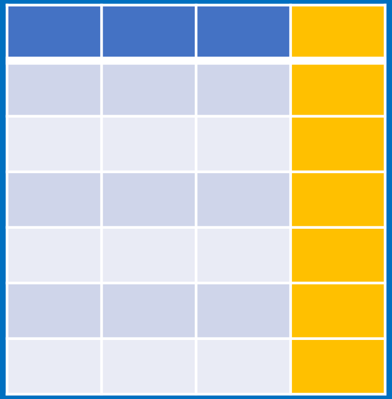
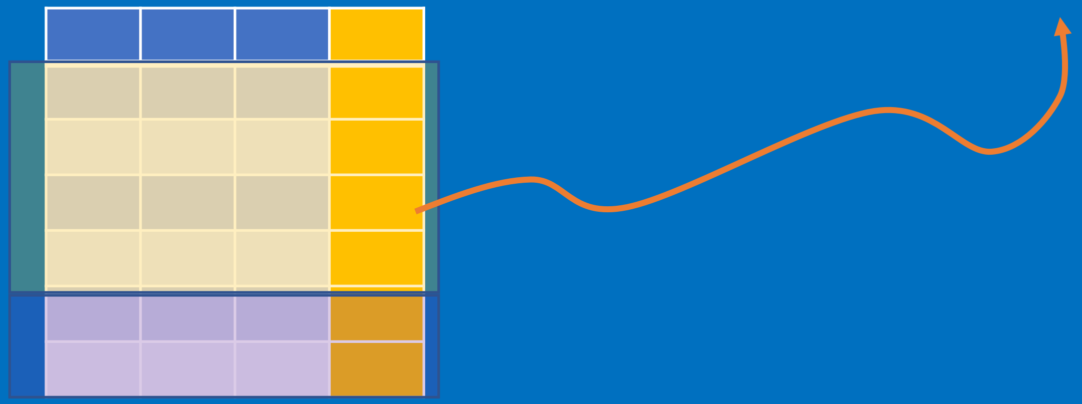
Permutation importance

Sijin Zhang

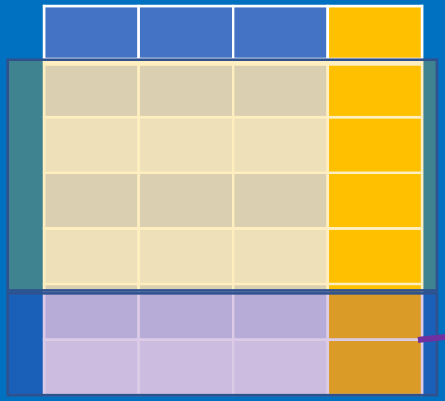


Assuming that we have the above dataset, with 3 features and 1 target

Step 1: use the training dataset to fit a model



Assuming that we have the above dataset, with 3 features and 1 target

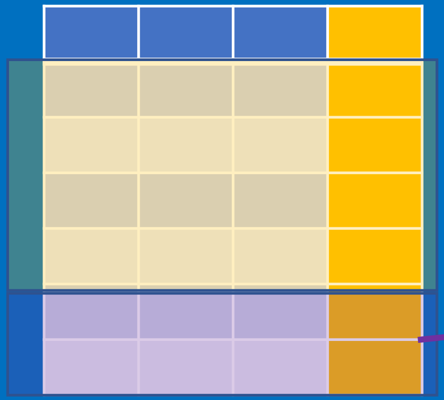


Assuming that we have the above dataset, with 3 features and 1 target

Step 1: use the training dataset to fit a model

Step 2: use test data to evaluate the trained model’s performance (e.g., we can use ACC or something else) 0.9

accuracy = 0.9



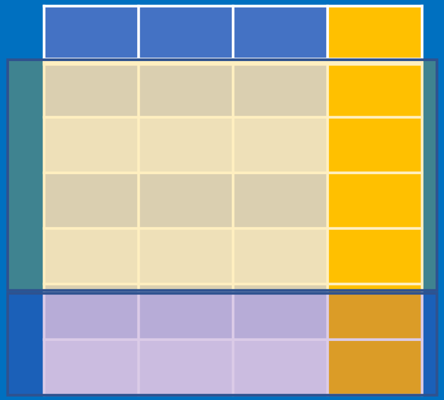
Assuming that we have the above dataset, with 3 features and 1 target

