



ΓΛΩΣΣΙΚΗ ΤΕΧΝΟΛΟΓΙΑ

Κάτια Κερμανίδου

kerman@ionio.gr

Αποσπάσματα κάποιων διαφανειών είναι δανεισμένα από τα
slides του Speech and Language Processing

ΛΟΓΙΣΤΙΚΗ
ΠΑΛΙΝΔΡΟΜΗΣΗ

LOGISTIC
REGRESSION

Η ταξινόμηση κειμένου (text classification)

Πολλές εφαρμογές ΓΤ μπορούν να μοντελοποιηθούν σαν απόδοση μιας ετικέτας σε ένα κείμενο.

- ❑ Κατηγοριοποίηση ενός κειμένου σε θεματικές κατηγορίες
- ❑ Ανίχνευση spam
- ❑ Αναγνώριση συγγραφέα (Authorship Attribution)
- ❑ Ανάλυση Συναισθήματος (Sentiment Analysis)
- ❑ Αναγνώριση Γλώσσας (Language Identification)

Who wrote which *Federalist Papers*?

1787-8: essays anonymously written by:

Alexander Hamilton, James Madison, and John Jay

to convince New York to ratify U.S Constitution

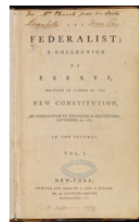
Authorship of 12 of the letters unclear between:



Alexander Hamilton



James Madison



1963: solved by Mosteller and Wallace using Bayesian methods

Is this spam?

Dear Sir,

Please find attached the forms you requested.

Special Note: Student needs to provide the hand signed forms on the very first day of registration- (Forms attached)

IMPORTANT NOTE:

We will be needing additional bank statements and proof for justification if students have completed education outside their home state. (Legitimate documents- electricity bill, Rental Agreement, Bank Statement, College/University ID)

Please feel free to reach us in case any queries arise.

Study Abroad Partners

Η Ταξινόμηση Κειμένου ως Εργασία Επιβλεπόμενης Μάθησης

Εκπαίδευση (Training)

Είσοδος: ένα σετ N κειμένων $(x_1, x_2, x_3, \dots, x_N)$ επισημειωμένο ως προς την τιμή τη εξόδου $(y_1, y_2, y_3, \dots, y_N)$

Η μεταβλητή εξόδου παίρνει μια τιμή από ένα σετ πεπερασμένων τιμών.

Ένας ταξινομητής μαθαίνει να ταιριάζει (map) την είσοδο με την μεταβλητή εξόδου.

Ο ταξινομητής μπορεί να είναι

- ☐ Λογιστική Παλινδρόμηση
- ☐ K-NN
- ☐ Δέντρα Απόφασης
- ☐ Νευρωνικά Δίκτυα
- ☐ LLMs,
 - ☐ Είτε με προσαρμογή σε δικά μου δεδομένα (fine-tuning)
 - ☐ Είτε με prompting

Δοκιμή (Testing)

Είσοδος: Ένα καινούριο κείμενο

Έξοδος: Μια τιμή της μεταβλητής εξόδου

Η Λογιστική Παλινδρόμηση

Είσοδος:

Κάθε κείμενο αναπρίσταται μέσα από ένα σύνολο χαρακτηριστικών $x = [x_1, x_2, \dots, x_n]$
Κάθε χαρακτηριστικό έχει ένα βάρος $W = [w_1, w_2, \dots, w_n]$ – αυτά υπολογίζονται κατά την εκπαίδευση

Έξοδος:

Στην Δυαδική (Binary) Λογιστική Παλινδρόμηση

Η έξοδος είναι 0 ή 1, $y = \{0, 1\}$

Στην Πολυωνυμική (Multinomial) Λογιστική Παλινδρόμηση

Η έξοδος είναι μια από ένα σετ πεπερασμένων τιμών $y = \{0, 1, 2, 3, 4\}$

Πώς γίνεται η ταξινόμηση;

Έρχεται ένα παράδειγμα προς ταξινόμηση.

