

# AI Technologies Empower Professional Services Insurance in Ping An Life

Dr. Haiqin Yang

AI Department, Ping An Life

# Who Are We?

- **Ping An Life Insurance Co.** sells life insurance products
  - long period of protection
  - comprehensive scope of insurance coverage
  - ...

Fin. + Tech.

Agents



Customers



# Outline

- Introduction
- Two pieces of representative work
  - emotion recognition in dialogues
  - entity alias discovery from knowledge graphs
- Conclusion

# Our AI Technologies

## Computational Intelligence



Deep Learning,  
Recommendation

## Perceptual Intelligence



Computer Vision

## Cognitive Intelligence



NLP, Knowledge Graphs,  
Chatbots

# Core Applications

- **To Agents**
  - Visiting assistant
  - Training assistant
  - AskBob for agents
  - ...
- **To Customers**
  - AskBob for financial customer services
  - Video follow-up chatbot
  - ...



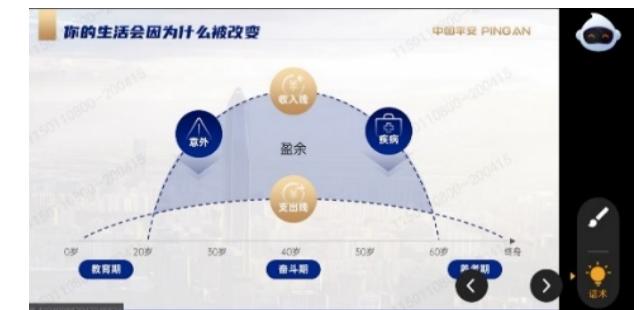
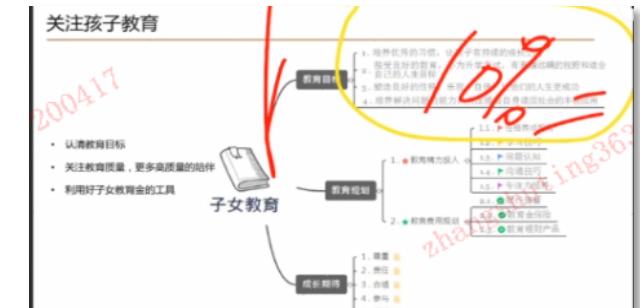
# Professional Insurance Services: Visiting Assistant

