



Data Platform
Information Technologies

Sayın AisanKheiri

2024-10-02 17:02:10 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.

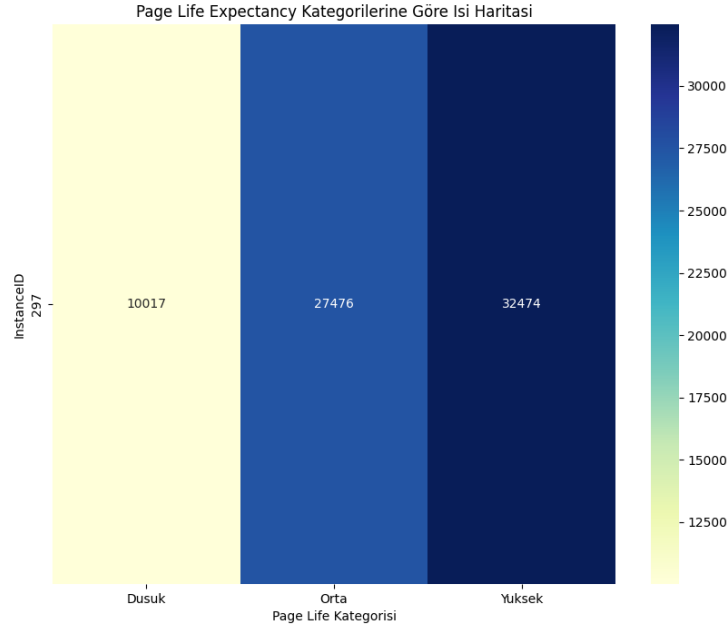
Seçtiğiniz KNN Modeli parametreleri arasında `n_neighbors:None` , `metric:None` ve `weights:None` ,`sutun1:2` , `sutun2:7` , `test_size=0.2` ,`random_state=42` ve Hedef Etiketi `y = df['page life expectancy']` > 300 değerleri bulunmaktadır, bu parametrelerin sonuçları da raporda detaylı bir şekilde açıklanmıştır.

Focus on your business
We'll manage your databases

✉ info@datapatform.com.tr

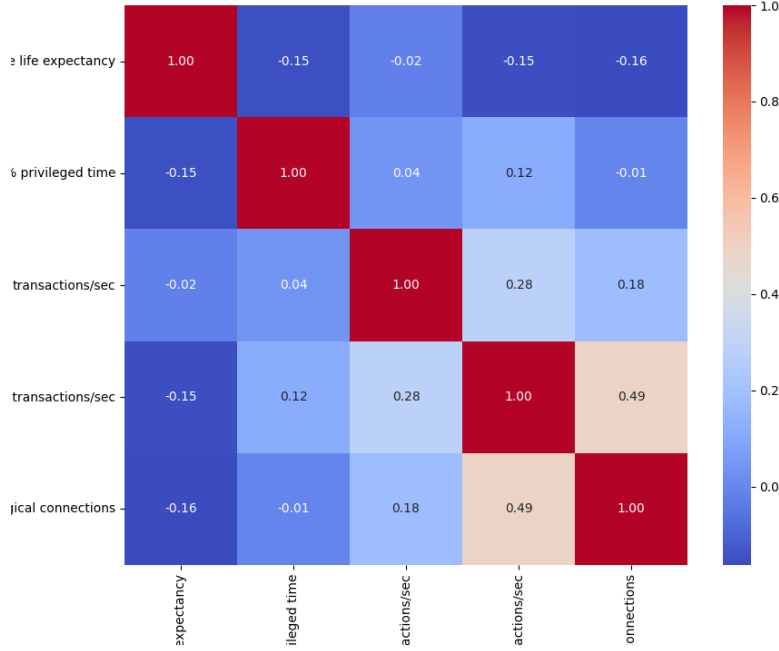
📍 Data Platform Bilgi Sistemleri A.Ş. Bağlarbaşı Mah.
Irmak Sok. No: 15 D:2 K:2 Maltepe/İstanbul

