Learning to learn neural networks

12/12/2015

Tom Bosc, Inria Sophia-Antipolis, Wimmics team

Meta-learning with RNNs

RNNs can learn supervised, online learning algorithms to train FNNs

Method : the state of the RNN contains the parameters of the FNN.

At each timestep :

- The model sees new data and makes a prediction
- The RNN updates the parameters of the model

Learning to learn using gradient descent, Sepp Hochreiter, A. Steven Younger, Peter R. Conwell, 2001

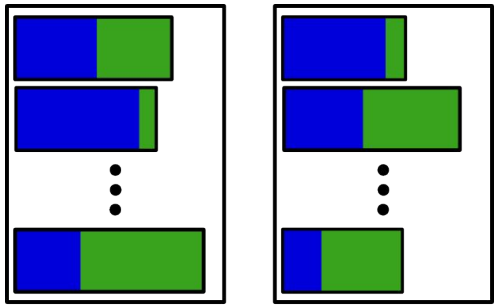
- Learning coefficients of quadratic functions from samples
- Slow meta-learning but fast learning

Data

Learning: one dataset

training objective: average loss over train set

Meta-learning: set of datasets

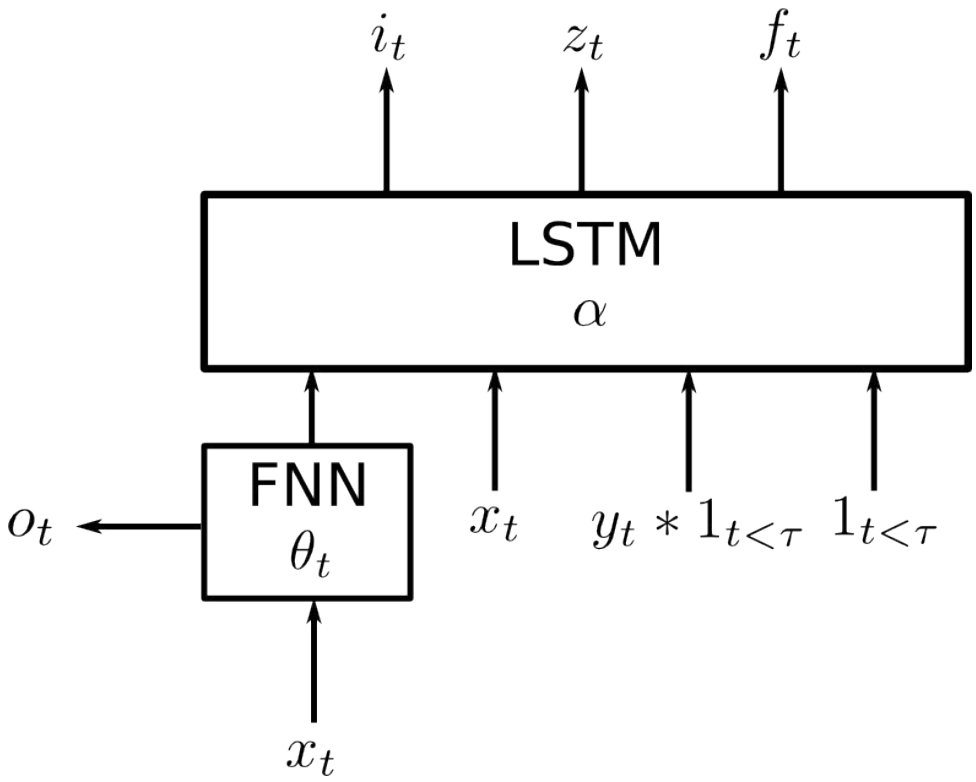


training objective: average loss over test sets in D_{train}
penalizes overfitting and underfitting

Train set
Test set

Architecture

$$\theta_{t+1} = i_t * z_t + f_t * \theta_t$$



Experiments

Binary classification task

Artificial data, variable noise, includes non-linearly separable datasets

FNN: 225 parameters: 5 inputs, 32 hidden units, 1 output

LSTM: 207651 parameters

	LSTM	Logistic Regression	SVM (linear)	SVM (RBF)
μ MCE	0.540	0.574	0.573	0.507
σ MCE	0.139	0.208	0.159	0.164

Conclusion

Recurrent neural nets can train feedforward neural nets.

Works with potentially deep nets, on noisy data.

Please see the paper on the RAM website for more about the data, cost function, illustrations and future work.