- The **first Online Visiting Assistant**

- AI guest room

- Features

- One-click explanation
- Full accompany
- Content creation



Dialog Q&A

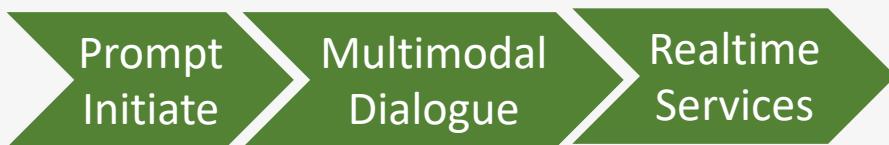
Real-time Prediction

Real-time Conversation Assistance

Content Generation



# Professional Insurance Services: Video Follow-up Chatbot

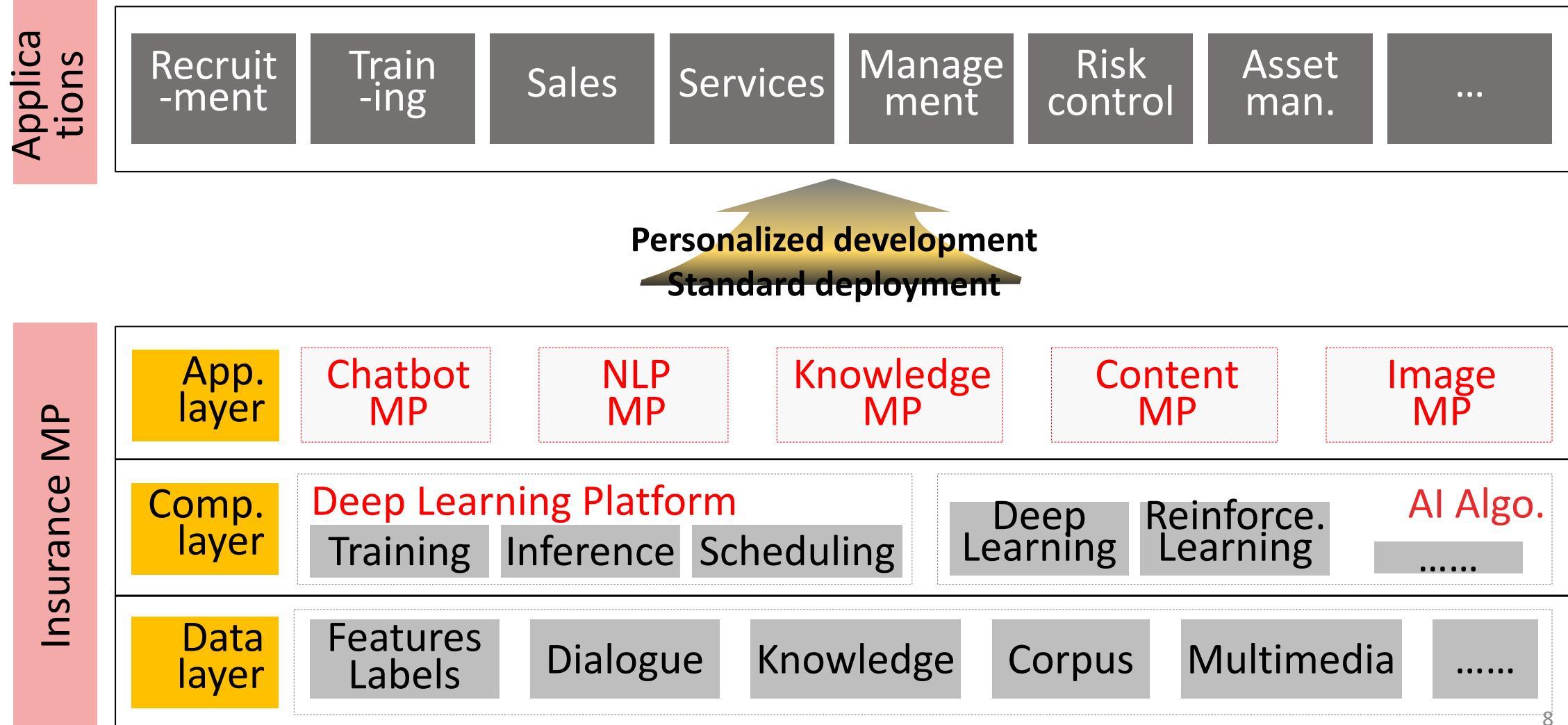


- **Impersonate video interaction**
  - User-friendly
- **Flexible follow-up**
  - 7\*24 online service
- **Efficient access to insurance policy**
  - Real-time Q&A

- The **first industry multi-modal** follow-up chatbot
  - Face-to-face interaction
  - Time reduces from 5 days to 2.8 minutes

Image Generation	Face Recognition
Lips Matching	Active Dialogue Guidance

# Insurance Middle Platform (MP)



# Chatbots

## Text



## Speech



## Video



**Champions**



7

DSTC8, SemEval2020

**Papers**



10 +

IJCAI, CIKM, ...

**Patents**



300 +

System, Technology, ...

**20+**  
**Chatbots**

**10+**  
**Platforms**

**100+**  
**Online Services**

A unified multi-modal chatbot platform



# Representative Work I

- HiGRU: Hierarchical Gated Recurrent Units for Utterance-level Emotion Recognition
- Joint work with CUHK



# Motivation: Emotion Recognition in Dialogue Systems

- The **same** word delivers **different** emotions
- Some emotions **rarely** appear
- **Long-range contextual** information is hard to captured

Role	Utterance	Emotion
Rachel	Oh okay, I'll fix that to. What's her e-mail address?	Neutral
Ross	Rachel!	Anger
Rachel	All right, I promise. I'll fix this. I swear. I'll-I'll- I'll-I'll talk to her.	Non-neutral
Ross	<i>Okay!</i>	<i>Anger</i>
Rachel	<i>Okay.</i>	<i>Neutral</i>
Nurse	This room's available.	Neutral
Rachel	<i>Okay!</i>	<i>Joy</i>
Rachel	Okay wait!	Non-neutral
Rachel	You listen to me!	Anger

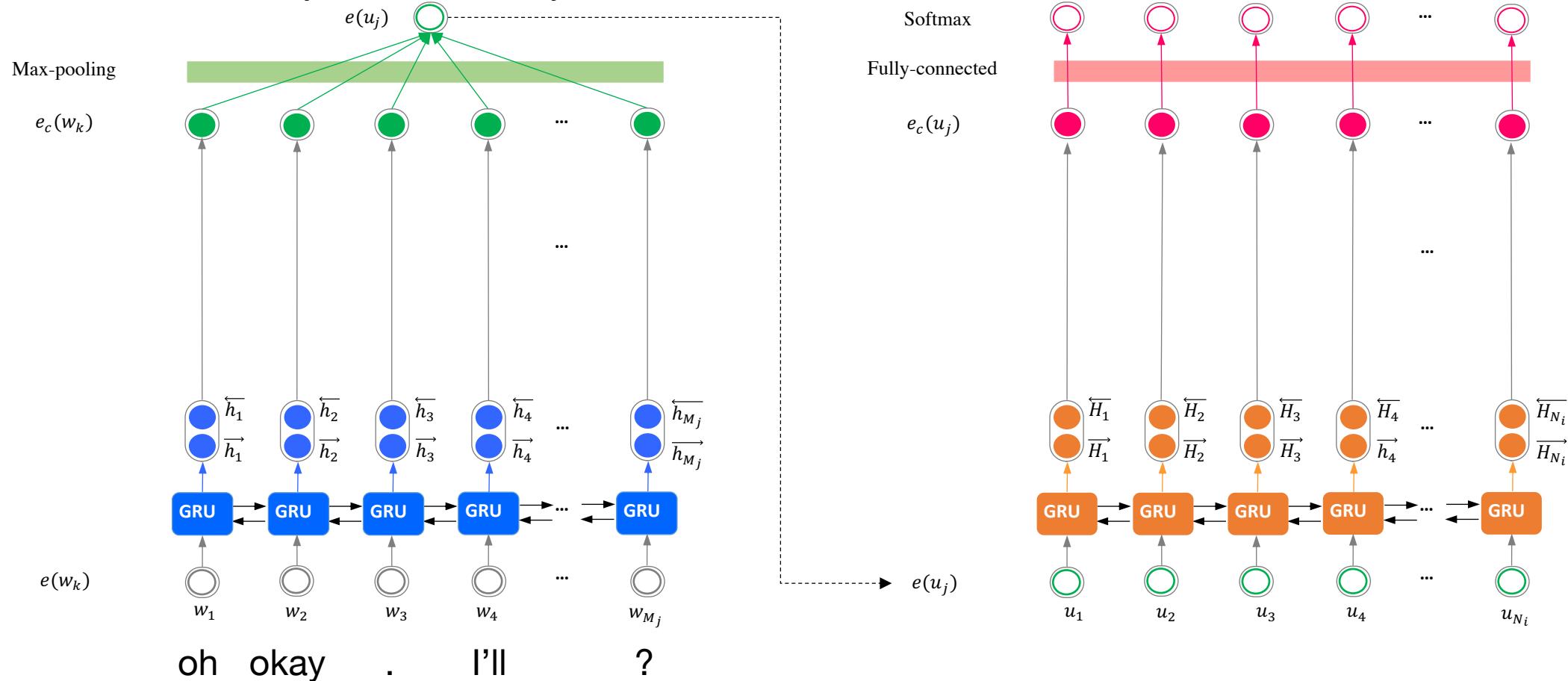
# Contribution

- A HiGRU framework to better learn both the **individual** utterance embeddings and the **contextual** information of utterances
- Two progressive variants:
  - a) HiGRU-f with residual connection to sufficiently incorporate the individual word/utterance-level information; and
  - b) HiGRU-sf with self-attention to capture the long-range contextual information
- Extensive experiments are conducted to demonstrate the superior performance of our proposal than SOTA methods

