

Chadi HELWE

✉ chadi.helwe@telecom-paris.fr | ☎ +33 7 54 17 39 45 | in chadihelwe | 🌐 ChadiHelwe

EDUCATION

MAY 2024	Doctor of Philosophy in Artificial Intelligence - Institut Polytechnique de Paris	
2021	Ph.D.'s Thesis: Evaluating and Improving Reasoning Abilities of Large Language Models Advisors: Prof. Fabian Suchanek and Prof. Chloé Clavel	
2017	Master of Science in Computer Science - American University of Beirut	GPA: 3.68/4
2015	Master's Thesis: Arabic Named Entity Recognition via Deep Co-learning Advisor: Prof. Shady Elbassuoni	
2014	Bachelor of Science in Computer Science - Notre Dame University - Louaize	GPA: 3.63/4
2010	Final Year Project: Educally - An Educational Social Network Advisor: Prof. Marie Khair	

SELECTED WORK EXPERIENCE

PRESENT	Graduate Research Assistant - Institut Polytechnique de Paris, France	
JAN. 2021	<ul style="list-style-type: none">• MAFALDA: A Benchmark and Comprehensive Study of Fallacy Detection and Classification:<ul style="list-style-type: none">– Designed a novel taxonomy of fallacies.– Proposed an annotation framework customized for fallacy classification and introduced an evaluation metric to address subjectivity within fallacy classification.– Built a benchmark to assess the performance of language models in detecting and categorizing fallacies.• TINA: Textual Inference with Negation Augmentation [1]:<ul style="list-style-type: none">– Proposed a probabilistic definition of textual entailment (TE) and used it to augment TE datasets with new entailment relationships automatically.– Explored various models fine-tuned with augmented datasets using an unlikelihood loss to enhance language model robustness in textual entailment tasks, particularly those involving negation examples.• Logitorch: A PyTorch-based Library for Logical Reasoning on Natural Language [2]:<ul style="list-style-type: none">– Developed a Python library on top of Pytorch for logical reasoning with diverse benchmarks, models, and utility functions, simplifying dataset utilization and model training.• Reasoning with Transformer-based Models: Deep Learning, but Shallow Reasoning [3]:<ul style="list-style-type: none">– Wrote as survey paper discussing the performance of transformers on different reasoning tasks, including mathematical reasoning, commonsense reasoning, and logical reasoning. This paper highlights both the successes and limitations of these models, covering empirical and theoretical aspects.	
DEC 2020	Research Engineer - Institut Polytechnique de Paris, France	
SEP 2020	Investigated performance of transformers on different reasoning tasks [3].	
AUG 2020	Research Assistant - American University of Beirut, Lebanon	
OCT 2017	<ul style="list-style-type: none">• Retrieving Textual Evidence for Knowledge Graph Facts using Deep Learning:<ul style="list-style-type: none">– A project in collaboration with Aalborg University (Denmark) and Hacettepe University (Turkey).– Investigated transformer-based models trained with a dataset generated using distant supervision to rank passages based on their relevance to a given fact in the form of a Resource Description Framework (RDF) triple.• Automated Detection and Measurement of Corneal Haze and Demarcation Line in OCT Images [4, 5, 6, 7]:<ul style="list-style-type: none">– A project in collaboration with the Department of Ophthalmology (American University of Beirut Medical Center) and the ELZA Institute in Zurich (Switzerland).– Redesigned an outdated software to detect and measure corneal haze and demarcation line in different types of Optical Coherence Tomography (OCT) images.– Added new features and introduced deep learning methods to the software.– Developed OCTAnalysis.com a web interface of the software in Django/Python and Postgres SQL.– Proposed and implemented a semi-weakly supervised learning approach to segment the area between the top boundary of a cornea and the demarcation line in OCT images.– Developed a SegNet neural network to detect the boundaries of a cornea in OCT images and a VGG-16 neural network to detect artifacts in OCT images.– Supervised two undergrad students who annotated a large dataset of OCT images.– Developed an image segmentation tool that is used for labeling.• Predicting Arabic Blog Credibility using Deep Co-learning [8]:<ul style="list-style-type: none">– Implemented and evaluated a novel semi-supervised learning approach based on an algorithm called Co-training, which was adapted to the context of deep learning for the task of Arabic blog's credibility prediction and which can be trained using a small labeled dataset and a large unlabeled dataset.	

