

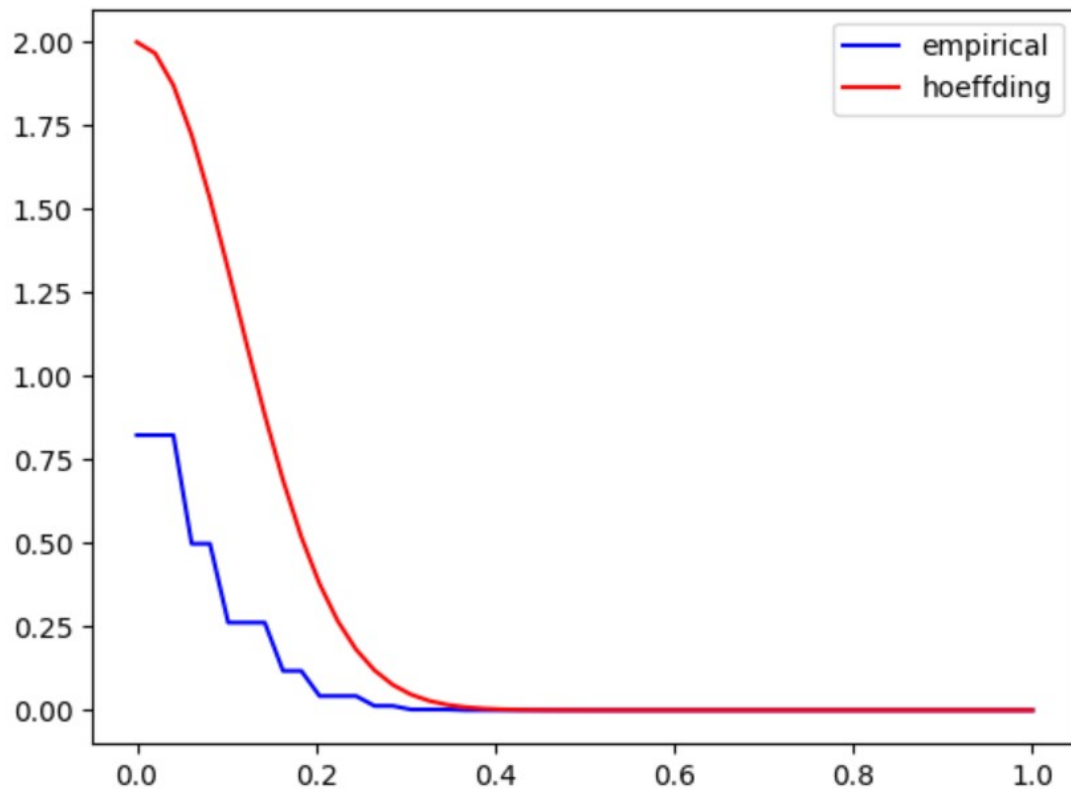
(a)

```

N = 20000
n = 20

mat = np.random.randint(2, size=(N,n))