# Problem Definition: Utterance-level Emotion Recognition

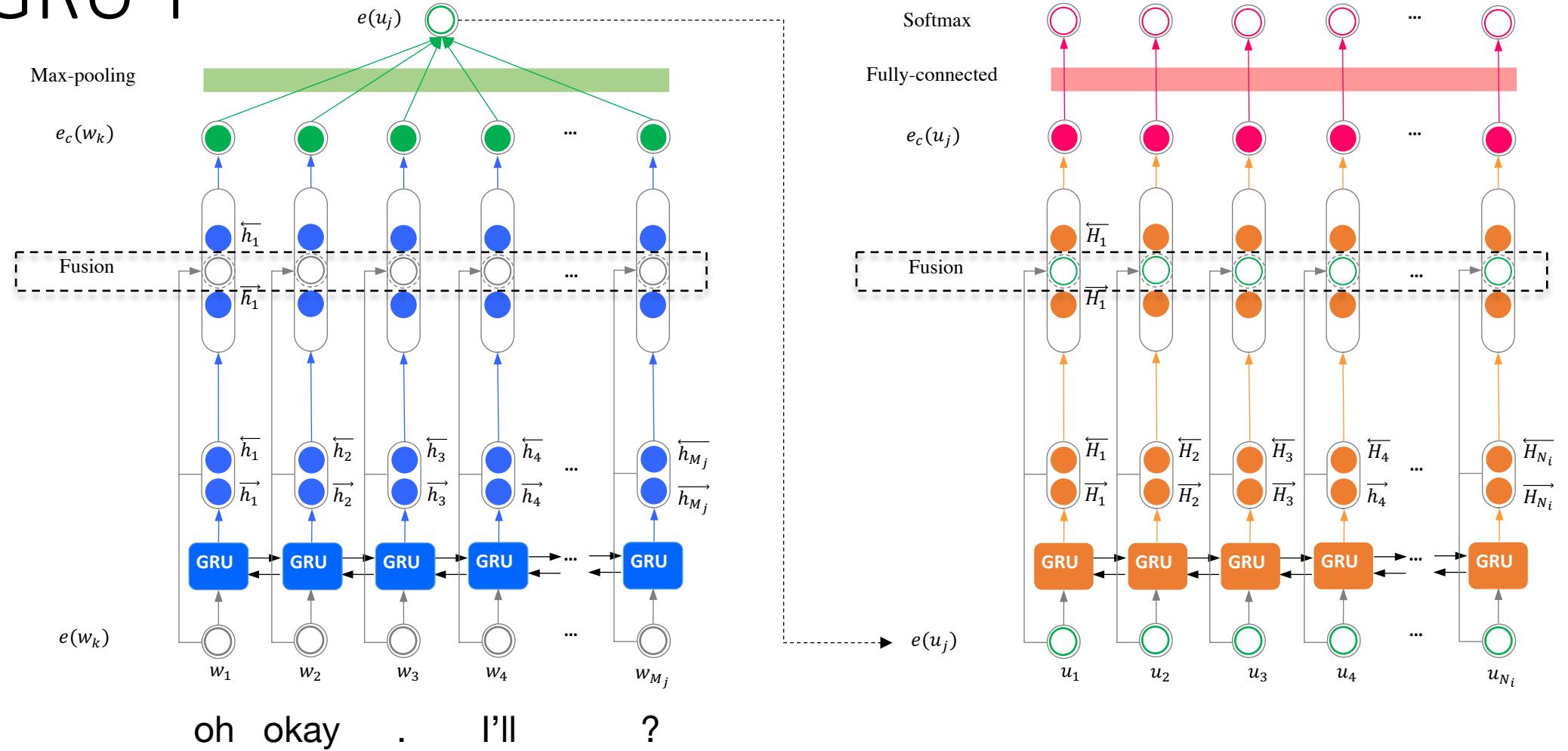
- **Input:** a set of dialogues,  $\mathcal{D} = \{D_i\}_{i=1}^L$ 
  - $L$ : the number of dialogues
  - $D_i = \{(u_j, s_j, c_j)\}_{j=1}^{N_i}$ : a dialogue
    - $u_j$ : utterance
    - $s_j$ : speaker
    - $c_j$ : emotion
- **Goal:** to train a model to tag each new utterance with an emotion label as accurately as possible

