

تمرین پنجم عملی

دانیال ملک محمد 94100092

تمرین پنجم

(۱)

k= 1 , Accuracy : 0.9774092078924793
k= 2 , Accuracy : 0.9739776951672863
k= 4 , Accuracy : 0.976551329711181
k= 8 , Accuracy : 0.9754074921361167
k= 16 , Accuracy : 0.9708321418358593

(۲)

Accuracy : 0.7057477838146983

Confusion Matrix:

```
[[322.  0.  0.  0.  8.  0.  1. 27. 109. 10.]  
 [ 0. 170.  0.  2.  5.  0.  0.  9.  0. 21.]  
 [ 2. 145. 360.  0.  1.  0.  0. 35.  5.  0.]  
 [ 0.  45.  4. 333. 25. 162. 17. 24.  1. 197.]  
 [ 1.  4.  0.  1. 314.  0.  4.  0.  0. 18.]  
 [ 0.  0.  0.  0.  0. 167.  0.  0. 40.  1.]  
 [35.  0.  0.  0. 11.  3. 313.  0.  0.  0.]  
 [ 3.  0.  0.  0.  0.  0.  1. 268. 48.  0.]  
 [ 0.  0.  0.  0.  0.  1.  0.  1. 132.  0.]  
 [ 0.  0.  0.  0.  0.  2.  0.  0.  0. 89.]]
```

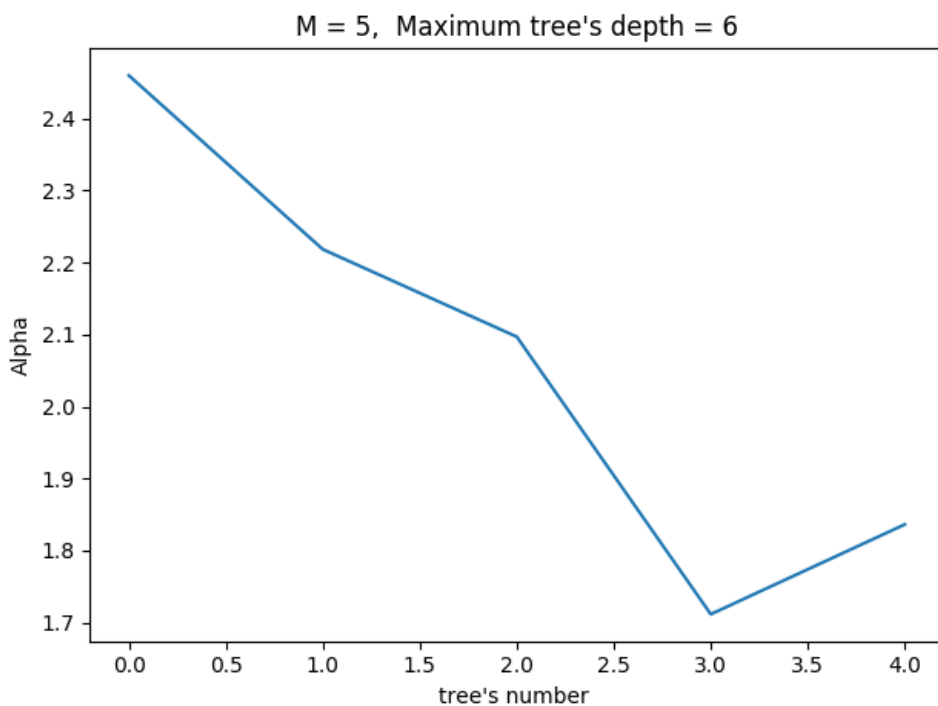
(۳)

حداکثر عمق را ۴ قرار دادیم.

