REPORT OF OPPONENT OF BACHELOR'S THESIS

Title: Machine learning for recognition of simple physical systems
Author: Jan Benda

SUMMARY OF THE WORK

This thesis studies a deep learning method for the solution of Hamiltonian systems. The thesis provides a good summary of the basics of deep learning, the Hamiltonian systems studied, and an existing technique for using deep learning to solve these systems — Direct Poisson Neural Network (DPNN). The thesis provides some initial numerical experiments for the solution, and then explores several potential extensions and issues, and how they could be countered — notably, the effect of noise on the training data, how the solutions from the neural network behave outside of the training regions, and the implementation of Ehrenfest regularisation for dissipation.

The method studied is interesting and fairly complicated. The amount, and quality, of the mathematical and computational work is excellent. However, the thesis does lack in some areas, especially in explanation of the results.

OVERALL EVALUATION OF THE WORK

Thesis topic. The thesis topic was suitable for a batchelor thesis, and the mathematical and computational work appears to be been completed to a good standard.

Author's contribution. The author explored several issues for DPNN, with both mathematical background and numerical experiments, such as the effect of Gaussian noise on training data and the effect on the solution when extended outside the region studied by the training set. Additionaly they suggested potential fixes for these problems, and also added in dissipative behaviour, via Ehrenfest regularisation, to the studied problem and the DPNN.

Mathematical level. The mathematical content is generally good and very well structured. However, there are several things mentioned which are not explained or referenced.

The first two chapters give an excellent overview of the background of the mathematics behind deep learning with neural networks and the studied Hamiltonian systems, respectively. The only issues in these chapters is the description of mini-batch gradient descent algorithm, where it is not clear how the iteration of the gradient descent algorithm and the "mini" batches are related — essentially, what is the relationship between the mini-batch index k and the gradient descent iteration i?

The third chapter has a good summary of how DPNN can be used to solve the Hamiltonian system, and provides some good numerical experiments; while the four chapter looks at a number of issues and extensions. There are several issues here. It is not clear from the flow charts exactly how the DPNN works for the various methods; for example, what does the "@" symbol mean joining two of the processes. A more detailed description of how the process work would have been useful. Additionally, several things are mentioned without reference or description — for example, what is the *softplus* activation function and what is *autograd*? It is also not always clear what is being shown by the figures, and how they relate to the description of the results. In fact, the figure references within the text appear to be often incorrect.

Sources. The provided sources are excellent. However, they are several things mentioned in the thesis which should have additional references — notably, the *softplus* activation function and *autograd*.

Formal preparation. The level of English in the thesis could do with some improvement. Generally the structure and presentation is good. The main issue appears to be in cross referencing, especially to result figures, where it is not always clear if an equation, section or figure is being referenced (numbers used without a prefix); Additionally often the figure referenced appears to be incorrect; for example, in the first paragraph of page 18 figures 3.6 and 3.7 are referenced, but it should be 3.3 and 3.4.

Comments and questions

1. On page 20, it is stated that the "paraboloid is upside-down and shifted but that is again due to the degrees of freedom discussed in the demo with harmonic oscillator". I was unable to find any such discussion which mentions degrees of freedom. Can you explain more fully why the paraboloid is inverted and shifted?
2. *autograd* is mentioned several times, without description. Additionally, it is described as "computationally expensive". Why is it computationally expensive, and what effect does this have?
3. In the discussion of reducing the error introduced by the noise it is mentioned that dropout with a value of $p = 0.3$ obtains the best results for the harmonic oscillator. Is this value independent on the variance σ^2 used in the Gaussian noise?

Conclusion

I **recommend** this thesis as a Bachelor's Thesis

Scott Congreve, Ph.D.
Prague, 19.06.2024