

Data Platform
Information Technologies

Sayın AisanKheiri

2024-10-02 13:20:25 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.

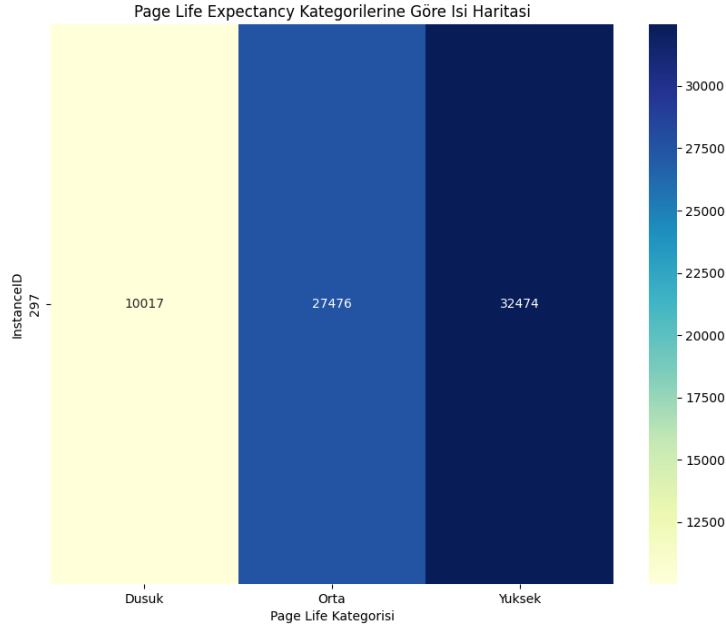
Seçtiğiniz SVM Modeli parametreleri arasında Kernel:sigmoid , C:2 , sutun1:2 , sutun2:7 , test_size=0.2 ,random_state=42 ve Hedef Etiketi y = df['page life expectancy'] > 300 değerleri bulunmaktadır, bu parametrelerin sonuçları da raporda detaylı bir şekilde açıklanmıştır.

