

Μηχανική μάθηση

Ιατρική Διάγνωση, ομάδα Α3

Κυριακή, 8 Ιουνίου 2025

Ομάδα: Κατωπόδης Οδυσσέας, Α.Μ.: 7115112400019

Γκούμας Βασίλειος, Α.Μ.: 7115132400008

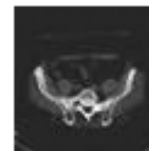
Χελβατζή Δήμητρα, Α.Μ.: 7115132400007

Εισαγωγή και περιγραφή του προβλήματος

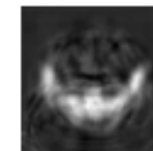
1. Classification στο dataset [Medical MNIST](#)
(διαχωρισμός εικόνων σε 6 κλάσεις)
2. Classification στο dataset [BreastMNIST 224x224](#)
(διαχωρισμός εικόνων σε έχουν/δεν έχουν καρκίνο)

(Dataset [CT Medical Images](#) χωρίς labels, μόνο προεπεξεργασία και εξερεύνηση)

Original: AbdomenCT



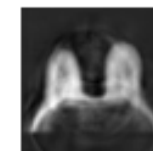
Reconstructed: AbdomenCT



Original: BreastMRI



Reconstructed: BreastMRI



Original: Hand



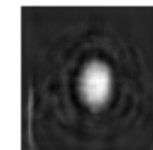
Reconstructed: Hand



Original: HeadCT



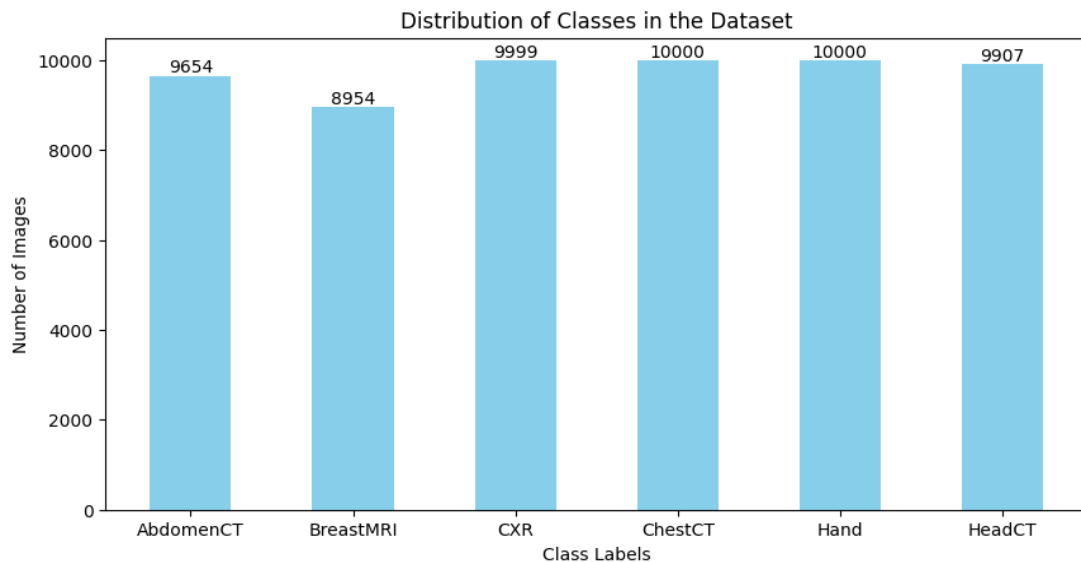
Reconstructed: HeadCT



Προεπεξεργασία και Εξερεύνηση Δεδομένων

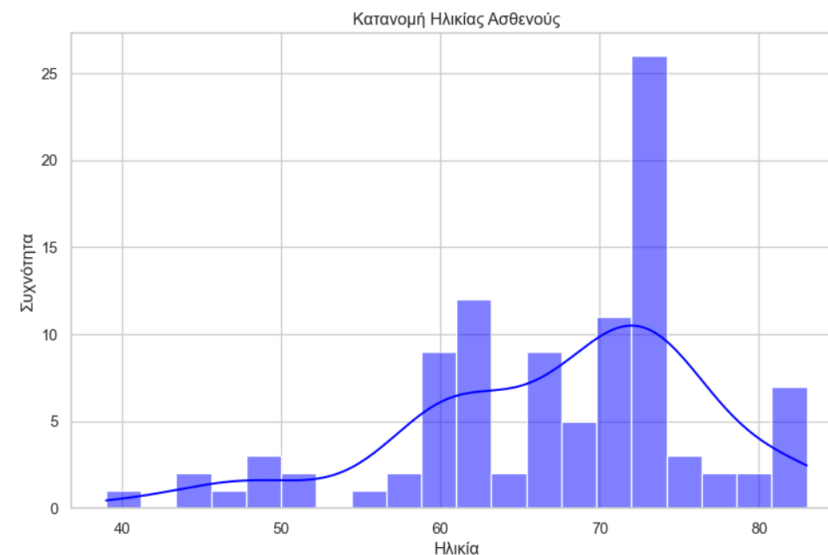
Medical MNIST

- 58954 εικόνες κατανεμημένες σε φακέλους (κλάσεις)
- Όλες οι εικόνες είναι 64x64 pixels και δεν χρειάστηκε να γίνει resize σε καμία, ασπρόμαυρες,
- Βρέθηκαν συνολικά 440 διπλότυπα, 346 στην κλάση 'AbdomenCT', 1 στην κλάση 'CXR', και 93 στην κλάση 'HeadCT'. Ευτυχώς το μέγεθος των μοναδικών εικόνων ανά κλάση δεν επηρεάστηκε πολύ και κρίνεται επαρκές για να προχωρήσουμε δίχως να χρειάζεται άλλο χειρισμό από εμάς.



CT Medical Images (παλιό dataset Δήμητρας)

- Πραγματοποιήθηκε EDA στο αρχείο eda_cleaned.csv. Το dataset full_archive.npz δεν αξιοποιήθηκε λόγω έλλειψης labels. Ενδεικτικά χρησιμοποιήθηκαν 100 εικόνες από τις 475 με labels και μεταδεδομένα. Έγινε προσπάθεια για EDA, ενδεικτικά βρέθηκε ότι η μέση ηλικία ασθενών: 67.48 έτη, τιμές pixel: μεγάλη διακύμανση → ανάγκη κανονικοποίησης, (ωστόσο θα ήταν χρήσιμο να πραγματοποιηθεί εάν αξιοποιούνταν τελικά για την εργασία), ContrastTag: 2 ελλείψεις → αντικαταστάθηκαν με "NONE".
- Για το λόγο αυτό η Δήμητρα με τον Οδυσσέα μοιράστηκαν το Medical MNIST dataset (Οδυσσέας κλάσεις :AbdomenCT, BreastMRI, Hand, and HeadCT, Δήμητρα κλάσεις: BreastMRI, ChestCT, CXR, and Hand)



Logistic regression

Κλάσεις : AbdomenCT, BreastMRI, Hand, and HeadCT

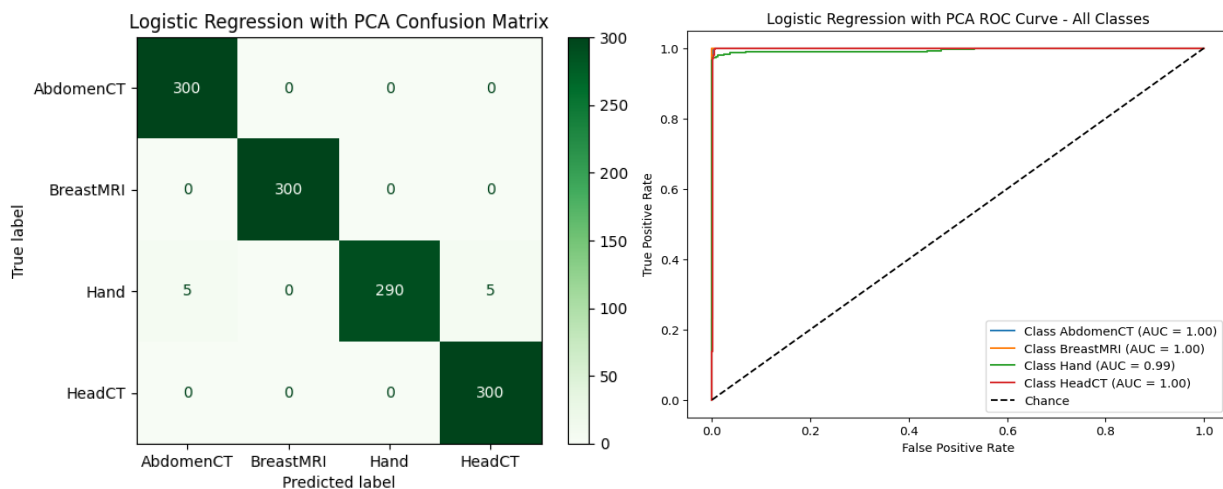
- Χωρίς καμία μείωση των διαστάσεων το μοντέλο αποδίδει πολύ καλά με τιμές τις ακρίβειας ταξινόμησης $>0,99$. Συνολικά 10 λάθη γίνονται και όλα παρατηρούνται στην κλάση “hand”

PCA

- Η ακρίβεια παραμένει πρακτικά ίδια, αλλά μειώνεται σίγουρα το υπολογιστικό κόστος οπότε είναι δόκιμη χρήση της.
- Οι μέθοδοι βελτιστοποίησης των παραμέτρων δεν δίνουν κάποια αξιόλογη αύξηση ήδη τα νούμερα της ακρίβειας προσεγγίζουν το 1.

LDA

- Αποτελέσματα με ακρίβεια της τάξης του 0,9175
- Βελτιστοποίηση των υπερ παραμέτρων η ακρίβεια φτάνει το 0,94 όμως το αυξημένο υπολογιστικό κόστος δεν δικαιολογείται.



