

Τεχνητή Νοημοσύνη Μηχανική Μάθηση και εφαρμογές - Μέρος Α

Μαρία Κουρούσια
Παναγιώτα Πρέζα

10 Ιουνίου 2021

1 Περιγραφή γνώσης του προβλήματος

Στο παρόν πρόβλημα διατίθεται ένα σύνολο δεδομένων 270 ασθενών με μετρήσεις και άλλα χαρακτηριστικά για τον καθένα τους καθώς και μια ένδειξη για το αν εμφανίζουν ή όχι καρδιακή νόσο χαρακτηριστικό το οποίο αποτελεί την κλάση των ασθενών. Στόχος είναι η δημιουργία ενός συστήματος το οποίο δοθέντος ενός νέου ασθενή με τις αντίστοιχες μετρήσεις να μπορεί να τον κατηγοριοποιήσει σε αυτούς που εμφανίζουν ή όχι καρδιακή νόσο. Στο συγκεκριμένο πρόβλημα χρησιμοποιείται το σύνολο δεδομένων **StatLog** από το **UCI - Machine Learning Repository** [2]. Τα χαρακτηριστικά για κάθε ασθενή που περιλαμβάνονται στο σύνολο δεδομένων περιγράφονται ακολούθως.

1. **age**: Ηλικία του ασθενή. Στο σύνολο εμφανίζονται άτομα ηλικίας από 34 έως 76 ετών με μέση ηλικία του συνόλου περίπου στα 55 έτη.
2. **sex**: Φύλο του ασθενή, με τιμές 0 και 1 για τα δύο φύλα αντίστοιχα.
3. **chest pain type**: Είδος πόνου στο στήθος με τιμές 1, 2, 3 και 4 για τέσσερα διαφορετικά είδη.
4. **resting blood pressure**: Τιμή αρτηριακής πίεσης, με τιμές από 94 έως και 200.
5. **serum cholestoral in mg /dl**: χοληστερόλη ορού σε mg /dl με εύρος τιμών από 126 έως 564.
6. **fasting blood sugar > 120 mg/dl**: Το χαρακτηριστικό αυτό εξετάσει αν το σάκχαρο αίματος νηστείας είναι μεγαλύτερο από 120 mg/dl με τιμή 0 όταν δεν είναι και 1 όταν είναι.

7. **resting electrocardiographic results:** Αποτελέσματα καρδιογραφήματος με τρεις διαφορετικούς τύπους 0, 1, 2.
8. **maximum heart rate achieved :** Μέγιστος ρυθμός καρδιακών παλμών με εύρος από 76 έως και 195.
9. **exercise induced angina:** Εμφάνιση ή όχι στηθάγχης κατά τη άσκηση. Η τιμή 0 υποδεικνύει τη μη εμφάνιση ενώ 1 την εμφάνιση στηθάγχης.
10. **oldpeak:** βαθμός αύξησης καρδιακού ρυθμού τύπου ST που προκαλείται από άσκηση σε σχέση με την ανάπαυση. Το χαρακτηριστικό αυτό παίρνει πραγματικές τιμές από 0 έως 5.6 στο παρόν σύνολο δεδομένων [3].
11. **the slope of the peak exercise ST segment:** Η μέγιστη κλίση του που εμφανίζει η αύξηση του ST [3].
12. **number of major vessels (0-3) colored by flourosopy:** Ο αριθμός κύριων αγγείων με τιμές 0, 1, 2 ή 3 που χρωματίζονται από φθοροσκόπηση.
13. **thal:** Αποτέλεσμα μεσογειακής αναιμίας: 3 = κανονικό, 6 = σταθερό ελάττωμα, 7 = αναστρέψιμο ελάττωμα [1].

