POZNAŃ UNIVERSITY OF ECONOMICS AND BUSINESS

FACULTY OF INFORMATICS AND ELECTRONIC ECONOMY

Selected Properties of Artificial Neural Networks in Application to Prediction of an Equity Price Index

Wybrane własności sztucznych sieci neuronowych w zastosowaniu do prognozowania pewnego indeksu cen akcji

Masters thesis in Financial Engineering

Author: **Vincent Ibia**

Thesis supervisor: **Prof. Witold Jurek**

Outline

1. Artificial Neural Networks
2. Types of ANNs
3. Selected Properties of ANNs
4. Hypothesis
5. Details of Compared Networks
6. Results from Compared Networks
7. Network Comparisons
8. Conclusions

Artificial Neural Networks

Based initially on biological structures of the same name, ANNs are computational models capable of solving complicated engineering tasks[1].

Several designs for ANNs originate from Machine Learning (ML), and in supervised learning tasks from ML, ANNs can serve as universal approximators for non-linear problems[1].

1. Basterrech, S., & Rubino, G. (2017). Echo State Queueing Networks: A combination of Reservoir Computing and Random Neural Networks. *Probability in the Engineering and Informational Sciences* (31), 457-476.

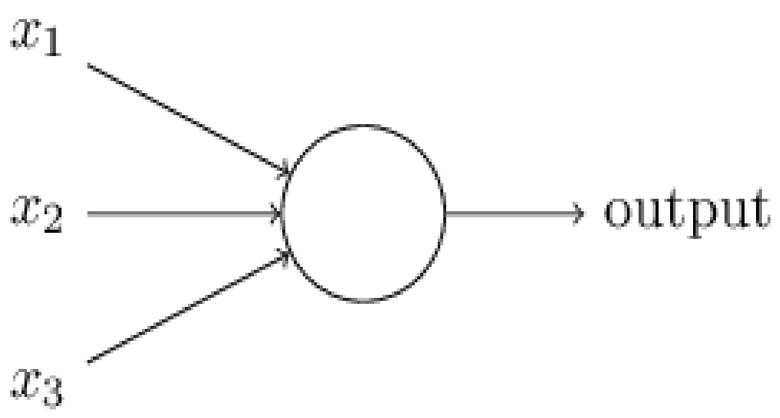
Types of ANNs

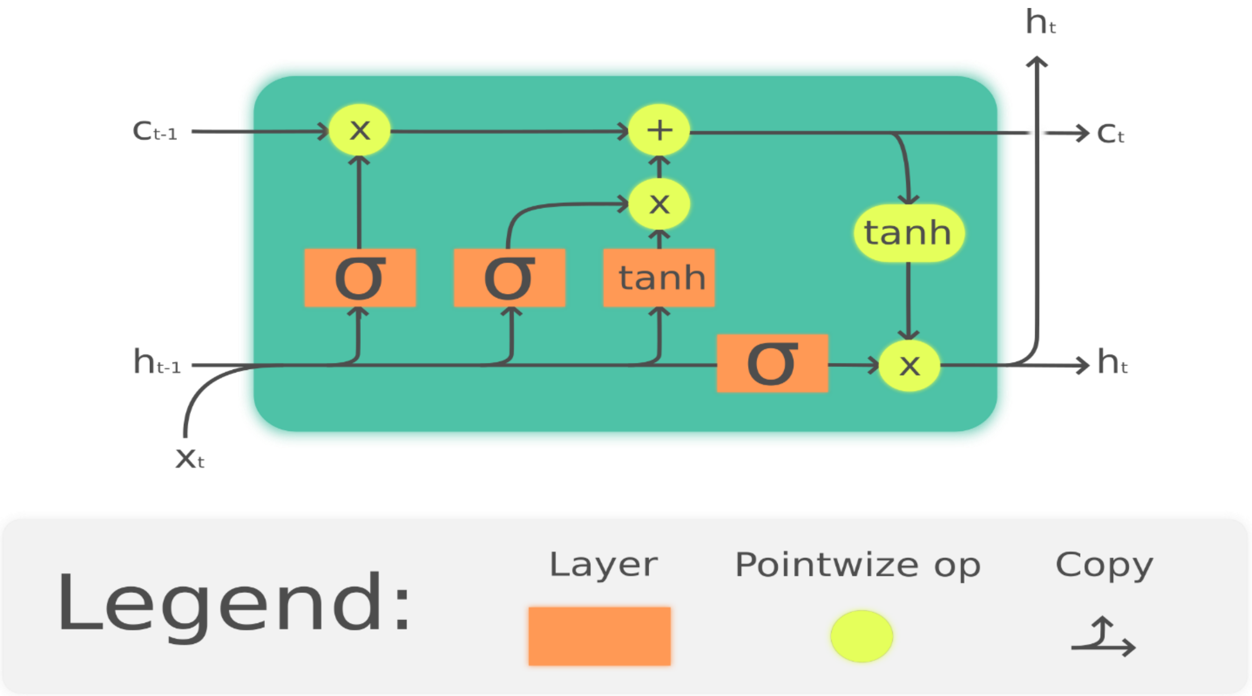
ANN Types have to do with the way in which their nodes/layers are organized, and how learning takes place within the network. Broadly, we have (at the least) Feed-Forward networks, and Recurrent Networks.

