

# **4th Workshop on Open-Source Arabic Corpora and Processing Tools with a Shared Task on Offensive Language Detection 2020**

Marseille, France  
11-16 May 2020

## **Editors:**

**Hend Al-Khalifa  
Walid Magdy  
Kareem Darwish**

**Tamer Elsayed  
Hamdy Mubarak**

ISBN: 978-1-7138-1277-2

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2020) by the Association for Computational Linguistics  
All rights reserved.

Copyright for individual papers remains with the authors and are licensed under a Creative Commons 4.0 license, CC-BY-NC. (<https://creativecommons.org/licenses/by-nc/4.0/>)

Printed with permission by Curran Associates, Inc. (2020)

For permission requests, please contact the Association for Computational Linguistics at the address below.

Association for Computational Linguistics  
209 N. Eighth Street  
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006

Fax: 1-570-476-0860

[acl@aclweb.org](mailto:acl@aclweb.org)

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2633  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

<i>An Arabic Tweets Sentiment Analysis Dataset (ATSAD) using Distant Supervision and Self Training</i> Kathrein Abu Kwaik, Stergios Chatzikyriakidis, Simon Dobnik, Motaz Saad and Richard Johansson . . . . .	1
<i>AraBERT: Transformer-based Model for Arabic Language Understanding</i> wissam antoun, Fady Baly and Hazem Hajj . . . . .	9
<i>AraNet: A Deep Learning Toolkit for Arabic Social Media</i> Muhammad Abdul-Mageed, Chiyu Zhang, Azadeh Hashemi and El Moatez Billah Nagoudi . . . . .	16
<i>Building a Corpus of Qatari Arabic Expressions</i> Sara Al-Mulla and Wajdi Zaghouani . . . . .	24
<i>From Arabic Sentiment Analysis to Sarcasm Detection: The ArSarcasm Dataset</i> Ibrahim Abu Farha and Walid Magdy . . . . .	32
<i>Understanding and Detecting Dangerous Speech in Social Media</i> Ali Alshehri, El Moatez Billah Nagoudi and Muhammad Abdul-Mageed . . . . .	40
<i>Overview of OSACT4 Arabic Offensive Language Detection Shared Task</i> Hamdy Mubarak, Kareem Darwish, Walid Magdy, Tamer Elsayed and Hend Al-Khalifa . . . . .	48
<i>OSACT4 Shared Task on Offensive Language Detection: Intensive Preprocessing-Based Approach</i> Fatemah Husain . . . . .	53
<i>ALT Submission for OSACT Shared Task on Offensive Language Detection</i> Sabit Hassan, Younes Samih, Hamdy Mubarak, Ahmed Abdelali, Ammar Rashed and Shammur Absar Chowdhury . . . . .	61
<i>ASU_OPTO at OSACT4 - Offensive Language Detection for Arabic text</i> Amr Keleg, Samhaa R. El-Beltagy and Mahmoud Khalil . . . . .	66
<i>OSACT4 Shared Tasks: Ensembled Stacked Classification for Offensive and Hate Speech in Arabic Tweets</i> Hafiz Hassaan Saeed, Toon Calders and Faisal Kamiran . . . . .	71
<i>Arabic Offensive Language Detection with Attention-based Deep Neural Networks</i> Bushr Haddad, Zoher Orabe, Anas Al-Abood and Nada Ghneim . . . . .	76
<i>Offensive language detection in Arabic using ULMFiT</i> Mohamed Abdellatif and Ahmed Elgammal . . . . .	82
<i>Multitask Learning for Arabic Offensive Language and Hate-Speech Detection</i> Ibrahim Abu Farha and Walid Magdy . . . . .	86
<i>Combining Character and Word Embeddings for the Detection of Offensive Language in Arabic</i> Abdullah I. Alharbi and Mark Lee . . . . .	91
<i>Multi-Task Learning using AraBert for Offensive Language Detection</i> Marc Djandji, Fady Baly, wissam antoun and Hazem Hajj . . . . .	97

<i>Leveraging Affective Bidirectional Transformers for Offensive Language Detection</i>	
AbdelRahim Elmadany, Chiyu Zhang, Muhammad Abdul-Mageed and Azadeh Hashemi . . . . .	102
<i>Quick and Simple Approach for Detecting Hate Speech in Arabic Tweets</i>	
Abeer Abuzayed and Tamer Elsayed . . . . .	109