

Τεχνητή Νοημοσύνη

7η διάλεξη (2025-26)

Ίων Ανδρουτσόπουλος

<http://www.aueb.gr/users/ion/>

Οι διαφάνειες αυτής της διάλεξης βασίζονται σε ύλη των βιβλίων (α) *Artificial Intelligence – A Modern Approach* των S. Russel και P. Norvig, 2^η και 4^η έκδοση, Prentice Hall, 2003 και 2020, (β) *Τεχνητή Νοημοσύνη των Βλαχάβα κ.ά.*, 3^η έκδοση, Β. Γκιούρδας Εκδοτική, 2006 και (γ) *Machine Learning* του T. Mitchell, McGraw-Hill, 1997. Τα περισσότερα σχήματα των διαφανειών βασίζονται σε αντίστοιχα σχήματα των διαφανειών που συνοδεύουν τα πρώτα δύο βιβλία.

Τι θα ακούσετε σήμερα

- Είδη συλλογιστικής
 - Παραγωγική, απαγωγική, επαγωγική συλλογιστική.
- Μηχανική μάθηση
 - Εισαγωγή και είδη μηχανικής μάθησης.
 - Παράσταση δεδομένων και υποθέσεων.
 - Η μηχανική μάθηση ως πρόβλημα αναζήτησης.
- Προβλήματα κατάταξης κειμένων σε κατηγορίες.
- Εντροπία και κέρδος πληροφορίας.
- Επιλογή ιδιοτήτων μέσω κέρδους πληροφορίας και μέσω αναζήτησης υποσυνόλων ιδιοτήτων.

Είδη συλλογιστικής

- **Παραγωγική** συλλογιστική (deduction).
 - Παραγωγή ορθών συμπερασμάτων με **κανόνες λογικής**.
 - Π.χ. modus ponens, modus tollens (επόμενες διαλέξεις)
 - Π.χ. $\{\forall x_1 (\text{Dog}(x_1) \Rightarrow \text{Animal}(x_1)), \text{Dog}(\text{Fido})\} \vdash \text{Animal}(\text{Fido})$
- **Απαγωγική** συλλογιστική (abduction).
 - Προσπάθεια εύρεσης **πιθανής** υπόθεσης που να εξηγεί παρατηρήσεις. Η υπόθεση μπορεί να **μην ισχύει**.
 - Π.χ. γνώση: $\forall x_1 (\text{Has}(x_1, \text{Grippe}) \Rightarrow \text{Fever}(x_1, \text{High}))$
 - Παρατήρηση: $\text{Fever}(\text{John}, \text{High})$
 - Πιθανή εξήγηση: $\text{Has}(\text{John}, \text{Grippe})$
 - Πολύ σημαντικός μηχανισμός σε συστήματα **διάγνωσης**.

Είδη συλλογιστικής (συνέχεια)

- **Επαγωγική** συλλογιστική (induction).
 - Προσπάθεια εύρεσης **γενίκευσης**.
 - *Has(Patient323, Grippe), Fever(Patient323, High),*
Has(Patient357, Pneumonia), Fever(Patient357, High),
Has(Patient389, Grippe), Fever(Patient389, High),
Has(Patient456, Grippe), Fever(Patient456, High),
Has(Patient498, Grippe), Fever(Patient498, Medium).
 - Γενίκευση (αγνοώντας σπάνιες περιπτώσεις):
$$\forall x_1 (\text{Has}(x_1, \text{Grippe}) \Rightarrow \text{Fever}(x_1, \text{High}))$$
 - Ιδιαίτερα σημαντική στη **μηχανική μάθηση**.

Μηχανική μάθηση

- Η **χειρωνακτική** εισαγωγή γνώσεων σε ένα σύστημα είναι συχνά **δύσκολη**.
 - Δυσκολία/κόστος εκμαίευσης γνώσης από ειδικούς.
 - Δυσκολία προσδιορισμού των απαιτούμενων γνώσεων.
- **Αλγόριθμοι μηχανικής μάθησης**: εξάγουν αυτόματα **νέες γνώσεις από εμπειρικά δεδομένα**, βελτιώνοντας έτσι τη συμπεριφορά ενός συστήματος.
 - Π.χ. πρόγραμμα που μαθαίνει να κάνει **ιατρικές διαγνώσεις** από προηγούμενες διαγνώσεις ιατρών και τα αντίστοιχα αποτελέσματα εργαστηριακών εξετάσεων.
 - Π.χ. πρόγραμμα που μαθαίνει να **κατατάσσει κριτικές προϊόντων** σε θετικές, αρνητικές, μικτές, ουδέτερες.

