Stock Market prediction using Echo State Networks
Practical Training

Andrei Nicolae, group 407

July 10, 2022

Abstract — This report focuses on predicting the Close price of 5 NYSE stocks. The Bidirectional Echo State Networks (ESN) will receive as input a multivariate time series. The bidirectional property of the reservoir will improve the memorization capability. Results presented in several papers show that ESNs perform almost as better as LSTMs or attention mechanisms when learning stock market data. In this study, we will use an enhanced model of the ESN model and compare the results with the classical unidirectional ESN model.

Introduction — Reservoir computing (RC) is an established paradigm for modeling nonlinear temporal sequences [1]. In machine learning tasks, echo state networks (ESNs) are the most common RC models, wherein the input sequence is projected to a high-dimensional space through the use of a (fixed) nonlinear recurrent reservoir [1]. Learning is performed by applying simple linear techniques in the high-dimensional reservoir space. The lack of flexibility in the recurrent part is balanced by a range of advantages, including faster training compared to other recurrent neural networks (RNNs) [1]. In tasks requiring a limited amount of temporal memory, ESNs achieve state-of-the-art results in many real-world scenarios constrained by time budgets, low-power hardware and limited data [1]. On the other hand, fully-trained RNNs trade architectural and training complexity with more accurate representations and a larger memory capability [1].

Methodology — The basic architecture of ESN with K input units, N neurons in the dynamic reservoir, and L output units are shown in Figure 1. The solid arrows in Figure 1 represent the synaptic connections whose weights are randomly generated and fixed, while the dashed arrows are the trainable output weights [3].

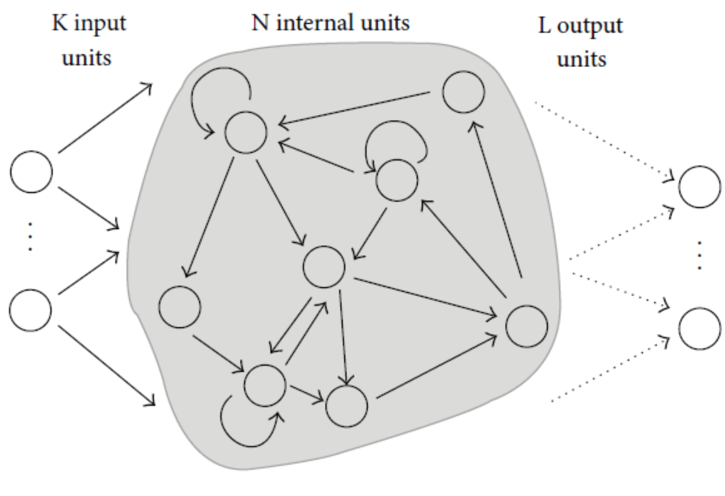

Figure 1: *Architecture of an unidirectional ESN.*

Bidirectional architectures have been successfully applied in RNNs to extract temporal features from the time series that accounts also for dependencies very far in time [2]. The reservoir is governed by the following equation:

$h_t = f(W^h h_{t-1} + W^i x_t)$ (1),

where h_t is the internal time-dependent state, which combines the current input x_t and the previous computation h_{t-1}. The function f is a non linear activation(usually a tanh), W^h is a sparse matrix that defines the recurrent self-connections in the reservoir, and W^i defines input connections. Both matrices are randomly generated and elft untrained, and the reservoir behavior is controlled mainly by three hyperparameters. These are the state size N, the spectral radius ρ of W^h, and scaling of the inputs ω. Through an optimal tuning of these hyperparameters, the reservoir produces rich dynamic and its internal states can be used to solve many prediction and regression tasks [1]. The last state h_t generated by the reservoir, after the whole input input x is processed, is a high-level representation of fixed size that embeds the temporal dependencies of x [1]. Since the reservoir trades its internal stability with a vanishing memory of past inputs [4], at time T the state maintains scarce information about the first inputs. To aleviate this issue, we feed to the same reservoir also the multivariate time series in reverse order, $x' = \{x_{T-t}\}_{t=0}^{T}$ and we generate a new representation h'_T that is more influenced by the first inputs [1]. A final representation is obtained by concatenating the two states $\bar{h}_T = [h_T; h'_T]^T$ [1].

