Deep learning for mental health disorders

Ana-Sabina Uban

(+Paolo Rosso, Berta Chulvi, Petr Lorenc)

BioMedical NLP

2021/2022

Mental health

Depression

Depression (major depressive disorder) is a common and serious medical illness that negatively affects how you feel, the way you think and how you act.

Depression causes feelings of sadness and/or a loss of interest in activities you once enjoyed. It can lead to a variety of emotional and physical problems and can decrease your ability to function at work and at home.

Source: American Psychiatric Association website

Psycho-linguistic categories (LIWC)

II. PSYCHOLOGICAL PROCESSES

Social Processes	talk, us, amigo	Cognitive Processes	cause, know, ought
Friends	pal, buddy, coworker	Insight	think, know, consider
Family	mom, brother, cousin	Causation	because, effect, hence
Humans	boy, woman, group	Discrepancy	should, would, could
Affective Processes	happy, ugly, bitter	Tentative	maybe, perhaps, conjetura
Positive Emotions	happy, pretty, good	Certainty	always, never
Negative Emotions	hate, worthless, enemy	Inhibition	block, constrain
Anxiety	nervous, afraid, tense	Inclusive	with, and, include
Anger	hate, kill, pissed	Exclusive	but, except, without
Sadness	grief, cry, sad		

Cognitive styles
The hopelessness theory of depression
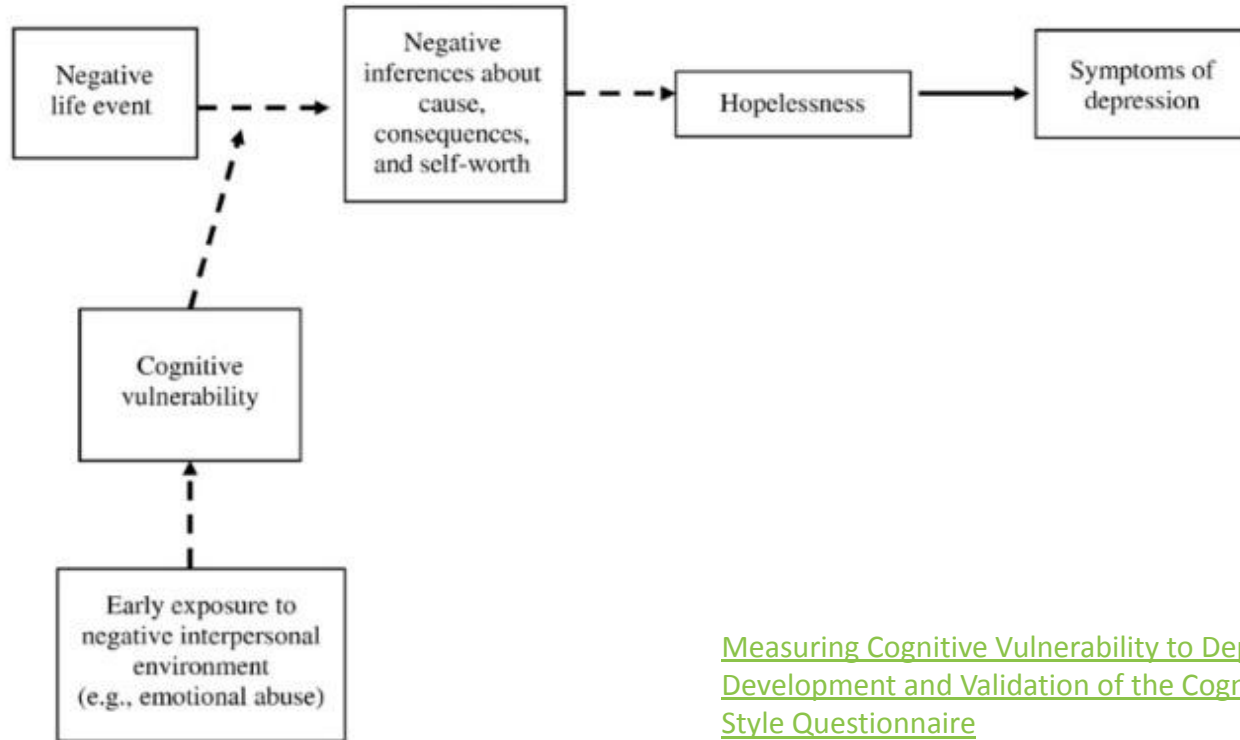
Measuring Cognitive Vulnerability to Depression: Development and Validation of the Cognitive Style Questionnaire

Emotions over time

Not only the static expression of certain emotions or discussion of topics is relevant, but their **evolution over time**

Track evolution of emotion expression over time

Track evolution of usage of different psycho-linguistic categories over time (LIWC)

Analyze their correlations

=> Understand how emotions relate to different psycho-linguistic categories (e.g. causation, society, self etc) for users suffering from a mental disorders

Emotions over time

Method:

Measure emotion usage in texts posted per day - separately for positive vs negative and users + average across users

Measure psycho-linguistic categories usage per day for each user, ...

Rolling average of 100 days

Pearson correlations between obtained time series for every (emotion, psycho-linguistic category) pair

Compare correlations between positive users and negative users

Select pairs with significantly different correlations between the two groups (z-test)

Emotions over time
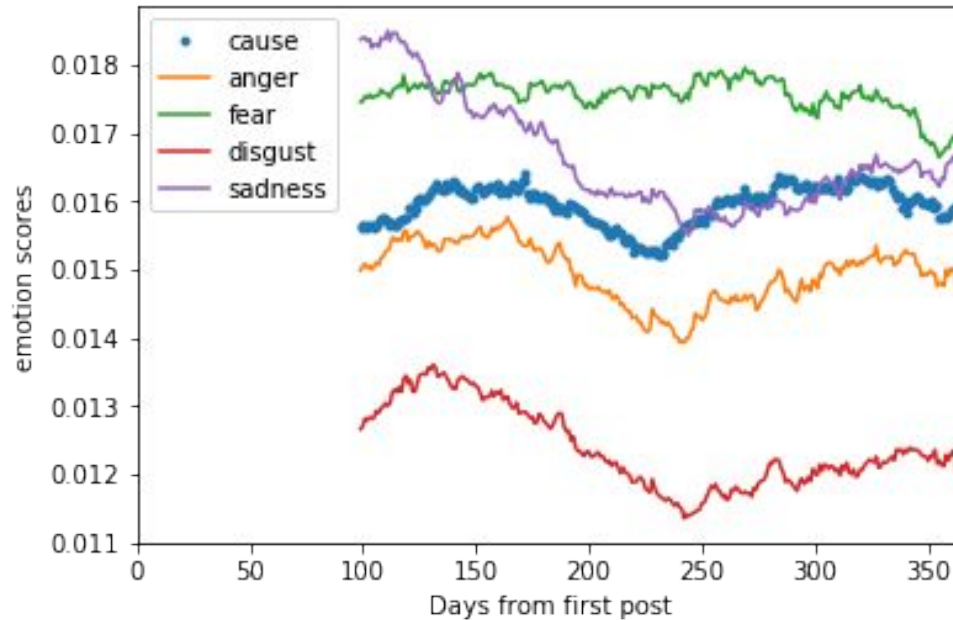
Depressed users expressing causation & negative emotions over time

