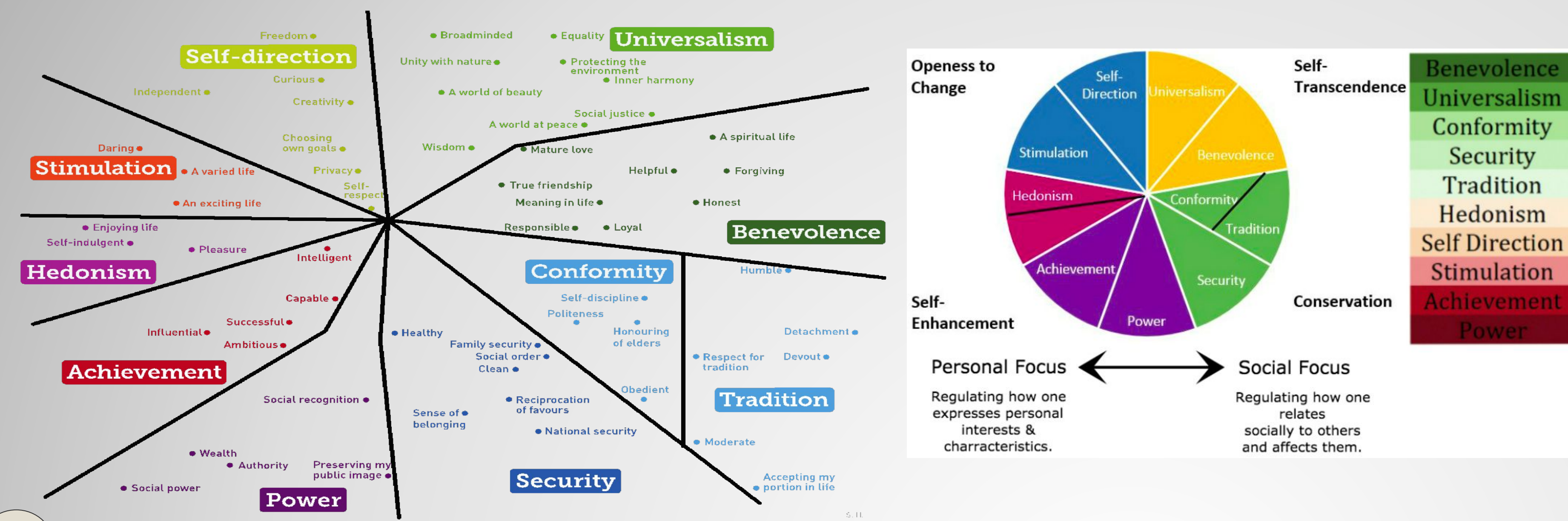


Upendra Kumar, Vishal Singh Rana, Srinivas PYKL & Amitava Das  
Indian Institute of Information Technology (IIIT), Sri City, AP, India, 517541  
{upendra.k14, vishal.s14, srinivas.p, amitava.das}@iiits.in

1

## Psycho-sociological Factors: Personality and Values



2

## Objective

- To study how psycho-sociological factors influence on **optimism/pessimism** at the individual level.
- Optimists are people who tend to have favourable outlook of their life whereas pessimists tend to derive negative interpretations from the events around them.
- Identify optimist/pessimist users from Twitter, analyze their personality (psychological) and values & ethics (sociological) at the community level.
- Empirical analysis reveals some interesting insights and behavioral patterns related to user level optimism/pessimism in different combinations of psychological and sociological factors.

2

## Data Sources

TABLE 1: An example of the instructions and format of the Portrait Values Questionnaire (PVQ). For each statement, the respondents should answer the question “How much like you is this person?” by checking one of the six boxes.

Here we briefly describe some people. Please read each description and think about how much each person is or is not like you. Tick the box to the right that shows how much the person in the description is like you.

HOW MUCH LIKE YOU IS THIS PERSON?	Very much like me	Like me	Some-what like me	A little like me	Not like me	Not like me at all
1. Thinking up new ideas and being creative is important to her. She likes to do things in her original way. SD	6	5	4	3	2	1
2. It is important to her to be rich. She wants to have a lot of money and expensive things. PO	6	5	4	3	2	1
3. She thinks it is important that every person in the world be treated equally. She believes everyone should have equal opportunities in life. UN	6	5	4	3	2	1
4. Its important to her to show her abilities. She wants people to admire what she does. AC	6	5	4	3	2	1
5. It is important to her to live in secure surroundings. She avoids anything that might endanger her safety. SE	6	5	4	3	2	1

## Optimism/Pessimism

- 500 potential optimists and 500 potential pessimists.
- Randomly selected 15 tweets per user.
- Each tweet was annotated for optimism/pessimism
- A dataset of 2000 tweets was created after filtering stage.

