KNOCKOFFGAN: GENERATING KNOCKOFFS FOR FEATURE SELECTION USING GENERATIVE ADVERSARIAL NETWORKS

James Jordon, Jinsung Yoon, Mihaela van der Schaar

Overview

The paper builds on the improvement of a previously successful method called knockoff for feature selection. This paper extends the idea with the Generative Adversarial Networks and proves that their approach gives better results on the non-gaussian distribution keeping the FDR maintained below a certain threshold. The paper is smooth to read and it provides an ease to the reader to understand how the concept of GAN can improve the existing method. There were some problems regarding the lack of dataset information and the number of comparisons that were made. The idea behind choosing the model architecture was also not mentioned. Although later the author convinced most of these by adding a 13-page appendix section at the end in order to fulfill previous faults. But even after it, there are certain comparisons which are not general and that might be made less specific. Overall Paper is nice and is full of results and analysis, it motivates the reader to understand the capability of Generative adversarial networks.

Motivation

The author gives emphasis on how feature selection in a certain analysis is as important as knowing the true labels. He also gives force that how modern machine learning approaches can help in choosing better features. The authors understood the severity of the problem with previous approaches that have to be provided with information about the distribution. In general, seems to work only with gaussian distribution whereas leading to a lowering in accuracy when coming to the non-gaussian distribution.

Concept

The Model Consists of four networks

Generator

A simple fully connected neural network with some weights that takes data as input and outputs knockoff features.

Discriminator

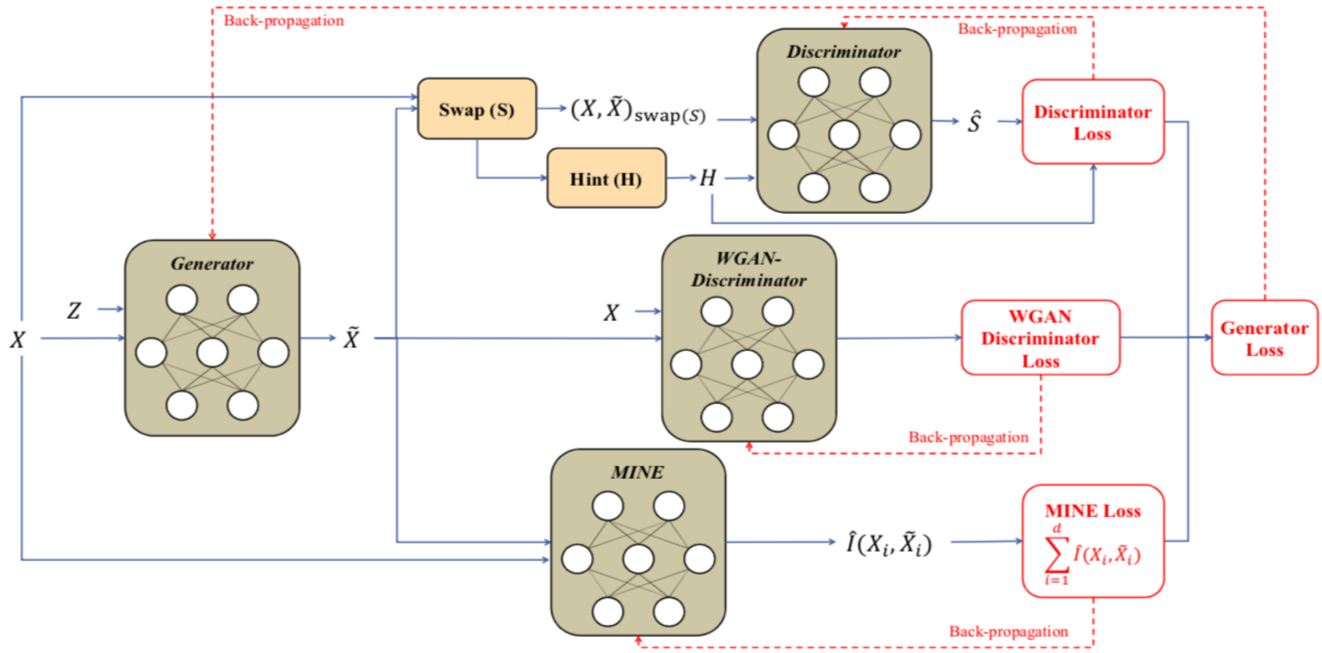
This network is the soul of the paper, It takes swapped knockoff-features as input to determine which variables are swappped where the cross-entropy loss was used to train it. For better results, a Hint variable was also taken as input which describes some information of swapping.

WGAN Discriminator

This network is used as a stabilizer to provide the regularization term to our final loss.

MINE Network

Mutual Information Neural Estimation is used to maximize the power of the knockoff selection procedure by estimating the mutual information between each knockoff pair.



The Block Diagram of KnockoffGAN

$$\boxed{\min_{G} \left(\max_{D}(\mathcal{L}_D) + \lambda \max_{P}(\mathcal{L}_P) + \mu \max_{f}(\mathcal{L}_f) \right)}$$

The final objective is to optimize the equation given in the side box

Finally, the method is tested on a synthetic dataset with both Gaussian and non-gaussian distributions to prove the capability of the network and finally also tested on a real dataset.

Strengths

Firstly I will like to praise the authors for their writing, The paper is well written, reads smoothly and the ideas are well exposed.

Most of the decisions are backed by theorems and lemmas and their proofs.

Modern ideas of using the MINE network as well as providing a separate network for regularization seem interesting.

The equations in the paper walk through how the loss function is developing by adding each network.

Through graphs, the importance of the WGAN network is also shown while comparing the TPR(True positive rate) with and without WGAN.

Although the target was to overcome the original knockoff feature selection, other standard networks are also compared to give a general comparison.

Weakness

All the information about the synthetic dataset is not given in the paper to maintain anonymity

Although the target was just to keep FDR below a certain threshold you can notice that both in gaussian and non-gaussian distributions The FDR of Knockoff-GAN is higher than the original knockoff method.

For synthetic dataset lasso is used whereas for real dataset random forest is used, no satisfactory justification is provided in the paper for this decision.

Scope of improvement

Modern convolutional networks could be used rather than using simple fully connected layers to reduce the parameters and convolutions also provide the surrounding information which can help in deciding more relevant features. The author might also experiment on an Image dataset.

As the author himself suggests, this model can also be made for time series data so as to generalize the approach

Reviewer-
Rishab Khantwal
180100095