

דו"ח עבודה 4 בבינה עסקית

בעבודה שלנו יצרנו 3 מחלקות שאחראיות על התהליך.

הקובץ שמכיל את המחלקה GUI הוא גם הקובץ שמכיל את תחילת התכנית, כלומר main.

מחלקה GUI:

מחלקה זו אחראית על העיצוב של ה-GUI שאותו המשתמש רואה ברגע שמריץ את התוכנה. מחלקה זו מכילה תכונות כגון Labels / Textboxes / buttons שאחראיות על חיבור המשתמש עם הפונקציות.

פונקציות:

__init__ - פעולה בונה שיוצרת את כל gui.

validate - פונקציה אשר אחראית לבדוק כאשר מכניסים ערך לשדות ה-Textbox, שערך זה יהיה מספרי. במידה והוא לא מספרי היא מחזירה False ובכך לא נכניס את ערך זה.

browse - פונקציה שאחראית לפתוח חלון של בחירת קובץ נתונים ולשמור את הבחירה ב-Textbox.

preprocess - פונקציה שאחראית ליצור instance של processed data class שעליה נפרט בהמשך.

kmeans - פונקציה שאחראית לקבל את המשתנים שהמשתמש הכניס, לבדוק את תקינותם, וליצור את המודל. לאחר יצירת המודל היא מציירת את הגרפים: scatter / choropleth-map.

מחלקה PreProcess:

מחלקה זו אחראית לטפל בקובץ נתונים ולעשות לו את כל תהליך ה-Preprocessing.

פונקציות:

__init__ - פעולה בונה שמקבלת נתיב של קובץ נתונים. היא קוראת את הקובץ ומבצעת השלמת שורות חסרות של עמודות לפי ממוצע העמודה. נוסף על כך היא מנרמלת כל עמודה לפי ממוצע ולפי סטיית התקן. לאחר מכן היא יוצרת dataframe חדש של ממוצע התכונות של כל מדינה.

מחלקה Kmeans:

מחלקה זו אחראית על לבנות את המודל ולחזות את הקבוצות של הנתונים. בנוסף היא גם אחראית על יצירת הגרפים של התוצאות.

פונקציות:

__init__ - פעולה בונה שמקבלת את הפרמטרים שהמשתמש הכניס ובונה מודל ושומרת את התוצאות בתוך dataframe שקיבלה. הגבולות של מספר הקלסטרים הוא בין 2 למספר הארצות, והגבולות של מספר הריצות הוא כל מספר int חיובי.

draw - פעולה שיוצרת את כל הגרפים (חשוב לציין שהיא לא מציירת אותם אלא רק יוצרת) לפי התוצאות שקיבלנו בפעולה הבונה.