☎ 0532 320 22 39

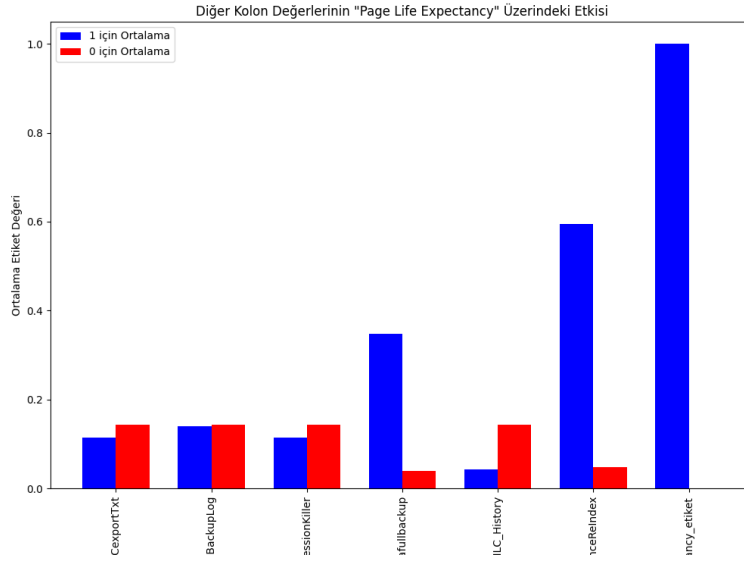


Page_life_category kategorisine göre, InstanceID 297 için elde edilen değerlere göre şu yorumu yapabilirim:

Chat-GPT tarafından üretilen analiz: "InstanceID 297'nin sayfa işlem süresi düşük kategoride olan 10017, orta kategoride 27476 ve yüksek kategoride 32474 değeri bulunmaktadır. Bu verilere göre, sayfa işlem süresi yüksek kategoride en fazla değere sahiptir."



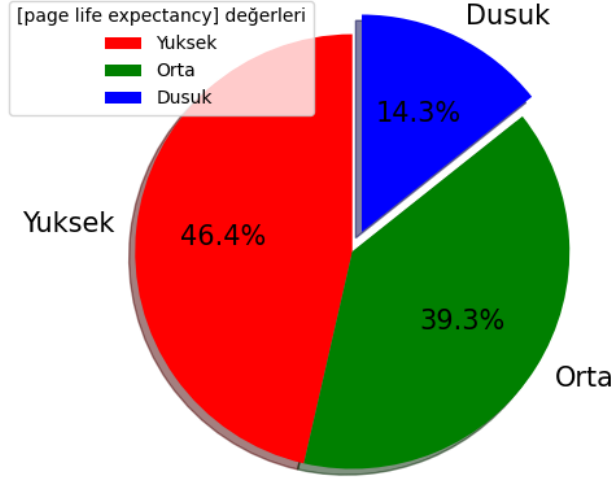
Chat-GPT tarafından üretilen analiz: "Veri setindeki belirli sütunlar arasındaki korelasyon değerleri incelendiğinde, page life expectancy ile diğer sütunlar arasında şu ilişkiler dikkat çekiyor: % privileged time ile düşük negatif bir korelasyon mevcut (-0.150469), transactions/sec ile neredeyse hiç korelasyon bulunmamakta (-0.019418), write transactions/sec ile orta düzeyde negatif bir korelasyon var (-0.154409), logical connections ile düşük negatif bir korelasyon görülüyor (-0.162372). Ayrıca, page life expectancy ile write transactions/sec arasında orta düzeyde pozitif bir korelasyon (0.490174) ve page life expectancy ile logical connections arasında düşük negatif bir korelasyon (-0.162372) gözlemlenmektedir."



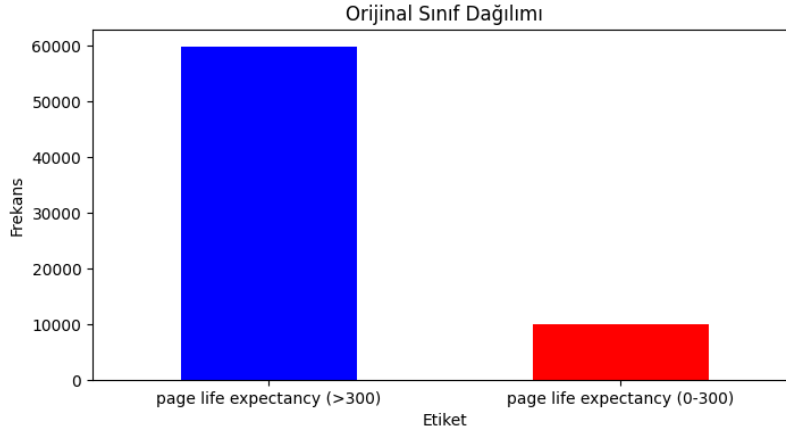
Chat-GPT tarafından üretilen analiz: Veri setindeki belirli sütunların 'page life expectancy' sütunu üzerindeki etkileri incelendiğinde, farklı sütunların bu değer üzerinde farklı etkilere sahip olduğu görülmektedir. Örneğin, 'dbaHCDOCexportTxt' sütunu için ortalama etki 0.113636 iken, 'dbafullbackup' sütunu için bu değer çok daha yüksek olan 0.347987 şeklindedir. Ayrıca, 'OTS_Instrument_OHLC_History' ve 'dbaMaintenanceReIndex' sütunları da farklı etkilere sahiptir.