2 Δημιουργία δέντρου αποφάσεων και κανόνων

2.1 Προεπεξεργασία δεδομένων

Στο σύνολο που διατίθεται υπάρχουν 270 οντότητες από ασθενείς, 13 χαρακτηριστικά και η κλάση αυτών. Στα δεδομένα δεν υπάρχουν απώλειες τιμών ή μη έγκυρες εκχωρήσεις που χρειάζονται περαιτέρω επεξεργασία. Ωστόσο για την εισαγωγή τους στο **Weka** όπου θα γίνει η ανάλυση χρειάζεται να διαμορφωθούν σε μορφή .csv ή .arff. Στην τελική τους μορφή το **header** του .arff αρχείου διαμορφώνεται ώστε το κάθε χαρακτηριστικό να είναι τέτοιου τύπου (**nominal, numeric, categorical** κλπ) που να αντιπροσωπεύει το είδος τιμών που λαμβάνει.

```
@relation heart
```

```
@attribute age numeric
@attribute sex {0,1}
@attribute chest_pain_type {1,2,3,4}
@attribute blood_pressure numeric
@attribute serum_cholesterol numeric
@attribute blood_sugar_gt_120 {0,1}
@attribute electrocardio_rslt {0,1,2}
@attribute max_heart_rate numeric
@attribute exercise_angina {0,1}
@attribute oldpeak numeric
@attribute slope_ST numeric
@attribute vessels_flourosopy {0,1,2,3}
@attribute thal {3,6,7}
@attribute class {1,2}
```

@data

70,1,4,130,322,0,2,109,0,2.4,2,3,3,2

67,0,3,115,564,0,2,160,0,1.6,2,0,7,1

...

2.2 Δημιουργία δέντρου αποφασής

Για την δημιουργία δέντρου αποφάσεων χρειάστηκε να χωρίσουμε το σύνολο δεδομένων σε σύνολο εκπαίδευσης που είναι το 65% απο το σύνολο δεδομένων, δηλαδή (176 ασθενείς) και σε σύνολο ελέγχου που είναι το 35%, δηλαδή 94 ασθενείς. Για το σκοπό αυτό χρησιμοποιήθηκε το φίλτρο του Weka, **RemovePercentage** από τα **unsupervised filters** και επιλέχθηκε 35%. Αφού έγινε αποθήκευση του αρχείου έγινε αναίρεση των αλλαγών και χρησιμοποιήθηκε το ίδιο φίλτρο με **invertSelection** για να μείνουν ως **test set** οι τιμές που απομακρύνθηκαν πριν. Η διαδικασία αυτή δοκιμάστηκε για διάφορα χωρίσματα και κρατήθηκε το χωρίσμα που έδινε την καλύτερη αναλογία των δύο κατηγοριών της κλάσης στα επιμέρους σύνολα δεδομένων.

Έπειτα απο το εργαλείο Weka επιλέχθηκε ως classifier ο αλγορίθμος J48 με **reduced-error pruning** και **Use training set**.

The screenshot shows the Weka Explorer interface with the Classifier tab selected. The classifier chosen is J48 - R - B - N 3 - Q 3 - M 2. The Test options are set to Use training set. The Classifier output window displays the following results:

Time taken to build model: 0 seconds
=== Evaluation on training set ===
Time taken to test model on training data: 0 seconds
=== Summary ===

Metric	Value	Percentage
Correctly Classified Instances	155	88.0682 %
Incorrectly Classified Instances	21	11.9318 %
Kappa statistic	0.7574	
Mean absolute error	0.1792	
Root mean squared error	0.3224	
Relative absolute error	36.2107 %	
Root relative squared error	64.8246 %	
Total Number of Instances	176	

