

**CHENNAI – PONDICHERRY**

**Predicting Microblog Sentiments via Weakly Supervised Multi-Modal Deep Learning**

**Abstract**

Predicting sentiments of multimodal microblogs composed of text, image, and emoticon have attracted ever-increasing research focus recently. The key challenge lies in the difficulty of collecting a sufficient amount of training labels to train a discriminative model for multimodal prediction. One potential solution is to exploit the labels collected from social media users, which is, however, restricted by the negative effect of label noise. Besides, we have quantitatively found that sentiments in different modalities may be independent, which disables the usage of previous multimodal sentiment analysis schemes in our problem. In this paper, we introduce a weakly supervised multimodal deep learning (WS-MDL) scheme toward robust and scalable sentiment prediction. WS-MDL learns convolutional neural networks iteratively and selectively from “weak” emoticon labels, which are cheaply available and noise containing. In particular, to filter out the label noise and to capture the modality dependency, a probabilistic graphical model is introduced to simultaneously learn discriminative multimodal descriptors and infer the confidence of label noise. Extensive evaluations are conducted in a million-scale, real-world microblog sentiment dataset crawled from Sina Weibo. We have validated the merits of the proposed scheme by quantitatively showing its superior performance over several state-of-the-art and alternative approaches.