Ancak, 'page_life_expectancy_etiket' sütununun 1.0 etkiye sahip olması dikkat çekicidir. Bu durum, bu sütunun 'page life expectancy' üzerinde büyük bir etkisi olduğunu göstermektedir. Diğer yandan, 0.0 etkiye sahip sütunlar ise bu değer düşük olduğu anlamına gelmektedir.

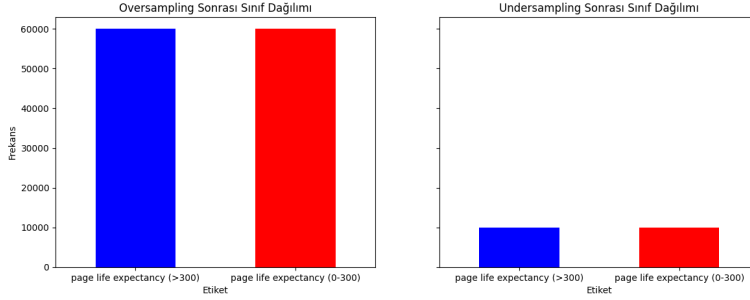
Sonuç olarak, veri setindeki belirli sütunların 'page life expectancy' sütunu üzerindeki etkileri dikkate alındığında, farklı sütunların bu değeri nasıl etkilediği açık bir şekilde görülmektedir.



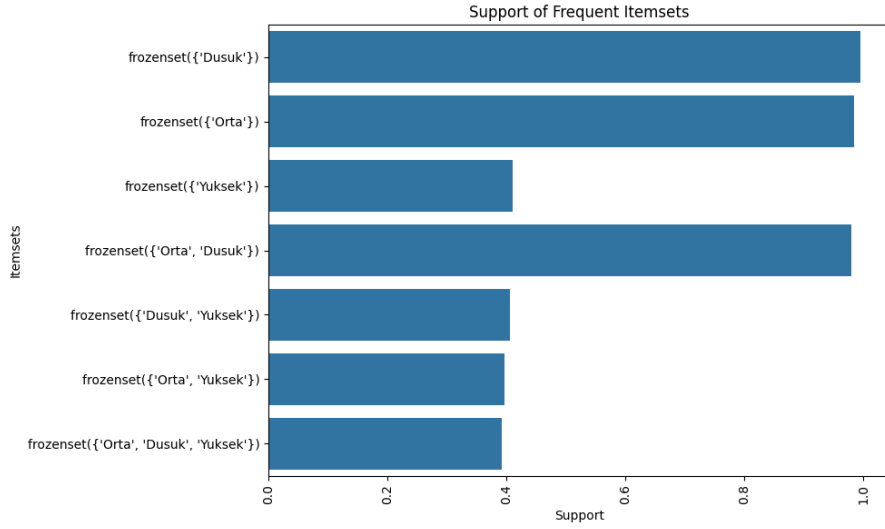
Chat-GPT tarafından üretilen analiz: "Bu pasta grafiğine göre, veri setindeki sayfaların yaşama süreleri çeşitli kategorilere ayrılmıştır. En yüksek kategori olan 'Yüksek' sayfaların gözlem sayısı 32,474 iken, orta kategoride olan sayfaların gözlem sayısı 27,476 ve en düşük kategorideki sayfaların gözlem sayısı ise 10,017'dir. Bu da bize 'Yüksek' kategorideki sayfaların yaşama süreleri genellikle diğer kategorilere göre daha fazla olduğunu göstermektedir."



