# ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ 1<sup>Η</sup> ΣΕΙΡΑ ΑΣΚΗΣΕΩΝ ΚΟΥΡΕΑΣ ΑΘΑΝΑΣΙΟΣ ΑΜ:4392 ΓΕΩΡΓΙΟΥ ΚΩΝΣΤΑΝΤΙΝΟΣ ΑΜ:4333

1η Μέθοδος Ταξινόμησης:

A)Nearest Neighbor k-NN με Ευκλείδεια Απόσταση Παρακάτω θα παραθέσω τον κώδικα που χρειάστηκα για να βγάλω τα αποτελέσματα των δοκίμων

### Κώδικας σε python

```
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers, models
from tensorflow.keras.models import load model
import os
import pandas as pd
import time
from sklearn.neighbors import KNeighborsClassifier
import numpy as np
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy score
from sklearn.metrics import precision score, recall score
from sklearn.metrics import fl score
fashion mnist = keras.datasets.fashion mnist
(train images, train labels), (test images, test labels) = fashion mnis
t.load data() #Edw fortwnoume ta arxeia apo to sunolo dedomenwn pou exo
ume to fashion mnist
train images = train images.reshape((60000, 28, 28, 1)) #Edw katalabain
oume oti oi eikones mas exoun 4 diastaseis opou i prwti diastasi einai
oi test kai oi train eikones
                            #Deuteri kai triti diastasi einai to ipsos
kai to platos tis eikonas kai i tetarti diastasi einai 1 giati
test images = test images.reshape((10000, 28, 28, 1)) #exoume asproma
vres eikones
```

```
train images = train images / 255.0
test images = test images / 255.0
training size = 6000
test size = 1000
train images = train images.reshape(training size*10, 784) #28*28 #Gia
na ftiaxoume tis methodous kai na tis xrisimopoiisoume xreiazomaste oi
eikones na einai mia diastasi
test images = test images.reshape(test size*10, 784) #gia auto tis k
anoume monodiastates
x train, y train, x test, y test =train images, train labels, test imag
es, test labels
x train = x train[:10000] #Edw xrisimopoioume ligotera dedomena giati
pairnei polu xrono gia na exetasei ola ta dedomena
y train = y train[:10000]
start=time.time()
classifier= KNeighborsClassifier(n neighbors=5, metric='euclidean')
#Edw xrisimopoioume tin methodo K-
NN pou ypologizei tin euklideia apostasi
classifier.fit(x train, y train)
y pred knn = classifier.predict(x test)
knn accuracy = accuracy score(y test, y pred knn) #Ypologizoume accurac
knn_recall = recall_score(y_test, y_pred_knn,average= "weighted") #Ypol
ogizoume recall
knn precision = precision score(y test, y pred knn,average= "weighted")
#Ypologizoume precision
knn f1 = f1 score(y test, y pred knn, average= "weighted") #Ypologizoum
e f1 score
end=time.time()
knn time=end-start
print("-----K-NN Report for Euclidean Distance-----
print("F1 score: {}".format(knn f1))
print("Accuracy score: {}".format(knn_accuracy))
print("Precision Score: {}".format(knn precision))
print("Recall Score: {}".format(knn recall))
print("Estimated Time {}".format(knn time))
```

