Twitter corpus about technology for sentiment and emotion analysis with automatic and manual annotation

Building Language Resources

Oihane Cantero, Julen Etxaniz and Jose Javier Saiz

Dataset slice example

id	text	adversarial_text	sentiment_auto	sentiment_annotator	emotion_auto	emotion_annotator
1488301 0603347 59939	just put a CD into my MacBook to burn it and m...	just put a CD into my MacBook to burn it and m...	positive (0.532)	positive	joy (0.921)	joy
1488301 0576504 38144	oh yeah tesla well what about a car that just ...	oh yeah tesla well what about a automobiles th...	neutral (0.655)	neutral	joy (0.450)	anger
1488301 0401600 43008	#100DaysOfCode Haven’t updated in a while due ...	#100DaysOfCode Haven updated in a while due to...	neutral (0.500)	neutral	sadness (0.691)	joy

Index

- Description
- Methodology
- Data Statement
- License
- Guidelines
- Inter Annotator Agreement
- Adversarial examples
- Automatic analysis
- Conclusions

Description

- 140 monolingual tweets in English
- Topic: Technology
- Sentiment analysis (positive, negative, neutral)
- Emotion detection (joy, sadness, anger, optimism)
- Automatic and manual annotation
- Regular text and adversarial text

Methodology

- 140 tweets about technology
- Exclude retweets, replies, quotes, links to avoid repetitions or spam
- 20 annotations in common for ITA and guidelines
- 40 annotations individually after deciding guidelines
- Compare manual and automatic annotations
- Compare automatic annotations of adversarial examples

Data Statement

Curation Rationale	Obtain opinions with pronounced sentiments
Language Variety	English BCP 47 code (en)
Speaker Demographic	Male (70.4%)
	18 - 49 years old (78.7%)
	United States
Speech Situation	Spontaneous; Time sensitive text
Text Characteristics	Context “Interests and Hobbies: Technology"

License

- The resource and code are publicly available at GitHub.
- Problems with new Twitter policy
- Remove tweet text and only keep ID
- The code is licensed under the MIT open source license.
- Non-code materials provided under terms of the CC BY 4.0 license.

Initial Guidelines

- Select only one label for each task
- Read one tweet and annotate sentiment and emotion
- 3 labels for sentiment (positive, negative, neutral)
- 4 labels for emotion (joy, anger, sadness, optimism)

Inter Annotator Agreement

- 20 annotations in common for ITA
- Calculate agreement and kappa in sentiment and emotions
- Look at individual examples for disagreement
- Update guidelines if necessary
- Repeat annotation with new guidelines
- Recalculate agreement and kappa to see improvement
- Select the majority class for the final corpus

Inter Annotator Agreement

- Initial agreement and kappa values for each task
- Better than expected with very few guidelines
- Not good enough, update guidelines

	sentiment agreement	**sentiment kappa**	**emotion agreement**	**emotion kappa**
julen - oihane	0.70	0.54	0.55	0.33
julen - javier	0.60	0.40	0.40	0.13
javier - oihane	0.60	0.39	0.65	0.46
average	0.63	0.44	0.53	0.31

Updated Guidelines

- Misunderstanding some words (burn CD)
- Misunderstanding some symbols (<3 = <3) (> = >)
- Look or ask for the meaning

EMOTION		**SENTIMENT**	
Problem	**Solution**	**Problem**	**Solution**
If not sure / No emotion	Less specific: Joy	Irony / Sarcasm	Sentiment opposed to the overall sentiment. Either “positive” or “negative”
Joy / Optimism	More general: Joy	Contradicting sentiments	Overall sentiment

Adversarial examples

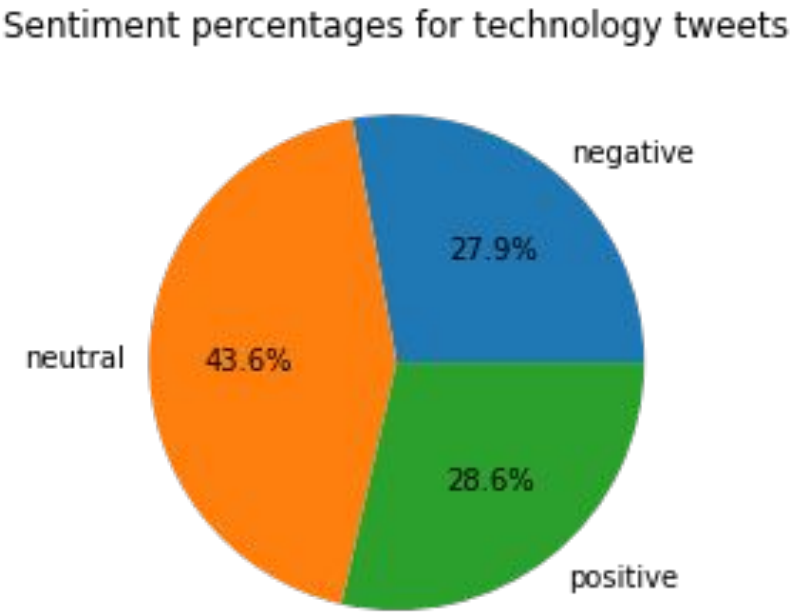
- Create adversarial tweets automatically if possible
- Create adversarial tweets manually if necessary
- Measure the impact in the accuracy of the models
- The aim is to confuse the models
- Keep the manual label, no need to annotate again
- For example, substituting words with their synonyms
- Add these adversarial examples to our final resource

