元智大學 Yuan Ze University

Dimensional Sentiment Analysis — Resources, Methods, and Applications

Liang-Chih Yu
Department of Information Management
Yuan Ze University, Taiwan, R.O.C.

YUAN ZE UNIVERSITY

Outline

- Introduction
  - Categorical and Dimensional Sentiment Analysis
  - Valence-Arousal (VA) Space
- Resources
  - Lexicons and Corpora
- Methods
  - Word/Phrase-level
  - Sentence/Text-level
- Conclusions

Introduction

- Sentiment Analysis
  - Identify and extract opinion/sentiment/subjective information from texts
  - Categorical Representation (discrete class)
    - Positive or Negative
    - Six basic emotions (anger, happiness, fear, sadness, disgust and surprise) (Ekman, 1992)
  - Dimensional Representation (continuous value)
    - Valence-Arousal (VA) (Russell, 1980)
    - Pleasure-Arousal-Dominance (PAD) (Mehrabian, 1996)

Categorical Sentiment Analysis

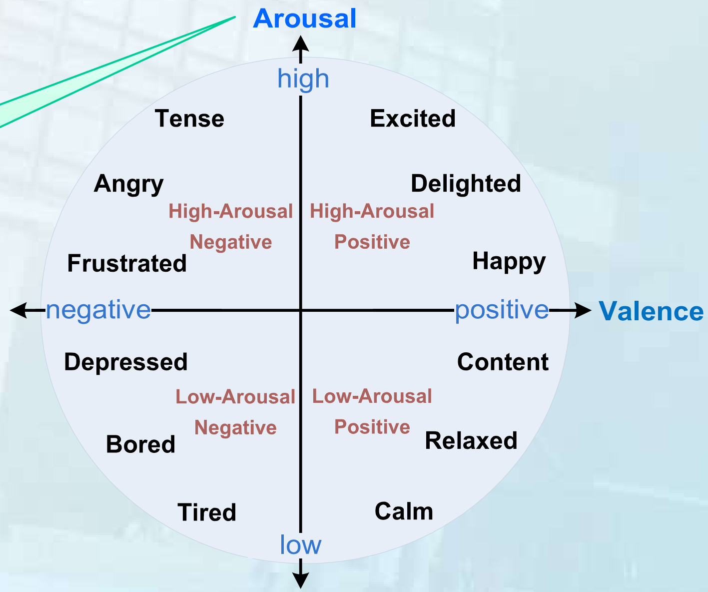
- Classify given texts into a set of predefined categories

資料來源: https://dailyview.tw/

Dimensional Sentiment Analysis

- Dimensional representation represents emotion states as continuous numerical values for multiple dimensions
  - Valence-Arousal (VA) (Russell, 1980)
  - Pleasure-Arousal-Dominance (PAD) (Mehrabian, 1996)

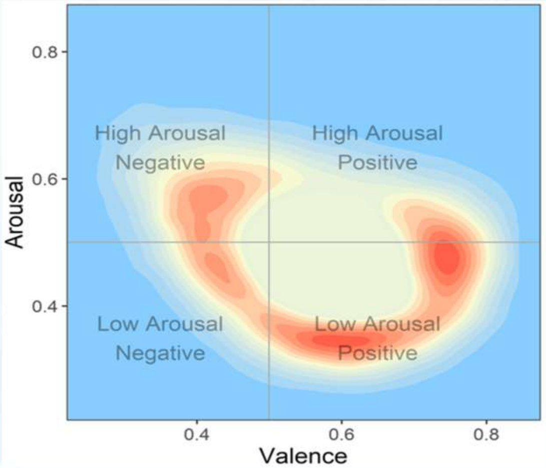
The **arousal** represents the degree of excitement and calm



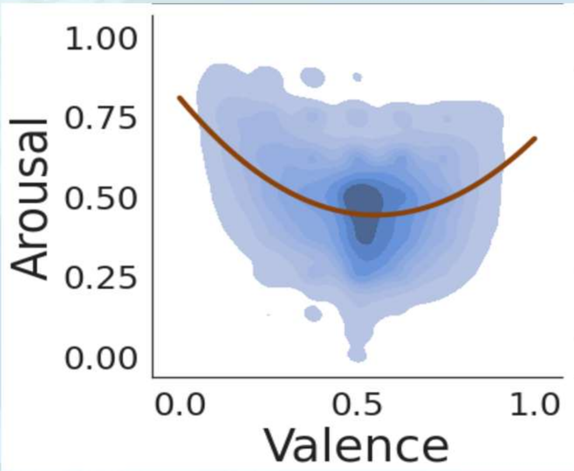
The **valence** represents the degree of pleasant and unpleasant

Categorical vs Dimensional Representation

Categorical	Dimensional
• Coarse-grain (e.g., positive, negative) • Difficult to enumerate all possible sentiment labels before analysis • Different researchers may propose different sentiment labels	• fine-grained (e.g., VA space) • Each word/sentence/document can be represented as a point in the VA space • Emotions can be compared across multiple dimensions.



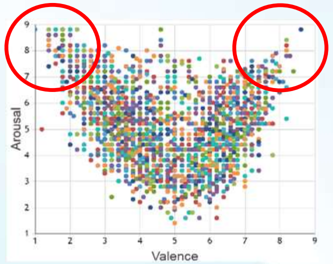
(Hipson and Mohammad, 2021)



(Mendes and Martins, 2023)

Applications of Dimensional Sentiment Analysis

- Misinformation identification
- Differentiating between mental conditions
- Emotion dynamics tracking
- Stance detection in climate change, political, and COVID-related discussions
- Personalized/ dialogue generation



Misinformation Identification

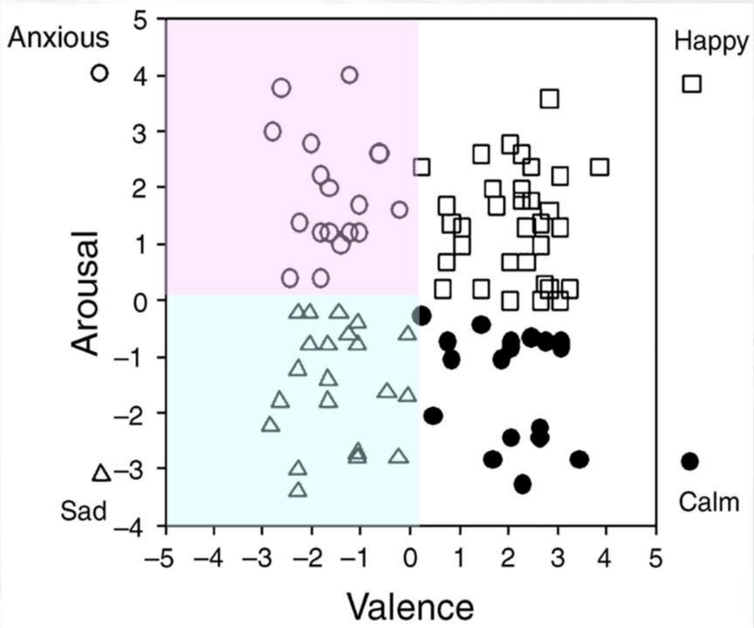
- Misinformation identification: High-arousal posts are likely to propagate and even contain misinformation (Liu et al., 2024; Yun et al., 2024)

Table 1
Relationships between emotions and misinformation. ED/SA: Emotion Detection/Sentiment Analysis method, RAM: Relationship analysis method. MANOVA: Multivariate Analysis of Variance, MANCOVA: Multivariate Analysis of Covariance, ANOVA: Analysis of Variance. ANCOVA: Analysis of covariance.

