A STUDY ON DEEP AUTOENCODER BASED FAULT DETECTION IN TENNESSEE EASTMAN PROCESS

Zhongying Xiao[1*], Arthur Kordon[2] and Subrata Sen[1]
1. Georgia Pacific LLC, Atlanta, GA 30303
2. Kordon Consulting, Fort Lauderdale, Florida 33308

Abstract Overview

Data-driven modeling has been considered as an attractive approach for fault detection in chemical processes. Of special interest to industry are methods that represent nonlinear phenomena and detect complex faults. The growing area of deep learning offers new opportunities to fill this gap. In this paper, a semi-supervised deep learning method - deep autoencoder (DAE) for fault detection in Tennessee Eastman Process (TEP) is proposed. The TEP process is a simulated benchmark for evaluating process control and monitoring method. The performance of the proposed method is evaluated and compared with Principal Component Analysis (PCA). The experimental results demonstrate that the proposed optimized five-layers DAE model for fault detection outperforms the standard PCA. Of special importance to real-world applications is its capability for automatic variable selection. In comparison to PCA it demonstrated higher prediction accuracy for most of the generated faults. Deep autoencoder has the potential to become an excellent approach for process monitoring and fault detection in chemical processes.

Keywords

Deep Autoencoder, Deep Learning, Tennessee Eastman Process, Fault Detection, Process Monitoring, Non-linearity, Automatic Variable Selection.

Introduction

Data-driven process monitoring, and fault detection is becoming one of the most active field in chemical process control(Chiang, L. , Russell, E. , Braatz, 2002). Among them, multivariate statistical methods, such as principal component analysis (PCA), partial least squares (PLS), and Fisher discriminant analysis (FDA) have been extensively studied for fault detection (Yin, Ding, Haghani, Hao, & Zhang, 2012). Most of these methods, however, are limited by the assumption that fault data could be distinguished with linear transformations.

Another class of fault detection methods is based on non-linearity of features in data. For instance, Support Vector Machines (SVM) were applied to fault detection in Tennessee Eastman Process (Chiang, Kotanchek, & Kordon, 2004). They can capture nonlinear features embedded in the data and detect some challenging faults if the model structure is properly designed.

Despite of the progress of these two categories of data-driven methods, fault detection is still far from being widely used in industrial applications due to two major issues. First, these methods often require significant amount of domain expertise for variable selection and model validation. Second, the imbalance between normal and fault data makes model development process a real challenge.

Recently deep learning methods have shown significant progress in its capabilities and has been applied in the broad application areas of image and natural language processing (Goodfellow, Bengio, & Courville, 2016). The key advantage of this method is that it automatically discovers features with gradually increasing complexity. Recently, there is a growing interest in exploring deep learning for fault detection and diagnosis of chemical processes. A deep convolutional neural network (CNN) model was proposed for diagnosing the faults on the TE process(Wu & Zhao, 2018). However, this method still required tedious variable selection and models with very complex architecture, which could be a challenge in real-time process monitoring.

This paper proposes a deep learning neural network structure, called Deep Autoencoder (DAE) algorithm, to detect faults without tedious feature selection. The proposed DAE framework is trained based on time series data in normal process condition without manual variable selection. The paper demonstrates the model performance of the DAE through testing it for detecting different types of faults in Tennessee Eastman Process (TEP). To compare DAE with traditional statistical models, PCA method is used as a benchmark method.

[*] To whom all correspondence should be addressed
E-mail address: zhongying.xiao@gapac.com (Z. Xiao).

Deep Autoencoder

Autoencoder is a type of neural network which is adopted to transfer significant information of its input to its output. Historically, autoencoders have been used to denoise signals, extract features and reduce dimensionality (Goodfellow et al., 2016). As an unsupervised learning method, DAE consists of three components: an input layer, single or multiple hidden layers, and an output layer. At the middle of the architecture is a bottleneck layer where the information of data is most concentrated and represented. The parameters are optimized via backpropagation to minimize the loss function. In this paper, Adaptive Moment Estimation (Adam) gradient descent optimization algorithm is used for optimization.

Tennessee Eastman Process

TEP model is a realistic simulation program of a chemical plant which is recognized as a benchmark for process control and fault detection studies. The process is described in (Downs & Fogel, 1993) and the MATLAB code for process simulation is available over the website (https://depts.washington.edu/control/LARRY/TE/download.html).

