

AI Özeti

Bu veri seti, 2246 satır ve 12 kolon içermektedir. Toplamda yalnızca 1 eksik hücre bulunmaktadır. En fazla eksik veriye sahip kolon ise SAYAÇ NO'dur. Genel olarak veri seti, analiz ve modelleme için uygun bir yapı sunmaktadır.

Riskler

Veri kalitesi ile ilgili başlıca riskler şunlardır: 1. SAYAÇ NO kolonundaki eksik veri, analiz sonuçlarını etkileyebilir. 2. Eksik verinin nedenleri araştırılmadığı takdirde, modelin genelleme yeteneği zayıflayabilir. 3. Veri setinin dengesizliği, bazı sınıfların yeterince temsil edilmemesi durumunda model performansını olumsuz etkileyebilir.

Feature Engineering Önerileri

Özellik mühendisliği için öneriler: 1. SAYAÇ NO kolonundaki eksik verilerin doldurulması için uygun bir yöntem belirlenmelidir. 2. Diğer kolonlarla ilişkili yeni özellikler oluşturulabilir. 3. Zaman serisi analizi yapılacaksa, tarih ve saat bilgileri üzerinden ek özellikler türetilebilir.

Uygun ML Modelleri

Bu veri seti için uygun makine öğrenmesi modelleri arasında regresyon analizi, karar ağaçları ve rastgele ormanlar bulunmaktadır. Örneğin, SAYAÇ NO'y'u tahmin etmek için rastgele orman modeli kullanılabilir. Modelin eğitimi için eksik verilerin uygun bir şekilde işlenmesi gerekmektedir.

Aksiyon Önerileri

Şirket için öneriler: 1. Eksik verilerin nedenlerini araştırmak ve gerekli düzeltmeleri yapmak. 2. Veri setinin daha iyi analiz edilebilmesi için özellik mühendisliği çalışmalarına odaklanmak. 3. Modelleme sürecinde farklı algoritmaların performansını karşılaştırmak ve en iyi sonucu veren modeli seçmek.