Pub	Year	Data	ED/SA	RAM	Relationship (Partly)
[88]	2019	Demonetization related	LIWC	Logistic Regression	Posts with a higher level of anger, sadness, and anxiety are indicative of rumor.
[17]	2020	COVID-19 Related	Manual	Time-lagged Cross-correlation Analyses	The angrier, sadder, or more fear the public feels, the more rumors there are likely to be.
[82]	2020	News Headlines	Question-naire,PANAS	Linear Mixed-effects Analyses	Emotion plays a causal role in people's susceptibility to incorrectly perceiving fake news as accurate.
[95]	2020	[96]	EmoLex	SVM	Emotion-based features contribute more to rumor recognition capabilities than personality-based ones.
[11]	2020	Open-Source Data	Meaningt.Ioud, TextBlob, AFINN	Chi-square Test, P(T|S), Goodman and Kruskal's Gamma	Relationships exist between negative sentiment and fake news, and between positive sentiment and genuine news.
[15]	2021	Twitter	Questionnaire	Generalized Linear Model	Rumors conveying anticipation, anger, or trust, or which are highly offensive, generate more shares, are longer-lived, and more viral.
[16]	2021	Twitter	EmoLex	Generalized Linear Model	False rumors with a high proportion of terms conveying positive sentiment, trust, anticipation, or anger are more likely to go viral.
[97]	2021	COVID-19 Related	Decision Tree	SPSS 22.0, Granger Causality Test	The more negative people feel about COVID-19, the more likely it is that rumors will be generated.
[13]	2021	News Headlines	Questionnaire	MANOVA, MANCOVA, ANOVA	Emotional reactivity of participants is associated with response behavior intentions.

Differentiating between Mental Conditions

- High-arousal anxiety vs. low-arousal depression (both involve negative emotions) (Jefferies et al., 2008; Larson et al., 2013;Teodorescu et al., 2023)



		Average Emotion		
Dataset	**MHC–Control**	V	A	D
Twitter-STMHD	ADHD–control	↓	↓	↓
	Bipolar–control	–	↓	↓
	MDD–control	↓	–	↓
	OCD–control	–	↓	↓
	PPD–control	–	↓	↓
	PTSD–control	↓	–	↓
	Depression–control	–	↓	↓
Reddit eRisk	Depression–control	–	–	↓

Emotion Dynamics Tracking

(Hipson and Mohammad, 2021)

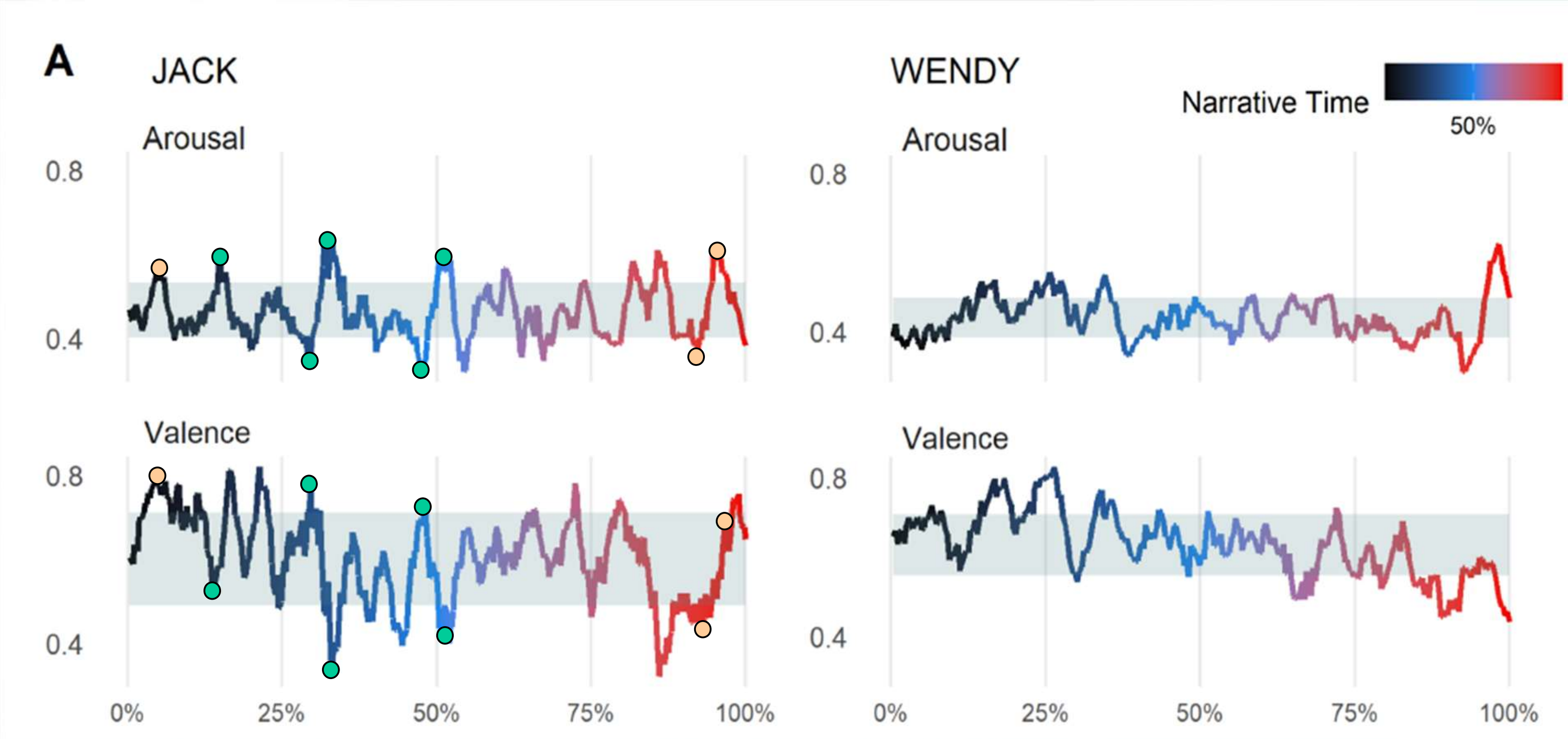
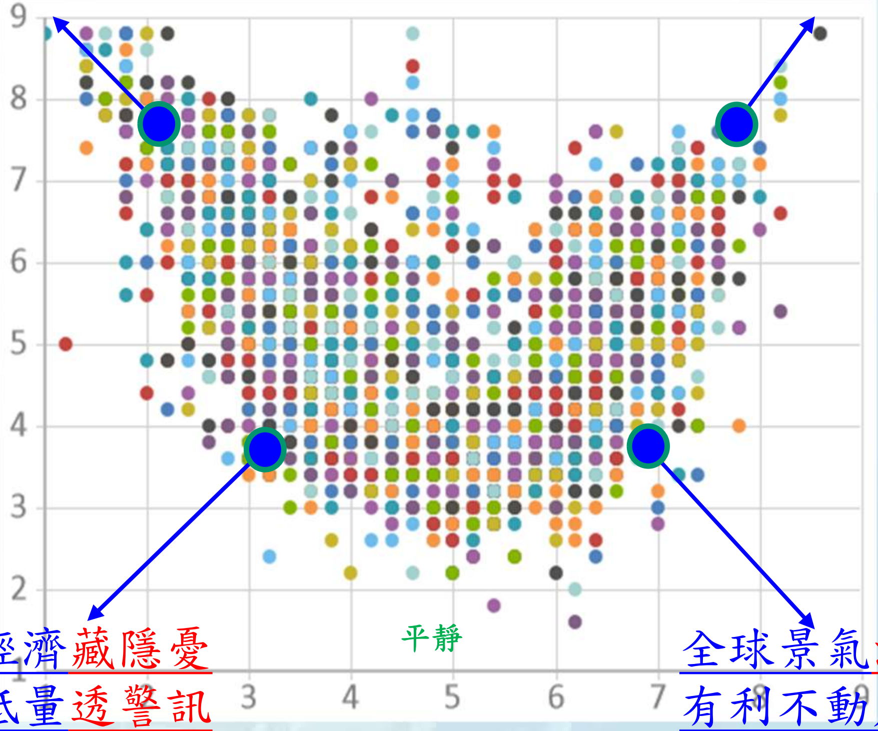


Fig 3. One dimensional and two dimensional state spaces for Jack (n = 389 words) and Wendy (n = 279 words), two main characters from *The Shining* (鬼店) (1980).

