A Deep Architecture for Depression Detection

Arbee L.P. Chen

Journal of Intelligent Information Systems
(to appear)

PART 1

Background & Motivation

Background

- The World Health Organization predicts depressive disorders will be widespread in the next 20 years
  - World wide: 300,000,000; Asia: 50,000,000; Taiwan: 1,500,000
- These disorders may affect a person’s general health and habits
- 40% of the patients have suicidal thoughts and 10%~15% die by suicide

Motivation

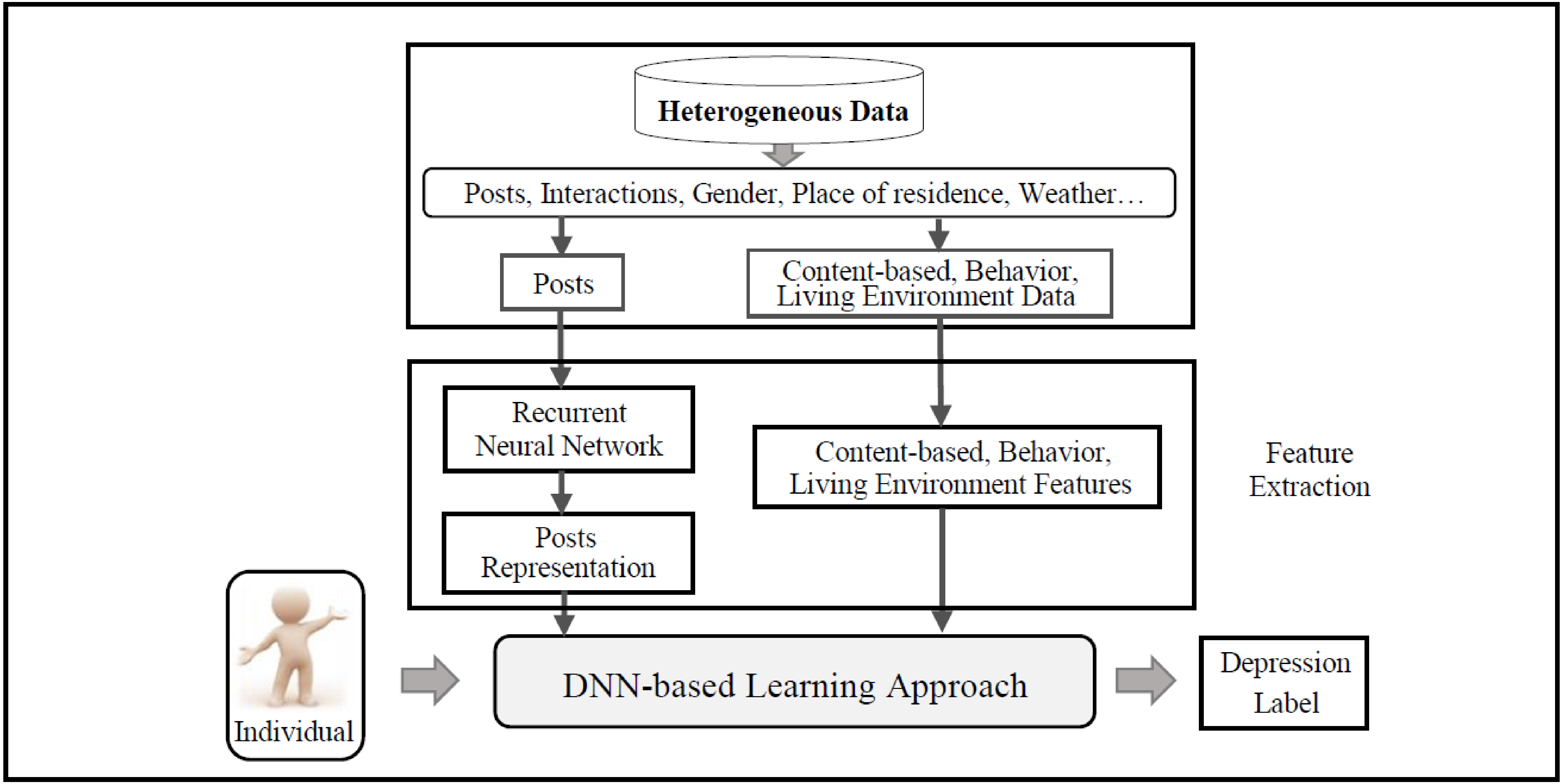
- 80% of the patients have it controlled after appropriate treatments
- Most patients are unaware of their illness and do not seek for clinical intervention until the symptoms become severe
- Our goal: Detecting depressive disorders in early stages to allow the patients to receive proper treatments

