T.C

KIRIKKALE ÜNİVESİTESİ

MÜHENDİSLİK FAKÜLTESİ

BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

MAKİNE ÖĞRENMESİ DERSİ VİZE ÖDEVİ

Hazırlayan Öğrenci

130255065

NECDET YAPICI

Ders Sorumlusu

Dr. Arş. Gör. Enes Ayan

Kırıkkale 2020

Sorular

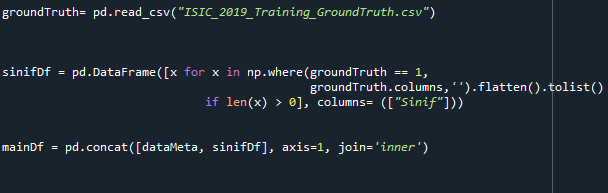
1- Analiz için ödevde belirtilen gerekli özellikler haricinde gerekli olmayan özellikleri silinmesi için yazdığınız kodu yazınız.



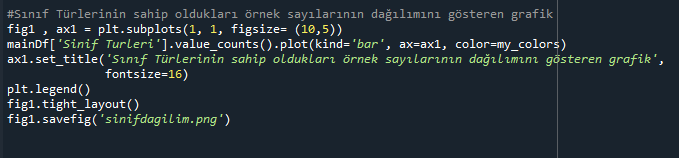
2- Tamamı boş olan ya da içinde herhangi bir boş veri içeren örneğin silinmesi için gereken kodu yazınız.

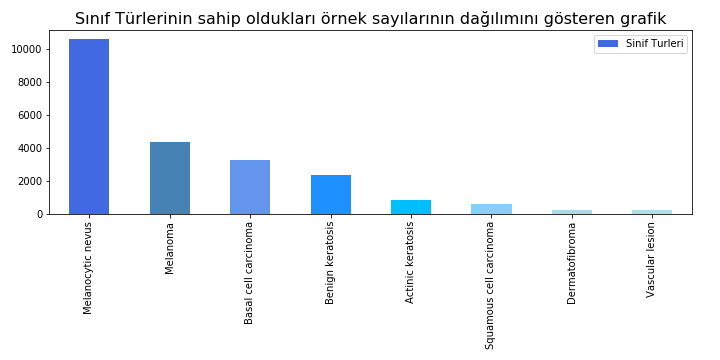


3- Meta-data dosyası ile etiketlerin olduğu (ground-truth) dosyayı birleştirin. Bu birleştirme sonucunda her bir örneğin karşısında ait olduğu sınıf yazmalıdır. (ISIC\_0000000, 55, anterior torso, female, NV). Bunu gerçekleştiren kodu yazınız.

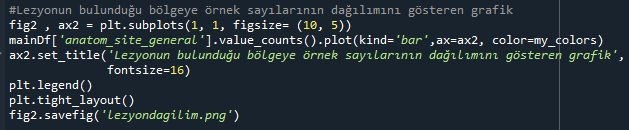


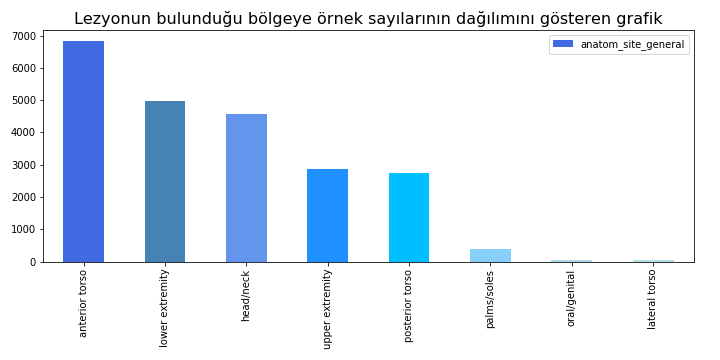
4- Verideki lezyon türlerinin dağılımını gösteren grafiği çizdiren kodu yazınız. Grafiğin görsel sonucunu ekleyiniz.



