

# Ιόνιο Πανεπιστήμιο



## Τμήμα Πληροφορικής

Μάθημα: Συστήματα Υποστήριξης Αποφάσεων

Επιτηρητής Καθηγητής: Έξαρχος Θέμης

Μέλη Ομάδας: Γιαμπούρης Δημήτριος, Π2015178, [p14giab@ionio.gr](mailto:p14giab@ionio.gr)

Τσιβιντζέλη Χρύσα, Π2015185, [p15tsiv@ionio.gr](mailto:p15tsiv@ionio.gr)

Εξάμηνο Μαθήματος: ΣΤ'

## Περιεχόμενα:

1. Εισαγωγή
2. Συστήματα Υποστήριξης Αποφάσεων
3. Weka & Εξαγωγή Μοντέλου
4. Υλοποίηση Κώδικα
5. Τελικό Αποτέλεσμα & Στιγμιότυπα
6. Πηγές & Βιβλιοθήκες
7. Κώδικας

## 1. Εισαγωγή

Το παρόν έγγραφο αποτελεί την τελική αναφορά της εξαμηνιαίας εργασίας (τύπου Β') με τίτλο "Συστήματα Υποστήριξης Αποφάσεων: Καρκίνος του Μαστού" των φοιτητών: Γιαμπούρης Δημήτριος (Π2014206) και Τσιβιντζέλη Χρύσα (Π2015185), στα πλαίσια του μαθήματος "Συστήματα Υποστήριξης Αποφάσεων" του ΣΤ' Εαρινού Εξαμήνου 2018 με διδάσκων τον κ. Έξαρχο.

Σκοπός της εργασίας είναι η εξοικείωση με το εργαλείο Weka και η υλοποίηση ενός Συστήματος Υποστήριξης Αποφάσεων και ενός user interface όπου θα αξιοποιείται ένα Weka Model.

Η εργασία υλοποιείται έχοντας ως βάση το εργαλείο του Weka για την εξαγωγή του μοντέλου, τις γλώσσες HTML, JavaScript και CSS και τη πλατφόρμα του Github.

## 2. Συστήματα Υποστήριξης Αποφάσεων

Η λήψη αποφάσεων (decision making) είναι αποτέλεσμα σύνθετων διαδικασιών, που έχουν σαν στόχο, αρχικά μεν να μελετήσουν και να αναλύσουν διεξοδικά τις επιπτώσεις όλων των εναλλακτικών επιλογών (αποφάσεων), στη συνέχεια δε να προσχωρήσουν σε μια προσπάθεια σύνθεσης και σύγκλισης των απαιτήσεων όλων των εμπλεκόμενων, στη διαδικασία απόφασης, μερών, ώστε να καταλήξουν τελικά στην εύρεση της πλέον κοινά αποδεκτής λύσης.

Τα Σ.Υ.Α. στο τομέα της Υγείας, αποτελούν ένα σημαντικό εργαλείο στα χέρια των γιατρών και των νοσοκόμων. Βελτιώνουν την ποιότητα της ιατρικής φροντίδας, αυξάνουν την ασφάλεια των ασθενών και καθιστούν την ιατρική τους εξυπηρέτηση αποτελεσματικότερη. Τα Σ.Υ.Α. βοηθούν σημαντικά όσους εργάζονται στο χώρο της Υγείας σε θέσεις όπου απαιτείται διαχείριση δεδομένων και γνώσης.