BreastMRI, ChestCT, CXR, Hand

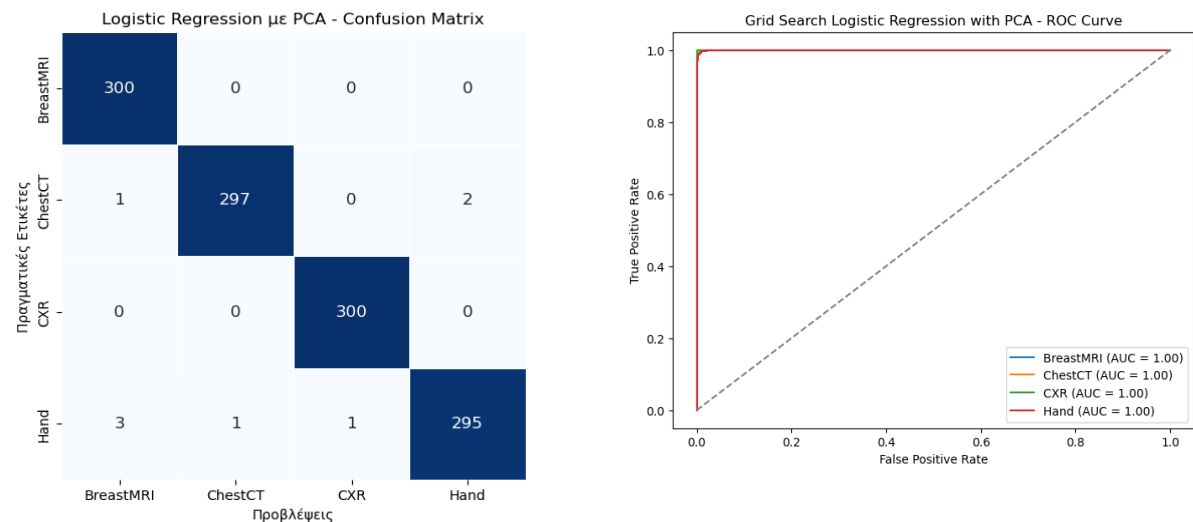
- Χωρίς μείωση διαστάσεων: ακρίβεια 99.25%

PCA

- Πέτυχε ακρίβεια 99.33%, ελαφρώς βελτιωμένη από το baseline χωρίς μείωση.
- Τα σφάλματα ήταν ελάχιστα, κυρίως σε ChestCT και Hand.
- Ο ROC-AUC ήταν 1.0 σε όλες τις κατηγορίες, δείχνοντας τέλεια διάκριση.
- Η βελτιστοποίηση με GridSearchCV δεν άλλαξε την ακρίβεια, επιβεβαιώνοντας τη σταθερότητα του μοντέλου.

LDA

- Η ακρίβεια μειώθηκε σε 97.17%, με ελαφρά βελτίωση σε 97.5% μετά από GridSearchCV.
- Παρατηρήθηκαν περισσότερα λάθη σε ChestCT, CXR, Hand.
- Παρόλο που οι διαστάσεις μειώθηκαν σε 3, η οπτική διάκριση ήταν καλή.
- Η συνολική απόδοση ήταν καλή, αλλά υστερούσε ελαφρώς έναντι του PCA.



K nearest neighbor

Κλάσεις : AbdomenCT, BreastMRI, Hand, and HeadCT

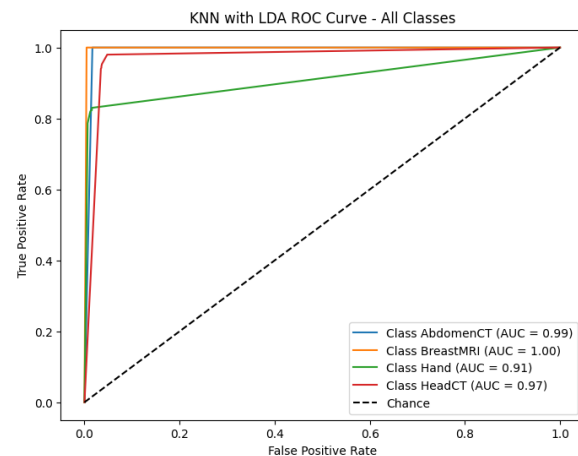
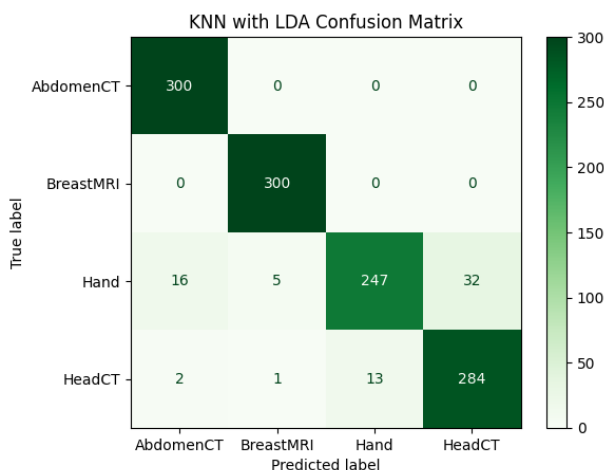
• Χωρίς καμία μείωση των διαστάσεων αποτελέσματα με ακρίβεια $> 0,99$ και συνολικά 11 λάθη στο σύνολο των 1200 εικόνων του τεστ σετ.

PCA

- Αύξηση της ακρίβειας στο 0,993 και σημαντική μείωση στο υπολογιστικό φορτίο.
- Με αναζήτηση βέλτιστων παραμέτρων μέσω της Grid search η ακρίβεια αυξάνεται στο 0,996 αλλά μαζί της αυξάνονται και οι απαραίτητοι υπολογισμοί. Λόγω της μικρής μόνο αύξησης της ακρίβειας δεν έχει ιδιαίτερο νόημα η χρήση του.

LDA

- Αποτελέσματα με ακρίβεια της τάξης του 0,9425
- Οι κλάσεις που το δυσκολεύουν περισσότερο είναι αυτές του χεριού και του κεφαλιού, πιθανώς λοιπόν οι τιμές των πίξελ σε αυτές τις εικόνες να μην ακολουθούν τόσο κανονική κατανομή και να μην είναι τόσο κατάλληλες για τη μέθοδο LDA.
- Η αναζήτηση για βέλτιστους υπερ παραμέτρους δεν ευδοκίμησε καθώς η ακρίβεια παρέμεινε πρακτικά σταθερή ($=0,9392$)



BreastMRI, ChestCT, CXR, Hand

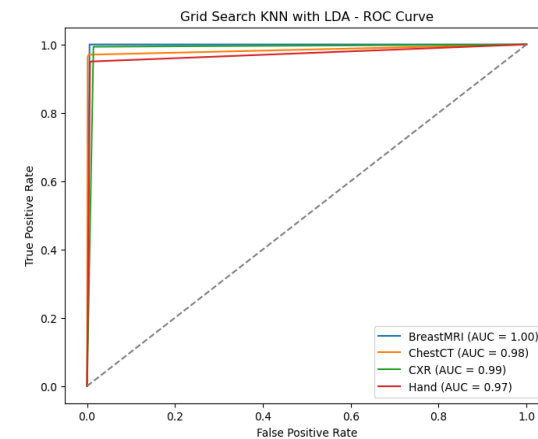
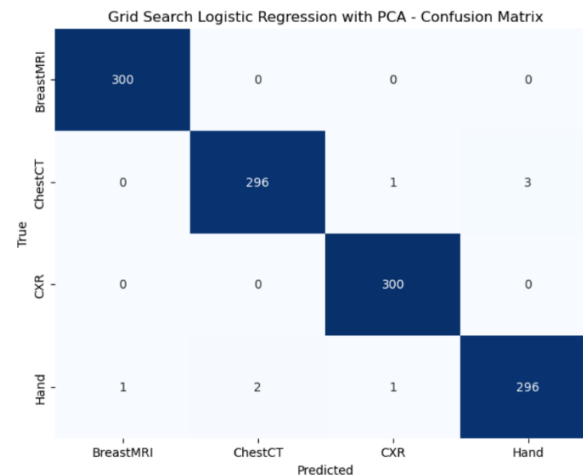
- Χωρίς μείωση διαστάσεων: ακρίβεια 99.25%

PCA

- Η ακρίβεια ανέβηκε από 99.25% σε 99.42%, και έφτασε 99.67% με GridSearchCV.
- Παρατηρήθηκαν λίγα σφάλματα, κυρίως σε ChestCT και Hand.
- Ο αλγόριθμος με PCA απέδωσε εξαιρετικά μετά το tuning.

LDA

- Η ακρίβεια ήταν 97.67%, οριακά βελτιωμένη σε 97.75% μετά από GridSearchCV.
- Λάθη εντοπίστηκαν σε ChestCT, CXR, Hand.
- Η ROC-AUC ήταν χαμηλότερη για την κατηγορία Hand (0.97).
- Συνολικά, η απόδοση με LDA ήταν υποδεέστερη από αυτή του PCA.



Support vector machine

Κλάσεις : AbdomenCT, BreastMRI, Hand, and HeadCT

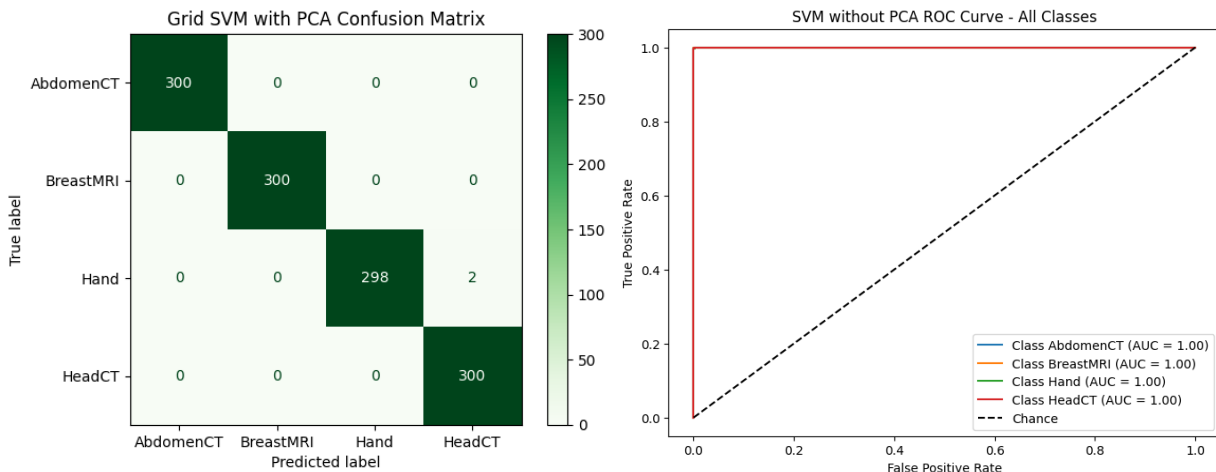
- Χωρίς καμία μείωση των διαστάσεων, το συγκεκριμένο μοντέλο αποδείχθηκε εξαιρετικό και ταχύτατο. Η πρώτη εκτέλεση δίχως καμία μείωση των διαστάσεων κατάφερε ακρίβεια ίση με 0,9975!

PCA

- Η ακρίβεια παρέμεινε ολόιδια (0,9975) και ταυτόχρονα μειώθηκε σημαντικά και ο υπολογιστικός φόρτος. Φυσικά και αξίζει η χρήση της.
- Βελτιστοποίηση των υπέρ παραμέτρων με τη μέθοδο Grid Search έφερε μια αύξηση στην ακρίβεια η οποία πλέον είναι ίση με 0,9983 αλλά φυσικά αύξησε το υπολογιστικό φορτίο με αποτέλεσμα να μην αξίζει η χρήση της.

LDA

- Το μοντέλο πετυχαίνει ακρίβεια 0,9425 ελάχιστα μικρότερη από αυτήν με το μοντέλο PCA. Πάλι και εδώ τη μεγαλύτερη δυσκολία την αντιμετωπίζει στις κλάσεις Hand και HeadCT.
- Με βελτιστοποίηση των παραμέτρων η ακρίβεια αυξάνεται ελάχιστα στο 0,9441 πράγμα που σημαίνει ότι δεν αξίζει ιδιαίτερα και η χρήση της πιστοποίησης.