Emotion Dynamics Tracking

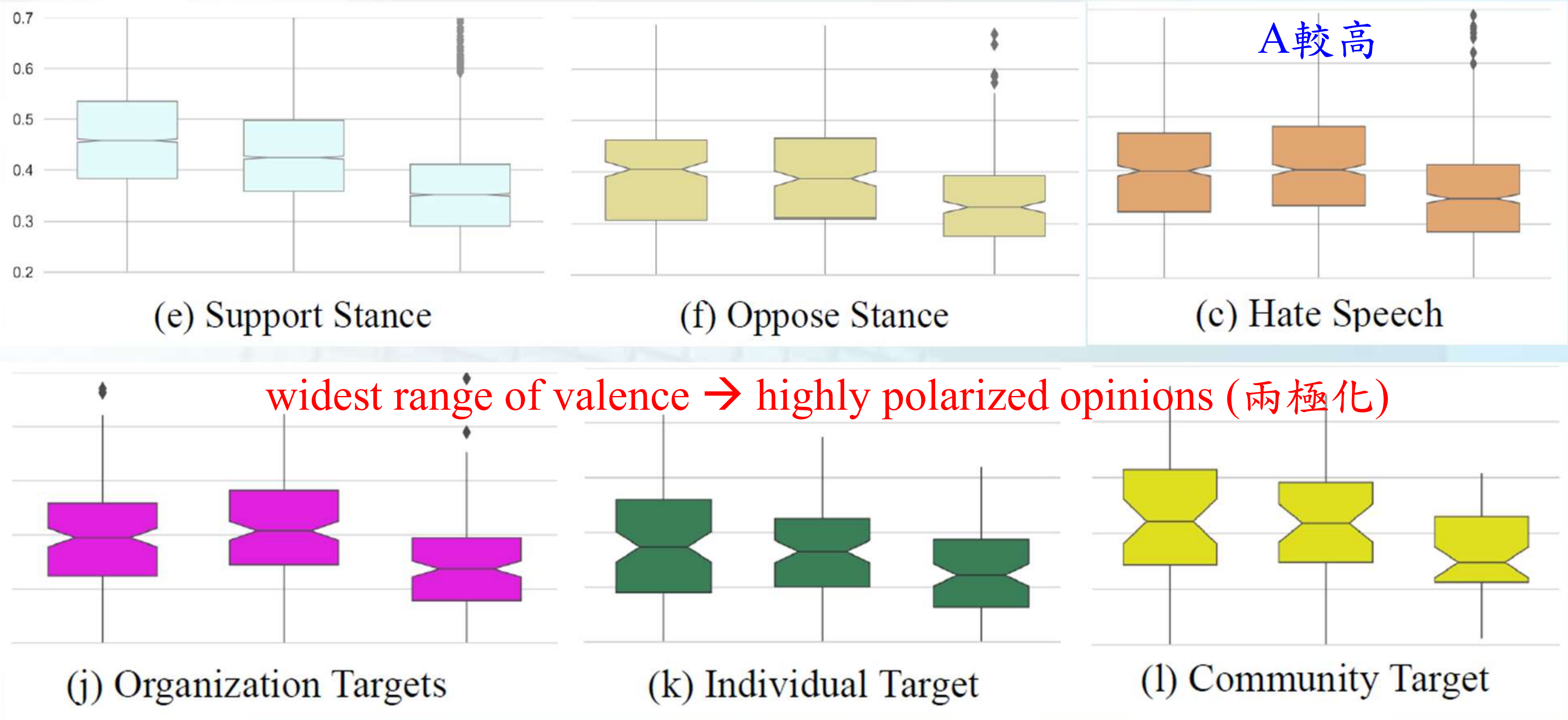


圖三、台股加權指數與多維度型投資溫度相關圖

圖四、S&P500 與多維度型投資溫度相關圖

(Peng and Yu, 2020)

Stance Detection in Climate Change



(Shiwakoti et al., 2024)

Dialogue Generation

- Generating appropriate emotional responses is crucial for dialogue systems to achieve human-like interactions
- Discrete emotions can be mapped into the VA space, enabling the generation of dialogue responses with varying degrees of valence and arousal

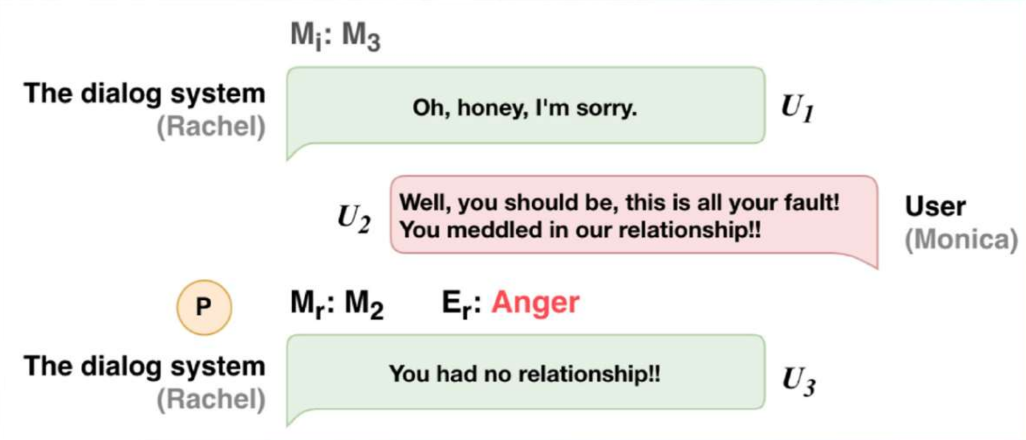


Table 3. Mood VAD Vectors Representing Different Mood States

Mood States	(Valence, Arousal, Dominance)
M_1	(1.0, 1.0, 0.0)
M_2	(−1.0, 1.0, 0.0)
M_3	(−1.0, −1.0, 0.0)
M_4	(1.0, −1.0, 0.0)
Neutral	(0.00, 0.00, 0.00)

(Wen et al., 2024)

Resources — Lexicons

Lexicon	Granularity	Size	Scale	Dimension
SentiWordNet (Baccianella et al., 2010)	Word	147,306	Continuous [0, 1]	Valence
SO-CAL (Taboada et al. 2011)	Word	5,042	Multi-point [-5, 5]	Valence
AFINN (Nielsen, 2011)	Word	2,477	Multi-point [-5, 5]	Valence
SentiStrength (Thelwall et al., 2012)	Word	2,609	Multi-point [-4, 4]	Valence
VADER (Hutto and Gilbert, 2014)	Word	7,520	Continuous [-4, 4]	Valence
NRC-EIL (Mohammad, 2018a)	Word	9,921	Continuous [0, 1]	Valence for Eight emotions
SemEval 2015 Task 10 (Rosenthal et al., 2015)	Word/Phrase	1,515 (subtask E)	Continuous [0, 1]	Valence
SemEval 2016 Task 7 (Kiritchenko et al., 2016)	Word/Phrase	3,207 (subtask 1)	Continuous [-1, 1]	Valence
ANEW (Bradley and Lang, 1999)	Word	1,034	Continuous [1,9]	Valence, Arousal, Dominance
Extended ANEW (Warriner et al., 2013)	Word	13,915	Continuous [1,9]	Valence, Arousal, Dominance
NRC-VAD (Mohammad, 2018b)	Word	20,007	Continuous [0, 1]	Valence, Arousal, Dominance

