

**02010044-NÖRAL BULANIK SİSTEMLER-1**  
**Yiğit ALIŞAN (24020010001)**

**ÖZET**

Bulanık sinir ağları (FNN'ler), bulanık mantık ve sinir ağlarının birleşimiyle, sınıflandırma, regresyon ve kontrol gibi alanlarda yüksek doğruluk ve yorumlanabilirlik sunmaktadır. Ancak, geleneksel Bulanık C-Ortalamlar (FCM) ve Sert C-Ortalamlar (HCM) tabanlı bulanık evrimsel sinir ağları (FCNN'ler), gürültü ve aykırı değerlere karşı hassasiyet ile hesaplama karmaşıklığı gibi sınırlamalarla karşı karşıyadır. Bu çalışma, bu sorunları ele almak amacıyla, destek vektör (SV) tabanlı kümeleme ve L2 norm regularizasyonunu entegre eden bir Düzenli Bulanık C-Ortalamlar Sinir Ağı (RFCNN) modelini test etmektedir. Kullanılan model, giriş uzayını alt uzaylara ayıran, polinom ağırlıklarını optimize eden ve çıktıyı birleştiren üç fazlı bir öğrenme süreciyle yapılandırılmıştır. 144 hiperparametre kombinasyonu üzerinde gerçekleştirilen deneyler, RFCNN'nin, özellikle gürültülü veri setlerinde, FCM ve HCM tabanlı FCNN'lere kıyasla daha düşük Ortalama Karekök Hata (RMSE) ve standart sapma değerleri sunduğunu göstermiştir. Bulgular, SV tabanlı kümelemenin gürültü etkisini azalttığını ve L2 norm regularizasyonunun model sağlamlığını artırdığını doğrulamaktadır. Bu çalışmada, FNN optimizasyonu için yenilikçi bir çerçeve sunmakta olan RFCNN modeli test edilmiş ve başarılı sonuçlar elde edilmiştir.

## 1. GİRİŞ

Bulanık sinir ağıları (FNN'ler), bulanık mantığın belirsizlikleri modelleme yeteneğini sinir ağlarının öğrenme kapasitesiyle birleştirerek, sınıflandırma, regresyon ve kontrol gibi alanlarda etkili çözümler sunmaktadır (Zhao et al., 2025). Geleneksel sinir ağlarına kıyasla, FNN'ler gürültülü ve karmaşık veri setlerinde daha yüksek doğruluk, yorumlanabilirlik ve sağlamlık sağlamakta, özellikle tıbbi teşhis, görüntü analizi ve endüstriyel kontrol sistemlerinde öne çıkmaktadır (Z. Wang et al., 2022). Ancak, FNN'lerin performansı, bulanık küme sayısı ve veri boyutluluğunun artmasıyla ortaya çıkan hesaplama karmaşıklığı nedeniyle sınırlanabilmektedir. Ayrıca, geleneksel kümeleme yöntemleri olan Bulanık C-Ortalamlar (FCM) ve Sert C-Ortalamlar (HCM), Öklidyen mesafe ölçütüne dayalı yapıları nedeniyle gürültü ve aykırı değerlere karşı hassasiyet göstermektedir (de Campos Souza & Lughofer, 2022; Sreedhara et al., 2023; Tran et al., 2023). Çalışmalar, FCM ve HCM'nin çeşitli hiperparametre kombinasyonları altında karşılaştırıldığını ve en iyi performans gösteren kombinasyonların düşük hata oranları sağladığını, ancak gürültü hassasiyetinin devam ettiğini ortaya koymuştur (Tran et al., 2023).

Bu sınırlamaları aşmak için, destek vektör (SV) tabanlı kümeleme ve L2 norm regularizasyonu gibi yenilikçi yaklaşımlar literatürde önerilmiştir (J. Wang et al., 2023). SV tabanlı kümeleme, yalnızca seçilen destek vektörlerle kümeler oluşturarak gürültü etkisini azaltırken, L2 norm regularizasyonu, model eğitimi sırasında varyans-yanlılık dengesini optimize ederek sağlamlığı artırmaktadır (Atashgahi et al., 2024; Xu et al., 2022). Bu çalışmada, geleneksel FCNN'lerin (Bulanık Evrişimsel Sinir Ağları) gürültüye duyarlılığı ve sınırlı sağlamlık sorunlarını ele almak amacıyla, SV tabanlı kümeleme ve L2 norm regularizasyonunu entegre eden bir Düzenli Bulanık C-Ortalamlar Sinir Ağı (RFCNN) modeli önerilmektedir. RFCNN, giriş uzayını alt uzaylara ayıran, polinom ağırlıklarını hesaplayan ve çıktıyı birleştiren üç fazlı bir öğrenme süreciyle yapılandırılmıştır. Deney sonuçları, RFCNN'nin gürültülü veri setlerinde geleneksel FCNN'lere kıyasla daha düşük hata oranları ve daha yüksek kararlılık sunduğunu göstermektedir. Bu çalışma, FNN optimizasyonu için yeni bir çerçeve sunmayı ve gelecekte farklı veri setleriyle genelleme testleri ile hesaplama karmaşıklığı analizlerini teşvik etmeyi amaçlamaktadır.

Yapılan çalışmada literatür tarama bölümünde FNN'lerin tarihçesi, mevcut yöntemlerin güçlü ve zayıf yönleri, SV tabanlı kümeleme ve regularizasyon tekniklerinin katkıları, materyal ve metot bölümünde Veri seti özellikleri, algoritmaların matematiksel temelleri, model mimarisi ve deneysel yöntemler, deneysel sonuçlar bölümünde Hiperparametre kombinasyonlarının performansı, FCM ve HCM'nin karşılaştırması, görselleştirme ve analiz ve son olarak sonuç ve değerlendirme bölümünde Bulguların özeti, yöntemlerin güçlü/zayıf yönleri, gelecekteki araştırma önerileri sunulmaktadır.

## 2. LİTERATÜR TARAMA

Bulanık sinir ağıları (FNN'ler), bulanık mantık ve sinir ağlarının sinerjisiyle, karmaşık veri setlerinde esnek ve yorumlanabilir modeller sunmaktadır. Zadeh [1] ve Jang [2], FNN'lerin belirsizlik modellemedeki üstünlüğünü vurgularken, Souza ve Lughofer [9], n-uniform tabanlı modellerin esnek kural bağlantıları sağladığını belirtmiştir. Hiyerarşik FNN yapıları, denetimsiz ve denetimli öğrenmeyi birleştirerek tıbbi teşhis ve görüntü analizi gibi alanlarda yüksek doğruluk elde etmiştir [10]. Ancak, FNN'lerin hesaplama karmaşıklığı, özellikle yüksek boyutlu veri setlerinde ve artan küme sayılarında önemli bir zorluk oluşturmaktadır [11]. Wang ve diğerleri [12], özellik seçimi ve kümeleme temelli yaklaşımlarla bu sorunu hafifletmeyi önerirken, Liu ve diğerleri [13], uyarlanabilir bilgi üreten FNN modelleriyle genelleme kabiliyetini artırmıştır.

Son yıllarda, bulanık sinir ağıları (FNN) belirsizlik ve bulanık verileri ele almada önemli ilerlemeler kaydetmiş olsa da bazı zorluklarla karşı karşıyadır. (Huang et al., 2022a), RBFNN'lerin sağlamlığını ve doğruluğunu artırmak için çift yinelemeli öğrenmeye dayalı polinomlu bir RBFNN modeli önermiştir. Bu model, destek vektörü çekirdek bulanık kümeleme ile en az mutlak küçülme sapmalarını birleştirmektedir. Yine (Huang et al., 2022b), sinir ağı modelleme doğruluğunu artırmak amacıyla destek vektörleriyle geliştirilmiş bulanık kümeleme sinir ağlarını incelemiş, bu süreçte destek vektörü kümeleme ve düzenleme tekniklerini kullanmıştır (Huang et al., 2022c). (S. Liu et al., 2024) ise dikkat mekanizmalarını ve ardışık öğrenmeyi entegre eden, sınıflandırma performansını önemli ölçüde artıran güçlendirilmiş aralıklı tip-2 bulanık kümelemeye dayalı bir sinir ağı sunmuştur. (Z. Wang et al., 2022), tahmin doğruluğunu ve kararlılığı artırmak için polinom nöronları ve özellik seçimiyle oluşturulan bulanık polinom nöronları tarafından yönlendirilen bulanık küme tabanlı hibrit sinir ağlarını araştırmıştır. Başka bir çalışmada araştırmacılar yüksek boyutlu verilerle başa çıkmak için güçlendirilmiş bulanık kural tabanlı sinir ağıları (FRNNs) geliştirmiştir. Bu yöntemde, belirleme/korelasyon katsayısına dayalı sadeleştirilmiş özellik seçimi ve standart sapma ile bulanık kümeleme kullanılarak hesaplama yükü azaltılmış, L2-norm düzenlemesiyle genelleme yeteneği artırılmıştır. FRNNs'nin üstünlüğü, 28 gerçek dünya veri seti üzerinde yapılan deneylerle ortaya konmuştur (Z. Wang et al., 2024). (Bu et al., 2024) regresyon problemleri için güçlendirilmiş Bulanık Kümeleme tabanlı Sinir Ağı (FCNN) önermiştir. Doğrusal regresyon kalıntılarından elde edilen farklılıklar bulanık bölünmeye entegre edilerek daha iyi bulanık kümeler oluşturulmuş, konvolüsyonel düzenleme stratejisiyle aşırı öğrenme sorunu azaltılarak modelin genelleme kapasitesi artırılmıştır. Bu yöntemin etkinliği, sentetik ve gerçek dünya veri setleriyle doğrulanmıştır. Liu vd. (2024) akıllı şebeke sistemlerinde kısa vadeli yük tahmini (STLF) için bulanık C-ortalama (FCM) kümeleme ve geliştirilmiş uzun kısa vadeli bellek (LSTM) sinir ağı birleştirilerek yeni bir model sunulmuştur. Yük profilleri PCA ile boyut indirgeme sonrası kümelenecek, her desen için LSTM tabanlı bir tahminci tasarlanmış ve MAPE metriğine göre diğer modellere kıyasla üstün performans elde edilmiştir (F. Liu et al., 2024).

