COMPUTATIONAL MECHANICS 2023
38th conference with international participation
Srní
October 23 - 25, 2023

Feature selection for high-order neural unit based identification

V. Pawlik[a], P. Steinbauer[a], I. Bukovský[a]

[a] Faculty of Mechanical Engineering, CTU in Prague, Technická 4, 160 00 Praha 6, Czech Republic

1. Introduction

The main challenge of using learning systems is the selection of input data. In the field of system identification, the input vector usually consists of a combination of the plant's inputs and outputs. With high-order neural units (HONUs) as the model, the input vector can also contain said quantities in higher powers and combinations of products, however, not all contribute to the model's output, and selecting the features to include by hand is inefficient. This paper shows the use of the Boruta package for feature selection in the task of dynamic system identification with HONUs as a model.

2. High-order neural units

The structure of HONU is as follows

$$y_m = \sum_{i=0}^{n} w_i x_i = \mathbf{w}^T \mathbf{x}, \tag{1}$$

where y_m is the model's output, $\mathbf{x}$ is the input vector and $\mathbf{w}$ is the neural weights vector [1].

The general HONU output formula is

$$y_m = \mathbf{w}^T \operatorname{col}\mathbf{x}, \tag{2}$$

where $\mathbf{w}$ and $\operatorname{col}\mathbf{x}$ are both column vectors. The neural weights $\mathbf{w}$ can be adapted using a gradient descent algorithm and optimizing the criterion

$$J = \frac{1}{2} e(k)^2, \tag{3}$$

where the error e is the difference between the desired and the current output

$$e(k) = y(k) - y_m(k). \tag{4}$$

The gradient of the criterion J with respect to $\mathbf{w}$, which is the steepest direction, is

$$\frac{\partial J}{\partial \mathbf{w}} = e(k)\left(\frac{\partial J}{\partial \mathbf{w}} y(k) - \frac{\partial J}{\partial \mathbf{w}} y_m(k)\right) = e(k)\left(0 - \operatorname{col}\mathbf{x}\right)) = -e(k)\operatorname{col}\mathbf{x}. \tag{5}$$

The weights are then adjusted toward the minimum of the criterion using the algorithm for stochastic optimization Adam [3].

To assess the performance of HONU, the sum of squared errors (SSE) over a certain horizon N_e might be used

$$\mathrm{SSE} = \sum_{i=k-N_e}^{k} e(i)^2. \tag{6}$$

3. Boruta

An approach of feature selection using the Boruta package [4] was explored in [5] for data with many input variables. During the process, the original features are shuffled randomly and a new dataset of shadow features is created. The new dataset is then used for training a Random Forest Regressor [2] and the feature importance is checked. The original features with higher importance than the most important shadow feature are marked. This is done in multiple iterations and the results are combined and the original features are ranked using statistical metrics.

4. Plant and simulation

The plant used for testing in simulation is a non-linear oscillating plant with time-variant parameters

$$\ddot{z} + b|\dot{z}|\dot{z} + cz = u. \tag{7}$$

The parameters of the plant start changing randomly at halftime, the input is a square wave with varying amplitude. The plant is simulated continuously with Runge-Kutta method integration and then sampled with zero-order-hold at 100 Hz, resulting in a history of $z(k\Delta t) = z_k$, $\Delta t = 0.01$ s, $k = 0,\ 1,\ \ldots$ An example of the plant's input and parameters is shown in Fig. 1.

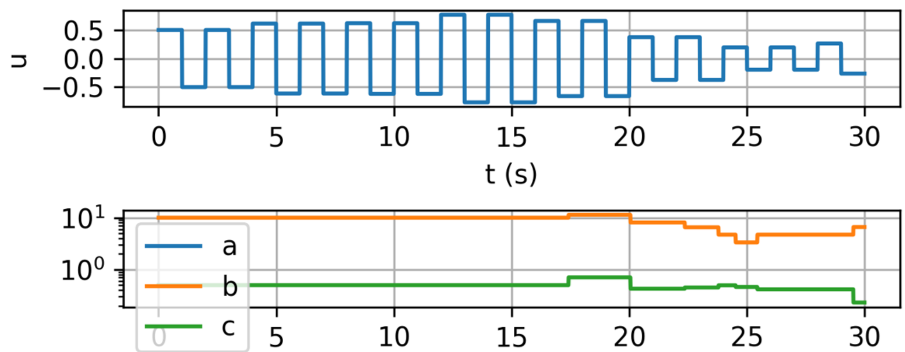

Fig. 1. The input and output of the testing plant

A dataset of plant outputs z_k and HONU inputs $\operatorname{col}\mathbf{x}_k$ from the simulation history was assembled and used for the feature selection. For ease of presentation, a quadratic cubic unit with $n_y = 2$ and $n_u = 2$ was selected. In Fig. 2, the inputs, outputs, and parameters of the plant and HONU process are shown. In Table 1, the results of the Boruta method are shown. The features with rank 1 are to be kept and others are to be discarded.