Step 1: use the training dataset to fit a model

Step 2: use test data to evaluate the trained model's performance (e.g., we can use ACC or something else)

0.9

This is used as the baseline performance

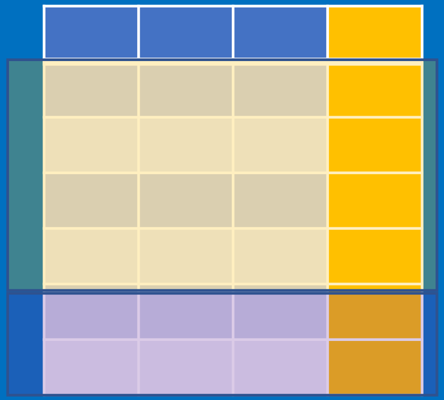


Assuming that we have the above dataset, with 3 features and 1 target

Step 1: use the training dataset to fit a model

Step 2: use test data to evaluate the trained model's performance (e.g., we can use ACC or something else) 0.9

Step 3: data permutation



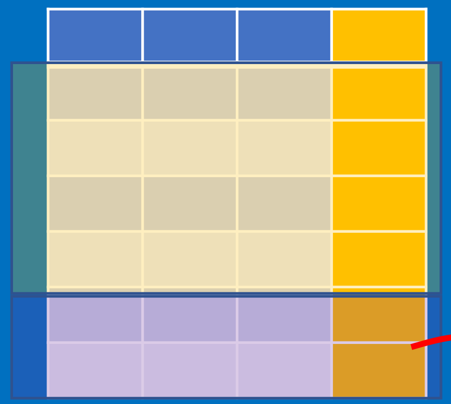
Assuming that we have the above dataset, with 3 features and 1 target

Step 1: use the training dataset to fit a model

Step 2: use test data to evaluate the trained model’s performance (e.g., we can use ACC or something else) 0.9

Step 3: test data permutation

For 1st feature, we randomly permute the column



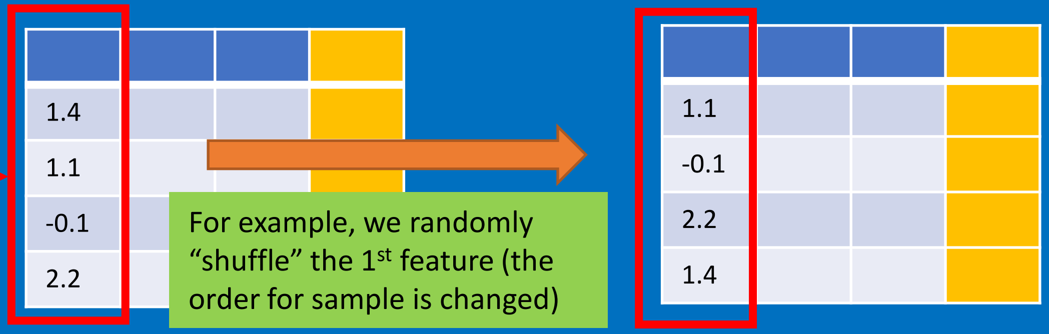
Assuming that we have the above dataset, with 3 features and 1 target

Step 1: use the training dataset to fit a model

Step 2: use test data to evaluate the trained model's performance (e.g., we can use ACC or something else) 0.9

Step 3: test data permutation

For 1st feature, we randomly permute the column



1.4			
1.1			
-0.1			
2.2			

For example, we randomly "shuffle" the 1st feature (the order for sample is changed)