BreastMRI, ChestCT, CXR, Hand

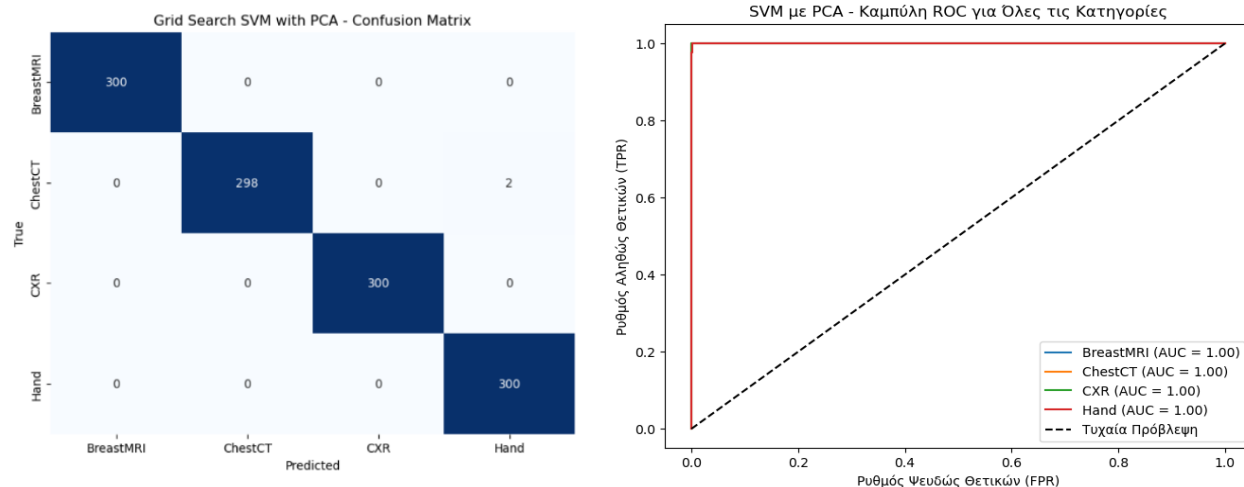
- Χωρίς μείωση διαστάσεων: ακρίβεια 99.75%

PCA

- Πέτυχε ακρίβεια 99.83%, ελαφρώς καλύτερη από χωρίς μείωση.
- Τα λάθη περιορίστηκαν σχεδόν μόνο στην κατηγορία ChestCT.
- Η χρήση GridSearchCV επιβεβαίωσε τη βέλτιστη απόδοση με τον πυρήνα RBF.

LDA

- Η ακρίβεια μειώθηκε σε 97.58%, με μικρή βελτίωση σε 97.67% μετά από tuning.
- Παρουσιάστηκαν λάθη στις ίδιες ευαίσθητες κατηγορίες (ChestCT, CXR, Hand).
- Η LDA πρόσφερε καλή οπτική διάκριση αλλά με μικρό κόστος στην ακρίβεια.
- Το μοντέλο ήταν λιγότερο σταθερό σε σύγκριση με το PCA.



Naive bayes

Κλάσεις : AbdomenCT, BreastMRI, Hand, and HeadCT

- Χωρίς καμία μείωση των διαστάσεων, το μοντέλο επιτυγχάνει αξιοσημείωτη ακρίβεια 0,9975. Αυτό πρακτικά μεταφράζεται σε 3 μόνο λάθη από τις 1200 φωτογραφίες του test set όλα στην κλάση HeadCT.

PCA

- Η ακρίβεια πέφτει στο 0,985, μικρή μείωση πράγμα που σημαίνει ότι δικαιολογείται η χρήση της μεθόδου καθώς συνοδεύεται με μείωση του υπολογιστικού φορτίου
- Η προσπάθεια βελτιστοποίησης των παραμέτρων δεν πετυχαίνει αύξηση της ακρίβειας = 0,985 και έτσι δεν δικαιολογεί την χρησιμοποίηση της.

LDA

- Η ακρίβεια πέφτει σημαντικά = 0,8583. Η κλάση Hand είναι αυτή που επηρεάζεται περισσότερο αρνητικά και περιέχει τις περισσότερες λάθος εκτιμήσεις από το μοντέλο.
- Βελτιστοποίηση παραμέτρων το μοντέλο που χρησιμοποιεί LDA βελτιώνεται σημαντικά με την ακρίβεια να ανεβαίνει στο 0,9425 και τα λάθη να μειώνονται σημαντικά.

BreastMRI, ChestCT, CXR, Hand

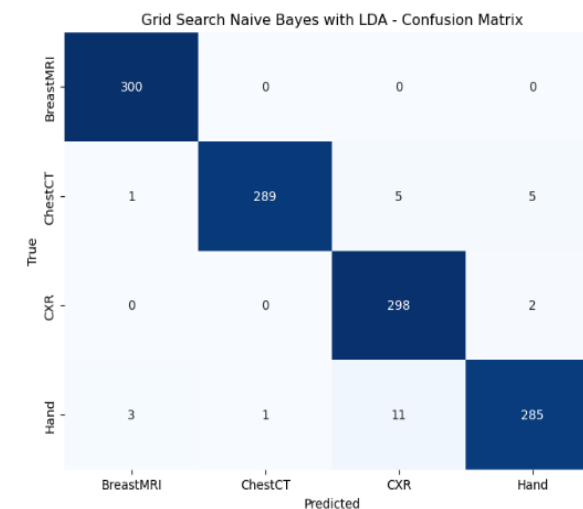
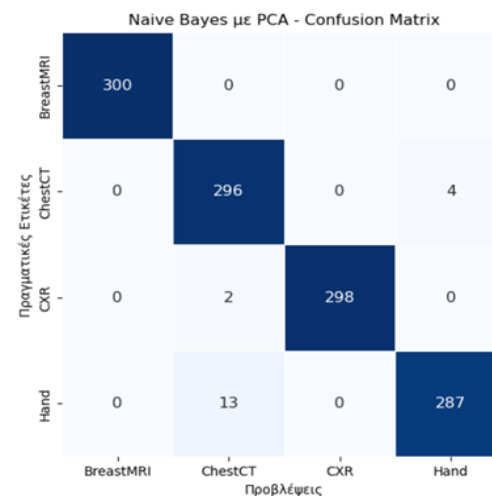
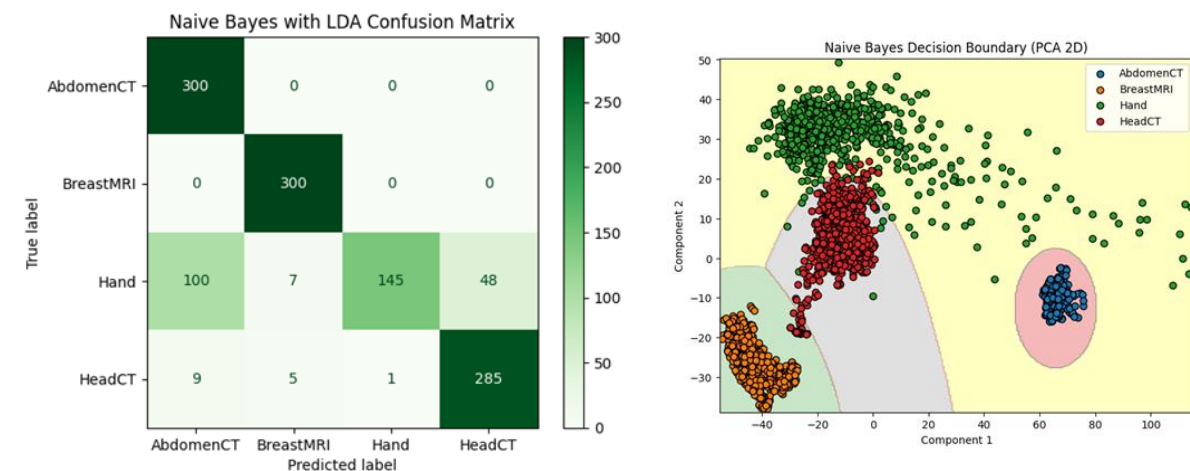
- Χωρίς μείωση διαστάσεων: Ακρίβεια 96.83%

PCA

- Η ακρίβεια αυξήθηκε από 96.83% σε 98.42%.
- Το confusion matrix έδειξε λιγότερα σφάλματα από ό,τι χωρίς μείωση.
- Η βελτιστοποίηση (var_smoothing) δεν άλλαξε την απόδοση, σταθερή στο 98.4%.

LDA

- Η ακρίβεια έφτασε 98.17%, και έπεσε ελαφρώς σε 97.67% με tuning.
- Παρατηρήθηκαν πιο συχνά λάθη σε ChestCT, CXR, Hand.
- Αν και προσέφερε καλή οπτική διάκριση, η συνολική απόδοση υστέρησε του PCA.



Multi-Layer Perceptron neural network

Κλάσεις : AbdomenCT, BreastMRI, Hand, and HeadCT

- Χωρίς καμία μείωση των διαστάσεων, το μοντέλο επιτυγχάνει ακρίβεια 0.993 για το υπολογιστικό κόστος τουλάχιστον σε χρόνο φάνηκε να μην διαφέρει ιδιαίτερα από τις άλλες μεθόδους.
- Η προσπάθεια βελτιστοποίησης των παραμέτρων όμως αύξησε **κατακόρυφα** το υπολογιστικό κόστος το οποίο χρονικά έφτασε τα 60 λεπτά περίπου με την ακρίβεια να αγγίζει το 0,994, ελάχιστη λοιπόν η διαφορά της και σίγουρα δεν ναι δικαιολογεί την χρήση της.