Emotions over time: findings
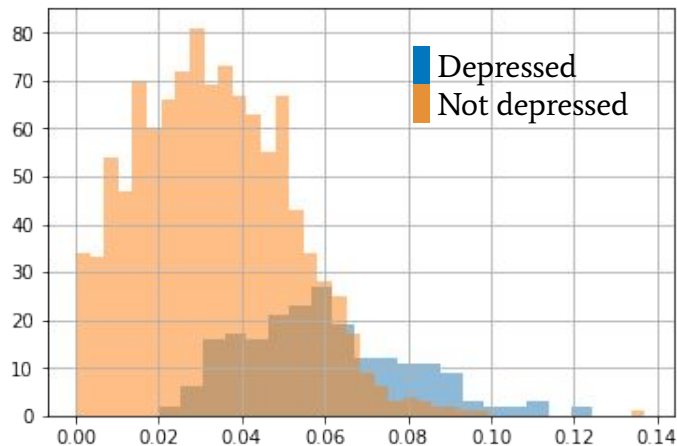
Causation and emotions

pos(P)/neg(N)	Anger		Disgust		Fear		Sadness		Trust		Anticip.		Joy	
	P	N	P	N	P	N	P	N	P	N	P	N	P	N
Depression	0.39	-0.36	0.33	-0.21	0.46	-0.43	-	-	-0.06	-0.31	-	-	-0.23	-0.03
Anorexia	0.48	0.07	0.40	0.02	0.39	0.04	0.45	-0.38	0.41	0.16	-0.13	-0.23	-0.08	-0.55
Self-harm	0.25	0.12	-	-	0.15	0.27	-0.15	0.03	-	-	0.23	-0.17	0.26	-0.19

Table 7: Correlation between "causation" and emotions in the three mental disorders for positive users (diagnosed with a mental disorders) and negative ones (healthy). Only correlations which are significantly different between the positive and negative classes are shown.
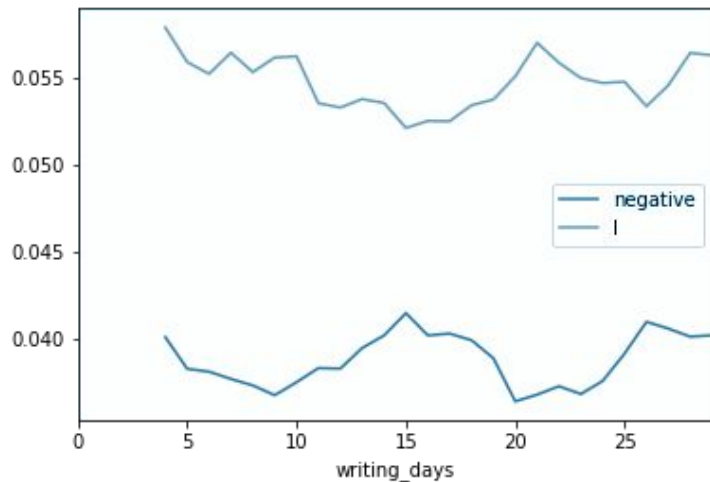
Feature analysis (over time) - "I"
Depression

The use of "I" in depressed vs non-depressed users

Use of "I" vs negative emotion in depressed users

Emotions over time: findings

The self and emotions

pos(P)/neg(N)	Anger		Disgust		Fear		Sadness		Trust		Anticip.		Joy	
	P	N	P	N	P	N	P	N	P	N	P	N	P	N
Depression	-	-	-	-	-0.11	-0.29	0.25	-0.04	0.15	-0.15	0.58	-0.62	0.50	-0.06
Anorexia	0.12	-0.16	0.08	-0.22	0.30	-0.16	0.24	0.06	0.27	-0.25	0.53	0.24	0.72	0.55
Self-harm	0.42	0.01	0.34	0.13	0.21	-0.28	0.34	-0.06	-	-	-0.16	0.31	-0.05	0.40

Table 8: Correlation between the use of "I" and emotions in the three mental disorder for positive users (diagnosed with a mental disorder) and negative ones (healthy). Only correlations which are significantly different between the positive and negative classes are shown.

Other tasks

- Detecting the **severity** of depression / suicide risk level
- Detecting specific symptoms (lack of sleep, loss of appetite, lack of energy…)
- Detecting **causes** of depression - helps with prevention, and with targeted management
- Detecting depression from video therapy sessions (based on video/audio signals)
- Analyze different disorders jointly (co-morbidities); transfer learning
- **Profiling** users suffering from a disorder: age, behavioral patterns, social media activity patterns (nocturnal, seasonal)
- Conversational data: therapy sessions, therapist chatbot (https://woebothealth.com/)
- **Multimodal** depression detection
- Social media: depression and **aggression**

In practice: eRisk 2021

Best results in overall level of depression prediction (some metrics) at Task 3:
http://ceur-ws.org/Vol-2936/paper-75.pdf

Transfer learning

Clinical evidence of comorbidity within mental disorders. ([Exploring Comorbidity Within Mental Disorders Among a Danish National Population](#))

- ❖ Improve performance on tasks with less data (depression → other disorders)
- ❖ Understand connection/compatibility between disorders and expression media (genre/platform)