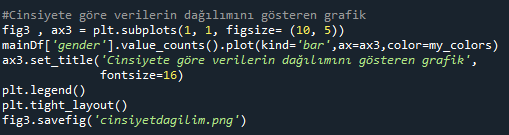


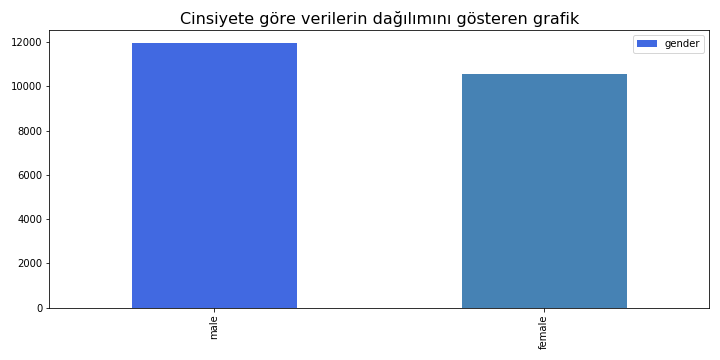
5- Verideki lezyonun bulunduğu bölgeye göre dağılımını gösteren grafiği çizdiren kodu yazınız. Grafiğin görsel sonucunu ekleyiniz.



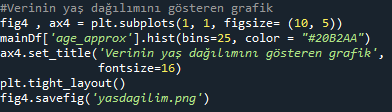


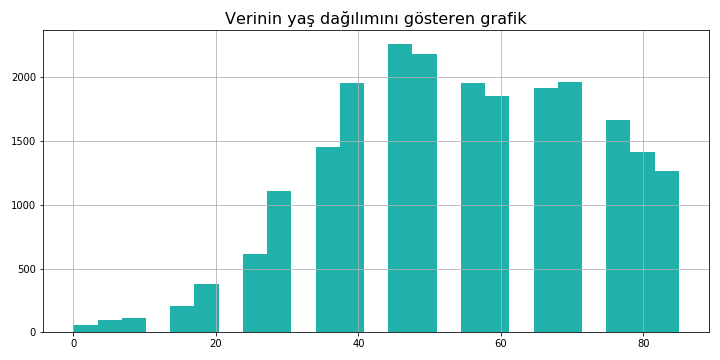
6- Verideki cinsiyet dağılımını gösteren grafiği çizdiren kodu yazınız. Grafiğin görsel sonucunu ekleyiniz.



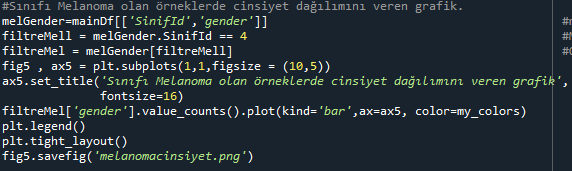


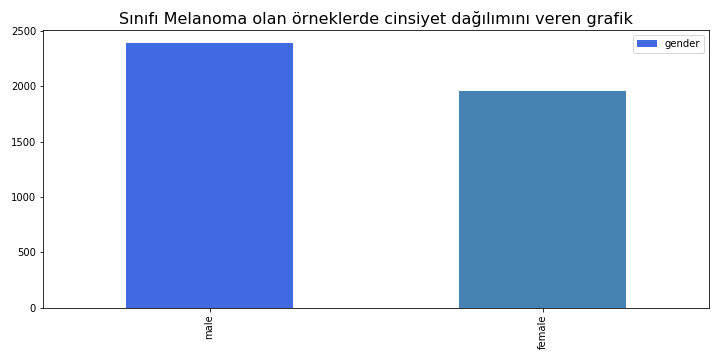
7- Verideki yaş dağılımını gösteren grafiği çizdiren kodu yazınız. Grafiğin görsel sonucunu ekleyiniz.



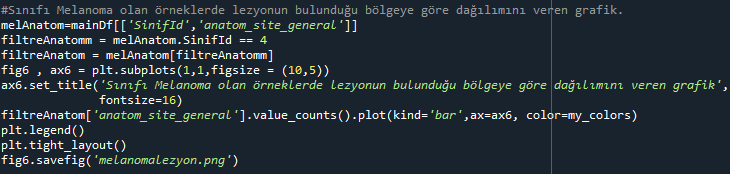


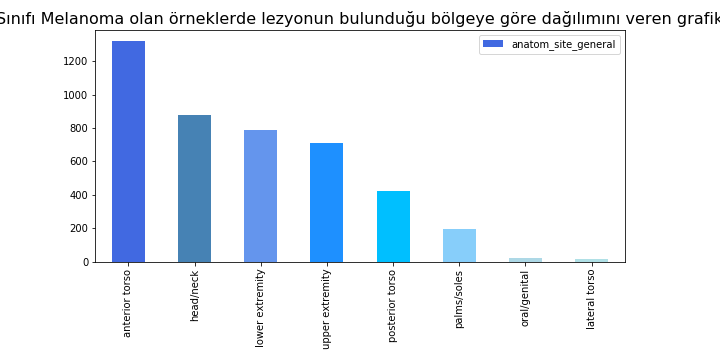
8- Verideki sınıfı MEL (melanoma) olup cinsiyet dağılımını gösteren grafiği çizdiren kodu yazınız. Grafiğin görsel sonucunu ekleyiniz.





