



Data Platform
Information Technologies

Sayın kullanıcı1

2024-08-16 10:37:23 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.

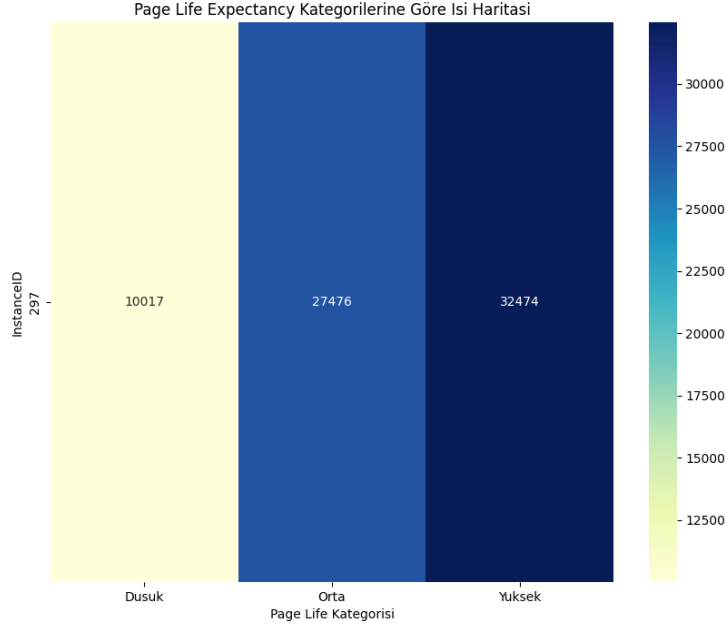
Seçtiğiniz KNN Modeli parametreleri arasında `n_estimators:100` , `max_depth:5` , `min_samples_split:2` , `criterion:gini` , `test_size=0.2` ve `random_state=42` değerleri bulunmaktadır, bu parametrelerin sonuçları da raporda detaylı bir şekilde açıklanmıştır.

