

DOI: <https://doi.org/10.36910/6775-2524-0560-2023-53-06>

UDC 004.8

Bortnyk Kateryna, PhD, Associate Professor

<https://orcid.org/0000-0001-5282-099X>

Yaroshchuk Bohdan, Master Student

Bahniuk Nataliia, PhD, Associate Professor

<https://orcid.org/0000-0002-7120-5455>

Pekh Petro, PhD, Associate Professor

<https://orcid.org/0000-0002-6327-3319>

Lutsk National Technical University, c. Lutsk, Ukraine

OVERCOMING CHALLENGES IN ARTIFICIAL INTELLIGENCE TRAINING: DATA LIMITATIONS, COMPUTATIONAL COSTS AND MODEL ROBUSTNESS

Bortnyk K., Yaroshchuk B., Bahniuk N., Pekh P. **Overcoming challenges in artificial intelligence training: data limitations, computational costs and model robustness.** This paper explores challenges in AI training, focusing on data limitations, computational costs, and the need for robust models. It discusses innovative solutions like synthetic data generation, efficient neural architectures, and robustness techniques, highlighting the importance of AI model interpretability.

Keywords: artificial intelligence, AI training, computational costs, environmental impact, model robustness, interpretability, energy efficiency, AI ethics, sustainable AI.

Бортник К.Я., Ярошчук Б.Р. **Подолання викликів у тренуванні штучного інтелекту: обмеження даних, обчислювальні витрати та надійність моделей.** Стаття аналізує виклики, пов'язані з тренуванням систем штучного інтелекту, зокрема проблеми з обмеженістю та якістю даних, високі обчислювальні витрати та необхідність забезпечення надійності моделей. Обговорюються інноваційні рішення, такі як синтез даних, ефективні архітектури нейронних мереж і методи забезпечення робастності, а також важливість інтерпретованості ШІ-моделей.

Ключові слова: штучний інтелект, тренування ШІ, обчислювальні витрати, екологічний вплив, надійність моделей, інтерпретованість, енергоефективність, етика ШІ, сталий ШІ.

The problem. The advent of artificial intelligence has spurred a technological revolution that has permeated nearly every sector of society. From predictive analytics in healthcare to autonomous vehicles, AI systems have shown immense potential to advance and innovate. However, as the appetite for more sophisticated AI grows, so too do the challenges associated with training these complex models. One of the most significant hurdles is the limited availability and quality of training data. Additionally, the computational resources required to process this data and train AI models are substantial, raising concerns over costs and energy consumption. Furthermore, ensuring the robustness of AI models to function reliably in diverse and often unpredictable real-world scenarios remains a paramount concern. This paper aims to dissect these critical challenges, explore the implications of current practices, and highlight innovative solutions that could redefine the future landscape of AI training. By understanding and overcoming these barriers, we can pave the way for more efficient, equitable, and sustainable AI development.

Artificial Intelligence training is confronted with critical challenges such as data scarcity, computational expenses, robustness, and complexity of models. This abstract discusses the implications of data scarcity and quality on AI training, the robustness of AI models, financial and environmental cost. It further examines the trade-offs between model complexity and performance, addressing the need for efficient algorithms that can operate with limited data while still ensuring model accuracy. Environmental concerns are also highlighted, as AI's carbon footprint becomes a growing issue, urging the development of more energy-efficient computing techniques. The financial implications are vast, encompassing not only the direct costs of hardware and electricity but also the investment in research and development of innovative solutions to these challenges. Overall, this discussion aims to shed light on the multifaceted nature of AI development and the necessity for a balanced approach to its advancement.

