

GİRİŞİMCİLİĞİN İLLERE GÖRE DAĞILIMI: BULANIK K ORTALAMALAR VE SOM KÜMELEME YÖNTEMLERİ*

Distribution of Entrepreneurship by Cities:
Fuzzy K Means and SOM Cluster Methods

Gönderim Tarihi / Received: 17.02.2020

Kabul Tarihi / Accepted: 24.11.2020

Doi: <https://doi.org/10.31795/baunsobed.690031>

Kasım Can IŞIK**1

Mesut POLATGİL²

ÖZ: Sürekli değişen pazar koşulları ve gelişen teknoloji bağlamında girişimcilik faaliyetlerin önemi özellikle Türkiye gibi gelişmekte olan ülkelerde giderek artmaktadır. Girişimcilik faaliyetlerinin benzer özellikleri taşıyan yerlere göre kümelenmesi de büyük önem arz etmektedir. Çünkü benzer özellikleri taşıyan ve sayısı; il, kalkınma bölgesi ve coğrafi bölgelere göre daha az olan bu kümeler gerek girişimcilik faaliyetlerinin teşvik edilmesi ve destek sağlanmasında gerekse girişimcilik faaliyetlerinde ortaya çıkan sorunların belirlenmesi ve çözümünde önemli bir veri kaynağıdır. Literatürde Türkiye'deki tüm illeri girişimcilik özelliklerine göre kümeleyen herhangi bir çalışma bulunmamakla birlikte bu çalışma ile literatürdeki bu açığı kapatmak amaçlanmıştır. Türkiye Sanayi ve Teknoloji Bakanlığı GBS (Girişimci Bilgi Sistemi)'nde yer alan "İllere göre girişim sayısı, net satışlar ve aktifler" raporları çalışmada veri olarak kullanılmıştır. Çalışma kapsamında 2017 yılına ilişkin GBS veri tabanında yayımlanan verilere kümeleme yöntemlerinden Bulanık k ortalamalar ile SOM YSA(Yapay Sinir Ağları) yöntemleri uygulanarak Türkiye'nin girişimcilik haritası çıkarılmıştır. Bulanık k ortalamalar yöntemine göre göre 19 il 3 farklı kümede, kalan 62 il tek kümede toplanmışken; SOM kümeleme yöntemine göre ise 8 il 4 farklı kümede, kalan 73 il ise tek kümede toplanmıştır. Son olarak ise dikkat çeken bulgulara yer verilmiş, 2 yönteme ait sonuçlar arasındaki benzerlik ve farklılıklar yorumlanmıştır.

Anahtar Kelimeler: Girişimcilik, Kümeleme Analizi, Yapay Sinir Ağları.

* Analizler Sanayi ve Teknoloji Bakanlığı - Girişimci Bilgi Sistemi'nden alınan verilere uygulandığı için etik kurul raporu alınmamıştır.

** Sorumlu Yazar / Corresponding Author

¹ Dr., Sivas Sanayi ve Teknoloji Müdürlüğü, canm001@gmail.com,
<https://orcid.org/0000-0002-9637-8889>

² Dr., Sivas Milli Eğitim Müdürlüğü, mesutbiyan@gmail.com, <https://orcid.org/0000-0002-7503-2977>

ABSTRACT: Due to constantly changing market conditions and developing technology; The importance of entrepreneurial activity is increasing, especially in developing countries such as Turkey. Likewise, the clustering of entrepreneurial activities according to the places with the same features is becoming more and more important. Because the number of these clusters is less than the provinces, development regions and geographical regions. But because of the common features they have, these clusters; It is an important data source both in promoting / supporting entrepreneurship activities and in identifying / resolving problems that arise in entrepreneurship activities. There are no studies in the literature that cluster all cities according to their entrepreneurial characteristics, with this study, it is aimed to close this gap in the literature. Turkey Ministry of Industry and Technology / Entrepreneur Information System (GBS) located in the “Number of enterprises by cities, net sales and assets” were used as data in the study. Fuzzy k means and SOM ANN (Artificial Neural Networks) methods, one of the clustering methods, have been applied to the published data of 2017 in the GBS database. and Turkey’s entrepreneurship is mapped. According to the fuzzy k means method, 19 cities were collected in 3 different clusters and 62 cities were collected in a single cluster. According to SOM clustering method, 8 cities were collected in 4 different clusters and 73 cities were collected in a single cluster. Finally, remarkable findings have been explained and the similarities / differences between the results of the 2 methods have been interpreted.

Keywords: Entrepreneurship, Cluster Analysis, Artificial Neural Networks.

GİRİŞ

Ekonomik rekabetin ön planda olduğu günümüzde, girişimciler ekonomik gelişimin motoru olarak görülmektedir. Bu nedenle girişimcilik pek çok araştırmanın da ilgi odağı olmuştur (Güney ve Çetin, 2003: 191). Girişimcilik köken olarak, Fransızca “entreprendre” ve Almanca’da “unternehmen” kelimelerinden gelmekteyken, Türkçe’de ise “üstlenmek” anlamına gelmektedir (Wickham, 1998: 4). Girişimcilik kavramının tanımları içerisinde en çok kullanılanlar şunlardır:

- Yenilikler oluşturarak, yaratıcı metodlar kullanarak veya başka bir yöntemle, yeni; ürün, teknoloji, piyasa, hizmet ve kaynaklara ulaşım, değerleri elde etmektir
- Değer yaratmak amacıyla, kar etmeyi amaçlayan yeni bir işletmeyi kurma, büyütme ya da yeni bir hizmet veya mal yaratma sürecine verilen isimdir
- Fırsat elde edebilmek için alternatif üretim adımlarını inceleme ve bunların optimizasyonunu sağlama faaliyetlerine verilen isimdir
- Bir fırsatı algılama ve o fırsatı elde etmek için bir organizasyon yaratma faaliyetidir

- Kişilere ve topluluklara değer yaratan, ekonomik fırsatları cevaplayan olan veya bunları yaratan, kişiler tarafından oluşturulan, elde ettiği yeniliklerle ekonomik yapıda bazı farklılıklar doğuran bir süreçtir (Yılmaz ve Sümbül, 2009: 196).

Başarılı bir girişimcide olması beklenen tipik özellikler: Yenilikçilik, risk almadaki yetenek, pazarlama becerisi, know-how üretebilme, piyasa fonksiyonlarının nasıl işlediği hakkında bilgi sahibi olma, işbirliği yapma (co-operate) yeteneği ve işletme yönetimi becerisi olarak sıralanabilir (Littunen, 2000:295).

Gelişmiş ülkelerde, günden güne artan girişimcilik faaliyetleri, durgun endüstrilere hareket getirmiş ve yeni iş imkanları oluşturarak, işsizlik problemlerini azaltma eğilimi oluşturmuştur. Hatta girişimcilik faaliyetleri teknolojik ilerlemeyi hızlandıran en büyük faktörlerden biri olarak yeniden keşfedilmiştir. Bundan dolayı, az gelişmiş veya gelişmekte olan ekonomilerde girişimcilik faaliyetlerinin desteklenerek geliştirilmesinin sağlanması, ekonomik büyümenin önemli bir adımı olarak kabul edilmeye başlanmıştır (Özkan, 2003: 149).

