Wissam **Antoun**

MACHINE LEARNING SCIENTIST · GRADUATE RESEARCHER

Beirut, Lebanon

☐ (+961) 71-778924 | ✉ wissam.antoun@gmail.com | ⌂ wissamantoun.github.io | ▦ wissamantoun | 🎓 Wissam Antoun

Summary

I'm an Artificial Intelligence and Machine Learning Scientist specialized in state-of-the-art Natural Language Processing (NLP) techniques. My research focuses mainly on language modeling, conversational agents and on the intersection of NLP and Information Retrieval. The output of my **2 years of Master's research** includes **7 peer-reviewed papers published in top NLP conferences** and a nomination for the Abdul Hadi Debs Endowment **Award for Academic Excellence**.

I recently **ranked first** among 74 teams in the largest Arabic Sentiment Analysis competition. Moreover, I have been reviewing papers for international peer-reviewed journals and venues in the AI and NLP domain. I gave multiple NLP workshops as an **AI instructor at Zaka**, in addition to my teaching experience that i gained as a **teaching assistant at the American University of Beirut**. I'm **currently a Data Scientist at Siren Analytics**, making use of my AI/NLP skills to solve various problems that directly impacts the public sector.

Education

Masters Of Engineering In Electrical And Computer Engineering *Beirut, Lebanon*

AMERICAN UNIVERSITY OF BEIRUT *Feb. 2018 - Sep. 2020*

- Major Area: Artificial Intelligence and Machine Learning systems
- Minro Area: Software, Networking and Security
- Awarded Graduate Fellowship with full tuition coverage
- Nominated for the Abdul Hadi Debs Endowment Award for Academic Excellence

Bachelor Of Engineering In Computer And Communication Engineering *Zouk Mosbeh, Lebanon*

NOTRE DAME UNIVERSITY - LOUAIZE *Sept. 2014 - Jul. 2017*

- GPA: 3.58 Dean's List for 6 semesters

Work Experience

Siren Analytics *Beirut, Lebanon*

DATA SCIENTIST *Jan. 2021 - Current*

- Worked on information retrieval, search engines, named-entity recognition, face-recognition, optical character recognition, anomaly detection systems for lockdown mobility permits, and data analysis as part of an effort to create the first e-gov platform in lebanon.

MIND Lab - American University of Beirut *Beirut, Lebanon*

GRADUATE RESEARCH ASSISTANT *Sep. 2018 - Jan. 2021*

- Co-founded AUB's Machine INtelligence Development Lab
- My research focused on state-of-the-art NLP technologies for natural language understanding, generation, and chatbots.
- Worked on pre-training various Arabic language models that ranges from RNN-based to the Transformer-based models, using Cloud TPU and GPU instances.
- Researched improving temporal and spacial resolution of remote sensing ML systems for soil moisture prediction.

ECE Dept. - American University of Beirut *Beirut, Lebanon*

GRADUATE TEACHING ASSISTANT *Feb. 2018 - Dec. 2020*

- EECE 435L: Software Tools Lab, Co-Instructor (3 semesters). Delivered a range of teaching and assessment activities including tutorials directed towards the delivery of modern Software Tools at the undergraduate level.
- EECE 696: Applied Parallel Programming, Lab instructor (1 semester). Delivered hands-on experience to graduate students on parallel programming and GPU computing using the CUDA C/C++ framework on local computers and on AWS.
- EECE 330: Data Structures and Algorithms, Lab instructor (1 semester). Delivered hands-on experience on fundamental algorithms and data structures used in software applications at the undergraduate level using C++.
- Introduced a plagiarism detection tool for computer code to the Electrical Engineering Department.

Academic Reviewer

REVIEWER FOR THE FOLLOWING JOURNALS: *2020 - Current*

- Journal of King Saud University - Computer and Information Sciences (IF: 13.47)
- Expert Systems with Applications (IF: 6.95)
- Workshop on Arabic Natural Language Processing (h-index: 51, i10-index: 50)
- ACM Transactions on Asian and Low-Resource Language Information Processing - TALLIP (IF: 1.42)
- Language Resources and Evaluation (IF: 1.36)

Zaka *Online*

AI INSTRUCTOR *May. 2020*

- Gave a 4-part online workshop series titled *A Comprehensive NLP series* that was focused on delivering the major NLP concepts through hands-on coding examples along with the NLP fundamentals theory.

Stars of Science *Doha, Qatar*

MACHINE LEARNING EXPERT *Feb. 2020 - Apr. 2020*

- Provided Machine Learning and Artificial Intelligence solutions as part of the support team that helps contestants throughout their Stars of Science journey

Huawei Technologies (Lebanon) S.A.R.L. *Beirut, Lebanon*

INTERN *June 2016 - Aug. 2016*

- Learned about various LTE principles, base station hardware, and Microwave transmissions systems.

