Opinion Mining from Customers' Reviews

XIE Yuxiao
Supervisor: Prof. Lam Wai

Department of Information Engineering, The Chinese University of Hong Kong

Introduction

Opinion mining is the computational study of people's opinions, attitudes, and emotions toward entities, topics and their attributes. In this project, a novel model based on Restricted Boltzmann Machines (RBMs) with prior information was applied to extract aspect and sentiment related words in an unsupervised setting.

Objective

- To extract and summarize salient aspects of entities and determining relevant sentiment polarities from reviews.

Always a fun place! → <Ambience><Positive>

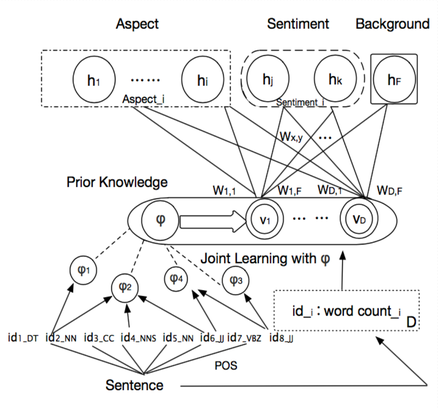
The food is delicious! → <Food><Positive>

The staff isn't the friendliest or most competent , and I am stickler for service , but everything else about this place makes up for it. → <Staff><Negative>

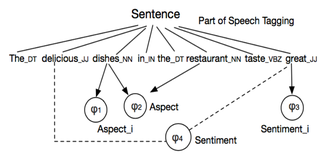
Great for groups , great for a date , great for early brunch or a nightcap. → <Ambience><Positive>

Model Architecture

Model



Priors



Objective function

$$\ln L_S = \ln \prod_{i=1}^{n_s} P(\mathbf{v}^i) - \sum_{i=1}^{n_s} \Bigg[\lambda_1 \ln \prod_{j=1}^{F_1-1} \prod_{k \in R_1} \Big[P(h_j = 1 \mid \hat{v}^k) - p_{A_j, v_k} \Big]^2 + \lambda_2 \ln \prod_{k \in R_2} \Big[\sum_{j=1}^{F_1} P(h_j = 1 \mid \hat{v}^k) - p_{A, v_k} \Big]^2 + \lambda_3 \ln \prod_{j=F_2}^{F_2+1} \prod_{k \in R_3} \Big[P(h_j = 1 \mid \hat{v}^k) - p_{S_j, v_k} \Big]^2 + \lambda_4 \ln \prod_{k \in R_4} \Big[\sum_{j=F_2}^{F_2+1} P(h_j = 1 \mid \hat{v}^k) - p_{S, v_k} \Big]^2 \Bigg]$$

Experiment

Dataset

Restaurant review dataset information

# of reviews	# of tokens
52624	179139

- Documents in this dataset are annotated with one or more labels from a gold standard label set
  S = {*Food, Staff, Ambience, Price, Anecdote, Miscellaneous*}

RBM training: Contrastive Divergence algorithm

Algorithm 1

RBMupdate(x_1, ϵ, W, b, c)
This is the RBM update procedure for binomial units. It can easily adapted to other types of units.
x_1 is a sample from the training distribution for the RBM
ϵ is a learning rate for the stochastic gradient descent in Contrastive Divergence
W is the RBM weight matrix, of dimension (number of hidden units, number of inputs)
b is the RBM biases vector for hidden units
c is the RBM biases vector for input units

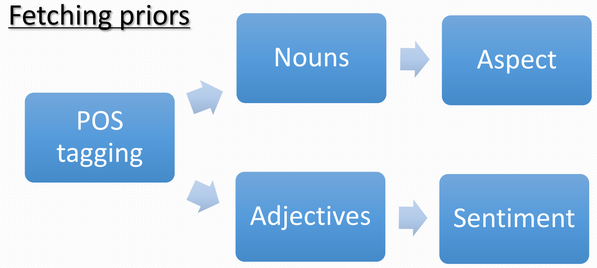
for all hidden units i **do**
- compute $Q(\mathbf{h}_{1i} = 1|x_1)$ (for binomial units, $\text{sigm}(b_i + \sum_j W_{ij} x_{1j})$)
- sample $\mathbf{h}_{1i}$ from $Q(\mathbf{h}_{1i}|x_1)$

end for
for all visible units j **do**
- compute $P(x_{2j} = 1|\mathbf{h}_1)$ (for binomial units, $\text{sigm}(c_j + \sum_i W_{ij} \mathbf{h}_{1i})$)
- sample x_{2j} from $P(x_{2j} = 1|\mathbf{h}_1)$

end for
for all hidden units i **do**
- compute $Q(\mathbf{h}_{2i} = 1|x_2)$ (for binomial units, $\text{sigm}(b_i + \sum_j W_{ij} x_{2j})$)

end for
- $W \leftarrow W + \epsilon(\mathbf{h}_1 x_1' - Q(\mathbf{h}_{2\cdot} = 1|x_2)x_2')$
- $b \leftarrow b + \epsilon(\mathbf{h}_1 - Q(\mathbf{h}_{2\cdot} = 1|x_2))$
- $c \leftarrow c + \epsilon(x_1 - x_2)$

Fetching priors



Conclusion

Evaluation

$$\mathrm{P} = \frac{TP}{TP + FP}, \mathrm{R} = \frac{TP}{TP + FN}, F_1 = \frac{2PR}{P + R}$$

Aspect	Precision	Recall	F1
Food	0.6765	0.7657	0.7183
Staff	0.5869	0.6009	0.5938
Ambience	0.3965	0.4388	0.4166

Comparison

Aspect	RBM	LDA	Our model
Food	0.6359	0.6579	0.7183
Staff	0.4028	0.5802	0.5938
Ambience	0.2065	0.2546	0.4166

Representative words

<Food> Sauce, pork, salad, flavour, dessert, outstanding, duck, meat, delicious, appetizer, cheese, bean, grilled, spinach, stick, entrée, filet, dish, pepper, excellent, small, amaze, chicken, tomato

<Staff> Service, staff, friendly, bad, attentive, help, accommodate, bartender, waiter, server, polite, nice, attitude, waitress, professional, quick, rude

<Ambience> Place, music, ambience, room, loud, portion, hip, try, look, village, atmosphere, romantic, space, cozy, area, east, casual, time, square, noisy

References

[1] Sentiment-Aspect Extraction based on Restricted Boltzmann Machines. L Wang, K Liu, Z Cao, J Zhao, G de Melo, ACL (1), 616-625, 2015
[2] Learning Deep Architectures for AI. Yoshua Bengio. Technical Report 1312, 26-27, 2009