Table of Contents

Is Whole Word Masking Always Better for Chinese BERT?: Probing on Chinese Grammatical Error Correction
Yong Dai, Linyang Li, Cong Zhou, Zhangyin Feng, Enbo Zhao, Xipeng Qiu, Piji Li and Duyu Tang .......................................................... 1

Compilable Neural Code Generation with Compiler Feedback
Xin Wang, Yasheng Wang, Yao Wan, Fei Mi, Yitong Li, Pingyi Zhou, Jin Liu, Hao Wu, Xin Jiang and Qun Liu ........................................... 9

Towards Unifying the Label Space for Aspect- and Sentence-based Sentiment Analysis
Yiming Zhang, Min Zhang, Sai Wu and Junbo Zhao .................................. 20

Input-specific Attention Subnetworks for Adversarial Detection
Emil Biju, Anirudh Sriram, Pratyush Kumar and Mitesh M Khapra.......................... 31

RelationPrompt: Leveraging Prompts to Generate Synthetic Data for Zero-Shot Relation Triplet Extraction
Yew Ken Chia, Lidong Bing, Soujanya Poria and Luo Si .................................. 45

Pre-Trained Multilingual Sequence-to-Sequence Models: A Hope for Low-Resource Language Translation?
En-Shiun Annie Lee, Sarubi Thillainathan, Shrvan Nayak, Surangika Ranathunga, David Ifeoluwa Adelani, Ruisi Su and Arya D. McCarthy ......................... 58

Multi-Scale Distribution Deep Variational Autoencoder for Explanation Generation
ZeFeng Cai, Linlin Wang, Gerard de Melo, Fei Sun and Liang He ....................... 68

Dual Context-Guided Continuous Prompt Tuning for Few-Shot Learning
Jie Zhou, Le Tian, Houjin Yu, Zhou Xiao, Hui Su and Jie Zhou .......................... 79

Extract-Select: A Span Selection Framework for Nested Named Entity Recognition with Generative Adversarial Training
Peixin Huang, Xiang Zhao, Minghao Hu, Yang Fang, Xinyi Li and Weidong Xiao .......... 85

Controlled Text Generation Using Dictionary Prior in Variational Autoencoders
Xianghong Fang, Jian Li, Lifeng Shang, Xin Jiang, Qun Liu and Dit-Yan Yeung .................. 97

Challenges to Open-Domain Constituency Parsing
Sen Yang, Leyang Cui, Ruoxi Ning, Di Wu and Yue Zhang ............................ 112

Going Deeper: Structured Sememe Prediction via Transformer with Tree Attention
Yining Ye, Fanchao Qi, Zhiyuan Liu and Maosong Sun .................................. 128

Table-based Fact Verification with Self-adaptive Mixture of Experts
Yuxuan Zhou, Xien Liu, Kaiyin Zhou and Ji Wu ........................................ 139

Investigating Data Variance in Evaluations of Automatic Machine Translation Metrics
Jiannan Xiang, Huayang Li, Yahui Liu, Lemao Liu, Guoping Huang, Defu Lian and Shuming Shi 150

Sememe Prediction for BabelNet Synsets using Multilingual and Multimodal Information
Fanchao Qi, Chuancheng Lv, Zhiyuan Liu, Xiaojun Meng, Maosong Sun and Hai-Tao Zheng 158
Read Top News First: A Document Reordering Approach for Multi-Document News Summarization
Chao Zhao, Tenghao Huang, Somnath Basu Roy Chowdhury, Muthu Kumar Chandrasekaran, Kathleen McKeown and Snigdha Chaturvedi .................................................. 613

Human Language Modeling
Nikita Soni, Matthew Matero, Niranjan Balasubramanian and H. Schwartz ....... 622

Inverse is Better! Fast and Accurate Prompt for Few-shot Slot Tagging
Yutai Hou, Cheng Chen, Xianzhen Luo, Bohan Li and Wanxiang Che ............ 637

Cross-Modal Cloze Task: A New Task to Brain-to-Word Decoding
Shuxian Zou, Shaonian Wang, Jiajun Zhang and Chengqing Zong ............. 648

Mitigating Gender Bias in Distilled Language Models via Counterfactual Role Reversal
Umang Gupta, Jwala Dhamala, Varun Kumar, Apurv Verma, Yada Pruksachatkun, Satyapriya Krishna, Rahul Gupta, Kai-Wei Chang, Greg Ver Steeg and Aram Galstyan .......................... 658

Domain Representative Keywords Selection: A Probabilistic Approach
Pritom Saha Akash, Jie Huang, Kevin Chang, Yunyao Li, Lucian Popa and ChengXiang Zhai 679

Hierarchical Inductive Transfer for Continual Dialogue Learning
Shaoxiong Feng, Xuancheng Ren, Kan Li and Xu Sun .................................. 693

Why Exposure Bias Matters: An Imitation Learning Perspective of Error Accumulation in Language Generation
Kushal Arora, Layla El Asri, Hareesh Bahuleyan and Jackie CK Cheung ....... 700

Question Answering Infused Pre-training of General-Purpose Contextualized Representations
Robin Jia, Mike Lewis and Luke Zettlemoyer ............................................. 711

Automatic Song Translation for Tonal Languages
Fenfei Guo, Chen Zhang, Zhirui Zhang, Qixin He, Kejun Zhang, Jun Xie and Jordan Lee Boyd-Grabner ................................................................. 729

Read before Generate! Faithful Long Form Question Answering with Machine Reading
Dan SU, Xiaoguang Li, Jindi Zhang, Lifeng Shang, Xin Jiang, Qun Liu and Pascale Fung ... 744

A Simple yet Effective Relation Information Guided Approach for Few-Shot Relation Extraction
Yang Liu, Jinpeng Hu, Xiang Wan and Tsung-Hui Chang ............................. 757

MIMICause: Representation and automatic extraction of causal relation types from clinical notes
Vivek Khetan, Md Imbesat Hassan Rizvi, Jessica Huber, Paige Bartusiak, Bogdan Eugen Sacaleanu and Andrew Fano ............................................................. 764

Compressing Sentence Representation for Semantic Retrieval via Homomorphic Projective Distillation
Xuandong Zhao, Zhiguo Yu, Ming Wu and Lei Li ........................................ 774

Debiasing Event Understanding for Visual Commonsense Tasks
Minji Seo, YeonJoon Jung, Seungtaek Choi, Seung-won Hwang and Bei Liu .... 782

Fact-Tree Reasoning for N-ary Question Answering over Knowledge Graphs
Yao Zhang, Peiyao Li, Hongru Liang, Adam Jatowt and Zhenglu Yang ............ 788

DeepStruct: Pretraining of Language Models for Structure Prediction
Chenguang Wang, Xiao Liu, Zui Chen, Haoyun Hong, Jie Tang and Dawn Song .......... 803

The Change that Matters in Discourse Parsing: Estimating the Impact of Domain Shift on Parser Error
Katherine Atwell, Anthony Sicilia, Seong Jae Hwang and Malihe Alikhani ........ 824

xxxv
Mukayese: Turkish NLP Strikes Back
Ali Safaya, Emirhan Kurtuluş, Arda Goktogan and Deniz Yuret ................................. 846