- Feed-Forward: information only move forward from earlier layers
  - Examples are Multi-layer perceptron (MLP) and Convolutional Neural Network (CNN)
- Recurrent: information from a layer may persist or move backward
  - Examples are Recurrent Neural Network (RNN) and Long Short-term Memory (LSTM)

Details of Compared Networks

Nodes Used: Perceptron Neuron (left) versus LSTM Unit (right)





1. Nielsen, M. A. (2015). *Neural Networks and Deep Learning.* Determination Press.
2. Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review , 65* (6), 386-408.
3. Wikipedia, c. (2019, 9 18). *Long short-term memory.* Retrieved 9 27, 2019, from Wikipedia: https://en.wikipedia.org/w/index.php?title=Long_short-term_memory&oldid=916374525

Selected Properties of ANNs

Fixed

1. Number of hidden layers
2. Number of nodes per hidden layer | trainable parameters
3. Number of epochs

Variable

1. (Set of) Activation Functions
2. (Type of) Training Algorithm

These were the initially desired settings for each network.

Hypothesis

When Selected Properties of ANNs are held constant, an LSTM will outperform an MLP in the areas of

- Loss Minimization and
- Forecasting Accuracy.

Details of Compared Networks

Plan:

3 hidden layers, 64 nodes, 1000 epochs

- MLP x (Sigmoid, ReLU) x (Adam, SGD)
- LTSM x (Adam, SGD)

Actual:

3 hidden layers, 64 nodes, 1000 epochs

- MLP x (Sigmoid, ReLU) x (Adam, MBGD*)

3 hidden layers, 20 nodes, epochs $\in [100,1000]$

- LSTM x $\mathbb{I}_{shuffled}$ x batch $\in [8,16,32]$ x (Adam, GD†)

