סיכום

במעבדה למדנו כלים חדשים מספריות המובילות לנו כבר. למדנו איך לאמן מודל סיווג של naive bayes בעצמנו ולפי דרך ספריות פייתון שונות בעזרת מודלים שונים. נכנסו יותר לעומק מה זה אלגוריתם naive bayes ואיזה סוגי מישוש שלו קיימים.

למדנו שלוש שיטות של feature selection ומתי כל אחת מהן יכולה להיות שימושית לנו ואיזה סוג naive bayes הכי מתאים.

במעבדה למדנו עשינו השוואה של כמה מודלים ומדדי סיווג והגענו למסקנה שמודל מולטינומיאלי של סיווג הכי מתאים ויעיל לסיווג לקסט. אותנו מתיחסים לכל המילים שלנו בכל המסמכים במילונים ממוימים. בנינו קבצים לאימון ולבדיקה והעברנו תמונים אלה על המודלים וראינו שדרך שיטה אחת מקבלים דיוק הכי גבוה. ולפי פחות לפי כידה שקיבלנו ראינו ששיטה זו יותר יעילה. MultinomialNB היה אלגוריתם הכי מתאים מכל המודלים שהשתמשנו בהם.

Сканировано с CamScanner