Orbit Structure Segmentation

Sanghyuck Lee

2021. 03. 16.

Outline

- Feature Selection for Medical Image Deep learning
  - Paper Study: DeepPINK: reproducible feature selection in deep neural networks
- 7 orbit structures segmentation (CAU MED)
  - Completed Segmentation Images

Paper Study

- Last week:

  Feature selection for medical image

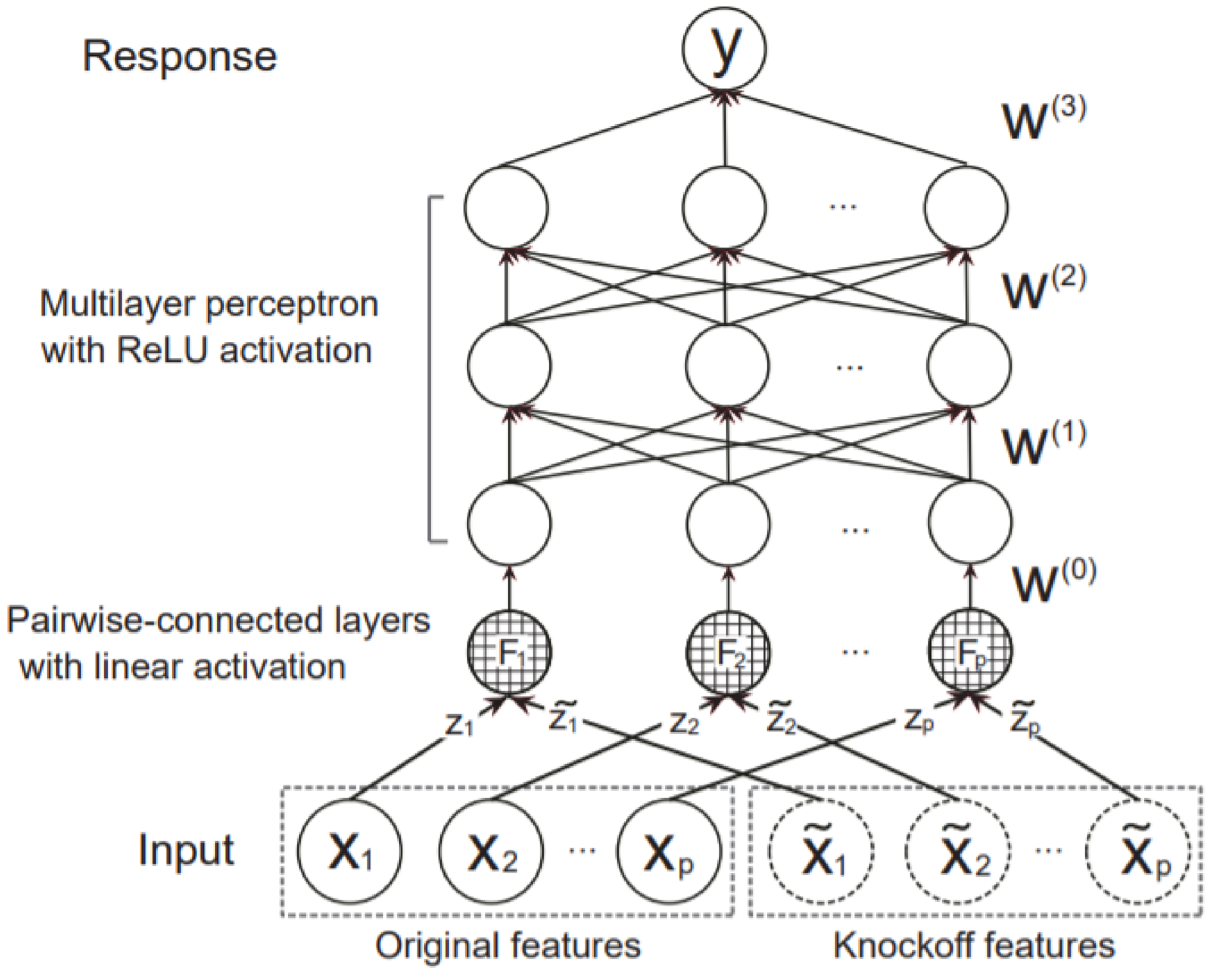
  => Deep learning approach

- DeepPINK: reproducible feature selection in deep neural networks

Paper Study

- DeepPINK: reproducible feature selection in deep neural networks (Deep feature selection using Paired-Input Nonlinear Knockoffs).