To investigate the ability of DAE for fault detection in this chemical process, the TEP simulator was used to generate three classes of faulty data, which correspond to TEP specification as: Faults 4 (step change in reactor cooling water inlet temperature), Fault 5 (step change in condenser cooling water inlet temperature), and Fault 11 (random variation in reactor cooling water inlet temperature). For each faulty case, two sets of data were generated. The training data containing only normal operations data were used to build the models and the test data containing both normal and faulty operations data were used for model validation. Both the training and test data contain 960 observations. In test data, the first 160 observations were based on normal operation and the corresponding faults occurred after the 161^{st} observation. Each dataset contains 52 process variables.

Results and Discussion

DAE Model Architecture

To find a proper architecture, we have tuned several models with various number of layers, neurons, and different activation functions with the best performance of Parametric Rectified Linear Units (PReLU).

With mean squared error (MSE) as loss function, model performance was evaluated by changing number of layers and number of moving windows. The optimized architecture has 5 neural layers and slide window with 3 data points, resulting in 156 neurons at the input layer. This architecture generated an excellent model with very low training and test errors, shown in Figure 1. As a result, this DAE structure was selected to train and test the explored datasets.

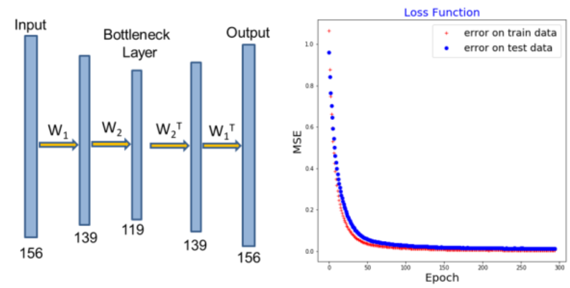

Figure 1. Optimized DAE Architecture and Model Training Process

Automatic Variable Selection

Unlike other machine learning methods, the explored deep autoencoder does not need additional variable selection step based on domain knowledge or statistical methods. Deep autoencoder was trained by normal operation scenario and its output was trying to preserve the information of the input, by minimizing the reconstruction error during model training. Variable-wise reconstruction errors at the output layer are also minimized in normal operation scenarios. In a faulty process, variables leading to or affected by faults would show huge differences compared with normal scenarios. When trained DAE model was mapped into data with faulty scenarios, these highly related variables would show large reconstruction errors relative to the other unrelated variables.

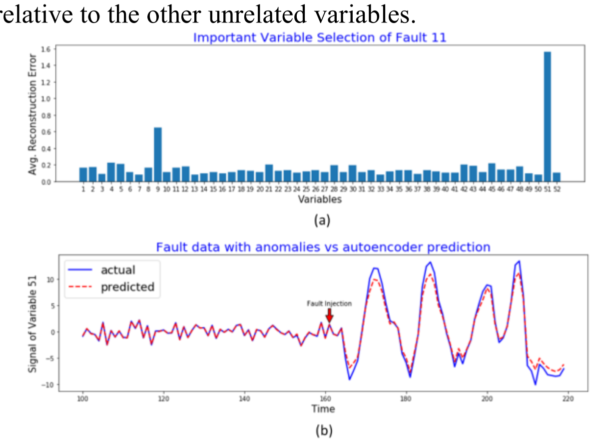

Figure 2. Automatic Variable Selection of Fault 11

An example of automatic variable selection for Fault 11 is shown in Figure 2. The reconstruction errors of all input variables for Fault 11 is plotted in Figure 2(a). Clearly, two spikes of high reconstruction errors are displayed for Variable 9 (Reactor temperature) and Variable 51 (Reactor cooling water flow), while the other variables have relatively small reconstruction errors. The signal of these two variables demonstrated significant changes after Fault 11 has occurred. Trained DAE cannot capture enough features from Variable 51 and the predicted values (red dash line) have huge differences with actual values (blue solid line) after Fault 11 was injected, resulting in a big reconstruction error for Variable 51 (Figure 2(b)). The results have been validated by published reference (Chiang et al., 2004). Automatic variable selection, based on reconstruction errors, is a major advantage of DAE compared with the other methods for fault detection. The selected important variables with DAE model give a very useful information for root-cause analysis of the faults.

