

Data Platform  
Information Technologies

Sayın AisanKheiri

2024-10-11 10:56:40 saatinde SVM Modeli üzerinde yapılan çalışmaya özel olarak hazırlanan rapor sizin için sunulmuştur. Data platform şirketimizi tercih ettiğiniz için de ayrıca teşekkür ederiz. Bu raporda, model eğitiminin başarıyla tamamlanmasının ardından oluşturulan grafikler detaylı bir şekilde incelenmiş ve analiz edilmiştir.

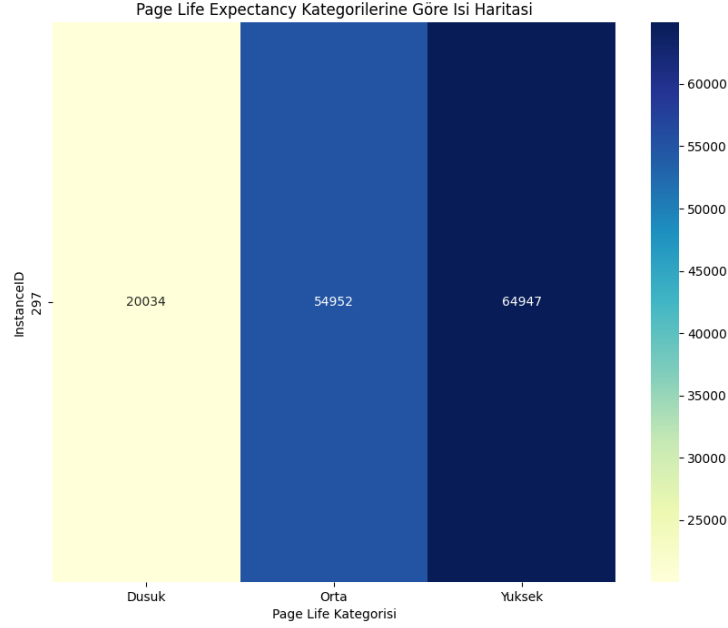
Seçtiğiniz SVM Modeli parametreleri arasında Kernel:sigmoid , C:2 , sutun1:2 , sutun2:7 , test\_size=0.2 ,random\_state=42 ve Hedef Etiketi  $y = df['page\ life\ expectancy'] > 300$  değerleri bulunmaktadır, bu parametrelerin sonuçları da raporda detaylı bir şekilde açıklanmıştır.

Focus on your business  
We'll manage your databases

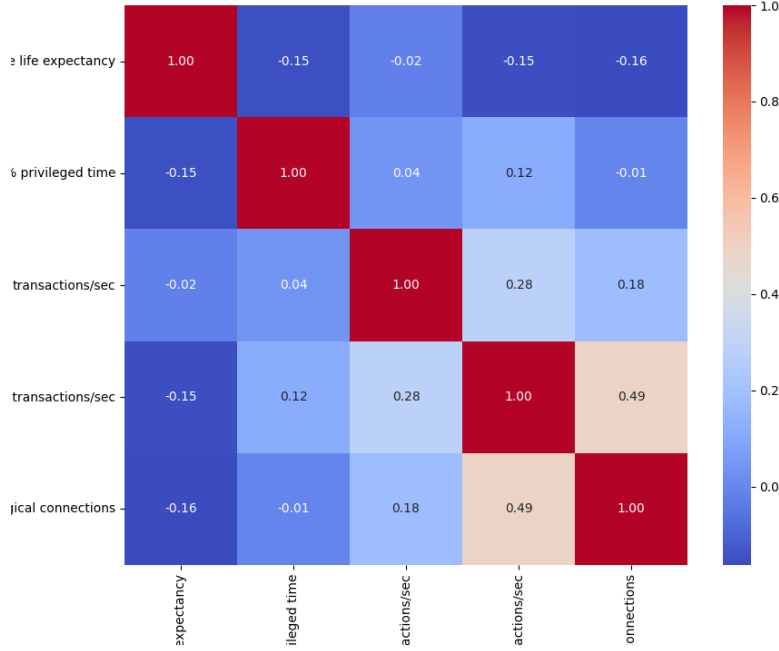
 info@dataplatfrom.com.tr

 Data Platform Bilgi Sistemleri A.Ş. Bağlarbaşı Mah.  
Irmak Sok. No: 15 D:2 K:2 Maltepe/İstanbul