1) Where the Knockoff features come from
2) How to actually select the features
3) How the pairwise-connected layers are used



Paper Study

- Ref: Lu et al. (2018)(DeepPINK) , Barber et al. (2015), Candes et al. (2016)

The False Discovery Rate (FDR) in Variable Selection:

$$FDR = E[\frac{|\hat{S} \cap S_0^c|}{|\hat{S}|}]$$

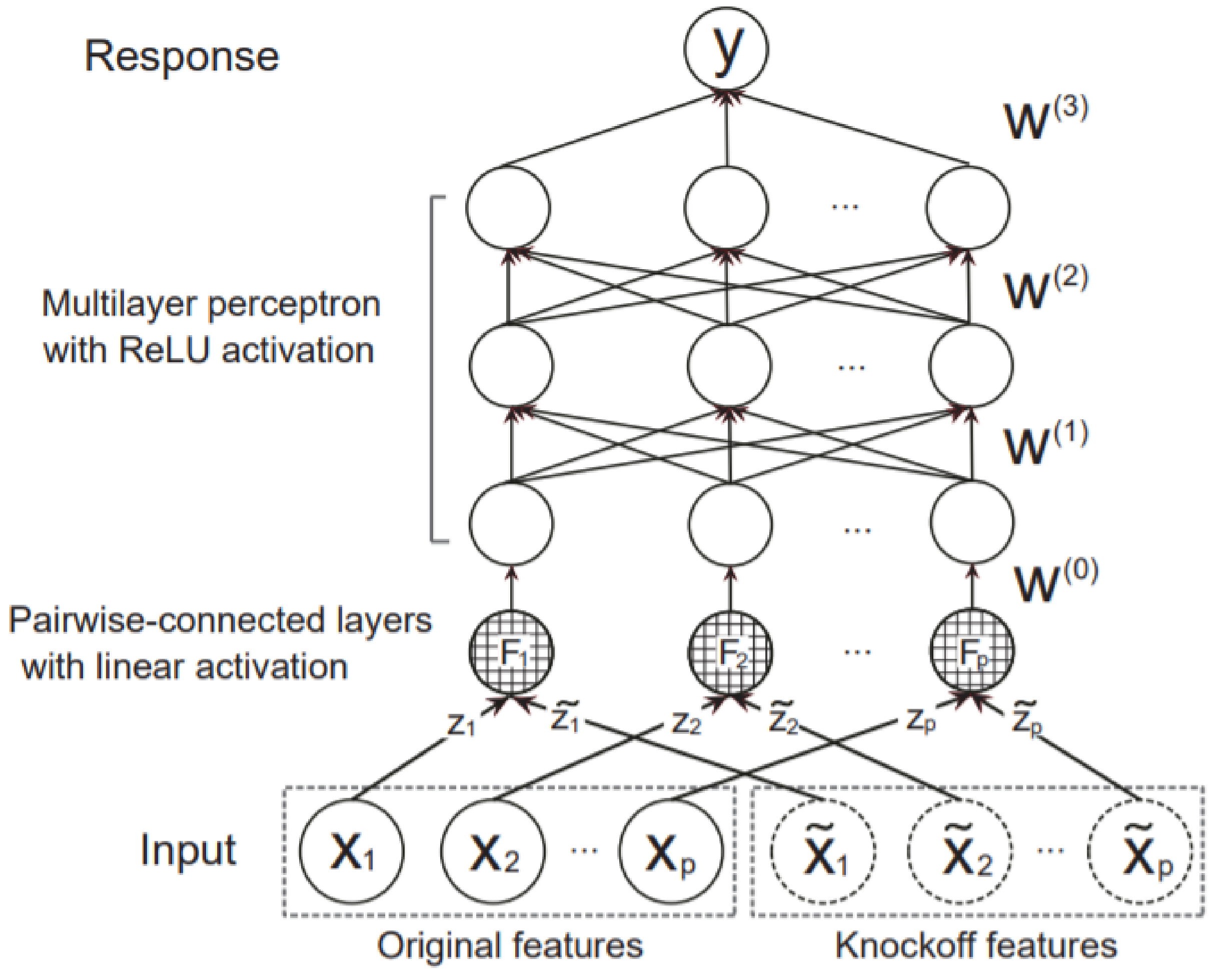
"Assume that there exists a subset $S_0 \subset \{1, \cdots, p\}$ such that, conditional on features in S_0, the response Y_i is independent of features in the complement S_0^C."

