



Data Platform
Information Technologies

Sayın AisanKheiri

2024-09-30 16:08:18 saatinde KNN 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.

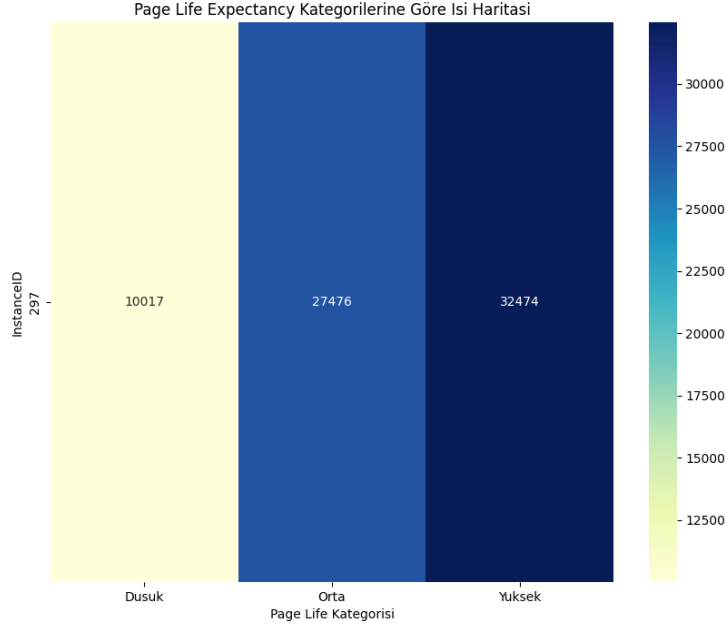
Seçtiğiniz KNN Modeli parametreleri arasında `n_neighbors:None` , `metric:None` ve `weights:None` ,`sutun1:2` , `sutun2:7` , `test_size=0.2` ,`random_state=42` ve Hedef Etiketi `y = df['page life expectancy']` > 300 değerleri bulunmaktadır, bu parametrelerin sonuçları da raporda detaylı bir şekilde açıklanmıştır.

Focus on your business
We'll manage your databases

✉ info@datapatform.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:

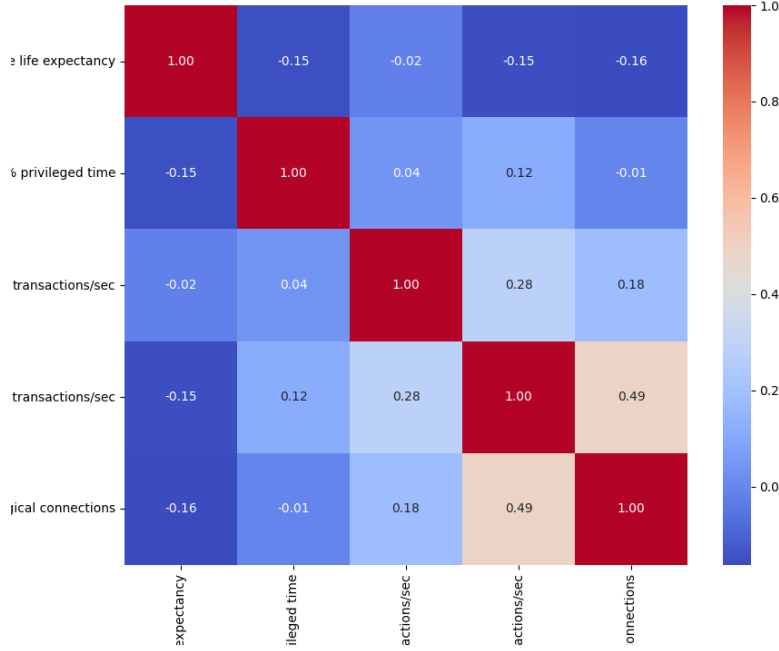
Bu veri setinde yaşam kategorisine göre örüntüler incelendiğinde, Düşük kategoride en fazla 10,017, Orta kategoride 27,476 ve Yüksek kategoride 32,474 değeri elde edilmiştir. Bu verilere dayanarak, genel olarak sayısal olarak daha yüksek olan kategorilerin daha fazla tercih edildiği gözlemlenmektedir. Bu durum, sayısal olarak daha yüksek kategorilere daha fazla ilgi gösterildiğini düşündürebilir.