PART 2

Approach

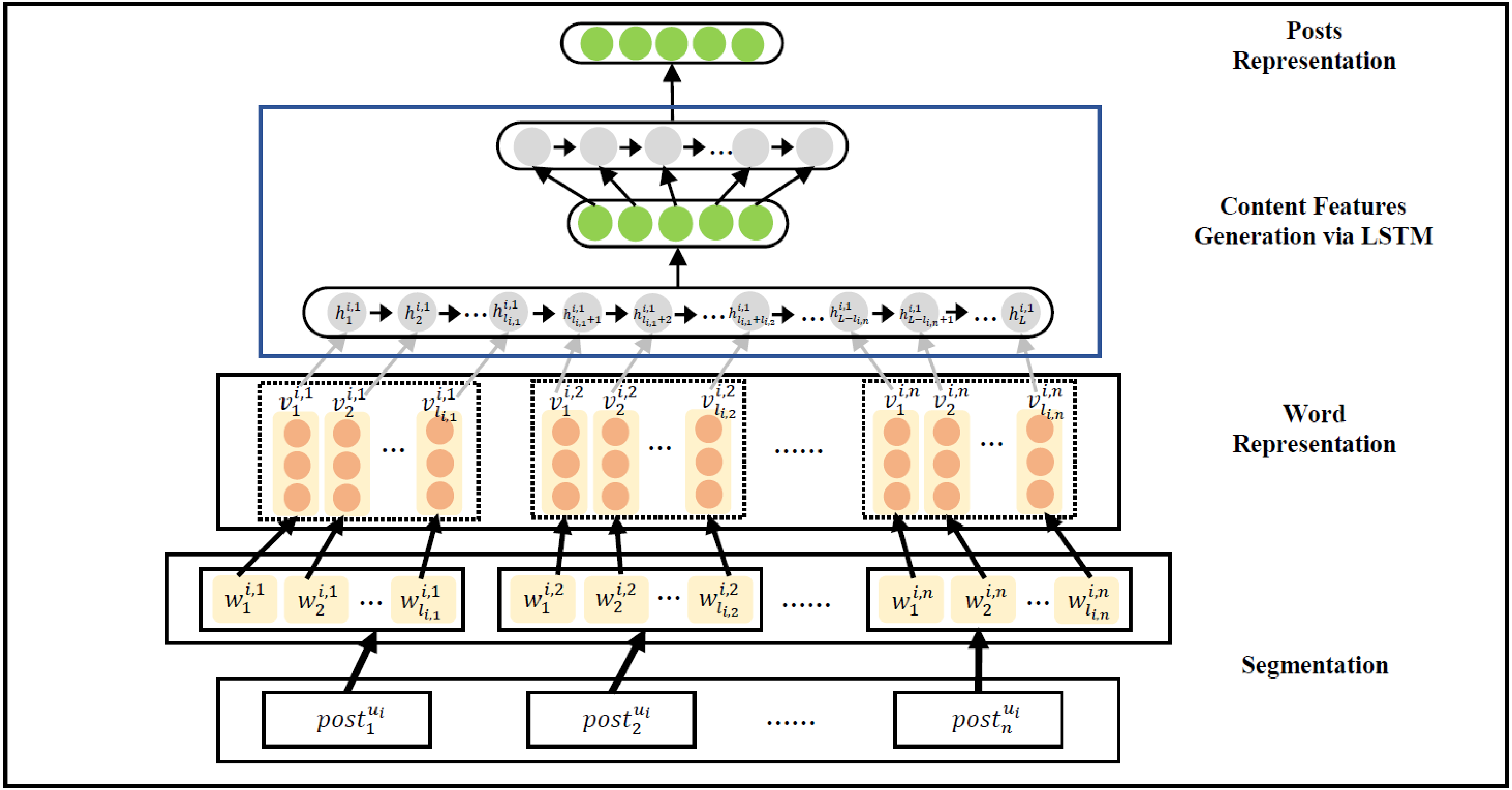
- 1,453 students from 58 universities in Taiwan were recruited
- These participants were asked to do CES-D depression screening test and their Facebook records collected
  - Center for Epidemiologic Studies Depression Scale
  - The CES-D results from the volunteers are used as the ground truth for building the prediction model
- Their living environment data were also collected from the government open data

