



## 5. Logistic Regression & ROC

# Table of Contents

- Logistic Regression with ROC & AUC.
  - With sklearn package.
  - With tensorflow.



# Sklearn.linear\_model.LogisticRegression

- `class sklearn.linear_model.LogisticRegression`
  - `(penalty='l2', dual=False, tol=0.0001, C=1.0, fit_intercept=True, intercept_scaling=1, class_weight=None, random_state=None, solver='liblinear', max_iter=100, multi_class='ovr', verbose=0, warm_start=False, n_jobs=1)`
    - `penalty`: regularization (l1 or l2)
    - `C`: Inverse of regularization strength
    - `multi_class`: multiclass option ('ovr' or 'multinomial')

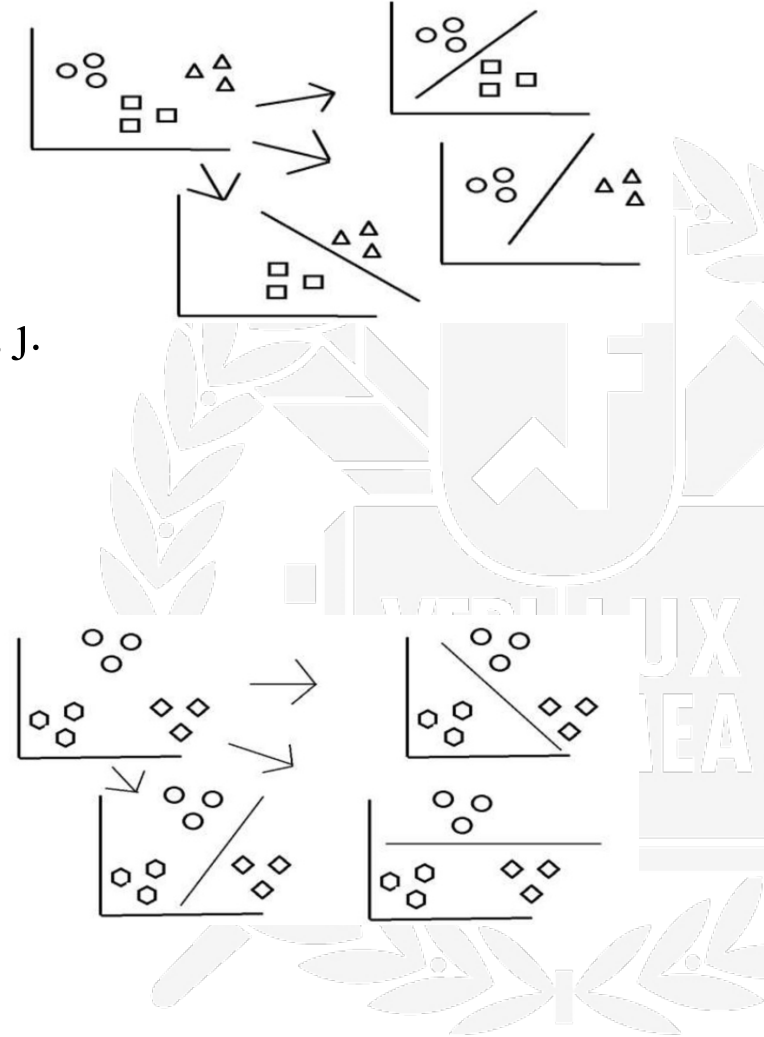
# One-versus-one & One-versus-all

- One-versus-one.

- Build  $\binom{n}{2}$  all pairwise models.
  - [ Class  $i$  vs Class  $j$  ] for all pairs of classes  $i, j$ .
- Data is less biased.
- Too many models.

- One-versus-all.

- Build  $n$  models.
  - [ Class  $i$  vs {not Class  $i$ } ] for all classes.
- True data could be insufficient.
- Reasonable number of models.



# Multinomial Logistic Regression

- Remind binary classification.

- $\text{logit} = \ln(\text{odds}) = \ln\left(\frac{P(Y=1)}{P(Y=0)}\right) = w_0 + \sum_{i=1}^k w_i x_i = WX$

- Extending for  $K$  classes.

- $\ln\left(\frac{P(Y=1)}{P(Y=K)}\right) = W_1 X \quad \Rightarrow \quad P(Y=1) = P(Y=K)e^{W_1 X}$

- ...

- $\ln\left(\frac{P(Y=K-1)}{P(Y=K)}\right) = W_{K-1} X \quad \Rightarrow \quad P(Y=K-1) = P(Y=K)e^{W_{K-1} X}$