9- Verideki sınıfı MEL (melanoma) olup lezyonun bulunduğu bölgeye göre dağılımını gösteren grafiği çizdiren kodu yazınız. Grafiğin görsel sonucunu ekleyiniz.





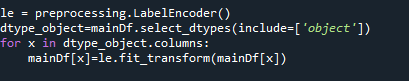
10- Veriyi %80 eğitim %20 test olacak şekilde bölen kodu yazınız.



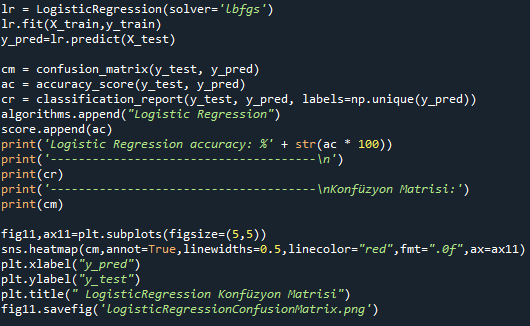
11- Veriyi ölçekleyen kodu yazınız (Bir özellik için yazmanız yeterli örneğin yaş)

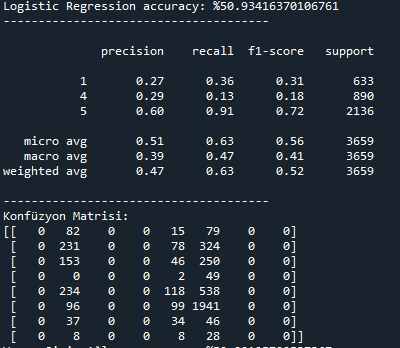


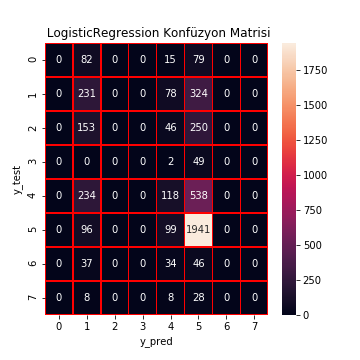
12- Verideki kategorik değerleri nümerik (sayısal) hale getiren kodu yazınız.



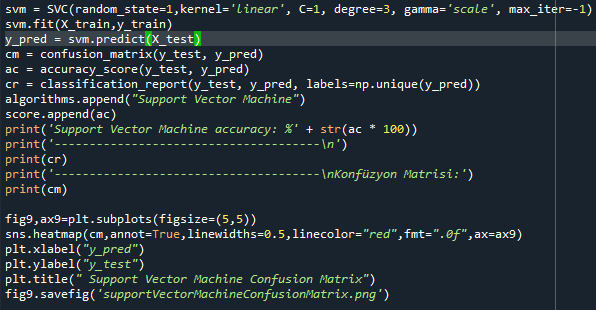
13- Logistic Regresyon ile veriyi sınıflandırınız ve accuracy, presicion, recall, f1 score ve confusion matrix sonuçlarınızı paylaşınız. (kodu da yazın)

