LSTM research

Bernhard Nornes Lotsberg

February 4, 2020

LSTM

- RNN with long and short term memory.
- Long term memory not possible in ordinary RNN's because of vanishing gradients.
- Long term memory is very important for runoff modelling, as some features take a long time to impact. Snow for example.

EA-LSTM

- Paper [2] introduces the EA-LSTM (Entity Aware), which is a modified version of the LSTM designed to better with this this problem.
- Beneficial for hydrological modelling because it is able to process information about the current catchment that it is modelling (?).

Usage of existing code

- Notebook tutorial for using LSTM used in paper [1] can be found at `https://github.com/kratzert/pangeo_lstm_example`
  - This repository includes a binder link for running the notebook in browser.
  - Notebook shows easy to understand examples of how to use Pytorch and how to load the CAMELS dataset.
  - Only the first experiment in the paper.
- Code for paper [2] can be found at `https://github.com/kratzert/ealstm_regional_modeling`
  - This github page contains links to datasets used as well as pre trained models.
  - Need stronger computer to recreate the models, but should hopefully be possible to use the pre trained models from laptop.
  - Runs fine on my home desktop.
  - All results in the paper and all relevant code is here, including pre-trained models and code for creating and training models yourself.

- Code for paper [3] can be found at https://github.com/kratzert/lstm_for_pub
  - Includes step-by-step guide on how to recreate results from article. (Including bash scripts for automation)
  - Need Matlab for a few plots, but unsure if this is important
- Spend most time on the first paper in the beginning, it is important to actually understand how an LSTM model works (not to mention RNNs in general!).
- Paper [3] seems to take a lot more computational resources to recreate. It performs tuning of hyperparameters as well as using an ensemble of ten independently trained LSTMs.

Purely data driven model

- A lot of code already exists
- First thing to try is to explicitly provide snow data (which is not done by papers cited in this document).
- Possibly unreliable as we have less physical intuition, but unsure if a hybrid model performs better.
- LSTM or EA-LSTM
- There is room for improvement in a basic manner as well, as the papers state that a proper hyperparameter search hasn't been performed. This is probably time consuming and boring, though. (Kratzert has done this better in later research according to the talk)
- Also, only simple single and double layer LSTMs have been used, with few hidden cells. Parameters in the lower hundreds.
- Humidity should be considered, as it is not provided to the model right now according to the talk given my Kratzert.
- Integrated gradients can be used to interpret model. https://github.com/hiranumn/IntegratedGradients

Hybrid model

- Could be more intuitive, possibly also more reliable.
- Trained using Shyft in the training process. Using the output from Shyft as an input for instance.
- Probably more difficult to train.
- LSTM or EA-LSTM

Choice of model

- I think starting with a purely data driven model and then try introducing output from Shyft along with data not used by Shyft.
- Several configurations need to be tested.
- Before any of this I need to actually learn how to use Pytorch and how LSTMs work.
- Paper [3] suggests that physical constraints applied to the LSTM models could improve results.
- Hybrid model should at least be attempted!
- All papers referenced use the CAMELS dataset, could be interesting to try data from Statkraft.

References

[1] F. Kratzert, D. Klotz, C. Brenner, K. Schulz, and M. Herrnegger, "Rainfall–runoff modelling using long short-term memory (lstm) networks," *Hydrology and Earth System Sciences*, vol. 22, no. 11, pp. 6005–6022, 2018.

[2] F. Kratzert, D. Klotz, G. Shalev, G. Klambauer, S. Hochreiter, and G. Nearing, "Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets," *Hydrology and Earth System Sciences*, vol. 23, no. 12, pp. 5089–5110, 2019.

[3] F. Kratzert, D. Klotz, M. Herrnegger, A. K. Sampson, S. Hochreiter, and G. S. Nearing, "Toward improved predictions in ungauged basins: Exploiting the power of machine learning," *Water Resources Research*, vol. n/a, no. n/a.