W-NUT 2020

The Sixth Workshop on Noisy User-generated Text (W-NUT 2020)

Proceedings of the Workshop

Nov 19, 2020
Online

©2020 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-952148-76-7

Introduction

The W-NUT 2020 workshop focuses on a core set of natural language processing tasks on top of noisy user-generated text, such as that found on social media, web forums and online reviews. Recent years have seen a significant increase of interest in these areas. The internet has democratized content creation leading to an explosion of informal user-generated text, publicly available in electronic format, motivating the need for NLP on noisy text to enable new data analytics applications.

This year, in addition to the main workshop track, we have three shared tasks: (1) Entity and relation recognition over wet-lab protocols, (2) Identification of informative COVID-19 English Tweets, and (3) COVID-19 Event Extraction from Twitter. We accepted 33 regular workshop papers and 47 shared-task papers. The workshop will be held online and live in two different time zones (GMT– and GMT++). There are two invited speakers for each time zone, Eduardo Blanco (University of North Texas) and Manaal Faruqui (Google) in time zone GMT– and Robert Munro (Machine Learning Consulting; former CTO of Figure Eight) and Irwin King (The Chinese University of Hong Kong) in time zone GMT++ with each of their talks covering a different aspect of NLP for user-generated text. We have the best paper award(s) sponsored by Twitter this year, for which we are thankful. We would like to thank the Program Committee members who reviewed the papers and the shared task organizers who enriched our workshop this year. We would also like to thank the workshop participants.

Wei Xu, Alan Ritter, Tim Baldwin and Afshin Rahimi
Co-Organizers

Organizers:

Wei Xu, Ohio State University
Alan Ritter, Ohio State University
Tim Baldwin, University of Melbourne
Afshin Rahimi, University of Melbourne

Program Committee:

Muhammad Abdul-Mageed (University of British Columbia)
Željko Agić (Corti)
Sweta Agrawal (University of Maryland)
Gustavo Aguilar (University of Houston)
Nikolaos Aletras (University of Sheffield)
Rahul Aralikatte (University of Copenhagen)
Eiji Aramaki (NAIST)
JinYeong Bak (Sungkyunkwan University)
Francesco Barbieri (Universitat Pompeu Fabra)
John Beieler (ODNI Science and Technology)
Eric Bell (PNNL)
Anya Belz (University of Brighton)
Adrian Benton (JHU)
Eduardo Blanco (University of North Texas)
Su Lin Blodgett (UMass Amherst)
Julian Brooke (University of British Columbia)
Cornelia Caragea (University of Illinois at Chicago)
Tuhin Chakrabarty (Columbia University)
Stevie Chancellor (Northwestern University)
Mingda Chen (Toyota Technological Institute at Chicago)
Sihao Chen (University of Pennsylvania)
Dhivya Chinnappa (Thomson Reuters)
Colin Cherry (Google)
Zewei Chu (University of Chicago)
Manuel R. Ciosici (IT University of Copenhagen)
Oana Cocarascu (Imperial College London)
Nigel Collier (University of Cambridge)
Çağrı Çöltekin (University of Tübingen)
Paul Cook (University of New Brunswick)
Marina Danilevsky (IBM Research)
Pradipto Das (Rakuten Institute of Technology)
Leon Derczynski (IT University of Copenhagen)
Jay DeYoung (Northeastern University)
Bhuwan Dhingra (Carnegie Mellon University)
Seza Doğruöz (Tilburg University)
Xinya Du (Cornell University)
Jacob Eisenstein (Google)
Heba Elfardy (Amazon)
Micha Elsner (Ohio State University)
Alexander Fabbri (Yale University)

Manaal Faruqui (Google)
Yansong Feng (Peking University)
Catherine Finegan-Dollak (IBM Research)
Tim Finin (UMBC)
Lucie Flek (Mainz University of Applied Sciences)
Lisheng Fu (New York University)
Yoshinari Fujinuma (University of Colorado, Boulder)
Juri Ganitkevitch (Google)
Dan Garrette (Google)
Sahil Garg (University of Southern California)
Spandana Gella (Amazon)
Debanjan Ghosh (MIT)
Kevin Gimpel (TTIC)
Amit Goyal (Amazon)
Yvette Graham (Dublin City University)
Chulaka Gunasekara (IBM Research)
Mika Hämäläinen (University of Helsinki)
William L. Hamilton (McGill University/MILA)
Xiaochuang Han (Carnegie Mellon University)
Devamanyu Hazarika (National University of Singapore)
Hua He (Amazon)
Jack Hessel (Cornell University)
Graeme Hirst (University of Toronto)
Nathan Hodas (PNNL)
Junjie Hu (Carnegie Mellon University)
Dirk Hovy (Bocconi University)
Binxuan Huang (Carnegie Mellon University)
Sarthak Jain (Northeastern University)
Nanjiang Jiang (Ohio State University)
Lifeng Jin (Ohio State University)
Ishan Jindal (IBM Research)
Kristen Johnson (Michigan State University)
Kenny Joseph (University at Buffalo)
Katharina Kann (University of Colorado, Boulder)
David Kauchak (Pomona College)
Ashique KhudaBukhsh (Carnegie Mellon University)
Roman Klinger (University of Stuttgart)
Hayato Kobayashi (Yahoo! Research)
Ekaterina Kochmar (University of Cambridge)
Reno Kriz (University of Pennsylvania)
Sachin Kumar (Carnegie Mellon University)
Vivek Kulkarni (Stanford University)
Jonathan Kummerfeld (University of Michigan)
Ophélie Lacroix (Siteimprove)
Wuwei Lan (Ohio State University)
Jiwei Li (ShannonAI)
Jessy Junyi Li (University of Texas Austin)
Jing Li (Hong Kong Polytechnic University)
Yitong Li (University of Melbourne)
Nut Limsopatham (University of Glasgow)
Zhiyuan Liu (Tsinghua University)

