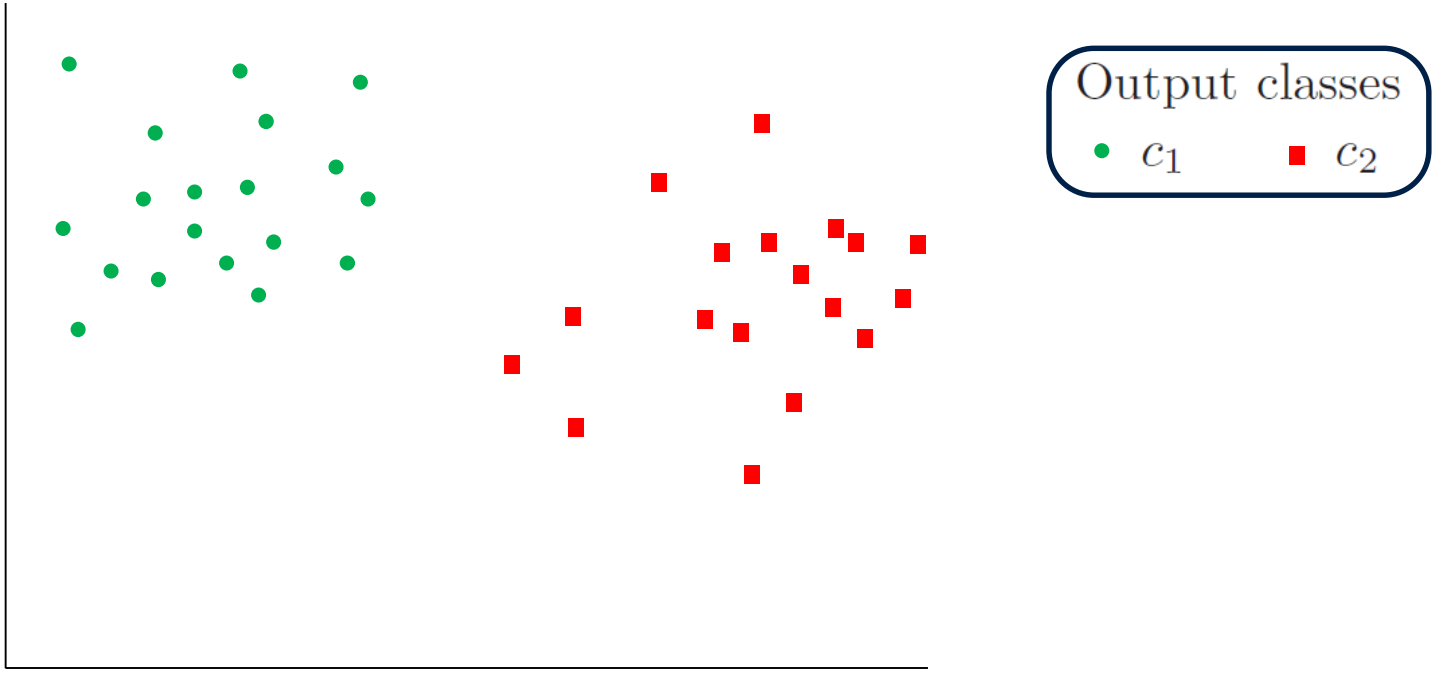
Minimum distance classifier

Dripta Mj
Department of Mathematics
Ramakrishna Mission Vivekananda Educational and Research Institute
Belur Math, India