Be positive. Think positive and live positive. A new day is always a promise of the best things to come.

I am soooooo fed up!!! just leave me alone! my life!!!!

This is a bad time for me.

Pettiest phone ever. I love it.

Xianzhi Ruan, Steven R Wilson, and Rada Mihalcea. 2016. Finding optimists and pessimists on twitter. In the 54th Annual Meeting of the Association for Computational Linguistics.

## Values

- 367 users' data using PVQ
- Highest no. of tweets: 15K
- Lowest no. of tweets: 100
- Avg. no. tweets 1,608

## Personality

- 250 users' data
- 10K Facebook posts

## SNAP

- 1562 ground truth communities
- 80K nodes and 17.5 lakh edges
- Communities having size less than 5 nodes and n(tweets) < 100 were discarded.

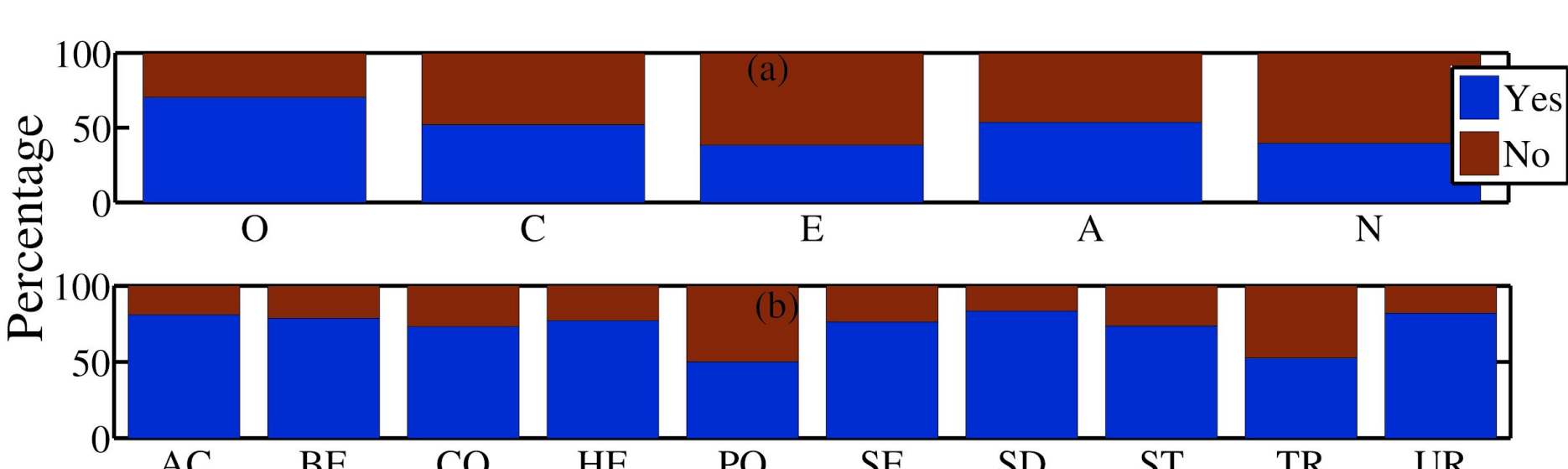
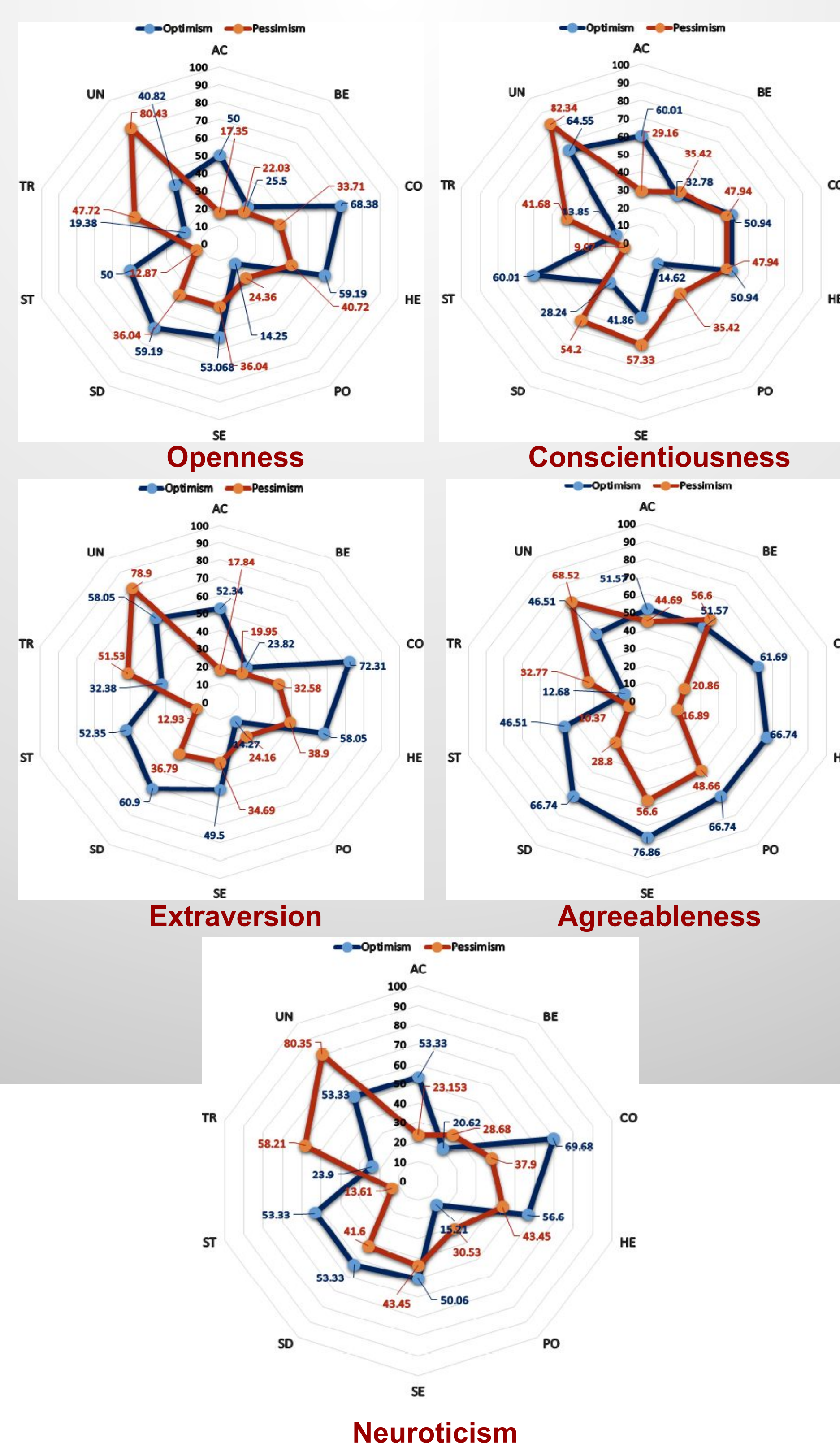