# Our Proposal: Hierarchical Gated Recurrent Units (HiGRU)



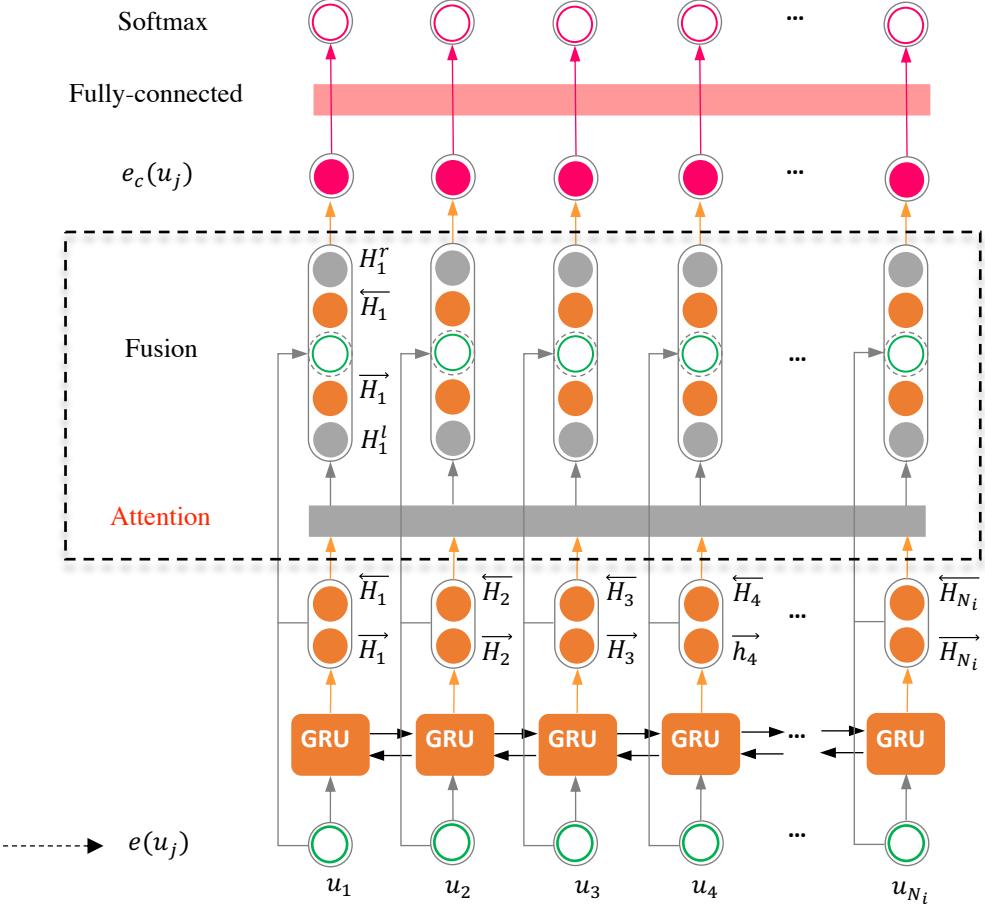
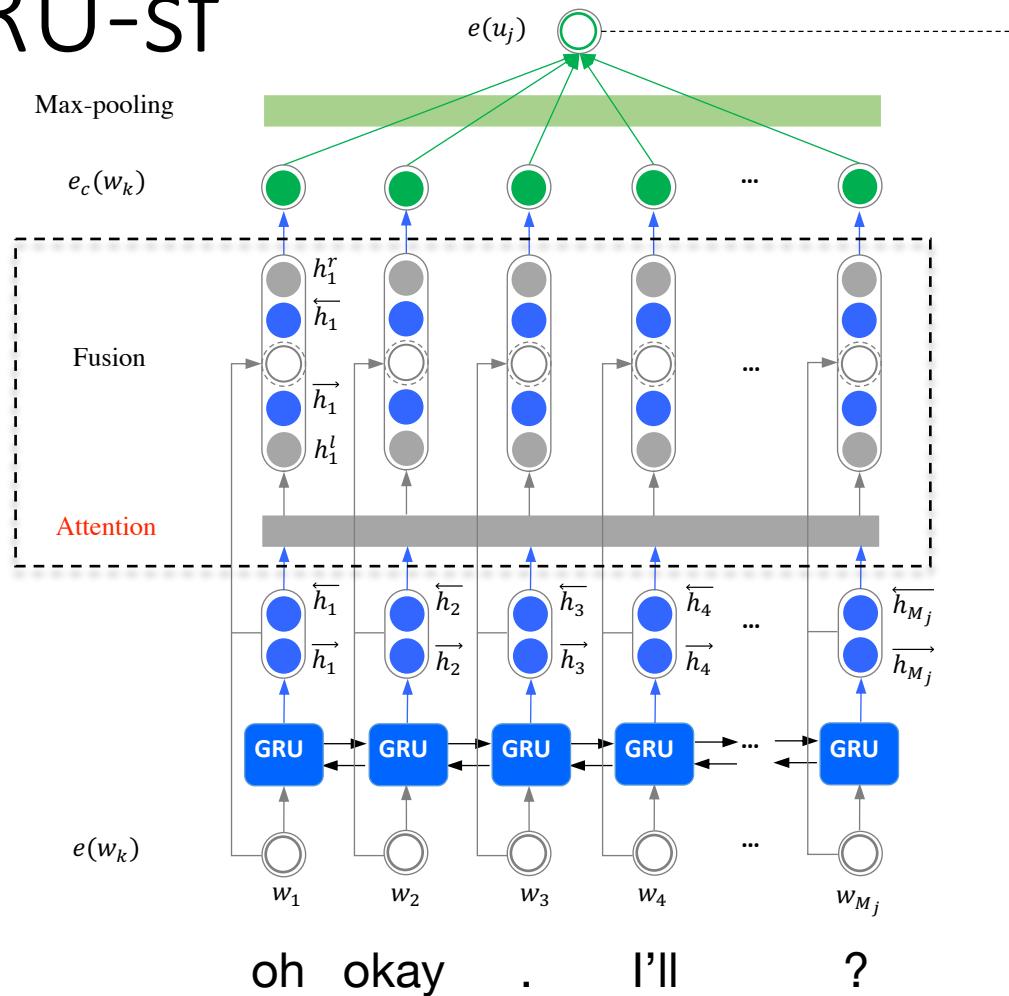
- **Lower-level bi-GRU:** individual utterance embedding
- **Upper-level bi-GRU:** contextual utterance embedding

# HiGRU-f



- **Fuse individual word/utterance embeddings to strengthen individual information**

# HiGRU-sf



- Self-Attention + Feature Fusion

# Model Training

- **Minimize weighted categorical cross-entropy**

$$loss = -\frac{1}{\sum_{i=1}^L N_i} \sum_{i=1}^L \sum_{j=1}^{N_i} \omega(c_j) \sum_{c=1}^{|\mathcal{C}|} y_j^c \log_2(\hat{y}_j^c)$$

