Structured Denoising Autoencoder For Fault Detection and Analysis

Hillol Datta (20162092)
Sayantan Gangopadhyay (20162126)
Mainak Bhattacharjee (20162101)

Introduction

- Fault detection and analysis problem
  - Normal data to learn a model
  - Test data are evaluated to detect faults and analysis of their cause
- Data driven methods
- Model driven methods

Problems of other methods

- *Conventional Data-driven approach* suffer from the problem of overfitting and results in high rates of false positives.
- *Model driven approach* suffer from a lack of specified information about complex systems.

Proposed solution

This paper has proposed a new fault detection and analysis approach. Structured Denoising(StrDA) is a version of denoising autoencoder(DA) which can:

- Utilize incomplete prior information.
- It does not require specific informations and does not suffer from overfitting
- It performs better than DA even there is partially incorrect or abstract information.

Contribution Analysis (Fault Detection)

Given a faulty sample of M observed variables $\mathbf{x} = \{x_1, . . . , x_M\}$

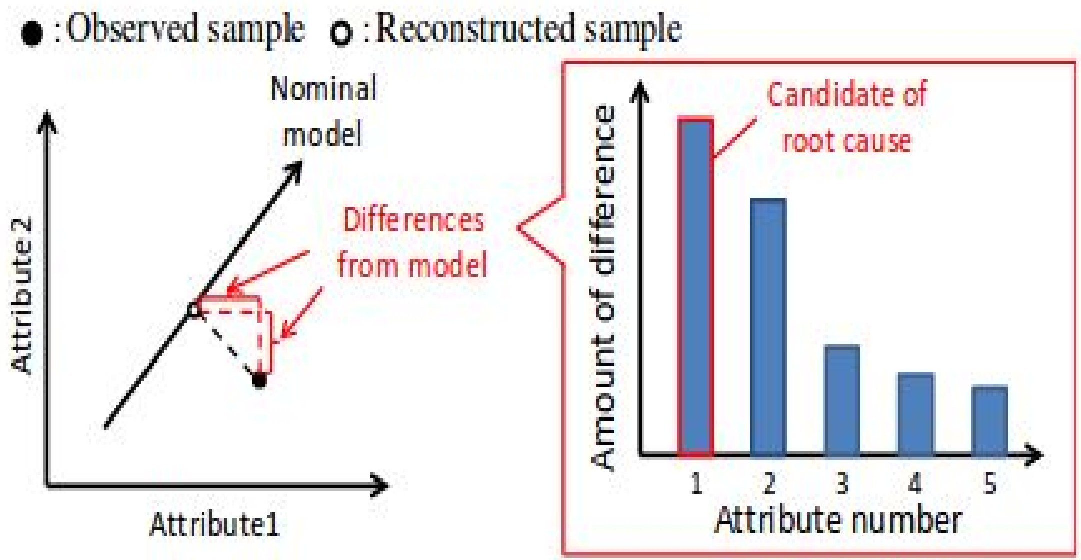
$\mathbf{x_r}$ = corresponding variables predicted by a learned normal model

CA considers the errors $e = \{e_1, . . . , e_M\}$

$\mathbf{e} = x - x_r$ [e_i represents the contribution of the variable $i \in 1...M$]

The variables are then ranked in the order of their contributions to fault and variable with higher contribution is the candidate for the cause.

Contribution Analysis



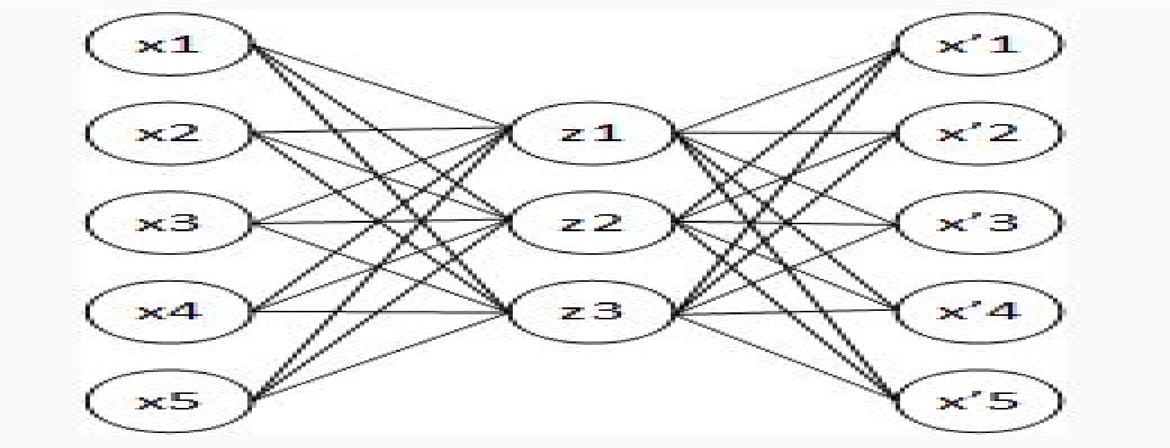
PCA

- A statistical procedure represents a higher dimensional feature vectors into a space of lower dimension.
- The data should lies near a linear manifold in a higher dimensional space.
- Uses orthogonal transformation.

Autoencoder

- It is an unsupervised learning algorithm (like PCA)
- It minimizes the same objective function as PCA
- It is a neural network
- The neural network's target output is its input

Autoencoder Architecture



So the dimensionality of the input is the same as the dimensionality of the output, and essentially what we want is x' = x

Why not PCA?