Focus on your business
We'll manage your databases

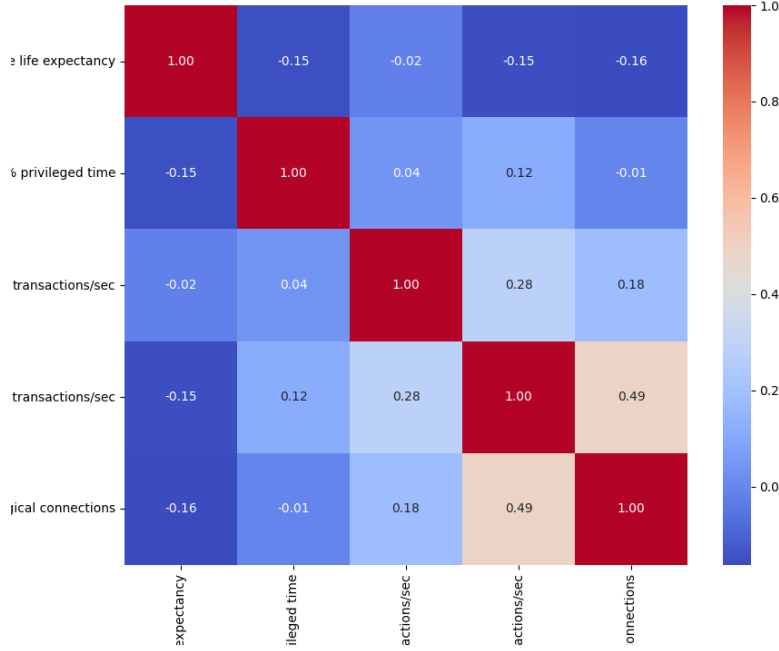
 info@dataplatfrom.com.tr

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

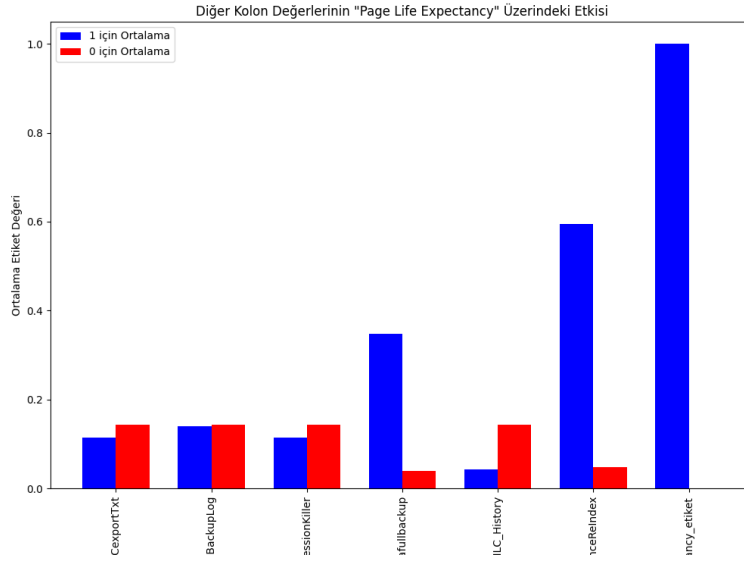
 0532 320 22 39



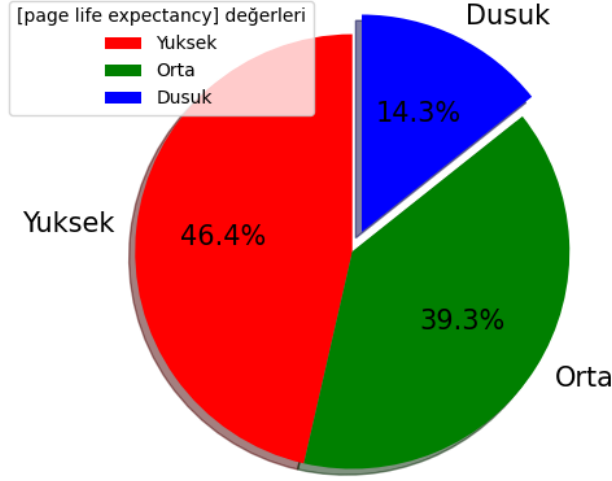
Chat-GPT tarafından üretilen analiz: "Veri setindeki ilişkileri incelediğimde, Düşük kategorideki sayıların diğer kategorilere göre daha düşük olduğunu görüyorum. Orta kategoride ise orta seviyede bir dağılım olduğunu söyleyebilirim. Yüksek kategoride ise en yüksek değerlerin bulunduğunu fark ediyorum. Bu verilere göre, sayısal olarak yüksek kategorideki örüntülerin diğer kategorilere göre belirgin bir şekilde farklılık gösterdiğini söyleyebilirim."



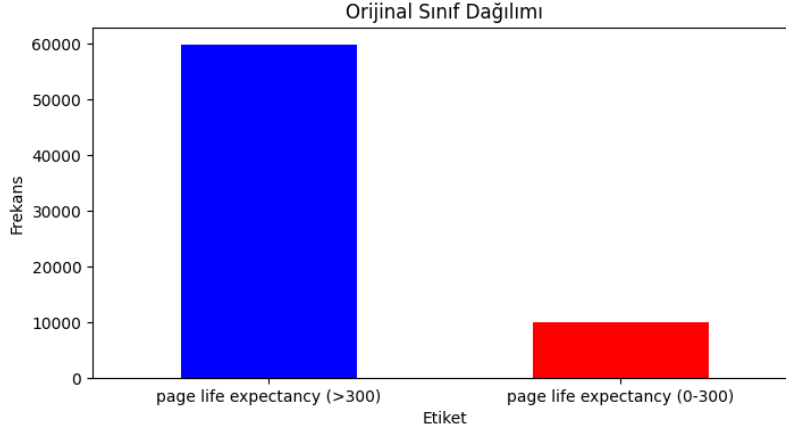
Chat-GPT tarafından üretilen analiz: Bu veri setinde page life expectancy ile diğer sütunlar arasındaki korelasyon incelendiğinde, belirli ilişkiler göze çarpıyor. % privileged time ile page life expectancy arasında negatif yönlü ve zayıf bir korelasyon (-0.150469) gözlemlenirken, transactions/sec ve write transactions/sec gibi sütunlar arasında ise pozitif yönlü ve orta düzeyde korelasyonlar tespit edilmiştir. En ilginç ilişki ise write transactions/sec ile logical connections arasındaki güçlü ve pozitif yönlü korelasyondur (0.490174). Sonuç olarak, veri setindeki farklı sütunlar arasındaki ilişkiler incelendiğinde belirli bağlantılar ortaya çıkmaktadır.



