

パターン情報学 プログラミング課題 1

03-170312
航空宇宙工学科
新幡 駿

2017 年 10 月 14 日

1 アルゴリズムについて

指示された k 近傍法を用いた. 近傍の判断基準は四次元の空間のユークリッド距離を用いた. また授業で指示された一つ抜き法を用いて k を評価した.

2 結果

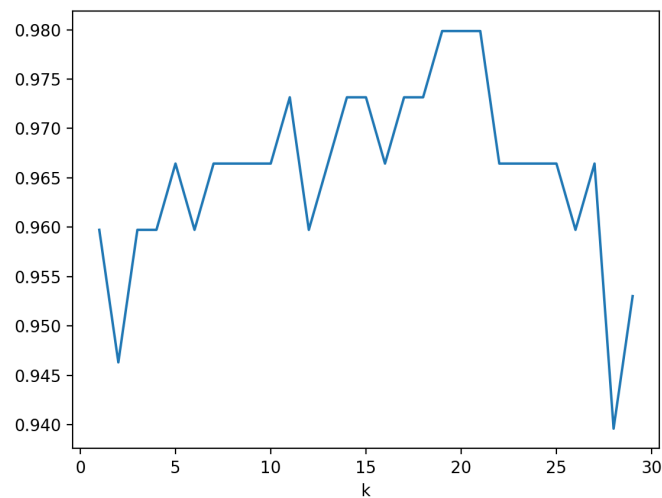


図 1 一つ抜き法での識別率

最適な k の値は 19 であった.

3 考察

k の値は 20 を越えると識別率は単調現象を始めた. 授業では k の range を 1 から 30 と指示されていたのでそのグラフを掲載してるが試しに 50 までやってみるとはっきりとその傾向を確認することができた. k の値がある程度を越えると近傍に同じクラスの点が少ないためだと思う. 今回はデータ全体の点の数が 150 だったので 20 を越えると識別率が下がり始めたが, データの点の数が多いと識別率が減少し始める数は大きくなると思う. また k の値が 1 から 20 の間はあまり識別率が変動しないが大きい値の方がノイズに強いと思われる.