Μερικές από τις πολλές εφαρμογές της MM

- Αναγνώριση χαρακτήρων σε **χειρόγραφα**.
 - Προγράμματα οπτικής αναγνώρισης χαρακτήρων (OCR).
 - Προγράμματα αναγνώρισης χαρακτήρων για κινητά.
- Αναγνώριση **ομιλίας**.
 - Αναγνώριση προφορικών εντολών (π.χ. στο αυτοκίνητο).
 - Μετατροπή συνεχούς ομιλίας σε γραπτή μορφή (υπαγόρευση).
- Προγράμματα που παίζουν **τάβλι, σκάκι, Go, ...**
 - TD-GAMMON (Tesauro 1992, 95): εκπαιδεύτηκε παίζοντας πάνω από 1 εκατομμύριο παρτίδες με τον εαυτό του.
 - AlphaGo (Google DeepMind, <https://en.wikipedia.org/wiki/AlphaGo>).
- **Οδήγηση** αυτοκινήτου.
 - ALVINN (Pomerlau 1989): μετά από εκπαίδευση οδήγησε σε αυτοκινητόδρομο επί 90 μίλια με ταχύτητα 70 μιλίων/ώρα.
 - Self-driving cars (<https://waymo.com/>).
- **Μεγάλα γλωσσικά μοντέλα (LLMs)**.
 - Π.χ. GPT-x, Claude, Llama, Gemini, Quen, Mixtral, DeepSeek...

Επιβλεπόμενη/μη επιβλεπόμενη μάθηση

- Αλγόριθμοι **επιβλεπόμενης μάθησης**.
 - Προϋποθέτουν ότι υπάρχουν **παραδείγματα εκπαίδευσης** για τα οποία είναι γνωστές (ή μπορούν να αποκτηθούν) οι **ορθές απαντήσεις**.
 - Π.χ. σύστημα που μαθαίνει να κάνει **ιατρικές διαγνώσεις** από προηγούμενες περιπτώσεις ασθενών και διαγνώσεων.
- Αλγόριθμοι **μη επιβλεπόμενης μάθησης**.
 - Προσπαθούν να ανακαλύψουν νέες γνώσεις από δεδομένα που **δεν περιέχουν τις επιθυμητές απαντήσεις**.
 - Π.χ. οργάνωση πελατών ή ειδήσεων σε **ομάδες (clusters)** ή εύρεση συσχετισμών της μορφής «οι πελάτες που αγοράζουν το X την Παρασκευή αγοράζουν και ...».
 - Πολλές εφαρμογές στην **εξόρυξη γνώσης** και την **επιστήμη δεδομένων** (παρακολουθήστε το αντίστοιχο μάθημα).

Παραδείγματα εκπαίδευσης

- Τραπεζικό σύστημα που θα αποφασίζει αν πρέπει να δοθεί **δάνειο** σε έναν πελάτη.

ιδιότητες

Πελάτης	Οφειλές	Εισόδημα	Παντρεμένος	Καλός;
1	Υψηλές (1)	Υψηλό (1)	Ναι (1)	Καλός (1)
2	Χαμηλές (0)	Υψηλό (1)	Όχι (0)	Κακός (0)
3	Χαμηλές (0)	Υψηλό (1)	Ναι (1)	Καλός (1)
4	Υψηλές (1)	Χαμηλό (0)	Ναι (1)	Κακός (0)
5	Χαμηλές (0)	Χαμηλό (0)	Ναι (1)	Κακός (0)

- Διανυσματική παράσταση εμπειρίας:

$\{ \langle 1, 1, 1, \mathbf{1} \rangle, \langle 0, 1, 1, \mathbf{1} \rangle, \langle 0, 1, 0, \mathbf{0} \rangle, \langle 1, 0, 1, \mathbf{0} \rangle, \langle 0, 0, 1, \mathbf{0} \rangle \}$

επιθυμητές απαντήσεις

Η μάθηση ως πρόβλημα αναζήτησης

- Αναζήτηση μιας **συνάρτησης** («υπόθεσης»).
- $h(x, y, z) = ?$
- $h: \{0, 1\} \times \{0, 1\} \times \{0, 1\} \rightarrow \{0, 1\}$
- γενικότερα $h: \mathbf{D}_1 \times \mathbf{D}_2 \times \dots \times \mathbf{D}_n \rightarrow \mathbf{C}$
- \mathbf{D}_i : οι δυνατές τιμές της i -στής ιδιότητας.
- \mathbf{C} : οι δυνατές **απαντήσεις**.
- Αναζήτηση σε ένα **χώρο συναρτήσεων**.
- Θεωρούμε ότι τα δεδομένα εκπαίδευσης είναι ένα **δείγμα** από έναν πληθυσμό που παράγεται σύμφωνα με μια **συγκεκριμένη άγνωστη συνάρτηση**.
- Αναζητούμε αυτή τη συνάρτηση μέσα σε ένα **χώρο συναρτήσεων**.

