E-ISSN: 2706-8927
P-ISSN: 2706-8919
www.allstudyjournal.com
IJAAS 2024; 6(SP-6): 194-197
Received: 27-04-2024
Accepted: 30-05-2024

Palak Kalsi
Assistant Professor, DPG College, Haryana, India

Corresponding Author:
Palak Kalsi
Assistant Professor, DPG College, Haryana, India

Globalization and Sustainable Development: Interdisciplinary Perspectives (GSD IP-24) (Conference proceedings of DPG DEGREE College)

Meta learning techniques for efficient working of NLP

Palak Kalsi

DOI: https://doi.org/10.33545/27068919.2024.v6.i6c.1241

Abstract
In the realm of learning models, there exists an ongoing opportunity for advancement. Initially, within machine learning, the notion involved applying task-specific algorithms to machines, enabling them to execute limited tasks. Subsequently, the emergence of deep learning empowered machines to learn from a diverse range of available data and analyze it using neural networks. However, this approach faces limitations, particularly regarding the availability of diverse data and resources. In scenarios where data variety and resources are lacking, results may vary. The introduction of meta-learning revolutionizes the learning process by examining experiences from multiple learning tasks and utilizing them to enhance future learning performances. Meta-learning, defined as the ability to learn how to learn, brings benefits by improving data and computational efficiency. In the global success of AI, each state has a significant role to play: machine learning improves engineering features, while deep learning enhances feature representation. Meta-learning within neural networks aims to replace manually designed algorithms with learned learning algorithms. This paper aims to explore the objectives and features of meta-learning and how its attributes address the limitations of traditional learning models, including those of deep learning and machine learning. It delves into issues such as computational challenges, efficiency requirements, and the significant fields where meta-learning finds application. Additionally, it discusses the potential applications of meta-learning across various domains such as natural language processing (NLP), speech recognition, design, and beyond. The primary emphasis is on how meta-learning contributes to the evolution of NLP and enhances the efficiency of NLP models.

Keywords: Machine learning, meta learning, NLP, deep learning

Introductions

Meta Learning Basics

Meta learning is an approach for multitasking scenarios where task-oriented knowledge is extracted from given multiple tasks and used to improve learning of future tasks. In easy words we can define meta learning as “learning to learn”. In conventional ML models learning take place at base level algorithms to solve a task where a single problem with defined dataset and objective. In meta learning there is upper-level algorithm which update base level algorithm. The major concern is to improve upper-level objective by generalizing performance or increasing learning speed of base level algorithms. The two major goals of meta learning is to “solving it and improving the strategies employed to solve it” (Schmidhuber). According to Bengio, Samy and Gloutier, meta learning based on genetical algorithms and gradient descent.

Meta-learning is like learning how to learn. Instead of just focusing on one specific task, it's about getting better at learning a variety of tasks. So, when you encounter new things to learn, you can pick them up more easily. (Debmalya Mandal, 2021) [4] Imagine you're trying to get better at doing different kinds of puzzles. Instead of just practicing one type of puzzle over and over again, you practice various types. (Debmalya Mandal, 2021) [4] This helps you develop skills that can be useful across different puzzles. Meta-learning involves looking back at your past experiences with different puzzles and using what you've learned to tackle new ones more effectively. The more similar the puzzles you've practiced before, the more helpful those experiences will be in learning new ones. However, figuring out which puzzles are similar enough to each other can be a bit tricky.

Meta-learning has turned out to be an important strategy the problem of limited data in various machine learning applications. The main idea behind meta-learning is to design learning algorithms that can leverage prior learning experience to adapt to a new problem quickly, and learn a useful algorithm with few samples

Difference between Conventional learning methods and meta learning

Conventional learning and meta-learning are two distinct paradigms in machine learning, each with its own principles, techniques, and applications. Here's differences between conventional learning and meta-learning:

- **Conventional Learning:** In contrast, meta-learning transcends the confines of single-task optimization and delves into the realm of learning how to learn. Rather than fixating on a singular task or dataset, meta-learning endeavors to equip models with the capability to swiftly adapt to new tasks or domains with limited data. It operates on two fundamental levels: the inner loop and the outer loop. (Timothy Hospedales, Meta-Learning in Neural Networks: A Survey, 2022) The inner loop entails rapid parameter updates based on insights gleaned from small datasets pertaining to individual tasks, facilitating quick adaptation. Meanwhile, the outer loop orchestrates the overarching learning process, honing the model's ability to generalize across diverse tasks or domains. Unlike its conventional counterpart, meta-learning thrives in scenarios where data may be scarce or where the demand for rapid adaptation reigns supreme. Its applications span the spectrum, ranging from few-shot learning and transfer learning to hyperparameter optimization and domain adaptation. (Timothy Hospedales, Meta-Learning in Neural Networks: A Survey, 2022) By leveraging knowledge gleaned from multiple tasks, meta-learning empowers models to transcend the constraints of individual datasets and exhibit robust performance across a plethora of real-world applications.
- **Meta-learning:** Meta-learning, in contrast, sets its sights on the acquisition of learning itself. Rather than fixating on singular tasks, meta-learning casts its net across a multitude of tasks or domains. Its mission: to cultivate algorithms or models capable of swift adaptation to novel tasks or domains, even in the face of sparse data. Meta-learning often demands less data per task, drawing upon insights garnered from a tapestry of related tasks. Crafted to transcend task-specific constraints, meta-learning algorithms boast a prowess in generalization, enabling seamless adaptation to fresh tasks with scant data. (Timothy Hospedales, Meta-Learning in Neural Networks: A Survey, 2022) Meta-learning operates on dual fronts: the inner loop and the outer loop. Within the inner loop, models glean wisdom from scant task-specific data to swiftly refine their parameters. Meanwhile, the outer loop orchestrates overarching parameter updates across tasks, guided by insights gleaned from validation or meta-training data. (Timothy Hospedales, Meta-Learning in Neural Networks: A Survey, 2022) Emphasizing broad generalization across tasks and domains, meta-learned models stand poised to swiftly adapt to novel tasks with minimal data, charting paths into uncharted territories with aplomb. In realms where data is scant or where rapid adaptation reigns supreme, meta-learning shines brightest. Its applications span the gamut, from the realms of few-shot learning and transfer learning to the domains of hyperparameter optimization and domain adaptation.

Overview of NLP

Natural Language Processing (NLP) stands at the crossroads of computer science and linguistics, harnessing the power of machine learning to facilitate seamless human-computer interaction. Its primary aim is to streamline communication between humans and machines, ensuring both ease and efficiency in the exchange of information. Through the lens of NLP, machines imbibe the intricacies of human language-its syntax, semantics, and pragmatics-processing it to deliver meaningful outputs to users.

The realm of NLP is dedicated to empowering computer systems to execute tasks of significance using natural, human-understandable language. Its significance lies in its ability to furnish us with models and methodologies capable of ingesting diverse forms of input-be it in the form of spoken words, written text, or even images-and manipulating them in accordance with predefined algorithms. Consequently, NLP systems adeptly process inputs in various modalities, including speech and text, yielding outputs that cater to the needs of users. As we march towards the future, the indispensability of NLP becomes increasingly evident. Its transformative potential lies in its capacity to pave the way for the development of sophisticated models and streamlined processes. By embracing NLP, we unlock the ability to harness vast swathes of information and seamlessly navigate the ever-expanding landscape of human-computer interaction.

"Natural Language Processing (NLP) is a multidisciplinary field at the intersection of linguistics, computer science, and artificial intelligence. It focuses on enabling machines to understand, interpret, and generate human language in a way that is both meaningful and contextually relevant. NLP involves a range of tasks, such as text analysis, language translation, sentiment analysis, and information retrieval. The ultimate goal of NLP is to bridge the communication gap between humans and machines, enabling seamless interaction and understanding." (Abu Rayhan1, 2023) [1]

Challenges of NLP

The ambiguity and creativity of human language are just two of the characteristics that make NLP a demanding area to work in. (Sowmya Vajjala, 2020) [7] This section explores each characteristic in more detail, starting with ambiguity of language. Ambiguity means uncertainty of meaning. Most human languages are inherently ambiguous. Consider the following sentence: "I made her duck." This sentence has multiple meanings. The first one is: I cooked a duck for her. The second meaning is: I made her bend down to avoid an object. (There are other possible meanings, too; we'll leave them for the reader to think of.) Here, the ambiguity comes from the use of the word "made." Which of the two meanings applies depends on the context in which the sentence appears. (Sowmya Vajjala, 2020) [7].

