MAI4CAREU

Master programmes in Artificial Intelligence 4 Careers in Europe

Πανεπιστήμιο Κύπρου - Τεχνητή Νοημοσύνη

MAI612 - ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

Διάλεξη 4: Ταξινόμηση

Βασίλης Βασιλειάδης, PhD

Χειμερινό Εξάμηνο 2022/23

CYENS CENTRE OF EXCELLENCE

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Διάλεξη 4: Ταξινόμηση

Μαθησιακά αποτελέσματα

Θα καταλάβετε:

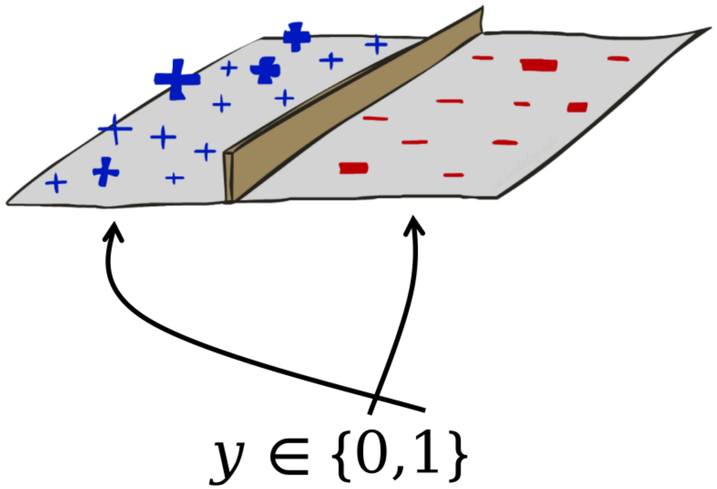
1. ο αλγόριθμος ταξινόμησης k-κοντινότερου γείτονα
2. πώς λειτουργεί το logistic regression
3. την έννοια του ορίου της απόφασης
4. συνάρτηση σφάλματος διασταυρούμενης εντροπίας για δυαδική και πολλαπλή ταξινόμηση
5. ανάλυση σφαλμάτων: μετρήσεις, πίνακας σύγχυσης και καμπύλες ROC
6. ταξινόμηση πολλαπλών κατηγοριών με χρήση του OneVsRest και του softmax ταξινομητή.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

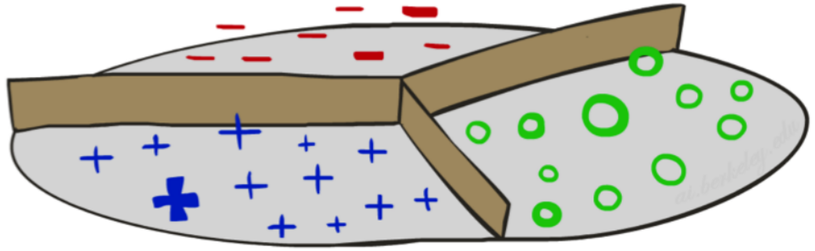
Ταξινόμηση

Binary classification



$y \in \{0,1\}$

Multiclass classification



$y \in \{0,1,2\}$

Πηγή εικόνων: ***UC Berkeley CS188 – Εισαγωγή στο μάθημα AI***

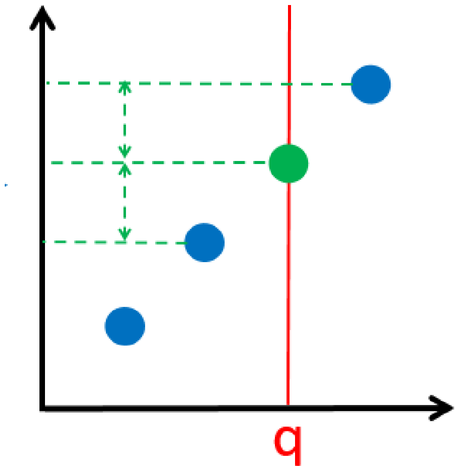
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Ταξινόμηση k-κοντινότερου γείτονα

Δίνονται:

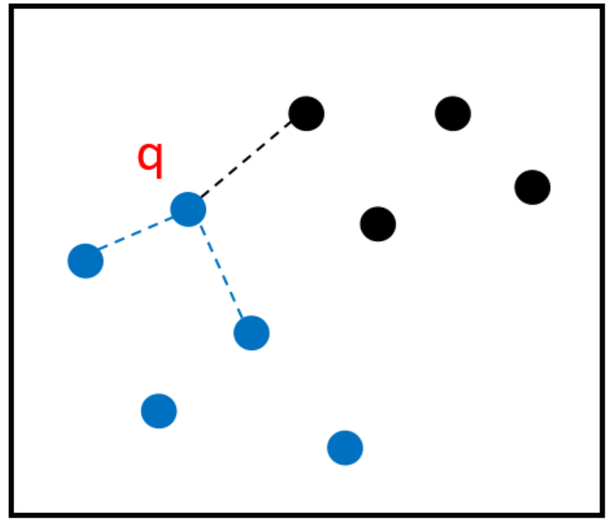
- Δεδομένα Εκπαίδευσης $D = \{x^{(i)}, y^{(i)}\}_{i=1:m}$
- Σημείο ερωτήματος q
- Μετρική απόστασης $d(q, x^{(i)})$
- Αριθμός γειτόνων k



NN = $\{i: d(q, x^{(i)})\ k$ μικρότερο$\}$ (k πλησιέστερα στο σημείο ερωτήματος)

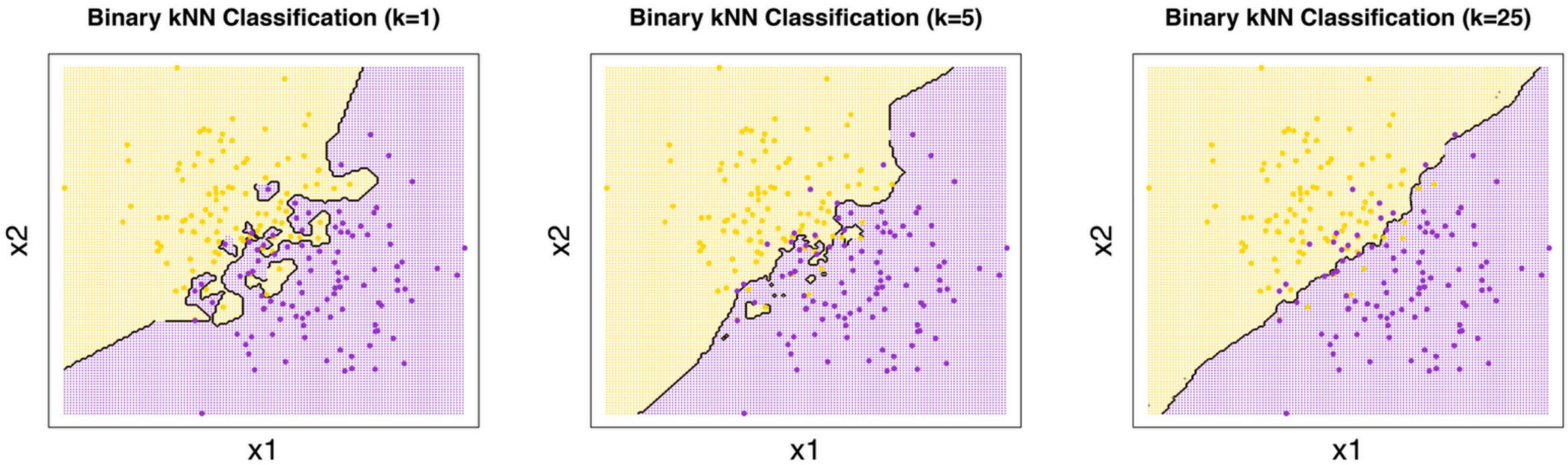
Επιστρέφει:

- Παλινδρόμηση: Μέση τιμή του $y^{(i)}$ στο NN
- Ταξινόμηση: Ψήφος του $y^{(i)}$ του NN



This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Ταξινόμηση k-κοντινότερου γείτονα



Τι γίνεται αν k = m; ⟶ **Πρόβλεψε πάντα την κλάση της πλειοψηφίας**

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Παράδειγμα

```
In [1]: X = [[0], [1], [2], [3]]
   ...: y = [0, 0, 1, 1]

In [2]: from sklearn.neighbors import KNeighborsClassifier

In [3]: neigh = KNeighborsClassifier(n_neighbors=3)

In [4]: neigh.fit(X, y)
Out[4]:
KNeighborsClassifier(algorithm='auto', leaf_size=30, metric='minkowski',
           metric_params=None, n_jobs=None, n_neighbors=3, p=2,
           weights='uniform')

In [5]: print(neigh.predict([[1.1]]))
[0]

In [6]: print(neigh.predict_proba([[0.9]]))
[[0.66666667 0.33333333]]
```


This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

MAI4CAREU

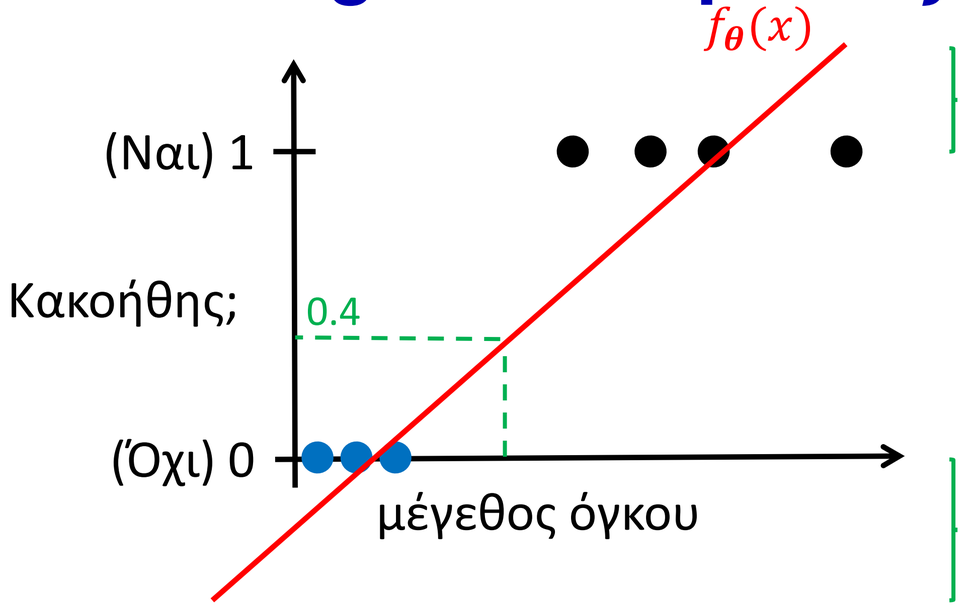
Master programmes in Artificial Intelligence 4 Careers in Europe

Logistic Regression

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Linear regression για ταξινόμηση



Η γραμμική παλινδρόμηση θα μπορούσε να λειτουργήσει, αλλά:

- χρειαζόμαστε μια άλλη παράμετρο (κατώτατο όριο): εάν $f_\theta(x) \geq 0.4$,προβλέψτε «y=1» εάν $f_\theta(x) < 0.4$,προβλέψτε «y=0»
- είναι απεριόριστο $f_\theta(x)$ μπορεί να είναι > 1 και < 0

Logistic Regression

εάν$f_\theta(x) \geq 0.5$, προβλέψτε το «y=1»
εάν$f_\theta(x) < 0.5$, προβλέψτε «y=0» $0 \leq f_\theta(x) \leq 1$

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Logistic regression

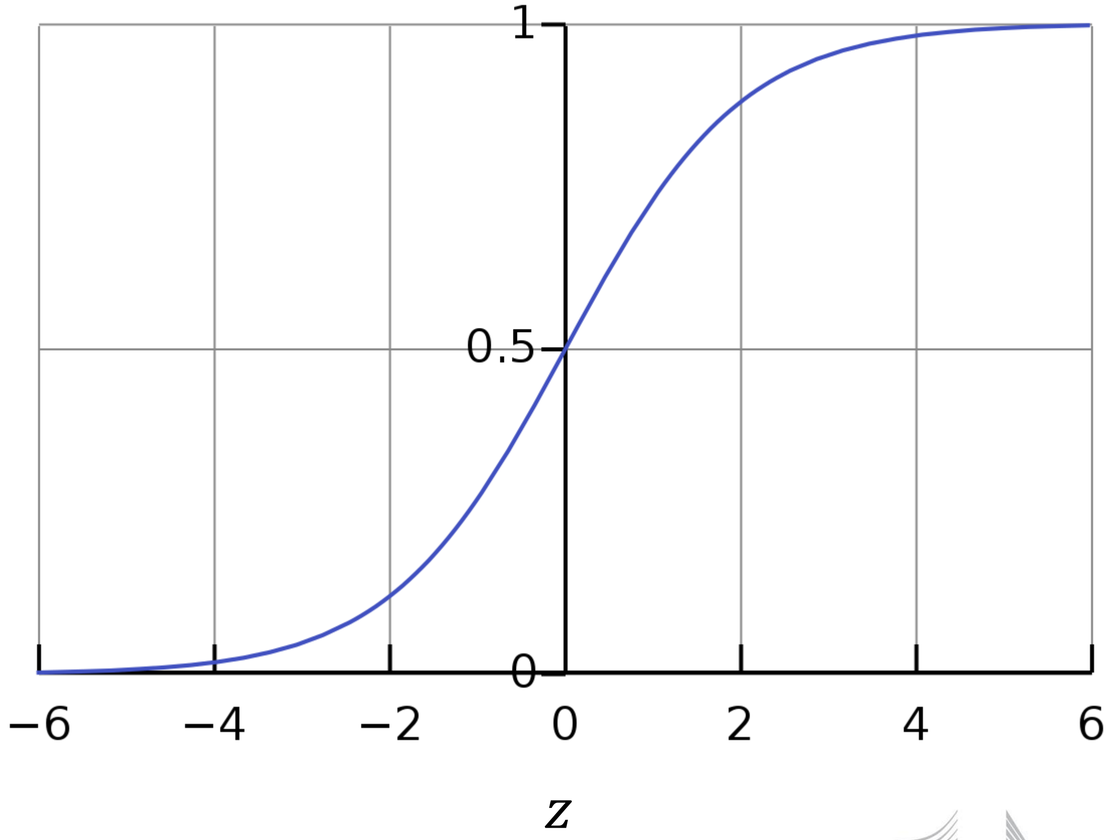
$$0 \leq f_{\boldsymbol{\theta}}(\boldsymbol{x}) \leq 1$$

$$f_{\boldsymbol{\theta}}(\boldsymbol{x}) = \boldsymbol{\theta}^T \boldsymbol{x}$$

$$f_{\boldsymbol{\theta}}(\boldsymbol{x}) = \frac{1}{(1 + e^{-\boldsymbol{\theta}^T \boldsymbol{x}})}$$

εκτιμώμενη πιθανότητα ότι «y=1» στην είσοδο x

Logistic function
ή
Sigmoid function



This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Κουίζ

Ας υποθέσουμε ότι θέλουμε να προβλέψουμε αν ένας όγκος είναι κακοήθης ή όχι με βάση το μέγεθός του.