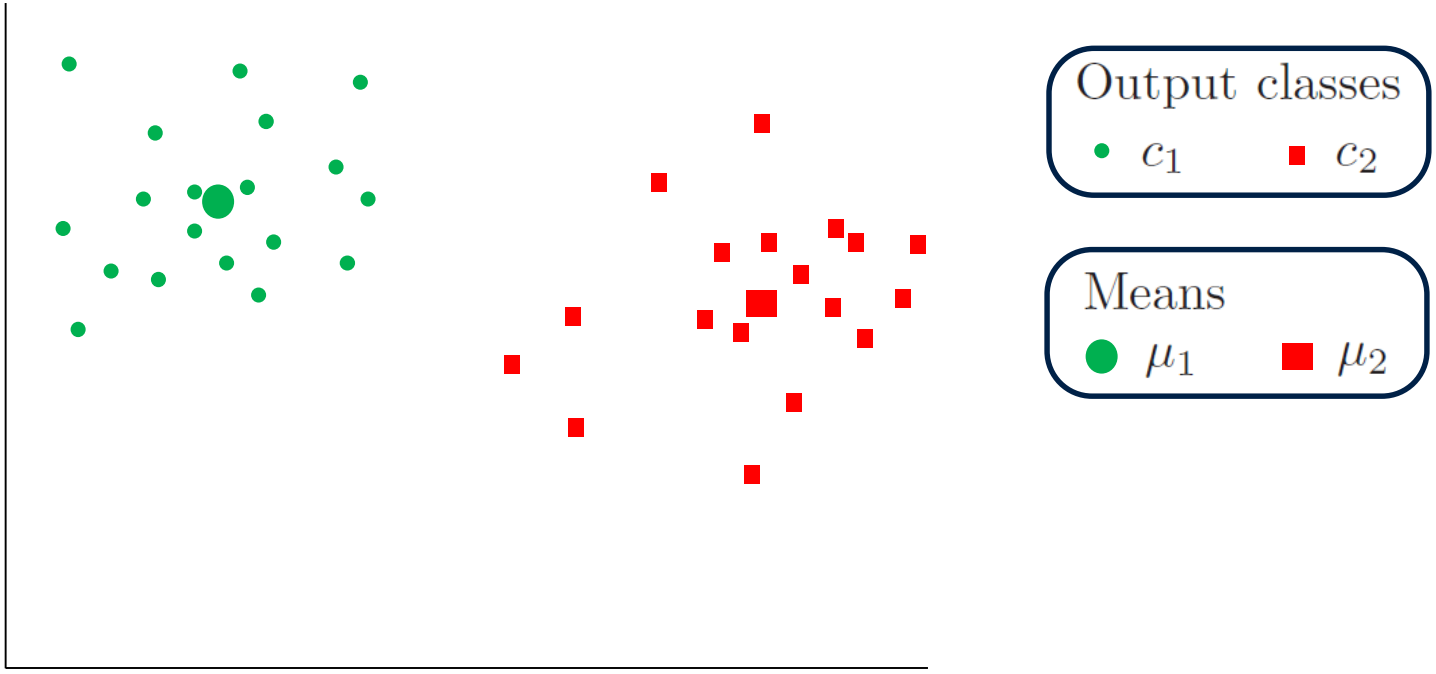
Machine Learning
CS230
Sem 3, 2018-19

Training data



- Segregation of the training data based on the class they belong to.