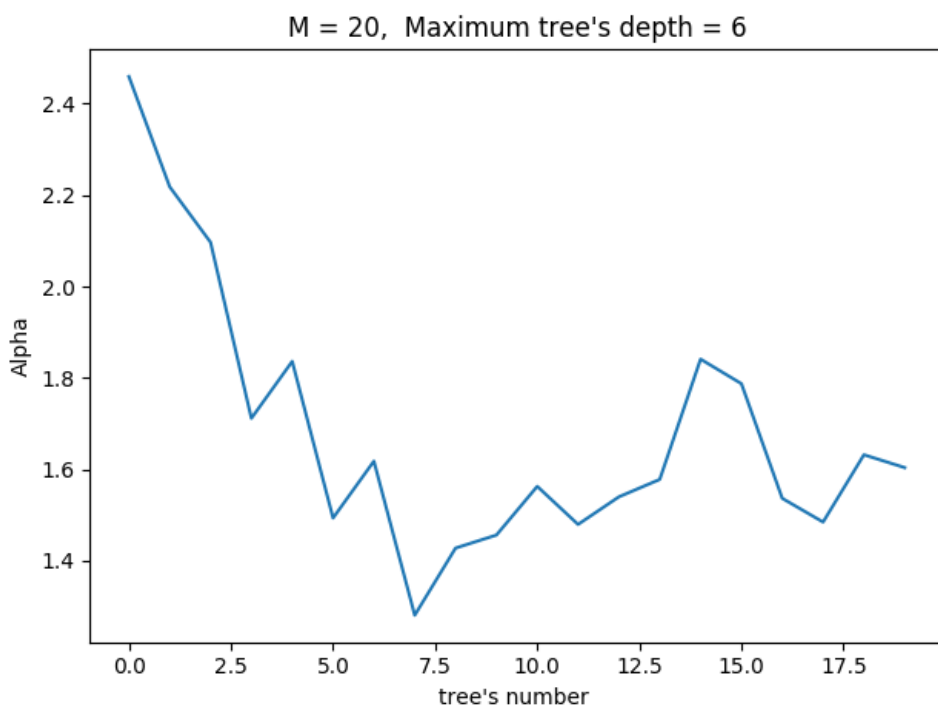
t= 0 Accuracy : 0.10380325993708893 epsilon: 0.19404777792606429
t= 1 Accuracy : 0.7454961395481842 epsilon: 0.3301711730241282
t= 2 Accuracy : 0.7454961395481842 epsilon: 0.2871709048394986
t= 3 Accuracy : 0.822419216471261 epsilon: 0.36481646578437793
t= 4 Accuracy : 0.8281384043465828 epsilon: 0.44043167811874917
t= 5 Accuracy : 0.8278524449528167 epsilon: 0.4633791854939084
t= 6 Accuracy : 0.8350014297969688 epsilon: 0.4606559222336713
t= 7 Accuracy : 0.8358593079782671 epsilon: 0.420334938690118
t= 8 Accuracy : 0.8564483843294253 epsilon: 0.41982492760067547
t= 9 Accuracy : 0.8730340291678582 epsilon: 0.43915567040128084

(۴)

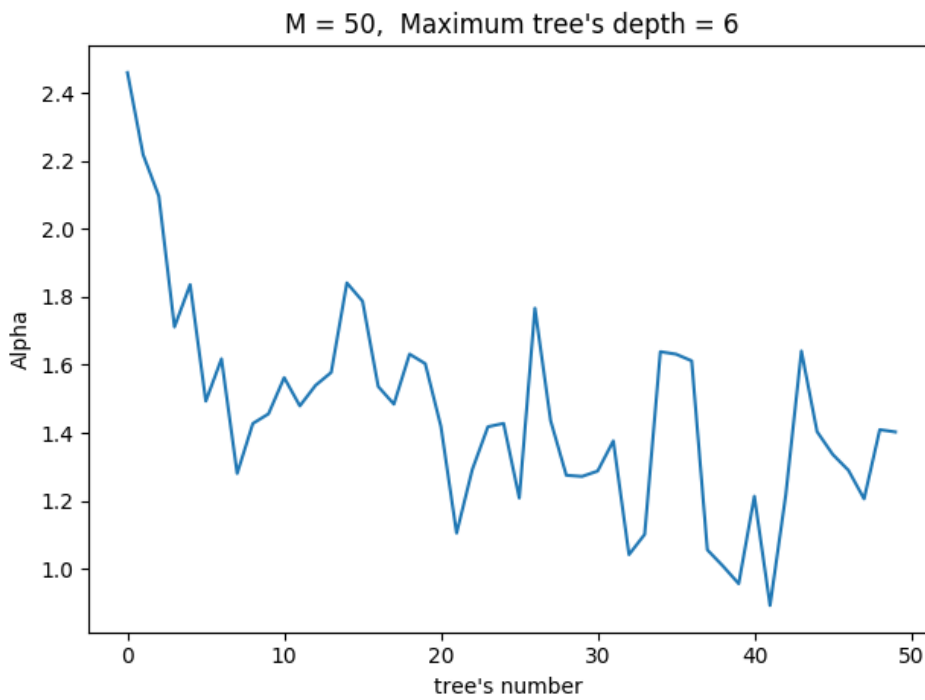
Number of decision trees = 5 Accuracy : 0.9436659994280812