Τι είναι το x (η είσοδος) και τι εννοούμε με το «y=1» και το «y=0»;

Χ: μέγεθος, «y=1»: ο όγκος είναι κακοήθης, «y=0»: ο όγκος δεν είναι κακοήθης

Ας υποθέσουμε τώρα ότι $f_{\boldsymbol{\theta}}(\boldsymbol{x}) = 0.8$

Τι σημαίνει αυτό;

Η εκτιμώμενη πιθανότητα ότι το x είναι ένας όγκος όπως προβλέπεται από το μοντέλο μας είναι 0,8.

Ομοίως, η εκτιμώμενη πιθανότητα ότι το x ΔΕΝ είναι όγκος είναι 0,2 (= 1-0,8).

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Όριο απόφασης

$$f_{\boldsymbol{\theta}}(\boldsymbol{x}) = g(\boldsymbol{\theta}^T \boldsymbol{x})$$

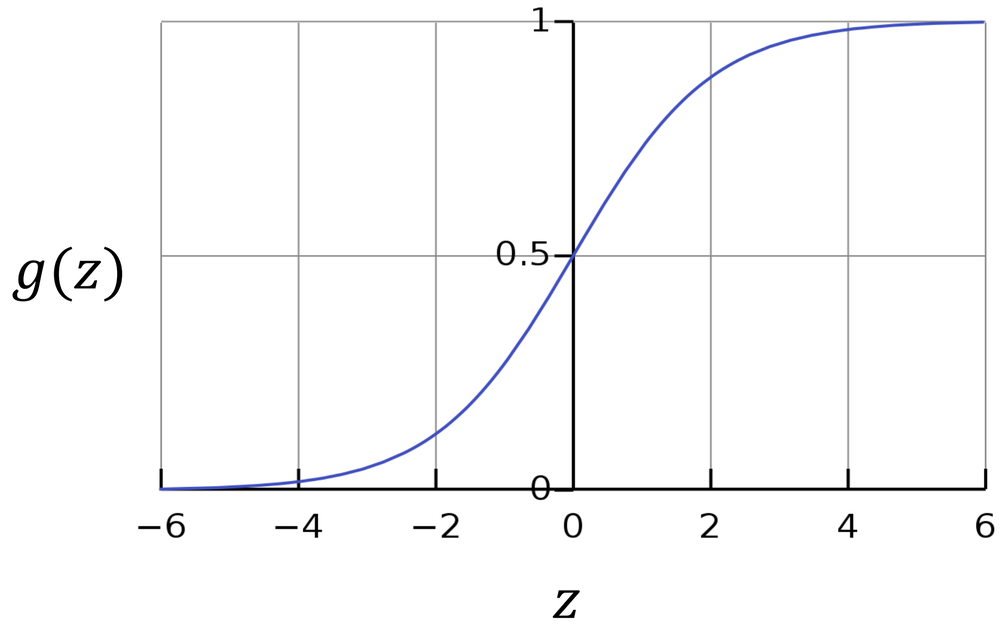
$$g(z) = \frac{1}{(1 + e^{-z})}$$

εάν $f_{\boldsymbol{\theta}}(\boldsymbol{x}) \geq 0.5$, πρόβλεψε το «y=1»

$$\boldsymbol{\theta}^T \boldsymbol{x} \geq 0$$

εάν $f_{\boldsymbol{\theta}}(\boldsymbol{x}) < 0.5$, πρόβλεψε «y=0»

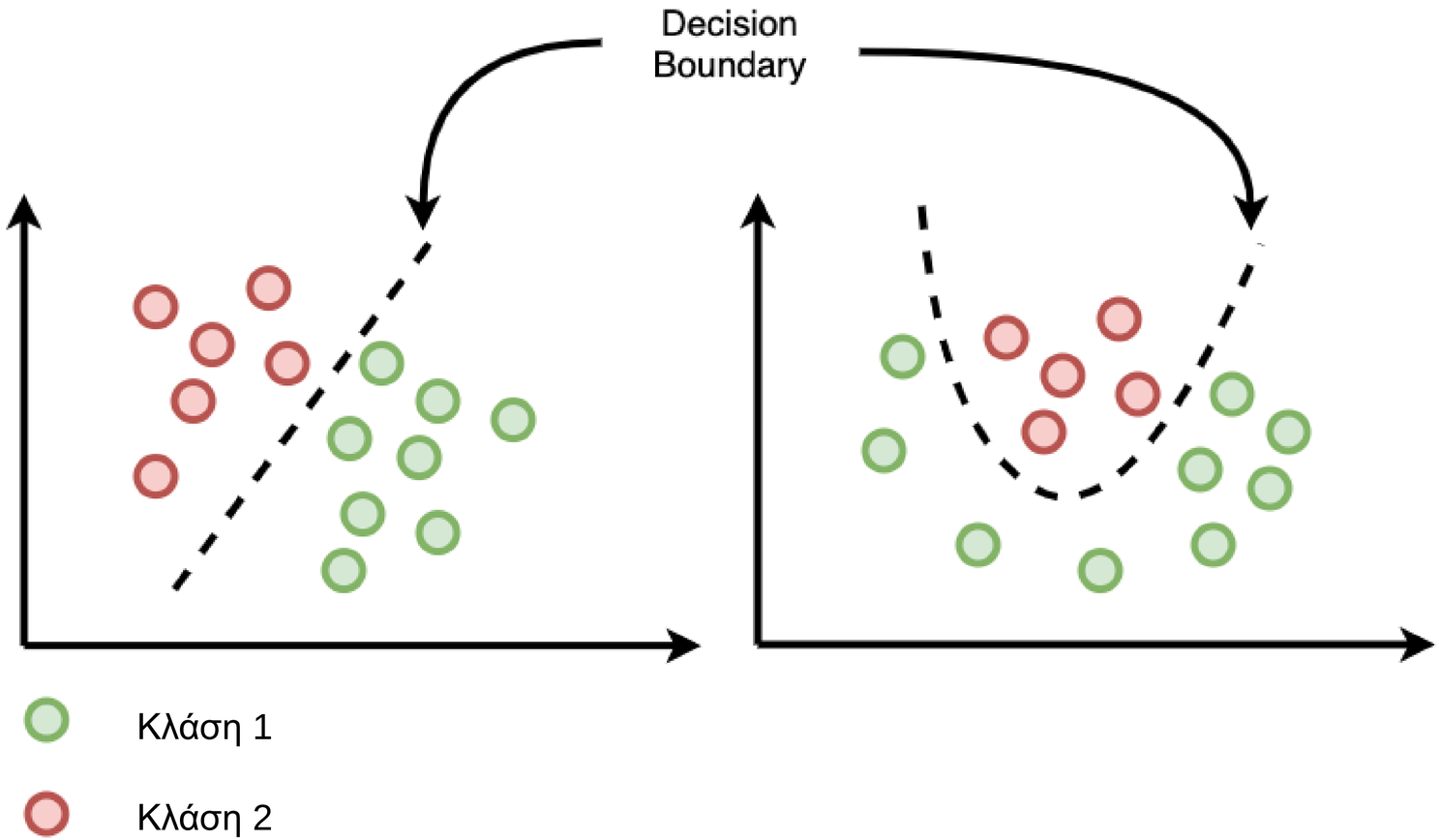
$$\boldsymbol{\theta}^T \boldsymbol{x} < 0$$



Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Όριο απόφασης



This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Συνάρτηση κόστους

Δεν ελαχιστοποιούμε το Μέσο Τετραγωνικό Σφάλμα

Αντίθετα, ελαχιστοποιούμε **το Σφάλμα Διασταυρούμενης Εντροπίας:**

- Διαισθητικά: απόσταση μεταξύ των κατανομών δύο πιθανοτήτων (στόχος και εκτιμώμενη)
- Το τετραγωνικό σφάλμα έχει ένα μη κυρτό τοπίο όταν χρησιμοποιείται με logistic regression
- Το σφάλμα ΔΕ έχει ένα κυρτό τοπίο

Για να **βρούμε το θ** χρησιμοποιούμε gradient descent (ή πιο προηγμένους αλγόριθμους βελτιστοποίησης, π.χ. conjugate gradient, L-BFGS).