```

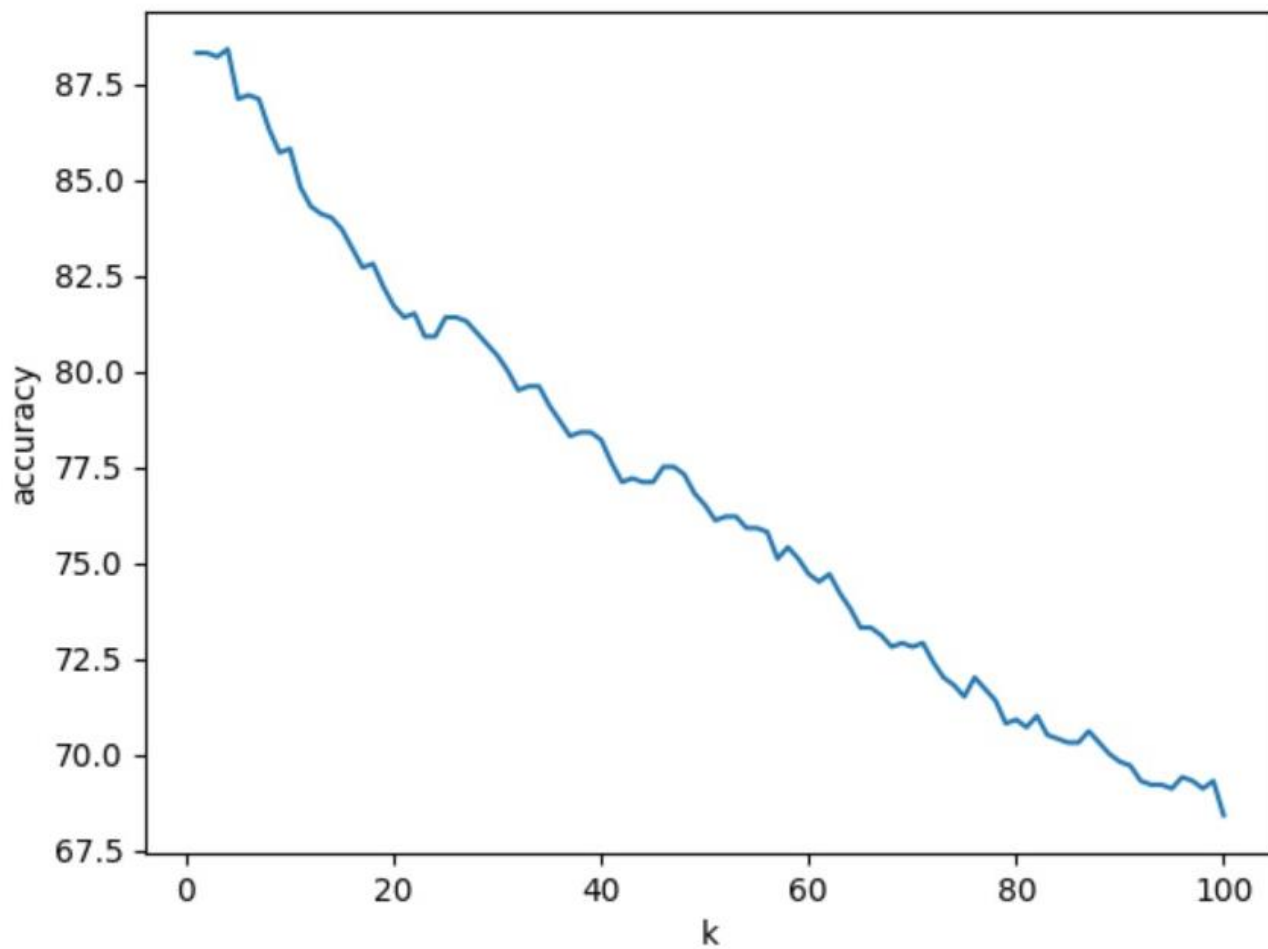


.4

(b)