SEPT 2017	Graduate Research Assistant - American University of Beirut, Lebanon
JUNE 2016	<ul style="list-style-type: none"> • Arabic Named Entity Recognition via Deep Co-learning [9, 10]: <ul style="list-style-type: none"> – Built a supervised deep learning model that infers the name entities' class in a Wikipedia article by classifying their Wikipedia pages into one of four classes: person, location, organization, or miscellaneous. – Generated a large dataset of partially annotated Wikipedia articles for the task of Arabic Named Entity Recognition (NER). – Proposed a novel semi-supervised learning approach based on an algorithm called Co-training, which was adapted to the context of deep learning for the task of Arabic NER and which can be trained using a small fully annotated dataset and a large partially annotated dataset. – Evaluated our proposed approach, the Deep Co-learning algorithm, on three different Arabic NER datasets. • ICD and CCS Coding using Deep Learning [11]: <ul style="list-style-type: none"> – A project in collaboration with the Department of Emergency Medicine (American University of Beirut Medical Center). – Designed and implemented a deep neural network architecture to predict the International Classification of Diseases (ICD) code and Clinical Classifications Software (CCS) single level code of a discharge diagnosis. • Methodical Evaluation of Arabic Word Embeddings [12]: <ul style="list-style-type: none"> – A project in collaboration with Qatar University. – Built the first word analogy benchmark designed specifically for Arabic word embeddings. – Implemented different Long Short-term Memory recurrent neural network architectures to evaluate Arabic word embeddings on two NLP tasks: Document Classification and Named Entity Recognition. • Adaptive QoS for Spark Applications [13]: <ul style="list-style-type: none"> – Developed an adaptive quality management/selection method for Spark applications. – Implemented different QoS policies in Java.

TEACHING EXPERIENCE

AUG 2023	Teaching Assistant - Institut Polytechnique de Paris, France
JULY 2021	<ul style="list-style-type: none"> • Bases de données • Données du Web • Mise en pratique Données du Web • Mining of Large Datasets
MAY 2020	Teaching Assistant - American University of Beirut, Lebanon
FEB 2015	<ul style="list-style-type: none"> • Artificial Intelligence • Introduction to Programming • Compiler Construction (graduate course) • Machine Learning (graduate course)
JUNE 2013	Teaching Assistant - Notre Dame University - Louaize, Lebanon
FEB 2013	<ul style="list-style-type: none"> • Program Design and Data Abstraction I • Program Design and Data Abstraction II

OPEN SOURCE PROJECTS

PRESENT	LogiTorch - Creator and Maintainer
DEC 2021	LogiTorch [2] is a PyTorch-based library that includes different logical reasoning benchmarks, different models, as well as utility functions such as co-reference resolution. The library allows researchers and developers to easily use a logical reasoning dataset and train logical reasoning models with just a few lines of code.

SKILLS

Python, Java, C++, PyTorch, Pandas, Numpy, Keras, L^AT_EX

REVIEWER

AI Review, ACL 2023, EMNLP 2022, EACL 2022 (External Reviewer)

AWARDS

2018	Recipient of the Best Computer Science Graduate Student Award from the American University of Beirut.
2015	Awarded a full graduate assistantship from the American University of Beirut.
2014	Graduated from Notre Dame University - Louaize with high distinction.
2013	Awarded a scholarship from Notre Dame University - Louaize.
2013	Dean's List for the Spring semester.
2012	Dean's List for the Spring and the Fall semesters.