Cross-task - transfer knowledge between labels for different disorders

Cross-genre - transfer knowledge between different data platforms (reddit/Twitter)
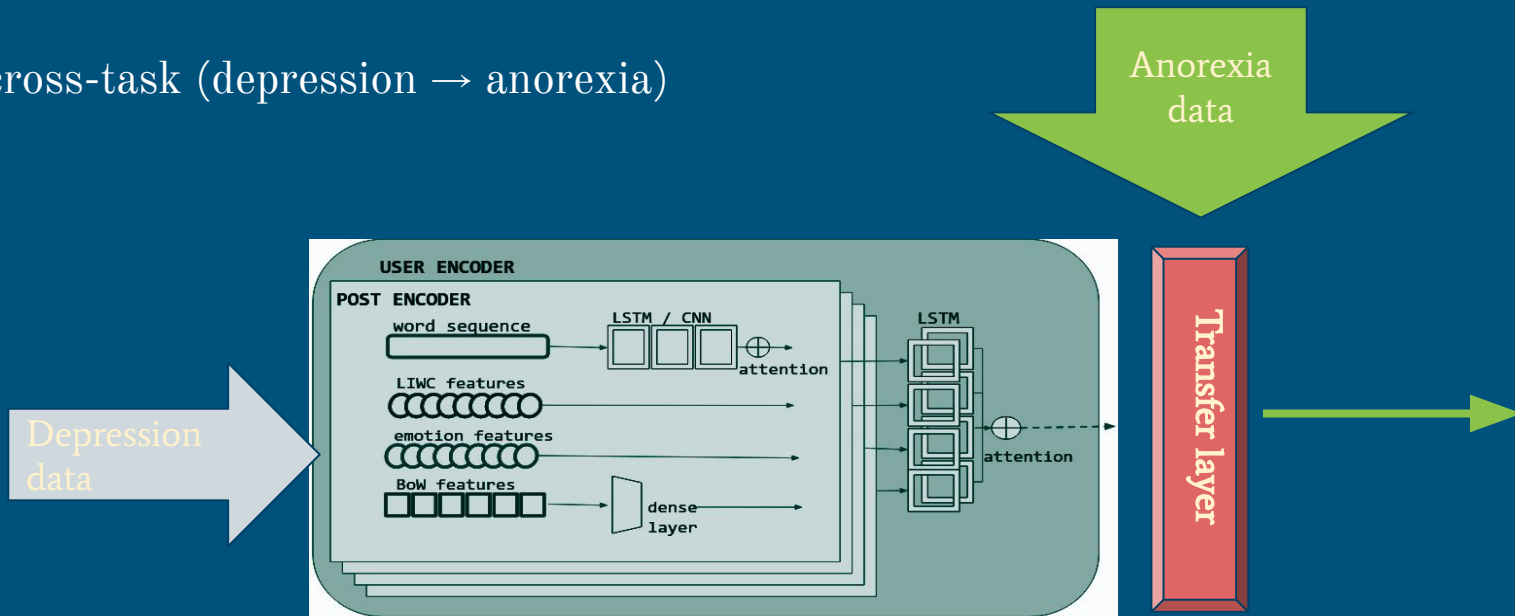
Transfer learning

Strategy 0. No pre-training
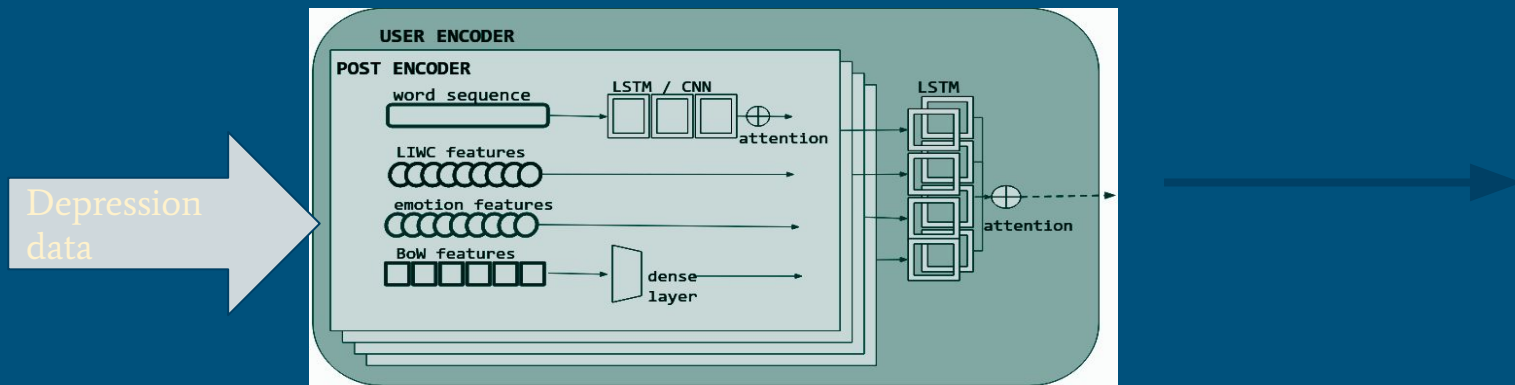
Transfer learning

Strategy 1. Transfer layer
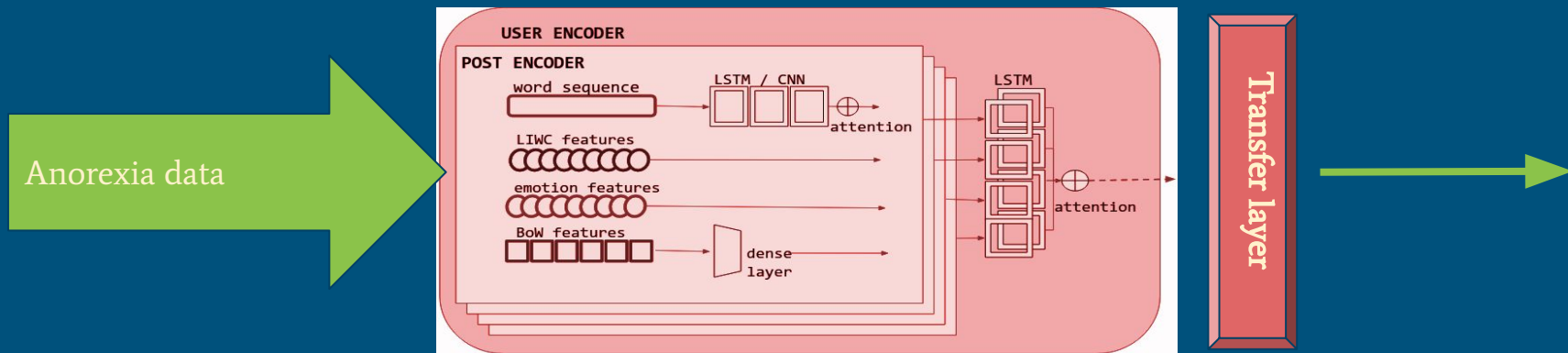
Example: cross-task (depression → anorexia)

