

SpringerBriefs in Computer Science

Hua Xu · Hanlei Zhang · Ting-En Lin



# Intent Recognition for Human-Machine Interactions



清华大学出版社  
TSINGHUA UNIVERSITY PRESS



Springer

# **SpringerBriefs in Computer Science**

SpringerBriefs present concise summaries of cutting-edge research and practical applications across a wide spectrum of fields. Featuring compact volumes of 50 to 125 pages, the series covers a range of content from professional to academic.

Typical topics might include:

- A timely report of state-of-the art analytical techniques
- A bridge between new research results, as published in journal articles, and a contextual literature review
- A snapshot of a hot or emerging topic
- An in-depth case study or clinical example
- A presentation of core concepts that students must understand in order to make independent contributions

Briefs allow authors to present their ideas and readers to absorb them with minimal time investment. Briefs will be published as part of Springer's eBook collection, with millions of users worldwide. In addition, Briefs will be available for individual print and electronic purchase. Briefs are characterized by fast, global electronic dissemination, standard publishing contracts, easy-to-use manuscript preparation and formatting guidelines, and expedited production schedules. We aim for publication 8–12 weeks after acceptance. Both solicited and unsolicited manuscripts are considered for publication in this series.

**\*\*Indexing:** This series is indexed in Scopus, Ei-Compendex, and zbMATH\*\*

Hua Xu • Hanlei Zhang • Ting-En Lin

# Intent Recognition for Human-Machine Interactions

 Springer

 清华大学出版社  
TSINGHUA UNIVERSITY PRESS

Hua Xu   
Department of Computer Science  
and Technology  
Tsinghua University  
Beijing, China

Hanlei Zhang   
Department of Computer Science  
and Technology  
Tsinghua University  
Beijing, China

Ting-En Lin   
DAMO Academy  
Alibaba Group  
Beijing, China

ISSN 2191-5768 ISSN 2191-5776 (electronic)  
SpringerBriefs in Computer Science  
ISBN 978-981-99-3884-1 ISBN 978-981-99-3885-8 (eBook)  
<https://doi.org/10.1007/978-981-99-3885-8>

Jointly published with Tsinghua University Press  
The print edition is not for sale in China (Mainland). Customers from China (Mainland) please order the  
print book from: Tsinghua University Press.

© Tsinghua University Press 2023

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publishers, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publishers nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publishers remain neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.  
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

# Preface

The natural interaction ability between human and machine mainly involves human-machine dialogue ability, multi-modal sentiment analysis ability, human-machine cooperation ability, and so on. In order to realize the efficient dialogue ability of intelligent computer, it is necessary to make the computer own strong user intention understanding ability in the process of human-computer interaction. This is one of the key technologies to realize efficient and intelligent human-computer dialogue.

Currently, the understanding of the objects to be analyzed requires different levels of ability such as recognition, cognition, and reasoning. The current research on human-computer interaction intention understanding is still focused on the level of recognition. The research and application of human-computer natural interaction intention recognition mainly includes the following levels: intention classification, unknown intention detection, and open intention discovery. Intention understanding for natural interaction is a comprehensive research field involving the integration of natural language processing, machine learning, algorithms, human-computer interaction, and other aspects. In recent years, our research team from State Key Laboratory of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University has conducted a lot of pioneering research and applied work, which have been carried out in the field of intention understanding for natural interaction, especially in the field of intention classification, unknown intention detection and open intention discovery based on text information of human-machine dialogue based on deep learning models. Related achievements have also been published in the top academic international conferences in the field of artificial intelligence in recent years, such as *ACL*, *AAAI*, *ACM MM*, and well-known international journals, such as *Pattern Recognition* and *Knowledge-based Systems*. In order to systematically present the latest achievements in intention classification, unknown intention detection, and open intention discovery in academia in recent years, the relevant work achievements are systematically sorted out and presented to readers in the form of a complete systematic discussion.

Currently, the research on intention understanding in natural interaction develops quickly. The author's research team will timely sort out and summarize the latest

achievements and share them with readers in the form of a series of books in the future. This book can not only be used as a professional textbook in the fields of natural interaction, intelligent question answering (customer service), natural language processing, human-computer interaction, etc., but also as an important reference book for the research and development of systems and products in intelligent robots, natural language processing, human-computer interaction, etc.

As the natural interaction is a new and rapidly developing research field, limited by the author's knowledge and cognitive scope, mistakes and shortcomings in the book are inevitable. We sincerely hope that you can give us valuable comments and suggestions for our book. Please contact [xuhua@tsinghua.edu.cn](mailto:xuhua@tsinghua.edu.cn) or a third party in the open-source system platform <https://thuiar.github.io/> to give us a message. All of the related source codes and datasets for this book have also been shared on the following websites <https://github.com/thuiar/Books>.