Fei Liu (University of Melbourne)
Nikola Ljubešić (Jožef Stefan Institute)
Wei-Yun Ma (Academia Sinica)
Mounica Maddela (Georgia Institute of Technology)
Peter Makarov (University of Zurich)
Héctor Martínez Alonso (Apple)
Aaron Masino (The Children's Hospital of Philadelphia)
Nitika Mathur (University of Melbourne)
Ahmed Mourad (RMIT University)
Yasuhide Miura (Fuji Xerox)
Hamdy Mubarak (Qatar Computing Research Institute)
Graham Mueller (Leidos)
Maria Nadejde (Grammarly)
Guenter Neumann (German Research Center for Artificial Intelligence)
Vincent Ng (University of Texas at Dallas)
Thien Huu Nguyen (University of Oregon)
Eric Nichols (Honda Research Institute)
Tong Niu (University of North Carolina at Chapel-Hill)
Benjamin Nye (Northeastern University)
Alice Oh (KAIST)
Naoaki Okazaki (Tohoku University)
Naoki Otani (CMU)
Myle Ott (Facebook AI)
Symeon Papadopoulos (CERTH-ITI)
Umashanthi Pavalanathan (Georgia Tech)
Yuval Pinter (Georgia Tech)
Christopher Potts (Stanford University)
Vinodkumar Prabhakaran (Stanford University)
Daniel Preoţiuc-Pietro (Bloomberg)
Ella Rabinovich (University of Toronto)
Dianna Radpour (University of Colorado Boulder)
Preethi Raghavan (IBM Research)
Afshin Rahimi (University of Queensland)
Revanth Rameshkumar (Microsoft Research)
Adithya Renduchintala (JHU)
Carolyn Rose (CMU)
Alla Rozovskaya (City University of New York)
Derek Ruths (McGill University)
Koustuv Saha (Georgia Tech)
Keisuke Sakaguchi (Allen Institute for Artificial Intelligence)
Maarten Sap (University of Washington)
Amirreza Shirani (University of Houston)
Dan Simonson (BlackBoiler)
Kevin Small (Amazon)
Jan Šnajder (University of Zagreb)
Xingyi Song (University of Sheffield)
Evangelia Spiliopoulou (Carnegie Mellon University)
Gabriel Stanovsky (Allen Institute for Artificial Intelligence)
Ian Stewart (Georgia Tech)
Nadiya Straton (Copenhagen Business School)
Shivashankar Subramanian (University of Melbourne)

Jeniya Tabassum (Ohio State University)
Yi Tay (Google)
Zhiyang Teng (Westlake University)
Joel Tetreault (Dataminr)
James Thorne (University of Cambridge)
Rob van der Goot (University of Groningen)
Vasudeva Varma (IIIT Hyderabad)
Daniel Varab (IT University of Copenhagen)
Olga Vechtomova (University of Waterloo)
Nikhita Vedula (Ohio State University)
Alakananda Vempala (Bloomberg)
Rob Voigt (Northwestern University)
Soroush Vosoughi (Dartmouth University)
Xiaojun Wan (Peking University)
Zeerak Waseem (University of Sheffield)
Zhongyu Wei (Fudan University)
Hong Wei (University of Maryland)
Steven Wilson (University of Edinburgh)
Zach Wood-Doughty (Johns Hopkins University)
Ning Yu (Leidos)
Marcos Zampieri (Rochester Institute of Technology)
Guido Zarrella (MITRE)
Vicky Zayats (University of Washington)
Justine Zhang (Cornell University)
Xiao Zhang (Purdue University)
Shi Zong (Ohio State University)

Invited Speakers:

Eduardo Blanco (University of North Texas)
Manaal Faruqui (Google)
Robert Munro (Machine Learning Consulting; former CTO of Figure Eight)
Irwin King (The Chinese University of Hong Kong)

Table of Contents

May I Ask Who's Calling? Named Entity Recognition on Call Center Transcripts for Privacy Law Compliance
Micaela Kaplan 1