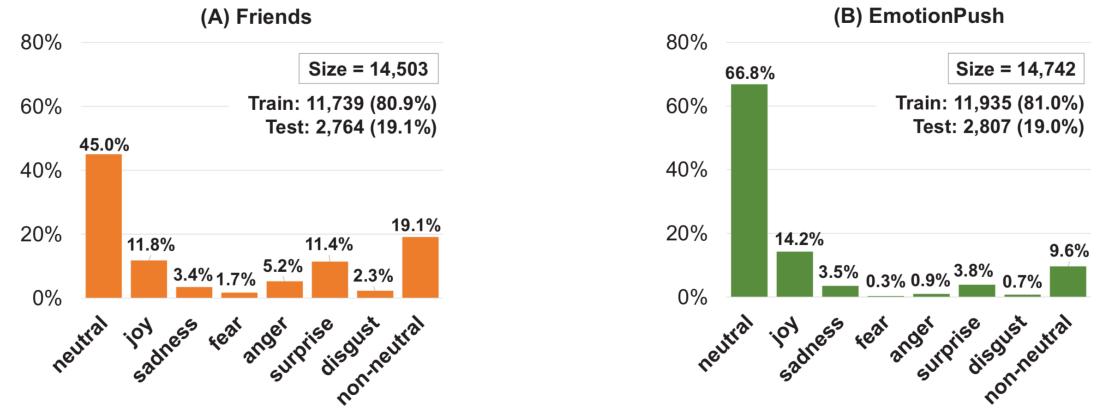
- predicted emotion:  $\hat{y}_j = \text{softmax}(W_{fc} \cdot e_c(u_j) + b_{fc})$
- weight:  $\frac{1}{\omega(c)} = \frac{I_c^\alpha}{\sum_{c'=1}^{|\mathcal{C}|} I_{c'}^\alpha}$

# Experiments

- **Datasets**

Dataset	#Dialogue (#Utterance)			#Emotion				
	Train	Val	Test	Ang	Hap/Joy	Sad	Neu	Others
IEMOCAP	96 (3,569)	24 (721)	31 (1,208)	1,090	1,627	1,077	1,704	0
Friends	720 (10,561)	80 (1,178)	200 (2,764)	759	1,710	498	6,530	5,006
EmotionPush	720 (10,733)	80 (1,202)	200 (2,807)	140	2,100	514	9,855	2,133

- IEMOCAP: ~12 hours of audiovisual data, including video, speech, motion capture of face, and text transcriptions
- Friends: Friends TV show transcripts
- EmotionPush: Facebook messenger logs



- **Metrics**

- Weighted Accuracy (WA)
- Unweighted Accuracy (UWA)

$$\text{WA} = \sum_{c=1}^{|C|} p_c \cdot a_c, \quad \text{UWA} = \frac{1}{|C|} \sum_{c=1}^{|C|} a_c,$$

# Methods and Setup

- **Methods**
  - Existing methods: bcLSTM, CMN, SA-BiLSTM, CNN-DCNN
  - Our implementation: bcLSTM<sub>\*</sub>(weighted loss), bcGRU(weighted loss), **HiGRU**, **HiGRU-f**, **HiGRU-sf**
- **Parameters of HiGRU, HiGRU-f, HiGRU-sf**
  - # hidden states: 300
  - FC layer: two sub-layers with 100 neurons each
- **Training:** Adam, Anneal strategy, early stop, gradient clipping, dropout

# Results on IEMOCAP

Model (Feat)	Ang	Hap	Sad	Neu	WA	UWA
bcLSTM <sup>1</sup> (T) (T+V+A)	76.07	78.97	76.23	67.44	73.6	<u>74.6</u>
	77.98	79.31	78.30	69.92	76.1	<u>76.3</u>
CMN <sup>2</sup> (T) (T+V+A)	-	-	-	-	74.1	-
	<b>89.88</b>	81.75	77.73	67.32	77.6	<u>79.1</u>
bcLSTM <sub>*</sub> (T)	75.29	79.40	78.07	76.53	77.7 <sub>(1.1)</sub>	77.3 <sub>(1.4)</sub>
bcGRU(T)	77.20	80.99	76.26	72.50	76.9 <sub>(1.6)</sub>	76.7 <sub>(1.3)</sub>
HiGRU(T)	75.41	<b>91.64</b>	79.79	70.74	80.6 <sub>(0.5)</sub>	79.4 <sub>(0.5)</sub>
HiGRU-f(T)	76.69	88.91	80.25	75.92	81.5 <sub>(0.7)</sub>	80.4 <sub>(0.5)</sub>
HiGRU-sf(T)	74.78	89.65	<b>80.50</b>	<b>77.58</b>	<b>82.1</b> <sub>(0.4)</sub>	<b>80.6</b> <sub>(0.2)</sub>

<sup>1</sup> by (Poria et al., 2017); <sup>2</sup> by (Hazarika et al., 2018).