#### Αποτελέσματα 1<sup>ης</sup> Μεθόδου Ευκλείδεια Απόσταση για k=1):

-----K-NN Report for Euclidean Distance-----

F1 score: 0.8050389193711215

Accuracy score: 0.8038

Precision Score: 0.810551225467716 Recall Score: 0.8038

Estimated Time 172.49013328552246

#### Αποτελέσματα 1<sup>ης</sup> Μεθόδου Ευκλείδεια Απόσταση για k=5):

-----K-NN Report for Euclidean Distance-----

F1 score: 0.8176757982415854

Accuracy score: 0.8179

Precision Score: 0.8231371964518323

Recall Score: 0.8179

Estimated Time 177.45053791999817

#### Αποτελέσματα 1<sup>ης</sup> Μεθόδου Ευκλείδεια Απόσταση για k=10):

-----K-NN Report for Euclidean Distance-----

F1 score: 0.812096817392384

Accuracy score: 0.8128

Precision Score: 0.8176699674598117

Recall Score: 0.8128

Estimated Time 180.2935791015625

### A)Nearest Neighbor k-NN με Συνημιτονοειδή Απόσταση

```
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers, models
from tensorflow.keras.models import load model
import os
import pandas as pd
import time
from sklearn.neighbors import KNeighborsClassifier
import numpy as np
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy score
from sklearn.metrics import precision score, recall score
from sklearn.metrics import f1 score
fashion_mnist = keras.datasets.fashion_mnist
(train_images, train_labels), (test_images, test_labels) = fashion_mnis
t.load data()
```

```
train images = train images.reshape((60000, 28, 28, 1))
test images = test images.reshape((10000, 28, 28, 1))
train images = train images / 255.0
test_images = test_images / 255.0
training size = 6000
test size = 1000
train images = train images.reshape(training size*10, 784) #28*28
test images = test images.reshape(test size*10, 784)
x_train, y_train, x_test, y_test = train_images, train_labels,test_ima
ges, test labels
x train = x train[:10000]
y_train = y_train[:10000]
start=time.time()
classifier2= KNeighborsClassifier(n neighbors=1, metric='cosine')
classifier2.fit(x train, y train)
y pred knn2 = classifier2.predict(x test)
knn accuracy2 = accuracy score(y test, y pred knn2)
knn recall2 = recall score(y test, y pred knn2,average= "weighted")
knn precision2 = precision score(y test, y pred knn2,average= "weighted
")
knn f12 = f1 score(y test, y pred knn2, average= "weighted")
end=time.time()
knn2 time=end-start
print("-----K-NN Report for Cosine Distance----")
print("F1 score: {}".format(knn f12))
print("Accuracy score: {}".format(knn_accuracy2))
print("Precision Score: {}".format(knn precision2))
print("Recall Score: {}".format(knn_recall2))
print("Estimated Time {}".format(knn2 time))
```

Αποτελέσματα 1<sup>ης</sup> Μεθόδου Συνημιτονοειδή Απόσταση για k=1 :

-----K-NN Report for Cosine Distance-----

F1 score: 0.8135041691925162

Accuracy score: 0.814

Precision Score: 0.8202003296970454

Recall Score: 0.814

Estimated Time 7.01285529136657

### Αποτελέσματα 1<sup>ης</sup> Μεθόδου Συνημιτονοειδή Απόσταση για k=5 :

-----K-NN Report for Cosine Distance-----

F1 score: 0.8128422152699226

Accuracy score: 0.8168

Precision Score: 0.824690657867799

Recall Score: 0.8168

Estimated Time 6.864148855209351

### Αποτελέσματα 1ης Μεθόδου Συνημιτονοειδή Απόσταση για k=10:

F1 score: 0.805449478862958

Accuracy score: 0.8106

Precision Score: 0.8219570381992511

Recall Score: 0.8106

Estimated Time 6.810957670211792

## 2η Μέθοδος Ταξινόμησης:

A)Neural Networks με 1 κρυμμένο επίπεδο και 500 κρυμμένους νευρώνες

Παρακάτω θα παραθέσω τον κώδικα που χρειάστηκα για να βγάλω τα αποτελέσματα των δοκίμων