Chat-GPT tarafından üretilen analiz: "Veri setindeki belirli sütunların 'page life expectancy' sütunu üzerindeki etkilerini incelemek için yapılan analizde, farklı sütunların bu değeri etkilediği görülmektedir. Örneğin, 'dbaHCDOCexportTxt' sütunu için ortalama etki düşük iken 'dbaBackupLog' sütunu için biraz daha yüksek bir etki gözlemlenmiştir. 'dbaSessionKiller' sütunu da düşük bir etki gösterirken, 'dbafullbackup' sütunu bu değeri oldukça yüksek etkilemektedir. Diğer yandan, 'OTS_Instrument_OHLC_History' ve 'dbaMaintenanceReIndex' sütunlarının etkileri ise diğer sütunlara kıyasla daha düşük seviyededir. En güçlü etki 'page_life_expectancy_etiket' sütunuyla elde edilmiştir, yani bu sütunun 'page life expectancy' üzerinde belirgin bir etkisi olduğu görülmektedir."

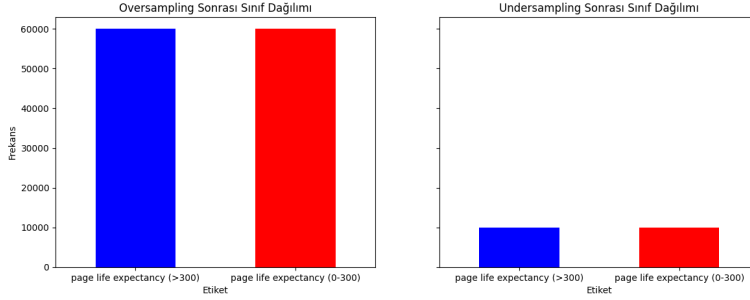


Chat-GPT tarafından üretilen analiz: 'Page Life Expectancy değerlerine bakıldığında, veri setindeki gözlemlerin çoğunun orta seviyede olduğu görülmektedir. Orta seviyede olan gözlem sayısı 27,476'dır. Yüksek seviyede olan gözlem sayısı 32,474 iken, düşük seviyede olan gözlem sayısı ise 10,017'dir.'

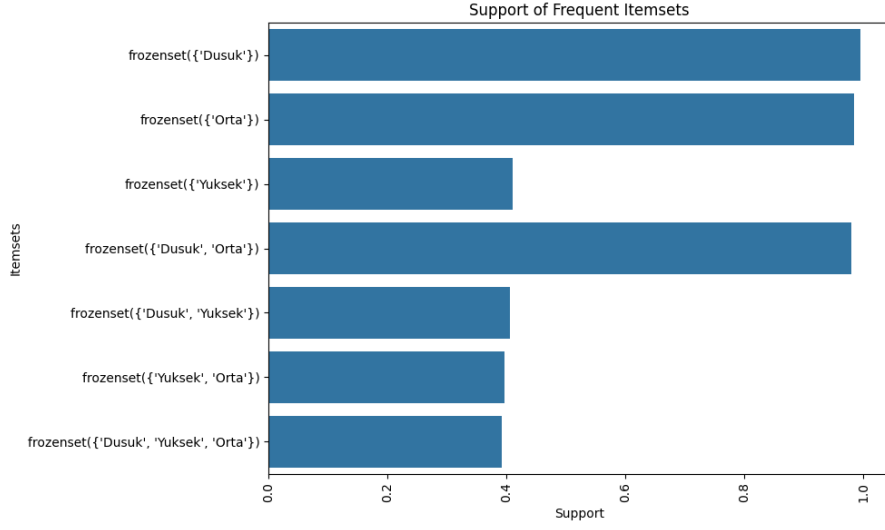


Chat-GPT tarafından üretilen analiz: "Senin analizin"