Paper Study

- 2 steps for FDR-control,

  1) construction of knockoff features

  2) filtering using knockoff statistics



Paper Study

1) construction of knockoff features

Original Features: $x = (\mathrm{X}_1, \ldots, \mathrm{X}_p)^{\mathrm{T}}$

Knockoff Features: $\tilde{x} = (\tilde{\mathrm{X}}_1, \ldots, \tilde{\mathrm{X}}_p)^{\mathrm{T}}$

(1) $(x, \tilde{x})_{\mathrm{swap(S)}} \overset{d}{=} (x, \tilde{x})$ for any subset $S \subset \{1, \ldots, p\}$

(2) $\tilde{x} \perp\!\!\!\perp Y | x$, i.e. $\tilde{x}$ is independent of response Y given feature x

Paper Study

(1) $(x, \tilde{x})_{\mathrm{swap}(S)} \stackrel{d}{=} (x, \tilde{x})$ for any subset $S \subset \{1, \dots, p\}$

for example, in Candes et al. (2016):

$p = 3$ and S = {2, 3}. but, "$x = (x_1, \dots, x_p)$, $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_p)$" in this case.

$$(x_1, x_2, x_3, \tilde{x}_1, \tilde{x}_2, \tilde{x}_3)_{\mathrm{swap}(S)} \stackrel{d}{=} (x_1, \tilde{x}_2, \tilde{x}_3, \tilde{x}_1, x_2, x_3)$$

Paper Study

1) construction of knockoff features

In some cases such as $x \sim N(0, \Sigma)$, where Σ is the covariance matrix

$$\tilde{x} | x \sim N(x - \mathrm{diag}\{s\} \Sigma^{-1} x,\ 2\mathrm{diag}\{s\} - \mathrm{diag}\{s\} \Sigma^{-1} \mathrm{diag}\{s\}) \qquad (1)$$

"Here $\mathrm{diag}\{s\}$ with all components of s being positive is a diagonal matrix the requirement that the conditional covariance matrix in Equation 1 is positive definite", more study in Barber et al. (2015).

$$(x, \tilde{x}) \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma & \Sigma - \mathrm{diag}\{s\} \\ \Sigma - \mathrm{diag}\{s\} & \Sigma \end{pmatrix}\right) \qquad (2)$$