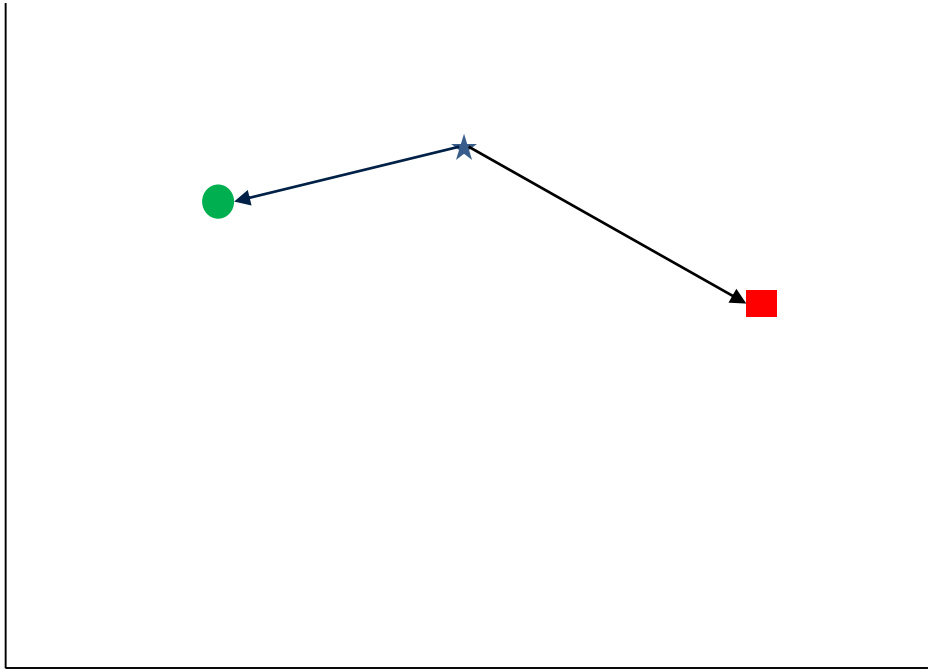
Mean of each class



- Compute the mean μ of each class using training data points belonging to that class.

Test data

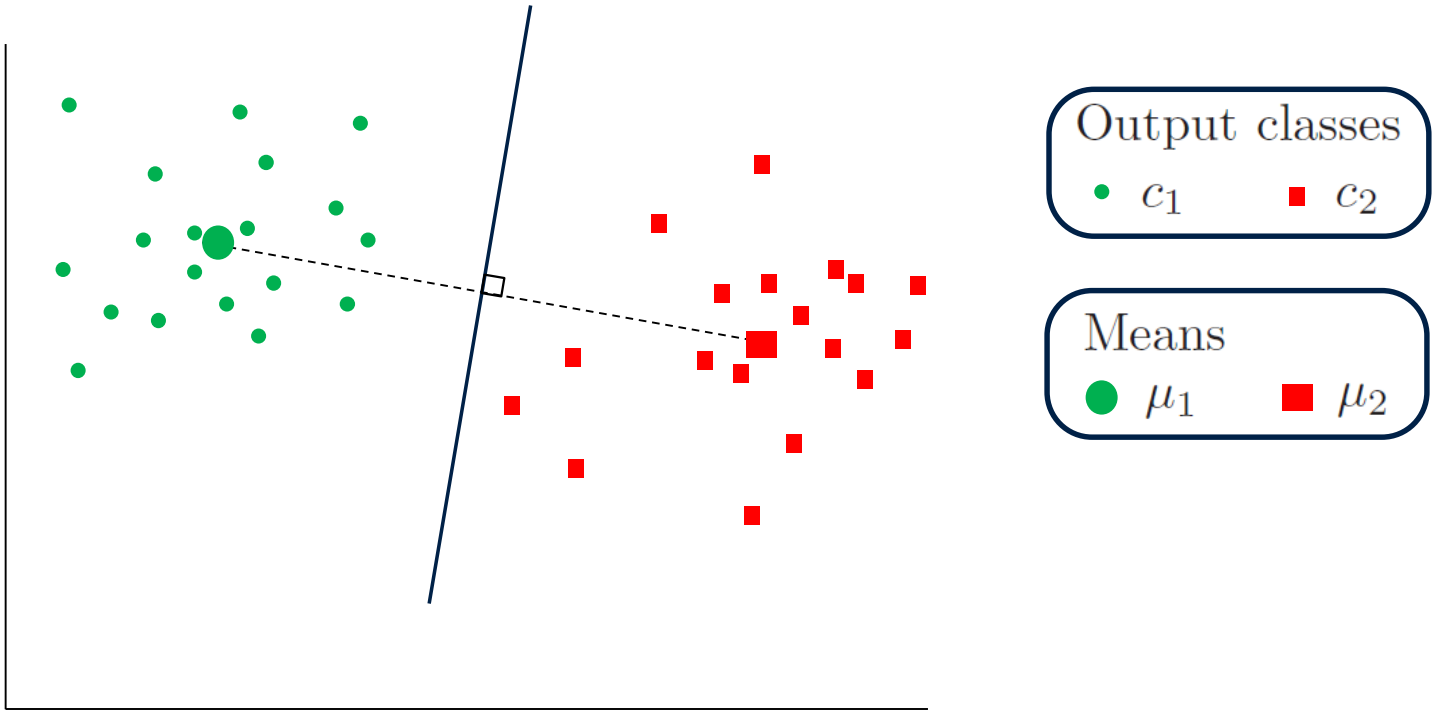
- Classify new input data.