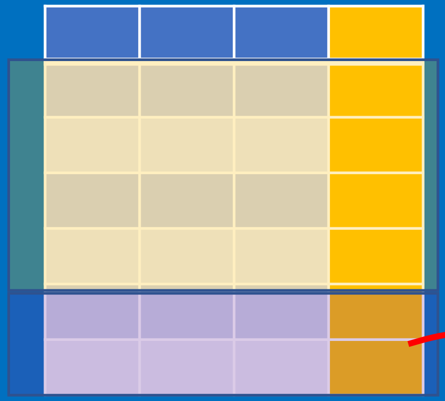
1.1			
-0.1			
2.2			
1.4			

Step 1: use the training dataset to fit a model

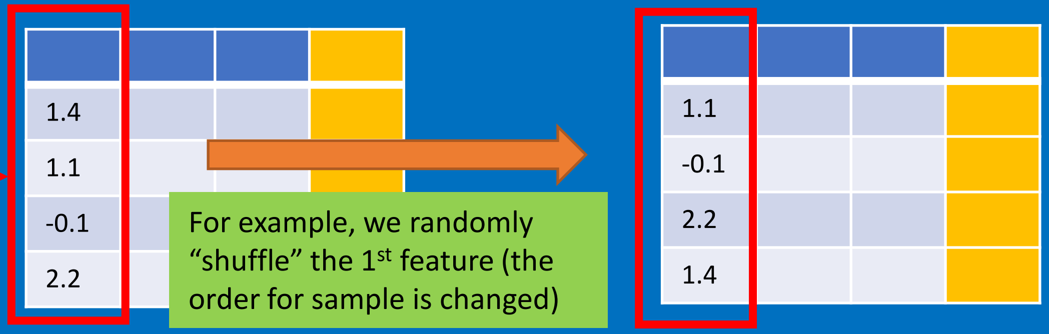
Step 2: use test data to evaluate the trained model's performance (e.g., we can use ACC or something else) 0.9

Step 3: test data permutation

For 1st feature, we randomly permute the column



Assuming that we have the above dataset, with 3 features and 1 target



1.4			
1.1			
-0.1			
2.2			

For example, we randomly "shuffle" the 1st feature (the order for sample is changed)

1.1			
-0.1			
2.2			
1.4			

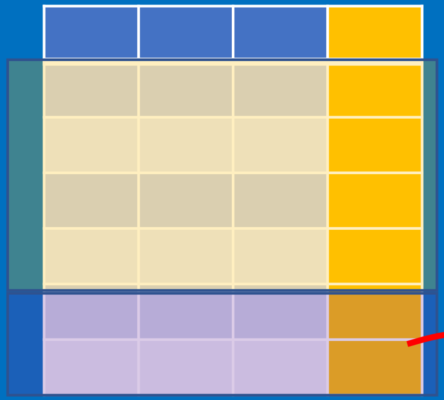
Note that we are not touching any other features and the target

Step 1: use the training dataset to fit a model

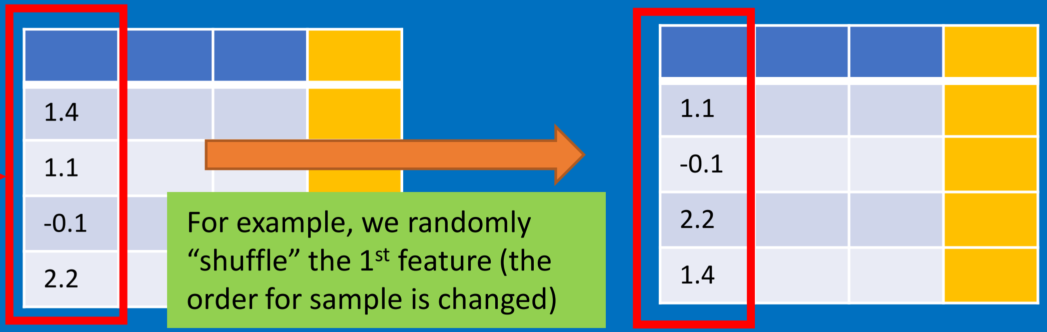
Step 2: use test data to evaluate the trained model’s performance (e.g., we can use ACC or something else) 0.9

