

MAAZ AMJAD

Manuela Senz 136, Coapa, Culhuacan CTM VI, Coyoacan, 04480 Mexico City, Mexico
Github, LinkedIn, Personal Webpage, Google Scholar, Email: h.maazamjad@gmail.com

RESEARCH INTERESTS

- Deception Detection
- Fake News Detection
- Fact Checking
- Rumor Detection
- Computational Linguistics
- Natural Language Processing

EDUCATION

- Ph.D. in Computer Science** *2018 - 2022*
Thesis: Automatic Detection of Fake Content in Digital Media Texts
Computing Research Center, Instituto Politécnico Nacional, Mexico CUM LUADE
Advisors: **Dr.Alexander Gelbukh, Dr.Grigori Sidorov, Dr.Alisa Zhila**
- Master of Science in Applied Mathematics and Informatics (with honors)** *2015 - 2017*
Thesis: Text Classification with Deep Learning Neural Networks
Moscow Institute of Physics and Technology, Russia Overall GPA: 4.87 / 5
- BS(HON) in Mathematics** *2011 - 2015*
Thesis: Applications of Linear Algebra
Government College University, Lahore, Pakistan Overall CGPA: 3.03 / 4

PUBLICATIONS AND CONFERENCE PROCEEDINGS

Journal Publications

- **Maaz Amjad**, Sabur Butt, Alisa Zhila, Grigori Sidorov, Alexander Gelbukh
Survey of Fake News Datasets and Detection Methods for Asian Languages and European Languages. (Accepted, International Journal of Applied and Computational Mathematics (IJACM), **2022**).
- **Maaz Amjad**, Noman Ashraf, Hamza Imam Amjad, Alisa Zhila, Grigori Sidorov
Abusive language detection in Urdu language in Tweets. (Accepted, journal of Acta Polytechnica Hungarica, **2022**).
- **Maaz Amjad**, Noman Ashraf, Alisa Zhila, Grigori Sidorov, Arkaitz Zubiaga, Alexander Gelbukh
Threatening Language Detecting and Threatening Target Identification in Urdu Tweets. (IEEE Access, **2021**).
- **Maaz Amjad**, Grigori Sidorov, Alisa Zhila, Helena Gómez-Adorno, Ilia Voronkov, Alexander Gelbukh
“Bend The Truth”: Benchmark Dataset for Fake News Detection in Urdu Language and its Evaluation. (Journal of Intelligent & Fuzzy Systems, ISSN 1064-1246, **2020**).
- Mubarak Alkhatnai, Hamza Imam Amjad, **Maaz Amjad**, Alexander Gelbukh
Methods and Trends of Machine Reading Comprehension in the Arabic Language. (Computación y Sistemas, Vol. 24, pp. 1607-1615, **2020**).

Conference Publications

- **Maaz Amjad**, Oxana Vitman, Grigori Sidorov, Alisa Zhila, Alexander Gelbukh.
Analysis of Fake News Detection Methods.
8th World Conference on Soft Computing, Baku, springer publication, **2022**.

- **Maaz Amjad**, Sabur Butt, Hamza Imam Amjad, Alisa Zhila, Grigori Sidorov, Alexander Gelbukh. UrduFake@FIRE2021: Shared Track on Fake News Identification in Urdu. In Forum for Information Retrieval Evaluation (FIRE 21). ACM, New York, USA, **2021**.
- **Maaz Amjad**, Alisa Zhila, Grigori Sidorov, Andrey Labunets, Sabur Butt, Hamza Imam Amjad, Oxana Vitman, Alexander Gelbukh. UrduThreat@ FIRE2021: Shared Track on Abusive Threat Identification in Urdu. In Forum for Information Retrieval Evaluation (FIRE 21). ACM, New York, USA, **2021**.
- **Maaz Amjad**, Sabur Butt, Hamza Imam Amjad, Alisa Zhila, Grigori Sidorov, Alexander Gelbukh. Overview of the Shared Task on Fake News Detection in Urdu at FIRE 2021. In CEUR Workshop Proceedings, **2021**.
- **Maaz Amjad**, Alisa Zhila, Grigori Sidorov, Andrey Labunets, Sabur Butt, Hamza Imam Amjad, Oxana Vitman, Alexander Gelbukh. Overview of the shared task on threatening and abusive detection in Urdu at Fire 2021. In CEUR Workshop Proceedings, **2021**.
- **Maaz Amjad**, Grigori Sidorov, Alisa Zhila, Alexander Gelbukh, Paolo Rosso. UrduFake@FIRE2020: Shared Track on Fake News Identification in Urdu. In Forum for Information Retrieval Evaluation (FIRE 20). ACM, New York, USA, **2020**.
- **Maaz Amjad**, Grigori Sidorov, Alisa Zhila, Alexander Gelbukh, Paolo Rosso. Overview of the Shared Task on Fake News Detection in Urdu at FIRE 2020. In CEUR Workshop Proceedings, **2020**.
- **Maaz Amjad**, Yuriria Rodriguez Chavez, Zaryyab Nayab, Alisa Zhila, Grigori Sidorov, Alexander Gelbukh. Analysis of COVID-19 Pandemic Using Artificial Intelligence. In Mexican International Conference on Artificial Intelligence (pp. 65-73). Springer, Cham, **2020**.
- **Maaz Amjad**, Alisa Zhila, Grigori Sidorov. Data Augmentation through Machine Translation for Fake News Detection in the Urdu Language. Proceedings of The 12th Language Resources and Evaluation Conference (LREC 2020), France, **2020**.
- **Maaz Amjad**, Ilia Voronkov, Anna Saenko, Alexander Gelbukh. Comparison of Text Classification Methods using Deep Learning Neural Networks. Proceedings of the 20th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing 2019), France, **2019**.
- **Maaz Amjad**, Zhanibek Kaimuldenov, Ilia Voronkov. Text Classification with Deep Neural Networks. Text Classification with Deep Neural Networks. Proceedings of the 5th International Conference on Actual Problems of System and Software Engineering (APSSE 2017), Russia, **2017**.