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Συνάρτηση κόστους

Σφάλμα διασταυρούμενης εντροπίας

$$L(\boldsymbol{\theta}) = -\frac{1}{m}\sum_{i=1}^{m}[y^{(i)} log f_{\boldsymbol{\theta}}(x^{(i)}) + (1 - y^{(i)}) log(1 - f_{\boldsymbol{\theta}}(x^{(i)}))]$$

Gradient

$$\nabla_{\theta_j} L(\boldsymbol{\theta}) = \frac{1}{m}\sum_{i=1}^{m}[f_{\boldsymbol{\theta}}(x^{(i)}) - y^{(i)})]x_j^{(i)}$$

όπως και στην

Γραμμική παλινδρόμηση: $f_{\boldsymbol{\theta}}(\boldsymbol{x}) = \boldsymbol{\theta}^T \boldsymbol{x}$

Logistic Regression: $f_{\boldsymbol{\theta}}(\boldsymbol{x}) = \frac{1}{(1 + e^{-\boldsymbol{\theta}^T \boldsymbol{x}})}$

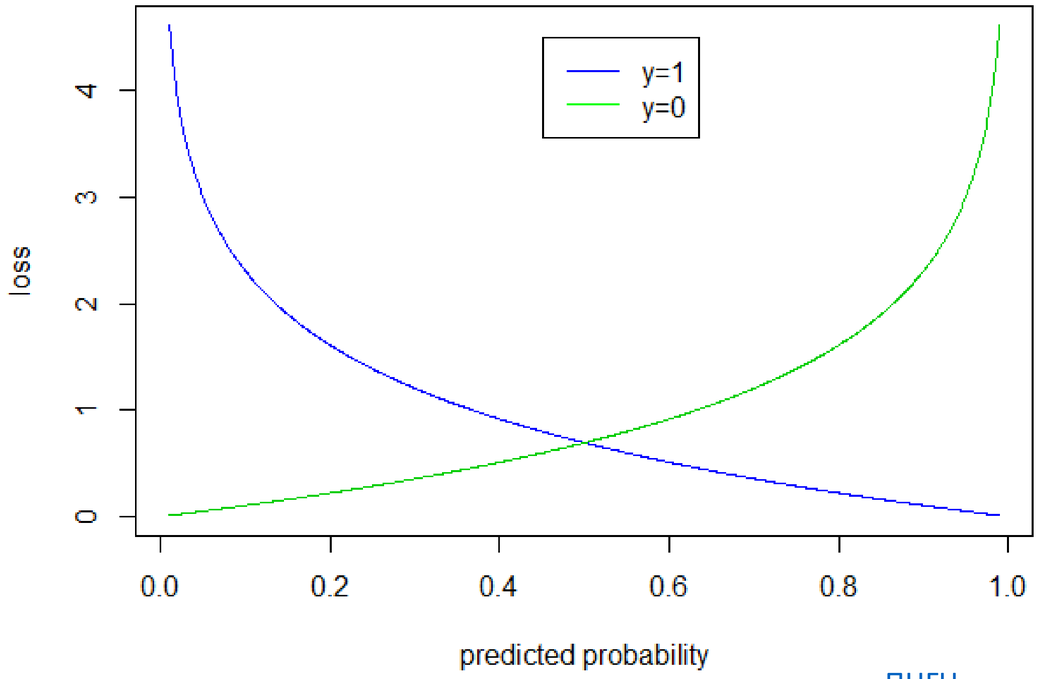
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Σφάλμα διασταυρούμενης εντροπίας

$$L(\boldsymbol{\theta}) = -\frac{1}{m}\sum_{i=1}^{m}[y^{(i)} log f_{\boldsymbol{\theta}}(x^{(i)}) + (1 - y^{(i)}) log(1 - f_{\boldsymbol{\theta}}(x^{(i)}))$$

$$L(\boldsymbol{\theta}) = \frac{1}{m}\sum_{i=1}^{m}[loss(f_{\boldsymbol{\theta}}(x^{(i)}), y^{(i)})]$$

$$loss(f_{\boldsymbol{\theta}}(x^{(i)}), y^{(i)}) = \begin{cases} -log(f_{\boldsymbol{\theta}}(x^{(i)})) & if\ y^{(i)} = 1 \\ -log(1 - f_{\boldsymbol{\theta}}(x^{(i)})) & if\ y^{(i)} = 0 \end{cases}$$



ΠΗΓΗ

Το κόστος αυξάνεται προς στο άπειρο, καθώς η προβλεπόμενη πιθανότητα είναι πολύ μακριά από την πραγματική

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Παράδειγμα

```
In [1]: X = [[0], [1], [2], [3]]
   ...: y = [0, 0, 1, 1]
   ...:
   ...: from sklearn.linear_model import LogisticRegression
   ...: clf = LogisticRegression(penalty='none').fit(X, y)

In [2]: print(clf.predict([[1.1]]))
[0]

In [3]: print(clf.predict_proba([[0.9]]))
[[9.99979559e-01 2.04410118e-05]]

In [4]: print(clf.score(X, y))
1.0
```


This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

MAI4CAREU

Master programmes in Artificial Intelligence 4 Careers in Europe

Ανάλυση σφαλμάτων

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Ανάλυση σφαλμάτων

Στη δυαδική ταξινόμηση μπορούμε να χαρακτηρίσουμε την απόδοση ενός ταξινομητή με βάση μια μετρική που ονομάζεται **ακρίβεια**

- αριθμός σωστών προβλέψεων διαιρούμενος με τον αριθμό των δειγμάτων στο σύνολο
- δίνει μια τιμή μεταξύ 0 και 1.

Ωστόσο, μπορούμε να προχωρήσουμε πέρα από την ακρίβεια ερευνώντας τους τύπους σφάλματος που κάνει ο ταξινομητής.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Ανάλυση σφαλμάτων

Παράδειγμα (πηγή): Λαμβάνοντας υπόψη ένα δείγμα 12 ατόμων, 8 που έχουν διαγνωστεί με καρκίνο και 4 που είναι χωρίς καρκίνο, όπου τα άτομα με καρκίνο ανήκουν στην κλάση 1 (θετικά) και τα άτομα χωρίς καρκίνο ανήκουν στην κλάση 0 (αρνητικά), μπορούμε να παρουσιάσουμε τα δεδομένα αυτά ως εξής:

Individual Number	1	2	3	4	5	6	7	8	9	10	11	12
Actual Classification	1	1	1	1	1	1	1	1	0	0	0	0

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Ανάλυση σφαλμάτων

Συνέχεια παραδείγματος: Ας υποθέσουμε ότι έχουμε έναν ταξινομητή που διακρίνει μεταξύ ατόμων με και χωρίς καρκίνο με κάποιο τρόπο, μπορούμε να πάρουμε τα 12 άτομα και να εκτελέσουμε τον ταξινομητή για αυτά. Στη συνέχεια, ο ταξινομητής κάνει 9 ακριβείς προβλέψεις και χάνει 3: 2 άτομα με καρκίνο τα οποόα προβλέπεται λανθασμένα ότι δεν έχουν καρκίνο (δείγμα 1 και 2) και 1 άτομο χωρίς καρκίνο που λανθασμένα προβλέπεται να έχει καρκίνο (δείγμα 9).

