

CURRICULUM VITAE OF HALA MULKI

Contact Information Date of Birth: January 1, 1980 Nationality: Syrian/Turkish
Email: hallamulki@gmail.com
Phone: +90 537 048 46 01
Profile: [Google Scholar](#), [ResearchGate](#), [Linkedin](#)

Summary Computational Linguist with a specialization in social media analysis for multilingual contents, under-represented languages (Arabic, Turkish) and Arabic dialects (Levantine, Egyptian, Maghrebi and Gulf). Research skills include leveraging natural language processing approaches within the socio-political domain focusing on sentiment analysis, multiple emotion classification, Author profiling and hate speech and abusive language detection. Throughout years of research, I performed automatic collection/normalization of textual data (comments, tweets) from Facebook, Twitter, Telegram and Youtube, built and annotated Arabic dialectal resources (corpora, lexicons) and provided the proper text analysis using different visualization, machine learning and deep learning tools.

Educational Background **Ph.D. (Computer Science and Engineering)** **2015 - 2019**
Selcuk University, Fen Bilimler Institute, Konya, Turkey

Dissertation title: *Sentiment Classification of Arabic Tweets using a Novel Learning of a Sentiment-Specific Word Embeddings Technique*

Supervisor: Dr. Ismail Babaoglu

M.Sc. (Computer Science and Engineering) **2006 - 2009**

Aleppo University, Computer Engineering Faculty, Aleppo, Syria

Thesis title: *Distributed Collections for Fragment Assembly of Long DNA Strands*

Supervisor: Dr. Amer Bouchi and Dr. Ahmad Alkhodre

B.S. (Computer Engineering) **1998 - 2003**

Aleppo University, Computer Engineering Faculty, Aleppo, Syria

Professional Experience **The Center for Middle Eastern Studies (ORSAM), Ankara, Turkey.**
Data Scientist **2020 - Present**

World of Opinions, Kuwait.

Freelance Data Scientist **2019 - 2020**

Aleppo University, High Technical Institute of Mechanical & Electronic Engineering at Medical, Telecommunication and Control Engineering departments, Aleppo, Syria.

Teaching & Research Assistant **2004 - 2012**

Previous Courses Taught and Coordinated

- Undergraduate courses at Aleppo University (2004 - 2012):
 - Databases Design and Management: Fall 2004, Fall 2006, Fall 2010, Fall 2012
 - Introduction to Programming (C++, Visual C): Spring 2005, Spring 2007, Spring 2009, Spring 2011
 - Algorithms and Data Structures: Spring 2009, Spring 2011
 - Micro-controllers Applications and Programming: Fall 2010, Fall 2012

Current Research Interests

- Sentiment Analysis & Text Visualization
- Hate speech and abusive language detection
- Natural Language understanding & Chatbots

Skills

Programming Languages:

Proficient in Python, R, C/C++ and familiar with Java, Perl and Prolog.

Databases:

MySQL, MS SQL server 2012.

Machine Learning & Data Analysis:

Keras, Tensorflow, Pandas, NumPy, Scikit-learn, Jupyter notebook, LibSVM, Tweepy and MS Azure.

NLP & Text Visualization:

NLTK, Gensim, Gephi, t-SNE, Matplotlib, Wordcloud, Networkx and Bokeh.

Awards and Grants

Winner of the **Turkiye Bursları** Scholarship for Ph.D. study at Selcuk University, Konya, Turkey.

Within the context of **SemEval-2018 Shared Task 1**, I developed a multi-label emotion classification system (Tw-StAR) for Arabic, English and Spanish languages utilizing novel combinations of preprocessing tasks and a binary relevance-based classifier. Using this system, I achieved the THIRD place for Arabic and Spanish among 14 participants [[LINK](#)].

Languages

Arabic: Native

English: Fluent

Turkish: Fluent

Publications

International peer-reviewed journals

H. Mulki, H. Haddad, M. Gridach and I. Babaoglu. Syntax-Ignorant N-Gram Embeddings for Dialectal Arabic Sentiment Analysis. *Natural Language Engineering, Cambridge University Press*. pp.1–39, ISSN:1469-8110, 2020.

H. Mulki, H. Haddad, M. Gridach and I. Babaoglu. Empirical Evaluation of Leveraging Named Entities for Arabic Sentiment Analysis. *The International Arab Journal of Information Technology (IAJIT), Zarqa University, Jordan*. Vol.17, No.2. ISSN: 1683-3198. 2020.

H. Mulki, C. Bechikh Ali, H. Haddad, I. Babaoglu. Tunisian Dialect Sentiment Analysis: A Natural Language Processing-based Approach. In *Computación y Sistemas*, Vol.22, No.4. ISSN: 1405-5546. 2018.