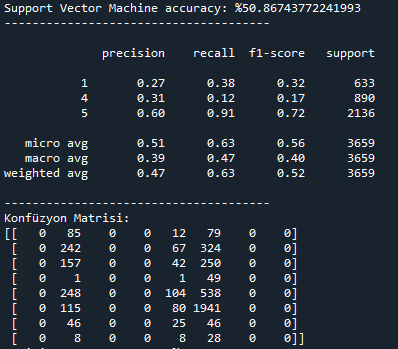


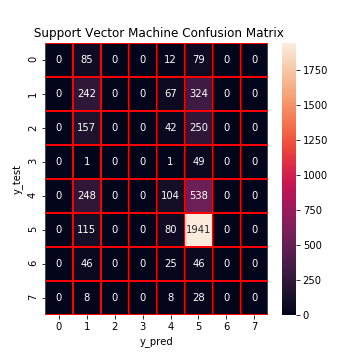




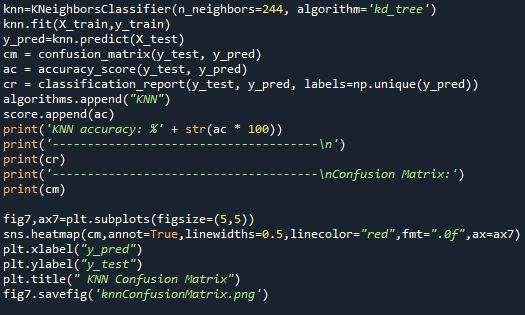
14- SVM ile veriyi sınıflandırınız ve accuracy, presicion, recall, f1 score ve confusion matrix sonuçlarınızı paylaşınız. (kodu da yazın)

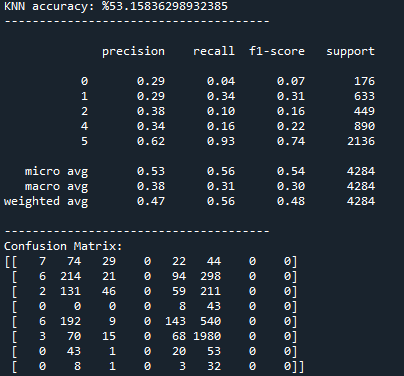


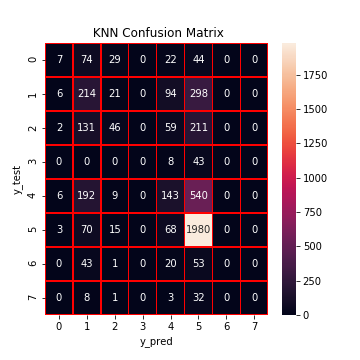




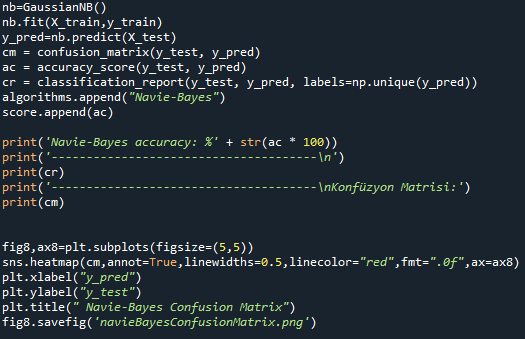
15- K-NN ile veriyi sınıflandırınız ve accuracy, presicion, recall, f1 score ve confusion matrix sonuçlarınızı paylaşınız. (kodu da yazın)

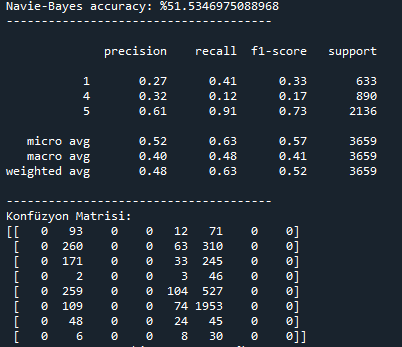


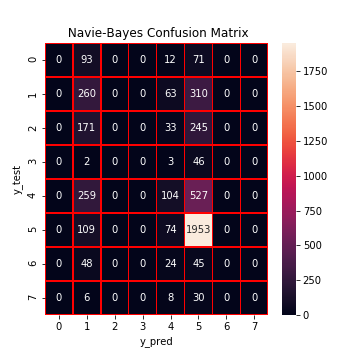




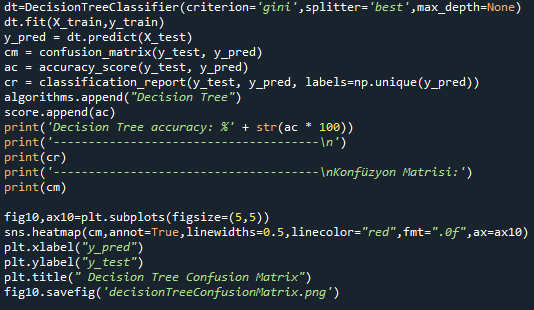
16- Naive Bayes ile veriyi sınıflandırınız ve accuracy, presicion, recall, f1 score ve confusion matrix sonuçlarınızı paylaşınız. (kodu da yazın)

