Mohamed Salem Mohamed Elaraby

Pittsburgh PA , USA
Mobile: (+1) 412-214-2403
Email: `mse30@pitt.edu`
Github: github page
Scholar: scholar profile

CURRICULUM EMPHASIS

Natural Language Processing - Summarization- Computational Argumentation.

EDUCATION AND ACADEMIC RECORDS

2021–Present	School of Computer Science **PhD: University of Pittsburgh, USA**
2014–2017	Department of Electronics and Electrical Communications **MSc: Cairo University, Egypt** Thesis: *Advanced Techniques in Speaker Diarization for TV Broadcast*
2009–2014	Department of Communications and Computer Engineering, Credit Hours System **BSc, Cairo University, Egypt** **Graduation Project**: *OBD On Board Diagnostics* OBD systems give the vehicle owner or repair technician access to the status of the various vehicle subsystems. The project was Funded by Valeo Inc.

PROFESSIONAL EXPERIENCE

Sep21–Present	**PETAL Lab: Department of Computer Science** *Graduate Research Assistant.* **CASUM:** Building argument aware summarization models for legal opinions. **CourseMirror:** Improving the summarization of students reflective writing (students opinions about lectures).
May25-Aug25	**Zillow Inc.** *Applied Scientist Intern* Working on hallucination detection in longform text.
May24-Aug24	**Microsoft** *Research Intern at **Research and Open AI Research (ROAR)** team* Worked on automatic redteaming of Large Language Models.
May23-Aug23	**Tiktok Inc. (ByteDance)** *Research Intern (NLP) at **SAMI** team* Developed HaLo: Hallucination and reduction framework in open source LLMs focusing on BLOOM7B

Apr21–Aug21 **Agolo**
Senior Natural Language Processing Engineer.
Headline Generation: Built a key-phrase based Headline Generation.
Evaluation Framework: Built an evaluation API using python to standardize summary evaluation.

Jun 20 – Apr 21 **Microsoft (Advanced Technology Lab Cairo)**
Data Scientist (Contract).
Working on Improving Acoustic Modeling for Arabic Speech Recognition Systems using transfer learning, knowledge distillation and domain adaptation techniques.
• **Arabic Yemeni:** Provided the models for Microsoft new locale Arabic Yemeni.
• **Gulf Assistant Scenario:** Provided the models for assistant scenario in speech recognition.
• **Language Model Augmentation** Provided new techniques to augment Language Models.

Dec 18 – Jun 20 **Raisa Energy LLC.**
Data Scientist II.
Building analytical and predictive models for the oil and gas domain
• *DSU Extraction:* Developing an Information Extraction tool based on state of the art Natural Language Processing techniques to extract entities in legal documents.
• *Parent-Child:* Used Data analytic techniques to study the effect of neighboring wells on the production of the well to be drilled
• *Well-Spacing:* Built an algorithm that automates calculating spacing between wells and return each well neighbor's within 1 Mile.
• *TC-module:* Built a hybrid feature-based and Deep Learning time series forecasting algorithm to predict oil production in undeveloped areas.
• *ML-Interpret:* Supervised an intern in building a tool that explains machine learning model's predictions.

Sept 17–Dec 18 **School of Information The University of British Columbia, Canada**
Research Assistant (Natural Language Processing Lab).
• Arabic Dialect Identification: Introduced a new dataset for city level dialect identification, introduced two new models in Arabic dialect identification.
• Emotion Detection: Contributed in the 2018 WASSA competition , ranked 4th out of 30 in final evaluation

Oct 15–Aug 17 **RDI Research and Development International, Egypt**
Researcher at Speech Processing team at RDI.
• Hafss: Built a Computer Aided Pronunciation learning system for Arabic based on Deep Neural Network acoustic Model.
• MGB Arabic 2016 Challenge:Contributed in building RDI Large Vocabulary system, helped introducing new DNN based Features for Speech Recognition.
• Contributed in RDI Large Vocabulary Speech Recognition Engine.

Sep 14–Oct 15	**German University In Cairo, Egypt**

Teaching Assistant at the Faculty of Information Engineering and Technology.
- Teach to ∼20 students per section in the following courses:

- Digital Logic Design	Fall 2014
- Electric Circuits I	Fall 2014
- Electric Circuits II	Spring 2015
- Laboratory	Spring 2015

- Supervise several undergraduate projects and lab experiments.

July–Sep 14	**Valeo Inc., Egypt**

Graduate Intern — GEEDS Team AutoSAR service line.
- Self-studied many Automotive Concepts (e.g. AutoSaR, CAN protocol and LIN protocol).
- Developed an Automatic tool for testing on different ECUs using configurable hardware automated for testing

PUBLICATIONS
* indicates equal contribution

Preprints

1. **Mohamed Elaraby**, Mengyin Lu, Jacob Dunn, Xueying Zhang, Yu Wang and Shizhu Liu "Halo: Estimation and Reduction of Hallucinations in Open-Source Weak Large Language Models" arxiv 2023

Conference Papers

1. **Mohamed S. Elaraby**, Mostafa Abdallah, Sherif Abdou, and Mohsen Rashwan "A Deep Neural Networks (DNN) Based models for a Computer Aided Pronunciation Learning System ,"in *Speech and Computer Conference SPECOM-2016 (SPECOM) 2016*

