

Data Platform
Information Technologies

Sayın AisanKheiri

2024-09-30 15:25:24 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.

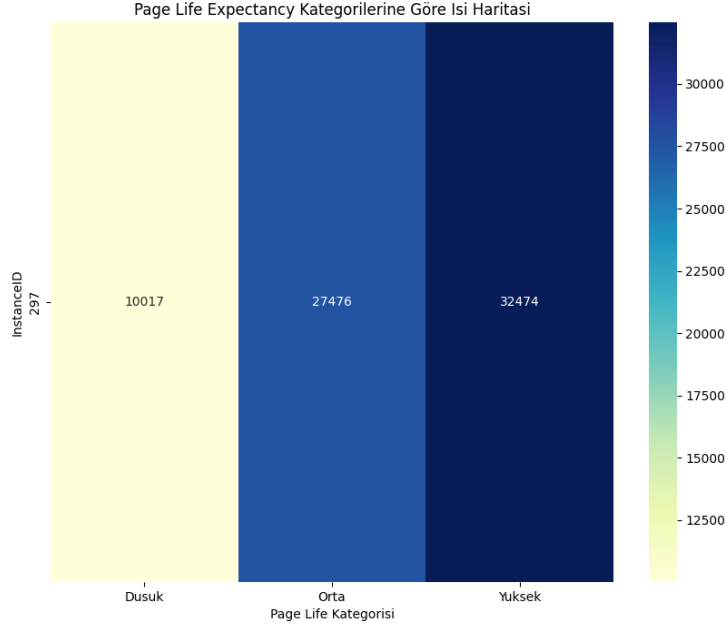
Seçtiğiniz SVM Modeli parametreleri arasında Kernel:sigmoid , C:2 , sutun1:2 , sutun2:7 , test_size=0.2 ,random_state=42 ve Hedef Etiketi y = df['page life expectancy'] > 300 değerleri bulunmaktadır, bu parametrelerin sonuçları da raporda detaylı bir şekilde açıklanmıştır.