Πολλαπλασιάζω κάθε χαρακτηριστικό (x_i) του παραδείγματος με το βάρος του (w_i), και προσθέτω όλα αυτά τα γινόμενα $\sum w_i x_i$. Επίσης προσθέτω ένα επιπλέον βάρος b , την πόλωση (bias).

Αν το αποτέλεσμα $z = \sum w_i x_i + b$ είναι μεγάλο, τότε λέμε $y = 1$, αν είναι μικρό τότε $y = 0$.

Τι θα πει «μεγάλο» και «μικρό» άθροισμα;

□ Θέλουμε να μετατρέψουμε την διαδικασία σε **πιθανοτική**, δηλαδή θέλουμε ένα μοντέλο το οποίο θα μπορεί να μας πει ποια είναι η πιθανότητα η έξοδος του παραδείγματός μου να πάρει τιμή 0 και ποια τιμή 1;

□ δεδομένων των συγκεκριμένων χαρακτηριστικών εισόδου που έχει το παράδειγμα και

□ του βάρους που έχω υπολογίσει για το κάθε χαρακτηριστικό από την εκπαίδευση

$$p(y=1 \mid x; w) = ?$$

$$p(y=0 \mid x; w) = ?$$

Το προηγούμενο άθροισμα z είναι ένας αριθμός, δεν είναι πιθανότητα.

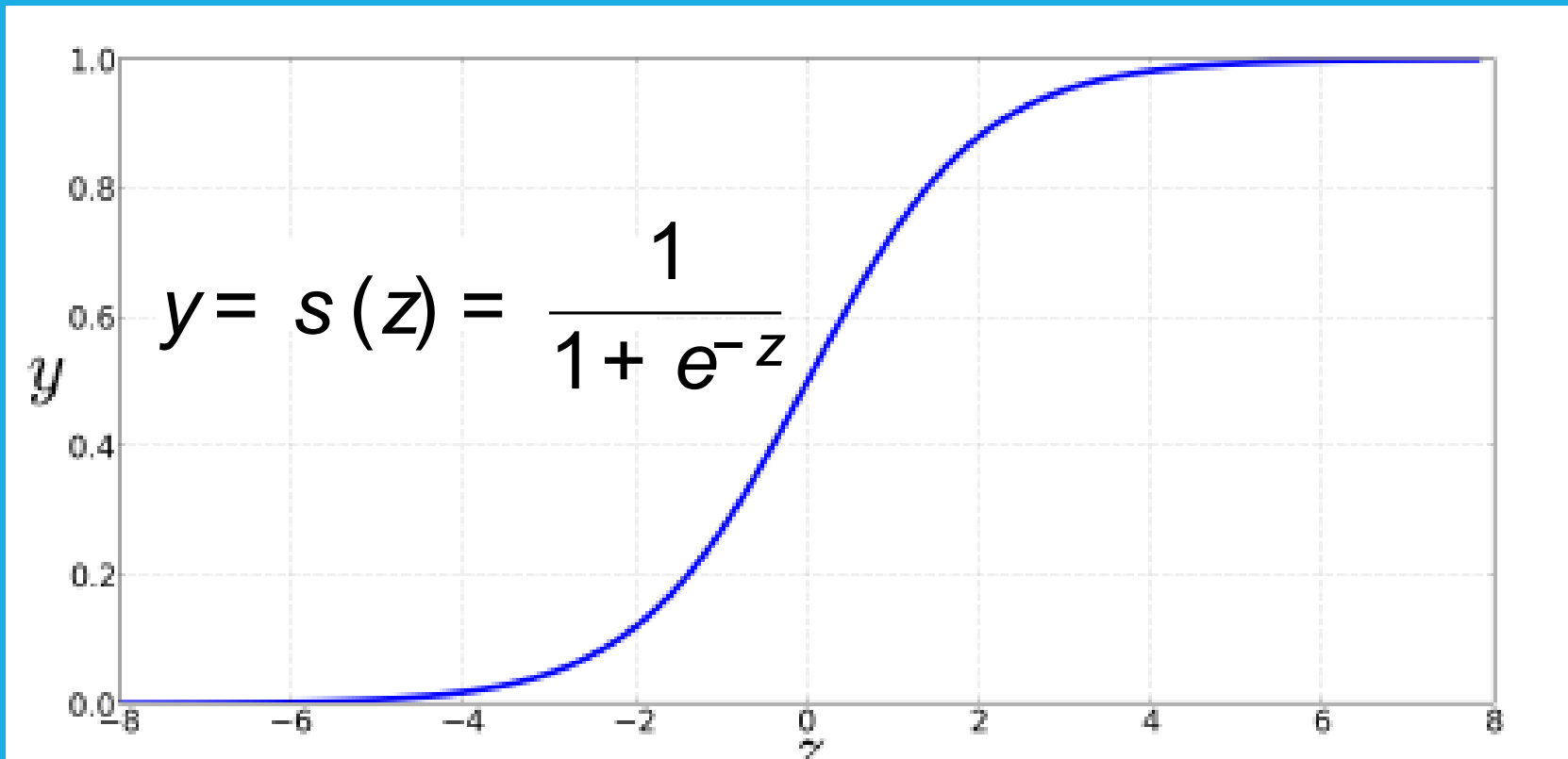
Για να τον μετατρέψω σε πιθανότητα χρειαζομαι μια συνάρτηση που θα παίρνει σαν όρισμα το z και θα βγάζει τιμές στο πεδίο $[0,1]$.