Literatürde Türkiye'deki tüm illeri girişimcilik özelliklerine kümeleyen herhangi bir çalışma bulunmamakla birlikte bu çalışma ile literatürdeki bu açığı kapatmak amaçlanmıştır.

Literatür incelendiğinde konuyla ilgili benzer olan az sayıda çalışma olduğu görülmüştür. Bu çalışmalarda örneklemin sadece bölgesel olarak ele alındığı, ülkelerin girişimci davranışlara göre kümelendiği, Türkiye çapında Ar-ge ve yenilik performansının değerlendirildiği ve 2 bölgenin girişimcilerinin karşılaştırıldığı görülmüştür. Bu çalışmaya benzer örnek çalışmalar aşağıda yer almaktadır.

- Güneş ve İncekırık 2016 yılındaki çalışmalarında; Ege bölgesinde faaliyet gösteren KOSGEB kapsamındaki farklı ölçeklerdeki şirketleri bulanık kümeleme analizi ile gruplandırmışlardır.
- Tokatlıoğlu ve Yalçın 2019 yılındaki çalışmalarında; toplamda 60 ülkeyi girişimci davranışına ilişkin bireysel algılara göre Hiyerarşik kümeleme analizi (Ward) ve Hiyerarşik olmayan kümeleme analizi (k ortalama) yöntemleriyle kümelemişlerdir
- Belgin ve Avşar 2019 yılındaki çalışmalarında; Türkiye'de yer alan bölgeler ve iller seviyesinde Ar-ge ve yenilik performansı kıyaslaması yapılmış olup yöntem olarak ise çok kriterli karar verme tekniklerinden biri olan gri ilişkisel analiz yöntemi kullanılmıştır.
- Özkan 2003 yılındaki çalışmasında; Türkiye'de girişimcilik ve belirleyicileri Marmara ve Doğu Anadolu bölgesi üzerinden

değerlendirilmiştir. Türkiye'deki en gelişmiş ve en gelişmemiş bu 2 bölgesindeki girişimcilerin ortak ve farklı yönleri ortaya konmuştur.

YÖNTEM

Çok değişkenli analiz tekniklerinden biri olan kümeleme analizi, birey, nesne, ya da herhangi türde birimlerin birbirlerine olan yakınlık ya da uzaklıkları esas alınarak farklı gruplara ayrılması işlemidir. Yani kümeleme analizinin amacı birimleri tüm değişkenler bakımından benzer kümelere atamak, bu kümeleri tanımlamak ve istenildiğinde yeni bir birimin hangi kümede olduğunu tespit edebilmektir. Kümeleme işlemi sonucunda oluşturulan kümeler yüksek seviyede küme içi homojenlik içerirken yüksek seviyede kümeler arası heterojenlik içerirler (Işık ve Çamurcu, 2007: 32-33; Alpar, 2017: 305-317).

Kümeleme analiz yöntemleri genellikle 3 grup altında toplanmaktadır. Bunlar; hiyerarşik kümeleme, hiyerarşik olmayan kümeleme ve kohonen ağlarıdır. Hiyerarşik kümeleme yöntemleri genellikle benzerlik ya da uzaklık değerlerine göre birimleri bir araya getirme ya da ayırıştırma yöntemini kullanarak kümeleme işlemini gerçekleştirir. Hiyerarşik olmayan kümeleme yöntemlerinde ise başlangıçta belirlenen küme sayısı kadar kümeye birimler uzaklıklarına göre kümelendirilme yapılmaktadır. Kohonen ağları ise yapay sinir ağlarının kümeleme analizi için özelleştirilmiş bir biçimidir (Alpar, 2017: 305-317).

Bu çalışmada hiyerarşik olmayan kümeleme yöntemlerinden bulanık k ortalamalar yöntemi ile, yapay sinir ağlarının özel bir çeşidi olan SOM ağlar kullanılmıştır. Bunun nedeni ise literatürde gerçekleştirilen birçok çalışmada karşılaştırılan bu yöntemler arasındaki farklılıklar olmasıdır

Bulanık k Ortalamalar (Fuzzy k Means)

Klasik mantık işlemlerinde olduğu gibi klasik kümeleme yönteminde, bir birim bir kümeye ya tamamen aittir (üyelik 1'e eşittir) ya da tamamen ait değildir (üyelik 0'a eşittir). Gerçek hayatta birçok alanda olduğu gibi kesinlik yoktur ve kümeleme işleminde de bir birim kesin olarak bir kümeye aittir ya da ait değildir demek de doğru değildir. Bu birimin ilgili kümeye bir aitlik derecesi (üyelik değeri) vardır. Üyelik değeri denilen bu değer 0 ile 1 arasında sonsuz tane değer olabilmektedir ve bu üyelik değerleri toplamı her zaman 1'dir. Bu yüzden kümelenecek birim hangi kümeye daha çok üyelikle aitse o kümeye ait olacaktır. Birimlerin bir küme içerisinde yer almasında öncelik birimlerin birbirine oldukça benzemesi şartıdır (Murat ve Şekerler, 2009: 4763).

Bulanık mantık algoritmalarında, bir elemana kümeleme veya sınıflandırma işlemi yapılırken; elemanın sınıfını belirlemeye ek olarak elemanın o sınıfa ne

kadar ait olduğuna ilişkin bir bilgi de verildiğinden bulanık algoritmalar, klasik algoritmalarla kıyaslandığında daha fazla bilgi içerirler (Höppner vd, 1999).

Bulanık kümeleme yönteminde, öncelikle (1) eşitliğine göre U üyelik matrisine rastgele değer atanır.

$$\sum_{i=1}^c u_{ij} = 1, \forall_j = 1, \dots, n \quad (1)$$

Bulanık k-ortalamlar metodunda kullanılan farklılık fonksiyonu,

$$J(U, c_1, c_2, \dots, c_c) = \sum_{i=1}^c J_i = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2 \quad (2)$$

u_{ij} : 0 ile 1 aralığındaki değer,

c_i : i. kümenin merkezi,

d_{ij} : i. merkez c_i ile j. veri noktası arasında yer alan öklid, karesel öklid, binary öklid, pearson, manhattan, mahalanobis, canberra, minkowski, chebyshev uzaklıkları,

m: Minimum değeri 1 fakat genel olarak 2 değeri kullanan ağırlıklandırma faktörü.

Maliyet fonksiyonunu en az yapmak için iki şart aşağıdaki (3) ve (4) denklemlerinde yer almaktadır:

$$c_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m} \quad (3)$$

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}} \right)^{2/(m-1)}} \quad (4)$$

Bezdek (1973) bulanık k-ortalamlar algoritmasını aşağıdaki şekilde açıklamıştır:

1. (1) denklemini sağlamak koşuluyla U üyelik matrisine rastgele değerler ata,
2. (3) denklemleri aracılığıyla c_i merkezleri hesapla,

3. (2) denklemi aracılığıyla merkezlerle veri noktaları arasındaki farkı bul. Önceki iterasyon için artış eşik değerini altına inerse işlemi durdur,
4. (4) denklemi aracılığıyla tekrardan U matrisini hesapla. Ardından tekrardan adım 2'ye git.

