

ΔΔΠΜΣ “Προηγμένα Συστήματα Υπολογιστών και Επικοινωνιών”, 2019-2020
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Τεχνικές Μηχανικής Μάθησης

Ταξινόμηση αναπνευστικών ασθενειών και ήχων

Μέλη της ομάδας

Τσαλίδης Γεώργιος, AEM:472
Κασπαρίδης Παναγιώτης, AEM:448

Email

gtsalidis@ece.auth.gr
pkasparid@ece.auth.gr

Εισαγωγή και περιγραφή προβλήματος

Οι αναπνευστικοί ήχοι είναι σημαντικοί δείκτες της αναπνευστικής υγείας και των αναπνευστικών διαταραχών. Ο ήχος που εκπέμπεται όταν ένα άτομο αναπνέει σχετίζεται άμεσα με την κίνηση του αέρα, τις αλλαγές στον πνευμονικό ιστό και τη θέση των εκκρίσεων μέσα στον πνεύμονα. Ένας ήχος συριγμού, για παράδειγμα, είναι ένα κοινό σημάδι ότι ένας ασθενής έχει μια αποφρακτική ασθένεια των αεραγωγών όπως το άσθμα ή η χρόνια αποφρακτική πνευμονοπάθεια (COPD).

Αυτοί οι ήχοι μπορούν να καταγραφούν χρησιμοποιώντας ψηφιακά στηθοσκόπια και άλλες τεχνικές καταγραφής. Αυτά τα ψηφιακά δεδομένα ανοίγουν τη δυνατότητα χρήσης μηχανικής μάθησης για την αυτόματη διάγνωση αναπνευστικών διαταραχών όπως η χρόνια αποφρακτική πνευμονοπάθεια, το άσθμα, η πνευμονία ή βρογχιολίτιδα, η λοίμωξη της άνω αναπνευστικής οδού κ.α.

Οι αναπνευστικές ασθένειες προκαλούν τεράστια επιβάρυνση στην ζωή ενός ανθρώπου και αποτελούν την τρίτη κύρια αιτία θανάτου παγκοσμίως και σημαντική επιβάρυνση για τα δημόσια συστήματα υγείας. Ως εκ τούτου, γίνονται σημαντικές ερευνητικές προσπάθειες με στόχο τη βελτίωση της έγκαιρης διάγνωσης και της αυτοματοποιημένης παρακολούθησης ρουτίνας ασθενών με αναπνευστικές νόσους για να μπορέσουν να κάνουν έγκαιρες παρεμβάσεις. Η αυτόματη ανάλυση των αναπνευστικών ήχων είναι ένα πεδίο με μεγάλο ερευνητικό ενδιαφέρον κατά τις τελευταίες δεκαετίες. Η αυτοματοποιημένη ταξινόμηση των αναπνευστικών ήχων έχει τη δυνατότητα ανίχνευσης ανωμαλιών ενός αναπνευστικού συστήματος και την δημιουργία ενός μοντέλου για την ταξινόμηση αναπνευστικών ασθενειών.

Διερεύνηση και οπτικοποίηση δεδομένων

Για την επίλυση του προβλήματος και τη δημιουργία ταξινομητή ήχων, χρησιμοποιήθηκε σύνολο δεδομένων ([RespiratorySoundDatabase\(ICBHI 2017 Challenge\)](#)) που δημιουργήθηκε από δύο ερευνητικές ομάδες στην Πορτογαλία και την Ελλάδα και περιλαμβάνει 920 εγγραφές από 126 άτομα.

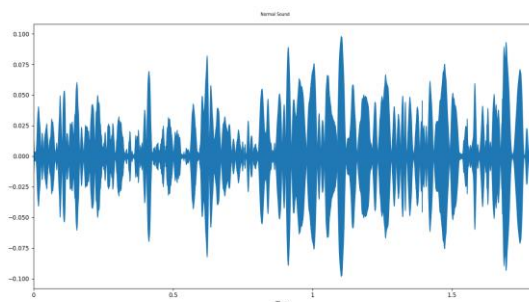
Πιο συνοπτικά το σύνολο δεδομένων αποτελείται από:

- 920 .wav αρχεία ήχου
- 920.txt αρχεία με τους αναπνευστικούς κύκλους για κάθε αρχείο ήχου
- Ένα αρχείο κειμένου που περιέχει τη διάγνωση για κάθε ασθενή
- Ένα αρχείο κειμένου που εξηγεί τη μορφή ονομασίας των αρχείων
- Ένα αρχείο κειμένου που περιέχει δημογραφικές πληροφορίες για κάθε ασθενή

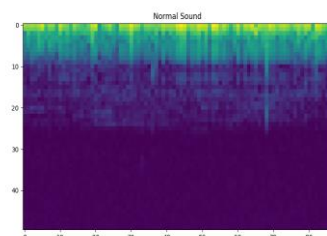
Σε αυτά τα αρχεία ήχου υπάρχουν κύκλοι αναπνοής. Οι κύκλοι σχολιάστηκαν από εμπειρογνώμονες της αναπνευστικής οδού που τα χώρισαν σε τέσσερις κατηγορίες, ρωγμές(crackles), συριγμούς(wheezes), και τα 2 προηγούμενα ή του τυχαίου αναπνευστικού ήχου, δηλαδή χωρίς την ύπαρξη κάποιων από τα προηγούμενα. Οι ηχογραφήσεις συλλέχθηκαν χρησιμοποιώντας ετερογενή εξοπλισμό και η διάρκεια κυμάνθηκε από 10 έως 90 δευτερόλεπτα. Οι θέσεις του θώρακα από τις οποίες οι ηχογραφήσεις είχαν αποκτηθεί ήταν επίσης διαθέσιμες. Οι ανεπιθύμητες ενέργειες σε ορισμένους κύκλους αναπνοής ήταν υψηλές, οι οποίες προσομοιάζουν σε πραγματικές συνθήκες ζωής και καθιστούν πιο δύσκολη τη διαδικασία ταξινόμησης.