Ayrıca, (Z. Wang et al., 2023), kapılı tekrarlayan birim yapıları ve korelasyon/olasılıksal seçim stratejileriyle tasarlanmış ilerici bulanık polinom sinir ağlarını önererek modelin farklı veri ölçeklerine uyarlanabilirliğini artırmıştır. Ancak bu modeller, tek tip verilerle (örneğin finansal veriler) iyi performans gösterse de, karma verilerle (örneğin finansal veriler ve metin duygusal tonu) çalışırken belirgin sınırlamalar sergiler (Latifi et al., 2022).

### 3. MATERYAL VE METOT

#### 3.1. Veri Seti

Çalışmada kullanılan veriseti UCI Makine Öğrenmesi Deposundan alınmıştır. Boston konut verileri 1978'de toplanmıştır ve 506 girdinin her biri Massachusetts, Boston'daki çeşitli banliyölerdeki evler için 14 özellik hakkında toplanmış verileri temsil eder. Verisetisine ait örnek bir kesit Tablo X' te gösterilmektedir.

Veri setindeki 14 özellik sırasıyla:

1. **crim**: Mahalledeki kişi başına suç oranı.
2. **zn**: 25,000 fit kareden büyük arsalar için ayrılmış alan oranı.
3. **indus**: Mahalledeki ticari olmayan iş alanlarının oranı.
4. **chas**: Charles Nehri'ne yakınlık (1: yakın, 0: değil).
5. **nox**: Azot oksit konsantrasyonu (hava kirliliği göstergesi).
6. **rm**: Konut başına ortalama oda sayısı.
7. **age**: 1940'tan önce inşa edilmiş evlerin oranı.
8. **dis**: Boston'daki beş istihdam merkezine ağırlıklı mesafe.
9. **rad**: Radyal otoyollara erişim indeksi.
10. **tax**: 10,000 dolar başına emlak vergisi oranı.
11. **ptratio**: Öğrenci-öğretmen oranı.
12. **b**: Siyah nüfus oranının dönüştürülmüş hali ( $1000(Bk - 0.63)^2$ , Bk: siyah nüfus oranı).
13. **lstat**: Nüfusun düşük sosyo-ekonomik statüsü oranı.
14. **medv**: Mahalledeki evlerin medyan değeri (hedef değişken, bin USD) şeklindedir.

Tablo 1 Veri seti örnek kesit

crim	zn	indus	chas	nox	rm	age	dis	rad	tax	ptratio	b	lstat	medv
0.0063	18	2.31	0	0.54	6.58	65	4.09	1	296	15.3	396.9	4.98	24
0.0273	0	7.07	0	0.47	6.42	79	4.967	2	242	17.8	396.9	9.14	21.6
0.0273	0	7.07	0	0.47	7.19	61	4.967	2	242	17.8	392.8	4.03	34.7
0.0324	0	2.18	0	0.46	7	46	6.062	3	222	18.7	394.6	2.94	33.4
0.0691	0	2.18	0	0.46	7.15	54	6.062	3	222	18.7	396.9	5.33	36.2

#### 3.2. (FCNN) Fully Connected Neural Network

Tam Bağlantılı Sinir Ağı (Fully Connected Neural Network - FCNN), yapay sinir ağlarının temel bir türü olarak kullanılmaktadır ve her katmandaki nöronların bir sonraki katmandaki tüm nöronlara bağlanmasıyla karakterize edilmektedir. FCNN, giriş katmanı, bir veya daha fazla gizli katman ve çıkış katmanından oluşmakta; her nöron, önceki katmandan gelen ağırlıklı girdileri toplamakta, bir aktivasyon fonksiyonundan geçirmekte ve sonucu bir sonraki katmana iletmektedir.

Sınıflandırma, regresyon ve örüntü tanıma gibi görevlerde yaygın olarak kullanılan bu ağlar, yoğun sinir ağı (Dense Neural Network) veya çok katmanlı algılayıcı (Multilayer Perceptron - MLP) olarak da adlandırılmaktadır. Ancak, tam bağlantılı yapıları nedeniyle büyük boyutlu verilerde yüksek hesaplama maliyeti doğurabilmekte ve görüntü veya dizi işleme gibi özel görevlerde özelleşmiş mimariler (CNN, RNN gibi) kadar etkin olmamaktadır.

➤ *Sert C-Ortalamlar* (Hard C-Means - HCM), veri noktalarını belirli sayıda kümeye ayırmak için kullanılan klasik bir kümeleme algoritması olarak bilinmektedir ve genellikle K-Means kümelemesiyle eş anlamlıdır. HCM, her veri noktasının yalnızca bir kümeye tam üyelikle (0 veya 1) ait olmasını sağlamak ve bu kesin atama özelliğiyle "sert" kümeleme olarak tanımlanmaktadır. Algoritma, her veri noktası ile ait olduğu küme merkezi arasındaki mesafelerin karelerinin toplamını minimize eden bir amaç fonksiyonunu optimize etmektedir. HCM, basitliği ve hesaplama verimliliği nedeniyle yaygın bir şekilde tercih edilmekte, ancak bulanık C-ortalamlar (FCM) gibi yöntemlerin aksine, belirsizlik içeren veya karmaşık veri yapılarında sınırlı kalabilmektedir.

FCNN ve HCM, sırasıyla denetimli ve denetimsiz öğrenme alanlarında temel yaklaşımlar sunmaktadır. FCNN, makine öğrenmesine başlangıç için uygun bir mimari olarak öne çıkarken, HCM, veri kümelemede hızlı ve anlaşılır bir çözüm sağlamaktadır. Bununla birlikte, her iki yöntemin de karmaşık veri yapılarında veya özel görevlerde daha gelişmiş alternatiflere ihtiyaç duyabileceği anlaşılmaktadır. Bu nedenle, uygulama bağlamına bağlı olarak, daha özelleşmiş algoritmaların değerlendirilmesi önem taşımaktadır.

➤ *Bulanık C-Ortalamlar* (Fuzzy C-Means - FCM) algoritması (Denklem 1), sert kümeleme yöntemlerinin (HCM/K-Means) aksine, bulanık mantık prensiplerini temel alan esnek bir kümeleme yaklaşımı sunar. FCM'nin en belirgin özelliği, bir veri noktasının birden fazla kümeye 0 ile 1 arasında değişen üyelik dereceleriyle ait olabilmesidir. Bu, her veri noktasının tüm kümelere olan üyelik derecelerinin toplamının 1'e eşit olması koşuluyla gerçekleştirilir. Algoritma, bulanıklık parametresi (m) ile kümelemenin bulanıklık seviyesini kontrol eder; m=1 olduğunda FCM, geleneksel HCM'ye dönüşürken, m'nin artması daha bulanık bir kümeleme sağlar. Bu esneklik, FCM'yi belirsizlik içeren veri setlerinde güçlü bir araç haline getirir.

$$J_{FCM} = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m \|x_k - v_i\|^2 \quad (1)$$

FCM'nin temel amacı, bulanık üyelik dereceleri ile ağırlıklandırılmış mesafe karelerinin toplamını minimize eden bir amaç fonksiyonunu optimize etmektir. Burada:

- $u_{ik}^m$ :  $x_k$  veri noktasının i kümeye üyelik derecesi
- m: Bulanıklaştırma katsayısı
- n: Veri noktası sayısını temsil etmektedir.

Algoritma, rastgele seçilen küme merkezleriyle başlar, her veri noktası için üyelik derecelerini hesaplar ve bu dereceleri kullanarak küme merkezlerini günceller. Bu süreç, üyelik dereceleri ve küme merkezleri sabitlenene veya belirlenen maksimum iterasyon sayısına ulaşılan kadar tekrar

etmektedir. FCM'nin avantajları, özellikle küme sınırlarının net olmadığı durumlarda öne çıkmaktadır. Geleneksel sert kümelemeye kıyasla, FCM belirsizliği daha iyi modeller, geçiş bölgelerindeki veri noktalarını daha etkin bir şekilde ele alır ve aykırı değerlere karşı daha dayanıklıdır. Bu özellikler, FCM'yi görüntü segmentasyonu, örüntü tanıma, veri madenciliği, tıbbi görüntüleme ve kontrol sistemleri gibi çeşitli uygulama alanlarında etkili bir yöntem haline getirir. Özellikle, küme sınırlarının kesin olmadığı veri setlerinde, veri noktalarının birden fazla kümeye farklı derecelerde ait olabilmesi, daha gerçekçi ve esnek bir modelleme sağlamaktadır.