"Did you really mean what you said?" : Sarcasm Detection in Hindi-English Code-Mixed Data using Bilingual Word Embeddings
Akshita Aggarwal, Anshul Wadhawan, Anshima Chaudhary and Kavita Maurya 7

Noisy Text Data: Achilles' Heel of BERT
Ankit Kumar, Piyush Makhija and Anuj Gupta 16

Determining Question-Answer Plausibility in Crowdsourced Datasets Using Multi-Task Learning
Rachel Gardner, Maya Varma, Clare Zhu and Ranjay Krishna 22

Combining BERT with Static Word Embeddings for Categorizing Social Media
Israa Alghanmi, Luis Espinosa Anke and Steven Schockaert 28

Enhanced Sentence Alignment Network for Efficient Short Text Matching
Zhe Hu, Zuohui Fu, Cheng Peng and Weiwei Wang 34

PHINC: A Parallel Hinglish Social Media Code-Mixed Corpus for Machine Translation
Vivek Srivastava and Mayank Singh 41

Cross-lingual sentiment classification in low-resource Bengali language
Salim Sazzed 50

The Non-native Speaker Aspect: Indian English in Social Media
Rupak Sarkar, Sayantan Mahinder and Ashiqur KhudaBukhsh 61

Sentence Boundary Detection on Line Breaks in Japanese
Yuta Hayashibe and Kensuke Mitsuzawa 71

Non-ingredient Detection in User-generated Recipes using the Sequence Tagging Approach
Yasuhiro Yamaguchi, Shintaro Inuzuka, Makoto Hiramatsu and Jun Harashima 76

Generating Fact Checking Summaries for Web Claims
Rahul Mishra, Dhruv Gupta and Markus Leippold 81

Intelligent Analyses on Storytelling for Impact Measurement
Koen Kicken, Tessa De Maesschalck, Bart Vanrumste, Tom De Keyser and Hee Reen Shim 91

An Empirical Analysis of Human-Bot Interaction on Reddit
Ming-Cheng Ma and John P. Lalor 101

Detecting Trending Terms in Cybersecurity Forum Discussions
Jack Hughes, Seth Aycock, Andrew Caines, Paula Buttery and Alice Hutchings 107

Service registration chatbot: collecting and comparing dialogues from AMT workers and service's users
Luca Molteni, Mittul Singh, Juho Leinonen, Katri Leino, Mikko Kurimo and Emanuele Della Valle 116

Automated Assessment of Noisy Crowdsourced Free-text Answers for Hindi in Low Resource Setting
Dolly Agarwal, Somya Gupta and Nishant Baghel 122

Punctuation Restoration using Transformer Models for Resource-Rich and -Poor Languages
Tanvirul Alam, Akib Khan and Firoj Alam 132

Truecasing German user-generated conversational text
Yulia Grishina, Thomas Gueudre and Ralf Winkler 143

Fine-Tuning MT systems for Robustness to Second-Language Speaker Variations
Md Mahfuz Ibn Alam and Antonios Anastasopoulos 149

Impact of ASR on Alzheimer's Disease Detection: All Errors are Equal, but Deletions are More Equal than Others
Aparna Balagopalan, Ksenia Shkaruta and Jekaterina Novikova 159

Detecting Entailment in Code-Mixed Hindi-English Conversations
Sharanya Chakravarthy, Anjana Umapathy and Alan W Black 165

Detecting Objectifying Language in Online Professor Reviews
Angie Waller and Kyle Gorman 171

Annotation Efficient Language Identification from Weak Labels
Shriphani Palakodety and Ashiqur KhudaBukhsh 181

Fantastic Features and Where to Find Them: Detecting Cognitive Impairment with a Subsequence Classification Guided Approach
Ben Eyre, Aparna Balagopalan and Jekaterina Novikova 193

Quantifying the Evaluation of Heuristic Methods for Textual Data Augmentation
Omid Kashefi and Rebecca Hwa 200

An Empirical Survey of Unsupervised Text Representation Methods on Twitter Data
Lili Wang, Chongyang Gao, Jason Wei, Weicheng Ma, Ruibo Liu and Soroush Vosoughi 209

Civil Unrest on Twitter (CUT): A Dataset of Tweets to Support Research on Civil Unrest
Justin Sech, Alexandra DeLucia, Anna L Buczak and Mark Dredze 215

Tweeki: Linking Named Entities on Twitter to a Knowledge Graph
Bahareh Harandizadeh and Sameer Singh 222

Representation learning of writing style
Julien Hay, Bich-Lien Doan, Fabrice Popineau and Ouassim Ait Elhara 232

"A Little Birdie Told Me ... " - Social Media Rumor Detection
Karthik Radhakrishnan, Tushar Kanakagiri, Sharanya Chakravarthy and Vidhisha Balachandran 244

Paraphrase Generation via Adversarial Penalizations
Gerson Vizcarra and Jose Ochoa-Luna 249