Virtual Augmentation Supported Contrastive Learning of Sentence Representations
Dejiao Zhang, Wei Xiao, Henghui Zhu, Xiaofei Ma and Andrew Arnold .......................... 864

MoEfcation: Transformer Feed-forward Layers are Mixtures of Experts
Zhengyan Zhang, Yankai Lin, Zhiyuan Liu, Peng Li, Maosong Sun and Jie Zhou ............. 877

DS-TOD: Efficient Domain Specialization for Task-Oriented Dialog
Chia-Chien Hung, Anne Lauscher, Simone Paolo Ponzetto and Goran Glavaš ...................... 891

Distinguishing Non-natural from Natural Adversarial Samples for More Robust Pre-trained Language Model
Jiayi Wang, Rongzhou Bao, Zhuosheng Zhang and Hai Zhao ................................. 905

Learning Adaptive Axis Attentions in Fine-tuning: Beyond Fixed Sparse Attention Patterns
Zihan Wang, Jiuxiang Gu, Jason Kuen, Handong Zhao, Vlad I Morariu, Ruiyi Zhang, Ani Nenkova, Tong Sun and Jingbo Shang ................................................................. 916

Using Interactive Feedback to Improve the Accuracy and Explainability of Question Answering Systems Post-Deployment
Zichao Li, Prakhar Sharma, Xing Han Lu, Jackie CK Cheung and Siva Reddy ................. 926

To be or not to be an Integer? Encoding Variables for Mathematical Text
Deborah Ferreira, Mokanarangan Thayaparan, Marco Valentino, Julia Rozanova and Andre Freitas ............................................................ 938

GRS: Combining Generation and Revision in Unsupervised Sentence Simplification
Mohammad Dehghan, Dhruv Kumar and Lukasz Golab .................................................. 949

BPE vs. Morphological Segmentation: A Case Study on Machine Translation of Four Polysynthetic Languages
Manuel Mager, Arturo Oncevay, Elisabeth Mager, Katharina Kann and Thang Vu .......................... 961

Distributed NLI: Learning to Predict Human Opinion Distributions for Language Reasoning
Xiang Zhou, Yixin Nie and Mohit Bansal ................................................................. 972

Morphological Processing of Low-Resource Languages: Where We Are and What’s Next
Adam Wiemerslage, Miikka Silfverberg, Changbing Yang, Arya D. McCarthy, Garrett Nicolai, Eliana Colunga and Katharina Kann ......................................................... 988

Learning and Evaluating Character Representations in Novels
Naoya Inoue, Charuta Pethe, Allen Kim and Steven Skiena ........................................... 1008

Answer Uncertainty and Unanswerability in Multiple-Choice Machine Reading Comprehension
Vatsal Raina and Mark Gales ................................................................. 1020

Measuring the Language of Self-Disclosure across Corpora
Ann-Katrin Reuel, Sebastian Peralta, João Sedoc, Garrick Sherman and Lyle Ungar ........ 1035

When Chosen Wisely, More Data Is What You Need: A Universal Sample-Efficient Strategy For Data Augmentation
Ehsan Kamalloo, Mehdi Rezagholizadeh and Ali Ghodsi ................................. 1048

Explaining Classes through Stable Word Attributions
Samuel Rönnqvist, Aki-Juhani Kyröläinen, Amanda Myntti, Filip Ginter and Veronika Laippala 1063

xxxvi
Multi-Stage Prompting for Knowledgeable Dialogue Generation
Zihan Liu, Mostofa Patwary, Ryan Prenger, Shrimai Prabhumoye, Wei Ping, Mohammad Shoeybi and Bryan Catanzaro ................................................................. 1317

DuReader_vis: A Chinese Dataset for Open-domain Document Visual Question Answering
Le Qi, Shangwen Lv, Hongyu Li, Jing Liu, Yu Zhang, Qiaoqiao She, Hua Wu, Haifeng Wang and Ting Liu ................................................................. 1338

Coloring the Blank Slate: Pre-training Imparts a Hierarchical Inductive Bias to Sequence-to-sequence Models
Aaron Mueller, Robert Frank, Tal Linzen, Luheng Wang and Sebastian Schuster .... 1352

C³KG: A Chinese Commonsense Conversation Knowledge Graph
Dawei Li, Yanran Li, Jiayi Zhang, Ke Li, Chen Wei, Jianwei Cui and Bin Wang ....... 1369

Graph Neural Networks for Multiparallel Word Alignment
Ayyoob Imani, Lütfi Kerem Senel, Masoud Jalili Sabet, François Yvon and Hinrich Schuetze 1384

Sentiment Word Aware Multimodal Refinement for Multimodal Sentiment Analysis with ASR Errors
Yang Wu, Yanyan Zhao, Hao Yang, Song Chen, Bing Qin, Xiaohuan Cao and Wenting Zhao 1397

A Novel Framework Based on Medical Concept Driven Attention for Explainable Medical Code Prediction via External Knowledge
Tao Wang, Linhai Zhang, Chenchen Ye, Junxi Liu and Deyu Zhou ........ 1407

Effective Unsupervised Constrained Text Generation based on Perturbed Masking
Yingwen Fu, Wenjie Ou, Zhou Yu and Yue Lin .................. 1417

Combining (Second-Order) Graph-Based and Headed-Span-Based Projective Dependency Parsing
Songlin Yang and Kewei Tu ........................................ 1428

End-to-End Speech Translation for Code Switched Speech
Orion Weller, Matthias Sperber, Telmo Pires, Hendra Setiawan, Christian Gollan, Dominic C Telaar and Matthias Paulik ........................................ 1435

A Transformational Biencoder with In-Domain Negative Sampling for Zero-Shot Entity Linking
Kai Sun, Richong Zhang, Samuel Mensah, Yongyi Mao and Xudong Liu .......... 1449

Finding the Dominant Winning Ticket in Pre-Trained Language Models
Zhuocheng Gong, Di He, Yelong Shen, Tie-Yan Liu, Weizhu Chen, Dongyan Zhao, Ji-Rong Wen and Rui Yan ........................................ 1459

Thai Nested Named Entity Recognition Corpus
Weerayut Buaphet, Can Udomcharoenchaikit, Peerat Limkonchotiwat, Attapol Rutherford and Sarana Nutanong ........................................ 1473

Two-Step Question Retrieval for Open-Domain QA
Yeon Seonwoo, Juhee Son, Jiho Jin, Sang-Woo Lee, Ji-Hoon Kim, Jung-Woo Ha and Alice Oh 1487

Semantically Distributed Robust Optimization for Vision-and-Language Inference
Tejas Gokhale, Abhishek Chaudhary, Pratyay Banerjee, Chitta Baral and Yezhou Yang ...... 1493

