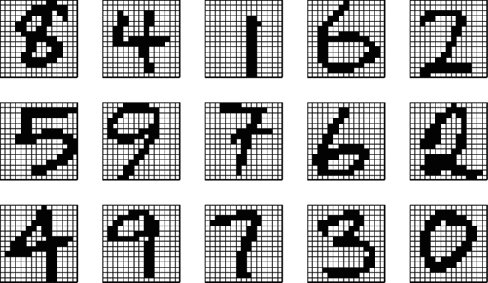
|  |
| --- |
| *ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ*  *ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ* |
| **3η Εργαστηριακή Άσκησης** |
| **Aναγνώριση Συναισθήματος στη Μουσική** |

**Μάθημα : Αναγνώριση Προτύπων**

**Ροή Σ**

Συνεργάτες :

* Βαβουλιώτης Γεώργιος ( **Α.Μ. : 03112083** )
* Σταυρακάκης Δημήτριος ( **Α.Μ. : 03112017** )

**Σκοπός**: Σκοπός της τρίτης εργαστηριακής άσκησης είναι η αυτόματη ταξινόμηση μουσικών αποσπασμάτων με βάση το συναίσθημα που δημιουργεί στον ακροατή. Πραγματοποιείται η εξαγωγή χαρακτηριστικών από από τα διαθέσιμα δεδομένα, δηλαδή τα μουσικά σήματα. Τα συναισθήματα των μουσικών σημάτων θα περιγραφούν με μια δισδιάστατη αναπαράσταση, η οποία έχει στον έναν άξονα το valence (χαρά/λύπη) και στον άλλο το activation (ενεργοποίηση/απενεργοποίηση). Το τελικό πακέτο χαρακτηριστικών θα περιλαμβάνει ένα συνδυασμό χαρακτηριστικών όπως χαρακτηριστικά εμπνευσμένα από τη μουσική και στατιστικά μεγέθη των Mel-Frequency Cepstral Coefficients (MFCCs). Τα βήματα 1-9 αγνοούνται διότι εξηγήθηκαν στην προπαρασκευή.

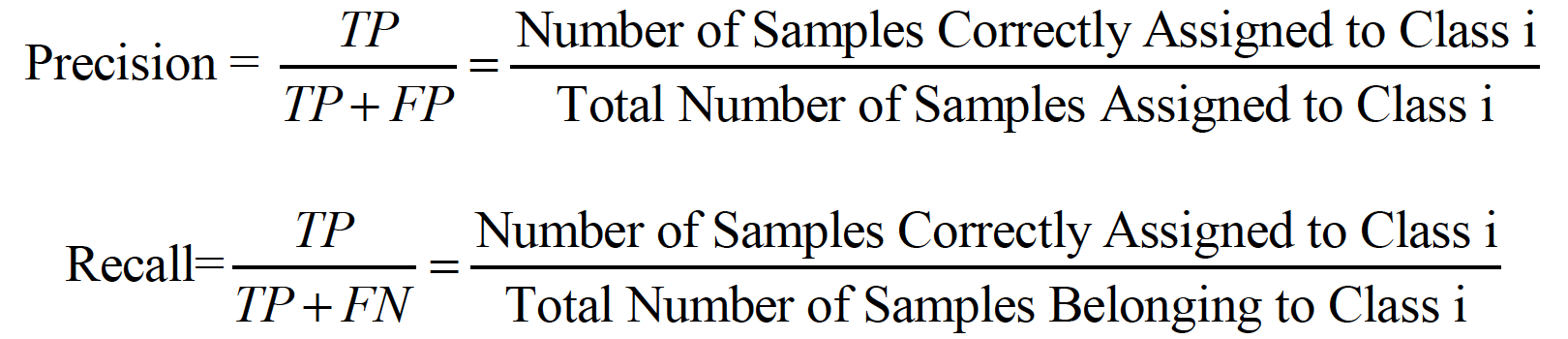
**Εκτέλεση Άσκησης**

**Βήμα 10** : Στο βήμα αυτό για το πρόβλημα της ταξινόμησης θα θεωρήσουμε τα εξής : Για την ταξινόμηση του valence τα δείγματα με valence > 3 αντιστοιχίζονται στην κλάση -1, τα δείγματα με valence <3 στην κλάση 1 και τα δείγματα με valence = 3 αγνοούνται. Για την ταξινόμηση του activation τα δείγματα με activation > 3 αντιστοιχίζονται στην κλάση 1, τα δείγματα με activation <3 στην κλάση -1 και τα δείγματα με activation = 3 αγνοούνται. Με αυτό τον τρόπο έχουμε πλέον 359 και 355 δείγματα για τον valence και το activation αντίστοιχα.