Kohonen Ağları (SOM-Self Organizing Map)

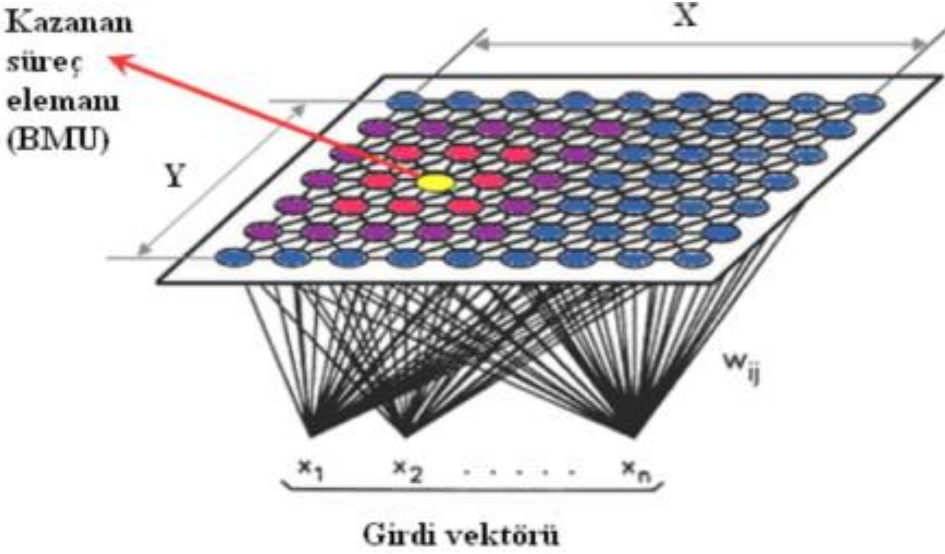
Yapay sinir ağları (YSA), insan beynindeki sinir hücrelerinin (nöron) bilgisayarlarda çeşitli araçlarla matematiksel bir modellenmesidir. SOM (özdüzenleyici haritalar) günümüzde en çok kullanılan Yapay Sinir Ağları'ndan biridir. SOM ağları, 1981 yılında Teuvo Kohonen (1998) tarafından geliştirilmiş ve sinir ağı topluluğuna tanıtılmıştır.

SOM ağları kümeleme işlemini gerçekleştirirken herhangi bir denetime ihtiyaç duymaz yani denetimsiz bir öğrenme gerçekleştirilmektedir. Denetimsiz öğrenme yönteminde birimler sadece kümelendirilmek istendiği için bir çıkış vektörünün olması şart değildir. SOM ağları, veri setindeki elemanları kümelendirmekle kalmaz aynı zamanda bu kümelendirmeyi harita üzerinde de gösterebilir. Bu ağlar veri kaç boyutlu olursa olsun onu iki boyutlu bir haritada gösterebilir.

Şekil 1'de özdüzenleyici haritalardaki girdi ve çıktı katmanı gösterilmiştir. X girdi vektörünü gösterirken; bu vektörde değişkenler yer almaktadır (x_i ifadesinde n adet değişken olduğu ifade edilmektedir). Ağırlık vektörü ise w ile temsil edilmiştir ve her bir değişkenden bütün nöronlara giden ağırlıkları içermektedir. Sarı renkli olan nöron ise kazanan nöronu temsil etmekle birlikte çevresinde yer alan nöronlar ise kazanan nöronun komşularıdır.

Özdüzenleyici haritaların en önemli avantajı, çok boyutlu bir veri setini indirgeyerek iki boyuta dönüştürmesidir. Bu durum karar vericilerin veri setinin tamamını kolayca değerlendirmelerine olanak sağlamaktadır.

Şekil 1: Özdüzenleyici Haritalar Yönteminde Girdi ve Çıktı Katmanı



Kaynak: Bikari, 2017: 125.

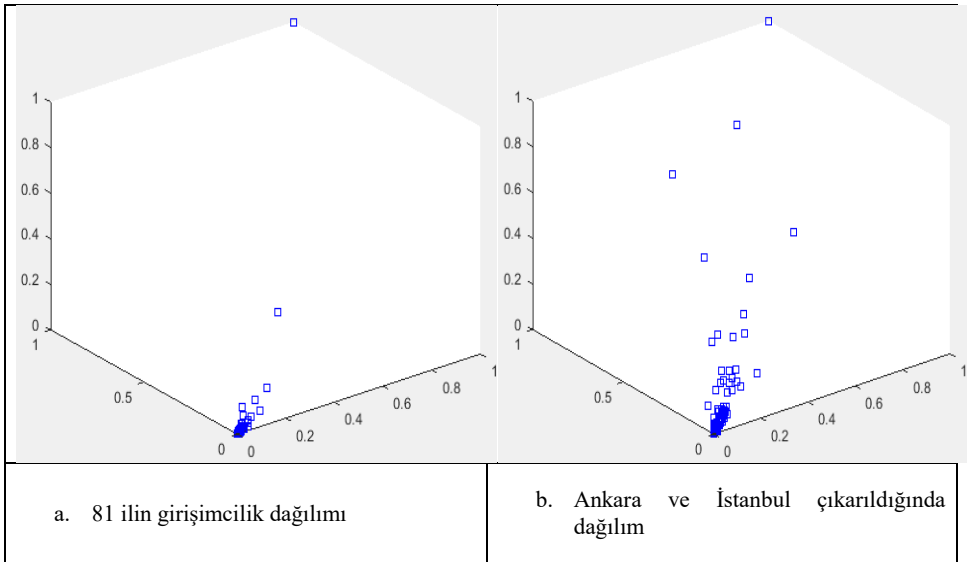
Kullanıcının karar vermesi gereken parametreler arasında, çıktı katmanının boyutu, çıktı katmanının şekli, eğitim algoritması da sayılabilir. Çıktı katmanının boyutu herhangi bir değer alabileceğinden dolayı deneme yanılma yoluyla karar verilebilmektedir.

BULGULAR

Çalışma kapsamında 2017 yılında ülkemizde gerçekleştirilen girişimcilik faaliyetlerine ait veriler GBS' den elde edilmiştir (<https://gbs.sanayi.gov.tr>). GBS; farklı kamu kurum ve kuruluşların idari kayıtlarında bulunan işletmelerin ekonomik faaliyetlerine ilişkin bilgilerin, ortak standartlar içerisinde bir veri tabanında toplanması ve bu verilerin entegrasyon projesi olarak tanımlanmaktadır. GBS'nin temel amacı, kamu kurum ve kuruluşları, araştırma enstitüleri, üniversiteler ve sivil toplum kuruluşlarındaki uzmanlar, karar alıcılar ve araştırmacılara ekonomi, sektörel ve bölgesel bazlı politikaların oluşturulması, uygulanması ve etkinlik ölçümü için mikro ve makro düzeyde güvenilir, geçerli, doğru ve ihtiyacı karşılayabilir verilerin / bilgilerin sunulacağı bir ortam yaratmaktır. Bu sistem sayesinde Türkiye çapında yaklaşık 3 milyon girişimcinin ekonomik faaliyetlerine ilişkin bilgilerine ulaşılabilmektedir (gbs.sanayi.gov.tr, 01.02.2020). Bu kapsamı ve önemi nedeniyle veri seti olarak GBS verileri kullanılmıştır.