Αρχικά, γίνεται διερεύνηση και οπτικοποίηση των δεδομένων για καλύτερη προσέγγιση του προβλήματος.

Normal Sound



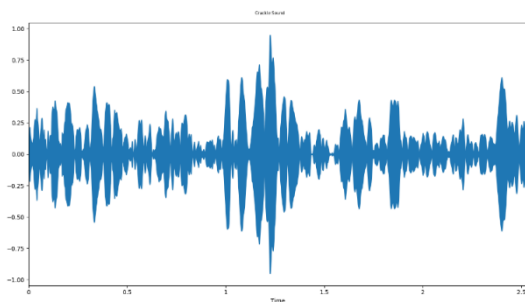
(a)



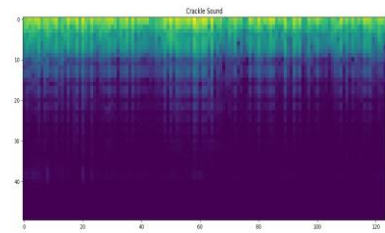
(b)

Εικόνα 1 Παράδειγμα κανονικού αναπνευστικού ήχου, (a)κυματομορφή και (b)MFCC

Crackle Sound



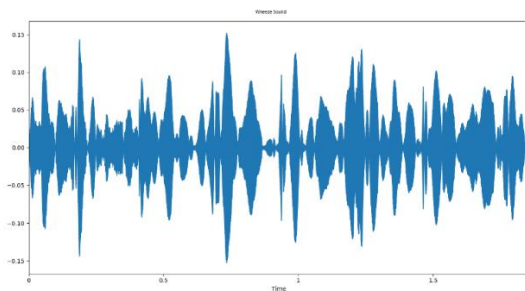
(a)



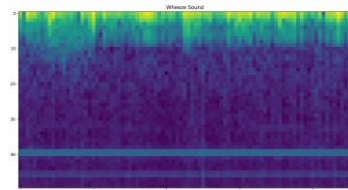
(b)

Εικόνα 2 Παράδειγμα αναπνευστικού ήχου με crackle, (a)κυματομορφή και (b)MFCC

Wheeze sound



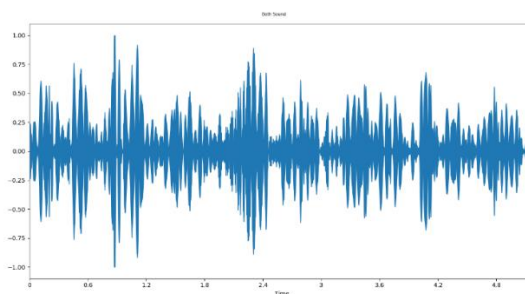
(a)



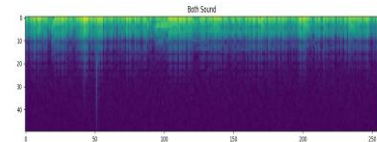
(b)

Εικόνα 3 Παράδειγμα αναπνευστικού ήχου με wheeze, (a)κυματομορφή και (b)MFCC

None Sound



(a)

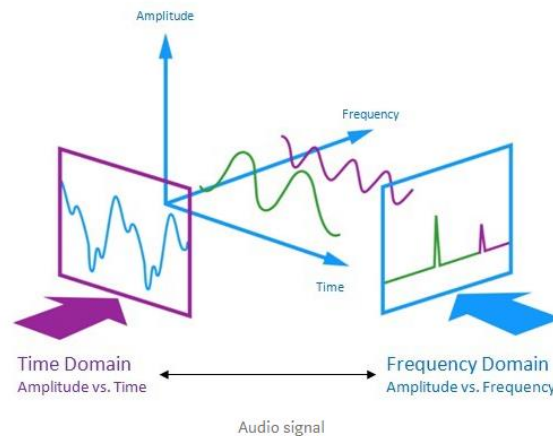


(b)

Εικόνα 4 Παράδειγμα αναπνευστικού ήχου με crackle και wheeze, (a)κυματομορφή και (b)MFCC