Τέτοια συνάρτηση είναι η σιγμοειδής (sigmoid):

$$y = s(z) = \frac{1}{1 + e^{-z}} = \frac{1}{1 + \exp(-z)}$$

Η σιγμοειδής ή Λογιστική Συνάρτηση (Sigmoid)

Σχεδόν γραμμική για εισόδους κοντά στο 0, αλλά τις μεγαλύτερες ή τις μικρότερες εισόδους τις στριμώνει στο 1 και το 0 αντίστοιχα.



Πιθανότητες με τη Σιγμοειδή

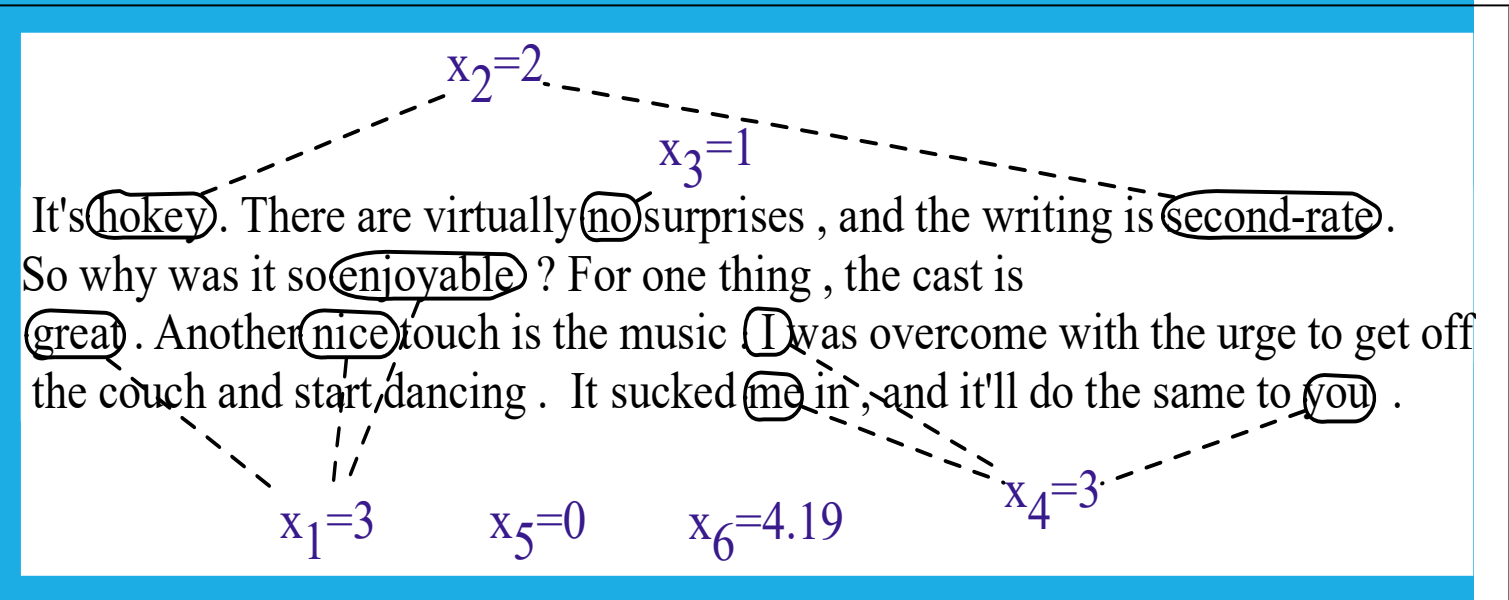
$$\begin{aligned} P(y = 1) &= \sigma(w \cdot x + b) \\ &= \frac{1}{1 + \exp(-(w \cdot x + b))} \end{aligned}$$

