1. Short CV of the advisor

Mauro Castelli

Mauro Castelli has a Ph.D. in Computer Science obtained at the Università di Milano Bicocca (Italy). He is currently an Associate Professor at the Universidade Nova de Lisboa, and a member of the Scientific Council of NOVA Information Management School (NOVA IMS). He is also a researcher at the Information Management Research Center of this university. He participated as a principal investigator, co-principal investigator, or work package leader in different research projects at national and international levels. Finally, in 2022 and 2023, he has been ranked in the top 2% world scientists by a study of the Stanford University.

It has international collaborations with researchers in more than twenty different countries and with universities recognized as leaders in the area of artificial intelligence. He is a member of ACM (Association for Computing Machinery) and SIGEVO (ACM Special Interest Group on Genetic and Evolutionary Computation). He is the author of more than 180 scientific publications and has presented about fifty seminars, conferences, and communications. He was awarded in 2013 and 2014, in the framework of the main European conference of Artificial Intelligence and Evolutionary Computing (Evo*), for the quality and contribution of his scientific research. He participated in research projects concerning the implementation of machine learning solutions with companies and organizations, such as Critical Software (Portugal), the Champalimaud Foundation (Portugal), the Directorate General of Education and Science Statistics (Portugal), and the Portuguese Ministry of Tourism.

At NOVA IMS, he is/was responsible for the following master-level courses: Deep Learning, Machine Learning, Data Mining, Decision Support Systems, and Big Data. He is also responsible for the following undergraduate courses: Algorithms Design, and Object-Oriented Programming. Finally, he is responsible for the PhD course on Genetic Programming. He was the supervisor of more than 130 master theses and eight Ph.D. theses in the field of Machine Learning.

2. Description of the research group

MagIC

The Information Management Center (MagIC) is the research center of Nova Information Management School (NOVA IMS), which is one of the worldwide leading schools in the area of Information Management. NOVA IMS is also ranked in the TOP 5 of the Best Masters and Postgraduate Programs in the world by the Eduniversal and the first European Institution to obtain the ABET accreditation.

With over 50 integrated researchers and more than 1200 published articles since 2011 (800 articles WOS or Scopus indexed), the primary objective of the MagIC working group is to contribute to the advancement of the fields of Information Management and Data Science. The research center is organised around 4 research streams: Data Science, Information Systems, Geoinformatics, and Data-Driven Marketing.

The working group of Data Science has two main branches. One is dedicated to the study of data generation algorithms for imbalanced learning. The other branch, in which the student will be working, is focused on Evolutionary Computation (EC). This includes the theoretical and practical study of EC, the development of EC models and innovative applications of these models.

In 2020, NOVA IMS and MagIC have launched a Data Analytics Lab, with the participation and co-supervision of Professor Mauro Castelli. This lab is a location where students working in Artificial Intelligence and Data Science can share a very friendly environment, with comfortable space and facilities, ideal for social networking and collaborations. The majority of the students and young researchers of the Data Analytics Lab work in Evolutionary Computation.

3. Description of the work to be carried out by the student

Model merging involves consolidating multiple Large Language Models (LLMs) into one unified model. Essentially, the main aim is the integration of two or more pretrained models, whether trained on similar tasks or employing distinct architectures, into a single model that preserves the strengths and capabilities of the originals. This approach, which is still in its early stages, offers a cost-effective alternative for model creation (for instance, by eliminating the need for GPU usage). Notably, the process of model merging demonstrates remarkable effectiveness, yielding numerous cutting-edge models [1,2].

The fundamental reasoning behind this strategy is that diverse LLMs, each producing excellent performance in particular tasks, can be combined to achieve better performance by harnessing the collective expertise of the models.
Within the domain of LLMs, different approaches for model merging have been proposed. The linear approach utilizes a weighted average to merge two (or more) models, allowing users to adjust the contribution of each model's characteristics to the resultant merged model [3]. The Spherical Linear Interpolation (SLERP) technique harmonizes the functionalities of LLMs by interpolating between their parameters within a high-dimensional space. By looking at the models' parameters as points on a hypersphere, SLERP determines the shortest path connecting them, thus ensuring a blend that more accurately embodies the characteristics of the parent models (see https://github.com/Digitous/LLM-SLERP-Merge). Other popular techniques ate the Task Arithmetic method [5] and TIES-merging, probably the most popular technique nowadays. All these methods do not require training data for merging operations, but a relevant requirement is that the models to be merged must have identical architectures.
Recently, evolutionary techniques started to appear with LLMs. Brahmachary et al. [6] proposed an evolutionary approach (LEO - Language-model-based Evolutionary Optimizer) in which an LLM is used to generate new candidate solutions. Concerning model merging, to the best of our knowledge, the only existing attempts are the works of Akiba et al [7] and the work of Jiang [8]. Akita and co-authors proposed an evolutionary-based method to automate the creation of foundation models. In particular, moving from existing model merging methods, they proposed an algorithm that works in both parameter space and data flow space, thus allowing for optimization beyond just the weights of the individual models.
Jiang introduced EvoMerge, an evolutionary-based approach to LLMs training and merging.
These contributions showed the possibility of relying on evolutionary approaches for model merging. Still, they represent a first proposal for the considered task: thus, different alternatives must be explored.