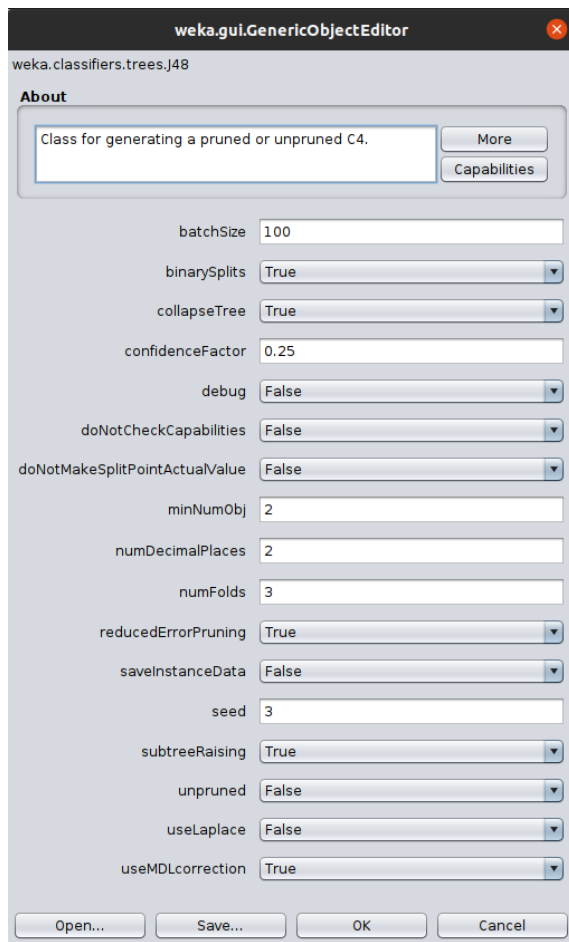
=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
Weighted Avg.	0.918	0.165	0.873	0.918	0.894	0.759	0.907	0.898	1
	0.835	0.082	0.892	0.835	0.863	0.759	0.907	0.863	2

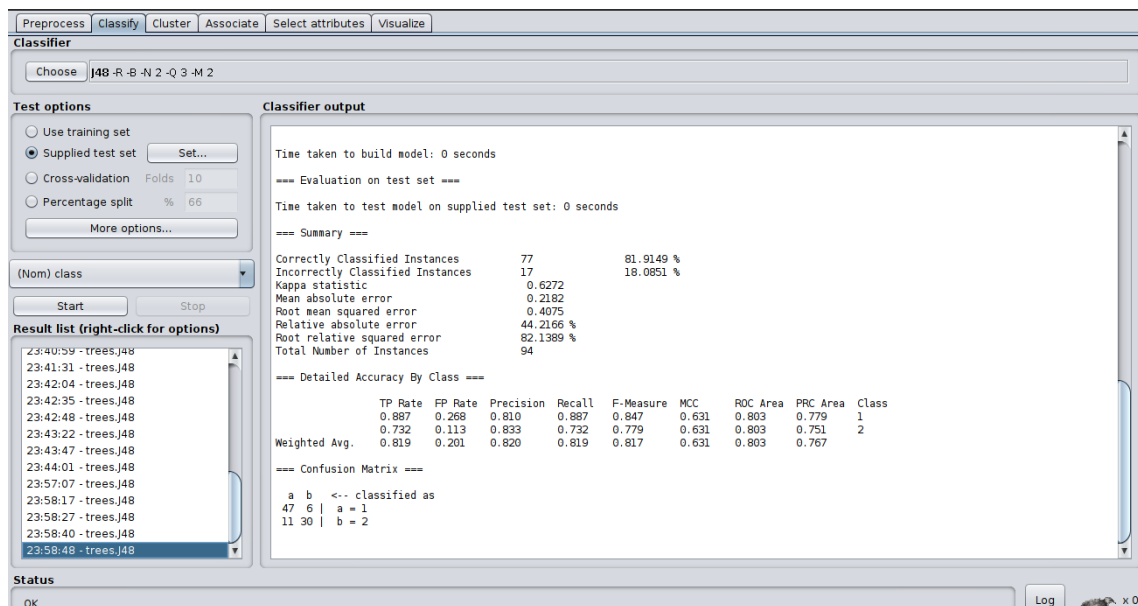
=== Confusion Matrix ===

a	b	<- classified as
89	8	a = 1
13	66	b = 2

Παρακάτω φαίνονται οι επιλογές των παραμέτρων για τις οποίες είχαμε τον καλύτερο συνδυασμό Recall=0.918 και Precision=0.873 καθώς και το καλύτερο ποσοστό Correctly Classified Instances = 88.1%.