Transfer learning

Strategy 2. Fine-tuning

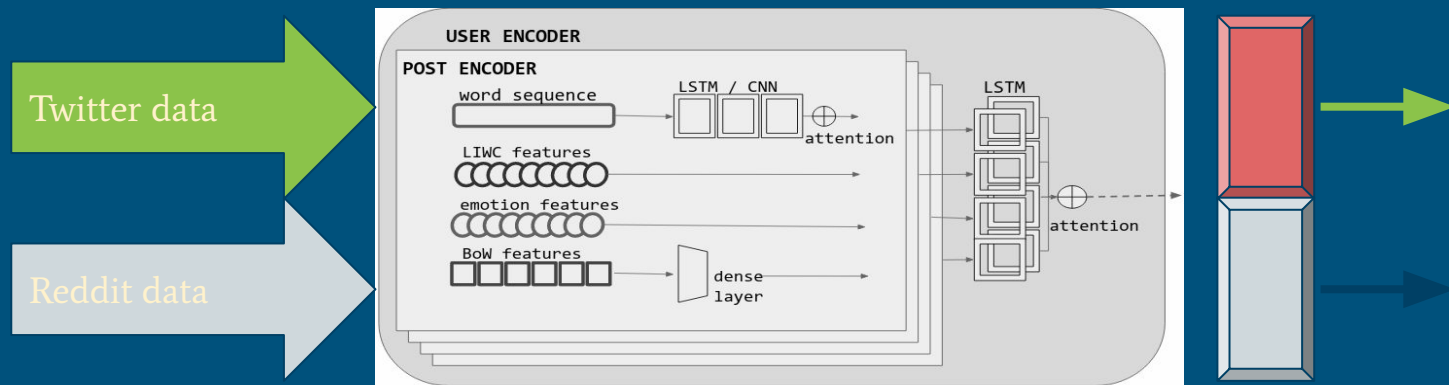Example: cross-task (depression → anorexia)

Transfer learning

Strategy 3. Multi-task learning

Example: cross-genre (reddit / Twitter)

Transfer learning experiments: Results

Source	CROSS-TASK						CROSS-GENRE			
	eRisk depression				CLPsych depression		eRisk depression			
Target	eRisk Anorexia		eRisk Self-harm		CLPsych PTSD		(Shen et al.) depression		CLPsych depression	
	F1	AUC	F1	AUC	F1	AUC	F1	AUC	F1	AUC
Strategy 0	.17	.62	.13	.69	.31	.60	.69	.59	.38	.57
Strategy 1	.64	.90	.54	.87	.43	.73	.65	.74	.61	.72
Strategy 2	.63	.93	.67	.87	.58	.78	.86	.94	.60	.74
Baseline BiLSTM	.62	.93	.62	.84	.55	.78	.75	.83	.56	.72

| Source | All depression | | | | | |
| Target | eRisk | | (Shen et al.) | | CLPsych | |
	F1	AUC	F1	AUC	F1	AUC
Strategy 3	.39	.81	.74	.83	.56	.82
Single-task	.40	.83	.75	.83	.56	.72

Depression and offensive language

Depression and aggression linked in psychology studies (especially self-aggression:)

Frontiers | Aggressive and Disruptive Behavior Among Psychiatric Patients With Major Depressive Disorder, Schizophrenia, or Alcohol Dependency and the Effect of Depression and Self-Esteem on Aggression | Psychiatry

Computational linguistics:

An Exploratory Analysis of the Relation between Offensive Language and Mental Health - ACL Anthology
(Ana-Maria Bucur, Marcos Zampieri and Liviu P. Dinu)

Depression and offensive language

RQ1: Are posts from individuals suffering from depression more likely to contain offensive language in existing datasets?

RQ2: Are there differences in the nature of offensive language used by individuals with depression compared to control groups?

Depression and offensive language

- Fine-tuned a BERT model on the OLID dataset (offensive language + subtypes according to target)
- Applied it for prediction on corpora of depressed users on social media
- Also selecting posts showing signs of depression based on emotions and polarity
- There are more posts with signs of depression labeled as offensive, the majority of them are untargeted (containing swears, profanity)
- Depressed individuals tend to use more self-deprecating content and less deprecation of others

Depression and offensive language

	Self-reported		Signs of depression	
Dataset	**Depression**	**Control**	**Showing**	**Not showing**
eRisk 2018	8.24%	5.91%	18.50%	7.40%
RSDD	11.31%	8.91%	24.33%	10.10%

Table 1: Percentage of posts labeled as offensive from total posts of self-reported individuals and of individuals showing/not-showing signs of depression measured with the H_s heuristic.
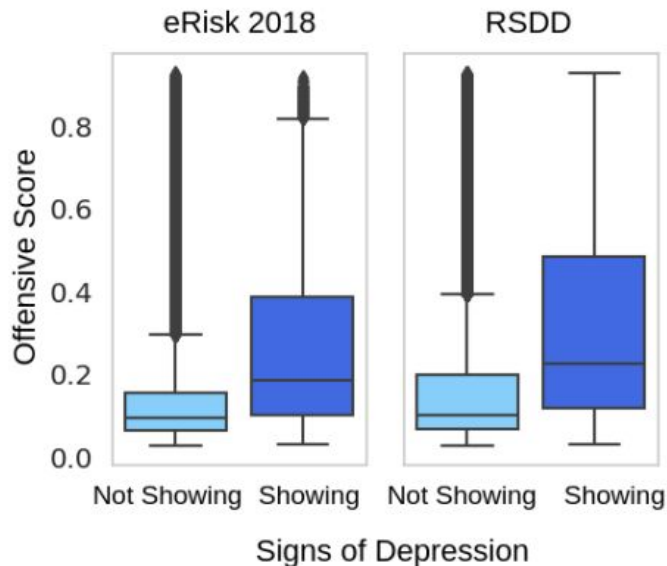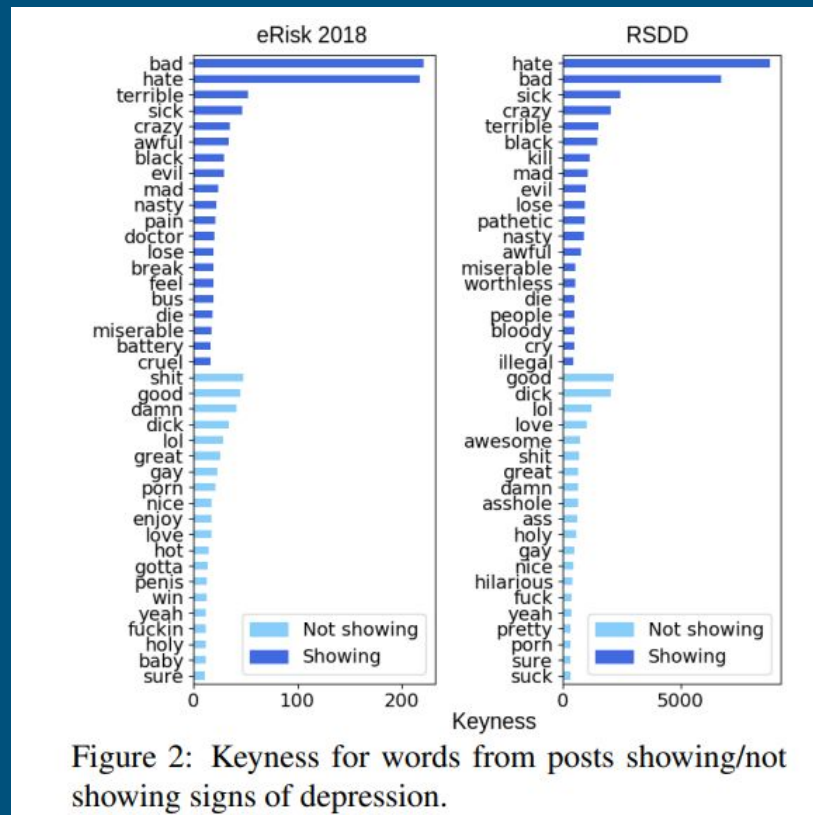
Depression and offensive language