The main part. A primary obstacle in AI training is the scarcity of high-quality, annotated datasets[1]. These datasets are the bedrock upon which machine learning models learn and develop an understanding of the world. When data is scarce, unrepresentative, or biased, it can lead to models that perform poorly or unfairly when deployed. Diverse and extensive datasets are particularly critical for deep learning models, which require vast amounts of data to achieve high levels of accuracy. Moreover, the creation and maintenance of such datasets are often resource-intensive and can be constrained by privacy concerns and data protection regulations. To mitigate these issues, researchers have turned to synthetic data

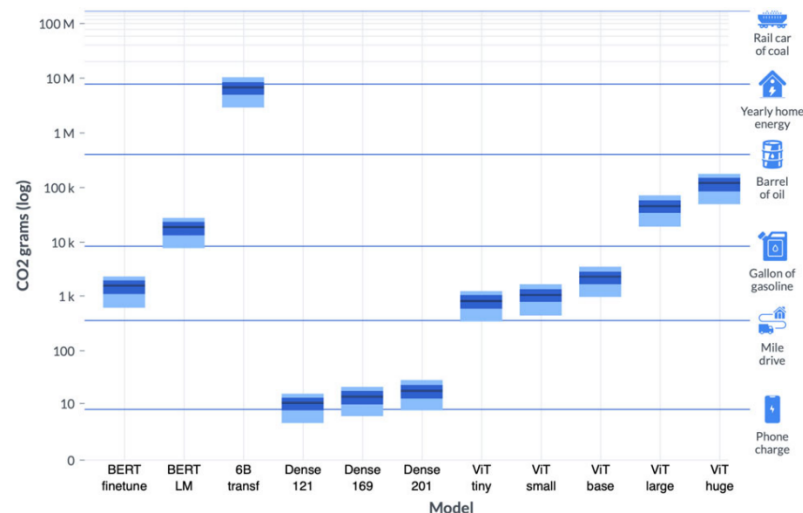
generation and data augmentation techniques. Synthetic data, created through simulations or generative models, can expand training datasets and introduce a wider range of scenarios that may not be present in the original data. Data augmentation[3], which involves altering existing data to create new variations, enhances the diversity within a dataset without the need to collect new data.

These methods not only augment the quantity of data available but also improve the robustness and generalizability of AI models. Further, the adoption of federated learning approaches allows for the utilization of decentralized data, enabling the model to learn from a wide array of sources while maintaining data privacy. Transfer learning is another strategy that leverages pre-trained models on large datasets to perform new tasks with limited data. This not only conserves resources but also accelerates the development cycle of AI models. The challenges of data scarcity are not solely technical but also ethical and legal. Responsible AI development necessitates adherence to ethical guidelines and legal frameworks, particularly regarding data privacy and the potential biases in AI systems. Continuous efforts in developing fair, accountable, and transparent AI systems are crucial to address these concerns. As the field advances, interdisciplinary collaboration becomes key, with insights from social sciences, law, and ethics guiding the responsible creation and use of AI technologies. Each of the mentioned approaches has its pros and cons (table 1).

Table 1. Comparison of dataset enhancement techniques

Technique	Pros	Cons	Best Use Cases
Synthetic Data Generation	Generates large datasets, Useful when real data is limited or unavailable	May not accurately represent real-world scenarios	Scenarios with limited or sensitive real data
Data Augmentation	Enhances dataset variety, Improves model generalization	Can introduce noise, Limited by original data diversity	Image and speech recognition tasks
Transfer Learning	Leverages pre-trained models, Reduces training time and resources	May carry biases from source data, Limited by relevance of pre-trained models	Tasks with similar data to pre-trained models
Semi-Supervised Learning	Utilizes both labeled and unlabeled data, Reduces labeling costs	Dependent on quality of labeled data, Complex implementation	Large datasets with limited labeled data

Training state-of-the-art AI models typically demands substantial computational power, which can be cost-prohibitive and environmentally unsustainable. The carbon footprint of training a single deep learning model can be equivalent to the lifetime emissions of several cars (picture 1). The financial cost of these computations is also significant, often running into thousands or even millions of dollars for cutting-edge models, limiting the ability to train such models to well-funded organizations. This creates a barrier to entry for smaller entities and can stifle innovation. Efforts to reduce computational costs include the development of more efficient neural network architectures that require less processing power without compromising performance. For example, techniques such as pruning[4], quantization, and knowledge distillation can reduce the size of a neural network and its computational needs. Moreover, there is a push towards utilizing cloud-based platforms and specialized hardware such as tensor processing units (TPUs) and field-programmable gate arrays (FPGAs), which can offer better performance for AI-specific computations. Companies are also exploring the use of renewable energy sources to power data centers, seeking to mitigate the environmental impact of large-scale computation. Despite these advances, the quest for more computationally efficient AI training methods continues, as the demand for AI capabilities grows in tandem with global computational infrastructure.