Learning from Missing Relations: Contrastive Learning with Commonsense Knowledge Graphs for Commonsense Inference
Yong-Ho Jung, Jun-Hyung Park, Joon-Young Choi, Mingyu Lee, Junho Kim, Kang-Min Kim and SangKeun Lee ........................................ 1514

xxxviii
Capture Human Disagreement Distributions by Calibrated Networks for Natural Language Inference
Yuxia Wang, Minghan Wang, Yimeng Chen, Shimin Tao, Jiaxin Guo, Chang Su, Min Zhang and Hao Yang ................................................................. 1524

Efficient, Uncertainty-based Moderation of Neural Networks Text Classifiers
Jakob Smedegaard Andersen and Walid Maalej ........................................................ 1536

Revisiting Automatic Evaluation of Extractive Summarization Task: Can We Do Better than ROUGE?
Mousumi Akter, Naman Bansal and Shubhra Kanti Karmaker ........................................... 1547

Open Vocabulary Extreme Classification Using Generative Models
Daniel Simig, Fabio Petroni, Pouya Yanki, Kashyap Popat, Christina Du, Sebastian Riedel and Majid Yazdani ................................................................. 1561

Decomposed Meta-Learning for Few-Shot Named Entity Recognition
Tingting Ma, Huiqiang Jiang, Qianhui Wu, Tiejun Zhao and Chin-Yew Lin ......................... 1584

TegTok: Augmenting Text Generation via Task-specific and Open-world Knowledge
Chao-Hong Tan, Jia-Chen Gu, Chongyang Tao, Zhen-Hua Ling, Can Xu, Huang Hu, Xiubo Geng and Daxin Jiang ................................................................. 1597

EmoCaps: Emotion Capsule based Model for Conversational Emotion Recognition
Zaijing Li, Fengxiao Tang, Ming Zhao and Yusen Zhu .................................................... 1610

Logic-Driven Context Extension and Data Augmentation for Logical Reasoning of Text
Siyuan Wang, Wanjun Zhong, Duyu Tang, Zhongyu Wei, Zhihao Fan, Daxin Jiang, Ming Zhou and Nan Duan ................................................................. 1619

Transfer Learning and Prediction Consistency for Detecting Offensive Spans of Text
Amir Pouran Ben Veyseh, Ning Xu, Quan Hung Tran, Varun Manjunatha, Franck Dernoncourt and Thien Huu Nguyen ................................................................. 1630

Learning Reasoning Patterns for Relational Triple Extraction with Mutual Generation of Text and Graph
Yubo Chen, Yunqi Zhang and Yongfeng Huang .......................................................... 1638

Document-Level Event Argument Extraction via Optimal Transport
Amir Pouran Ben Veyseh, Minh Van Nguyen, Franck Dernoncourt, Bonan Min and Thien Huu Nguyen ................................................................. 1648

N-Shot Learning for Augmenting Task-Oriented Dialogue State Tracking
Ibrahim Taha Aksu, Zhengyuan Liu, Min-Yen Kan and Nancy F. Chen ............................ 1659

Document-Level Relation Extraction with Adaptive Focal Loss and Knowledge Distillation
Qingyu Tan, Ruidan He, Lidong Bing and Hwee Tou Ng ............................................. 1672

Calibration of Machine Reading Systems at Scale
Shehzaad Zuzar Dhuliawala, Leonard Adolphs, Rajarshi Das and Mrinmaya Sachan ...... 1682

Towards Adversarially Robust Text Classifiers by Learning to Reweight Clean Examples
Jianhan Xu, Cenyuan Zhang, Xiaqing Zheng, Linyang Li, Cho-Jui Hsieh, Kai-Wei Chang and Xuanjing Huang ................................................................. 1694

Morphosyntactic Tagging with Pre-trained Language Models for Arabic and its Dialects
Go Inoue, Salam Khalifa and Nizar Habash .......................................................... 1708
How Pre-trained Language Models Capture Factual Knowledge? A Causal-Inspired Analysis
Shaobo Li, Xiaoguang Li, Lifeng Shang, Zhenhua Dong, Chengjie Sun, Bingshan Liu, Zhenzhou Ji, Xin Jiang and Qun Liu ................................................................. 1720

Simran Arora, Sen Wu, Enci Liu and Christopher Re ........................................... 1733

Enhancing Natural Language Representation with Large-Scale Out-of-Domain Commonsense
Wanyun Cui and Xingran Chen ................................................................. 1746

Weighted self Distillation for Chinese word segmentation
Rian He, Shubin Cai, Zhong Ming and Jialei Zhang ........................................... 1757

Sibylvariant Transformations for Robust Text Classification
Fabrice Y Harel-Canada, Muhammad Ali Gulzar, Nanyun Peng and Miryung Kim .... 1771

DaLC: Domain Adaptation Learning Curve Prediction for Neural Machine Translation
Cheonbok Park, Hantae Kim, Ioan Calapodescu, Hyun Chang Cho and Vassilina Nikoulina 1789

Hey AI, Can You Solve Complex Tasks by Talking to Agents?
Tushar Khot, Kyle Richardson, Daniel Khashabi and Ashish Sabharwal ............... 1808

Modality-specific Learning Rates for Effective Multimodal Additive Late-fusion
Yiqun Yao and Rada Mihalcea ................................................................. 1824

BiSyn-GAT+: Bi-Syntax Aware Graph Attention Network for Aspect-based Sentiment Analysis
Shuo Liang, Wei Wei, Xian-Ling Mao, Fei Wang and Zhiyong He .......................... 1835

IndicBART: A Pre-trained Model for Indic Natural Language Generation
Raj Dabre, Himani Shrotriya, Anoop Kunchukuttan, Ratish Puduppully, Mitesh M Khapra and Pratyush Kumar ................................................................. 1849

Sentence-T5: Scalable Sentence Encoders from Pre-trained Text-to-Text Models
Jianmo Ni, Gustavo Hernandez Abrego, Noah Constant, Ji Ma, Keith Hall, Daniel Cer and Yinfei Yang ................................................................. 1864

Improving Relation Extraction through Syntax-induced Pre-training with Dependency Masking
Yuanhe Tian, Yan Song and Fei Xia ............................................................. 1875

Striking a Balance: Alleviating Inconsistency in Pre-trained Models for Symmetric Classification Tasks
Ashutosh Kumar and Aditya Joshi ............................................................. 1887

Diversifying Content Generation for Commonsense Reasoning with Mixture of Knowledge Graph Experts
Wenhao Yu, Chenguang Zhu, Lianhui Qin, Zhihan Zhang, Tong Zhao and Meng Jiang ..... 1896

Dict-BERT: Enhancing Language Model Pre-training with Dictionary
Wenhao Yu, Chenguang Zhu, Yuwei Fang, Donghan Yu, Shuohang Wang, Yichong Xu, Michael Zeng and Meng Jiang ................................................................. 1907

A Feasibility Study of Answer-Unaware Question Generation for Education
Liam Dugan, Eleni Miltsakaki, Shriyash Kaustubh Upadhyay, Etan Jacob Ginsberg, Hannah Gonzalez, DaHyeon Choi, Chunhyung Yuan and Chris Callison-Burch ................................................................. 1919

