## 7.1. Anomali Tespiti:

Veri analizinde, anomali tespiti (aynı zamanda aykırı değer tespiti), verilerin çoğunluğundan önemli ölçüde farklılaşarak şüphe uyandıran nadir öğelerin, olayların veya gözlemlerin tanımlanmasıdır. Tipik olarak anormal öğeler, banka dolandırıcılığı, yapısal bir kusur, tıbbi sorunlar veya bir metindeki hatalar gibi bir tür soruna dönüşecektir. Anormallikler ayrıca aykırı değerler, yenilikler, gürültü, sapmalar ve istisnalar olarak da adlandırılmaktadır.

Özellikle, kötüye kullanım ve ağa izinsiz giriş tespiti bağlamında, ilginç nesneler genellikle nadir nesneler değil, beklenmedik etkinlik patlamalarıdır. Bu model, bir aykırı değerin nadir bir nesne olarak genel istatistiksel tanımına uymaz ve uygun şekilde bir araya getirilmediği sürece birçok aykırı değer algılama yöntemi (özellikle denetimsiz yöntemler) bu tür verilerde başarısız olmaktadır. Bunun yerine, bir küme analizi algoritması, bu modellerin oluşturduğu mikro kümeleri tespit edebilmektedir.

Üç geniş anomali tespit tekniği kategorisi mevcuttur[4]. Denetimsiz anomali tespit teknikleri, veri setindeki örneklerin çoğunluğunun normal olduğu varsayımı altında, veri setinin geri kalanına en az uyan örnekleri arayarak etiketlenmemiş bir test veri setindeki anormallikleri tespit etmektedir. Denetimli anomali tespit teknikleri, "normal" ve "anormal" olarak etiketlenmiş bir veri seti gerektirir ve bir sınıflandırıcının eğitimini içermektedir (diğer birçok istatistiksel sınıflandırma probleminden temel fark, aykırı değer tespitinin doğal dengesiz doğasıdır). Yarı denetimli anomali tespit teknikleri, belirli bir normal eğitim veri setinden normal davranışı temsil eden bir model oluşturur ve ardından kullanılan model tarafından bir test örneğinin oluşturulma olasılığını test etmektedir.