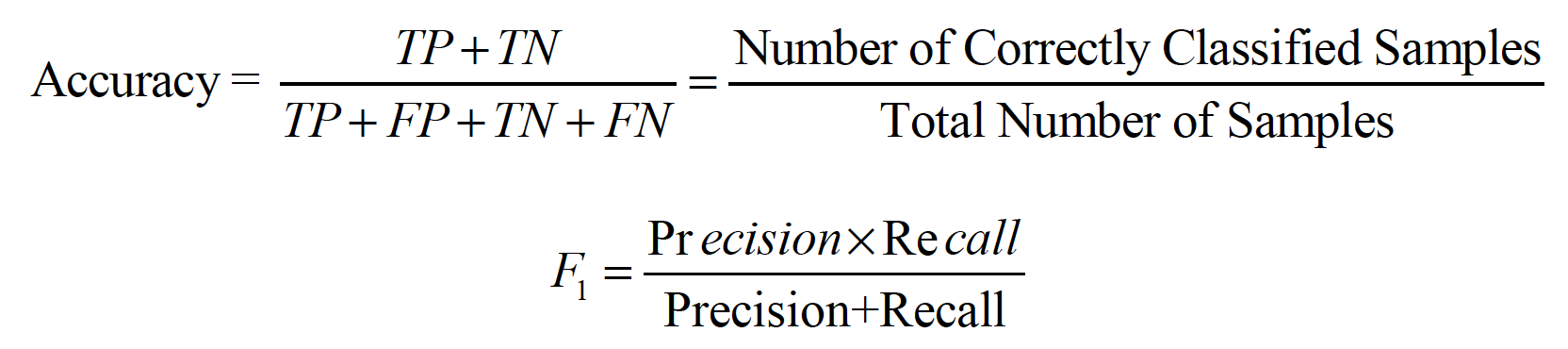
**Βήμα 11** : Στο βήμα αυτό γίνεται πειραματισμός για κάθε διάσταση ξεχωριστά στα παρακάτω σύνολα χαρακτηριστικών:  
1. Χαρακτηριστικά από το βήμα 6. Το πλήθος αυτού του συνόλου είναι 10.  
2. Χαρακτηριστικά από το βήμα 7. Το πλήθος αυτού του συνόλου είναι 156.   
3. Χαρακτηριστικά από το βήμα 6 και βήμα 7. Το πλήθος αυτού του συνόλου είναι 166.

Για την εκπαίδευση των ταξινομητών τα δεδομένα χωρίζονται με τυχαίο τρόπο σε train (80%) και test (20%), υπολογίζονται τα ποσοστά επιτυχίας για το συγκεκριμένο διαχωρισμό και η διαδικασία αυτή επαναλαμβάνεται 3 φορές. Τα τελικά αποτελέσματα προκύπτουν ως ο μέσος όρος των επιμέρους αποτελεσμάτων.

**Βήμα 12** : Στο βήμα αυτό θα γίνει ταξινόμηση του 20% των δεδομένων, χρησιμοποιώντας στο 80% των δεδομένων με χρήση του αλγορίθμου NNR-1 και στη συνέχεια του αλγορίθμου NRR-k για k = 3,5,7. Για να αξιολογήσουμε την επίδοση κάθε fold χρησιμοποιούνται οι παρακάτω μετρικές, όπως προτείνεται και στην εκφώνηση της άσκησης:



Τα αποτελέσματα τα οποία πήραμε φαίνονται παρακάτω :



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **KNN Algorithm - Features’ Set 1** | | | | |
| **Activation** | | | | |
| **Number of Neighbours** | Accuracy | Precision | Recall | F1-score |
| **1** | 0.666666666666667 | 0.631559290382820 | 0.613333333333333 | 0.618139992053036 |
| **3** | 0.652582159624413 | 0.615395021645022 | 0.607619047619048 | 0.605770963356103 |
| **5** | 0.610328638497653 | 0.569932297889287 | 0.533333333333333 | 0.545486234195912 |
| **7** | 0.610328638497653 | 0.569047619047619 | 0.516190476190476 | 0.538613587670191 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **KNN Algorithm - Features’ Set 2** | | | | |
| **Activation** | | | | |
| **Number of Neighbours** | Accuracy | Precision | Recall | F1-score |
| **1** | 0.544600938967136 | 0.488719124013242 | 0.521904761904762 | 0.502596320887175 |
| **3** | 0.568075117370892 | 0.515179265179265 | 0.470476190476191 | 0.488573715907644 |
| **5** | 0.577464788732394 | 0.524389233954451 | 0.464761904761905 | 0.488407224958949 |
| **7** | 0.582159624413146 | 0.530138204707170 | 0.480000000000000 | 0.495682960857090 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **KNN Algorithm - Features’ Set 3** | | | | |
| **Activation** | | | | |
| **Number of Neighbours** | Accuracy | Precision | Recall | F1-score |
| **1** | 0.652582159624413 | 0.607215836526181 | 0.605714285714286 | 0.604186562872009 |
| **3** | 0.666666666666667 | 0.642889784946237 | 0.598095238095238 | 0.610951621477937 |
| **5** | 0.610328638497653 | 0.569932297889287 | 0.533333333333333 | 0.545486234195912 |
| **7** | 0.610328638497653 | 0.569047619047619 | 0.516190476190476 | 0.538613587670191 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **KNN Algorithm - Features’ Set 1** | | | | |
| **Valence** | | | | |
| **Number of Neighbours** | Accuracy | Precision | Recall | F1-score |
| **1** | 0.610328638497653 | 0.710357195139804 | 0.684914361001317 | 0.697184759684760 |
| **3** | 0.563380281690141 | 0.659100073531156 | 0.682516469038208 | 0.670588235294118 |
| **5** | 0.577464788732394 | 0.667136752136752 | 0.715138339920949 | 0.689245911130410 |
| **7** | 0.615023474178404 | 0.682121872317951 | 0.779934123847167 | 0.726544082979727 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **KNN Algorithm - Features’ Set 2** | | | | |
| **Valence** | | | | |
| **Number of Neighbours** | Accuracy | Precision | Recall | F1-score |
| **1** | 0.671361502347418 | 0.782818113064852 | 0.693399209486166 | 0.734816369647830 |
| **3** | 0.633802816901409 | 0.733333333333333 | 0.701805006587615 | 0.715249691045950 |
| **5** | 0.629107981220657 | 0.726167471819646 | 0.701554677206851 | 0.712873362505933 |
| **7** | 0.671361502347418 | 0.746004728132388 | 0.760513833992095 | 0.751878092667566 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **KNN Algorithm - Features’ Set 3** | | | | |
| **Valence** | | | | |
| **Number of Neighbours** | Accuracy | Precision | Recall | F1-score |
| **1** | 0.586854460093897 | 0.690382081686430 | 0.670671936758893 | 0.680116381203338 |
| **3** | 0.558685446009390 | 0.654217762913415 | 0.682187088274045 | 0.667908960274417 |
| **5** | 0.586854460093897 | 0.672975857009471 | 0.721805006587615 | 0.696105383734250 |
| **7** | 0.610328638497653 | 0.679738562091503 | 0.773267457180501 | 0.722385667138143 |

