Tutorial for Project 2

Đorđe Miladinović, Aytunc Şahin

October 30. - November 1

Advanced Machine Learning, Autumn 2019

Outline

- Multiclass Classification
- Class imbalance
- Task description

Multiclass Classification

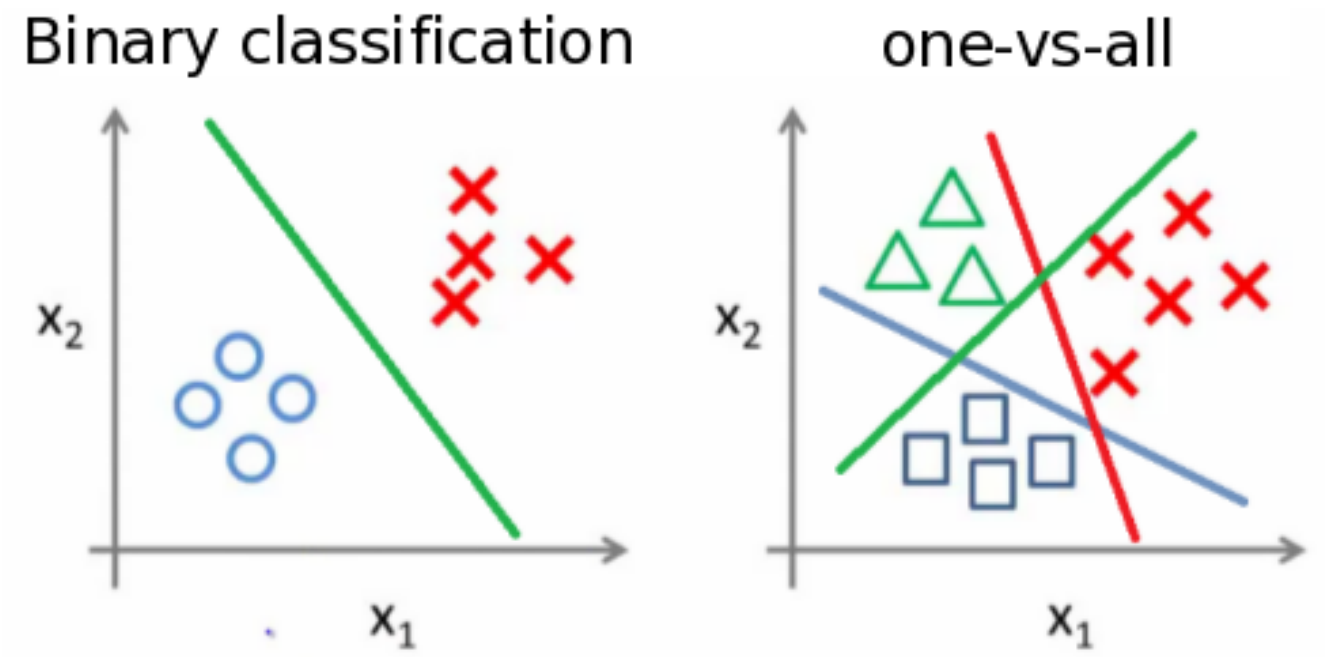
Multiclass Classification

- Classification when there are more than two classes.
- Several ways to approach this:

    1. one-vs-all strategy
    2. one-vs-one strategy
    3. multinomial logistic regression
    4. ...

