

EVALITA 2016

**The Fifth Evaluation Campaign of Natural Language
Processing and Speech Tools for Italian**

Proceedings of the Workshop

December 7, 2016

Copyright ©2016 for the individual papers by the papers' authors.
Copyright ©2016 for the volume as a collection by its editors.
This volume and its papers are published under the Creative Commons License Attribution 4.0 International (CC BY 4.0).

These proceedings have been published in the CEUR Workshop Proceedings series.
The original papers are available at: <https://ceur-ws.org/Vol-1749>.
The papers are mirrored in the ACL Anthology.

ISBN 978-88-99982-09-6

Table of Contents

Preface

Pierpaolo Basile, Franco Cutugno, Malvina Nissim, Viviana Patti and Rachele Sprugnoli 1

EVALITA 2016: Overview of the 5th Evaluation Campaign of Natural Language Processing and Speech Tools for Italian

Pierpaolo Basile, Franco Cutugno, Malvina Nissim, Viviana Patti and Rachele Sprugnoli 3

Raising Interest and Collecting Suggestions on the EVALITA Evaluation Campaign

Rachele Sprugnoli, Viviana Patti and Franco Cutugno 7

Keynote: Profiling the Personality of Social Media Users

Walter Daelemans 14

The ArtiPhon Task at Evalita 2016

Leonardo Badino 15

Phone Recognition Experiments on ArtiPhon with KALDI

Piero Cosi 21

The EVALITA 2016 Event Factuality Annotation Task (FactA)

Anne-Lyse Minard, Manuela Speranza and Tommaso Caselli 27

Overview of the EVALITA 2016 Named Entity rEcognition and Linking in Italian Tweets (NEEL-IT) Task

Pierpaolo Basile, Annalina Caputo, Anna Lisa Gentile and Giuseppe Rizzo 35

Using Embeddings for Both Entity Recognition and Linking in Tweets

Giuseppe Attardi, Daniele Sartiano, Maria Simi and Irene Sucameli 43

UNIMIB@NEEL-IT: Named Entity Recognition and Linking of Italian Tweets

Flavio Massimiliano Cecchini, Elisabetta Fersini, Pikakshi Manchanda, Enza Messina, Debora Nozza, Matteo Palmonari and Cezar Sas 49

MicroNeel: Combining NLP Tools to Perform Named Entity Detection and Linking on Microposts

Francesco Corcoglioniti, Alessio Palmero Aprosio, Yaroslav Nechaev and Claudio Giuliano . . 55

sisinflab: an Ensemble of Supervised and Unsupervised Strategies for the NEEL-IT Challenge at Evalita 2016

Vittoria Cozza, Wanda La Bruna and Tommaso Di Noia 62

FBK-NLP at NEEL-IT: Active Learning for Domain Adaptation

Anne-Lyse Minard, Mohammed R. H. Qwaider and Bernardo Magnini 69

Overview of the EVALITA 2016 Part Of Speech on TWitter for ITALian Task

Cristina Bosco, Fabio Tamburini, Andrea Bolioli and Alessandro Mazzei 75

Character Embeddings PoS Tagger vs HMM Tagger for Tweets

Giuseppe Attardi and Maria Simi 82

Building the state-of-the-art in POS tagging of Italian Tweets

Andrea Cimino and Felice Dell’Orletta 86

Building a Social Media Adapted PoS Tagger Using FlexTag – A Case Study on Italian Tweets

Tobias Horsmann and Torsten Zesch 92

<i>Mivoq Evalita 2016 PosTwITA tagger</i>	
Giulio Paci	96
<i>NLP-NITMZ: Part-of-Speech Tagging on Italian Social Media Text using Hidden Markov Model</i>	
Partha Pakray and Goutam Majumder	101
<i>When silver glitters more than gold: Bootstrapping an Italian part-of-speech tagger for Twitter</i>	
Barbara Plank and Malvina Nissim	105
<i>bot.zen @ EVALITA 2016 - A minimally-deep learning PoS-tagger (trained for Italian Tweets)</i>	
Egon W. Stemle	111
<i>A BiLSTM-CRF PoS-tagger for Italian tweets using morphological information</i>	
Fabio Tamburini	117
<i>Overview of the EVALITA 2016 Question Answering for Frequently Asked Questions (QA4FAQ) Task</i>	
Annalina Caputo, Marco de Gemmis, Pasquale Lops, Francesco Lovcchio and Vito Manzari	121
<i>Question Answering System for Frequently Asked Questions</i>	
Divyanshu Bhardwaj, Partha Pakray, Jereemi Benthams, Saurav Saha and Alexander Gelbukh	126
<i>Tweaking Word Embeddings for FAQ Ranking</i>	
Erick R. Fonseca, Simone Magnolini, Anna Feltracco, Mohammed R. H. Qwaider and Bernardo Magnini	131
<i>ChiLab4It System in the QA4FAQ Competition</i>	
Arianna Pipitone, Giuseppe Tirone and Roberto Pirrone	137
<i>Overview of the Evalita 2016 SENTiment POLarity Classification Task</i>	
Francesco Barbieri, Valerio Basile, Danilo Croce, Malvina Nissim, Nicole Novielli and Viviana Patti	143
<i>Convolutional Neural Networks for Sentiment Analysis on Italian Tweets</i>	
Giuseppe Attardi, Daniele Sartiano, Chiara Alzetta and Federica Semplici	154
<i>IRADABE2: Lexicon Merging and Positional Features for Sentiment Analysis in Italian</i>	
Davide Buscaldi and Delia Irazú Hernandez-Farias	159
<i>Context-aware Convolutional Neural Networks for Twitter Sentiment Analysis in Italian</i>	
Giuseppe Castellucci, Danilo Croce and Roberto Basili	164
<i>Tandem LSTM-SVM Approach for Sentiment Analysis</i>	
Andrea Cimino and Felice Dell’Orletta	171
<i>Lacam&Int@UNIBA at the EVALITA 2016-SENTIPOLC Task</i>	
Vito Vincenzo Covella, Berardina De Carolis, Stefano Ferilli and Domenico Redavid	177
<i>Sentiment Analysis using Convolutional Neural Networks with Multi-Task Training and Distant Supervision on Italian Tweets</i>	
Jan Deriu and Mark Cieliebak	183
<i>Tweet2Check evaluation at Evalita Sentipolc 2016</i>	
Emanuele Di Rosa and Alberto Durante	188
<i>Computational rule-based model for Irony Detection in Italian Tweets</i>	
Simona Frenda	194

<i>On the performance of B4MSA on SENTIPOLC'16</i>	
Daniela Moctezuma, Eric S. Tellez, Mario Graff and Sabino Miranda-Jiménez	200
<i>Exploiting Emotive Features for the Sentiment Polarity Classification of tweets</i>	
Lucia C. Passaro, Alessandro Bondielli and Alessandro Lenci	205
<i>Samskara Minimal structural features for detecting subjectivity and polarity in Italian tweets</i>	
Irene Russo and Monica Monachini	211