Group

Title	Latent Weights Do Not Exist: Rethinking Binarized Neural Network Optimization
Student 1	Nik Vaessen
Student 2	Peter Mastnak
Student 3	Sri Datta Budaraju

Main criteria

1. Report readability

The paper is clearly written and is easy to follow. The reproduced experiments and plots from the original paper are clearly referenced, making the reproducibility study easy to follow. Some paragraphs (highlighted in the annotated report) are very close to the original wording from the authors and should be properly referenced.

2. Paper summary

The paper is summarized clearly and relevant notions from the literature (such as pseudo-gradients) are described to aide readability, including the architectures and hyperparameters used in the experiments.

3. Reproducibility thoroughness

The students reproduce most of the authors' results and further perform experiments on the role of different optimizers, learning rate scales and initialization scales on MNIST. The HP choices are taken from the paper and ad-hoc choices made by the students are mostly motivated.

4. Reproducibility discussion

The students are able to qualitatively match almost all of the results reported in the NeurIPS paper (besides ImageNet, which was skipped for computational reasons). Quantitatively, the obtained results do not always match those reported by the original paper, but the students argue for why this might be the case.

5. Difficulty (theory)

The chosen paper proposed an optimization method designed for binary neural networks. Since the training algorithm and experimental setup are explained clearly and the number of hyperparameters is relatively low, I believe the paper has a theoretical difficulty of 1-2.

6. Difficulty (implementation)

Similarly to what described in the previous point, the well documented paper and the availability of the authors' source code should make the implementation straightforward. Challenges may arise due to the computational resources required for training large networks on ImageNet. The paper should further lend itself to empirical experimentation, making it a good target for a reproducibility study (difficulty of 2).

Bonus points

- **Different framework**: the original source code is released for Tensorflow while the student reimplemented the experiments from scratch in Pytorch.
- **Extra dataset**: the students runs a preliminary experiment on MNIST.