Resources — Corpora

Corpus	Granularity	Size	Scale	Dimension
Stanford Sentiment Treebank (Socher et al., 2013)	Sentence	11,855	Continuous [0, 1]	Valence
SemEval-2017 Task 5 (Cortis et al., 2017)	Tweets (subtask 1) Headlines (subtask 2)	2,510 (subtask 1) 1,647 (subtask 2)	Continuous [-1, 1]	Valence
WASSA-2017 (Mohammad and Bravo-Marquez, 2017)	Tweets	7,097	Continuous [0, 1]	Valence for four emotions
SemEval-2018 Task 1 (Mohammad et al., 2018)	Tweets	12,634 (EI-reg) 2,567 (V-reg)	Continuous [0, 1]	Valence for four emotions
ANET (Bradley and Lang, 2007)	Text	120	Continuous [1,9]	Valence, Arousal, Dominance
IEMOCAP (Busso et al., 2008)	Sentences/ Dialogues	10,039	Continuous [1,5]	Valence, Arousal, Dominance
Facebook posts (Preoţiuc-Pietro et al., 2016)	Sentence	2,895	Continuous [1,9]	Valence, Arousal
EmoBank (Buechel and Hahn, 2017)	Sentence	10,062	Continuous [1,9]	Valence, Arousal, Dominance
Chinese VAI (Xie et al., 2021)	Sentence	1,465	Continuous [1,9]	Valence, Arousal, Irony
Chinese EmoBank (Yu et al., 2016a; Lee et al, 2022)	Word/Phrase/ Sentence/Text	5,512/2,998/ 2,582/2,969	Continuous [1,9]	Valence, Arousal

Methods－Shared Tasks

- Word Level
  - SemEval 2015 Task 10 Subtask E for Determining strength of Twitter terms
  - IALP 2016 Shared Task: Dimensional Sentiment Analysis for Chinese Words
- Phrase Level
  - SemEval-2016 Task 7: Determining Sentiment Intensity of English and Arabic Phrases
  - IJCNLP 2017 Task 2: Dimensional Sentiment Analysis for Chinese Phrases
- Sentence Level
  - WASSA-2017 Shared Task on Emotion Intensity
  - SemEval-2018 Task 1: Affect in Tweets
  - ROCLING-2021 Shared Task: Dimensional Sentiment Analysis of Educational Texts
  - SIGHAN 2024 Shared Task for Chinese dimensional aspect-based sentiment analysis
  - SemEval-2025 Task 11 Track B: Emotion Intensity

SemEval-2015 Task 10: Sentiment Analysis in Twitter

(Rosenthal et al., 2015)

- Subtask E: Determining Strength Twitter Terms with Positive Sentiment
- Goal: Given a word/phrase, propose a score between 0 (lowest) and 1 (highest) that is indicative of the strength of association of that word/phrase with positive sentiment
- Top 1 INESC-ID: SVR + word embeddings (Amir et al., 2015)

Type	Sample words
words	*sweetest, giggle, sleazy, broken*
slang	*bday, lmao, kewl, pics*
negations	*can't cope, don't think, no probs*
interjections	*weee, yays, woooo, eww*
emphasized	*goooood, loveeee, cuteeee, excitedddd*
hashtags	*#gorgeous, #smelly, #fake, #classless*
multiword hashtag	*#goodvibes, #everyonelikesitbutme*
emoticons	*:o): -.- :') <33*

Team	Kendall's τ coefficient	Spearman's ρ coefficient
INESC-ID	**0.6251**	0.8172
lsislif	**0.6211**	0.8202
ECNU	**0.5907**	0.7861
CLaC-SentiPipe	**0.5836**	0.7771
KLUEless	**0.5771**	0.7662
UMDuluth-CS8761-10	**0.5733**	0.7618
IHS-RD-Belarus	**0.5143**	0.7121
sigma2320	**0.5132**	0.7086
iitpsemeval	**0.4131**	0.5859
RGUSentminers123	**0.2537**	0.3728
Baseline	**0.5842**	0.7843

IALP 2016 Shared Task on Dimensional Sentiment Analysis for Chinese Words (DSA_W) (Yu et al., 2016b)

- Motivation: Few Chinese VA lexicons exist
- Goal: Determine the VA ratings of sentiment words between 1-9

Example 1:

Input: 0001, 勝利

Output: 0001, 7.8, 7.2

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

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

SemEval-2016 Task 7: Determining Sentiment Intensity of English and Arabic Phrases

(Kiritchenko et al., 2016)

- Goal: Given a list of terms (single words and multi-word phrases), propose a score between 0 and 1 that is indicative of the term's strength of association with positive sentiment
- Top 1 ECNU: Learning to rank (Wang et al., 2016)

Dataset / **Term**	**Sentiment score**
General English Sentiment Modifiers Set	
favor	0.826
would be very easy	0.715
did not harm	0.597
increasingly difficult	0.208
severe	0.083
English Twitter Mixed Polarity Set	
best winter break	0.922
breaking free	0.586
isn't long enough	0.406
breaking	0.250
heart breaking moment	0.102

Team	**Overall**	
	Kendall's τ	Spearman's ρ
ECNU	**0.704**	0.863
UWB	0.659	0.854
LSIS	0.350	0.508

Team	**Overall**	
	Kendall's τ	Spearman's ρ
ECNU	**0.523**	0.674
LSIS	0.422	0.591
UWB	0.414	0.578

IJCNLP 2017 Task 2: Dimensional Sentiment Analysis for Chinese Phrases (DSA_P) (Yu et al., 2017)

- Motivation: Few Chinese VA lexicons exist
- Goal: Determine the VA ratings of sentiment phrases between 1-9
- Top 1 THU_NGN: Deep LSTM (Wu et al., 2017)

Example 1:
 Input: 1, 好
 Output: 1, 6.8, 5.2
Example 2:
 Input: 2, 非常好
 Output: 2, 8.500, 6.625

All-Level	V-MAE	V-MAE Rank	V- PCC	V- PCC Rank	A-MAE	A-MAE Rank	A-PCC	A-PCC Rank	Mean Rank
THU_NGN-Run2	0.427	1	0.9345	1	0.6245	1	0.7985	1	1
THU_NGN-Run1	0.4795	2	0.9085	2	0.6645	4	0.766	3	2.75
AL_I_NLP-Run2	0.5355	3	0.8965	3	0.661	3	0.766	2	2.75
AL_I_NLP-Run1	0.539	4	0.8955	4	0.659	2	0.761	4	3.5
CKIP-Run1	0.547	7	0.8895	6	0.6655	5	0.742	5	5.75

WASSA-2017 Shared Task on Emotion Intensity

(Mohammad and Bravo-Marquez, 2017)

- Goal: Given a tweet, determine the emotion intensity (between 0 to 1) of the tweet for anger, fear, joy, or sadness
- Top 1 Prayas: Ensemble of DNNs (Mohammad and Bravo-Marquez, 2017)

Tweet	Emotion	Score
How the fu*k! Who the heck! moved my fridge!... should I knock the landlord door. #angry #mad ##	anger	0.94
So my Indian Uber driver just called someone the N word. If I wasn't in a moving vehicle I'd have jumped out #disgusted	anger	0.90
@DPD_UK I asked for my parcel to be delivered to a pick up store not my address #fuming #poorcustomerservice	anger	0.90
so ef whichever butt wipe pulled the fire alarm in davis bc I was sound asleep #pissed #angry #upset #tired #sad #tired #h..	anger	0.90

Team Name	r avg. (rank)	r fear (rank)	r joy (rank)	r sadness (rank)	r anger (rank)
1. Prayas	0.747 (1)	0.732 (1)	0.762 (1)	0.732 (1)	0.765 (2)
2. IMS	0.722 (2)	0.705 (2)	0.726 (2)	0.690 (4)	0.767 (1)
3. SeerNet	0.708 (3)	0.676 (4)	0.698 (6)	0.715 (2)	0.745 (3)
4. UWaterloo	0.685 (4)	0.643 (8)	0.699 (5)	0.693 (3)	0.703 (7)
5. IITP	0.682 (5)	0.649 (7)	0.713 (4)	0.657 (7)	0.709 (5)

SemEval-2018 Task 1: Affect in Tweets

(Mohammad et al., 2018)

- Subtask EI-reg: Given a tweet, determine the emotion intensity (between 0 to 1) of the tweet for anger, fear, joy, or sadness
- Subtask V-reg: Given a tweet, determine the valence (between 0 to 1) of the tweet
- Top 1 SeerNet: Ensemble of XG Boost and Random Forest (Duppada et al., 2018)