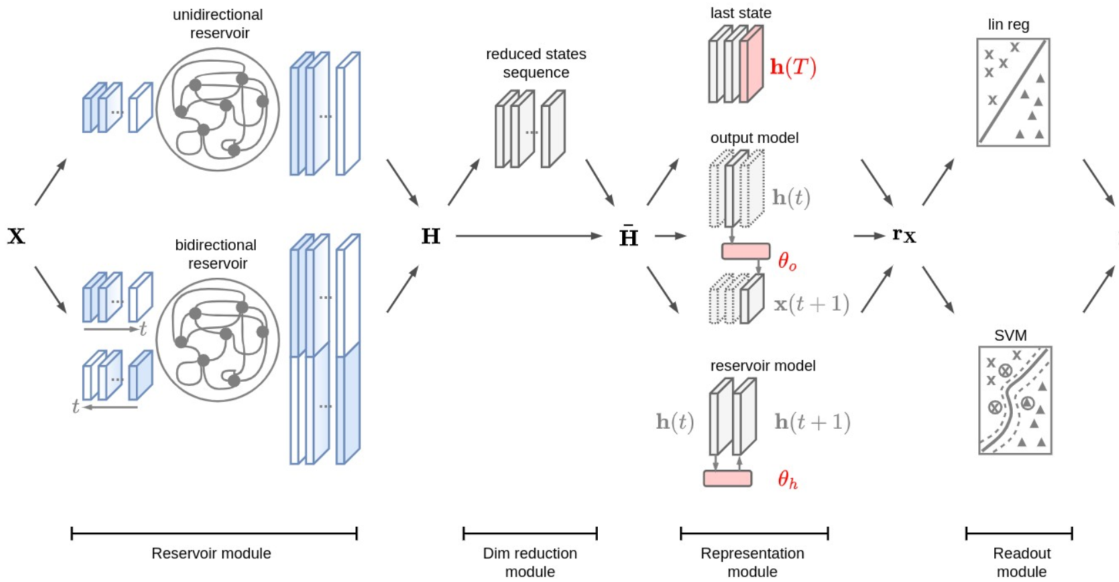

Figure 2: *Architecture of a bidirectional ESN.*

Experimental Results — As described in the *Abstract* section, the multivariate time series used as input consisted in 5 different indicators: Open Price, High Price, Low Price, Adj Close Price and Volume. The 5 NYSE STOCK on wich the ESNs are used for a 5-day Close Price prediction are: 'GS', 'C', 'WFC', 'BAC', 'JPM'. The results for every stock are presented in Figure 3.

Stock symbol	Unidirectional ESN MSE [-]	Bidirectional ESN MSE [-]	Unidirectional ESN MAPE [%]	Bidirectional ESN MAPE [%]
GS	6.42	15.08	1.73	2.84
C	2.61	2.94	3.46	4.55
WFC	1.68	2.64	3.26	3.51
BAC	1.34	0.77	3.06	1.98
JPM	4.61	3.19	3.21	2.02

Figure 3: *MSE and MAPE for the 5-Day Forecast*

In this paragraph, the key hyperparameters of the ESN will be explained together with the optimal values obtained after performing a scikit-style grid search.The number of processing units in the reservoir is 400, while the amount of leakage in the reservoir state update is equal to 0.1, which translates to 10% leakage. The largest eigenvalue (spectral radius) of the reservoir matrix connection of weights is 0.89. In order for the echo state property guarantee, the spectral must be less or equal to 1.0. The deviation of the Gaussian noise injected in the state update is equal with 0.01. At last, the number of iterations during the optimization (epochs) is 500. Additionally, because of the random weights of the reservoir, the ESN yields good results on time series which seem chaotic, like the stock market.

Conclusion— As seen in Figure 3, the difference in errors, between bidirectional and unidirectional ESN, is not significant. For our dataset, the bidirectional connections in the reservoir did not significantly improve the memorization capability. Because the reservoir doesn't contain learnable weights, the Echo State Network's computation time is relatively short. This property is important when dealing with a multivariate time series, consisting of a large number of lags.

In conclusion, the results presented in this report confirm the existence of an alternative model to LSTMs, Attention Models or Autoregressive models, when dealing with stock market data.

References

[1] Filippo M. Bianchi, Simone Scardapane, Sigurd Løkse and Robert Jenssen. Bidirectional deep-readout echo state networks, 2018.

[2] A. Graves and J. Schmidhuber. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Networks, 2005

[3] Jingpei Dan, Wenbo Guo, Weiren Shi, Bin Fang and Tingping Zhang. Deterministic Echo State Networks Based Stock Price Forecasting, 2014.

[4] A. Rodan, A. F. Sheta, and H. Faris. Bidirectional reservoir networks trained using SVM+ privileged information for manufacturing process modeling. Soft Computing, 2017.