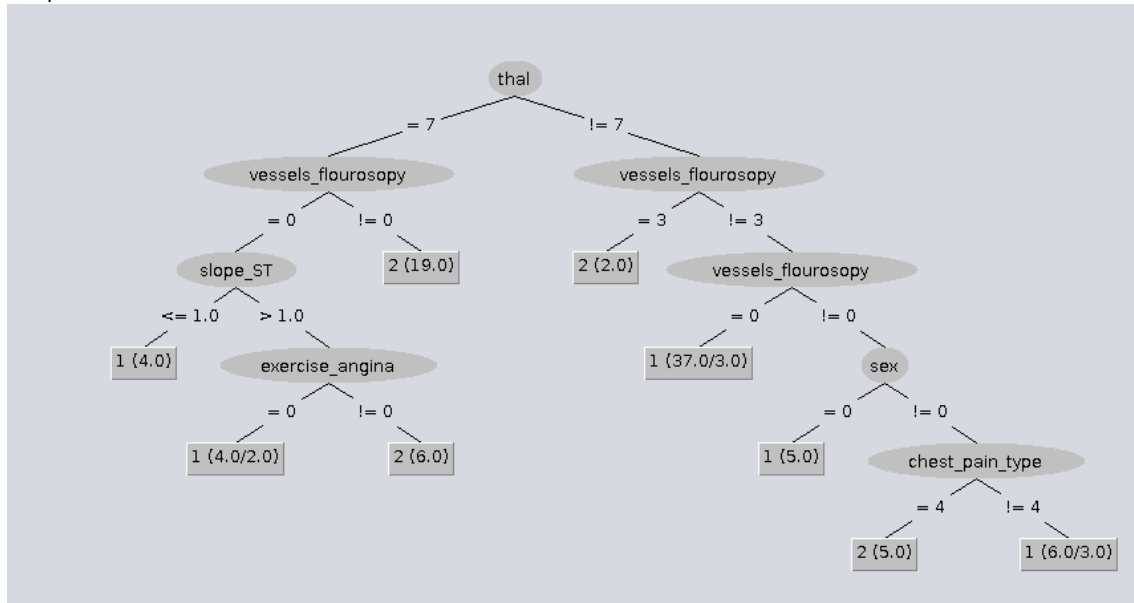


Στη συνέχεια ο ίδιος classifier εξετάστηκε με με την επιλογή **Supplied test set** όπου φορτώθηκε το αρχείο που διαχωρίστηκε ως σύνολο ελέγχου στην αρχή. Σε αυτή την περίπτωση επιτυγχάνουμε λιγότερο καλά αποτελέσματα για τις μετρικές Recall, Precision και το ποσοστό των σωστά ταξινομημένων είναι περίπου 82%.



Απο το τελευταίο run, γίνεται οπτικοποίηση και του δέντρου απόφασης όπως φαίνεται και

στην εικόνα που ακολουθεί.



2.3 Εξαγωγή κανόνων

Από το δέντρο πο παρουσιάστηκε παραπάνω προκύπτουν οι κανόνες κατηγοριοποίησης που είναι όσοι και τα φύλλα του δέντρου δηλαδή 9. Κάθε κανόνας αντιστοιχεί σε έναν συνδυασμό τιμών για τα διαφορετικά attributes που συμμετέχουν.

1. όταν $thal \neq 7$ και $vessels_flourosopy = 3$ τότε 2.
2. όταν $thal \neq 7$ και $vessels_flourosopy \neq 3$ και $vessels_flourosopy \neq 0$ και $sex = 1$ και $chest_pain \neq 4$ τότε 1.
3. όταν $thal \neq 7$ και $vessels_flourosopy \neq 3$ και $vessels_flourosopy \neq 0$ και $sex = 1$ και $chest_pain = 4$ τότε 2.
4. όταν $thal \neq 7$ και $vessels_flourosopy \neq 3$ και $vessels_flourosopy \neq 0$ και $sex = 0$ τότε 1.
5. όταν $thal \neq 7$ και $vessels_flourosopy = 0$ τότε 1.
6. όταν $thal = 7$ και $vessels_flourosopy \neq 0$ τότε 2.
7. όταν $thal = 7$ και $vessels_flourosopy = 0$ και $vessels_flourosopy = 0$ και $slope_ST > 1.0$ και $exercise_angina \neq 0$ τότε 2.
8. όταν $thal = 7$ και $vessels_flourosopy = 0$ και $vessels_flourosopy = 0$ και $slope_ST > 1.0$ και $exercise_angina = 0$ τότε 2.
9. όταν $thal = 7$ και $vessels_flourosopy = 0$ και $vessels_flourosopy = 0$ και $slope_ST \leq 1.0$ τότε 1.