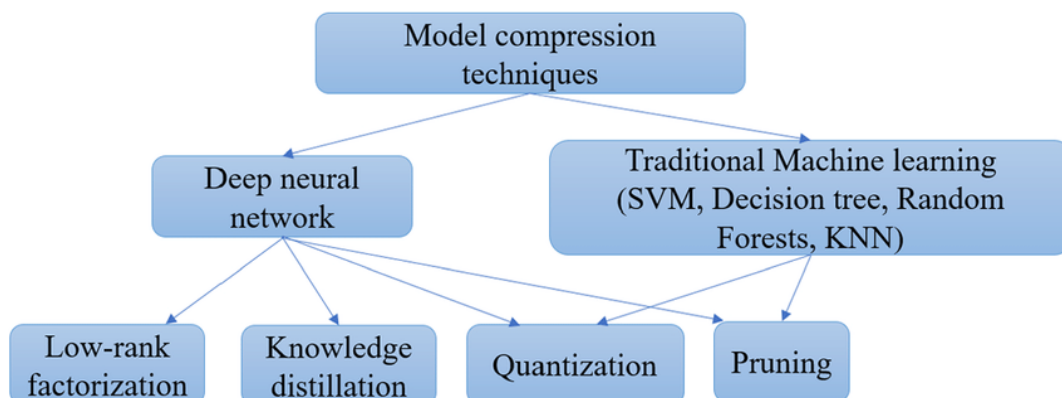
Pic. 1. Carbon footprint of AI models

Achieving robustness in AI models means ensuring they make reliable and accurate predictions, even when faced with new and unforeseen data in real-world applications. To combat this, techniques such as regularization and cross-validation[2] are employed to improve generalization. Regularization methods, like dropout or L1 and L2 regularization, discourage complex models that can overfit by adding a penalty to the loss function during training. Moreover, recent advances in machine learning focus on transfer learning and few-shot learning, where a model trained on one task is adapted to perform different but related tasks with limited additional data. This approach leverages pre-existing knowledge and can greatly reduce the amount of data required for training robust models. Another promising area is adversarial training, where models are deliberately exposed to challenging conditions during training, such as inputs designed to confuse the model. This can help AI systems become more resilient to errors and adversarial attacks when deployed in the real world. Ensuring robustness in AI models is not only critical for maintaining performance standards but also for building trust with users. When people understand that an AI system can handle a wide array of scenarios reliably, they are more likely to adopt it in their daily lives and work.

The complexity of artificial intelligence models, especially those based on deep learning, often results in a lack of interpretability, which is a significant issue for applications where understanding the decision-making process is crucial. The so-called "black box" nature of these models can hinder their adoption in fields that require transparent decision-making, such as healthcare and justice. Interpretable AI seeks to address this by creating models whose workings can be understood by humans. This involves developing methods that can explain, in human terms, how a model arrived at a particular decision or prediction. Techniques like feature importance scores, which highlight which inputs most influenced the output, and model-agnostic methods, which can be applied to any machine learning model to increase transparency, are key tools in this endeavor. There's also a push for "explainable AI" (XAI), where the goal is not only to make AI decisions more transparent but also to provide explanations that are meaningful to the end-users. This could involve visual explanations, textual rationales, or even interactive interfaces that allow users to query the model and understand the rationale behind specific decisions. Ensuring the interpretability of AI models is essential not just for trust and adoption, but also for regulatory compliance. As laws and regulations begin to demand greater accountability in automated decision-making, the ability to interpret and explain AI models becomes increasingly important.