Figure 1: Distribution of the offensive language score for posts written by users with self-reported depression diagnosis and showing or not showing signs of depression measured with the H_s heuristic.

Depression and offensive language

Figure 2: Keyness for words from posts showing/not showing signs of depression.

Mental disorders in conversational domain

Data:
> Possible source: Therapy sessions transcripts (…)

Applications:
> Chatbot "therapist"

> Notable example: https://woebothealth.com/

Mental disorders in conversational domain

Woebot: https://woebothealth.com/

Clinically tested therapeutic approaches

Cognitive Behavioral Therapy (CBT), Interpersonal Psychotherapy (IPT), and Dialectical Behavioral Therapy (DBT) provide the foundation for Woebot's therapeutic support.

AI-powered delivery

Through the use of AI and NLP, Woebot forms trusted bonds and delivers clinically validated techniques in an approachable, conversational manner.

Mental disorders in conversational domain

Woebot: https://woebothealth.com/

Efficiency proved in studies:

Evidence of Human-level Bond Established with a Digital Conversational Agent: An Observational Study

Acceptability of Postnatal Mood Management Through a Smartphone-based Automated Conversational Agent

Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial

Mental disorders in conversational domain

Woebot: https://woebothealth.com/

Efficiency proved in studies:

Satisfaction with the chatbot was assessed by the Client Satisfaction Questionnaire (CSQ-8)

Acceptability of the chatbot as a mood management tool was assessed with the Working Alliance Inventory - Short Revised (WAI-SR) survey, a measure of therapeutic alliance that correlates with favorable psychotherapy outcomes.

Datasets on mental health in conversational domain + availability

- Not readily available
  - [Crisis Counseling](#) [webpage](#)
    - Real SMS Conversations
    - 5 millions conversation, several labels
    - Need to do a long process of starting cooperation with Crisis counseling (months)
  - [Motivational Interviewing Dataset](#)
    - Transcripts
    - 22,719 counselor utterances
    - Get after request? Various labels/disorders
  - [Annotating Reflections for Health Behavior Change Therapy](#)
    - Transcripts
    - 324 sessions

Datasets on mental health in conversational domain + availability

- Publicly available...
  - [Counsel Chat](#)
    - Single talk turn
    - <1k utterances
  - **Counseling and Psychotherapy Transcripts Series ([CPTS](#))**
    - Transcripts
    - 4,000 transcripts of real therapy sessions (various labels/disorders)
    - (requires account)
  - **Distress Analysis Interview Corpus ([MIT paper](#)) (DAIC-WOZ)**
    - Transcripts
    - 142 interactions, 3K words

Mental disorders in conversational domain

Logistic regression model on conversational data - depressed vs non-depressed feature analysis:

environment (-7.5), open-minded (-6.3), or accomplish (-4.7) - associated to non-depressed patients.

insignificant (5.36), television (5.66), or pollution (+6.1) - associated to depressed patients

Mental disorders in conversational domain
Transfer learning from non-conversational domain (social media data)

Table 3. Results - biLSTM stands for bidirectional LSTM, use5 - Universal Sentence Encoder based on Transformers, "f" stands for additional features.

	DAIC-WOZ	eRisk/GPC without fine-tuning	eRisk/GPC with fine-tuning
biLSTM + use5	**0.660**	**0.651** / 0.440	**0.803** / 0.690
biLSTM + use5 + f	0.565	0.541 / 0.410	0.597 / 0.511

Mental disorders in conversational domain
Transfer learning from non-conversational domain (social media data)

Is the size of the source domain dataset more critical than domain relatedness?

Using GPC as the source domain underperforms the setting in which eRisk data is used as the source domain, even though GPC is a more similar type of data to our target dataset =>

The size of the source domain dataset is more important than domain closeness.

Multimodality and mental disorder detection

Multimodal approach = approach involving multiple modes of communication

Multimodal data:

Text +

Images
Video
Speech

Behavioral data

Sensor data

Physiologic testing

Social media metadata

Medical information

etc

Features types for mental disorder detection

- **Text** posted by user on social media
- Audio/video data from interviews (therapy sessions) <- AVEC Workshop
- Sensor data from smartphones, eHealth: heart rate, eye sensor, accelerometer, calls, location, dermal activity, microphone, video cameras, gps, …
- Behavioral data: social media metadata (login times, interactions, followers, …)
- Demographic data (gender, age, …)
- Medical statistics (psychometric and physiologic testing, indices of awareness, consciousness, insight and anosognosia, unawareness, and awareness of illness)
- Images posted by user on social media / profile pictures
- Disorders approached:
  - Images features: suicide, depression, anxiety
  - General: depression, types of PTSD (behavioral), bipolar, stress, others (audio/video)

Multimodal mental health: images

- Types of image features:
  - their five-color combinations, brightness, saturation, cool color ratio, and clear color ratio
  - Image embeddings, features extracted directly from images
  - Image captions/tags
- Platforms: Twitter, Instagram
- Types of images: avatar images/profile (Twitter), shared images on Instagram (selfies?), shared images on Twitter (generic)

Research questions

➔ Do users suffering from depression/a mental disorder post a different kind of images (various perspectives)?
➔ Is the content in the images (objects depicted) different than in healthy users?
➔ Can the combination of text and image features help with detection of mental disorders?

