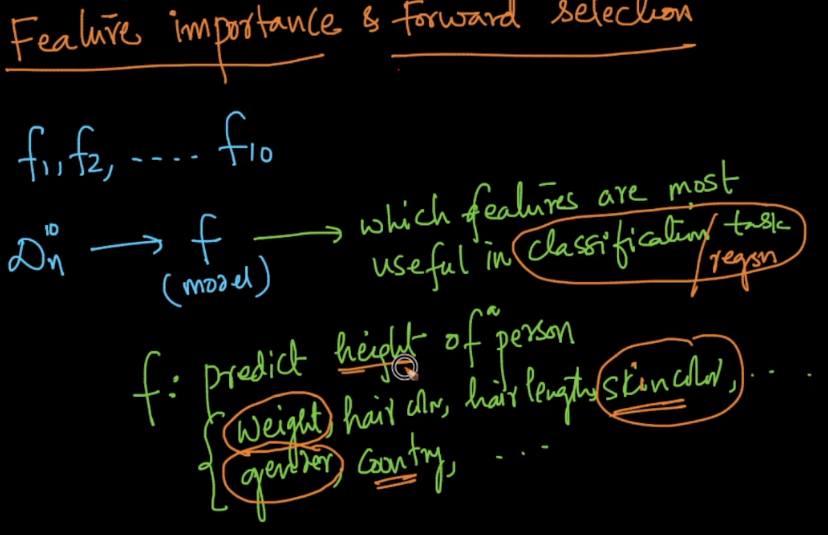
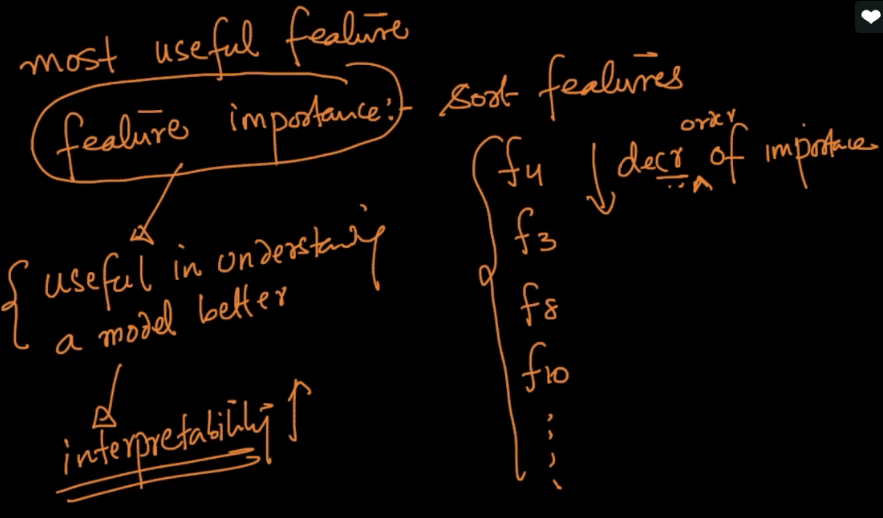
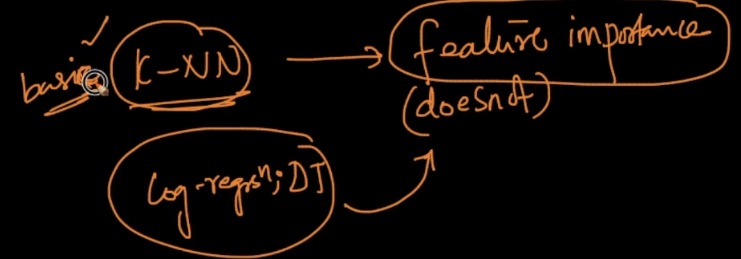
Given a dataset, It’s not important that all features are important, so who can we know relevant feature and keep them and remove rest of the features.