Focus on your business
We'll manage your databases

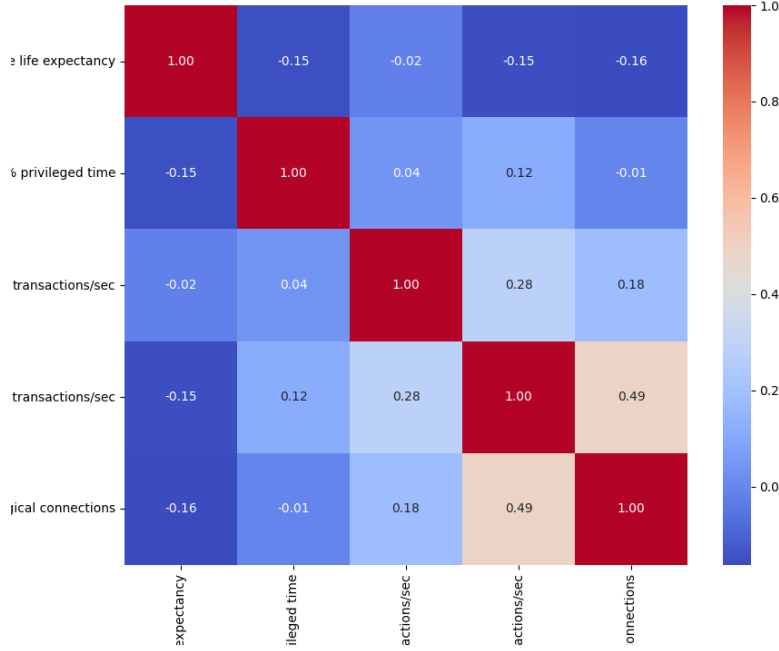
✉ info@dataplatfrom.com.tr

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

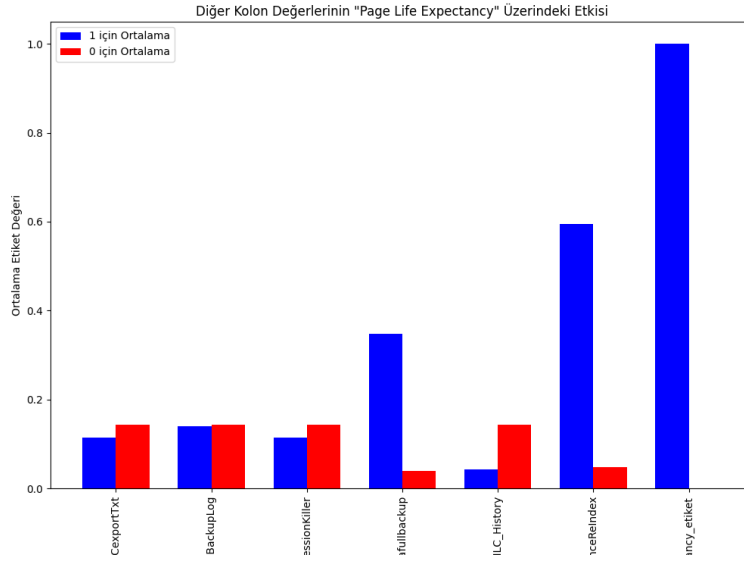
☎ 0532 320 22 39



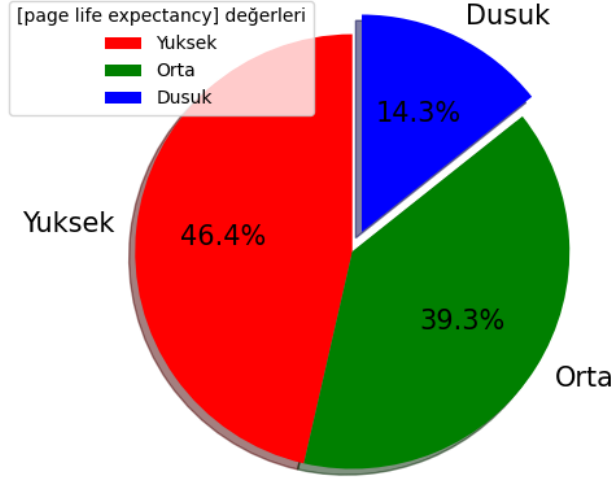
Chat-GPT tarafından üretilen analiz: "Veri setindeki ilişkiler incelendiğinde, düşük sayıda olan `Düşük` kategorisine ait değerlerin 10017, `Orta` kategorisine ait değerlerin 27476 ve `Yüksek` kategorisine ait değerlerin ise 32474 olduğu görülmektedir. Bu verilere göre, `Orta` ve `Yüksek` kategoriler arasında belirgin bir fark olduğu gözlemlenmektedir."



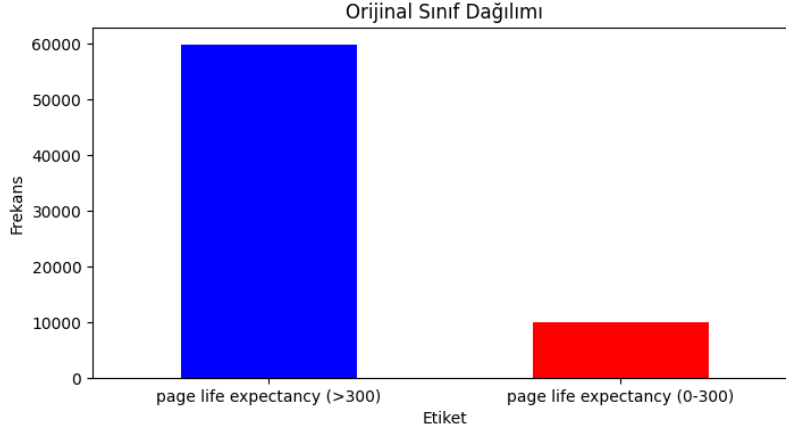
Chat-GPT tarafından üretilen analiz: Bu veri setindeki belirli sütunlar arasındaki korelasyon incelendiğinde, 'page life expectancy' ile 'logical connections' arasında güçlü bir negatif ilişki bulunmaktadır (-0.162372). 'page life expectancy' ile diğer sütunlar arasında ise çok düşük veya zayıf ilişkiler gözlemlenmektedir. 'write transactions/sec' ile 'logical connections' arasında ise orta düzeyde pozitif bir ilişki (0.490174) belirlenmiştir. Genel olarak, veri setindeki sütunlar arasındaki korelasyonlar incelendiğinde 'page life expectancy' ve 'logical connections' arasındaki ilişkinin diğerlerine göre daha belirgin olduğu görülmektedir.



