

Summary and Contributions

The authors give an overview and review of the emerging field of learned optimizers. Learned optimizers would alleviate the need for handcrafted optimization algorithms and costly hyperparameter tuning.

The authors give an introduction to the supervised deep learning setting. They explain the task of optimization in such a setting and give formal context. Popular, *traditional* gradient based methods are described in detail.

The authors then continue by giving a similarly detailed introduction to the notion of learned optimizers. They also support their formal explanation with figures to help the reader build an intuition.

With a formal background established, different approaches to tackle the challenge of learning optimizers are introduced. Special emphasis is layed on a strain of work by Metz et al. which is discussed further. The results of this work are described in detail and are critically evaluated.

Strengths

The reviewer positively notes that

- the *Fundamentals* section gives a detailed formal background,
- the written language is precise and free of errors,
- the relevance of the topic is stated clearly,
- visuals are used well,
- the idea of learned optimizers is explained in detail,
- literature is well selected
- weaknesses of the discussed work are assessed critically,
- and specific problems are pointed out.

Weaknesses

The reviewer negatively notes that

- the *No-free-lunch-theorem* could use more context,
- and that some commas were wrong.

Correctness

While the reviewer is not particularly familiar with the relevant literature, all references appear to be cited properly. The reviewer could not find any false formulations or equations.

Clarity

The work is easy to follow. The structure is well composed. Language is precise and detailed. Equations and text support each other. Visuals provide further intuition where needed. Overall, the work is clear and precise.

Further Comments

The work provides a good overview over the emerging field of learned optimizers. While clear motivation is given, all described work is discussed critically, making it a very good review. The reviewer recommends to publish it (with maybe less background but additional ablation studies that are missing in the Metz et al. study).

Rating and Confidence

- Overall Score: 9.9
- Confidence: 7.5