Relevant CommonSense Subgraphs for What if... Procedural Reasoning
Chen Zheng and Parisa Kordjamshidi ........................................................... 1927
Combining Feature and Instance Attribution to Detect Artifacts
Pouya Pezeshkpour, Sarthak Jain, Sameer Singh and Byron C Wallace .......................... 1934

Leveraging Expert Guided Adversarial Augmentation For Improving Generalization in Named Entity Recognition
Aaron Reich, Jiaao Chen, Aastha Agrawal, Yanzhe Zhang and Diyi Yang ......................... 1947

Label Semantics for Few Shot Named Entity Recognition
Jie Ma, Miguel Ballesteros, Srikanth Doss, Rishita Anubhai, Sunil Mallya, Yaser Al-Onaizan and Dan Roth ............................................................... 1956

Detection, Disambiguation, Re-ranking: Autoregressive Entity Linking as a Multi-Task Problem
Khalil Mrini, Shaoliang Nie, Jiatao Gu, Sinong Wang, Maziar Sanjabi and Hamed Firooz . 1972

VISITRON: Visual Semantics-Aligned Interactively Trained Object-Navigator
Ayush Shrivastava, Karthik Gopalakrishnan, Yang Liu, Robinson Piramuthu, Gokhan Tur, Devi Parikh and Dilek Hakkani-Tur .................................................................. 1984

Investigating Selective Prediction Approaches Across Several Tasks in IID, OOD, and Adversarial Settings
Neeraj Varshney, Swaroop Mishra and Chitta Baral ......................................................... 1995

Unsupervised Natural Language Inference Using PHL Triplet Generation
Neeraj Varshney, Pratyay Banerjee, Tejas Gokhale and Chitta Baral ............................... 2003

Data Augmentation and Learned Layer Aggregation for Improved Multilingual Language Understanding in Dialogue
Evgeniia Razumovskiaia, Ivan Vulić and Anna Korhonen .................................................. 2017

Ranking-Constrained Learning with Rationales for Text Classification
Juanyan Wang, Manali Sharma and Mustafa Bilgic ......................................................... 2034

CaM-Gen: Causally Aware Metric-Guided Text Generation
Navita Goyal, Roodram Paneri, Ayush Agarwal, Udit Kalani, Abhilasha Sancheti and Niyati Chhaya .......................................................... 2047

Training Dynamics for Text Summarization Models
Tanya Goyal, Jiacheng Xu, Junyi Jessy Li and Greg Durrett ............................................ 2061

Richer Countries and Richer Representations
Kaitlyn Zhou, Kawin Ethayarajh and Dan Jurafsky ...................................................... 2074

BBQ: A hand-built bias benchmark for question answering
Alicia Parrish, Angelica Chen, Nikita Nangia, Vishakh Padmakumar, Jason Phang, Jana Thompson, Phu Mon Htut and Samuel R. Bowman ............................................. 2086

Zero-shot Learning for Grapheme to Phoneme Conversion with Language Ensemble
Xinjian Li, Florian Metze, David R Mortensen, Shinji Watanabe and Alan Black ............... 2106

Dim Wihl Gat Tun: The Case for Linguistic Expertise in NLP for Under-Documented Languages
Clarissa Forbes, Farhan Samir, Bruce Harold Oliver, Changbing Yang, Edith Coates, Garrett Nicolai and Miikka Silfverberg ...................................................... 2116

Question Generation for Reading Comprehension Assessment by Modeling How and What to Ask
Bilal Ghanem, Lauren Lutz Coleman, Julia Rivard Dexter, Spencer McIntosh von der Ohe and Alona Fyshe ................................................................. 2131
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABi: Type-Aware Bi-Encoders for Open-Domain Entity Retrieval</td>
<td>Megan Leszczynski, Daniel Y Fu, Mayee F Chen and Christopher Re</td>
<td>2147</td>
</tr>
<tr>
<td>Hierarchical Recurrent Aggregative Generation for Few-Shot NLG</td>
<td>Giulio Zhou, Gerasimos Lampouras and Ignacio Iacobacci</td>
<td>2167</td>
</tr>
<tr>
<td>Training Text-to-Text Transformers with Privacy Guarantees</td>
<td>Natalia Ponomareva, Jasmijn Bastings and Sergei Vassilvitskii</td>
<td>2182</td>
</tr>
<tr>
<td>Revisiting Uncertainty-based Query Strategies for Active Learning with Transformers</td>
<td>Christopher Schröder, Andreas Niekler and Martin Potthast</td>
<td>2194</td>
</tr>
<tr>
<td>The impact of lexical and grammatical processing on generating code from natural language</td>
<td>Nathanaël Beau and Benoit Crabbé</td>
<td>2204</td>
</tr>
<tr>
<td>Seq2Path: Generating Sentiment Tuples as Paths of a Tree</td>
<td>Yue Mao, Yi Shen, Jingchao Yang, Xiaoying Zhu and Longjun Cai</td>
<td>2215</td>
</tr>
<tr>
<td>Mitigating the Inconsistency Between Word Saliency and Model Confidence with Pathological Contrastive Training</td>
<td>Pengwei Zhan, Yang Wu, Shaolei Zhou, Yunjian Zhang and Liming Wang</td>
<td>2226</td>
</tr>
<tr>
<td>Your fairness may vary: Pretrained language model fairness in toxic text classification</td>
<td>Ioana Baldini, Dennis Wei, Karthikeyan Natesan Ramamurthy, Moninder Singh and Mikhail Yurochkin</td>
<td>2245</td>
</tr>
<tr>
<td>ChartQA: A Benchmark for Question Answering about Charts with Visual and Logical Reasoning</td>
<td>Ahmed Masry, Do Xuan Long, Jia Qing Tan, Shafiq Joty and Enamul Hoque</td>
<td>2263</td>
</tr>
<tr>
<td>A Novel Perspective to Look At Attention: Bi-level Attention-based Explainable Topic Modeling for News Classification</td>
<td>Dairui Liu, Derek Greene and Ruihai Dong</td>
<td>2280</td>
</tr>
<tr>
<td>Learn and Review: Enhancing Continual Named Entity Recognition via Reviewing Synthetic Samples</td>
<td>Yu Xia, Quan Wang, Yajuan Lyu, Yong Zhu, Wenhao Wu, Sujian Li and Dai Dai</td>
<td>2291</td>
</tr>
<tr>
<td>Phoneme transcription of endangered languages: an evaluation of recent ASR architectures in the single speaker scenario</td>
<td>Gilles Boulianne</td>
<td>2301</td>
</tr>
<tr>
<td>Does BERT really agree? Fine-grained Analysis of Lexical Dependence on a Syntactic Task</td>
<td>Karim Lasri, Alessandro Lenci and Thierry Poibeau</td>
<td>2309</td>
</tr>
<tr>
<td>Combining Static and Contextualised Multilingual Embeddings</td>
<td>Katharina Hämmerl, Jindřich Libovický and Alexander Fraser</td>
<td>2316</td>
</tr>
<tr>
<td>An Accurate Unsupervised Method for Joint Entity Alignment and Dangling Entity Detection</td>
<td>Shengxuan Luo and Sheng Yu</td>
<td>2330</td>
</tr>
<tr>
<td>Square One Bias in NLP: Towards a Multi-Dimensional Exploration of the Research Manifold</td>
<td>Sebastian Ruder, Ivan Vulić and Anders Søgaard</td>
<td>2340</td>
</tr>
<tr>
<td>Systematicity, Compositionality and Transitivity of Deep NLP Models: a Metamorphic Testing Perspective</td>
<td>Edoardo Manino, Julia Rozanova, Danilo Carvalho, Andre Freitas and Lucas Carvalho Cordeiro</td>
<td>2355</td>
</tr>
</tbody>
</table>
Improving Neural Political Statement Classification with Class Hierarchical Information
Erenay Dayanik, Andre Blessing, Nico Blokker, Sebastian Haunss, Jonas Kuhn, Gabriella Lapesa and Sebastian Pado .......................................................... 2367

