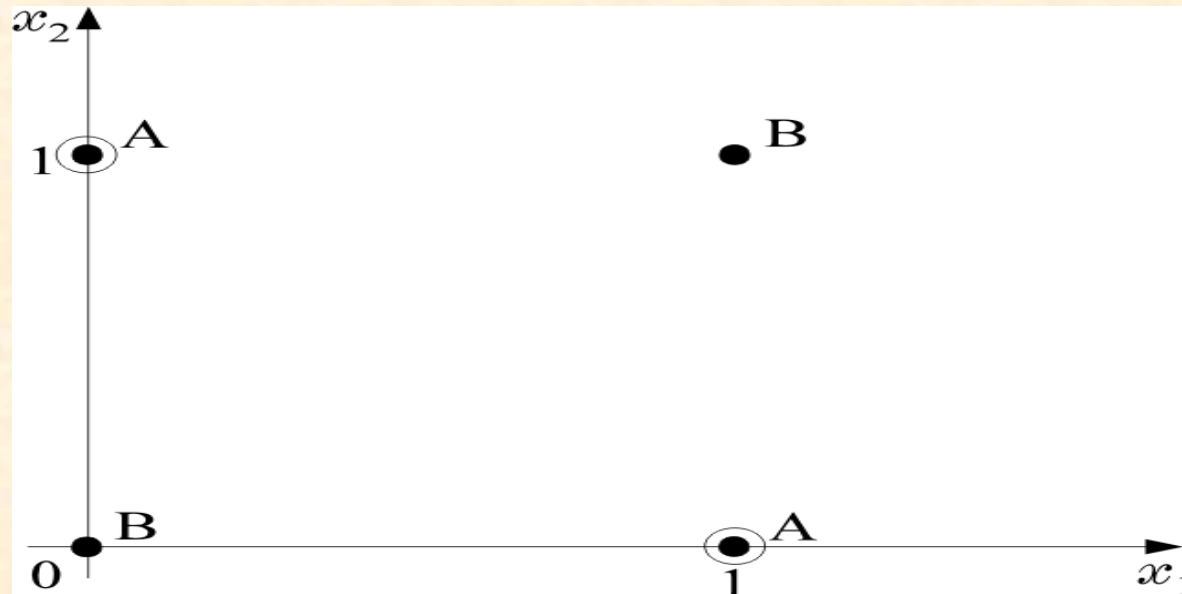


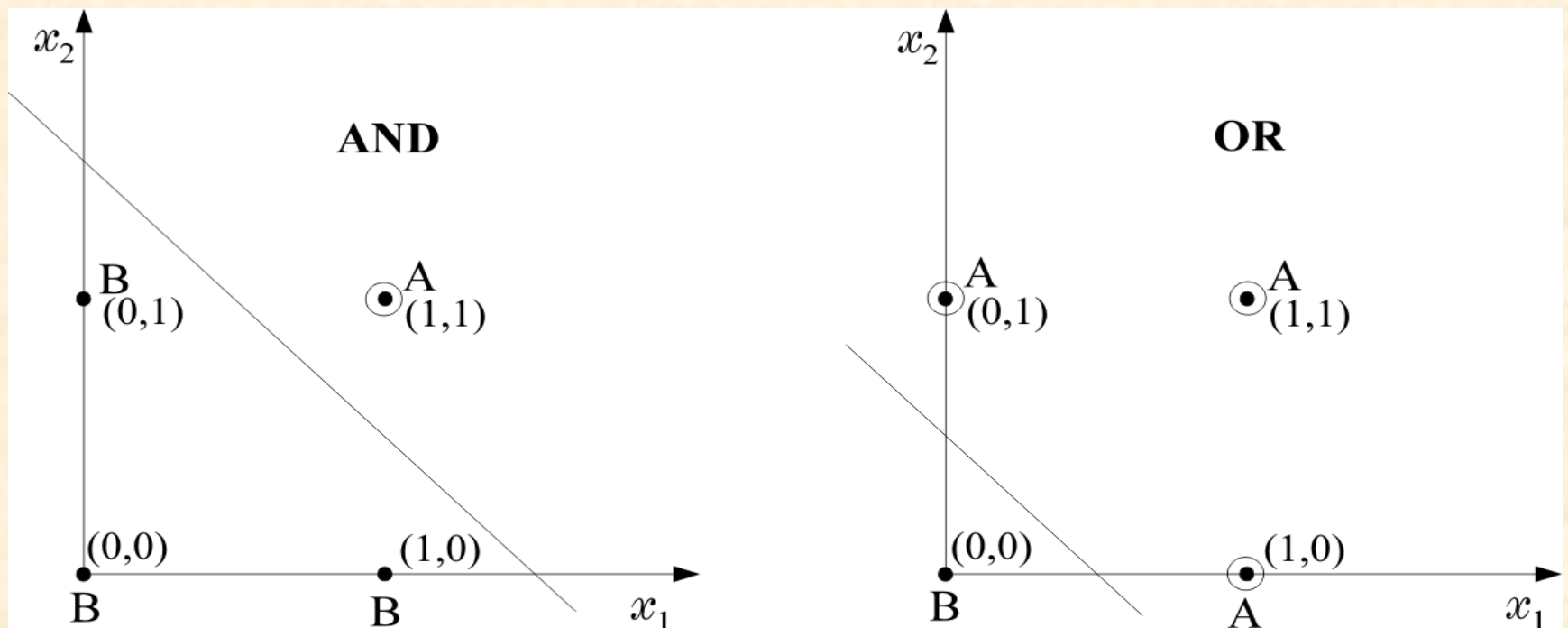
Μη γραμμικοί ταξινομητές

❖ Το πρόβλημα XOR

x_1	x_2	XOR	Class
0	0	0	B
0	1	1	A
1	0	1	A
1	1	0	B

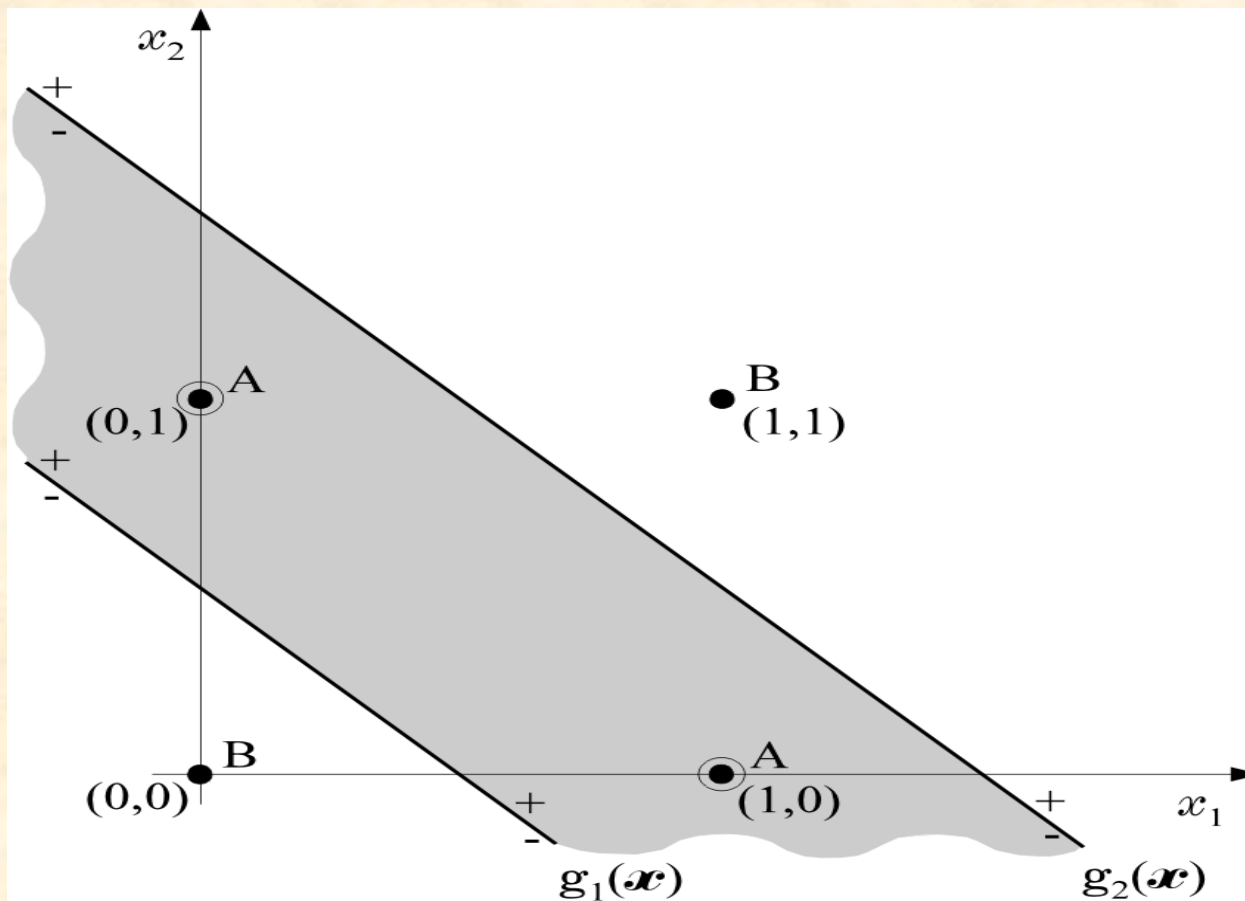


- ❖ Δεν υπάρχει μία γραμμή (υπερεπίπεδο) που να διαχωρίζει την κλάση A από την κλάση B. Αντίθετα, οι λειτουργίες