Mesay Gemeda Yigezu - PhD

Contact Information

- **Address:** 6021 Zante Cir, Aurora, 80015, Colorado, USA
- **Phone:** +1 720-415-2422
- **Personal Email:** messay.gemeda@gmail.com
- **Institutional Emails:** mgemedak2022@cic.ipn.mx
- **LinkedIn:** linkedin.com/in/mesay-gemeda
- **Twitter:** twitter.com/Mesay_Gemeda
- **Website:** mesay-gemeda.github.io

Personal Profile

I hold a Ph.D. in Computer Science, supervised by Prof. Alexander Gelbukh and Prof. Olga Kolesnikova. My research focuses on NLP, with a specialization in multilingual hate speech detection in low-resource languages, LLMs, and explainable AI (XAI). I am passionate about leveraging machine learning, deep learning, and transformer-based architectures to address challenges in low-resource language processing. I am a motivated individual with strong organizational, collaborative, and independent work skills.

Skills

- **Languages:** English (Proficient), Amharic (Native), Wolaita (Native)
- **Programming Languages:** Python, C++, C#, SQL, Java, JavaScript
- **Databases:** MySQL, Oracle
- **Web Development:** HTML, PHP, Apache Web Server, CSS
- **Technologies & Frameworks:** PyTorch, Keras, TensorFlow, Git, HuggingFace, NLTK

Work Experience

- **2022-03 – 2022-09**
  Part-Time- IT Expert
  Awash Bank, Sodo, Ethiopia
  - Managed and maintained IT systems, ensuring operational efficiency.
  - Provided technical support and troubleshooting for banking software and hardware.
- **2017-03 – 2022-08**
  Lecturer
  Wachemo University, Hosanna, Ethiopia
  - Taught courses in Information Technology and Computer Science.
  - Supervised student projects and research activities.

- **2019-01 – 2020-01**
  Part-Time Lecturer
  Wolaita Sodo University, Sodo, Ethiopia
  - Delivered lectures and workshops on IT-related topics.
  - Assisted in curriculum development and academic planning.

Education

- **2024-04 – 2024-08**
  Research Stay
  University of Colorado Colorado Springs (UCCS), Colorado, USA
  - Conducting research in NLP and machine learning applications.

- **2022-08 – 2025-11**
  PhD in Computer Science
  Instituto Politécnico Nacional (IPN-CIC), Mexico City, Mexico
  - Research focus: Natural Language Processing (NLP), Machine Learning, and Deep Learning.
  - I am conducting research in NLP, focusing on multilingual hate speech detection for low-resource languages.
  - Developing machine learning, deep learning, and transformer-based models for text classification and disinformation detection.

- **2019-02 – 2021-03**
  Master of Science in Information Technology
  Wolaita Sodo University, Wolayta Sodo, Ethiopia

- **2013-02 – 2016-06**
  Bachelor of Science in Information Technology
  Arba Minch University, Arba Minch, Ethiopia

Research Publications List

- **Evaluating the Effectiveness of XAI Techniques for Encoder-Based Language Models.** *Knowledge-Based Systems*, 2025.
  - MA. Mersha, MG. Yigezu, and Kalita (2025).

- **Odio-BERT: Evaluating domain task impact in hate speech detection.** *Journal of Intelligent & Fuzzy Systems*, 2024.
  - MG. Yigezu, Kolesnikova, Gelbukh, and Sidorov (2024).

- **Detecting multilingual hate speech targeting immigrants and women on Twitter.** *Journal of Intelligent & Fuzzy Systems*, 2024.
  - O. Kolesnikova, MG. Yigezu, Gelbukh, and Sidorov (2024).

- **Ethio-LLM: Multilingual Large Language Models for Five Ethiopian Languages and English, and Ethio Benchmark – A New Benchmark Dataset for Various Downstream NLP Tasks.** *LREC-COLING*, 2024.
  - A. L. Tonja, Azime, MG. Yigezu et al. (2024).

- **Developing MasakhaNEWS – A New Benchmark Dataset for News Topic Classification Covering 16 Languages Widely Spoken in Africa.** *IJCNLP*, 2023.
  - Adelani, A.L. Tonja, MG. Yigezu et al. (2023).

- **Evaluating the Effectiveness of Hybrid Features in Fake News Detection on Social Media.** *IEEE*, 2023.
  - MG. Yigezu, Mehamed, Kolesnikova, Guge, Gelbukh, and Sidorov (2023).

- **Ethio-Fake: Cutting-Edge Approaches to Combat Fake News in Under-Resourced Languages Using Explainable AI.** *ACLing*, 2024.
  - MG. Yigezu, Mersha, Bade, Kalita, Kolesnikova, and Gelbukh (2024).

- **Semantic-Driven Topic Modeling Using Transformer-Based Embeddings and Clustering Algorithms.** *ACLing*, 2024.
  - Mersha, MG. Yigezu, Bade, and Kalita (2024).

- **Bilingual Word-Level Language Identification for Omotic Languages.** *ICAST*, 2023.
  - MG. Yigezu, GY. Bade, AL. Tonja, Kolesnikova, Gelbukh, and Sidorov (2023).

- **Transformer-Based Hate Speech Detection for Multi-Class and Multi-Label Classification.** *IberLEF@ SEPLN*, 2023.
  - MG. Yigezu, O. Kolesnikova, G. Sidorov, and A. Gelbukh (2023).

- **Detecting Fake News Detection in Dravidian Languages Using Deep Learning.** *Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, 2024.
  - MG. Yigezu, AL. Tonja, O. Kolesnikova, G. Sidorov, and A. Gelbukh (2024).