Figure 1: Values Class Distribution in the Twitter Corpus.

7

## Understanding Correlations between Optimism/Pessimism with Psycho-sociological factors



- Openness positively correlated with optimism irrespective of communities.
- People high in conscientiousness are positively correlated with optimism in stimulation or achievement oriented settings.
- Extroversion has positive correlation with very optimistic people in communities highly oriented towards achievement, conformity, hedonism, security, self-direction and stimulation owing to their high energy and positive emotions.
- For optimism, certain social factors dominate over all user-level factors : AC > CO > ST.
- Traditional or Power oriented settings generally seem to be positively correlated with Pessimism owing to sense of responsibility coupled with pressure for maintaining stability and security.

3

## Fuzzy/Overlapping Distributions

Personality Traits	O	C	E	A	N
Openness (O)			52.27	40.90	60.22
Conscientiousness (C)	70.76		46.92	57.69	28.46
Extraversion (E)	75.00	63.54		59.37	22.91
Agreeableness (A)	79.10	55.97	42.53		25.37
Neuroticism (N)	68.68	37.38	22.25	34.56	

Table 1: Fuzzy distributions of Big5 Personality traits in the Personality corpus.

Schwartz Values	AC	BE	CO	HE	PO	SE	SD	ST	TR	UN
Achievement(AC)	24.12	19.84	19.49	29.41	41.54	15.81	11.77	19.85	41.91	17.28
Benevolence(BE)	18.59	23.42	19.84	31.12	52.97	18.68	10.51	22.18	42.8	7
Conformity(CO)	17.63	24.04	25.32	47.58	12.64	17.47	24.91	35.32	15.99	12.88
Hedonism(HE)	12.64	33.52	22.53	27.47	43.35	21.03	9.01	14.6	45.92	12.88
Power(PO)	17.63	24.82	15.47	33.81	46.04	13.74	17.03	41.21	20.33	12.88
Security(SE)	21.05	24.34	26.97	30.26	48.35	20.72	20.72	47.04	12.5	12.5
Stimulation(ST)	18.66	25.37	24.63	25.75	43.66	19.03	10.08	42.91	16.04	16.04
Tradition(TR)	18.13	23.83	9.84	34.72	44.56	11.4	16.58	20.73	42.91	17.1
Universalism(UN)	24.75	20.07	24.41	32.11	51.51	20.4	11.04	24.75	46.49	46.49