The research work and writing of this book were supported by the National Natural Science Foundation of China (Project No. 62173195). We deeply appreciate the following student from State Key laboratory of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University for her hard preparing work: Xiaofei Chen. We also deeply appreciate the following students for the related research directions of cooperative innovation work: Ting-en Lin, Hanlei Zhang, Wenmeng Yu, and Xin Wang. Without the efforts of the members of our team, the book could not be presented in a structured form in front of every reader.

Beijing, China  
Beijing, China  
Beijing, China  
November 2022

Hua Xu  
Hanlei Zhang  
Ting-En Lin

# Contents

## Part I Overview

<b>1</b>	<b>Dialogue System</b>	3
1.1	Review of Dialogue System	3
	References	6
<b>2</b>	<b>Intent Recognition</b>	7
2.1	Review of the Literature on Intent Representation	8
2.1.1	Discrete Representation	8
2.1.2	Distributed Representation	10
2.1.3	Summary	16
2.2	Review of Known Intent Classification	17
2.2.1	Review of Single-Model Intent Classification	17
2.2.2	Review of Bi-model Intent Classification	17
2.2.3	Summary	18
2.3	Review of Unknown Intent Detection	18
2.3.1	Unknown Intent Detection Based on Traditional Discriminant Model	18
2.3.2	Unknown Intent Detection Based on Open Set Recognition in Computer Vision	19
2.3.3	Unknown Intent Detection Based on Out-of-Domain Detection	20
2.3.4	Unknown Intent Detection Based on Other Methods	21
2.3.5	Summary	21
2.4	Review of New Intent Discovery	21
2.4.1	New Intent Discovery Based on Unsupervised Clustering	22



2.4.2	New Intent Discovery Based on Semi-Supervised Clustering . . . . .	23
2.4.3	Summary . . . . .	24
2.5	Conclusion . . . . .	25
	References . . . . .	26

## Part II Intent Classification

<b>3</b>	<b>Intent Classification Based on Single Model . . . . .</b>	<b>33</b>
3.1	Introduction . . . . .	33
3.2	Comparison Systems . . . . .	35
3.2.1	Baseline Systems . . . . .	35
3.2.2	NNLM-Based Utterance Classifier . . . . .	35
3.2.3	RNN-Based Utterance Classifier . . . . .	36
3.2.4	LSTM- and GRU-Based Utterance Classifier . . . . .	37
3.3	Experiments . . . . .	38
3.3.1	Datasets . . . . .	38
3.3.2	Experiment Settings . . . . .	38
3.3.3	Experiment Results . . . . .	40
3.4	Conclusion . . . . .	43
	References . . . . .	44
<b>4</b>	<b>A Dual RNN Semantic Analysis Framework for Intent Classification and Slot . . . . .</b>	<b>45</b>
4.1	Introduction . . . . .	46
4.2	Intent Classification and Slot Filling Task Methods . . . . .	46
4.2.1	Deep Neural Network for Intent Detection . . . . .	46
4.2.2	Recurrent Neural Network for Slot Filling . . . . .	47
4.2.3	Joint Model for Two Tasks . . . . .	47
4.3	Bi-Model RNN Structures for Joint Semantic Frame Parsing . . . . .	47
4.3.1	Bi-model Structure with a Decoder . . . . .	49
4.3.2	Bi-Model Structure without a Decoder . . . . .	49
4.3.3	Asynchronous Training . . . . .	50
4.4	Experiments . . . . .	51
4.4.1	Datasets . . . . .	51
4.4.2	Experiment Settings . . . . .	51
4.4.3	Experiment Results . . . . .	51
4.5	Conclusion . . . . .	52
	References . . . . .	53

## Part III Unknown Intent Detection

<b>5</b>	<b>Unknown Intent Detection Method Based on Model Post-Processing . . . . .</b>	<b>57</b>
5.1	Introduction . . . . .	57
5.2	A Post-Processing for New Intent Detection . . . . .	58