### 3.3. (SVs-based Clustering) Support Vector Tabanlı Kümeleme

Destek Vektör Tabanlı Kümeleme (SVs-based Clustering), destek vektör makinelerinin (SVM) prensiplerini kümeleme problemlerine entegre eden yenilikçi bir yaklaşımdır. Geleneksel kümeleme algoritmalarından farklı olarak, bu yöntem SVM'lerin veri noktalarını ayırma ve sınıflandırma yeteneklerini kullanarak karmaşık veri yapılarında etkili sonuçlar üretir. Çekirdek fonksiyonları ve hiperdüzlem kavramlarını temel alan bu yaklaşım, özellikle doğrusal olmayan küme sınırlarının bulunduğu durumlarda güçlü bir alternatif sunar. SVs Tabanlı Kümelemenin çalışma mekanizması, veri noktalarını yüksek boyutlu bir özellik uzayına dönüştürmekle başlar. Bu uzayda, en uygun ayırıcı hiperdüzlemler belirlenerek küme sınırları oluşturulur. Yöntemin yaygın türleri arasında Support Vector Clustering (SVC), One-Class SVM kümeleme ve kernel-based kümeleme yer alır. Çekirdek fonksiyonları, orijinal veri uzayında doğrusal olmayan sınırların yüksek boyutlu uzayda doğrusal hale getirilmesini sağlayarak, farklı şekil ve boyutlardaki kümelerin tespitini mümkün kılar. Bu özellik, yöntemin esnekliğini ve etkinliğini artırır. SV'ler, eğitim verilerinden seçilen ve küme merkezlerini belirlemede (Denklem 4) kullanılan bir alt kümedir. Yapılacak seçim,  $\varepsilon$ -duyarsız kayıp fonksiyonu (Denklem 2) ile belirlenmektedir:

$$L_{\varepsilon}(y, f(x)) = \begin{cases} 0 & \text{eğer } |y - f(x)| \leq \varepsilon, \\ |y - f(x)| - \varepsilon & \text{eğer } |y - f(x)| > \varepsilon, \end{cases} \quad (2)$$

Burada:

- $y$ : Gerçek çıktı (hedef değer),
- $f(x)$ : Model tahmini,
- $\varepsilon$ : Hata tolerans parametresi (SV'lerin seçimini etkiler).

SV'ler,  $|y - f(x)| > \varepsilon$  koşulunu sağlayan veri noktalarıdır. Bu, gürültü ve aykırı değerlerin dışarıda bırakılmasını sağlar.

SVs Tabanlı Kümelemenin avantajları, aykırı değerlere karşı sağlamlık, yüksek boyutlu verilerde iyi performans ve karmaşık küme yapılarını tespit etme yeteneği olarak öne çıkar. Bu

özellikler, yöntemi anomali tespiti, görüntü segmentasyonu, müşteri segmentasyonu ve biyoenformatik veri analizi gibi çeşitli uygulama alanlarında değerli kılar. Ancak, yöntemin hesaplama açısından yoğun olması, özellikle büyük veri setlerinde uygulanabilirliğini sınırlandırabilir. Buna rağmen, geleneksel yöntemlerin (K-means, FCM gibi) başarısız olduğu karmaşık veri yapılarında, SVs Tabanlı Kümeleme daha sağlam ve güvenilir sonuçlar sunar.

➤ *Destek Vektör Tabanlı Sert C-Ortalamlar* (SVs-based Hard C-Means, HCM) algoritması, geleneksel HCM algoritmasının sınırlamalarını aşmak amacıyla destek vektör makinelerinin (SVM) prensiplerini entegre etmektedir. Bu hibrit yaklaşımda, SVM'lerin çekirdek fonksiyonları kullanılarak veri noktaları yüksek boyutlu bir özellik uzayına dönüştürülmektedir. Bu uzayda, geleneksel HCM algoritması (Denklem 3) uygulanmakta ve her veri noktası yalnızca tek bir kümeye kesin üyelikle (0 veya 1) atanmaktadır. Söz konusu yöntem, doğrusal olmayan veri yapılarında daha etkin kümeleme performansı sergilemekte ve karmaşık küme sınırlarının belirlenmesinde üstünlük sağlamaktadır. Çekirdek fonksiyonu seçimine bağlı olarak farklı veri yapılarına uyarlanabilen bu algoritma, yüksek boyutlu verilerin işlenmesi konusunda da kapasite sunmaktadır.

$$J_{HCM} = \sum_{i=1}^c \sum_{x_k \in C_i} \|x_k - v_i\|^2 \quad (3)$$

Geleneksel HCM'de tüm veriler kullanılırken, burada yalnızca SV'ler (SS) dikkate alınmaktadır.  $v_i$ , i-inci küme merkezidir ve  $C_i$ , i-inci kümeye ait SV'leri içermektedir. Bu denklemde amaç, SV'ler arasındaki mesafeyi minimize etmektir.

$$v_i = \frac{1}{|S \cap C_i|} \sum_{x_k \in S \cap C_i} x_k \quad (4)$$

Denklem de  $v_i$  i-inci kümenin merkezidir ve yalnızca SV'lerin ortalaması alınarak güncellenmekte ve bu durum gürültü ve aykırı değerlerin etkisini azaltmaktadır.

➤ *Destek Vektör Tabanlı Bulanık C-Ortalamlar* (SVs-based Fuzzy C-Means, FCM) algoritması ise, FCM'nin bulanık mantık temelli esnekliğini SVM'lerin güçlü sınıflandırma yetenekleriyle birleştirmektedir. Bu yöntemde, veri noktaları yine çekirdek fonksiyonları aracılığıyla yüksek boyutlu bir uzaya dönüştürülmekte ve bu uzayda FCM algoritması, veri noktalarının birden fazla kümeye 0 ile 1 arasında değişen üyelik dereceleriyle atanmasını sağlamaktadır. Bulanıklık parametresi (m) aracılığıyla kümelemenin bulanıklık düzeyi ayarlanmakta, böylece belirsizlik içeren veri yapılarında daha etkin sonuçlar elde edilmektedir. Bu algoritma, küme sınırlarında yumuşak geçişler sağlamakta, aykırı değerlere karşı dayanıklılık göstermekte ve doğrusal olmayan karmaşık küme yapılarının belirlenmesinde başarılı olmaktadır. SV'ler tabanlı FCM'de, yalnızca SV'ler üzerinden hesaplama yapılmaktadır.

$$J_{SV-FCM} = \sum_{i=1}^c \sum_{x_k \in S} u_{ik}^m \|x_k - v_i\|^2 \quad (5)$$

$u_{ik}$ ,  $x_k$  veri noktasının i-inci kümeye üyelik derecesidir ve m bulanıklaştırma katsayısını temsil etmektedir. Sadece SV'ler (S) üzerinden hesaplama yapılmaktadır.



- Üyelik derecelerinin güncellenmesi (Denklem 6):

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left( \frac{\|x_k - v_i\|}{\|x_k - v_j\|} \right)^{\frac{2}{m-1}}} \quad (6)$$

$u_{ik}$ ,  $x_k$  ve  $v_i$  arasındaki mesafeye bağlı olarak hesaplanır.  $m=2$  için bu, bulanık kümeler arasındaki geçişi yumuşatmaktadır.

- Küme merkezleri (Denklem 7):

$$v_i = \frac{\sum_{x_k \in S} u_{ik}^m x_k}{\sum_{x_k \in S} u_{ik}^m} \quad (7)$$

$v_i$ , SV'lerin ağırlıklı ortalamasıdır; ağırlıklar  $u_{ik}^m$  ile belirlenir. Bu, gürültüden arındırılmış küme merkezleri sağlamaktadır.

Burada:

- $u_{ik}^m$ :  $x_k$  veri noktasının  $i$  kümeye üyelik derecesi
- $m$ : Bulanıklaştırma katsayısı
- $n$ : Veri noktası sayısını temsil etmektedir.

Her iki algoritma da, geleneksel C-Ortalamlar algoritmalarının kısıtlamalarını aşarak, destek vektör makinelerinin yüksek boyutlu uzaylarda veri ayrımı yapma kapasitesinden faydalanmaktadır. SVs-based HCM, kesin küme atamalarıyla daha net sınırlar oluştururken, SVs-based FCM, belirsizlik içeren veri setlerinde esnek ve gerçekçi bir modelleme sunmaktadır. Bu hibrit yaklaşımlar, özellikle doğrusal olmayan veri yapılarının bulunduğu ve geleneksel kümeleme yöntemlerinin yetersiz kaldığı senaryolarda, daha sağlam ve güvenilir kümeleme sonuçları elde edilmesini sağlamaktadır.