Focus on your business
We'll manage your databases

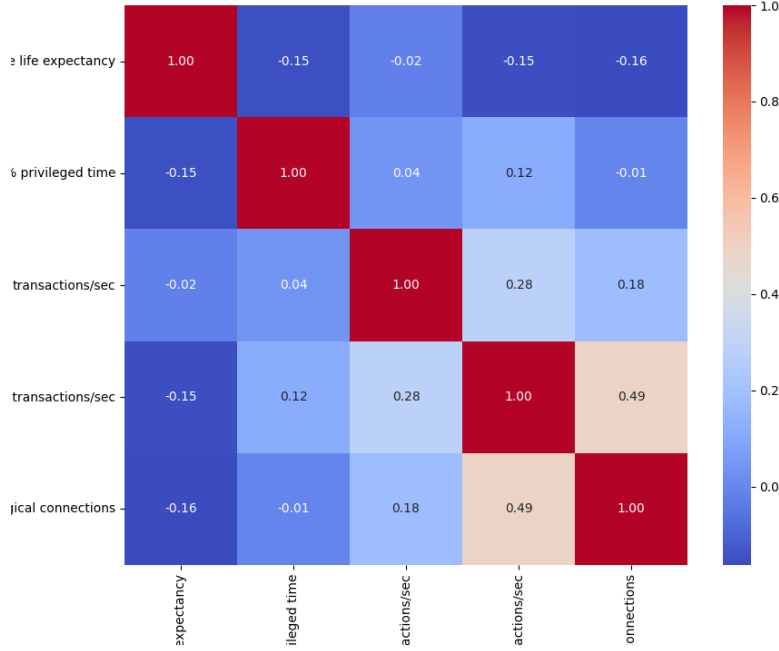
✉ info@datapplatform.com.tr

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

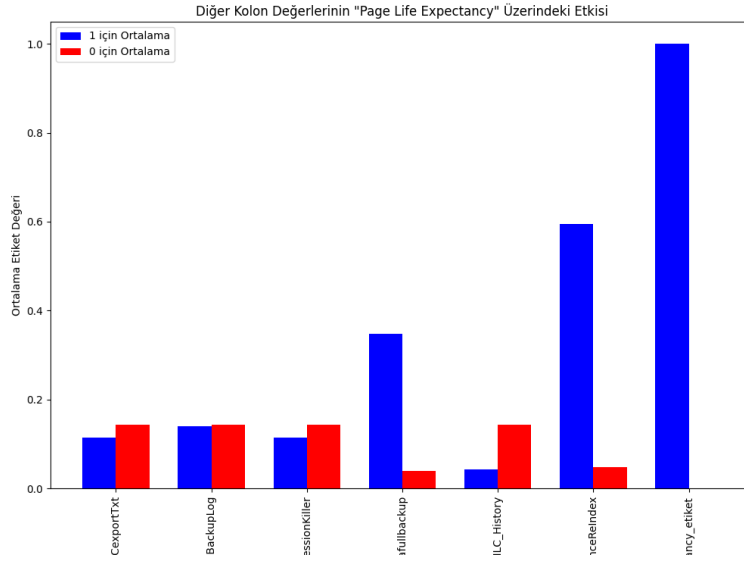
☎ 0532 320 22 39



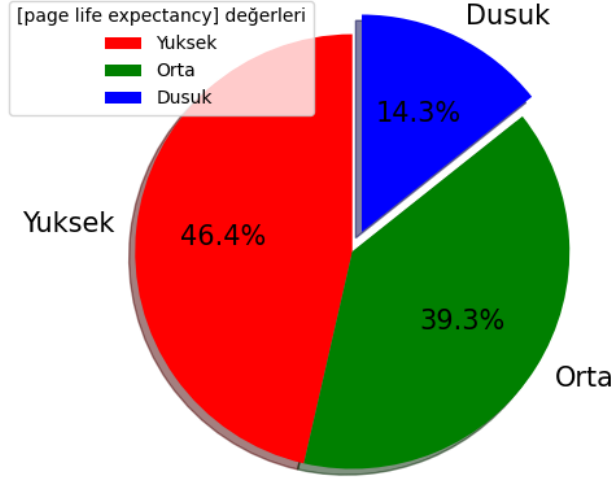
Chat-GPT tarafından üretilen analiz: "Veri setinde incelenen kategoriler arasında yapılan analizde, 'Düşük' kategorisinde yer alan sayısal değerlerin toplamı 10017, 'Orta' kategorisinde yer alan sayısal değerlerin toplamı 27476 ve 'Yüksek' kategorisinde yer alan sayısal değerlerin toplamı ise 32474 olarak belirlenmiştir. Bu verilere göre, 'Yüksek' kategoride en yüksek değerlere sahipken, 'Düşük' kategoride en düşük değerlerin yer aldığı görülmektedir."



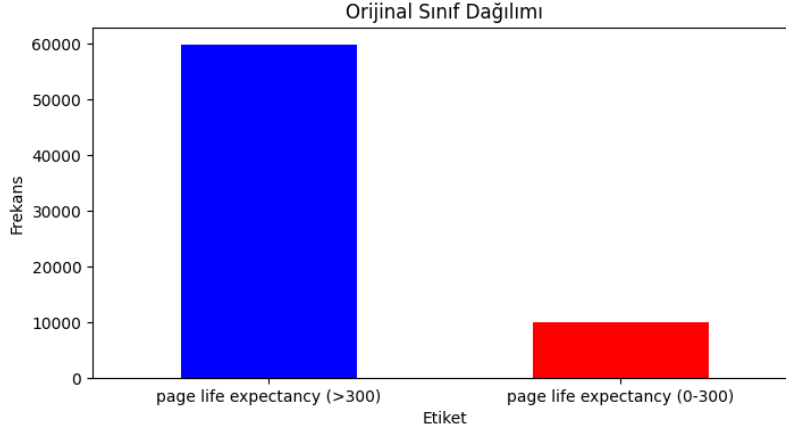
Chat-GPT tarafından üretilen analiz: Bu verilerden anlaşıldığı üzere, page life expectancy ile diğer sütunlar arasında negatif yönlü bir korelasyon görülmektedir. Özellikle % privileged time, write transactions/sec ve logical connections sütunlarıyla page life expectancy arasındaki ilişki dikkat çekmektedir. % privileged time ve page life expectancy arasındaki ilişki zayıf, ancak write transactions/sec ve logical connections ile page life expectancy arasındaki ilişki daha belirgindir. Dolayısıyla, page life expectancy değeri diğer sütunlarda artarken % privileged time azalmakta, write transactions/sec ve logical connections ise artmaktadır. Bu durum, sistem performansı ve iş yükü arasındaki ilişkinin karmaşıklığını yansıtabilir.