GBS’de yer alan “İllere göre girişim sayısı, net satışlar ve aktifler” raporları çalışmada veri olarak kullanılmıştır. 81 ilin verileri GBS’nden anlık olarak alınmadığından araştırmada 2017 yılının verileri tüm iller için eksiksiz olarak alınıp kullanılmıştır. 81 ilin “girişim sayısı”, “net satışlar” ve “aktifler” verilerine göre üç boyutlu olarak çizilen grafik Şekil 3’de verilmiştir. Şekil 3-a’da görüldüğü gibi 2 il diğer bütün illerden ayrılmakta ve oldukça farklılık göstermektedir. Bu illerin Ankara ve İstanbul olduğu tespit edilmiştir. Kümeleme analizi sonuçlarını ciddi oranda etkileyecekleri düşünülerek bu iller analiz başlangıcında zaten ayrı bir küme olarak değerlendirilmiş ve analize geriye kalan 79 il ile devam edilmiştir. Ankara ve İstanbul şehirlerinin analizin başlangıcında ayrı tutulmasında yer alan bir diğer unsur da veri setinde aykırı bulunan gözlemlerin genellikle veriden çıkarılmasıdır. Bu çalışmada kümeleme algoritmaları kullanıldığı için verilerin çıkarılması mümkün değildir bu nedenle aykırı gözlemler ayrı bir küme olarak başlangıçta kabul edilmiştir. Diğer illere göre oldukça farklı özelliklere sahip olan Ankara ve İstanbul veri setinden ayrıldığına geriye kalan 79 ilin dağılımı ise Şekil 2-b’de gösterilmiştir.

Şekil 2: Türkiye İllerin Girişimcilik Verilerine Göre Dağılımı



Veriler arasında ölçüm birimleri olarak ciddi farklılıklar olduğu için kümeleme analizi başlatılmadan önce verilere min-max normalizasyonu denklem (5)’e göre uygulanmıştır. Böylece ölçüm farklılıkları verilerin 0 ile 1 arasında olması sağlanarak ortadan kaldırılmıştır. Bu normalizasyon işlemi literatürde kümelendirme çalışmaları da dahil olmak üzere birçok çalışmada kullanılmıştır (Atalay ve Öztürk, 2016; Çınaroğlu ve Bulut, 2018).

$$z = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (5)$$

Bulanık K Ortalamalar Kümeleme Yöntemi Bulguları

Türkiye illerinin girişimcilik verilerine göre kümelendirilmesi işlemi öncelikle Bulanık k Ortalamalar yöntemi ile matlab programı aracılığıyla gerçekleştirilmiştir. Küme sayısının belirlenmesinde Silhouette ölçütünden faydalanılmıştır. Bu ölçüt Denklem 6'da gösterildiği gibi küme içerisindeki her birimin küme elemanları uzaklığı ile diğer kümelerde bulunan elemanlara uzaklığını baz alan bir ölçüttür. -1 ile 1 arasında yer almaktadır. Bu yöntemde başlangıçta kaç küme olacağına algoritmaya girilmesi gerekmektedir. Bu yüzden algoritma farklı küme sayıları ile çalıştırılmış ve Tablo 1'de görüldüğü gibi Silhouette değeri 1'e en yakın olan küme sayısı 3 olduğundan başlangıçtaki küme sayısı 3 olarak seçilmiştir.

a(i): i. birimin kendi kümesindeki tüm noktalara olan ortalama uzaklıkları

b(i): i. birimin diğer kümelerdeki tüm noktalara olan ortalama uzaklıkların minimumu olmak üzere

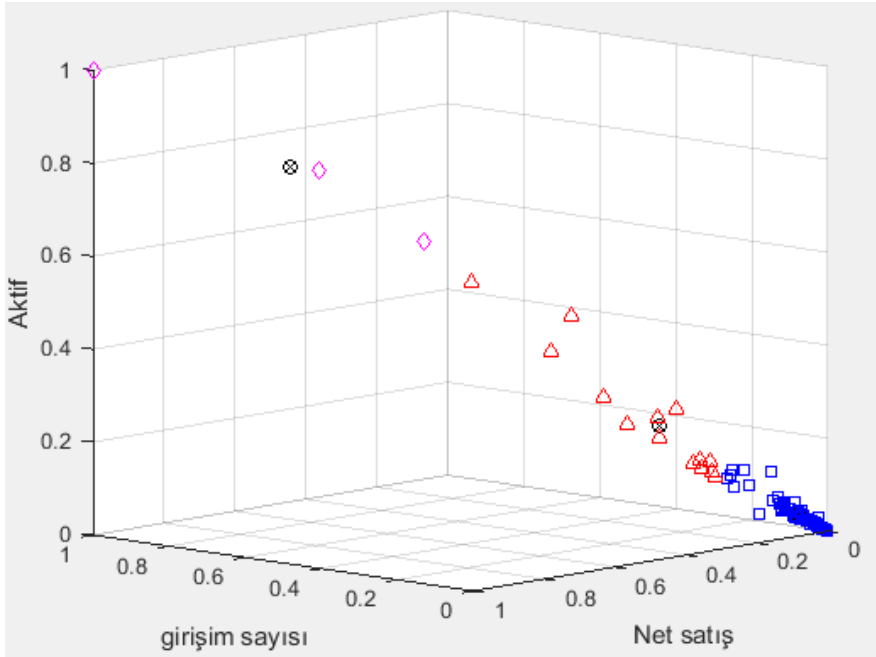
$$sil(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (6)$$

Tablo 1: Farklı Küme Sayılarına Göre Küme Uzaklık Değerleri

Küme sayısı	İterasyon	Silhouette Ölçütü
2	5	0.66571
3	2	0.82961
4	2	0.78235
5	2	0.51156

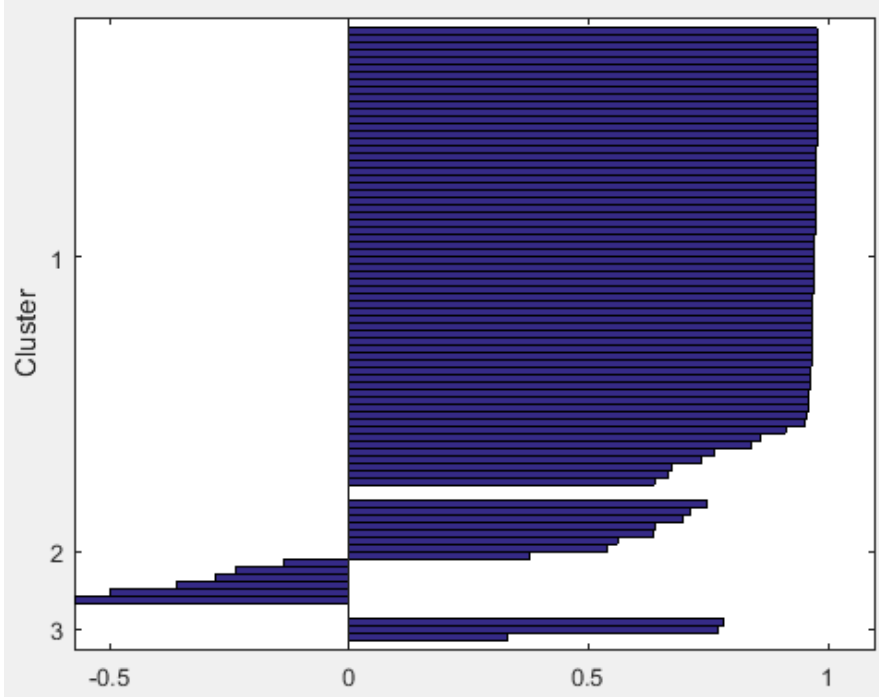
Bulanık k ortalamalar algoritması başlangıç küme sayısı 3 olarak çalıştırılmış ve veriler 3 kümeye ayrılmıştır. Verilerin bu 3 kümeye göre dağılımı Şekil 3'de verilmiştir.