This research proposes the application of evolutionary algorithms, to automatically generate and optimize fused model architectures. The approach should evolve novel model architectures that integrate the capabilities of multiple pre-trained LLMs, without the need for extensive retraining.
The focus is on the use of genetic programming, but other evolutionary approaches can be considered.
In particular, the main research contributions are:
(1) The development of a genetic programming framework tailored to the model merging task, capable of automatically generating fused model architectures.

(2) Investigate the feasibility of evolving model architectures that effectively merge the strengths of multiple pre-trained LLMs, while preserving their individual characteristics.
(3) Explore evolutionary approaches to optimize the performance and efficiency of the fused models, without resorting to retraining or gradient-based optimization.
(4) Assess the generalization capabilities of the evolved model architectures across various tasks and datasets.

The research is expected to provide insights into the effectiveness of evolutionary algorithms for the model fusion task.

References

[1] Goddard, C., Siriwardhana, S., Ehghaghi, M., Meyers, L., Karpukhin, V., Benedict, B., ... & Solawetz, J. (2024). Arcee's MergeKit: A Toolkit for Merging Large Language Models. arXiv preprint arXiv:2403.13257.

[2] Labrak, Y., Bazoge, A., Morin, E., Gourraud, P. A., Rouvier, M., & Dufour, R. (2024). BioMistral: A Collection of Open-Source Pretrained Large Language Models for Medical Domains. arXiv preprint arXiv:2402.10373.

[3] Wortsman, M., Ilharco, G., Gadre, S. Y., Roelofs, R., Gontijo-Lopes, R., Morcos, A. S., et al. (2022, June). Model soups: averaging weights of multiple fine-tuned models improves accuracy without increasing inference time. In International conference on machine learning (pp. 23965-23998). PMLR.

[4] Yadav, P., Tam, D., Choshen, L., Raffel, C. A., & Bansal, M. (2024). Ties-merging: Resolving interference when merging models. Advances in Neural Information Processing Systems, 36.

[5] Ilharco, G., Ribeiro, M. T., Wortsman, M., Gururangan, S., Schmidt, L., Hajishirzi, H., & Farhadi, A. (2022). Editing models with task arithmetic. arXiv preprint arXiv:2212.04089.

[6] Brahmachary, S., Joshi, S. M., Panda, A., Koneripalli, K., Sagotra, A. K., Patel, H., ... & Kalyanaraman, K. (2024). Large Language Model-Based Evolutionary Optimizer: Reasoning with elitism. arXiv preprint arXiv:2403.02054.

[7] Akiba, T., Shing, M., Tang, Y., Sun, Q., & Ha, D. (2024). Evolutionary optimization of model merging recipes. arXiv preprint arXiv:2403.13187.

[8] Jiang, Y. (2024). EvoMerge: Neuroevolution for Large Language Models. arXiv preprint arXiv:2402.00070.

4. Other relevant information

Lisbon is the capital and the largest city of Portugal. It is recognised as an alpha-level global city because of its importance in finance, commerce, media, entertainment, arts, international trade, education and tourism. It is a very multifaceted and multi skilled city. Morevoer, it is a safe town, which foster the dissemination of culture all around - a perfect environment for students. The climate of Lisbon is great and the city was considered the 23rd World's Best City to Live by Monocle Magazine (July / August 2011).

NOVA IMS is located in an excellent neighborhood of Lisbon, with public transportation and all facilities nearby. The School was founded in 1989 and currently it provides education at the highest level to more than 3.000 students from more than 80 countries. They are undergraduates, post-graduates, masters and PhD degree students. The vocation of NOVA IMS is to contribute in a pioneering and leading way to the conversion of data into value through Data Science, by an integrated and properly articulated commitment to teaching, research, and third mission initiatives. NOVA IMS é driven by what is genuinely central to the best higher education institutions. In the journey of affirmation for the quality of NOVA IMS is formed by the attraction of great students, teachers, and researchers, accompanied by support services provided by great professionals; a human community that through an agile, uncomplicated, vibrant, motivating environment, based on meritocracy and a strong internal cohesion, is encouraged to dream and fulfill dreams in a context of great intellectual and creative freedom.