Κατηγοριοποίηση Κρίνων με

Χρήση Νευροασαφούς Μοντέλου

Άνι Χατσατριάν

Τμήμα Μηχανικών Πληροφορικής, ΑΤΕΙΘ

Θεσσαλονίκη, Ελλάδα

achatsat@it.teithe.gr

Μιχαήλ Κοσματόπουλος

Τμήμα Μηχανικών Πληροφορικής, ΑΤΕΙΘ

Θεσσαλονίκη, Ελλάδα

mkosm@it.teithe.gr

*Ο κύριος σκοπός αυτής της εργασίας είναι η επίλυση ενός προβλήματος κατηγοριοποίησης. Ειδικότερα, με τη χρήση ενός ασαφούς μοντέλου Sugeno, επιτυγχάνεται ο διαχωρισμός κρίνων σε κατηγορίες. Παράλληλα, αναλύονται όλα όσα χρειάζεται να γνωρίζει κανείς, έτσι ώστε να καταφέρει να λύσει ένα πρόβλημα κατηγοριοποίησης, αφού εξηγούνται σε ένα ικανοποιητικό βαθμό όλες οι απαραίτητες έννοιες που έχουν να κάνουν με τα νευρωνικά δίκτυα, τα ασαφή συστήματα κα τους τρόπους διαχωρισμού των δεδομένων.*

Λέξεις κλειδιά: κατηγοριοποίηση; ασάφεια; νευρώνες; συσταδοποίηση; sugeno; νευροασάφεια; συστημα συμπερασμού; συναρτήσεις συμμετοχής

# Εισαγωγη

# Νευρωνικα Δικτυα

## Ορισμός

Τα νευρωνικά δίκτυα είναι μια ιδιαίτερη προσέγγιση στη δημιουργία συστημάτων με νοημοσύνη καθώς αποφεύγουν να αναπαραστήσουν ρητά τη γνώση και να υιοθετήσουν ειδικά σχεδιασμένους αλγόριθμους αναζήτησης. Αντίθετα, βασίζονται σε βιολογικά πρότυπα καθώς χρησιμοποιούν δομές και διαδικασίες που μιμούνται τις αντίστοιχες του ανθρωπίνου εγκέφαλου [βλαχαβας].

Ένα τεχνητό νευρωνικό δίκτυο αποτελεί ένα σύστημα επεξεργασίας πληροφοριών το οποίο έχει συγκεκριμένα χαρακτηριστικά απόδοσης κοινά με τα βιολογικά νευρωνικά δίκτυα. Τα τεχνητά νευρωνικά δίκτυα έχουν αναπτυχθεί για την γενίκευση των μαθηματικών μοντέλων που διέπουν τον ανθρώπινο νου ή τη νευρωνική βιολογία και είναι βασισμένα στις εξής αρχές:

* Η επεξεργασία της πληροφορίας συμβαίνει σε πολλά απλά στοιχεία που ονομάζονται νευρώνες
* Ανάμεσα στους νευρώνες περνούν σήματα μέσα από τα συνδετικά τους σημεία
* Κάθε συνδετικό σημείο είναι σχετισμένο με ένα βάρος, το οποίο σε ένα τυπικό νευρωνικό δίκτυο ενισχύει το σήμα που εκπέμπεται
* Κάθε νευρώνας εφαρμόζει μια μέθοδο ενεργοποίησης στην είσοδο του δικτύου του για να προσδιορίσει το σήμα εξόδου

## Γενικά χαρακτηριστικά

Ένα νευρωνικό δίκτυο χαρακτηρίζεται από την αρχιτεκτονική του, δηλαδή από το τρόπο με τον οποίο συνδέονται οι νευρώνες μεταξύ τους, από τη μέθοδο εκπαίδευσής του, δηλαδή από τον τρόπο που υπολογίζονται τα βάρη στις συνδέσεις και από τη συνάρτηση ενεργοποίησης του.

Τα στοιχεία που αποτελούν ένα νευρωνικό δίκτυο ονομάζονται νευρώνες. Κάθε νευρώνας είναι συνδεδεμένος με άλλους νευρώνες μέσω καθοδηγούμενων συνδέσεων. Η κάθε σύνδεση έχει το δικό της σχετικό βάρος. Τα βάρη εκπροσωπούν τη πληροφορία που χρησιμοποιείται από το δίκτυο για την επίλυση ενός προβλήματος.

Κάθε νευρώνας έχει μια εσωτερική κατάσταση η οποία ονομάζεται ενεργοποίηση και είναι μια συνάρτηση των εισόδων που έχει δεχτεί. Τυπικά, ένας νευρώνας στέλνει το σήμα ενεργοποίησης του σε διάφορους άλλους νευρώνες. Είναι αρκετά σημαντικό να σημειωθεί ότι ο νευρώνας μπορεί να στέλνει μόνο ένα σήμα κάθε φορά, αλλά το σήμα αυτό μεταφέρεται σε πολλούς άλλους νευρώνες.

## Τα οφέλη των νευρωνικών δικτύων

Είναι προφανές ότι ένα νευρωνικό δίκτυο αντλεί την υπολογιστική ισχύ του από τη μεγάλη παράλληλα διαμοιρασμένη δομή του, καθώς και από τη δυνατότητα μάθησης που κατέχει, δηλαδή στην ικανότητά του να γενικεύσει. Η γενίκευση αφορά τη παραγωγή λογικών εξόδων για εισόδους που δεν αντιμετωπίστηκαν κατά της διάρκεια της εκπαίδευσης.