PCA

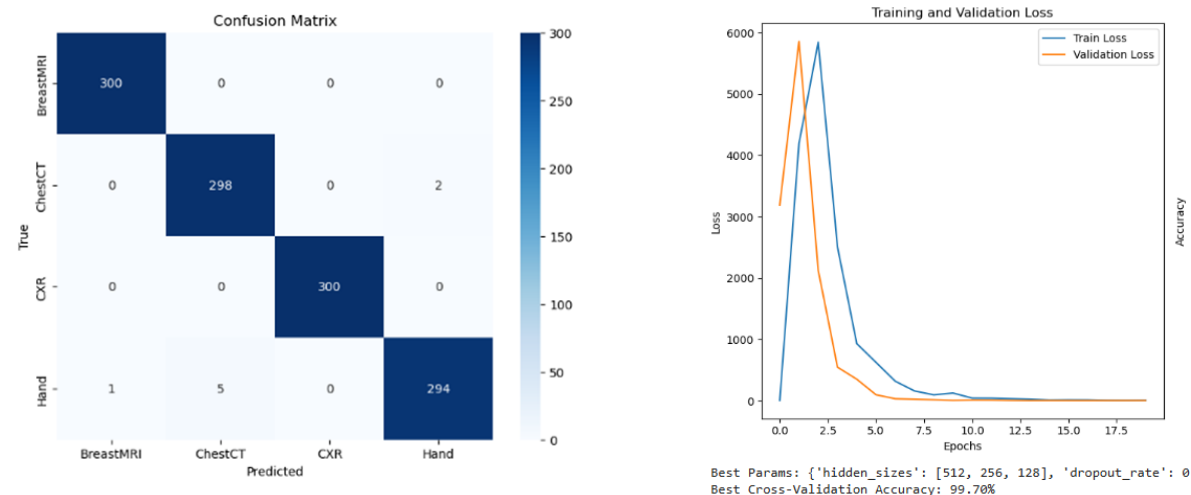
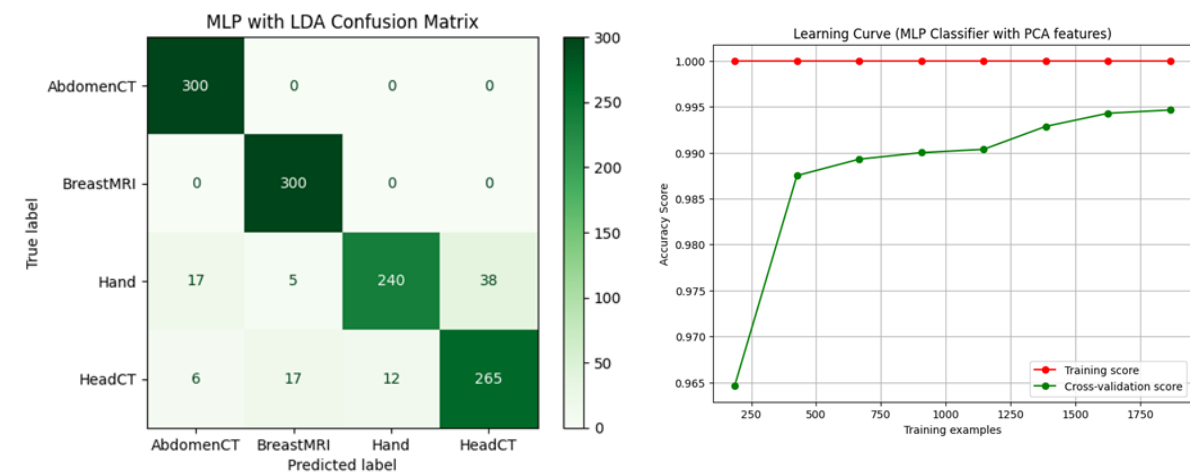
- Πετυχαίνει ακρίβεια 0,994 μετά από 34 iterations με $\text{loss} = 0.0009947$, ακόμα όμως και από τη 3η iteration, το $\text{loss} = 0.08$. Επομένως πολύ γρήγορα φτάνει σε πολύ καλά επίπεδα. Και ταυτόχρονα δεν είναι ιδιαίτερα υπολογιστικά «βαρύ».
- Προσπάθεια βελτιστοποίησης υπέρ παραμέτρων δεν αλλάζει την τιμή της ακρίβειας, υπάρχει μεγάλη διαφορά όμως στο υπολογιστικό κόστος (~60 min)

LDA

- επηρεάζει ελαφρώς αρνητικά την ακρίβεια του μοντέλου = 0.92.
- Βελτιστοποίηση παραμέτρων αυξάνει ελαφρώς την ακρίβεια στην τιμή 0,93, με χρόνο εκτέλεσης (~10min)

BreastMRI, ChestCT, CXR, Hand

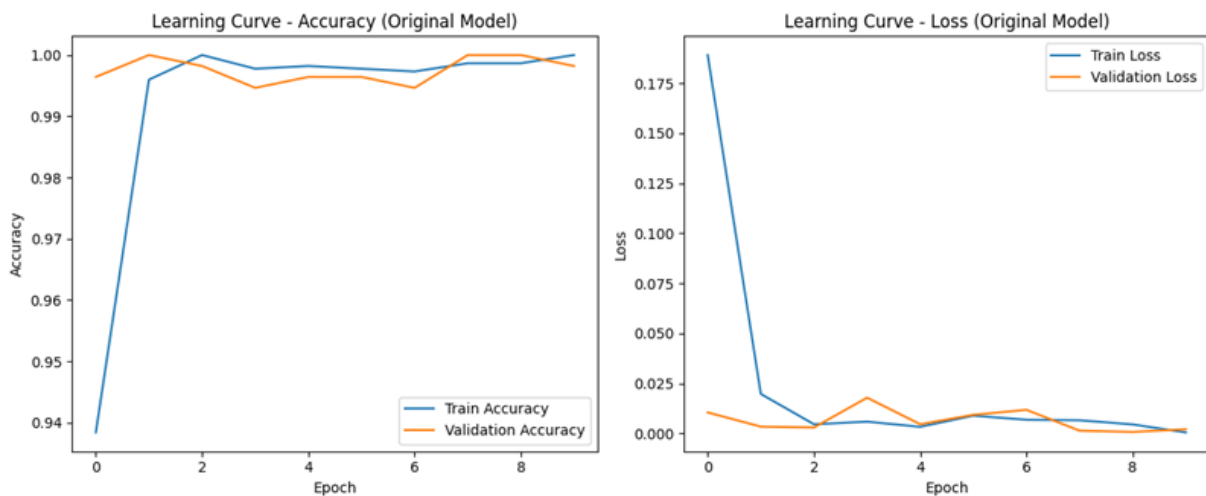
- Τα δεδομένα εκπαίδευσης και δοκιμής μετατράπηκαν σε PyTorch tensors χωρίς να γίνει μείωση διαστάσεων.
- Αξιοποιήθηκαν για Baseline μοντέλο δύο κρυφά επίπεδα των 256 νευρώνων, με ReLU και dropout 0.3, ως ανάγκη για μια σχετικά απλή αλλά αποδοτική λύση για την κατηγοριοποίηση των εικόνων. Η εκπαίδευση πραγματοποιήθηκε για 20 εποχές
- Ακρίβεια στο test set: 99.17% . Πολύ ικανοποιητική για απλό MLP, με ελάχιστα σφάλματα (κυρίως σε ChestCT και Hand).
- Χρήση GridSearch & Stratified Cross-Validation (5 folds): Εξασφαλίστηκε σταθερότητα και γενίκευση σε διαφορετικά subsets.
- Πειραματιστήκαμε με κρυφά επίπεδα με διαφορετικά μεγέθη νευρώνων, ρυθμούς dropout, batch normalization και διαφορετικές συναρτήσεις ενεργοποίησης (ReLU, Sigmoid, Tanh), διαφορετικά μεγέθη κρυφών επιπέδων, συναρτήσεις ενεργοποίησης (ReLU, Sigmoid, Tanh), dropout, batch norm, learning rates.
- Βρέθηκε για βέλτιστο μοντέλο: κρυφά επίπεδα: [512, 256, 128], χωρίς dropout/batch norm, learning rate 0.001, 20 εποχές. Και ακρίβεια CV: 99.70%, test accuracy : 99.33%.



Convolutional neural network

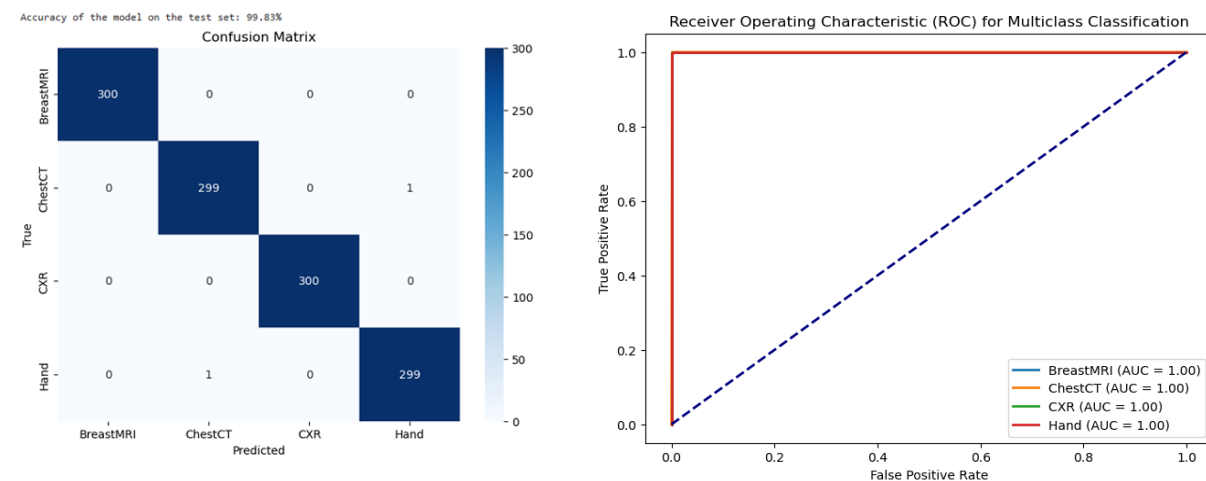
Κλάσεις : AbdomenCT, BreastMRI, Hand, and HeadCT

- Επιλέχθηκε η βιβλιοθήκη tensorflow για την δημιουργία του CNN, παρότι έγιναν δοκιμές και με την Pytorch, η επιλογή έγινε από προσωπική προτίμηση.
- Χρησιμοποιήθηκαν 10 εποχές και από την 3η κιόλας η ακρίβεια είχε γίνει πρακτικά≈1. Ταυτόχρονα και ο χρόνος εκτελέσεως ήταν ικανοποιητικός.
- Η αρχιτεκτονική του μοντέλου βασίστηκε σε ένα βασικό πρότυπο από το βιβλίο “Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow”.
- Χρησιμοποιήθηκε και το keras tuner για τη βελτιστοποίηση των υπερ παραμέτρων του μοντέλου όμως όπως είδαμε το μοντέλο ήδη ήταν σε άριστο επίπεδο και έτσι η βελτιστοποίηση δεν απέφερε κάποια αύξηση στην ακρίβεια παρά μόνο αύξησε τον όγκο υπολογισμών. Εξ αυτού η χρήση της δεν έχει κάποιο νόημα.
- Για εκπαιδευτικούς σκοπούς έγινε και μια δοκιμή διαφορετικών configurations.



BreastMRI, ChestCT, CXR, Hand

- Τα δεδομένα χρησιμοποιήθηκαν χωρίς να γίνει μείωση διαστάσεων.
- Για Baseline μοντέλο, αξιοποιήθηκαν 3 συνελκτικά επίπεδα με φίλτρα (32, 64, 128), ReLU + MaxPooling + Dropout 0.5. Η επιλογή της συγκεκριμένης αρχιτεκτονικής βασίστηκε στην ικανότητά της να μαθαίνει σταδιακά όλο και πιο σύνθετα χαρακτηριστικά των εικόνων. Ακολουθούν 2 πλήρως συνδεδεμένα επίπεδα για την τελική ταξινόμηση.
- Εκπαίδευση: 20 εποχές με Adam optimizer και Cross-Entropy Loss.
- Μεγαλύτερη από 99% ακρίβεια εκπαίδευσης από νωρίς (ήδη από τη 2η εποχή).
- Ακρίβεια στο test set: 99.67%, με χαμηλό loss (0.0126), με σφάλματα μόνο σε 1 δείγμα ChestCT και 1 Hand.
- ROC-AUC: 1.0 σχεδόν παντού (0.99 μόνο στο Hand), άριστη ικανότητα διάκρισης κατηγοριών.
- Grid Search + Stratified Cross-Validation (5 folds): Εξετάστηκαν αριθμός επιπέδων, νευρώνες, dropout, learning rate, συναρτήσεις ενεργοποίησης.
- Καλύτερες υπερπαραμέτροι: Learning_rate: 0.001, dropout_rate: 0.3, layers: 2, neurons/layer: 64, activation: ReLU.
- Τελική ακρίβεια test set: 99.83% Με precision & recall = 1.00 για όλες τις κατηγορίες.



