2021 International Conference on Asian Language Processing (IALP 2021)

Singapore
11 – 13 December 2021

IEEE Catalog Number: CFP2144I-POD
ISBN: 978-1-6654-8312-4

Copyright © 2021 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP2144I-POD
ISBN (Print-On-Demand):	978-1-6654-8312-4
ISBN (Online):	978-1-6654-8311-7

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

Table of Content

About the Conference IX

Message from the Conference Chairs X

Message from the Program Chairs XI

Conference Committees XII

Keynote Speeches XV

Author Index XVIII

Papers

1: Natural Language Understanding & Question Answering

P1-1 The Construction and Application of Reading Comprehension Model
Kai Zhang, Fumeng Ma and Weili Wang..... 1

P1-2 Mirror Distillation Model with Focal Loss for Chinese Machine Reading Comprehension
Fulai Nan, Jin Wang and Xuejie Zhang 7

P1-3 Multi-granularity Interaction Fusion for Neural Answer Source Selection
Hongyu Zhu, Zhilin Jin, Yan Bai, Yu Hong and Min Zhang 13

P1-4 Knowledge Graph Embedding Based on Semantic Hierarchical Spatial Rotation
Liangcheng Yin, Jie Zhu, Enshuai Hou and Ma Ni 19

P1-5 Automatic Scoring Method of Short-Answer Questions in the Context of Low-Resource Corpora
Tao-Hsing Chang, Ju-Ling Chen, Hui-Min Chou, Ming-Hong Bai, Fu-Yuan Hsu and Yu-Chi Chen 25

2: Spoken Language Processing

P2-1 End-to-end Indonesian Speech Synthesis Based On Transfer Learning And Alternate Training
Yu Lu, Jian Yang and Ruolin Yang 30

P2-2 Indonesian Speech Recognition Based on Deep Neural Network
Ruolin Yang, Jian Yang and Yu Lu 36

P2-3 Depressive Tendency Recognition Using the Gated Recurrent Unit From Speech and Text Features
Jingyi Yuan, Xiaoyong Lu, Yang Liu, Daimin Shi, Tao Pan and Yanqin Li 42

P2-4 Perception of Synthesized Flectional Tones in Mandarin Chinese by Native Chinese Speakers and Korean L2 Learners
Jung-yueh Tu and Chen Li 47

P2-5 Research on Depression Recognition Using Machine Learning from Speech
Daimin Shi, Xiaoyong Lu, Yang Liu, Jingyi Yuan, Tao Pan and Yanqin Li 52

P2-6 Improved Depression Recognition Using Attention and Multitask Learning of Gender Recognition
Yang Liu, Xiaoyong Lu, Daimin Shi, Jingyi Yuan, Tao Pan and Haizhen An 57

P2-7 Towards the Use of Pretrained Language Model GPT-2 for Testing the Hypothesis of Communicative Efficiency in the Lexicon
Yuqing Zhang, Zhu Li and Jinsong Zhang 62

P2-8 A Preliminary Study on the Gender Differences of Mandarin Focal Accent
Yingyang Wang and Jinsong Zhang 67

P2-9 Audio-Based Hate Speech Classification from Online Short-Form Videos
Michael Ibañez, Ranz Sapinit, Lloyd Antonie Reyes, Mohammed Hussien, Joseph Marvin Imperial and Ramon Rodriguez 72

P2-10 Mongolian emotional speech synthesis based on transfer learning and emotional embedding
Aihong Huang, Feilong Bao, Guanglai Gao, Yu Shan and Rui Liu 78

3: Information Extraction and Information Retrieval

P3-1 Employing Adversarial Training to Detect Rumors on Social Media
Xiaojia Li and Zhong Qian 84

P3-2 Deep-Learning-based Cross-Modal Luxury Microblogs Retrieval
Menghao Ma, Wuying Liu and Wenhe Feng 90

P3-3 Pseudo Label based Contrastive Sampling for Long Text Retrieval
Le Zhu, Shumin Shi and Heyan Huang 95

P3-4 Joint Extraction of Clinical Entities and Relations Using Multi-head Selection Method
Xintao Fang, Yuting Song and Akira Maeda 99

P3-5 Joint Model of Triple Relation Extraction with Label Embeddings
Mengjiao Xie, Jin Wang and Xuejie Zhang 105

P3-6 Causality Extraction in Futures Domain
Yuxiang Jia, Le Kang, Huiyong Sun, Xiaojing Du, Shuai Cao, Hongying Zan and Fang Li 111