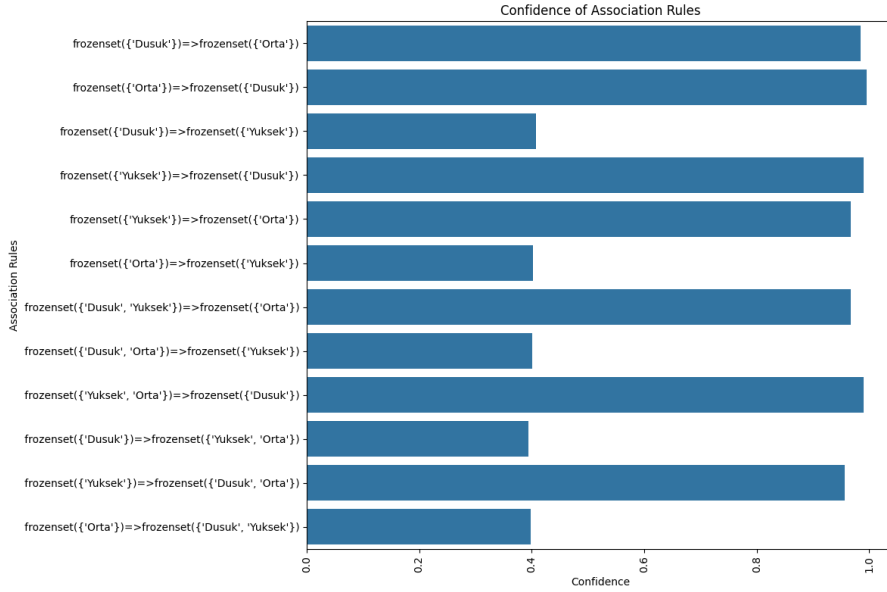
Bu analiz, veri setindeki page life expectancy değerlerine göre yapılan etiketlemeyi temsil etmektedir. Veri setindeki değerlere göre elde edilen etiketler, 0 ve 1 değerleridir. Page life expectancy değerleri incelendiğinde, ortalama olarak 59950 olan bir grup ve 10017 olan diğer bir grup belirlenmiştir. Bu durum, veri setindeki farklı page life expectancy değerleri arasında belirgin bir ayrımın olduğunu göstermektedir. Bu etiketlemeye göre sayısal değerler, iki farklı kategoriye ayrılmıştır ve bu sayısal ayrımın veri seti hakkında önemli bir bilgi sağladığını göstermektedir.



Chat-GPT tarafından üretilen analiz: "Dengesiz sınıf dağılımı problemine yönelik çözümler arasında kullanılan örnekleme teknikleri olan oversampling ve undersampling yöntemleri kullanılmış. Oversampling sonucunda her iki sınıfın da eşit sayıda örneğe sahip olduğu bir dağılım elde edilmiş. Undersampling yöntemi ise her iki sınıfın da örnek sayısının dengelendiği bir dağılım oluşturmuş. Bu sonuçlar, veri setindeki sınıf dengesizliğini gidermek için etkili yöntemlerin uygulandığını göstermektedir."



Chat-GPT tarafından üretilen analiz: "Veri setindeki öge kümelerinin sık görülen kombinasyonlarını FP-Growth algoritması ile analiz ederek elde ettiğim sonuçlara göre, en yüksek destek değerine sahip olan öge kümeleri sırasıyla (Düşük), (Orta) ve (Yüksek) olarak belirlendi. Bu ögeler arasındaki kombinasyonlara baktığımızda, en yüksek destek değerine sahip olan kombinasyonlar (Düşük, Orta) ve (Düşük, Yüksek) olarak sıralandı. Ayrıca, üç öge kombinasyonu olan (Düşük, Yüksek, Orta) kombinasyonu da önemli bir destek değerine sahip olarak belirlendi."



Chat-GPT tarafından üretilen analiz: 'senin analizin'

Bu veri setinde, FP-Growth algoritması kullanılarak çeşitli item setlerinin ilişkileri incelenmiştir.

