EVALITA
Evaluation
of NLP
and Speech Tools
for Italian

Proceedings of the
Seventh Evaluation Campaign
of Natural Language Processing and Speech Tools for Italian
Final Workshop (EVALITA 2020)

December 17th, 2020

Editors:

Valerio Basile
Danilo Croce
Maria Di Maro
Lucia C. Passaro

aA Accademia university press

aA

AILC

Accademia University Press

EVALITA Evaluation of NLP and Speech Tools for Italian - December 17th, 2020

Proceedings of the Seventh Evaluation Campaign of Natural Language Processing and Speech Tools for Italian Final Workshop

Under the direction of Valerio Basile, Danilo Croce, Maria Maro and Lucia C. Passaro

DOI: 10.4000/books.aaccademia.6732
Publisher: Accademia University Press
Place of publication: Torino
Year of publication: 2020
Publication date: 11 mai 2021
Series: Collana dell'Associazione Italiana di Linguistica Computazionale
Digital ISBN: 979-12-80136-32-9

https://books.openedition.org

Print edition
Number of pages: 521

DIGITAL REFERENCE

Basile, Valerio, et al., editors. *EVALITA Evaluation of NLP and Speech Tools for Italian - December 17th, 2020*. Accademia University Press, 2020, https://doi.org/10.4000/books.aaccademia.6732.

This text was automatically generated on 10 février 2025.

The PDF format is issued under the Creative Commons - Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International - CC BY-NC-ND 4.0 license unless otherwise stated.

The ePub format is issued under the Creative Commons - Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International - CC BY-NC-ND 4.0 license unless otherwise stated.

ABSTRACT

Welcome to EVALITA 2020! EVALITA is the evaluation campaign of Natural Language Processing and Speech Tools for Italian. EVALITA is an initiative of the Italian Association for Computational Linguistics (AILC, http://www.ai-lc.it) and it is endorsed by the Italian Association for Artificial Intelligence (AIxIA, http://www.aixia.it) and the Italian Association for Speech Sciences (AISV, http://www.aisv.it).

© 2020 by AILC - Associazione Italiana di Linguistica Computazionale
sede legale: c/o Bernardo Magnini, Via delle Cave 61, 38122 Trento
codice fiscale 96101430229
email: info@ai-lc.it

Pubblicazione resa disponibile
nei termini della licenza Creative Commons
Attribuzione – Non commerciale – Non opere derivate 4.0



Accademia University Press
via Carlo Alberto 55
I-10123 Torino
info@aAccademia.it

isbn PDF 9791280136275
www.aAccademia.it/EVALITA_2020

Accademia University Press è un marchio registrato di proprietà
di LEXIS Compagnia Editoriale in Torino srl

Table of Contents

Preface 1

EVALITA 2020: Overview of the 7th Evaluation Campaign of Natural Language Processing and Speech Tools for Italian
Valerio Basile, Danilo Croce, Maria Di Maro, Lucia C. Passaro 5

Keynote Talk
Flattening the Curve of the COVID-19 Infodemic: These Evaluation Campaigns Can Help!
Preslav Nakov 15

Track "Affect, Hate, and Stance"

AMI: Automatic Misogyny Identification

AMI @ EVALITA2020: Automatic Misogyny Identification
Elisabetta Fersini, Debora Nozza and Paolo Rosso 21

UniBO @ AMI: A Multi-Class Approach to Misogyny and Aggressiveness Identification on Twitter Posts Using AlBERTo
Arianna Muti and Alberto Barrón-Cedeño 29

fabsam @ AMI: A Convolutional Neural Network Approach
Samuel Fabrizi 35

Jigsaw @ AMI and HaSpeeDe2: Fine-Tuning a Pre-Trained Comment-Domain BERT Model
Alyssa Lees, Jeffrey Sorensen and Ian Kivlichan 40

PoliTeam @ AMI: Improving Sentence Embedding Similarity with Misogyny Lexicons for Automatic Misogyny Identification in Italian Tweets
Giuseppe Attanasio and Eliana Pastor 48

MDD @ AMI: Vanilla Classifiers for Misogyny Identification
Samer El Abassi and Sergiu Nisioi 55