Αυτά τα δύο στοιχεία επεξεργασίας δεδομένων δίνουν την δυνατότητα στα νευρωνικά δίκτυα να λύνουν περίπλοκα προβλήματα. Στην ουσία όμως τα νευρωνικά δίκτυα δεν παρέχουν λύσεις δουλεύοντας το καθένα ξεχωριστά.

Η χρήση των νευρωνικών δικτύων προσφέρει πολλά χρήσιμα χαρακτηριστικά και δυνατότητες, για αυτό το λόγο εφαρμόζονται σε πολλά προβλήματα. Μερικά από τα χαρακτηριστικά και τις δυνατότητες τους αναφέρονται παρακάτω.

### Μη γραμμικότητα

Ένας τεχνητός νευρώνας μπορεί να είναι γραμμικός ή μη-γραμμικός. Μη-γραμμικό είναι ένα νευρωνικό δίκτυο το οποίο αποτελείται από τη διασύνδεση μη-γραμμικών νευρώνων.

### Αντιστοίχηση εισόδων/εξόδων

Ένα δημοφιλές σύστημα εκμάθησης είναι η εκμάθηση υπό επίβλεψη. Η τεχνική αυτή εμπλέκει τη τροποποίηση των βαρών του νευρωνικού δικτύου με την εφαρμογή διάφορων δεδομένων εκπαίδευσης.

### Προσαρμοστικότητα

Τα νευρωνικά δίκτυα έχουν ενσωματωμένα τη δυνατότητα της προσαρμογής των συναπτικών βαρών τους σε αλλαγές που συμβαίνουν στο περιβάλλον τους. Ειδικότερα, ένα νευρωνικό δίκτυο που είναι εκπαιδευμένο να λειτουργεί σε ένα συγκεκριμένο περιβάλλον μπορεί με εύκολο τρόπο να επανεκπαιδευτεί και να αντιμετωπίσει μικρές αλλαγές που συνέβησαν στο περιβάλλον.

### Άμεση απόκριση

Στην κατηγοριοποίηση προτύπων, ένα νευρωνικό δίκτυο μπορεί να σχεδιαστεί με τέτοιο τρόπο, έτσι ώστε όχι μόνο να είναι σε θέση να επιλέξει ένα πρότυπο, αλλά έχει επίσης τη δυνατότητα να παρέχει πληροφορίες για τη εγκυρότητα της απόφασής του. Αυτό το χαρακτηριστικό μπορεί να χρησιμοποιηθεί και για την απόρριψη άλλων προτύπων.

### Συνάφεια της πληροφορίας

Η γνώση παρουσιάζεται με τη δομή και τη κατάσταση ενεργοποίησης του νευρωνικού δικτύου. Κάθε νευρώνας στο δίκτυο επηρεάζεται από τη δραστηριότητα των άλλων νευρώνων που υπάρχουν στο δίκτυο. Επομένως, η συνάφεια της πληροφορίας είναι φυσικό χαρακτηριστικό που κατέχουν τα νευρωνικά δίκτυα.

### Ανοχή στα σφάλματα

Το υλικό κομμάτι του νευρωνικού δικτύου έχει τη δυνατότητα να είναι έμφυτα ανεκτικό στα σφάλματα. Η απόδοση του δικτύου πέφτει ομαλά όταν αρχίζει να αντιμετωπίζει πολύπλοκα προβλήματα που απαιτούν πολλούς υπολογισμούς.

Για παράδειγμα, εάν ένας νευρώνας πάθει κάποια βλάβη, το πρότυπο που ανακαλείται από αυτό θα έχει επηρεαστεί. Παρόλα αυτά, εξαιτίας της κατανεμημένης φύσης της πληροφορίας στο δίκτυο, θα πρέπει οι βλάβες να είναι πολλές και σοβαρές για να σταματήσει να ανταποκρίνεται.

### Δυνατότητα υλοποίησης VLSI αρχιτεκτονικής

H μαζικά παράλληλη φύση ενός νευρωνικού δικτύου το καθιστά γρήγορο στον υπολογισμό διάφορων αποτελεσμάτων. Αυτό το χαρακτηριστικό ορίζει τα νευρωνικά δίκτυα κατάλληλα για την υλοποίηση αρχιτεκτονικών τύπου VLSI (Very Large Scale Intergraded).

### Νευροβιολογική αναλογία

Η αρχιτεκτονική και η λειτουργία του νευρωνικού δικτύου είναι εμπνευσμένη από τον εγκέφαλό του ανθρώπου, το οποίο αποδεικνύει ότι η παράλληλη επεξεργασία με ανοχή στα σφάλματα, όχι μόνο είναι δυνατή, αλλά εξαιρετικά γρήγορα και ισχυρή.

Οι νευροβιολόγοι ερευνούν τα τεχνητά νευρωνικά δίκτυα με την ίδια περιέργεια που ερευνούν τα βιολογικά νευρωνικά δίκτυα για να μπορέσουν να ερμηνεύσουν τα νευροβιολογικά φαινόμενα που αντιμετωπίζουν οι άνθρωποι. Από την άλλη πλευρά, οι μηχανικοί μελετούν τη νευροβιολογία για να βρουν νέες ιδέες που θα είναι θέση να λύνουν μεγάλα και πολύπλοκα προβλήματα με πιο απλούς τρόπου από τους ήδη γνωστούς.