Step 3: test data permutation

For 1st feature, we randomly permute the column



Assuming that we have the above dataset, with 3 features and 1 target



1.4			
1.1			
-0.1			
2.2			

For example, we randomly “shuffle” the 1st feature (the order for sample is changed)

1.1			
-0.1			
2.2			
1.4			

Note that we are not touching any other features and the target

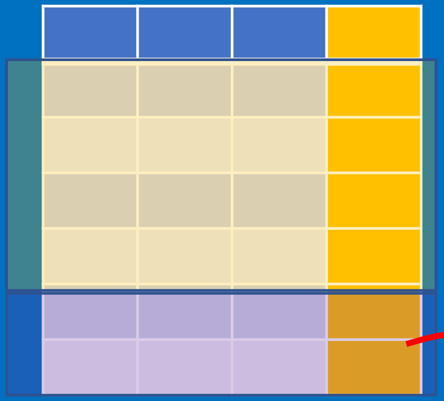
Then we use the permuted data to evaluate the trained model’s performance

Step 1: use the training dataset to fit a model

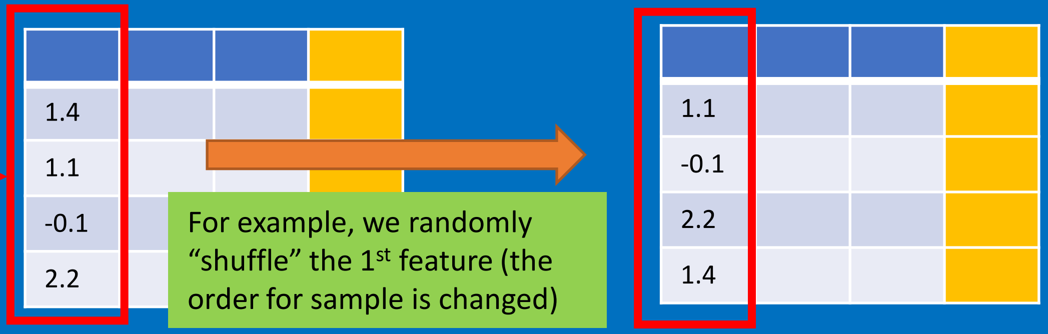
Step 2: use test data to evaluate the trained model’s performance (e.g., we can use ACC or something else) 0.9

Step 3: test data permutation

For 1st feature, we randomly permute the column



Assuming that we have the above dataset, with 3 features and 1 target



1.4			
1.1			
-0.1			
2.2			

For example, we randomly “shuffle” the 1st feature (the order for sample is changed)

1.1			
-0.1			
2.2			
1.4			

Note that we are not touching any other features and the target

accuracy = 0.7

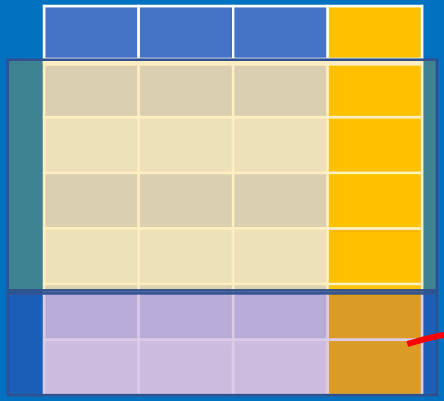
Then we use the permuted data to evaluate the trained model’s performance 0.7