To tackle the challenges discussed above, there are few innovations. Meta-learning, or "learning to learn," which involves training a model on a variety of learning tasks, such that it can solve new learning tasks using only a small number of training samples. Meta-learning models are particularly promising for their ability to adapt quickly to new and varied datasets. Another one is development of transfer learning frameworks, where a model developed for a particular task is reused as the starting point for a model on a second task. This approach has significantly reduced the amount of data and computation needed to develop effective models. Neural architecture search (NAS) is also an area of active research, automating the design of neural networks to discover the best-performing models for a given dataset, which can result in more efficient and less computationally expensive architectures. Another approach gaining traction is meta-

learning, or "learning to learn," which involves training a model on a variety of learning tasks, such that it can solve new learning tasks using only a small number of training samples. Meta-learning models are particularly promising for their ability to adapt quickly to new and varied datasets. Efforts in model compression techniques (picture 2), like network pruning, where redundant or non-informative connections are removed without sacrificing accuracy, help in deploying models on edge devices with limited computational capacity. Quantization reduces the precision of the model's parameters, thereby decreasing the size of the model and speeding up inference while maintaining performance. Lastly, reinforcement learning, especially in complex environments, has shown potential for training models that make decisions in dynamic settings, pushing the boundaries of what AI can achieve in real-world applications. These innovations are not only enhancing the capabilities of AI systems but are also crucial for ensuring that the benefits of AI can be realized broadly, democratically, and ethically.



Pic. 2. Model compression techniques

As the capabilities of artificial intelligence (AI) continue to expand, so does the ethical landscape in which these technologies are developed and deployed. AI training processes not only need to be technically sound but also ethically responsible. This section delves into the key ethical considerations that must be integrated into the lifecycle of AI training. One of the most pressing ethical concerns in AI training is the potential for embedded biases within algorithms. Biases can originate from various sources, including prejudiced data sets, biased human annotators, or even the design of the AI system itself. To counteract this, it is imperative to implement diversity and inclusion protocols in the data collection and model training phases. Ensuring that datasets are representative of all demographics and that AI systems are tested for fairness across different groups is crucial. AI systems often operate as "black boxes," with decision-making processes that are opaque to users and stakeholders. This lack of transparency can lead to trust issues, especially in critical domains like healthcare or criminal justice. It is essential to develop AI with explainable algorithms, where the decision-making process is clear, understandable, and able to be interrogated. Explainable AI enables accountability and provides insights into how and why certain AI decisions are made. Training AI models frequently involves large volumes of personal data. Protecting the privacy of individuals whose data is used in AI training is a fundamental ethical obligation. Techniques like differential privacy, federated learning, and homomorphic encryption can enable AI training while safeguarding user data. Moreover, adherence to privacy regulations such as the General Data Protection Regulation (GDPR) is not just a legal necessity but an ethical commitment to individual rights. Assigning responsibility for AI decisions is a complex ethical challenge. In cases where an AI system causes harm, it is necessary to have clear lines of accountability. This includes the developers, the data scientists, the companies deploying the AI, and even the regulatory bodies. As AI systems become more autonomous, the mechanisms for accountability must evolve to ensure that moral and legal responsibility is clearly defined and enforced. AI technologies can have profound effects on society. They can influence job markets, personal freedoms, and social interactions. Ethical AI training must consider the long-term societal impacts, striving to minimize potential harm and maximize benefits. This involves engaging with various stakeholders, including policymakers, to ensure that the deployment of AI aligns with societal values and contributes positively to the community. AI should be designed with the end-user in mind, ensuring that human dignity is respected and that AI acts to augment rather than replace human capabilities. This involves understanding the context in which AI operates and designing systems that enhance human decision-