Şekil 3: Bulanık k Ortalamalar Yöntemine Göre İllerin Kümelenmesi Grafiği Sonuçları



Bulanık k ortalamalar yöntemine göre illerin kümelenmesi grafiği sonuçları Şekil 4'de verilmiştir. Bulanık k ortalamalar yönteminin yaptığı kümelemenin başarısını göstermek amacı ile Silhouette grafiği ise Şekil 5'de yer almaktadır. Bu grafik kümeleme işleminin başarısının görsel olarak görülmesini sağlar. Bu grafik kümeleme işlemi sonuçları bakımından incelendiğinde; bir küme hariçinde genel olarak kümelemenin başarılı bir şekilde gerçekleştiği görülmüştür

Şekil 4: Bulanık K Ortalamalar Kümeleme Yöntemine Göre Silhouette Grafiği



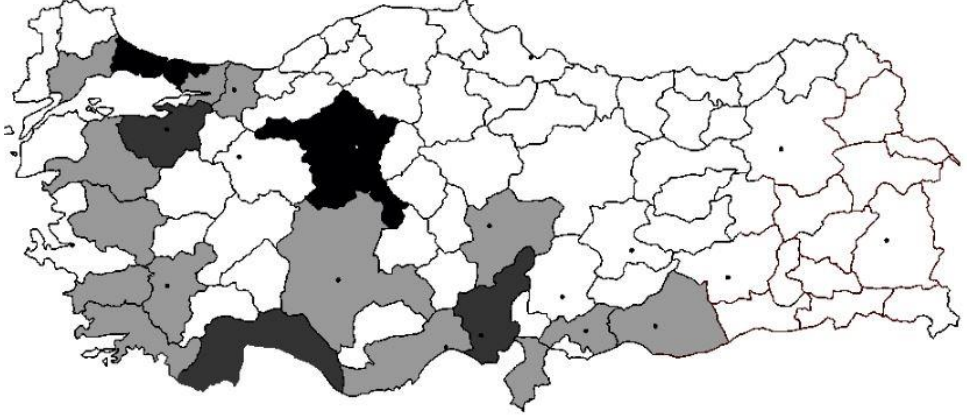
Bulanık k ortalamalar kümeleme yöntemine göre illerin girişimcilik verilerine göre kümelenme sonuçları ise aşağıda Tablo 2’de yer almaktadır.

Tablo 2: İllerin Bulanık K Ortalamalar Yöntemine Göre Kümelenme Sonuçları

1 Nolu Küme	2 Nolu Küme	3 Nolu Küme	4 Nolu Küme
Ankara ve İstanbul	Adana, Antalya ve Bursa	Şanlıurfa, Sakarya, Tekirdağ, Aydın, Balıkesir, Denizli, Hatay, Manisa, Kayseri, Muğla, Gaziantep, Mersin, Kocaeli, Konya	Geri Kalan Tüm İller

Tablo 2’ye göre analizin başında diğer illerden ayrılan Ankara ve İstanbul ayrı bir küme; Adana, Antalya ve Bursa ayrı bir küme; Şanlıurfa, Sakarya, Tekirdağ, Aydın, Balıkesir, Denizli, Hatay, Manisa, Kayseri, Muğla, Gaziantep, Mersin, Kocaeli, Konya ayrı bir küme kalan 62 il ise ayrı bir küme oluşturmuştur. Şekil 5’de ise bu illerin Türkiye haritasında dağılımı gösterilmektedir.

Şekil 5: Bulanık k Ortalamalar Kümeleme Yöntemine Göre İllerin Türkiye Haritasında Dağılımı



Bulanık k ortalamalar kümeleme yöntemine göre İllerin Türkiye haritasında dağılımına bakıldığında Türkiye'nin en büyük ve genel anlamda en gelişmiş iki ili olan Ankara ve İstanbul'un diğer illerden ayrı bir küme olması tahmin edilebilen bir sonuç olmuştur.

Adana, Antalya ve Bursa illerinin diğer illerden ayrılarak farklı bir küme oluşturması da bu illerin genç nüfusu, sanayisi, iş potansiyeli, iş gücüne katılım oranı, kalifiye iş gücü gibi parametreler göz önüne alındığında yine tahmin edilebilen bir sonuç olmuştur. Burada tahmin edilemeyen sonuç ise üstte yer alan parametreler açısından en yukarılarda yer alan İzmir'in bu kümede yer almamasıdır. Bu durumun nedeni çevre illerdeki sanayileşmenin yüksek olmasının iş gücünü İzmir'den çekmesi olabilir.

Şanlıurfa, Sakarya, Tekirdağ, Aydın, Balıkesir, Denizli, Hatay, Manisa, Kayseri, Muğla, Gaziantep, Mersin, Kocaeli ve Konya illerinin diğer illerden ayrılarak farklı bir küme oluşturması da bu illerin genç nüfusu, sanayisi, iş potansiyeli, iş gücüne katılım oranı, kalifiye iş gücü gibi parametreler göz önüne alındığında yine tahmin edilebilen bir sonuç olmuştur. Burada tahmin edilemeyen sonuç ise üstte yer alan parametreler açısından en yukarılarda yer alan İzmir'in 2. kümede yer almadığı gibi bu kümede de yer almamasıdır. Ayrıca başka bir tahmin edilemeyen sonuç da yine bu parametrelere göre bu kümede yer alması beklenen Diyarbakır ilinin bu kümede yer almaması olmuştur. Bu durumun nedeni Diyarbakır ilinin coğrafi konumu ve dar bir sektörel dağılım sergilemesi olabilir.

Yukarıda iller dışında kalan 62 il ise son kümede yer almıştır. Genç nüfusu, sanayisi, iş potansiyeli, iş gücüne katılım oranı, kalifiye iş gücü gibi parametreler göz önüne alındığında İzmir ve Diyarbakır haricinde diğer illerin ayrı bir kümede olması tahmin edilebilen bir sonuç olmuştur.

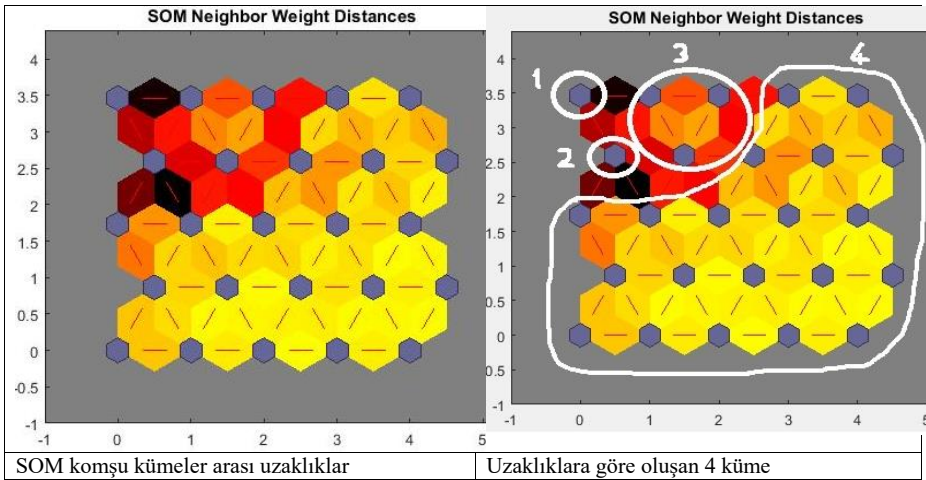
Kümelerin kendi içinde ve genel olarak dağılımına bakıldığında ise bölgesel bir dağılıma rastlanmamıştır. Fakat Karadeniz haricinde diğer denizlere kıyısı olan illerden çoğunun genelde son kümeye dahil olmadığı 1. 2. ve 3. kümelerde yer aldığı görülmüştür.

