Not Just Depressed:
Bipolar Disorder Prediction on Reddit

Ivan Sekulić, Matej Gjurković, Jan Šnajder

Text Analysis and Knowledge Engineering Lab
University of Zagreb

TakeLab

WASSA @ EMNLP, Brussels
31 October, 2018

Motivation

How does everyday language reflect basic social and personality processes?

Motivation

How does everyday language reflect our mental health?

Motivation

Can text and emotion analysis help with bipolar disorder detection?

Overview

Mental health & language

The Mental Health issue

- Up to 27% of Europe population suffer or have suffered from some kind of mental disorder (WHO, 2017; Wykes et al., 2015)
- 35–50% go undiagnosed

Language analysis for **deeper understanding** and **early detection** of illnesses:

- Depression (Rude et al., 2004; De Choudhury et al., 2013), anxiety (Shen and Rudzicz, 2017), PTSD (Coppersmith et al., 2015), schizophrenia (Benton et al., 2017), bipolar disorder (Kramer et al., 2004)

Bipolar Disorder

- Uncontrolled changes in mood and energy levels
- Manic and depressive episodes
- Recurrent phases that differ in intensity and duration
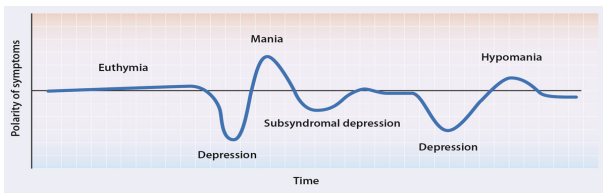
Why is research on bipolar disorder important?
- Affects more than 60M people worldwide
- High suicide rate (more than 6%)

Reddit: a gold mine for user-generated text

Reddit

- One of the largest social media sites in the world
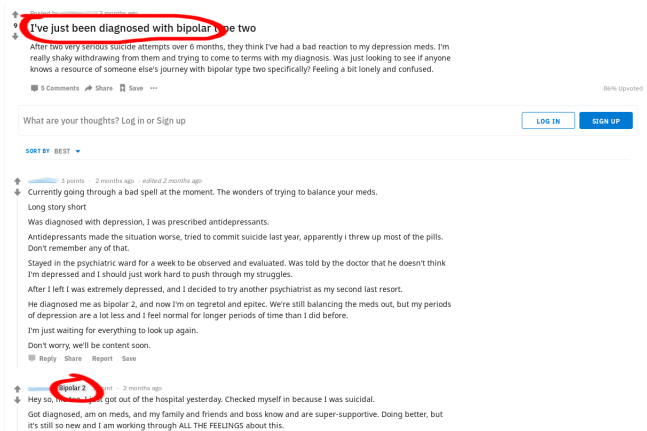- User anonymity
- Wide range of topics

Self-reported mental disorders

- Bipolar-related subreddits: *bipolar, bipolar2, BipolarReddit, BipolarSOs, bipolarart*
- Search for *"I am diagnosed with bipolar"* in user's comments
- Inspect users' *flairs* (following Gjurković and Šnajder, 2018): *"Bipolar"*, *"BP1"*, etc.

Self-reported Bipolar Disorder

Self-reported Bipolar Disorder

- Compiled a **control group** to represent general Reddit population

Dataset pruning

Cleaning the dataset:

- Removed all comments on bipolar-related subreddits, as well as on the general *mentalhealth* subreddit
- Removed all comments that had *"bipolar"* or *"BP"* in them
- Keep only users with > 1000 words left resulting in $\sim$**3.5k bipolar users**

Dataset pruning

Cleaning the dataset:

- Removed all comments on bipolar-related subreddits, as well as on the general *mentalhealth* subreddit
- Removed all comments that had *"bipolar"* or *"BP"* in them
- Keep only users with > 1000 words left resulting in $\sim$**3.5k bipolar users**

Parallel work by Cohan et al. (2018): gathered 36k users labeled with 9 mental disorders

Topical Categories

- Additional topic **bias reduction**
- Manually grouped top ~400 subreddits in the categories below:

Category	# bipolar	# control
Animals	397	898
AskReddit	1797	2767
Gaming	489	1501
Jobs and finance	293	586
Movies/music/books	502	1606
Politics	332	2445
Religion	264	700
Sex and relationships	948	1000
Sports	156	785
All	3488	3931

Feature Extraction

(1) Psycholinguistic features:

- **LIWC** (Pennebaker et al., 2015) – classify the words into dictionary-defined categories (syntactic, topical, psychological features)
- **Empath** (Fast et al., 2016) – based on neural embeddings (200 manually curated categories)

(2) Lexical features – tf-idf weighted bag-of-words

(3) Reddit user features – post/comment ratio, *#gilded* posts, *ups–downs*, time intervals between comments