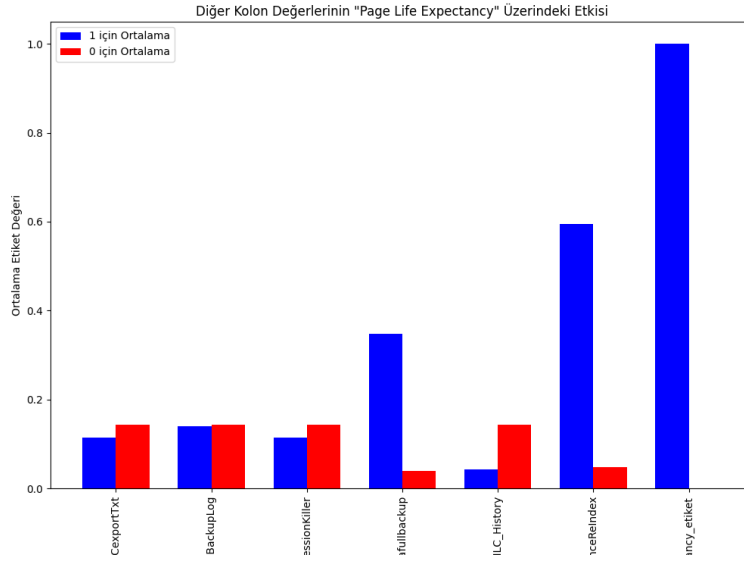
 0532 320 22 39



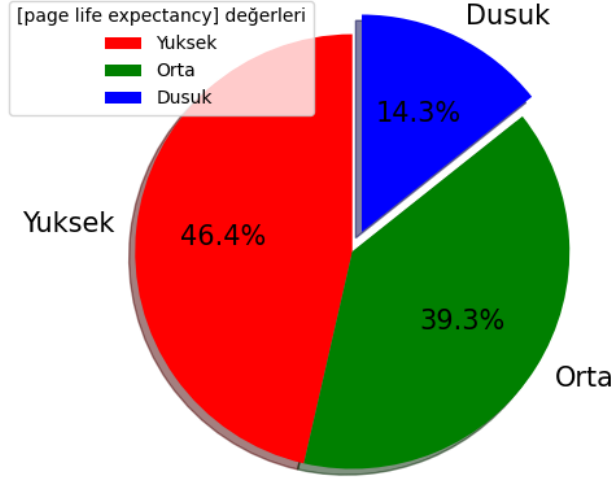
Chat-GPT tarafından üretilen analiz: "Veri setindeki 'Düşük', 'Orta' ve 'Yüksek' kategorileri arasında gözlemlenen ilişkiler incelendiğinde, genel eğilimler şu şekilde belirlendi: 'Orta' kategorideki sayılar, 'Düşük' kategoriye göre belirgin bir şekilde daha yüksekken, 'Yüksek' kategorideki sayılar da 'Orta' kategoriye kıyasla biraz daha yüksektir. Bu durum, sayı değerlerinin kategorilere göre dağılımının yaklaşık olarak doğrusal bir şekilde artış gösterdiğini göstermektedir. Özellikle 'Düşük' ve 'Orta' kategorileri arasındaki farkın belirgin olduğu gözlemlenmektedir. Bu analiz, veri setindeki sayısal ilişkilerin kategorilere göre belirgin örüntüler sergilediğini ortaya koymaktadır."



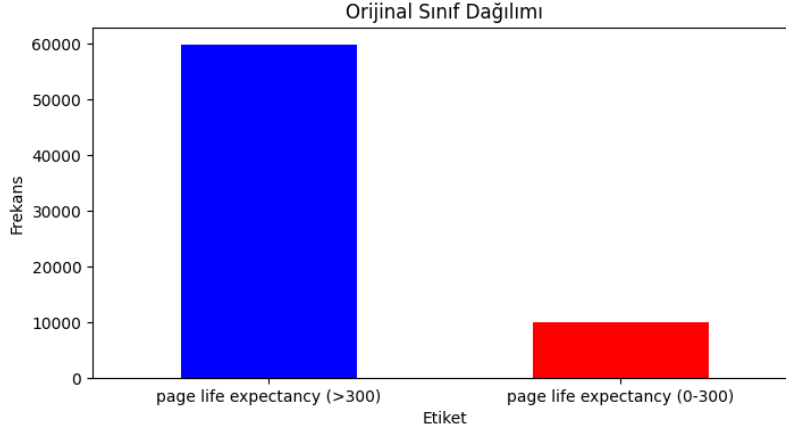
Chat-GPT tarafından üretilen analiz: Bu veri setindeki page life expectancy ile logical connections arasında güçlü bir negatif korelasyon görülmektedir (-0.162372). Yani, birinin değeri artarken diğerinin azaldığı bir ilişki bulunmaktadır. Aynı şekilde, write transactions/sec ile logical connections arasında ise güçlü bir pozitif korelasyon gözlemlenmektedir (0.490174). Yani, bu iki değişken arasında birisinin değeri artarken diğerinin de arttığı bir ilişki vardır. Diğer sütunlar arasındaki korelasyonlar ise daha zayıf ya da neredeyse hiçbir ilişki göstermemektedir.