Παρατηρείται ότι στους κανόνες που προκύπτουν από το δέντρο συμμετέχουν έξι μόνο από τα αρχικά attributes που δόθηκαν στο σύνολο δεδομένων. Αυτό είναι αποτέλεσμα δύο αιτιών.

Αρχικά, φαίνεται ότι κάποια από αυτά τα **attributes** δεν σχετίζονται τόσο σημαντικά με την εμφάνιση καρδιακής νόσου όσο αυτά που παρέμειναν. Επιπλέον, το δέντρο έχει προκύψει από τον αλγόριθμο J48 με παράμετρο **reducedErrorPruning** που σημαίνει ότι το δέντρο έχει 'κλαδεύει' από τον αλγόριθμο αφαιρώντας κόμβους με τέτοιο τρόπο ώστε να μην επηρεάζεται η τελική κατάταξη. Συγκεκριμένα, ξεκινώντας από τα φύλλα, κάθε κόμβος αντικαθίσταται με την πιο δημοφιλή κατηγορία του. Εάν η ακρίβεια της πρόβλεψης δεν επηρεάζεται, τότε η αλλαγή διατηρείται. Το **pruning** αν και αφελές έχει το πλεονέκτημα της απλότητας και της ταχύτητας στην εξαγωγή συμπερασμάτων [5].

3 Δημιουργία έμπειρου συστήματος σε CLIPS

3.1 Κανόνες CLIPS

Όπως αναφέρθηκε οι κανόνες που προκύπτουν είναι 9 όσα και τα φύλλα του δέντρου απόφασης. Έτσι στο CLIPS [ζλιπς] δημιουργήθηκε ένα σύστημα από αυτούς τους κανόνες και ορίζονατς το προφιλ του ασθενή **Patient** με τα **attributes** που συμμετέχουν σε αυτό.

```
(deftemplate Patient
  (slot id (type INTEGER))
  (slot sex)
  (slot chest_pain_type)
  (slot serum_cholesterol)
  (slot exercise_angina)
  (slot slope_ST)
  (slot vessels_flourosopy)
  (slot thal)
  (slot class))
```

Επιπλέον ορίζεται ένα **template Diagnosis** όπου θα εισαχθούν τα αποτελέσματα από το έμπειρο σύστημα και οι τιμές του οποίου θα χρησιμοποιηθούν έπειτα για τον υπολογισμό των μετρικών αξιολόγησης.

```
(deftemplate Diagnosis
  (slot id (type INTEGER))
  (slot diagnosis (type INTEGER) (range 1 2))
  (slot realClass (type INTEGER) (range 1 2)))
```

Οι κανόνες για το σύστημα περιλαμβάνονται στο αρχείο **rules.clp**. Όταν ξεκινάει το σύστημά, εμφανίζει ένα μενού με επιλογή για το αν ο χρήστης επιθυμεί να φορτώσει το σύνολο εκπαίδευσης ή το σύνολο ελέγχου. Έπειτα φορτώνει το αντίστοιχο αρχείο και εκτυπώνει στην κονσόλα την κλάση στην οποία κατηγοριοποιήθηκε το κάθε **instance** του συνόλου που δόθηκε καθώς και οι μετρικές που περιγράφονται στη συνέχεια.

```
CLIPS> (run)
Selection 1: Load training set
Selection 2: Load test set
1
Training set loaded
2
1
1
```