Step 1: use the training dataset to fit a model

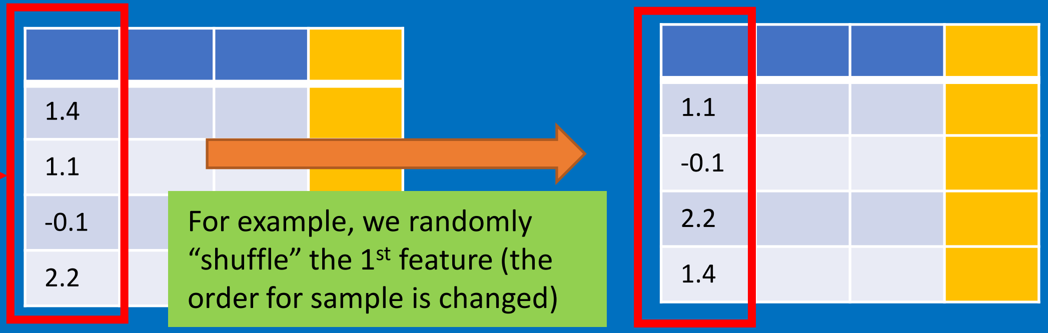
Step 2: use test data to evaluate the trained model’s performance (e.g., we can use ACC or something else) 0.9

Step 3: test data permutation

For 1st feature, we randomly permute the column



Assuming that we have the above dataset, with 3 features and 1 target



1.4			
1.1			
-0.1			
2.2			

For example, we randomly “shuffle” the 1st feature (the order for sample is changed)

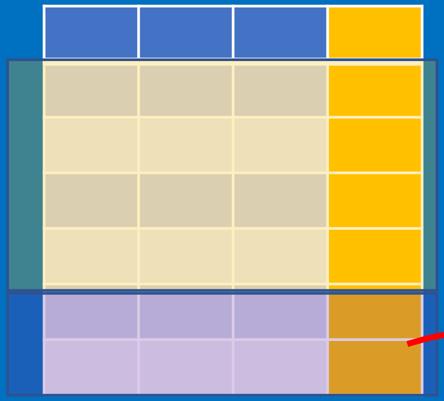
1.1			
-0.1			
2.2			
1.4			

Note that we are not touching any other features and the target

accuracy = 0.7

Then we use the permuted data to evaluate the trained model’s performance 0.7

(usually we shuffle the dataset many times and get the average performance …)



Assuming that we have the above dataset, with 3 features and 1 target

Step 1: use the training dataset to fit a model

Step 2: use test data to evaluate the trained model's performance (e.g., we can use ACC or something else) 0.9

Step 3: test data permutation

For 1st feature, we randomly permute the column

1.4			
1.1			
-0.1			
2.2			

For example, we randomly "shuffle" the 1st feature (the order for sample is changed)

1.1			
-0.1			
2.2			
1.4			

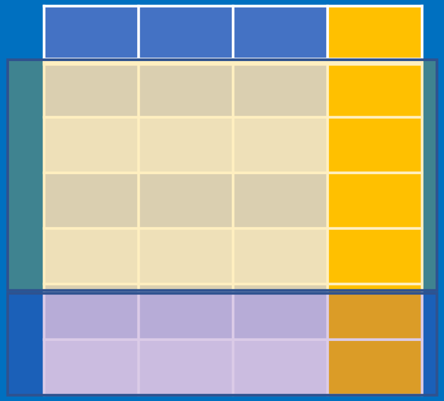
Note that we are not touching any other features and the target

accuracy = 0.7

Then we use the permuted data to evaluate the trained model's performance 0.7

(usually we shuffle the dataset many times and get the average performance …)

We would repeat the above process for all the features (in this case there are three features)



Assuming that we have the above dataset, with 3 features and 1 target

Step 1: use the training dataset to fit a model

Step 2: use test data to evaluate the trained model's performance (e.g., we can use ACC or something else) 0.9