- In WA: HiGRU achieves at least
  - 8.7% improvement over CMN (T) and
  - 3.8% improvement over CMN (T+V+A)

# Results on Friends and EmotionPush

Model	Train	Friends (F)						EmotionPush (E)					
		Ang	Joy	Sad	Neu	WA	UWA	Ang	Joy	Sad	Neu	WA	UWA
SA-BiLSTM <sup>1</sup>	F+E	49.1	68.8	30.6	<b>90.1</b>	-	59.6	24.3	70.5	31.0	<b>94.2</b>	-	55.0
CNN-DCNN <sup>2</sup>	F+E	55.3	71.1	55.3	68.3	-	62.5	45.9	76.0	51.7	76.3	-	62.5
bcLSTM_*	F(E)	64.7	69.6	48.0	75.6	72.4(4.2)	64.4(1.6)	32.9	69.9	47.1	78.0	74.7(4.4)	57.0(2.1)
bcGRU	F(E)	69.5	65.4	52.9	74.7	71.7(4.7)	65.6(1.2)	33.7	71.1	57.2	76.1	73.9(2.9)	59.5(1.8)
bcLSTM_*	F+E	54.5	75.6	43.4	73.0	70.5(4.5)	61.6(1.6)	52.4	79.1	54.7	73.3	73.4(3.8)	64.9(2.1)
bcGRU	F+E	59.0	78.6	42.3	71.4	70.2(5.1)	62.8(1.4)	49.4	74.8	61.9	72.4	72.1(4.3)	64.6(1.8)
HiGRU	F(E)	66.9	73.0	51.8	77.2	<b>74.4</b> (1.7)	67.2(0.6)	55.6	78.1	57.4	73.8	73.8(2.0)	66.3(1.7)
HiGRU-f	F(E)	69.1	72.1	<b>60.4</b>	72.1	71.3(2.9)	68.4(1.0)	55.9	78.9	60.4	72.4	73.0(2.2)	66.9(1.2)
HiGRU-sf	F(E)	<b>70.7</b>	70.9	57.7	76.2	74.0(1.4)	<b>68.9</b> (1.5)	57.5	78.4	64.1	72.5	73.0(1.6)	68.1(1.2)
HiGRU	F+E	55.4	81.2	51.4	64.4	65.8(4.2)	63.1(1.5)	50.8	76.9	69.0	75.7	75.3(1.7)	68.1(1.2)
HiGRU-f	F+E	54.9	78.3	55.5	68.7	68.5(3.0)	64.3(1.2)	<b>58.3</b>	79.1	<b>69.6</b>	70.0	71.5(2.5)	69.2(0.9)
HiGRU-sf	F+E	56.8	<b>81.4</b>	52.2	68.7	69.0(2.0)	64.8(1.3)	57.8	<b>79.3</b>	66.3	77.4	<b>77.1</b> (1.0)	<b>70.2</b> (1.1)

<sup>1</sup> by (Luo et al., 2018); <sup>2</sup> by (Khosla, 2018).

- HiGRU gains at least 6.0% improvement over CNN-DCNN, the best performance
- HiGRU-f and HiGRU-sf usually perform better than HiGRU
- Training with mixed datasets can only help the imbalanced dataset, EmotionPush

# Successful Cases

- **Scene-1:** both success
- **Scene-2:**
  - bcGRU: Joy → Ang
- **Scene-3:**
  - bcGRU:
    - Sad → Hap
    - Sad → Neu
  - HiGRU-sf:
    - Hap → Sad

Role	Utterance	Truth	bcGRU	HiGRU-sf
<b>Scene-1</b>				
Phoebe	Okay. Oh but don't tell them Monica's pregnant because they frown on that.	Neu	Neu	Neu
Rachel	Okay.	Neu	Neu	Neu
Phoebe	Okay.	Neu	Neu	Neu
<b>Scene-2</b>				
Phoebe	Yeah! Sure! Yep! Oh, y'know what? If I heard a shot right now, I'd throw my body on you.	Joy	Ang	Joy
Gary	Oh yeah? Well maybe you and I should take a walk through a bad neighborhood.	Other	/	/
Phoebe	Okay!	Joy	Ang	Joy
Gary	All right.	Neu	Neu	Neu
<b>Scene-3</b>				
Female	Can I send you, like videos and stuff? What about when they start walking.	Other	/	/
Male	Yeah yeah yeah.	Sad	Hap	Sad
Male	You you record every second. You record every second because I want to see it all. Okay?	Hap	Hap	Sad
Male	If I don't get to see it now, I get to see it later at least, you know? You've got to keep it all for me; all right?	Other	/	/
Female	Okay.	Sad	Neu	Sad

# Failed Cases

- **Scene-4:**

- bcGRU: Joy → Sad
- HiGRU: Joy → Neu

- **Scene-5:**

- bcGRU:
  - Neu → Sad
  - Joy → Neu
- HiGRU-sf
  - Neu → Sad
  - Joy → Sad

Role	Utterance	Truth	bcGRU	HiGRU-sf
<b>Scene-4</b>				
Ross	Hi.	Neu	Neu	Neu
Rachel	Hi.	Neu	Neu	Neu
Ross	Guess what?	Neu	Neu	Neu
Rachel	What?	Neu	Neu	Neu
Ross	They published my paper.	Joy	Sad	Neu
Rachel	Oh, really, let me see, let me see.	Joy	Neu	Neu
Phoebe	Rach, look! Oh, hi! Where is my strong Ross Skywalker to come rescue me. There he is.	Other	/	/
<b>Scene-5</b>				
Speaker-1	Sorry for keeping you up	Sad	Sad	Sad
Speaker-2	Lol don't be	Joy	Joy	Joy
Speaker-2	I didn't have to get up today	Neu	Sad	Sad
Speaker-1	:p	Joy	Joy	Joy
Speaker-2	It's actually been a really lax day	Joy	Neu	Sad