Working Papers

- **Maaz Amjad**, Alisa Zhila, Grigori Sidorov. Feature Engineering: The most prominent features to detect Fake News
- **Maaz Amjad**, Jessica López Espejel, Alisa Zhila, Grigori Sidorov. Generate Headlines from a given text using transformers

RESEARCH PROJECTS

- **2021: Threat and Abusive language detection using Tweeter postings**
 - Collected, and annotated tweets in Urdu language, which is a low-resource language

- Designed and developed socially-aware Artificial Intelligence systems on Urdu corpus
- Evaluated the performance of the system on unseen dataset
- Investigated the hidden pattern according to linguistic point of view and features engineering
- Published the obtained results in a high impact factor journal
- **2020: Word-embeddings for Urdu Language**
 - Assembled Urdu corpus for threatening tweets and developed the guidelines for data annotation
 - Data pre-processing and text features creations
 - Implemented the machine learning pipeline on Urdu corpus containing threatening tweets
 - Data and linguistic /NLP analysis to examine the hidden pattern according to the linguistic point of view
 - Features engineering using machine learning algorithms
 - Published the obtained results in a high impact factor journal
- **2019: Fake News Detection in Urdu**
 - Data Collection, and annotation of news corpus
 - Designed and developed machine learning pipeline to identify fake news
 - Evaluated the performance of the system on unseen dataset
 - Features engineering using machine learning to determine false positives
 - Investigated the hidden pattern according to linguistic point of view and features engineering
 - Published the obtained results in a high impact factor journal
- **2020: Exploring Data Augmentation through MT approach for Fake News detection in low-resource languages**
 - Data Augmentation through machine translation is used to increase the dataset size
 - Analyze whether the translation of the news article from one language to another changes the veracity of the news article
 - Designed and developed machine learning pipeline to identify fake news
 - Evaluated the performance of the system on unseen dataset
 - Features engineering using machine learning algorithms
 - Published the obtained results in a well-known conference
- **2017: Text classification with Deep Machine Learning**
 - IMDB dataset was used for sentiment classification task
 - Analyzed whether the sentiment of the movie reviews were positive or negative
 - Designed and developed machine learning pipeline to solve the binary classification problem
 - Evaluated the performance of the different systems on unseen dataset
 - Published the obtained results in a well-known conference

ORGANIZATION OF INTERNATIONAL SHARED TASKS

- **Fake News Detection Competition 2021:** “Fake News Detection in the Urdu Language”, CICLing 2021 UrduFake track @ FIRE.
I am the main organizer of this competition, I set up evaluation; 34 teams reached out from many countries for datasets and 18 teams submitted the results on test dataset. The publications are in progress about this task are in progress. **Link:** <https://www.urdufake2021.cicling.org>
- **Threatening and Abusive Language Detection 2021:** “Abusive and Threatening Language Detection Task in Urdu”, CICLing 2021 UrduFake track @ FIRE.
I am the main organizer of this competition, I set up evaluation; 21 teams reached out from many countries for datasets and 9 teams submitted the results on test dataset. The publications are in progress about this task are in progress. **Link:** <https://www.urduthreat2021.cicling.org>
- **Fake News Detection Competition 2020:** “Fake News Detection in the Urdu Language”, CICLing 2020 UrduFake track @ FIRE.
I am the main organizer of this competition, I set up evaluation; 39 teams reached out from many countries for datasets and 9 teams submitted the results on test dataset. The publications are in progress about this task are in progress. **Link:** <https://www.urdufake2020.cicling.org>

