Αλγοριθμική Επιχειρησιακή Έρευνα
Τρίτη Εργασία

Σιώρος Βασίλειος
Ανδρινοπούλου Χριστίνα

Οκτώβριος 2019

Examples of Linear Programming Problems: Pattern classification

Έστω ότι έχουμε m αντικείμενα και θέλουμε να τα διαχωρίσουμε σε δύο διακριτές ομάδες. Κάθε αντικέιμενο ανήκει αυστηρά σε μία μόνο ομάδα.

Για να το επιτύχουμε αυτό πρέπει να είμαστε σε θέση να περιγράψουμε κάθε αντικείμενο. Επιλέγουμε καθολικά για όλα τα αντικείμενα n το πλήθος χαρακτηριστικά, τα οποία οργανώνουμε σε ένα vector. Συνεπώς, κάθε αντικείμενο περιγράφεται από ένα vector μεγέθους n.

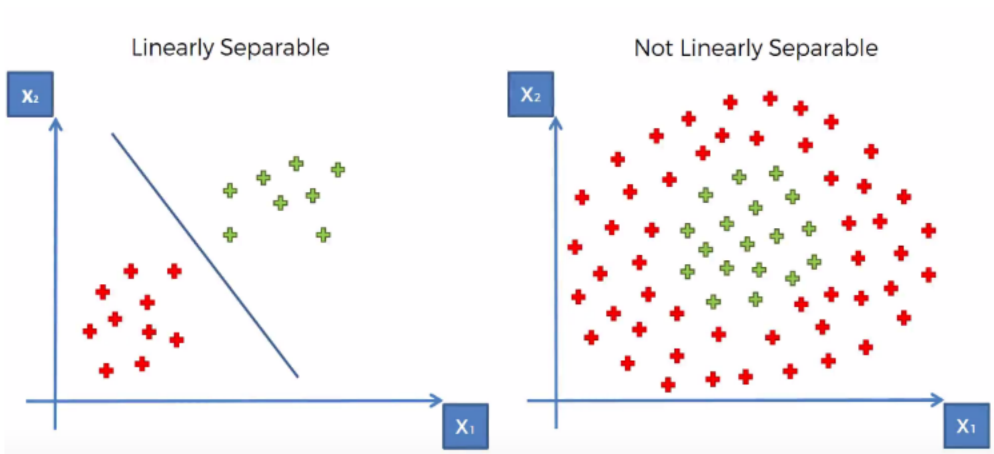
Ωστόσο, για να είμαστε σε θέση να μιλάμε για διαχωρισμό των αντικειμένων σε δύο διακριτές ομάδες, θα πρέπει να γνωρίζουμε ότι τα αντικείμενά μας όντως μπορούν να διαχωριστούν.

Έστω ότι έχουμε δύο set:

$$K = \{K^1, K^2, ..., K^k\} \subseteq \mathbb{R}^n$$

$$N = \{N^1, N^2, ..., N^n\} \subseteq \mathbb{R}^n$$

Τα K και N είναι διαχωρίσιμα αν υπάρχει hyperplane που τα διαχωρίζει.



Στην παραπάνω εικόνα παρατηρούμε ότι στην πρώτη περίπτωση τα αντικείμενα μπορούν να διαχωριστούν γραμμικά επιτυχώς, ενώ στη δεύτερη περίπτωση όχι.

Ένας πιο αυστηρός ορισμός για τον γραμμικό διαχωρισμό είναι ο εξής:
Δύο set αντικειμένων $K \subseteq \mathbb{R}^n$ και $N \subseteq \mathbb{R}^n$ είναι γραμμικά διαχωρίσιμα αν $\exists a \in \mathbb{R}^n, b \in \mathbb{R}$: $K \subseteq \{x \in \mathbb{R}^n : a^T \cdot x \geq b\}$ και $N \subseteq \{x \in \mathbb{R}^n : a^T \cdot x < b\}$.

Ένας linear classifier εκπαιδεύεται αρχικά με τα ήδη υπάρχοντα αντικείμενα κι έπειτα για κάθε νέο αντικείμενο αποφασίζει σε ποια από τις δύο ομάδες ανήκει. Πιο αναλυτικά, κατασκευάζεται κατάλληλο hyperplane τέτοιο ώστε τα αντικείμενα της ομαδας A να βρίσκονται εξ ολοκλήρου από τη μία πλευρά και τα αντικείμενα της ομάδας B από την άλλη.

Στόχος, λοιπόν, είναι να βρεθεί μία γραμμική συνάρτητη f τέτοια ώστε

$$f(K^i) \geq 0 \text{ για την ομάδα A}$$

και

$$f(N^i) < 0 \text{ για την ομάδα B}$$

όπου f ορίζεται ως $f(\vec{x}) = a_1 \cdot x_1 + a_2 \cdot x_2 + \dots + a_n \cdot x_n$.

Το παραπάνω μπορεί να οριστεί και ως πρόβλημα γραμμικού προγραμματισμού:
αν y είναι το περιθώριο μεταξύ του αντικειμένου K^i που βρίσκεται πιο κοντά στα αντικείμενα της ομάδας B και του αντικειμένου N^j που βρίσκεται πιο κοντά στα αντικείμενα της ομάδας A, σκοπός είναι να μεγιστοποιήσουμε το περιθώριο αυτό. Άρα

$$\max y$$

υπό τις προυποθέσεις:

$$a_1 \cdot k_1 + a_2 \cdot k_2 + \dots + a_n \cdot k_n + b - y \geq 0 \text{ για τα αντικείμενα της ομάδας A}$$

$$a_1 \cdot n_1 + a_2 \cdot n_2 + \dots + a_n \cdot n_n + b + y \leq 0 \text{ για τα αντικείμενα της ομάδας B}$$