Χώρος αναζήτησης

- Ο χώρος αναζήτησης εξαρτάται από τις **ιδιότητες**.
 - **Ποιες ιδιότητες** θα χρησιμοποιήσουμε;
 - Αντιστοιχούν στα ορίσματα των υποθέσεων.
 - Π.χ. οικογ. κατάσταση, οφειλές, ηλικία, επάγγελμα, ύψος;
 - **Τιμές** από πεπερασμένο σύνολο, πραγματικοί αριθμοί;
- Και από το **μοντέλο παράστασης** των υποθέσεων.
 - Π.χ. **γραμμική συνάρτηση** των ιδιοτήτων;
 - Περιορίζουμε την αναζήτηση στις (π.χ. γραμμικές) **υποθέσεις που μπορεί να παραστήσει το μοντέλο**.
 - Δεν μπορούμε να μάθουμε μια υπόθεση που δεν μπορεί να παρασταθεί από το μοντέλο που διαλέξαμε.
 - Οι περιορισμοί του μοντέλου, όμως, **μειώνουν το μέγεθος του χώρου αναζήτησης**.
 - Δυστυχώς συχνά δεν ξέρουμε τι περιορισμούς να επιβάλουμε.

Μοντέλο παράστασης υποθέσεων

- Απλοϊκό παράδειγμα για το πρόβλημα των δανείων:
 - Παριστάνουμε κάθε συνάρτηση-υπόθεση με μια **τριάδα** που αντιστοιχεί στις περιπτώσεις που δίνουμε δάνειο.
 - Π.χ. $\mathbf{h}_1 = \langle 0, ?, 1 \rangle$, $\mathbf{h}_2 = \langle 1, 1, 1 \rangle$
 - «?» σημαίνει για **οποιαδήποτε** τιμή.

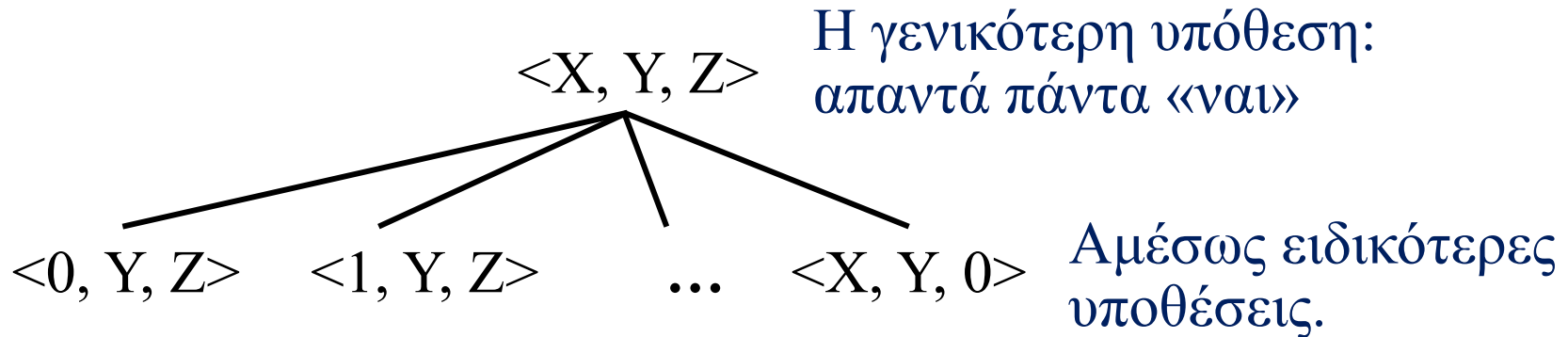
Αν χαμηλές οφειλές και παντρεμένος/η, δώσε δάνειο ανεξαρτήτως εισοδήματος. Διαφορετικά μη δώσεις

Αν υψηλές οφειλές και υψηλό εισόδημα και παντρεμένος/η, δώσε δάνειο. Διαφορετικά μη δώσεις.

- Αδύνατον, όμως, να μάθουμε την $(\mathbf{h}_1 \vee \mathbf{h}_2)$, γιατί δεν περιλαμβάνεται στο χώρο αναζήτησης.