INVITED TALKS

- **February 2020:** Deceiving AI is not that easy: a case for fake news detection.
IT University of Copenhagen, Denmark
- **March 2020:** Pragmatist Media Ethics and the Challenges of Fake News.
DPS & The Technical University of Munich (TUM), Germany
- **August 2019:** Automatic Detection of Fake Content in Digital Media Texts.
International Workshop on Computational Linguistics, towards the deep analysis of digital documents, Mexico

GRANTS FOR RESEARCH PROJECTS

- **2021:** Financial Support for project “Automatic Learning Applied To Clinical Notes Texts For Skin Diseases Detection”, project ID: 20211784, worked as a participant, Sponsored by
 - Secretary for Research and Postgraduate Studies (IPN), Mexico
- **2020:** Financial Support for project “Automatic Analysis Of Changes In Writing Style Of Text Authors”, project ID: 20200797, worked as a participant, Sponsored by
 - Secretary for Research and Postgraduate Studies (IPN), Mexico

HONOUR AND ACADEMIC AWARDS

- **2018 - 2022:** Full Tuition Waiver Government Scholarship for PhD from
 - National Council of Science and Technology of Mexico (CONACYT), Mexico
- **2020 - 2022:** Institutional Stimulus Scholarship for Researcher Training (BEIFI) from
 - Ministry of Research and Postgraduate, Mexico
- **2021:** NeurIPS Financial Support, Sponsored by
 - Naver Labs Europe, DeepMind and Google AI
- **2021:** NAACL 2021 Financial Support, Sponsored by
 - Microsoft, Twitter, Amazon, Apple and Association for Computational Linguistics

- **2020:** NeurIPS Financial Support, Sponsored by
 - Amazon
- **2015 - 2017:** Full Tuition Waiver Government Scholarship for Masters from
 - Ministry of Education and Science of Russian Federation, Russia
- **2016 - 2017:** \$1,000 as additional funding on master's performance (PGAS) from
 - Ministry of Education and Science of Russian Federation, Russia
- **2011 - 2015:** Punjab benevolent fund scholarship for BS (HON) from
 - The Government of Punjab, Pakistan
- **2014 - 2015:** Certificate of Merit from GC University, Lahore, Pakistan as
 - President of Chawala Mathematical Society, Pakistan
- **2013 - 2014:** Certificate of Merit from GC University, Lahore, Pakistan as
 - General Secretary of GC Library Society, Pakistan
- **2012 - 2014:** Certificate of Merit from GC University, Lahore, Pakistan as
 - Associate Editor of Scientific Ravi, Pakistan
- **2011 - 2012:** Got Laptop on HSC Marks base from
 - The Government of Punjab, Pakistan

PROFESSIONAL EXPERIENCE AND RESEARCH INTERNSHIPS

- **November 2021:** Data Scientist, VALOREO, Mexico
 - Data analysis
 - Future sales prediction
 - Applied data science, machine learning, and engineering skills to developed a prototype to identify top selling products
- **October 2020 - December 2020:** Artificial Intelligence Engineer, EFINFO, Mexico
 - Data collection (print and electronic media) and designed pipeline for data warehouse
 - Data analysis and created dashboards
 - Landing page prototyping for the proof-of-concept
 - Features extraction to determine public opinion about personalities /organisations
 - Applied data science, machine learning, and engineering skills to develop a prototype for opinion mining.
- **January 2020 - March 2020:** Artificial Intelligence Engineer, UnternehmerTUM, Germany
 - Developed various algorithms to identify the hidden data patterns and dependencies
 - Develop pioneering proof-of-concept (POC) for data-driven applications to detect deep fakes
 - Applied data science, machine learning, and engineering skills to developed a prototype to identify fake content
 - Agile product development (Scrum, retrospectives, planning, sprint reviews) and lean startup approach for risk validation through experiments

- Interdisciplinary and cross-functional teamwork for AI based product development
- **August 2019 - December 2019:** Moscow Institute of Physics and Technology, Russia
 - Reviewed Literature of abusive and threatening language detection
 - Defined procures for dataset collection and annotation
 - Designed and developed machine learning algorithms
 - Features engineering using machine learning algorithms
 - Evaluated the performance of the system on unseen dataset
- **July 2016 - August 2016:** Summer School on Neuro-science, brain structure, Russia
 - Developed understanding of neural networks
 - Attended seminars on neuro-science, brain structure
 - Developed understanding of how to develop machine learning algorithms pipeline