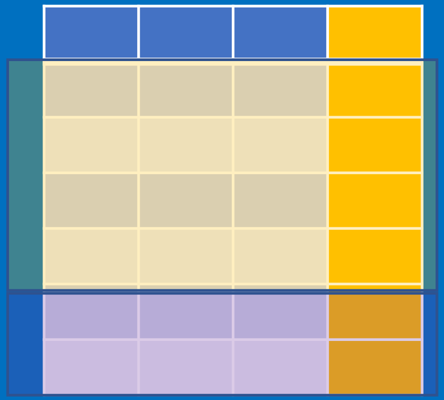
Step 3: test data permutation

For 1st feature 0.7

For 2nd feature 0.4

For 3rd feature 0.6

⇦ Model accuracies for shuffled/permuted test data



Assuming that we have the above dataset, with 3 features and 1 target

Step 1: use the training dataset to fit a model

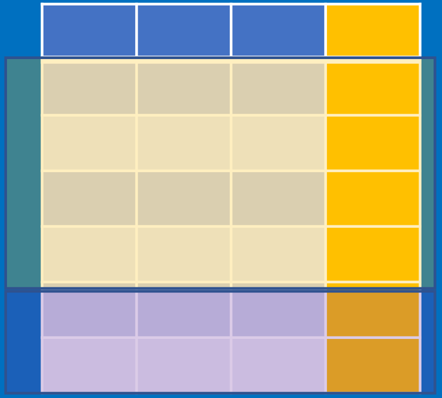
Step 2: use test data to evaluate the trained model's performance (e.g., we can use ACC or something else) 0.9

Step 3: test data permutation

For 1st feature	0.7	0.2
For 2nd feature	0.4	0.5
For 3rd feature	0.6	0.3

e.g., 0.9 – 0.7 = 0.2

Difference between the base model performance (0.9) and the one with permuted dataset



Assuming that we have the above dataset, with 3 features and 1 target

Step 1: use the training dataset to fit a model

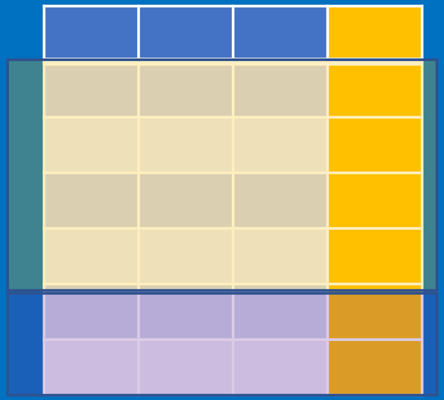
Step 2: use test data to evaluate the trained model’s performance (e.g., we can use ACC or something else) 0.9

Step 3: test data permutation

For 1st feature	0.7	0.2
For 2nd feature	0.4	0.5
For 3rd feature	0.6	0.3

Difference between the base model performance (0.9) and the one with permuted dataset

We can think about the value as how much the model relies on this feature



Assuming that we have the above dataset, with 3 features and 1 target

Step 1: use the training dataset to fit a model

Step 2: use test data to evaluate the trained model's performance (e.g., we can use ACC or something else) 0.9

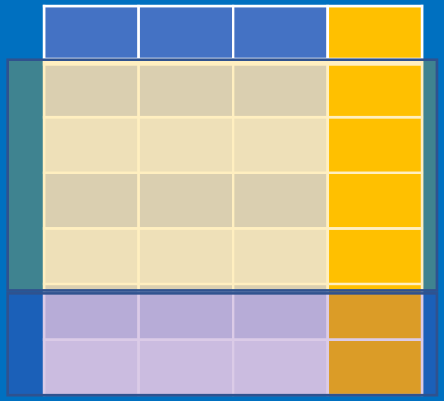
Step 3: test data permutation

For 1st feature	0.7	0.2
For 2nd feature	0.4	0.5
For 3rd feature	0.6	0.3

Difference between the base model performance (0.9) and the one with permuted dataset