- Employ Recurrent Neural Networks (RNN) to capture the semantics of the posts in social media
- Extract the key features which may affect depression from the content-based, behavior, and living environment data
- Construct a Deep Neural Network (DNN) model to predict a person’s label of depression


Heterogeneous Data
Posts, Interactions, Gender, Place of residence, Weather…
Posts
Content-based, Behavior, Living Environment Data
Recurrent Neural Network
Posts Representation
Content-based, Behavior, Living Environment Features
Feature Extraction
Individual
DNN-based Learning Approach
Depression Label

PART 3

Posts Representation



Posts Representation
Content Features Generation via LSTM
Word Representation
Segmentation
$post_1^{u_i}$
$post_2^{u_i}$
$post_n^{u_i}$

Data Collection

- The depression screening tests were done in December 2015
- All the posts published by the volunteers before 2015 are used to generate word representations
    - Totally 1,156,241 posts with 88,649 words

Word Representation

- The concept of word vector is employed to represent a word
  - Word2vec is an efficient tool for word representation, which was launched by Google in 2013
  - Take as its input a large corpus of text and produce a vector space, typically of several hundred dimensions, with each unique word in the corpus being assigned a corresponding vector in the space
  - Word vectors are positioned in the vector space such that words that share common contexts in the corpus are located in close proximity to one another in the space

Content Features Generation

- Sequentially concatenate the word vectors into a fixed length L of word vectors
- For example
  - Set L = 6
  - User ui published two posts (represented as word vectors) in order: $< v_1^{i,1}, v_2^{i,1}, v_3^{i,1} >$ and $< v_1^{i,2}, v_2^{i,2} >$
  - Concatenate all word vectors into an ordered vector set $< 0, v_1^{i,1}, v_2^{i,1}, v_3^{i,1}, v_1^{i,2}, v_2^{i,2} >$

Posts Representation

Embedding Layer	Input	(1000)
	Output	(1000,300)

LSTM	Input	(1000,300)
	Output	(1000,1500)

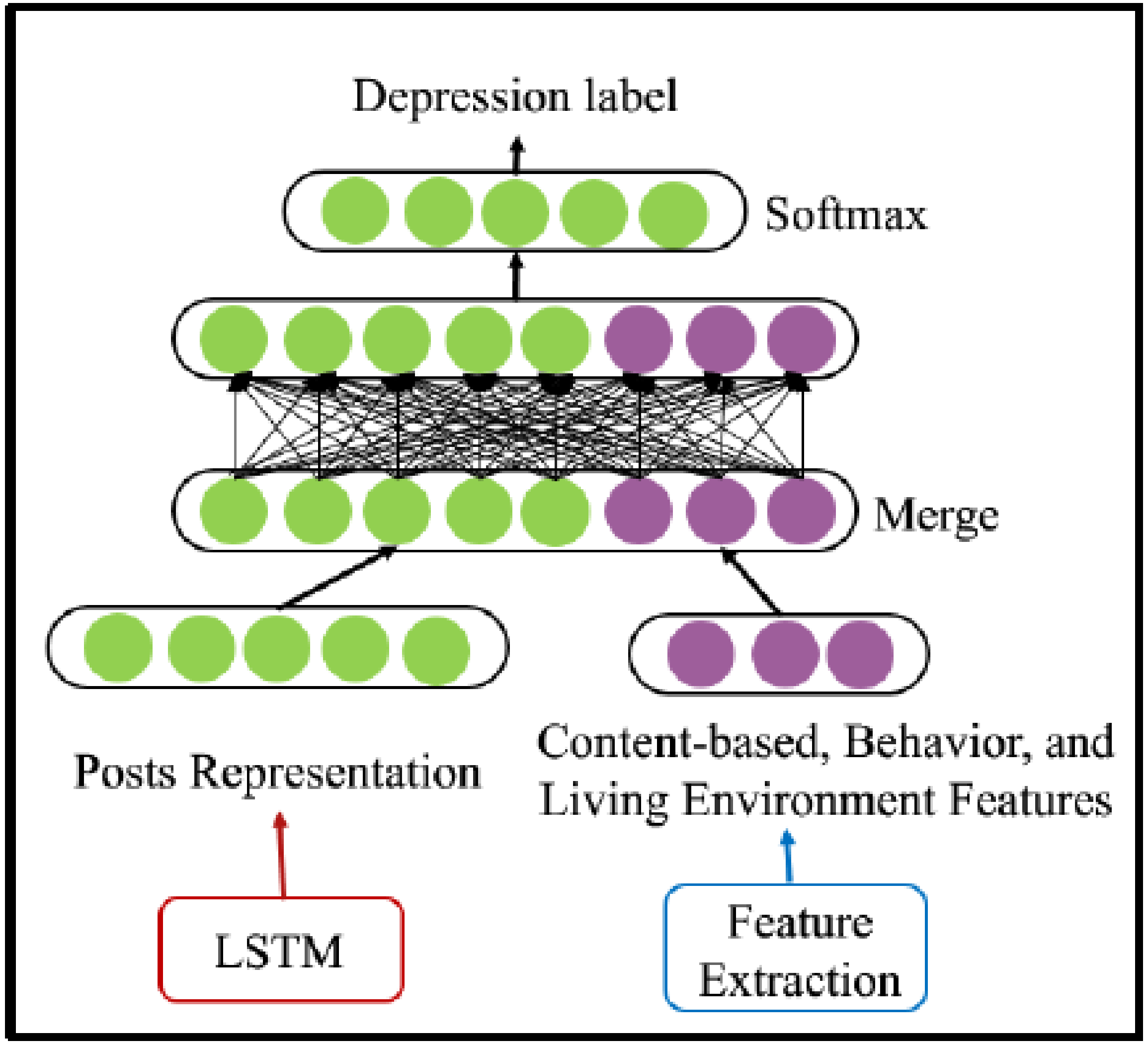
LSTM	Input	(1000,1500)
	Output	(500)