There are two techniques for forward selection which are explained below:

1. Forward selection.
2. Backward selection.







**Forward Selection:**

In forward selection we do following things. Suppose we have ‘d’ no of features(f1, f2,….. fd), then we do following things

Iteration 1:

we train model with only f1 feature and check accuracy.

We train model with only f2 feature and check accuracy.

Similarly we do it for all d features and check accuracy.

Then we keep the feature which has highest accuracy. Let’s f5 gets highest accuracy so we add f5 in final features list.

Iteration 2:

Since we have f5 in our final features list. So now we do following things.

Train model with only f1 & f5 feature and check accuracy.

Train model with only f2 & f5 feature and check accuracy.

Similarly we do it for all d features and check accuracy.

Then we keep the feature which has highest accuracy in combination with feature f5. Let’s f10&f5 gets highest accuracy so we add f10 also along with f5 in final features list.

So after 2 Iteration we have following features in our final feature list = { f5, f10 }.

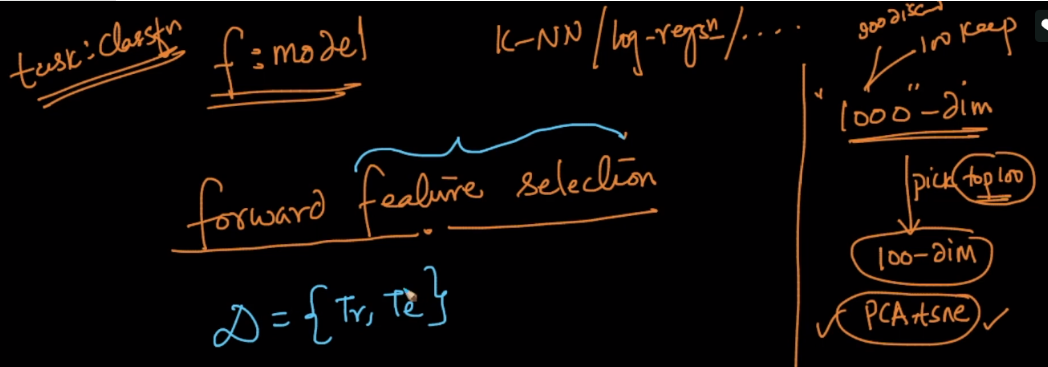
Now Each time again we perform this iteration one new feature would be added to the final feature list, which is helping in getting highest accuracy along with existing features in final feature list.

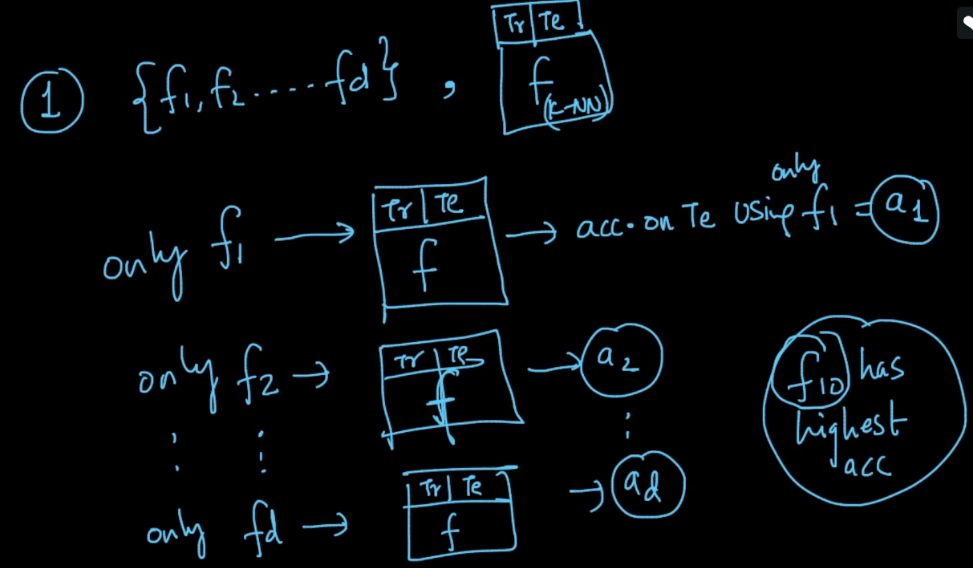
So In this way we get all the relavent features.