Κώδικας σε python

```
from tensorflow import keras
from tensorflow.keras import layers, models
from tensorflow.keras.models import load model
from keras.models import Sequential
from keras.layers import Dense, Activation, Flatten, Dropout, BatchNorm
alization
from keras.optimizers import SGD, Adagrad, RMSprop, Adam
import time
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy score
from sklearn.metrics import precision score, recall score
from sklearn.metrics import f1 score
fashion mnist = keras.datasets.fashion mnist
(train images, train labels), (test images, test labels) = fashion mnis
t.load data()
train images = train images.reshape((60000, 28, 28, 1))
test images = test images.reshape((10000, 28, 28, 1))
input shape = (28, 28, 1)
train images = train images / 255.0
test images = test images / 255.0
x train, y train, x test, y test =train images, train labels, test imag
es, test labels
x train = x train[:10000]
y train = y train[:10000]
model = Sequential()
model.add(Flatten(input shape=input shape))
model.add(Dense(500, activation='relu'))
model.add(Dense(10, activation='softmax'))
model.summary()
model.compile(optimizer=SGD(),
              loss='sparse categorical crossentropy',
              metrics=['accuracy'])
start-time.time()
fit = model.fit(train images, train labels, epochs=30, verbose=0)
y pred NN = model.predict classes(test images, verbose=0)
```

```
NN accuracy = accuracy score(test labels, y pred NN)
NN precision = precision score(test labels, y pred NN, average= "weighte
d")
NN recall = recall score(test labels, y pred NN, average= "weighted")
NN_f1 = f1_score(test_labels, y_pred_NN,average= "weighted")
end=time.time()
NN time=end-start
print("-----
Neural Networks Report for 1 Hidden Layer with K=500-----")
print("Accuracy:{}".format(NN accuracy))
# precision tp / (tp + fp)
print("Precision:{}".format(NN precision))
# recall: tp / (tp + fn)
print("Recall:{}".format(NN recall))
# f1: 2 tp / (2 tp + fp + fn)
print("F1 Score:{}".format(NN f1))
print("Estimated Time:{}".format(NN_time))
Αποτελέσματα 2<sup>ης</sup> Μεθόδου Α):
-----Neural Networks Report for 1 Hidden Layer with K=500-----
Accuracy: 0.881
Precision: 0.8815372075418735
Recall:0.881
F1 Score: 0.8809905653878153
Estimated Time: 1408.963793039322
```

# B)Neural Networks με 2 κρυμμένα επίπεδα και 500 κρυμμένους νευρώνες και 200 κρυμμένους νευρώνες

```
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers, models
from tensorflow.keras.models import load_model
from keras.models import Sequential
```

```
from keras.layers import Dense, Activation, Flatten, Dropout, BatchNorm
alization
from keras.optimizers import SGD, Adagrad, RMSprop, Adam
import time
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy score
from sklearn.metrics import precision score, recall score
from sklearn.metrics import f1_score
fashion mnist = keras.datasets.fashion mnist
(train images, train labels), (test images, test labels) = fashion mnis
t.load data()
train images = train images.reshape((60000, 28, 28, 1))
test images = test images.reshape((10000, 28, 28, 1))
input shape = (28, 28, 1)
train images = train images / 255.0
test images = test images / 255.0
x_train, y_train, x_test, y_test =train_images, train labels,test imag
es, test labels
x train = x train[:10000]
y train = y train[:10000]
model = Sequential()
model.add(Flatten(input shape=input shape))
model.add(Dense(500, activation='relu'))
model.add(Dense(200, activation='relu'))
model.add(Dense(10, activation='softmax'))
model.summary()
model.compile(optimizer=SGD(),
              loss='sparse categorical crossentropy',
              metrics=['accuracy'])
start=time.time()
fit = model.fit(train images, train labels, epochs=30, verbose=0)
y pred NN2 = model.predict classes(test images, verbose=0)
```

```
NN2 accuracy = accuracy score(y test, y pred NN2)
NN2 precision = precision score(y test, y pred NN2, average= "weighted")
NN2 recall = recall score(y test, y pred NN2,average= "weighted")
NN2 f1 = f1 score(y test, y pred NN2, average= "weighted")
end=time.time()
NN2 time=end-start
print("-----
Neural Networks Report for 2 Hidden Layers with K1=500 and K2=200-----
print("Accuracy:{}".format(NN2 accuracy))
# precision tp / (tp + fp)
print("Precision:{}".format(NN2 precision))
# recall: tp / (tp + fn)
print("Recall:{}".format(NN2 recall))
# f1: 2 tp / (2 tp + fp + fn)
print("F1 Score:{}".format(NN2 f1))
print("Estimated Time:{}".format(NN2_time))
```