5.2.1	Classifiers . . . . .	59
5.2.2	SofterMax . . . . .	62
5.2.3	Deep Novelty Detection . . . . .	65
5.2.4	SMDN . . . . .	66
5.3	Experiments . . . . .	67
5.3.1	Datasets . . . . .	67
5.3.2	Baselines . . . . .	68
5.3.3	Experiment Settings . . . . .	68
5.3.4	Experiment Results . . . . .	70
5.4	Conclusion . . . . .	75
	References . . . . .	76
<b>6</b>	<b>Unknown Intent Detection Based on Large-Margin Cosine Loss . . .</b>	<b>77</b>
6.1	Introduction . . . . .	77
6.2	New Intent Detection Model Based on Large Margin Cosine Loss Function . . . . .	78
6.2.1	Large Margin Cosine Loss (LMCL) . . . . .	79
6.3	Experiments . . . . .	79
6.3.1	Datasets . . . . .	79
6.3.2	Baselines . . . . .	80
6.3.3	Experiment Setting . . . . .	80
6.3.4	Experiment Results . . . . .	81
6.4	Conclusion . . . . .	82
	References . . . . .	82
<b>7</b>	<b>Unknown Intention Detection Method Based on Dynamic Constraint Boundary . . . . .</b>	<b>83</b>
7.1	Introduction . . . . .	83
7.2	The Frame Structure of the Model . . . . .	84
7.3	The Main Approach . . . . .	85
7.3.1	Intent Representation . . . . .	85
7.3.2	Pre-Training . . . . .	85
7.3.3	Adaptive Decision Boundary Learning . . . . .	86
7.3.4	Open Classification with Decision Boundary . . . . .	88
7.4	Experiments . . . . .	88
7.4.1	Datasets . . . . .	88
7.4.2	Baselines . . . . .	89
7.4.3	Experiment Settings . . . . .	89
7.4.4	Experiment Results . . . . .	90
7.5	Discussion . . . . .	92
7.5.1	Boundary Learning Process . . . . .	92
7.5.2	Effect of Decision Boundary . . . . .	92
7.5.3	Effect of Labeled Data . . . . .	93
7.5.4	Effect of Known Classes . . . . .	94

7.6	Conclusion . . . . .	95
	References . . . . .	96

## Part IV Discovery of Unknown Intents

<b>8</b>	<b>Discovering New Intents Via Constrained Deep Adaptive Clustering with Cluster Refinement . . . . .</b>	<b>99</b>
8.1	Introduction . . . . .	99
8.2	New Intent Discovery Model Based on Self-Supervised Constrained Clustering . . . . .	100
8.2.1	Intent Representation . . . . .	100
8.2.2	Pairwise-Classification with Similarity Loss . . . . .	102
8.2.3	Cluster Refinement with KLD Loss . . . . .	104
8.3	Experiments . . . . .	105
8.3.1	Datasets . . . . .	105
8.3.2	Baseline . . . . .	105
8.3.3	Experiment Settings . . . . .	107
8.3.4	Experiment Results . . . . .	107
8.4	Conclusion . . . . .	111
	References . . . . .	112
<b>9</b>	<b>Discovering New Intents with Deep Aligned Clustering . . . . .</b>	<b>115</b>
9.1	Introduction . . . . .	115
9.2	Deep Aligned Clustering . . . . .	117
9.2.1	Intent Representation . . . . .	117
9.2.2	Transferring Knowledge From Known Intents . . . . .	117
9.2.3	Deep Aligned Clustering . . . . .	118
9.3	Experiments . . . . .	120
9.3.1	Datasets . . . . .	120
9.3.2	Baselines . . . . .	120
9.3.3	Experiment Settings . . . . .	122
9.4	Conclusion . . . . .	126
	References . . . . .	127

## Part V Dialogue Intent Recognition Platform

<b>10</b>	<b>Experiment Platform for Dialogue Intent Recognition Based on Deep Learning . . . . .</b>	<b>131</b>
10.1	Introduction . . . . .	131
10.2	Open Intent Recognition Platform . . . . .	133
10.2.1	Data Management . . . . .	133
10.2.2	Models . . . . .	133
10.2.3	Training and Evaluation . . . . .	136
10.2.4	Result Analysis . . . . .	137

10.3	Pipeline Framework . . . . .	138
10.4	Experiments . . . . .	140
10.5	Conclusion . . . . .	142
	References . . . . .	143
 <b>Part VI Summary and Future Work</b>		
<b>11</b>	<b>Summary . . . . .</b>	<b>147</b>
	<b>Appendix . . . . .</b>	<b>151</b>