➤ *L2 norm regularizasyonu*, makine öğrenmesi ve özellikle derin öğrenme modellerinde aşırı öğrenmeyi (overfitting) önlemek amacıyla yaygın bir şekilde kullanılmaktadır. Ağırlık azaltma (weight decay) veya ridge regularizasyonu olarak da adlandırılan bu yöntem, modelin maliyet fonksiyonuna bir ceza terimi eklenmesiyle işlev görmektedir. Ceza terimi, model parametrelerinin (ağırlıkların) karelerinin toplamı olarak hesaplanmakta ve matematiksel olarak (Denklem 8):

$$L2 = \lambda * \Sigma(w^2) \quad (8)$$

şeklinde ifade edilmektedir. Burada,  $\lambda$  (lambda) hiperparametresi regularizasyonun etkisini kontrol etmekte,  $w$  ise model ağırlıklarını temsil etmektedir. Böylece, yeni maliyet fonksiyonu (Denklem 9)

$$J(w) = \text{Orijinal\_Maliyet} + \lambda * \Sigma(w^2) \quad (9)$$

şekline dönüşmektedir.

Yapılan regularizasyon tekniğinin temel avantajları arasında, modelin tüm özelliklerini kullanmaya teşvik etmesi ve genelleme performansını artırması yer almaktadır. L2 norm regularizasyonu (Denklem 10), ağırlıkları sıfıra yakın değerlere doğru küçültmekte, ancak tam sıfır yapmamaktadır; bu da özellik sıkıştırmasını desteklemekte ve özellik seçiminden ziyade tüm özelliklerin dengeli bir şekilde kullanılmasını sağlamaktadır. Özellikle yüksek boyutlu veri setleri ve derin sinir ağlarında etkili olan bu yöntem, gradyan inişi sırasında ağırlıkların daha hızlı küçülmesine olanak tanımakta ve böylece modelin karmaşıklığını azaltarak aşırı öğrenme riskini en aza indirmektedir. L2 norm regularizasyonunun uygulanması, modelin veri üzerindeki genelleme yeteneğini güçlendirmekte ve yüksek varyanslı modellerde sağlam bir çözüm sunmaktadır. Hata fonksiyonu (Denklem 11):

$$J_{RFCNN} = \sum_{k=1}^n (y_k - t_k)^2 + \lambda \sum_{i=1}^c \sum_{j=0}^d a_{ij}^2 \quad (10)$$

$\lambda$  : Düzenlileştirme parametresi.

$$J_{LSE} = \sum_{k=1}^n (y_k - t_k)^2 \quad (11)$$

Burada  $t_k$ , gerçek hedef değerdir; amaç, model tahmini  $y_k$  ile  $t_k$  arasındaki farkı minimize etmektir.

Uygun bir  $\lambda$  değeri seçimi, modelin öğrenme kapasitesini korurken aşırı öğrenme sorununu etkili bir şekilde ele almaktadır. Bu özellikleriyle, L2 norm regularizasyonu, modern makine öğrenmesi uygulamalarında vazgeçilmez bir araç olarak kabul edilmektedir.

### 3.4. (RFCNN) Robust Fuzzy Clustering-Based Neural Networks

Sağlam Bulanık Kümeleme Tabanlı Sinir Ağları (Robust Fuzzy Clustering-Based Neural Networks - RFCNN), makine öğrenmesi ve yapay zeka alanında kullanılan hibrit bir modeldir. Bu model, bulanık mantık ile sinir ağlarının güçlü yönlerini birleştirerek daha sağlam ve etkili bir sistem oluşturmayı amaçlar. Örüntü tanıma, sınıflandırma, zaman serisi tahmini, görüntü işleme ve kontrol sistemleri gibi birçok alanda kullanılabilir. RFCNN' i temel olarak 3 katmanlı bir sinir yapısı olarak düşünülebilmektedir. İlk katman olan "Giriş Katmanı" nda özellik vektörleri bulunmaktadır. "Gizli Katman" da ise Küme merkezleri ve üyelik dereceleri bulunmaktadır. Ayrıca bu katmanda her küme için bir düğüm bulunmaktadır. Son katman olan "Çıkış Katmanı"nda ise tahmin edilen değer (her kümenin yerel lineer modelinin üyelik dereceleriyle ağırlıklandırılmış toplamıdır) elde edilmektedir.

Çalışmada kullanılan modele ait temel akış diyagramı Şekil 1' de gösterilmektedir. Model ilk aşamada girdi olarak veri setini almakta ve kullanılabilir hale getirmektedir. Daha sonra deneme yapılacak parametrelere ait sınırlar belirlenmektedir. Bu aşamada 10 katlı çapraz doğrulama

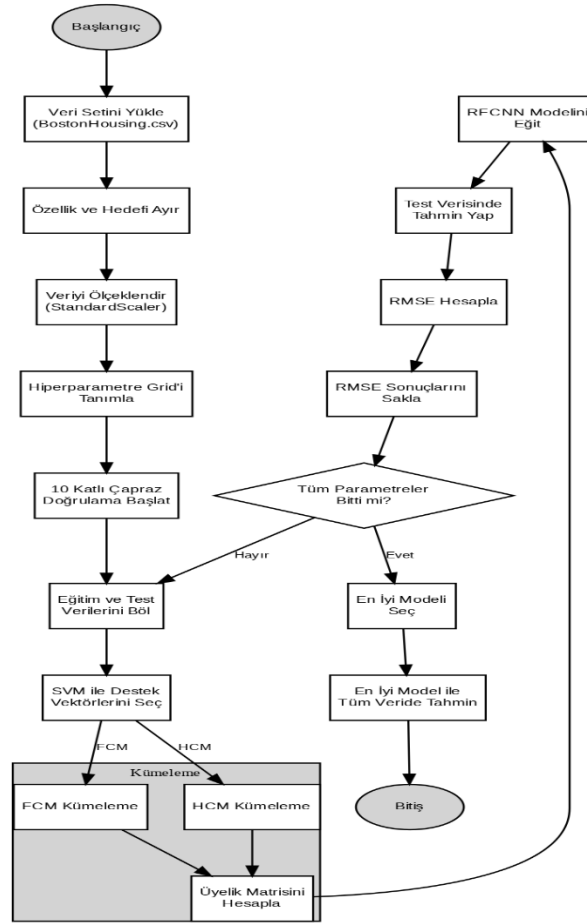
kullanılmaktadır. Sonraki aşamada SVM ile birlikte HCM ve FCM yöntemleri kullanılarak denemeler gerçekleştirilmektedir. Elde edilen değerlendirme sonuçlarına göre RMSE değerleri hesaplanmaktadır. Sonuçlara göre nihai olarak kullanılacak yöntem ve ağırlık değerleri belirlenmektedir. Son aşamada ise tüm veri seti kullanılarak tahmin işlemi gerçekleştirilmektedir. Genel RFCNN modeli aşağıda (Denkem 12) gösterilmektedir.

$$\hat{y}_i = \sum_{j=1}^C u_{ij}(\mathbf{w}_j^T \mathbf{x}_i + b_j) \quad (12)$$

Burada:

$u_{ij}$ : i-inci veri noktasının j-inci kümeye üyelik derecesi,

$\mathbf{w}_j, b_j$ : j-inci küme için lineer regresyon parametre değerlerini ifade etmektedir.



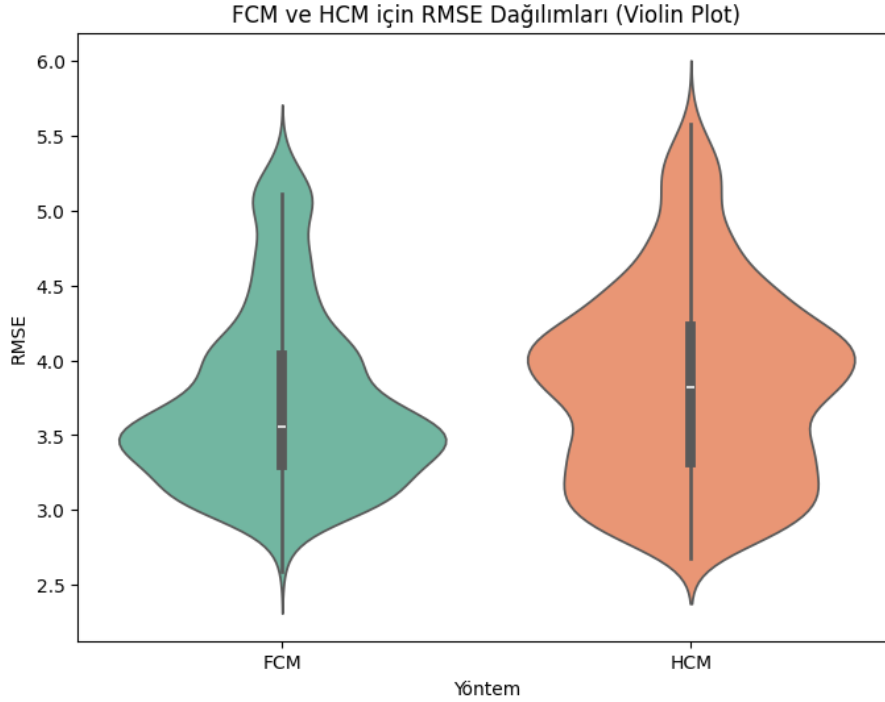
**Şekil 1** Önerilen yönteme ait akış şeması