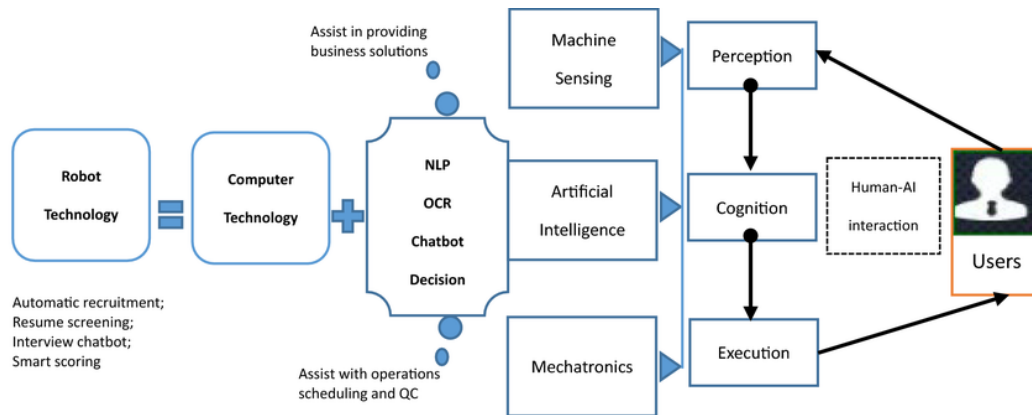
making, provide valuable assistance, and remain under human oversight. Addressing the ethical dimensions of AI training requires an interdisciplinary approach. Collaboration among technologists, ethicists, legal experts, sociologists, and other stakeholders is essential. This collective effort can help in formulating guidelines, frameworks, and best practices that align AI advancements with ethical standards. Ethical considerations in AI training are not a one-time effort but require ongoing attention. As AI systems are deployed and begin to interact with the real world, continuous monitoring is essential to ensure that they operate as intended and that any emergent ethical issues are promptly addressed.

Human-AI interaction (picture 3) and the ensuing feedback loops are pivotal in the training of artificial intelligence (AI) systems. This dynamic shapes the AI's learning process, ensuring that the systems we develop are not only technically competent but also aligned with human intent and usability. This section explores the facets of human-AI interaction and the importance of feedback mechanisms in AI training. AI systems are designed to serve human needs, and thus, incorporating human feedback is essential. Human input provides the contextual nuances that raw data cannot capture, guiding AI systems to better understand and interact with the world as humans do. This involves iterative processes where human trainers provide continuous feedback to AI systems, refining their outputs and enhancing their decision-making capabilities. Interactive machine learning is a paradigm where AI training involves a real-time feedback loop with human users. In this framework, the AI system learns directly from human interaction, adapting to corrections and suggestions. This approach can significantly improve the learning rate and the relevance of the AI's outputs to real-world applications. The design of AI systems must prioritize user experience, ensuring that interactions are intuitive, efficient, and satisfying. Incorporating principles of human-computer interaction (HCI) into AI training helps in creating systems that are user-friendly and accessible to a broad audience. This includes the design of interfaces, feedback mechanisms, and interaction protocols that are user-centered. Feedback loops are integral to refining AI models.

They allow for the incremental improvement of AI behavior through cycles of performance review and adjustment. By analyzing how humans interact with AI systems and the issues they encounter, developers can fine-tune models to better meet user expectations and needs. Human feedback must be ethically sourced and utilized, ensuring that it does not perpetuate biases or unethical outcomes. This is particularly important when feedback is used to train AI in sensitive domains, such as personalization algorithms or content moderation systems. Ethical oversight is necessary to ensure that the feedback used enhances the AI's societal value. Accurate data annotation and labeling by human experts play a crucial role in supervised learning. The quality of human-labeled data directly impacts the AI's performance. Therefore, developing robust and clear guidelines for data annotators is paramount, as is the recognition of the cognitive load and potential biases these human annotators may introduce. Feedback loops contribute to the robustness and adaptability of AI systems. By continuously testing and updating AI systems with new data and human insights, these systems can remain relevant and perform well even as the environment and tasks evolve. Looking forward, the goal is to develop a symbiotic relationship between humans and AI, where each enhances the capabilities of the other. AI systems should not only automate tasks but also provide insights and augment human decision-making. This synergy can lead to breakthroughs in how we solve complex problems and make decisions. In summary, human-AI interaction and feedback loops are foundational to the development of AI systems that are attuned to human needs and behaviors. By embracing these interactions, we can build AI that effectively complements human abilities and becomes an integral and trustworthy component of our daily lives.