# List of Figures

Fig. 2.1	Neural network language model .....	11
Fig. 2.2	Recurrent neural network language model .....	12
Fig. 2.3	CBOW model .....	13
Fig. 2.4	Skip-gram model .....	13
Fig. 2.5	Pre-trained language models. (a) ELMo. (b) GPT. (c) BERT ....	15
Fig. 2.6	An example for discovering new intents .....	23
Fig. 3.1	The standard neural network language model .....	35
Fig. 3.2	RNN classifier model .....	37
Fig. 3.3	DET curves for (a) intent detection from ATIS words (b) intent detection from ATIS autotags and (c) domain classification in Cortana .....	42
Fig. 4.1	Bi-model Structure. (a) Bi-model structure with a decoder. (b) Bi-model structure without a decoder .....	48
Fig. 5.1	The experiment flowchart for unknown intent detection in the single-turn dialogue system .....	58
Fig. 5.2	The experiment flowchart for unknown intent detection in the multi-turn dialogue system .....	59
Fig. 5.3	The difference between Softmax and SofterMax .....	62
Fig. 5.4	Macro F1-Score on SNIPS dataset with different proportion of classes treated as known intents .....	71
Fig. 5.5	Macro F1-Score on ATIS dataset with different proportion of classes treated as known intents .....	71
Fig. 5.6	Macro F1-Score on SwDA dataset with different proportion of classes treated as known intents .....	71
Fig. 5.7	Confusion matrix for the overall classification results with SMDN method on three different datasets .....	73
Fig. 5.8	Confidence score distribution of SofterMax .....	74
Fig. 5.9	Novelty score distribution of LOF .....	74
Fig. 5.10	The distribution of novelty probability after Platt scaling for SofterMax, LOF and SMDN .....	75

Fig. 6.1	A new intent detection model based on large marginal cosine loss function .....	78
Fig. 6.2	Visualization of deep features learned with Softmax and LMCL on SNIPS dataset .....	81
Fig. 7.1	The model architecture of dialogue intention discovery model based on dynamic constraint boundary .....	84
Fig. 7.2	The boundary learning process .....	92
Fig. 7.3	Influence of the learned decision boundary .....	93
Fig. 7.4	Influence of labeled ratio on three datasets with different known class proportions (25%, 50%, 75%) .....	93
Fig. 8.1	The model architecture of CDAC+ .....	101
Fig. 8.2	Influence of the number of clusters on three datasets .....	109
Fig. 8.3	Confusion matrix for the clustering results of CDAC+ on SNIPS datasets .....	109
Fig. 8.4	Influence of the labeled ratio on three datasets .....	110
Fig. 8.5	Influence of the unknown class ratio on three datasets .....	110
Fig. 8.6	Visualization of intent representation learned on StackOverflow dataset .....	112
Fig. 9.1	The model architecture of the approach presented .....	116
Fig. 9.2	Influence of the known class ratio on CLINC dataset .....	125
Fig. 9.3	Influence of the known class ratio on BANKING dataset .....	125
Fig. 9.4	Influence of the number of clusters on BANKING dataset .....	125
Fig. 9.5	Influence of the number of clusters on CLINC dataset .....	126
Fig. 10.1	An example for open intent recognition .....	132
Fig. 10.2	The architecture of the TEXTOIR platform .....	134
Fig. 10.3	The architecture of open intent recognition .....	135
Fig. 10.4	Known and open intent distributions with different confidence scores .....	138
Fig. 10.5	Visualization of the intent representations .....	139
Fig. 10.6	Intent center distribution .....	140
Fig. 10.7	The pipeline framework of open intent recognition .....	141

# List of Tables

Table 3.1	ATIS intent classification results .....	40
Table 3.2	Average, best held out, and oracle errors on ATIS intent classification .....	40
Table 3.3	Cortana domain classification results .....	41
Table 4.1	Performance of different models on ATIS dataset .....	52
Table 5.1	Statistics of ATIS, SNIPS, and SwDA dataset .....	67
Table 5.2	Classifier performance of using all classes for training on different datasets (%) .....	68
Table 5.3	Macro F1-Score of unknown intent detection on SNIPS, ATIS and SwDA dataset with different proportion of classes treated as known intents .....	70
Table 5.4	Expected calibration error of the model(ECE) .....	72
Table 6.1	Macro f1-score of unknown intent detection with different proportion (25%, 50% and 75%) of classes are treated as known intents on SNIPS and ATIS dataset .....	81
Table 7.1	Statistics of BANKING, OOS and StackOverflow datasets. (# indicates the total number of sentences) .....	89
Table 7.2	Results of open classification with different known class proportions(25%, 50% and 75%) on BANKING, OOS and StackOverfflow datasets .....	90
Table 7.3	Results of open classification with different known class ratios(25%, 50% and 75%) on BANKING, OOS and StackOverfflow datasets .....	91
Table 8.1	Statistics of SNIPS, DBPedia, and StackOverfflow dataset . . .	106
Table 8.2	The clustering results on three datasets .....	108
Table 8.3	The clustering results of CDAC+ and its variant methods .....	109
Table 9.1	Statistics of CLINC and BANKING datasets .....	121
Table 9.2	The clustering results on two datasets .....	123
Table 9.3	Effectiveness of the pre-training and the alignment strategy on two datasets .....	124

Table 9.4    The results of predicting K with an unknown number  
              of clusters ..... 124

Table 10.1   The open intent recognition results of ADB + DeepAligned  
              on four datasets ..... 142