Enabling Multimodal Generation on CLIP via Vision-Language Knowledge Distillation
Wenliang Dai, Lu Hou, Lifeng Shang, Xin Jiang, Qun Liu and Pascale Fung ............... 2383

Co-VQA: Answering by Interactive Sub Question Sequence
Ruonan Wang, Yuxi Qian, Fangxiang Feng, Xiaojie Wang and Huixing Jiang ................ 2396

A Simple Hash-Based Early Exiting Approach For Language Understanding and Generation
Tianxiang Sun, Xiangyang Liu, Wei Zhu, Zhichao Geng, Lingling Wu, Yilong He, Yuan Ni, Guotong Xie, Xuanjing Huang and Xipeng Qiu ............................................. 2409

Auxiliary tasks to boost Biaffine Semantic Dependency Parsing
Marie Candito ................................................................................................. 2422

Syntax-guided Contrastive Learning for Pre-trained Language Model
Shuai Zhang, Wang Lijie, Xinyan Xiao and Hua Wu ........................................ 2430

Improved Multi-label Classification under Temporal Concept Drift: Rethinking Group-Robust Algorithms in a Label-Wise Setting
Ilias Chalkidis and Anders Søgaard ............................................................. 2441

ASCM: An Answer Space Clustered Prompting Method without Answer Engineering
Zhen Wang, Yating Yang, Zhou Xi, Bo Ma, Lei Wang, Rui Dong and Azmat Anwar ...... 2455

Why don’t people use character-level machine translation?
Jindřich Libovický, Helmut Schmid and Alexander Fraser .................................. 2470

Seeking Patterns, Not just Memorizing Procedures: Contrastive Learning for Solving Math Word Problems
Zhongli Li, Wenxuan Zhang, Chao Yan, Qingyu Zhou, Chao Li, Hongzhi Liu and Yunbo Cao 2486

xGQA: Cross-Lingual Visual Question Answering
Jonas Pfeiffer, Gregor Geigle, Aishwarya Kamath, Jan-Martin O. Steitz, Stefan Roth, Ivan Vulić and Iryna Gurevych ............................................................... 2497

Automatic Speech Recognition and Query By Example for Creole Languages Documentation
Cécile Macaire, Didier Schwab, Benjamin Lecouteux and Emmanuel Schang .......... 2512

MReD: A Meta-Review Dataset for Structure-Controllable Text Generation
Chenhui Shen, Liying Cheng, Ran Zhou, Lidong Bing, Yang You and Luo Si .......... 2521

Sho Takase, Tatsuya Hiroyuki and Naoaki Okazaki ........................................... 2536

Detecting Various Types of Noise for Neural Machine Translation
Christian Herold, Jan Rosendahl, Joris Vanwinckelenre and Hermann Ney ............ 2542

DU-VLG: Unifying Vision-and-Language Generation via Dual Sequence-to-Sequence Pre-training
Luyang Huang, Guoqiang Niu, Jiachen Liu, Xinyan Xiao and Hua Wu .................... 2552

HiCLRE: A Hierarchical Contrastive Learning Framework for Distantly Supervised Relation Extraction
Dongyang Li, Taolin Zhang, Nan Hu, Chengyu Wang and Xiaofeng He .................. 2567
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prompt-Driven Neural Machine Translation</td>
<td>Yafu Li, Yongjing Yin, Jing Li and Yue Zhang</td>
<td>2579</td>
</tr>
<tr>
<td>On Controlling Fallback Responses for Grounded Dialogue Generation</td>
<td>Hongyuan Lu, Wai Lam, Hong Cheng and Helen M. Meng</td>
<td>2591</td>
</tr>
<tr>
<td>CRAFT: A Benchmark for Causal Reasoning About Forces and Interactions</td>
<td>Tayfun Ates, M. Şamil Ateşoğlu, Çağatay Yiğit, Ilker Keser, Mert Kobas, Erkut Erdem, Aykut Erdem, Tilbe Goksun and Deniz Yuret</td>
<td>2602</td>
</tr>
<tr>
<td>A Graph Enhanced BERT Model for Event Prediction</td>
<td>Li Du, Xiao Ding, Yue Zhang, Ting Liu and Bing Qin</td>
<td>2628</td>
</tr>
<tr>
<td>Lacking the Embedding of a Word? Look it up into a Traditional Dictionary</td>
<td>Elena Sofia Ruzzetti, Leonardo Ranaldi, Michele Mastromattei, Francesca Fallucchi, Noemi Scarpato and Fabio Massimo Zanzotto</td>
<td>2651</td>
</tr>
<tr>
<td>MTRec: Multi-Task Learning over BERT for News Recommendation</td>
<td>Qiwei Bi, Jian Li, Lifeng Shang, Xin Jiang, Qun Liu and Hanfang Yang</td>
<td>2663</td>
</tr>
<tr>
<td>Cross-domain Named Entity Recognition via Graph Matching</td>
<td>Junhao Zheng, Haibin Chen and Qianli Ma</td>
<td>2670</td>
</tr>
<tr>
<td>Assessing Multilingual Fairness in Pre-trained Multimodal Representations</td>
<td>Jialu Wang, Yang Liu and Xin Eric Wang</td>
<td>2681</td>
</tr>
<tr>
<td>More Than Words: Collocation Retokenization for Latent Dirichlet Allocation Models</td>
<td>Jin Cheevaprawatdomrong, Alexandra Schofield and Attapol Rutherford</td>
<td>2696</td>
</tr>
<tr>
<td>Generalized but not Robust? Comparing the Effects of Data Modification Methods on Out-of-Domain Generalization and Adversarial Robustness</td>
<td>Tejas Gokhale, Swaroop Mishra, Man Luo, Bhavdeep Singh Sachdeva and Chitta Baral</td>
<td>2705</td>
</tr>
<tr>
<td>ASSIST: Towards Label Noise-Robust Dialogue State Tracking</td>
<td>Fanghua Ye, Yue Feng and Emine Yilmaz</td>
<td>2719</td>
</tr>
<tr>
<td>Graph Refinement for Coreference Resolution</td>
<td>Lesly Miculicich and James Henderson</td>
<td>2732</td>
</tr>
<tr>
<td>ECO v1: Towards Event-Centric Opinion Mining</td>
<td>Ruoxi Xu, Hongyu Lin, Meng Liao, Xianpei Han, Jin Xu, Wei Tan, Yingfei Sun and Le Sun</td>
<td>2743</td>
</tr>
<tr>
<td>Deep Reinforcement Learning for Entity Alignment</td>
<td>Lingbing Guo, Yuqiang Han, Qiang Zhang and Huajun Chen</td>
<td>2754</td>
</tr>
<tr>
<td>Breaking Down Multilingual Machine Translation</td>
<td>Ting-Rui Chiang, Yi-Pei Chen, Yi-Ting Yeh and Graham Neubig</td>
<td>2766</td>
</tr>
<tr>
<td>Mitigating Contradictions in Dialogue Based on Contrastive Learning</td>
<td>Weizhao Li, Junsheng Kong, Ben Liao and Yi Cai</td>
<td>2781</td>
</tr>
<tr>
<td>ELLE: Efficient Lifelong Pre-training for Emerging Data</td>
<td>Yujia Qin, Jiajie Zhang, Yankai Lin, Zhiyuan Liu, Peng Li, Maosong Sun and Jie Zhou</td>
<td>2789</td>
</tr>
</tbody>
</table>
EnCBP: A New Benchmark Dataset for Finer-Grained Cultural Background Prediction in English
Weicheng Ma, Samiha Datta, Lili Wang and Soroush Vosoughi ............................................. 2811

