

הסברים על המחברות לעבודת גמר - קורס הדמיית נתונים:

תבור לוי, 315208439

מחברת ראשונה: שיפור סמסטר א

במחברת זאת התבקשנו לשפר את המודל שבנינו בסמסטר הקודם. למרות כל ניסיונותי לא הצלחתי לשפר את המודל, אך מודל voting stacking הגיע לאותה תוצאה כמו המודל הקודם. עשיתי ניסיון של הדאטא מחדש על מנת לדייק יותר את המודל והשתמשתי בשיטות שונות של למידת מכונה (voting (hard and soft), stacking, bagging, xgboost, adaboost, gradient boosting))

מחברת שניה : Fashion Mnist

במחברת זאת התבקשנו לטעון את הדאטא ולסווג ל multy class. לאחר הצגה של הדאטא ובניית פונקציות עזר השתמשתי במודלים על מנת להגיע לדיוק מקסימלי בסוף השתמשתי ב10000 הדוגמאות של test

מחברת שלישית: dog vs cat

במחברת זאת בעזרת פקודות של pandas ופונקציות עזר הצלחתי להמיר את התמונות לקובץ csv ומשם הדרך הייתה הרבה יותר פשוטה. בדקתי את המודלים כמו בשתי המחברות הקודמות עד לדיוק מירבי

מחברת רביעית: hand motion

במחברת זאת עיקר העבודה הייתה ליצור את קובץ האימון והמבחן. לאחר שהעליתי את כל המידע השארתי אדם אחד בצד בשביל validation וכל שאר האנשים הכנסתי לקובץ test באמצעות פקודות של Pandas ופונקציות עזר. בסוף המחברת הכנתי שוב קובץ test משלושת האנשים הנותרים ועליו בדקתי את המודל הסופי

Models:

KNN

Logistic Regression

XGBoost

Ada Boost

Voting (hard and soft)

Stacking

Bagging

Gradient Boosting

Method:

Pca – dimensionality reduction

Pandas, Numpy, Seaborn, Matplot

Classification report

Accuracy

Scaling

Ensamble learnning