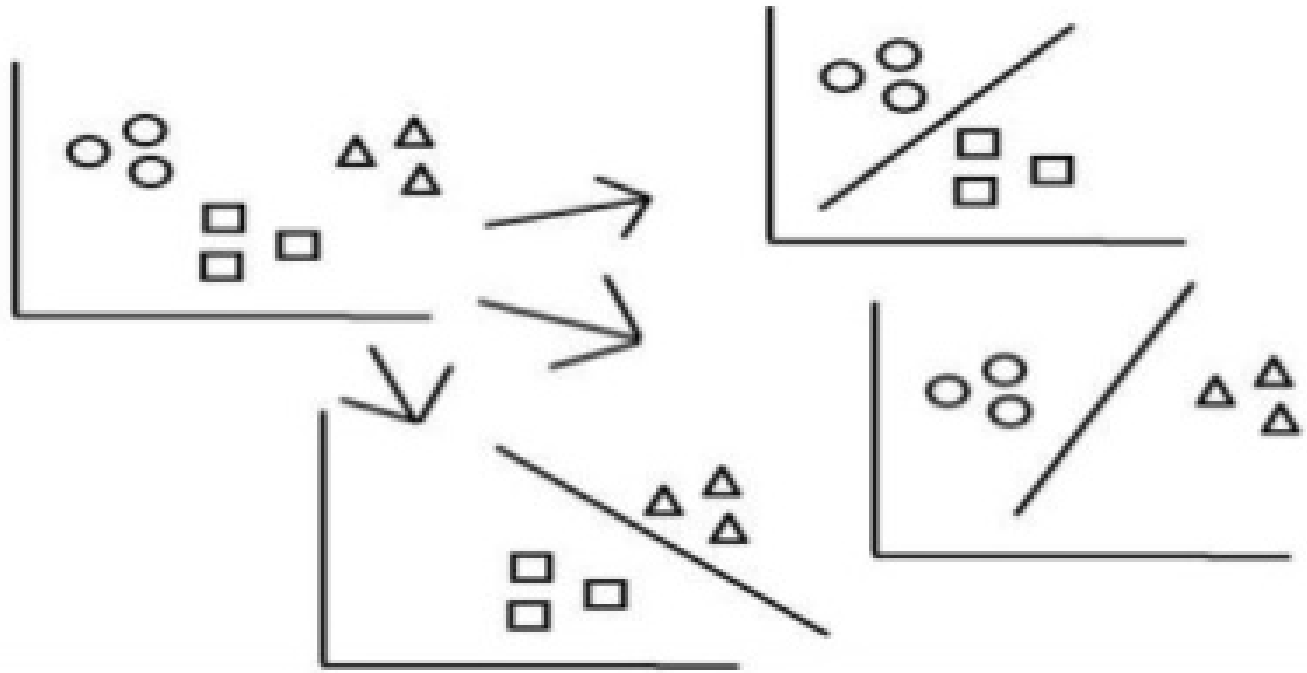
Multiclass Classification – One-vs-All

- Create C *binary* classifiers.
- Each classifier separates one of the classes with respect to *the union of all other classes*.
- Classifier with highest value / confidence is chosen in test time.



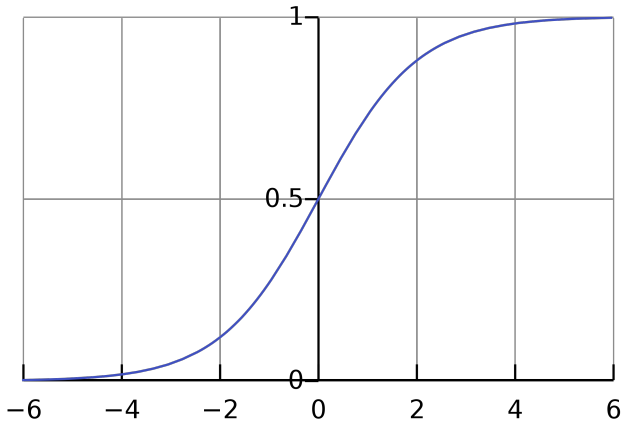
Multiclass Classification – One-vs-One

- Create $\binom{C}{2}$ binary classifiers, one for every pair of classes.
- In test time, the class with "highest number of votes wins", according to some voting schema and some tie-breaking rule.



Multiclass Classification – Multinomial Logistic Regression

Logistic function is given as $p_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$. For $\theta = \vec{1}$ we have:



- Binary classification loss is defined as follows:
- $$L(\theta) = -\sum_{i=1}^{N} y_i log(p_\theta(x_i)) + (1 - y_i) log(1 - p_\theta(x_i))$$

where (x_i, y_i) are N pairs of type (data point, binary label).

Multiclass Classification – Multinomial Logistic Regression

Logistic regression for multiclass classification:

$$L(\theta) = -\sum_{i=1}^{N}\sum_{c=1}^{C} \mathbb{1}_{\{y_i=c\}} \log\left(\frac{e^{\theta_c^T x_i}}{\sum_{k=1}^{C} e^{\theta_k^T x_i}}\right)$$

Class imbalance

Class imbalance

- Refers to the case when the classes are unequally distributed across a data set.
- We may have a situation in which certain classes are under- or over-represented.
- Can be more or less important depending on what we optimize for.
- Example:

		Predicted	
	Classifier I	Cancer	Normal
True	Cancer	0	5
	Normal	0	1000

	Predicted	
Classifier II	Cancer	Normal
Cancer	5	0
Normal	95	900

Which classifier performs better?

Class imbalance

Possible solutions:

1. Oversampling.
2. Downsampling.
3. Balanced training with weighed loss function.

Task description

Task description

Data:

- Pre-extracted features from medical images.
- Each image given as a continuous feature vector of size 1000.
- 4800 training and 4100 test feature vectors.
- 3 classes labeled as $\{0, 1, 2\}$.
- The same class distribution for training and test data.

Task:

- Multiclass image classification.
- Learn to map feature vectors into labels.

Evaluation:

- Class balance-sensitive metric.
- $\boxed{BMAC = \frac{1}{C}\sum_{c=1}^{C} TPR_c}$ where $C \in \{0, 1, 2\}$ and $TPR_c = \frac{\#correct_c}{\#all_c}$

Questions?