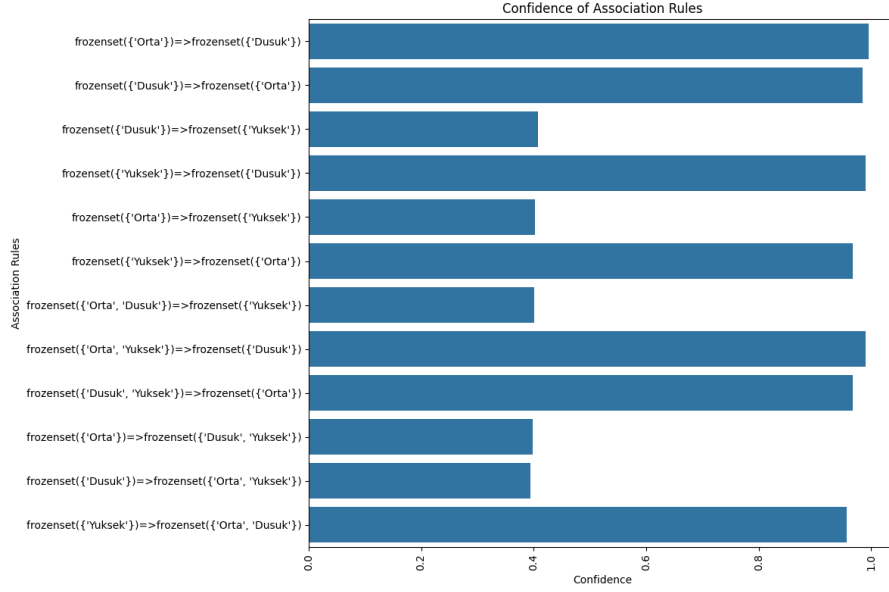
Chat-GPT tarafından üretilen analiz: "Page life expectancy değerlerine göre yapılan etiketlemede, 0 etiketiyle 59950 veri noktası, 1 etiketiyle ise 10017 veri noktası belirlenmiştir. Bu sonuçlar, veri setindeki sayısal dağılımın net bir şekilde gösterilmesini sağlamaktadır. Page life expectancy değerlerine göre elde edilen bu etiketler, veri setindeki farklı değer aralıklarının gruplandırılmasına ve analiz edilmesine yardımcı olabilir. Genel olarak, yüksek sayıda veri noktasının 0 etiketiyle belirlenmiş olması, bu değer aralığına daha fazla örnek veriye sahip olduğunu göstermektedir."



Chat-GPT tarafından üretilen analiz: 'Sınıf dengesizliği sorununu ele almak için gerçekleştirilen örnekleme teknikleri incelendiğinde, Oversampling yapıldıktan sonra her iki sınıfın da eşit sayıda veriye sahip olduğu görülmektedir (0: 59950, 1: 59950). Bu sayede sınıflar arasındaki dengesizlik giderilmiş ve modelin daha adil bir şekilde eğitilmesi sağlanmış olabilir. Diğer yandan Undersampling sonucunda ise her iki sınıfın da veri sayısının azaltıldığı ancak hala dengeli bir dağılımın korunduğu görülmektedir (0: 10017, 1: 10017). Bu yöntem ile modelin ezici çoğunluktaki sınıfa karşı dengeli bir şekilde eğitilmesi hedeflenmiş olabilir.'



Chat-GPT tarafından üretilen analiz: Bu veri setinde en sık görülen kombinasyonlar incelendiğinde, "Dusuk" seviyesine sahip olan öğelerin %99.55'i, "Orta" seviyesine sahip olan öğelerin %98.48'i ve "Yuksek" seviyesine sahip olan öğelerin %41.03'ü bulunmaktadır. Ayrıca, "Orta, Dusuk" kombinasyonu %98.03, "Dusuk, Yuksek" kombinasyonu %40.62, "Orta, Yuksek" kombinasyonu %39.67 ve "Orta, Dusuk, Yuksek" kombinasyonu ise %39.26 support değerlerine sahiptir. Bu değerlerin göz önüne alındığında, "Dusuk" seviyesinin diğer seviyelerle sık olarak kombinasyon halinde olduğu gözlemlenmektedir.



Chat-GPT tarafından üretilen analiz: 'senin analizin'

Bu veri seti üzerinde yapılan FP-Growth algoritması analizi, Düşük, Orta ve Yüksek seviyedeki öğeler arasındaki ilişkileri değerlendirmektedir. Özellikle Orta seviye ile Düşük/Yüksek seviyeler arasında çeşitli ilişkilerin incelendiği görülmektedir.