**Βήμα 13** : Στο βήμα αυτό για την ταξινόμηση των δεδομένων γίνεται χρήση του αλγορίθμου Naive Bayes, όπως αυτός υλοποιήθηκε στο πρώτο εργαστήριο. Τα αποτελέσματα τα οποία πήραμε φαίνονται παρακάτω :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **BAYES Algorithm** | | | | |
| **Activation** | | | | |
| **Features’ set** | Accuracy | Precision | Recall | F1-score |
| **1** | 0.525821596244132 | 0.485802835937409 | 0.939047619047619 | 0.636165020569945 |
| **2** | 0.605633802816902 | 0.544588744588745 | 0.628571428571429 | 0.582179845470985 |
| **3** | 0.530516431924883 | 0.487512813138352 | 0.939047619047619 | 0.637930569598921 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **BAYES Algorithm** | | | | |
| **Valence** | | | | |
| **Features’ set** | Accuracy | Precision | Recall | F1-score |
| **1** | 0.582159624413146 | 0.662014730339748 | 0.746192358366272 | 0.697679845783346 |
| **2** | 0.652582159624413 | 0.766018534311217 | 0.681225296442688 | 0.719708207877408 |
| **3** | 0.619718309859155 | 0.683550428960034 | 0.790158102766798 | 0.730158730158730 |

**Βήμα 14** : Στο βήμα αυτό κάνουμε μείωση της διάστασης των χαρακτηριστικών με χρήση της τεχνικής της ανάλυσης κύριων συνιστωσών(Principal Component Analysis-PCA). Ακολουθούν τα αποτελέσματα για διαφορετικούς αριθμούς κύριων συνιστωσών και για τους 2 διαφορετικούς αλγορίθμους ταξινόμησης:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Activation** | | | | | | |
| **PCA Components:** | **Algorithm→**  **Metrics↓** | **NNR-1** | **NNR-3** | **NNR-5** | **NNR-7** | **BAYES** |
| **10** | **Accuracy** | 0.58216 | 0.63380 | 0.56808 | 0.57277 | 0.5774 |
| **Precision** | 0.53522 | 0.60577 | 0.52407 | 0.53625 | 0.51825 |
| **Recall** | 0.55083 | 0.56190 | 0.49056 | 0.47981 | 0.92736 |
| **F1-score** | 0.54186 | 0.58148 | 0.50667 | 0.50561 | 0.66445 |
| **50** | **Accuracy** | 0.59624 | 0.65258 | 0.56338 | 0.55869 | 0.54930 |
| **Precision** | 0.58004 | 0.65090 | 0.55556 | 0.55148 | 0.51680 |
| **Recall** | 0.58676 | 0.58527 | 0.50061 | 0.49311 | 0.86759 |
| **F1-score** | 0.58245 | 0.61599 | 0.52381 | 0.51518 | 0.64605 |
| **100** | **Accuracy** | 0.59155 | 0.62441 | 0.55399 | 0.53521 | 0.57746 |
| **Precision** | 0.60474 | 0.66657 | 0.58095 | 0.56098 | 0.55511 |
| **Recall** | 0.60115 | 0.56403 | 0.51106 | 0.49440 | 0.85639 |
| **F1-score** | 0.60024 | 0.60674 | 0.54086 | 0.52174 | 0.67304 |
| **150** | **Accuracy** | 0.61033 | 0.63380 | 0.54460 | 0.57277 | 0.56338 |
| **Precision** | 0.59700 | 0.64518 | 0.54391 | 0.58253 | 0.53512 |
| **Recall** | 0.64762 | 0.58095 | 0.46667 | 0.47619 | 0.90476 |
| **F1-score** | 0.62121 | 0.61070 | 0.49962 | 0.52300 | 0.67111 |

