Segun Taofeek Aroyehun

 +436704034591   •   ✉ aroyehun.segun@gmail.com

Education

Computing Research Center (CIC), Instituto Politécnico Nacional   **Mexico City, Mexico**
PhD in Computer Science, (till 27th August 2021)   *2017–2021*
Thesis: *Deep Learning Methods for Mining Social Media*
Advisor: Alexander Gelbukh

African University of Science and Technology   **Abuja, Nigeria**
MSc Computer Science   *2011*

Obafemi Awolowo University   **Ile Ife, Nigeria**
BSc Computer Science with Economics   *2008*

Experience

Nile University of Nigeria   **Abuja, Nigeria**
Lecturer II in the Department of Computer Science   *2014–2017*

Achievements and Awards

- First place English (Subtasks A and B), HASOC 2021 shared task
- First place (Subtask 1), MeOffendEs@IberLef 2021 shared task
- First place, ALTA 2020 shared task
- First place (out of 30 systems), shared task on aggression identification (English track, Facebook dataset), 1st Workshop on Trolling, Aggression and Cyberbullying (TRAC-2018) at COLING 2018
- Best graduating MSc. student in Computer Science class of 2011 at the African University of Science and Technology, Nigeria

Competitive Grants and Scholarships

- Microsoft Research Latin America PhD Award 2020
- Co-Recipient (with my advisor) Google Latin America Research Award 2020
- IPN Student Innovation Grant 2019–2020 (including a travel grant for a research visit to the RiTUAL lab at University of Houston), one of 81 winners out of 175,000 students
- FAPESP travel grant to attend the São Paulo Advanced School on Data Science 2019
- Volkswagen travel grant to attend the Summer School on Concepts: Acquisition, Representation, Learning and Applications (CARLA), 6–12 August 2018, University of Osnabrueck, Germany
- PhD scholarship of the CONACYT, Government of Mexico 2017-2021, for studying in the PhD program certified at the international quality level by the CONACYT (Ministry of Science)
- DAAD travel grant to attend Autumn School for Information Retrieval and Foraging (ASIRF), 3–8 October 2016, Schlos Dagstuhl, Germany
- Full MSc. scholarship for the top performing student at African University of Science and Technology entrance examination, reserved for two best students out of about 100 competing (2010)

Funded Research Projects

As Principal Investigator

o Detection of Implicit and Explicit Aggression in Social Media using Deep Learning, PI, institutional project, 2018
o Detection of Target and Type of Abusive Language, PI, institutional project, 2019–2020

As Participant

o Multilingual Analysis of Veracity and Author Profile in Social Networks, participant, national project A1-S-47854, 2019–2021
o Sentiment, Opinion, and Language Models with Machine-Learning and Deep-Learning Methods, participant, institutional project SIP 20196437, 20200859, 2019–2020
o Patent Classification using Lexical Semantics and Deep Learning Techniques, participant, institutional project, 2019–2020
o Detection of Adverse Drug Reaction in Tweets, participant, institutional project, 2018
o Opinion Mining and Authorship Attribution in Texts and Multimodal Data with Neural Networks, participant, institutional project SIP 20181792, 2018

Research Stays and Summer Schools Attended

🇺🇸 Research stay at the RiTUAL lab at University of Houston, USA, October 2019
🇧🇷 Sao Paulo Advanced School on Data Science, University of São Paulo, Brazil, July 2019
🇩🇪 Summer School on Concepts in Action: Representation, Learning, and Application, University of Osnabrueck, Germany, 2018
🇩🇪 Autumn School for Information Retrieval and Foraging at Schlos Dagstuhl, Germany, 2016

Professional Activities

Reviewer for journals: Cognitive Computation (IF: 4.29, JCR Q1), Computación y Sistemas (SJR Q3), POLIBITS, Research in Computing Science

Reviewer for international conferences: CSCW 2019, RANLP 2019, CICLing 2019, CICLing 2018

Academic activity: Coordinator of the NLP Research Seminar, Computing Research Center, IPN

Technical Skills

Programming Languages: Python, Java, Bash Script

Technologies: PyTorch, Keras, scikit-learn, NLTK, UNIX, Git

Publications

1. S. T. Aroyehun, J. Angel, N. Majumder, A. Gelbukh, and A. Hussain. "Leveraging label hierarchy using transfer and multi-task learning: A case study on patent classification". In: *Neurocomputing* (2021). ISSN: 0925-2312. DOI: https://doi.org/10.1016/j.neucom.2021.07.057.

2. S. T. Aroyehun and A. Gelbukh. "Evaluation of Intermediate Pre-training for the Detection of Offensive Language". In: *IberLEF@SEPLN*. 2021, to appear.

3. S. T. Aroyehun and A. Gelbukh. "NLP-CIC at HASOC 2020: Multilingual Offensive Language Detection using All-in-one Model." In: *FIRE (Working Notes)*. 2020, pp. 331–335.

4. S. T. Aroyehun and A. Gelbukh. "Automatically Predicting Judgement Dimensions of Human Behaviour". In: *Proceedings of the The 18th Annual Workshop of the Australasian Language Technology Association*. Virtual Workshop: Australasian Language Technology Association, Dec. 2020, pp. 131–134. URL: https://aclanthology.org/2020.alta-1.18.

5. J. Angel, S. T. Aroyehun, and A. Gelbukh. "NLP-CIC@ PRELEARN: Mastering prerequisites relations, from handcrafted features to embeddings". In: *CEUR Workshop Proceedings*. Vol. 2765. CEUR-WS. 2020.

6. J. Angel, S. T. Aroyehun, A. Tamayo, and A. Gelbukh. "NLP-CIC at SemEval-2020 Task 9: Analysing Sentiment in Code-switching Language Using a Simple Deep-learning Classifier". In: *Proceedings of the Fourteenth Workshop on Semantic Evaluation*. Barcelona (online): International Committee for Computational Linguistics, Dec. 2020, pp. 957–962. URL: https://aclanthology.org/2020.semeval-1.123.

7. S. T. Aroyehun and A. Gelbukh. "Detection of Adverse Drug Reaction in Tweets Using a Combination of Heterogeneous Word Embeddings". In: *Proceedings of the 4th Social Media Mining for Health Applications (# SMM4H) Workshop & Shared Task*. 2019, pp. 133–135.

8. S. T. Aroyehun and A. Gelbukh. "Aggression Detection in Social Media: Using Deep Neural Networks, Data Augmentation, and Pseudo Labeling". In: *Proceedings of the 1st Workshop on Trolling, Aggression and Cyberbullying (TRAC-2018)*. 2018, pp. 90–97.

9. S. T. Aroyehun, J. Angel, D. A. Pérez Alvarez, and A. Gelbukh. "Complex Word Identification: Convolutional Neural Network vs. Feature Engineering". In: *Proceedings of the 13th Workshop on Innovative Use of NLP for building educational applications*. 2018, pp. 322–327.

10. S. T. Aroyehun and A. Gelbukh. "Automatic Identification of Drugs and Adverse Drug Reaction Related Tweets". In: *Proceedings of the 2018 EMNLP Workshop SMM4H: The 3rd Social Media Mining for Health Applications Workshop & Shared Task*. 2018, pp. 54–55.