Colors and mental health: insights

The strong associations between color sensitivity and mood has been highlighted by several studies [...]. In an earlier research, a strong correlation between specific color selection such as yellow and depressive behavior has been reported by [...].

Reece and Danforth [...] analyzed uploaded images to Instagram and found that photos posted by depressed users were more likely to be bluer, grayer and darker.

In studies associating mood, color, and mental health, healthy individuals identified darker, grayer colors with negative mood, and generally preferred brighter, more vivid colors [...]. By contrast, depressed individuals were found to prefer darker, grayer colors

Multimodal studies on mental health with social media images

Depression:
- Depression Detection via Harvesting Social Media: A Multimodal Dictionary Learning Solution [1]
- What Twitter Profile and Posted Images Reveal about Depression and Anxiety [2]
- Multimodal mental health analysis in social media [6]
- Depression-related Imagery on Instagram [8]

Suicide:
- Detection of Suicidal Ideation on Social Media: Multimodal, Relational, and Behavioral Analysis [3]

Image-only/depression:
- Instagram photos reveal predictive markers of depression [4]

Emotion:
- Multimodal Classification for Analysing Social Media [5]

Multimodal studies on mental health with social media images (References)

[1] Shen, Guangyao, Jia Jia, Liqiang Nie, Fuli Feng, Cunjun Zhang, Tianrui Hu, Tat-Seng Chua, and Wenwu Zhu. "Depression Detection via Harvesting Social Media: A Multimodal Dictionary Learning Solution." In IJCAI, pp. 3838-3844. 2017.

[2] Guntuku, Sharath Chandra, Daniel Preotiuc-Pietro, Johannes C. Eichstaedt, and Lyle H. Ungar. "What twitter profile and posted images reveal about depression and anxiety." In Proceedings of the international AAAI conference on web and social media, vol. 13, pp. 236-246. 2019.

[3] Ramírez-Cifuentes, Diana, Ana Freire, Ricardo Baeza-Yates, Joaquim Puntí, Pilar Medina-Bravo, Diego Alejandro Velazquez, Josep Maria Gonfaus, and Jordi Gonzàlez. "Detection of suicidal ideation on social media: multimodal, relational, and behavioral analysis." Journal of medical internet research 22, no. 7 (2020): e17758.

[4] Reece, Andrew G., and Christopher M. Danforth. "Instagram photos reveal predictive markers of depression." EPJ Data Science 6 (2017): 1-12.

[5] Duong, Chi Thang, Remi Lebret, and Karl Aberer. "Multimodal classification for analysing social media." arXiv preprint arXiv:1708.02099 (2017).

Multimodal studies on mental health with social media images (References)

[6] Yazdavar, Amir Hossein, Mohammad Saeid Mahdavinejad, Goonmeet Bajaj, William Romine, Amit Sheth, Amir Hassan Monadjemi, Krishnaprasad Thirunarayan et al. "Multimodal mental health analysis in social media." Plos one 15, no. 4 (2020): e0226248.

[7] Garcia-Ceja, Enrique, Michael Riegler, Tine Nordgreen, Petter Jakobsen, Ketil J. Oedegaard, and Jim Tørresen. "Mental health monitoring with multimodal sensing and machine learning: A survey." Pervasive and Mobile Computing 51 (2018): 1-26.

[8] Nazanin Andalibi, Pinar Ozturk, and Andrea Forte. 2015. Depression-related Imagery on Instagram. In Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing (CSCW'15 Companion). Association for Computing Machinery, New York, NY, USA, 231–234. DOI:https://doi.org/10.1145/2685553.2699014

Depression Detection via Harvesting Social Media: A Multimodal Dictionary Learning Solution [1]

- Disorder: Depression
- Platform: Twitter
- Features - 6 modalities: social network features, user profile features, visual features, emotional features, topic-level features, and domain-specific features.
- Image types: user avatars
- Image features: their five-color combinations, brightness, saturation, cool color ratio, and clear color ratio
- Model: multimodal depressive dictionary learning model (MDL) to learn the sparse user representations + relations between the modalities

What Twitter Profile and Posted Images Reveal about Depression and Anxiety [2]

- Platform: Twitter, + Facebook (text only)
- Disorders: depression, anxiety
- Labels: survey-based ground-truth labels (regression)
- Features: text, images (avatars of users + posted images), demographic information
- Image features: Color, Facial (emotion, posture, …), Aesthetics (symmetry, lightning, harmony,…), Content, VGG-Net (bag-of-tags + dimensionality reduction; image vectors)
- Results indicate that multi-task learning gives significant improvements in performance for modeling mental health conditions jointly with demographics (here age and gender), factors which clinicians usually consider while diagnosing patients. Further, models trained on larger data sets using text-predicted outcomes show reliable performance when predicting more reliable survey based mental health outcomes.

What Twitter Profile and Posted Images Reveal about Depression and Anxiety [2]

- *For depression, we find that profile pictures suppress positive emotions rather than display more negative emotions, likely because of social media self-presentation biases. They also tend to show the single face of the user (rather than show her in groups of friends), marking increased focus on the self, emblematic for depression.*
- *Posted images are dominated by grayscale and low aesthetic cohesion across a variety of image features.*
- *Anxious users additionally seem to post marginally more content related to family and work*
- *Users high in depression post images which emphasize foreground objects and are low in brightness, (...) images containing text, animals, while users low in depression and anxiety post images of sports, nature, every day things*