Γενικότερες/ειδικότερες υποθέσεις

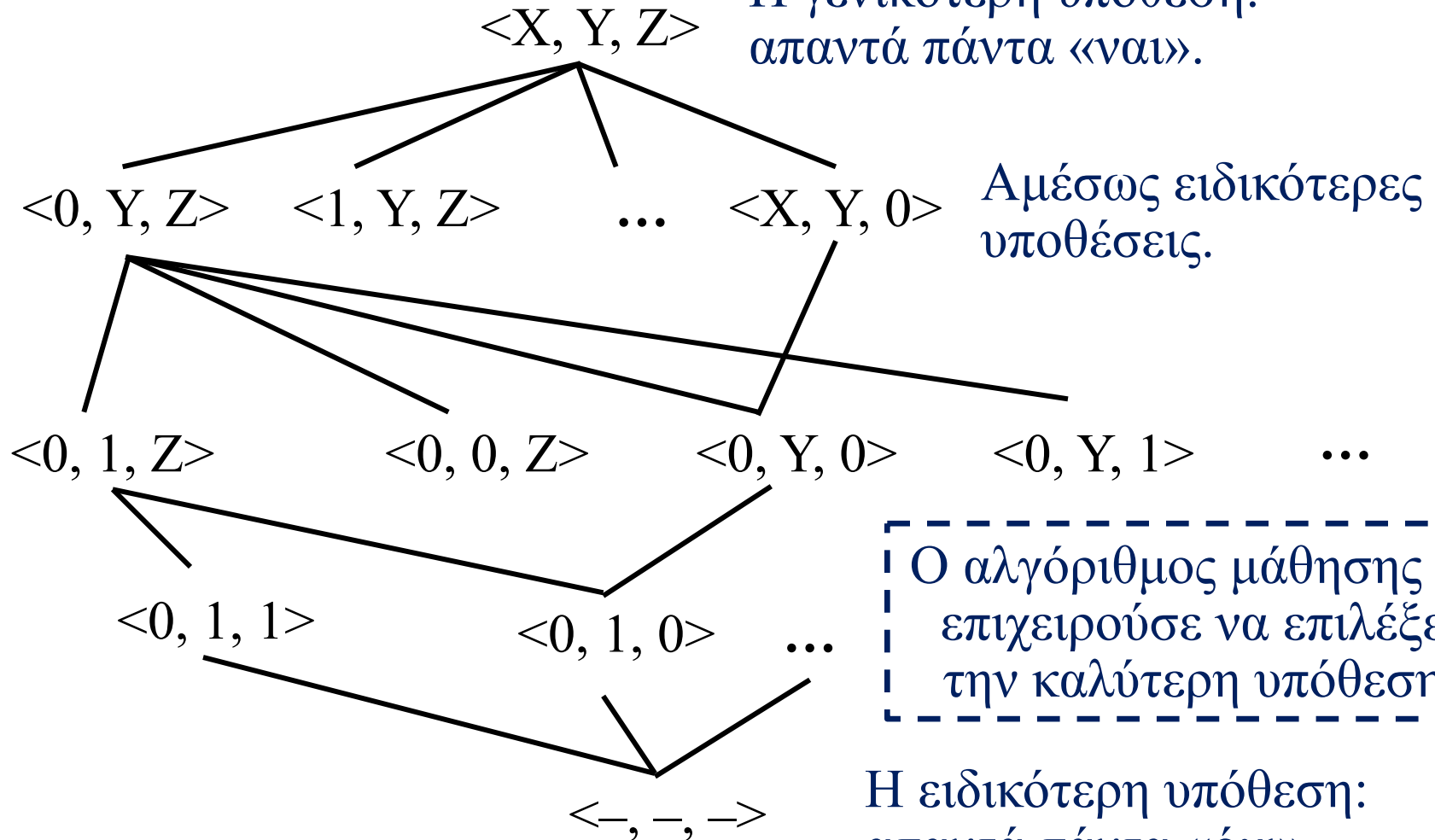
- Υποθέτουμε εδώ πάλι ότι οι τιμές των υποθέσεων είναι 0 ή 1 και οι ιδιότητες έχουν διακριτές τιμές.
- Η h_2 είναι ειδικότερη της h_1 αν $h_1 \neq h_2$ (διαφορετικές συναρτήσεις) και:
 - αν $h_2(\mathbf{x}_1, \dots, \mathbf{x}_n) = 1$, τότε $h_1(\mathbf{x}_1, \dots, \mathbf{x}_n) = 1$.
 - Π.χ. αν δίνει δάνειο η h_2 , τότε δίνει και η h_1 .
- Η h_1 είναι γενικότερη της h_2 αν η h_2 είναι ειδικότερη της h_1 .



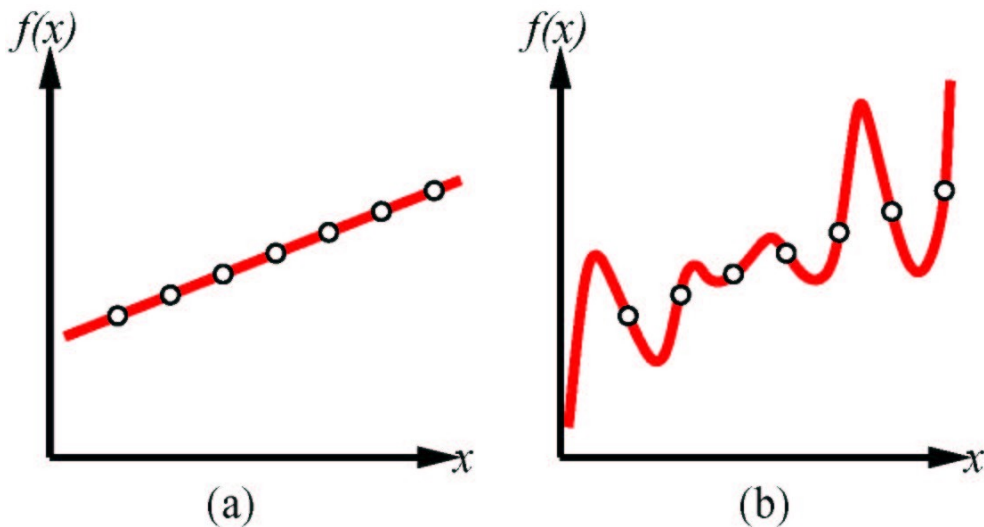
Εδώ χρησιμοποιούμε μεταβλητές αντί για «?».