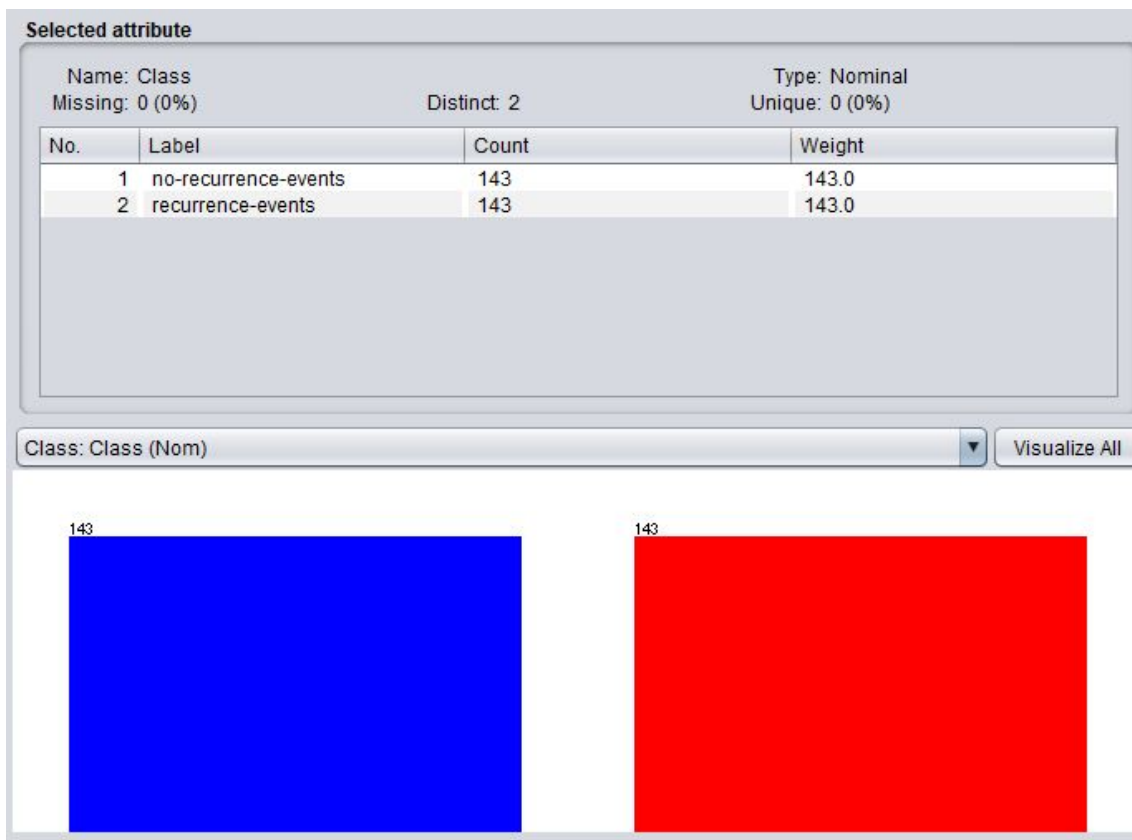
Το σύστημα που υλοποιήθηκε στην παρούσα εργασία αποτελείται από ένα ταξινομητή, ο οποίος βασίζεται σε κανόνες, δηλαδή χρησιμοποιεί ένα σύνολο κανόνων (statements) IF-THEN για ταξινόμηση. Οι κανόνες αυτοί προέκυψαν από ένα δέντρο ταξινόμησης (decision tree) που σχεδιάστηκε από το μοντέλο του WEKA, ανεβάζοντας το σύνολο δεδομένων μας (breast-cancer.arff).

## 3. Weka & Εξαγωγή Μοντέλου

Ανεβάζοντας το σύνολο δεδομένων στο WEKA, παρατηρήθηκε πως υπήρχε μεγάλη διαφορά στον αριθμό των δειγμάτων ανάμεσα στις δύο κλάσεις ("recurrence", "no-recurrence"). Για το λόγο αυτό, χρησιμοποιήθηκαν οι τεχνικές oversampling και undersampling, μέσα από το WEKA, για την προσαρμογή της ταξικής κατανομής του συνόλου. Έτσι, ο αριθμός και η κατανομή των δειγμάτων, από 201 για "no-recurrence", και 85 για "recurrence", μετατράπηκε σε 143 και για τις δυο κλάσεις (Εικόνα 1).

Έπειτα, μετά την προσαρμογή του συνόλου, χρησιμοποιήθηκε ο αλγόριθμος J48, για ταξινόμηση με την τεχνική cross-fold validation, με 10 folds, όπου το σύνολο δεδομένων χωρίστηκε σε 10 ίσα μέρη (folds) και 10 φορές έγινε εκπαίδευση στο μοντέλο χρησιμοποιώντας κάθε φορά 9 μέρη για εκπαίδευση (training) και 1 (σε κάθε μια από τις 10 φορές, χρησιμοποιούνταν διαφορετικό μέρος) για αξιολόγηση (testing). Ο πίνακας σύγχυσης που προέκυψε είναι ο εξής (Εικόνα 2).

Εικόνα 1



Εικόνα 2

**Classifier output**

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	240	83.9161 %
Incorrectly Classified Instances	46	16.0839 %
Kappa statistic	0.6783	
Mean absolute error	0.1949	
Root mean squared error	0.3604	
Relative absolute error	38.9716 %	
Root relative squared error	72.0714 %	
Total Number of Instances	286	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0,874	0,196	0,817	0,874	0,845	0,680	0,891	0,895	no-recurrence-events
	0,804	0,126	0,865	0,804	0,833	0,680	0,891	0,862	recurrence-events
Weighted Avg.	0,839	0,161	0,841	0,839	0,839	0,680	0,891	0,879	

=== Confusion Matrix ===

a	b	<-- classified as
125	18	a = no-recurrence-events
28	115	b = recurrence-events

#### 4. Υλοποίηση Κώδικα

(Ο κώδικας της εργασίας υπάρχει στην σελίδα: <https://github.com/p15tsiv/DSS-Breast-Cancer>)

Στην παρούσα εργασία, η εφαρμογή που αναπτύσσεται αξιοποιεί τα εξής:

- Τη γλώσσα HTML, με την οποία υλοποιήθηκε η σελίδα, που αποτελείται από μια φόρμα, σε μορφή ερωτηματολόγιου, όπου ο χρήστης βάζει τα δεδομένα. Τα δεδομένα που ζητούνται είναι τα εξής (μαζί με το όνομα της μεταβλητής και τις πιθανές τιμές της):
  1. ηλικία (age: 10-99)
  2. κατάσταση εμμηνόπαυσης ασθενούς (menopause: lt40, ge40, premeno)
  3. μέγεθος του όγκου (tumor\_size: 0-59)
  4. ο αριθμός των μασχαλαίων λεμφαδένων που περιέχουν μεταστατικό καρκίνο του μαστού ορατό σε ιστολογική εξέταση (inv\_nodes: 0-39)
  5. εάν υπάρχει μετάσταση εντός των λεμφαδένων (node\_caps: yes, no)
  6. βαθμός κακοήθειας (deg\_malign: 1, 2, 3)
  7. σε ποιο στήθος υπήρξε ο όγκος (breast: right, left)
  8. σε ποιο μέρος/”τεταρτημόριο” του στήθους (breast\_quad: left upper, left lower, right upper, right lower, central)
  9. εάν υπήρξε έκθεση σε ακτινοβολία (irradiat: yes, no)
- Τη γλώσσα JavaScript, με την οποία υλοποιήθηκαν κάποιες συναρτήσεις, έτσι ώστε τα δεδομένα από τις φόρμες να επιστρέφουν ως μεταβλητές, να περνάνε μέσα από μια άλλη συνάρτηση που περιέχει τα if then statements, και να επιστρέφει την ταξινόμηση (classification). Οι συναρτήσεις είναι οι εξής:
  1. getAge()
  2. getMenopause()
  3. getTumorSize()
  4. getInvNodes()
  5. getNodeCaps()
  6. getDegMalign()
  7. getBreast()
  8. getBreastQuad()
  9. getIrradiat()Αυτές οι συναρτήσεις, απλά αναθέτουν την επιλεγμένη τιμή του χρήστη, σε κάθε μεταβλητή, αντίστοιχα με το όνομα τους.
  - display(): περιέχει όλα τα if then statements και επιστρέφει το αποτέλεσμα (result)
- Τη γλώσσα CSS, με την οποία βελτιώθηκε η εμφάνιση της σελίδας, προστέθηκαν περισσότερες πληροφορίες σχετικά με τη λειτουργία της και κάποιοι σύνδεσμοι για εξωτερικές ιστοσελίδες, σχετικές με την εργασία

#### 5. Τελικό Αποτέλεσμα & Στιγμιότυπα



## Welcome to this Decision Support System regarding Breast Cancer.

This DSS is a health information technology system that is designed to provide physicians and other health professionals with clinical decision support (CDS), that is, assistance with clinical decision-making tasks.

Based on a set of rules produced by a data set provided by the Institute of Oncology (University Medical Center, Ljubljana, Yugoslavia), the system will process the data input and make a prediction on whether the patient may experience recurrence of the cancer.

It is advised, in order for the system to make the best "prediction" possible, to choose carefully your answer for every question and avoid blank answers.

### 1) AGE

Select your age: ▾

### 2) MENOPAUSE

☐ lt40 ☐ ge40 ☐ premeno

### 3) TUMOR SIZE

Select size: ▾

### 4) INV-NODES

over time, and with more aggressive disease, the tumor may replace the lymph node and then penetrate the capsule, allowing it to invade the surrounding tissues

☐ Yes ☒ No

### 6) DEGREE OF MALIGNANCY

The histological grade of the tumor

☐ 1 ☐ 2 ☒ 3

### 7) BREAST

☒ Right ☐ Left

### 8) BREAST QUADRANT

Right Lower Quadrant ▾

### 9) IRRADIATION

☐ Yes ☒ No

[Click here for results](#)

[GITHUB](#)

[CHRYSA](#)

[JIM](#)

[REPORT](#)

[LEARN MORE](#)

## Αποτελέσματα

over time, and with more aggressive disease, the tumor may re...  
node and then penetrate the capsule, allowing it to invade the st...

☐ Yes ☒ No

6) DEGREE OF MALIGNANCY

☐ 1 ☐ 2 ☒ 3

7) BREAST

☐ Right ☒ Left

8) BREAST QUADRANT

Left Lower Quadrant ▾

9) IRRADIATION

☐ Yes ☒ No

[Click here for results](#)

[GITHUB](#) [CHRYSA](#) [JIM](#) [REPORT](#) [LEARN MORE](#)

p15tsiv.github.io says  
RESULT: The patient may experience no recurrence.  
[OK](#)

## 6. Πηγές & Βιβλιοθήκες

Χρήσιμες σελίδες και βιβλιοθήκες αποτέλεσαν:

- <https://www.cs.waikato.ac.nz/ml/weka/>
- [https://en.wikipedia.org/wiki/Clinical\\_decision\\_support\\_system](https://en.wikipedia.org/wiki/Clinical_decision_support_system)
- <https://codepen.io/ericrasch/pen/zjDBx>
- <https://archive.ics.uci.edu/ml/datasets/Breast+Cancer>
- <https://www.medicalnewstoday.com/articles/37136>

## 7. Κώδικας

<https://github.com/p15tsiv/DSS-Breast-Cancer>