# http://blog.peltarion.com/img/fuzz/fuzz7.pngΕισαγωγή στα Ασαφή Συστήματα

Σχ. 2. Παράδειγμα φράκτη

## Ασαφή σύνολα

Το 1965 ο Lotfi Zadeh, ένας καθηγητής στο πανεπιστήμιο του Μπέρκλεϋ [1] επινόησε την ασαφή λογική σαν έναν τρόπο επεξεργασίας δεδομένων στο οποίο οι μεταβλητές μπορούν να ανήκουν μερικώς σε ένα ή περισσότερα σύνολα (ενώ στα παραδοσιακά δυαδικά σύνολα, οι μεταβλητές μπορούν είτε να ανήκουν, είτε να μην ανήκουν σε ένα σύνολο – τίποτα ενδιάμεσο). [2] Στα παραδοσιακά σύνολα, αν υποθέσουμε ότι οι άνθρωποι που έχουν ύψος από 1.9 m και πάνω, ανήκουν στο σύνολο των ψηλών ανθρώπων, ενώ οι υπόλοιποι ανήκουν στο σύνολο των κοντών ανθρώπων, τότε ένας άνθρωπος που έχει ύψος 1.89 m, θα ανήκει στους κοντούς ανθρώπους. Αυτό έρχεται ενάντια στην κοινή λογική. [3]

Η βασική ιδέα της θεωρίας αυτής είναι ότι η διαδικασία της μετατροπής διακριτών μεγέθων σε ασαφή (fuzzification) επιτρέπει τη γενίκευση μιας διακριτής (distinct) θεωρίας σε συνεχόμενη (continuous)[βλαχαβας].

## Ασαφής Συλλογιστική

Ο όρος ασαφής συλλογιστική (fuzzy reasoning) αφορά την εξαγωγή συμπερασμάτων (ενδεχομένως σε ασαφή μορφή) με χρήση ασαφών κανόνων. Στη γενική περίπτωση, η χρήση ασαφούς συλλογιστικής απαιτεί την ύπαρξη μια ασαφούς λεκτικής περιγραφής του προβλήματος και περιλαμβάνει τα εξής τέσσερα στάδια:

* Υπολογισμό της συνάρτησης συνεπαγωγής για κάθε εμπλεκόμενο κανόνα
* Παραγωγή επιμέρους αποτελεσμάτων μέσω κάποιας συλλογιστικής διαδικασίας
* Συνάθροιση των επιμέρους αποτελεσμάτων
* Αποσαφήνιση αποτελεσμάτων

## Αποσαφήνιση

Κατά την αποσαφήνιση (defuzzification) το ασαφές αποτέλεσμα του βήματος της συνάθροισης μετατρέπεται σε πραγματική (crisp) τιμή. Το στάδιο αυτό είναι υποχρεωτικό όταν το αποτέλεσμα χρησιμοποιείται για έλεγχο συσκευών οι οποίες απαιτούν αριθμητικές τιμές. Αν και υπάρχουν περισσότερες από 30 μέθοδοι αποσαφήνισης, περισσότερο γνωστές είναι οι maximum και centroid.

## Συναρτήσεις συμμετοχής

Ο τρόπος με τον οποίο ορίζεται το πόσο πολύ μια μεταβλητή ανήκει σε ένα σύνολο, γίνεται με την χρήση μιας συνάρτησης συμμετοχής για κάθε σύνολο. [4]

Το σύνολο τιμών μιας συνάρτησης συμμετοχής ανήκει στο πεδίο [0, 1]. Όσο μεγαλύτερη η τιμή της συνάρτησης συμμετοχής, τόσο μεγαλύτερο βαθμό συμμετοχής έχει η μεταβλητή. Όπως φαίνεται στο **Error! Reference source not found.**, μια μεταβλητή μπορεί να ανήκει σε παραπάνω από ένα σύνολα. Στο σημείο που υπάρχει η κάθετος, η μεταβλητή ανήκει 90% στο κρύο σύνολο, 10% στο θερμό σύνολο, και 0% στο ζεστό σύνολο.

Στον **Error! Reference source not found.** φαίνονται μερικές από τις πιο διαδεδομένες συναρτήσεις συμμετοχής. [5]

## Ασαφής Συλλογιστική

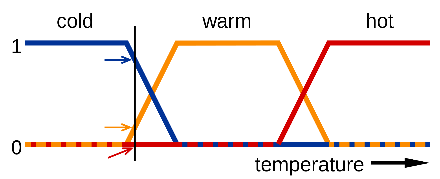
Ο όρος ασαφής συλλογιστική (fuzzy reasoning) αφορά την εξαγωγή συμπερασμάτων (ενδεχομένως σε ασαφή μορφή) με χρήση ασαφών κανόνων. Στη γενική περίπτωση, η χρήση ασαφούς συλλογιστικής απαιτεί την ύπαρξη μια ασαφούς λεκτικής περιγραφής του προβλήματος και περιλαμβάνει τα εξής τέσσερα στάδια:

* Υπολογισμό της συνάρτησης συνεπαγωγής για κάθε εμπλεκόμενο κανόνα
* Παραγωγή επιμέρους αποτελεσμάτων μέσω κάποιας συλλογιστικής διαδικασίας
* Συνάθροιση των επιμέρους αποτελεσμάτων
* Αποσαφήνιση αποτελεσμάτων

## Φράκτες

Στην καθημερινή ζωή χρησιμοποιούμε εκφράσεις όπως «λίγο θερμός», «πάρα πολύ θερμός», κτλ. Μπορούμε με ευκολία να ορίσουμε τέτοιους όρους βασιζόμενοι στις ήδη υπάρχουσες συναρτήσεις συμμετοχής. Οι φράκτες (hedges) μας παρέχουν αυτή τη δυνατότητα, και δεν είναι τίποτα άλλο, παρά κάποιοι τροποποιητές που εφαρμόζονται στις συναρτήσεις συμμετοχής.