$$\begin{aligned} P(y = 0) &= 1 - \sigma(w \cdot x + b) \\ &= 1 - \frac{1}{1 + \exp(-(w \cdot x + b))} \\ &= \frac{\exp(-(w \cdot x + b))}{1 + \exp(-(w \cdot x + b))} \end{aligned}$$

Παράδειγμα στην Ανάλυση Συναισθήματος

- Έστω ότι έχετε το παρακάτω κείμενο να το ταξινομήσετε ως θετικό ($\gamma=1$) ή ως αρνητικό ($\gamma=0$).
- *It's hokey. There are virtually no surprises, and the writing is second-rate. So why was it so enjoyable? For one thing, the cast is great. Another nice touch is the music. I was overcome with the urge to get off the couch and start dancing. It sucked me in, and it'll do the same to you.*

Παράδειγμα στην Ανάλυση Συναισθήματος



Var	Definition	Value in Fig. 5.2
x_1	count(positive lexicon) $\in$ doc)	3
x_2	count(negative lexicon) $\in$ doc)	2
x_3	$\begin{cases} 1 & \text{if "no" } \in \text{ doc} \\ 0 & \text{otherwise} \end{cases}$	1
x_4	count(1st and 2nd pronouns $\in$ doc)	3
x_5	$\begin{cases} 1 & \text{if "!" } \in \text{ doc} \\ 0 & \text{otherwise} \end{cases}$	0
x_6	log(word count of doc)	$\ln(66) = 4.19$

Ας υποθέσουμε τα εξής βάρη, και
πόλωση

$$\mathbf{W} = [2.5, -5.0, -1.2, 0.5, 2.0, 0.7]$$

$$b = 0.1$$

Πράξεις για την ταξινόμηση

$$\begin{aligned} p(+|x) = P(Y=1|x) &= s(w \cdot x + b) \\ &= s([2.5, -5.0, -1.2, 0.5, 2.0, 0.7] \cdot [3, 2, 1, 3, 0, 4.19] + 0.1) \\ &= s(.833) \\ &= 0.70 \end{aligned}$$

$$\begin{aligned} p(-|x) = P(Y=0|x) &= 1 - s(w \cdot x + b) \\ &= 0.30 \end{aligned}$$

Εκπαίδευση: Πώς υπολογίζονται τα βάρη w ;

