

## İris veri seti için yapılmış olan çalışmanın sonucu

İris veri seti için öğrenme Algoritmalarını karşılaştırırken test için ayırdığım veri setinin değerini test yüzde değerlerini artırarak öğrenme algoritmasındaki değişiklikleri değerlendirmenin programdaki bir hatayı ortaya çıkardığını fark ettim. Topluluk öğrenme algoritması öğrenmeyi gerçekleştirirken algoritmalar arasındaki doğru sınıflandırma karar mekanizmasını öğrenirken algoritmanın hatalarını da öğreniyor.

test\_size parametresi ile test için ne kadar bir veri ayrılacak onu belirtiyoruz. 0.2 verinin % 20'sini test için ayır demek. Bu parametreyi atamakla aslında train\_size'ı da dolaylı olarak 0.8 yapmış oluyoruz. veri setinin %20'sini test, % 80'ini eğitim olarak ayırmış bulunuyoruz.

Tüm veri setinin %20'sini test için kullandığımızda Algoritmaların bizlere verdiği doğruluk oranı

|                          | %20 | %40 | %60 | %80 |
|--------------------------|-----|-----|-----|-----|
| k-NN                     | 100 | 93  | 95  | 88  |
| Rastgele Orman           | 100 | 93  | 95  | 91  |
| Lojistik Regrasyon       | 100 | 93  | 94  | 86  |
| Naive Bayes              | 96  | 93  | 94  | 93  |
| Destek vektör makineleri | 100 | 93  | 96  | 94  |
| Topluluk Öğrenme         | 100 | 91  | 95  | 90  |

%20 test verisi için tahmin doğruluk oranı

|                          | Setosa | Versicolor | Virginica |
|--------------------------|--------|------------|-----------|
| k-NN                     | 100    | 100        | 100       |
| Rastgele Orman           | 100    | 100        | 100       |
| Lojistik Regrasyon       | 100    | 100        | 100       |
| Naive Bayes              | 100    | 96         | 91        |
| Destek vektör makineleri | 100    | 100        | 100       |
| Topluluk Öğrenme         | 100    | 100        | 100       |

%40 test verisi için tahmin doğruluk oranı

|                          | Setosa | Versicolor | Virginica |
|--------------------------|--------|------------|-----------|
| k-NN                     | 100    | 92         | 90        |
| Rastgele Orman           | 100    | 92         | 90        |
| Lojistik Regrasyon       | 100    | 92         | 89        |
| Naive Bayes              | 100    | 92         | 89        |
| Destek vektör makineleri | 100    | 92         | 90        |
| Topluluk Öğrenme         | 100    | 90         | 87        |

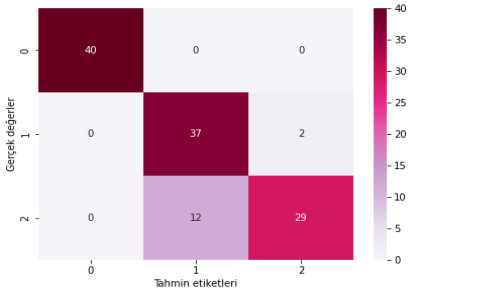
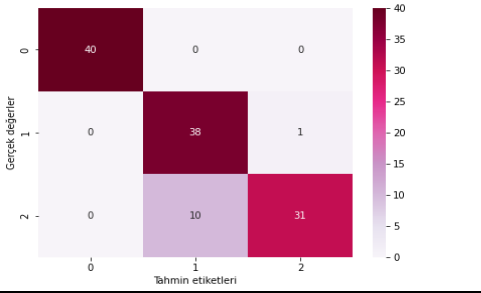
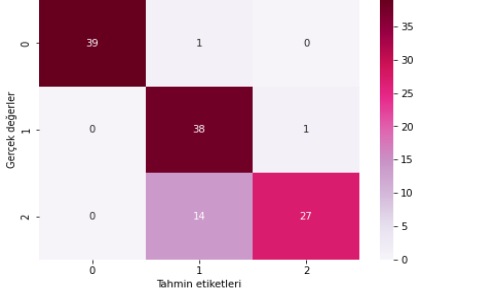
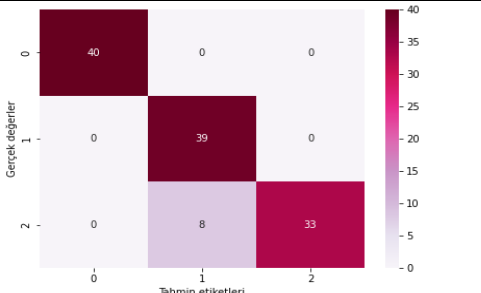
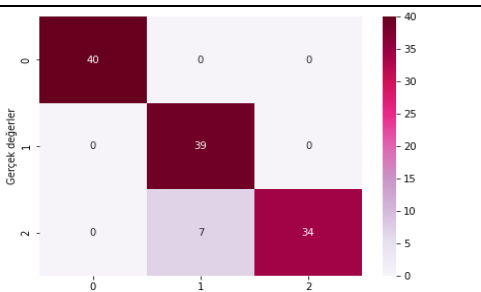
%60 test verisi için tahmin doğruluk oranı

|                          | Setosa | Versicolor | Virginica |
|--------------------------|--------|------------|-----------|
| k-NN                     | 100    | 94         | 93        |
| Rastgele Orman           | 100    | 94         | 93        |
| Lojistik Regrasyon       | 100    | 93         | 92        |
| Naive Bayes              | 100    | 93         | 91        |
| Destek vektör makineleri | 100    | 96         | 95        |
| Topluluk Öğrenme         | 100    | 96         | 95        |

