**הסבר על הפרוייקט**

**חילקנו לשלוש מחלקות:**

NaiveBayes – מחלקה זו היא המחלקה הראשית והיא בונה את ה-GUI בתוכה יש את הפונקציות הבאות:

**def** browse(self): הפוקציה פותחת אפשרות לבחירת נתיב

**def** load\_csv(self, filename): הפונקציה טוענת קובץ CSV

**def** load\_txt(self, filename): הפונקציה טוענת קובץ txt

**def** build(self):הפונקציה בונה את המודל על פי ה-train

**def** classify(self): הפונקציה מסווגת את ה-test לפי המודל

**def** \_\_init\_\_(self, master): הפונקציה בונה את ה-GUI

Analayser – מחלקה זו בונה את המודל לפי ה-train set

**def** average(self, key): הפונקציה משלימה לכל עמודה עם ערכים נומרים את הערכים החסרים על ידי ממוצע של העמודה

**def** mostCommon(self, key): הפונקציה משלימה לכל עמודה עם ערכים קטגורייאלים את הערכים החסרים על ידי הערך השכיח ביותר בעמודה

**def** handleMissingData(self): הפונקציה מטפלת בערכים חסרים

**def** EqualWidth(self, maxval, minval, cut\_points): הפונקציה מבצעת חלוקה לרוחב עבור עמודה

**def** binning(self, col, cut\_points, key): הפונקציה מבצעת חלוקה לרוחב ומחליפה את הערכים לתגיות

**def** Discretization(self): הפונקציה מבצעת דיסקרטיזציה לכל משתנה נומרי

**def** calculateProb(self, classValue, attribute, category): הפונקציה מחשבת את ההסתברות המותנת של כל ערך ב-train set בהנתן הערך של ה-class.

**def** probClass(self): הפונקציה מחשבת את ההסתברות של כל ערך ב-class

**def** createClassProbTable(self, attribute, category): הפונקציה מבצעת טבלת הסתברויות עבור על ערך שב-class

**def** createTableCategory(self, attributesCategory, category): הפונקציה בונה טבלת הסתברויות עבור כל תכונה אפשרית בכל קטגוריה

**def** buildTablesProb(self): הפונקציה בונה את טבלת כל ההסתברויות המותנות

**def** buildModel(self, sructureDic, train, numOfBins): הפונקציה בונה את המודל

classifier - מחלקה זו מסווגת את test set לפי המודל

**def** Discretization(self): הפונקציה מבצעת דיסקרטיזציה לכל משתנה נומרי

**def** writeToFile(self, dirpath, classification): הפונקציה כותבת לקובץ את התוצאות של הסיווג

**def** getMult(self, listProb): הפונקציה מבצעת כפל בין הסתברויות

**def** getProb(self, model, category, attribute, classAttribute): הפונקציה שולפת את ההסתברות המתאימה מטבלת ההסתברויות

**def** getClassificationRow(self,probByVal): הפונקציה שולפת את ההסתברות המתאימה עבור כל ערך של class

**def** getClassification(self,model): הפונקציה מחשבת את הסיווגים עבור כל ה-test

**def** classify(self, test, structIntervals, structDic, structLables, dirpath,model): הפונקציה מקבלת את כל הסיווגים ורושמת אותם לקובץ