

2.3.1. Veri Özgilemeye Genel Bakış

• Veri Temizleme (Data Cleaning / Cleansing)

- Gürcültölöl Veri (Noisy Data)

Örneğin; veri seti içerisinde anniyati erkek dan bir kızının aynı zamanda hamilelikle ilgili bir değışkende hamile olduğunun ifade edilmesi. Bu mesela gürcültölöl, yanlış bir veridir. Örneğin bir bebek kiberonunun fiyatının 1 milyon TL olması. Bu aykırı bir durum değil gürcültölöl, bozuk bir durumdur.

- Eksik Veri Analizi (Missing Data Analysis)

- Aykırı Gösterim Analizi (Outlier Analysis)

• Veri Standartizasyonu (Data Standardization, Feature Scaling)

- 0-1 Dönüşümü (Normalization)

- z-skora Dönüşüm (Standardization)

- Logaritmik Dönüşüm (Log Transformation)

gibi elimizdeki değışkenleri, verileri ceptli dönüştürme işlemlerine tabi tutuyoruz.

• Veri İndirgeme (Data Reduction)

- Gözlem (Observation) Sayısının Azaltılması
- Değişken (Variable) Sayısının Azaltılması

• Değişken Dönüşümleri (Variable Transformation)

- Sürekli Değişkeni Kategorik Değişkene Çevirmek
- Kategorik Değişkenlerin Sınıfları Arasında Değişimler Yapmak