Για παράδειγμα, έστω ότι έχουμε τρεις συναρτήσεις συμμετοχής που περιγράφουν το ύψος ενός ανθρώπου (κοντός, μέσος, ψηλός) και θέλουμε να ορίσουμε και τις συναρτήσεις συμμετοχής «πολύ κοντός» και «πολύ ψηλός». Μπορούμε σύμφωνα με τον **Error! Reference source not found.** να υψώσουμε τις αντίστοιχες συναρτήσεις στο τετράγωνο, για να πάρουμε το επιθυμητό αποτέλεσμα. [6]



Σχ. 1. Συναρτήσεις συμμετοχής

## Πράξεις συνόλων

Όπως και στα κανονικά σύνολα, έτσι και στα ασαφή μπορούμε να πραγματοποιήσουμε πράξεις μεταξύ των συνόλων. Έστω οι συναρτήσεις συμμετοχής και που ορίζουν τα ασαφή σύνολα Α και Β αντίστοιχα. Τότε μπορούμε να ορίσουμε τις πράξεις μεταξύ συνόλων ως εξής: [7]

* Ένωση (OR)



* Τομή (AND)

(2)

* Συμπλήρωμα (NOT)

(3)

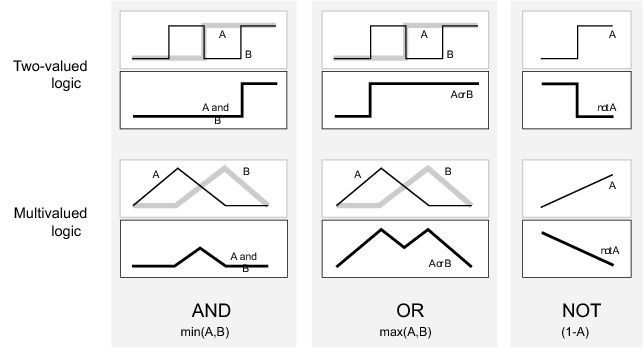
# Ασαφή συστήματα συμπερασμού

Ένα ασαφές σύστημα συμπερασμού (Fuzzy Interference System – FIS) είναι ένα σύστημα που χρησιμοποιεί θεωρία ασαφών συνόλων για την αντιστοίχηση τιμών από έναν χώρο εισόδου σε έναν χώρο εξόδου. Αυτό γίνεται με την χρήση απλών ΕΑΝ-ΤΟΤΕ κανόνων (IF-THEN rules) της ανθρώπινης γλώσσας. [8] Για παράδειγμα, ο αυτόματος πιλότος (cruise control) στα καινούργια αυτοκίνητα είναι ένα σύστημα το οποίο έχει σκοπό να κρατάει σταθερή την ταχύτητα του οχήματος ελέγχοντας το γκάζι. Ένα τέτοιο σύστημα θα μπορούσε να χρησιμοποιεί τον εξής κανόνα:

*Εάν η ταχύτητα είναι* ***μικρή*** *και η επιτάχυνση* ***μικραίνει****, τότε πάτα* ***λίγο*** *γκάζι.*

## Δομή ενός ασαφούς συστήματος

Όλα ξεκινούν από τον ασαφοποιητή. Ο ασαφοποιητής δέχεται την είσοδο του συστήματος (που είναι καθαροί (crisp) αριθμοί) και την μετατρέπει σε ασαφή σύνολα με την βοήθεια των συναρτήσεων συμμετοχής. Η έξοδος του ασαφοποιητή είναι η είσοδος της μηχανής συμπερασμού, η οποία προσομοιώνει τον ανθρώπινο τρόπο σκέψης βγάζοντας ασαφή συμπεράσματα βάσει των εισόδων και των ΕΑΝ-ΤΟΤΕ κανόνων που βρίσκονται αποθηκευμένοι στη βάση γνώσεων. Τέλος, το αποτέλεσμα της μηχανής συμπερασμού περνάει από τον αποσαφοποιητή για να βγει μια καθαρή τιμή.



Σχ. 3. Πράξεις συνόλων

1. Διάφοροι φράκτεσ

|  |  |  |
| --- | --- | --- |
| **Φράκτης** | **Τελεστής** | **Αποτέλεσμα** |
| Λίγο |  | http://blog.peltarion.com/img/fuzz/h_a_little.png |
| Ελαφρώς |  | http://blog.peltarion.com/img/fuzz/h_slightly.png |
| Πολύ |  | http://blog.peltarion.com/img/fuzz/h_very.png |
| Υπερβολικά |  | http://blog.peltarion.com/img/fuzz/h_extremely.png |
| Υπερβολικά πολύ |  | http://blog.peltarion.com/img/fuzz/h_very_very.png |
| Περίπου |  | http://blog.peltarion.com/img/fuzz/h_somewhat.png |