We can think about the value as how much the model relies on this feature

The larger the skill drops, the more the model relies on that particular feature



Assuming that we have the above dataset, with 3 features and 1 target

Step 1: use the training dataset to fit a model

Step 2: use test data to evaluate the trained model's performance (e.g., we can use ACC or something else) 0.9

Step 3: test data permutation

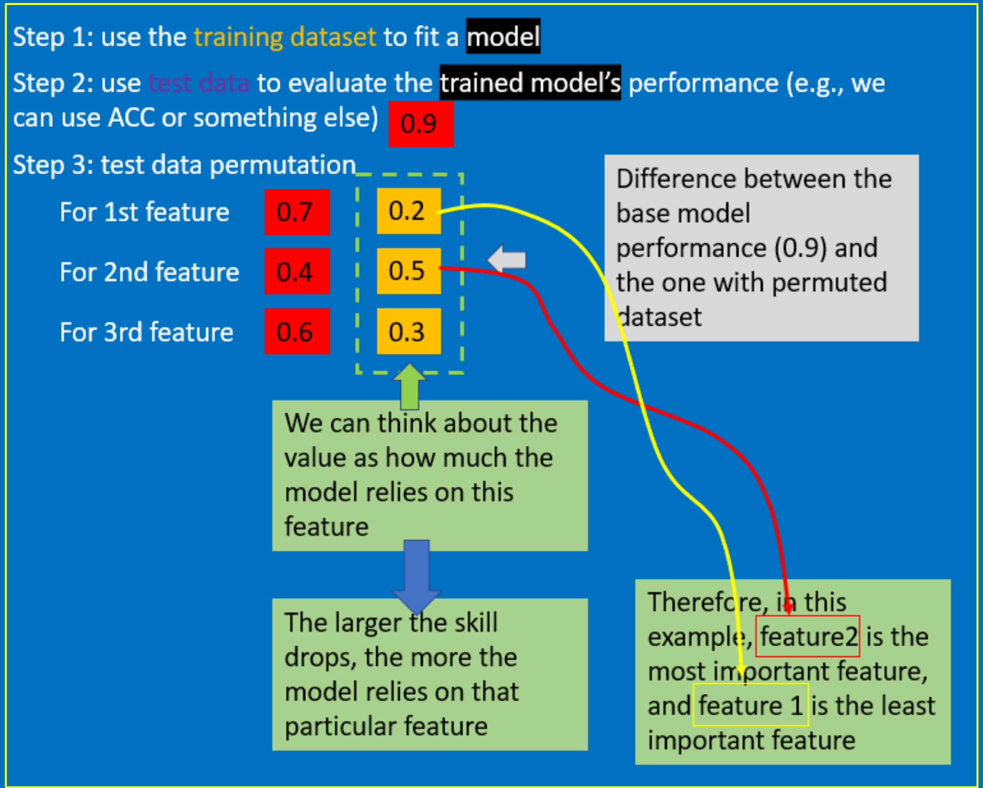
For 1st feature	0.7	0.2
For 2nd feature	0.4	0.5
For 3rd feature	0.6	0.3

Difference between the base model performance (0.9) and the one with permuted dataset

We can think about the value as how much the model relies on this feature

The larger the skill drops, the more the model relies on that particular feature

Therefore, in this example, feature2 is the most important feature, and feature 1 is the least important feature



Step 1: use the training dataset to fit a model

Step 2: use test data to evaluate the trained model's performance (e.g., we can use ACC or something else) 0.9

Step 3: test data permutation

For 1st feature	0.7	0.2
For 2nd feature	0.4	0.5
For 3rd feature	0.6	0.3

Difference between the base model performance (0.9) and the one with permuted dataset