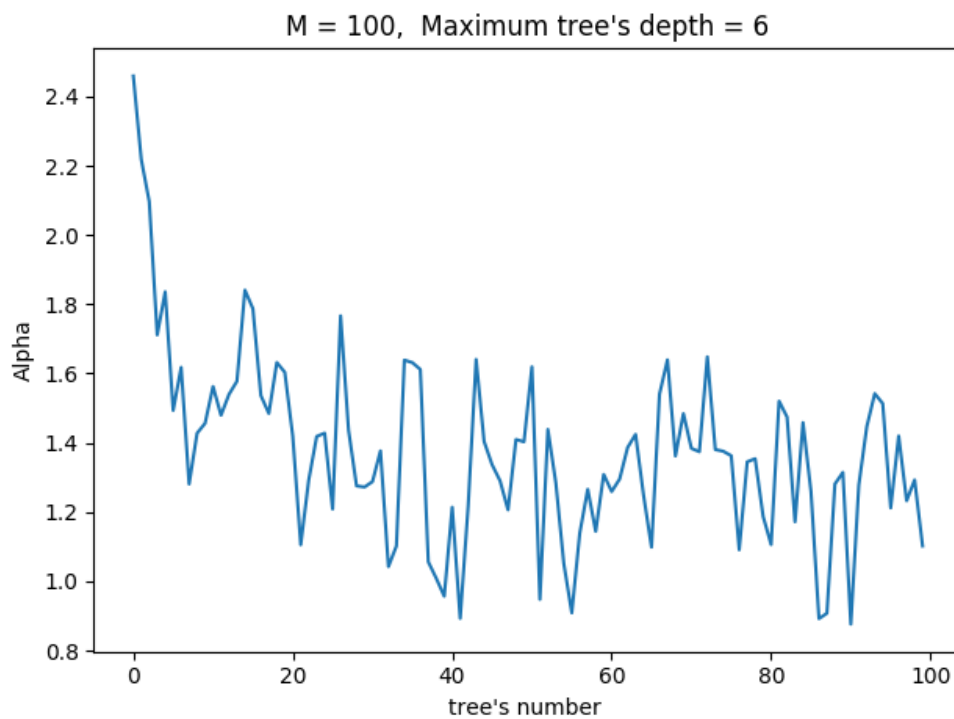
Number of decision trees = 20 Accuracy : 0.9639691163854732



Number of decision trees = 50 Accuracy : 0.9682585072919645



Number of decision trees = 100 Accuracy : 0.9694023448670289



با افزایش تعداد درخت ها، به طور کلی وزن دسته بند های جدید کم می شود یعنی دسته بند های اولیه بخش بزرگی از وظیفه ی ایجاد مدل مذکورمان را انجام می دهند و دسته بند های بعدی صرفاً برای افزایش دقت هستند. این می تواند یک مسئله دیگر را توجیه کند. اینکه وقتی دسته بند ها زیاد می شوند، نوسانات زیاد می شوند. در واقع در تعداد بالا دسته بند ها مدل به اندازه کافی تعمیم پیدا کرده و برای دقت بیشتر مدل شروع به اورفیت شدن می کند و نوسانات می کند.

سوال ۶)

Accuracy: 0.966256791535602

```
clf =  
xgb.XGBClassifier(objective='multi:softprop', lea  
rning_rate=0.4, max_depth=3, n_estimators=200)
```