More ideas

* Take the time, not just the accuracy
* use other evaluation methods (cfr. Medvet)
* do a table to compare different learning rates, etc.
* separate and fin the contribution of each improvement, maybe studying the combinations