**Σχολιασμός Αποτελεσμάτων** : Όσο αφορά τα αποτελέσματα του activation μετά την μείωση της διάστασης των χαρακτηριστικών, παρατηρούμε οτι ο βέλτιστος αλγόριθμος είναι ο NNR-3 για πλήθος χαρακτηριστικών 50. Επίσης τα αποτελέσματα για κάθε αλγόριθμο είναι καλύτερα από αυτά που προέκυπταν πριν τη μείωση διαστάσεων.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Valence** | | | | | | |
| **PCA Components:** | **Algorithm→**  **Metrics↓** | **NNR-1** | **NNR-3** | **NNR-5** | **NNR-7** | **BAYES** |
| **10** | **Accuracy** | 0.58216 | 0.52582 | 0.49765 | 0.54460 | 0.52582 |
| **Precision** | 0.62478 | 0.57255 | 0.54921 | 0.57482 | 0.56937 |
| **Recall** | 0.68045 | 0.69089 | 0.70677 | 0.78070 | 0.72682 |
| **F1-score** | 0.65137 | 0.62484 | 0.61687 | 0.66097 | 0.63305 |
| **50** | **Accuracy** | 0.64319 | 0.58216 | 0.54460 | 0.53521 | 0.58216 |
| **Precision** | 0.68166 | 0.62346 | 0.59486 | 0.59089 | 0.61563 |
| **Recall** | 0.78274 | 0.78274 | 0.77500 | 0.75149 | 0.83321 |
| **F1-score** | 0.72713 | 0.69365 | 0.67302 | 0.66153 | 0.70544 |
| **100** | **Accuracy** | 0.55869 | 0.56808 | 0.56338 | 0.55399 | 0.57746 |
| **Precision** | 0.61879 | 0.62901 | 0.62302 | 0.61140 | 0.62016 |
| **Recall** | 0.72170 | 0.71736 | 0.72511 | 0.74188 | 0.77432 |
| **F1-score** | 0.66533 | 0.66916 | 0.66911 | 0.66886 | 0.68742 |
| **150** | **Accuracy** | 0.59155 | 0.55869 | 0.53991 | 0.54930 | 0.52113 |
| **Precision** | 0.62630 | 0.59389 | 0.57527 | 0.57842 | 0.55855 |
| **Recall** | 0.72588 | 0.73837 | 0.75689 | 0.76938 | 0.78639 |
| **F1-score** | 0.66899 | 0.65620 | 0.65059 | 0.65783 | 0.65151 |

**Σχολιασμός Αποτελεσμάτων :** Όσο αφορά τα αποτελέσματα του valence μετά την μείωση της διάστασης των χαρακτηριστικών, παρατηρούμε οτι ο βέλτιστος αλγόριθμος είναι ο NNR-1 για πλήθος χαρακτηριστικών 50. Επίσης τα αποτελέσματα για κάθε αλγόριθμο είναι καλύτερα από αυτά που προέκυπταν πριν τη μείωση διαστάσεων.

Απο τα παραπάνω αποτελέσματα έχω οτι η PCA βοηθάει στο να πάρουμε πιο γρήγορα αποτελέσματα, με καλύτερη απόδοση γεγονός το οποίο είναι πολύ σημαντικό για την διαδικασία της ταξινόμησης.

**Βήμα 15** : Στο βήμα αυτό γίνεται χρήση του εργαλείου WEKA. Για να εισαχθούν στο WEKA τα χαρακτηριστικά που υπολογίστηκαν στο Matlab μετατρέπονται σε format .arff, με χρήση του πακέτου matlab2weka το οποίο παρέχει 5 την συνάρτηση που κάνει αυτό που θέλουμε.

**Βήμα 16** : Οι αλγόριθμοι που δοκιμάζονται με χρήση του WEKA είναι οι Multilayer Perceptron με πειραματισμό στις τιμές των hidden layers και learning rate, Suppoct Vector Machines (SVM) και Random Forest. Τα δεδομένα χωρίζονται σε 5 ισοπληθή υποσύνολα και σε κάθε μία από τις 5 επαναλήψεις χρησιμοποιείται το 1 από τα 5 υποσύνολα ως validation set ενώ τα άλλα 4 ως train set. Τα αποτελέσματα φαίνονται παρακάτω :