- Larger s, more power

Paper Study

2) filtering using knockoff statistics

knockoff statistic: $W_j = g_j(Z_j, \tilde{Z}_j)$

where Z is feature importance measure and g is an antisymmetric function

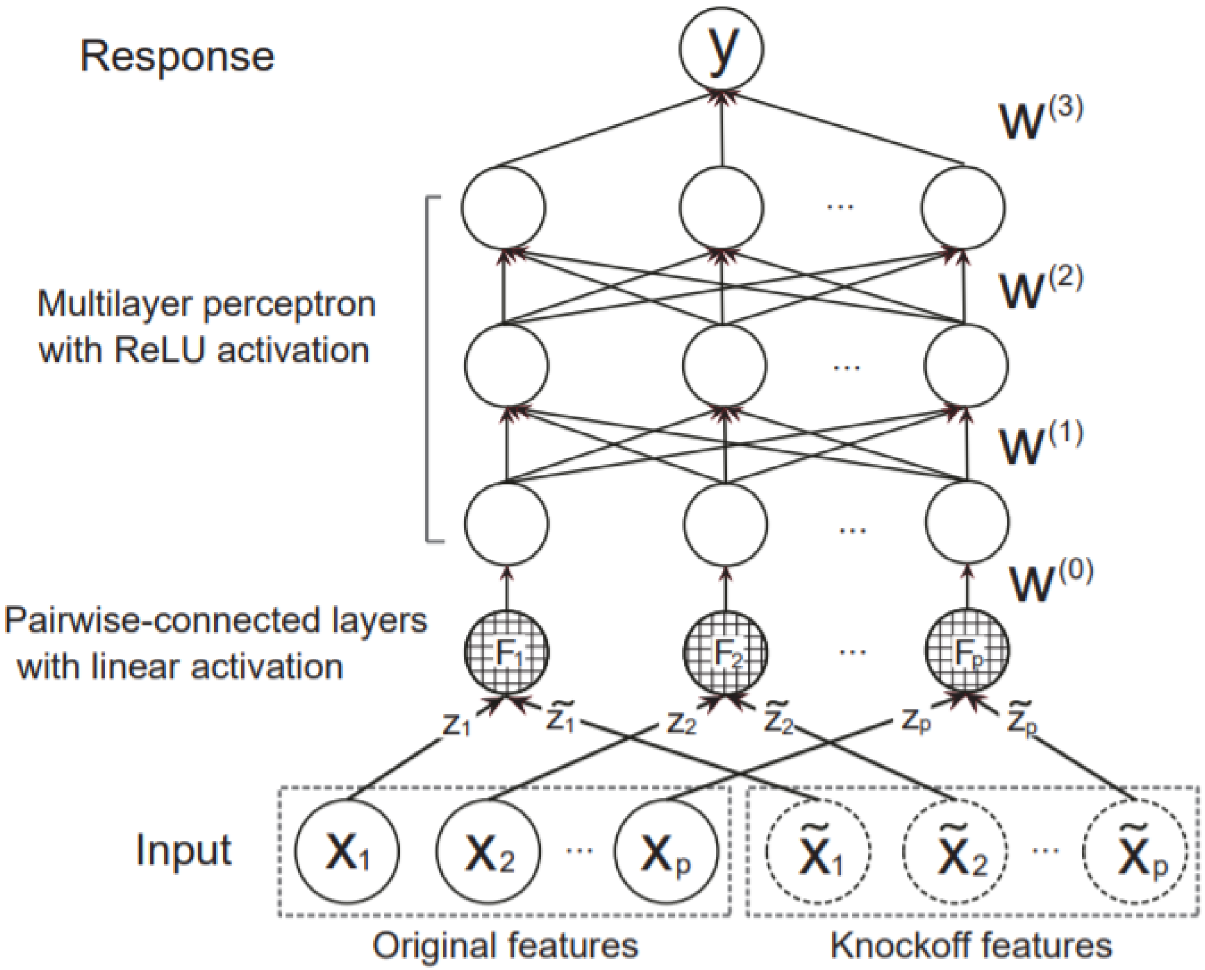
$\Rightarrow \; g_j(Z_j, \tilde{Z}_j) = -g_j(\tilde{Z}_j, Z_j)$

For example, $W_j = |Z_j| - |\tilde{Z}_j|$, Z_j and $\tilde{Z}_j$ are Lasso regression coefficients of X_j and $\tilde{X}_j$

sort W, use threshold

$$T = \min\{t \in W, \frac{|\{j: W_j \leq -t\}|}{|\{j: W_j \geq t\}|} \leq q\}, \qquad T_+ = \min\{t \in W, \frac{1 + |\{j: W_j \leq -t\}|}{1 \vee |\{j: W_j \geq t\}|} \leq q\}$$

Paper Study


Response
y
W(3)
W(2)
Multilayer perceptron
with ReLU activation
W(1)
W(0)
Pairwise-connected layers
with linear activation
F1
F2
Fp
z1
z̃1
z2
z̃2
zp
z̃p
Input
X1
X2
Xp
X̃1
X̃2
X̃p
Original features
Knockoff features

7 Orbit Structure Segmentation

- Total: 1492 patients
- Completed:

Target Type	Completed
AX 내직근	921
AX 눈	1049
AX 시신경	933
AX 외직근	870
CO 눈	899
SA 눈	922
SA 상안검	946
SA 상직근	1018
SA 시신경	994
SA 하직근	830

References

Barber, Rina Foygel, and Emmanuel J. Candès. "Controlling the false discovery rate via knockoffs." *Annals of Statistics* 43.5 (2015): 2055-2085.

Candes, Emmanuel, et al. "Panning for gold: Model-X knockoffs for high-dimensional controlled variable selection." *arXiv preprint arXiv:1610.02351* (2016).

Lu, Yang Young, et al. "DeepPINK: reproducible feature selection in deep neural networks." *arXiv preprint arXiv:1809.01185* (2018).