#### Αποτελέσματα 2<sup>ης</sup> Μεθόδου Β):

```
---Neural Networks Report for 2 Hidden Layers with K1=500 and K2=200---Accuracy:0.8806
Precision:0.8835198883243569
Recall:0.8806
F1 Score:0.8804790305037774
Estimated Time:202.74010276794434
```

### 3η Μέθοδος Ταξινόμησης:

A)Support Vector Machines (SVM): Μηχανές διανυσματικής στήριξης, χρησιμοποιώντας (α) γραμμική συνάρτηση πυρήνα (linear kernel)

Παρακάτω θα παραθέσω τον κώδικα που χρειάστηκα για να βγάλω τα αποτελέσματα των δοκίμων

### Κώδικας σε python

```
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers, models
```

```
from tensorflow.keras.models import load model
import os
import pandas as pd
import time
import numpy as np
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy score
from sklearn.metrics import precision score, recall score
from sklearn.metrics import fl score
fashion mnist = keras.datasets.fashion mnist
from sklearn.svm import SVC
from sklearn.preprocessing import StandardScaler
from keras.datasets import fashion mnist
from sklearn.multiclass import OneVsRestClassifier
(train images, train labels), (test images, test labels) = fashion mnis
t.load data()
train images = train images.reshape((60000, 28, 28, 1))
test images = test images.reshape((10000, 28, 28, 1))
train images = train images / 255.0
test images = test images / 255.0
training size = 6000
test size = 1000
train images = train images.reshape(training size*10, 784) #28*28
test_images = test_images.reshape(test_size*10, 784)
x train, y train, x test, y test =train images, train labels, test imag
es, test labels
scaler = StandardScaler()
x train = scaler.fit transform(x train.astype(np.float64))
x train = x train[:10000]
train = y train[:10000]
start=time.time()
clf = OneVsRestClassifier(SVC(C=1, kernel='linear'))
clf.fit(x train, y train)
y pred svm = clf.predict(x test)
```

Αποτελέσματα μεθόδου Support Vector Machines χρησιμοποιώντας γραμμική συνάρτηση πυρήνα (linear kernel)

B)Support Vector Machines (SVM): Μηχανές διανυσματικής στήριξης, χρησιμοποιώντας Gaussian συνάρτηση πυρήνα (kernel) δοκιμάζοντας διάφορες τιμές της παραμέτρου της

```
import tensorflow as tf
from tensorflow.keras import layers, models
from tensorflow.keras.models import load_model
import os
import pandas as pd
import time
import numpy as np
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy_score
from sklearn.metrics import precision_score, recall_score
from sklearn.metrics import f1_score
fashion_mnist = keras.datasets.fashion_mnist
from sklearn.svm import SVC
from sklearn.multiclass import OneVsRestClassifier
```

```
from sklearn.preprocessing import StandardScaler
from keras.datasets import fashion mnist
(train_images, train_labels), (test_images, test_labels) = fashion_mnis
t.load data()
train images = train images.reshape((60000, 28, 28, 1))
test images = test images.reshape((10000, 28, 28, 1))
train images = train images / 255.0
test images = test images / 255.0
training size = 6000
test size = 1000
train images = train images.reshape(training size*10, 784) #28*28
test images = test images.reshape(test size*10, 784)
x train, y train, x test, y test =train images, train labels, test imag
es, test labels
scaler = StandardScaler()
x train = scaler.fit transform(x train s.astype(np.float64))
x train = x train[:10000]
y train = y train[:10000]
start=time.time()
svm2 = OneVsRestClassifier(SVC(C=1, kernel='rbf'))
svm2.fit(x train, y train)
y pred svm2 = svm2.predict(x test)
svm2 accuracy = accuracy score(y test, y pred svm2)
svm2 recall = recall score(y test, y pred svm2,average= "weighted")
svm2_precision = precision_score(y_test, y_pred_svm2,average= "weighted")
svm2 f1 = f1 score(y test, y pred svm2, average= "weighted")
end=time.time()
svm2 time = end-start
print("-----")
print("F1 score: {}".format(svm2_f1))
print("Accuracy score: {}".format(svm2 accuracy))
print("Precision Score: {}".format(svm2 precision))
```

```
print("Recall Score: {}".format(svm2_recall))
print("Estimated Time:{}".format(svm2 time))
```