Chat-GPT tarafından üretilen analiz: 'Bu veriler incelendiğinde, "dbaHCDOCexportTxt", "dbaSessionKiller" ve "OTS_Instrument_OHLC_History" sütunlarının "page life expectancy" sütunu üzerinde negatif bir etkiye sahip olduğu gözlemlenmektedir. Bu sütunlar, 0 değerine yakın ortalama etkiye sahipken, diğer sütunlar ise 1'e çok daha yakındır. Özellikle "dbaMaintenanceReIndex" sütununun "page life expectancy" üzerinde belirgin bir etkisinin olduğu görülmektedir. Sonuçlar, belirli sütunların veritabanı performansı üzerindeki potansiyel etkilerini göstermektedir.'

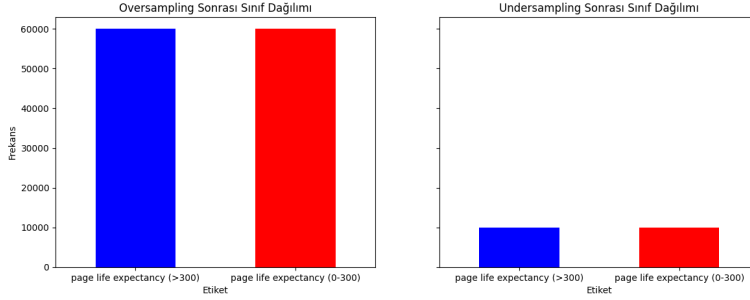


Chat-GPT tarafından üretilen analiz: 'Page life expectancy değerleri incelendiğinde, veri setindeki gözlemlerin çoğunun orta seviyede bulunduğu görülmektedir. Yüksek kategorisinde bulunan gözlem sayısı oldukça fazla iken, düşük kategorisindeki gözlem sayısı oldukça azdır. Bu durum, genel olarak server tarafında daha yeterli bir performans sunulduğunu ancak düşük seviyedeki page life expectancy değerlerinin de göz ardı edilmemesi gerektiğini düşündürmektedir.'

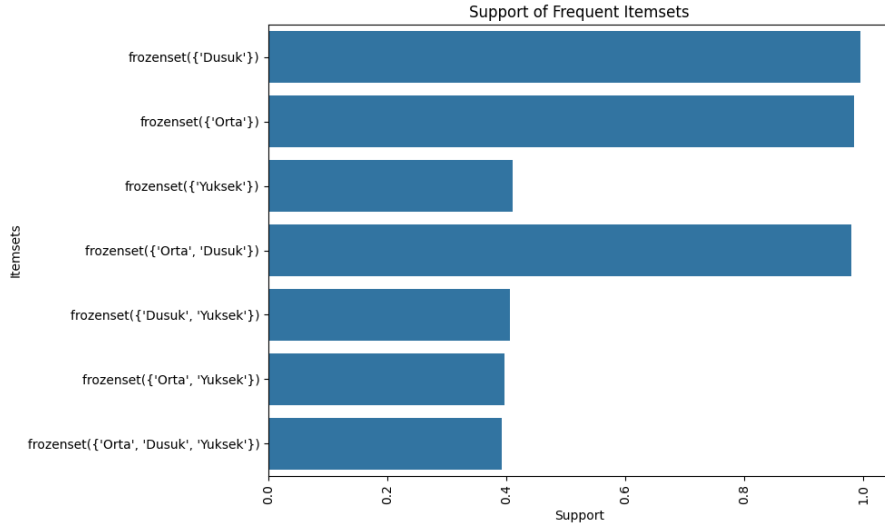


Chat-GPT tarafından üretilen analiz:

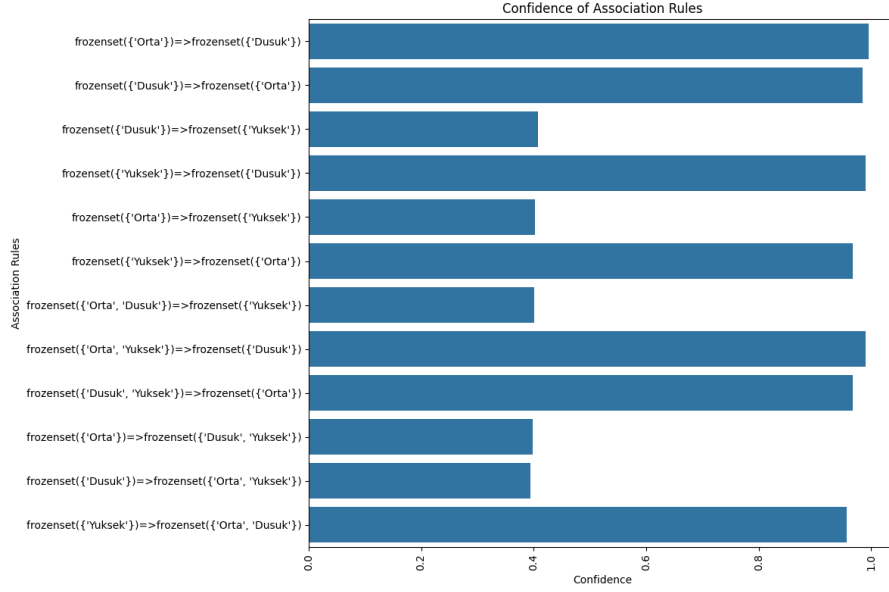
Bu veri setinde bulunan page life expectancy değerlerine göre yapılan etiketleme sonucunda 59950 veri "0" etiketiyle, 10017 veri ise "1" etiketiyle belirlenmiştir. Bu durum, belirli bir eşik değerinin altında veya üstünde bulunan veri noktalarının farklı kategorilere atanmasını sağlar. Veri setindeki farklı etiketlerin dağılımı görselleştirildiğinde, "0" ve "1" etiketlerine sahip veri noktalarının sayısal olarak farklı olduğu açıkça görülmektedir. Bu durum, page life expectancy değerlerinin belirlenen kriterlere göre farklı gruplara ayrıldığını ve belki de farklı performans seviyelerini yansıttığını göstermektedir. Bu etiketleme işlemi sayesinde veri setindeki düzensizlikler daha net bir şekilde anlaşılabilir ve analizler daha doğru bir şekilde yapılarak verilerin daha etkili bir şekilde değerlendirilmesi sağlanabilir.



Chat-GPT tarafından üretilen analiz: "Sınıf dengesizliği sorununu ele almak için uygulanan örnekleme tekniklerinin sonuçlarına göre, oversampling yapıldıktan sonra her iki sınıfın da 59950 etikete sahip olduğu görülmektedir. Undersampling yapıldıktan sonra ise her iki sınıfın da 10017 etikete sahip olduğu gözlemlenmektedir. Bu durum, sınıflar arasındaki dengesizliğin başarılı bir şekilde düzeltildiğini ve her iki sınıfın da eşit sayıda örneğe sahip olduğunu göstermektedir."