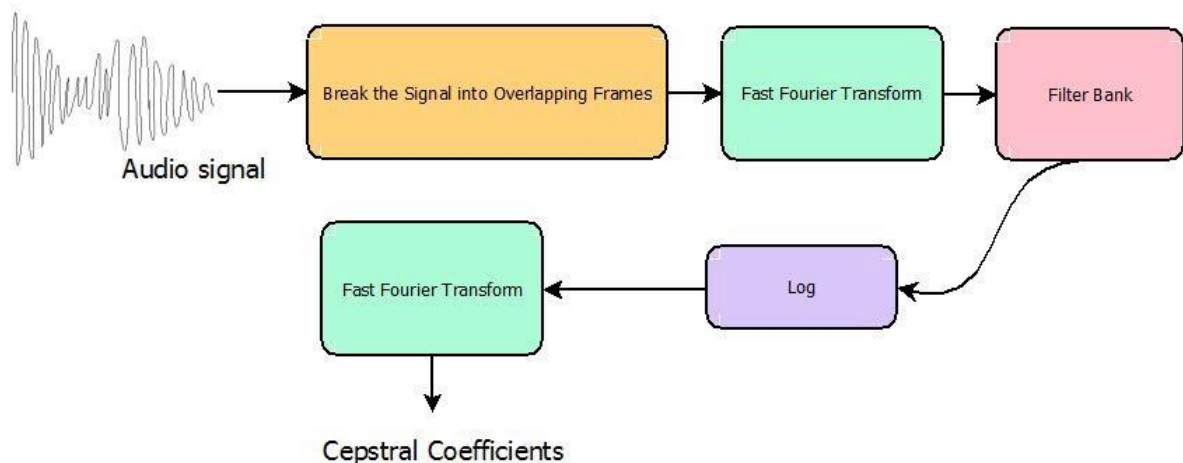
Εξαγωγή χαρακτηριστικών

Η εξαγωγή χαρακτηριστικών είναι ένα πολύ σημαντικό μέρος στην ανάλυση και την εύρεση σχέσεων μεταξύ διαφορετικών πραγμάτων. Τα δεδομένα που παρέχονται από τον ήχο δεν μπορούν να κατανοηθούν από τα μοντέλα απευθείας. Στον ήχο η πιο συνηθισμένη προ επεξεργασία των δεδομένων είναι η μετατροπή τους στο πεδίο της συχνότητας ή σε ένα συνδυασμό time-frequency domain. Πρόκειται για μια διαδικασία που εξηγεί περισσότερα στοιχεία αλλά με κατανοητό τρόπο. Για την ανάλυση του σήματος υπάρχουν διάφορες τεχνικές, Mel Frequency Cepstral Coefficients, Energy, Entropy, Spectral Centroid, Zero Cross Rate, Chroma. Ανάλογα την εφαρμογή ίσως να απαιτείται διαφορετικά χαρακτηριστικά του σήματος. Για παράδειγμα στη δική μας περίπτωση δεν υπάρχει νόημα να δοκιμάσουμε χρωματικά χαρακτηριστικά, αυτά θα ήταν καταλληλότερα σε μουσική ή ομιλία ίσως. Η ανάλυση του σήματος έγινε με MFCC γιατί είναι μια αναπαράσταση του ήχου πιο κοντά στον τρόπο με τον οποίο ακούει ο άνθρωπος τις συχνότητες.



Εικόνα 5 Αναπαράσταση ήχου στο πεδίο του χρόνου και της συχνότητας

MFCC – Mel Frequency Cepstral Coefficients



Εικόνα 6 Διάγραμμα εξαγωγής Cepstral Coefficients, όμοια για MFCC

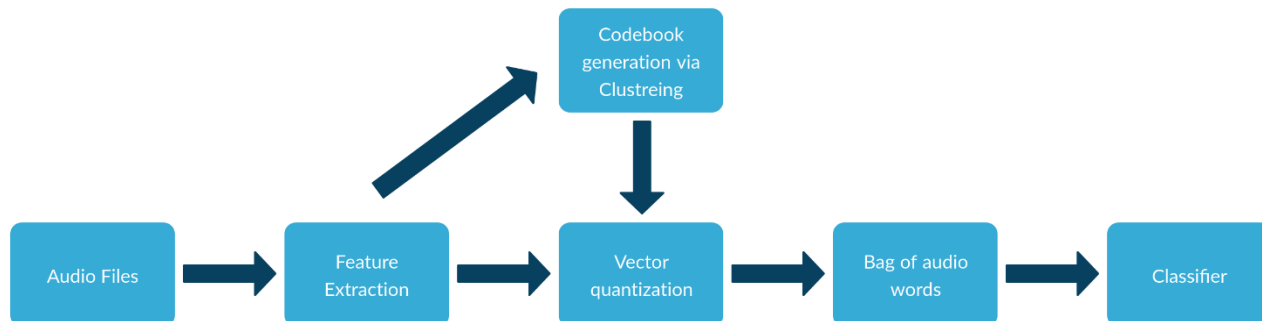
Τα MFCC είναι Cepstral Coefficients με filter banks να είναι τα mel filters δηλαδή ένας μετασχηματισμός στην κλίμακα mel. Η mel κλίμακα θεωρείται ότι είναι πιο κοντά στον τρόπο που ακούει ο άνθρωπος από την γραμμική αναπαράσταση της συχνότητας.

Για το πρόβλημα μας επειδή ο ήχος ενός αναπνευστικού κύκλου είναι μεταβλητού μεγέθους, κυμαίνεται από δέκατα έως μερικά δευτερόλεπτα, πήραμε στατιστικά, μέση τιμή και τυπική απόκλιση. Θα γίνει ανάλυση και μιας άλλης μεθόδου που μας είπατε Bag of Audio Words ωστόσο μια πρώτη προσπάθεια δεν έδειξε καλά αποτελέσματα. Σε αυτό ίσως να φταίει και η υλοποίησή μας.

Η εξαγωγή των mfcc έγινε με την βιβλιοθήκη για επεξεργασία ήχου librosa με παράθυρο των 25ms και ρυθμό με επικάλυψη ανά 10ms.

Bag of Audio Words

Η διαδικασία αυτή έγκειται να λύσει το πρόβλημα του μεταβλητού μεγέθους των δειγμάτων ήχου. Η εφαρμογή του είναι βασίζεται στην δημιουργία cluster χαρακτηριστικών από κάθε frame. Η γενική ιδέα παρουσιάζεται στην Εικόνα 7. Η λύση δεν έδωσε καλά αποτελέσματα. Για αυτό παραθέτουμε τα αποτελέσματα στο Παράρτημα μετά τις Αναφορές.



Εικόνα 7 Μπλοκ διάγραμμα χρήσης Bag of Audio Words για ταξινόμηση ήχου

Συγκεκριμένα παίρνουμε για κάθε δείγμα ήχου όλο τον πίνακα MFCC μέσω librosa. Αυτός έχει μέγεθος Mel_banks, (μεταβλητό ανάλογα το αρχείο ήχου). Στοιβαζουμε όλα τα mfcc από ολόκληρο το σέτ και το θεωρούμε ότι είναι το σέτ για το clustering. Αυτό έχει γνωστό αριθμό στην μια διάσταση και προσπαθούμε να αντιστοιχίσουμε cluster. Έπειτα για κάθε δείγμα ήχου αντιστοιχούμε τα frame σε κάποιο κέντρο του cluster. Έτσι έχουμε την τελική αναπαράσταση του ήχου σαν ένα διάνυσμα συχνοτήτων εμφάνισης του κάθε κέντρου του cluster. Επειδή έχουμε μεταβλητό μέγεθος δειγμάτων ήχου για κάθε γραμμή του συνόλου δεδομένων κάνουμε scale, γιατί προφανώς κομμάτια ήχου που διαρκούν παραπάνω θα έχουν περισσότερα frames και άρα είναι πιθανό ο αριθμός τους να είναι μεγαλύτερος σε σχέση με τα υπόλοιπα. Εμάς μας ενδιαφέρει ποσοστιαία πόσο συχνά εμφανίστηκε και όχι ο απόλυτος αριθμός. Αυτό αποτελεί την είσοδο στα μοντέλα για ταξινόμηση.

