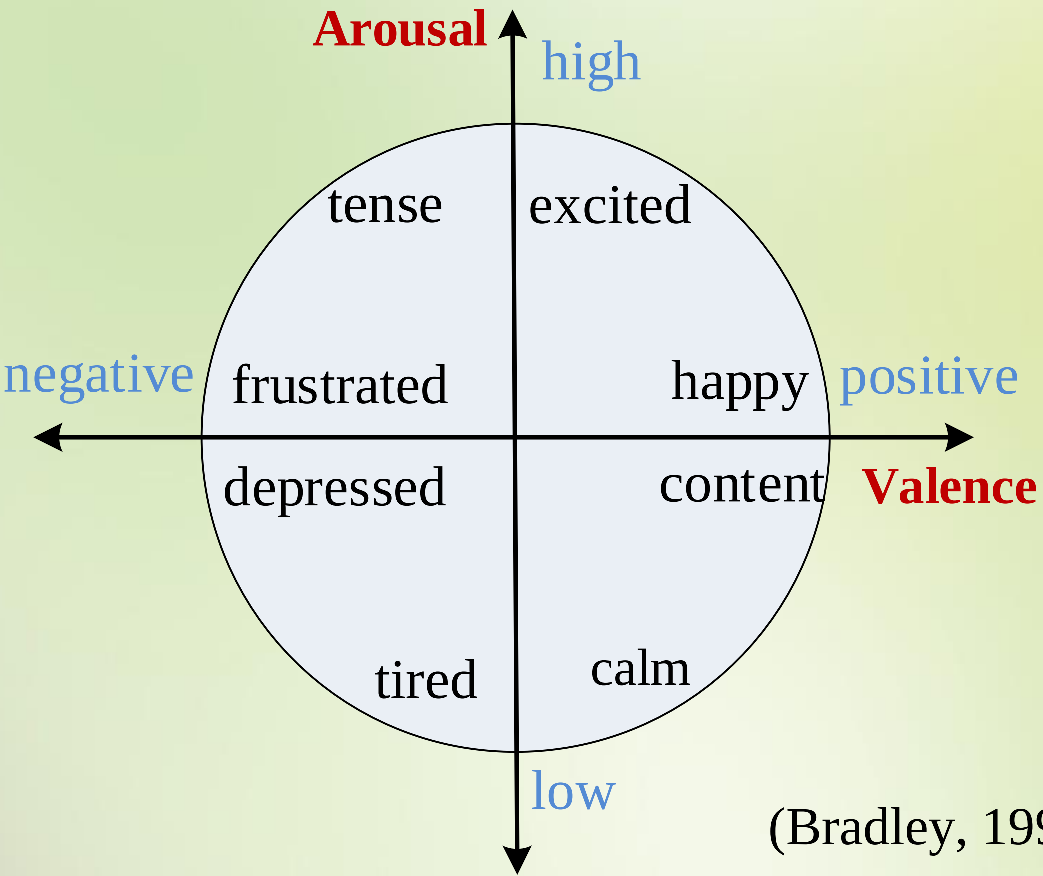
A Regression Approach to Valence-Arousal Ratings of Words from Word Embedding