WNUT-2020 Task 1 Overview: Extracting Entities and Relations from Wet Lab Protocols
Jeniya Tabassum, Wei Xu and Alan Ritter 260

IITKGP at W-NUT 2020 Shared Task-1: Domain specific BERT representation for Named Entity Recognition of lab protocol
Tejas Vaidhya and Ayush Kaushal 268

PublishInCovid19 at WNUT 2020 Shared Task-1: Entity Recognition in Wet Lab Protocols using Structured Learning Ensemble and Contextualised Embeddings
Janvijay Singh and Anshul Wadhawan 273

Big Green at WNUT 2020 Shared Task-1: Relation Extraction as Contextualized Sequence Classification
Chris Miller and Soroush Vosoughi 281

WNUT 2020 Shared Task-1: Conditional Random Field(CRF) based Named Entity Recognition(NER) for Wet Lab Protocols
Kaushik Acharya 286

mgsohrab at WNUT 2020 Shared Task-1: Neural Exhaustive Approach for Entity and Relation Recognition Over Wet Lab Protocols
Mohammad Golam Sohrab, Anh-Khoa Duong Nguyen, Makoto Miwa and Hiroya Takamura . . 290

Fancy Man Launches Zippo at WNUT 2020 Shared Task-1: A Bert Case Model for Wet Lab Entity Extraction
Qingcheng Zeng, Xiaoyang Fang, Zhexin Liang and Haoding Meng 299

BiTeM at WNUT 2020 Shared Task-1: Named Entity Recognition over Wet Lab Protocols using an Ensemble of Contextual Language Models
Julien Knafou, Nona Naderi, Jenny Copara, Douglas Teodoro and Patrick Ruch 305

WNUT-2020 Task 2: Identification of Informative COVID-19 English Tweets
Dat Quoc Nguyen, Thanh Vu, Afshin Rahimi, Mai Hoang Dao, Linh The Nguyen and Long Doan 314

TATL at WNUT-2020 Task 2: A Transformer-based Baseline System for Identification of Informative COVID-19 English Tweets
Anh Tuan Nguyen 319

NHK_STRL at WNUT-2020 Task 2: GATs with Syntactic Dependencies as Edges and CTC-based Loss for Text Classification
Yuki Yasuda, Taichi Ishiwatari, Taro Miyazaki and Jun Goto 324

NLP North at WNUT-2020 Task 2: Pre-training versus Ensembling for Detection of Informative COVID-19 English Tweets
Anders Giovanni Møller, Rob van der Goot and Barbara Plank 331

Siva at WNUT-2020 Task 2: Fine-tuning Transformer Neural Networks for Identification of Informative Covid-19 Tweets
Siva Sai 337

IIITBH at WNUT-2020 Task 2: Exploiting the best of both worlds
Saichethan Reddy and Pradeep Biswal 342

Phonemer at WNUT-2020 Task 2: Sequence Classification Using COVID Twitter BERT and Bagging Ensemble Technique based on Plurality Voting
Anshul Wadhawan 347

CXP949 at WNUT-2020 Task 2: Extracting Informative COVID-19 Tweets - RoBERTa Ensembles and The Continued Relevance of Handcrafted Features
Calum Perrio and Harish Tayyar Madabushi 352

InfoMiner at WNUT-2020 Task 2: Transformer-based Covid-19 Informative Tweet Extraction
Hansi Hettiarachchi and Tharindu Ranasinghe 359

BANANA at WNUT-2020 Task 2: Identifying COVID-19 Information on Twitter by Combining Deep Learning and Transfer Learning Models
Tin Huynh, Luan Thanh Luan and Son T. Luu 366

DATAMAFIA at WNUT-2020 Task 2: A Study of Pre-trained Language Models along with Regularization Techniques for Downstream Tasks
Ayan Sengupta 371

UPennHLP at WNUT-2020 Task 2 : Transformer models for classification of COVID19 posts on Twitter
Arjun Magge, Varad Pimpalkhute, Divya Rallapalli, David Siguenza and Graciela Gonzalez-Hernandez 378

UIT-HSE at WNUT-2020 Task 2: Exploiting CT-BERT for Identifying COVID-19 Information on the Twitter Social Network
Khiem Tran, Hao Phan, Kiet Nguyen and Ngan Luu Thuy Nguyen 383

Emory at WNUT-2020 Task 2: Combining Pretrained Deep Learning Models and Feature Enrichment for Informative Tweet Identification
Yuting Guo, Mohammed Ali Al-Garadi and Abeed Sarker 388

CSECU-DSG at WNUT-2020 Task 2: Exploiting Ensemble of Transfer Learning and Hand-crafted Features for Identification of Informative COVID-19 English Tweets
Fareen Tasneem, Jannatun Naim, Radiathun Tasnia, Tashin Hossain and Abu Nowshed Chy . . . 394

