Feature Selection Using Criterion Relaxation

Author: Lê Hoàng Sinh

University of Information Technology

What ?

We introduce a process to improve feature selection about speed and performance for linear models:

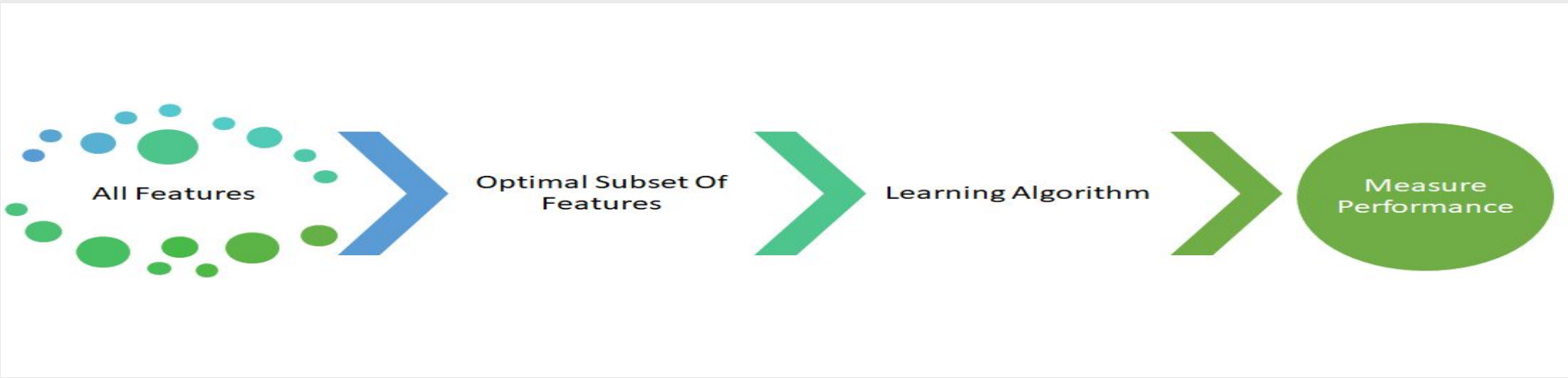
- Proposed a technique to help speeding up feature selection
- Implement the method to select features on the big dataset.
- Evaluated the feature selection method speed comparing to other methods.

Why ?

Benefit of feature selection:

- Improved model performance
- Reduced dimensionality
- Enhanced interpretability
- Faster model training and inference
- Data understanding and insights

Overview



Description

1. Feature Selection Using Criterion Relaxation

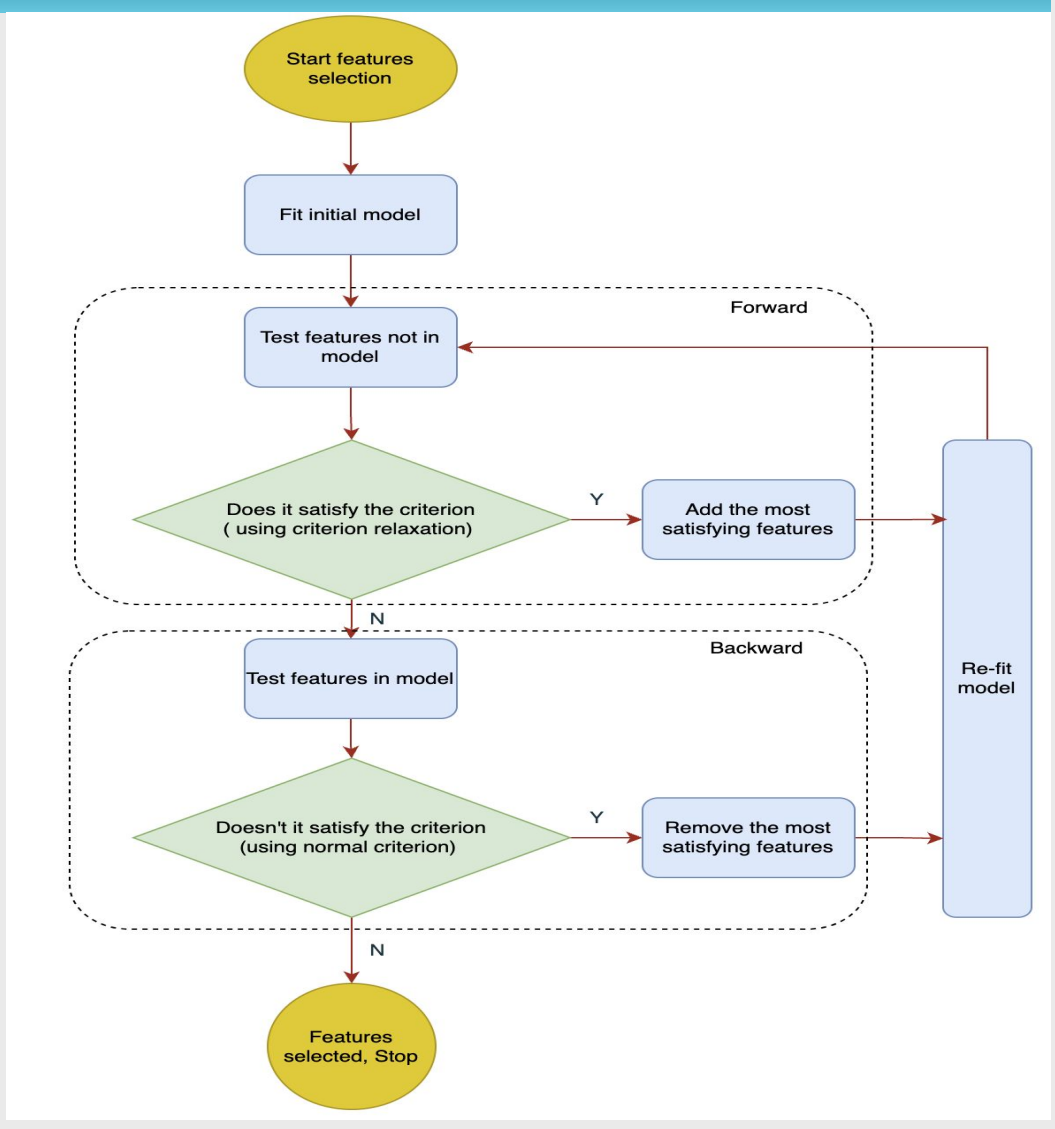
- The process includes two phase: the first for selecting the features from total features in dataset, the second phase for removing the bad features from features which chosen from first phase
- The criterion relaxation is put in the first phase, which replacing a criterion for feature selection with a closely related and simpler criterion

2. Forward Selection - First Phase

It involves a sequential process of building a model by iteratively adding one feature at a time based on relaxation criteria. The goal is to identify the most relevant features that contribute significantly to the predictive power of the model.

3. Backward Selection - Second Phase

It is the reverse process of forward selection, where the model starts with all features and iteratively eliminates one feature at a time based on certain criteria. The goal is to identify the most relevant features and remove any unnecessary or redundant ones from the model.



NII

Lê Hoàng Sinh – University of Information Technology

TEL : ********* Email : sinhlh.14@grad.uit.edu.vn