Mental health / personality multimodal studies

Study	# Users	Traits	Image Type	Image Features
Our Work	**887 + 4,132**	Depression & Anxiety	Twitter Posted & Profile Images	Color, Facial, Aesthetics, Content, VGG-Net
(Reece and Danforth 2017)	166	Depression	Instagram Photos	Colors
(Andalibi, Ozturk, and Forte 2015)	–	Depression	500 '#depression' Instagram Photos	Manual Annotation
(Ferwerda and Tkalcic 2018)	193	Personality (continuous)	Instagram Photos	Content
(Nie et al. 2018)	2238	Perceived Personality (continuous)	Web Portrait Images	Facial, Social information
(Samani et al. 2018)	300	Personality (continuous)	Twitter and Flickr Posts, Likes, & Profiles	Colors, Content, VGG-Net
(Farnadi et al. 2018)	5670	Personality (binary)	Facebook Profile Images	Facial, Text, Likes
(Guntuku et al. 2017a)	4132 + 161	Personality (continuous)	Posted, liked images and text on Twitter	Color, Bag of Imagga tags, VGG-Net
(Segalin et al. 2017)	11,736	Personality (continuous & binary)	Facebook Profile Images	Aesthetics, BOVW, VGG-Net, IATO
(Liu et al. 2016)	66,502	Personality (continuous)	Twitter Profile Images	Color, Facial
(Ferwerda, Schedl, and Tkalčič 2016)	113	Personality (binary)	Instagram Photos	Colors, #Faces, Filters
(Skowron et al. 2016)	62	Personality (binary)	Instagram Photos	Colors
(Guntuku et al. 2016)	300	Personality (continuous)	Liked ('Fave') images on Flickr	Colors, semantic features, aesthetics
(Guntuku et al. 2015)	123	Personality (continuous)	Selfies on Weibo	Color, Aesthetics, BOVW, Emotions
(Al Moubayed et al. 2014)	829	Personality (binary)	Face Images	Eigenfaces
(Celli, Bruni, and Lepri 2014)	112	Personality (binary)	Facebook Profile Images	Bag-of-Visual-Words (BOVW)

Table 1: Summary of data and methods used in previous work analyzing images of individual users.

Detection of Suicidal Ideation on Social Media: Multimodal, Relational, and Behavioral Analysis [3]

- Disorder: Suicide, platform: Twitter
- Collected dataset: 252 users annotated by clinicians
- Features: images, behavioral, texts
- Image features: CNN (ResneXt) pre-trained on Instagram suicide-related images => user score
- *The combination of textual, visual, relational, and behavioral data outperforms the accuracy of using each modality separately. We defined text-based baseline models based on bag of words and word embeddings, which were outperformed by our models, obtaining an increase in accuracy of up to 8% when distinguishing users at risk from both types of control users.*

Instagram photos reveal predictive markers of depression [4]

- Disorder: depression
- Only images; various hypotheses tested & statistical analyses
- Statistical features were computationally extracted from 43,950 participant Instagram photos
- Image features: color analysis, metadata components, and algorithmic face detection.
- *Human ratings of photo attributes (happy, sad, etc.) were weaker predictors of depression*
- *Our results supported Hypothesis 1, that markers of depression are observable in Instagram user behavior, and Hypothesis 2, that these depressive signals are detectable in posts made even before the date of first diagnosis.*

Depression-related Imagery on Instagram [8]

- Disorder: depression
- Platform: Instagram
- Dataset: images posted by users, #depression
- Quantitative analysis (no classification): Examined the distribution of themes in 500 images, uncovering themes posted by depressed users + relationship between image and caption
- Image features: manually coded image descriptions
- *Images alone were used to communicate about most topics; however, seeking or providing help/support/engagement, and positive emotions were shared more often through textual captions.*
- *In posts that depicted individuals and were not celebrities or screen shots, we identified 72% females and 17% males and around 4% minors*
- *13% of captions were descriptive of the image, 35% provided contextual and different information, 32% provided additional but similar information, and 20% were unrelated.*
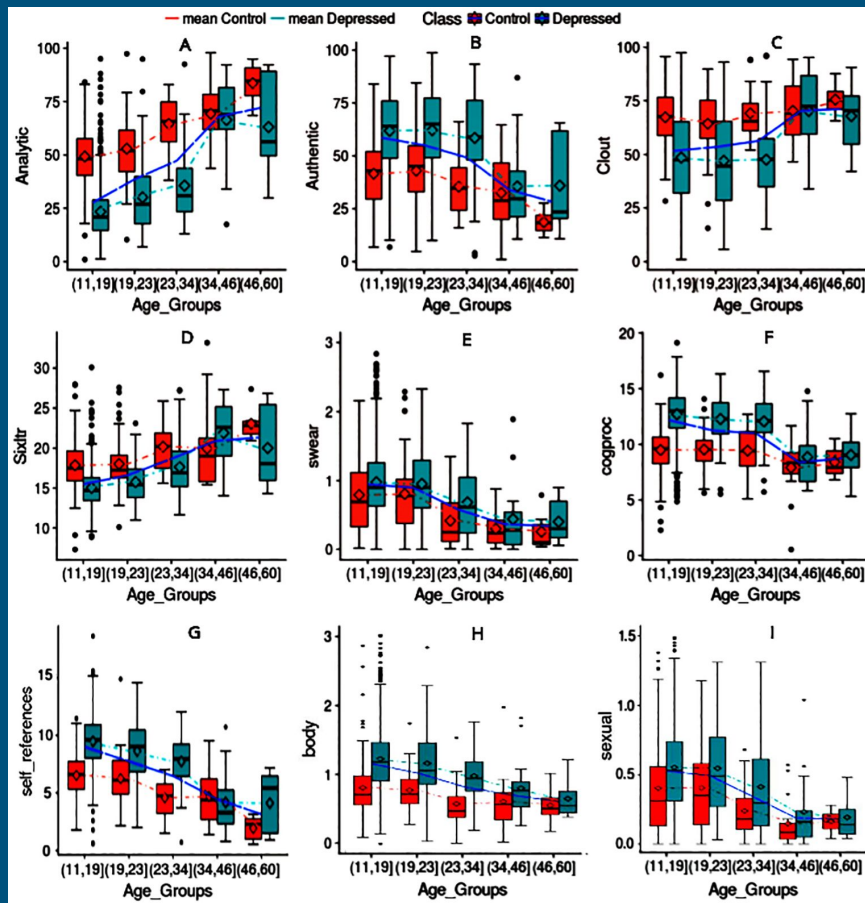
# Multimodal mental health analysis in social media [6]

- Disorder: depression
- Data: Twitter (collect the Twitter profiles of individuals with self-declared depressive symptoms + age, gender)
- Features: visual, textual, and user interaction data
- Analysis of the content of posted images in terms of colors, aesthetic, facial presentation, and their associations with depressive symptoms;
- Uncovering the underlying relationships between visual and contextual content
- Improve the state-of-the-art for depression detection on Twitter: average F1-Score increased by 5 percent
- Feature importance analysis
- Demographic analysis: predicting age and gender on basis of text + images
- Thinking styles features & analysis: authenticity, self-references, informal, certainty (LIWC-based)
- *(...) a strong association between female gender, and expression of depressive symptoms on Twitter.*

Multimodal mental health analysis in social media [6]

The Pearson correlation between the average emotions derived from facial expressions through the shared images and emotions from textual content for depressed-(a) and control users-(b).

