**Bu staj çalışması, Pavza bünyesinde 28/07/2025 – 22/08/2025 tarihleri arasında, bilgisayar mühendisliği 3. sınıf öğrencisi olarak gerçekleştirilmiştir. Stajın amacı, akademik süreçte edinilen teorik bilgileri gerçek veri setleri ve algoritmalar üzerinde uygulayarak pekiştirmek, veri bilimi ve makine öğrenmesi alanlarında pratik deneyim kazanmaktır.**

**Staj süresince, 10 farklı veri seti üzerinde çeşitli kümeleme (clustering) algoritmalarının performans analizleri gerçekleştirilmiştir. Bu kapsamda K-Means, Hiyerarşik Kümeleme (Hierarchical Clustering) ve DBSCAN gibi yöntemler kullanılarak her veri seti ayrı ayrı işlenmiş, elde edilen sonuçlar ARI (Adjusted Rand Index), V-Measure ve NMI (Normalized Mutual Information) gibi metriklerle değerlendirilmiş ve karşılaştırılmıştır. Analiz sonuçları, görselleştirme teknikleri ile desteklenerek görevli takım liderine sunulmuştur.**

**Ayrıca, literatür taraması yapılarak KNN (K-Nearest Neighbors) algoritmasında NaN değerleri dikkate alan bir yaklaşım incelenmiş ve bu yöntem kullanılarak ilgili veri setleri yeniden analiz edilmiştir. Elde edilen sonuçlar, farklı algoritmalar ve veri ön işleme teknikleri arasındaki performans farklarını ortaya koyacak şekilde raporlanmıştır.**

**Bu staj süreci, yalnızca teknik becerilerin değil, aynı zamanda araştırma yapma, verileri yorumlama, sonuçları raporlama ve sunum becerilerinin de gelişmesine önemli katkı sağlamıştır.**