No Place For Hate Speech @ AMI: Convolutional Neural Network and Word Embedding for the Identification of Misogyny in Italian
Adriano dos S. R. da Silva and Norton T. Roman 60

ATE_ABSITA: Aspect Term Extraction and Aspect-Based Sentiment Analysis

ATE_ABSITA @ EVALITA2020: Overview of the Aspect Term Extraction and Aspect-based Sentiment Analysis Task
Lorenzo de Mattei, Graziella De Martino, Andrea Iovine, Alessio Miaschi, Marco Polignano and Giulia Rambelli 67

SentNA @ ATE_ABSITA: Sentiment Analysis of Customer Reviews Using Boosted Trees with Lexical and Lexicon-based Features
Francesco Mele, Antonio Sorgente and Giuseppe Vettigli 75

ghostwriter19 @ ATE_ABSITA: Zero-Shot and ONNX to Speed up BERT on Sentiment Analysis Tasks at EVALITA 2020
Mauro Bennici 80

App2Check @ ATE_ABSITA 2020: Aspect Term Extraction and Aspect-based Sentiment Analysis
Emanuele Di Rosa and Alberto Durante 85

HaSpeeDe: Hate Speech Detection

HaSpeeDe 2 @ EVALITA2020: Overview of the EVALITA 2020 Hate Speech Detection Task
Manuela Sanguinetti, Gloria Comandini, Elisa Di Nuovo, Simona Frenda, Marco Stranisci, Cristina Bosco, Tommaso Caselli, Viviana Patti and Irene Russo 93

YNU_OXZ @ HaSpeeDe 2 and AMI : XLM-RoBERTa with Ordered Neurons LSTM for Classification Task at EVALITA 2020
Xiaozhi Ou and Hongling Li 102

DH-FBK @ HaSpeeDe2: Italian Hate Speech Detection via Self-Training and Oversampling
Elisa Leonardelli, Stefano Menini and Sara Tonelli 110

By1510 @ HaSpeeDe 2: Identification of Hate Speech for Italian Language in Social Media Data
Tao Deng, Yang Bai and Hongbing Dai 116

Venses @ HaSpeeDe2 & SardiStance: Multilevel Deep Linguistically Based Supervised Approach to Classification
Rodolfo Delmonte 121

UR NLP @ HaSpeeDe 2 at EVALITA 2020: Towards Robust Hate Speech Detection with Contextual Embeddings
Julia Hoffmann and Udo Kruschwitz 129

Fontana-Unipi @ HaSpeeDe2: Ensemble of transformers for the Hate Speech task at Evalita
Michele Fontana and Giuseppe Attardi 136

TheNorth @ HaSpeeDe 2: BERT-based Language Model Fine-tuning for Italian Hate Speech Detection
Eric Lavergne, Rajkumar Saini, György Kovács and Killian Murphy 142

UO @ HaSpeeDe2: Ensemble Model for Italian Hate Speech Detection
Mariano Jason Rodriguez Cisnero and Reynier Ortega Bueno 148

No Place For Hate Speech @ HaSpeeDe 2: Ensemble to Identify Hate Speech in Italian
Adriano dos S. R. da Silva and Norton T. Roman 154

Svandiela @ HaSpeeDe: Detecting Hate Speech in Italian Twitter Data with BERT
Svea Klaus, Anna-Sophie Bartle and Daniela Rossmann 159

CHILab @ HaSpeeDe 2: Enhancing Hate Speech Detection with Part-of-Speech Tagging
Giuseppe Gambino and Roberto Pirrone 165

Montanti @ HaSpeeDe2 EVALITA 2020: Hate Speech Detection in online contents
Elia Bisconti, Matteo Montagnani 171

SardiStance: Stance Detection

SardiStance @ EVALITA2020: Overview of the Task on Stance Detection in Italian Tweets
Alessandra Teresa Cignarella, Mirko Lai, Cristina Bosco, Viviana Patti and Paolo Rosso 177