			Pearson r (all instances)				
Test Set	Rank	Team Name	avg.	anger	fear	joy	sadness
English							
	1	SeerNet	79.9	82.7	77.9	79.2	79.8
	2	NTUA-SLP	77.6	78.2	75.8	77.1	79.2
	3	PlusEmo2Vec	76.6	81.1	72.8	77.3	75.3
	23	Median Team	65.3	65.4	67.2	64.8	63.5
	37	SVM-Unigrams	52.0	52.6	52.5	57.5	45.3
	46	Random Baseline	-0.8	-1.8	2.4	-5.8	2.0

Rank	Team Name	r (all)
English		
1	SeerNet	87.3
2	TCS Research	86.1
3	PlusEmo2Vec	86.0
18	Median Team	78.4
31	SVM-Unigrams	58.5
35	Random Baseline	3.1

ROCLING-2021 Shared Task: Dimensional Sentiment Analysis of Educational Texts

(Yu et al., 2021)

- Goal: Determine the VA ratings (between 1 and 9) of self- evaluation comments written by students
- Top 1 NTUST-NLP-2: Ensemble of BERT-based models (Lu and Chen, 2021)

Example: 今天教了許多以前沒有學過的東西，所以上起課來很新鮮

Valence: 6.8, Arousal: 5.2

Team	Valence MAE	Valence *r*	Arousal MAE	Arousal *r*
ntust-nlp-1-run1	0.684	0.912	0.906	0.607
ntust-nlp-1-run2	**0.586**	0.901	0.885	0.585
ntust-nlp-2-run1	0.654	0.905	0.880	0.581
ntust-nlp-2-run2	0.667	**0.913**	**0.866**	**0.616**

SIGHAN 2024 Shared Task for Chinese Dimensional Aspect-based Sentiment Analysis (dimABSA) (Lee et al., 2024)

- Subtask 1: Intensity Prediction
- Subtask 2: Triplet Extraction
- Subtask 3: Quadruple Extraction

Input: E0001:S001, 檸檬醬也不會太油，塔皮對我而言稍軟。, 檸檬醬#塔皮

Output: E0001:S001 (檸檬醬,5.67#5.5)(塔皮,4.83#5.0)

Input: E0002:S002, 不僅餐點美味上菜速度也是飛快耶！！

Output: E0002:S002 (餐點, 美味, 6.63#4.63) (上菜速度, 飛快, 7.25#6.00)

Input: E0003:S003, 這碗拉麵超級無敵霹靂難吃

Output: E0003:S003 (拉麵, 食物#品質, 超級無敵霹靂難吃, 2.00#7.88)

SIGHAN 2024 Shared Task for Chinese Dimensional Aspect-based Sentiment Analysis (dimABSA) (Lee et al., 2024)

- Top 1 HITSZ-HLT: BERT + LLM (Xu et al., 2024)

Subtask 1: Intensity Prediction					
Team	Evaluation Metrics				Overall Rank
	V-MAE	V-PCC	A-MAE	A-PCC	
HITSZ-HLT	**0.279** (1)	**0.933** (1)	**0.309** (1)	**0.777** (1)	**1**
CCIIPLab	0.294 (2)	0.916 (3)	**0.309** (1)	0.766 (3)	2
YNU-HPCC	0.294 (2)	0.917 (2)	0.318 (3)	0.771 (2)	2
DS-Group	0.460 (4)	0.858 (5)	0.501 (4)	0.490 (4)	4
yangnan	1.032 (5)	0.877 (4)	1.095 (5)	0.097 (5)	5

SemEval-2025 Task 11: Bridging the Gap in Text-Based Emotion (Muhammad et al., 2025)

- Track B: Emotion Intensity Detection
- Goal: Determine the emotion intensity (0, 1, 2, 3) of a text for joy, sadness, fear, anger, surprise, and disgust
- Top 1 PAI: Ensemble of LLMs (ChatGPT-4o, OpenAI, DeepSeek-V3, DeepSeek-AI, Gemma-9b, Qwen-2.5-32b, Mistral-Small-24B

eng	pai	**0.840**
	nycu-nlp	**0.837**
	R_{baseline}	0.641
	M_{baseline}	0.001

Summary of Word/Phrase-Level Methods

2015, 2016

- Neural Network (2)
  - ✓ NN, Boosted NN
- Regression (7)
  - ✓ SVR (3), Linear (2), Kernel, Gaussian, Ensemble
- *k*-Nearest Neighbor (*k*NN) (6)
- Others (4)
  - ✓ Ranking, PMI, CRF, Rule (2)

2017

- Neural Network (6)
  - ✓ NN, Boosted NN, Ensembles, BiLSTM, Deep LSTM, CNN
- Regression (3)
  - ✓ SVR (2), Linear
- Others (2)
  - ✓ Rule-based, *k*NN

English		
Task	Spearman	Method
SemEval-2015	0.817	**SVR**
SemEval-2016	0.863	**Ranking**

Chinese			
Task	V-PCC	A-PCC	Method
IALP-2016	0.865	0.671	**Boost NN**
IJCNLP-2017	0.935	0.799	**Deep LSTM**
Deng et al., 2022	0.948	0.865	**MacBERT**

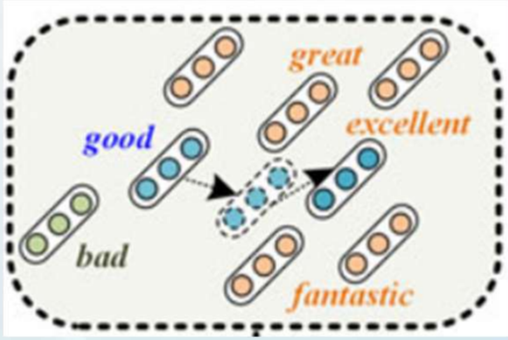
Sentiment Embeddings

	GloVe
satisfied	satisfaction, satisfy, satisfactory, **dissatisfied**, reasonable, **unsatisfied**, pleased, **disappointed**, satisfying, confident
wealthy	millionaire, rich, wealth, aristocratic, billionaire, prosperous, **impoverished**, **greedy**, privileged, businessman
strong	strength, **weak**, good, robust, solid, **tough**, consistent, powerful, confident, tremendous

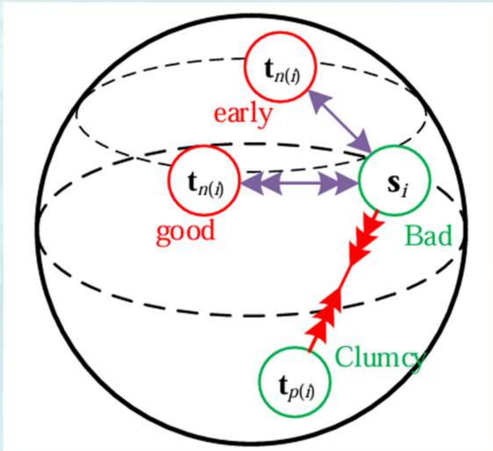
NOISE@10 FOR DIFFERENT WORD EMBEDDINGS

Method		noise@10 (%)
Conventional Embeddings	word2vec	24.3
	GloVe	24.0
Sentiment Embeddings	HyRank	18.5
Refined Embeddings	Re(word2vec)	14.4
	Re(GloVe)	13.8
	Re(HyRank)	17.2

(Yu et al., 2018)



refinement



Contrastive Learning

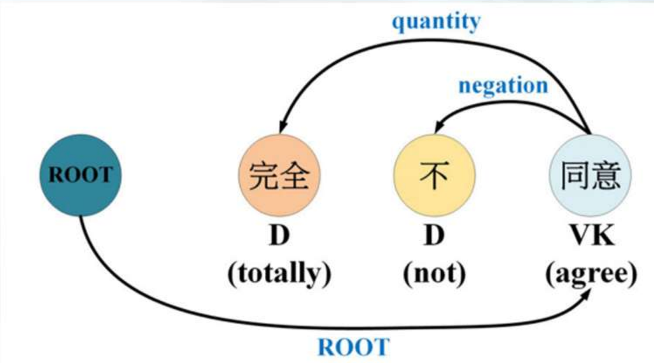
(Wang et al., 2024)