- More flexibility
- PCA can only represent linear transformations
- Autoencoders can use non linear encoding using tanh, sigmoid, etc.
- Various models like:
  - Denoising Autoencoder,etc

Denoising Autoencoder

- Adding noise to the input (corrupted input)
- Force hidden layer to discover more robust features
- Reconstruction of input from a corrupted version

x_n = corrupted input

$$
\begin{aligned}
J(\boldsymbol{\theta}) &= \sum_{p=1}^{N} \frac{1}{2} \|\mathbf{x}^{(p)} - \mathbf{x}_r^{(p)}\|_2^2, \\
\mathbf{x}_r^{(p)} &= \mathbf{W}_2 \mathbf{h}_1^{(p)} + \mathbf{b}_2, \\
\mathbf{h}_1^{(p)} &= g(\mathbf{W}_1 \mathbf{x}_n^{(p)} + \mathbf{b}_1),
\end{aligned}
$$

x = original input
x_n = corrupted input
x_r = reconstructed input

Structured Denoising Autoencoder

- Utilize incomplete prior information/knowledge about structure of the given data and/or of possible faults

$$\boldsymbol{\alpha} = \{\alpha_{ij} \in \{0,1\} | i,j \in 1,\ldots,M\}.$$

- Set α_{ij} to 1 to focus on the relation between ith and jth variable

Structured Denoising Autoencoder(Continued..)

Prior knowledge of important relations

- i/o structure of the variable
- Wish to focus on certain variables
- Suspect certain variables based on sensory analysis

StrDA objective function

We extend the DA to include the prior information α by modifying the objective function

$$\tilde{J}(\boldsymbol{\theta}) = \frac{1}{2}\sum_{p=1}^{N}\sum_{i<j}\left\{\left(e_{ij}^{(p)}\right)^2 + \left(e_{ji}^{(p)}\right)^2\right\},$$

$$e_{ij}^{(p)} = x_{r,i}^{(p)} - \left\{x_i^{(p)} + (1-\alpha_{ij})(x_{n,i}^{(p)} - x_i^{(p)})\right\},$$

$x_{n,i}$=ith variable of the corrupted input

$x_{r,i}$=ith variable of the reconstructed output

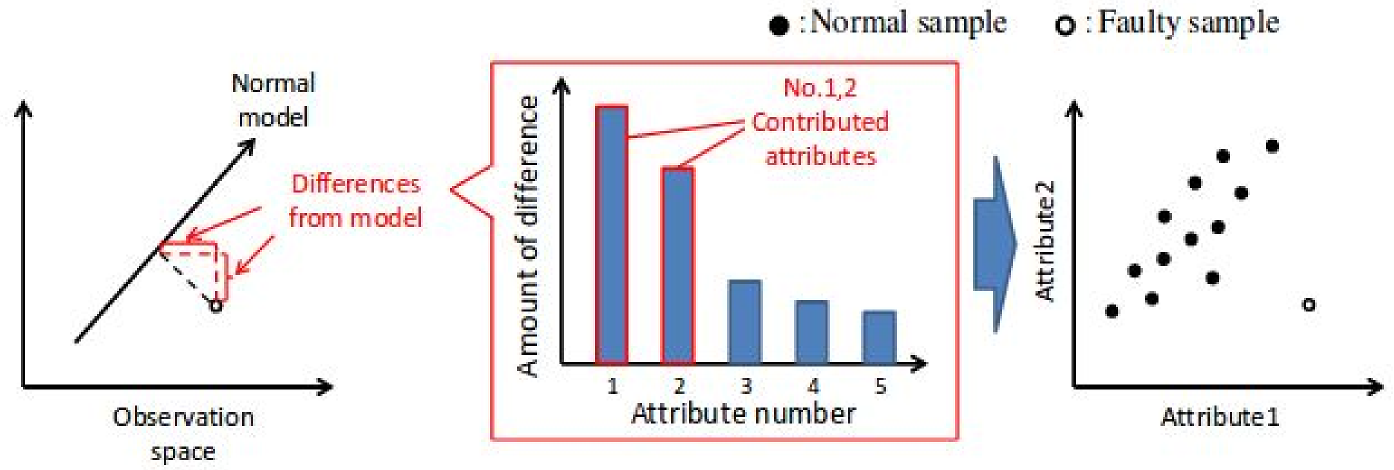
Fault detection and Analysis using StrDA

- Same equation as used in DA to compute x_r

$$\mathbf{x}_r^{(p)} = \mathbf{W}_2\mathbf{h}_1^{(p)} + \mathbf{b}_2,$$
$$\mathbf{h}_1^{(p)} = g(\mathbf{W}_1\mathbf{x}_n^{(p)} + \mathbf{b}_1),$$

- Reconstruction error $e=x-x_r$ to conduct CA and calculate the rms as an anomaly score
- Main Advantage(StrDA over DA)-StrDA reflects prior knowledge in its reconstruction errors,but the DA does not.

FAULT DETECTION ANALYSIS(CONTD.)



Conclusion

- The StrDA is an extension of the conventional DA
- It utilizes prior information about the data structure.
- Experiments using synthetic data suggest that using the prior information can improve the detection of faults.
- Fault analysis, the StrDA extracted and represented the true causes of the changes
- Effective for detecting and analyzing faults