- **Extractive Text Summarization for Wolaytta Language Using Recurrent Neural Network.** *ICTD4A-IEEE*, 2024.
  - Eyobed et al. and MG. Yigezu (2024).

- **Design and Develop Automatic Parts of Speech Tagger for Siltigna Language.** *ICTD4A-IEEE*, 2024.
  - Eyobed et al. and MG. Yigezu (2024).

- **Abusive Comment Detection Using a Deep Learning Approach.** *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, 2023.
  - MG. Yigezu, SA. Kanta, Kolesnikova, G. Sidorov, and A. Gelbukh (2023).

- **Multilingual Hope Speech Detection Using Machine Learning.** *IberLEF@ SEPLN*, 2023.
  - MG. Yigezu, O. Kolesnikova, G. Sidorov, and A. Gelbukh (2023).

- **Word-Level Language Identification in Code-Mixed Kannada-English Texts Using Deep Learning Approach.** *Proceedings of the 19th International Conference on Natural Language Processing (ICON)*, 2022.
  - MG. Yigezu, AL. Tonja, O. Kolesnikova, G. Sidorov, and A. Gelbukh (2022).

- **Transformer-Based Model for Word-Level Language Identification in Code-Mixed Kannada-English Texts.** *Proceedings of the 19th International Conference on Natural Language Processing (ICON)*, 2022.
  - A. L. Tonja, MG. Yigezu, O. Kolesnikova, G. Sidorov, and A. Gelbukh (2022).

- **Ginger disease detection using a computer vision pre-trained model.** *Springer Nature Journal Article*, 2023.
  - Kolesnikova, MG. Yigezu, A.L. Tonja, MM. Woldeyohannis, G. Sidorov, and A. Gelbukh (2023).

- **Word-Level Language Identification in Code-Mixed Kannada-English Texts Using Traditional Machine Learning Algorithms.** *Proceedings of the 19th International Conference on Natural Language Processing (ICON)*, 2022.
  - Tash, MG. Yigezu, Tonja, O. Kolesnikova, G. Sidorov, and A. Gelbukh (2022).

- **Utilizing Deep and Transfer Learning Approaches for Sentiment Analysis.** *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, 2023.
  - MG. Yigezu, O. Kolesnikova, G. Sidorov, and A. Gelbukh (2023).

- **Multilingual Neural Machine Translation for Low-Resource Languages: Ometo-English.** *International Conference on Information and Communication Technology for Development for Africa (ICT4DA)*, 2021.
  - MG. Yigezu, MM. Woldeyohannis, and AL. Tonja (2021).

- **A Parallel Corpora for Bi-Directional Neural Machine Translation for Low-Resource Ethiopian Languages.** *International Conference on Information and Communication Technology for Development for Africa (ICT4DA)*, 2021.
  - Atnafu Lambebo Tonja, Michael Melese Woldeyohannis, and Mesay Gemeda Yigezu (2021).

- **Early Ginger Disease Detection Using Deep Learning Approach.** *Advances of Science and Technology: 9th EAI International Conference, ICAST 2021*, 2022.
  - Mesay Gemeda Yigezu, Michael Melese Woldeyohannis, and Atnafu Lambebo Tonja (2022).

- **Early Ginger Disease Detection Using Deep Learning Approach**. In *Advances of Science and Technology: 9th EAI International Conference, ICAST 2021* (pp. XX-XX). Springer.
  - M. G. Yigezu, Woldeyohannis, M. M., & Tonja, A. L. (2022).

- **Multilingual neural machine translation for low resourced languages: Ometo-english**. *2021 International Conference on Information and Communication Technology for Development for Africa (ICT4DA)* (pp. 89-94). IEEE.
  - M. G. Yigezu, Woldeyohannis, M. M., & Tonja, A. L. (2021, November).

- **A parallel corpora for bi-directional neural machine translation for low resourced Ethiopian languages**. *2021 International Conference on Information and Communication Technology for Development for Africa (ICT4DA)* (pp. 71-76). IEEE.
  - Tonja, A. L., Woldeyohannis, M. M., & M. G. Yigezu, (2021, November).

Professional Reference

1. **Prof. Alexander Gelbukh**
   Full Professor
   Centro de Investigación en Computación (CIC)
   Instituto Politécnico Nacional (IPN)
   Mexico City, Mexico
   Email: gelbukh@gelbukh.com
   PhD Supervisor

2. **Prof. Olga Kolesnikova**
   Professor
   Centro de Investigación en Computación (CIC)
   Instituto Politécnico Nacional (IPN)
   Mexico City, Mexico

Email: okolesnikova@cic.ipn.mx
Phone: +52 55 2177 4791
Co-Advisor

3. **Prof. Grigori Sidorov**
Full Professor
Centro de Investigación en Computación (CIC)
Instituto Politécnico Nacional (IPN)
Mexico City, Mexico
Email: sidorov@cic.ipn.mx
Phone: +52 55 9188 7293
Thesis Committee Member

4. **Prof. Jugal Kalita**
Professor of Computer Science
Department of Computer Science
University of Colorado, Colorado Springs
Colorado, USA
Email: jkalita@uccs.edu
Phone: +1 (719) 255-3432
Research Stay Supervisor

5. **Dr. Atnafu Lambebo Tonja**
Assistant Professor
Department of Computer Science
MBZUAI, Lelapa AI
Abu Dhabi, UAE
Email: atnafu.tonja@mbzuai.ac.ae
Phone: +971 58 517 3270
Research Collaborator