**236501**

**מבוא לבינה**

**תרגיל בית 3**

חלק ב'

3ד .2. חשוב לשמור על עקביות זו כוון שנרצה להשוות בין מסווגים. לפיכך נרצה להיות עקביים עם המבחנים איתם אנו בוחנים את המסווגים, כדי למנוע מצב בו מסווג מקבל ציון נמוך בגלל חלוקה קשה יותר של הנתונים הגולמיים ולא בגלל איכות סיווג נמוכה יותר.

5. 2.

5.3. כפי שניתן לראות ערכי השגיאה המינימאליים והדיוק המקסימלי התקבלו עבור .

5.4. ניתן להבחין כי המסווגים באופן כללי הניבו סיווג טוב של הדוגמאות. הופתענו לראות שהמסווג שקיבל את הציון הגבוה ביותר הוא דווקא זה עם הפרמטר , כלומר המסווג שמסווג את הדגימה על סמך אובייקט יחיד שהוא הקרוב ביותר ב- DataSet לדגימה. ציפינו שדוגמאות רועשות ב-DataSet יובילו ליותר סווגים שגויים במקרה של החלטה על סמך שכן יחיד קרוב ביותר. נראה שבחלוקה של מסד הנתונים אותה ביצענו הדוגמא הסמוכה ביותר לכל דגימה באופן מובהק יותר דוגמא טובה, ואילו רק בהרחבת החיפוש לדוגמאות רחוקות יותר מסתננות דוגמאות רועשות אשר משבשות את ההחלטה.

7.4. התוצאות הטובות ביותר התקבלו עבור מסווג KNN עם פרמטר .