Modelin performansı, Denklem 13' de matematiksel formülü gösterilen Kök Ortalama Kare Hata (RMSE) ile ölçülmektedir.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (13)$$

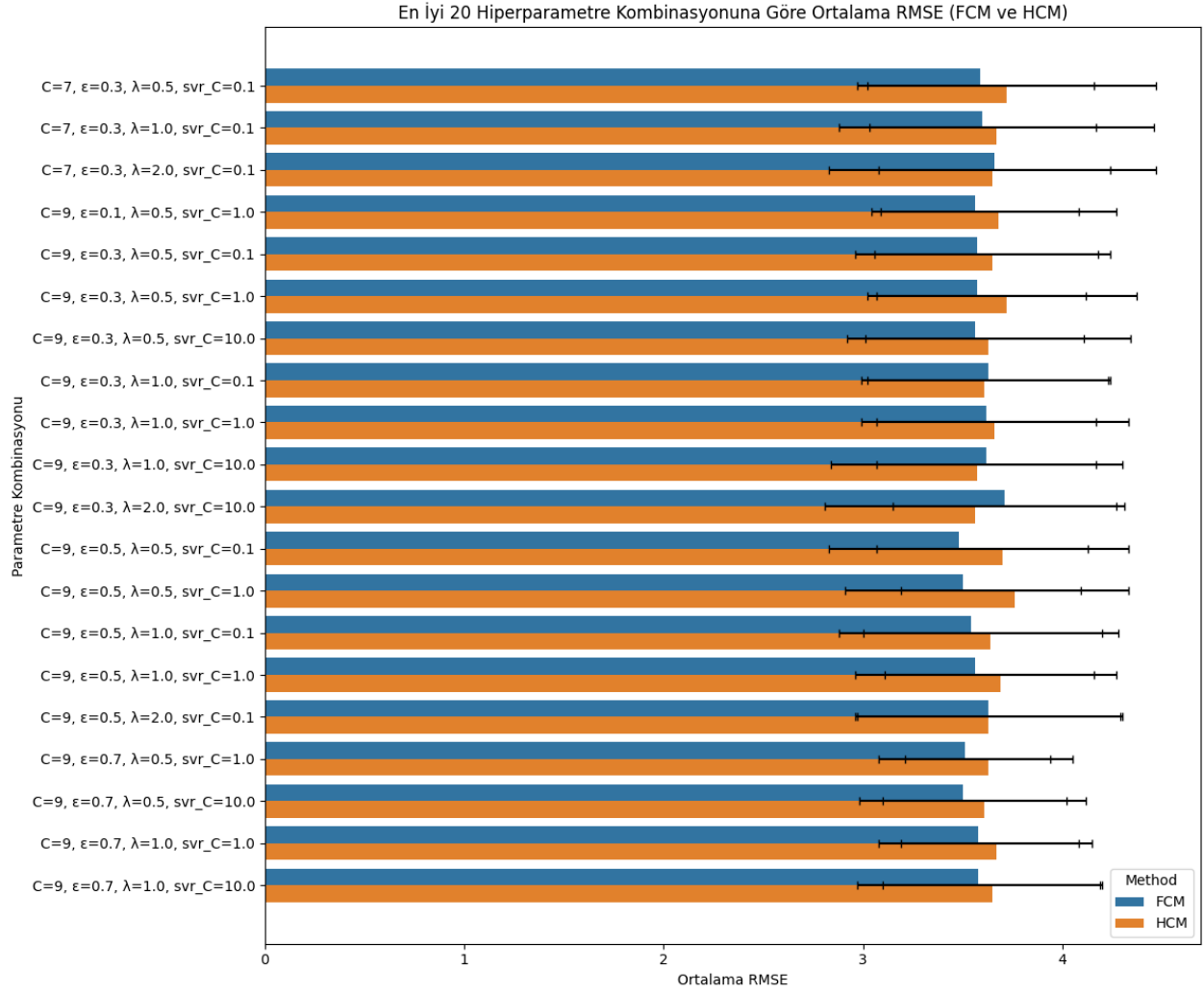
#### 4. DENEYSEL SONUÇLAR

İki farklı yöntem (FCM - Fuzzy C-Means ve HCM - Hard C-Means) kullanılmıştır. FCM, bulanık kümeleme yaparken, HCM sert kümeleme yapar. Her kombinasyon için bu iki yöntemin performansları ayrı ayrı hesaplanmıştır. Her kombinasyon için, çapraz doğrulama (cross-validation) kullanılarak ortalama RMSE ve standart sapma hesaplanmıştır. FCM ve HCM yöntemlerinin farklı parametreler ile elde etmiş oldukları kombinasyonlara ait RMSE dağılımlarının genel görünümü Şekil 2’de gösterilmektedir. Tablo 2’de ve Tablo 3’te sırasıyla FCM ve HCM modellerinin her hiperparametre kombinasyonu için ne kadar iyi performans gösterdiğini ve sonuçların ne kadar tutarlı olduğuna ait değerler içerisindeki en düşük RMSE sonuçlarına sahip modellerin parametre değerleri gösterilmektedir.



**Şekil 2** FCM ve HCM yöntemlerinin elde etmiş oldukları RMSE dağılımları

En iyi RMSE değerlerine sahip ilk 20 hiperparametre kombinasyonları için FCM ve HCM'nin ortalama RMSE değerlerini karşılaştıran çubuk grafik Şekil 3’te gösterilmektedir. Grafikte X eksi, y eksenini, hata çubukları sırasıyla ortalama RMSE değerlerini, Hiperparametre kombinasyonlarını ve RMSE'nin standart sapmasını göstermektedir.



**Şekil 3 2** Yönteme ait en iyi ilk 20 parametre değeri

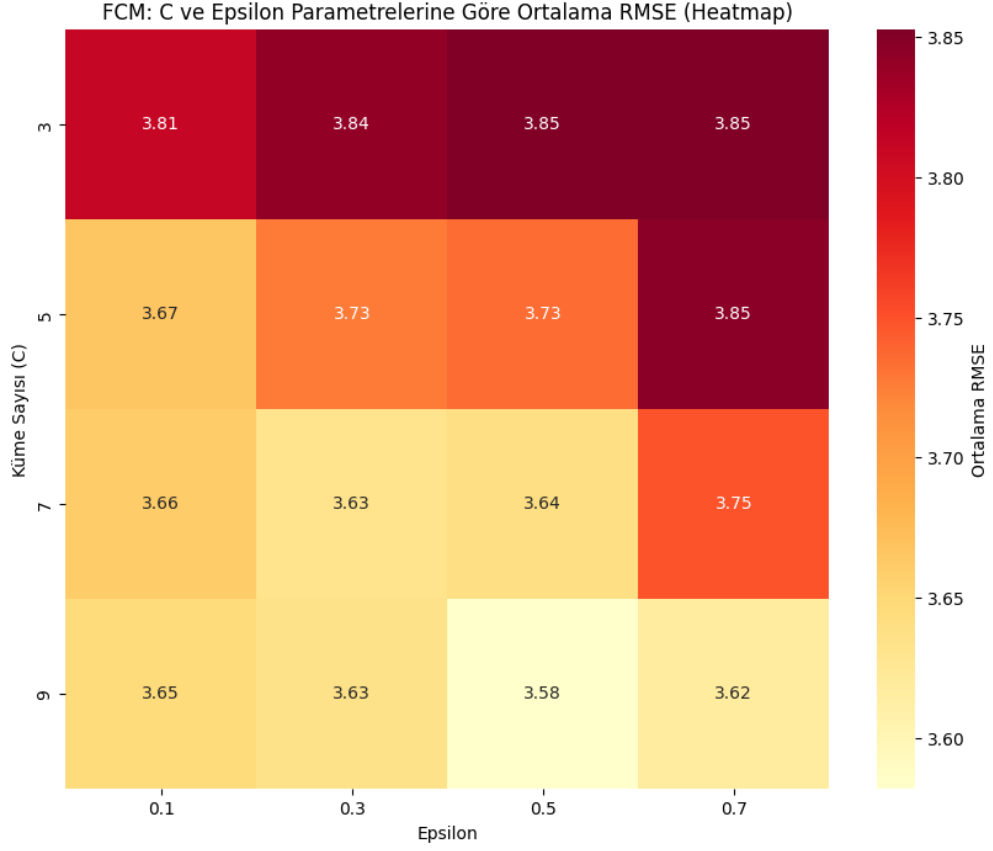
Tablo 2 ve Tablo 3’te bulunan c, epsilon, lambda reg., Svr. C, Mean RMSE ve Std. RMSE sütunları sırasıyla küme sayısını (3,5,7,9), SVR modelindeki tüp genişliğini (0.1, 0.3, 0.5, 0.7), Düzenleştirme parametresini (0.5, 1.0, 2.0), SVR modelindeki ceza parametresini (0.1, 1.0, 10.0), 10 katlı çapraz doğrulama ile hesaplanan ortalama RMSE (Kök Ortalama Kare Hata) değerini ve son olarak RMSE'nin standart sapmasını yani modelin kararlılığına ait değerleri göstermektedir.