P3-7 Multi-Strategy Fusion for Medical Named Entity Recognition
Jian Su, Rongchuan Tang, Lu Xiang, Feifei Zhai and Yu Zhou 117

P3-8 Chinese Medical Named Entity Recognition Based on Chinese Character Radical Features and Pre-trained Language Models
Haixin Tan, Zhihao Yang, Jinzhong Ning, Zeyuan Ding and Qiming Liu 121

4: Natural Language Generation, Machine Translation & Summarization

P4-1 Automatic Generation of Multiple-choice Cloze Test Questions for Lao Language Learning
Xinying Qiu, Haiwei Xue, Lingfeng Liang, Zexin Xie, Shuxuan Liao and Guofeng Shi 125

P4-2 Topic-to-Text Generation with PMI-IR Additional Semantic Information
Xinyi Ning *131*

P4-3 Speech Stimulus Generation: A Deep Learning Approach
Zhu Li, Yanlu Xie and Dengfeng Ke *137*

P4-4 English-Manipuri Machine Translation: An empirical study of different Supervised and Unsupervised Methods
Telem Joyson Singh, Sanasam Ranbir Singh and Priyankoo Sarmah *142*

P4-5 Latent Attribute Control for Story Generation
Yu-Siou Tang and Chung-Hsien Wu *148*

P4-6 Predicting Post-Editing Effort for English-Chinese Neural Machine Translation
Na Ye, Ling Jiang, Dandan Ma, Yingxin Zhang, Sanyuan Zhao and Dongfeng Cai *154*

P4-7 RESA: Relation Enhanced Self-Attention for Low-Resource Neural Machine Translation
Xing Wu, Shumin Shi and Heyan Huang *159*

5: Embedding, Semantics & Text Classification

P5-1 Efficient Language Identification for All-Language Internet News
Jian Tang, Xiaojiang Chen and Wuying Liu *165*

P5-2 Negative Siamese Network for Classifying Semantically Similar Sentences
Haibei Zhu and Baoyan Wang *170*

P5-3 An Exploratory Study to Extract Analogies from Arabic Text
Rawan N. Al-Matham and Hend S. Al-Khalifa *174*

P5-4 Rule-Based Semantic Role Labeling of Bei-sentences
Mengxiang Wang *178*

P5-5 An Interdisciplinary Literature Classifier Based on Multi-task Multi-label Learning
Lianxi Wang, Zhuolin Chen, Nankai Lin and Xixuan Huang *183*

P5-6 A Deep Learning Model Fused with Word Sense Knowledge for Textual Entailment Recognition
Yalei Liu, Lingling Mu and Hongying Zan *189*

P5-7 Manipuri-English Cross-lingual Word Embeddings using a Temporally Aligned Comparable Corpus
Lenin Laitonjam and Sanasam Ranbir Singh *195*

P5-8 Evaluation of Arabic-Based Contextualized Word Embedding Models
Sane Mo Yagi, Youssef Mansour, Firuz Kamalov and Ashraf Elnagar *200*

P5-9 Profiling Celebrity Profession from Twitter Data
Kumar Gourav Das, Braja Gopal Patra and Sudip Kumar Naskar *207*

6: Resources and Dialect

P6-1 The Affiliation of Yunxi Dialect in Hubei Province
Yan Li *213*

P6-2 Contextual Vowel Nasalization in Mongolian
Yin Bao and Rian Bao *217*

P6-3 The Quest for NLP Applications and Tools: The Case of Standard Arabic and the Dialects
Rawdah Abu Hashem *222*

P6-4 Corpus construction for topic-based summarization of multi-party conversation
Yuri Nakayama, Tsukasa Shiota and Kazutaka Shimada *229*

P6-5 Chinese Paraphrase Dataset and Detection
Bo An *235*

P6-6 HinPLMs: Pre-trained Language Models for Hindi
Xixuan Huang, Nankai Lin, Kexin Li, Lianxi Wang and Suifu Gan *241*

P6-7 Negative Transfer of Sichuan Dialect in English Vowel Acquisition
Jingya Jin and Donghong Liu *247*

P6-8 Coda Nasal Perception in Mandarin by Fuqing Min Speakers
Ying Chen, Yaping Chen and Yan Li *252*

7: Segmentation, Tagging, Syntax & Linguistics

P7-1 The Domain Value and Mental Space of Ba in the Mid-sentence
Minfeng Wang *256*

P7-2 The Calculation of Verb Similarity Based on Ternary Morpheme Collocation
Tian Shao, Endong Xun, Guirong Wang, Chengwen Wang, Gaoqi Rao and Bo Xia *262*

