

סיכום עבודה:

במחברת הראשונה ניסיתי לשפר את הדיוק מאז הסמסטר הקודם באמצעות המודל KMeans,

ומודלים של ensemble learning.

אחרי הריצה על המודלים ניסיתי להוריד מיימדים עי PCA ולבדוק אם זה משפיע על הדיוק, הפחתת המיימדים לא השפיע על הדיוק אז השארתי אותה לטובת היעילות.

מודלים של ensemble learning שהשתמשתי בהם:

AdaBoostClassifier.1

GradientBoostingClassifier.2

VotingClassifier.3

StackingClassifier.4

(ביצעתי סטנדרזציה על הדטה באמצעות standarscler כבר בסמסטר הקודם, אחרת הייתי חושבת להוסיף גם את הפעולה הזאת)

במחברת השניה והשלישית ביצעתי סטנדרזציה לדטה

(לאחר שהפכתי את קובץ התמונות באמצעות pandas ל DataFrame עם ערכים שלמדנו לעבוד איתם)

הורדת מימדים באמצעות PCA

חילקתי את הדטה ל 3 חלקים , trine validation test

את האימון של המחברות ביצעתי על trinen וה validation והשתמשתי ב test רק בריצה האחרונה והצגת הדיוק הסופי אחרי שבחרתי את המודל האידיאלי עבור הדטה.

(ניסיתי לבצע cross validation אבל זה לא הועיל לדיוק בצורה משמעותית)

במחברת הרביעית התחלתי ביצירת שלושה דטה סטים שונים, trine validation test

איחדתי כל frame left and right, דגמתי כל שורה חמישית, איחדתי כל שלוש שורות, השמטתי עמודות

מיותרות וניקיתי את הדטה לפי הצורך עם pandas

אחרי שהדטה היה מוכן לעבודה פעלתי בצורה דומה לדרך שפעלתי על מחברות שתיים ושלוש.

המודלים שהשתמשתי בהם:

DecisionTreeClassifier

LogisticRegression

KNeighborsClassifier

AdaBoostClassifier

XGBClassifier

GradientBoostingClassifier

RandomForestClassifier

VotingClassifier

StackingClassifier