AND και OR είναι γραμμικώς διαχωρίσιμα προβλήματα



❖ Το δι-επίπεδο (Two-Layer) Perceptron

- Για το πρόβλημα XOR, ζωγράφισε **δύο**, αντί, για μία γραμμή



➤ Τότε, η κλάση B βρίσκεται **εκτός** της σκιασμένης περιοχής ενώ η κλάση A **εντός αυτής**. Πρόκειται για μία σχεδίαση **δύο φάσεων** (**two-phase design**).

- Φάση 1: χάραξε δύο γραμμές (υπερεπίπεδα)

$$g_1(\underline{x}) = g_2(\underline{x}) = 0$$

Καθένα από αυτά υλοποιείται από ένα perceptron. Οι έξοδοι των perceptrons θα είναι

$$y_i = f(g_i(\underline{x})) = \begin{cases} 0 \\ 1 \end{cases} \quad i = 1, 2$$

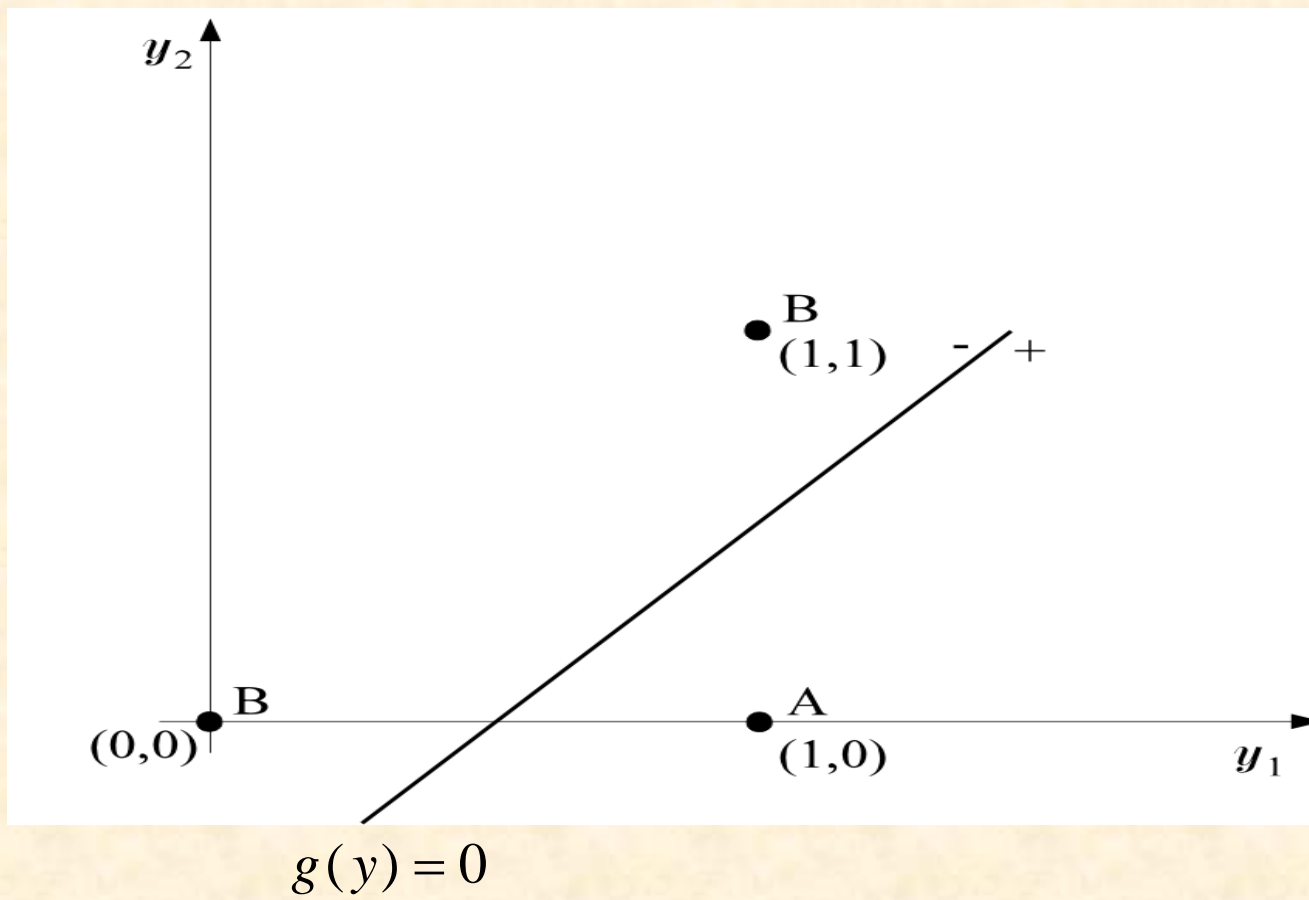
ανάλογα με τη θέση του \underline{x} .