2. Muhammad Abdul-Mageed, Hassan Alhuzali, **Mohamed Elaraby** "You Tweet What You Speak: A City-Level Dataset of Arabic Dialects" in *The International Conference on Language Resources and Evaluation LREC 2018*

3. Ahmed Magooda, **Mohamed Elaraby**, Diane Litman "Exploring Multitask Learning for Low Resource Abstractive Summarization" *Findings of the Association for Computational Linguistics: EMNLP 2021*

4. **Mohamed Elaraby**, Diane Litman "ArgLegalSumm: Improving Abstractive Summarization of Legal Documents with Argument Mining" *COLING 2022*

5. **Mohamed Elaraby**, Yang Zhong, Diane Litman "Towards Argument-Aware Abstractive Summarization of Long Legal Opinions with Summary Reranking" *Findings of ACL 2023*

6. **Mohamed Elaraby***, Yang Zhong*, Diane Litman, Ahmed Ashraf Butt and Muhsin Menekse "ReflectSumm: A Benchmark for Course Reflection Summarization" *LREC-COLING 2024*

7. **Mohamed Elaraby**, Diane Litman, Xiang Lorraine Li, Ahmed Magooda "Persuasiveness of Generated Free-Text Rationales in Subjective Decisions: A Case Study on Pairwise Argument Ranking" *Findings of the Association for Computational Linguistics: EMNLP 2024*

Workshop Papers

1. Hany Ahmed, **Mohamed S. Elaraby**, Mostafa Abdallah, Abdullah Mousa, Sherif Abdou, and Mohsen Rashwan "An Unsupervised Speaker Clustering Technique based on SOM and I-vectors for Speech Recognition Systems ,"in *The third Arabic Natural Language Processing Workshop ("WANLP 2017) co-located with ("EACL 2017)*

2. **Mohamed Elaraby**, Muhammad Abdul-Mageed "Deep Models for Arabic Dialect Identification on Benchmarked Data" *"Fifth Workshop on NLP for Similar Languages, Varieties and Dialects VarDial 2018 co-located with COLING 2018*

3. Hassan Alhuzali **Mohamed Elaraby** Muhammad Abdul-Mageed "UBC-NLP at IEST 2018: Learning Implicit Emotion With an Ensemble of Language Models" *9th Workshop on Computational Approaches to Subjectivity, Sentiment Social Media Analysis co-located with EMNLP 2018*

4. **Mohamed Elaraby** Ahmed Zahran "A Character level Convolutional BiLSTM for Arabic Dialect Identification" *In The Fourth Arabic Natural Language Processing Workshop. (WANLP 2019). co-located with ACL 2019, Florence, Italy, July 28-Aug 2*

5. **Mohamed Elaraby**, Diane Litman "Self-trained Pretrained Language Models for Evidence Detection" *Accepted in Argument Mining Workshop Colocated with EMNLP 2021*

6. Zheixong Liu, **Mohamed Elaraby***,Yang Zhong*, Diane Litman "Overview of ImageArg-2023: The first shared task in multimodal argument mining" *Argument Mining Workshop Colocated with EMNLP 2023*

7. **Mohamed Elaraby**, Huihui Xu, Morgan Gray, Kevin Ashley, Diane Litman "Adding Argumentation into Human Evaluation of Long Document Abstractive Summarization: A Case Study on Legal Opinions" *HumanEval 2024 Colocated with LREC-COLING 2024*

8. **Mohamed Elaraby**, Diane Litman "Lessons Learned in Assessing Student Reflections with LLMs" BEA @ ACL2025

Abstracts

1. Hadeer ElAshhab **Mohamed Elaraby** Motaz Elsaban "Modeling hydrocarbons flow from reservoir rocks using Deep Learning" *AI for Earth Science Workshop @ICLR 2020*

Community Services
•Coroganizer of the **ImageArg** shared task.

•Reviewer for LREC 2020, LREC 2022,ArgMining 2023, ARR 2024-present

SELECTED HONORS AND AWARDS

1. LRDC Council Award Grant 2024.

2. Nominated for the **CS50 fellowship** at the department of computer science, the University of Pittsburgh.

3. Undergraduate **Half scholarship** to study in the Credit Hours System (CHS) from Faculty of Engineering, Cairo University, Egypt. [2009-2011]

4. Award of **Excellence** from the Ministry of Education, Egypt. [2009]

REFERENCES

- Dr. Diane J. LITMAN, *Ph.D, Rutgers University.* (Ph.D advisor)

  Department of Computer Science, University of Pittsburgh

  Email: `dlitman@pitt.edu`

- Dr. Mohsen A. RASHWAN, *Ph.D, the University of Queens.* (MSc thesis supervisor, RDI Chariman)

  Department of Electrical and Electronics Engineering, Cairo University, Giza, Egypt

  Email: `mrashwan@rdi-eg.com`

- Dr. Sherif M. ABDOU, *Ph.D, University of Miami , Florida.* (MSc thesis supervisor)

  Department of Information Technology, Cairo University, Egypt.

  Email: `sabdou@rdi-eg.com`

- Dr. Motaz ELSABAN, *Ph.D, University of California at Santa Barbra.* (Chef Technology Officer at Raisa Energy LLC)

  Department of Information Technology, Cairo University, Egypt.

  Email: `melsaban@raisaenergy.com`

- Dr. Ahmed TAWFIK, *Ph.D, University of Saskatoon.* (Direct Manager)

  Microsoft Advanced Technology Lab, Cairo Egypt.

  Email: `atawfik@microsoft.com`