Minglei Li, Yunfei Long, Qin Lu

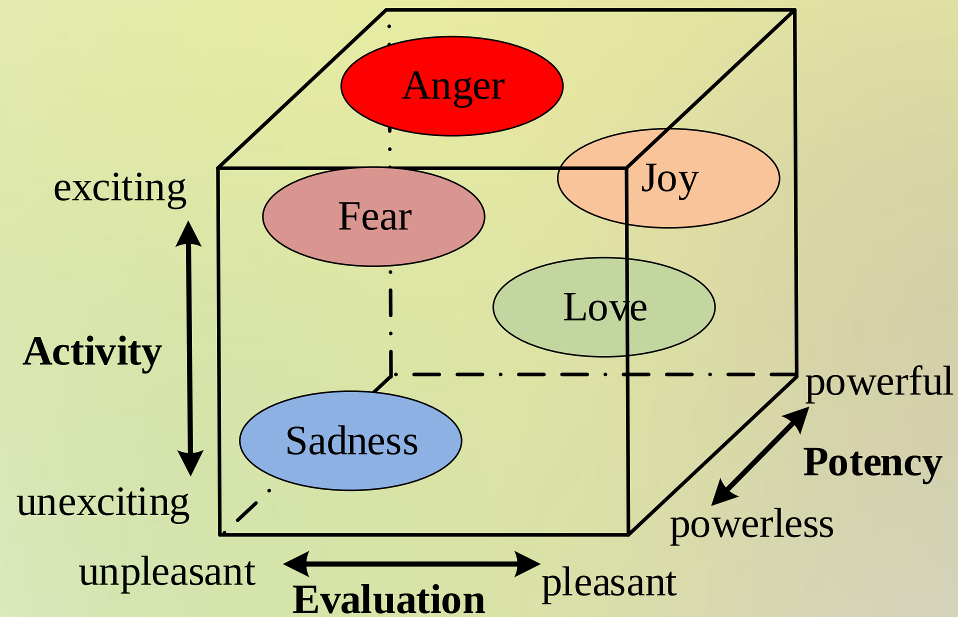
The Hong Kong Polytechnic University

1. Introduction

- **Emotion lexicon:** a word annotated with emotional information.
  - Discrete emotion model
    - *surprise, happiness, anger, fear, disgust, sadness* (Ekamn 1993)
  - Dimensional emotion model



(Bradley, 1999)



(Osgood, 1957)

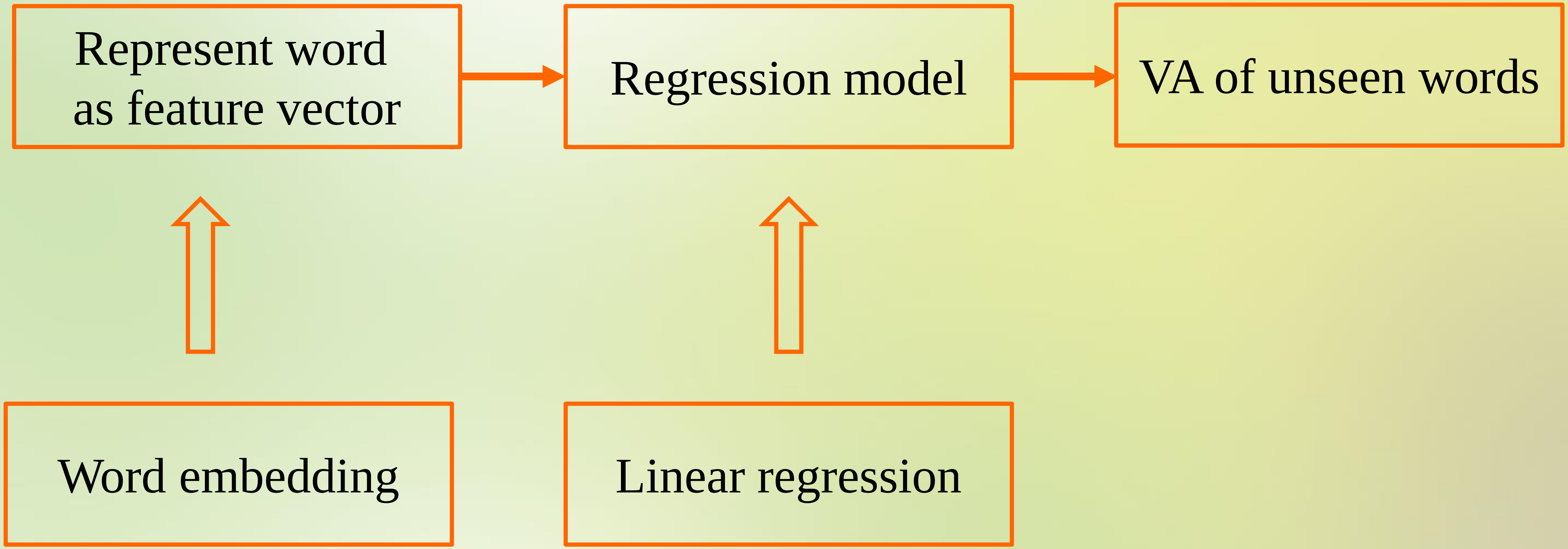
- **Motivation:**
  - Dimensional model is more expressive
  - Limited resources
- **Previous Methods:**
  - Manual annotation (Bradley 1999, Yu 2016): time-consuming
  - Crowdsourcing (Warriner, 2013): expensive
  - Linear regression
    - Source ratings (English) to target ratings (Chinese) in each cluster based on SUMO (Wei, 2011)
    - Weighted addition of seed words (Malandrakis, 2011)
    - Limitation: Based on knowledge base, not extendable