Higher Prediction Accuracy

Another advantage of DAE is the non-linear relationships between predictors and outcome represented by this method. As a result, it is assumed that the DAE can detect differences between normal and fault scenarios with much higher accuracy than corresponding linear approaches. The results from a performance comparison between DAE and PCA for three selected faults (Fault 4, 5, and 11) are given in this section. The full paper will include the results for 15 TEP faults out of the total 21 faults. Based on PCA, Hotelling's T^2 and Squared Prediction Error (SPE) were calculated as benchmarks for fault detection. With the same training and test dataset, the accuracy of fault detection of the DAE with the optimized architecture is evaluated. Table 1 shows the fault detection rate (FDR) of the three different methods. Among the methods tested, DAE generated better results with much higher FDR.

Table 1. FDR of The Three Methods

Faults	T^2	SPE	DAE
Fault 4	18.1%	99.8%	99.8%
Fault 5	26.6%	31.0%	100.0%
Fault 11	33.1%	77.0%	96.6%

In Figure 3, three methods were utilized to detect Fault 5. Both Hotelling's T^2 and SPE can only detect errors at early stages and their statistic became similar with normal scenarios at later stages after sample 350-400. Their FDRs are 26.6% and 31.0%, respectively. With DAE, however, we can conduct much simpler and more accurate process monitoring for Fault 5. Due to non-linear transformation of deep neural network, DAE can preserve more detailed features from data and detect derivations from trained data with higher sensitivity when fault occurs. For Fault 5, the misclassification rate is 0.

Conclusions

An important branch of deep learning neural networks - deep autoencoder has been studied for fault detection in Tennessee Eastman Process benchmark. The performance of an optimized five-layers DAE model for fault detection of TEP-generated faults is compared with an established linear method, such as PCA. A big advantage of the proposed DAE is the automatic variable selection it provides, based on reconstruction errors. The important variables, selected by DAE algorithm, is a vital information for root-cause analysis of the faults by engineers and data analysts. Compared with linear PCA method, nonlinear transformation of features embedded in the dataset by DAE can capture more useful information when fault occurs, resulting in a higher fault detection rate. The higher rates have been demonstrated for most of the explored faults. The next step will be focusing on designing a proper DAE architecture for real-world applications.

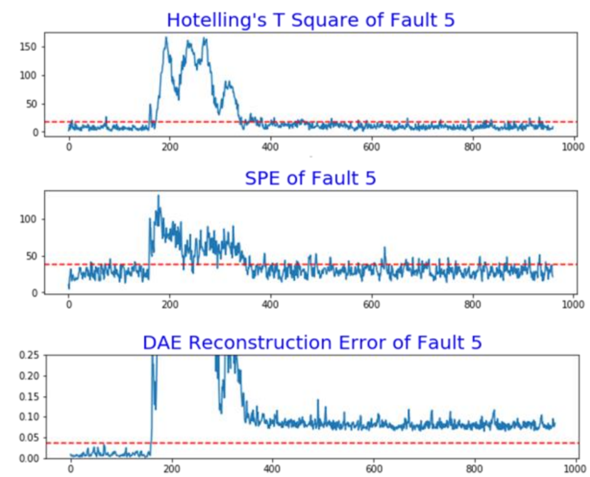

Figure 3. Process Monitoring with Hotelling's T^2, SPE and DAE in case of Fault 5

Acknowledgments

Sincere acknowledgements go to Dr. N. Lawrence Ricker in making the dataset available for research and Shakir Ali with meaningful discussion of TEP dataset. The authors gratefully acknowledge support from Georgia Pacific LLC.

References

Chiang, L. , Russell, E. , Braatz, R. (2002). Fault Detection and Diagnosis in Industrial Systems. In *Springer-Verlag*.

Chiang, L. H., Kotanchek, M. E., & Kordon, A. K. (2004). Fault diagnosis based on Fisher discriminant analysis and support vector machines. *Computers and Chemical Engineering*. https://doi.org/10.1016/j.compchemeng.2003.10.002

Downs, J. J., & Fogel, E. (1993). *A plant-wide industrial process control problem*. *17*(3), 245–255. https://doi.org/10.1016/0098-1354(93)80018-I

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Wu, H., & Zhao, J. (2018). Deep convolutional neural network model based chemical process fault diagnosis. *Computers and Chemical Engineering*, *115*, 185–197. https://doi.org/10.1016/j.compchemeng.2018.04.009

Yin, S., Ding, S. X., Haghani, A., Hao, H., & Zhang, P. (2012). A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. *Journal of Process Control*, *22*(9), 1567–1581. https://doi.org/10.1016/j.jprocont.2012.06.009