# Summary

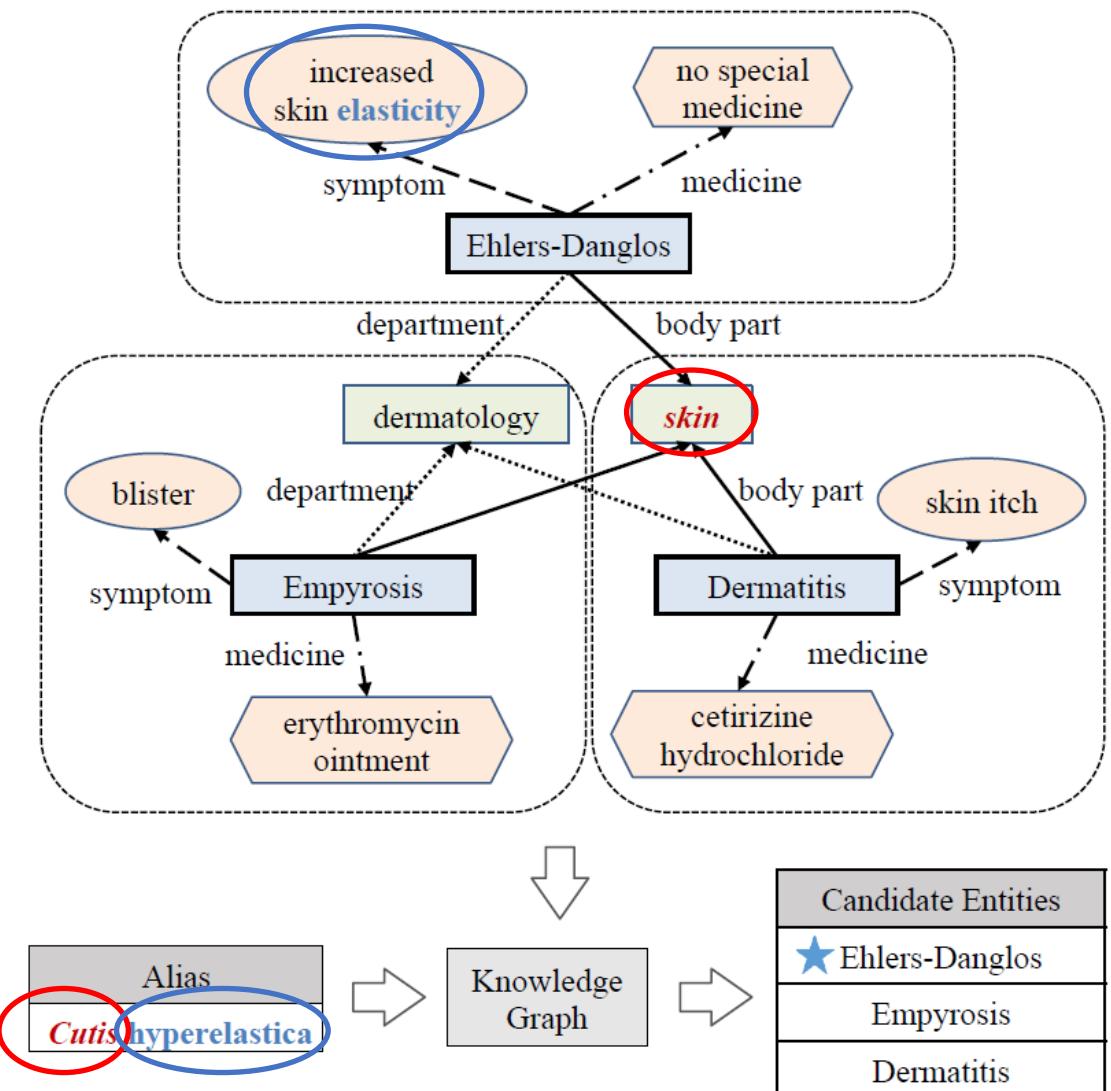
- A hierarchical Gated Recurrent Unit (HiGRU) framework
  - to tackle the utterance-level emotion recognition in dialogue systems
  - Lower-level GRU: learn the individual utterance embeddings
  - Upper-level GRU: capture the contexts of utterances
- Two variants
  - HiGRU-f: capture the word/utterance-level inputs, and
  - HiGRU-sf: capture the long-range contextual information
- Demonstrate the superior performance on three public datasets

# Representative Work II

KGSynNet: A Medical Entity Alias Discovery Framework with Knowledge Graphs

# Motivation

- KG entity alias (synonym) discovery aims to find synonymous aliases for an entity in knowledge graphs
- Challenges
  - Only query terms, no context
  - Only entities in knowledge graph
- Existing methods via *surface string matching* or *word/char embedding* **cannot** capture external knowledge



# Contribution

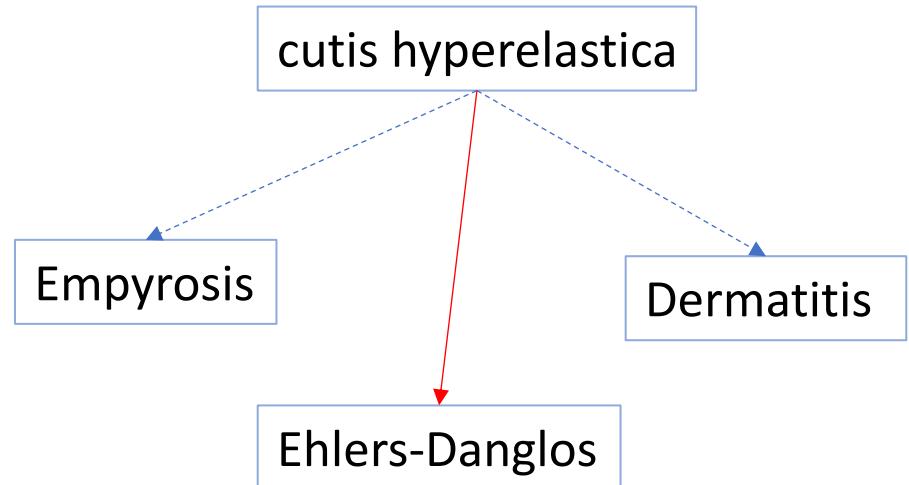
- A novel framework, **KGSynNet**, to
  - jointly learn both semantic feature and knowledge representation of entities from knowledge graphs
  - craftily design **fusion gate** to enhance information interaction
  - demonstrate the effectiveness through experiments on both offline and online test
- The **first health insurance benchmark** consists of
  - a Chinese cross-domain knowledge graph: occupations, diseases, and insurance products
  - a dataset of annotated alias-to-entity pairs of diseases

