

## ÖZET

### SAĞLIK ALANINDA YAPILAN ARAŞTIRMALARDA KÜMELEME ALGORİTMALARININ KULLANIMI: BİR UYGULAMA

Özge PASİN

Yüksek Lisans Tezi, Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı

Tez Danışmanı: Prof. Dr. Handan ANKARALI

EYLÜL 2015, 240 Sayfa

Kümeleme yöntemleri ile benzer özelliklere sahip değişken ve bireyler bir grupta toplanabilmektedir. Birçok uygulama alanına sahip olmasına rağmen kümeleme yöntemi ülkemizde sağlık araştırmalarında nadir olarak kullanılmaktadır. Bu tez çalışmasının amacı, farklı kümeleme algoritmalarını tanıtmak ve bu algoritmaların nasıl ve hangi durumlarda doğru bir şekilde kullanılabileceğini göstermektir. Aynı zamanda sağlık alanından elde edilmiş gerçek bir veri seti üzerinde uygulanabilir olan farklı kümeleme algoritmalarının sonuçlarını karşılaştırmaktır. Yapılan değerlendirmeler sonucunda kullanılan iki farklı veri seti için hesaplanan kappa katsayıları istatistiksel olarak orta düzeyde anlamlı bulundu. Gerçekleştirilen uygulama sonucunda her iki veri seti için de kappa katsayısı bakımından en uygun ve en hızlı sonuçlar üreten algoritmanın *En Uzak İlk Kümeleme Yöntemi* olduğu sonucuna varıldı. Framingham risk grupları ile oluşturulan kümeler arasında çapraz tablolar oluşturularak grupların dağılımı incelendiğinde ise, en isabetli kararların *Make Density Based* ve *EM algoritmalarıyla* elde edilen kümeleme sonuçları olduğu görüldü. Sonuç olarak kümeleme yöntemlerinin hastalıklara ait risk faktörlerinin incelenmesinde, klinik bilgileri de dikkate alarak hastalık gruplarının oluşturulmasında ve buna bağlı olarak da doğru hastalık teşhislerinin konulmasında önemli bir rol oynayacağı düşünülmektedir. Ayrıca veri dağılımı ve özellikleri dikkate alınarak kullanıldığında kümeleme algoritmalarının, sağlık alanında her türlü planlama ve hastalık teşhisi için bir tanı aracı olarak kullanılabileceği kanısındayız.

**Anahtar Sözcükler:** Algoritma, Framingham risk skoru, Kümeleme Analizi, Veri Madenciliği