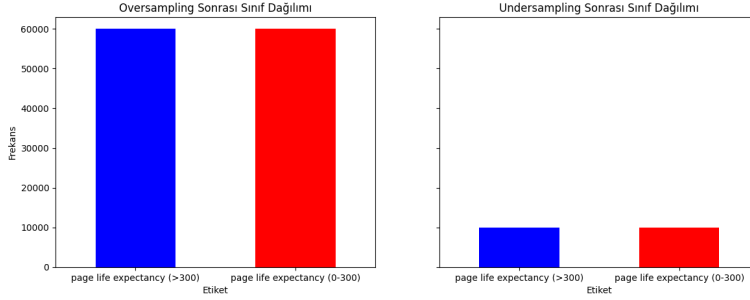
Chat-GPT tarafından üretilen analiz: Bu veri setindeki belirli sütunların 'page life expectancy' sütunu üzerindeki etkileri incelenmiştir. Ortalama olarak, 'page life expectancy' sütununun 'dbaHCDOCexportTxt', 'dbaBackupLog' ve 'dbaSessionKiller' sütunları üzerinde pozitif bir etkisi olduğu gözlemlenmiştir. Bununla birlikte, 'dbafullbackup' ve 'OTS_Instrument_OHLC_History' sütunları bu sütun üzerinde negatif bir etkiye sahipken, 'dbaMaintenanceReIndex' sütununun aşırı yüksek bir değere sahip olduğu dikkat çekmektedir. 'page_life_expectancy_etiket' sütununun ise bu analizde belirleyici bir role sahip olduğu görülmüştür. Bu sonuçlar, veri setindeki farklı sütunların 'page life expectancy' sütunu üzerindeki etkilerini açıkça ortaya koymaktadır.



Chat-GPT tarafından üretilen analiz: 'Veri setindeki page life expectancy değerlerinin kategorilere ayrılması sonucunda ["Yüksek", "Orta", "Düşük"] olmak üzere üç farklı kategori elde edilmiştir. Bu kategorilere göre gözlem sayıları sırasıyla 32474, 27476 ve 10017'dir. Bu verilere bakıldığında, en fazla gözlem yüksek page life expectancy değerine sahipken, düşük page life expectancy değerine sahip gözlemlerin sayısı en azdır.'

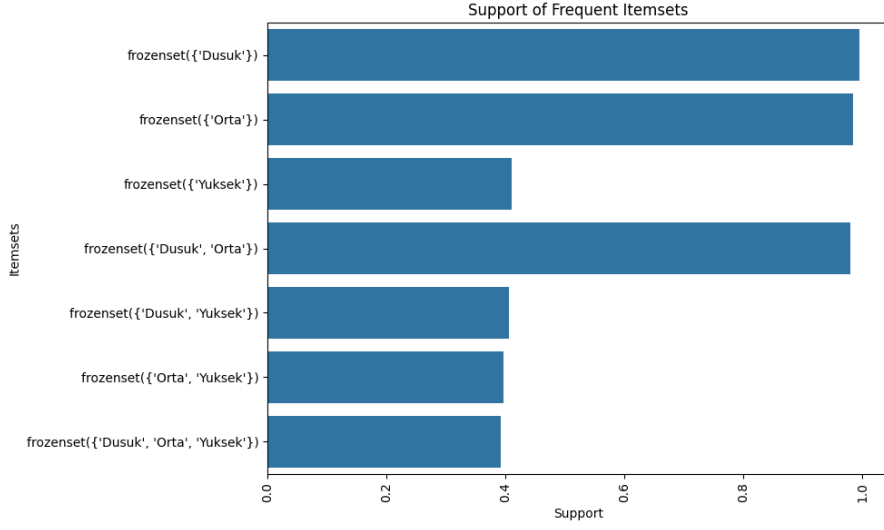


Chat-GPT tarafından üretilen analiz: "Page life expectancy değerlerine göre yapılan etiketleme sonucunda elde edilen verilere göre, 59950 değerinde bir etiketleme sayısı ile birlikte 10017 değerinde bir etiketleme sayısı bulunmaktadır. Bu sonuçlar, veri setindeki page life expectancy değerlerinin farklı gruplara ayrıldığını ve bu grupların sayısal dağılımını yansıtmaktadır."