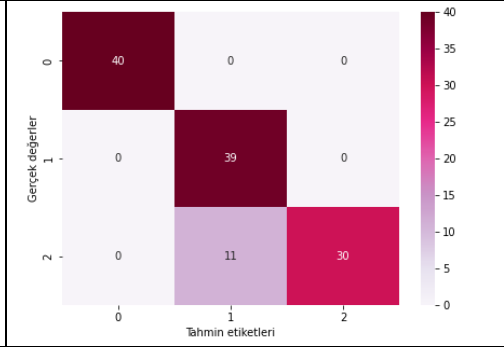
%80 test verisi için tahmin doğruluk oranı

|                          | Setosa | Versicolor | Virginica |
|--------------------------|--------|------------|-----------|
| k-NN                     | 100    | 84         | 81        |
| Rastgele Orman           | 100    | 87         | 85        |
| Lojistik Regrasyon       | 99     | 83         | 78        |
| Naive Bayes              | 100    | 91         | 89        |
| Destek vektör makineleri | 100    | 92         | 91        |
| Topluluk Öğrenme         | 100    | 88         | 85        |

## %80 test verisi için Sınıflandırma hataları

| k-NN                     |  <p>Confusion matrix for k-NN classifier. The y-axis is labeled 'Gerçek değerler' (Actual values) and the x-axis is labeled 'Tahmin etiketleri' (Predicted labels). The matrix shows counts for three classes (0, 1, 2). A color scale on the right ranges from 0 to 40.</p> <table><tr><th>Gerçek \ Tahmin</th><th>0</th><th>1</th><th>2</th></tr><tr><th>0</th><td>40</td><td>0</td><td>0</td></tr><tr><th>1</th><td>0</td><td>37</td><td>2</td></tr><tr><th>2</th><td>0</td><td>12</td><td>29</td></tr></table>  | Gerçek \ Tahmin | 0  | 1 | 2 | 0 | 40 | 0 | 0 | 1 | 0 | 37 | 2 | 2 | 0 | 12 | 29 |
|--------------------------|---|-----------------|----|---|---|---|----|---|---|---|---|----|---|---|---|----|----|
| Gerçek \ Tahmin          | 0   | 1               | 2  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 0                        | 40  | 0               | 0  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 1                        | 0   | 37              | 2  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 2                        | 0   | 12              | 29 |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| Rastgele Orman           |  <p>Confusion matrix for Rastgele Orman (Random Forest) classifier. The y-axis is labeled 'Gerçek değerler' (Actual values) and the x-axis is labeled 'Tahmin etiketleri' (Predicted labels). The matrix shows counts for three classes (0, 1, 2). A color scale on the right ranges from 0 to 40.</p> <table><tr><th>Gerçek \ Tahmin</th><th>0</th><th>1</th><th>2</th></tr><tr><th>0</th><td>40</td><td>0</td><td>0</td></tr><tr><th>1</th><td>0</td><td>38</td><td>1</td></tr><tr><th>2</th><td>0</td><td>10</td><td>31</td></tr></table>                      | Gerçek \ Tahmin | 0  | 1 | 2 | 0 | 40 | 0 | 0 | 1 | 0 | 38 | 1 | 2 | 0 | 10 | 31 |
| Gerçek \ Tahmin          | 0   | 1               | 2  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 0                        | 40  | 0               | 0  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 1                        | 0   | 38              | 1  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 2                        | 0   | 10              | 31 |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| Lojistik Regrasyon       |  <p>Confusion matrix for Lojistik Regrasyon (Logistic Regression) classifier. The y-axis is labeled 'Gerçek değerler' (Actual values) and the x-axis is labeled 'Tahmin etiketleri' (Predicted labels). The matrix shows counts for three classes (0, 1, 2). A color scale on the right ranges from 0 to 35.</p> <table><tr><th>Gerçek \ Tahmin</th><th>0</th><th>1</th><th>2</th></tr><tr><th>0</th><td>39</td><td>1</td><td>0</td></tr><tr><th>1</th><td>0</td><td>38</td><td>1</td></tr><tr><th>2</th><td>0</td><td>14</td><td>27</td></tr></table>           | Gerçek \ Tahmin | 0  | 1 | 2 | 0 | 39 | 1 | 0 | 1 | 0 | 38 | 1 | 2 | 0 | 14 | 27 |
| Gerçek \ Tahmin          | 0   | 1               | 2  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 0                        | 39  | 1               | 0  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 1                        | 0   | 38              | 1  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 2                        | 0   | 14              | 27 |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| Naive Bayes              |  <p>Confusion matrix for Naive Bayes classifier. The y-axis is labeled 'Gerçek değerler' (Actual values) and the x-axis is labeled 'Tahmin etiketleri' (Predicted labels). The matrix shows counts for three classes (0, 1, 2). A color scale on the right ranges from 0 to 40.</p> <table><tr><th>Gerçek \ Tahmin</th><th>0</th><th>1</th><th>2</th></tr><tr><th>0</th><td>40</td><td>0</td><td>0</td></tr><tr><th>1</th><td>0</td><td>39</td><td>0</td></tr><tr><th>2</th><td>0</td><td>8</td><td>33</td></tr></table>  | Gerçek \ Tahmin | 0  | 1 | 2 | 0 | 40 | 0 | 0 | 1 | 0 | 39 | 0 | 2 | 0 | 8  | 33 |
| Gerçek \ Tahmin          | 0   | 1               | 2  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 0                        | 40  | 0               | 0  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 1                        | 0   | 39              | 0  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 2                        | 0   | 8               | 33 |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| Destek vektör makineleri |  <p>Confusion matrix for Destek vektör makineleri (Support Vector Machines) classifier. The y-axis is labeled 'Gerçek değerler' (Actual values) and the x-axis is labeled 'Tahmin etiketleri' (Predicted labels). The matrix shows counts for three classes (0, 1, 2). A color scale on the right ranges from 0 to 40.</p> <table><tr><th>Gerçek \ Tahmin</th><th>0</th><th>1</th><th>2</th></tr><tr><th>0</th><td>40</td><td>0</td><td>0</td></tr><tr><th>1</th><td>0</td><td>39</td><td>0</td></tr><tr><th>2</th><td>0</td><td>7</td><td>34</td></tr></table> | Gerçek \ Tahmin | 0  | 1 | 2 | 0 | 40 | 0 | 0 | 1 | 0 | 39 | 0 | 2 | 0 | 7  | 34 |
| Gerçek \ Tahmin          | 0   | 1               | 2  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 0                        | 40  | 0               | 0  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 1                        | 0   | 39              | 0  |   |   |   |    |   |   |   |   |    |   |   |   |    |    |
| 2                        | 0   | 7               | 34 |   |   |   |    |   |   |   |   |    |   |   |   |    |    |

