Back to CV

Download thesis

Looking back at my bachelors thesis (about 2 years after I wrote it), there is no longer much I am proud of.
I choose my topic "Entwicklung eines Deep Learning Verfahrens zur Ladungsrekonstruktions mit dem Übergangsstrahlungsdetektor des AMS-Experiments"(German for "Development of a Deep Learning Method to reconstruct charges from the Transition radiation detector of the AMS experiment") because I did not know anything about deep learning, and I thought my bachelor thesis would be a great opportunity to learn about it.
And even though this worked quite well, everything else didnÂ´t go that well. I talk more about what I think went wrong and what IÂ´ve learned from this here.

I was working in the AMS group of my university, and my goal was to create a model, that could take in hits in the AMS TRD (Transition Radiation Detector) and try to reconstruct the (absolut value of the) charge of an input particle.
Ordinarily you can do this, by looking at the total energy deposition, which should proportional to the square of the particles charge.
This was not used in the group, mostly because other parts of AMS, like for example the RICH (Ring Image Cherenkov counter) detector is better suited for this task, and the TOF (Time Of Flight) detector is able to also get the sign of the particles charge.
My task was now, to try to use the TRD data, to get a better charge reconstruction (using machine learning) than using rutherfords formula.
Because of the abundance of measurements in this group, I was able to work on actual detector measurements, and thus tried to reconstruct the charge that the group generated from the whole of AMS from just the TRD.
The only limitation was, that I had to limit my model to charges between 1 and 8 electron charges (particles that are not charged leave no marks in a TRD).
This task is fairly limited for machine learning, since the input data is not (completely) translation invariant, which restricted me to use simple dense networks.
These migth not be very fast to train, but the group had neirly infinite computational ressources for their other analysis tasks, so this was not a problem. A more interresting question was the type of machine learning task: Since charges of incomming nuclei are quantizied to integer mulplicates of electron charges, it seems to be appropiate to use a classification network, but since we prefer a particle of charge 9 to be identified as a particle of charge 8 than of a charge 2, using a regression task has also benefits.
I tried both to compare them later. Sadly this comparison is not so easy, since finding a clear loss function is not trivial (If you also take a look at my master thesis (or its abstract), there I write multiple subchapters on how I defined my loss function for a choice between 2 classes, and here I have 8 classes, that are also distributed drastically different: While basically all particles have a low charge, those particles with a high charge are usually the most interresting).
I prefered the regression task, since even though a regression network allows for basically every charge, by training it on quantizied outputs, you also get mostly quantized outputs, and the deviation from this quantized expections is a good indication of the uncertainity of the classification (Even though it is not simply the uncertainity).
The fact that defining a loss function is not so trivial, also makes giving you final numbers here not so easy.
My network was able to reconstruct more than 99% of all expected particles correctly, but this is mostly a consequence of low charge particles beeing reconstructed better than high particles (even though in my training set each charge was represented the same).
For high particles, the accuracy never falls below 60% and most errors are only a few charges away. The average difference between the charge measured by AMS and by my network, amounts to about #0.25# which is considerably lower, than by just using rutherfords formula, which resulted in an average difference of #0.4#