SOM Kümeleme Yöntemi Bulguları

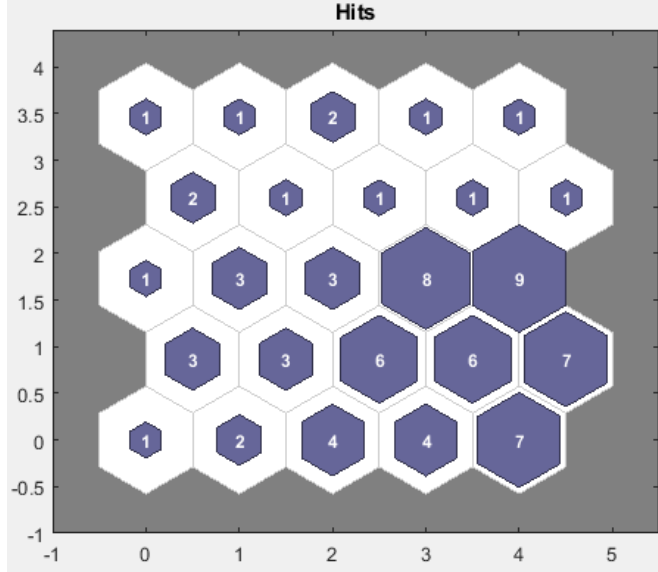
Matlab programı kümeleme araç kutusu kullanılarak SOM yöntemine göre illerin girişimcilik verileri kümelendirilmiştir. Kümeleme işlemi için 5*5 yapı-sındaki bir harita düzeni seçilmiştir. Yani iller 25 kümeden oluşan bir haritaya göre kümeleneceklerdir.

SOM yöntemine göre yapılan kümeleme işlemi sonucunda Şekil 6'da görüldüğü gibi uzaklık grafiği elde edilmiştir. Kümeler arasındaki uzaklıklarda koyu renkler uzaklığın arttığını göstermektedir. Buna göre bu 25 küme de kendi aralarındaki uzaklıklar baz alındığında 4 farklı kümeye ayrılmaktadır. Burada 4 nolu kümenin oldukça geniş olduğu dikkat çekmektedir.

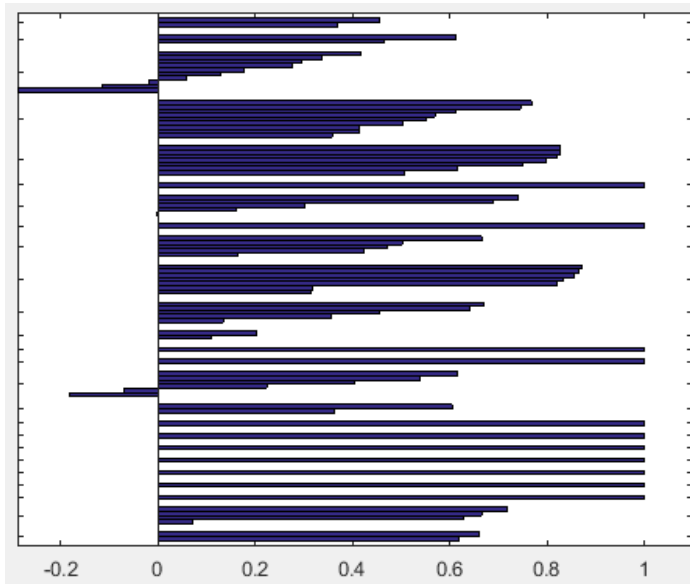
Şekil 6: SOM Kümeleme Yöntemi Göre Uzaklık Grafiği



SOM yöntemine göre oluşturulan 25 kümenin kaç şehirden oluştuğunu gösteren hit grafiği ise Şekil 7'de gösterilmektedir.

Şekil 7: SOM Kümeleme Yöntemi Göre Hit Grafiği

SOM yöntemine göre kümelemenin başarısını göstermek amacı ile Silhouette grafiği Şekil 8'de verilmiştir. Bu grafik kümeleme işleminin başarısının görsel olarak görülmesini sağlar. Grafikte görüldüğü gibi birkaç küme haricinde SOM kümeleme yöntemi verileri başarılı bir şekilde kümeleyebilmiştir.

Şekil 8: SOM Kümeleme Yöntemi Göre Silhouette Grafiği

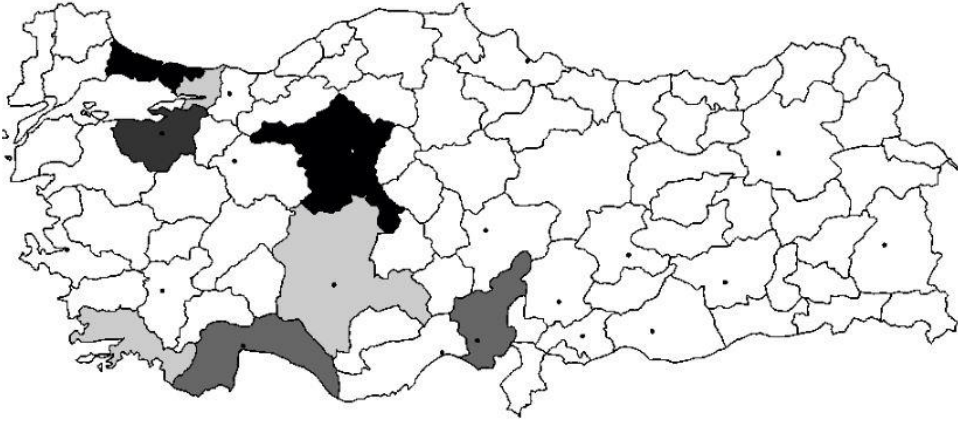
SOM kümeleme yöntemine göre şehirlerin girişimcilik durumuna göre kümeleme sonuçları Tablo 3’de verilmiştir.

Tablo 3: İllerin SOM Kümeleme Yöntemi Göre Kümelene Sonuçları

1 Nolu Küme	2 Nolu Küme	3 Nolu Küme	4 Nolu Küme	5 Nolu Küme
Ankara ve İstanbul	Bursa	Adana ve Antalya	Kocaeli, Konya ve Muğla	Geri Kalan İller

Tablo 3’e göre analizin başında diğer illerden ayrılan Ankara ve İstanbul illeri ayrı bir küme; Bursa ili tek başına ayrı bir küme; Adana ve Antalya illeri ayrı bir küme, Kocaeli, Konya ve Muğla illeri ayrı bir küme, kalan 73 il ise ayrı bir küme oluşturmuştur. Şekil 9’da ise bu illerin Türkiye haritasında dağılımı gösterilmektedir.