Συγκριτική Ανάλυση για όλα τα μοντέλα

Κλάσεις : AbdomenCT, BreastMRI, Hand, and HeadCT

- Κατά γενική ομολογία η μέθοδος μείωσης διαστάσεων PCA έκανε τα μοντέλα να εκτελούνται σε σημαντικά μικρότερο χρόνο και είτε επηρέαζε θετικά είτε δεν επηρέαζε την ακρίβεια, πράγμα που στην πραγματικότητα μάλλον την κάνει απαραίτητη για οποιαδήποτε εφαρμογή.
- Ταυτόχρονα και η μέθοδος LDA έδωσε πολύ καλά αποτελέσματα και η μείωση του υπολογιστικού φορτίου ήταν αισθητή. Η κλάσεις Hand και HeadCT είναι αυτές που την «δυσκόλεψαν» περισσότερο.
- Το μοντέλο SVM είχε πολύ εντυπωσιακά αποτελέσματα ακόμα και χωρίς κάποια βελτιστοποίηση και με ταχύτατους χρόνους εκτέλεσης. Έδωσε την αίσθηση ότι ήταν το μόνο σε παρόμοιο επίπεδο με τα νευρωνικά δίκτυα.
- Τα νευρωνικά δίκτυα MLP και CNN, πέτυχαν εντυπωσιακή ακρίβεια ακόμα και στο basic configuration τους. Το υπολογιστικό φορτίο δεν ήταν ιδιαίτερα μεγαλύτερο από άλλες μεθόδους στην εκδοχή δίχως μείωση διαστάσεων.
- Η βελτιστοποίηση παραμέτρων για το νευρωνικά δίκτυα MLP δίχως μείωση διαστάσεων επέφερε πολύ μεγάλο υπολογιστικό κόστος!
- Η βελτιστοποίηση παραμέτρων για το νευρωνικά δίκτυα CNN δίχως μείωση διαστάσεων επέφερε και αυτή υπολογιστικό κόστος, αλλά όχι με τόση μεγάλη διαφορά όπως η αντίστοιχη για τα MLP.

Accuracy	Χωρίς μείωση διαστάσεων	PCA	Grid search PCA	LDA	Grid search LDA
Logisitic regression	0.992	0.992	0.991	0.918	0.936
SVM	0.998	0.998	0.998	0.943	0.944
KNN	0.991	0.993	0.995	0.943	0.940
Naïve Bayes	0.998	0.985	0.985	0.858	0.943
MLP (10 iter.)	0.993	0.994	0.994	0.921	0.930
CNN	1	-	-	-	-

BreastMRI, ChestCT, CXR, Hand

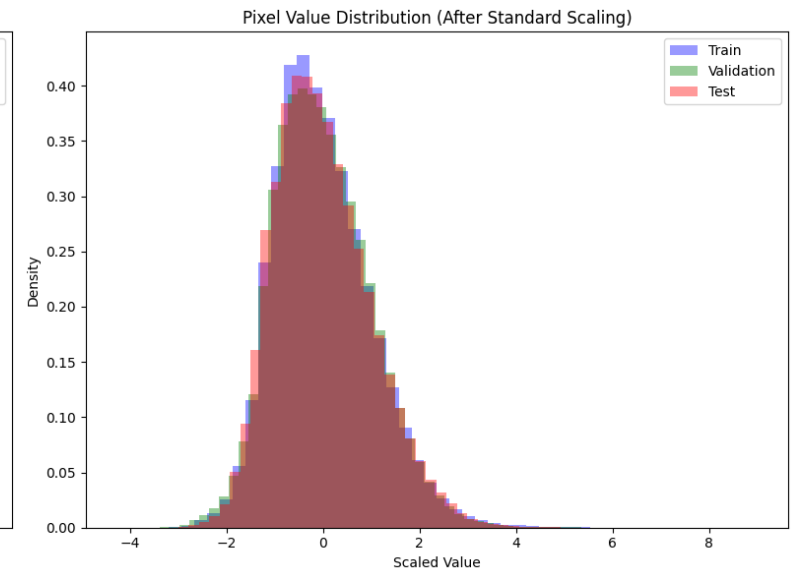
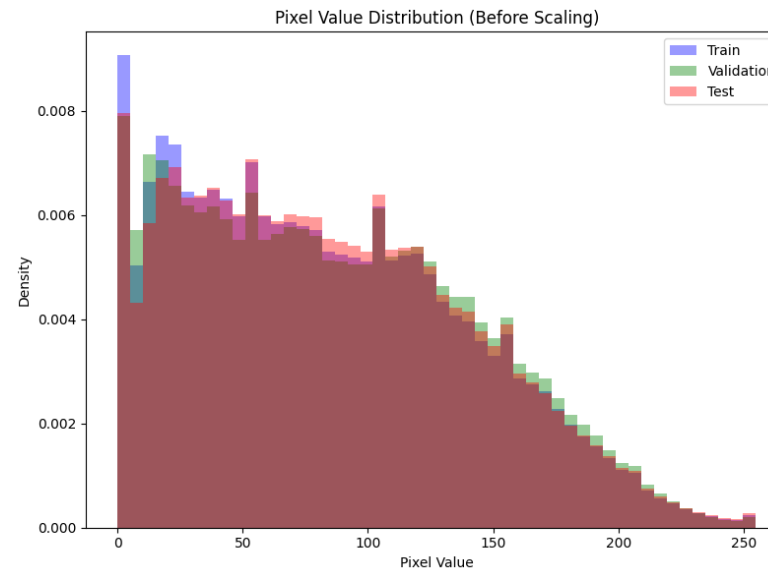
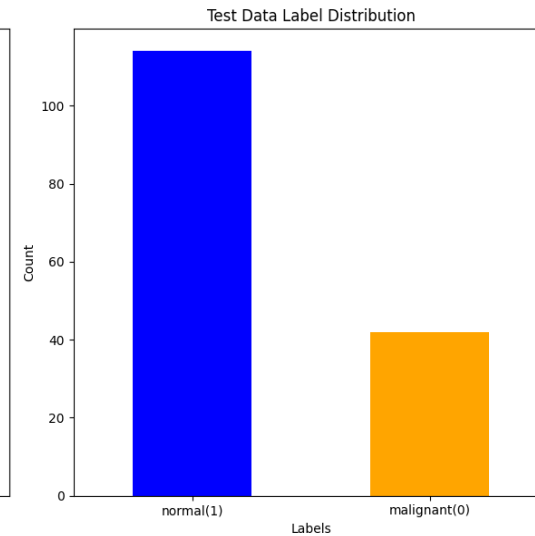
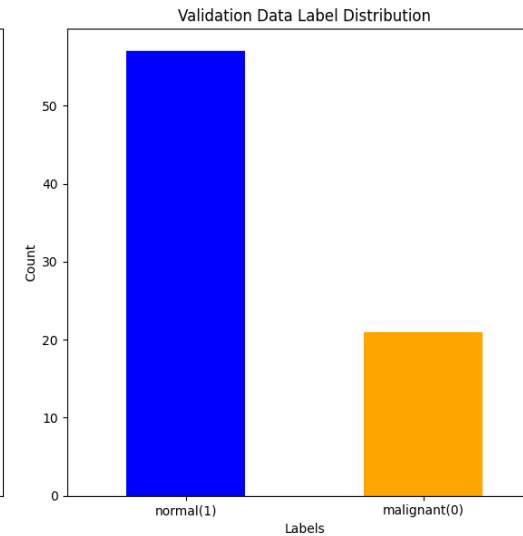
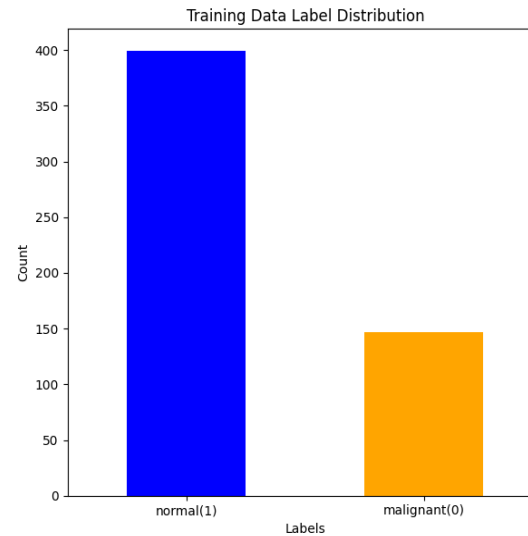
- Όλα τα μοντέλα πέτυχαν υψηλή ακρίβεια, με καλύτερες επιδόσεις τα SVM, CNN και MLP.
- Η εφαρμογή PCA βελτίωσε την ακρίβεια σε όλα σχεδόν τα μοντέλα και μείωσε την πολυπλοκότητα.
- Το CNN είναι το πιο αποδοτικό για εικόνες, αλλά έχει υψηλό υπολογιστικό κόστος.
- Απλά μοντέλα (Logistic, KNN) είχαν καλή απόδοση και είναι κατάλληλα για ταχύτερη ταξινόμηση.
- Η πολυπλοκότητα βελτιώνει την απόδοση, αλλά χρειάζεται σωστή ρύθμιση για αποφυγή υπερπροσαρμογής.
- Η σωστή επιλογή υπερπαραμέτρων και η Cross-Validation οδηγούν σε πιο σταθερά και γενικεύσιμα μοντέλα.
- Η μείωση διαστάσεων βοηθά στη σταθερότητα και απόδοση, χωρίς σημαντική απώλεια πληροφορίας.