REVIEWER FOR JOURNALS AND CONFERENCES

Journals

- Computer Speech & Language, Reviewer from 2022
- Journal of Ambient Intelligence and Humanized Computing, Reviewer from 2022
- Language Resources and Evaluation, Reviewer from 2022
- IEEE Intelligent Systems, Reviewer from 2022
- Jordanian Journal of Computers and Information Technology, Reviewer from 2022
- International Journal of Data Science and Analytics (JDSA), Reviewer from 2020
- Journal of Intelligent & Fuzzy Systems, Reviewer from 2020
- Transactions on Asian and Low-Resource Language Information Processing, Reviewer from 2020
- HKS Misinformation Review (Harvard Kennedy School), Reviewer from 2020
- Information Processing and Management, Reviewer from 2019
- Computación y Sistemas (English translation: Computing and Systems), Reviewer from 2018
- Research in Computing Science, Reviewer from 2018

Conferences

- NAACL, Reviewer from 2022
- COLING, Reviewer from 2022
- EMNLP, Reviewer from 2020
- LKE, Reviewer from 2020
- MICAI, Reviewer from 2018
- CICLing, Reviewer from 2018
- CORE, Reviewer from 2019

CO-CURRICULAR ACHIEVEMENTS

2018 - CURRENT	President of CIC International Students Association (CIC-ISA), Mexico https://nlp.cic.ipn.mx/isa
2017-2018	President of Skoltech Badminton Club, Russia
2015-2017	International student representative at MIPT, Russia
2014-2015	President of Chawala Mathematical Society, Pakistan
2014-2015	Senior Vice President of Environmental Protection Society, Pakistan
2014-2015	General Secretary of Library Society, Pakistan
2014-2015	Member of Parliament at Youth General Assembly of Pakistan, Pakistan
2013-2014	Associate Editor of Scientific Ravi, Pakistan
2013-2014	Senior Vice president of Horticultural Society, Pakistan
2012-2013	Vice president of Blood Donor Society, Pakistan

TEACHING EXPERIENCE

SEP - DEC 2019	Course Instructor for “Text Processing” Moscow Institute of Physics and Technology (State University), Russia
DEC 2017 - JAN 2018	Course Instructor for “Communication Skills for Engineering Leaders” Skolkovo Institute of Science and Technology, Russia
2012-2015	Teaching of Applied Math (Under-Graduate Level)
2010-2013	Mathematics instructor at The Punjab City School
2010-2012	Instructing Mathematics at Little Angels Pilot School

MENTORSHIP

- **Artificial Intelligence Mentorship** - I advise students about problem identification, find appropriate dataset for the task, machine and deep learning algorithms, machine learning model deployment at Digital Product School, Germany teams in exploring AI applications, **Germany**.
- **Natural Language Processing /Text processing Mentorship** - I advise Masters students about natural language processing, text processing. I advise students how to formulate research problem, identify research gaps, find appropriate dataset for the task, develop machine and deep learning algorithms at:
 - Computing Research Center, Instituto Politécnico Nacional, Mexico
 - Moscow Moscow Institute of Physics and Technology (State University), Russia

MEMBERSHIPS

- Mexican Association for Natural Language Processing
- Association for Computational Linguistics (ACL)
- Institute of Electrical and Electronics Engineers (IEEE)
- Association for Computing Machinery (ACM)

ORGANIZATION AND ADMINISTRATION

- **2020** Program Committee of CORE conference
- **2020** Program Committee of MICAI conference
- **2019** International Committee chair of CORE conference
- **2016 - 2017** Brand Ambassador at Moscow Institute of Physics and Technology - MIPT, Russia

LANGUAGES

URDU: Mothertongue
ENGLISH: Proficient, C2 Level
SPANISH: Intermediate Knowledge, B1 Level
RUSSIAN: Intermediate Knowledge, B1 Level
GERMAN: Basic Knowledge, A1 Level
ARABIC: Basic Knowledge, A1 Level

TECHNICAL SKILLS

Competence: PYTHON, L^AT_EX, MAPLE, MATHEMATICA, LINUX, HTML, CSS, GIT, LIWC
Intermediate Knowledge: C++, MATLAB, TABLEAU, SQL
Basic Knowledge: FORTRAN, UXPIN
Libraries: **Pandas, NumPy, spaCy, SciPy, Seaborn, scikit-learn, Keras, TensorFlow, Transformers, PyTorch, PyMongo, Plotly**