P7-3 A GRU-based Pipeline Approach for Word-Sentence Segmentation and Punctuation Restoration in English
Jasivan Sivakumar, Jake Muga, Flavio Spadavecchia, Daniel White and Burcu Can *268*

P7-4 Cognitive Motivation for the Definition of Chinese Four-Character Idioms
Qianqian Wang *274*

P7-5 A Hybrid of Rule-based and HMM-based Part-of-Speech Tagger for Indonesian
Muhammad Ridho Ananda, Muhammad Yudistira Hanifmuti and Ika Alfina *280*

P7-6 Building a Part of Speech tagger for the Tamil Language
Kengatharaiyer Sarveswaran and Gihan Dias *286*

P7-7 Ancient Tibetan Word Segmentation based on Deep Learning
Bo An and Congjun Long *292*

P7-8 Influence of Syntactic Structures on the Acquisition of Wh-questions: In the case of Tianjin English Learners
Binbin Sun and Hui Feng *298*

P7-9 The Cartography of Suffering Construction in a Broad Sense
Huibin Zhuang, Qionglei Lin and Jie Li *304*

P7-10 A Review of Morphological Analysis Methods on Uyghur Language
Gvzelnur Imin, Mijit Ablimit and Askar Hamdulla *310*

8: Sentiment Analysis and Opioin Mining

P8-1 NFCMF: Noise Filtering and CrossModal Fusion for Multimodal Sentiment Analysis
Guixin Su, Junyi He, Xia Li, Meixiu Lu and Hanqun Yang *316*

P8-2 Japanese Election Prediction Based on Sentiment Analysis of Twitter Replies to Candidates
Yoshio Okimoto, Yuki Hosokawa, Jianwei Zhang and Lin Li *322*

P8-3 Aspect-based Sentiment Classification with Dependency Relation and Structured Attention
Yuxiang Jia, Yadong Wang, Hongying Zan, Qi Xie, Yingjie Yan and Yalei Liu *328*

P8-4 Trans-CNN: A Sentiment Analysis Model Which Focuses on Key Information of Sentence
Jianlong Li, Yangsen Zhang, Gaijuan Huang, Zhuofan Huang, Renjie Wang and Xiang Chen *334*

P8-5 Tibetan-Chinese Cross-Lingual Sentiment Classification Based on Adversarial Network
Tingting Zhang, Tao Jiang and Ruikang Shan *340*

P8-6 DFNM: Dynamic Fusion Network of Intra- and Inter-modalities for Multimodal Sentiment Analysis
Zebin Li, Junteng Ma, Xia Li and Xin Pan *346*

P8-7 THANOS: The Aspect Classification Model for Imbalanced Vietnamese E-commerce Review Data
Minh-Son Cong, Tuan-Minh Dam, Anh-My Phuong, Thi-Quyen Nguyen, Duy-Cat Can, Hoang-Quynh Le, Long Hoang and Mai-Vu Tran *352*

P8-8 A Text Similarity Measurement Employs Semantic Dictionary-based Sentiment Analysis
Turdi Tohti, Shasha Li and Askar Hamdulla *358*

9: Natural Language Processing & Application

P9-1 EA-MLM: Error-aware Masked Language Modeling for Grammatical Error Correction
Junyi He, Xia Li, Han Su, Xinyin Chen, Hao Yang and Minghao Chen *363*

P9-2 A Sentence Classification Method for Chinese Spelling Error Detection Based on BERT
Jin Jiang and Yanquan Zhou *369*

P9-3 Quantitative Research on the Differences of Indonesian News Texts
Lixian Xiao, Nankai Lin and Shengyi Jiang *373*

P9-4 Scientific Landscape of Publications in Natural Language Processing in the ASEAN Region on COVID-19: A Bibliometric Approach
Rachel Edita Roxas, Rogelio Ruzcko Tobias and Johanna Minglana 379

P9-5 Analyses of Hate and Non-Hate Expressions during Election using NLP
Nico T. Solitana and Charibeth K. Cheng 385

P9-6 Gender differences in the production of sibilant contrast by native Chinese speakers
Kai Sheng, Jian Gong, Yuhong Sun, Weizhong Zhang, William Bellamy and Xiaoli Ji 391

P9-7 Predicting User Code-Switching Level from Sociological and Psychological Profiles
Injy Hamed, Alia El Bolock, Nader Rizk, Cornelia Herbert, Slim Abdennadher and Ngoc Thang Vu . 395