AI Technologies Empower Professional Services Insurance in Ping An Life

Dr. Haiqin Yang
Al 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
  • ...

Agents

Customers

1.2M+

200M+
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

<table>
<thead>
<tr>
<th>Dialog Q&amp;A</th>
<th>Real-time Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real-time Conversation Assistance</td>
<td>Content Generation</td>
</tr>
</tbody>
</table>
Professional Insurance Services: Video Follow-up Chatbot

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

- **Impersonate video interaction**
  - User-friendly

- **Flexible follow-up**
  - 7*24 online service

- **Efficient access to insurance policy**
  - Real-time Q&A

- **Realtime Services**
  - Image Generation
  - Face Recognition
  - Lips Matching
  - Active Dialogue Guidance
**Insurance Middle Platform (MP)**

**Applications**
- Recruit-ment
- Train-ing
- Sales
- Services
- Manage-ment
- Risk control
- Asset man.
- ...

**Insurance MP**

**App. layer**
- Chatbot MP
- NLP MP
- Knowledge MP
- Content MP
- Image MP

**Comp. layer**
- Deep Learning Platform
  - Training
  - Inference
  - Scheduling
- Deep Learning
- Reinforce. Learning
- Al Algo.

**Data layer**
- Features
- Labels
- Dialogue
- Knowledge
- Corpus
- Multimedia
- ......
Chatbots

- **Champions**: 7
  - DSTC8, SemEval2020

- **Papers**: 10 +
  - IJCAI, CIKM, ...

- **Patents**: 300 +
  - System, Technology, ...

- **Platforms**: 10+

- **Online Services**: 100+

- **Chatbots**: 20+

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
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

Max-pooling

\[ e(x) \]

\[ e_u(w_k) \]

\[ h_1 \]

\[ h_2 \]

\[ h_3 \]

\[ h_4 \]

\[ h_{M_j} \]

\[ e(u_j) \]

Softmax

Fully-connected

\[ e_u(u_j) \]

\[ u_1 \]

\[ u_2 \]

\[ u_3 \]

\[ u_4 \]

\[ u_{N_I} \]

\[ w_1 \]

\[ w_2 \]

\[ w_3 \]

\[ w_4 \]

\[ w_{M_j} \]
HiGRU-f

- Fuse individual word/utterance embeddings to strengthen individual information
HiGRU-sf

- Self-Attention + Feature Fusion
Model Training

• Minimize weighted categorical cross-entropy

\[
loss = -\frac{1}{\sum_{i=1}^{L} N_i} \sum_{i=1}^{L} \sum_{j=1}^{N_i} \omega(c_j) \sum_{c=1}^{|C|} y_{ij}^c \log_2(\hat{y}_{ij}^c)
\]

• predicted emotion: \( \hat{y}_j = \text{softmax}(W_{fc} \cdot e_c(u_j) + b_{fc}) \)

• weight: \( \frac{1}{\omega(c)} = \frac{I_c^\alpha}{\sum_{c'=1}^{|C|} I_{c'}^\alpha} \)
Experiments

• Datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>#Dialogue (#Utterance)</th>
<th>#Emotion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Train</td>
<td>Val</td>
</tr>
<tr>
<td>IEMOCAP</td>
<td>96 (3,569)</td>
<td>24 (721)</td>
</tr>
<tr>
<td>Friends</td>
<td>720 (10,561)</td>
<td>80 (1,178)</td>
</tr>
<tr>
<td>EmotionPush</td>
<td>720 (10,733)</td>
<td>80 (1,202)</td>
</tr>
</tbody>
</table>

• 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)
  \[ WA = \sum_{c=1}^{|C|} p_c \cdot a_c, \quad UWA = \frac{1}{|C|} \sum_{c=1}^{|C|} a_c, \]
Methods and Setup

• **Methods**
  - Existing methods: bcLSTM, CMN, SA-BiLSTM, CNN-DCNN
  - Our implementation: bcLSTM+(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

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

\(^1\) by (Poria et al., 2017); \(^2\) 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