Individual Number	1	2	3	4	5	6	7	8	9	10	11	12
Actual Classification	1	1	1	1	1	1	1	1	0	0	0	0
Predicted Classification	0	0	1	1	1	1	1	1	1	0	0	0

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Ανάλυση σφαλμάτων

Συνέχεια παραδείγματος :

Individual Number	1	2	3	4	5	6	7	8	9	10	11	12
Actual Classification	1	1	1	1	1	1	1	1	0	0	0	0
Predicted Classification	0	0	1	1	1	1	1	1	1	0	0	0
Result	FN	FN	TP	TP	TP	TP	TP	TP	FP	TN	TN	TN

Αληθώς θετικό (TP): Θετική πραγματική ταξινόμηση, θετική προβλεπόμενη ταξινόμηση

Αληθώς αρνητικό (TN): Αρνητική πραγματική ταξινόμηση, αρνητική προβλεπόμενη ταξινόμηση

Ψευδώς θετικό (FP): Αρνητική πραγματική ταξινόμηση, θετική προβλεπόμενη ταξινόμηση

Ψευδώς αρνητικό (FN): Θετική πραγματική ταξινόμηση, αρνητική προβλεπόμενη ταξινόμηση

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Πίνακας σύγχυσης

		Predicted condition	
	Total population = P + N	Positive (PP)	Negative (PN)
Actual condition	Positive (P)	True positive (TP)	False negative (FN)
	Negative (N)	False positive (FP)	True negative (TN)

		Predicted condition	
	Total 8 + 4 = 12	Cancer 7	Non-cancer 5
Actual condition	Cancer 8	6	2
	Non-cancer 4	1	3

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Μετρικές

Accuracy = (ΑΘ + ΑΑ)/(Θ + Α)

Precision = ΑΘ/(ΑΘ + ΨΑ)

Αναλογία Πραγματικών θετικών = ΑΘ/(ΑΘ + ΨΑ)
- Επίσης γνωστό ως: **ανάκληση**, ευαισθησία, ρυθμός πρόσκρουσης ή πιθανότητα ανίχνευσης

Αναλογία Ψευδώς θετικών = ΨΘ/(ΨΘ + ΑΑ)
- Επίσης γνωστό ως: πιθανότητα ψευδούς συναγερμού

F1-score (αρμονικός μέσος όρος precision και ευαισθησίας) = 2ΑΑ/(2ΑΑ + ΨΘ + ΨΑ)

		Predicted condition	
	Total population = P + N	Positive (PP)	Negative (PN)
Actual condition	Positive (P)	True positive (TP)	False negative (FN)
	Negative (N)	False positive (FP)	True negative (TN)

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

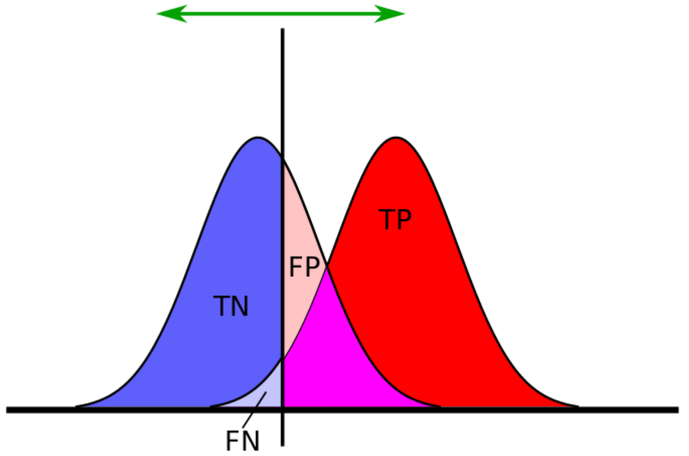
Καμπύλη ROC

Ένας δυαδικός ταξινομητής χρειάζεται ένα κατώτατο όριο για να αποφασίσει πού να θέσει το όριο απόφασης

Η επιλογή κατωφλίου επηρεάζει τόσο τον ΑΑΘ όσο και τον ΑΨΑ

Μπορούμε να διαφοροποιήσουμε τα όρια και να σχεδιάσουμε το ΑΑΘ σε σχέση με το ΑΨΑ

Αυτό είναι γνωστό ως Receiver operating characteristic (ROC)



TP	FP
FN	TN

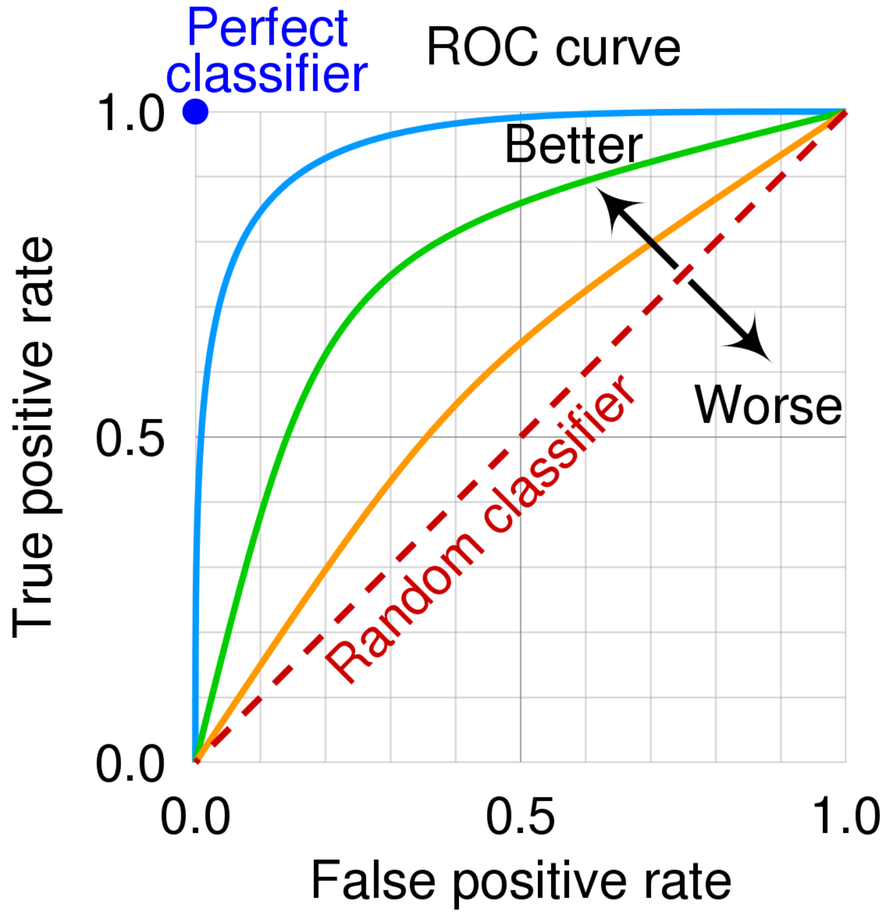
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Καμπύλη ROC

- **Καμπύλη ROC**: γράφημα που δείχνει τις επιδόσεις ενός δυαδικού μοντέλου ταξινόμησης σε **όλα τα** κατώτατα όρια ταξινόμησης.

- Η μείωση του κατώτατου ορίου ταξινόμησης ταξινομεί περισσότερα στοιχεία ως θετικά, αυξάνοντας έτσι τόσο τα ΨΘ όσο και τις ΑΘ.

