTRAL	0.52	0.38	0.43	40
POSITIVE	0.50	0.62	0.56	48
avg / total	0.56	0.55	0.55	148

Sentiment Results

```
In [348]: model.predict(tfidf_transformer.transform(vec.transform(["Arre chalo kisiko mera sarcasm samajh aaya"])))
Out[348]: array([2])

In [349]: model.predict(tfidf_transformer.transform(vec.transform(["kya bakwaas is this"])))
Out[349]: array([0])

In [350]: model.predict(tfidf_transformer.transform(vec.transform(["This is kya hai"])))
Out[350]: array([1])
```

Classes:
  0- Neutral
  1 - Positive
  2 - Negative

Thank You