Automatic Analysis

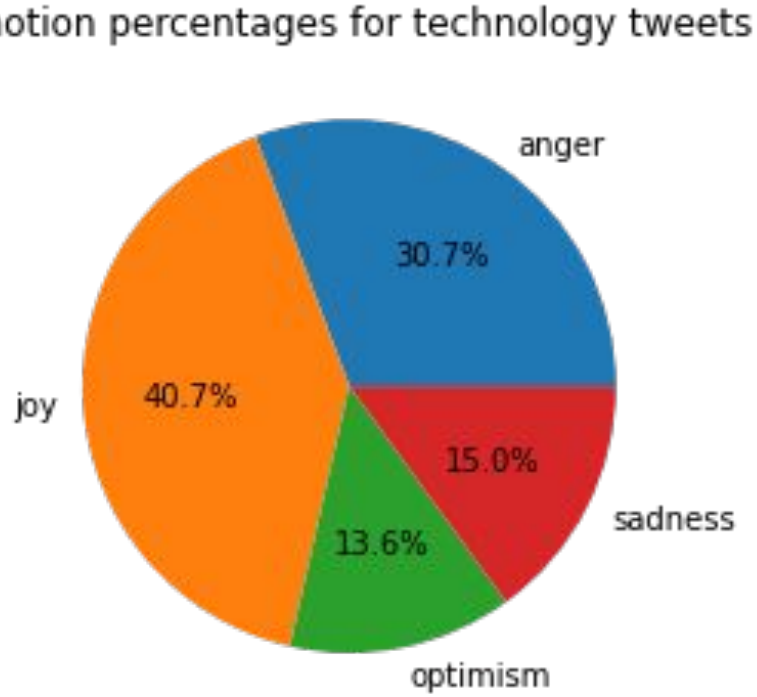
- Automatic sentiment and emotion analysis
- Huggingface Transformers fine tuned for those tasks (TweetEval)
- We select the most probable tag
- Pie charts for tag counts
- Wordclouds for word frequencies
- Comparison with manual annotation
- Comparison with adversarial tweets

Pie charts

Sentiments

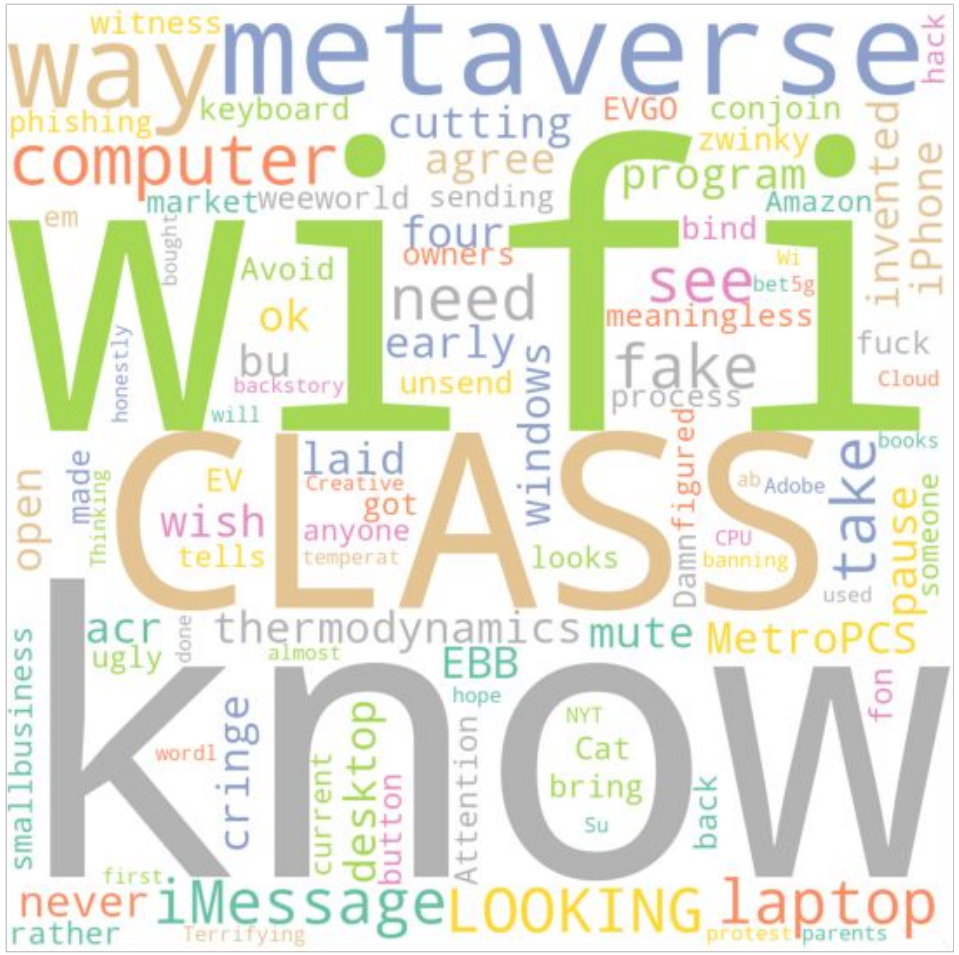


Emotions



Sentiment wordclouds (all vs neutral)

Sentiment wordclouds (negative vs positive)



Emotion wordclouds (anger vs joy)

Emotion wordclouds (sadness vs optimism)

Remaining work

- Annotate ITA tweets with new guidelines
- Recalculate ITA measures
- Annotate 40 tweets individually
- Create adversarial tweets
- Compare manual and automatic annotation
- Compare adversarial results
- Pie charts and wordclouds for manual annotation

Conclusions

- Annotation is more difficult than we thought
- Sentiment annotation easier than emotion annotation
- Defining good guidelines is important
- Creating good adversarial examples automatically is difficult
- Visualization is very helpful to identify patterns