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Activation - Features’ Set 1** | | | | | | |
| **Multilayer Perceptron** | | | | | | |
| **Number of Hidden Layers:** | **Metrics→**  **Learning Ratio↓** | **Accuracy** | **Precision** | | **Recall** | **F1-score** |
| **1** | **0.1** | 72.3944% | 0.725 | | 0.724 | 0.724 |
| **0.4** | 72.9577% | 0.730 | | 0.730 | 0.730 |
| **0.7** | 72.9577% | 0.730 | | 0.730 | 0.730 |
| **0.9** | 72.6761% | 0.728 | | 0.727 | 0.727 |
| **2** | **0.1** | 73.2394% | 0.732 | | 0.732 | 0.732 |
| **0.4** | 74.0845% | 0.740 | | 0.741 | 0.740 |
| **0.7** | 71.2676% | 0.714 | | 0.713 | 0.713 |
| **0.9** | 71.2676% | 0.715 | | 0.713 | 0.713 |
| **4** | **0.1** | 71.831% | 0.718 | | 0.718 | 0.718 |
| **0.4** | 72.3944% | 0.724 | | 0.724 | 0.724 |
| **0.7** | 71.831% | 0.719 | | 0.718 | 0.718 |
| **0.9** | 70.4225% | 0.703 | | 0.704 | 0.703 |
| **6** | **0.1** | 71.2676% | 0.738 | | 0.420 | 0.803 |
| **0.4** | 73.5211% | 0.755 | | 0.467 | 0.803 |
| **0.7** | 71.831% | 0.735 | | 0.435 | 0.784 |
| **0.9** | 71.2676% | 0.745 | | 0.418 | 0.782 |
| **8** | **0.1** | 72.6761% | 0.726 | | 0.727 | 0.726 |
| **0.4** | 71.5493% | 0.716 | | 0.715 | 0.716 |
| **0.7** | 71.831% | 0.718 | | 0.718 | 0.718 |
| **0.9** | 73.8028% | 0.738 | | 0.738 | 0.738 |
| **Support Vector Machines(SVM)** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **73.8028 %** | **0.739** | **0.738** | | **0.738** | | |
| **Random Forest** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **74.9296 %** | **0.749** | **0.749** | | **0.749** | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Activation - Features’ Set 2** | | | | | | |
| **Multilayer Perceptron** | | | | | | |
| **Number of Hidden Layers:** | **Metrics→**  **Learning Ratio↓** | **Accuracy** | **Precision** | | **Recall** | **F1-score** |
| **1** | **0.1** | 71.5493% | 0.718 | | 0.715 | 0.716 |
| **0.4** | 73.2394% | 0.737 | | 0.732 | 0.733 |
| **0.7** | 76.0563% | 0.764 | | 0.761 | 0.761 |
| **0.9** | 74.3662% | 0.746 | | 0.744 | 0.744 |
| **2** | **0.1** | 69.5775% | 0.698 | | 0.696 | 0.696 |
| **0.4** | 71.831% | 0.718 | | 0.718 | 0.718 |
| **0.7** | 71.2676% | 0.712 | | 0.713 | 0.711 |
| **0.9** | 74.3662% | 0.744 | | 0.744 | 0.744 |
| **4** | **0.1** | 72.9577% | 0.733 | | 0.730 | 0.730 |
| **0.4** | 69.5775% | 0.696 | | 0.696 | 0.696 |
| **0.7** | 72.3944% | 0.724 | | 0.724 | 0.724 |
| **0.9** | 74.9296% | 0.750 | | 0.749 | 0.749 |
| **6** | **0.1** | 72.1127% | 0.722 | | 0.721 | 0.721 |
| **0.4** | 71.831% | 0.719 | | 0.718 | 0.719 |
| **0.7** | 73.8028% | 0.738 | | 0.738 | 0.738 |
| **0.9** | 71.831% | 0.719 | | 0.718 | 0.718 |
| **8** | **0.1** | 71.831% | 0.721 | | 0.718 | 0.719 |
| **0.4** | 70.9859% | 0.710 | | 0.710 | 0.710 |
| **0.7** | 72.1127% | 0.722 | | 0.721 | 0.721 |
| **0.9** | 69.5775% | 0.696 | | 0.696 | 0.696 |
| **Support Vector Machines(SVM)** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **75.7746 %** | **0.757** | **0.758** | | **0.757** | | |
| **Random Forest** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **75.7746 %** | **0.757** | **0.758** | | **0.757** | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Activation - Features’ Set 3** | | | | | | |
| **Multilayer Perceptron** | | | | | | |
| **Number of Hidden Layers:** | **Metrics→**  **Learning Ratio↓** | **Accuracy** | **Precision** | | **Recall** | **F1-score** |
| **1** | **0.1** | 72.3944% | 0.739 | | 0.447 | 0.779 |
| **0.4** | 72.3944% | 0.743 | | 0.445 | 0.787 |
| **0.7** | 73.2394% | 0.751 | | 0.462 | 0.777 |
| **0.9** | 73.2394% | 0.748 | | 0.464 | 0.773 |
| **2** | **0.1** | 73.5211% | 0.735 | | 0.735 | 0.735 |
| **0.4** | 73.5211% | 0.736 | | 0.735 | 0.735 |
| **0.7** | 73.5211% | 0.735 | | 0.735 | 0.735 |
| **0.9** | 73.8028% | 0.738 | | 0.738 | 0.738 |
| **4** | **0.1** | 71.5493% | 0.717 | | 0.715 | 0.716 |
| **0.4** | 71.5493% | 0.716 | | 0.715 | 0.716 |
| **0.7** | 72.3944% | 0.725 | | 0.724 | 0.724 |
| **0.9** | 71.831% | 0.718 | | 0.718 | 0.718 |
| **6** | **0.1** | 71.831% | 0.719 | | 0.718 | 0.719 |
| **0.4** | 71.831% | 0.719 | | 0.718 | 0.718 |
| **0.7** | 73.5211% | 0.737 | | 0.735 | 0.736 |
| **0.9** | 71.2676% | 0.714 | | 0.713 | 0.713 |
| **8** | **0.1** | 73.8028% | 0.739 | | 0.738 | 0.738 |
| **0.4** | 73.8028% | 0.738 | | 0.738 | 0.738 |
| **0.7** | 71.2676% | 0.712 | | 0.713 | 0.712 |
| **0.9** | 72.1127% | 0.721 | | 0.721 | 0.721 |
| **Support Vector Machines(SVM)** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **75.7746 %** | **0.758** | **0.758** | | **0.758** | | |
| **Random Forest** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **79.1549 %** | **0.792** | **0.792** | | **0.792** | | |