- Φάση 2: Βρες τη θέση του \underline{x} ως προς **αμφότερες** τις γραμμές, με βάση τις τιμές των y_1, y_2 .

1 st phase				2 nd phase
x_1	x_2	y_1	y_2	
0	0	0(-)	0(-)	B(0)
0	1	1(+)	0(-)	A(1)
1	0	1(+)	0(-)	A(1)
1	1	1(+)	1(+)	B(0)

- Ισοδύναμα: Οι υπολογισμοί της πρώτης φάσης υλοποιούν μία απεικόνιση $\underline{x} \rightarrow \underline{y} = [y_1, y_2]^T$

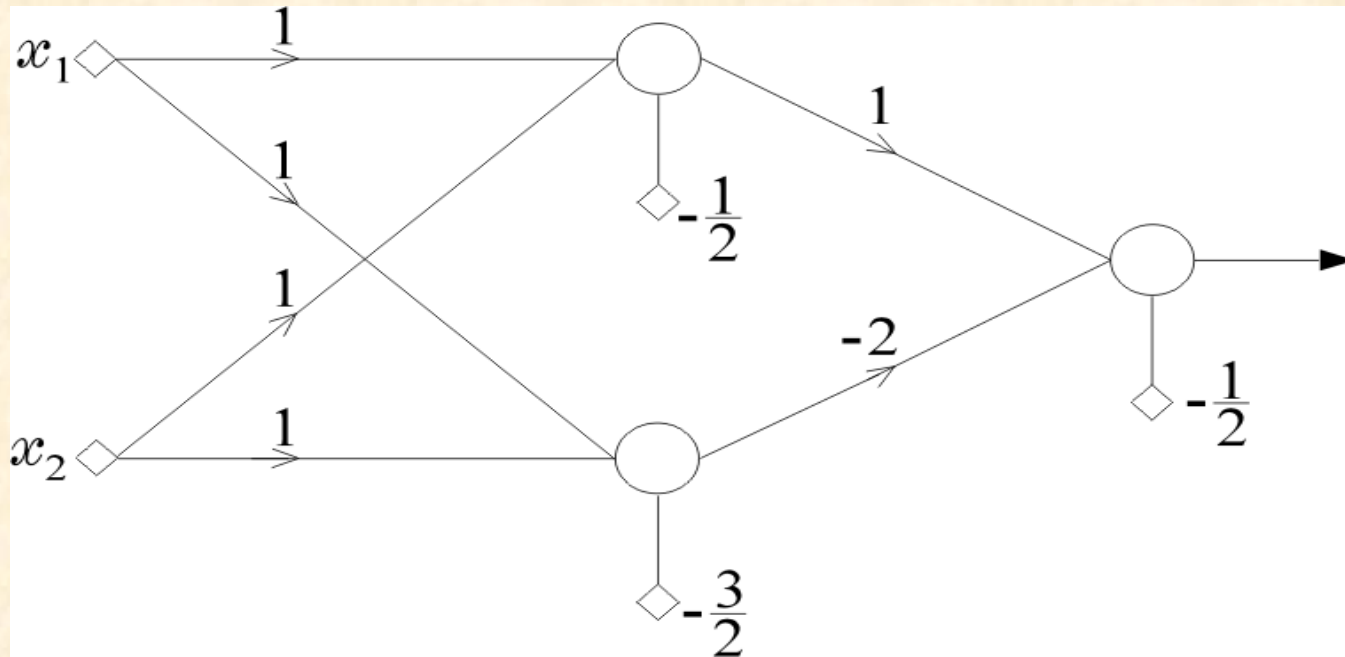
Η απόφαση παίρνεται τώρα με βάση τα **μετασχηματισμένα** δεδομένα. \underline{y}