IRLab@IITBHU at WNUT-2020 Task 2: Identification of informative COVID-19 English Tweets using BERT
Supriya Chanda, Eshita Nandy and Sukomal Pal 399

NutCracker at WNUT-2020 Task 2: Robustly Identifying Informative COVID-19 Tweets using Ensembling and Adversarial Training
Priyanshu Kumar and Aadarsh Singh 404

DSC-IIT ISM at WNUT-2020 Task 2: Detection of COVID-19 informative tweets using RoBERTa
Sirigireddy Dhana Laxmi, Rohit Agarwal and Aman Sinha 409

Linguist Geeks on WNUT-2020 Task 2: COVID-19 Informative Tweet Identification using Progressive Trained Language Models and Data Augmentation
Vasudev Awatramani and Anupam Kumar 414

NLPRL at WNUT-2020 Task 2: ELMo-based System for Identification of COVID-19 Tweets
Rajesh Kumar Mundotiya, Rupjyoti Baruah, Bhavana Srivastava and Anil Kumar Singh 419

SU-NLP at WNUT-2020 Task 2: The Ensemble Models
Kenan Fayoumi and Reyyan Yeniterzi 423

IDSOU at WNUT-2020 Task 2: Identification of Informative COVID-19 English Tweets
Sora Ohashi, Tomoyuki Kajiwara, Chenhui Chu, Noriko Takemura, Yuta Nakashima and Hajime Nagahara 428

ComplexDataLab at W-NUT 2020 Task 2: Detecting Informative COVID-19 Tweets by Attending over Linked Documents
Kellin Pelrine, Jacob Danovitch, Albert Orozco Camacho and Reihaneh Rabbany 434

NEU at WNUT-2020 Task 2: Data Augmentation To Tell BERT That Death Is Not Necessarily Informative
Kumud Chauhan 440

LynyrdSkynyrd at WNUT-2020 Task 2: Semi-Supervised Learning for Identification of Informative COVID-19 English Tweets
Abhilasha Sancheti, Kushal Chawla and Gaurav Verma 444

NIT_COVID-19 at WNUT-2020 Task 2: Deep Learning Model RoBERTa for Identify Informative COVID-19 English Tweets
Jagadeesh M S and Alphonse P J A 450

EdinburghNLP at WNUT-2020 Task 2: Leveraging Transformers with Generalized Augmentation for Identifying Informativeness in COVID-19 Tweets
Nickil Maveli 455

#GCDH at WNUT-2020 Task 2: BERT-Based Models for the Detection of Informativeness in English COVID-19 Related Tweets
Hanna Varachkina, Stefan Ziehe, Tillmann Dönicke and Franziska Pannach 462

Not-NUTs at WNUT-2020 Task 2: A BERT-based System in Identifying Informative COVID-19 English Tweets
Thai Hoang and Phuong Vu 466

CIA_NITT at WNUT-2020 Task 2: Classification of COVID-19 Tweets Using Pre-trained Language Models
Yandrapati Prakash Babu and Rajagopal Eswari 471

UET at WNUT-2020 Task 2: A Study of Combining Transfer Learning Methods for Text Classification with RoBERTa
Huy Dao Quang and Tam Nguyen Minh 475

Dartmouth CS at WNUT-2020 Task 2: Fine tuning BERT for Tweet classification
Dylan Whang and Soroush Vosoughi 480

SunBear at WNUT-2020 Task 2: Improving BERT-Based Noisy Text Classification with Knowledge of the Data domain
Linh Doan Bao, Viet Anh Nguyen and Quang Pham Huu 485

ISWARA at WNUT-2020 Task 2: Identification of Informative COVID-19 English Tweets using BERT and FastText Embeddings
Wava Carissa Putri, Rani Aulia Hidayat, Isnaini Nurul Khasanah and Rahmad Mahendra 491

COVCOR20 at WNUT-2020 Task 2: An Attempt to Combine Deep Learning and Expert rules
Ali Hürriyetoğlu, Ali Safaya, Osman Mutlu, Nelleke Oostdijk and Erdem Yörük 495

TEST_POSITIVE at W-NUT 2020 Shared Task-3: Cross-task modeling
Chacha Chen, Chieh-Yang Huang, Yaqi Hou, Yang Shi, Enyan Dai and Jiaqi Wang 499

imec-ETRO-VUB at W-NUT 2020 Shared Task-3: A multilabel BERT-based system for predicting COVID-19 events
Xiangyu Yang, Giannis Bekoulis and Nikos Deligiannis 505

UCD-CS at W-NUT 2020 Shared Task-3: A Text to Text Approach for COVID-19 Event Extraction on Social Media
Congcong Wang and David Lillis 514

Winners at W-NUT 2020 Shared Task-3: Leveraging Event Specific and Chunk Span information for Extracting COVID Entities from Tweets
Ayush Kaushal and Tejas Vaidhya 522

HLTRI at W-NUT 2020 Shared Task-3: COVID-19 Event Extraction from Twitter Using Multi-Task Hopfield Pooling
Maxwell Weinzierl and Sanda Harabagiu 530