Posts representation

PART 4

Classification for Depression Labels

Merging Features to Build DNN



Feature Extraction

Content-based

A

Behavior

B

Living

Environment

C

Content-based features

- Positive words (PW): number of positive words used per post
- Negative words (NW): number of negative words used per post
- First person pronouns (FP): number of “I” used per post
- Links (LK): number of hyperlinks made per post
- Vocabulary (VB): number of words used in a time period
- Specific linguistic styles: number of nouns, verbs, adjectives, conjunctions, etc. used per post

Behavior features

- Number of posts (NP)
- Post time (PT) : number of posts posted during 0:00 – 6:00am / NP
- Number of uploaded photos (NU)
- Ratio of game posts (GP): number of game posts / NP
- Number of friends (NF)
- Average number of thumbs-ups (CT)
- Average number of profile photo changes (PC)
- Number of actions (NA): friend addition, posting, commenting, etc.

Living environment features

- Average sunshine hours (SH)
- Average rainfall days (RD)
- Average temperature (TP)
- Population density of the administrative districts (PD)
- Number of traffic accidents normalized by PD (TA)
- Number of fire accidents normalized by PD (FA)

Model Construction

- In total, we extract 48 features from publicly available data and social media, these features are merged with the posts representation to build a deep learning classifier
- The posts published in the three months prior to the depression screening tests are used as the training and testing data
  - Totally 1,294 volunteers with 121,767 posts

Label Acquisition

- The CES-D results are used to obtain the depression labels of the participants
- Existing research indicates that the cut-off point in the CES-D scale for depression in the binary label setting is 16
  - Participants with CES-D scores smaller than 16 are labeled as negative, and those with CES-D scores ≥ 16 are labeled as positive
  - There are 20 questions, each with maximum 3 points

