**עבודת הגשה #4 בינה עסקית ומדעי הנתונים:**

**מבנה הפרויקט:**

בחלון הנפתח בעת הרצת התוכנית יש לבחור את הנתיב המלא אל התיקייה בה נמצאים קבצי ה- train, ה- test וה- Structure. לאחר מכן יש לכתוב את גודל ה- bind הרצוי ואז תתאפשר לחיצה על כפתור ה- Build. ברגע שבניית המודל נגמרת תעלה התראה שהבנייה הסתיימה ולאחר מכן יהיה ניתן ללחוץ על כפתור ה- classify ולקבל את פלט הסיווג.

**תיאור המחלקות:**

* המחלקה NaïveBayes: מחלקה זו הינה ה"מוח" של התוכנית, בה מתבצעים שלבי ה- build וה- classify כנדרש.
* המחלקה GUI: כשמה כן היא מכילה את קוד ה- GUI של התוכנית.

**תיאור השיטות:**

מחלקת NaïveBayes:

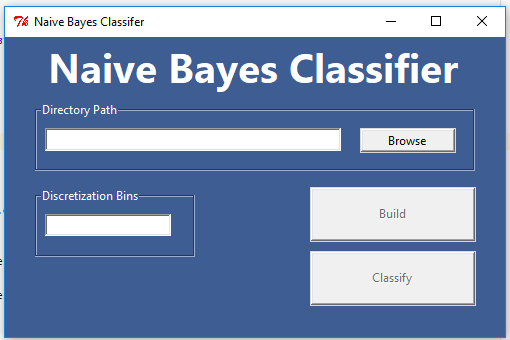
* Build (folderPath, bind): פונקציה זו ראשית קוראת את קבצי ה- train ה- structure וה-test ובונה מהם את מבני הנתונים בהם אנו משתמשים בתוכנית (=רשימות אשר מכילות את התצפיות השונות וכדומה..).  
  לאחר מכן מתבצעת השלמת ערכים חסרים בכל רשומה:  
   fillEmptyValues(): מוצאת ומעדכנת ערכים בתאים חסרים.   
   missingValues(): פונקציה אשר יוצרת רשימה של הערכים החסרים אותם יש לשים בכל עמודה.

ולבסוף מתבצעת דיסקרטיזציה לערכי התכונות הנומריות (בעזרת הפונקציה discretization).

* Classify(): פונקציה בה מתבצע תהליך הסיווג ע"י חישוב הסתברויות. בתחילת הפונקציה מתבצעת השלמת ערכים חסרים ודיסקרטיזציה בעבור כל רשומה בקובץ ה-test. לאחר מכן נעשה חישוב ההסתברויות בעבור כל רשומה בקובץ ה- test לפי נוסחת ה- M-Estimator (המתבצעת בפונקציה mEstimator והיא נעזרת בפונקציה ncValue) ומודפס את קובץ ה- output הסיווג הנבחר של אותה רשומה.

מחלקת ה-GUI:

* מחלקה זו מריצה את ממשק המשתמש, ובתוכו קיימות פונקציות (כדוגמת valuesCheckButtonsAbillity, startBuild, startClasiffy) אשר בודקות את טיב הערכים אותם הזין המשתמש ובמידת הצורך מקפיצות התראות ודרישות שונות.  
  בנוסף, קיימים listenters (כדוגמת bindListenerVal) אשר מוודאים גם כן את הערכים אשר הוזנו ומאפשרים/לא מאפשרים לחיצות על כפתורי הממשק.

דוגמת ממשק המערכת:

בנוסף: קישור ל-GIT של הפרוייקט : <https://github.com/ilayFriedman/naiveBayes>