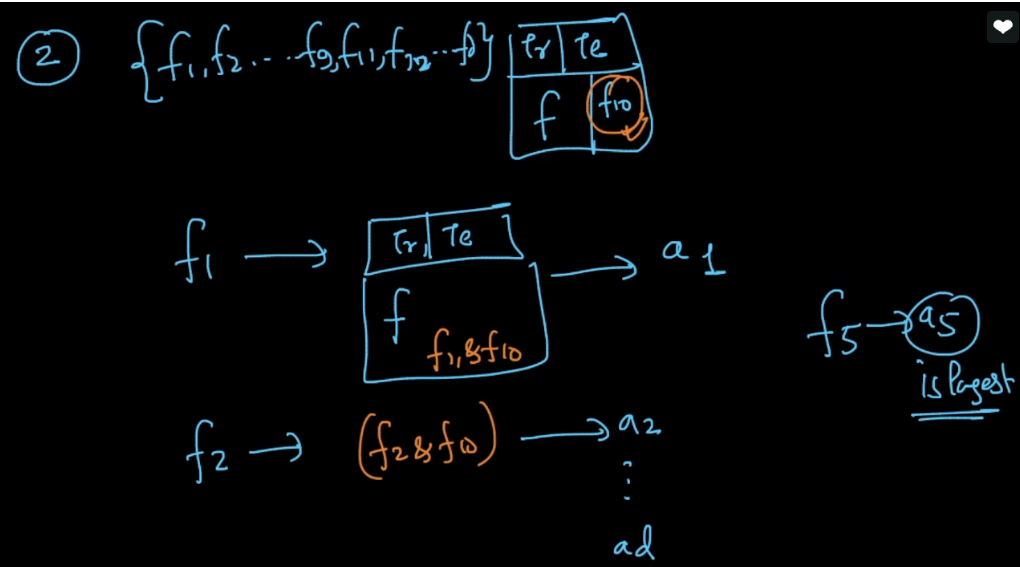
Here one question arise, why are doing so much iteration, where in 1st iteration itself we’ll get all the features with their accuracy. So we can select top N features from 1st iteration.

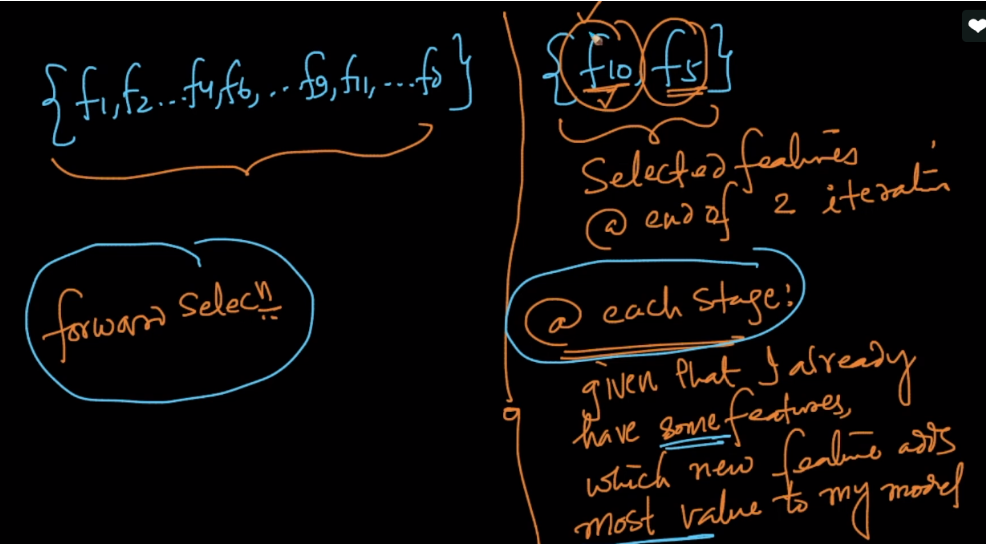
Here look we need to find features which will gives more accuracy along with the feature selected in 1st iteration instead of those features are giving high accuracy sololey.