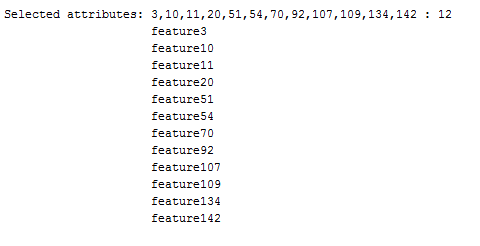
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Valence - Features’ Set 1** | | | | | | |
| **Multilayer Perceptron** | | | | | | |
| **Number of Hidden Layers:** | **Metrics→**  **Learning Ratio↓** | **Accuracy** | **Precision** | | **Recall** | **F1-score** |
| **1** | **0.1** | 73.8162% | 0.735 | | 0.738 | 0.736 |
| **0.4** | 72.9805% | 0.724 | | 0.730 | 0.725 |
| **0.7** | 72.9805% | 0.726 | | 0.730 | 0.727 |
| **0.9** | 73.2591% | 0.732 | | 0.733 | 0.732 |
| **2** | **0.1** | 74.0947% | 0.736 | | 0.741 | 0.736 |
| **0.4** | 73.8162% | 0.733 | | 0.738 | 0.731 |
| **0.7** | 73.5376% | 0.730 | | 0.735 | 0.729 |
| **0.9** | 73.8162% | 0.733 | | 0.738 | 0.733 |
| **4** | **0.1** | 75.2089% | 0.749 | | 0.752 | 0.750 |
| **0.4** | 72.1448% | 0.718 | | 0.721 | 0.719 |
| **0.7** | 72.7019% | 0.724 | | 0.727 | 0.725 |
| **0.9** | 73.8162% | 0.735 | | 0.738 | 0.736 |
| **6** | **0.1** | 73.5376% | 0.630 | | 0.426 | 0.762 |
| **0.4** | 71.8663% | 0.601 | | 0.386 | 0.732 |
| **0.7** | 70.7521% | 0.578 | | 0.359 | 0.730 |
| **0.9** | 68.2451% | 0.581 | | 0.325 | 0.691 |
| **8** | **0.1** | 74.9304% | 0.745 | | 0.749 | 0.746 |
| **0.4** | 68.8022% | 0.685 | | 0.688 | 0.687 |
| **0.7** | 71.3092% | 0.709 | | 0.713 | 0.710 |
| **0.9** | 68.8022% | 0.684 | | 0.688 | 0.685 |
| **Support Vector Machines(SVM)** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **74.0947 %** | **0.742** | **0.741** | | **0.724** | | |
| **Random Forest** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **71.5877 %** | **0.711** | **0.716** | | **0.712** | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Valence - Features’ Set 2** | | | | | | |
| **Multilayer Perceptron** | | | | | | |
| **Number of Hidden Layers:** | **Metrics→**  **Learning Ratio↓** | **Accuracy** | **Precision** | | **Recall** | **F1-score** |
| **1** | **0.1** | 77.4373% | 0.775 | | 0.774 | 0.775 |
| **0.4** | 77.7159% | 0.775 | | 0.777 | 0.776 |
| **0.7** | 77.7159% | 0.777 | | 0.777 | 0.777 |
| **0.9** | 78.273% | 0.783 | | 0.783 | 0.783 |
| **2** | **0.1** | 76.8802% | 0.767 | | 0.769 | 0.767 |
| **0.4** | 77.4373% | 0.771 | | 0.774 | 0.772 |
| **0.7** | 77.4373% | 0.772 | | 0.774 | 0.773 |
| **0.9** | 77.4373% | 0.771 | | 0.774 | 0.771 |
| **4** | **0.1** | 77.1588% | 0.769 | | 0.772 | 0.770 |
| **0.4** | 77.7159% | 0.774 | | 0.777 | 0.775 |
| **0.7** | 76.3231% | 0.762 | | 0.763 | 0.763 |
| **0.9** | 74.6518% | 0.745 | | 0.747 | 0.746 |
| **6** | **0.1** | 77.1588% | 0.770 | | 0.772 | 0.771 |
| **0.4** | 75.4875% | 0.755 | | 0.755 | 0.755 |
| **0.7** | 75.4875% | 0.754 | | 0.755 | 0.754 |
| **0.9** | 75.766% | 0.760 | | 0.758 | 0.759 |
| **8** | **0.1** | 76.3231% | 0.762 | | 0.763 | 0.762 |
| **0.4** | 77.9944% | 0.779 | | 0.780 | 0.779 |
| **0.7** | 76.3231% | 0.762 | | 0.763 | 0.763 |
| **0.9** | 74.9304% | 0.752 | | 0.749 | 0.750 |
| **Support Vector Machines(SVM)** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **79.1086 %** | **0.789** | **0.791** | | **0.789** | | |
| **Random Forest** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **77.7159 %** | **0.776** | **0.777** | | **0.769** | | |