Şekil 9: SOM Kümeleme Yöntemine Göre İllerin Türkiye Haritasında Dağılımı



SOM kümeleme yöntemine göre İllerin Türkiye haritasında dağılımına bakıldığında Türkiye’nin en büyük ve genel anlamda en gelişmiş iki ili olan Ankara ve İstanbul’un diğer illerden ayrı bir küme olması tahmin edilebilen bir sonuçtur.

Bursa ilinin diğer tüm illerden ayrılarak farklı bir küme oluşturması genç nüfus, sanayi, iş potansiyeli, iş gücüne katılım oranı, kalifiye iş gücü gibi parametreler göz önüne alındığında beklenen bir sonuç olmamıştır. Bursa ilinin yanında yukarıdaki parametreler açısından üstlerde yer alan illerin de bu kümede yer alması daha beklenebilir bir sonuçtur. Bu durum Bursa ilinin coğrafi konumu, pazarlara yakınlığı, farklı tüm sektörlerde yüksek iş imkanı gibi nedenlerden kaynaklanmış olabilir.

Adana ve Antalya illerinin diğer illerden ayrılarak farklı bir küme oluşturması da bu illerin genç nüfusu, sanayisi, iş potansiyeli, iş gücüne katılım oranı, kalifiye iş gücü gibi parametreler göz önüne alındığında beklenen bir sonuç olmuştur. Fakat başka illerin de bu kümede yer alması daha tahmin edilebilir bir sonuç olarak kabul edilebilir.

Kocaeli, Konya ve Muğla illerinin diğer illerden ayrılarak farklı bir küme oluşturmuştur. İllerin genç nüfusu, sanayisi, iş potansiyeli, iş gücüne katılım oranı, kalifiye iş gücü gibi parametreler göz önüne alındığında Muğla ilinin bu kümede yer alması tahmin edilebilir bir sonuç olmamıştır. Bu durumun nedeni Muğla ilinin turizm konusunda lokomotif bir rol üstlenerek günden güne artan hizmet sektöründeki açık olabilir.

Yukarıda iller dışında kalan 73 il ise son kümede yer almıştır. Genç nüfusu, sanayisi, iş potansiyeli, iş gücüne katılım oranı, kalifiye iş gücü gibi parametreler göz önüne alındığında üstlerde yer alan İzmir, Şanlıurfa, Gaziantep vb. illerin bu kümeye dahil olması tahmin edilebilir bir sonuç olmamıştır.

Kümelerin kendi içinde ve genel olarak dağılımına bakıldığında ise bölgesel bir dağılıma rastlanmamıştır.

SONUÇ VE ÖNERİLER

Sürekli değişen pazar koşulları ve gelişen teknoloji bağlamında girişimcilik faaliyetlerin önemi özellikle Türkiye gibi gelişmekte olan ülkelerde giderek artmaktadır. Girişimcilik faaliyetlerinin benzer özellikleri taşıyan yerlere göre gruplanması da aynı şekilde büyük önem teşkil etmektedir. Çünkü aynı özellikleri taşıyan ve sayısı; il, kalkınma bölgesi ve coğrafi bölgelere göre daha az olan bu gruplar gerek girişimcilik faaliyetlerinin teşvik edilmesi ve destek sağlanmasında gerekse girişimcilik faaliyetlerinde ortaya çıkan sorunların belirlenmesi ve çözümünde önemli bir veri kaynağı olacaktır. Bu nedende bu çalışmada ülkemizdeki illeri girişimcilik faaliyetlerine göre kümeleyerek literatürdeki bu boşluğun kapatılması amaçlanmıştır.

Türkiye Sanayi ve Teknoloji Bakanlığı Girişimci Bilgi Sistemi tarafından sağlanan veriler kullanılarak GBS'nde yer alan "İllere göre girişim sayısı, net satışlar ve aktifler" raporları çalışmada veri olarak kullanılmıştır. Kümeleme yöntemlerinden Bulanık k ortalamalar ile SOM yöntemi uygulanmıştır.

Bu verileri içeren üç boyutlu olarak çizilen grafiğe göre İstanbul ve Ankara illerinin diğer bütün illerden ayrılmış olduğu ve oldukça farklılık gösterdiği görülmüştür. Bu illerinin verilerinin aykırı veri oldukları, kümeleme analizi sonuçlarını ciddi oranda etkileyecekleri düşünülerek bu iller analiz başlangıcında zaten ayrı bir küme olarak değerlendirilmiş ve analizden çıkarılarak araştırmaya geriye kalan 79 il ile devam edilmiştir.

Bulanık k ortalamalar yöntemine göre Ankara ve İstanbul ayrı bir küme; Adana, Antalya ve Bursa ayrı bir küme; Şanlıurfa, Sakarya, Tekirdağ, Aydın, Balıkesir, Denizli, Hatay, Manisa, Kayseri, Muğla, Gaziantep, Mersin, Kocaeli, Konya ayrı bir küme kalan 62 il ise ayrı bir küme oluşturmuştur. İzmir ve Di-

yarbakır illerinin son kümede yer alması genç nüfus, sanayi, iş potansiyeli, iş gücüne katılım oranı, kalifiye iş gücü gibi parametreler göz önüne alındığında dikkat çekici bir sonuç olmuştur. Kümelerin kendi içinde ve genel olarak dağılımına bakıldığında ise bölgesel bir dağılıma rastlanmamıştır. Fakat Karadeniz haricinde diğer denizlere kıyısı olan illerden çoğunun genelde son kümeye dahil olmadığı 1. 2. ve 3. kümelerde yer aldığı görülmüştür.

Tablo 4: Bulanık K Ortalamalar ve SOM Kümeleme Yöntemi Sonuçları Arasındaki Farklar

Kümeleme Yöntemi	1 Nolu Küme	2 Nolu Küme	3 Nolu Küme	4 Nolu Küme	5 Nolu Küme
Bulanık K Ortalamalar	Ankara ve İstanbul	Adana, Antalya ve Bursa	Şanlıurfa, Sakarya, Tekirdağ, Aydın, Balıkesir, Denizli, Hatay, Manisa, Kayseri, Muğla, Gaziantep, Mersin, Kocaeli, Konya	Geri Kalan Tüm İller	-
SOM	Ankara ve İstanbul	Bursa	Adana ve Antalya	Kocaeli, Konya ve Muğla	Geri Kalan İller

SOM kümeleme yöntemine göre Ankara ve İstanbul illeri ayrı bir küme; Bursa ili tek başına ayrı bir küme; Adana ve Antalya illeri ayrı bir küme, Kocaeli, Konya ve Muğla illeri ayrı bir küme, kalan 73 il ise ayrı bir küme oluşturmuştur. İzmir, Şanlıurfa ve Gaziantep illerinin son kümede yer alması genç nüfus, sanayi, iş potansiyeli, iş gücüne katılım oranı, kalifiye iş gücü gibi parametreler göz önüne alındığında dikkat çekici bir sonuç olmuştur. Ayrıca Muğla ilinin son kümede yer almaması da yukarıdaki parametreler doğrultusunda dikkat çekici bir sonuç olmuştur. Kümelerin kendi içinde ve genel olarak dağılımına bakıldığında ise bölgesel bir dağılıma rastlanmamıştır.