Accuracy	Χωρίς μείωση διαστάσεων	PCA	Grid Search PCA	LDA	Grid Search LDA
Logistic Regression	99.25%.	99.33%	99.33%	97.17%	97.5%.
Support Vector Machines	99.75%	99.83%	99.83%	97.58%	97.67%.
k-Κοντινότεροι Γείτονες (k-NN)	99.25%	99.42%	99.67%	97.67%	97.75%
Naive Bayes	96.83%	98.42%	98.4%	98.17%	97.67%
MLP / MLP (GridSearch)	99.17% / 99.33%	-	-	-	-
CNN / CNN (GridSearch)	99.67% / 99.83%	-	-	-	-

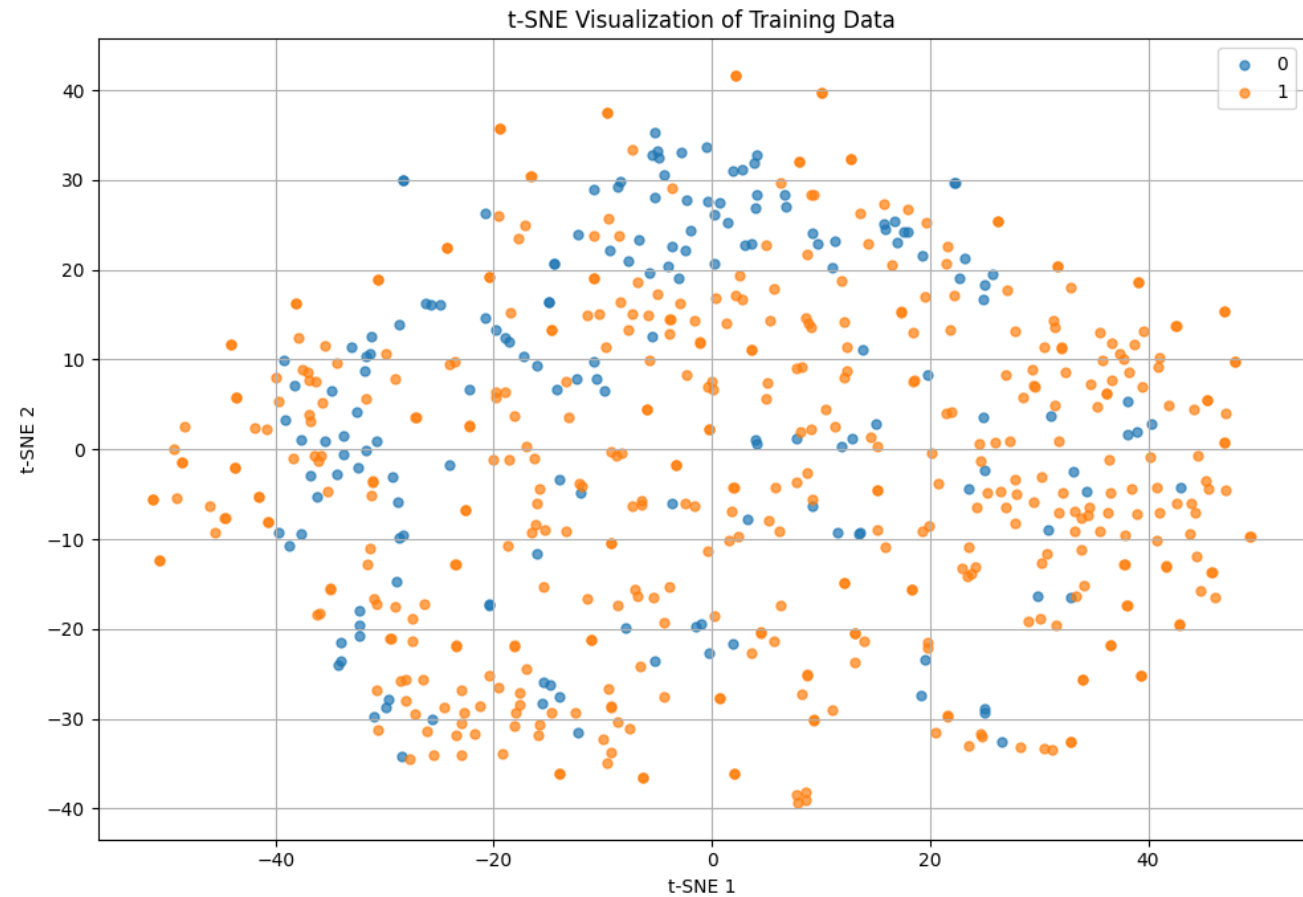
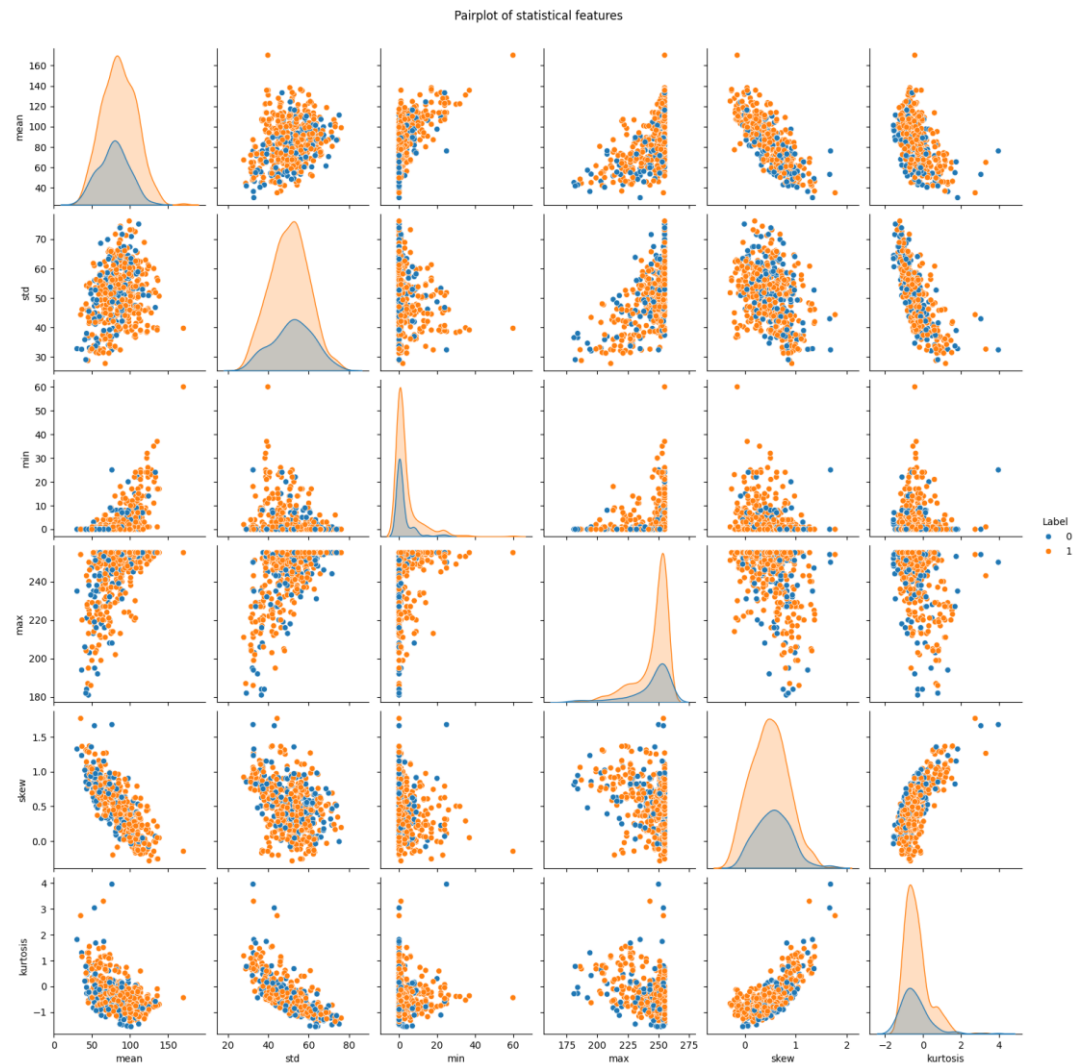
BreastMNIST 224x224: Προεπεξεργασία και Εξερεύνηση Δεδομένων

BreastMNIST 224x224

- Το dataset αποτελείται από 780 εικόνες οι οποίες χωρίζονται σε 3 σετ (train, validation, test) των 546 (70.00%), 78 (10.00%) και 156 (20.00%) αντιστοίχως. Μεταξύ τους ήταν labeled με σταθερό αλλά μη ισορροπημένο τρόπο, 73.8% του κάθε σετ να αντιστοιχεί σε υγιείς (normal/benign) εικόνες και το άλλο 26.92% σε εικόνες καρκινοπαθών (malignant) σε κάθε υποσύνολο
- Δεν υπήρχαν ελλείπουσες ή corrupted εικόνες· έγινε έλεγχος φωτεινότητας & μέσου όρου pixel.
- Εντοπίστηκε μία outlier εικόνα, διατηρήθηκε λόγω μικρού μεγέθους dataset και μη φθοράς
- Εφαρμόστηκε feature engineering: στατιστικά, υφή (LBP), σχήμα (area, perimeter, compactness).
- Έγινε κανονικοποίηση, αφαίρεση διακύμανσης και αποφεύχθηκε εφαρμογή πολυωνυμικών χαρακτηριστικών.

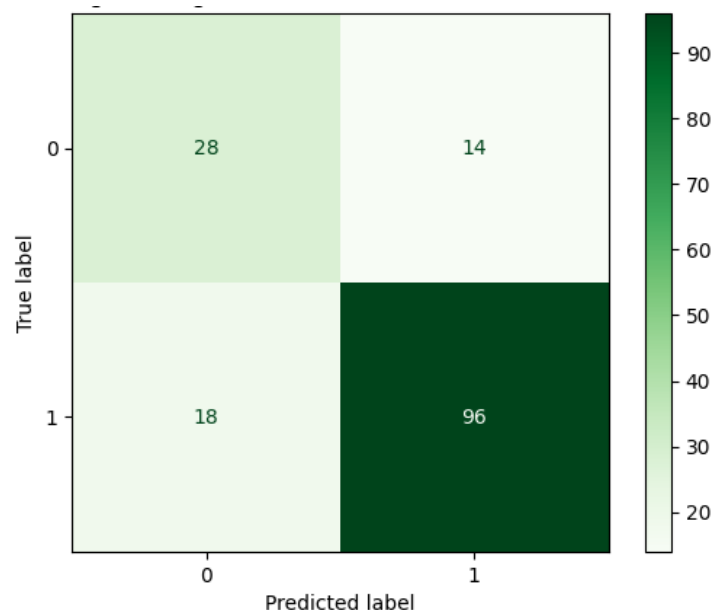


BreastMNIST 224x224: Εξερεύνηση δεδομένων



BreastMNIST 224x224: Logistic regression

LDA



Classification Report:

	precision	recall	f1-score	support
0	0.61	0.67	0.64	42
1	0.87	0.84	0.86	114
accuracy			0.79	156
macro avg	0.74	0.75	0.75	156
weighted avg	0.80	0.79	0.80	156

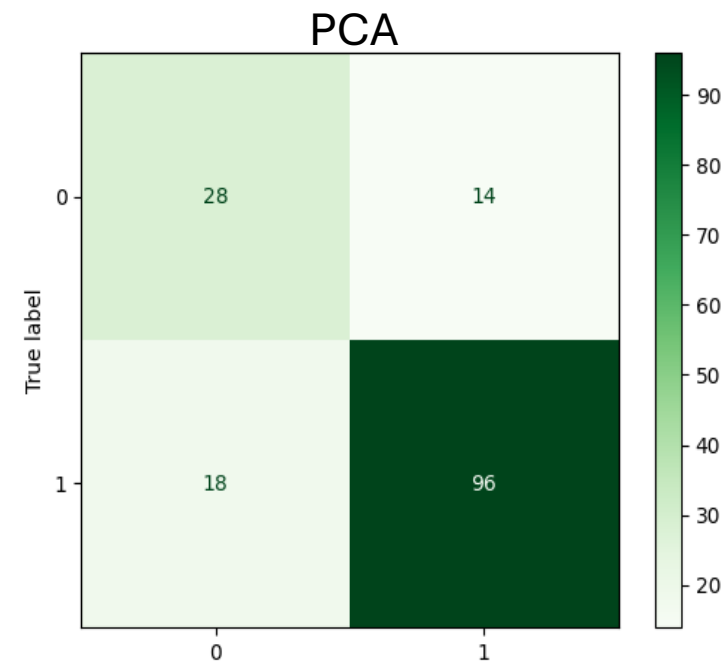
Accuracy: 0.7949

Precision (macro): 0.7407

Recall (macro): 0.7544

F1 Score (macro): 0.7468

PCA



Classification Report:

	precision	recall	f1-score	support
0	0.61	0.67	0.64	42
1	0.87	0.84	0.86	114
accuracy			0.79	156
macro avg	0.74	0.75	0.75	156
weighted avg	0.80	0.79	0.80	156