# Problem Definition: Entity alias discovery

- **Inputs (assumption: aliases is given)**
  - a set of disease query terms (aliases)
    - e.g., cutis hyperelastica
  - a cross-domain knowledge graph
    - occupations, diseases, and insurance products
  - a number of annotated alias-to-entity pairs of diseases
- **Output**
  - determine a list of synonymous entities for the disease query term

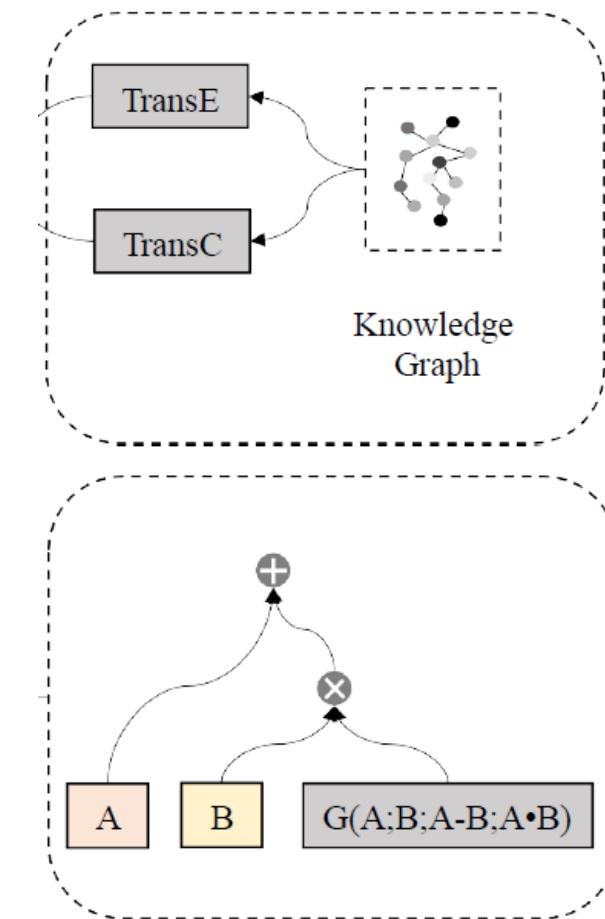
# Our Proposal

1. Input representations
  - Char embeddings (semantic information)
  - Entity knowledge embedding (pretrained)
2. Embedding space alignment
  - Shared weights in FC
3. Fusion of Entity's Semantic and Knowledge Representations
4. Similarity Matching: noise- contrastive estimation



# Properties: Adding External Knowledge

- Knowledge embedding
  - to represent triples in the knowledge graph via jointly TransC-TransE learning
- Adaptive knowledge integration
  - **Fusion gate** to adaptively incorporate the amount of knowledge with the learned semantics embeddings



# Results

- Difficult: no char overlap in aliases and entities
- Regular: at least one char overlap
- All: Difficult + Regular

Model variants	Hits@3			Hits@5			Hits@10		
	All	Regular	Difficult	All	Regular	Difficult	All	Regular	Difficult
JACCARD [19]	52.28%	56.61%	0.00%	58.03%	62.83%	0.00%	63.76%	69.04%	0.00%
w2v [3]	47.00%	50.88%	0.00%	52.28%	56.59%	2.30%	58.31%	63.10%	4.60%
CNN [16]	51.76%	55.69%	4.33%	57.75%	61.98%	6.38%	65.13%	69.72%	9.34%
BERT [4]	54.60%	58.87%	2.96%	60.41%	65.02%	4.78%	66.50%	71.39%	7.52%
DNorm [20]	56.23%	59.78%	12.76%	63.79%	67.58%	17.77%	71.89%	75.64%	26.42%
SurfCon [30]	58.29%	62.02%	12.98%	66.27%	70.11%	19.59%	75.20%	79.03%	28.93%
KGSynNet	<b>66.84%</b>	<b>70.81%</b>	18.91%	<b>73.09%</b>	<b>77.13%</b>	24.37%	<b>79.41%</b>	<b>83.35%</b>	31.89%
-TransC (TransE only)	65.80%	69.92%	15.95%	71.44%	75.79%	18.91%	78.94%	83.18%	27.80%
->Direct Addition	63.51%	67.19%	<b>19.13%</b>	70.85%	74.47%	<b>27.10%</b>	78.13%	81.77%	<b>34.17%</b>
->Ernie Fusion	61.98%	65.85%	15.26%	68.63%	72.54%	21.41%	76.28%	80.29%	27.79%
-Knowledge Embedding	64.91%	69.07%	14.58%	71.56%	75.77%	20.73%	79.12%	83.14%	30.52%

- KGSynNet beats all baselines
- Knowledge embedding plays a significant role in improving performance

# Summary

- A novel framework, KGSynNet,
  - captures both semantic meaning and knowledge information
  - effectively leverage the knowledge information via fusion gate
  - end-to-end implementation to learn entity representation and to discover an entity aliases
- The **first health insurance benchmark** for
  - Chinese cross-domain knowledge graphs, and
  - an annotated dataset for alias-to-entity pairs of diseases

# Conclusion

- Briefly review AI technologies development and applications in Ping An Life
- Present two representative work on
  - emotion recognition in dialogues
  - medical entity alias discovery in knowledge graphs
- Many potential applications and research problems exist ...

<https://iconip2020.apnns.org/>

A large, colorful word cloud centered around the words "thank you" in various languages. The word "thank" is in red, "you" is in yellow, and "you" is in green. The background is white with a subtle grid pattern. The word cloud includes many other words related to gratitude and thanks in different languages, such as "danke", "gracias", "merci", "mochchakkeram", and "arigato".

# AI Department

# Ping An Life

# **AI Empowers Insurance Services**



[hqyang.github.io](https://hqyang.github.io)  
[hqyang@ieee.org](mailto:hqyang@ieee.org)