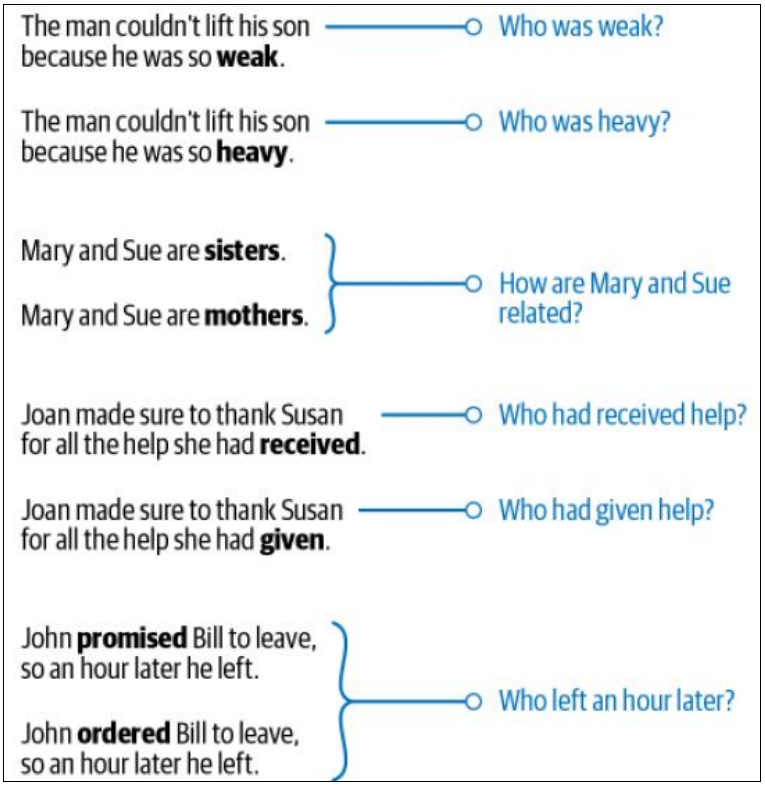

Fig 1: Examples of ambiguity in language from the Winograd Schema Challenge (Sowmya Vajjala, 2020) [7]

How approaches of Meta Learning helpful in NLP Tasks

The capacity of meta-learning techniques to enhance model adaption, efficiency, and generalization makes them more and more useful for natural language processing (NLP) applications. In NLP, meta-learning is advantageous in the following ways:

Adjustment to Novel Tasks: By learning how to draw lessons from past experiences, meta-learning helps models to quickly adjust to novel tasks with little data. In NLP, this indicates that a model doesn't require intensive task-specific training because it can generalize well across several task types (Such as sentiment analysis and named item identification).

Two-shot and one-shot Learning: By using meta-knowledge gained during training, models are able to generate predictions or complete tasks with very few or even no instances thanks to meta-learning techniques that enable few-shot and zero-shot learning scenarios. This is especially helpful in situations where labeled data is available.

Applications of Meta Learning in NLP

There is a growing number of studies applying metalearning techniques to NLP applications and achieving excellent results. In the second part of the tutorial, we will review these studies. Here we summarize these studies by categorizing their applications

1. Text Classification Text classification has a vast spectrum of applications, such as sentiment classification and intent classification. The meta-learning algorithms developed for image classification can be applied to text classification with slight modification to incorporate domain knowledge in each application.
2. Sequence Labeling Using a meta-learning algorithm to make the model fast adapt to new languages or domains is also useful for sequence labeling like name-entity recognition (NER) (Wu *et al*., 2020) and slot tagging. However, the typical meta-learning methods developed on image classification may not be optimal for sequence labeling because sequence labeling benefits from modeling the dependencies between labels, which is not leveraged in typical meta-learning methods. Techniques, such as the collapsing labeling mechanism, are proposed to optimize meta-learning for sequence labeling problem.
3. Automatic Speech Recognition and Neural Machine Translation Automatic speech recognition (ASR), Neural machine translation (NMT), and speech translation require a large amount of labeled training data. Collecting such data is cost-prohibitive. To facilitate the expansion of such systems to new use cases, meta learning is applied in these systems for the fast adaptation to new languages in NMT (Gu *et al*., 2018) and ASR, and fast adaptation to new accents, new speakers, code-switching (Winata *et al*., 2020a) in ASR.
4. Classification of Texts There is a wide range of applications for text classification, including intent and sentiment categorization. To incorporate domain knowledge in each application, the meta-learning algorithms created for image classification can be slightly modified and applied to text classification.
5. Labeling Sequences for sequence labeling, such as name-entity recognition (NER) (Wu *et al*., 2020) and slot tagging, using a meta-learning approach to help the model quickly adapt to new languages or domains is also useful. On the other hand, sequence labeling may not benefit as much from standard meta-learning techniques since they do not take advantage of the

benefits that sequence labeling derives from modeling label dependencies. These techniques were created for picture classification. Methods are suggested to optimize meta-learning for sequence labeling problems, such as the collapsing labeling mechanism.