Honors & Awards

INTERNATIONAL

2021	**First Place**, Arabic Sentiment Analysis 2021 @ KAUST (wti.kaust.edu.sa/solve/Arabic-Sentiment-Analysis-Challenge)	*Kaggle, Online*
2020	**Second Place**, OSACT4-Shared task on Offensive Language Detection	*Marseille, France*
2019	**Best Domain Detection system** , Qatar International Fake News Detection and Annotation Contest	*Doha, Qatar*

DOMESTIC

2021	**Nominee**, Abdul Hadi Debs Endowment Award for Academic Excellence	*AUB, Lebanon*
2019	**Second Place**, Mindsets-OLX Datathon, Best car price prediction score.(www.kaggle.com/c/mindsets)	*AUB, Lebanon*
2019	**First place**, Big Data, AI, and Media Hackathon Powered by Touch and Anghami	*AUB, Lebanon*

Publications

My full publication list with links and up-to-date citation scores is available on my Google Scholar page (linked above).

My Research metrics:

- **Citation score: 261**
- **h-index: 6**
- **i10-index: 5**

AraGPT2: Pre-Trained Transformer for Arabic Language Generation
Antoun, **Wissam**, Fady Baly, and Hazem Hajj — 2021
Proceedings of the Sixth Arabic Natural Language Processing Workshop co-located with EACL, Kyiv, Ukraine — **Cited by: 6**

AraELECTRA: Pre-Training Text Discriminators for Arabic Language Understanding
Antoun, **Wissam**, Fady Baly, and Hazem Hajj — 2021
Proceedings of the Sixth Arabic Natural Language Processing Workshop co-located with EACL, Kyiv, Ukraine — **Cited by: 6**

Empathetic BERT2BERT Conversational Model: Learning Arabic Language Generation with little data
Naous, Tarek, **Wissam Antoun**, Reem Mahmoud, and Hazem Hajj — 2021
Proceedings of the Sixth Arabic Natural Language Processing Workshop co-located with EACL, Kyiv, Ukraine — **Cited by: 1**

AraBERT: Transformer-based Model for Arabic Language Understanding
Antoun, **Wissam**, Fady Baly, and Hazem Hajj — 2020
Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools co-located with LREC 2020, Marseille, France — **Cited by: 177**

Multi-Task Learning using AraBert for Offensive Language Detection
Djandji, Marc, Fady Baly, **Wissam Antoun**, and Hazem Hajj — 2020
Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools co-located with LREC, Marseille, France — **Cited by: 11**

State of the Art Models for Fake News Detection Tasks
Antoun, **Wissam**, Fady Baly, Rim Achour, Amir Hussein, and Hazem Hajj — 2020
2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT'2020), Doha, Qatar — **Cited by: 8**

hULMonA: The Universal Language Model in Arabic
ElJundi, Obeida, **Wissam Antoun**, Nour El Droubi, Hazem Hajj, Wassim El-Hajj, and Khaled Shaban — 2019
Proceedings of the Fourth Arabic Natural Language Processing Workshop co-located with ACL, Florence, Italy — **Cited by: 26**

Real-Time Fire Detection Using CUDA (Poster)
Antoun, **Wissam** and Manal Jalloul — 2019
ISC High Performance, Frankfurt, Germany — **Cited by: 1**

A Comprehensive Approach for A Smart Medication Dispenser
Kassem, Abdallah, **Wissam Antoun**, Mustapha Hamad, and Chady El-Moucary — 2019
International Journal of Computing and Digital Systems, University of Bahrain — **Cited by: 10**

Smart Medicine Dispenser (SMD)

Antoun, **Wissam**, Ali Abdo, Suleiman Al-Yaman, Abdallah Kassem, Mustapha Hamad, and Chady El-Moucary   2018
2018 IEEE 4th Middle East Conference on Biomedical Engineering (MECBME), IEEE   **Cited by: 15**

Presentations

AraBERT: pretraining, use cases and future improvements   *Online*

PYDATA KHOBAR AND KSU's IWAN RESEARCH GROUP MEETUPS   *May. 2020*

- Introduced the AraBERT model for the Arab community and explained the pretraining and fine-tuning procedure

High Performance Computing for Lebanon (HPC4L) Brown Bag   *AUB, Lebanon*

PRESENTER FOR CURRENT AND FUTURE/POTENTIAL USE CASES FOR HPC4L   *Feb. 2020*

- Introduced the use case of HPC for pretraining transformer-based model for Arabic language Understanding
- Explained parallelization and scaling techniques for transformer models

Skills

Python | TensorFlow | PyTorch | Transformers | Elastic Stack | Keras | Docker | Rasa | Bash | Git | CUDA | Parallel Programming
C++ | Matlab | Java | SQL | LaTex | VSCode