Χώρος αναζήτησης

Η γενικότερη υπόθεση:
απαντά πάντα «ναι».



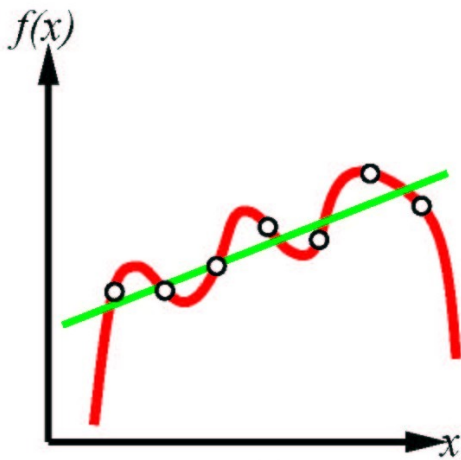
Συνεπείς υποθέσεις και γενίκευση



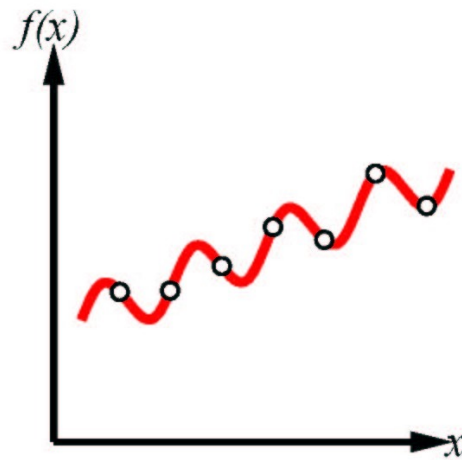
Οι τελείες αντιστοιχούν σε δεδομένα εκπαίδευσης. Εδώ έχουμε μία ιδιότητα και οι τιμές των υποθέσεων είναι πραγματικοί. Έστω ότι αναζητούμε πολυωνυμικές υποθέσεις.

- **Συνεπείς υποθέσεις:** συμφωνούν με τα δεδομένα εκπαίδευσης.
 - (a): πολυώνυμο 1ου βαθμού. (b): πολυώνυμο μεγαλύτερου βαθμού.
- **Ξυράφι του Ockham:** Προτιμότερη είναι η **απλούστερη** συνεπής υπόθεση, εδώ το πολυώνυμο μικρότερου βαθμού.
 - Απλούστερες υποθέσεις είναι ευκολότερο να κατασκευαστούν (π.χ. γραμμική παρεμβολή) και οι αποκρίσεις τους υπολογίζονται ευκολότερα.
- **Ικανότητα γενίκευσης:** Να δίνει σωστές απαντήσεις και για νέες περιπτώσεις που δεν ανήκουν στα παραδείγματα εκπαίδευσης.
 - Απλούστερες υποθέσεις συνήθως γενικεύουν καλύτερα τις παρατηρήσεις.

Ασυνεπείς υποθέσεις ίσως είναι προτιμότερες



(c)



(d)

Τα δεδομένα εκπαίδευσης ακολουθούν μια συνάρτηση της μορφής $a \cdot x + b + c \cdot \sin x$ (διάγραμμα d), που είναι αδύνατον να παρασταθεί ως πολυώνυμο πεπερασμένου βαθμού.

- Αν περιορίσουμε την αναζήτηση σε **πολυώνυμο πεπερασμένου βαθμού** (έστω $\leq k$), η συνάρτηση-στόχος του (d) **δεν περιλαμβάνεται** στο χώρο αναζήτησης.
- Αν επιμείνουμε στην εξεύρεση **συνεπούς** υπόθεσης, καταλήγουμε στο **πολυώνυμο βου βαθμού** της κόκκινης γραμμής του (c), που **δεν επιτυγχάνει καλή γενίκευση**.
 - Μας ενδιαφέρει οι προβλέψεις να είναι καλές για όλα τα x .
- Οι προβλέψεις της **ευθείας γραμμής** (c) είναι εν γένει **πιο κοντά** στις επιθυμητές τιμές (d), αν και η (c) είναι **ασυνεπής**.

Παράδειγμα: φίλτρα ανεπιθύμητων μηνυμάτων

our highly successful multi – national company gives you an exclusive business that generates an extra weekly income of up to \$ 600 or more ... anyone can easily make money ... if you wish to be removed from our list ...

call for papers 9 th european workshop on natural language generation ... is a subfield of natural language processing that generates texts in human languages from non-linguistic data or knowledge ... for the systems to be successful ...

