

# Wissam Antoun

DATA SCIENTIST · GRADUATE RESEARCHER

Beirut, Lebanon

☎ (+961) 71-778924 | ✉ [wissam.antoun@gmail.com](mailto:wissam.antoun@gmail.com) | 🏠 [wissamantoun.github.io](https://github.com/wissamantoun) | 💻 [wissamantoun](#) | 🎓 Wissam Antoun

## Summary

Artificial Intelligence and Machine Learning Engineer specialised in state-of-the-art Natural Language Processing techniques. Within 1 year, I had developed my knowledge of NLP and published one of the most famous and powerful AI models for Arabic text representation. I also ranked in top places in almost all AI competitions that I participated in. In addition, I have 2 years of experience in teaching as a Graduate Teaching Assistant at the Eng. faculty at AUB and through online sessions. I'm currently a Data Scientist at Siren Analytics, working on various problems that directly impacts the public sector. I'm self-motivated, self-taught and a performer when working on tight deadlines. I'm also a tech geek who loves computer hardware and technology, and enjoys DIY projects and gaming.

## Work Experience

### Siren Analytics

Beirut, Lebanon

DATA SCIENTIST

Jan. 2021

- Worked on Anomaly Detection systems for lockdown mobility permits, Chatbots, and Smart Search systems.

### 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. Concepts ranged from Bag-of-Words, Word Embeddings, Advanced Deep-learning Models for text, and transformer-based architectures

### 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 each contestant throughout their Stars of Science journey

### Academic Reviewer

REVIEWER FOR THE FOLLOWING JOURNALS:

2020 - Current

- Expert Systems with Applications (IF: 5.45)
- Journal of King Saud University - Computer and Information Sciences (IF: 3.98)
- 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: 1.014)

### MIND Lab (Prof. Hazem Hajj) - American University of Beirut

Beirut, Lebanon

GRADUATE RESEARCH ASSISTANT

Sep. 2018 - Current

- My research focuses on state-of-the-art NLP technologies for natural language understanding, generation, and social 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.

### 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. Prepared lab tutorials, graded assignments and provided support for the student projects.
- 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. Provided office hours, assignment grading and support for course project.
- 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++. Provided weekly assessments, feedback, and support.
- Introduced a plagiarism detection tool for computer code to the Electrical Engineering Department.

### Huawei Technologies (Lebanon) S.A.R.L.

Beirut, Lebanon

INTERN

June 2016 - Aug. 2016

## Publications

AraGPT2: Pre-Trained Transformer for Arabic Language Generation

**Antoun, Wissam, Fady Baly, and Hazem Hajj**

*Proceedings of the Sixth Arabic Natural Language Processing Workshop at EACL 2021, 2020*

AraELECTRA: Pre-Training Text Discriminators for Arabic Language Understanding

**Antoun, Wissam, Fady Baly, and Hazem Hajj**

*Proceedings of the Sixth Arabic Natural Language Processing Workshop at EACL 2021, 2020*

Empathetic BERT2BERT Conversational Model: Learning Arabic Language Generation with little data

Naous, Tarek, **Wissam Antoun**, Reem Mahmoud, and Hazem Hajj

*Proceedings of the Sixth Arabic Natural Language Processing Workshop at EACL 2021, 2021*

AraBERT: Transformer-based Model for Arabic Language Understanding

**Antoun, Wissam**, Fady Baly, and Hazem Hajj

*Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools at LREC 2020, 2020, Marseille, France*

Multi-Task Learning using AraBert for Offensive Language Detection

Djandji, Marc, Fady Baly, **Wissam Antoun**, and Hazem Hajj

*Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools at LREC 2020, 2020, Marseille, France*

State of the Art Models for Fake News Detection Tasks

**Antoun, Wissam**, Fady Baly, Rim Achour, Amir Hussein, and Hazem Hajj

*2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT) (ICIoT'2020), 2020, Doha, Qatar*

hULMonA: The Universal Language Model in Arabic

ElJundi, Obeida, **Wissam Antoun**, Nour El Droubi, Hazem Hajj, Wassim El-Hajj, and Khaled Shaban

*Proceedings of the Fourth Arabic Natural Language Processing Workshop at ACL, 2019, Florence, Italy*

A Comprehensive Approach for A Smart Medication Dispenser

Kassem, Abdallah, **Wissam Antoun**, Mustapha Hamad, and Chady El-Mou Cary

*International Journal of Computing and Digital Systems 8.02 (2019) PP. 131–141. UNIVERSITY OF BAHRAIN, 2019*

Smart Medicine Dispenser (SMD)

**Antoun, Wissam**, Ali Abdo, Suleiman Al-Yaman, Abdallah Kassem, Mustapha Hamad, and Chady El-Mou Cary

*2018 IEEE 4th Middle East Conference on Biomedical Engineering (MECBME), 2018*

Real-Time Fire Detection Using CUDA

**Antoun, Wissam** and Manal Jalloul

*ISC High Performance, Poster Presentation, 2019, Frankfurt, Germany*

## Honors & Awards

### INTERNATIONAL

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

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*

## Presentations

### High Performance Computing for Lebanon (HPC4L) Brown Bag

*Beirut, 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

### PyData Khobar meetup

*Online*

PRESENTER FOR ARABERT: PRETRAINING, USE CASES AND FUTURE IMPROVEMENTS

*May. 2020*

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

## 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
- Minor Area: Software, Networking and Security
- Awarded Graduate Fellowship with full tuition coverage

### Bachelor Of Engineering In Computer And Communication Engineering

*Zouk Mosbeh, Lebanon*

NOTRE DAME UNIVERSITY - LOUAIZE

*Fall. 2014 - Summer 2017*

- GPA: 3.58 Dean's List for 6 semesters

### CCNA Routing And Switching: Introduction To Networks

*Zouk Mosbeh, Lebanon*

CISCO NETWORKING ACADEMY- NOTRE DAME UNIVERSITY

*Sep 2016 - Jan 2017*

## Skills

Python TensorFlow PyTorch Keras Docker Rasa Bash Git CUDA Parallel Programming C++ C# Matlab Java  
SQL LaTeX VSCode