- **Σκοπός**:
  - Ανάλυση της δύναμης/προβλεπόμενης ισχύος ενός ταξινομητή
  - Σύγκριση ταξινομητών
  - Καθορισμός του «βέλτιστου» ορίου με βάση τις προτιμήσεις του χρήστη



ΠΗΓΗ

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

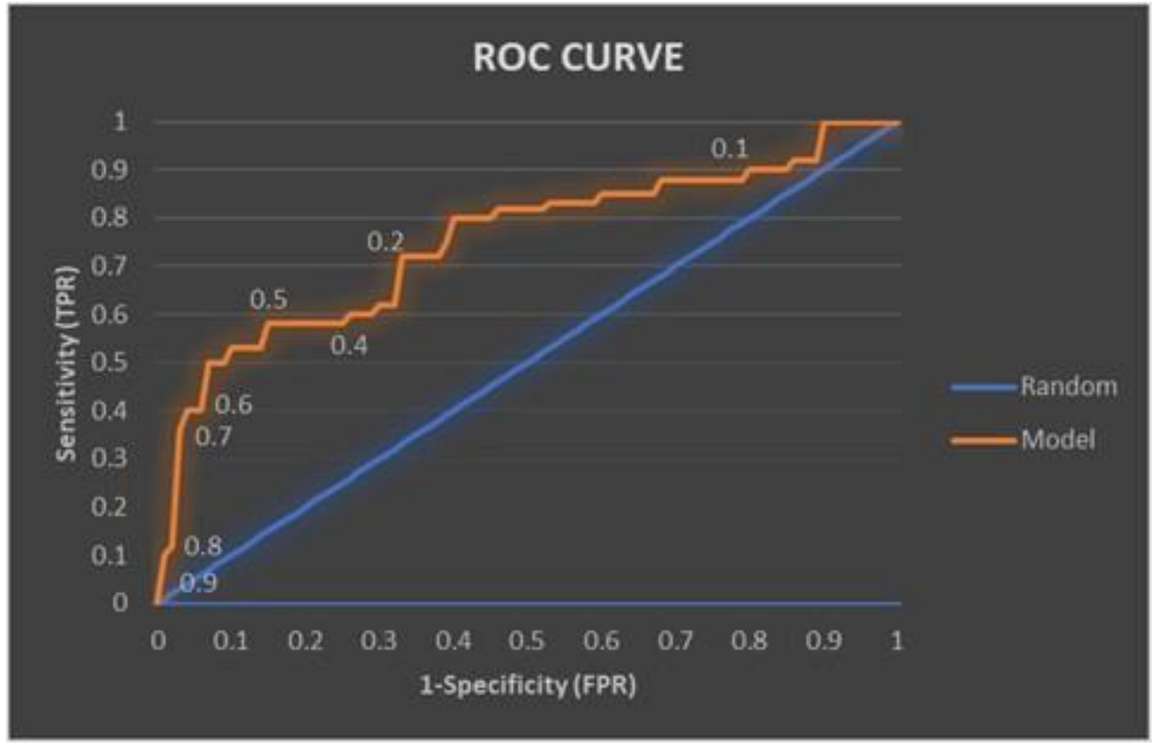
Καμπύλη ROC: Είναι ένας συμβιβασμός.

Ποιο κατώφλι θα επέλεγες;

- Αν ενδιαφέρεστε περισσότερο για το ΑΨΑ ≤ 0.1 (αποκτήθηκε με κατώτατα όρια 0,9, 0,8, 0.7 και 0,6)

- Αν σας ενδιαφέρει περισσότερο το υψηλό ΑΠΘ (≥ 0.7)

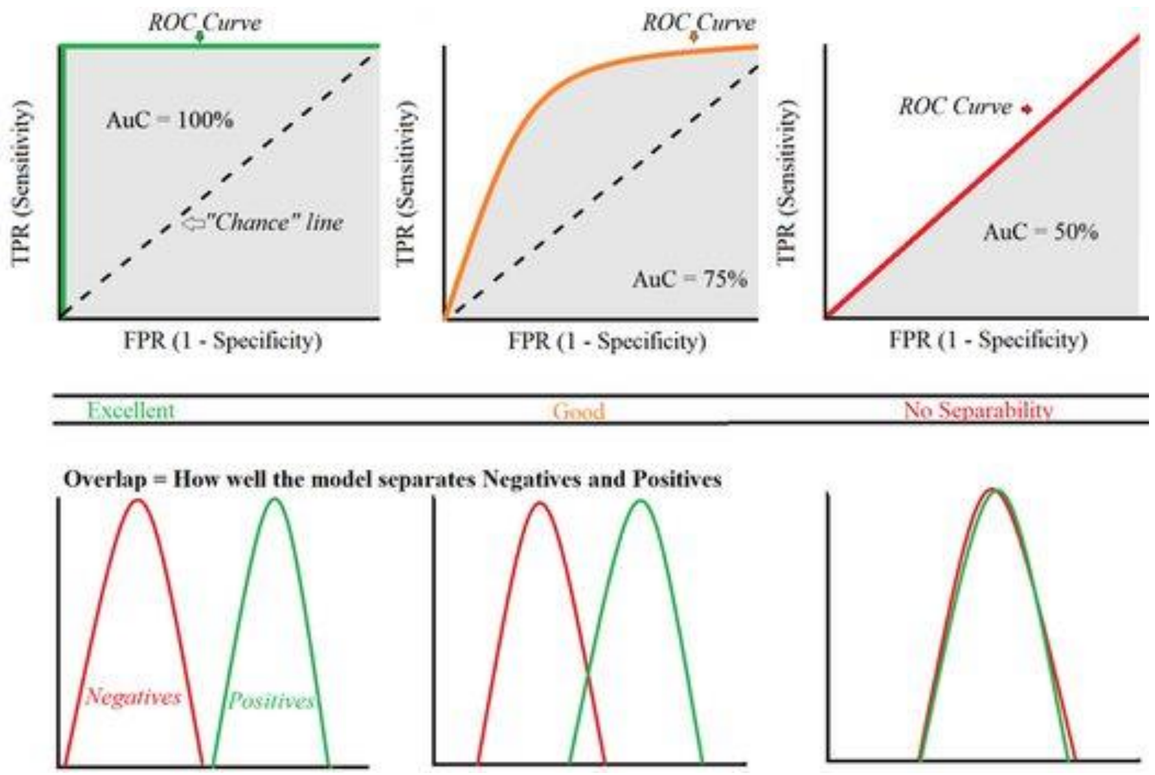
- Αν είστε ευχαριστημένοι με μια ΑΠΘ ≈ 0.6 (αποκτήθηκε με κατώφλια 0.5 και 0.4)



ΠΗΓΗ

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Καμπύλη ROC: Σύγκριση ταξινομητών



ΠΗΓΗ

Περιοχή κάτω από την καμπύλη (ΠΚΚ)

- ενιαία μέτρηση (από 0 έως 1)
- συνολική μέτρηση της απόδοσης σε όλα τα πιθανά κατώτατα όρια ταξινόμησης
- μέτρηση του κατά πόσο το μοντέλο είναι ικανό να διακρίνει τις δύο κατηγορίες

- **Αμετάβλητη Κλίμακα**: μετράει πόσο καλά κατατάσσονται οι προβλέψεις και όχι οι απόλυτες τιμές τους
- **Αμετάβλητο Ταξινόμηση-κατώφλι**: μέτρηση της ποιότητας των προβλέψεων του μοντέλου ανεξάρτητα από το κατώτατο όριο ταξινόμησης

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Κουίζ

Ας υποθέσουμε ότι έχετε έναν ταξινομητή που προβλέπει όλα τα θετικά ως αρνητικά και όλα τα αρνητικά ως θετικά.