Previous Methods (cont.)

- Label propagation (seed words + word graph + propagation agorithm)
  - WordNet based graph + label propagation (Alhothali, 2015)
  - Weighted graph based on word embedding + PageRanking (Yu, 2015)
  - Graph based on word embedding + label propagation (Hamilton, 2016)
  - Limitations:
    - Scale not controllable
    - Seed word sensitive
    - Computation expensive

2. Our Proposed Method

- **Word feature vector + regression:**



2.1 Word Embedding

- **Word representation**
  - One-hot representation (high dimension)
    - dog [0,0,1,0,…,0]
    - cat [0,1,0,.., 0]
  - Word embedding (Based on distributional hypothesis, low dimension)
    - dog [0.1,0.3, 0.5, …, 1.2]
    - cat [0.2, 0.4, 0.7, ..., 1.1]

$$vec(king) - vec(queen) = \mathrm{vec}(\mathrm{man}) - \mathrm{vec}(\mathrm{woman})$$

$$vec(France) - vec(Pairs) = \mathrm{vec}(\mathrm{China}) - \mathrm{vec}(\mathrm{Beijing})$$

2.1 Word Embedding

- **Skip-Gram with Negative Sampling (SGNS, Mikolov 2013)**

$$P(D=1|w,c) = \sigma(\vec{w}\cdot\vec{c}) = \frac{1}{1+e^{-\vec{w}\cdot\vec{c}}}$$

$$\log\sigma(\vec{w}\cdot\vec{c}) + k\cdot\mathbb{E}_{c_N\sim P_D}\left[\log\sigma(-\vec{w}\cdot\vec{c}_N)\right]$$

⇩

Word vector $\vec{w}$
Context vector $\vec{c}$

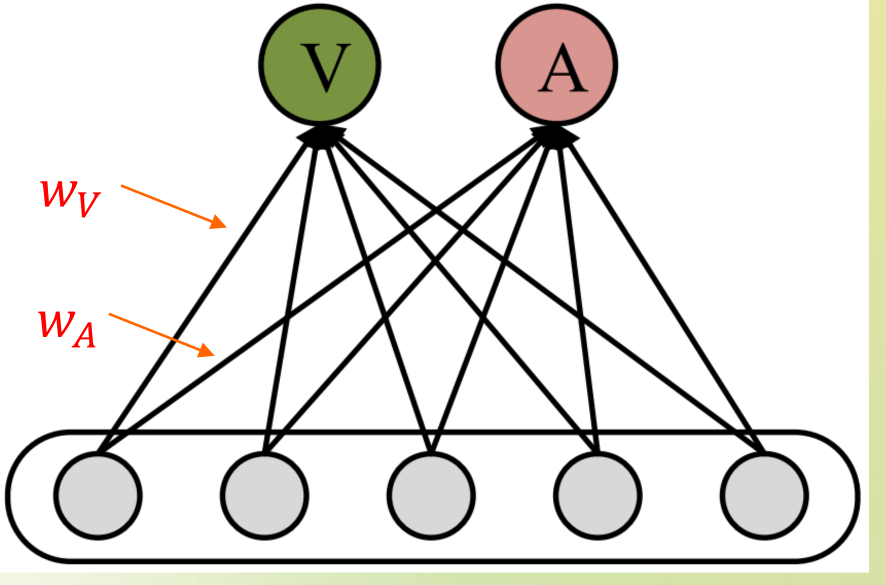
2.2 Regression Model

- **Assumption**
  - Word embedding encodes general word meaning (*father vs dad*)
  - Each dimension contributes differently to different meanings