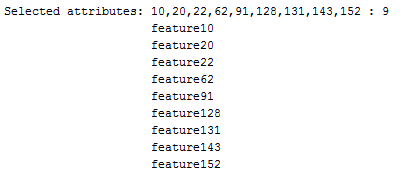
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Valence - Features’ Set 3** | | | | | | |
| **Multilayer Perceptron** | | | | | | |
| **Number of Hidden Layers:** | **Metrics→**  **Learning Ratio↓** | **Accuracy** | **Precision** | | **Recall** | **F1-score** |
| **1** | **0.1** | 77.1588% | 0.689 | | 0.509 | 0.822 |
| **0.4** | 77.1588% | 0.687 | | 0.508 | 0.803 |
| **0.7** | 78.5515% | 0.714 | | 0.542 | 0.796 |
| **0.9** | 79.9443% | 0.739 | | 0.577 | 0.820 |
| **2** | **0.1** | 77.4373% | 0.772 | | 0.774 | 0.773 |
| **0.4** | 76.6017% | 0.762 | | 0.766 | 0.763 |
| **0.7** | 75.766% | 0.755 | | 0.758 | 0.755 |
| **0.9** | 76.8802% | 0.766 | | 0.769 | 0.766 |
| **4** | **0.1** | 78.8301% | 0.790 | | 0.788 | 0.789 |
| **0.4** | 76.8802% | 0.768 | | 0.769 | 0.768 |
| **0.7** | 75.766% | 0.758 | | 0.758 | 0.758 |
| **0.9** | 76.0446% | 0.759 | | 0.760 | 0.760 |
| **6** | **0.1** | 77.7159% | 0.778 | | 0.777 | 0.777 |
| **0.4** | 77.7159% | 0.776 | | 0.777 | 0.776 |
| **0.7** | 77.9944% | 0.778 | | 0.780 | 0.779 |
| **0.9** | 76.6017% | 0.764 | | 0.766 | 0.764 |
| **8** | **0.1** | 77.9944% | 0.780 | | 0.780 | 0.780 |
| **0.4** | 77.9944% | 0.779 | | 0.780 | 0.779 |
| **0.7** | 78.273% | 0.781 | | 0.783 | 0.782 |
| **0.9** | 78.5515% | 0.784 | | 0.786 | 0.785 |
| **Support Vector Machines(SVM)** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **79.1086 %** | **0.789** | **0.791** | | **0.790** | | |
| **Random Forest** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **75.2089 %** | **0.748** | **0.752** | | **0.743** | | |

**Βήμα 17** : Στο τελευταίο βήμα για να πετύχουμε καλύτερη επίδοση χρησιμοποιούμε την μέθοδο επιλογής χαρακτηριστικών(feature selection), μέσω του εργαλείου Weka. Για την επιλογή χαρακτηριστικών πραγματοποιούνται οι διαδικασίες Attribute Evaluation,Search. Για Attribute Evaluator επιλέγεται η Wrapper, κατά την οποία γίνεται αξιολόγηση των υποσυνόλων των χαρακτηριστικών και επιλογή ως τελικού συνόλου αυτού που διαχωρίζει καλύτερα τα δεδομένα. Για classifier επιλέχθηκαν τα SVMs. Για Search Method επιλέχθηκε η BestFirst ώστε να γίνει αναζήτηση σε όλα τα πιθανά υποσύνολα χαρακτηριστικών με αποδοτικό τρόπο. Τα αποτελέσματα φαίνονται παρακάτω:

Attributes selected for Activation : 3 10 11 20 51 54 70 92 107 109 134 142 (total 12)



Attributes selected for Valence : 10 20 22 62 91 128 131 143 152 (total 9)