Focus on your business
We'll manage your databases

✉ info@dataplatfrom.com.tr

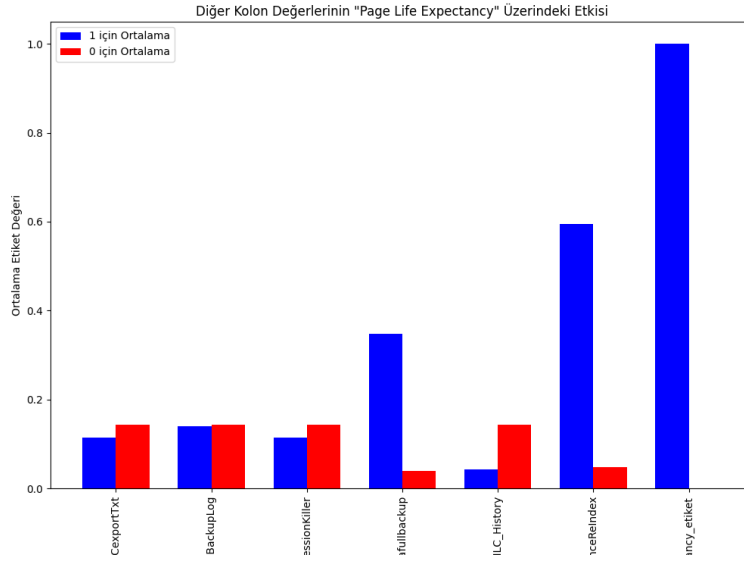
📍 Data Platform Bilgi Sistemleri A.Ş. Bağlarbaşı Mah.
Irmak Sok. No: 15 D:2 K:2 Maltepe/İstanbul

☎ 0532 320 22 39

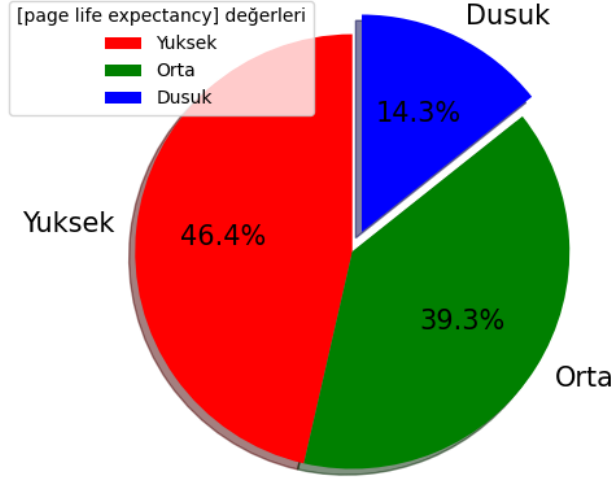


Chat-GPT tarafından üretilen analiz: "Veri setindeki belirli sütunlar arasındaki korelasyon analizine göre, `page life expectancy` ile diğer özellikler arasında şu ilişkiler gözlemlenmiştir:

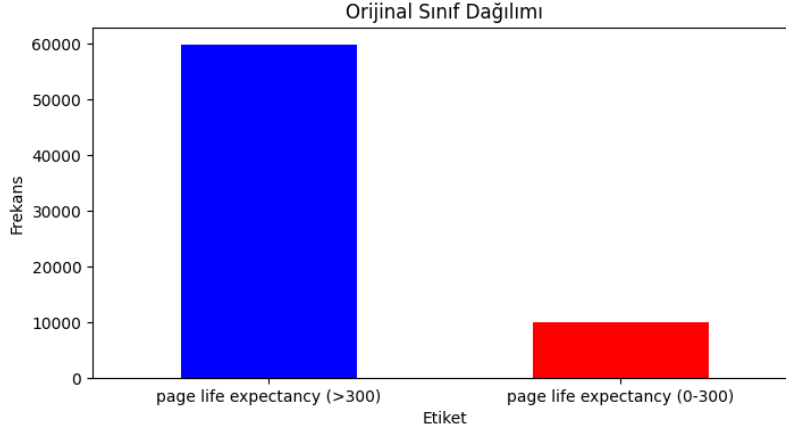
- `% privileged time` ile negatif, zayıf bir korelasyon vardır (-0.150469).
- `transactions/sec` ile zayıf bir korelasyon vardır (-0.019418).
- `write transactions/sec` ile orta düzeyde pozitif bir korelasyon vardır (0.490174).
- `logical connections` ile negatif, zayıf bir korelasyon vardır (-0.162372)."