# Αποτελέσματα μεθόδου Support Vector χρησιμοποιώντας Gaussian συνάρτηση πυρήνα (kernel)

# Γ)Support Vector Machines (SVM): Μηχανές διανυσματικής στήριξης, χρησιμοποιώντας συνημιτονοειδή συνάρτηση πυρήνα (cosine kernel)

```
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers, models
from tensorflow.keras.models import load model
import os
import pandas as pd
import time
import numpy as np
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy score
from sklearn.metrics import precision score, recall score
from sklearn.metrics import fl score
fashion mnist = keras.datasets.fashion mnist
from sklearn.svm import SVC
from sklearn.preprocessing import StandardScaler
from keras.datasets import fashion mnist
from sklearn.multiclass import OneVsRestClassifier
import math
from numpy import linalg as LA
def my kernel(X, Y):
#Συναρτηση για να υπολογίζει την συνημιτονοειδή συνάρτηση
   norm = LA.norm(X) * LA.norm(Y)
    return np.dot(X, Y.T)/norm
(train_images, train_labels), (test_images, test labels) = fashion mnis
t.load data()
```

```
train images = train images.reshape((60000, 28, 28, 1))
test_images = test_images.reshape((10000, 28, 28, 1))
train images = train images / 255.0
test images = test images / 255.0
training size = 6000
test size = 1000
train images = train images.reshape(training size*10, 784) #28*28
test images = test images.reshape(test size*10, 784)
x train, y train, x test, y test =train images, train labels, test imag
es, test labels
scaler = StandardScaler()
x train = scaler.fit transform(x train.astype(np.float64))
x train = x train[:10000]
y_train = y_train[:10000]
start=time.time()
svm3 = OneVsRestClassifier(SVC(kernel = my kernel,C=1))
svm3.fit(x train, y train)
y pred svm = svm3.predict(x test)
svm3 accuracy = accuracy score(y test, y pred svm)
svm3 recall = recall score(y test, y pred svm,average= "weighted")
svm3_precision = precision_score(y_test, y_pred_svm,average= "weighted"
svm3 f1 = f1 score(y test, y pred svm, average= "weighted")
end=time.time()
svm3 time=end-start
print("-----SVM Report for Cosine----")
print("F1 score: {}".format(svm3 f1))
print("Accuracy score: {}".format(svm3 accuracy))
print("Precision Score: {}".format(svm3 precision))
print("Recall Score: {}".format(svm3 recall))
print("Estimated Time:{}".format(svm3 time))
```

