




Areg Mikael Sarvazyan

✉ areg.sarvazyan@gmail.com |  areg-sarvazyan |  Areg Mikael Sarvazyan |  asarvazyan

EXPERIENCE

Junior Research Scientist

Oct 2022 – Present

Symanto

Valencia, Spain

- Developing task-specific models for bot detection, machine generated text detection and attribution, aspect-based sentiment analysis, topic discovery, and image classification, resulting in contracts of over 1M € and decreases in project time delivery by up to 70%.
- Building dialogue assistants for different use cases and domains yielding contracts of up to 1.6M €.
- Creating new techniques to improve LLM inference, zero and few-shot task performance, and dataset generation.

Research Intern

Feb 2022 – Jul 2022

MLLP – VRain

Valencia, Spain

- Developed various state of the art transformer-based models for multilingual online neural machine translation.

AI Engineer

Aug 2021 – Nov 2021

Daidalonic Student Group

Valencia, Spain

- Built models for electromyographic signal classification for the development of a transradial prosthesis.

Software Engineering Intern

Apr 2021 – Oct 2021

StadlerRail

Valencia, Spain

- Created automatic speech recognition and parsing systems for automatic train control via voice commands.
- Designed and developed a full-stack application for subjective evaluation.

EDUCATION

MSc in Artificial Intelligence, Pattern Recognition and Digital Imaging

Sep 2022 – Jul 2023

Universitat Politècnica de València

Valencia, Spain

- GPA: 9.5/10 with honors in 1 course. Thesis: 10/10.

BSc in Computer Engineering

Sep 2018 – Jul 2022

Universitat Politècnica de València

Valencia, Spain

- GPA: 9.3/10 with honors in 16 courses. Thesis: 10/10 cum laude.

ACHIEVEMENTS AND AWARDS

First place at SemEval 2024: Task 8 on Machine Generated Text Detection

2024

Awarded by the 18th International Workshop on Semantic Evaluation (SemEval 2024)

Mexico City, Mexico

Kaggle Competitions, Notebooks, Datasets and Discussions Expert

2023

Best BSc Thesis in the Computer Engineering Field - Data Science Branch

2023

Awarded by the Cátedra Stadler

Valencia, Spain

First place at the lablab.ai Reinforcement Learning Hackathon

2023

Sapiens Award to the best BSc Thesis

2022

Awarded by the Official College of Computer Engineering of the Valencian Community (COIICV)

Valencia, Spain

ACIA Award to the best BSc Thesis in the field of Artificial Intelligence

2022

Awarded by the Catalan Association for Artificial Intelligence (ACIA)

Barcelona, Spain

Third prize winner in the 3rd edition of the ACM RSA CryptoHack

2019

Best First Year Student of the UPV's BSc in Computer Engineering

2019

SKILLS

Artificial Intelligence: Large Language Models, Generative AI, Deep Learning, Machine Learning, Natural Language Processing, Computer Vision

Languages: Python, Bash, SQL, C/C++, Java

Tools: PyTorch, Scikit-Learn, HuggingFace, TensorFlow, Keras, Docker, VLLM, TensorRT, ONNX, FastAPI, Streamlit

Software Development: Git, Azure DevOps, Visual Studio, GitHub

LANGUAGES

Mother Tongue: Spanish, Armenian
Proficiency - C2 Level: English
Basic Competence - A1 Level: German

PUBLICATIONS

1. **Sarvazyan, A. M.**, González, J. Á. & Franco-Salvador, M. TextMachina: Seamless Generation of Machine-Generated Text Datasets. *Procedia Computer Science* **246**. 28th International Conference on Knowledge Based and Intelligent information and Engineering Systems (KES 2024), 566–575. ISSN: 1877-0509. <https://www.sciencedirect.com/science/article/pii/S1877050924025031> (2024).
2. **Sarvazyan, A. M.**, González, J. Á. & Franco-salvador, M. *Genaios at SemEval-2024 Task 8: Detecting Machine-Generated Text by Mixing Language Model Probabilistic Features* in *Proceedings of the 18th International Workshop on Semantic Evaluation (SemEval-2024)* (Association for Computational Linguistics, Mexico City, Mexico, June 2024), 101–107. <https://aclanthology.org/2024.semeval-1.17/>.
3. **Sarvazyan, A. M.**, González, J. Á., Rangel, F., Rosso, P. & Franco-Salvador, M. Overview of IberAuTexTification at IberLEF 2024: Detection and Attribution of Machine-Generated Text on Languages of the Iberian Peninsula. *Procesamiento del Lenguaje Natural* **73**, 421–434. ISSN: 1989-7553. <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/6628> (2024).
4. **Sarvazyan, A. M.**, González, J. Á., Franco-Salvador, M., Rangel, F., Chulvi, B. & Rosso, P. Overview of AuTexTification at IberLEF 2023: Detection and Attribution of Machine-Generated Text in Multiple Domains. *Procesamiento del Lenguaje Natural* **71**, 275–288. ISSN: 1989-7553. <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/6559> (2023).
5. **Sarvazyan, A. M.**, González, J. Á., Rosso, P. & Franco-Salvador, M. *Supervised Machine-Generated Text Detectors: Family and Scale Matters* in *Experimental IR Meets Multilinguality, Multimodality, and Interaction* (Springer Nature Switzerland, Cham, 2023), 121–132. ISBN: 978-3-031-42448-9.
6. Tamayo, R. L. & **Sarvazyan, A. M.** *Reshape or Update? Metric Learning and Fine-tuning for Low-Resource Influencer Profiling* in *CLEF (Working Notes)* (2023), 2679–2690.

TALKS

1. González, J. Á. & **Sarvazyan, A. M.** *Symanto Research's Advances in Machine Generated Detection* in *Master's Degree in Artificial Intelligence, Pattern Recognition and Digital Imaging* (Universitat Politècnica de València, 2024).
2. **Sarvazyan, A. M.** *AuTexTification 2023 and more: Detection and Attribution of Machine-Generated Text in Multiple Domains* in *hOUPSh: Humain ou pas humain* (The 18th Conference on Information Research, Applications (CORIA), and the 30nd Conference on Natural Language Processing (TALN), 2023).

SERVICE

Program Committee Member
Reviewer
Program Committee Member
Reviewer
Reviewer
Reviewer

IberLEF 2025 at SEPLN
CLEF 2025
DAIGC 2025 at COLING 2025
LREC-COLING 2024
Information Processing and Management
EMNLP 2023