Cutting Down on Prompts and Parameters: Simple Few-Shot Learning with Language Models
Robert L. Logan IV, Ivana Balazevic, Eric Wallace, Fabio Petroni, Sameer Singh and Sebastian Riedel ................................................................. 2824

uFACT: Unfaithful Alien-Corpora Training for Semantically Consistent Data-to-Text Generation
Tisha Anders, Alexandru Coca and Bill Byrne ................................................................. 2836

Good Night at 4 pm?! Time Expressions in Different Cultures
Vered Shwartz ................................................................. 2842

Extracting Person Names from User Generated Text: Named-Entity Recognition for Combating Human Trafficking
Yifei Li, Pratheeksha Nair, Kellin Pelrine and Reiheaneh Rabbany ........................................... 2854

OneAligner: Zero-shot Cross-lingual Transfer with One Rich-Resource Language Pair for Low-Resource Sentence Retrieval
Tong Niu, Kazuma Hashimoto, Yingbo Zhou and Caiming Xiong ........................................... 2869

Suum Cuique: Studying Bias in Taboo Detection with a Community Perspective
Osama Khalid, Jonathan Rusert and Padmini Srinivasan ......................................................... 2883

Modeling Intensification for Sign Language Generation: A Computational Approach
Mert Inan, Yang Zhong, Sabit Hassan, Lorna Quandt and Malihe Alikhani .................................. 2897

Controllable Natural Language Generation with Contrastive Prefixes
Jing Qian, Li Dong, Yelong Shen, Furu Wei and Weizhu Chen .............................................. 2912

Revisiting the Effects of Leakage on Dependency Parsing
Nathaniel Krasner, Miriam Wanner and Antonios Anastasopoulos ........................................... 2925

Learning to Describe Solutions for Bug Reports Based on Developer Discussions
Sheena Panthaplackel, Junyi Jessy Li, Milos Gligoric and Ray Mooney ...................................... 2935

Perturbations in the Wild: Leveraging Human-Written Text Perturbations for Realistic Adversarial Attack and Defense
Thai Le, Jooyoung Lee, Kevin Yen, Yifan Hu and Dongwon Lee ............................................. 2953

Improving Chinese Grammatical Error Detection via Data augmentation by Conditional Error Generation
Tianchi Yue, Shulin Liu, Huihui Cai, Tao Yang, Shengkang Song and TingHao YU ................... 2966

Modular and Parameter-Efficient Multimodal Fusion with Prompting
Sheng Liang, Mengjie Zhao and Hinrich Schuetze .............................................................. 2976

Synchronous Refinement for Neural Machine Translation
Kehai Chen, Masao Utiyama, Eiichiro Sumita, Rui Wang and Min Zhang .................................. 2986

HIE-SQL: History Information Enhanced Network for Context-Dependent Text-to-SQL Semantic Parsing
Yanzhao Zheng, Haibin Wang, Baohua Dong, Xingjun Wang and Changshan Li ....................... 2997

CRASpell: A Contextual Typo Robust Approach to Improve Chinese Spelling Correction
Shulin Liu, Shengkang Song, Tianchi Yue, Tao Yang, Huihui Cai, TingHao YU and Shengli Sun 3008
Gaussian Multi-head Attention for Simultaneous Machine Translation
Shaolei Zhang and Yang Feng .............................................................. 3019

Composing Structure-Aware Batches for Pairwise Sentence Classification
Andreas Waldis, Tilman Beck and Iryna Gurevych ................................. 3031

Factual Consistency of Multilingual Pretrained Language Models
Constanza Fierro and Anders Søgaard .................................................. 3046

Selecting Stickers in Open-Domain Dialogue through Multitask Learning
Zhexin Zhang, Yeshuang Zhu, Zhengcong Fei, Jinchao Zhang and Jie Zhou .................................................. 3053

ZiNet: Linking Chinese Characters Spanning Three Thousand Years
Yang Chi, Fausto Giunchiglia, Daqian Shi, Xiaolei Diao, Chuntao Li and Hao Xu .......... 3061

How Can Cross-lingual Knowledge Contribute Better to Fine-grained Entity Typing?
Hailong Jin, Tiansi Dong, Lei Hou, Juanzi Li, Hui Chen, Zelin Dai and Qu Yincen .......... 3071

AMR-DA: Data Augmentation by Abstract Meaning Representation
Ziyi Shou, Yuxin Jiang and Fangzhen Lin ............................................... 3082

Using Pre-Trained Language Models for Producing Counter Narratives Against Hate Speech: a Comparative Study
Serra Sinem Tekiroğlu, Helena Bonaldi, Margherita Fanton and Marco Guerini .......... 3099

Improving Robustness of Language Models from a Geometry-aware Perspective
Bin Zhu, Zhaoyuan Gu, Le Wang, Jinyin Chen and Qi Xuan .............................. 3115

Task-guided Disentangled Tuning for Pretrained Language Models
Jiali Zeng, Yufan Jiang, Shuangzhi Wu, Yongjing Yin and Mu Li .................... 3126

Exploring the Impact of Negative Samples of Contrastive Learning: A Case Study of Sentence Embedding
Rui Cao, Yihao Wang, Yuxin Liang, Ling Gao, Jie Zheng, Jie Ren and Zheng Wang .......... 3138

The Inefficiency of Language Models in Scholarly Retrieval: An Experimental Walk-through
Shruti Singh and Mayank Singh .......................................................... 3153

Fusing Heterogeneous Factors with Triaffine Mechanism for Nested Named Entity Recognition
Zheng Yuan, Chuanqi Tan, Songfang Huang and Fei Huang .......................... 3174

UNIMO-2: End-to-End Unified Vision-Language Grounded Learning
Wei Li, Can Gao, Guocheng Niu, Xinyan Xiao, Hao Liu, Jiachen Liu, Hua Wu and Haifeng Wang 3187