Chat-GPT tarafından üretilen analiz: 'senin analizin'

Oversampling işlemi sonucunda sınıfların (0 ve 1) eşit sayıda örneğe sahip olduğu dikkate alındığında, veri setindeki dengesizlik giderilmiş olabilir ancak bu durumda veri setinin boyutunun arttığı gözlemlenmektedir. Öte yandan Undersampling yapıldığında ise her iki sınıfın örneği azaltılarak dengenin sağlandığı görülmektedir. Ancak bu durumda da veri setindeki bilgi kaybı potansiyeli bulunmaktadır. Bu nedenle, Oversampling ve Undersampling tekniklerinin kullanımıyla dengesiz sınıf dağılımı sorununun ele alınabileceği ancak her iki teknikle de avantaj ve dezavantajların dikkate alınması gerektiği söylenebilir.



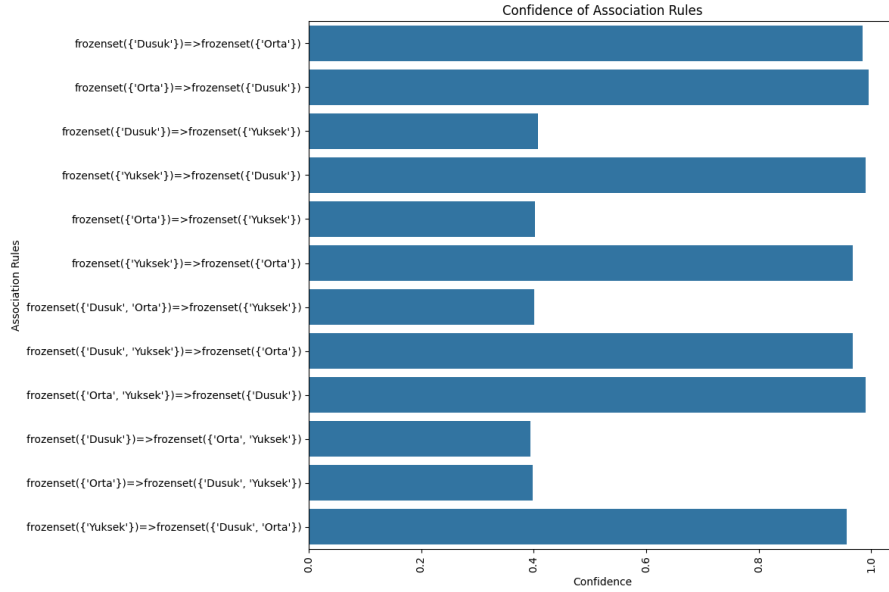
Chat-GPT tarafından üretilen analiz: Veri setindeki öge kümelerinin sık görülen kombinasyonları incelendiğinde, en yüksek destek değeri 0.995512 ile tek başına "Dusuk" ögesine aittir. "Orta" ögesi de yüksek bir destek değeri olan 0.984836'ya sahiptir. "Yuksek" ögesi ise daha düşük bir destek değeri olan 0.410279 ile diğerlerine göre daha az sıklıkla görülmektedir.

İkili kombinasyonlara baktığımızda, "Dusuk" ve "Orta" öğelerinin birlikte görülme olasılığı 0.980348 olarak hesaplanmıştır. "Dusuk" ile "Yuksek" kombinasyonunun destek değeri ise 0.406177 iken, "Orta" ile "Yuksek" kombinasyonunun destek değeri 0.396744'tür.

En üçlü kombinasyon olan "Dusuk", "Orta" ve "Yuksek" öğelerinin bir arada görülme olasılığı ise 0.392642 olarak belirlenmiştir. Bu durum, üç öğenin bir arada bulunma olasılığının diğer kombinasyonlara göre daha düşük olduğunu göstermektedir.