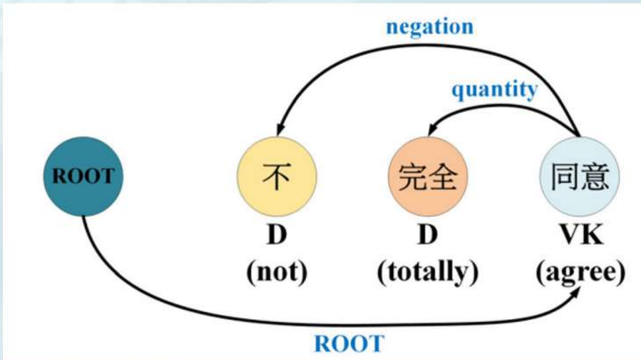
Phrase-Level Sentiment Intensity Prediction

List of Modifiers and Their Training Samples

Category		Modifier	Training examples $(mod, Int(w), Int(mod\ w))$	
Negator		cannot, could not, did not, does not, had no, have no, may not, never, no, not, nothing, was no, was not, will not, would not	accept/never accept: difficult/no difficult:	(never, 0.604, 0.292) (no, 0.354, 0.458)
Intensifier	Amplifier	certainly, especially, extremely, fairly, highly, increasingly, more, most, much, much more, particularly, pretty, quite, rather, really, so, too, very	good/extremely good: trouble/much trouble:	(extremely, 0.814, 0.924) (much, 0.252, 0.167)
	Downtoner	less, probably, relatively	free/less free:	(less, 0.869, 0.368)
Modal		can, could, may, might, must, should, would	interest/should interest: doubt/would doubt:	(should, 0.681, 0.597) (would, 0.392, 0.403)

(Yu et al., 2020)





(Deng et al., 2022)

Summary of Sentence/Text-Level Methods

WASSA-2017 and SemEval-2018

- Neural Network (12)
  - ✓ Boosted NN, CNN, RNN, LSTM, BiLSTM, CNN-LSTM, BiLSTM-CNN, Attention-based, LSTM
- Regression (20)
  - ✓ SVR, Boosting, Linear Regression, Logistic Regression, Random Forest, Ensemble

Emotion Intensity (*r*)						
Task	Avg	Fear	joy	Sadness	anger	Method
WASSA-2017	0.747	0.732	0.762	0.732	0.765	**Ensemble of DNNs**
SemEval-2018	0.799	0.779	0.792	0.798	0.827	**Ensemble of Regressors**

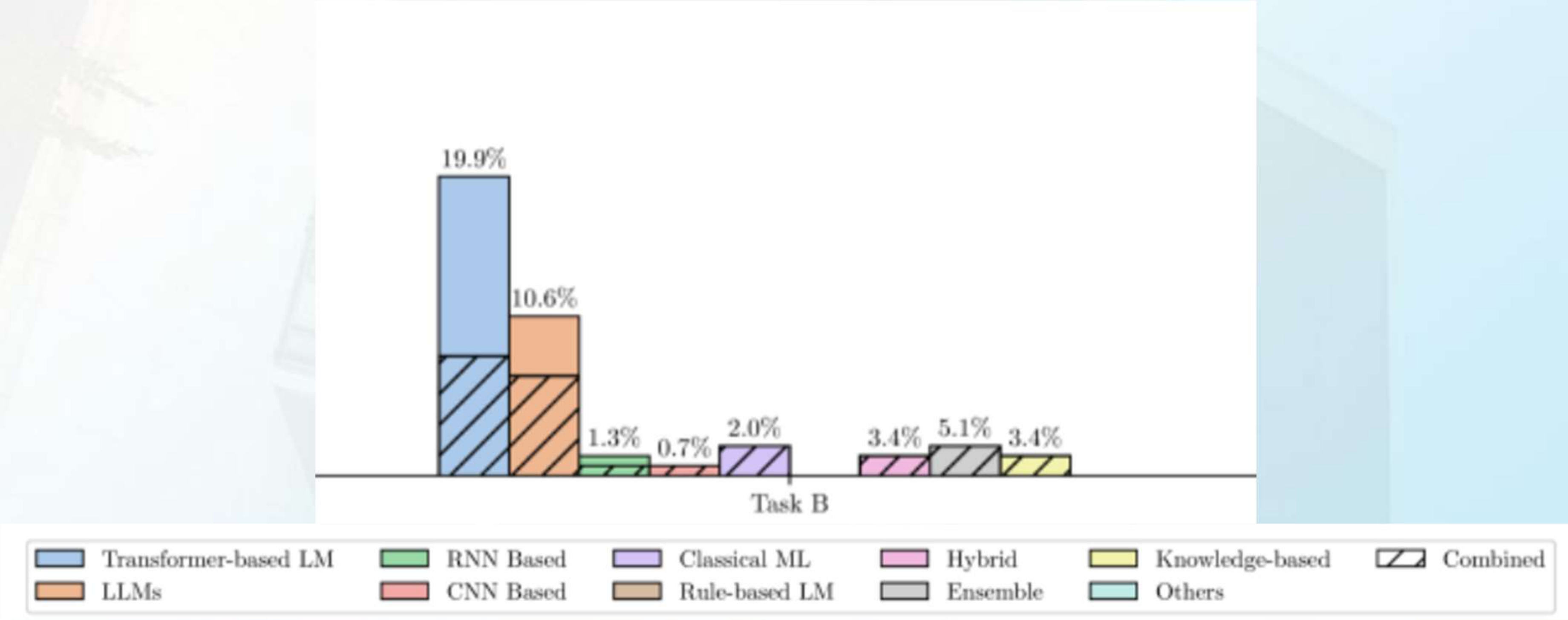
Summary of Sentence/Text-Level Methods

SIGHAN-2024 and SemEval-2025

- Fine-tuned BERT-based transformers
- Instruction-fine-tune using LoRA in combination with prompt design and on LLMs

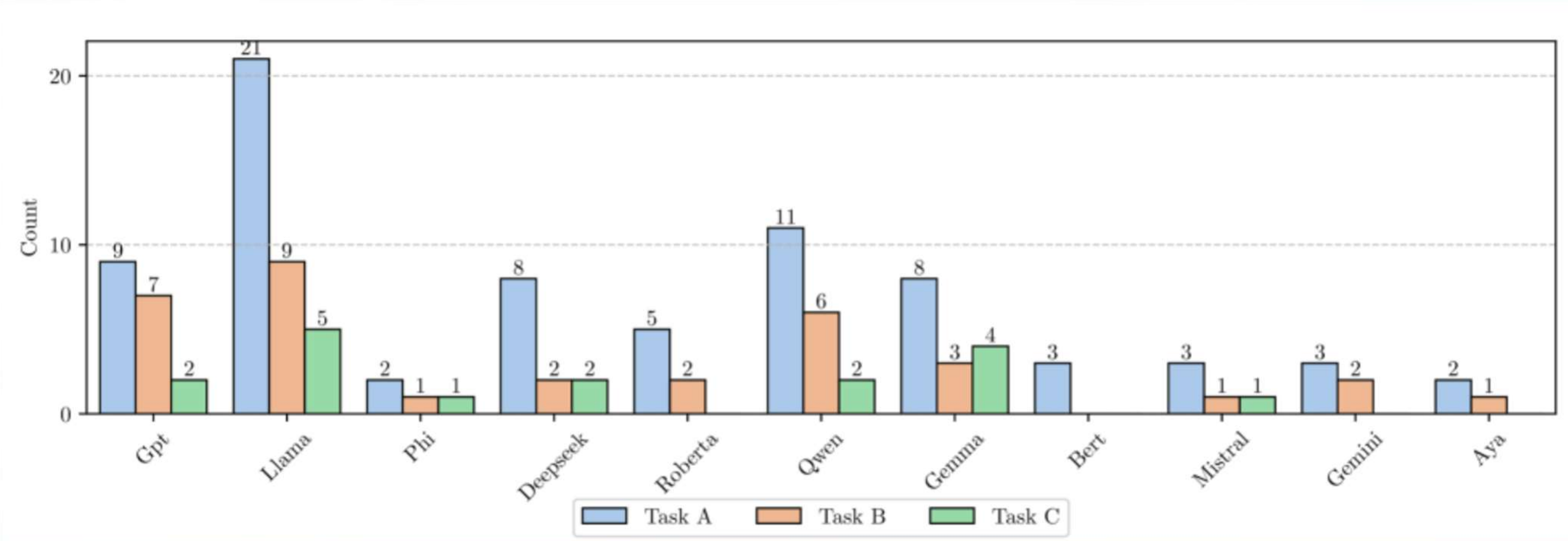
Valence (*r*)		
Task	Valence (*r*)	Method
SemEval-2018	0.873	**Ensemble of DNNs**
SIGHAN-2024	0.933	**BERT+LLM**
SemEval-2025	0.840	**Ensemble of LLMs**