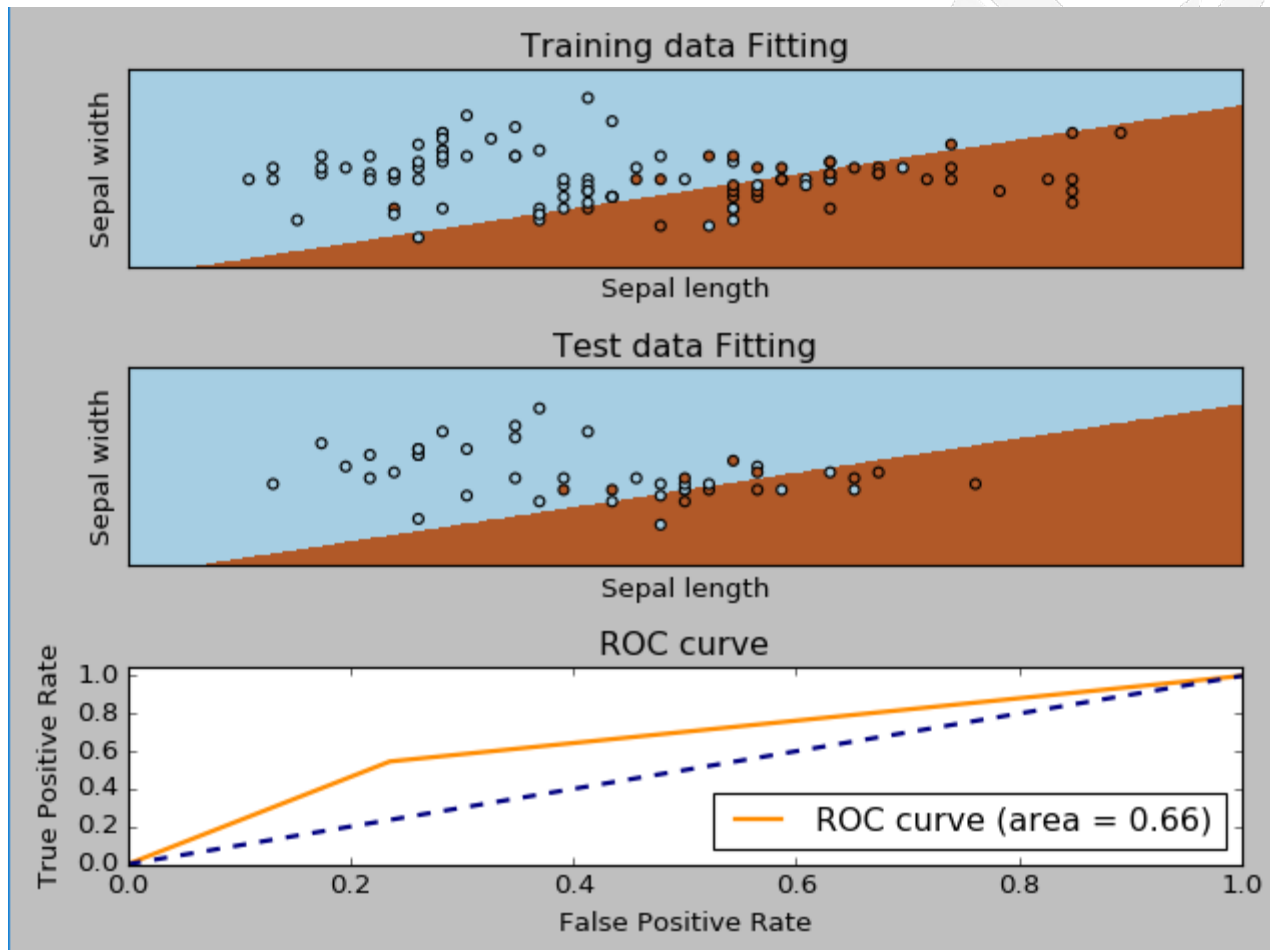
- $P(Y=K) = 1 - \sum_{i=1}^{K-1} P(Y=i) = 1 - \sum_{i=1}^{K-1} P(Y=K)e^{W_i X}$

- $P(Y=K) = \frac{1}{1 + \sum_{i=1}^{K-1} e^{W_i X}}$

- $P(Y=1) = \frac{e^{W_1 X}}{1 + \sum_{i=1}^{K-1} e^{W_i X}}$

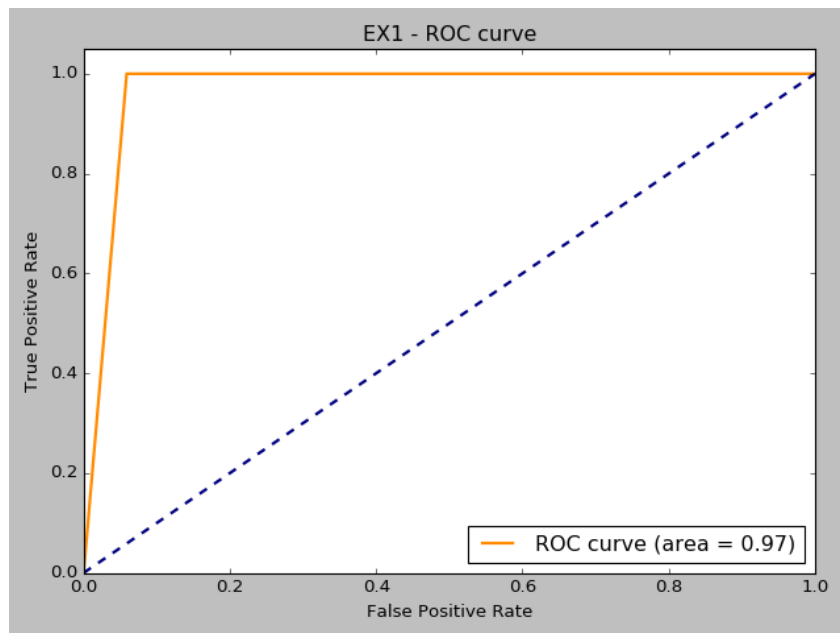
# Code Example - sklearn

- Uploaded code.