Table 1. The result of the Boruta algorithm. The features with the lowest rank are the most important

Feature	Rank	Feature	Rank	Feature	Rank	Feature	Rank	Feature	Rank
u0u0y0	1	u1u1y0	1	y0y1y1	1	u1y0	1	u1u1u1	5
y0y0y1	1	u0y1y1	1	y1y1y1	1	u0y1	1	u0u0u1	6
y0y0y0	1	u0y0y1	1	y1y1	1	y1	1	u1	7
u1y1y1	1	u0y0y0	1	y0y1	1	u0y0	1	u0	8
u1y0y1	1	u0u1y1	1	y0	1	u0u0	2	u0u0u0	9
u1y0y0	1	u0u1y0	1	y0y0	1	u0u1	3	u0u1u1	10
u1u1y1	1	u0u0y1	1	u1y1	1	u1u1	4	1	11

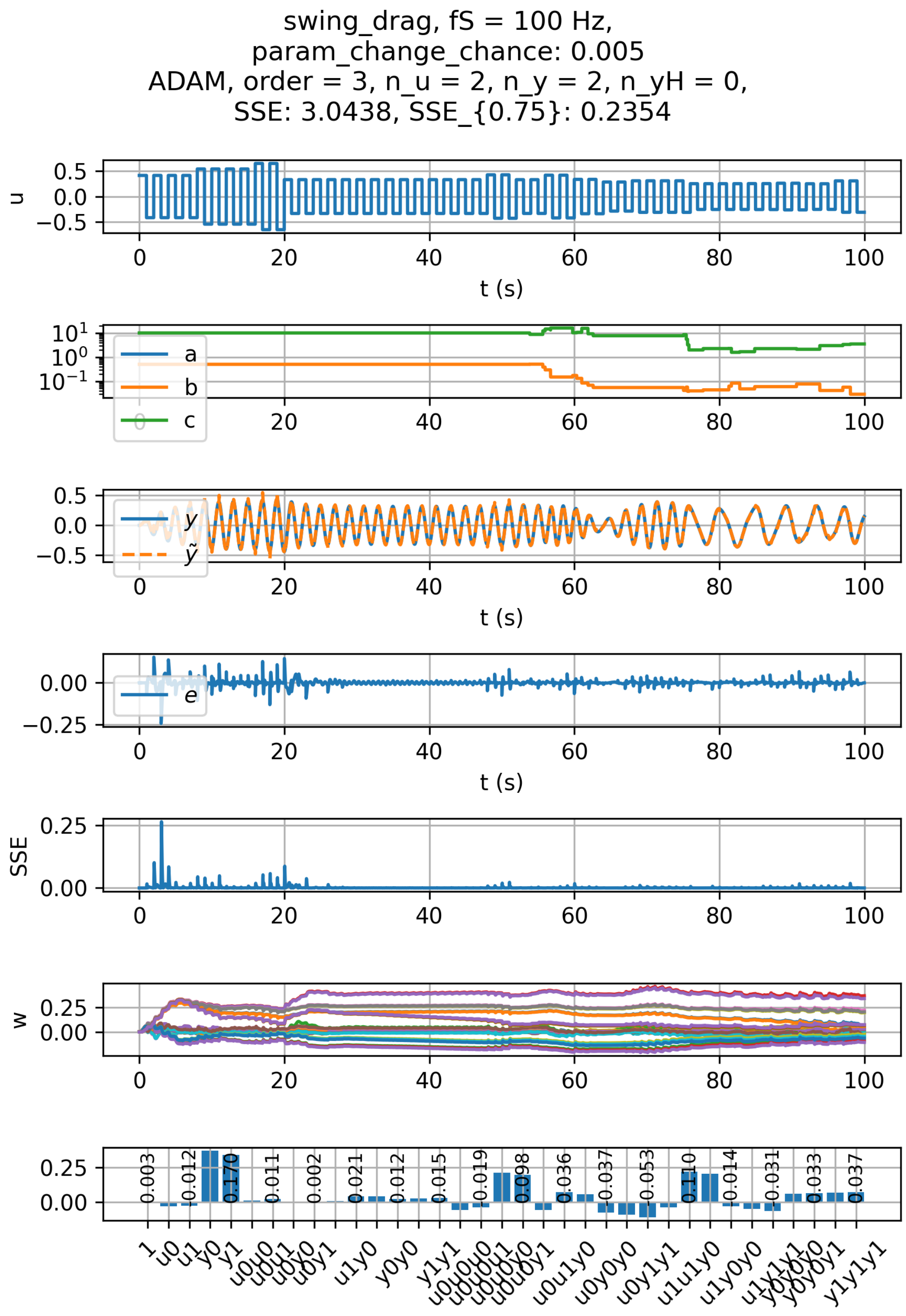

Fig. 2. Training of a HONU. The plant input u, plant's parameters b and c, plant's and HONU's outputs y and $\tilde{y}$ and the error e and criterion SSE are shown. In the last plot two plots, the history of neural weights w and the final weights of the HONU are shown

5. Conclusion

The Boruta algorithm was able to select features to be disregarded. This selection is confirmed by comparing the neural weight values after the training process with the rank from Boruta. The features with high rank have low weight and do not contribute to the output of the HONU.

Therefore, Boruta can be used to a priory disregard features and speed up the computationally intensive learning process of HONUs with higher orders.

Further research shall focus on evaluating the speed-up by feature elimination and feature selection for HONUs used in a control process.

Acknowledgement

The work has been supported by the project SGS22/150/OHK2/3T/12 “Mechatronics and adaptronics 2022” of Czech Technical University in Prague.

References

[1] Bukovsky, I., Voracek, J., Ichiji, K., Homma, N., Higher order neural units for efficient adaptive control of weakly nonlinear systems, Proceedings of the 9th International Joint Conference on Computational Intelligence, Funchal, Madeira, Portugal, 2017, pp. 149-157.

[2] Breiman, L., Random forests, Machine Learning 45 (2001) 5-32.

[3] Kingma, D. P., Ba, J., Adam: A method for stochastic optimization, 2014. doi: 10.48550/arXiv.1412.6980

[4] Kursa, M. B., Rudnicki, W. R., Feature selection with the Boruta package, Journal of Statistical Software 36 (11) (2010) 1-13.

[5] Marzova, K., Bukovsky, I., Feature selection and uncertainty analysis for bubbling fluidized bed oxy-fuel combustion data, Processes 9 (10) (2021) No. 1757.