Scalability is a cornerstone in the evolution of artificial intelligence (AI) systems from experimental models to real-world applications. The capacity to expand AI functionalities and manage increased loads without sacrificing performance is crucial. This section outlines the challenges and strategies inherent in scaling AI systems. Scalability in AI refers to the ability to maintain or improve performance as the size of the dataset and complexity of tasks increase. An AI system's scalability is determined by its architecture, the efficiency of algorithms, and the adaptability of its learning capabilities. Scaling AI systems involves technical, computational, and practical considerations to meet growing demands. As AI systems scale, they require more data to maintain accuracy and functionality. Efficient data management becomes critical, necessitating robust databases, data pipelines, and storage solutions. This also involves ensuring data quality and integrity throughout the AI lifecycle. Scaling AI systems often means increased computational demands. Optimizing resource allocation is essential to avoid bottlenecks and ensure timely processing. This may involve leveraging cloud resources, distributed computing, and edge computing solutions to

manage computational loads effectively. Deploying AI models to serve predictions at scale is a multifaceted challenge. It requires reliable model serving architectures that can handle high throughput and low latency demands. Techniques like model quantization and optimization are vital for efficient model deployment, especially in resource-constrained environments. As AI systems scale, the complexity of the models and the number of parameters often increase. Managing this complexity without compromising the speed or accuracy of the system requires innovative approaches to model design, such as modular architectures and hierarchical systems that can distribute the workload.



Pic. 3. Human AI interaction

Scalable AI systems must be capable of continuous learning, adapting to new data without the need for frequent retraining from scratch. Incremental learning and online learning algorithms enable AI systems to update their knowledge base continuously, making them more resilient to changes over time. Automating the deployment, monitoring, and management of AI systems is crucial for scalability. Orchestration tools and platforms can help manage the lifecycle of AI applications, including version control, testing, and continuous integration/continuous deployment (CI/CD) pipelines. Considering scalability at the outset of AI system design can preempt many of the challenges encountered during expansion. Designing with scalability in mind involves selecting the right algorithms, data structures, and infrastructure that can grow with the system's needs. As AI systems scale, so do their ethical and governance implications. It is vital to establish guidelines and monitoring mechanisms to ensure that the AI remains fair, transparent, and accountable, even as it grows in scope and influence. Ensuring the scalability of AI systems is not just a technical challenge; it is also about democratizing access to AI. Scalable systems can provide AI benefits to a broader range of users and developers, fostering innovation and inclusivity in the AI field.

The pursuit of leveraging artificial intelligence (AI) for social good is an inspiring facet of AI development. This section examines how AI can be a force for positive change, addressing societal challenges and enhancing the well-being of communities globally. AI has transformative potential in healthcare, from improving diagnostic accuracy to personalizing treatment plans. Machine learning models can analyze medical images, predict disease outbreaks, and optimize resource allocation in hospitals. AI-driven initiatives can bridge the gap in healthcare services, especially in underserved regions, by providing scalable solutions like telemedicine and AI-assisted diagnostics. Educational technologies powered by AI can provide personalized learning experiences, adapt to individual student needs, and identify gaps in knowledge. By analyzing data on how students learn, AI can assist in creating more effective educational content and teaching strategies, making quality education more accessible. AI applications are proving to be invaluable in environmental conservation, from monitoring wildlife populations to predicting climate change impacts. AI can optimize energy consumption, reduce waste through smart recycling systems, and assist in the design of sustainable infrastructures. The predictive power of AI aids in disaster response and management, providing critical information to mitigate the effects of natural calamities. AI is revolutionizing social services by enhancing the efficiency and reach of aid programs. Predictive models can identify areas of need, optimize the distribution of resources, and monitor the effectiveness of social interventions. AI tools can also assist in combating human trafficking and support legal aid services, providing support to the most vulnerable populations. AI-driven financial technologies can enable financial inclusion, providing access to credit and banking services for the unbanked. AI can also help small

businesses with market insights, automate administrative tasks, and offer personalized advice to foster economic growth in developing regions. AI has the potential to promote equality and inclusion by identifying and combating discriminatory practices. It can help in developing policies that ensure fair treatment across all sectors of society and assist in monitoring the enforcement of these policies. While AI has the potential to contribute significantly to social welfare, it is crucial to approach its deployment ethically. This includes ensuring the privacy of individuals, avoiding biases in decision-making, and engaging with stakeholders to understand the social context of AI applications. Realizing the potential of AI for social good requires collaboration between technologists, policymakers, non-profits, and community organizations. These partnerships ensure that AI solutions are grounded in real-world needs and are implemented responsibly. To understand the effectiveness of AI in social good initiatives, robust frameworks for measuring impact are essential. This involves setting clear objectives, collecting relevant data, and using AI itself to analyze the success and areas for improvement in various projects.