Table 2 : Fuzzy distributions of Schwartz' values in the Twitter corpus. Schwartz' theory explains how the values are interconnected and influence each other, since the pursuit of any of the values results in either an accordance with one another (e.g., Conformity and Security) or a conflict with at least one other value (e.g., Benevolence and Power)

4

## Psycholinguistic Features and Classifiers

- f1: Word n-grams
- f2: Linguistic Features (LIWC, Harvard General Inquirer, MRC, Sensicon)
- f3: Speech Act features
- f4: Sentiment Amplifiers (Exclamation, Quotes, Emoticons)
- f5: NRC Lexicon, Sentiwordnet
- f6: Topic words obtained from topic model (LDA)

Models	f1	f2	f3	f4	f5	f6	F-Score
Personality	+	+	+	-	-	-	79.35%
Values	-	+	+	-	-	+	80.10%
Optimism/Pessimism	+	+	-	-	+	-	77.89%

Table 3: Features used for developing different models

5

## Extracting Correlations between Psycho-sociological Factors and Degree of Optimism/Pessimism

- Extract communities homogeneous with respect to a particular value.
- Shannon's Entropy is used to measure the degree of randomness/degree of homogeneity.
- For each user, the scores for all ten values obtained as confidence scores from Values classifier, are converted to probability values using softmax formulation.
- For each community, entropy scores are computed across all ten values.
- A community is assumed to be more homogeneous in a particular class/value if entropy across that value is lower.

	AC	BE	CO	HE	PO	SE	SD	ST	TR	UN
$u_{(1)}$	0.91	0.47	0.02	0.07	0.32	0.24	0.65	0.78	0.94	0.10
$u_{(2)}$	0.97	0.40	0.49	0.50	0.56	0.83	0.62	0.73	0.04	0.08
$u_{(3)}$	0.99	0.75	0.50	0.72	0.38	0.60	0.75	0.02	0.57	0.62
$u_{(4)}$	0.77	0.44	0.40	0.16	0.19	0.55	0.73	0.08	0.53	0.25
$u_{(5)}$	0.29	0.02	0.26	0.56	0.41	0.23	0.95	0.02	0.79	0.86
$X_{(i)}$	1.54	1.40	1.40	1.39	1.55	1.50	1.59	0.99	1.42	1.28
$S_{(i)}$	0.87	-0.12	-0.12	-0.19	0.95	0.57	1.26	-2.35	0.00	-0.87
$T_{(i)}$	1.0	0.0	0.0	0.0	1.0	1.0	1.0	0.0	0.0	0.0

Table 4 : Estimating classes in which a community is more homogenous as compared to other classes.

$$N_{(i,j,k)} = \frac{\exp(A_{(i,j,k)})}{\|\exp(A_{(i,.,k)})\|_1}$$

Here  $A_{(i,j,k)}$  represents confidence score for user  $U(k)$  in community  $C(i)$  and value  $V(j)$ .  $N_{(i,j,k)}$  represents corresponding probability values for latter confidence scores.

$$X_{(i,k)} = - \sum_{j=1}^n N_{(i,j,k)} * \log N_{(i,j,k)}$$

$X_{(i,k)}$  represents the entropy score of community  $C(i)$  across value  $V(k)$ . Further using zero-mean unit variance method  $X_{(i,k)}$  is scaled.

Based on 0-threshold, values greater than 0 are set to 1 and less than 0 are set to 0 to obtain  $T_{(i,k)}$ .

## Aggregate Analysis

Further, using personality classifier,  $P(i)$  for each user  $U(i)$  is obtained across five personality traits.

In order to obtain, the correlation values between optimism/pessimism with personality and values, the analysis is done separately for five personality traits.

$$I_{(i,j,0)} = \sum_{u_{(k)} | (P_{(k,i)}=1) \cap (T_{(k,j)}=1)} O_{(k,j,0)}$$

Here  $I_{(i,j,0)}$  represents aggregate degree of correlation of optimism with  $i$ th personality trait and  $j$ th value. Similarly,  $I_{(i,j,1)}$  represents aggregate degree of correlation of pessimism with socio-psychological factors.