Results From Compared Networks

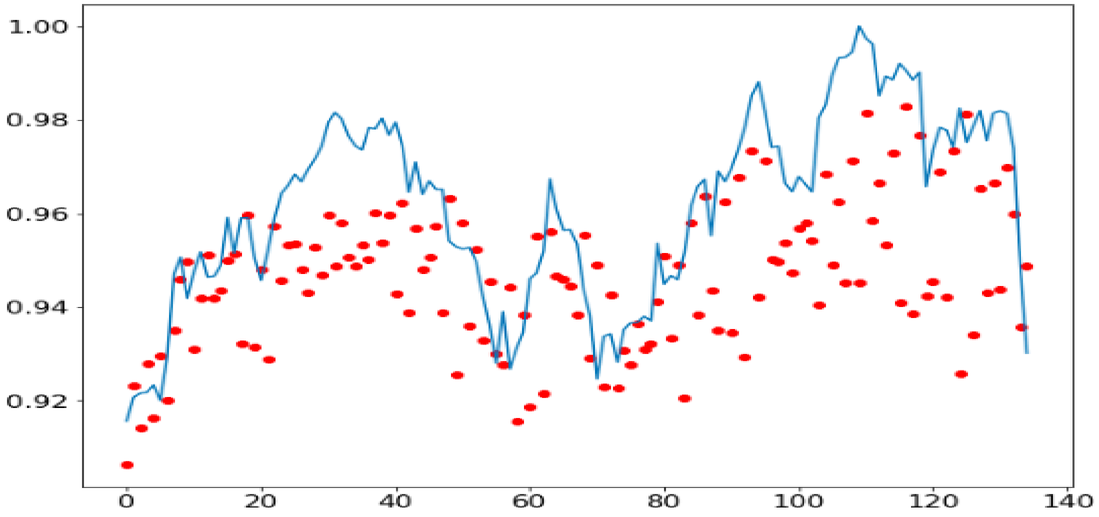
Note the MAPE and its Rank. 3 of 4 MLPs beat all LSTMs

ID	Optimizer	Activation	Shuffled	Batch	Epochs	MSE	MSE Rank	MAPE	MAPE Rank
01	Adam	ReLU	True	32	1000	153.9	2nd	0.555	2nd
02	MBGD	ReLU	True	32	1000	259.2	5th	0.610	3rd
03	Adam	Sigmoid	True	32	1000	206.8	3rd	0.553	1st
04	MBGD	Sigmoid	True	32	1000	777.0	8th	3.416	8th
04c	MBGD	Sigmoid	True	1	1000	306.1	6th	0.649	4th
1i1	Adam	LSTM	False	32	1000	241.2	4th	1.271	6th
2i1	MBGB	LSTM	False	32	1000	1232.5	9th	4.896	9th
6i2	Adam	LSTM	False	16	250	519.4	7th	0.695	5th
8i0	Adam	LSTM	False	8	1000	121.5	1st	1.806	7th

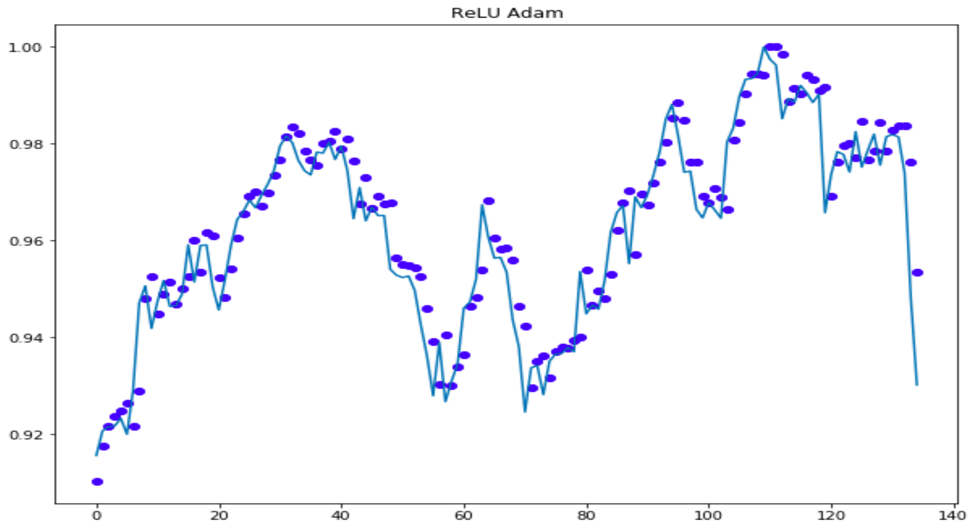
Poznań University of Economics and Business

Selected Properties of Artificial Neural Networks
in Application to Prediction of an Equity Price Index