Accuracy: 0.7949

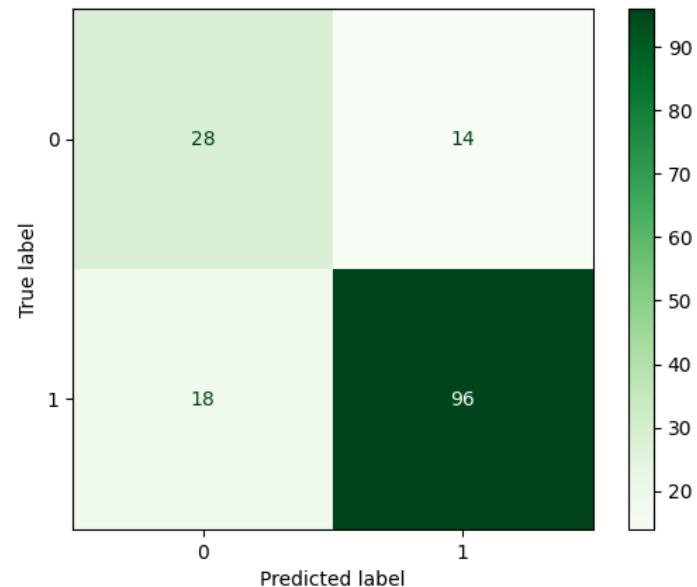
Precision (macro): 0.7407

Recall (macro): 0.7544

F1 Score (macro): 0.7468

BreastMNIST 224x224: K Nearest Neighbor

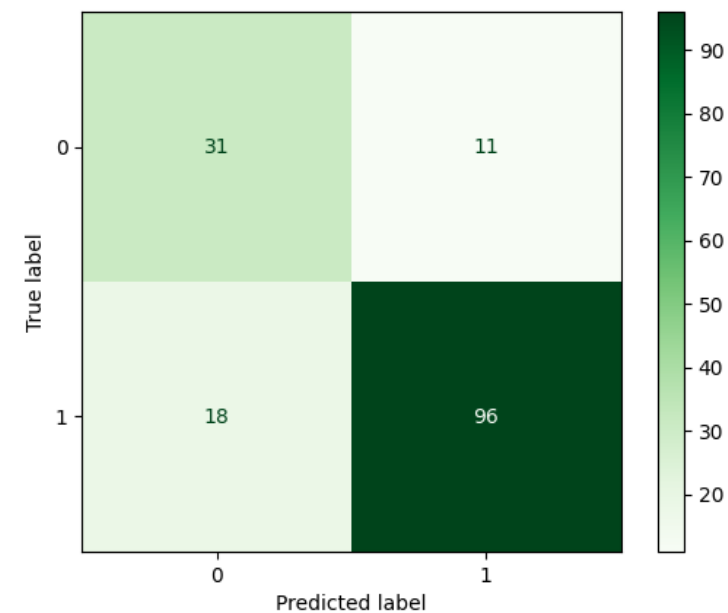
LDA



SVM Classification Report:

	precision	recall	f1-score	support
0	0.33	0.93	0.49	42
1	0.92	0.32	0.47	114
accuracy			0.48	156
macro avg	0.63	0.62	0.48	156
weighted avg	0.76	0.48	0.48	156

PCA



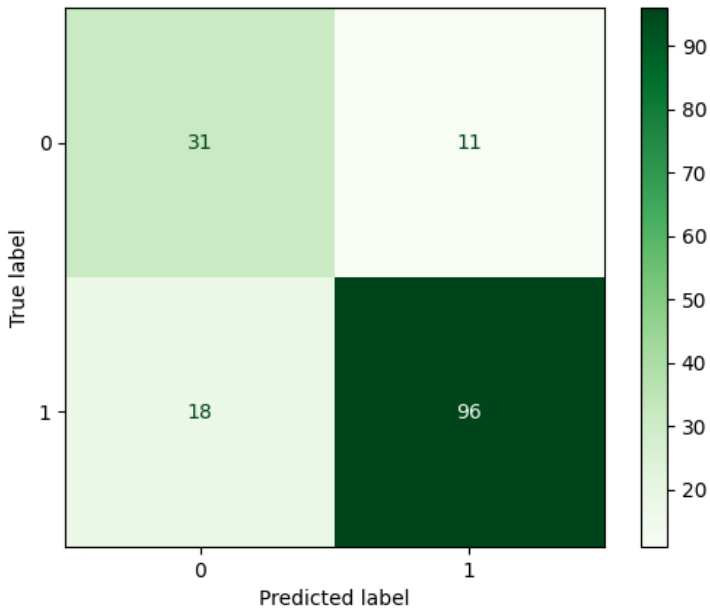
SVM Classification Report:

	precision	recall	f1-score	support
0	0.61	0.67	0.64	42
1	0.87	0.84	0.86	114
accuracy			0.79	156
macro avg	0.74	0.75	0.75	156
weighted avg	0.80	0.79	0.80	156

BreastMNIST 224x224: Support Vector Machine

LDA

SVM with LDA Confusion Matrix



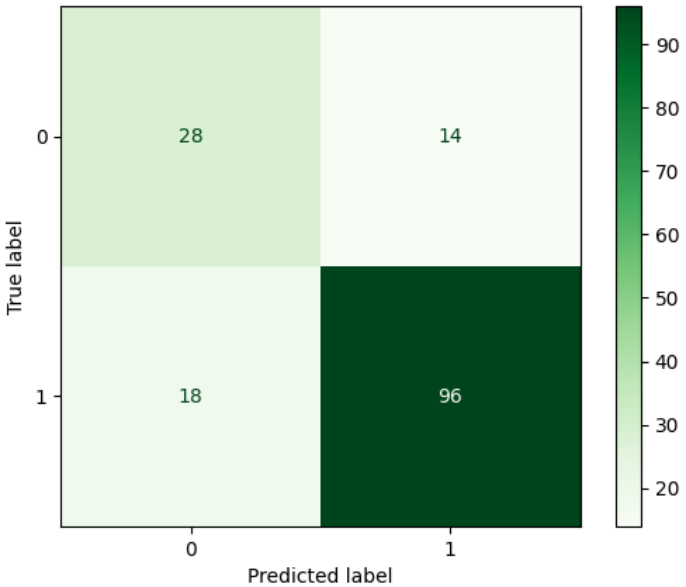
Classification Report:

	precision	recall	f1-score	support
0	0.63	0.74	0.68	42
1	0.90	0.84	0.87	114
accuracy			0.81	156
macro avg	0.76	0.79	0.78	156
weighted avg	0.83	0.81	0.82	156

Accuracy: 0.8141
Precision (macro): 0.7649
Recall (macro): 0.7901
F1 Score (macro): 0.7750

PCA

SVM with PCA Confusion Matrix



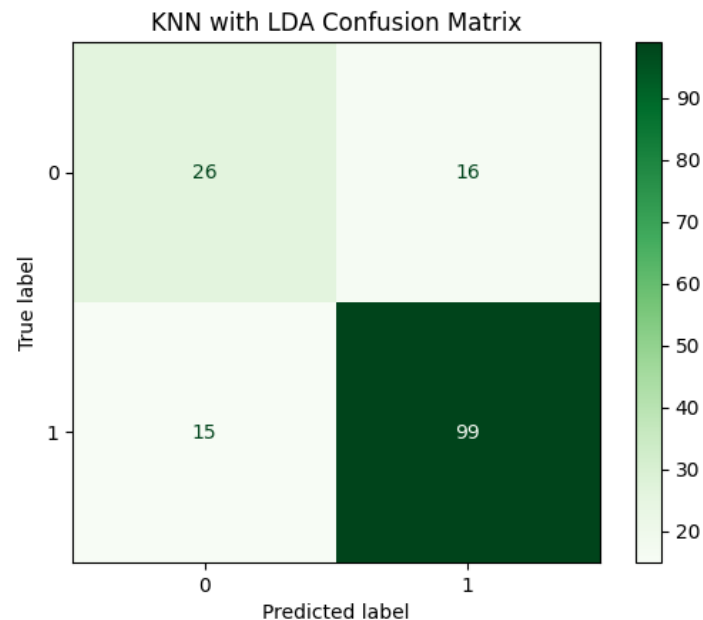
Classification Report:

	precision	recall	f1-score	support
0	0.61	0.67	0.64	42
1	0.87	0.84	0.86	114
accuracy			0.79	156
macro avg	0.74	0.75	0.75	156
weighted avg	0.80	0.79	0.80	156

Accuracy: 0.7949
Precision (macro): 0.7407
Recall (macro): 0.7544
F1 Score (macro): 0.7468

BreastMNIST 224x224: k-NN

LDA

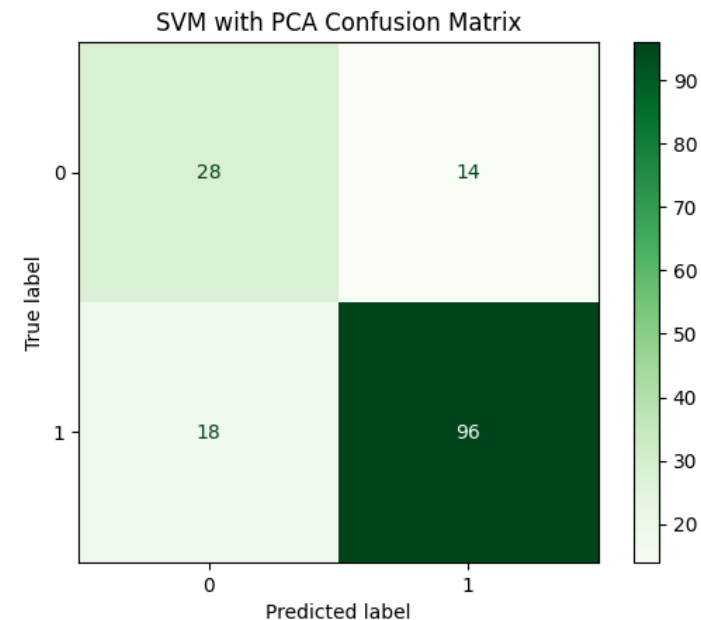


Classification Report:

	precision	recall	f1-score	support
0	0.63	0.62	0.63	42
1	0.86	0.87	0.86	114
accuracy			0.80	156
macro avg	0.75	0.74	0.75	156
weighted avg	0.80	0.80	0.80	156

Accuracy: 0.8013
Precision (macro): 0.7475
Recall (macro): 0.7437
F1 Score (macro): 0.7456