## Ασαφές σύστημα συμπερασμού Mamdani

Το Mamdani FIS είναι μία από τις πιο διαδεδομένες μεθοδολογίες στον χώρο των ασαφών συστημάτων και ήταν από τα πρώτα συστήματα ελέγχου που χρησιμοποίησαν θεωρία ασαφών συνόλων. Προτάθηκε το 1975 από τον Ebrahim Mamdani σαν μια προσπάθεια ελέγχου ενός συνδυασμού ατμομηχανής και βραστήρα χρησιμοποιώντας μια ομάδα από

1. Συνηθισμένεσ συναρτήσεισ συμμετοχήσ

|  |  |  |
| --- | --- | --- |
| Ονομασία | Τύπος | Γράφημα |
| Τριγωνική | Description of triangular function | triangular function graph |
| Τραπεζοειδή | Description of trapezoidal function | trapezoidal function graph |
| Γκαουσιανή | Description of gaussian function | gaussian function graph |

γλωσσικούς κανόνες που προμηθεύτηκαν από κάποιον ειδικό. [http://www.mathworks.com/help/fuzzy/what-is-mamdani-type-fuzzy-inference.html]

Για την δημιουργία και τον υπολογισμό ενός συστήματος Mamdani πρέπει να ακολουθηθούν τα εξής βήματα:

### Δημιουργία ασαφών κανόνων

Όπως περιγράφηκαν παραπάνω, οι ασαφείς κανόνες είναι μια συλλογή από προτάσεις που περιγράφουν πως θα πρέπει να παρθούν οι αποφάσεις από το FIS όσων αφορά την είσοδο ή τον έλεγχο της εξόδου.

Με τη χρήση των συναρτήσεων συμμετοχής προσδιορίζονται οι διάφορες τιμές των προτάσεων που προκύπτουν (υψηλό, χαμηλό, κτλ).

### Ασαφοποίηση

Στο βήμα της ασαφοποίησης γίνεται η αντιστοίχηση των εισόδων σε τιμές από 0 μέχρι 1 με την χρήση των συναρτήσεων συμμετοχής εισόδου.

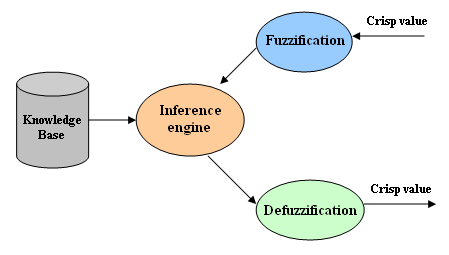
### Ασαφείς συνδυασμοί

Οι κανόνες μπορούν να περιέχουν εκφράσεις όπως «or», «and», και «not». Οι πράξεις αυτές, εκτός από την κλασική τους υλοποίηση που αναφέρθηκε παραπάνω στην ενότητα των πράξεων ασαφών συνόλων, μπορούν να οριστούν και με διαφορετικό τρόπο.

Συγκεκριμένα, στην πράξη AND, αντί για τη χρήση του ελάχιστου των δύο τιμών των συναρτήσεων συμμετοχής, μπορεί να χρησιμοποιηθεί το γινόμενό τους (product). Αντίστοιχα, στην πράξη OR, στη θέση του μεγίστου των τιμών των συναρτήσεων συμμετοχής, μπορεί να χρησιμοποιηθεί το άθροισμά τους μείον το γινόμενό τους.

### Εξαγωγή αποτελέσματος

Μετά τον υπολογισμό της δύναμης του ασαφούς κανόνα, γίνεται η αποκοπή της εξόδου στο σημείο της δύναμης του κανόνα (βλ. σχήμα).



Σχ. 4. Δομή ασαφούς συστήματος

### Συνδυασμός των αποτελεσμάτων

Δημιουργείται η τελική έξοδος από τον συνδυασμό των εξόδων των κανόνων. Συχνά, αλλά όχι πάντα, γίνεται η χρήση της πράξης OR.

### Αποσαφοποίηση

Τέλος, για να παραχθεί μια καθαρή τιμή εξόδου από το FIS, πρέπει να χρησιμοποιηθεί μια μέθοδος αποσαφοποίησης. Υπάρχουν διάφορες τέτοιες τεχνικές, με τις πιο διαδεδομένες να είναι το κέντρο βάρους και το μέσο των μεγίστων τιμών.

## Πλεονεκτήματα για το μοντέλο Mamdani

* Είναι διαισθητική μέθοδος
* Είναι κοινά αποδεκτό
* Είναι κατάλληλη για την είσοδο από ανθρώπους

## Ασαφές σύστημα συμπερασμού Sugeno

Το μοντέλο Sugeno (γνωστό και ως TSK από τα αρχικά των κατασκευαστών Takagi-Sugeno-Kang) παρουσιάζει πολλές ομοιότητες με το μοντέλο Mamdani. Η κύρια διαφορά βρίσκεται στη διαδικασία της εξαγωγής του αποτελέσματος, καθώς στο μοντέλο Sugeno δεν υπάρχει συνάρτηση συμμετοχής για την έξοδο. Στη θέση αυτής, υπάρχει ένας καθαρός αριθμός, ο οποίος προκύπτει από τον πολλαπλασιασμό κάθε εισόδου με μία σταθερά, και στη συνέχεια το άθροισμα των αποτελεσμάτων τους. Ένας τυπικός κανόνας στο μοντέλο Sugeno έχει την εξής μορφή:

Εάν Είσοδος και Είσοδος ,

τότε Έξοδος

Επίσης, υπάρχουν αλγόριθμοι που μπορούν να χρησιμοποιηθούν για την βελτιστοποίηση του FIS, ενώ στο Mamdani δεν υπάρχουν. Ένας αλγόριθμος για την βελτιστοποίηση του FIS είναι το προσαρμοστικό νευρο-ασαφές σύστημα συμπερασμού (adaptive neuro-fuzzy inference system – ANFIS).