# Αποτελέσματα μεθόδου Support Vector χρησιμοποιώντας συνημιτονοειδή συνάρτηση πυρήνα (kernel)

-----SVM Report for Cosine-----

F1 score: 0.4976564483010756

Accuracy score: 0.541

Precision Score: 0.577944951511452

Recall Score: 0.541

Estimated Time: 56.812220096588135

### Σύγκριση Μεθόδων

Για τα K-NN παρατηρούμε ότι για την Ευκλείδεια Απόσταση για k=5 δηλαδή 5 κοντινότερους γείτονες έχει καλυτέρα αποτελέσματα (F1 Score, Accuracy, Recall, Precision) από την ίδια μέθοδο για k=1 η k=10. Επισης για τα K-NN παρατηρούμε ότι για την συνημιτονοειδή απόσταση πάλι για k=5 έχουμε τα καλυτέρα αποτελέσματα (F1 Score, Accuracy, Recall, Precision)) από την ίδια μέθοδο για k=1 η k=10. Αλλα αναμεσά στην μέθοδο με την ευκλείδεια απόσταση και την μέθοδο με την συνημιτονοειδή απόστασή ευκλείδεια απόσταση έχει καλυτέρα αποτελέσματα αλλά παίρνει πολύ περισσότερο χρόνο από την συνημιτονοειδή απόσταση δηλαδή Ευκλείδεια Απόσταση:

Estimated Time 177.45053791999817

Συνημιτονοειδής Απόσταση: Estimated Time 6.864148855209351

Αλλά τα αποτελέσματα των δυο μεθόδων είναι αρκετά κοντά οπότε καλύτερη μέθοδος είναι η Κ-ΝΝ με συνημιτονοειδή απόσταση γιατί κάνει πολύ λιγότερο χρόνο

Για τα Νευρωνικά Δίκτυα παρατηρούμε ότι όταν έχουμε δυο κρυμμένα επίπεδα η ταξινόμηση γίνεται πιο γρηγορά αλλά έχει μικρότερο accuracy(0.8806) από όταν έχουμε ένα κρυμμένο επίπεδο Επομένως παρατηρούμε ότι στα νευρωνικά δίκτυα χρειαζόμαστε περισσότερο χρόνο για να ταξινομήσουμε τα δεδομένα μας με ένα κρυμμένο επίπεδο αλλά έχουμε καλύτερο accuracy(0.881),όμως δεν είναι τόσο μεγάλη η διαφορά . Επίσης παρατηρούμε ότι από τα αλλά αποτελέσματα (F1 score,Precision,Recall) ,οι δυο μέθοδοι των νευρωνικών δικτυών είναι πολύ κοντά στα αποτελέσματα αλλά η μια χρειάζεται πάρα πολύ χρόνο . Επομένως η καλύτερη μέθοδος είναι όταν έχουμε δυο κρυμμένα επίπεδα στα νευρωνικά δίκτυα γιατί η ταξινόμηση γίνεται σε χρόνο (202.74010276794434 δευτερόλεπτα) που είναι περίπου το 1/7 του χρόνου των νευρωνικών δικτυών με 1 κρυμμένο επίπεδο

Για τα SVM παρατηρούμε ότι καλύτερη ακρίβεια έχει η γραμμική μέθοδος με σχετικά μεγάλη διαφορά δηλαδή SVM Linear accuracy 0.7212, SVM Gaussian accuracy 0.5688, SVM Cosine accuracy 0.541
Αλλά το SVM Linear χρειάζεται 335.7148184776306 δευτερόλεπτα ενώ το

Αλλα το SVM Linear χρειαζεται 335./148184//6306 οευτερολεπτα ενώ το SVM Cosine χρειάζεται 56.812220096588135 δευτερόλεπτα και το

SVM Gaussian χρειάζεται 239.29149317741394 δευτερόλεπτα

Αλλά εφόσον έχουν τόσο μεγάλη διαφορά στην ακρίβεια(accuracy) η καλύτερη μέθοδος είναι η SVM Linear

Επίσης αν κρίνουμε και από τα αλλά αποτελέσματα (F1 score, Precision, Recall) η SVM Linear είναι η καλύτερη από τις άλλες μεθόδους.

Τέλος παρατηρούμε ότι η καλύτερη μέθοδος από όλες είναι η K-NN με συνημιτονοειδή απόσταση γιατί έχει πολύ καλά αποτελέσματα και κάνει και τον λιγότερο χρόνο από τις υπόλοιπες μεθόδους Αλλά τα καλυτέρα αποτελέσματα μας τα δίνουν τα νευρωνικά δίκτυα με δυο κρυμμένα επίπεδα