**Tablo 2** En düşük RMSE değerine ait ilk 20 FCM Sonuçları

C	Epsilon	Lambda Reg.	Svr. C	Mean RMSE	Std RMSE
9	0.5	0.5	0.1	3.483731	0.6528
9	0.5	0.5	1	3.500019	0.592589
9	0.7	0.5	10	3.500409	0.520057

9	0.7	0.5	1	3.510927	0.432477
7	0.5	0.5	1	3.531731	0.547023
9	0.5	1	0.1	3.5362	0.656153
9	0.1	0.5	10	3.538107	0.533904
9	0.5	1	1	3.55624	0.598082
9	0.1	0.5	1	3.556668	0.524176
7	0.1	0.5	10	3.557621	0.533312
9	0.3	0.5	10	3.559533	0.55305
9	0.5	0.5	10	3.563025	0.556848
9	0.3	0.5	1	3.567895	0.54632
7	0.5	1	1	3.570918	0.573605
9	0.3	0.5	0.1	3.573473	0.614499
9	0.7	1	10	3.579305	0.609528
9	0.7	1	1	3.581949	0.504337
7	0.3	0.5	1	3.583451	0.579147
7	0.3	0.5	0.1	3.588739	0.573435
9	0.7	0.5	0.1	3.594697	0.701967

En İyi FCM parametreleri  $C=9.0$ ,  $\epsilon=0.5$ ,  $\lambda_{reg}=0.5$ ,  $svr\_C=0.1$ , Ortalama RMSE:  $3.48 \pm .65$  değerlerini elde etmiştir. En İyi HCM parametreleri  $C=9.0$ ,  $\epsilon=0.3$ ,  $\lambda_{reg}=2.0$ ,  $svr\_C=10.0$ , Ortalama RMSE:  $3.56 \pm 0.75$  değerlerini elde etmiştir. Elde edilen sonuçlara bakıldığında, daha düşük RMSE değeri nedeniyle FCM yöntemi seçilmiştir. Ayrıca Şekil 4'te FCM için küme sayısı (C) ve epsilon hiperparametrelerinin ortalama RMSE üzerindeki etkisini bir ısı haritası ile gösterilmektedir. Grafikte X eksenini, y eksenini sırasıyla epsilon değerlerini ve küme sayılarını göstermektedir. Koyu renkler daha yüksek RMSE, açık renkler daha düşük RMSE anlamına gelmektedir. Ek olarak Hangi C ve epsilon kombinasyonlarının daha iyi performans gösterdiğini (düşük RMSE) hızlıca gösterir. Örneğin, açık renkli bölgeler en iyi hiperparametre kombinasyonlarını işaret eder.



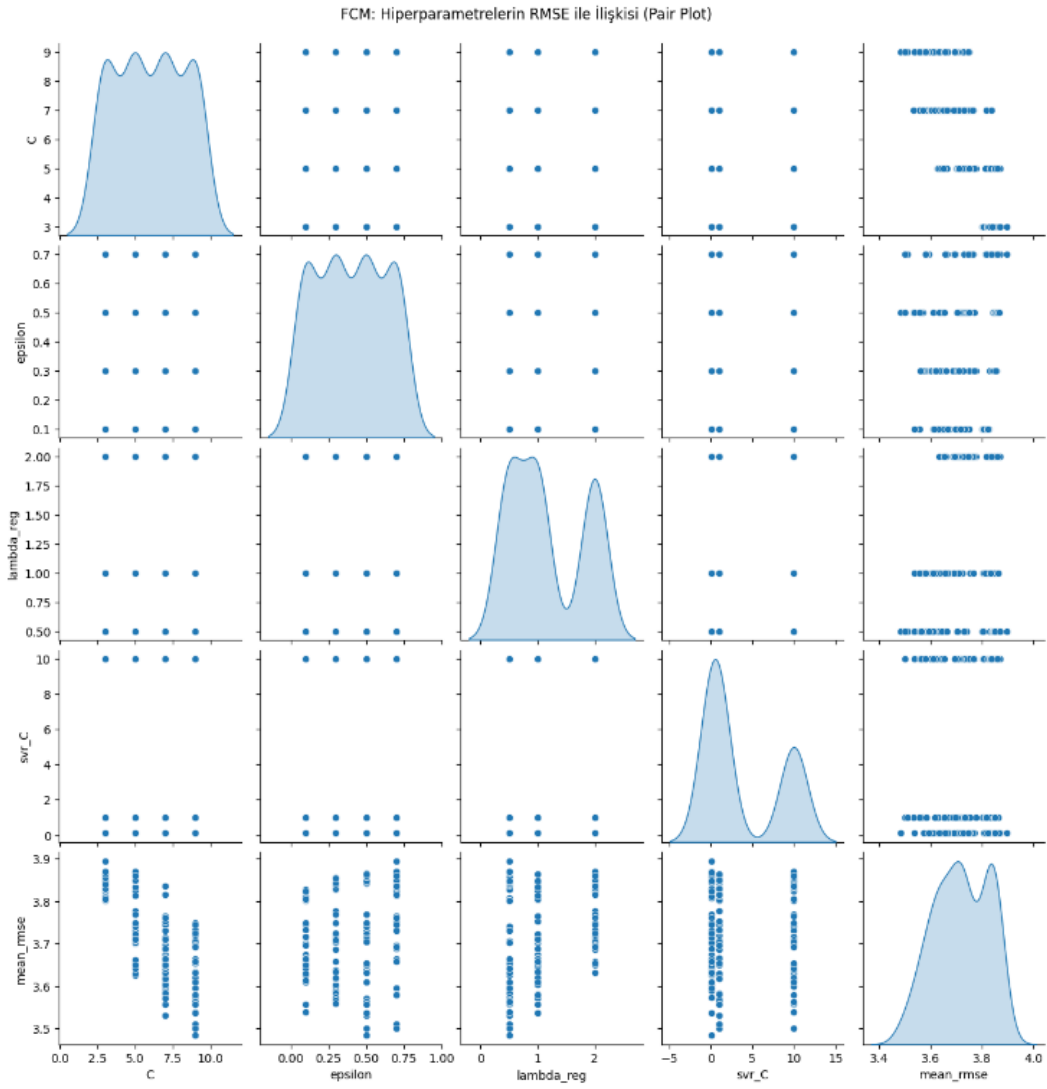
**Şekil 4** FCM için Hiperparametrelerin RMSE Üzerindeki Etkisi

*Tablo 3 En düşük RMSE değerine ait ilk 20 HCM Sonuçları*

C	Epsilon	Lambda Reg.	Svr. C	Mean RMSE	Std RMSE
9	0.3	2	10	3.564834	0.753732
9	0.3	1	10	3.570923	0.728228
9	0.3	2	0.1	3.603298	0.66495
9	0.3	1	0.1	3.60645	0.618006
9	0.7	0.5	10	3.607488	0.506056
9	0.5	2	0.1	3.626566	0.671904
9	0.3	0.5	10	3.626862	0.708932
9	0.7	0.5	1	3.628659	0.423697
9	0.5	1	0.1	3.637204	0.640888
7	0.3	2	0.1	3.64612	0.820137
9	0.7	1	10	3.646716	0.552059
9	0.3	2	1	3.653621	0.700818
9	0.3	0.5	0.1	3.654086	0.588835
9	0.3	1	1	3.663218	0.67086
7	0.3	1	0.1	3.665445	0.785362

9	0.7	1	1	3.667082	0.477175
9	0.1	1	1	3.677685	0.629226
9	0.5	2	1	3.680103	0.622782
9	0.1	0.5	1	3.680856	0.592674
9	0.5	1	1	3.694887	0.58246
9	0.1	2	1	3.696089	0.660652

En iyi FCM modeli ( $C=9.0$ ,  $\epsilon=0.5$ ,  $\lambda_{reg}=0.5$ ,  $svr\_C=0.1$ ), tüm veri seti üzerinde eğitilmiş ve test edilmiştir. RMSE değeri 2.95 olarak hesaplanmıştır, bu da çapraz doğrulama sonuçlarından (3.48) daha iyidir. Bu, modelin tüm veri setinde daha iyi genelleme yaptığını gösterebilir.



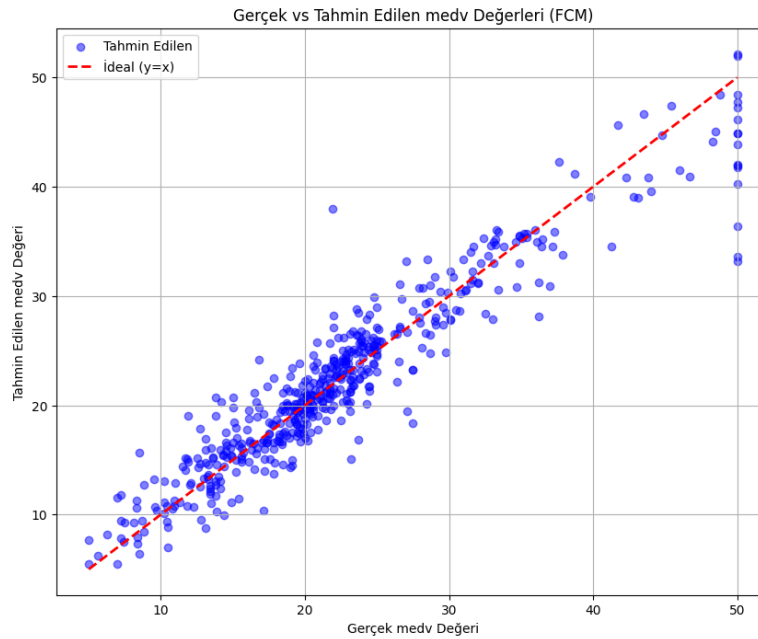
Şekil 5 Hiperparametrelerin RMSE Üzerindeki Etkisi



FCM için hiperparametrelerin (C, epsilon, lambda\_reg, svr\_C) ve ortalama RMSE arasındaki ilişkileri bir çift grafik ile Şekil 5' te gösterilmektedir. Grafikte diyagonal, her hiperparametrenin ve RMSE'nin yoğunluk grafikleri, diğer Hücreler ise hiperparametreler ve RMSE arasındaki dağılım grafikleri göstermektedir. Örnek olarak, belirli bir C veya epsilon değerinin düşük RMSE ile ilişkili olup olmadığını gösterilmektedir.