The Past Mistake is the Future Wisdom: Error-driven Contrastive Probability Optimization for Chinese Spell Checking
Yinghui Li, Qingyu Zhou, Yangning Li, Zhongli Li, Ruiyang Liu, Rongyi Sun, Zizhen Wang, Chao Li, Yunbo Cao and Hai-Tao Zheng ................................................. 3202

XFUND: A Benchmark Dataset for Multilingual Visually Rich Form Understanding
Yiheng Xu, Tengchao Lv, Lei Cui, Guoxin Wang, Yijuan Lu, Dinei Florencio, Cha Zhang and Furu Wei .......................................................... 3214

Type-Driven Multi-Turn Corrections for Grammatical Error Correction
Shaopeng Lai, Qingyu Zhou, Jiali Zeng, Zhongli Li, Chao Li, Yunbo Cao and Jinsong Su .. 3225
Leveraging Knowledge in Multilingual Commonsense Reasoning
Yuwei Fang, Shuohang Wang, Yichong Xu, Ruochen Xu, Siqi Sun, Chenguang Zhu and Michael Zeng ................................................................. 3237

Encoding and Fusing Semantic Connection and Linguistic Evidence for Implicit Discourse Relation Recognition
Wei Xiang, Bang Wang, Lu Dai and Yijun Mo ........................................ 3247

One Agent To Rule Them All: Towards Multi-agent Conversational AI
Christopher Clarke, Joseph J Peper, Karthik Krishnamurthy, Walter Talamonti, Kevin Leach, Walter Lasecki, Yiping Kang, Lingjia Tang and Jason Mars ........................................ 3258

Word-level Perturbation Considering Word Length and Compositional Subwords
Tatsuya Hiraoka, Sho Takase, Kei Uchiumi, Atsushi Keyaki and Naoaki Okazaki ........ 3268

Bridging Pre-trained Language Models and Hand-crafted Features for Unsupervised POS Tagging
Houquan Zhou, Yang Li, Zhenghua Li and Min Zhang ..................................... 3276

Controlling the Focus of Pretrained Language Generation Models
Jiabao Ji, Yoon Kim, James R. Glass and Tianxing He ..................................... 3291

Comparative Opinion Summarization via Collaborative Decoding
Hayate Iso, Xiaolan Wang, Stefanos Angelidis and Yoshihiko Suhara .............. 3307

IsoScore: Measuring the Uniformity of Embedding Space Utilization
William Rudman, Nate Gillman, Taylor Rayne and Carsten Eickhoff .......... 3325

A Natural Diet: Towards Improving Naturalness of Machine Translation Output
Markus Freitag, David Vilar, David Grangier, Colin Cherry and George Foster ........ 3340

From Stance to Concern: Adaptation of Propositional Analysis to New Tasks and Domains
Brodie Mather, Bonnie J Dorr, Adam Dalton, William de Beaumont, Owen Rambow and Sonja M. Schmer-Galunder .......... 3354

CUE Vectors: Modular Training of Language Models Conditioned on Diverse Contextual Signals
Scott Novotney, Sreeparna Mukherjee, Zeeshan Ahmed and Andreas Stolcke ........ 3368

Cross-Lingual UMLS Named Entity Linking using UMLS Dictionary Fine-Tuning
Rina Galperin, Shachar Schnapp and Michael Elhadad ................................ 3380

Aligned Weight Regularizers for Pruning Pretrained Neural Networks
James O’Neill, Sourav Dutta and Haytham Assem ...................................... 3391

Consistent Representation Learning for Continual Relation Extraction
Kang Zhao, Hua Xu, Jiangong Yang and Kai Gao ...................................... 3402

Event Transition Planning for Open-ended Text Generation
Qintong Li, Piji Li, Wei Bi, Zhaochun Ren, Yuxuan Lai and Lingpeng Kong .......... 3412

Comprehensive Multi-Modal Interactions for Referring Image Segmentation
Kanishk Jain and Vineet Gandhi ............................................................. 3427

MetaWeighting: Learning to Weight Tasks in Multi-Task Learning
Yuren Mao, Zekai Wang, Weiwei Liu, Xuemin Lin and Pengtao Xie .................. 3436

