

TÜRKİYE İLLERİNİN SOSYO-EKONOMİK GELİŞMİŞLİK DÜZEYLERİNE GÖRE MAKİNE ÖĞRENMESİ YÖNTEMLERİYLE KÜMELENMESİ

Emircan Demir (241307109) | Danışman: Prof. Dr. Hikmet Hakan Gürel | Bilişim Sistemleri Müh.

ÖZET

Bu çalışmada, Türkiye'deki 81 ilin sosyo-ekonomik gelişmişlik düzeyleri, TÜİK ve TCMB verileri kullanılarak makine öğrenmesi yöntemleriyle analiz edilmiştir. Çalışmada nüfus, eğitim, sağlık, ekonomi, maliye, altyapı ve yaşam kalitesi kategorilerinde toplam 24 değişken kullanılmıştır. Veri analizi sürecinde K-Means ve Hiyerarşik Kümeleme algoritmaları uygulanmış, optimum küme sayısı Elbow ve Silhouette yöntemleriyle 5 olarak belirlenmiştir. Sonuçlar; coğrafi bölgelerden bağımsız olarak illerin gelişmişlik düzeylerine göre anlamlı gruplar oluşturduğunu ve resmi SEGE-2022 raporlarıyla yüksek tutarlılık gösterdiğini ortaya koymuştur. Elde edilen bulgular, bölgesel kalkınma politikalarının oluşturulmasında veri odaklı bir karar destek mekanizması sunmaktadır.

GİRİŞ VE AMAÇ

Bölgelgelmişlik farklarının doğru tespiti, sürdürülebilir kalkınma politikalarının etkinliği için kritiktir. Geleneksel yöntemler genellikle tek boyutlu göstergelere dayanırken, bu proje veri madenciliği ve Makine Öğrenmesi (ML) algoritmaları ile çok boyutlu ve nesnel bir sınıflandırma sunmayı amaçlar.

Çalışmanın Temel Amaçları:

- Çok boyutlu sosyo-ekonomik veriyi işleyerek gizli örüntülerini ortaya çıkarmak.
- İllerin benzerliklerine göre gruplamak ve gelişmişlik harmasını çıkarmak.
- Her kümenin (cluster) karakteristik özelliklerini belirleyerek politika önerileri sunmaktır.

METODOLOJİ

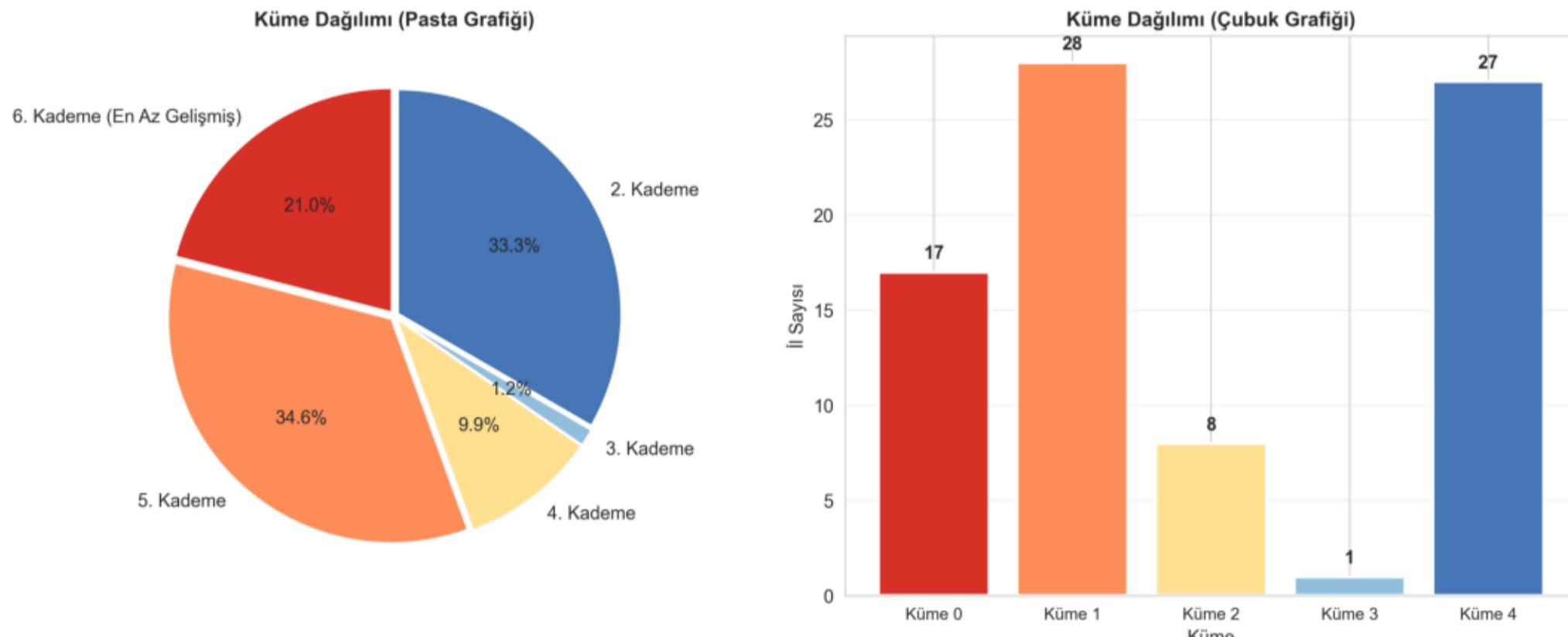
Veri Kaynakları ve Seti:

- Kaynaklar: TÜİK (2023-2024), TCMB, Sanayi ve Teknoloji Bakanlığı.
- Değişkenler: GSYH, Okulleşme Oranı, Hastane Yatak Sayısı, İhracat, Elektrik Tüketimi vb. (Toplam 24 Değişken).

Yöntemler:

- Ön İşleme: Eksik Veri Tamamlama, Aykırı Değer (IQR) Analizi, Min-Max Normalizasyon.
- Algoritmalar: K-Means (Euclidean Distance), Hiyerarşik Kümeleme (Ward Linkage).
- Değerlendirme: Elbow Metodu (Kırılma Noktası), Silhouette Skoru (0.348 - Orta/İyi Ayışma).

Türkiye İilleri Kümeleme Sonuçları



Şekil 1: K-Means ile İllerin Küme Dağılımı

BULGULAR VE SONUÇ

Analiz sonucunda iller 5 farklı gelişmişlik kümeye ayrılmıştır:

- Küme 0 (1. Derece - Metropoller): İstanbul, Ankara, İzmir. (Sanayi, Hizmet ve Eğitim Merkezi - Ülkenin Lokomotifleri)
- Küme 1 (2. Derece - Sanayi/Turizm): Bursa, Kocaeli, Antalya, Eskeşehir, Tekirdağ, Muğla. (Yüksek Üretim ve İhracat Potansiyeli)
- Küme 2 (3. Derece - Gelişen İller): Konya, Kayseri, Gaziantep, Manisa, Denizli, Sakarya, Balıkesir, Aydın, Mersin, Adana. (Sanayi-Tarım Dengesi)
- Küme 3 (4. Derece - Orta Gelir): Trabzon, Samsun, Çanakkale, İsparta, Bolu, Edirne, Kırklareli, Afyon, Kütahya, Sivas, Malatya, Elazığ. (Tarım ve Hizmet Odaklı)
- Küme 4 (5. ve 6. Derece - Öncelikli): Erzurum, Van, Diyarbakır, Şanlıurfa, Mardin, Batman, Ağrı, Kars, İğdır, Hakkari, Şırnak, Muş, Bitlis. (Altyapı Yatırımı Öncelikli)

SONUÇ: Elde edilen kümeleme yapısı, SEGE-2022 sonuçları ile yüksek korelasyon göstermektedir. İllerin sadece coğrafi konumlarına göre değil, yapısal sosyo-ekonomik özelliklerine göre de benzetiği istatistiksel olarak kanıtlanmıştır.

KAYNAKÇA

- [1] T.C. Sanayi ve Teknoloji Bakanlığı, SEGE-2022 Raporu. [2] TÜİK, İl Göstergeleri, 2024. [3] MacQueen, J. (1967). Classification Methods.