SemEval-2025 Methods



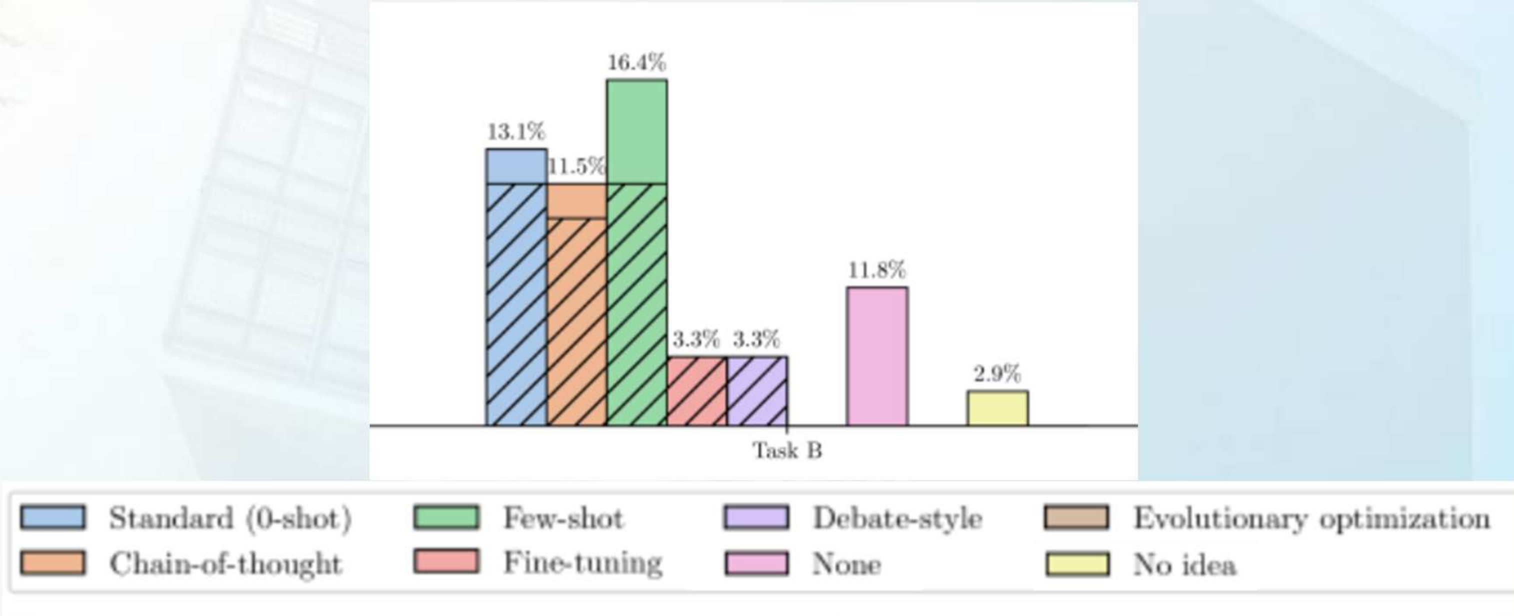
(Muhammad et al., 2025)

LLMs



(Muhammad et al., 2025)

Prompting Strategies



(Muhammad et al., 2025)

Conclusions

- Sentiment representation methods can be classified as
  - Categorical (e.g., positive and negative)
  - Dimensional (e.g., valence and arousal)
- We provide a survey of
  - Potential applications
  - Dimensional lexicons and corpora
  - Methods (traditional ML, neural networks, Transformers, LLMs)

Reference

- S. Amir, R. F. Astudillo, W. Ling, B. Martins, M. Silva, and I. Trancoso, "INESC-ID: A regression model for large scale Twitter sentiment lexicon induction," in *Proc. of SemEval-15*, pp. 613-618, 2015.
- S. Baccianella, A. Esuli, and F. Sebastiani, "SentiWordNet 3.0: an enhanced lexical resource for sentiment Analysis and Opinion Mining," in *Proc. of LREC-10*, pp. 2200-2204, 2010.
- V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, no. 10, pp. 10008, 2008.
- M. M. Bradley and P. J. Lang., 1999. Affective norms for English words (ANEW): Instruction manual and affective ratings. Technical Report C-1, University of Florida.
- M. M. Bradley and P. J. Lang. 2007. Affective norms for English text (ANET): Affective ratings of text and instruction manual. Technical Report. D-1, University of Florida.
- S. Buechel and U. Hahn. 2017. EmoBank: Studying the impact of annotation perspective and representation format on dimensional emotion analysis. In *Proc. of EACL-17*, pages 578–585.
- C. Busso, M. Bulut, C.C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J.N. Chang, S. Lee, and S.S. Narayanan, "IEMOCAP: Interactive emotional dyadic motion capture database," *Journal of Language Resources and Evaluation*, vol. 42, no. 4, pp. 335-359, December 2008.

- K. Cortis, A. Freitas, T. Daudert, M. Huerlimann, M. Zarrouk, S. Handschuh, and B. Davis, "SemEval-2017 Task 5: Fine-grained sentiment analysis on financial microblogs and news," in *Proc. of SemEval-17*, pp. 519-535, 2017.
- Y.-C. Deng, C.-Y. Tsai, Y.-R. Wang, S.-H. Chen and L. H. Lee. 2022. Predicting Chinese Phrase-level Sentiment Intensity in Valence-Arousal Dimensions with Linguistic Dependency Features. *IEEE Access*, 10, 126612-126620.
- V. Duppada, R. Jain, S. Hiray. 2018. SeerNet at SemEval-2018 Task 1: Domain Adaptation for Affect in Tweets. in *Proc. of SemEval-18*, pp. 18-23, 2018.
- P. Ekman, "An argument for basic emotions," *Cognition and emotion*, vol. 6, no. 3-4, pp. 169-200, 1992.
- A. Esuli, and F. Sebastiani F, "Pageranking wordnet synsets: An application to opinion mining," in *Proc. ACL*, 2007, pp. 442-431.
- P. Goel, D. Kulshreshtha, P. Jain, and K. K. Shukla, "Prayas at EmoInt 2017: An ensemble of deep neural architectures for emotion intensity prediction in tweets," in *Proc. WASSA*, 2017, pp. 58-65.
- W. E. Hipson and S. M. Mohammad. 2021. Emotion dynamics in movie dialogues. *PLOS ONE*, 16(9):e0256153.
- C. J. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," in *Proc. of 8th International AAAI Conference on Weblogs and Social Media*, pp. 216-225, 2014.