1. **Είναι αυτός ο ταξινομητής χειρότερος από τον τυχαίο ταξινομητή;**

   ΝΑΙ

2. **Ποια είναι η βαθμολογία της ΠΚΚ;**

   0

3. **Μπορούμε να το διορθώσουμε κάπως;**

   Ναι. Απλά αναπογυρίζουμε τις προβλέψεις του και έχουμε ένα τέλειο ταξινομητή!

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

MAI4CAREU

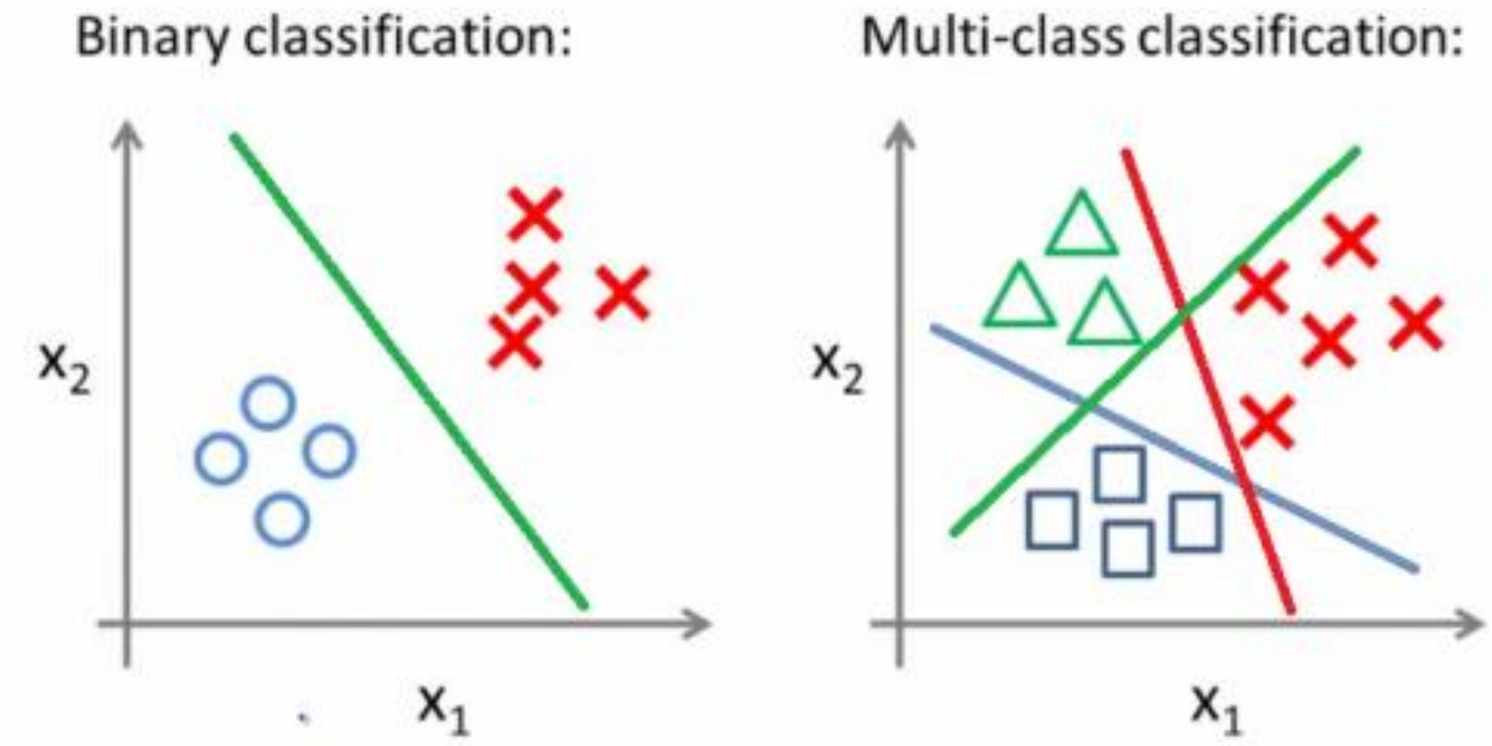
Master programmes in Artificial Intelligence 4 Careers in Europe

Ταξινόμηση πολλαπλών κλάσεων

Co-financed by the European Union
Connecting Europe Facility

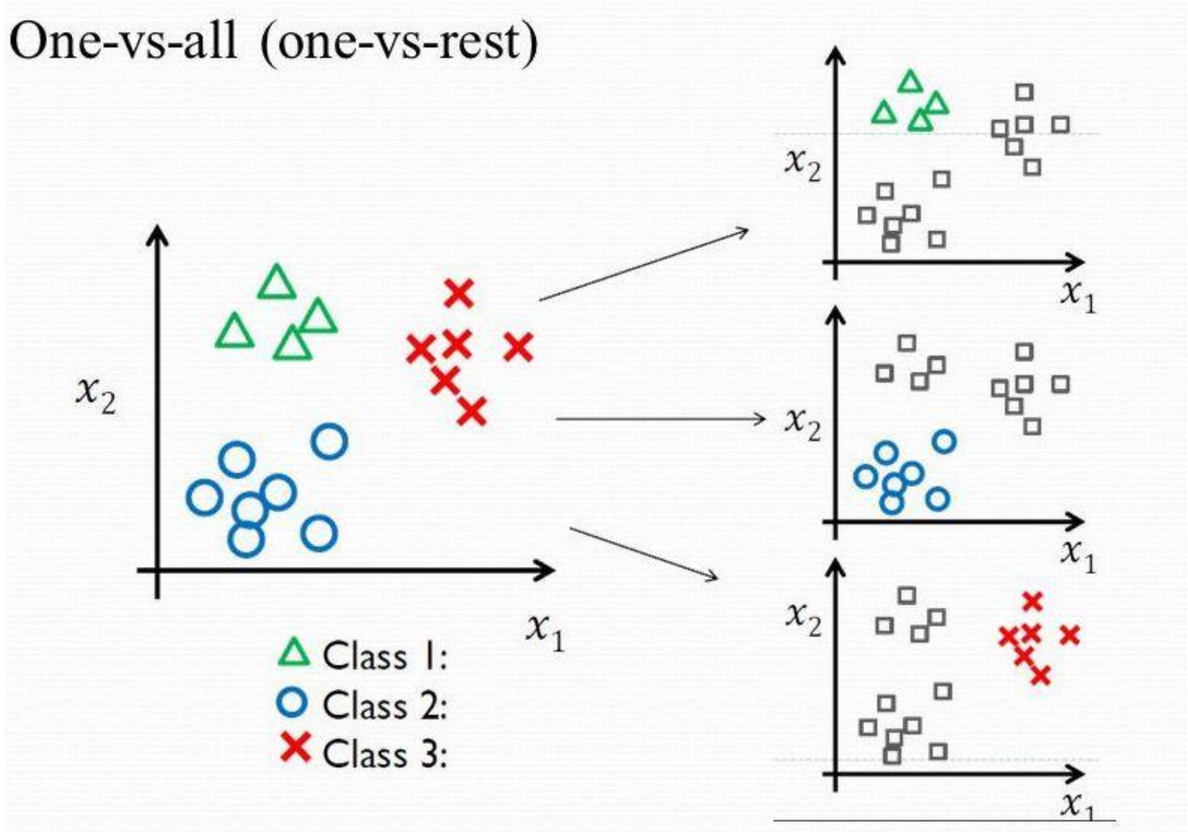
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Ταξινόμηση πολλαπλών κλάσεων



This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Ταξινόμηση πολλαπλών κλάσεων



Μετατροπή του προβλήματος σε K δυαδικά προβλήματα ταξινόμησης (K αριθμός κλάσεων)

Εκπαιδεύστε έναν ταξινομητή logistic regression για κάθε κατηγορία.