SUMMER SCHOOLS ATTENDED

- 2023 Oxford Machine Learning Summer School (OxML), Oxford, United Kingdom.
- 2021 Machine Learning Summer School (MLSS), Taipei, Taiwan.
- 2021 4th Advanced Course on Data Science and Machine Learning (ACDL), Tuscany, Italy.

PUBLICATIONS

- [1] Chadi Helwe, Simon Coumes, Chloé Clavel, and Fabian Suchanek. “TINA: Textual Inference with Negation Augmentation”. In: *Findings of the Association for Computational Linguistics: EMNLP 2022*. 2022.
- [2] Chadi Helwe, Chloé Clavel, and Fabian Suchanek. “LogiTorch: A PyTorch-based library for logical reasoning on natural language”. In: *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. 2022.
- [3] Chadi Helwe, Chloé Clavel, and Fabian M Suchanek. “Reasoning with Transformer-based Models: Deep Learning, but Shallow Reasoning”. In: *3rd Conference on Automated Knowledge Base Construction*. 2021.
- [4] Shady Awwad, Lily Chacra, Chadi Helwe, Ahmad Dhaini, Farhad Hafezi, Emilio Torres, and Madeleine Yehia. “Accelerated Corneal Cross-linking Using 20 Minutes Riboflavin With Hydroxypropyl Methylcellulose Soaking Time Versus Conventional Cross-linking”. In: *International CXL Experts Meeting 2019*. 2019.
- [5] Shady Awwad, Lily Chacra, Chadi Helwe, Ahmad Dhaini, Farhad Hafezi, Emilio Torres, and Talar Telvizian. “Corneal Haze After Cross-linking for Keratoconus Eyes With and Without Mitomycin C Application”. In: *International CXL Experts Meeting 2019*. 2019.
- [6] Chadi Helwe, Shady Elbassuoni, Ahmad Dhaini, Lily Chacra, and Shady Awwad. “A Deep Learning Approach to Detect the Demarcation Line in OCT Images”. In: *Annual Conference on Medical Image Understanding and Analysis*. Springer. 2020.
- [7] Shady T Awwad, Lily M Chacra, Chadi Helwe, Ahmad R Dhaini, Talar Telvizian, Julien Torbey, Maamoun Abdul Fattah, Emilio A Torres-Netto, Farhad Hafezi, and Rohit Shetty. “Mitomycin C application after corneal cross-linking for keratoconus increases stromal haze”. In: *Journal of Refractive Surgery* (2021).
- [8] Chadi Helwe, Shady Elbassuoni, Ayman Al Zaatari, and Wassim El-Hajj. “Assessing Arabic Weblog Credibility via Deep Co-learning”. In: *Proceedings of the Fourth Arabic Natural Language Processing Workshop*. Association for Computational Linguistics, 2019.
- [9] Chadi Helwe and Shady Elbassuoni. “Arabic named entity recognition via deep co-learning”. In: *Artificial Intelligence Review* (2019).
- [10] Chadi Helwe, Ghassan Dib, Mohsen Shamas, and Shady Elbassuoni. “A Semi-Supervised BERT Approach for Arabic Named Entity Recognition”. In: *Proceedings of the Fifth Arabic Natural Language Processing Workshop*. Association for Computational Linguistics, 2020.
- [11] Chadi Helwe, Shady Elbassuoni, Mirabelle Geha, Eveline Hitti, and Carla Makhoul Obermeyer. “CCS coding of discharge diagnoses via deep neural networks”. In: *Proceedings of the 2017 International Conference on Digital Health*. 2017.
- [12] Mohammed Elrazzaz, Shady Elbassuoni, Khaled Shaban, and Chadi Helwe. “Methodical evaluation of arabic word embeddings”. In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. 2017.
- [13] Bilal Abi Farraj, Wael Al Rahal Al Orabi, Chadi Helwe, Mohamad Jaber, Mohamad Omar Kayali, and Mohamed Nassar. “Reconfigurable and Adaptive Spark Applications”. In: *Proceedings of the 7th International Conference on Cloud Computing and Services Science*. 2017.