Bu iki yöntemin oluşturduğu kümeler kendi arasında kıyaslandığında ise bulanık k ortalamalar yöntemine göre toplamda 4 küme SOM kümeleme yöntemine göre ise toplamda 5 küme oluşmuştur. Gerek küme sayısının farklı olması gerekse bu yöntemlerin çalışma algoritmalarının farklı olması nedeniyle iller de farklı kümelerde yer almışlardır. Bulanık k ortalamalar yöntemine göre 19 il 3 farklı kümede kalan 62 il tek kümede toplanmışken, SOM kümeleme yöntemine göre ise 8 il 4 farklı kümede kalan 73 il ise tek kümede toplanmıştır. Kümeleme analizi sonuçlarının bu şekilde farklı çıkmasının en önemli nedeni yöntemlerin yaklaşımlarının ve matematiğinin farklı olmasıdır. Özellikle SOM kümeleme yöntemi yapay sinir ağlarını temel alan bir yöntem olduğu için bütün kümeleme algoritmalarından ayrılmaktadır bu nedenle farklı sonuçlar vermesi de beklenen bir sonuçtur çalışmada da bu şekilde çıkmıştır.

Bazı gruplara düşen il sayısının az, bazı gruplara düşen il sayısının fazla olmasının nedeni ise girişimciliğe verilen desteklerin belirli bölgelerde daha benimsenerek uygulanması, özellikle kültürel yapısı birbirine benzeyen ve daha az gelişmiş olan illerin girişimcilik açısından tutuk kalması olabilir.

Bu çalışma ile 81 il GBS'nde yer alan "İllere göre girişim sayısı, net satışlar ve aktifler" verilerine göre Bulanık k ortalamalar ile SOM yöntemleriyle kümelene- miş ve bunun sonucunda bu 2 yöntemle göre Türkiye'nin girişimcilik haritası çıkarılmıştır. Girişimcilik faaliyetlerinin aynı özellikleri taşıyan yerlere göre kümelene- mesi büyük bir önem arz etmektedir. Çünkü aynı özellikleri taşıyan ve sayısı il, kalkınma bölgesi ve coğrafi bölgelere göre daha az olan bu gruplar gerek girişimcilik faaliyetlerinin teşvik edilmesi ve destek sağlanmasında gerekse girişimcilik faaliyetlerinde ortaya çıkan sorunların belirlenmesi ve çözümünde önemli bir veri kaynağı olacaktır. Sanayi ve Teknoloji Bakanlığı, KOSGEB, TÜBİTAK, Kalkınma Ajansları, Sivil Toplum Kuruluşları vb. di- ğer karar alıcılar girişimcilikle ilgili bölgesel politikalarını oluştururken bu çalışmadan yararlanabilirler.

Bundan sonraki çalışmalarda illerde girişimciliğe ait daha değişik veriler de kullanılarak çalışma genişletilebilir, veriler GBS'nin dışında başka yerlerden de alınarak çeşitlendirilebilir bulanık k ortalamalar ile SOM yöntemleri dışın- da daha farklı ve çeşitli kümeleme yöntemleri kullanılabilir, kümeleme sadece iller bazında değil kalkınma bölgesi, teşvik bölgesi vb. gibi çeşitli alanlarda da yapılabilir ve sadece tek bir yıl için değil yıllara göre girişimciliğin Türkiye'de- ki değişimi de incelenebilir.

KAYNAKÇA

- Alpar, R. (2017). *Uygulamalı Çok Değişkenli İstatistiksel Yöntemler*. Ankara: Detay Yayıncılık.
- Atalay, M., ve Öztürk, Ş. (2016). Türkiye'deki İllerin Göç ve İşsizlik İstatistiklerine Göre Kümelenmesi. In *2nd International Congress on Applied Sciences: "Migration, Poverty and Employment"*: 23-25 Eylül 2016, Konya: Bildiriler (520-527).
- Bikari, S. (2017). *Tüketicilerin Çevreye Yönelik Tutumları ve Kişilik Özellikleri Açısından Kohonen Ağları (Self-Organizing Map-Som) ve Kümeleme Analizi İle Karşılaştırılması*, Doktora Tezi, Atatürk Üniversitesi Sosyal Bilimler Enstitüsü, Erzurum.
- Belgin, Ö. ve Avşar, B. A. (2019). Türkiye'de Bölgeler ve İller Düzeyinde Ar-Ge ve Yenilik Performansının Gri İlişkisel Analiz Yöntemi İle Ölçülmesi, *Verimlilik Dergisi*, (2), 27-48.
- Bezdek, J. C. (1973). *Fuzzy Mathematics in Pattern Classification*. PhD Thesis, Cornell University, New York.
- Çınaroğlu, S., & Bulut, H. (2018). K-Ortalamlar ve Parçacık Sürü Optimizasyonu Tabanlı Kümeleme Algoritmaları İçin Yeni İklendirme Yaklaşımları, *Journal of the Faculty of Engineering & Architecture of Gazi University*, 33(2), 413-422.
- Güneş, M. ve İncekırık, A. (2016). Ege Bölgesinde Faaliyet Gösteren KOSGEB Kapsamındaki Farklı Ölçeklerdeki Şirketlerin (KOBİ) Bulanık Kümeleme Analizi İle Gruplandırılması, *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 22(4), 314-323.
- Güney, S. ve Çetin, A. (2003). Kültürün Girişimciliğe Etkisi ve Türkiye'de Girişimcilik Kültürü, *Hacettepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 21(1), 189-210.
- Höppner, F., Klawonn, F., Rudolf, K. ve Runkler, T. (1999). *Fuzzy Cluster Analysis*. Wolfenbüttel: Wiley Publishers.
- Işık, M. ve Çamurcu, A.Y. (2007). K-Means, K-Medoids ve Bulanık C- Means Algoritmalarının Uygulamalı Olarak Performanslarının Tespiti, *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 6(11), 31-45.
- Kohonen, T. (1998). The Self-Organizing Map, *Neurocomputing*, 21, 1-6.
- Littunen, H. (2000). Entrepreneurship and The Characteristics of The Entrepreneurial Personality, *International Journal of Entrepreneurial Behavior & Research*, 6(6), 295-299.

- Murat, Y.Ş. ve Şekerler, A. (2009). Trafik Kaza Verilerinin Kümelenme Analizi Yöntemi ile Modellenmesi, *Teknik Dergi*, 20 (98), 4759-4777.
- Özkan, Ş. (2003). Türkiye’de Girişimcilik ve Belirleyiciler: Marmara ve Doğu Anadolu Bölgesi Üzerine Bir Uygulama, *Ankara Üniversitesi SBF Dergisi*, 58(04), 145-172.
- Tokathoğlu, S., ve Yalçın, E. C. (2019). Girişimci Davranışa İlişkin Bireysel Algılar: GEM Verileriyle Kümeleme Analizi, *International Journal of Innovative Approaches in Social Sciences*, 3(1), 1-19.
- Wickham, P. (1998). *Strategic Entrepreneurship. A Decision-Making Approach To New Venture Creation And Management*. UK: London: Pitman Publishing.
- Yılmaz, E., ve Sünbül, A.M. (2009). Üniversite Öğrencilerine Yönelik Girişimcilik Ölçeğinin Geliştirilmesi, *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (21), 195-203.