Προ επεξεργασία χαρακτηριστικών

Πολλοί αλγόριθμοι μηχανικής μάθησης λειτουργούν καλύτερα όταν τα χαρακτηριστικά γνωρίσματα είναι σε σχετικά παρόμοια κλίμακα και κοντά σε κανονική κατανομή. Τα εργαλεία MinMaxScaler, RobustScaler, StandardScaler και Normalizer είναι μέθοδοι του scikit-learn που δοκιμάστηκαν για την προ επεξεργασία δεδομένων για την ταξινόμηση.

Εν τέλει παρατηρήθηκε ότι τα αποτελέσματα επιτυγχάνονταν με τη χρήση του StandardScaler. Ο StandardScaler τυποποιεί ένα χαρακτηριστικό αφαιρώντας τη μέση τιμή και μεταβάλλεται στη μονάδα. Η διακύμανση μονάδας σημαίνει το διαχωρισμό όλων των τιμών από την τυπική απόκλιση. Το StandardScaler δεν πληροί τον αυστηρό ορισμό της κλίμακας που αναφέρθηκε νωρίτερα. Η χρήση του καταλήγει σε μια κατανομή με τυπική απόκλιση ίση με 1 και τον μέσο όρο της κατανομής 0. Η διακύμανση είναι ίση με 1, επειδή η διακύμανση είναι το τετράγωνο της τυπικής απόκλισης.

Διερεύνηση αλγορίθμων μηχανικής μάθησης

Εφαρμογή μοντέλων

- SVMs (Οι προσεγγίσεις one-vs-one (OVO) και one-vs-rest (OVR))
- Random Forests
- Gradient boosting
- Bagging
- Logistic Regression
- Convolutional Neural Network (CNN)

Μέθοδος πολύπτυχης διασταύρωσης (k-fold cross-validation)

Parameter Tuning

Για τα παραπάνω μοντέλα έγινε προσπάθεια εξαγωγής βέλτιστων υπερπαραμέτρων. Για το λόγο αυτό σχεδόν σε όλους τους αλγορίθμους κάναμε GridSearch για έναν συνδυασμό παραμέτρων ώστε τελικά να συγκρίνουμε τα καλύτερα μοντέλα μεταξύ τους.

Επειδή το πρόβλημα είναι imbalanced οι μετρικές ακρίβειας που εξάγουμε πρέπει να λάβουν αυτή την ιδιομορφία του σετ. Για το λόγο αυτό στις περιπτώσεις όπου υπήρχε χρησιμοποιήσαμε `class_weight = balanced` ώστε να μην βάλει όλα τα δείγματα στην πρώτη κλάση που είχε και τα παραπάνω στοιχεία. Για κάθε μοντέλο το grid παραμέτρων παρουσιάζεται παρακάτω μαζί με κάποια αποτελέσματα.

Decision Trees : `min_samples_split` : 1 έως 21 με βήμα 5, `max_depth`: 10 έως 101 με βήμα 10, `max_features` 40 έως 251 με βήμα 10 και `min_samples_leaf` 1 έως 17 με βήμα 3. Οι καλύτερες παράμετροι ήταν 6, 80, 140 και 1 αντίστοιχα.

Random Forest: `n_estimators`: 1 έως 110 με βήμα 40, `max_depth`: 10 έως 101 με βήμα 40, `max_features`: 40 έως 251 με βήμα 40, `min_samples_split`: 1 έως 21 με βήμα 8 και `min_samples_leaf`: 1,10,21. Οι καλύτερες παράμετροι ήταν 81, 50, 160, 9 και 1.

Για logistic Regression βάλουμε μόνο για `C=100` και για Gradient boosting επειδή δεν είχε `class_balanced` δεν ερευνήσαμε σε grid search γιατί θα μάθαινε μόνο την πρώτη κλάση. Επίσης για bagging δοκιμάσαμε μόνο σαν `base_estimator` decision trees και για ένα συγκεκριμένο μόνο σετ παραμέτρων.

Τέλος εξετάσαμε και SVM με kernel κυρίως rbf είχαμε δοκιμάσει linear και polynomial. Ωστόσο τα καλύτερα αποτελέσματα είχαν βγει με rbf πυρήνα και οπότε προχωρήσαμε σε grid search με μόνο αυτόν. Οι παράμετροι έρευνας ήταν το `C` από 2^{-5} έως 2^{15} , `gamma` από 2^{-15} έως 2^3 και τα βάρη κάθε κλάσης ,είτε `balanced` είτε `custom`.

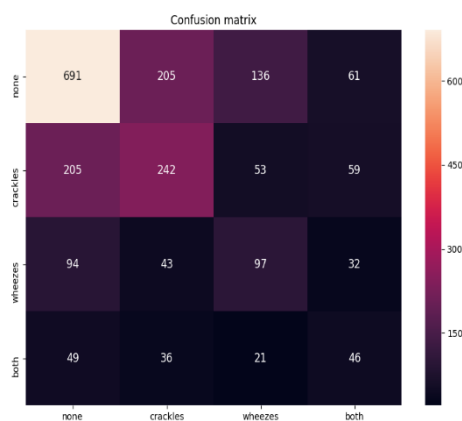
Αξιολόγηση επιδόσεων

- Η ορθότητα (accuracy) ως μετρική αξιολόγησης προβλημάτων ταξινόμησης
- Πίνακας σύγχυσης (confusion matrix)
- Ακρίβεια (precision), Ανάκληση/ευαισθησία (recall/sensitivity), specificity, F1-score
- Η περιοχή κάτω από την καμπύλη ROC (Area Under the Curve, AUC) ως αξιόπιστο μέτρο αξιολόγησης