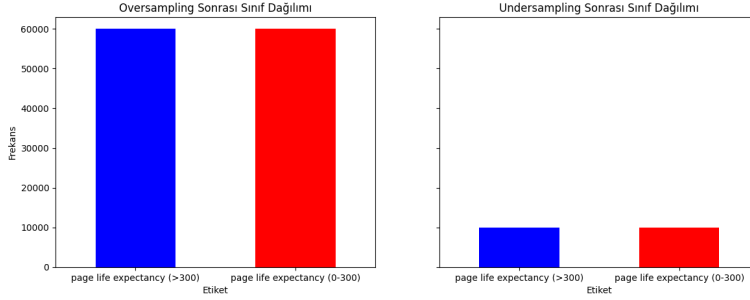
Chat-GPT tarafından üretilen analiz: 'Veri setindeki belirli sütunların 'page life expectancy' sütunu üzerindeki etkileri incelendi. 'dbaHCDOCexportTxt', 'dbaSessionKiller' ve 'OTS\_Instrument\_OHLC\_History' sütunlarının ortalama etkileri düşük seviyededir (%0,11 - %0,14 arasında). 'dbaBackupLog' sütununun etkisi ise biraz daha yüksek seviyededir (%0,14). 'dbafullbackup' sütununun 'page life expectancy' üzerindeki etkisi ise oldukça yüksektir (%0,35). 'dbaMaintenanceReIndex' sütunu ise en yüksek etkiye sahip sütundur (%0,59). Son olarak, 'page\_life\_expectancy\_etiket' sütununun etkisi diğer sütunlara göre en yüksektir ve %1'dir. Bu değerler, 'page life expectancy' üzerinde farklı sütunların ne kadar etkili olduğunu göstermektedir.



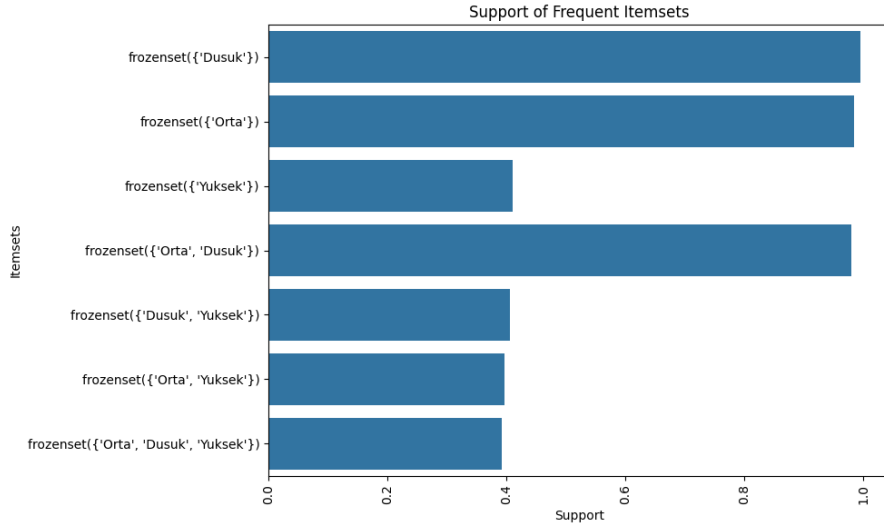
Chat-GPT tarafından üretilen analiz: "Page life expectancy değerleri yüksek olan gözlemlerin sayısı 32,474, ortalama olan gözlemlerin sayısı 27,476 ve düşük olan gözlemlerin sayısı 10,017 olarak belirlenmiştir. Bu veri setine göre genel olarak sayıca en fazla gözlem yüksek seviyede(page life expectancy yüksek) iken sayıca en az gözlem düşük seviyede(page life expectancy düşük) bulunmaktadır."