We can think about the value as how much the model relies on this feature

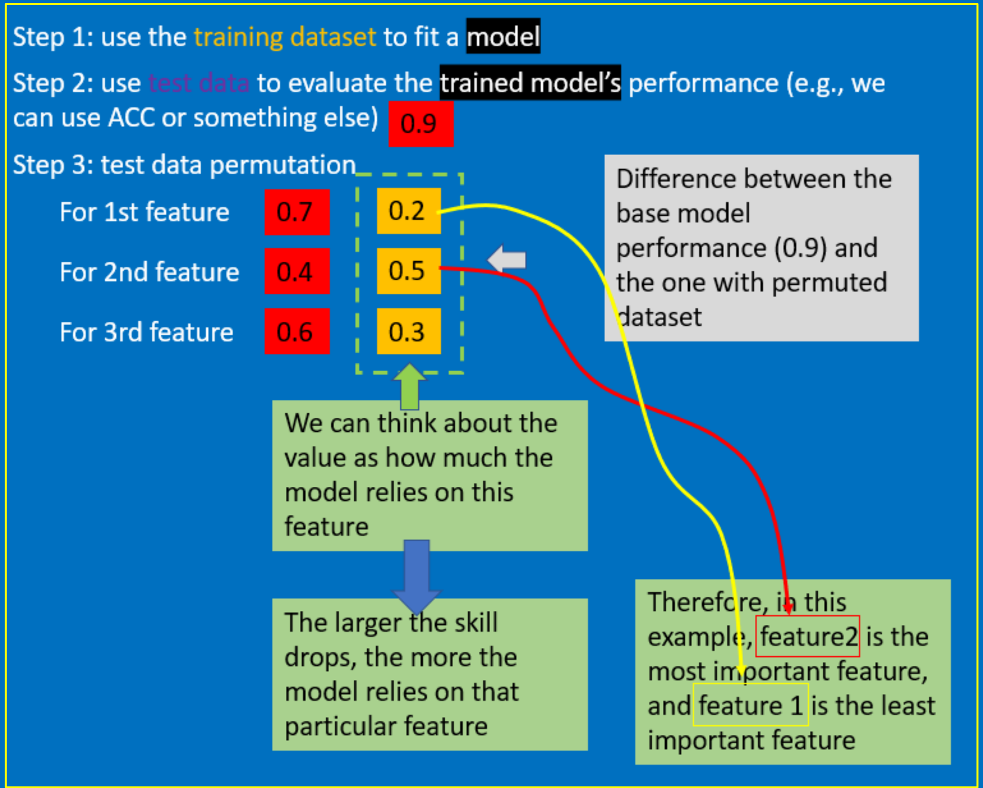
The larger the skill drops, the more the model relies on that particular feature

Therefore, in this example, feature2 is the most important feature, and feature 1 is the least important feature

Note that there is a similar method called "column-drop" method

The difference is that:

- Instead of shuffle the feature, we remove the entire feature, and "refit" a model
- We compare the base model performance (the one with entire dataset) and the "refitted" model performance (the one with one less feature)
- The feature corresponds to the most dropped skill is usually considered the most important feature



Step 1: use the training dataset to fit a model

Step 2: use test data to evaluate the trained model's performance (e.g., we can use ACC or something else) 0.9

Step 3: test data permutation

For 1st feature	0.7	0.2
For 2nd feature	0.4	0.5
For 3rd feature	0.6	0.3

Difference between the base model performance (0.9) and the one with permuted dataset

We can think about the value as how much the model relies on this feature

The larger the skill drops, the more the model relies on that particular feature

Therefore, in this example, feature2 is the most important feature, and feature 1 is the least important feature

Note that there is a similar method called “column-drop” method

The difference is that:

- Instead of shuffle the feature, we remove the entire feature, and “refit” a model
- We compare the base model performance (the one with entire dataset) and the “refitted” model performance (the one with one less feature)
- The feature corresponds to the most dropped skill is usually considered the most important feature

Compared to the permutation method, “column-drop” method is usually more accurate, but more expensive (we need to refit the model)