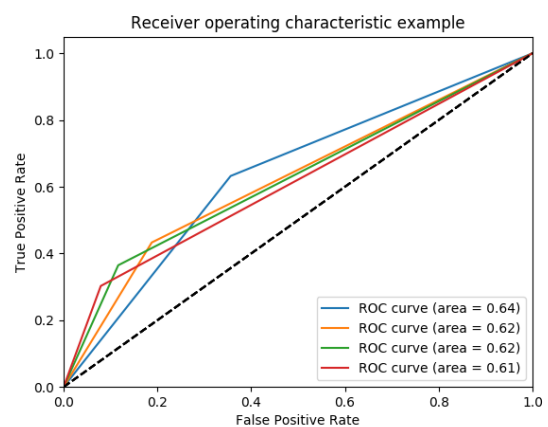
Decision Tree

Πίνακας 1 Classification Report για Decision Tree

	Precision	Recall	F1-score
0	0.67	0.63	0.65
1	0.46	0.43	0.45
2	0.32	0.36	0.34
3	0.23	0.3	0.26
Accuracy			0.52
Macro avg	0.42	0.43	0.42
Weighted avg	0.53	0.52	0.53



(a)



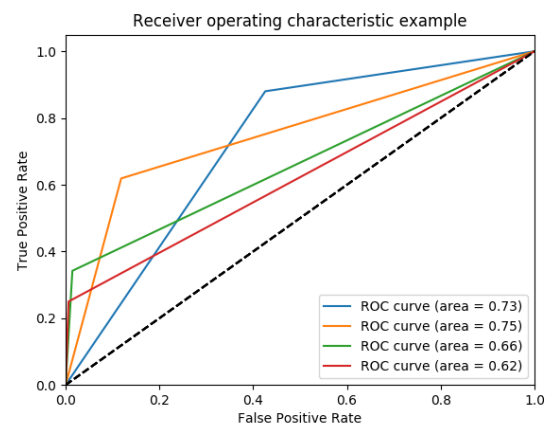
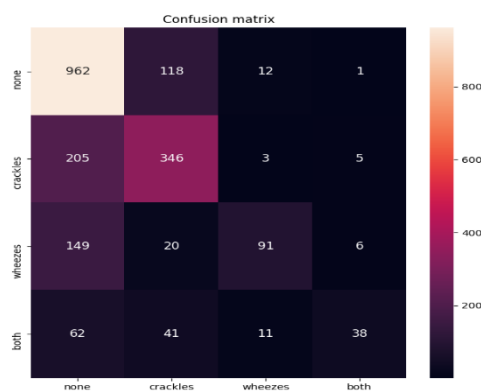
(b)

Εικόνα 8 Confusion Matrix and ROC for Decision Tree

Random Forest

Πίνακας 2 Classification Report για Random Forest

	Precision	Recall	F1-score
0	0.7	0.88	0.78
1	0.66	0.62	0.64
2	0.78	0.34	0.48
3	0.76	0.25	0.38
Accuracy			0.69
Macro avg	0.72	0.52	0.57
Weighted avg	0.70	0.69	0.67



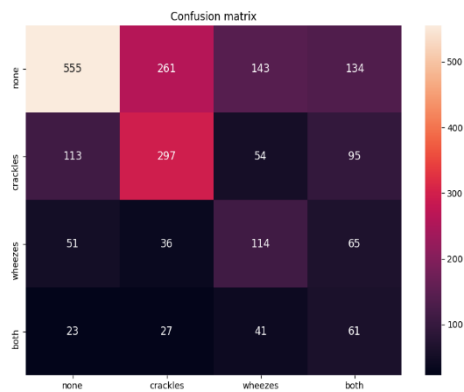
(a)

Εικόνα 9 Confusion Matrix and ROC for Random Forest

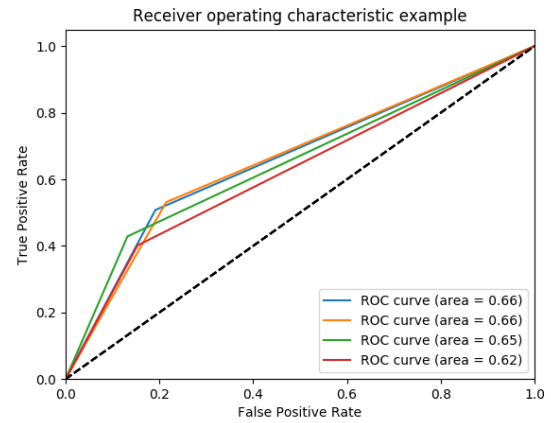
Logistic Regression

Πίνακας 3 Classification Report για Logistic Regression

	Precision	Recall	F1-score
0	0.75	0.51	0.60
1	0.48	0.53	0.50
2	0.32	0.43	0.37
3	0.17	0.40	0.24
Accuracy			0.5
Macro avg	0.43	0.47	0.43
Weighted avg	0.58	0.5	0.52



(a)



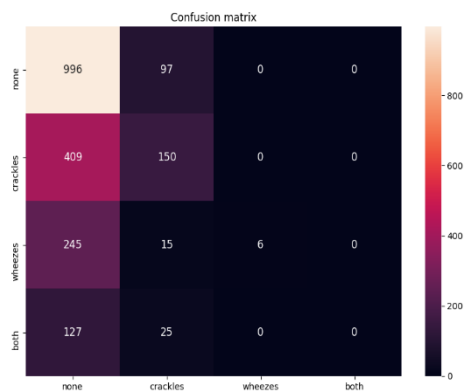
(b)

Εικόνα 10 Confusion Matrix and ROC for Logistic Regression

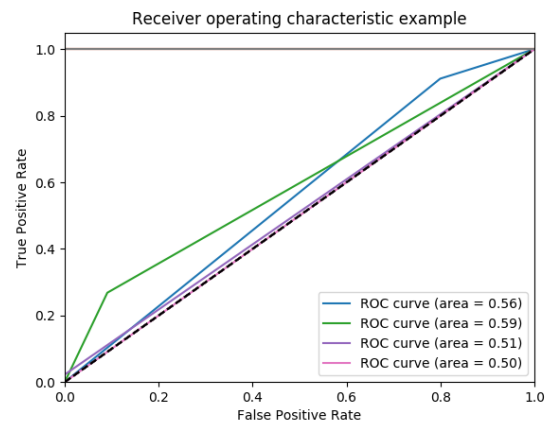
Bagging

Πίνακας 4 Classification Report για Bagging

	Precision	Recall	F1-score
0	0.56	0.91	0.69
1	0.52	0.27	0.35
2	1	0.02	0.04
3	0	0	0
Accuracy			0.56
Macro avg	0.52	0.3	0.27
Weighted avg	0.57	0.56	0.47