Chat-GPT tarafından üretilen analiz: 'Sayın kullanıcının sunmuş olduğu veri seti incelendiğinde, page life expectancy değerlerine göre yapılan etiketleme sonucunda, 0 ve 1 olarak iki farklı grupta toplamda 69967 veri bulunduğu belirlenmiştir. Bu durum, page life expectancy değerlerinin genel olarak iki farklı kategoriye ait olduğunu göstermektedir. Veri setindeki dağılım incelendiğinde, 0 etiketine sahip veri sayısı 59950 iken, 1 etiketine sahip veri sayısı ise 10017 olarak belirlenmiştir. Bu analiz doğrultusunda, page life expectancy değerlerine göre yapılan etiketleme sonucunda, belirli bir eşik değer aralığında ayırım yapıldığı ve verilerin iki farklı grup olarak sınıflandırıldığı yorumu yapılabilir.'



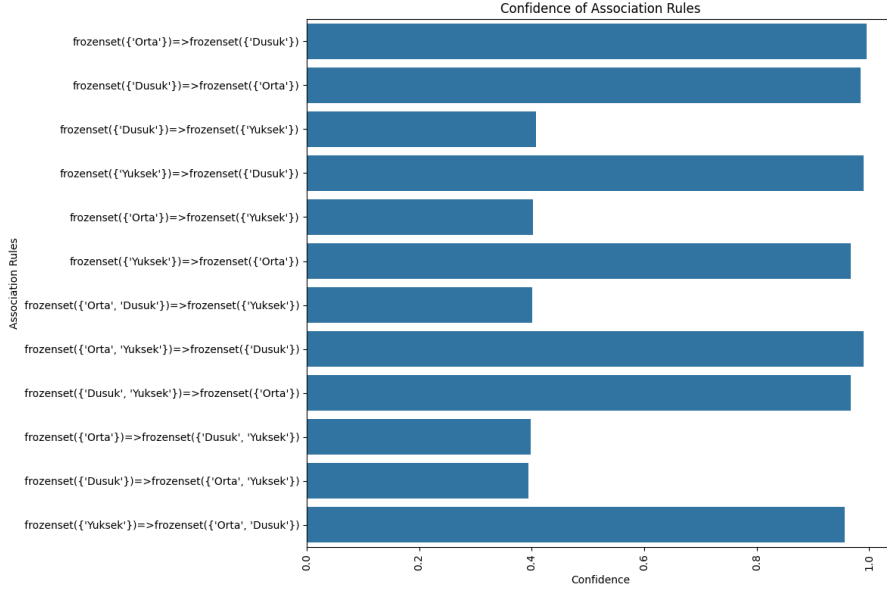
Chat-GPT tarafından üretilen analiz: 'Bu veriler, sınıf dengesizliği sorununu ele almak için örnekleme tekniklerinin (sampling) etkili bir şekilde nasıl kullanılabileceğini gösteriyor. Oversampling yapıldıktan sonra etiket 0 ve etiket 1 sınıflarının eşit sayıda örneğe sahip olduğunu görüyoruz. Bu, modelin daha dengeli bir şekilde eğitilmesine ve daha iyi performans göstermesine yardımcı olabilir. Ayrıca undersampling yapıldıktan sonra da benzer bir dengenin sağlandığı gözlemleniyor. Bu teknikler, sınıf dengesizliği sorununu ele almak için başarılı bir şekilde uygulanmış gibi görünmektedir.'



Chat-GPT tarafından üretilen analiz: "FP-Growth algoritması kullanılarak veri setindeki öge kümelerinin sık görülen kombinasyonları incelendi. En sık rastlanan kombinasyonlar şunlardır:

- Düşük sıklıkta ve %99,55 destek değeri ile (Düşük) kategorisi,
- Orta sıklıkta ve %98,48 destek değeri ile (Orta) kategorisi,
- Yüksek sıklıkta ve %41,03 destek değeri ile (Yüksek) kategorisi,
- Orta ve Düşük sıklıkların birlikte görüldüğü kombinasyon, %98,03 destek değeriyle,
- Düşük ve Yüksek sıklıkların birlikte görüldüğü kombinasyon, %40,62 destek değeriyle,
- Orta ve Yüksek sıklıkların birlikte görüldüğü kombinasyon, %39,67 destek değeriyle,
- Orta, Düşük ve Yüksek sıklıkların birlikte görüldüğü kombinasyon, %39,26 destek değeriyle elde edilmiştir."