Για να κάνετε μια πρόβλεψη επιλέξτε την τάξη της οποίας ο ταξινομητής είναι πιο σίγουρος.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Softmax ταξινομητής

- Φυσική γενίκευση του δυαδικού ταξινομητή logistic regression σε πολλαπλές κλάσεις
- Ονομάζεται επίσης πολυωνυμική logistic regression
- **Έξοδος**: κανονικοποιημένες πιθανότητες κλάσης

- **Συνάρτηση softmax:** μετατρέπει ένα διάνυσμα K πραγματικών αριθμών σε μια κατανομή πιθανότητας των K πιθανών αποτελεσμάτων.

$$\sigma(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \quad \text{for } i = 1, \ldots, K \text{ and } \mathbf{z} = (z_1, \ldots, z_K) \in \mathbb{R}^K.$$

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Σφάλμα διασταυρούμενης εντροπίας για πολλαπλές κλάσεις

- Κατηγορικό Σφάλμα Διασταυρούμενης Εντροπίας

- Κάθε $\boldsymbol{y}^{(i)}$ είναι one-hot encoded για τις κλάσεις K
- Η εκτιμώμενη $\hat{\boldsymbol{y}}^{(i)}$ είναι η έξοδος του softmax ταξινομητή

$$loss(\hat{\boldsymbol{y}}^{(i)}, \boldsymbol{y}^{(i)}) = -\sum_{c=1}^{K}[y_c^{(i)}\ log\ \hat{y}_c^{(i)}]$$

$$L(\boldsymbol{\theta}) = -\frac{1}{m}\sum_{i=1}^{m}\sum_{c=1}^{K}[y_c^{(i)}\ log\ \hat{y}_c^{(i)}]$$

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Softmax ταξινομητής: ζητήματα αριθμητικής σταθερότητας

- Κατά τη γραφή κώδικα για τον υπολογισμό της συνάρτησης Softmax, ο όρος e^{z_j} και $\sum_{c=1}^{K} e^{z_c}$ μπορεί να είναι πολύ μεγάλος λόγω των εκθετικών.
- Η διαίρεση μεγάλων αριθμών μπορεί να είναι αριθμητικά ασταθής, γι' αυτό είναι σημαντικό να χρησιμοποιηθεί ένα κόλπο εξομάλυνσης.

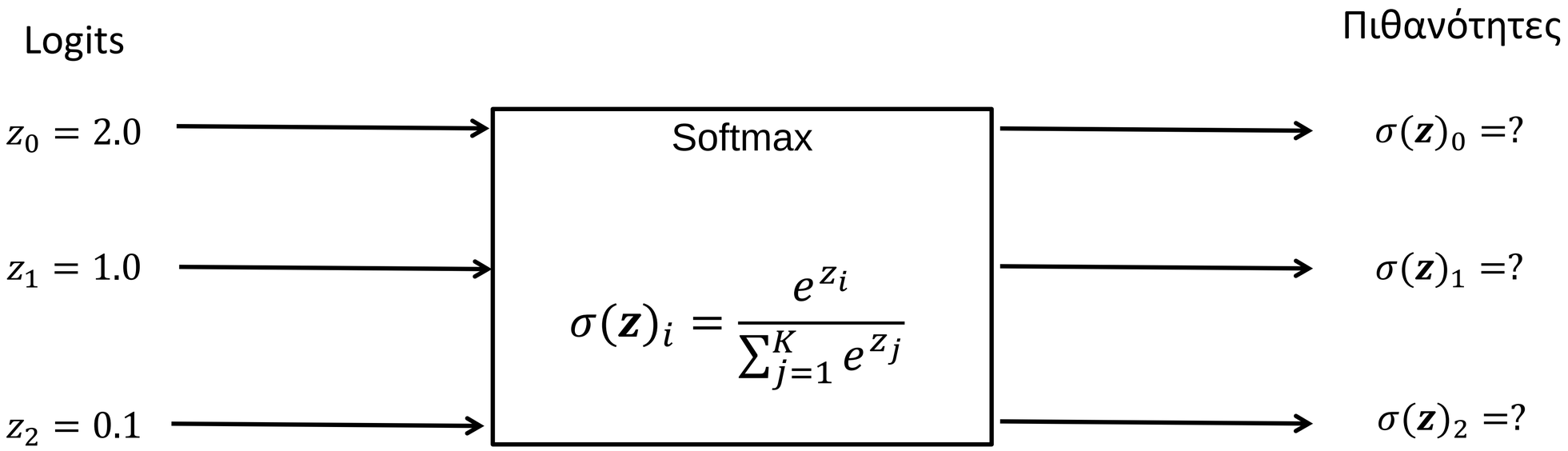
$$\frac{e^{z_j}}{\sum_{c=1}^{K} e^{z_c}} = \frac{A\, e^{z_j}}{A \sum_{c=1}^{K} e^{z_c}} = \frac{e^{z_j + logA}}{\sum_{c=1}^{K} e^{z_c + logA}}$$

$$logA = -max_c z_c$$

Μετατοπίζει τις τιμές z_j έτσι ώστε το υψηλότερο να είναι 0

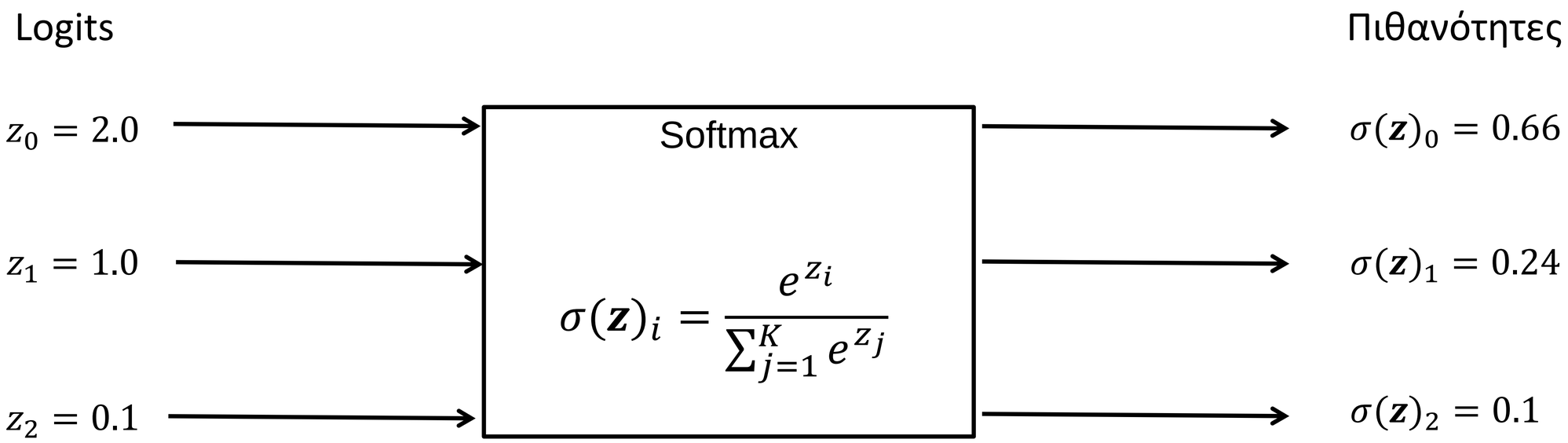
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Ταξινομητής Softmax: παράδειγμα



This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Ταξινομητής Softmax: παράδειγμα



This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Επόμενη Διάλεξη

- Αξιολόγιση και βελτίωση μοντέλου
- Δέντρα και δάση

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

MAI4CAREU

Master programmes in Artificial Intelligence 4 Careers in Europe

Σας ευχαριστούμε

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423