(a)



(b)

Εικόνα 11 Confusion Matrix and ROC for Bagging

Gradient Boosting

Πίνακας 5 Classification Report για Gradient Boosting

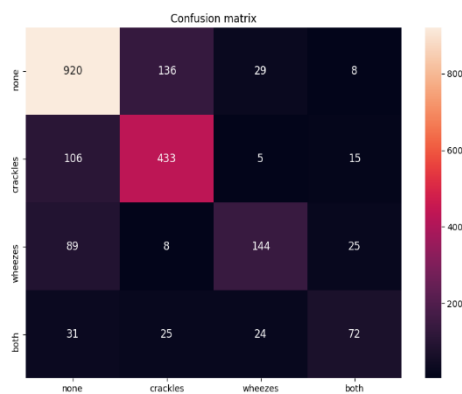
	Precision	Recall	F1-score
0	0.53	1	0.69
1	0	0	0
2	0	0	0

3	0	0	0
Accuracy			0.53
Macro avg	0.13	0.25	0.17
Weighted avg	0.28	0.53	0.36

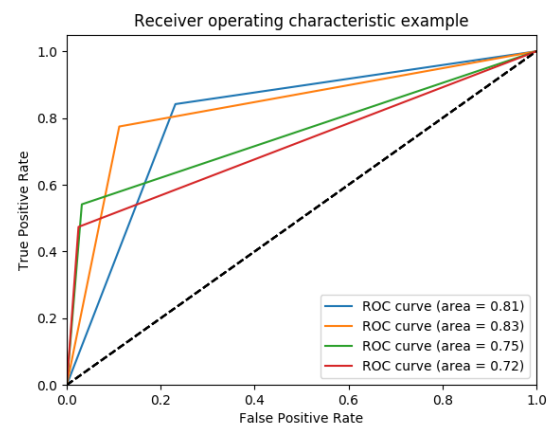
SVM

Πίνακας 7 Classification Report για SVM

	Precision	Recall	F1-score
0	0.8	0.84	0.82
1	0.72	0.77	0.75
2	0.71	0.54	0.62
3	0.6	0.47	0.53
Accuracy			0.76
Macro avg	0.71	0.66	0.68
Weighted avg	0.75	0.76	0.75



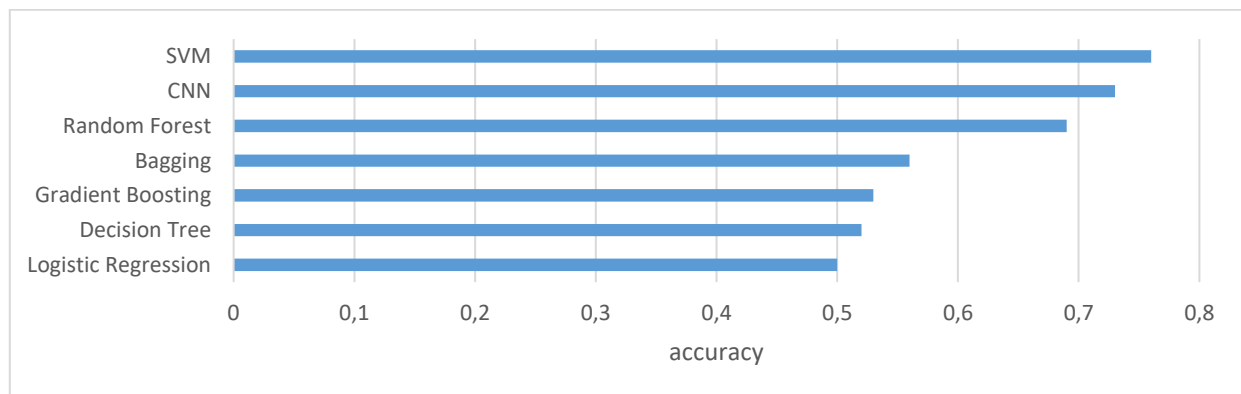
(a)



(b)

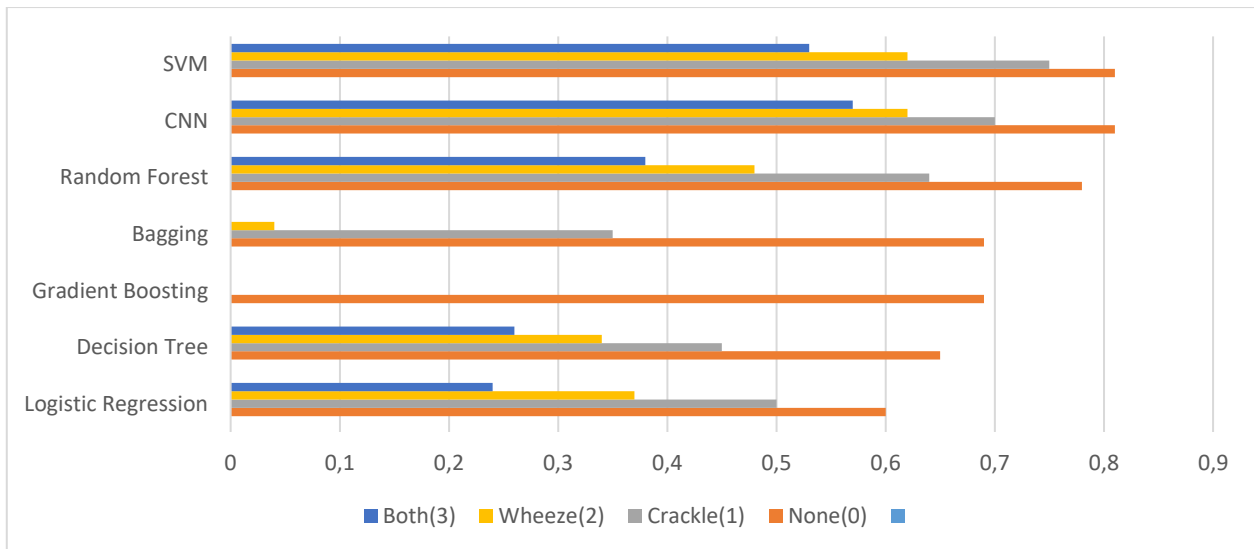
Εικόνα 12 Confusion Matrix and ROC for SVM