PCA



Classification Report:

	precision	recall	f1-score	support
0	0.77	0.48	0.59	42
1	0.83	0.95	0.89	114
accuracy			0.82	156
macro avg	0.80	0.71	0.74	156
weighted avg	0.81	0.82	0.81	156

Accuracy: 0.8205
Precision (macro): 0.8000
Recall (macro): 0.7118
F1 Score (macro): 0.7367

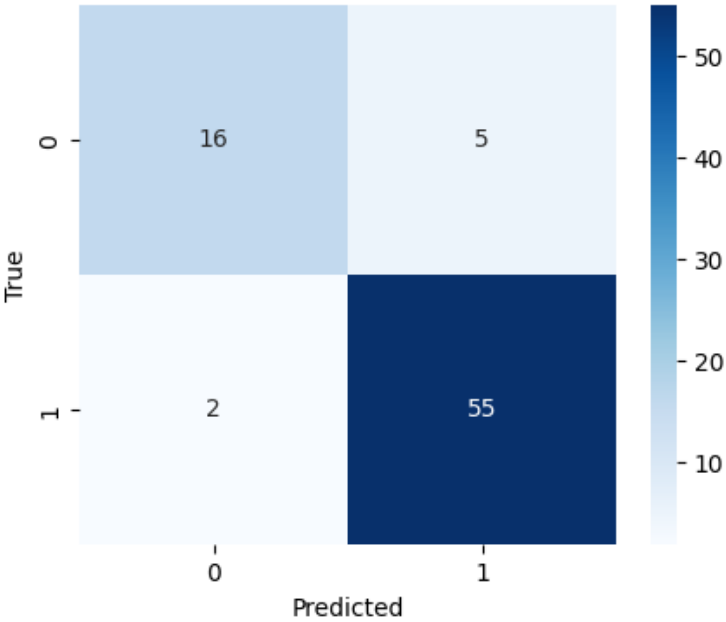
BreastMNIST 224x224: CNN

Validation Accuracy

Test Loss: 0.4457, Accuracy: 91.03%, Recall 0: 0.7619

	precision	recall	f1-score	support
0	0.89	0.76	0.82	21
1	0.92	0.96	0.94	57
accuracy			0.91	78
macro avg	0.90	0.86	0.88	78
weighted avg	0.91	0.91	0.91	78

Test Confusion Matrix

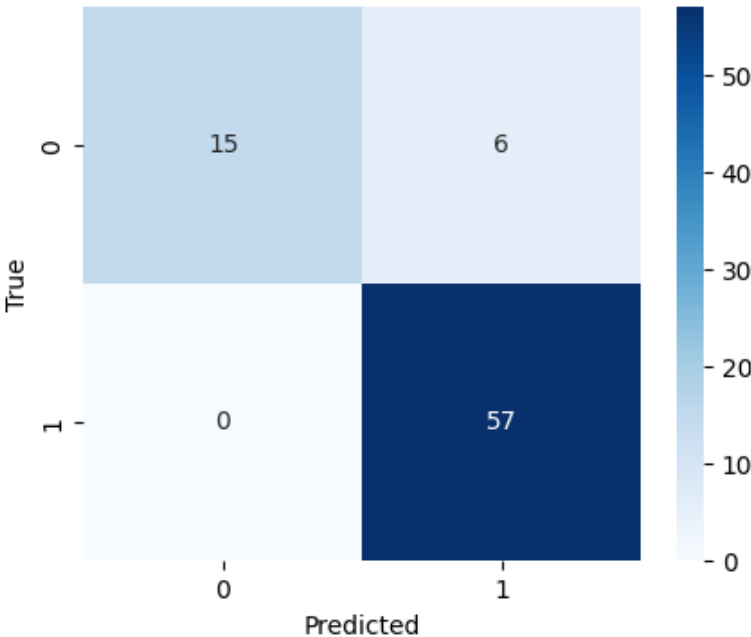


Recall 0

Test Loss: 0.2699, Accuracy: 92.31%, Recall 0: 0.7143

	precision	recall	f1-score	support
0	1.00	0.71	0.83	21
1	0.90	1.00	0.95	57
accuracy			0.92	78
macro avg	0.95	0.86	0.89	78
weighted avg	0.93	0.92	0.92	78

Test Confusion Matrix

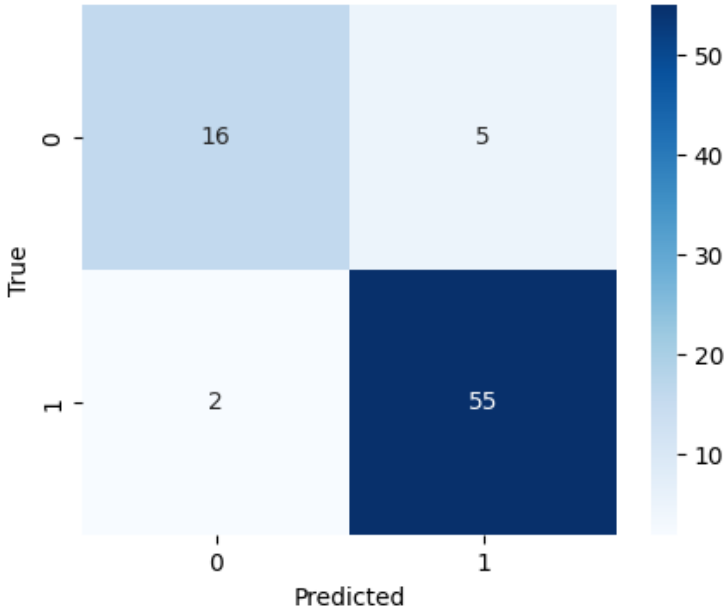


Custom

Test Loss: 0.3670, Accuracy: 91.03%, Recall 0: 0.7619

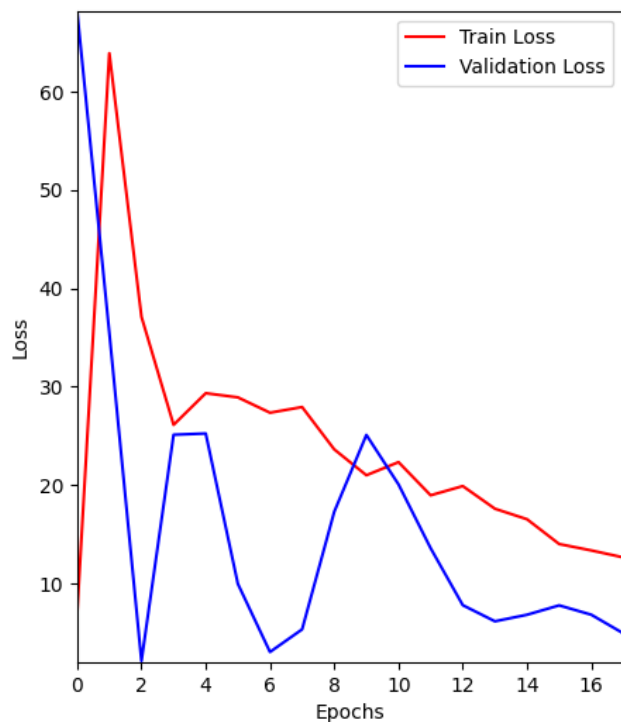
	precision	recall	f1-score	support
0	0.89	0.76	0.82	21
1	0.92	0.96	0.94	57
accuracy			0.91	78
macro avg	0.90	0.86	0.88	78
weighted avg	0.91	0.91	0.91	78

Test Confusion Matrix

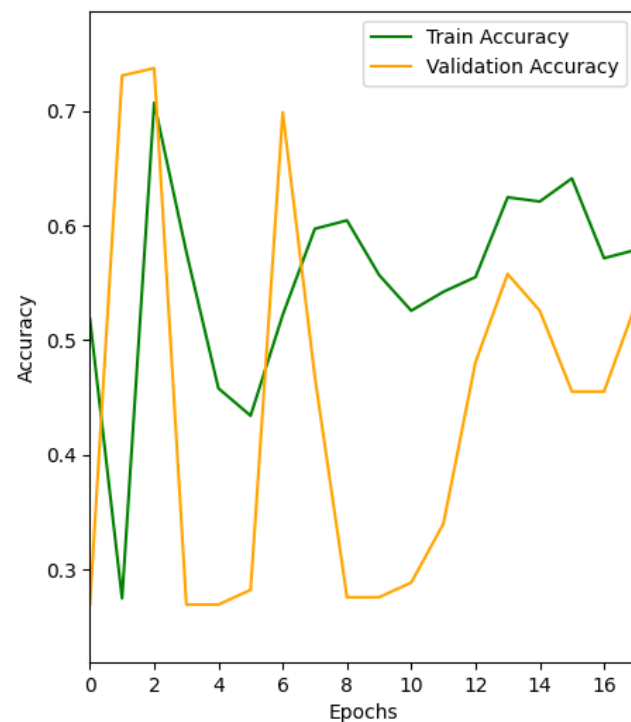


BreastMNIST 224x224: MLP

Losses for Best Model



Accuracies for Best Model



Classification Report for Train Set:

	precision	recall	f1-score	support
0	0.42	0.78	0.55	147
1	0.88	0.61	0.72	399
accuracy			0.65	546
macro avg	0.65	0.69	0.63	546
weighted avg	0.76	0.65	0.67	546

Train Accuracy: 0.6538

Classification Report for Validation Set:

	precision	recall	f1-score	support
0	0.46	0.71	0.56	42
1	0.87	0.69	0.77	114
accuracy			0.70	156
macro avg	0.66	0.70	0.67	156
weighted avg	0.76	0.70	0.71	156

Validation Accuracy: 0.6987

Classification Report for Test Set:

	precision	recall	f1-score	support
0	0.45	0.81	0.58	21
1	0.90	0.63	0.74	57
accuracy			0.68	78
macro avg	0.67	0.72	0.66	78
weighted avg	0.78	0.68	0.70	78

Test Accuracy: 0.6795

BreastMNIST 224x224: Συμπεράσματα

	PCA	LDA	Custom Metric	Best Recall 0	Best Validation Accuracy
Logistic Regression	Recall_0: 0.67 Test Accuracy: 0.7948717948717948	Recall_0: 0.67 Test Accuracy: 0.8141025641025641	-	-	-
SVM	Recall_0: 0.67 Test Accuracy: 0.7948717948717948	Recall_0: 0.93 Test Accuracy: 0.4807692307692308	-	-	-
K-NN	Recall_0: 0.76 Test Accuracy: 0.7051282051282052	Recall_0: 0.76 Test Accuracy: 0.7435897435897436	-	-	-
Naive Bayes	Recall_0: 0.76 Test Accuracy: 0.6602564102564102	Recall_0: 0.60 Test Accuracy: 0.8076923076923077	-	-	-
MLP	-	-	Test Accuracy: 0.6795 Recall 0: 0.81	Test Accuracy: 34.62% Recall 0: 1.0	-
CNN	-	-	Test Loss: 0.3670, Accuracy: 91.03%, Recall 0: 0.7619	Test Loss: 0.2699, Test Accuracy: 92.31% Recall 0: 0.83	Test Loss: 0.4457, Accuracy: 91.03%, Recall 0: 0.7619



Ευχαριστούμε
για τον χρόνο
σας