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Activation - Features’ Set 3** | | | | | | |
| **Multilayer Perceptron** | | | | | | |
| **Number of Hidden Layers:** | **Metrics→**  **Learning Ratio↓** | **Accuracy** | **Precision** | | **Recall** | **F1-score** |
| **1** | **0.1** | 76.6197% | 0.766 | | 0.766 | 0.766 |
| **0.4** | 76.338% | 0.763 | | 0.763 | 0.763 |
| **0.7** | 74.3662% | 0.743 | | 0.744 | 0.743 |
| **0.9** | 76.6197% | 0.766 | | 0.766 | 0.766 |
| **2** | **0.1** | 76.338% | 0.767 | | 0.763 | 0.764 |
| **0.4** | 74.9296% | 0.754 | | 0.749 | 0.750 |
| **0.7** | 73.5211% | 0.741 | | 0.735 | 0.736 |
| **0.9** | 73.5211% | 0.740 | | 0.735 | 0.736 |
| **4** | **0.1** | 54.3662% | 0.296 | | 0.544 | 0.383 |
| **0.4** | 52.3944% | 0.500 | | 0.524 | 0.476 |
| **0.7** | 52.3944% | 0.500 | | 0.524 | 0.476 |
| **0.9** | 52.3944% | 0.500 | | 0.524 | 0.476 |
| **6** | **0.1** | 54.3662% | 0.296 | | 0.544 | 0.383 |
| **0.4** | 52.3944% | 0.500 | | 0.524 | 0.476 |
| **0.7** | 52.3944% | 0.500 | | 0.524 | 0.476 |
| **0.9** | 52.3944% | 0.500 | | 0.524 | 0.476 |
| **8** | **0.1** | 54.3662% | 0.296 | | 0.544 | 0.383 |
| **0.4** | 52.3944% | 0.500 | | 0.524 | 0.476 |
| **0.7** | 52.3944% | 0.500 | | 0.524 | 0.476 |
| **0.9** | 52.3944% | 0.500 | | 0.524 | 0.476 |
| **Support Vector Machines(SVM)** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **84.507 %** | **0.845** | **0.845** | | **0.845** | | |
| **Random Forest** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **80.4648 %** | **0.801** | **0.804** | | **0.801** | | |

Πριν την επιλογή χαρακτηριστικών: Max Accuracy = 79,1549 % , Max F1 = 79,2%

Μετά την επιλογή χαρακτηριστικών: Max Accuracy = 84,4648 % , Max F1 = 84,5%

Παρατηρούμε ότι στην περίπτωσητου Activation τα αποτελέσματα βελτιώνονται αισθητά με την επιλογή χαρακτηριστικών.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Valence - Features’ Set 3** | | | | | | |
| **Multilayer Perceptron** | | | | | | |
| **Number of Hidden Layers:** | **Metrics→**  **Learning Ratio↓** | **Accuracy** | **Precision** | | **Recall** | **F1-score** |
| **1** | **0.1** | 82.4513% | 0.823 | | 0.825 | 0.823 |
| **0.4** | 79.6657% | 0.794 | | 0.797 | 0.794 |
| **0.7** | 77.9944% | 0.777 | | 0.780 | 0.777 |
| **0.9** | 79.6657% | 0.795 | | 0.797 | 0.795 |
| **2** | **0.1** | 82.1727% | 0.820 | | 0.822 | 0.820 |
| **0.4** | 79.6657% | 0.797 | | 0.797 | 0.797 |
| **0.7** | 81.0585% | 0.811 | | 0.811 | 0.811 |
| **0.9** | 81.0585% | 0.809 | | 0.811 | 0.809 |
| **4** | **0.1** | 62.6741% | 0.393 | | 0.627 | 0.483 |
| **0.4** | 62.6741% | 0.393 | | 0.627 | 0.483 |
| **0.7** | 62.6741% | 0.393 | | 0.627 | 0.483 |
| **0.9** | 62.6741% | 0.393 | | 0.627 | 0.483 |
| **6** | **0.1** | 62.6741% | 0.393 | | 0.627 | 0.483 |
| **0.4** | 62.6741% | 0.393 | | 0.627 | 0.483 |
| **0.7** | 62.6741% | 0.393 | | 0.627 | 0.483 |
| **0.9** | 62.6741% | 0.393 | | 0.627 | 0.483 |
| **8** | **0.1** | 62.6741% | 0.393 | | 0.627 | 0.483 |
| **0.4** | 62.6741% | 0.393 | | 0.627 | 0.483 |
| **0.7** | 62.6741% | 0.393 | | 0.627 | 0.483 |
| **0.9** | 62.6741% | 0.393 | | 0.627 | 0.483 |
| **Support Vector Machines(SVM)** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **84.9582 %** | **0.853** | **0.850** | | **0.845** | | |
| **Random Forest** | | | | | | |
| **Accuracy** | **Precision** | **Recall** | | **F1-score** | | |
| **79.3872 %** | **0.791** | **0.794** | | **0.791** | | |

Πριν την επιλογή χαρακτηριστικών: Max Accuracy = 79,9443% , Max F1 = 82%

Μετά την επιλογή χαρακτηριστικών: Max Accuracy = 84,9582% , Max F1 = 84,5%

Παρατηρούμε ότι στην περίπτωσητου Valence τα αποτελέσματα βελτιώνονται αισθητά με την επιλογή χαρακτηριστικών.