Reviewer 1

The authors study preconditioning issues in the Lipschitz-smooth optimization problems and propose Preconditioned Nesterov method (PN) and Preconditioned Heavy Ball method (PHB) with theoretical guarantees of convergence under unified assumption on scaling matrix. The performance of the proposed algorithms is illustrated on a couple of test datasets.
The obtained theoretical results may be of interest for various machine learning techniques.
I recommend publishing the paper in the conference proceedings. Probably, the amount of pages (16 pages) is a bit excessive and should be reduced to the required maximum of 15 pages.

Prof. Dmitri Kvasov.
Univ. of Calabria, Italy

Reviewer 2

Machine learning and deep learning are well-known approaches that allow solving many modern problems. In optimization theory, the Nesterov method and the heavy ball method are used for these purposes. But minimization problems can be ill-conditioned, which greatly affects the efficiency of the above-mentioned methods. In this situation, an approach based on the preconditioning matrix can be a solution to the problem.
The following results were obtained in the paper: two algorithms are presented, which are scaled versions of the heavy ball method and the Nesterov method. Guaranteed justifications for these methods were obtained, numerical experiments were carried out, and the advantages of the obtained methods in comparison with unscaled methods were shown.
The manuscript is proposed for publication in the conference proceedings.

Prof. Anatoly Antipin,
FRC Informatics and Control of RAS, Russia

Reviewer 3

This paper evaluates optimization methods for machine learning problems. The authors propose Preconditioned Nesterov (PN) and Preconditioned Heavy Ball (PHB) methods and present their proofs of convergence. Therefore, the objective of this study is interesting for the audience of the conference. Also, a strong point of the manuscript is the fact that, the authors perform some encouraging comparative computational results. The paper is fairly-well presented, describing in sufficient detail the problem, related work, proposed approach and evaluation. Thus, I believe that, the paper can be accepted for presentation at the conference.

My only minor remark is that it would be beneficial for the readers, if the authors
i) could provide at least a short discussion of their findings, after the computational experiments, and also
ii) a short section with conclusions at the end of the manuscript.

Prof. Angelo Sifaleras,
Univ. of Macedonia, Macedonia