- **Regression on Word Embedding (RoWE)**

Affective meaning

Word embedding



3. Experiment

- **Settings:**
  - VA lexicons:

Lexicon	Number	Language
ANEW (Bradley 1999)	1,034	English
E-ANEW (Warriner 2013)	13,915	English
CVAW (Yu 2016)	1,653	Chinese

  - 5-fold cross validation
  - Word Embedding
    - English trained on Wikipedia
    - Chinese trained on Baidu Baike
    - Dimension: 300
    - Trained using word2vec tool
  - Baseline: Weighted Graph (**WG**, Yu, 2015): iter (10), α (grid search)

3. Experiment

Evaluation metrics:

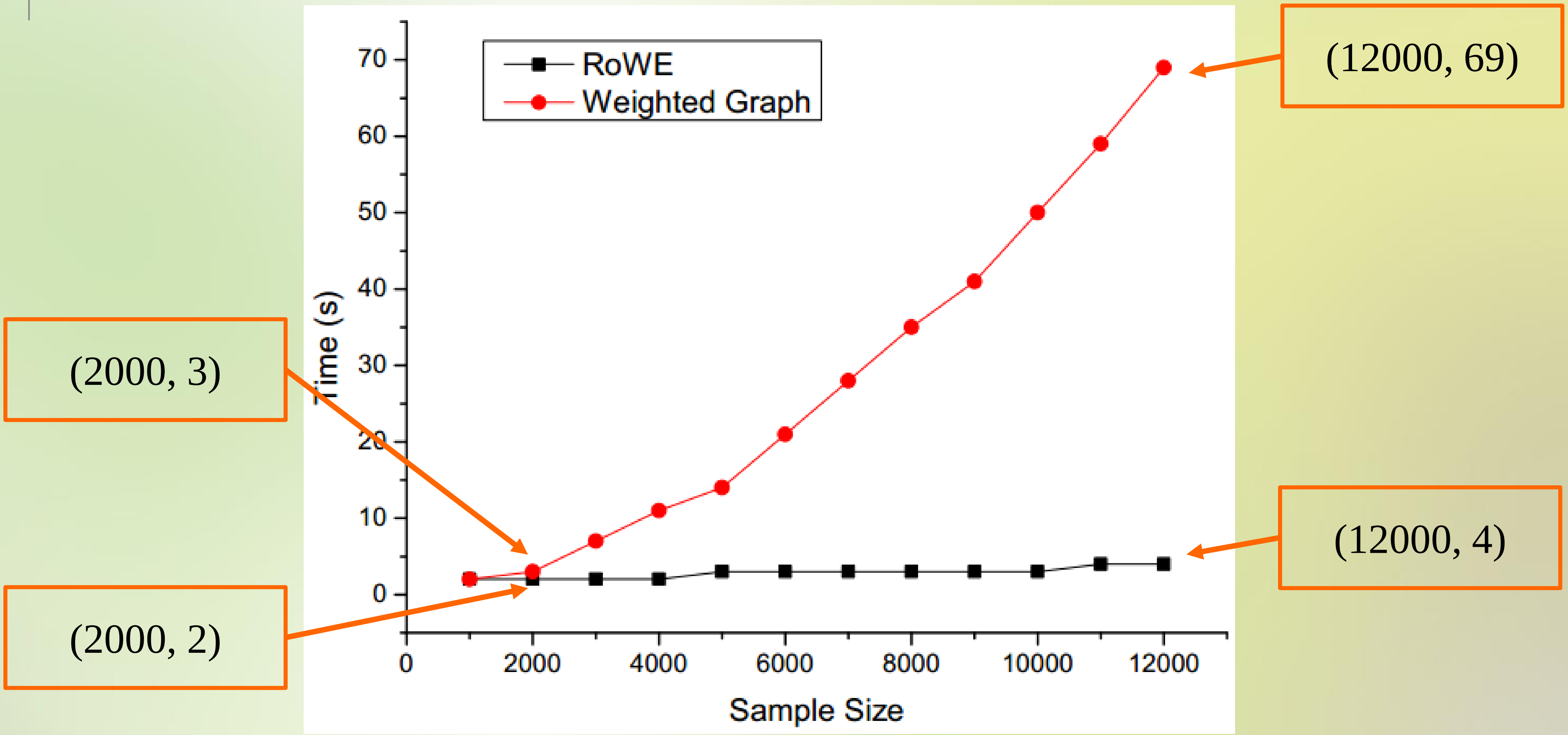
- Root mean squared error (RMSE): $RMSE = \sqrt{\sum_{i=1}^{n} (A_i - P_i)^2 / n}$
- Mean absolute error (MAE) $MAE = \frac{1}{n}\sum_{i=1}^{n} |A_i - P_i|$
- Mean absolute percentage error (MAPE) $MAPE = \frac{1}{n}\sum_{i=1}^{n} |A_i - P_i| / A_i$
- Pearson correlation coefficient $\rho = \frac{\sum_{i=1}^{n} (A_i - \bar{A})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^{n} (A_i - \bar{A})^2 (P_i - \bar{P})^2}}$