3.2 Υπολογισμός μετρικών

Για τον υπολογισμό των μετρικών, επεκτάθηκε το προηγούμενο σύστημα ώστε να γίνεται αυτόματα αμέσως μετά την εξαγωγή του συμπεράσματος από τους κανόνες. Συγκεκριμένα δημιουργήθηκε ένα νέο αρχείο **metrics.clp** στο οποίο έχουμε μεταβλητές για τα στοιχεία του **confusion matrix**, TP, TN, FP, FN και τους κανόνες για τον υπολογισμό τους καθώς και έναν κανόνα, στον οποίο ορίζουμε το ελάχιστο (**salience -1**) ώστε να εκτελεστεί με τη χαμηλότερη προτεραιότητα από το σύστημα, όπου υπολογίζονται οι μετρικές **Accuracy**, **Sensitivity**, **Specificity**, **Precision**, **Recall**, **F-Measure**. Για τους κανόνες όπως ακριβώς προκύπτουν από το δέντρο απόφασης, πάνω στο **training set** λαμβάνονται οι ακόλουθες τιμές των μετρικών.

```
:::*****
Accuracy: 0.869318181818182

Sensitivity: 0.979381443298969

Specificity: 0.734177215189873

Precision: 0.818965517241379

Recall: 0.979381443298969

F-Measure: 0.892018779342723

CLIPS> _
```

3.3 Βελτιστοποίηση και αξιολόγηση στο σύνολο ελέγχου

Στο παρόν σύστημά είναι κρίσιμο να βελτιωθεί η τιμή της μετρικής **Precision** που λαμβάνει υπόψη των αριθμό των TP, FP. Η μετρική αυτή δηλαδή, δείχνει κατά πόσο χάνονται περιπτώσεις ασθενών που ενώ εμφανίζουν καρδιακή νόσο (κλάση 2) το σύστημα τους κατηγοριοποιεί στην κλάση 1. Σαν σενάριο της πραγματικότητας κάτι τέτοιο θα μπορούσε να αποβεί επικίνδυνο καθώς θα είχε παραλειφθεί ενδεχομένως η χορήγηση κάποιας θεραπείας.

Το σημείο βελτίωσης των κανόνων εντοπίζεται στο χαρακτηριστικό **chest_pain** το οποίο καταλήγει σε εύρεση καρδιακής νόσου μόνο για την τιμή 4 που αντιστοιχεί σε επίπεδο 4 του πόνου στο στήθος. Η συνθήκη μεταβλήθηκε σε δύο από τους κανόνες ώστε να καταλήγουν σε κλάση 2, τιμές από 3 και πάνω του **chest_pain**. Σε αυτή την αλλαγή κατευθύνει το ίδιο το δέντρο από το **Weka** αφού στο αντίστοιχο φύλλο φαίνεται ότι 3 στα 6 από τα **instances** που συγκρούονται με αυτούς τους κανόνες δεν έχουν κατηγοριοποιηθεί σωστά.

Για τους βελτιωμένους κανόνες, πάνω στο **training set** λαμβάνονται οι ακόλουθες τιμές των μετρικών.

```
:::*****
Accuracy: 0.875

Sensitivity: 0.969072164948454

Specificity: 0.759493670886076

Precision: 0.831858407079646

Recall: 0.969072164948454

F-Measure: 0.895238095238095

CLIPS> _
```

Στους βελτιωμένους κανόνες για το test set οι μετρικές διαμορφώνονται όπως φαίνονται στο ακόλουθο screenshot.

```
::::*****
Accuracy: 0.797872340425532

Sensitivity: 0.849056603773585

Specificity: 0.731707317073171

Precision: 0.803571428571429

Recall: 0.849056603773585

F-Measure: 0.825688073394495

CLIPS> _
```

Για το τελευταίο παρατηρούνται λιγότερο καλές τιμές για τις μετρικές ωστόσο αρκετά κοντά στα αποτελέσματα του συνόλου εκπαίδευσης δεδομένου του μικρού πλήθους των δεδομένων και της φύσης τους. Στον πίνακα που ακολουθεί συνοψίζονται όλα τα αποτελέσματα των μετρικών:

Metric	Initial Rules - Training Set	Optimized Rules - Training Set	Optimized Rules - Test Set
Accuracy	0.869	0.875	0.797
Specificity	0.734	0.759	0.731
Precision	0.818	0.831	0.803
Recall	0.979	0.969	0.849
F-Measure	0.892	0.895	0.825