Conference Program

May I Ask Who's Calling? Named Entity Recognition on Call Center Transcripts for Privacy Law Compliance
Micaela Kaplan

"Did you really mean what you said?" : Sarcasm Detection in Hindi-English Code-Mixed Data using Bilingual Word Embeddings
Akshita Aggarwal, Anshul Wadhawan, Anshima Chaudhary and Kavita Maurya

Noisy Text Data: Achilles' Heel of BERT
Ankit Kumar, Piyush Makhija and Anuj Gupta

Determining Question-Answer Plausibility in Crowdsourced Datasets Using Multi-Task Learning
Rachel Gardner, Maya Varma, Clare Zhu and Ranjay Krishna

Combining BERT with Static Word Embeddings for Categorizing Social Media
Israa Alghanmi, Luis Espinosa Anke and Steven Schockaert

Enhanced Sentence Alignment Network for Efficient Short Text Matching
Zhe Hu, Zuohui Fu, Cheng Peng and Weiwei Wang

PHINC: A Parallel Hinglish Social Media Code-Mixed Corpus for Machine Translation
Vivek Srivastava and Mayank Singh

Cross-lingual sentiment classification in low-resource Bengali language
Salim Sazzed

The Non-native Speaker Aspect: Indian English in Social Media
Rupak Sarkar, Sayantan Mahinder and Ashiqur KhudaBukhsh

Sentence Boundary Detection on Line Breaks in Japanese
Yuta Hayashibe and Kensuke Mitsuzawa

Non-ingredient Detection in User-generated Recipes using the Sequence Tagging Approach
Yasuhiro Yamaguchi, Shintaro Inuzuka, Makoto Hiramatsu and Jun Harashima

Generating Fact Checking Summaries for Web Claims
Rahul Mishra, Dhruv Gupta and Markus Leippold

No Day Set (continued)

Intelligent Analyses on Storytelling for Impact Measurement
Koen Kicken, Tessa De Maesschalck, Bart Vanrumste, Tom De Keyser and Hee Reen Shim

An Empirical Analysis of Human-Bot Interaction on Reddit
Ming-Cheng Ma and John P. Lalor

Detecting Trending Terms in Cybersecurity Forum Discussions
Jack Hughes, Seth Aycock, Andrew Caines, Paula Buttery and Alice Hutchings

Service registration chatbot: collecting and comparing dialogues from AMT workers and service's users
Luca Molteni, Mittul Singh, Juho Leinonen, Katri Leino, Mikko Kurimo and Emanuele Della Valle

Automated Assessment of Noisy Crowdsourced Free-text Answers for Hindi in Low Resource Setting
Dolly Agarwal, Somya Gupta and Nishant Baghel

Punctuation Restoration using Transformer Models for Resource-Rich and -Poor Languages
Tanvirul Alam, Akib Khan and Firoj Alam

Truecasing German user-generated conversational text
Yulia Grishina, Thomas Gueudre and Ralf Winkler

Fine-Tuning MT systems for Robustness to Second-Language Speaker Variations
Md Mahfuz Ibn Alam and Antonios Anastasopoulos

Impact of ASR on Alzheimer's Disease Detection: All Errors are Equal, but Deletions are More Equal than Others
Aparna Balagopalan, Ksenia Shkaruta and Jekaterina Novikova

Detecting Entailment in Code-Mixed Hindi-English Conversations
Sharanya Chakravarthy, Anjana Umapathy and Alan W Black

Detecting Objectifying Language in Online Professor Reviews
Angie Waller and Kyle Gorman

Annotation Efficient Language Identification from Weak Labels
Shriphani Palakodety and Ashiqur KhudaBukhsh

No Day Set (continued)

Fantastic Features and Where to Find Them: Detecting Cognitive Impairment with a Subsequence Classification Guided Approach
Ben Eyre, Aparna Balagopalan and Jekaterina Novikova

Quantifying the Evaluation of Heuristic Methods for Textual Data Augmentation
Omid Kashefi and Rebecca Hwa

An Empirical Survey of Unsupervised Text Representation Methods on Twitter Data
Lili Wang, Chongyang Gao, Jason Wei, Weicheng Ma, Ruibo Liu and Soroush Vosoughi

Civil Unrest on Twitter (CUT): A Dataset of Tweets to Support Research on Civil Unrest
Justin Sech, Alexandra DeLucia, Anna L Buczak and Mark Dredze

Tweeki: Linking Named Entities on Twitter to a Knowledge Graph
Bahareh Harandizadeh and Sameer Singh

Representation learning of writing style
Julien Hay, Bich-Lien Doan, Fabrice Popineau and Ouassim Ait Elhara

"A Little Birdie Told Me ... " - Social Media Rumor Detection
Karthik Radhakrishnan, Tushar Kanakagiri, Sharanya Chakravarthy and Vidhisha Balachandran