Αυτό μπορεί να γίνει μέσω μιας δεύτερης γραμμής, η οποία μπορεί επίσης να υλοποιηθεί από ένα perceptron.

- Οι υπολογισμοί της πρώτης φάσης πραγματοποιούν μία απεικόνιση που μετασχηματίζει το μη γραμμικώς διαχωρίσιμο πρόβλημα σε ένα γραμμικώς διαχωρίσιμο.

- Η αρχιτεκτονική



- Αυτή είναι γνωστή ως perceptron δύο επιπέδων με ένα κρυφό (hidden) και ένα επίπεδο εξόδου (output layer). Οι συναρτήσεις ενεργοποίησης (activation functions) είναι

$$f(.) = \begin{cases} 0 \\ 1 \end{cases}$$

- Οι νευρώνες (κόμβοι) του σχήματος υλοποιούν τις ακόλουθες γραμμές (υπερέπιπεδα)

$$g_1(\underline{x}) = x_1 + x_2 - \frac{1}{2} = 0$$

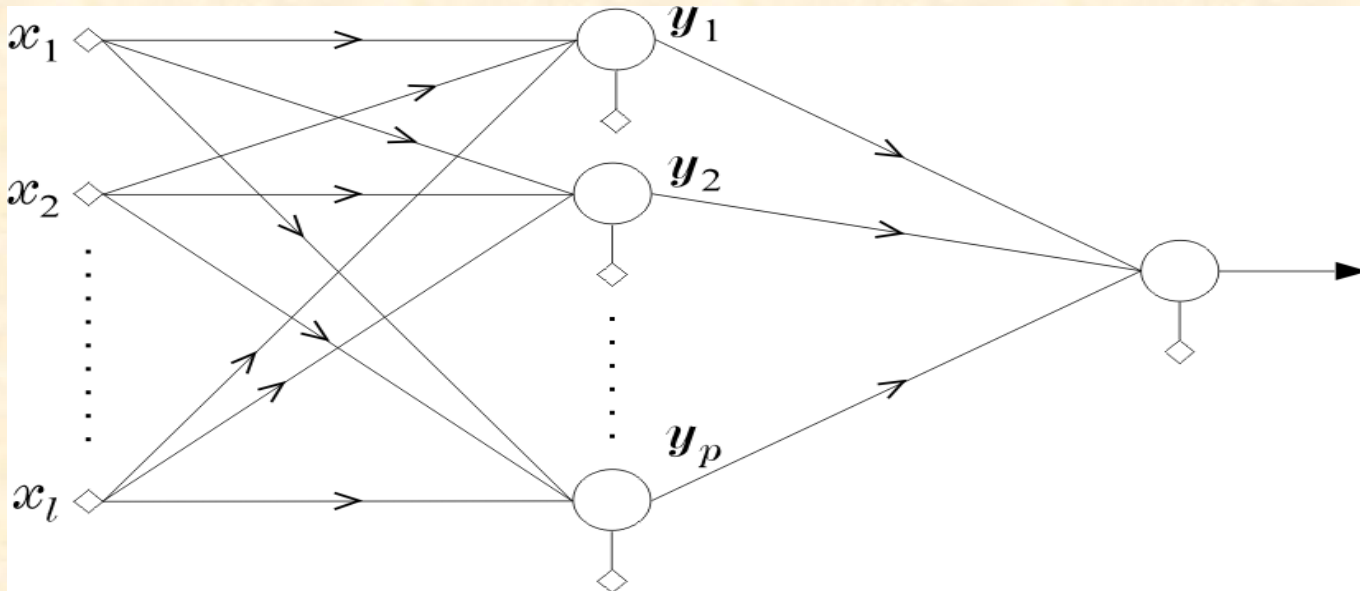
$$g_2(\underline{x}) = x_1 + x_2 - \frac{3}{2} = 0$$

$$g(\underline{y}) = y_1 - 2y_2 - \frac{1}{2} = 0$$

❖ Δυνατότητες ταξινόμησης δικτύου perceptron δύο επιπέδων

➤ Η απεικόνιση που πραγματοποιείται από τους νευρώνες του 1^{ου} επιπέδου είναι **πάνω στις κορυφές** του τετραγώνου πλευράς 1, e.g., $(0, 0), (0, 1), (1, 0), (1, 1)$.