- **Πρόβλημα κατάταξης με δύο κατηγορίες.**
 - **Spam** ($C = 1$), **ham** ($C = 0$).
- Γενικότερα μπορεί να υπάρχουν n κατηγορίες.
 - **Πιθανώς επικαλυπτόμενες** (multi-label): π.χ. οικονομικά, πολιτικά, κοινωνικά, αθλητικά νέα.
 - **Ή μη επικαλυπτόμενες** (single-label): π.χ. θετική, αρνητική, ουδέτερη κριτική προϊόντος.

Φίλτρο ανεπιθύμητων μηνυμάτων με επιβλεπόμενη MM

παραδείγματα **επιθυμητών** και **ανεπιθύμητων** μηνυμάτων + σωστές απαντήσεις

εκπαίδευση

προεπεξεργασία

μονάδα εκπαίδευσης

συνάρτηση ταξινόμησης

προεπεξεργασία

μονάδα ταξινόμησης

νεοεισερχόμενο μήνυμα

απόφαση: **επιθυμητό** ή **ανεπιθύμητο**

χρήση

Προεπεξεργασία κειμένων

our highly successful multi – national company gives you an exclusive business that generates an extra weekly income of up to \$ 600 or more ... anyone can easily make money ... if you wish to be removed from our list ...

call for papers 9 th european workshop on natural language generation ... is a subfield of natural language processing that generates texts in human languages from non-linguistic data or knowledge ... for the systems to be successful ...

< money:1, language:0, natural:0, \$:1, adult:0, call:0, exclusive:1, **successful:1**, removed:1, **generates:1**, ... >

< money:0, language:1, natural:1, \$:0, adult:0, call:1, exclusive:0, **successful:1**, removed:0, **generates:1**, ... >

- Εναλλακτικά οι ιδιότητες μπορεί να παριστάνουν **συχνότητες εμφάνισης λέξεων** στο μήνυμα και/ή να αντιστοιχούν σε **φράσεις** (π.χ. «to be removed»), **χαρακτηριστικά εικόνων** κλπ.

Απλές διανυσματικές παραστάσεις κειμένων

- **Δυαδικά διανύσματα** (τιμές 0, 1): ποιες λέξεις ενός λεξιλογίου εμφανίζονται ή όχι στο κείμενο;
- **Διανύσματα συχνότητων λέξεων** (term frequencies, TF): πόσες φορές εμφανίζεται κάθε λέξη στο κείμενο;
 - Συχνά διαιρούμε με το μήκος (σε λέξεις) του κειμένου.
- **Διανύσματα $TF-IDF$** : για κάθε λέξη w_i του λεξιλογίου το διάνυσμα περιέχει ένα σκορ $TF_i \cdot IDF_i$:
 - Θέλουμε **μεγάλα σκορ** για λέξεις που είναι **συχνές στο συγκεκριμένο κείμενο** αλλά **όχι κοινές στη γλώσσα**.

$$IDF_i = \log\left(\frac{N_{doc}}{DF_i}\right)$$

Πλήθος εγγράφων σε μια μεγάλη συλλογή.

Πλήθος εγγράφων που περιέχουν την w_i .

- IDF_i (inverse document frequency): δείχνει πόσο σπάνια είναι η w_i σε μια γλώσσα.

Επιλογή ιδιοτήτων

- Για **ποιες λέξεις** (ή φράσεις ή ...) θα υπάρχουν **ιδιότητες** στα διανύσματα;
 - 1ο βήμα: μόνο λέξεις που εμφανίζονται **τουλάχιστον** k φορές στα παραδείγματα εκπαίδευσης (π.χ. $k = 3$).
 - Συνήθως παραμένουν αρκετές **χιλιάδες** λέξεις (και πολύ περισσότερες φράσεις).
 - Με χιλιάδες ιδιότητες: προβλήματα **ταχύτητας**, **ορθότητας** και **υπερεφαρμογής** με πολλούς αλγορίθμους μηχανικής μάθησης.
- Πόσο **αξίζει** κάθε υποψήφια ιδιότητα X ;
 - Αν ξέρουμε την τιμή της X , **πόση πληροφορία** κερδίζουμε;
 - Πόσο μειώνεται η **αβεβαιότητά** μας για τη σωστή απάντηση;
 - **$C = 1$** (ανεπιθύμητο) ή **$C = 0$** (επιθυμητό).
 - Πόσο μειώνεται η **αβεβαιότητά** μας για την τιμή της **τυχαίας μεταβλητής** C , αν ξέρουμε την τιμή της X ;