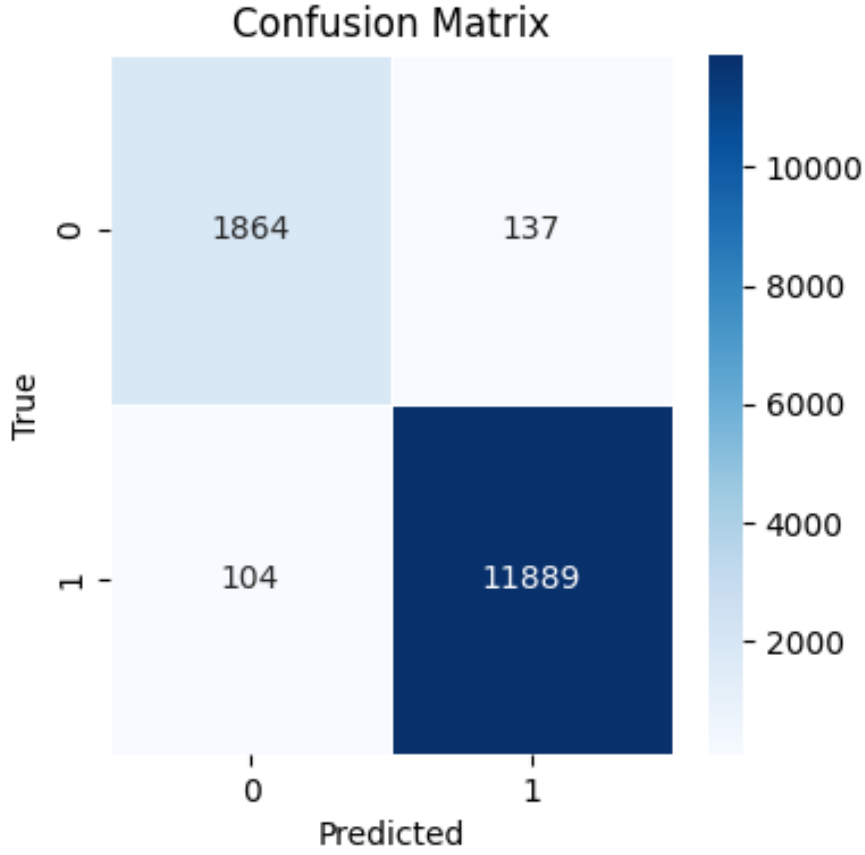


Chat-GPT tarafından üretilen analiz: Bu veri setinde, öncelikle Düşük, Orta ve Yüksek seviyelerinde yapılan analizlerin destek değerleri incelendiğinde, en yüksek destek değeri %99,55 ile Düşük seviyesinin olduğu görülmektedir. Ardından Orta seviyesi %98,48 ile ikinci sırada gelmekte ve son olarak da Yüksek seviyesi %41,03 destek değeri ile en düşük destek değeri olarak kaydedilmiştir. Eğer kombinasyonlar incelenirse, Orta ve Düşük seviyelerinin bir arada görülme olasılığı %98,03, Düşük ve Yüksek seviyelerinin bir arada görülme olasılığı ise %40,62 olarak bulunmuştur. Orta, Yüksek ve Düşük seviyelerinin bir arada görülme olasılığı ise %39,26 olarak hesaplanmıştır. Bu veriler ışığında, genel olarak Düşük seviyesinin diğer seviyelerle kombinasyonlarının daha sık olarak gözlemlendiği söylenebilir.



Chat-GPT tarafından üretilen analiz: 'Senin analizin'

Bu veriler, FP-Growth algoritması kullanılarak elde edilen sık kullanılan item setlerine dayalı olarak kuralların çıkarıldığını göstermektedir. Örneğin, "Orta" seviyesinde olan bir durumun "Düşük" seviyesinde olma olasılığı %98.48 olarak bulunmuştur. Aynı şekilde, "Düşük" seviyesinde olan bir durumun "Orta" seviyesine geçiş olasılığı %99.55 olarak hesaplanmıştır. Benzer şekilde, diğer kombinasyonlar için de benzer olasılık ve metrik değerleri hesaplanmıştır. FP-Growth algoritmasının çıkardığı bu kurallar, veri集中的 item'lar arasındaki ilişkileri anlamamıza ve gelecekteki davranışları tahmin etmemize yardımcı olabilir.



Çalışan algoritma: SVM

Chat-GPT tarafından üretilen analiz: Bu SVM algoritması oldukça başarılı bir performans sergilemiş. Confusion Matrix'e bakıldığında, toplam 13994 veri örneğinin 1303'ü yanlış sınıflandırılmış. Recall değeri yüksek olduğundan, modelin gerçek pozitifleri doğru şekilde tanıma yüzdesi oldukça iyi (0.991). Accuracy değeri de oldukça yüksektir (0.982), yani modelin doğru sınıflandırma oranı oldukça başarılıdır. F1 Score ise 0.989 olarak hesaplanmış, bu da modelin hem precision hem de recall metriklerine göre dengeli bir performans gösterdiğini gösterir. Classification Report'a göre, model False sınıfını %94.71 doğrulukla sınıflandırırken, True sınıfını %98.86 doğrulukla sınıflandırmış. Genel olarak, weighted avg metriğine bakıldığında modelin genel doğruluk düzeyi

0.983 olarak belirlenmiş. Bu sonuçlar, SVM algoritmasının bu veri seti üzerinde etkili bir şekilde

çalıştığını 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