UNITOR @ Sardistance2020: Combining Transformer-based Architectures and Transfer Learning for Robust Stance Detection
Simone Giorgioni, Marcello Politi, Samir Salman, Roberto Basili and Danilo Croce 187

ghostwriter19 @ SardiStance: Generating New Tweets to Classify SardiStance EVALITA 2020 Political Tweets
Mauro Bennici 193

QMUL-SDS @ SardiStance: Leveraging Network Interactions to Boost Performance on Stance Detection using Knowledge Graphs
Rabab Alkhalifa and Arkaitz Zubiaga 198

DeepReading @ SardiStance 2020: Combining Textual, Social and Emotional Features
Maria S. Espinosa, Rodrigo Agerri, Alvaro Rodrigo and Roberto Centeno 204

TextWiller @ SardiStance, HaSpeede2: Text or Con-text? A Smart Use of Social Network Data in Predicting Polarization
Federico Ferraccioli, Andrea Sciandra, Mattia Da Pont, Paolo Girardi, Dario Solari and Livio Finos 210

UninaStudents @ SardiStance: Stance Detection in Italian Tweets - Task A
Maurizio Moraca, Gianluca Sabella and Simone Morra 215

SSN NLP @ SardiStance : Stance Detection from Italian Tweets using RNN and Transformers
Kayalvizhi S, Thenmozhi D and Aravindan Chandrabose 220

SSNCSE-NLP @ EVALITA2020: Textual and Contextual Stance Detection from Tweets Using Machine Learning Approach
Bharathi B, Bhuvana J and Nitin Nikamanth Appiah Balaji 224

Track "Creativity and Style"

CHANGE-IT: Style Transfer

CHANGE-IT @ EVALITA 2020: Change Headlines, Adapt News, GEnerate
Lorenzo De Mattei, Michele Cafagna, Felice Dell'Orletta, Malvina Nissim and Albert Gatt 235

TAG-it: Topic, Age and Gender Prediction

TAG-it @ EVALITA2020: Overview of the Topic, Age, and Gender Prediction Task for Italian
Andrea Cimino, Felice Dell'Orletta and Malvina Nissim 243

UO_4to @ TAG-it 2020: Ensemble of Machine Learning Methods
Maria Fernanda Artigas Herold and Daniel Castro Castro 252

UOBIT @ TAG-it: Exploring a Multi-faceted Representation for Profiling Age, Topic and Gender in Italian Texts.
Roberto Labadie Tamayo, Daniel Castro Castro and Reynier Ortega Bueno 256

ItaliaNLP @ TAG-IT: UmBERTo for Author Profiling at TAG-it 2020
Daniela Occhipinti, Andrea Tesei, Maria Iacono, Carlo Aliprandi and Lorenzo De Mattei 263

Track "Semantics and Multimodality"

DANKMEMES: Multimodal Artefacts Recognition

DANKMEMES @ EVALITA 2020: The Memeing of Life: Memes, Multimodality and Politics
Martina Miliani, Giulia Giorgi, Ilir Rama, Guido Anselmi and Gianluca E. Lebani 275

SNK @ DANKMEMES: Leveraging Pretrained Embeddings for Multimodal Meme Detection
Stefano Fiorucci 284

UPB @ DANKMEMES: Italian Memes Analysis - Employing Visual Models and Graph Convolutional Networks for Meme Identification and Hate Speech Detection
George-Alexandru Vlad, George-Eduard Zaharia, Dumitru-Clementin Cercel and Mihai Dascalu 288

ArchiMeDe @ DANKMEMES: A New Model Architecture for Meme Detection
Jinen Setpal and Gabriele Sarti 294

UNITOR @ DANKMEME: Combining Convolutional Models and Transformer-based architectures for accurate MEME management
Claudia Breazzano, Edoardo Rubino, Danilo Croce and Roberto Basili 301

CONcreTEXT: Concreteness in Context

CONcreTEXT @ EVALITA2020: The Concreteness in Context Task
Lorenzo Gregori, Maria Montefinese, Daniele P. Radicioni, Andrea Amelio Ravelli and Rossella Varvara 311

ANDI @ CONcreTEXT: Predicting Concreteness in Context for English and Italian using Distributional Models and Behavioural Norms
Armand Stefan Rotaru 319