Chat-GPT tarafından üretilen analiz: 'Bu verilere göre, page life expectancy skoru 1 olan 'page_life_expectancy_etiket' adlı sütun, diğer sütunlara göre en yüksek ortalama etkiyi gösteriyor. Özellikle 'dbaMaintenanceReIndex' sütunuyla arasında büyük bir fark bulunmaktadır. Diğer sütunlar ise genel olarak daha düşük ortalama etkiye sahiptir. Bu durum, 'page_life_expectancy_etiket' sütununun, sayısal veri setinde belirgin bir etki potansiyeline sahip olduğunu göstermektedir.'

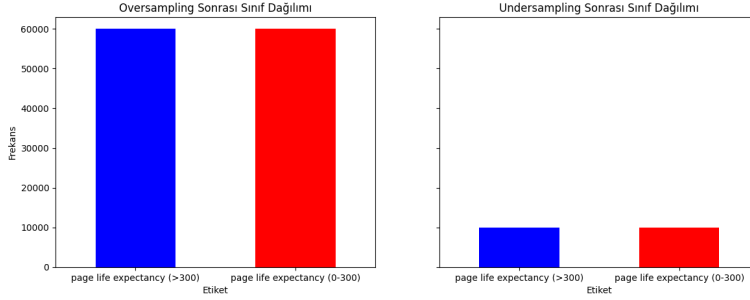


Chat-GPT tarafından üretilen analiz: "Veri setindeki sayfa yaşam süresi değerlerini kategorilere ayırarak gerçekleştirilen analizde yüksek, orta ve düşük olmak üzere üç farklı kategori elde edilmiştir. Bu kategorilere ait gözlem sayıları sırasıyla 32.474, 27.476 ve 10.017 olarak belirlenmiştir. Görselleştirme sonucunda, en fazla gözlemin yüksek page life expectancy kategorisinde olduğu görülmektedir."



Chat-GPT tarafından üretilen analiz: Veri setindeki sayfa yaşam süresi değerlerine dayanarak yapılan etiketleme, 59950 veriyi 0 etiketiyle ve 10017 veriyi 1 etiketiyle belirlemiştir. Bu durum, sayfa yaşam süresi değerlerinin iki farklı kategoriye ayrıldığını göstermektedir.

Veri seti hakkında genel bir bakış açısıyla, sayfa yaşam süresi değerlerinin iki farklı grup arasında belirgin bir ayrım olduğunu söyleyebiliriz. Bu ayrımın analiz ve görselleştirme süreçlerinde dikkate alınarak daha detaylı incelemeler yapılabilir. Bu veriler, sayfa yaşam süresi değerlerinin performans açısından farklılık gösterebileceğini düşündürmektedir. Bu nedenle, performansı etkileyen faktörleri araştırmak ve daha iyi bir veri analizi yapmak önemli olacaktır. Bu sonuçlar, veri setindeki sayfa yaşam süresi değerlerinin kapsamlı bir analizini gerektirebilir.