Tablo 4 İlk 10 Veri Noktası için Gerçek ve Tahmin Edilen Değerler

Index	Gerçek Medv	Tahmin Edilen Medv	Hata
0	24	25.453536	1.453536
1	21.6	22.496893	0.896893
2	34.7	30.803722	3.896278
3	33.4	30.600739	2.799261
4	36.2	31.236294	4.963706
5	28.7	24.797766	3.902234
6	22.9	20.506899	2.393101
7	27.1	19.457122	7.642878
8	16.5	14.168424	2.331576
9	18.9	18.368304	0.531696



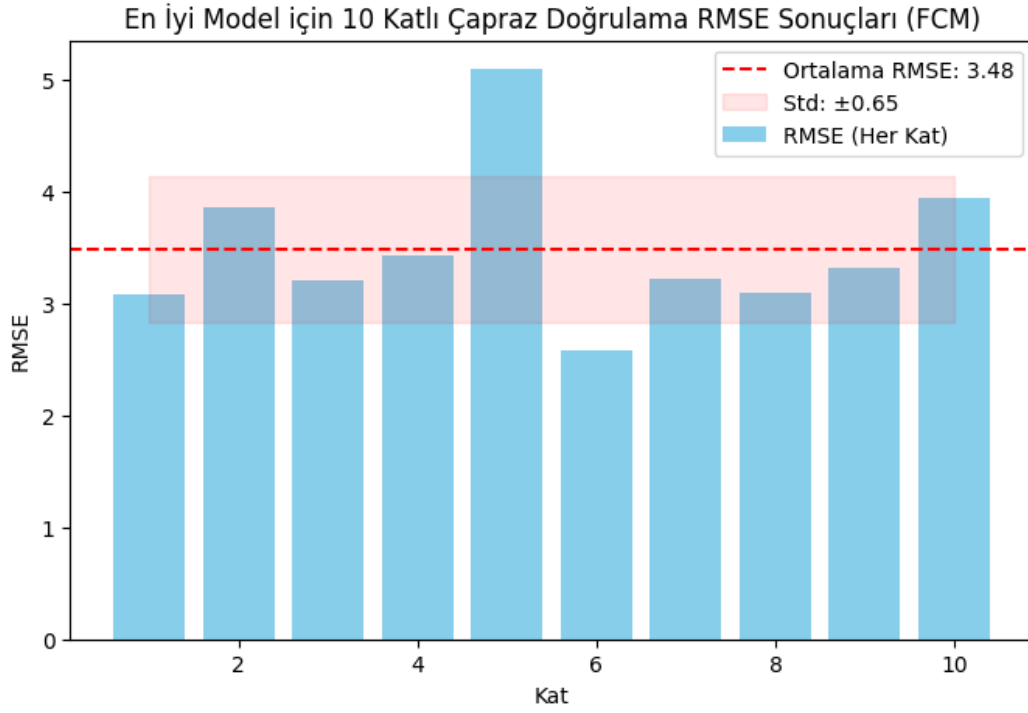
Şekil 6 FCM Yöntemine ait gerçek ve tahmin değer dağılımları

Tüm veri seti için gerçek medv değerleri ile en iyi modelin tahmin ettiği medv değerlerini karşılaştıran dağılım Şekil 6’te gösterilmektedir. X eksenini gerçek medv değerini, y eksenini ise tahmin edilen medv değerini göstermektedir. Grafikte kırmızı çizgi  $x=y$  denklemini (ideal tahmin) göstermektedir. Bu nedenle noktalar kırmızı çizgiye ( $y=x$ ) ne kadar yakınsa, model o kadar doğru tahmin yapmıştır. Dağılımın çizgiye yakınlığı, modelin doğruluğunu gösterir.



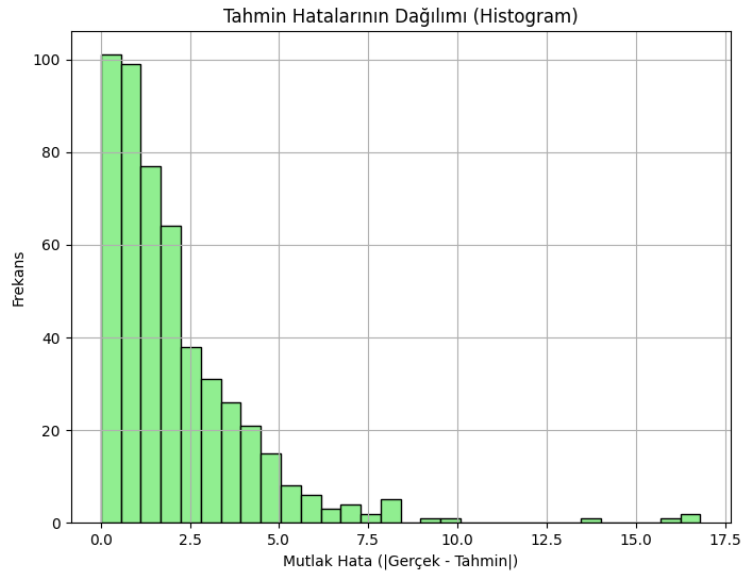
**Şekil 7** Tahmin hatalarının dağılımı

Şekil 7’de elde edilen tahmin verileri ile gerçek veriler arasındaki ilişki, tahmin hatalarının ( $|Gerçek - Tahmin|$ ) dağılımını bir kutu grafiği ile gösterilmektedir. Y eksenini mutlak hatayı göstermektedir. Hataların merkezi eğilimini (medyan), yayılımını ve aykırı değerlerini göstermektedir. Dar bir kutu ve az aykırı değer, modelin tutarlı ve doğru olduğunu belirtir.



**Şekil 8** En İyi Modelin Çapraz Doğrulama RMSE Grafiği

En iyi model için 10 katlı çapraz doğrulama RMSE değerleri Şekil 8’te gösterilmektedir. Şekil 8’de kırmızı çizgi ortalama RMSE değerini, kırmızı alan ise ortalama RMSE’ye ait standart sapma aralığını temsil etmektedir. Düşük ve tutarlı RMSE çubukları, modelin güvenilir olduğunu belirtmektedir.



**Şekil 9** Hata Dağılımı

Şekil 9’da hataların nasıl dağıldığı gösterilmektedir. Histogramın sıfıra yakın yoğunlaşması, modelin doğru tahminler yaptığını göstermektedir.