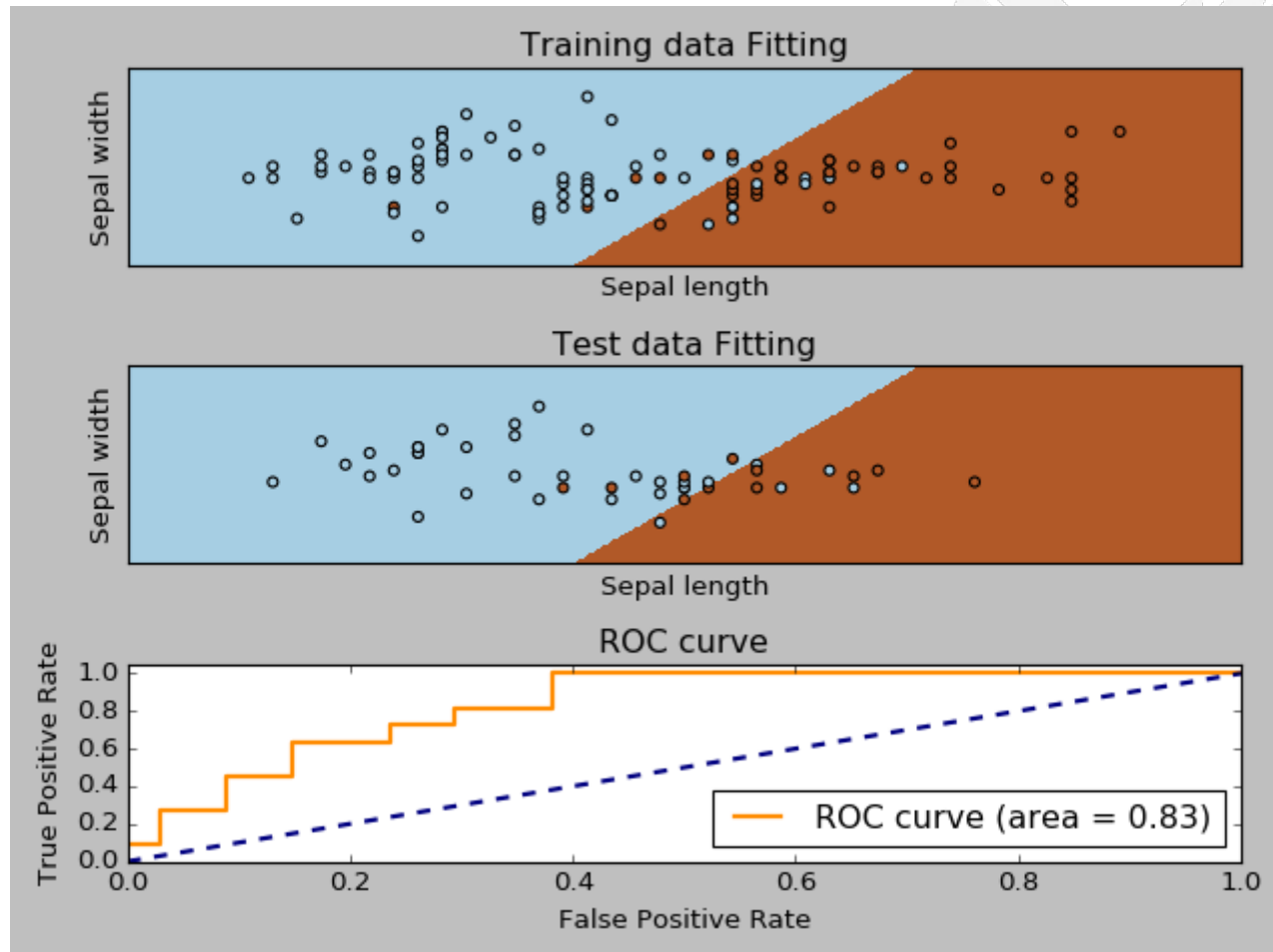
# Sklearn Exercise.

1. Use all features to fitting model with same setting of sample code.
  1. Don't plot data point scatter plots. (It is not feasible)
  2. Only plot ROC curve.
2. Do multiclass classification (3 classes) with same setting of sample code.
  1. Don't plot ROC curve. (It is not feasible).



# Code Example - tensorflow

- Uploaded code.





# Tensorflow Exercise.

1. Use all features to fitting model with same setting of sample code.
  1. Don't plot data point scatter plots. (It is not feasible)
  2. Only plot ROC curve.
2. Do multiclass classification (3 classes) with same setting of sample code.
  1. Don't plot ROC curve. (It is not feasible).