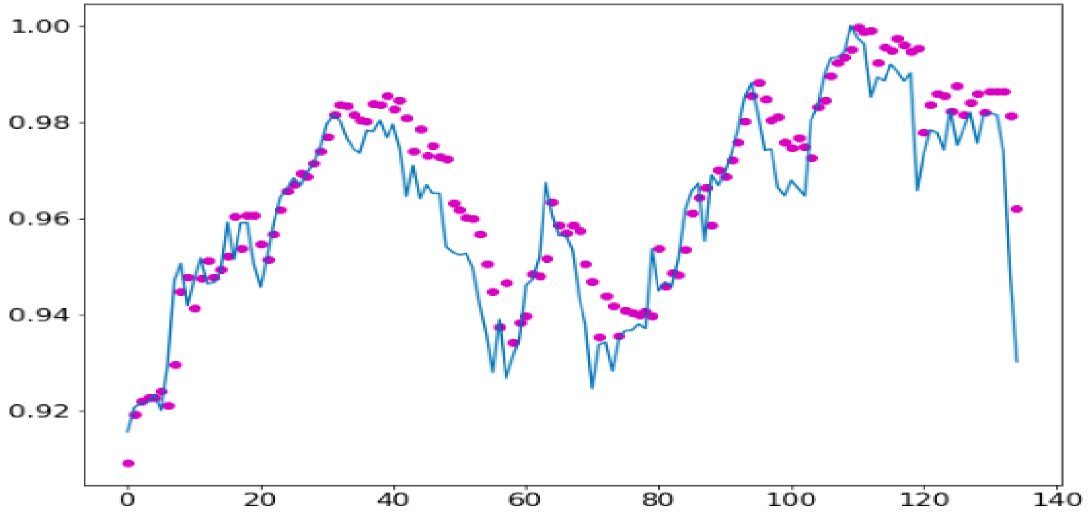
Network Comparisons



←LSTM

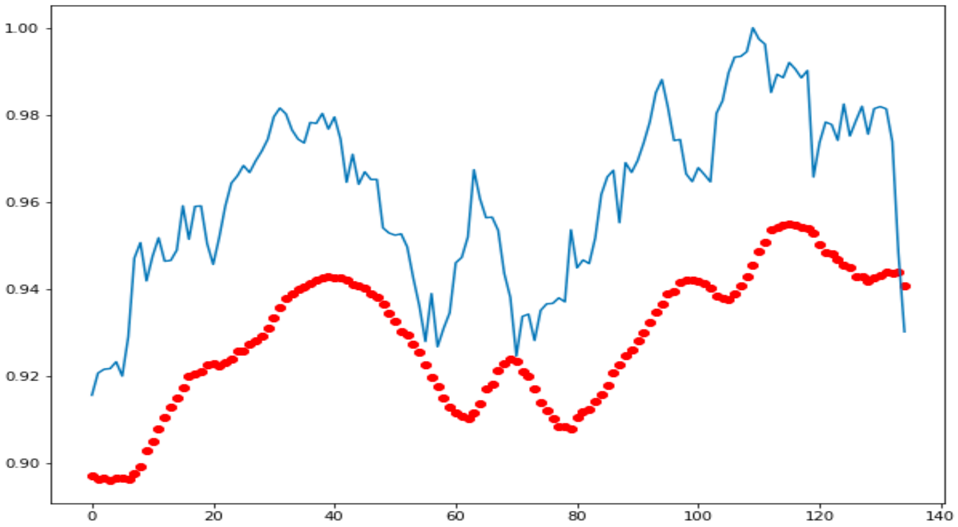


MLP→



← LSTM

MLP→



Conclusion

The MLP was simpler to build, and tune, and produced better results across various factors.

The LSTM seems to have over-trained quickly, and was difficult to retrain and tune.

The Adam optimizer, using 2nd order approximations vastly outperformed various forms of Gradient descent, especially in the LSTM case.

For our problem, ReLU seemed superior to Sigmoid, and worked well with both optimizers.

Thank You