Smarter, Not Just Harder Machine Learning: New Pathways for Artificial Intelligence

Contemporary Artificial Intelligence (AI) can perform amazing tasks, from speech recognition, through text and image analysis, to assisting in diagnosing diseases. However, the process of "learning" for machines is often inefficient and does not resemble the way humans learn. Machines receive training examples in random order and must build a general understanding of the problem based on that. This project aims to change this approach by developing innovative methods of so-called Curriculum Learning (CL). We want computers to learn similarly to children, starting with simpler tasks and gradually moving to more difficult ones. The main goal of the project is to create reliable, adaptive, and universal machine learning strategies that will significantly improve the effectiveness and performance of AI models.

Small steps towards smarter AI: Currently, computers often learn in a chaotic way, trying to understand all the data at once, which can lead to errors or "getting stuck" in a dead end. Our project proposes an integrated approach, combining three innovative research directions to make the learning process more natural and effective:

1. **Dynamic learning goals:** Imagine that we are teaching a machine to recognize images. Instead of immediately showing complicated scenes, we first teach it to distinguish general shapes (e.g., whether it is an animal or a building), and only then move on to identifying specific details (e.g., dog breed, architectural style). As part of the project, we will develop methods in which the computer focuses on different aspects of the problem at successive stages of learning, gradually building a more complete understanding. Preliminary results, for example in the task of speech recognition, where the model first learned general categories of sounds (vowels, consonants) and then individual phonemes, are very promising, showing faster convergence and better performance on new data.
2. **Contextual and adaptive learning:** We want the machine, when learning new information, to take into account its "neighborhood," that is, other similar data. In addition, the way "errors" made by the machine during learning are assessed will be dynamically adjusted. At the beginning of learning, the system can be treated more leniently, especially with difficult examples, and as progress is made, the learning "bar" will be raised.
3. **Intelligent "brains" (Kolmogorov-Arnold Networks):** In the project, we will use a new type of neural network, the so-called Kolmogorov-Arnold networks. Their unique feature is "learning" activation functions, i.e., fundamental elements processing information in a neural network, which can dynamically adjust their complexity. The planned research will include developing methods for the "development" of these functions so that at the beginning the AI model has a simpler structure for learning easier tasks, and its capabilities grow as the difficulty of the problem increases.

Why are we undertaking this research topic? Traditional methods of Curriculum Learning, despite their theoretical advantages, often fail in practice because they are very sensitive to settings and the type of problem. In an era of increasingly large and complex AI models (such as the models powering chatbots), we need smarter and more efficient training strategies. Current approaches often require enormous computing resources and energy, which is not only costly but also not environmentally friendly. Developing more effective CL methods will allow for faster implementation of innovations, reduce costs and energy consumption, and make advanced AI more accessible to smaller research centers and companies that do not have unlimited computing resources.

What significant results do we expect? The implementation of the project will bring a number of benefits to science and society. We expect that the developed methods will allow for training more effective AI models that will achieve better results, make fewer mistakes, and be better at handling new data. The new Curriculum Learning strategies have the potential to shorten the time needed to train advanced models, which will translate into savings in computing resources and energy. Models trained using these methods should be more resistant to poor quality data, information noise, or attempts to deliberately mislead AI models. Additionally, the project will provide scientists and engineers with new, more reliable tools for designing machine learning processes, which can accelerate progress in many fields of science and technology. These studies will also contribute to a deeper understanding of the dynamics of learning in complex AI models.