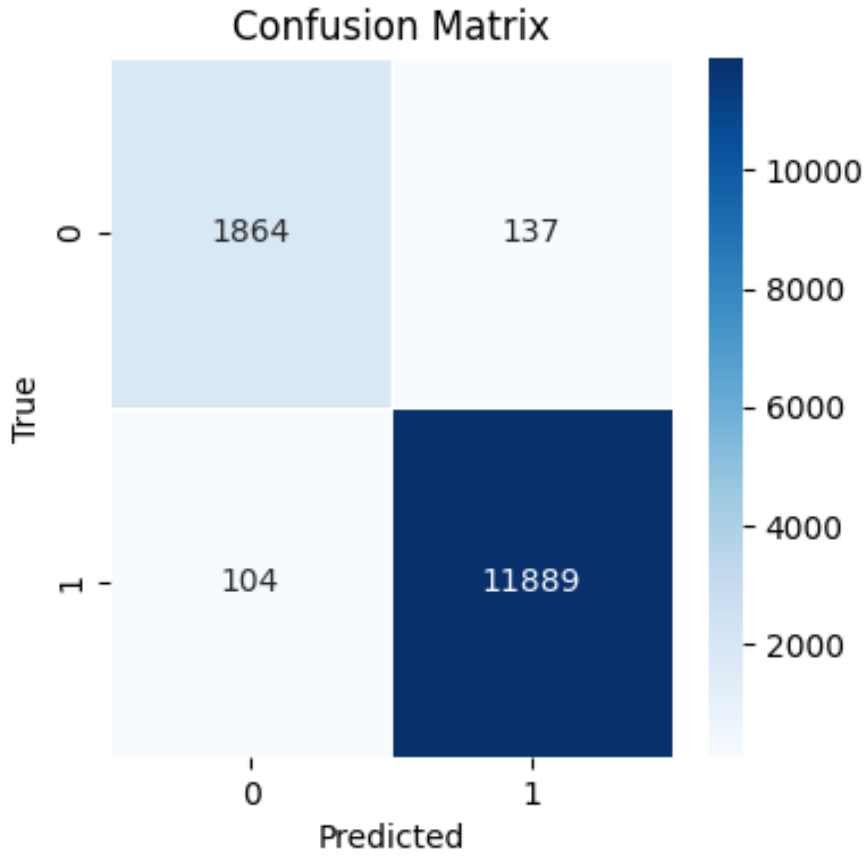


### Chat-GPT tarafından üretilen analiz: "Senin analizin"

Bu veri seti üzerinde yaptığım analize göre, öncelikle orta seviyede olan ürünlerin genellikle düşük seviyede ürünlerle birlikte alındığını gösteriyor. Diğer yandan, düşük seviyede olan ürünlerin yüksek seviyedeki ürünlerle birlikte alındığında ise alışveriş yapma olasılığının düşük olduğu bir ilişki olduğu gözlemlenmiştir.

Orta ve düşük seviyede olan ürünlerin birlikte alındığında genellikle yüksek seviyedeki ürünler de alınmaktadır. Ancak yüksek seviyedeki ürünler ile düşük seviyedeki ürünler bir arada alındığında orta seviyedeki ürünlerin alınma olasılığı daha düşüktür.

Bu analizler, alışveriş sepetindeki ürünler arasındaki ilişkileri anlamak ve müşteri davranışlarını daha iyi anlamak için kullanılabilir.



Çalışan algoritma: SVM

Chat-GPT tarafından üretilen analiz: Bu algoritmanın performansı oldukça yüksektir. Confusion Matrix'e göre, sınıflandırma modeli, 1864 doğru pozitif ve 11889 doğru negatif tahmin yapmıştır. Recall değeri (duyarlılık) 0.991 olarak oldukça yüksektir, yani modelin gerçek pozitifleri doğru tahminleme başarısı yüksektir. Accuracy değeri ise 0.982 olarak oldukça yüksektir, yani modelin doğru sınıflandırma oranı yüksektir. F1 Score'un yüksek olması da modelin hem precision'a hem de recall'a dengeli şekilde önem verdiğini gösterir. Sonuç olarak, SVM algoritması oldukça güçlü ve verimli bir şekilde çalışmıştır.