Paraphrase Generation via Adversarial Penalizations
Gerson Vizcarra and Jose Ochoa-Luna

WNUT-2020 Task 1 Overview: Extracting Entities and Relations from Wet Lab Protocols
Jeniya Tabassum, Wei Xu and Alan Ritter

IITKGP at W-NUT 2020 Shared Task-1: Domain specific BERT representation for Named Entity Recognition of lab protocol
Tejas Vaidhya and Ayush Kaushal

PublishInCovid19 at WNUT 2020 Shared Task-1: Entity Recognition in Wet Lab Protocols using Structured Learning Ensemble and Contextualised Embeddings
Janvijay Singh and Anshul Wadhawan

Big Green at WNUT 2020 Shared Task-1: Relation Extraction as Contextualized Sequence Classification
Chris Miller and Soroush Vosoughi

No Day Set (continued)

WNUT 2020 Shared Task-1: Conditional Random Field(CRF) based Named Entity Recognition(NER) for Wet Lab Protocols
Kaushik Acharya

mgsohrab at WNUT 2020 Shared Task-1: Neural Exhaustive Approach for Entity and Relation Recognition Over Wet Lab Protocols
Mohammad Golam Sohrab, Anh-Khoa Duong Nguyen, Makoto Miwa and Hiroya Takamura

Fancy Man Launches Zippo at WNUT 2020 Shared Task-1: A Bert Case Model for Wet Lab Entity Extraction
Qingcheng Zeng, Xiaoyang Fang, Zhexin Liang and Haoding Meng

BiTeM at WNUT 2020 Shared Task-1: Named Entity Recognition over Wet Lab Protocols using an Ensemble of Contextual Language Models
Julien Knafou, Nona Naderi, Jenny Copara, Douglas Teodoro and Patrick Ruch

WNUT-2020 Task 2: Identification of Informative COVID-19 English Tweets
Dat Quoc Nguyen, Thanh Vu, Afshin Rahimi, Mai Hoang Dao, Linh The Nguyen and Long Doan

TATL at WNUT-2020 Task 2: A Transformer-based Baseline System for Identification of Informative COVID-19 English Tweets
Anh Tuan Nguyen

NHK_STRL at WNUT-2020 Task 2: GATs with Syntactic Dependencies as Edges and CTC-based Loss for Text Classification
Yuki Yasuda, Taichi Ishiwatari, Taro Miyazaki and Jun Goto

NLP North at WNUT-2020 Task 2: Pre-training versus Ensembling for Detection of Informative COVID-19 English Tweets
Anders Giovanni Møller, Rob van der Goot and Barbara Plank

Siva at WNUT-2020 Task 2: Fine-tuning Transformer Neural Networks for Identification of Informative Covid-19 Tweets
Siva Sai

IIITBH at WNUT-2020 Task 2: Exploiting the best of both worlds
Saichethan Reddy and Pradeep Biswal

Phonemer at WNUT-2020 Task 2: Sequence Classification Using COVID Twitter BERT and Bagging Ensemble Technique based on Plurality Voting
Anshul Wadhawan

CXP949 at WNUT-2020 Task 2: Extracting Informative COVID-19 Tweets - RoBERTa Ensembles and The Continued Relevance of Handcrafted Features
Calum Perrio and Harish Tayyar Madabushi

No Day Set (continued)

InfoMiner at WNUT-2020 Task 2: Transformer-based Covid-19 Informative Tweet Extraction
Hansi Hettiarachchi and Tharindu Ranasinghe

BANANA at WNUT-2020 Task 2: Identifying COVID-19 Information on Twitter by Combining Deep Learning and Transfer Learning Models
Tin Huynh, Luan Thanh Luan and Son T. Luu

DATAMAFIA at WNUT-2020 Task 2: A Study of Pre-trained Language Models along with Regularization Techniques for Downstream Tasks
Ayan Sengupta

UPennHLP at WNUT-2020 Task 2 : Transformer models for classification of COVID19 posts on Twitter
Arjun Magge, Varad Pimpalkhute, Divya Rallapalli, David Siguenza and Graciela Gonzalez-Hernandez

UIT-HSE at WNUT-2020 Task 2: Exploiting CT-BERT for Identifying COVID-19 Information on the Twitter Social Network
Khiem Tran, Hao Phan, Kiet Nguyen and Ngan Luu Thuy Nguyen

Emory at WNUT-2020 Task 2: Combining Pretrained Deep Learning Models and Feature Enrichment for Informative Tweet Identification
Yuting Guo, Mohammed Ali Al-Garadi and Abeed Sarker

CSECU-DSG at WNUT-2020 Task 2: Exploiting Ensemble of Transfer Learning and Hand-crafted Features for Identification of Informative COVID-19 English Tweets
Fareen Tasneem, Jannatun Naim, Radiathun Tasnia, Tashin Hossain and Abu Nowshed Chy