Summary.

The journey of training artificial intelligence systems is fraught with significant challenges, from the scarcity of quality data to the computational resources required and the need for robust and interpretable models. This paper has outlined the complexity of these challenges while also highlighting the innovative solutions that are emerging within the field. As AI continues to evolve, the importance of addressing these challenges head-on grows. The future of AI training lies in the balance of leveraging cutting-edge techniques to reduce data and computational demands and in ensuring the ethical and equitable development of technology. Embracing solutions such as transfer learning, meta-learning, and efficient model architectures will be crucial in democratizing AI, making it accessible to a wider range of users and developers. At the same time, the AI community must strive to make these systems as transparent and interpretable as possible, fostering trust and understanding among the public.

Список бібліографічного опису

1. Conditional Synthetic Data Generation for Robust Machine Learning Applications with Limited Pandemic Data / H. P. Das et al. ResearchGate. URL: https://www.researchgate.net/publication/354597494_Conditional_Synthetic_Data_Generation_for_Robust_Machine_Learning_Applications_with_Limited_Pandemic_Data (date of access: 09.11.2023).
2. Bex T. 5 Powerful Cross-Validation Methods to Skyrocket Robustness of Your ML Models. Medium. URL: <https://pub.towardsai.net/5-powerful-cross-validation-methods-to-skyrocket-robustness-of-your-ml-models-5edaab564cd8> (date of access: 09.11.2023).
3. Carneiro G., Santana E., R. Cordeiro F. A Study on the Impact of Data Augmentation for Training Convolutional Neural Networks in the Presence of Noisy Labels. URL: https://www.researchgate.net/publication/362907986_A_Study_on_the_Impact_of_Data_Augmentation_for_Training_Convolutional_Neural_Networks_in_the_Presence_of_Noisy_Labels (date of access: 18.11.2023).
4. Neural Network Pruning by Cooperative Coevolution. Thirty-First International Joint Conference on Artificial Intelligence (IJCAI-22). URL: <https://www.ijcai.org/proceedings/2022/0667.pdf> (date of access: 09.11.2023).
5. Mirzsoleiman B., Yu Liu T. Data-Efficient Augmentation for Training Neural Networks. URL: https://www.researchgate.net/publication/364443224_Data-Efficient_Augmentation_for_Training_Neural_Networks.

References

1. Conditional Synthetic Data Generation for Robust Machine Learning Applications with Limited Pandemic Data / H. P. Das et al. ResearchGate. URL: https://www.researchgate.net/publication/354597494_Conditional_Synthetic_Data_Generation_for_Robust_Machine_Learning_Applications_with_Limited_Pandemic_Data (date of access: 09.11.2023).
2. Bex T. 5 Powerful Cross-Validation Methods to Skyrocket Robustness of Your ML Models. Medium. URL: <https://pub.towardsai.net/5-powerful-cross-validation-methods-to-skyrocket-robustness-of-your-ml-models-5edaab564cd8> (date of access: 09.11.2023).
3. Carneiro G., Santana E., R. Cordeiro F. A Study on the Impact of Data Augmentation for Training Convolutional Neural Networks in the Presence of Noisy Labels. URL: https://www.researchgate.net/publication/362907986_A_Study_on_the_Impact_of_Data_Augmentation_for_Training_Convolutional_Neural_Networks_in_the_Presence_of_Noisy_Labels (date of access: 18.11.2023).
4. Neural Network Pruning by Cooperative Coevolution. Thirty-First International Joint Conference on Artificial Intelligence (IJCAI-22). URL: <https://www.ijcai.org/proceedings/2022/0667.pdf> (date of access: 09.11.2023).
5. Mirzsoleiman B., Yu Liu T. Data-Efficient Augmentation for Training Neural Networks. URL: https://www.researchgate.net/publication/364443224_Data-Efficient_Augmentation_for_Training_Neural_Networks.