CAPISCO @ CONcreTEXT 2020: (Un)supervised Systems to Contextualize Concreteness with Norming Data
Alessandro Bondielli, Gianluca E. Lebani, Lucia C. Passaro and Alessandro Lenci 327

KonKretiKa @ CONcreTEXT: Computing Concreteness Indexes with Sigmoid Transformation and Adjustment for Context
Yulia Badryzlova 334

Ghigliottin-AI: Evaluating Artificial Players for the Language Game "La Ghigliottina"

Ghigliottin-AI@EVALITA2020: Evaluating Artificial Players for the Language Game ``La Ghigliottina''
Pierpaolo Basile, Marco Lovetere, Johanna Monti, Antonio Pascucci, Federico Sangati and Lucia Siciliani 345

"Il Mago della Ghigliottina" @ GhigliottinAI: When Linguistics meets Artificial Intelligence
Federico Sangati, Antonio Pascucci and Johanna Monti 349

GUL.LE.VER @ GhigliottinAI: A Glove based Artificial Player to Solve the Language Game "La Ghigliottina"
Nazareno de Francesco 356

PRELEARN: Prerequisite Relation Learning

PRELEARN @ EVALITA 2020: Overview of the Prerequisite Relation Learning Task for Italian
Chiara Alzetta, Alessio Miaschi, Felice Dell'Orletta, Frosina Koceva and Ilaria Torre 363

B4DS @ PRELEARN: Ensemble Method for Prerequisite Learning
Giovanni Puccetti, Luis Bolanos, Filippo Chiarello and Gualtiero Fantoni 371

UNIGE_SE @ PRELEARN: Utility for Automatic Prerequisite Learning from Italian Wikipedia (short paper)
Alessio Moggio and Andrea Parizzi 376

NLP-CIC @ PRELEARN: Mastering Prerequisites Relations, from Handcrafted Features to Embeddings
Jason Angel and Segun Aroyehun 381

Track "Time and Diachrony"

DaDoEval: Dating Documents

DaDoEval @ EVALITA 2020: Same-Genre and Cross-Genre Dating of Historical Documents
Stefano Menini, Giovanni Moretti, Rachele Sprugnoli and Sara Tonelli 391

matteo-brv @ DaDoEval: An SVM-based Approach for Automatic Document Dating
Matteo Brivio 398

rmassidda @ DaDoEval: Document Dating Using Sentence Embeddings at EVALITA 2020
Riccardo Massidda 403

DIACR-Ita: Diachronic Lexical Semantics

DIACR-Ita @ EVALITA2020: Overview of the EVALITA2020 Diachronic Lexical Semantics (DIACR-Ita) Task
Pierpaolo Basile, Annalina Caputo, Tommaso Caselli, Pierluigi Cassotti and Rossella Varvara 411

University of Padova @ DIACR-Ita
Benyou Wang, Emanuele Di Buccio and Massimo Melucci 420

UWB @ DIACR-Ita: Lexical Semantic Change Detection with CCA and Orthogonal Transformation
Ondřej Pražák, Pavel Přibáň, Stephen Taylor 426

QMUL-SDS @ DIACR-Ita: Evaluating Unsupervised Diachronic Lexical Semantics Classification in Italian
Rabab Alkhalifa, Adam Tsakalidis, Arkaitz Zubiaga and Maria Liakata 432

CL-IMS @ DIACR-Ita: Volente o Nolente: BERT is still not Outperforming SGNS on Semantic Change Detection
Severin Laicher, Gioia Baldissin, Enrique Castaneda, Dominik Schlechtweg and Sabine Schulte Im Walde 438

OP-IMS @ DIACR-Ita: Back to the Roots: SGNS+OP+CD still Rocks Semantic Change Detection
Jens Kaiser, Dominik Schlechtweg and Sabine Schulte Im Walde 444

UNIMIB @ DIACR-Ita: Aligning Distributional Embeddings with a Compass for Semantic Change Detection in the Italian Language
Federico Belotti, Federico Bianchi and Matteo Palmonari 451

NLP-CIC @ DIACR-Ita: POS and Neighbor Based Distributional Models for Lexical Semantic Change in Diachronic Italian Corpora
Jason Angel 456