- L. N. Jefferies, D. Smilek, E. Eich, J. T. Enns. 2008. Emotional valence and arousal interact in attentional control. *Psychological Science*, 19(3):290–295.
- S. Kiritchenko and S. M. Mohammad, and M. Salameh. 2016. SemEval-2016 Task 7: Determining sentiment intensity of English and Arabic phrases. In *Proc. of SemEval-16*, pages 42-51.
- M. J. Larson, A. C. Gray, P. E. Clayson, R. Jones and C. B. Kirwan. 2013. What are the influences of orthogonally-manipulated valence and arousal on performance monitoring processes? The effects of affective state. *International Journal of Psychophysiology*, 87(3):327-339.
- Z. Liu, T. Zhang, K. Yang, P. Thompson, Z. Yu, S. Ananiadou. 2024. Emotion detection for misinformation: A review. *Information Fusion*, 107:102300.
- L. H. Lee, J. H. Li, and L. C. Yu. 2022. Chinese EmoBank: Building Valence-Arousal Resources for Dimensional Sentiment Analysis. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 21(4): article 65.
- L. H. Lee, L. C. Yu, S. Wang and J. Liao, "Overview of the SIGHAN 2024 Shared Task for Chinese Dimensional Aspect-based Sentiment Analysis," in *Proc. of SIGHAN-10*, pp. 165-174, 2024.
- K. H Lu and K. Y. Chen. 2021. ntust-nlp-2 at ROCLING-2021 Shared Task: BERT-based semantic analyzer with word-level information. In *Proc. of ROCLING-21*, pp. 360-366, 2021.
- A. Mehrabian, Pleasure-Arousal-Dominance: A General Framework for Describing and Measuring Individual, *Current Psychology*, vol. 15, no. 4, pp. 505-525, 1996.

- G. A. Mendes and B. Martins. 2023. Quantifying Valence and Arousal in Text with Multilingual Pre-trained Transformers. In *Proc. of ECIR-23*, pages 84-100.
- S. M. Mohammad. 2018a. Word Affect Intensities. In *Proc. of LREC-18*, pages 174-183.
- S. M. Mohammad. 2018b. Obtaining Reliable Human Ratings of Valence, Arousal, and Dominance for 20,000 English Words. In *Proc. of ACL-18*, pages 174–184.
- S. M. Mohammad, F. Bravo-Marquez, M. Salameh, and S. Kiritchenko. 2018. Semeval-2018 Task 1: Affect in Tweets. In *Proc. of SemEval-18*, pages 1-17.
- S. H. Muhammad et al. 2025. SemEval-2025 Task 11: Bridging the Gap in Text-Based Emotion Detection. In Proc. of SemEval-25.
- M. E. Newman, "Modularity and community structure in networks," in *Proc. National Academy of Sciences*, vol. 103, no. 23, pp. 8577-8582, 2006.
- F. Å. Nielsen, "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs," arXiv preprint arXiv: 1103.2903, 2011.
- C. H. Peng and L. C. Yu. Sentiment Analysis for Investment Atmosphere Scoring. In *Proc. of ROCLING-20*, pp. 275-289.
- D. Preoţiuc-Pietro, A. H. Schwartz, G. Park, J. Eichstaedt, M. Kern, L. Ungar, and E. Shulman, "Modelling valence and arousal in Facebook posts," in *Proc. of WASSA-16*, pp. 9-15, 2016.
- S. Rosenthal, P. Nakov, S. Kiritchenko, S. M. Mohammad, A. Ritter, and V. Stoyanov, "Semeval-2015 task 10: Sentiment analysis in twitter," in *Proc. of SemEval-15*, pp. 451-463, 2015.

- J. A. Russell, “A circumplex model of affect,” *Journal of personality and social psychology*, vol. 39, no. 6, pp. 1161, 1980.
- S. Shiwakoti, S. Thapa, K. Rauniyar, A. Shah, A. Bhandari and U. Naseem. 2024. Analyzing the Dynamics of Climate Change Discourse on Twitter: A New Annotated Corpus and Multi-Aspect Classification. In *Proc. of LREC/COLING-24*, pages 984–994.
- R. Socher, A. Perelygin, J. Y. Wu, J. Chuang, , C. D. Manning, A. Ng, and C. Potts, “Recursive deep models for semantic compositionality over a sentiment treebank,” in *Proc. of EMNLP-13*, pp. 1631-1642, 2013.
- M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, “Lexicon-based methods for sentiment analysis.” *Computational linguistics*, vol. 37, no. 2, pp. 267-307, 2011.
- M. Thelwall, K. Buckley, and G. Paltoglou. 2012. Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology*, 63(1):163–173.
- J. Wang, L. C. Yu, K. R. Lai and X. Zhang, "Community-based Weighted Graph Model for Valence-Arousal Prediction of Affective Words," *IEEE/ACM Trans. Audio, Speech and Language Processing*, vol. 24, no. 11, pp. 1957-1968, 2016b.
- A. B. Warriner, V. Kuperman, and M. Brysbaert, “Norms of valence, arousal, and dominance for 13,915 English lemmas,” *Behavior research methods*, vol. 45, no. 4, pp. 1191-1207, 2013.
- F. Wang, Z. Zhang and M. Lan. 2016. ECNU at SemEval-2016 Task 7: An Enhanced Supervised Learning Method for Lexicon Sentiment Intensity Ranking. In Proc. of SemEval-16, 491-496.

- Z. Wen, J. Cao, J. Shen, R. Yang, S. Liu and M. Sun. 2024. Personality-affected Emotion Generation in Dialog Systems. *ACM Trans. Information Systems*, 42(5):134.
- D. Teodorescu, T. Cheng, A. Fyshe and S. M. Mohammad. 2023. Language and Mental Health: Measures of Emotion Dynamics from Text as Linguistic Biosocial Markers. In *Proc. of EMNLP-23*, pages 3117-3133.
- J. Wang, L. C. Yu, X. Zhang. 2024. SoftMCL: Soft Momentum Contrastive Learning for Fine-grained Sentiment-aware Pre-training. In *Proc. of LREC/COLING-24*, pages 15012-15023.
- C. Wu, F. Wu, Y. Huang, S. Wu, and Z. Yuan. 2017. THU_NGN at IJCNLP-2017 Task 2: Dimensional sentiment analysis for Chinese phrases with deep LSTM. In *Proc. of IJCNLP-17*, Shared Tasks, pages 47-52.
- H. Xie, W. Lin, S. Lin, J. Wang, L.-C. Yu, “A Multi-dimensional Relation Model for Dimensional Sentiment Analysis,” *Information Sciences*, vol. 579, pp. 832-844, 2021.
- H. Xu, D. Zhang, Y. Zhang, R. Xu. 2024. HITSZ-HLT at SIGHAN-2024 dimABSA Task: Integrating BERT and LLM for Chinese Dimensional Aspect-Based Sentiment Analysis. In *Proc. of SIGHAN-24*, pages 175-185.
- L. C. Yu et al., “Building Chinese Affective Resources in Valence-Arousal Dimensions,” in *Proc. NAACL-HLT,* 2016a, pp. 540-545.
- L. C. Yu, L. H. Lee and K. F. Wong, "Overview of the IALP 2016 Shared Task on Dimensional Sentiment Analysis for Chinese Words," in *Proc. IALP*, 2016b, pp. 156-160.

- L. C. Yu, L. H. Lee, J. Wang and K. F. Wong, "IJCNLP-2017 Task 2: Dimensional Sentiment Analysis for Chinese Phrases," in *Proc. IJCNLP*, 2017, pp. 9-16.
- L. C. Yu, J. Wang, K. R. Lai and X. Zhang, "Refining Word Embeddings Using Intensity Scores for Sentiment Analysis," *IEEE/ACM Trans. Audio, Speech and Language Processing*, vol. 26, no. 3, pp. 671-681, 2018.
- L. C. Yu, J. Wang, K.R. Lai, and X. Zhang. 2020. Pipelined Neural Networks for Phrase-level Sentiment Intensity Prediction. *IEEE Transactions on Affective Computing*, 11(3):447-458.
- L. C. Yu, J. Wang, B. Peng and C. R. Huang, “ROCLING-2021 Shared Task: Dimensional Sentiment Analysis for Educational Texts,” in *Proc. of ROCLING-21*, pp. 385-388, 2021.
- S. Yun, S. Jun and J. Woo Kim. 2024. The impact of high arousal and displeasure on online review helpfulness: Exploring the moderating role of product involvement. *Electronic Commerce Research and Applications*, 67:101436.