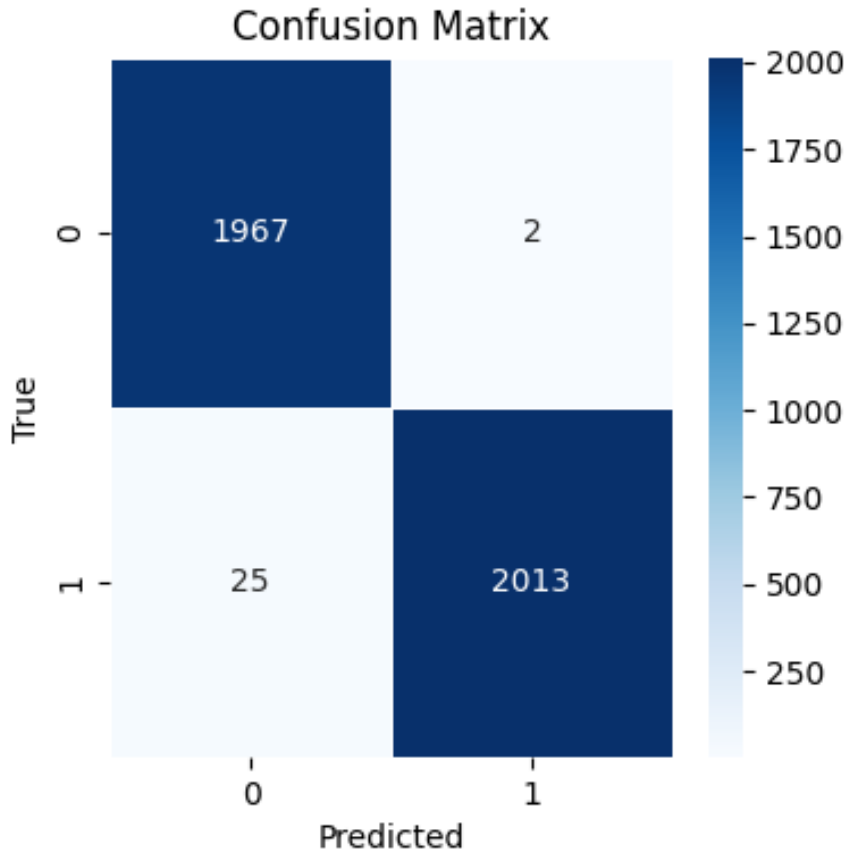
Analiz sonuçlarına göre, Orta seviye ile Düşük seviye arasında bir karşılıklı ilişki belirlenmiştir. Bu ilişki, bir öğenin diğerini teşvik etme eğiliminde olduğunu göstermektedir. Benzer şekilde, Düşük seviyeden Yüksek seviyeye geçişte de yüksek bir güven değeri tespit edilmiştir, bu da Düşük seviyedeki bir öğenin Yüksek seviyedeki bir öğeyi etkileyebileceğini göstermektedir.

Orta seviyenin Yüksek seviye ile ilişkisinde de yüksek bir güven değeri tespit edilmiştir. Bu durum,

Orta seviyedeki öğelerin Yüksek seviyedeki öğelerle güçlü bir ilişki içinde olduğunu göstermektedir.

Ayrıca, Orta ve Düşük seviyelerin bir arada olduğu durumlarda Yüksek seviyeye geçişte de güven değerleri yüksek bulunmuştur. Bu durum, Orta ve Düşük seviyedeki öğelerin birlikte Yüksek seviyedeki öğeleri etkileyebileceğini göstermektedir.

Sonuç olarak, analiz sonuçlarından çıkarılan kuralların, veri setindeki farklı seviyelerdeki öğeler arasındaki ilişkileri gösterdiği ve bu ilişkilerin güven düzeylerinin de hesaplandığı görülmektedir. Bu kuralların işletmelerin pazarlama stratejileri, ürün yerleşimleri ve fiyatlandırmaları gibi konularda değerli içgörüler sağlayabileceği düşünülmektedir.



Çalışan algoritma: KNN

Chat-GPT tarafından üretilen analiz: 'Verilerle yapılan analize göre, KNN algoritması yüksek bir doğruluk oranına sahiptir. Confusion Matrix sonuçlarına göre, 2004 örneği doğru bir şekilde sınıflandırırken sadece 25 örneği yanlış sınıflandırmıştır. Recall değeri yani gerçek pozitif oranı oldukça yüksektir (%98.75) ve Accuracy değeri de oldukça yüksektir (%99.38). F1 Score ise %99.37 olarak hesaplanmıştır. Bu sonuçlar, KNN algoritmasının bu veri seti üzerinde başarılı bir şekilde çalıştığını göstermektedir. Precision, Recall ve F1 Score değerleri sınıflandırma performansı hakkında daha detaylı bilgi vermektedir. Veri setindeki 'False' ve 'True' sınıflarına ait metrikler incelendiğinde, precision ve recall değerlerinde yüksek oranlar gözlemlenmektedir. Sonuç olarak, KNN algoritması bu veri setinde başarılı bir şekilde çalışmıştır.'

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