Track "New Challenges in Long-standing Tasks"

AcCompl-it: Acceptability & Complexity evaluation

AcCompl-it @ EVALITA2020: Overview of the Acceptability & Complexity Evaluation Task for Italian
Dominique Brunato, Cristiano Chesi, Felice Dell'Orletta, Simonetta Montemagni, Giulia Venturi and Roberto Zamparelli 465

UmBERTo-MTSA @ AcCompl-It: Improving Complexity and Acceptability Prediction with Multi-task Learning on Self-Supervised Annotations
Gabriele Sarti 473

Venses @ AcCompl-It: Computing Complexity vs Acceptability with a Constituent Trigram Model and Semantics
Rodolfo Delmonte 479

KIPoS: Part-of-speech Tagging on Spoken Language

KIPoS @ EVALITA2020: Overview of the Task on KIParla Part of Speech Tagging
Cristina Bosco, Silvia Ballarè, Massimo Cerruti, Eugenio Goria and Caterina Mauri 489

UniBO @ KIPoS: Fine-tuning the Italian "BERTology" for PoS-tagging Spoken Data
Fabio Tamburini 497

UniBA @ KIPoS: A Hybrid Approach for Part-of-Speech Tagging
Giovanni Luca Izzi and Stefano Ferilli 501

KLUMSy @ KIPoS: Experiments on Part-of-Speech Tagging of Spoken Italian
Thomas Proisl and Gabriella Lapesa 507

Preface to the EVALITA 2020 Proceedings *

Welcome to EVALITA 2020! EVALITA is the evaluation campaign of Natural Language Processing and Speech Tools for Italian. EVALITA is an initiative of the Italian Association for Computational Linguistics (AILC, http://www.ai-lc.it) and it is endorsed by the Italian Association for Artificial Intelligence (AIxIA, http://www.aixia.it) and the Italian Association for Speech Sciences (AISV, http://www.aisv.it).

This volume includes the reports of both task organisers and participants to all of the EVALITA 2020 challenges. In the 2020 edition, we coordinated the organization of 14 different tasks belonging to five research areas, being: (i) *Affect, Hate, and Stance*, (ii) *Creativity and Style*, (iii) *New Challenges in Long-standing Tasks*, (iv) *Semantics and Multimodality*, *Time and Diachrony*.

The volume is opened by an overview to the EVALITA 2020 campaign, in which we describe the tasks, provide statistics on the participants and task organizers as well as our supporting sponsors. The abstract of the keynote speech made by Preslav Nakov titled *"Flattening the Curve of the COVID-19 Infodemic: These Evaluation Campaigns Can Help!"* is also included in this collection.

Due to the 2020 COVID-19 pandemic, the traditional workshop was held online, where several members of the Italian NLP Community presented the results of their research. Despite the circumstances, the workshop represented an occasion for all participants from both academic institutions and private companies to disseminate their work and results and to share ideas through online sessions dedicated to each task and a general discussion during the plenary event.

We carried on with the tradition of the "Best system across tasks" award. As in 2018, it represented an incentive for students, IT developers and researchers to push the boundaries of the state of the art by facing tasks in new ways, even if not winning.

We would like to thank our sponsors Amazon Science (https://www.amazon.science/), Bnova (https://www.bnova.it/), CELI (https://www.celi.it/), European Language Resources Association (ELRA, http://elra.info/en/) and Google Research (https://research.google/) for their support to the virtual event and to the prize for the best system award.

Our gratitude goes also to University of Turin for hosting the online events. In addition, we sincerely thank the Best System across Tasks committee for providing their expertise and experience. Moreover, we acknowledge the AILC Board members for their trust and support, our mentor Nicole Novielli and all the chairs of the 2018 edition, who helped us during all the organization process of EVALITA 2020.

We warmly thank our invited speaker Preslav Nakov for having shared his knowledge and insights with his talk.

Last but not least, we would like to thank all the task organizers and participants, who made this edition special with their enthusiasm and creativity.