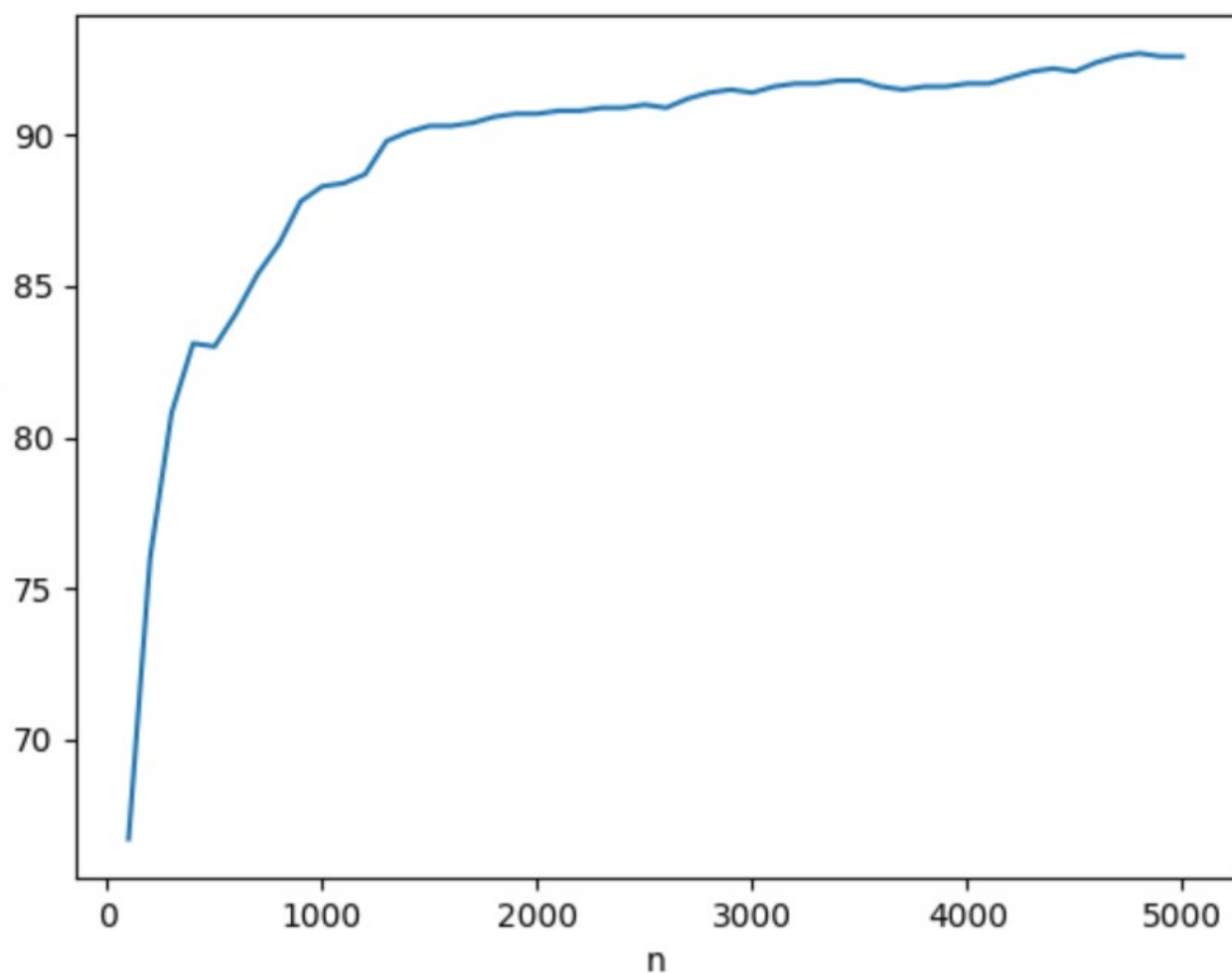
הדיוק שמתקבל עבור $n = 1000$, $k=10$ הוא 85.8 אחוז גובה בהרבה מניחוש בהתפלגות אחידה על 10 אפשרויות. כלומר פרידקטור רנדומלי בהנחה שהתמונות מגיעות בהתפלגות אחידה אמור להצליח לזהות בערך אחד מעשר תמונות כלומר עשר אחוז הצלחה נמוך בהרבה מהאחוז שהגענו אליו.

(c)



מהגרף ניתן לראות שככל שאנו משווים ליותר שכנים כך הדיוק יורד מכיוון שכאשר אנו מכנסים הרבה שכנים יכול להיות שנכניס קבוצה גדולה של תמונות שהם במרחק מאוד גדול ולכן גם בעלי ליבל שונה מהתמונה שעכשיו אנחנו מנסים לחזות עבורה ולכן סיכוי גובה יותר שנטעה בחיזוי.

(d)



מהגרף עולה שכלל שיש יותר דוגמאות להשוואה הדיוק עולה הסבר לכך הוא שכל שיש יותר דוגמאות כך הסיכוי שיש דוגמאות דומות מאוד (כלומר מרחק אוקלידי קטן בניהם) לתמונה שאנו מנסים לבצע חיזוי עולה ומכיוון ש $k=1$ מספיק למצוא דוגמא האחת המרחק שלה מהתמונה שאותה אנו רוצים לבצע עליה חיזוי כדי שנבצע חיזוי שזהה לליבל של הדוגמה הקרובה ביותר.