3.1 Experiment Result

			RMSE	MAE	MAPE	ρ
ANEW	Valence	**WG**	1.763	1.523	40.68	0.809
		RoWE	**0.957**	**0.915**	**22.11**	**0.828**
	Arousal	**WG**	0.993	0.815	17.11	0.602
		RoWE	**0.840**	**0.705**	**14.52**	**0.634**
E-ANEW	Valence	**WG**	1.175	0.938	22.27	0.702
		RoWE	**0.774**	**0.609**	**13.57**	**0.794**
	Arousal	**WG**	0.860	0.687	17.19	0.508
		RoWE	**0.706**	**0.559**	**13.92**	**0.616**
CVAW	Valence	**WG**	1.594	1.403	35.77	0.829
		RoWE	**0.827**	**0.684**	**18.26**	**0.877**
	Arousal	**WG**	1.204	0.977	19.18	0.563
		RoWE	**0.872**	**0.762**	**14.41**	**0.661**

3.1 Experiment

- **Computation Complexity:**
  - Using E-ANEW, varying the data size



3.2 IALP 2016 Shared Task

- **IALP 2016 Shared Task:**
  - Predict Chinese words valence arousal ratings.
  - 1,653 words for training. At least 1,000 for testing.
  - Both word embedding and context embedding trained on Baidu Baike.

	Valence			Arousal		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE
Word vec[1]	0.913	0.700	18.7	0.971	0.778	14.9
Context vec[1]	0.912	0.694	18.3	0.978	0.775	14.8
W&C Ens[1]	0.889	0.678	18.0	0.963	0.763	14.6
word2vec[2]	0.879	0.676	18.0	0.967	0.764	14.4
All Ens	**0.848**	**0.649**	**17.3**	**0.894**	**0.707**	**13.5**

Result on training data using 5-fold cross validation.
1: Trained using Hyperword tool (https://bitbucket.org/omerlevy/hyperwords)
2: Trained using word2vec tool (https://code.google.com/archive/p/word2vec/)

3.2 IALP 2016 Shared Task

- **Official result on valence:**

Submission	Valence MAE (rank)	Valence PCC (rank)	Mean Rank
CKIP-Run2	0.583 (4)	0.862 (3)	3.5
Aicyber-Run1	0.577 (1)	0.848 (8)	4.5
CKIP-Run1	0.601 (6)	0.854 (5)	5.5
NCTU+NTUT-Run1	0.613 (8)	0.854 (5)	6.5
Aicyber-Run2	0.581 (3)	0.843 (11)	7
SCAU-Run2	0.768 (15)	0.865 (1)	8
PolyU-Marine-Run1	**0.658 (13)**	**0.857 (4)**	**8.5**
NCTU+NTUT-Run2	0.621 (10)	0.853 (7)	8.5
ECNUCS-Run1	0.577 (1)	0.811 (17)	9
SCAU-Run1	0.774 (17)	0.864 (2)	9.5
KUAS-IsLab-Run2	0.583 (4)	0.817 (16)	10

