Binxuan Huang
binxuanh@cs.cmu.edu

Kathleen M. Carley
kathleen.carley@cs.cmu.edu

Carnegie Mellon University
www.casos.cs.cmu.edu

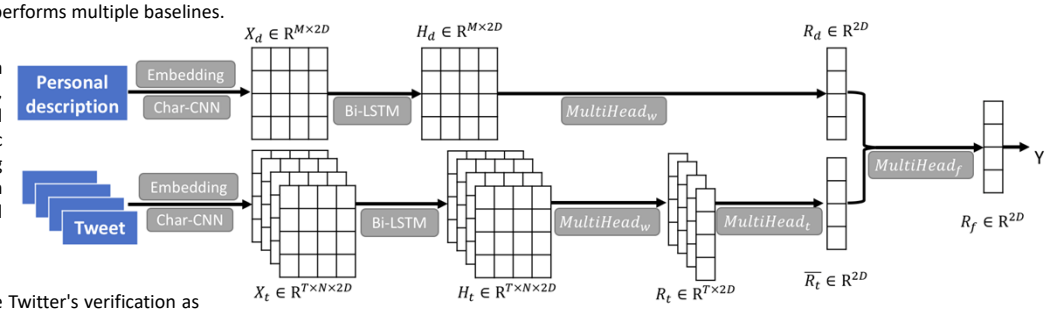
You Are What You Tweet: A Content-Based Approach for Twitter User Identity Classification

Introduction

- An identity is a characterization of the role an individual takes on. It is often described as the social context specific personality of an individual actor or a group of people. Identities can be things like jobs (e.g. ``lawyer'', ``teacher''), gender (man, woman), or a distinguishing characteristic (e.g. ``a shy boy'', ``a kind man''). There are many different kinds of actors using social media, e.g., people, organizations, and bots. Each type of actors has different motivations, different resources at their disposal, and may be under different internal policies. If we want to understand who is controlling the conversation and whom is being impacted, it is important to know what types of actors are doing what.
- In this paper, our goal is to classify Twitter users based on their identities. We first collect a coarse-grained public figure dataset automatically, then manually label a more fine-grained identity dataset. We propose a hierarchical self-attention neural network for Twitter user identity classification. Our experiments demonstrate that the proposed model significantly outperforms multiple baselines.

Method

- Our model first maps each word into a low dimension word embedding space, then it uses a Bidirectional LSTM neural network to extract context specific semantic representations for words. Using several layers of multi-head attention neural networks, it generates a final classification feature vector.



Datasets

- The first is a public figure dataset. We use Twitter's verification as a proxy for public figures. These verified accounts include users in music, government, sports, business, and etc. We sampled 156746 verified accounts and 376371 unverified accounts through Twitter's sample stream data.
- In addition, we introduce another human labeled identity dataset for more fine-grained identity classification, which contains seven identity classes: ``news media'', ``news reporter'', ``government official'', ``celebrity'', ``company'', ``sport'', and ``normal people''.

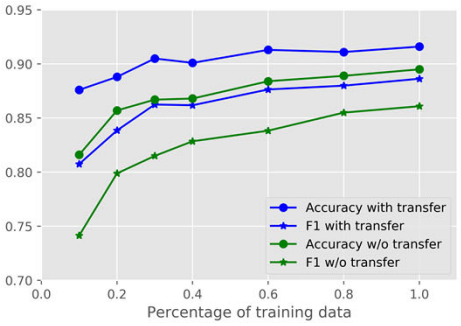
	Public Figure		Identity						
	Verified	Unverified	Media	Reporter	Celebrity	Government	Company	Sport	Normal
Train	152368	365749	1140	614	876	844	879	870	6623
Dev.	1452	3548	52	23	38	40	35	43	269
Test	2926	7074	97	39	75	81	66	74	568

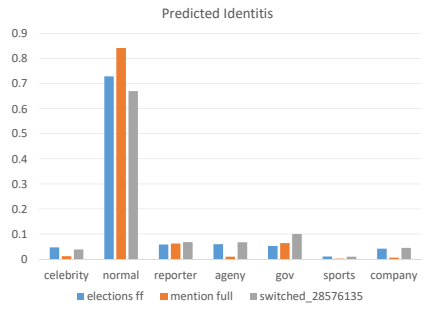
Table 1: A brief summary of our two datasets.

Results

		Public Figure		Identity	
		Accuracy	Macro-F1	Accuracy	Macro-F1
Baselines	MNB	81.81	82.79	82.9	75.91
	SVM	90.60	88.59	85.9	80.19
	fastText	90.93	89.01	85.7	80.01
	CNN	91.45	89.85	85.9	81.24
	Bi-LSTM	93.10	91.84	86.5	84.25
	Bi-LSTM-ATT	93.23	91.94	87.3	83.35
Ablated Models	w/o attentions	93.78	92.45	87.0	83.26
	w/o charcnn	93.47	92.23	89.0	85.39
	w/o description	92.39	90.90	86.7	81.56
	w/o tweets	91.62	89.77	84.2	78.41
	Full Model	**94.21**	**93.07**	**89.5**	**86.09**
	Full Model-transfer			**91.6**	**88.63**

Table 3: Comparisons between our methods and baselines.





Conclusion

We introduce two datasets for online user identity classification. One is automatically extracted from Twitter, the other is a manually labelled dataset. We present a novel content-based method for classifying social media users into a set of identities (social roles) on Twitter. Our experiments on two datasets show that our model significantly outperforms multiple baseline approaches. Using one personal description and up to twenty tweets for each user, we can identify public figures with accuracy 94.21% and classify more fine-grained identities with accuracy 89.5%.

This work was supported in part by the Office of Navel Research/UALR through, Grant ONR N00014-17-1-2605 and the Center for Computational Analysis of Social and Organization Systems (CASOS). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing official policies, either expressed or implied, of the Office of Navel Research/UALR or the U.S. government.

isr institute for SOFTWARE RESEARCH