## 5. SONUÇ VE DEĞERLENDİRME

Bu çalışma, FCM ve HCM kümeleme yöntemlerinin, çeşitli hiperparametre kombinasyonları altında tahmin performanslarını değerlendirmek amacıyla gerçekleştirilmiştir. Toplamda 144 farklı hiperparametre kombinasyonu ( $C \in \{3, 5, 7, 9\}$ ,  $\varepsilon \in \{0.1, 0.3, 0.5, 0.7\}$ ,  $\lambda \in \{0.5, 1.0, 2.0\}$ ,  $\text{svr\_C} \in \{0.1, 1.0, 10.0\}$ ) test edilmiş ve her bir kombinasyon için ortalama RMSE (Root Mean Square Error) ve standart sapma ( $\text{std\_rmse}$ ) değerleri hesaplanmıştır. Performans karşılaştırması, tüm kombinasyonları kapsayan bir çubuk grafik (Şekil 1) ve en iyi 20 kombinasyonu gösteren daha odaklı bir grafik (Şekil 2) ile görselleştirilmiştir. Bulgular, her iki yöntemin de belirli hiperparametre setlerinde yüksek performans sergilediğini, ancak HCM'nin bazı kombinasyonlarda daha düşük hata oranları ve daha yüksek kararlılık sunduğunu ortaya koymuştur. En iyi 20 kombinasyonun analizi, özellikle yüksek küme sayıları ( $C=7$  ve  $C=9$ ) ve orta düzey  $\varepsilon$  değerleri (0.3 ve 0.5) ile  $\lambda=1.0$  gibi regularizasyon parametrelerinin hem FCM hem de HCM için optimal sonuçlar ürettiğini göstermiştir. Örneğin,  $C=9$ ,  $\varepsilon=0.5$ ,  $\lambda=1.0$ ,  $\text{svr\_C}=1.0$  kombinasyonu, FCM için 3.50 ve HCM için 3.61 ortalama RMSE değerleriyle en düşük hata oranlarından birini sergilemiştir. Ayrıca,  $\text{svr\_C}=1.0$  ve  $\text{svr\_C}=10.0$  gibi daha yüksek regularizasyon değerlerinin genellikle daha iyi performansla ilişkili olduğu gözlemlenmiştir. Standart sapma analizleri, HCM'nin bazı kombinasyonlarda (örneğin,  $C=9$ ,  $\varepsilon=0.7$ ,  $\lambda=0.5$ ,  $\text{svr\_C}=1.0$  için  $\text{std\_rmse}=0.42$ ) FCM'ye kıyasla daha düşük varyans gösterdiğini ve dolayısıyla daha kararlı tahminler sunduğunu ortaya koymuştur. Buna karşılık, FCM'nin bulanık kümeleme yapısı, bazı kombinasyonlarda daha esnek ancak daha yüksek varyanslı sonuçlar üretmiştir (örneğin,  $C=3$ ,  $\varepsilon=0.7$ ,  $\lambda=2.0$ ,  $\text{svr\_C}=10.0$  için  $\text{std\_rmse}=0.70$ ). FCM ve HCM arasındaki performans farkları, genellikle küçük olsa da, HCM'nin sert kümeleme yaklaşımının belirli hiperparametre setlerinde daha kesin tahminler sağladığı görülmüştür. Örneğin,  $C=7$ ,  $\varepsilon=0.3$ ,  $\lambda=1.0$ ,  $\text{svr\_C}=0.1$  kombinasyonunda HCM (RMSE=3.67), FCM'ye (RMSE=3.60) kıyasla daha iyi performans göstermiştir. Bu bulgular, HCM'nin veri noktalarını kesin kümelerle ayırma yeteneğinin, özellikle düzenli veri yapılarında avantaj sağlayabileceğini öne sürmektedir. Bununla birlikte, FCM'nin esnek yapısı, daha karmaşık veya belirsiz veri setlerinde potansiyel olarak daha iyi sonuçlar verebilir. Her iki yöntemin performansı, hiperparametre seçimine güçlü bir şekilde bağlıdır ve bu çalışma, optimal parametre aralıklarını belirlemede önemli bir rehber sunmaktadır.

Sonuç olarak, bu çalışma, FCM ve HCM'nin tahmin performanslarını karşılaştırmalı olarak değerlendirmiş ve her iki yöntemin de uygun hiperparametrelerle yüksek doğruluk sağladığını göstermiştir. HCM, belirli kombinasyonlarda daha düşük RMSE ve standart sapma değerleriyle hafif bir üstünlük sağlarken, FCM'nin bulanık kümeleme yaklaşımı farklı senaryolarda avantajlı olabilir. Gelecekteki çalışmalar, bu bulguları temel alarak daha dar hiperparametre aralıklarında ince ayar yapabilir veya farklı veri setleriyle yöntemlerin genelleme kabiliyetlerini test edebilir. Ayrıca, hesaplama karmaşıklığı ve eğitim süreleri gibi ek metriklerin dahil edilmesi, bu yöntemlerin pratik uygulanabilirliğini daha iyi anlamak için faydalı olacaktır. Bu çalışma,

kümeleme tabanlı tahmin modellerinin optimizasyonu için sağlam bir temel sunmakta ve hiperparametre ayarının önemini vurgulamaktadır.

## KAYNAKÇA

- Atashgahi, Z., Liu, T., Pechenizkiy, M., Veldhuis, R., Mocanu, D. C., & Schaar, M. van der. (2024). *Unveiling the Power of Sparse Neural Networks for Feature Selection* (No. arXiv:2408.04583). arXiv. <https://doi.org/10.48550/arXiv.2408.04583>
- Bu, F., Zhang, C., Kim, E.-H., Yang, D., Fu, Z., & Pedrycz, W. (2024). Fuzzy clustering-based neural network based on linear fitting residual-driven weighted fuzzy clustering and convolutional regularization strategy. *Applied Soft Computing*, 154, 111403. <https://doi.org/10.1016/j.asoc.2024.111403>
- de Campos Souza, P. V., & Lughofer, E. (2022). An advanced interpretable Fuzzy Neural Network model based on uni-nullneuron constructed from n-uninorms. *Fuzzy Sets and Systems*, 426, 1–26. <https://doi.org/10.1016/j.fss.2020.11.019>
- Huang, H., Oh, S.-K., Wu, C.-K., & Pedrycz, W. (2022a). Double iterative learning-based polynomial based-RBFNNs driven by the aid of support vector-based kernel fuzzy clustering and least absolute shrinkage deviations. *Fuzzy Sets and Systems*, 443, 30–49. <https://doi.org/10.1016/j.fss.2021.10.007>
- Huang, H., Oh, S.-K., Wu, C.-K., & Pedrycz, W. (2022b). Fuzzy clustering-based neural networks modelling reinforced with the aid of support vectors-based clustering and regularization technique. *Neurocomputing*, 482, 139–153. <https://doi.org/10.1016/j.neucom.2021.11.029>
- Huang, H., Oh, S.-K., Wu, C.-K., & Pedrycz, W. (2022c). Fuzzy clustering-based neural networks modelling reinforced with the aid of support vectors-based clustering and regularization technique. *Neurocomputing*, 482, 139–153. <https://doi.org/10.1016/j.neucom.2021.11.029>
- Latifi, S., Jannach, D., & Ferraro, A. (2022). Sequential recommendation: A study on transformers, nearest neighbors and sampled metrics. *Information Sciences*, 609, 660–678. <https://doi.org/10.1016/j.ins.2022.07.079>
- Liu, F., Dong, T., Liu, Q., Liu, Y., & Li, S. (2024). Combining fuzzy clustering and improved long short-term memory neural networks for short-term load forecasting. *Electric Power Systems Research*, 226, 109967. <https://doi.org/10.1016/j.epsr.2023.109967>
- Liu, S., Oh, S.-K., Pedrycz, W., Yang, B., Wang, L., & Yoon, J. H. (2024). Reinforced Interval Type-2 Fuzzy Clustering-Based Neural Network Realized Through Attention-Based Clustering Mechanism and Successive Learning. *IEEE Transactions on Fuzzy Systems*, 32(3), 1208–1222. <https://doi.org/10.1109/TFUZZ.2023.3321197>
- Sreedhara, S. H., Kumar, V., & Salma, S. (2023). Efficient Big Data Clustering Using Adhoc Fuzzy C Means and Auto-Encoder CNN. In S. Smys, K. A. Kamel, & R. Palanisamy (Eds.), *Inventive Computation and Information Technologies* (pp. 353–368). Springer Nature. [https://doi.org/10.1007/978-981-19-7402-1\\_25](https://doi.org/10.1007/978-981-19-7402-1_25)
- Tran, A. T., Tran, B. Q., & Luong, K. T. (2023). Improvement for Large-Scale Image Data using Fuzzy Rough C-Mean Based Unsupervised CNN Clustering: An Empirical Study on designbyhumans.com. *Proceedings of the 2023 12th International Conference on Software and Computer Applications*, 1–7. <https://doi.org/10.1145/3587828.3587829>
- Wang, J., Wang, H., Nie, F., & Li, X. (2023). Feature selection with multi-class logistic regression. *Neurocomputing*, 543, 126268. <https://doi.org/10.1016/j.neucom.2023.126268>
- Wang, Z., Kim, E.-H., Oh, S.-K., Pedrycz, W., Fu, Z., & Yoon, J. H. (2024). Reinforced Fuzzy-Rule-Based Neural Networks Realized Through Streamlined Feature Selection Strategy and Fuzzy Clustering With Distance Variation. *IEEE Transactions on Fuzzy Systems*, 32(10), 5674–5686. <https://doi.org/10.1109/TFUZZ.2024.3422414>
- Wang, Z., Oh, S.-K., Fu, Z., & Pedrycz, W. (2022). *Fuzzy Set-based Hybrid Neural Networks Driven with the Aid of Polynomial Neuron/Fuzzy Polynomial Neuron Constructed through Feature Selection*. Research Square. <https://doi.org/10.21203/rs.3.rs-709504/v1>
- Wang, Z., Oh, S.-K., Wang, Z., Fu, Z., Pedrycz, W., & Yoon, J. H. (2023). Design of progressive fuzzy polynomial neural networks through gated recurrent unit structure and correlation/probabilistic selection strategies. *Fuzzy Sets and Systems*, 470, 108656. <https://doi.org/10.1016/j.fss.2023.108656>
- Xu, X., Wu, X., Wei, F., Zhong, W., & Nie, F. (2022). A General Framework for Feature Selection Under Orthogonal Regression With Global Redundancy Minimization. *IEEE Transactions on Knowledge and Data Engineering*, 34(11), 5056–5069. <https://doi.org/10.1109/TKDE.2021.3059523>
- Zhao, M., Song, Y., Huang, H., & Kim, E.-H. (2025). Attention-based fuzzy neural networks designed for early warning of financial crises of listed companies. *Information Sciences*, 686, 121374. <https://doi.org/10.1016/j.ins.2024.121374>