- Determine the mean closest to the test point.
- Prediction: The class of the mean closest to the test point.

★ ⟶ c_1

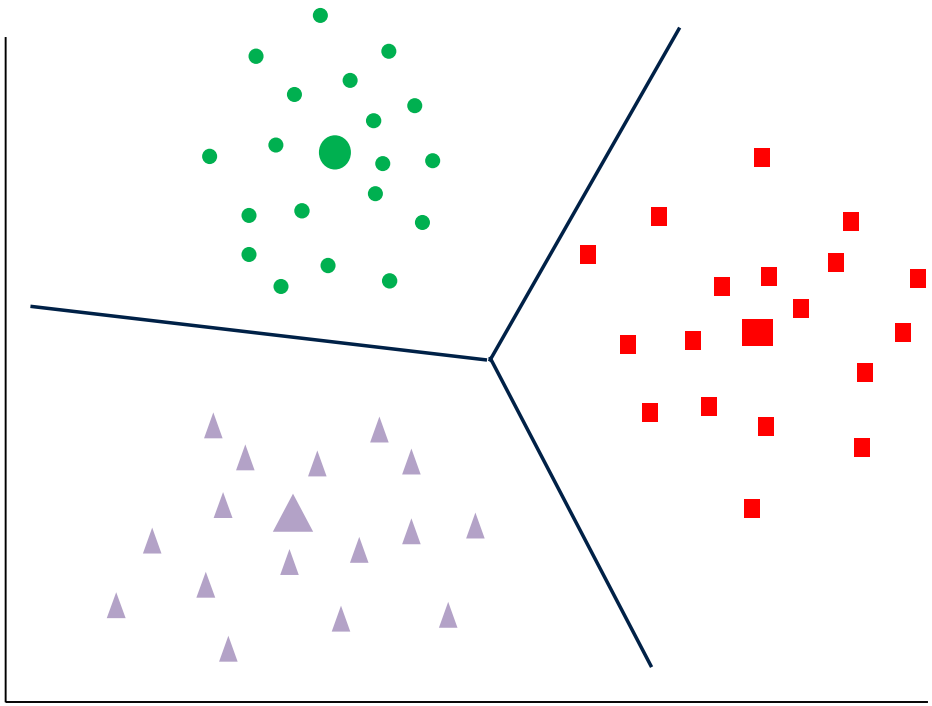
Decision boundary

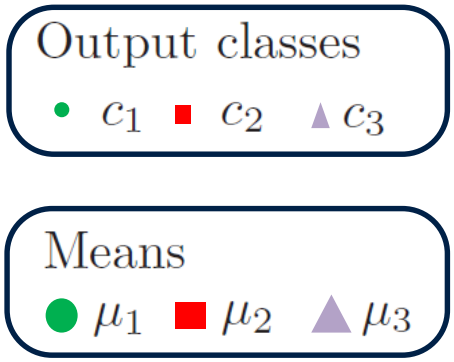


- Decision boundary: Locus of the line lying halfway between the two means.
- Any point lying on this line is equidistant to both the means.