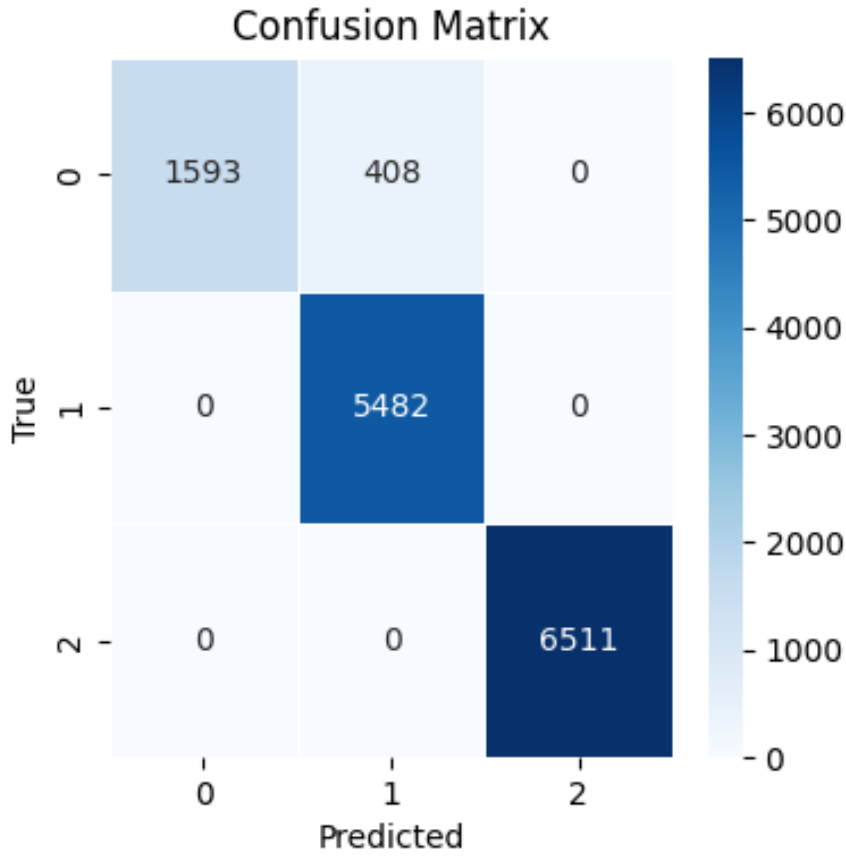
- Düşük seviyedeki verilerin %99.55'i Orta seviyede yer alırken, Orta seviyedeki verilerin %98.48'i Düşük seviyede yer almaktadır. Bu durum, bu iki seviyenin birlikte sıkça görüldüğünü göstermektedir.
- Düşük seviyedeki verilerin %99.55'i Yüksek seviyede yer alırken, Yüksek seviyedeki verilerin %41.03'ü Düşük seviyede yer almaktadır. Yüksek seviye ile Düşük seviye arasında negatif bir ilişki gözlemlenmektedir.

- Yüksek seviyedeki verilerin %41.03'ü Orta seviyede yer alırken, Orta seviyedeki verilerin %98.48'i

Yüksek seviyede yer almaktadır. Bu durum da Yüksek ve Orta seviyelerin birlikte sıkça görüldüğünü göstermektedir.

- Düşük ve Yüksek seviyelerin birlikte görülme oranı %40.62 iken, bu durumun Orta seviyede bulunma olasılığı %41.11 olarak belirlenmiştir.
- Düşük ve Orta seviyelerin birlikte görülme oranı %98.03 iken, bu durumun Yüksek seviyede bulunma olasılığı %41.03 olarak belirlenmiştir.
- Yüksek ve Orta seviyelerin birlikte görülme oranı %39.67 iken, bu durumun Düşük seviyede bulunma olasılığı %41.03 olarak belirlenmiştir.

Bu analiz sonuçlarına göre, Düşük seviye ile Orta seviye arasında en güçlü ilişki gözlemlenirken, Düşük ve Yüksek seviyeler arasında daha zayıf bir ilişki olduğu görülmektedir.



Çalışan algoritma: Random Forest

Chat-GPT tarafından üretilen analiz: Bu sonuçlara göre, Random Forest algoritması oldukça başarılı bir performans sergilemiştir. Confusion matrix değerlerine bakıldığında sınıflandırıcı modelin 3 sınıfı da yüksek doğrulukla tahmin ettiği görülmektedir. Recall, Accuracy ve F1 Score değerleri 0.97 olarak yüksek bir doğruluk seviyesini işaret etmektedir. Precision, Recall ve F1-Score değerleri sınıf bazında incelendiğinde ise genel olarak başarılı bir performans sergilenmiştir. Modelin 1. sınıfı tüm metriklerde yüksek değerlerle tahmin ettiği görülmektedir. Özetle, Random Forest algoritması veri setini etkili bir şekilde sınıflandırmış ve yüksek performans göstermiştir.