➤ Η πιο γενική περίπτωση,



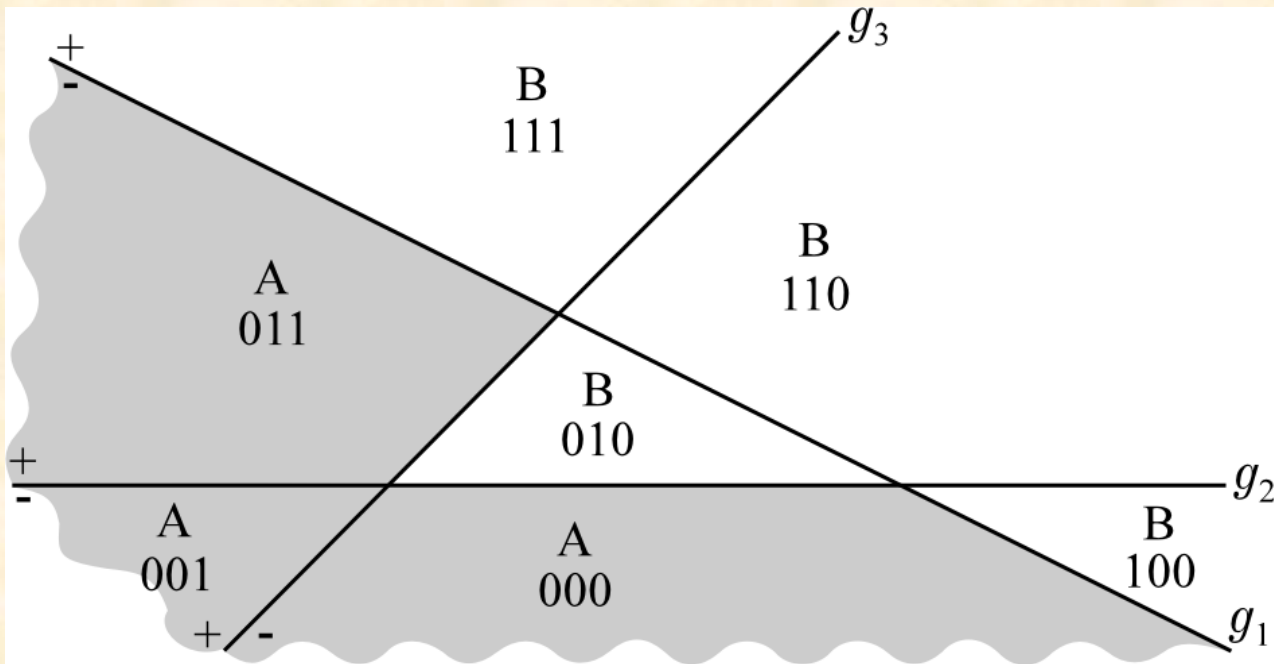
$$\underline{x} \in R^l$$

$$\underline{x} \rightarrow \underline{y} = [y_1, \dots, y_p]^T, y_i \in \{0, 1\} \quad i = 1, 2, \dots, p$$

πραγματοποιεί μία απεικόνιση ενός διανύσματος στις κορυφές του υπερκύβου H_p μοναδιαίας ακμής.

- Η απεικόνιση πραγματοποιείται με p νευρώνες καθένας από τους οποίους υλοποιεί ένα υπερεπίπεδο. Η έξοδος καθενός από αυτούς τους νευρώνες είναι 0 ή 1 ανάλογα με τη **σχετική θέση** του \underline{x} ως προς το υπερεπίπεδο.

- Τομές αυτών των υπερεπιπέδων ορίζουν περιοχές στον l -διάστατο χώρο. Κάθε περιοχή αντιστοιχεί σε μία κορυφή του μοναδιαίου υπερκύβου H_p .