IRLab@IITBHU at WNUT-2020 Task 2: Identification of informative COVID-19 English Tweets using BERT
Supriya Chanda, Eshita Nandy and Sukomal Pal

NutCracker at WNUT-2020 Task 2: Robustly Identifying Informative COVID-19 Tweets using Ensembling and Adversarial Training
Priyanshu Kumar and Aadarsh Singh

DSC-IIT ISM at WNUT-2020 Task 2: Detection of COVID-19 informative tweets using RoBERTa
Sirigireddy Dhana Laxmi, Rohit Agarwal and Aman Sinha

Linguist Geeks on WNUT-2020 Task 2: COVID-19 Informative Tweet Identification using Progressive Trained Language Models and Data Augmentation
Vasudev Awatramani and Anupam Kumar

NLPRL at WNUT-2020 Task 2: ELMo-based System for Identification of COVID-19 Tweets
Rajesh Kumar Mundotiya, Rupjyoti Baruah, Bhavana Srivastava and Anil Kumar Singh

No Day Set (continued)

SU-NLP at WNUT-2020 Task 2: The Ensemble Models
Kenan Fayoumi and Reyyan Yeniterzi

IDSOU at WNUT-2020 Task 2: Identification of Informative COVID-19 English Tweets
Sora Ohashi, Tomoyuki Kajiwara, Chenhui Chu, Noriko Takemura, Yuta Nakashima and Hajime Nagahara

ComplexDataLab at W-NUT 2020 Task 2: Detecting Informative COVID-19 Tweets by Attending over Linked Documents
Kellin Pelrine, Jacob Danovitch, Albert Orozco Camacho and Reihaneh Rabbany

NEU at WNUT-2020 Task 2: Data Augmentation To Tell BERT That Death Is Not Necessarily Informative
Kumud Chauhan

LynyrdSkynyrd at WNUT-2020 Task 2: Semi-Supervised Learning for Identification of Informative COVID-19 English Tweets
Abhilasha Sancheti, Kushal Chawla and Gaurav Verma

NIT_COVID-19 at WNUT-2020 Task 2: Deep Learning Model RoBERTa for Identify Informative COVID-19 English Tweets
Jagadeesh M S and Alphonse P J A

EdinburghNLP at WNUT-2020 Task 2: Leveraging Transformers with Generalized Augmentation for Identifying Informativeness in COVID-19 Tweets
Nickil Maveli

#GCDH at WNUT-2020 Task 2: BERT-Based Models for the Detection of Informativeness in English COVID-19 Related Tweets
Hanna Varachkina, Stefan Ziehe, Tillmann Dönicke and Franziska Pannach

Not-NUTs at WNUT-2020 Task 2: A BERT-based System in Identifying Informative COVID-19 English Tweets
Thai Hoang and Phuong Vu

CIA_NITT at WNUT-2020 Task 2: Classification of COVID-19 Tweets Using Pre-trained Language Models
Yandrapati Prakash Babu and Rajagopal Eswari

UET at WNUT-2020 Task 2: A Study of Combining Transfer Learning Methods for Text Classification with RoBERTa
Huy Dao Quang and Tam Nguyen Minh

Dartmouth CS at WNUT-2020 Task 2: Fine tuning BERT for Tweet classification
Dylan Whang and Soroush Vosoughi

No Day Set (continued)

SunBear at WNUT-2020 Task 2: Improving BERT-Based Noisy Text Classification with Knowledge of the Data domain
Linh Doan Bao, Viet Anh Nguyen and Quang Pham Huu

ISWARA at WNUT-2020 Task 2: Identification of Informative COVID-19 English Tweets using BERT and FastText Embeddings
Wava Carissa Putri, Rani Aulia Hidayat, Isnaini Nurul Khasanah and Rahmad Mahendra

COVCOR20 at WNUT-2020 Task 2: An Attempt to Combine Deep Learning and Expert rules
Ali Hürriyetoğlu, Ali Safaya, Osman Mutlu, Nelleke Oostdijk and Erdem Yörük

TEST_POSITIVE at W-NUT 2020 Shared Task-3: Cross-task modeling
Chacha Chen, Chieh-Yang Huang, Yaqi Hou, Yang Shi, Enyan Dai and Jiaqi Wang

imec-ETRO-VUB at W-NUT 2020 Shared Task-3: A multilabel BERT-based system for predicting COVID-19 events
Xiangyu Yang, Giannis Bekoulis and Nikos Deligiannis

UCD-CS at W-NUT 2020 Shared Task-3: A Text to Text Approach for COVID-19 Event Extraction on Social Media
Congcong Wang and David Lillis

Winners at W-NUT 2020 Shared Task-3: Leveraging Event Specific and Chunk Span information for Extracting COVID Entities from Tweets
Ayush Kaushal and Tejas Vaidhya

HLTRI at W-NUT 2020 Shared Task-3: COVID-19 Event Extraction from Twitter Using Multi-Task Hopfield Pooling
Maxwell Weinzierl and Sanda Harabagiu