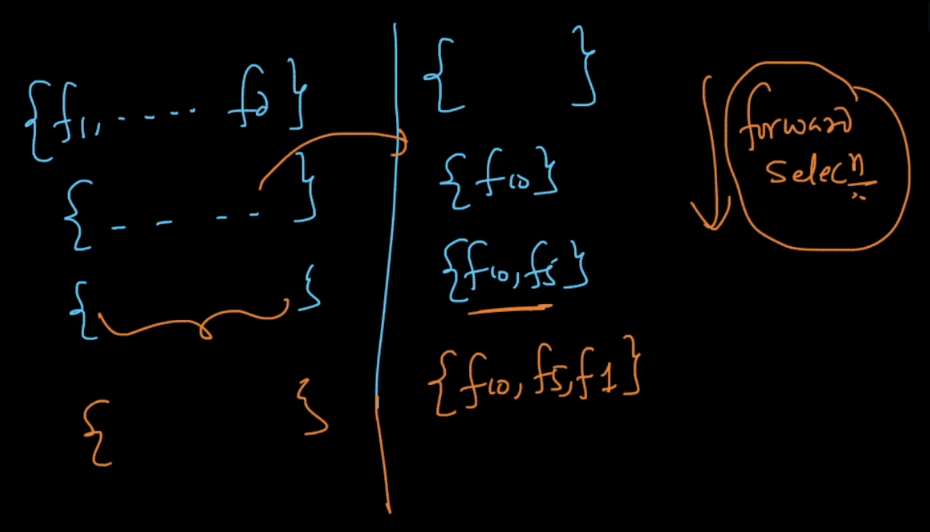
Or we can say that we want new feature will generate high accuracy, with features that I already have.