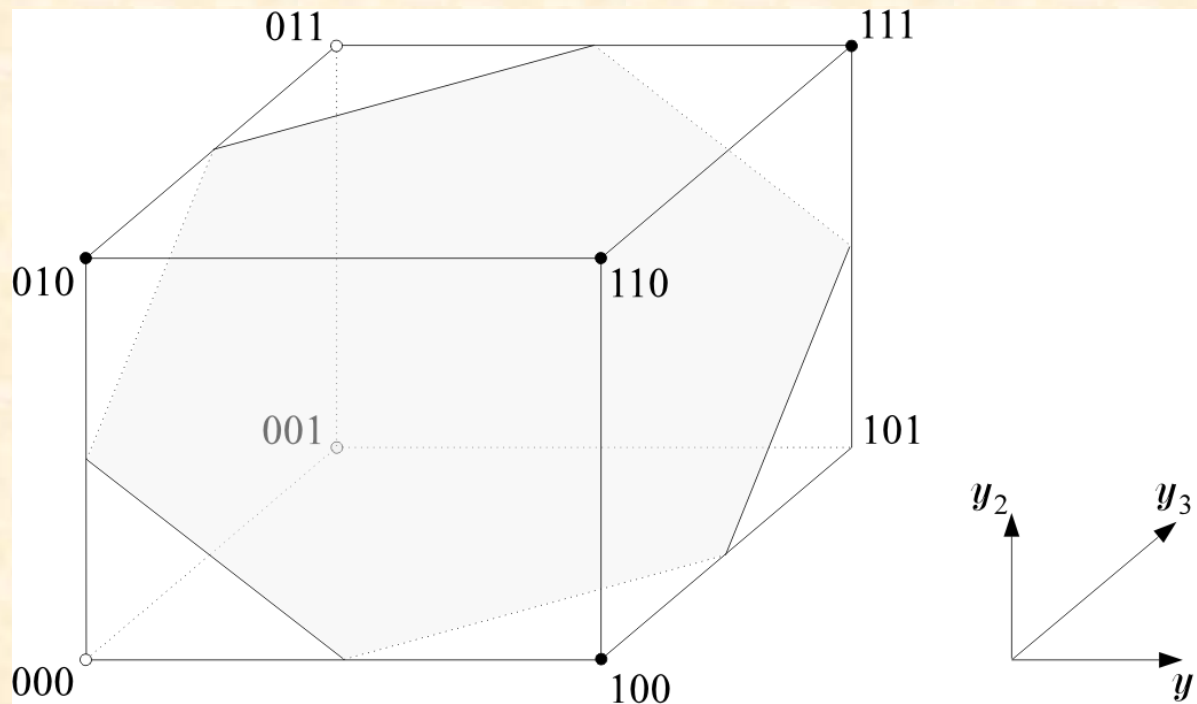


Για παράδειγμα, η κορυφή 001 αντιστοιχεί στην περιοχή που βρίσκεται

στην (-) πλευρά του $g_1(\underline{x})=0$

στην (-) πλευρά του $g_2(\underline{x})=0$

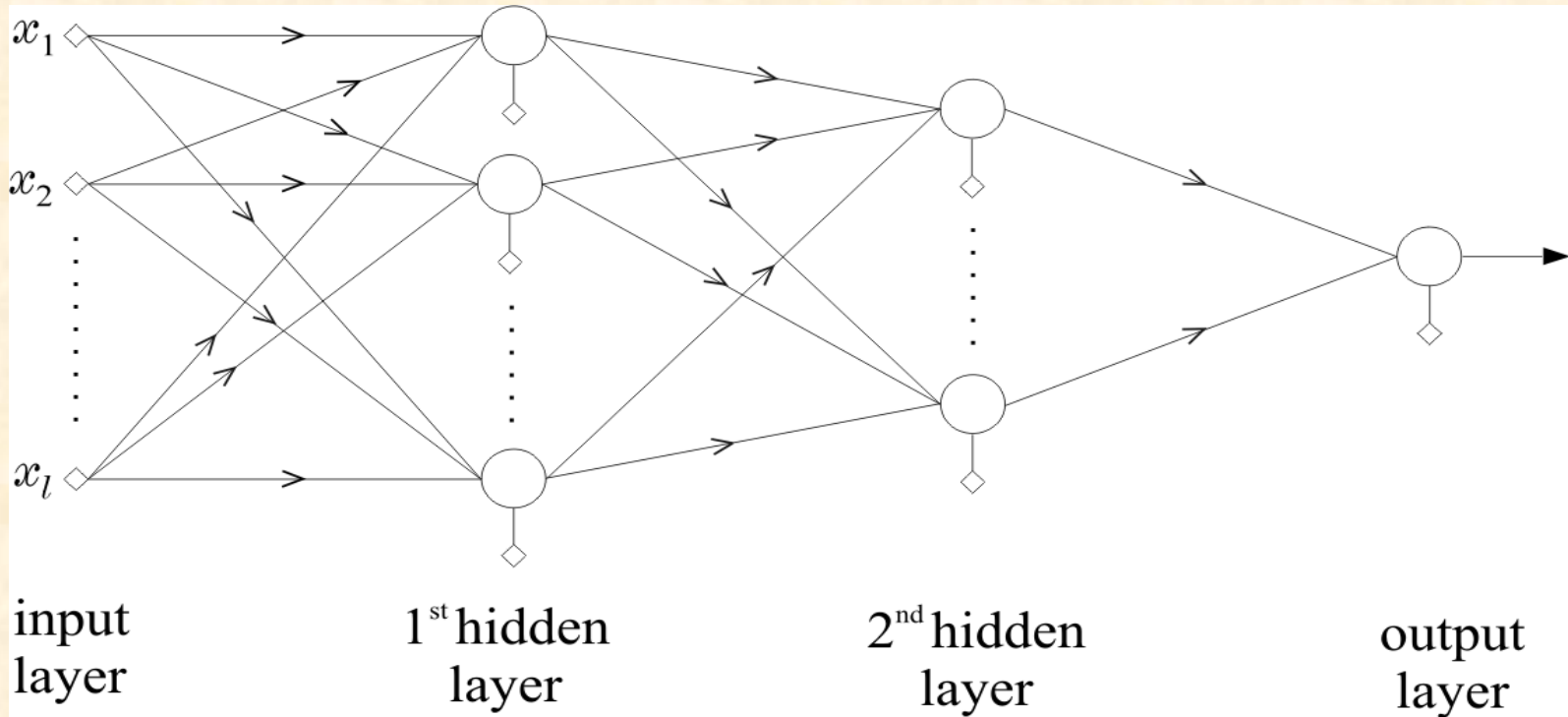
στην (+) πλευρά του $g_3(\underline{x})=0$



- Ο νευρώνας εξόδου υλοποιεί ένα υπερεπίπεδο στο μετασχηματισμένο χώρο, ο οποίος διαχωρίζει μερικές κορυφές από μερικές άλλες. Έτσι, το δίκτυο perceptron δύο επιπέδων έχει τη δυνατότητα να διαχωρίζει **κλάσεις που αποτελούνται από ενώσεις πολυεδρικών περιοχών.** Αλλά **ΟΧΙ ΟΠΟΙΕΣΔΗΠΟΤΕ** ενώσεις. Εξαρτάται από τη σχετική θέση των αντίστοιχων κορυφών.

❖ Δίκτυα perceptron τριών επιπέδων

➤ Η δομή τους



➤ Αυτά είναι ικανά να διαχωρίζουν κλάσεις οι οποίες αποτελούνται από **ΟΠΟΙΑΔΗΠΟΤΕ** ένωση πολυεδρικών περιοχών.