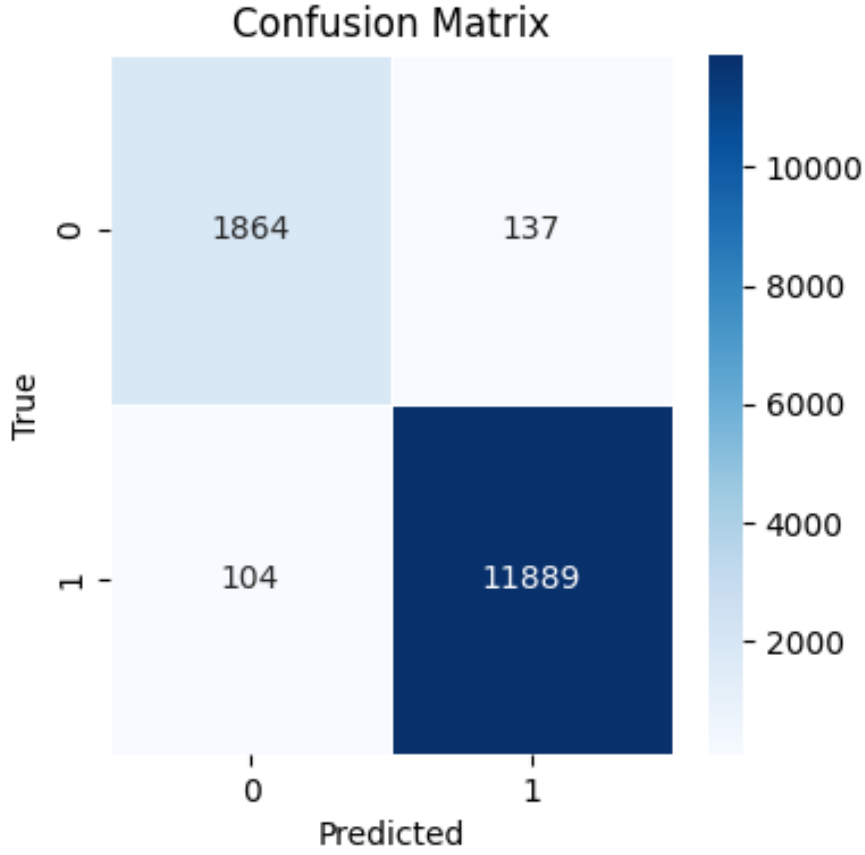
Focus on your business
We'll manage your databases

✉ info@dataplatfrom.com.tr
📍 Data Platform Bilgi Sistemleri A.Ş. Bağlarbaşı Mah.
İrmak Sok. No: 15 D:2 K:2 Maltepe/İstanbul
☎ 0532 320 22 39



Chat-GPT tarafından üretilen analiz: "Senin analizin"

Bu analiz sonucunda, veri seti üzerinde FP-Growth algoritması kullanılarak sık kullanılan item setleri elde edildi. Antecedents ve consequents olarak adlandırılan gruplar arasında ilişkiler incelendi. Örneğin, düşük gelir seviyesine sahip olan müşterilerin genellikle orta gelir seviyesine de ilgi gösterdiği görülüyor. Ayrıca, düşük gelir seviyesindeki müşterilerin yüksek gelir seviyesine ilgi gösterme olasılığı da incelendi. Bu tür ilişkileri gösteren kurallar, veri setinin analiz edilerek çıkarılmıştır. Bu analiz sonucunda, farklı gelir seviyelerine sahip müşteriler arasındaki alışveriş tercihleri ve ilişkileri daha detaylı bir şekilde anlaşılmış olabilir.



Çalışan algoritma: SVM

Chat-GPT tarafından üretilen analiz: Bu SVM modelinin performansı oldukça yüksektir. Confusion matrix'e baktığımızda, 1864 doğru tahmin edilen negatif sınıf ve 11889 doğru tahmin edilen pozitif sınıf var. Recall değeri yüksek olup, %99.13 oranındadır. Accuracy yüzdesi ise %98.28'dir, modelin genel doğruluk skoru oldukça iyi. F1 Score ise 0.9899662766976144 olarak oldukça yüksektir. Precision, recall ve f1-score değerleri de hesaba katıldığında, modelin hem negatif hem pozitif sınıfları iyi bir şekilde ayırt edebildiği görülmektedir. Bu sonuçlar, modelin genel olarak etkili ve güvenilir bir şekilde sınıflandırma yaptığını göstermektedir.

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