4 Δημιουργία ασαφούς έμπειρου συστήματος σε FUZZYCLIPS

4.1 Κανόνες FuzzyCLIPS

Για το ασαφές έμπειρο σύστημα FuzzyCLIPS [4] ορίσαμε δύο ασαφείς μεταβλητές από τους κανόνες, οι οποίες είναι το chest_pain_type και το vessels_flourosopy. Για το chest_pain_type οι τιμές είναι: low για τις τιμές 1, 2 και 3, και high για την τιμή 4. Για το vessels_flourosopy οι τιμές είναι: few για την τιμή 0, και many για τις τιμές 1, 2 και 3. Ορίστηκε το προφίλ του ασθενούς του ασαφούς συστήματος Patient_fuzzy με τα attributes που συμμετέχουν σε αυτό ως εξής:

```
(deftemplate Patient_fuzzy
  (slot id (type INTEGER))
  (slot sex (type NUMBER))
  (slot chest_pain_type-fuzzy (type FUZZY-VALUE chest_pain_type_fuzzy))
  (slot serum_cholesterol (type NUMBER))
  (slot exercise_angina (type NUMBER))
  (slot slope_ST (type NUMBER))
  (slot vessels_flourosopy-fuzzy (type FUZZY-VALUE vessels_flourosopy_fuzzy))
  (slot thal (type NUMBER))
  (slot class (type INTEGER) (range 1 2))
)
```


Επιπλέον ορίζεται ένα **template Diagnosis** όπου θα εισαχθούν τα αποτελέσματα από το ασαφές έμπειρο σύστημα και οι τιμές του θα χρησιμοποιηθούν έπειτα για τον υπολογισμό των μετρικών αξιολόγησης.

```
(deftemplate Diagnosis
  (slot id (type INTEGER))
  (slot diagnosis (type INTEGER) (range 1 2))
  (slot realClass (type INTEGER) (range 1 2)))
```

Οι κανόνες για το ασαφές έμπειρο σύστημα περιλαμβάνονται στο αρχείο **fuzzyclips.clp**.

4.2 Υπολογισμός μετρικών ασαφούς συστήματος

Για τους κανόνες του ασαφούς έμπειρου συστήματος πάνω στο **training set** λαμβάνονται οι ακόλουθες τιμές των μετρικών.

```
Accuracy: 0.66477272727273
Sensitivity: 0.4742268041237113
Specificity: 0.8987341772151899
Precision: 0.8518518518518519
Recall: 0.4742268041237113
F-Measure: 0.6092715231788079
```

Αντίστοιχα στο **test set** λαμβάνονται οι ακόλουθες τιμές των μετρικών:

```
Accuracy: 0.6276595744680851
Sensitivity: 0.3584905660377358
Specificity: 0.975609756097561
Precision: 0.95
Recall: 0.3584905660377358
F-Measure: 0.5205479452054794
```

4.3 Βελτιστοποίηση και αξιολόγηση στο σύνολο ελέγχου

Για την βελτίωση του συστήματος αλλάζουμε τις τιμές των ασαφών μεταβλητών **chest_pain** που είναι **low** για τις τιμές 1 και 2 και **high** για τις τιμές 3 και 4.

Σε διαφορετικούς συνδυασμούς του εύρους των ασαφών μεταβλητών για το **training set** οι μετρικές διαμορφώνονται όπως φαίνονται στο ακόλουθο **screenshot** όπου συμπερνούμε ότι είναι ίδιες με τις μετρικές που παρουσιάστηκαν παραπάνω με αποτέλεσμα να μην επιδέχεται το ασαφές έμπειρο σύστημα περαιτέρω βελτιστοποίηση με **tuning**.