PART 5

Experiment

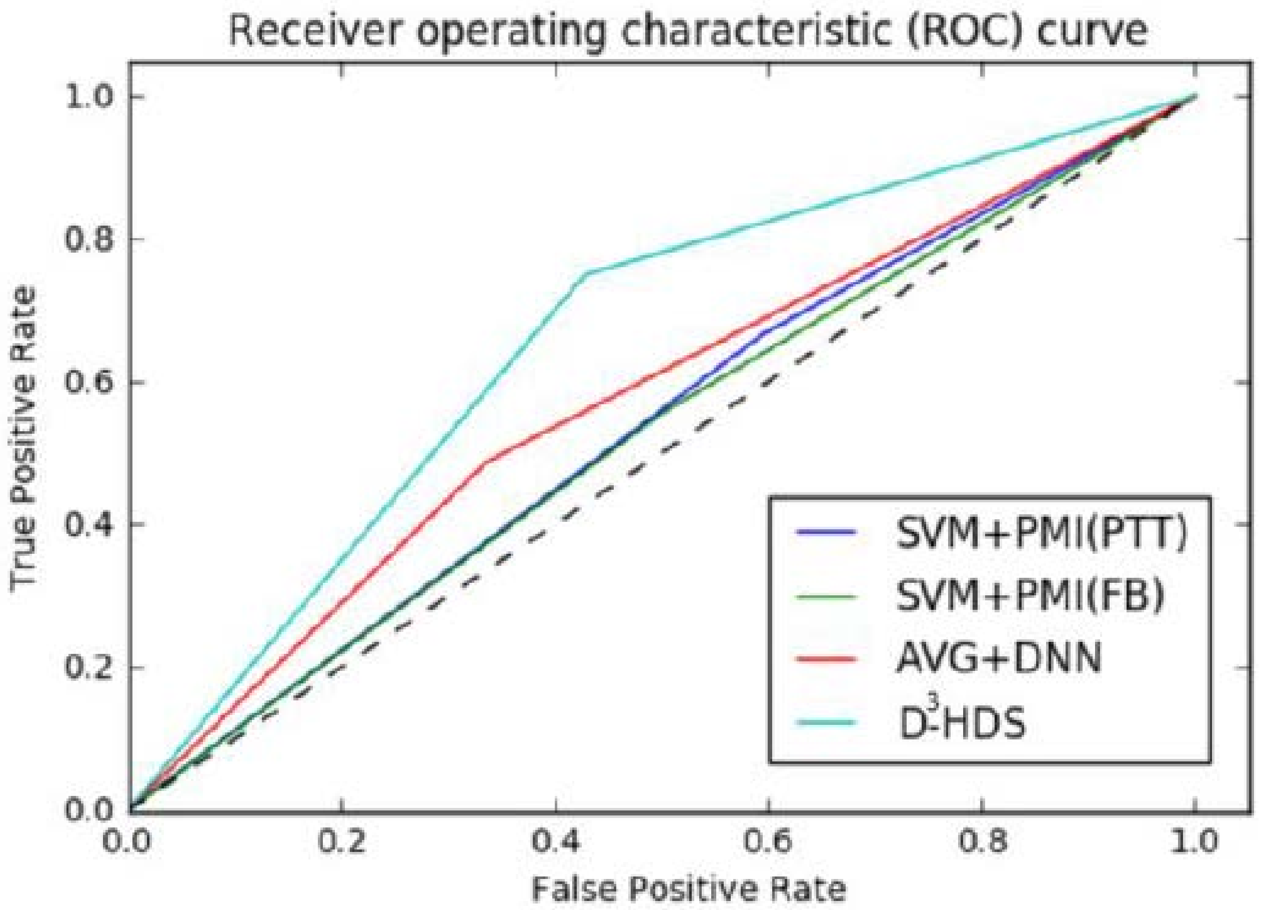
Parameter Tuning and Model Training

Hyper-parameters	Choice	Experiment Range
Word vector dim. (d)	300	50-500
LSTM input dim. (L)	1000	50-2000
LTSM output dim.	500	50-1500
LTSM layers	2	1-6

Other Parameters

- Word2vec window size: [2-8]
- UKW: [yes, no]
- Period of posts: [3 months, 6 months, 9 months, 1 year]
- Batch size: [1-817]
- Training, validation and testing split: [5.6:1.4:3]
- Activation function: [ReLu, Sigmoid]
- Loss function: [Cross entropy]
- Optimizer: [Adam, RMSprop]
- Epochs: [5-1000]

	Precision	Recall
SVM+PMI(PTT)	57.8%	52.4%
SVM+PMI(FB)	56.8%	50%
AVG+DNN	59.6%	66.3%
D^3-HDS	83.3%	71.4%



- True positive rate: correctly predicted in all positive cases
- False positive rate: incorrectly predicted in all negative cases

Thank you for your attention!