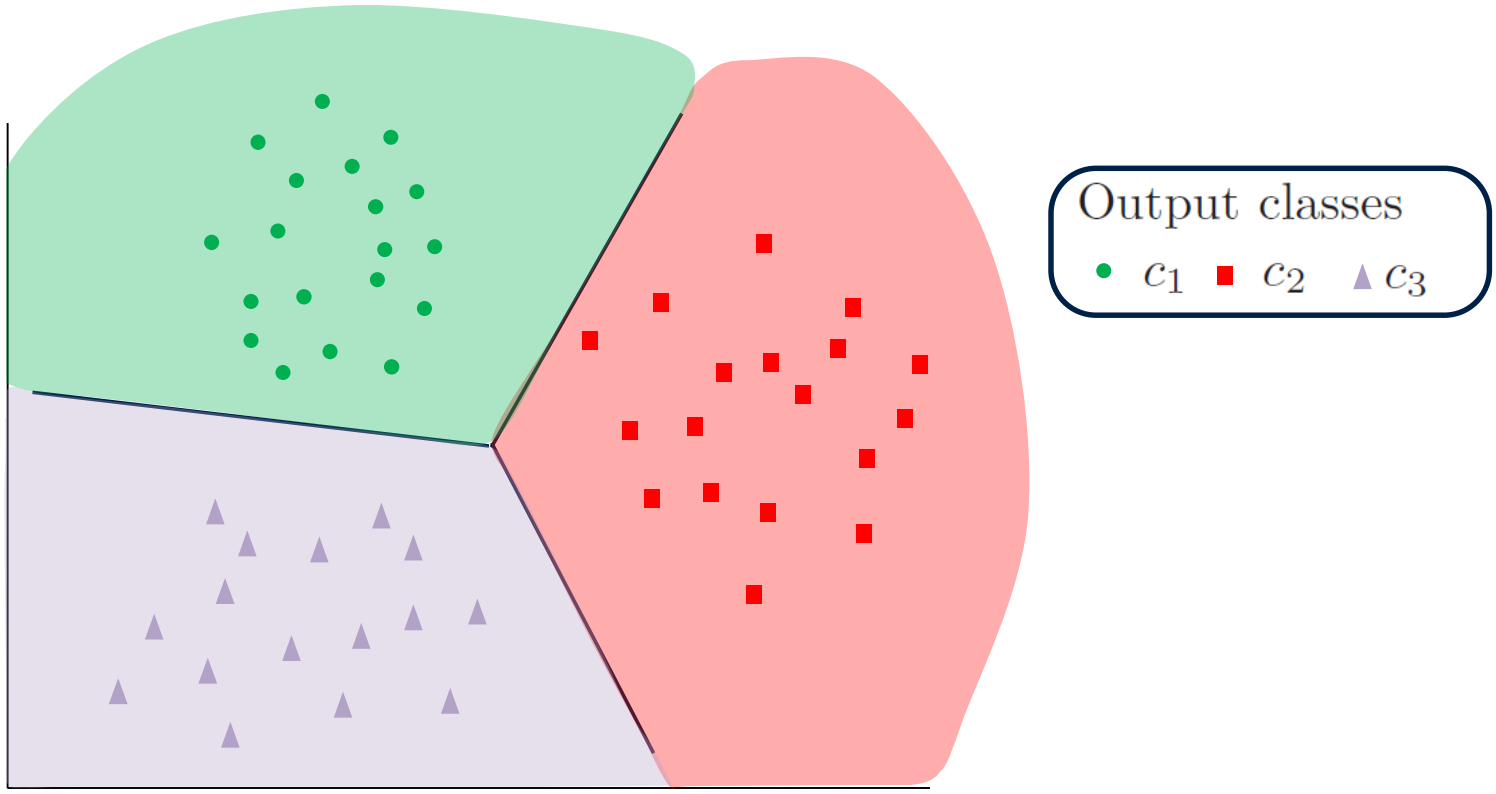
Data with 3 classes

- Feature space divided by decision boundaries.





Decision boundary



- Decision boundaries divide the feature space into decision regions $\{\mathcal{R}_1, \mathcal{R}_2, .., \mathcal{R}_M\}$, where M is the number of classes.
- Points located in region $\mathcal{R}_j$ is assigned class c_j.