<table>
<thead>
<tr>
<th>Model</th>
<th>Train</th>
<th>Friends (F)</th>
<th>EmotionPush (E)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Ang Joy Sad Neu</td>
<td>WA UWA</td>
</tr>
<tr>
<td>SA-BiLSTM¹</td>
<td>F+E</td>
<td>49.1 68.8 30.6 90.1</td>
<td>- 59.6</td>
</tr>
<tr>
<td>CNN-DCNN²</td>
<td>F+E</td>
<td>55.3 71.1 55.3 68.3</td>
<td>- 62.5</td>
</tr>
<tr>
<td>bcLSTMₙ</td>
<td>F(E)</td>
<td>64.7 69.6 48.0 75.6</td>
<td>72.4(4.2) 64.4(1.6)</td>
</tr>
<tr>
<td>bcGRU</td>
<td>F(E)</td>
<td>69.5 65.4 52.9 74.7</td>
<td>71.7(4.7) 65.6(1.2)</td>
</tr>
<tr>
<td>bcLSTMₙ</td>
<td>F+E</td>
<td>54.5 75.6 43.4 73.0</td>
<td>70.5(4.5) 61.6(1.6)</td>
</tr>
<tr>
<td>bcGRU</td>
<td>F+E</td>
<td>59.0 78.6 42.3 71.4</td>
<td>70.2(5.1) 62.8(1.4)</td>
</tr>
<tr>
<td>HiGRU</td>
<td>F(E)</td>
<td>66.9 73.0 51.8 77.2</td>
<td>74.4(1.7) 67.2(0.6)</td>
</tr>
<tr>
<td>HiGRU-f</td>
<td>F(E)</td>
<td>69.1 72.1 60.4 72.1</td>
<td>71.3(2.9) 68.4(1.0)</td>
</tr>
<tr>
<td>HiGRU-sf</td>
<td>F(E)</td>
<td>70.7 70.9 57.7 76.2</td>
<td>74.0(1.4) 68.9(1.5)</td>
</tr>
<tr>
<td>HiGRU</td>
<td>F+E</td>
<td>55.4 81.2 51.4 64.4</td>
<td>65.8(4.2) 63.1(1.5)</td>
</tr>
<tr>
<td>HiGRU-f</td>
<td>F+E</td>
<td>54.9 78.3 55.5 68.7</td>
<td>68.5(3.0) 64.3(1.2)</td>
</tr>
<tr>
<td>HiGRU-sf</td>
<td>F+E</td>
<td>56.8 81.4 52.2 68.7</td>
<td>69.0(2.0) 64.8(1.3)</td>
</tr>
</tbody>
</table>

¹ by (Luo et al., 2018); ² 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
Failed Cases

• **Scene-4:**
  - bcGRU: Joy → Sad
  - HiGRU: Joy → Neu

• **Scene-5:**
  - bcGRU:
    - Neu → Sad
    - Joy → Neu
  - HiGRU-sf
    - Neu → Sad
    - Joy → Sad

<table>
<thead>
<tr>
<th>Role</th>
<th>Utterance</th>
<th>Truth</th>
<th>bcGRU</th>
<th>HiGRU-sf</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scene-4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross</td>
<td>Hi.</td>
<td>Neu</td>
<td>Neu</td>
<td>Neu</td>
</tr>
<tr>
<td>Rachel</td>
<td>Hi.</td>
<td>Neu</td>
<td>Neu</td>
<td>Neu</td>
</tr>
<tr>
<td>Ross</td>
<td>Guess what?</td>
<td>Neu</td>
<td>Neu</td>
<td>Neu</td>
</tr>
<tr>
<td>Rachel</td>
<td>What?</td>
<td>Neu</td>
<td>Neu</td>
<td>Neu</td>
</tr>
<tr>
<td>Ross</td>
<td>They published my paper.</td>
<td>Joy</td>
<td>Sad</td>
<td>Neu</td>
</tr>
<tr>
<td>Rachel</td>
<td>Oh, really, let me see, let me see.</td>
<td>Joy</td>
<td>Neu</td>
<td>Neu</td>
</tr>
<tr>
<td>Phoebe</td>
<td>Rach, look! Oh, hi! Where is my strong Ross Skywalker to come rescue me. There he is.</td>
<td>Other</td>
<td>/</td>
<td>/</td>
</tr>
</tbody>
</table>

| **Scene-5**|                                               |       |       |          |
| 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

<table>
<thead>
<tr>
<th>Model variants</th>
<th>Hits@3</th>
<th>Hits@5</th>
<th>Hits@10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Regular</td>
<td>Difficult</td>
</tr>
<tr>
<td>JACCARD [19]</td>
<td>52.28%</td>
<td>56.61%</td>
<td>0.00%</td>
</tr>
<tr>
<td>w2v [3]</td>
<td>47.00%</td>
<td>50.88%</td>
<td>0.00%</td>
</tr>
<tr>
<td>CNN [16]</td>
<td>51.76%</td>
<td>55.69%</td>
<td>4.33%</td>
</tr>
<tr>
<td>BERT [4]</td>
<td>54.60%</td>
<td>58.87%</td>
<td>2.96%</td>
</tr>
<tr>
<td>DNorm [20]</td>
<td>56.23%</td>
<td>59.78%</td>
<td>12.76%</td>
</tr>
<tr>
<td>SurfCon [30]</td>
<td>58.29%</td>
<td>62.02%</td>
<td>12.98%</td>
</tr>
<tr>
<td>KGSynNet</td>
<td>66.84%</td>
<td>70.81%</td>
<td>18.91%</td>
</tr>
<tr>
<td>— TransC (TransE only)</td>
<td>65.80%</td>
<td>69.92%</td>
<td>15.95%</td>
</tr>
<tr>
<td>— Direct Addition</td>
<td>63.51%</td>
<td>67.19%</td>
<td>19.13%</td>
</tr>
<tr>
<td>— Ernie Fusion</td>
<td>61.98%</td>
<td>65.85%</td>
<td>15.26%</td>
</tr>
<tr>
<td>— Knowledge Embedding</td>
<td>64.91%</td>
<td>69.07%</td>
<td>14.58%</td>
</tr>
</tbody>
</table>

- 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/