Results

- Binary classification task
- Models: SVM, Logistic Regression, Random Forest
- 10x5 nested cross-validation

	Acc	F1
MCC	0.529	–
Random	0.546	0.453
SVM	0.865	0.853
LogReg	0.866	0.849
RF	**0.869**	**0.863**

Per-category results

* two-sided t-test, $p < 0.001$

	MCC	Our models
Animals	0.693	0.807*
AskReddit	0.606	0.856*
Gaming	0.754	0.777*
Jobs and finance	0.665	0.752*
Movies/music/books	0.761	0.817*
Politics	0.880	0.882*
Religion	0.724	0.784*
Sex and relationships	0.513	0.801*
Sports	0.832	0.837

Feature Analysis

Feature	Bipolar μ	Control μ
pronoun	**16.87**	14.86
ppron	**10.69**	8.66
i	**5.84**	3.38
article	5.88	**6.55**
health	**0.96**	0.50
feel	**0.69**	0.48
power	2.11	**2.58**
bio	**2.65**	1.90
Authentic	**52.65**	32.92
Clout	48.51	**58.03**

In line with previous research on depression: Personal pronoun *I*
(Stirman and Pennebaker, 2001; Rude et al., 2004)

Per-category analysis

Feature	Authentic	i	health	feel	power	bio	article
Animals	+	+	+	+	+	+	+
AskReddit	+	+	+	+	+	+	+
Gaming	+	+					+
Jobs	+	+	+	+	+	+	+
Movies	+	+	+	+	+	+	+
Politics	+	+	+	+		+	+
Religion	+	+	+	+	+		+
Sex	+	+	+	+	+	+	+
Sports	+	+		+			

+ signals significant difference in feature value between bipolar and control group

Emotion-expressive words analysis

LIWC emotion-related categories

Feature	Bipolar μ	Control μ
affect	**6.415**	6.074
posemo	**3.899**	3.442
negemo	2.432	**2.569**
anxiety	**0.367**	0.266
anger	0.818	**1.128**
sad	**0.455**	0.363

User-level emotion variance

- Split comments of 100 users (of both groups) with more than 100K words into **monthly chunks**
- Calculate LIWC features for each month and compute their **standard deviations**

User-level emotion variance

- Split comments of 100 users (of both groups) with more than 100K words into **monthly chunks**
- Calculate LIWC features for each month and compute their **standard deviations**

	Bipolar	Control	p-value
posemo	0.00272*	0.00166	0.00272
negemo	0.00583*	0.00379	0.00583
anxiety	0.00765*	0.00627	0.00765
anger	0.01745	0.01422	0.01745
sadness	0.00695*	0.00572	0.00695

* $p < 0.01$; averages of SDs in the use of emotion-expressive words

Wrapping up

Conclusion

- preliminary study of bipolar disorder
- benchmark classification results (86% accuracy)
- psycholinguistic and emotion-related feature analysis

Future work

- Detect manic and depressive episodes – compare to clinical depression
- Deep learning for prediction

Not Just Depressed: Bipolar Disorder Prediction on Reddit

Ivan Sekulić, Matej Gjurković, Jan Šnajder

ivan.sekulic@h-its.org
takelab@fer.hr

Thank you!

LIWC vs Empath vs tf-idf

	LIWC	Empath	Tf-idf	All
SVM	0.837	0.782	**0.865**	0.838
LogReg	0.841	0.819	**0.866**	0.862
RF	0.829	0.825	0.869	**0.869**

Topical categories 1

Animals: cats, Dogtraining , dogs, Aquariums, AnimalsBeingJerks;

AskReddit: AskReddit, IAmA, todayilearned, CasualConversation, Showerthoughts;

Gaming: gaming, DotA2, PS4, pcmasterrace, Games;

Jobs and finance: wallstreetbets, Economics, personalfinance, Frugal, TalesFromRetail;

Mental disorders: depression, Drugs, ADHD, mentalhealth, SuicideWatch;

Movies/music/books: movies, books, Music, harrypotter, BigBrother;

Politics: politics, Libertarian, PoliticalDiscussion, SandersForPresident, Conservative;

Topical categories 2

Religion: Christianity, atheism, islam, exmuslim,
DebateReligion;

Sex and relationships: relationships, AskWomen,
TwoXChromosomes, actuallesbians, childfree;

Sports: nfl, nba, clevelandcavs, MMA, olympics.

Insignificant features

Colon	Colons
Period	Periods
achieve	win, success, better
WPS	Words/sentence
quant	Quantifiers
see	view, saw, seen
differ	Differentiation
ipron	Impersonal pronouns
motion	arrive, car, go
hear	listen, hearing
compare	greater, best, after
netspeak	btw, lol, thx