

KENZA AMARA

PERSONAL INFORMATION

Born in France, 21 February 1997

webpage k-amara.github.io
email kenza.amara@ai.ethz.ch
phone (M) +41 77 265 54 38

INTEREST

Explainability for graph neural networks, deep learning on graphs, explainable Artificial Intelligence (xAI), and their applications for health, climate models, banks, and the e-commerce.

EDUCATION

<i>PhD in Computer Science</i>	<i>2021-2024</i> ETH AI Center, ETH University, Zurich Description: This thesis explores the research field of explainable AI and its specific use for graph neural networks. Current work consists in re-defining explainability and building a comparative framework of the existing methods. Advisors: Prof. Ce ZHANG & Assoc. Prof. Ulrik BRANDES, Prof. Sebastian SCHEMM
<i>Master in Environmental Sciences</i>	<i>2019-2021</i> ETH University, Zurich <i>Specialization: Environmental Systems and Policy</i> Description: This degree brought an in-depth understanding of environmental systems science with the knowledge and skills required to identify solution strategies that are environmentally, technically, and socially robust.
<i>Master in Computer Science</i>	<i>2016-2019</i> Ecole Polytechnique, Paris Description: Core areas in Computer Science and Applied Mathematics. Electives in Physics and Economy.
<i>Scientific Preparatory Program</i>	<i>2014-2016</i> Lycee Henri IV, Paris Description: Intensive undergraduate-level curriculum in advanced mathematics and physics, leading to competitive entrance examinations to the French Grandes Ecoles

RESEARCH EXPERIENCE

<i>Microsoft Research (MSR) Center</i>	<i>July-Oct 2022</i> Research Internship, MICROSOFT, Cambridge Improved feature attribution methods in drug discovery using contrastive learning and regularization techniques Reference: Jose SALVADOR JIMENEZ LUNA, Raquel RODRIGUEZ PEREZ
<i>Facebook AI Research Center</i>	<i>May-Sept 2021</i> Research Internship, FAIR, Paris Worked on data compression retrieval to increase the decoding accuracy for efficient vector indexing Reference: Hervé JÉGOU, Matthijs DOUZE

ETH University	June-Jan 2020/21	Master Thesis, CROWTHER LAB & DS3LAB, Zurich
		Developed OneForest, a tool that uses graph attention networks and optimal transport to merge drone imagery and citizen science, and build informative maps of our forests worldwide. Reference: Tom CROWTHER
Columbia University	Mar-Aug 2019	Research Internship, EARTH INSTITUTE, NYC
		Predicted the probability distribution of wildfires in Western United States with Mixture Density Networks. Reference: Pierre GENTINE
DAIKIN Japan	June-Sept 2018	Engineer Internship, DAIKIN, Osaka
		Predicted power consumption of air-conditioning systems with an optimized machine learning model regularized by thermodynamics laws. Reference: Haruo KOCHI
Ecole Polytechnique	Sept-Mar 2017/18	Scientific Project, LIX, Paris
		Developed a machine learning program embedded in a drone for endemic plants detection. Reference: Sylvie PUTOT

PUBLICATIONS

Neurips	June 2022	GraphFramEx: Towards Systematic Evaluation of Explainability Methods for Graph Neural Networks
		We propose the first systematic evaluation framework for GNN explainability: it uses a unique metric that combines the fidelity measures and classify explanations based on their quality of being sufficient or necessary. This work shows that structure-based methods give the best explanations on the widely used synthetic benchmarks because of their simple structure and node dependency. This paper and the webpage about GraphFramEx also aim to facilitate the assessment of future explainability methods that can be easily integrated to the framework. Authors: Kenza AMARA, Rex YING, Zitao ZHANG, Zhihao HAN, Yinan SHAN, Ulrik BRANDES, Sebastian SCHEMM, Ce ZHANG https://arxiv.org/abs/2206.09677 - preprint
International Conference on Multimedia Retrieval (ICMR)	March 2022	Nearest neighbor search with compact codes: A decoder perspective
		We re-interpret popular methods such as binary hashing or product quantizers as auto-encoders, and point out that they implicitly make suboptimal assumptions on the form of the decoder. We design backward-compatible decoders that improve the reconstruction of the vectors from the same codes, which translates to a better performance in nearest neighbor search. Our method significantly improves over binary hashing methods or product quantization on popular benchmarks. Authors: Kenza AMARA, Matthijs DOUZE, Alexandre SABLAYROLLES, Hervé JÉGOU https://dl.acm.org/doi/pdf/10.1145/3512527.3531408 - accepted to ICMR 2022
Association for the Advancement of Artificial Intelligence (AAAI)	January 2022	ReforesTree: A Dataset for Estimating Tropical Forest Carbon Stock with Deep Learning and Aerial Imagery
		We present ReforesTree, a benchmark dataset of forest carbon stock in six agro-forestry carbon offsetting sites in Ecuador. Furthermore, we show that a deep learning based end-to-end model using individual tree detection from low cost RGB-only drone imagery is accurately estimating forest carbon stock within official carbon offsetting certification standards.

Authors: Gyri REIERSEN, David DAO, Bjorn LÜTJENS, Konstantin KLEMMER,
 Kenza AMARA, Attila STEINEGGER, Ce ZHANG, Xiaoxiang ZHU
<https://arxiv.org/pdf/2201.11192.pdf>

August 2020 OneForest: Towards a Global Species Dataset by
 Fusing Remote Sensing and Citizen Science Data with Graph
 Neural Networks

*Knowledge
 Discovery and
 Data Mining
 (KDD) - Fragile
 Earth Workshop*

We propose ONEFOREST, a method to create a tree species dataset by fusing unlabeled data from remote sensing, i.e. drone images of forests, with labeled citizen science data, i.e. ground measurements. In this paper, we compare methods that map drone imagery and ground observations based on nearest neighbors, registration, and optimal transport techniques. Our proposed matching algorithm classifies species with 81.5% accuracy on a dataset of 5061 trees in Ecuador by leveraging optimal transport and an initial CNN-based matching step. We additionally test our methods on a variety of datasets composed of different species and locations from South America (Ecuador) and North America (NEON).

Authors: Kenza AMARA, David DAO, Charlotte BUNNE, Bjorn LÜTJENS, Dava NEWMAN, Ce ZHANG and Tom CROWTHER

SKILLS

<i>Code</i>	PYTHON, R, C/C++, JAVA
<i>Communication</i>	L ^A T _E X, HTML/CSS/PHP, OpenOffice, Linux, Microsoft Windows
<i>Languages</i>	Fluent · ENGLISH, FRENCH, GERMAN Intermediate · SPANISH, JAPANESE
<i>Presentation</i>	2022 · Panel discussion at the workshop on Practical ML for Developing Countries, PML₄DC ICLR 2022 2022 · Lightning talk on Advances of ML Approaches for Financial Decision Making & Time Series Analysis, AMLD EPFL 2022
<i>Leadership</i>	Oct-May 2016/17 · Zhuhai, China Teaching Assistant at the SINO-FRENCH INSTITUTE OF NUCLEAR ENGINEERING AND TECHNOLOGY.
<i>Interests</i>	Swimming · Cycling · Hiking · Skying

July 22, 2022