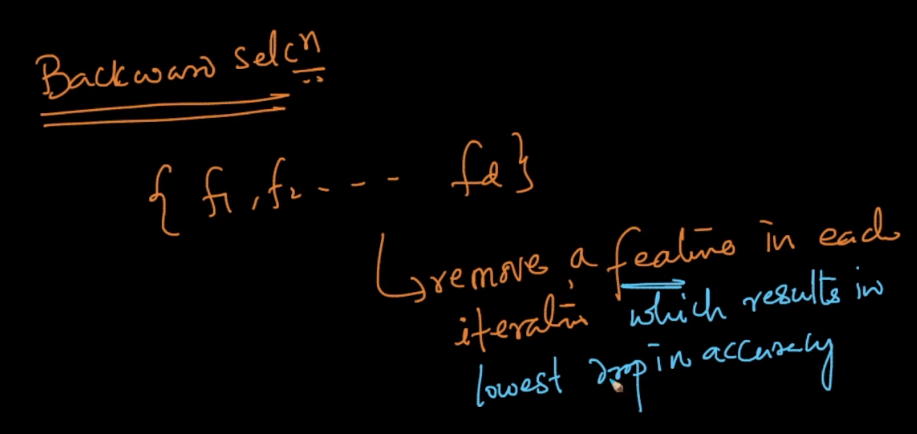






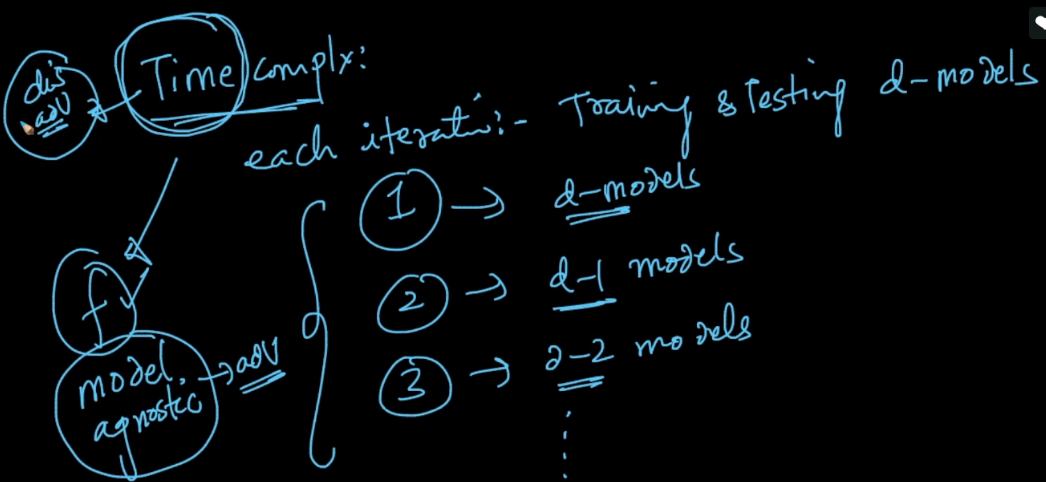
**Backward Selection:**

In backward selection also we perform Iterations, but in this first we keep all the features and remove one feature in each iteration without which our model is giving high accuracy or we are removing dumb feature.



**Advantage of Forward and Backward selection:** They are model independent, don’t need particular model to find relevant features.

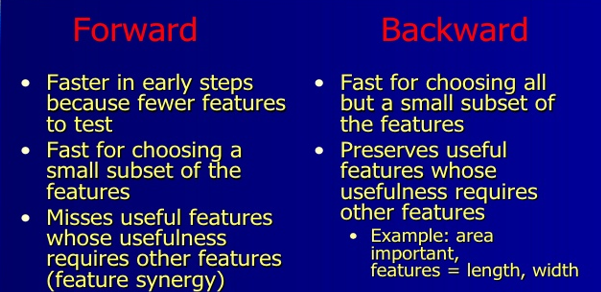
**Disadvantage of Forward and Backward selection:** They have very high time complexity.



**Diff b/w Forward and backward feature selection:**

Suppose we have a set of features {F1-F5}, upon which we run Forward Feature Selection. In the first iteration we get F3 as the most important feature and on consecutive iterations we get next most important features.  
But could there be a case where the best accuracy that can be achieved is by using the pair of F2 and F5 features? Now because we have already selected F3 as the most important feature, we might never check for such a possibility. Is this a valid case? If yes, then would it not be a case where this technique might fail?

it is one of the disadvantage of forward feature selection. backward feature selection overcomes this particular problem to some extent. it preserves useful features who's usefulness requires other features



**Comments:**

* differences between the PCA and Forward or backward selection.

TSNE and PCA are dimensionality reduction techniques not feature selection techniques. you can some times use to get some better feature representation in lower dimension but you won't select the features from these techniques.. It creates all new features and if we want to reduce the dimensionality from **a** to **b**where (b<a) then among all the **a** features given as an input to the dataset, those top **b**features are selected which could preserve maximum variance of the data

**PCA belongs to the category of Feature Engineering Techniques and Forward Feature Selection/Backward Eliminationbelong to the category of Feature Selection Techniques.**

* We should get rid of outliers before applying to PCA, because PCA is very much prone to outliers.
* Must read link: <https://www.analyticsvidhya.com/blog/2016/12/introduction-to-feature-selection-methods-with-an-example-or-how-to-select-the-right-variables/>