H. Mulki , H. Haddad, I. Babaoglu. Modern Trends in Arabic Sentiment Analysis: A Survey. In *RAITEMENT AUTOMATIQUE DES LANGUES*, Vol.58, No.3 , pp.15–39. ISSN: 12489433. 2017

A.A. Altun and **H. Mulki**. Multistage Filtering Algorithm for Salt and Pepper Noise Removal from Highly Corrupted Microscopic Blood Images. In *International Research Journal of Electronics and Computer Engineering*, Vol.2, No.4 , pp.11–16.2016. ISSN: 2412-4370. 2016

Articles in peer-reviewed international conference proceedings

H. Mulki LeT-Mi: A Levantine Twitter Dataset for Misogynistic Language. *Accepted at Inaugural Middle East Conference – HBKU, Doha, Qatar 2020.*

H. Haddad, **H. Mulki** and A. Oueslati. T-HSAB: A Tunisian Hate Speech and Abusive Dataset. *Proceedings of The 7th International Conference on Arabic Language Processing (ICALP 2019-CITALA)*, Nancy, France October 16-17, 2019.

H. Mulki , H. Haddad, C. Bechikh Ali, I. Babaoglu. Preprocessing Impact on Turkish Sentiment Analysis. In *IEEE 26th Signal Processing and Communications Applications Conference (SIU2018)*, Izmir, Turkey, May 02-05, 2018.

H. Mulki, C. Bechikh Ali, H. Haddad, I. Babaoglu. Tunisian Dialect Sentiment Analysis: A Natural Language Processing-based Approach. *18th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2018)*, LNCS, Hanoi, Vietnam, Mars 18-24, 2018.

M. Gridach, H. Haddad, **H. Mulki**. Empirical Evaluation of Word Representations on Arabic Sentiment Analysis. In *Proceedings of the 6th International Conference on Arabic Language Processing (ICALP 2017)*, CCIS, Fez, Morocco, pages 147-158, October 11-12 2017.

Articles in peer-reviewed workshop proceedings

H. Mulki, H. Haddad, C. Bechikh Ali and H. Alshabani. L-HSAB: A Levantine Twitter Dataset for Hate Speech and Abusive Language. *Proceedings of the Third Workshop on Abusive Language Online (ALW3) co-located with ACL 2019, Florence, Italy, August 1st, 2019.*

H. Mulki, H. Haddad, M. Gridach and I. Babaoglu. Syntax-Ignorant N-gram Embeddings for Sentiment Analysis of Arabic Dialects. *Proceedings of the Fourth Arabic Natural Language Processing Workshop (WANLP) co-located with ACL 2019, Florence, Italy, August 1st, 2019.*

H. Mulki, C. Bechikh Ali, H. Haddad, I. Babaoglu. Tw-StAR at SemEval-2019 Task 5: N-gram embeddings for Hate Speech Detection in Multilingual Tweets. *Proceedings of the 13th International Workshop on Semantic Evaluation (SemEval’19) co-located with NAACL-HLT 2019, Minneapolis, USA, June 6-7, 2019.*

H. Haddad, **H. Mulki**. Empirical Evaluation of Preprocessing on Tunisian Dialect (“Darija”) Sentiment Analysis. In *Proceedings of the second International Widening NLP Workshop (WiNLP ’18)*, co-located with NAACL-HLT 2018, New Orleans, USA, June 1, 2018.

H. Mulki, H. Haddad, C. Bechikh Ali, I. Babaoglu. Tw-StAR at SemEval-2018 Task 1: Affect in Tweets. In *Proceedings of the 12th International Workshop on Semantic Evaluation (SemEval’18) co-located with NAACL-HLT 2018, New Orleans, USA, June 5-6, 2018.*

M. Gridach, H. Haddad, **H. Mulki**. Churn Identification in Microblogs using Convolutional Neural Networks with Structured Logical Knowledge. In *Proceedings of the 3rd Workshop on Noisy User-generated Text (W-NUT) at EMNLP 2017, Copenhagen, Denmark, September 7, 2017.*

H. Mulki, H. Haddad, M. Gridach, I. Babaoglu. Tw-StAR at SemEval-2017 Task 4: Sentiment Classification of Arabic Tweets. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval’17)*. Association for Computational Linguistics (ACL), pages 655-660, Vancouver, Canada, August 3-4, 2017.

Other publications

H. Mulki, H. Haddad, M. Gridach. Polarity analysis of non figurative tweets : Tw-StAR participation on DEFT 2017. *In 24 Conférence sur le Traitement Automatique des Langues Naturelles (TALN), Orléans, France, pages 92-98, June 26, 2017.*

M. Gridach, H. Haddad, **H. Mulki**. FNER-BGRU-CRF at CAp 2017 NER challenge: Bidirectional GRU-CRF for French Named Entity Recognition in Tweets. *In Proceedings of machine learning french conference CAP 2017, Named Entity Recognition in Tweets competition, Grenoble, France, June 28-30, 2017.*

References

Ismail Babaoglu, Konya Technical University, Turkey

Phone: +90 532 513 83 79

Email: ibabaoglu@ktun.edu.tr

Ahmad Alkhodre, Islamic University of Medina, Saudi Arabia

Email: aalkhodre@iu.edu.sa