## Πλεονεκτήματα για το μοντέλο Sugeno

* Είναι υπολογιστικά αποδοτικό
* Δουλεύει καλά με τις γραμμικές τεχνικές
* Δουλεύει καλά με τις τεχνικές βελτιστοποίησης και προσαρμοστικότητας
* Παρέχει εγγύηση για την συνέχεια της εξόδου
* Είναι κατάλληλο για τη μαθηματική ανάλυση

# Συσταδοποίηση (Clustering)

## Απλός Ορισμός

Πολλά προβλήματα αφορούν την ομαδοποίηση αντικείμενων με βάση κάποιο κοινό χαρακτηριστικό. Η διαδικασία εύρεσης τέτοιων ομάδων (συστάδων), κατά την οποία κάθε αντικείμενο μιας συστάδας είναι όμοιο με κάποιο τρόπο με τα υπόλοιπα και διαφορετικό από τα αντικείμενα των άλλων συστάδων, ονομάζεται συσταδοποίηση.

Η συσταδοποίηση χαρακτηρίζεται και ως μάθηση χωρίς επίβλεψη, επειδή με τη χρήση του κατάλληλου αλγορίθμου επιτυγχάνεται η ομαδοποίηση των αντικειμένων χωρίς να είναι γνωστές από πριν οι ομάδες στις οποίες χωρίζονται τα αντικείμενα.

## Μαθηματικός Ορισμός

Έχοντας ένα σύνολο αντικειμένων ζητούνται ομάδες , με έτσι ώστε:

Έτσι οι ομάδες αποτελούν τη διαμέριση του συνόλου και κάθε αντικείμενο ανήκει αυστηρά σε μία και μόνο ομάδα.

Εναλλακτικά μπορεί να οριστεί η ασαφής συσταδοποίηση κατά την οποία, κάνοντας χρήση ασαφών συνόλων ορίζονται συναρτήσεις συμμετοχής για . Οι συναρτήσεις ποσοτικοποιούν τη βεβαιότητα που έχουμε για το αν κάποιο αντικείμενο ανήκει στην ομάδα .

## Στάδια Συσταδοποίησης

Τα βασικά βήματα της συσταδοποίησης είναι τα εξής:

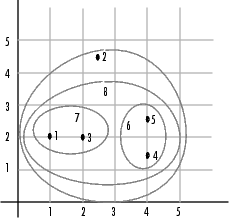
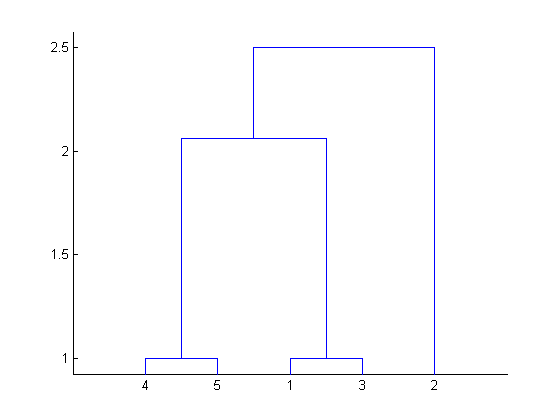
* Αναπαράσταση των προτύπων και εξαγωγή των χαρακτηριστικών τους για την βέλτιστη ομοιογένεια της κάθε συστάδας.
* Επιλογή του κατάλληλου αλγορίθμου με βάση την ομοιότητα των αντικειμένων και της συνάρτησης κόστους, καθώς και τη δυνατότητα καθορισμού του σχήματος της συσταδοποίησης.
* Ομαδοποίηση των δεδομένων και αξιολόγηση των αποτελεσμάτων του αλγορίθμου συσταδοποίησης σύμφωνα και τα κατάλληλα κριτήρια ορθότητας συσταδοποίησης. Σε κάποιες περιπτώσεις μπορεί να χρειαστεί να αφαιρεθούν κάποια δεδομένα.
* Αξιολόγηση του τελικού αποτελέσματος και ερμηνεία των αποτελεσμάτων.

## Αλγόριθμοι Συσταδοποίησης

Για την επίτευξη της συσταδοποίησης έχουν αναπτυχθεί εκατοντάδες αλγόριθμοι και συνεχίζονται να δημοσιεύονται νέοι μέχρι και σήμερα. Για αυτόν ακριβός το λόγο είναι αρκετά δύσκολη η επιλογή του καλύτερου αλγορίθμου για το κάθε πρόβλημα που αντιμετωπίζεται.

Για την διευκόλυνση της επιλογής του καταλληλότερου αλγορίθμου έχουν προταθεί κάποια κριτήρια που βασίζονται στο τρόπο σχηματισμού των ομάδων, τη δομή που έχουν τα δεδομένα προς επεξεργασία και στην ευαισθησία που έχει ο αλγόριθμος σε αλλαγές που δεν επηρεάζουν τη δομή των δεδομένων.

Παρακάτω γίνεται μια αναφορά στις πιο διαδεδομένες τεχνικές συσταδοποίησης που είναι χωρισμένες με βάση τα παραπάνω κριτήρια.