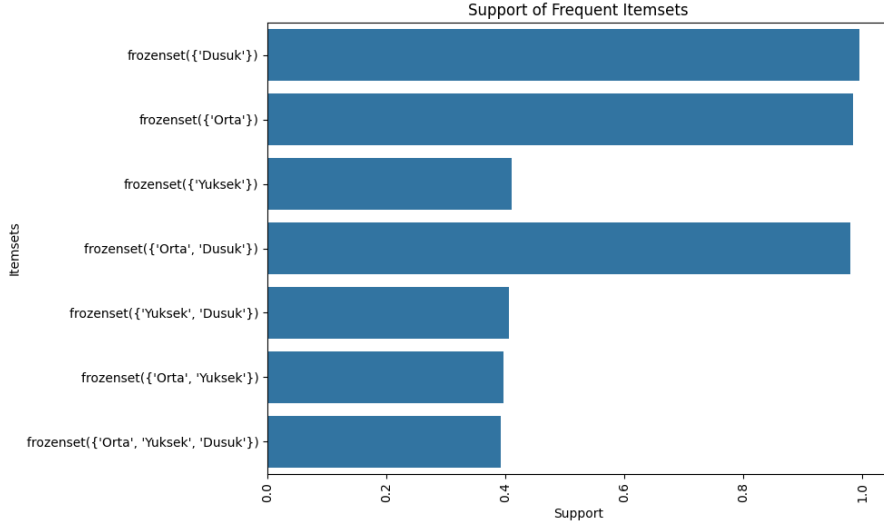


Chat-GPT tarafından üretilen analiz: 'Senin analizin'

Oversampling yapıldıktan sonra, sınıf dengesinin sağlandığı ve her iki sınıfın da 59950 veriye sahip olduğu görülüyor. Bu durum, eğitim verilerindeki dengesizlik sorununu gidermek için etkili bir yöntem olabilir.

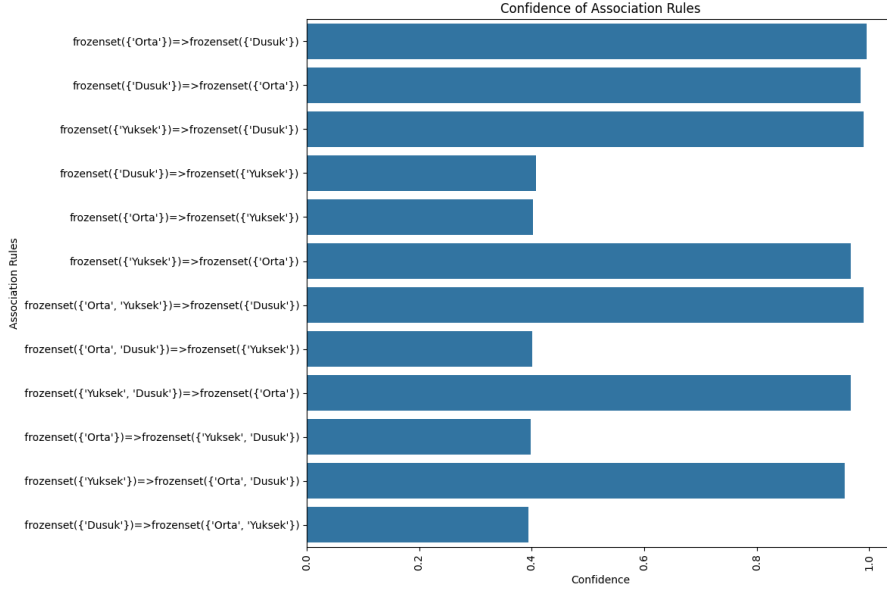
Undersampling yapıldıktan sonra ise, her iki sınıfın da 10017 veriye düştüğü görülüyor. Bu yöntemde azınlık sınıfının veri sayısının çoğunluk sınıfına eşitlendiği ve böylelikle sınıf dengesinin sağlandığı gözlemlenmektedir.

Her iki yöntem de amaçlarına ulaşmış gibi görünse de, hangi yöntemin daha iyi bir performans sağlayacağı veri setine ve probleme bağlı olacaktır. Örnekleme tekniklerinin sonuçlarına göre, model eğitiminde dikkate alınacak en uygun örnekleme yöntemi seçilmelidir.



Chat-GPT tarafından üretilen analiz:

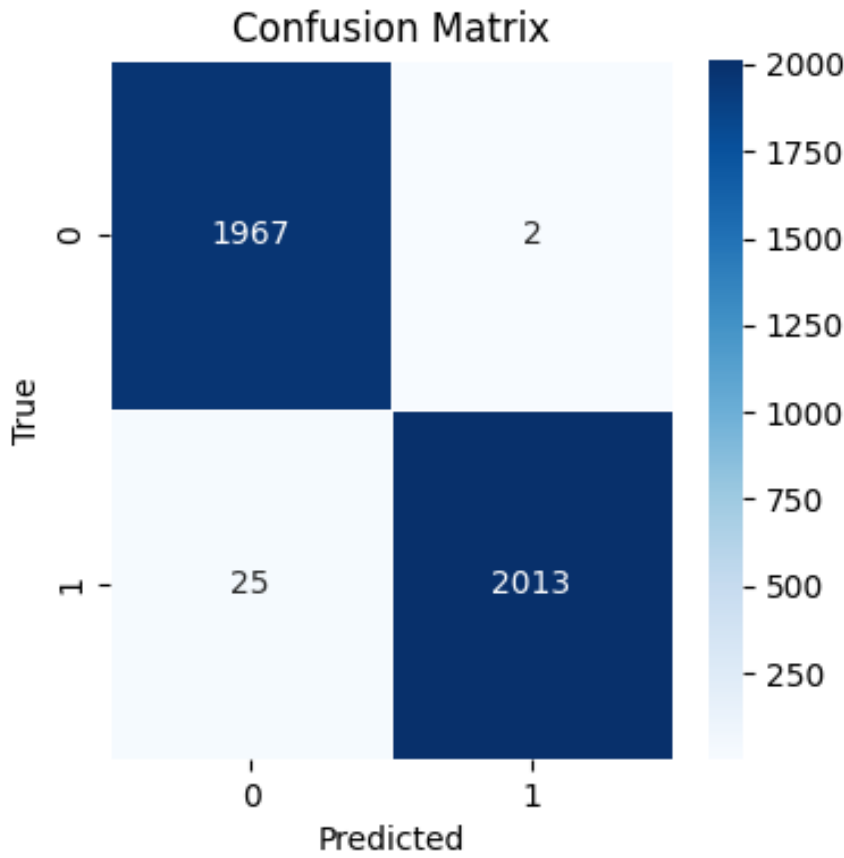
Bu veri setinde, genel olarak "Düşük", "Orta" ve "Yüksek" seviyeleri içeren veriler bulunmaktadır. En sık görülen kombinasyonlar incelendiğinde, "Orta" ve "Düşük" seviyelerinin bir arada görülme olasılığı oldukça yüksekken, "Yüksek" seviyesi diğerlerine göre daha az birlikte görülmektedir. Ayrıca, "Orta" ve "Yüksek" seviyelerinin aynı anda bulunma olasılığı da göz önüne alındığında, "Orta" seviyesi "Yüksek" seviyesi ile birlikte genellikle "Düşük" seviyesi ile bir arada bulunmaktadır. Bu veriler, seviyeler arasındaki ilişkileri ve kombinasyonların sıklığını analiz etmek için önemli ipuçları sağlamaktadır.



Chat-GPT tarafından üretilen analiz: 'Senin analizin'

Bu item setlerine göre, "Dusuk" seviyesinde olan ürün alımının "Orta" veya "Yuksek" seviyesindeki ürün alımlarına kıyasla daha yüksek olduğu görülmektedir. "Orta" ve "Yuksek" seviyedeki ürün alımları arasında da benzer bir ilişki söz konusudur. Ayrıca, "Orta" ve "Yuksek" seviyedeki alımların birlikte gerçekleştiği durumlarda genellikle "Dusuk" seviyedeki alımların da gerçekleşme olasılığı yüksektir.

Çıkarılan kurallar, alışveriş trendlerini anlamak ve stratejiler geliştirmek için değerli bir bilgi sunmaktadır. Bu bilgiler doğrultusunda "Orta" ve "Yuksek" seviyedeki alımların teşvik edilmesi veya birlikte satılması durumunda "Dusuk" seviyedeki alımların da artabileceği öngörülebilir.



Çalışan algoritma: KNN

Chat-GPT tarafından üretilen analiz: Bu KNN modelinin confusion matrix sonuçları oldukça başarılı görünüyor. Model, toplamda 4007 veriden 3978 tanesini doğru sınıflandırmış. Recall değeri %98.75 olarak oldukça yüksek, yani modelin gerçek pozitifleri kaçırma olasılığı oldukça düşük. Benzer şekilde, accuracy değeri de %99.38, yani modelin doğru tahmin oranı oldukça yüksek. F1 score ise %99.37 olarak modelin precision ve recall dengesini iyi gösteriyor. Sınıflandırma raporu da gösteriyor ki model veri setinde "False" ve "True" sınıflarını oldukça iyi ayırt etmiş. Sonuç olarak, bu KNN modeli genel olarak oldukça başarılı bir performans sergilemiş gibi görünmektedir.