6. Neural Machine Translation and Automatic Speech Recognition A significant amount of labeled training data is needed for voice translation, neural machine translation (NMT), and automatic speech recognition (ASR). Such data is prohibitively expensive to gather. Meta learning is applied in these systems for fast adaptation to new languages in NMT, ASR, and fast adaptation to new accents, new speakers, and code-switching (Winata *et al*., 2020a) in ASR. This allows the expansion of such systems to new use cases.

Future Directions

Within Natural Language Processing (NLP), meta learning is a new subject that seeks to enhance the effectiveness and efficiency of NLP models. It does this by making use of expertise from earlier assignments or datasets.

Effectiveness in a Variety of Ways: NLP meta learning techniques seek to improve algorithms in a number of ways, including boosting generalizability and data efficiency. Meta learning models are able to recognize patterns and structures that are pertinent to certain NLP tasks by learning from tasks or datasets.

Increasing Data Efficiency: By enabling models to learn from a variety of instances, meta learning helps to increase data efficiency. Even with a small amount of training data, meta learning algorithms may adapt to tasks by learning how to generalize from a series of tasks.

Conclusion

Meta-learning is a paradigm that has the potential to revolutionize machine learning by addressing data scarcity, rapid adaptation, and transfer learning. It enables models to learn from limited data and quickly adapt to new tasks, demonstrating remarkable generalization capabilities. Its applications span diverse domains, including computer vision, natural language processing, robotics, healthcare, and more. Benefits include faster adaptation, efficient knowledge transfer, improved generalization, and resource efficiency. However, challenges like task selection, hyperparameter tuning, and potential over fitting pose challenges to successful implementation. Despite these, the rewards of harnessing meta-learning's potential are substantial. As researchers push meta-learning boundaries, new techniques and advancements are likely to emerge, offering more robust models capable of adapting to ever-changing environments and learning from limited data. By embracing this paradigm, we can shape the future of AI, create models that perform well on specific tasks, and excel in a wide array of scenarios.

References

1. Abu Rayhan RK. Natural Language Processing: Transforming How Machines Understand Human Language. The Development of Artificial General Intelligence; c2023. p. 14.
2. Mandal SM. Meta-Learning with Graph Neural Networks; c2021 Mar 2. p. 10.
3. Francois F. The Little Book of Deep Learning; c2023.
4. Kelleher JD. Deep Learning. London: MIT Press; 2019. Schmidhuber J. Deep Learning. Available from: http://mitpress.mit.edu
5. Gao K, Sener O. Modeling and optimization tradeoff in meta-learning. Advances in Neural Information Processing Systems; c2020 Oct. ISSN 1049-5258. DOI: 10.48550/arxiv.2010.12916. Available from: https://arxiv.org/abs/2010.12916
6. Deya S, Al. SigNet: Convolutional Siamese Network for Writer Independent Offline Signature Verification. Elsevier; c2018. p. 7.
7. Vajjala SM. Practical Natural Language Processing. O'Reilly; c2020.
8. Hospedales AA. Meta-Learning in Neural Networks: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2022 Sep;44(9):21.
9. Hospedales AA. Meta-Learning in Neural Networks: A Survey. IEEE Xplore. 2022 Sep;21.
10. Vanschoren J. Meta-Learning: A Survey. Eindhoven University of Technology; c2018. p. 29.
11. Liang K, Zhang JY, Wang B, Yang Z, Koyejo O, Li B. Uncovering the Connections Between Adversarial Transferability and Knowledge Transferability; c2021.
12. Lin M, Lucas H, Shmueli G. Too big to fail: Large samples and the p-value problem. Inform Syst Res. 2013 Dec;24(4):906-917. DOI: 10.1287/isre.2013.0480.
13. Maji S, Rahtu E, Kannala J, Blaschko MB, Vedaldi A. Fine-Grained Visual Classification of Aircraft. 2013. Marr D. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. Phenomenology and the Cognitive Sciences. 1982;8(4):397. ISSN 1568-7759. Available from: http://mitpress.mit.edu/catalog/item/default.asp?ttype=2&tid=12242&ref=nf
14. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference. 2018 Oct;1:4171-4186. Available from: https://arxiv.org/abs/1810.04805v2
15. Dhillon GS, Chaudhari P, Ravichandran A, Soatto S. A Baseline for Few-Shot Image Classification; c2019. Available from: http://arxiv.org/abs/1909.02729
16. Finn C, Abbeel P, Levine S. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks; c2017. Available from: http://arxiv.org/abs/1703.03400. Available from: https://arxiv.org/abs/2010.12916v2