Σχ. 4. Παραδείγματα ιεραρχικής συσταδοποίησης

### Ιεραρχική Συσταδοποίηση

Σχ. 5. Αναπαράσταση αλγορίθμου k-means

Στόχος αυτής της μεθόδου συσταδοποίησης είναι η δημιουργία μιας ιεραρχίας από ομάδες. Για την εφαρμογή του χρησιμοποιούνται δύο βασικές προσεγγίσεις. Στη πρώτη προσέγγιση οι ομάδες συγχωνεύονται μεταξύ τους όσο ανεβαίνουμε τα σκαλοπάτια της ιεραρχίας, ενώ στη δεύτερη ξεκινώντας από μία ομάδα γίνονται αναδρομικά διαχωρισμοί καθώς κατεβαίνουμε την ιεραρχία. Οι δύο μέθοδοι μπορούν να περιγραφούν και ως «κάτω-προς-τα-επάνω» και «επάνω-προς-τα-κάτω».

Για την συγχώνευση ή το διαχωρισμό των ομάδων γίνεται η χρήση διάφορων συνδετικών κριτηρίων και μετρικών, όπως η Ευκλείδεια μετρική.

Γενικά η Ιεραρχική συσταδοποίηση είναι ιδανική για μικρό μέγεθος δεδομένων και τα αποτελέσματά της παρουσιάζονται με τη μορφή δενδροδιαγράμματος.

### Διαμεριστική Συσταδοποίηση

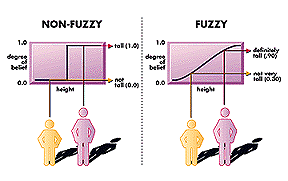
Στη διαμεριστική συσταδοποίηση δημιουργούνται αρχικά k ομάδες από το σύνολο των αντικειμένων Ν και στη συνέχεια βελτιστοποιούνται. Ο αλγόριθμος φτάνει στο τέλος του όταν δεν υπάρχει αλλαγή στη κατάσταση των αντικειμένων. Το κύριο πρόβλημα που έχει αυτή η τεχνική είναι η σωστή επιλογή των ομάδων και του αριθμού k.

### Ασαφής Συσταδοποίηση

Η ασαφής συσταδοποίηση είναι μια τεχνική που επιτρέπει το κάθε πρότυπο να ανήκει σε όλες τις ομάδες σε κάποιο βαθμό. Με τον ορισμό μια συνάρτησης συμμετοχής δηλώνονται οι βαθμοί συμμετοχής κάθε προτύπου στην κάθε ομάδα. Οι τιμές που μπορεί να πάρει ο βαθμός συμμετοχής είναι από μηδέν ως ένα.

Όσο πιο κοντά στο ένα είναι ο βαθμός συμμετοχής του προτύπου σε μια ομάδα, τόσο πιο μεγάλη σιγουριά υπάρχει για τη συμμετοχή του συγκεκριμένου προτύπου στη συγκεκριμένη ομάδα. Αντίθετα, όσο πιο κοντά στο μηδέν είναι ο βαθμός συμμετοχής του προτύπου σε μια ομάδα, τόσο μεγαλώνουν οι αμφιβολίες για τη συμμετοχή του σε αυτήν.

Φυσικά, δεν είναι απαραίτητο όλες οι ομάδες να έχουν ασαφή χαρακτήρα, καθώς είναι πιθανό να προκύψουν και απόλυτες ομάδες στη περίπτωση που όλα τα στοιχεία κάποιας ομάδας έχουν βαθμό συμμετοχής ένα.



Σχ. 6. Διαφορά ασαφών με μη-ασαφών τιμών

Η διαφορά της ασαφής συσταδοποίησης από τις κλασσικές τεχνικές οφείλεται στο γεγονός ότι η μεταβλητή που υποδηλώνει το βαθμό συμμετοχής ενός προτύπου σε κάποια ομάδα είναι συνεχής, ενώ στις κλασσικές τεχνικές παίρνει διακριτές τιμές.

# Ομαδοποίηση (Classification)

## Ορισμός

Με τον όρο ομαδοποίηση αναφερόμαστε στη διαδικασία της ενσωμάτωσης ενός νέου αντικειμένου σε μια ομάδα με τη βοήθεια της εκπαίδευσης στην οποία, γίνεται χρήση μιας συλλογής δεδομένων που προέκυψε από άλλα αντικείμενα των οποίων η ομαδοποίηση είναι γνωστή.

Η ομαδοποίηση που προκύπτει από μια σειρά δεδομένων των οποίων ο διαχωρισμός είναι γνωστός αναφέρεται και ως αναγνώριση προτύπων ή μάθηση υπό επίβλεψη (για τον διαχωρισμό του από την μάθηση χωρίς επίβλεψη ή αλλιώς συσταδοποίηση, στην οποία η ομαδοποίηση προκύπτει από τα δεδομένα).

# Παρουσιαση προβληματος

Ένα κλασσικό πρόβλημα ομαδοποίησης είναι ο διαχωρισμός λουλουδιών σε κατηγορίες. Συγκεκριμένα σε αυτήν την εργασία, το πρόβλημα που αντιμετωπίζεται είναι η κατηγοριοποίηση διαφόρων δειγμάτων κρίνων σε τρεις κατηγορίες (setosa, versicolor και virginica).

Το πρόβλημα αυτό είναι ιδιαίτερα γνωστό και χρησιμοποιείται για την αξιολόγηση διάφορων αλγορίθμων ταξινόμησης. Τα δεδομένα που χρησιμοποιούνται για την αξιολόγηση των αλγορίθμων, παρουσιάστηκαν από τον Ronald Fisher το 1936. Πρόκειται για μια συλλογή 150 δειγμάτων που έχουν συλλεχθεί από τον Edgar Anderson για την ανάλυση των μορφολογικών παραλλαγών που εμφανίζουν οι κρίνοι.