Παράδειγμα κωδικοποίησης τιμών

- Έστω οι ακόλουθες **δυνατές τιμές** της **C**:
 - c_1 με $P(c_1) = 1/4$, c_2 με $P(c_2) = 1/4$, c_3 με $P(c_3) = 1/2$.
- **Παράδειγμα** κωδικοποίησης:
 - Χρησιμοποιούμε **λιγότερα δυφία (bits) για τις πιο πιθανές τιμές**.
 - $c_1 \rightarrow 10$, $c_2 \rightarrow 11$. Χρησιμοποιούμε $-\log_2(1/4) = 2$ δυφία.
 - $c_3 \rightarrow 0$. Χρησιμοποιούμε $-\log_2(1/2) = 1$ δυφίο.
 - Αν ο παραλήπτης λάβει «1», ξέρει ότι ακολουθεί και δεύτερο δυφίο.
 - Αναμενόμενος αριθμός δυφίων: $1/4 \cdot 2 + 1/4 \cdot 2 + 1/2 \cdot 1 = 1.5$
- Γενικότερα, στην **ιδανική κωδικοποίηση** (ελάχιστος αναμενόμενος αριθμός μεταδιδόμενων δυφίων) χρησιμοποιούμε $-\log_2(P(c_i))$ **δυφία για κάθε τιμή c_i** .
 - Αν οι πιθανότητες $P(c_i)$ **δεν είναι δυνάμεις του 2**, ίσως είναι **αδύνατον** να χρησιμοποιήσουμε ακριβώς $-\log_2(P(c_i))$ **δυφία** για κάποια c_i στην πράξη. Αλλά δεν μας πειράζει αυτό, η ιδανική κωδικοποίηση είναι θεωρητική.
 - Στην πράξη, ο αναμενόμενος αριθμός μεταδιδόμενων δυφίων θα είναι γενικά **μεγαλύτερος ή ίσος** αυτού της ιδανικής κωδικοποίησης, τόσο μεγαλύτερος όσο απέχει η κωδικοποίηση από την ιδανική.

Εντροπία

Εδώ η **ορθή** απόκριση.

- **Εντροπία** της τυχαίας μεταβλητής C .
 - Δείχνει **πόσο αβέβαιοι** είμαστε για την **τιμή της C** .
 - Πόση είναι η **ελάχιστη ποσότητα πληροφορίας** που πρέπει να μας δοθεί για να γνωρίζουμε με βεβαιότητα την τιμή της C .
 - Ποιος είναι στην **καλύτερη περίπτωση** (με την ιδανική κωδικοποίηση) ο **αναμενόμενος αριθμός δυφίων** που πρέπει να μας μεταδοθεί για να καθοριστεί η τιμή της C ;

αναμενόμενη τιμή

$$H(C) = - \sum_c P(C = c) \log_2 P(C = c)$$

Δυφία που χρησιμοποιούμε στην ιδανική κωδικοποίηση για την παράσταση κάθε τιμής.

- Στην περίπτωση δύο κατηγοριών:

$$H(C) = -P(C = 1) \cdot \log_2 P(C = 1) - P(C = 0) \cdot \log_2 P(C = 0)$$

Εκτίμηση πιθανοτήτων από τα παραδείγματα εκπαίδευσης.

Παράδειγμα

- Συλλογή **800 μηνυμάτων** e-mail εκπαίδευσης.
 - Προηγούμενα μηνύματα που έχω λάβει.
 - **200 ανεπιθύμητα** διαφημιστικά μηνύματα (spam).
 - **600 επιθυμητά** μηνύματα.
- Εκτιμήστε την **εντροπία της C** (κατηγορία νέου μηνύματος) με βάση τα μηνύματα εκπαίδευσης.
 - **$C = 1$** (ανεπιθύμητο) ή **$C = 0$** (επιθυμητό).
 - $\log_2 3 = 1.585$
- **Πόση γίνεται η εντροπία της C** αν **όλα** τα μηνύματα εκπαίδευσης είναι **επιθυμητά** (ή ανεπιθύμητα);
- Αν είναι **400 επιθυμητά** και **400 ανεπιθύμητα**;

$$H(C) = -P(C = 1) \cdot \log_2 P(C = 1) - P(C = 0) \cdot \log_2 P(C = 0)$$

Κέρδος πληροφορίας

- Εντροπία της C αν μάθουμε ότι η τιμή της X είναι 1:

$$H(C | X = 1) = - \sum_c P(C = c | X = 1) \cdot \log_2 P(C = c | X = 1)$$

- Εντροπία της C αν μάθουμε ότι η τιμή της X είναι 0:

$$H(C | X = 0) = - \sum_c P(C = c | X = 0) \cdot \log_2 P(C = c | X = 0)$$

- **Κέρδος πληροφορίας IG** (information gain):

Αναμενόμενη μείωση της $H(C)$, αν μάθουμε την τιμή της X .

$$IG(C, X) = IG(X, C) = H(C) - \sum_x P(X = x) \cdot H(C | X = x)$$

αναμενόμενη τιμή \rightarrow x για κάθε δυνατή τιμή του X \rightarrow $H(C | X = x)$