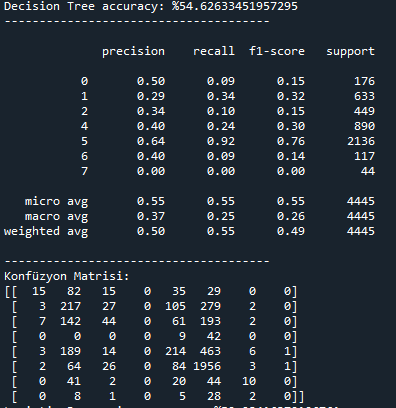


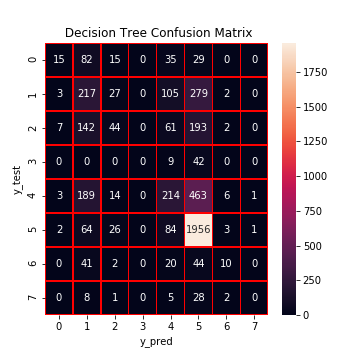




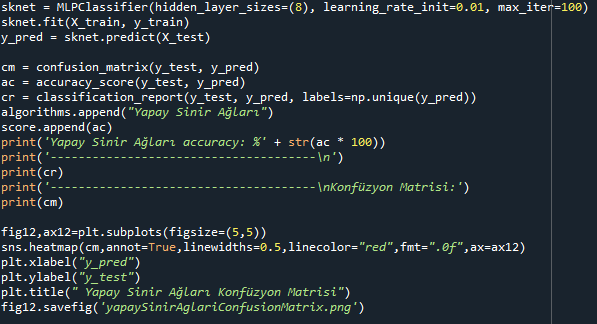
17- Karar ağaçları ile veriyi sınıflandırınız ve accuracy, presicion, recall, f1 score ve confusion matrix sonuçlarınızı paylaşınız. (kodu da yazın)

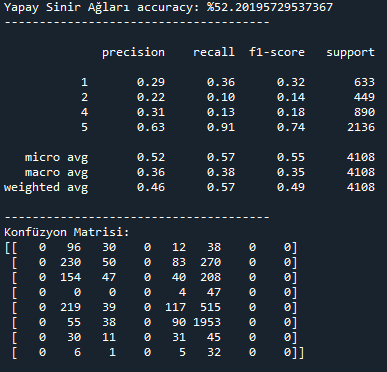


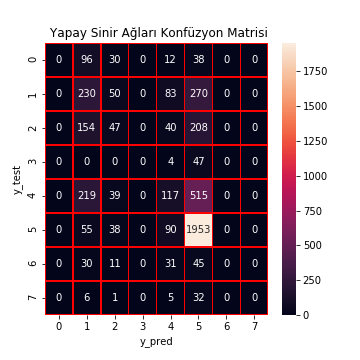




18- Yapay sinir ağları ile veriyi sınıflandırınız ve accuracy, presicion, recall, f1 score ve confusion matrix sonuçlarınızı paylaşınız. (kodu da yazın).

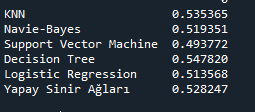






19- Sonuçları bireysel olarak yorumlayınız sonuçlarınız hangi yöntemde iyi çıktı sizce sebepleri veya sonuçlarınız neden kötü çıktı bunun sebepleri nelerdir?

**Algoritmaların gösterdiği başarı oranları**



Genel araştırmadan sonra en yüksek oranı Karar ağacı veriyor. En iyi sonuç diyebiliriz ama ulaştığımız veriler yeterli oranda değil. Gerek benim kullandığım bilgisayardan gerekse elimizdeki verilerden kaynaklanabilir bu dengesiz sonuçlar. Başka bir csv dosyası kullansaydık farklı sonuç alabilirdik. Kolon sayısı, satır sayısı sonuçları oldukça etkileyebilir. Ama bu durumda en iyisinin Karar Ağacı olduğunu söylemek zorundayız.

Not: Her bir sorunun altına çıktıları ve o çıktıyı elde ettiğiniz kod buluğunu yazınız. Haricinde tüm kodun yer aldığı Python dosyasını bu rapor ile birlikte rar yapıp gönderin gönderin. Rar’ı kendi isiminiz ile isimlendirin. Vize ödevini en geç 22.05.2020 tarihine kadar [enesayan@yandex.com](mailto:enesayan@yandex.com) mail adresine gönderiniz.