December 2020

Valerio Basile
Danilo Croce
Maria Di Maro
Lucia C. Passaro

* Originally published online by CEUR (CEUR-WS.org, ISSN 1613-0073, Vol-2765, urn:nbn:de:0074-2765-4)

Chairs

Valerio Basile, University of Turin
Danilo Croce, University of Rome “Tor Vergata”
Maria Di Maro, University of Naples “Federico II”
Lucia Passaro, University of Pisa
Advisor: Nicole Novielli, University of Bari “Aldo Moro”

Steering Committee

Chiara Alzetta, Università degli Studi di Genova
Guido Anselmi, University of Milan
Silvia Ballarè, Università degli Studi di Bologna
Pierpaolo Basile, University of Bari Aldo Moro
Cristina Bosco, University of Torino
Dominique Brunato, Istituto di Linguistica Computazionale “Antonio Zampolli” - CNR
Michele Cafagna, University of Pisa
Annalina Caputo, Dublin City University
Tommaso Caselli, University of Groningen, Netherlands
Pierluigi Cassotti, University of Bari Aldo Moro
Massimo Cerruti, University of Torino
Cristiano Chesi, NETS-IUSS
Alessandra Teresa Cignarella, University of Turin / Universitat Politècnica de València
Andrea Cimino, Istituto di Linguistica Computazionale “Antonio Zampolli” - CNR
Gloria Comandini, Università degli Studi di Trento
Graziella De Martino, University of Bari Aldo Moro
Lorenzo de Mattei, University of Pisa
Felice Dell’Orletta, Istituto di Linguistica Computazionale “Antonio Zampolli” - CNR
Elisa Di Nuovo, University of Torino
Elisabetta Fersini, University of Milano-Bicocca
Simona Frenda, University of Torino / Universitat Politècnica de València
Albert Gatt, University of Malta
Giulia Giorgi, University of Milan
Eugenio Goria, University of Torino
Lorenzo Gregori, University of Florence
Andrea Iovine, University of Bari Aldo Moro
Frosina Koceva, Università di Genova
Mirko Lai, University of Torino
Gianluca E. Lebani, Ca’ Foscari University of Venice
Marco Lovetere, Ghigliottiniamo
Caterina Mauri, Università degli Studi di Bologna
Stefano Menini, Fondazione Bruno Kessler
Alessio Miaschi, University of Pisa
Martina Miliani, University for Foreigners of Siena / University of Pisa
Maria Montefinese, University of Padua
Simonetta Montemagni, Istituto di Linguistica Computazionale “Antonio Zampolli” - CNR
Johanna Monti, University of Naples ”L’Orientale”
Giovanni Moretti, Università Cattolica del Sacro Cuore
Malvina Nissim, Faculty of Arts - CLCG University of Groningen
Debora Nozza, Bocconi University
Antonio Pascucci, L’Orientale University of Naples

Viviana Patti, University of Torino
Marco Polignano, University of Bari Aldo Moro
Daniele P. Radicioni, University of Turin
Ilir Rama, University of Milan
Giulia Rambelli, University of Pisa
Andrea Amelio Ravelli, ILC-CNR of Pisa
Paolo Rosso, Universitat Politècnica de València
Irene Russo, Istituto di Linguistica Computazionale "Antonio Zampolli" - CNR
Federico Sangati, Okinawa Institute of Science and Technology Graduate University
Manuela Sanguinetti, University of Torino
Lucia Siciliani, University of Bari Aldo Moro
Rachele Sprugnoli, Università Cattolica del Sacro Cuore
Marco Stranisci, Associazione Acmos
Sara Tonelli, Fondazione Bruno Kessler
Ilaria Torre, University of Genova
Rossella Varvara, University of Florence
Giulia Venturi, Istituto di Linguistica Computazionale "Antonio Zampolli" - CNR
Roberto Zamparelli, Università degli Studi di Trento

Best System Award committee

Giuseppe Attardi, University of Pisa
Francesca Chiusaroli, University of Macerata
Giuseppe Castellucci, Amazon Science
Gloria Gagliardi, "L'Orientale" University of Naples
Nicole Novielli, University of Bari "Aldo Moro"

Website

Manuela Speranza, Fondazione Bruno Kessler