Επιλογή ιδιοτήτων με IG

- Υπολογίζουμε το κέρδος πληροφορίας $IG(C, X)$ που παρέχει κάθε υποψήφια ιδιότητα X .
 - Π.χ. η X μπορεί να παριστάνει την ύπαρξη ($X = 1$) ή μη ($X = 0$) της λέξης «money» στο μήνυμα.
 - Οι πιθανότητες εκτιμούνται χρησιμοποιώντας τα παραδείγματα εκπαίδευσης.
- Επιλέγουμε τις ιδιότητες με τα m υψηλότερα $IG(C, X)$.
 - Π.χ. $m = 500$. Συχνά αφού παραλείψουμε πολύ συχνές λέξεις (stop-words, π.χ. άρθρα, συνδέσμους) και πολύ σπάνιες λέξεις.
- Παριστάνουμε τα παραδείγματα εκπαίδευσης και τα νεοεισερχόμενα μηνύματα ως διανύσματα m διαστάσεων.
 - Π.χ. $\langle X_1, X_2, X_3, \dots, X_m \rangle = \langle 0, 1, 1, \dots, 0 \rangle$

Παράδειγμα επιλεγμένων ιδιοτήτων

Λέξη του X_i	$P(X_i=1)$	$P(X_i=1 C=0)$	$P(X_i=1 C=1)$
!	0.484105	0.216129	0.828157
\$	0.257947	0.040322	0.538302
language	0.247956	0.440322	0.002070
money	0.163487	0.001612	0.372670
remove	0.146230	0.001612	0.333333
free	0.309718	0.104838	0.573498
university	0.219800	0.374193	0.022774

Αναζήτηση υπο-συνόλων ιδιοτήτων

- Ως τώρα θεωρούσαμε ότι **κάθε υποψήφια ιδιότητα αξιολογείται ξεχωριστά.**
 - Με *IG*, υπάρχουν και άλλα μέτρα (π.χ. χ^2).
- Μπορεί, όμως, μια ιδιότητα ουσιαστικά να παρέχει περίπου τις **ίδιες πληροφορίες με μια άλλη.**
 - Π.χ. αν οι λέξεις «money» και «rich» εμφανίζονται (ή όχι) σχεδόν πάντα μαζί.
 - Οπότε μπορεί και οι δύο να έχουν υψηλό **IG**, αλλά η μία να **περιττεύει.**
- Μπορεί, επίσης, δύο (ή περισσότερες) ιδιότητες **από μόνες τους να μην είναι πολύ χρήσιμες**, αλλά και οι δύο (ή περισσότερες) **μαζί να προβλέπουν καλά** την κατηγορία.
 - Π.χ. υψηλός πυρετός **ΚΑΙ** δύσπνοια...

Αναζήτηση υπο-συνόλων ιδιοτήτων – συνέχεια

- Μπορούμε να θεωρήσουμε το πρόβλημα **επιλογής του καλύτερου υποσυνόλου ιδιοτήτων** (μεταξύ των διαθέσιμων) ως ένα **πρόβλημα αναζήτησης**.
 - Π.χ. **αρχική κατάσταση**: όλες οι υποψήφιες ιδιότητες.
 - **Τελεστής μετάβασης**: αφαίρεση μιας υποψήφιας ιδιότητας.
 - **Συνάρτηση αξιολόγησης κατάστασης**: ποσοστό ορθότητας που επιτυγχάνουμε με το υπο-σύνολο ιδιοτήτων της κατάστασης εκπαιδευοντας στα δεδομένα εκπαίδευσης και αξιολογώντας σε διαφορετικά δεδομένα ανάπτυξης (ή «επικύρωσης»).
 - Αναζήτηση π.χ. με **αναρρίχηση λόφου, beam search...**
- Εναλλακτικά μπορούμε να χρησιμοποιήσουμε **μείωση διαστάσεων** μέσω **Singular Value Decomposition (SVD)**.
 - Βλ. μάθημα «Αριθμητική Γραμμική Άλγεβρα» (6ο εξάμηνο).
 - Απεικονίζει τα διανύσματα σε νέο χώρο λιγότερων διαστάσεων.
 - **Δεν επιλέγουμε** απλά ιδιότητες, **κατασκευάζουμε νέες** (λιγότερες) που **συνδυάζουν πληροφορίες** των αρχικών.

Βιβλιογραφία

- Russel & Norvig (4^η έκδοση): ενότητες 19.1, 19.2, 19.3.3.
 - Θα καλύψουμε αρκετές άλλες ενότητες του κεφαλαίου 19 στις επόμενες διαλέξεις.
- Βλαχάβας κ.ά: ενότητα 8.2, κεφάλαιο 18 ως και ενότητα 18.2, υπο-ενότητα «Κριτήριο διαχωρισμού με βάση την εντροπία» της ενότητας 18.5.1.
- Περισσότερα για το PCA (και SVD) θα βρείτε στην ενότητα 12.2 του βιβλίου «Machine Learning – A Probabilistic Perspective», του K.P. Murphy, MIT Press, 2012.
 - Υπάρχει στη βιβλιοθήκη του ΟΠΑ.