Multimodal mental health analysis in social media [6]
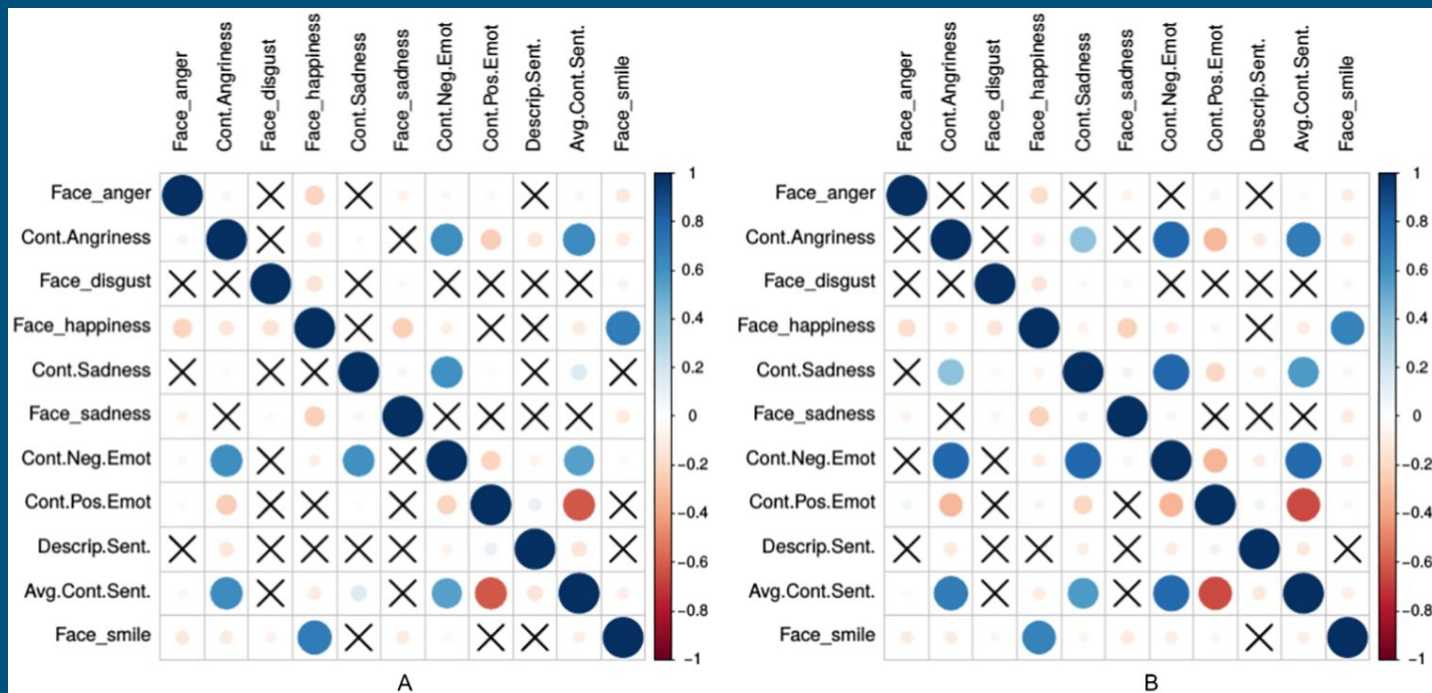
The Pearson correlation between the average emotions derived from facial expressions through the shared images and emotions from textual content for depressed-(a) and control users-(b).

Multimodal Classification for Analysing Social Media [5]

- Emotion classification task
- Datasets: reddit, flickr
- Approach: neural networks (CNN for images) with fusion for text + image classification, + emotions
- *For the reddit dataset, the difference is only 0.8% for common space fusion while it is 34% for joint fusion. The discrepancy for common space fusion is lower as it considers image and text vector of a post as equally important. As a result, using either image or text to classify, we get similar accuracy. On the other hand, joint fusion considers textual information as more important. This makes the classification using text significantly better than using only image*

multiRedditDep - a multimodal depression dataset

Collected images posted on Reddit + titles (text)
Positive users labeled based on self-stated diagnoses: "I was diagnosed with depression"...
Control users are random users who have posted at least once on the *r/depression* forum, but that have not mentioned any diagnosis of depression in the recorded history of their activity

	Images	Unique users
Depressed	48,121	1,419
Control	72,855	2,344

Classification experiments - multimodal

Models for classifying posts automatically based on text and visual features: feed-forward neural network

Text features (on image titles):

- ❖ word embeddings (word2vec)
- ❖ sentence embeddings (BERT)

Image features:

- ❖ image tags extracted with pretrained VGG model
- ❖ Text extracted from images which constitute screenshots
- ❖ + word2vec embeddings of tags

Classification experiments - multimodal

Future:

+ Visual features:
- ❖ Image embeddings
- ❖ Color information (color histogram, cool color ratio, brightness, saturation)
- ❖ Additional features from image (facial features, emotion/sentiment features,...)
- ❖ Fusion techniques

- ❖ Other disorders (anorexia!)

Classification experiments - multimodal

Features	Accuracy	AUC
Text only: word2vec embeddings	62.73	65.29
Images only: word2vec embeddings on image tags	61.88	61.63
Text and images: word2vec embeddings of text, image tags	**65.40**	**68.69**
Text only: BERT + word2vec embeddings	66.24	69.20
Text and images: BERT + word2vec embeddings	**66.39**	**69.37**

Multimodal dataset - image analysis

We build a taxonomy of objects depicted in the images in the dataset, based on clustering the image tags semi-automatically: word embeddings + manual curation => **6 categories** + sub-categories

	Depressed	Control
Animal	16.92	13.51
Vegetal	0.07	0.06
People	14.95	15.71
Actions & movement	11.97	14.64
Things	48.45	48.53
Scenery	7.61	7.52

Significant sub-categories
Dogs, cats, wild mammals, wild aquatic, insects&reptiles, wild birds, rodents
-
Body parts
shopping
Food, health objects, technology, musical instruments
Landscapes and buildings