Στον πίνακα που ακολουθεί συνοψίζονται όλα τα αποτελέσματα των μετρικών:

Metric	Initial Rules - Training Set	Test Set
Accuracy	0.664	0.627
Specificity	0.898	0.975
Precision	0.851	0.950
Recall	0.474	0.358
F-Measure	0.609	0.520

5 Σύγκριση Συστημάτων

Το FuzzyClips [6] παρέχει μια ασαφή ικανότητα δημιουργίας κανόνων που είναι πλήρως ενσωματωμένη με τα γεγονότα του CLIPS και επιτρέπει σε κάποιον να εκπροσωπεί και να χειρίζεται ασαφή γεγονότα και κανόνες. Στην περίπτωση του παρόντως συστήματος η ασάφεια των μεταβλητών δεν είναι το ίδιο ευνοϊκή όπως θα γινόταν σε άλλα σύστημα δημιουργώντας εύκολο χειρισμό χωρίς να θυσιάζεται η ακρίβεια του συστήματος. Το σύστημα με τους ασαφείς κανόνες προσδίδει απλότητα και είναι εύκολο στη διαχείριση, ωστόσο εδώ τα δύο **attributes** που επιλέχθηκε να είναι οι ασαφείς μεταβλητές, αν και παίρνουν πραγματικές τιμές δεν έχουν μεγάλο εύρος τιμών ώστε να έχει περισσότερο νόημα η ασάφεια. Η χρήση της εδώ έχει ως αποτέλεσμα να εξαλειφθούν κάποιες συνθήκες από τους κανόνες και να χαθούν περιπτώσεις κατηγοριοποίησης από αυτές. Συνεπώς τα ίδια **facts** εμπίπτουν σε κάποια κλάση έχοντας εξετάσει λιγότερα κριτήρια, το οποίο αυξάνει την πιθανότητα για λάθος κατηγοριοποίηση.

Metric	Clips Rules	Optimized Clips Rules	FuzzyClips Rules
Accuracy	0.869	0.875	0.664
Specificity	0.734	0.759	0.898
Precision	0.818	0.831	0.851
Recall	0.979	0.969	0.474
F-Measure	0.892	0.895	0.609

Όσον αφορά τον αριθμό των κανόνων παρέμεινε ο ίδιος, ωστόσο όπως αναφέρθηκε το σύστημα απλοποιήθηκε αφού εξαλείφθηκαν πολλές από τις συνθήκες μέσα στους ελέγχους. Από την παραπάνω σύγκριση προκύπτει ότι το **optimize CLIPS** σύστημα αποδίδει τις καλύτερες τιμές για τις μετρικές αξιολόγησης. Ωστόσο αξίζει να σημειωθεί ότι η μετρική **Precision** παρουσιάζει εξίσου καλή τιμή και στα ασαφή συστήματα. Η τιμή της **Recall** μειώθηκε σημαντικά που σημαίνει ότι με την ασάφεια αυξήθηκαν τα FN που είναι λογικό αφού οι κανόνες που επηρεάστηκαν είναι αυτοί στους οποίους εμπίπτουν τα περισσότερα **instances** για κατηγοριοποίηση στην κλάση 2.

REFERENCES

- [1] A. Aessopos, M. Kati, and D. Farmakis. “Heart disease in thalassemia intermedia: a review of the underlying pathophysiology”. In: *Haematologica* 92.5 (May 2007), pp. 658–665.
- [2] Dheeru Dua and Casey Graff. *UCI Machine Learning Repository*. 2017. URL: <http://archive.ics.uci.edu/ml>.
- [3] R. S. Finkelhor et al. “The ST segment/heart rate slope as a predictor of coronary artery disease: comparison with quantitative thallium imaging and conventional ST segment criteria”. In: *Am Heart J* 112.2 (Aug. 1986), pp. 296–304.
- [4] *FuzzyClips*. URL: <https://quentin.pradet.me/blog/fuzzyclips-downloads.html>.
- [5] W. Nor Haizan W. Mohamed, Mohd Najib Mohd Salleh, and Abdul Halim Omar. “A comparative study of Reduced Error Pruning method in decision tree algorithms”. In: (2012), pp. 392–397. DOI: 10.1109/ICCSCE.2012.6487177.
- [6] R. A. Orchard. *FuzzyCLIPS Version 6.04A User’s Guide*. URL: <https://profs.info.uaic.ro/~dcristea/cursuri/SE/fzdocs.pdf>.