Σύγκριση επιδόσεων, επιλογή μοντέλου και συμπεράσματα



Διάγραμμα 1: Ορθότητα(accuracy) των μοντέλων που δοκιμάστηκαν

* CNN: Είναι kernel στο Kaggle: CNN: Detection of wheezes and crackles [εδώ](#).



Διάγραμμα 2: F1-scores για κάθε κλάση, για κάθε μοντέλο

Όπως παρατηρείται και από το διάγραμμα με τα f1-scores για κάθε κλάση(Διάγραμμα 1), καθώς και από τα confusion matrix για κάθε μοντέλο που αναφέρθηκαν παραπάνω, τα μοντέλα SVM, Random Forests και CNN⁽¹⁾ τα καταφέρνουν καλύτερα στην κλάση “none”, καθώς και στην κλάση “crackle”, ενώ δυσκολεύονται αρκετά στην κλάση “wheeze”. Στην κλάση “both” όπου είναι και η πιο ασθενής κλάση τα f1-scores στα μοντέλα που δοκιμάστηκαν είναι χαμηλά, με υψηλότερο 0.57 από το CNN και με δεύτερο 0.53 από τον SVM.

Γενικότερα, τα μοντέλα SVM, Random Forests και CNN⁽²⁾ επιφέρουν τα καλύτερα αποτελέσματα. Τελικό μοντέλο επιλογής ήταν ο SVM με πυρήνα rbf και υπερ-παραμέτρους C=4 και gamma = 0.0078125 με βάρη balanced. Τα αποτελέσματα αναφέρθηκαν παραπάνω. Στον Πίνακα 6 βλέπουμε το Classification Report του SVM με ένα overall accuracy στα 76%. Στην Εικόνα 12 παρουσιάζονται ο πίνακας σύγχυσης και οι ROC καμπύλες για τις 4 κλάσεις.

Τα αποτελέσματα του svm παρουσιάζονται στα παραπάνω διαγράμματα. Το τελικό accuracy που πετυχαίνει είναι 76%. Ωστόσο κι εδώ υπάρχει πρόβλημα με τις πιο ασθενής κλάσεις ωστόσο το overall είναι πολύ καλύτερο σε σχέση με τα υπόλοιπα μοντέλα.

(1) Είναι kernel στο Kaggle CNN: Detection of wheezes and crackles [εδώ](#).

Αναφορές

- [1] N. Sengupta, M. Sahidullah και G. Saha, «Lung sound classification using cepstral-based statistical features,» Preprint submitted to Computers in Biology and Medicine, 2016.
- [2] B. Rocha, D. Filos, L. Mendes, I. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jácome, A. Marques και R. Paiva, «A Respiratory Sound Database for the Development of Automated Classification,» *Precision Medicine Powered by pHealth and Connected Health*, pp. 51-55, 2018.
- [3] M. & S. G. Sahidullah, «Design, analysis and experimental evaluation of block based transformation in MFCC computation for speaker recognition,» *Speech Communication*, τόμ. 54, pp. 543-565, 2012.
- [4] M. Schmitt, C. Janott, V. Pandit, K. Qian, C. Heiser, W. Hemmert και B. Schuller, «A Bag-of-Audio-Words Approach for Snore Sounds' Excitation Localisation,» *Speech Communication; 12. ITG Symposium*, pp. 1-5, 2016.

Παράρτημα

Αποτελέσματα Bag of Audio Words

Για το bag of audio words χρειάζεται η δημιουργία του vocabulary. Αυτό εδώ φτιάχνεται με clustering με έναν αριθμό κέντρων, δοκιμάσαμε από 500 έως 3000 κέντρα. Τα αποτελέσματα δεν ήταν καλά γι αυτό το παραθέτουμε εδώ σαν μια έξτρα υλοποίηση. Οι αλγόριθμοι clustering που δοκιμάσαμε είναι ο kmeans, με sklearn και βιβλιοθήκη MiniBatchKMeans, και ο bdscore από την βιβλιοθήκη hdbscan, με την βιβλιοθήκη του sklearn αργούσε υπερβολικά πολύ το clustering ή και δεν έτρεχε καθόλου λόγο μνήμης. Για τους αλγορίθμους έχει γίνει ένα grid search για την επιλογή υπερπαραμέτρων ωστόσο στην περίπτωση του kmeans παρατηρήθηκε κάτι περίεργο.

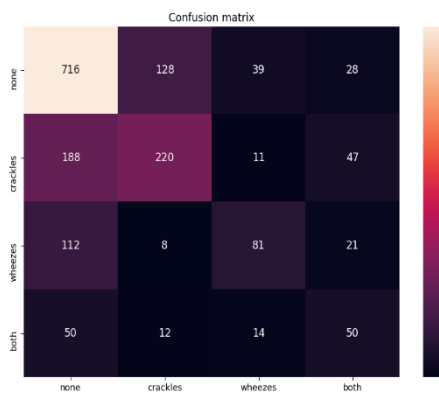
Kmeans

SVM(C=400000000,gamma=1000000*,kernel=rbf,weights balanced)*

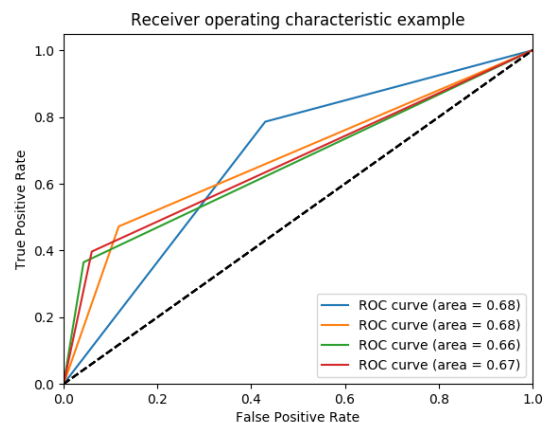
Αυτό που μας φάνηκε περίεργο είναι τα παρα πολύ μεγάλα νούμερα για C,gamma. Οτι αλλαγές στις τιμές τους και να κάναμε, λίγο προς τα πάνω ή προς τα κάτω, έβγαζε πάντα το ίδιο Classification Report. Ενώ για μικρές τιμές C,gamma τα αποτελέσματα αλλάζανε κανονικά. Για τον SVM έγινε grid search για το decision tree δεν έγινε.

