

## TÜBİTAK 2209-A ÜNİVERSİTE ÖĞRENCİLERİ YURT İÇİ ARAŞTIRMA PROJELERİ DESTEK PROGRAMI

<b>Proje Yürütücüsünün Adı Soyadı:</b> Ömer KÖROĞLU
<b>Araştırma Önerisinin Başlığı:</b> SOS-Tabanlı Meta-Sezgisel Kümeleme Algoritmasının Tasarımı ve Geliştirilmesi
<b>Danışmanın Adı Soyadı:</b> Doç. Dr. Hamdi Tolga KAHRAMAN
<b>Araştırmanın Yürütüleceği Kurum/Kuruluş:</b> Karadeniz Teknik Üniversitesi

### ÖZET

Nesnelerin interneti alanında yaşanan gelişmeler, elektronik ortamların/cihazların, bu ortamlardaki uygulamaların ve bu uygulamaları kullananların sayısındaki artış, verinin de beklenmedik bir hızla büyümesine yol açmaktadır. Bu gelişmelere bağlı olarak ortaya çıkan büyük veriyi bilgiye dönüştürme çabaları her geçen gün daha da artmakta ve stratejik öneme sahip bir konu haline gelmektedir. Veri madenciliği çalışmaları, ekonomiden siyasete günlük yaşantının her alanını etkilemektedir. Bu açılarından ele alındığında veri madenciliğinde kullanılan yöntemlerin önemi daha net anlaşılmaktadır.

Kümeleme konusu, veri madenciliğinin en önemli çalışma alanlarından biri olmasının da ötesinde gözetimsiz (denetimsiz/danışmansız) çalışan yapay zekâ tabanlı uygulamaları geliştirmenin de başlıca yoludur. Günümüzde verinin akışındaki süreklilik ve dinamik yapısı dikkate alındığında gözetimsiz öğrenmenin ve nesnelere bu yolla gruplandırmanın önemi de anlaşılmaktadır. Sektörde akan veri olarak da adlandırılan dinamik veri bankacılıktan, elektronik ticarete, sosyal medya uygulamalarından endüstriyel otomasyon sistemlerine kadar birçok alanda kümelenmeye çalışılmaktadır. Kümeleme sürecinde en çok ihtiyaç duyulan yöntem ise nesnelere benzerliklerine göre sınırlı ve belirli bir sayıda gruba ayrılmasıdır. Bu amaçla en sık kullanılan algoritmaların başında k-ortalamlar tekniği gelmektedir. k-ortalamlar algoritmasının kolay anlaşılabilir ve basit uygulama adımlarına sahip olması rakiplerine karşın üstünlük kurmasını sağlarken, veri sayısındaki ve verinin karmaşıklık düzeyindeki artışın algoritmanın performansını olumsuz yönde etkilemektedir. Bu handikapları ortadan kaldırmak ve kümeleme başarısını artırmak amacıyla k-ortalamlar tekniğinin farklı uzaklık metrikleri,

çeşitli ağırlıklandırma yöntemleri ve çeşitli teknikler ile birlikte kullanılarak melezleştirildiği görülmektedir.

Bu proje çalışmasının amacı k-ortalamlar yönteminin kümeleme performansını artırmak ve kararlı hale getirmektir. Özellikle, veri sayısındaki ve verinin karmaşıklık düzeyindeki artışa bağlı olarak ortaya çıkan performans düşüşünü azaltmak ve genel olarak algoritmanın daha başarılı bir kümeleme hassasiyeti göstermesini sağlamak amaçlanmaktadır. Bu amaçla, k-ortalamlar algoritmasında halihazırda kullanılmakta olan nesnelere uzaklık esaslı olarak (küme merkezlerine olan uzaklıklarına bağlı olarak) gruplandırma yöntemi terk edilmektedir. Bunun yerine proje çalışmasında, meta-sezgisel arama algoritması ile nesnelere gruplara atanması önerilmektedir. Meta-sezgisel arama (MSA) algoritmalarının en önemli özellikleri çok boyutlu ve karmaşık arama uzaylarında geleneksel matematik yöntemlerine kıyasla çok daha üstün bir performans sergilemeleridir. MSA algoritmalarının bu özelliğinden faydalanılarak, k-ortalamlar yönteminin çok boyutlu ve karmaşık arama uzaylarındaki handikapları ve performans kararsızlığı da ortadan kaldırılmaya ya da azaltılmaya çalışılacaktır. Geliştirilecek uygulama web tabanlı bir platform üzerinde çalışacaktır. Uygulamaya ait web sayfası hazırlanacak ve araştırmacıların erişimine açık olacaktır.

**Anahtar Kelimeler:** k-Ortalamlar Kümeleme Algoritması, SOS Meta- Sezgisel Arama Algoritması, Melez Kümeleme