Improving Controllable Text Generation with Position-Aware Weighted Decoding
Yuxuan Gu, Xiaocheng Feng, Sicheng Ma, Jiaming Wu, Heng Gong and Bing Qin .... 3449
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prompt Tuning for Discriminative Pre-trained Language Models</td>
<td>Yuan Yao, Bowen Dong, Ao Zhang, Zhengyan Zhang, Ruobing Xie, Zhiyuan Liu, Leyu Lin, Maosong Sun and Jianyong Wang</td>
<td>3468</td>
</tr>
<tr>
<td>Two Birds with One Stone: Unified Model Learning for Both Recall and Ranking in News Recommendation</td>
<td>Chuhan Wu, Fangzhao Wu, Tao Qi and Yongfeng Huang</td>
<td>3474</td>
</tr>
<tr>
<td>What does it take to bake a cake? The RecipeRef corpus and anaphora resolution in procedural text</td>
<td>Biaoyan Fang, Timothy Baldwin and Karin Verspoor</td>
<td>3481</td>
</tr>
<tr>
<td>MERI: Meta-Path Guided Contrastive Learning for Logical Reasoning</td>
<td>Fangkai Jiao, Yangyang Guo, Xuemeng Song and Liqiang Nie</td>
<td>3496</td>
</tr>
<tr>
<td>THE-X: Privacy-Preserving Transformer Inference with Homomorphic Encryption</td>
<td>Tianyu Chen, Hangbo Bao, Shaohan Huang, Li Dong, Binxing Jiao, Daxin Jiang, Haoyi Zhou, Jianxin Li and Furu Wei</td>
<td>3510</td>
</tr>
<tr>
<td>HLDC: Hindi Legal Documents Corpus</td>
<td>Arnav Kapoor, Mudit Dhawan, Anmol Goel, Arjun T H, Akshala Bhatnagar, Vibh Agrawal, Amul Agrawal, Arnab Bhattacharya, Ponnurangam Kumaraguru and Ashutosh Modi</td>
<td>3521</td>
</tr>
<tr>
<td>Rethinking Document-level Neural Machine Translation</td>
<td>Zewei Sun, Mingxuan Wang, Hao Zhou, Chengqi Zhao, Shujian Huang, Jiajun Chen and Lei Li</td>
<td>3537</td>
</tr>
<tr>
<td>Incremental Intent Detection for Medical Domain with Contrast Replay Networks</td>
<td>Guirong Bai, Shizhu He, Kang Liu and Jun Zhao</td>
<td>3549</td>
</tr>
<tr>
<td>LaPraDoR: Unsupervised Pretrained Dense Retriever for Zero-Shot Text Retrieval</td>
<td>Canwen Xu, Daya Guo, Nan Duan and Julian McAuley</td>
<td>3557</td>
</tr>
<tr>
<td>Do Pre-trained Models Benefit Knowledge Graph Completion? A Reliable Evaluation and a Reasonable Approach</td>
<td>Xin Lv, Yankai Lin, Yixin Cao, Lei Hou, Juanzi Li, Zhiyuan Liu, Peng Li and Jie Zhou</td>
<td>3570</td>
</tr>
<tr>
<td>EICO: Improving Few-Shot Text Classification via Explicit and Implicit Consistency Regularization</td>
<td>Lei Zhao and Cheng Yao</td>
<td>3582</td>
</tr>
<tr>
<td>Improving the Adversarial Robustness of NLP Models by Information Bottleneck</td>
<td>Cenyuan Zhang, Xiang Zhou, Yixin Wan, Xiaoqing Zheng, Kai-Wei Chang and Cho-Jui Hsieh</td>
<td>3588</td>
</tr>
<tr>
<td>Incorporating Dynamic Semantics into Pre-Trained Language Model for Aspect-based Sentiment Analysis</td>
<td>Kai Zhang, Kun Zhang, Mengdi Zhang, Hongke Zhao, Qi Liu, Wei Wu and Enhong Chen</td>
<td>3599</td>
</tr>
<tr>
<td>DARER: Dual-task Temporal Relational Recurrent Reasoning Network for Joint Dialog Sentiment Classification and Act Recognition</td>
<td>Bowen Xing and Ivor Tsang</td>
<td>3611</td>
</tr>
<tr>
<td>Divide and Conquer: Text Semantic Matching with Disentangled Keywords and Intents</td>
<td>Yicheng Zou, Hongwei Liu, Tao Gui, Junzhe Wang, Qi Zhang, Meng Tang, Haixiang Li and Daniell Wang</td>
<td>3622</td>
</tr>
<tr>
<td>Modular Domain Adaptation</td>
<td>Junshen K Chen, Dallas Card and Dan Jurafsky</td>
<td>3633</td>
</tr>
<tr>
<td>Title</td>
<td>Authors</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Detection of Adversarial Examples in Text Classification: Benchmark and Baseline via Robust Density Estimation</td>
<td>KiYoon Yoo, Jangho Kim, Jiho Jang and Nojun Kwak</td>
<td>3656</td>
</tr>
<tr>
<td>Platt-Bin: Efficient Posterior Calibrated Training for NLP Classifiers</td>
<td>Rishabh Singh and Shirin Goshtasbpour</td>
<td>3673</td>
</tr>
<tr>
<td>Improving Candidate Retrieval with Entity Profile Generation for Wikidata Entity Linking</td>
<td>Tuan Lai, Heng Ji and ChengXiang Zhai</td>
<td>3696</td>
</tr>
<tr>
<td>Local Structure Matters Most: Perturbation Study in NLU</td>
<td>Louis Clouatre, Prasanna Parthasarathi, Amal Zouaq and Sarath Chandar</td>
<td>3712</td>
</tr>
<tr>
<td>Probing Factually Grounded Content Transfer with Factual Ablation</td>
<td>Peter West, Chris Quirk, Michel Galley and Yejin Choi</td>
<td>3732</td>
</tr>
<tr>
<td>ED2LM: Encoder-Decoder to Language Model for Faster Document Re-ranking Inference</td>
<td>Kai Hui, Honglei Zhuang, Tao Chen, Zhen Qin, Jing Lu, Dara Bahri, Ji Ma, Jai Gupta, Cicero Nogueira dos Santos, Yi Tay and Donald Metzler</td>
<td>3747</td>
</tr>
<tr>
<td>Benchmarking Answer Verification Methods for Question Answering-Based Summarization Evaluation Metrics</td>
<td>Daniel Deutsch and Dan Roth</td>
<td>3759</td>
</tr>
<tr>
<td>Prior Knowledge and Memory Enriched Transformer for Sign Language Translation</td>
<td>Tao Jin, Zhou Zhao, Meng Zhang and Xingshan Zeng</td>
<td>3766</td>
</tr>
<tr>
<td>Discontinuous Constituency and BERT: A Case Study of Dutch</td>
<td>Konstantinos Kogkalidis and Gijs Wijnholds</td>
<td>3776</td>
</tr>
<tr>
<td>Probing Multilingual Cognate Prediction Models</td>
<td>Clémentine Fourrier and Benoît Sagot</td>
<td>3786</td>
</tr>
<tr>
<td>A Neural Pairwise Ranking Model for Readability Assessment</td>
<td>Justin Lee and Sowmya Vajjala</td>
<td>3802</td>
</tr>
<tr>
<td>First the Worst: Finding Better Gender Translations During Beam Search</td>
<td>Danielle Saunders, Rosie Sallis and Bill Byrne</td>
<td>3814</td>
</tr>
<tr>
<td>Unsupervised Preference-Aware Language Identification</td>
<td>Xingzhang Ren, Baosong Yang, Dayiheng Liu, Haibo Zhang, Xiaoyu Lv, Liang Yao and Jun Xie</td>
<td>3847</td>
</tr>
<tr>
<td>Using NLP to quantify the environmental cost and diversity benefits of in-person NLP conferences</td>
<td>Piotr Przybyla and Matthew Shardlow</td>
<td>3853</td>
</tr>
<tr>
<td>Interpretable Research Replication Prediction via Variational Contextual Consistency Sentence Masking</td>
<td>Tianyi Luo, Rui Meng, Xin Eric Wang and Yang Liu</td>
<td>3864</td>
</tr>
</tbody>
</table>
Structural Supervision for Word Alignment and Machine Translation
Lei Li, Kai Fan, Hongjia Li and Chun Yuan .......................................................... 4084

Focus on the Action: Learning to Highlight and Summarize Jointly for Email To-Do Items Summarization
Kexun Zhang, Jiaao Chen and Diyi Yang ............................................................ 4095

Exploring the Capacity of a Large-scale Masked Language Model to Recognize Grammatical Errors
Ryo Nagata, Manabu Kimura and Kazuaki Hanawa .......................................... 4107

Should We Trust This Summary? Bayesian Abstractive Summarization to The Rescue
Alexios Gidiotis and Grigorios Tsoumakas ...................................................... 4119

On the data requirements of probing
Zining Zhu, Jixuan Wang, Bai Li and Frank Rudzicz ........................................ 4132

Translation Error Detection as Rationale Extraction
Marina Fomicheva, Lucia Specia and Nikolaos Aletras ..................................... 4148

Towards Collaborative Neural-Symbolic Graph Semantic Parsing via Uncertainty
Zi Lin, Jeremiah Zhe Liu and Jingbo Shang ...................................................... 4160

Towards Few-shot Entity Recognition in Document Images: A Label-aware Sequence-to-Sequence Framework
Zilong Wang and Jingbo Shang ....................................................................... 4174

On Length Divergence Bias in Textual Matching Models
Lan Jiang, Tianshu Lyu, Yankai Lin, Meng Chong, Xiaoyong Lyu and Dawei Yin .... 4187

What is wrong with you?: Leveraging User Sentiment for Automatic Dialog Evaluation
Sarik Ghazarian, Behnam Hedayatnia, Alexandros Papangelis, Yang Liu and Dilek Hakkani-Tur 4194