Πίνακας 6 Classification Report για BoAW με Kmeans , SVM

	Precision	Recall	F1-score
0	0.67	0.79	0.72
1	0.60	0.47	0.53
2	0.56	0.36	0.44
3	0.34	0.40	0.37
Accuracy			0.62
Macro avg	0.54	0.50	0.52
Weighted avg	0.61	0.62	0.61



(a)



(b)

Εικόνα 13 Confusion Matrix and ROC for BoAW approach with kmeans, SVM

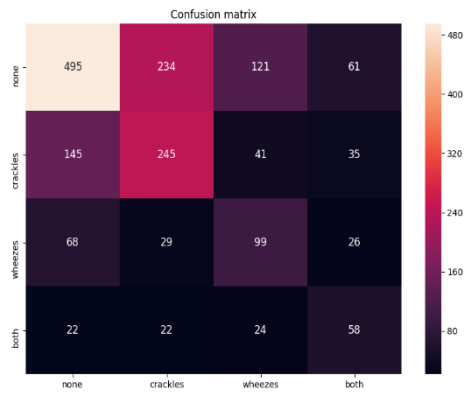
Random Forest

Εδώ δεν ψάξαμε υπερ παραμέτρους από την στιγμή που και ο SVM δεν έδινε καλά αποτελέσματα. Αλλάζοντας μόνοι μας τις παραμέτρους βρήκαμε τα παρακάτω αποτελέσματα.

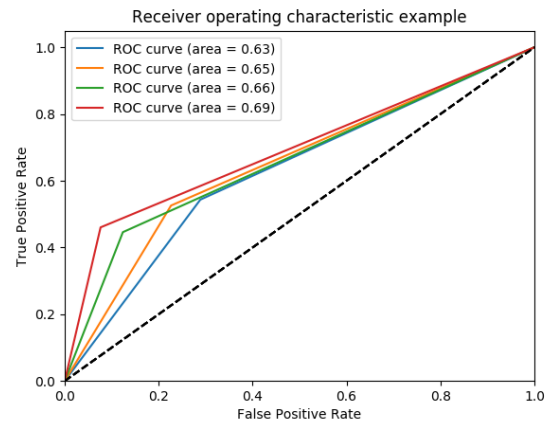
Πίνακας 7 Classification Report για BoAW με Kmeans , Random Forest

	Precision	Recall	F1-score
0	0.68	0.54	0.60
1	0.46	0.53	0.49
2	0.35	0.45	0.39
3	0.32	0.46	0.38

Accuracy			0.52
Macro avg	0.45	0.49	0.47
Weighted avg	0.55	0.52	0.53



(a)



(b)

Εικόνα 14 Confusion Matrix and ROC for BoAW approach with kmeans, Random Forest

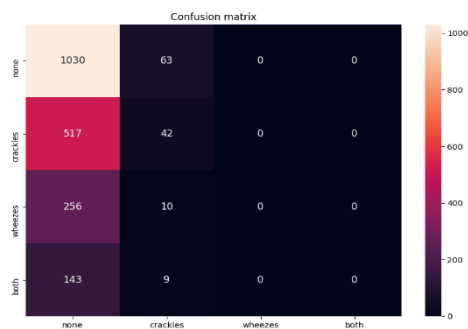
HDBSCAN

Το clustering με τον αλγόριθμο του hdbscan ήταν αποτυχία απλά για να υπάρχουν τα αποτελέσματα τα παραθέτουμε εδώ. Να πούμε ότι για την εκμάθηση στον SVM και Random Forest έχουμε και εδώ class_weights balanced. Ωστόσο τα αποτελέσματα δεν είναι balanced που ίσως σημαίνει ότι δεν τα διαχώρισε καλά τα κέντρα ο dbscan. Από τα παρακάτω αποτελέσματα τόσο ο SVM όσο ο Random Forest έχουν ακριβώς τα ίδια αποτελέσματα. Θεωρούμε ότι ο διαχωρισμός του clustering έχει αποτύχει σε αυτήν την περίπτωση.

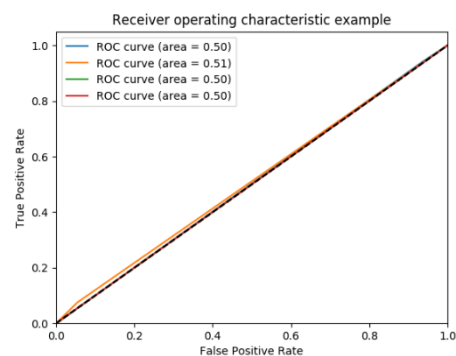
SVM ($C=0.25$, $\gamma=0.0078125$, $\text{kernel}= \text{rbf}$) και Random Forest

Πίνακας 8 Classification Report για BoAW με HDBSCAN, SVM & Random Forest

	Precision	Recall	F1-score
0	0.53	0.94	0.68
1	0.34	0.08	0.12
2	0	0	0
3	0	0	0
Accuracy			0.52
Macro avg	0.22	0.25	0.20
Weighted avg	0.37	0.52	0.39



(a)



(b)

Εικόνα 15 Confusion Matrix and ROC for BoAW approach with HDBSCAN, SVM & Random Forest