# Topluluk Öğrenme



%80 test verisi için Sınıflandırma algoritmalarına baktığımızda 120 adet veri üzerinde çalışılmış bu bağlamda sonuçlara baktığımızda ;

**setosa çiçek türü için** Lojistik Regrasyon algoritması 1 çiçeği yanlış sınıflandırarak versicolor olduğunu söylemiştir. Bunun dışında setosa çiçeği için diğer algoritmalar yanlış sınıflandırma yapmamış doğru sınıflandırma yapmıştır. Bu 4 algoritma setosa çiçeği için doğru değer döndürürken 1 algoritma setosa çiçeği için yanlış bir sınıflandırma yapmıştır Topluluk öğrenme algoritması da buna karşılık öğrenmeyi gerçekleştirirken 4 algoritma gibi hareket ederek doğru bir sınıflandırma yapmış ve yanlış sonuca varmamıştır.

**Versicolor çiçek türü için** Lojistik Regrasyon algoritması 1 çiçeği yanlış sınıflandırarak virginica olduğunu söylemiş, Rastgele Orman 1 çiçeği yanlış sınıflandırarak virginica olduğunu söylemiş k-NN algoritması 2 çiçeği yanlış sınıflandırarak virginica olduğunu söylemiş ve Destek vektör makineleri ile Naive Bayes versicolor çiçek türü için yanlış bir sınıflandırma yapmamıştır. Topluluk öğrenme algoritması da buna karşılık öğrenmeyi gerçekleştirirken doğru bir sınıflandırma yapmış ve yanlış sonuca varmamıştır.

**virginica çiçek türü için** Lojistik Regrasyon algoritması 14 çiçeği yanlış sınıflandırarak versicolor olduğunu, Rastgele Orman 10 çiçeği yanlış sınıflandırarak versicolor söylemiş , k-NN algoritması 12 çiçeği yanlış sınıflandırarak versicolor olduğunu, Destek vektör makineleri 7 çiçeği yanlış sınıflandırarak versicolor olduğunu , Naive Bayes 8 çiçeği yanlış sınıflandırarak versicolor olduğunu söylemiştir. Burada virginica çiçek türü sınıflandırmasında diğer çiçek türlerine göre daha fazla hata yapıldığı gözlenmiştir en düşük hatayı yapan destek vektör makineleri algoritması 7 hata yaparken en fazla hata yapan algoritma Lojistik Regrasyon 14 hata yapmıştır buna karşılık topluluk öğrenme algoritmamız 11 çiçeği yanlış sınıflandırarak versicolor olduğunu söyleyerek yüksek denilebilecek bir hata yapmıştır bunun nedeni topluluk öğrenme yöntemi uygulanırken algoritmaların çoğunluk ile verdiği doğru kararları alırken aynı zamanda algoritmanın çoğunluk ile birlikte verdiği yanlış kararların da algoritmanın kararını etkileyerek yanlış karar vermesine neden olmaktadır.