➤ Η ιδέα είναι παρόμοια μ' αυτή του προβλήματος XOR. Το δίκτυο υλοποιεί περισσότερα του ενός υπερεπίπεδα στο χώρο $\underline{y} \in R^p$

➤ Η λογική

- Για κάθε κορυφή, που αντιστοιχεί, ας πούμε, στην κλάση A κατασκεύασε ένα υπερεπίπεδο που αφήνει **ΑΥΤΗ** την **κορυφή** στην θετική πλευρά του (+) and **ΟΛΕΣ** τις άλλες κορυφές στην αρνητική πλευρά του (-).
- Ο νευρώνας εξόδου υλοποιεί μία πύλη OR

➤ Συνολικά:

Το πρώτο επίπεδο του δικτύου ορίζει τα **υπερεπίπεδα**, το δεύτερο επίπεδο ορίζει τις **περιοχές** και ο νευρώνας εξόδου ορίζει τις **κλάσεις**.

❖ Σχεδιάζοντας perceptrons πολλαπλών επιπέδων

- Ένας τρόπος είναι να υιοθετήσουμε την παραπάνω λογική.

Στην πράξη όμως σπάνια γνωρίζουμε ακριβώς τα όρια των κλάσεων

- Ο άλλος τρόπος είναι να επιλέξουμε μία δομή και να υπολογίσουμε τα συναπτικά βάρη της έτσι ώστε να **βελτιστοποιείται μία συνάρτηση κόστους**.