- ❑ Η διαδικασία εκπαίδευσης είναι επιβλεπόμενη, άρα για κάθε παράδειγμα ξέρω την πραγματική έξοδο ($y=1$ ή $y=0$).
- ❑ Ο αλγόριθμος για κάθε παράδειγμα θα παράξει μια εκτίμηση y' .
- ❑ Θέλουμε τα βάρη w και η πόλωση b να είναι τέτοια που να ελαχιστοποιούν την απόσταση ανάμεσα στο y και το y' , δηλ. Το σφάλμα.
- ❑ Χρειαζόμαστε
 - ❑ Μια συνάρτηση που θα μετράει αυτή την απόσταση (**loss function ή cost function**) - $L(y', y)$
 - ❑ Έναν αλγόριθμο που θα μεταβάλει (ανανεώνει) τα βάρη και την πόλωση ώστε να ελαχιστοποιείται η παραπάνω συνάρτηση – **optimization algorithm**

Μια συνάρτηση που μετράει αυτή την απόσταση (**loss function ή cost function**): **Cross-Entropy Loss**

- CE Loss = $L(y', y) = -\log p(y | x)$, όπου $p(y | x)$ είναι η πιθανότητα της σωστής ταξινόμησης
- $p(y | x) = y'^y (1-y')^{1-y}$
- $\log p(y | x) = \log [y'^y (1-y')^{1-y}] = y \log y' + (1-y) \log(1-y')$
- CE Loss = $-[y \log y' + (1-y) \log(1-y')]$

It's hokey. There are virtually no surprises, and the writing is second-rate. So why was it so enjoyable? For one thing, the cast is great. Another nice touch is the music. I was overcome with the urge to get off the couch and start dancing. It sucked me in, and it'll do the same to you.

$$\begin{aligned} L_{CE}(\hat{y}, y) &= -[y \log \sigma(w \cdot x + b) + (1 - y) \log (1 - \sigma(w \cdot x + b))] \\ &= -[\log \sigma(w \cdot x + b)] \\ &= -\log(.70) \\ &= .36 \end{aligned}$$

y=1 (θετικό)

$$\begin{aligned} L_{CE}(\hat{y}, y) &= -[y \log \sigma(w \cdot x + b) + (1 - y) \log (1 - \sigma(w \cdot x + b))] \\ &= -[\log (1 - \sigma(w \cdot x + b))] \\ &= -\log (.30) \\ &= 1.2 \end{aligned}$$

y=0 (αρνητικό)

➤ **Πράγματι το σφάλμα είναι μεγαλύτερο αν θεωρήσω ότι το παραπάνω κείμενο είναι αρνητικό, και άρα το μοντέλο κάνει λάθος!**

Optimization Algorithm - Gradient Descent

- Τα βάρη και η πόλωση είναι οι παράμετροί μου, και τα ονομάζω μαζί θ
- Θέλω εκείνα τα βάρη που ελαχιστοποιούν την συνάρτηση σφάλματος

$$\hat{\theta} = \operatorname{argmin}_{\theta} \frac{1}{m} \sum_{i=1}^m L_{\text{CE}}(f(x^{(i)}; \theta), y^{(i)})$$

$$w^{t+1} = w^t - h \frac{d}{dw} L(f(x; w), y)$$

Learning rate:

Αν είναι μεγάλο, το w μεταβάλλεται πολύ σε κάθε βήμα, αν είναι μικρό το w μεταβάλλεται λίγο

Η παράγωγος της συνάρτησης σφάλματος ως προς w

Optimization Algorithm - Gradient Descent

- Στην περίπτωση που έχω για Loss Function την Cross-entropy:

$$L_{\text{CE}}(\hat{y}, y) = -[y \log \sigma(w \cdot x + b) + (1 - y) \log (1 - \sigma(w \cdot x + b))]$$

- Η παράγωγος ως προς τα βάρη γίνεται

$$\frac{\partial L_{\text{CE}}(\hat{y}, y)}{\partial w_j} = [\sigma(w \cdot x + b) - y]x_j$$