3.2 IALP 2016 Shared Task

- **Official result on arousal:**

Submission	Arousal MAE (rank)	Arousal PCC (rank)	Mean Rank
NCTU+NTUT-Run2	1.165 (5)	0.631 (4)	4.5
Aicyber-Run1	1.212 (8)	0.671 (1)	4.5
Aicyber-Run2	1.215 (9)	0.662 (3)	6
PolyU-Marine-Run1	**1.241 (10)**	**0.664 (2)**	**6**
XMUT-Run1	1.265 (12)	0.622 (7)	9.5
YUN-ISE-HPC-Run1	1.084 (3)	0.495 (18)	10.5
HAUT-Run2	1.278 (13)	0.609 (10)	11.5
BIT1021-Run1	1.143 (4)	0.442 (21)	12.5
XMUT-Run2	1.247 (11)	0.543 (15)	13
KUAS-IsLab-Run2	0.953 (1)	0.295 (27)	14

4. Conclusion and Future Work

- **Conclusion**
  - Propose a word embedding + regression method for valence arousal ratings prediction
  - Promising results on three lexicons and IALP 2016 shared task
  - Model is simple but effective
  - Gain better results from ensemble on different embeddings

- **Future work**
  - Ensemble on more embeddings, such as multi-view embedding
  - Extend to other word meaning ratings (EPA, concreteness, perceptual strength)

Thank you!

Reference:

Ekman, Paul. "Facial Expression and Emotion." *American Psychologist* 48, no. 4 (1993): 384.

Bradley, Margaret M, and Peter J Lang. "Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings." Citeseer, 1999.

Warriner, Amy Beth, Victor Kuperman, and Marc Brysbaert. "Norms of Valence, Arousal, and Dominance for 13,915 English Lemmas." *Behavior Research Methods* 45, no. 4 (2013): 1191–1207.

Yu, Liang-Chih, Lung-Hao Lee, Shuai Hao, Jin Wang, Yunchao He, Jun Hu, K. Robert Lai, and Xuejie Zhang. "Building Chinese Affective Resources in Valence-Arousal Dimensions." In *Proceedings of NAACL-HLT*, 540–545, 2016.

Osgood, Charles Egerton, George J. Suci, and Percy H. Tannenbaum. *The Measurement of Meaning*. University of Illinois Press, 1957.

Reference:

Rastogi, Pushpendre, Benjamin Van Durme, and Raman Arora. “Multiview LSA: Representation Learning via Generalized CCA.” In *ACL*, 556–66. Association for Computational Linguistics, 2015. doi:10.3115/v1/N15-1058.

Yin, Wenpeng, and Hinrich Schütze. “Learning Word Meta-Embeddings,” 1351–60. Association for Computational Linguistics, 2016. doi:10.18653/v1/P16-1128.

Speer, Robert, and Joshua Chin. “An Ensemble Method to Produce High-Quality Word Embeddings.” *arXiv Preprint arXiv:1604.01692*, 2016.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. “Distributed Representations of Words and Phrases and Their Compositionality.” In *Advances in Neural Information Processing Systems*, 3111–19, 2013.

Pennington, Jeffrey, Richard Socher, and Christopher D. Manning. “Glove: Global Vectors for Word Representation.” In *Proceedings of EMNLP*, 1532--1543, 2014.

Reference:

Hamilton, William L., Kevin Clark, Jure Leskovec, and Dan Jurafsky. “Inducing Domain-Specific Sentiment Lexicons from Unlabeled Corpora.” In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 595–605, 2016.