Η συλλογή περιέχει 50 δείγματα από κάθε κατηγορία. Για κάθε δείγμα συλλέχθηκαν τέσσερα χαρακτηριστικά: το μήκος και το πλάτος των σεπάλων και των πετάλων τους. Επίσης καταγράφτηκε σε ποια κατηγορία ανήκει το κάθε δείγμα.

# Επίλυση Προβλήματος

Τα δεδομένα του προβλήματος φορτώθηκαν από την εντολή load iris.dat, με την οποία δημιουργείται ένας πίνακας ο οποίος περιέχει τις 150 εγγραφές των αντικειμένων ταξινομημένες με βάσει τη κατηγορία των δειγμάτων. Οι τέσσερις πρώτες στήλες του πίνακα περιέχουν το μήκος των πετάλων, πλάτος πετάλων, μήκος σεπάλων και πλάτος σεπάλων. Στην πέμπτη στήλη αναφέρεται ένας αριθμός που προσδιορίζει την κατηγορία στην οποία ανήκει το κάθε δείγμα.

Σύμφωνα με το πρόβλημα, το μοντέλο πρέπει να εκπαιδευτεί με 75 δείγματα και στη συνέχεια να γίνει η αξιολόγηση της εκπαίδευσης με τα υπόλοιπα 75. Για αυτό το λόγο χρειάστηκε να ανακατευτεί ο πίνακας, έτσι ώστε να χρησιμοποιηθούν στην εκπαίδευση δείγματα από όλες τις κατηγορίες. Στη συνέχεια έγινε ο διαχωρισμός των δεδομένων για εκπαίδευση και των δεδομένων για αξιολόγηση, αποθηκεύοντας τα πρώτα 75 δείγματα σε μία μεταβλητή (irisTrain) και τα υπόλοιπα σε μια δεύτερη (irisTest).

1. συγκριση αποτελεσματων

|  |  |  |
| --- | --- | --- |
| **Μεθοδολογία** | **Σφάλμα Εκπαίδευσης** | **Σφάλμα Ελέγχου** |
| Grid Partitioning | 0.3227 | 0.2504 |
| Sub. Clustering | 0.1929 | 0.1639 |
| Grid Partitioning με στρογγυλοποίηση | 0.3055 | 0.2828 |
| Sub. Clustering με στρογγυλοποίηση | 0.1155 | 0.1155 |

1. Grid partitioning: 500 εποχές, 3 gaussmf συναρτήσεις ενεργοποίησης, γραμμική έξοδος
2. Subtractive clustering: 2000 εποχές, 1.3 radius, 1.5 squash factor, 0.1 accept ratio, 0.1 reject ratio

Αρχικά έγινε η χρήση του γραφικού περιβάλλοντος του Fuzzy Toolbox. Με τη δημιουργία ενός FIS τύπου Sugeno, ορίστηκαν οι τέσσερις είσοδοι (μήκος και πλάτος σεπάλων και πετάλων) και η έξοδος (η κατηγορία στην οποία ανήκει το αντικείμενο). Μετά, με τη χρήση του Anfis editor, αφού φορτώθηκαν οι δυο μεταβλητές που περιέχουν τα δείγματα, διεκπαιρεώθηκε η εκπαίδευση και η αξιολόγηση του FIS.

Για την δημιουργία του FIS δοκιμάστηκαν δύο μέθοδοι, η μέθοδος grid partitioning και η μέθοδος subtractive clustering. Τα αποτελέσματα των παραπάνω μεθόδων παρουσίασαν μεγάλο μέσο τετραγωνικό σφάλμα, εξαιτίας του γεγονότος ότι οι παραγόμενες τιμές δεν προσδιόριζαν διακριτά την κατηγορία στην οποία ανήκει το δείγμα.

Ειδικότερα με τη χρήση του grid partitioning, το μέσο τετραγωνικό σφάλμα έβγαινε υπερβολικά μεγάλο, εξαιτίας μερικών μεμονωμένων τιμών που έβγαιναν έξω από τα όρια των κλάσεων. Στην περίπτωση του subtractive clustering, το μέσο τετραγωνικό σφάλμα ήταν σχετικά χαμηλότερο από το σφάλμα του grid partitioning. Παρόλα αυτά, δεν ήταν η πιο βέλτιστη λύση για τον λόγο που αναφέρθηκε παραπάνω.

Σε μια δεύτερη προσπάθεια για την αντιμετώπιση του προβλήματος των μη-ακέραιων τιμών στα αποτελέσματα, αποφασίστηκε να μην χρησιμοποιηθεί το γραφικό περιβάλλον του Fuzzy Toolbox. Στη θέση αυτού έγινε η συγγραφή του κώδικα που αντιστοιχεί σε όλες τις λειτουργίες του γραφικού περιβάλλοντος του Fuzzy Toolbox, με την διαφορά ότι στο τέλος έγινε στρογγυλοποίηση των τιμών της εξόδου. Με την στρογγυλοποίηση όλες οι τιμές γίνονται διακριτές, με αποτέλεσμα περισσότερες από αυτές να ταξινομηθούν σωστά. Έτσι, μειώνεται και το μέσο τετραγωνικό σφάλμα.