Viability of FF compared to standard backpropagation

This report shows the comparison between a standard model trained with backpropagation and the same model trained with the Forward Forward approach

Omar Facchini

This is a short version of the report that focuses solely on the results, for a more complete report please refer to this report.

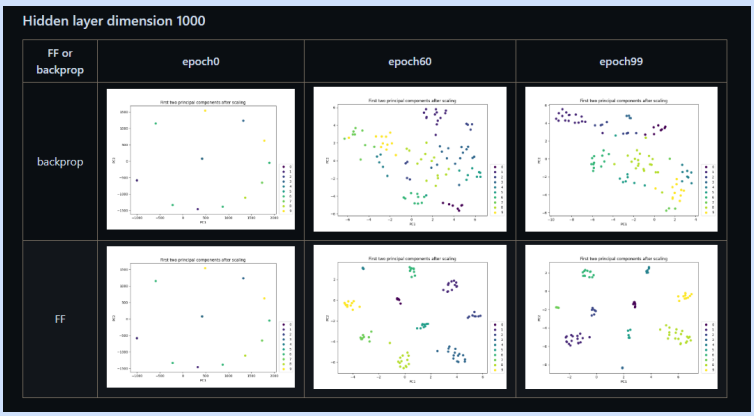
Classification performances

In this section the accuracy results are presented, showing that the Forward Forward approach is significantly more accurate while also maintaining a much lower loss, which is already better than what was expected.

It's also interesting to find out that while the model trained with backpropagation do not follow a pattern, the training with Forward Forward approach seems to have a consistent pattern and, especially with higher hidden layer dimensions, seems to saturate and maintain a consistent value.

Feature visualization

In this section the feature visualization results are presented, showing that the Forward-Forward approach is able to generate much more coherent and accurate clusters compared to the backpropagation approach using the t-SNE approach.



Here can be seen the comparison of the clustering obtained from the classifiers of the models trained with backpropagation and Forward-Forward with the dimension of the hidden layers set to 1000

These results only show the t-SNE approach but PCA was tested too, for a more in-depth showing of the comparison between t-SNE and PCA approach refer to this report.

As an extended comparison between backpropagation and Forward-Forward was needed, only the most relevant results were included in the report, for a more in-depth view on how the different layers, as well as the classifiers, evolve during the epochs with different dimensions and the application of normalization to the data refer to this report instead.

Conclusions

While the newly presented approach, in the Forward Forward, seems to be a viable option it is yet to be tested on complex datasets to better ensure the viability but for simple datasets and fairly small networks it would be a great option that also allows to have a clear visualization of the features.

So even in the unfortunate case in which the approach is not able to scale well, it could still be used to better understand the system and the data used with little cost.