



Extension to manifold learning methods via advanced regression methods

Gülşen Taşkın*

Institute of Disaster Management, Istanbul Technical University, Istanbul, 34469, Turkey

Highlights:

- Generalization of manifold learning methods
- Solving out-of-sample problem using regression methods
- Classification of hyperspectral datasets

Keywords:

- Manifold learning
- Regression methods
- Out-of-sample problem
- Hyperspectral image classification

Article Info:

Research Article
Received: 19.03.2020
Accepted: 28.06.2021

DOI:

10.17341/gazimmfd.704793

Acknowledgement:

This work is conducted under the project (MGA-2018-41152) supported by Istanbul Technical University.

Correspondence:

Author: Gülşen Taşkın
e-mail:
gulsen.taskin@itu.edu.tr
phone: +90 555 621 8177

Graphical/Tabular Abstract

Recently, nonlinear dimensionality reduction, also called manifold learning methods, has been studied in the context of classification tasks. Manifold learning methods are graph-based methods, and they assume there is a nonlinear manifold with low dimensional space, that is hidden in the high dimensional data. They aim at preserving the neighborhoods between the samples lying in the high-dimensional space when projecting the data. Most of the manifold learning methods embed the training data into low dimensional space, but neither provide any projection matrix or an embedding function representing the nonlinear transformation. Due to this, the test data cannot be mapped to the same low dimensional space. To map the test data, the entire data, including the test and the previous training data, are given to the manifold learning method, then the learning process is rerun. However, it should be noted that this process needs to be repeated in each of the cases when the test data become available, resulting in a very high computational burden process. Therefore, especially for the classification tasks, one needs to have a generalized solution of manifold learning methods to transform the test data into previously learned low-dimensional space. In this study, advanced regression methods are utilized to solve this problem, also known as the out-of-the-sample problem in the literature. The embedding function, representing the transformation between the high and the low dimensional space, is provided by modeling the corresponding manifold learning method via regression methods. The quality of each embedding is evaluated based on the classification of hyperspectral images in detail.

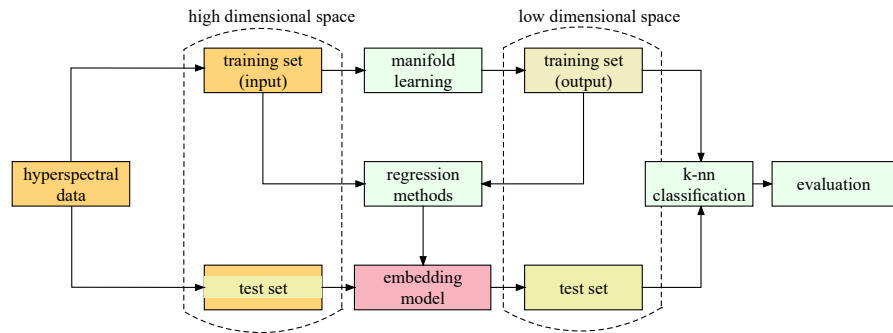


Figure A. Flow chart of the proposed method

Purpose: This study explores the effectiveness of the advanced regression methods in the out-of-sample problem of manifold learning methods. The schema of the proposed method is given in Figure A.

Theory and Methods:

The proposed method is based on the input and output set obtained via a specific manifold learning method. The nonlinear projection corresponding to manifold learning is modeled by advanced regression methods for hyperspectral image classification.

Results:

Extreme learning machines and support vector regression are very effective methods for out-of-sample methods, and they outperform the Nyström approach.

Conclusion:

The embedding function, representing the transformation between the high and the low dimensional space, is provided by modeling the corresponding manifold learning method via regression methods. The quality of each embedding is evaluated based on the classification of the hyperspectral images in detail.



Manifold öğrenme yöntemlerinin ileri seviye regresyon yöntemleri ile genelleştirilmesi

Gülşen Taşkın*

İstanbul Teknik Üniversitesi, Afet Yönetimi Enstitüsü, Deprem Mühendisliği Bölümü, 34469, İstanbul, Türkiye

Ö N E Ç İ K A N L A R

- Manifold öğrenme yöntemlerinin genelleştirilmesi
- Regresyon yöntemleri ile örneklem dışı veri probleminin çözümü
- Hiperspektral verilerin sınıflandırılması

Makale Bilgileri

Araştırma Makalesi
Geliş: 19.03.2020
Kabul: 28.06.2021

DOI:

10.17341/gazimmfd.704793

Anahtar Kelimeler:

Manifold öğrenme,
regresyon yöntemleri,
örneklem dışı veri sorunu,
hiperspektral veri
sınıflandırma

ÖZ

Son zamanlarda, doğrusal olmayan boyut indirgeme yöntemleri, diğer bir adı ile Manifold Öğrenme (MÖ) yöntemleri ile sınıflandırma konusunda çalışmalar yapılmıştır. MÖ yöntemleri, yüksek boyutlu verinin içinde aslında daha az boyutlu bir uzayda doğrusal olmayan bir manifoldun yer aldığı varsayımı üzerine çizge tabanlı bir dönüşüm yapmaktadır. Yüksek boyutlu uzayda yer alan verinin daha az boyutlu uzaya dönüştürülmesi için, veriler arası komşuluk ilişkilerinin korunması hedeflenir. MÖ yöntemlerinin birçoğu, eğitim verisinin tamamını birden alt uzaya dönüştür ve dönüşüme ait herhangi bir dönüşüm matrisi ya da analitik yapısı belli bir gömüleme fonksiyonu üretmezler. Bu sebepten ötürü, sonradan gelebilecek test verilerinin aynı alt uzaya dönüşümleri yapılamaz. Dönüşümün yapılabilmesi için, test verileri, önceki eğitim verileri ile birlikte, ilgili manifold öğrenme yöntemine tekrardan verilir ve öğrenme işlemi yeniden başlatılır. Ancak, her yeni test verisi geldikçe bu durumun tekrarlanması gerekeceğinden, hesaplama maliyeti artabilir. Bu nedenle, özellikle de sınıflandırma amaçlı çalışmalar için, manifold öğrenme yöntemlerinin, yeni gelecek test verisini alt uzaya dönüştürecek genel çözümlerine gereksinim vardır. Bu çalışmada, literatürde, örneklem dışı veri problemi olarak bilinen bu sorunun üstesinden gelmek için ileri seviye regresyon yöntemleri kullanılmıştır. İlgili manifold öğrenme yöntemi, regresyon yöntemleri ile modellenerek, dönüşüme ait gömüleme fonksiyonları üretilmiştir. Geliştirilen modellerin performansları hiperspektral görüntülerin sınıflandırılması üzerinde ayrıntılı bir biçimde analiz edilmiştir.

Extension to manifold learning methods via advanced regression methods

H I G H L I G H T S

- Generalization of manifold learning methods
- Solving out-of-sample problem using regression methods
- Classification of hyperspectral datasets

Article Info

Research Article
Received: 19.03.2020
Accepted: 28.06.2021

DOI:

10.17341/gazimmfd.704793

Keywords:

Manifold learning,
regression methods,
out-of-sample problem,
hyperspectral image
classification

ABSTRACT

Recently, nonlinear dimensionality reduction, also called manifold learning methods, has been studied in the context of classification tasks. Manifold learning methods are graph-based methods, and they assume there is a nonlinear manifold with low dimensional space, that is hidden in the high dimensional data. They aim at preserving the neighborhoods between the samples lying in the high-dimensional space when projecting the data. Most of the manifold learning methods embed the training data into low dimensional space, but neither provide any projection matrix or an embedding function representing the nonlinear transformation. Due to this, the test data cannot be mapped to the same low dimensional space. To map the test data, the entire data, including the test and the previous training data, are given to the manifold learning method, then the learning process is rerun. However, it should be noted that this process needs to be repeated in each of the cases when the test data become available, resulting in a very high computational burden process. Therefore, especially for the classification tasks, one needs to have a generalized solution of manifold learning methods to transform the test data into previously learned low-dimensional space. In this study, advanced regression methods are utilized to solve this problem, also known as the out-of-the-sample problem in the literature. The embedding function, representing the transformation between the high and the low dimensional space, is provided by modeling the corresponding manifold learning method via regression methods. The quality of each embedding is evaluated based on the classification of hyperspectral images in detail.

1. GİRİŞ (INTRODUCTION)

Hiperspektral görüntüler, yüksek spektral çözünürlüklerde algılama yapabilmeye kapasiteleri sayesinde, benzer özelliklere sahip nesnelerin birbirlerinden ayırt edilmesi gibi sınıflandırma çalışmaları için oldukça zengin bilgi içermektedirler [1]. Hiperspektral görüntülerin sınıflandırılmasında denetimli bir sınıflandırıcı kullanılması durumunda, yüksek bir sınıflandırma başarımı elde edebilmek için, oldukça fazla sayıda eğitim verisine gereksinim olduğu bilinmektedir [2-4]. Eğitim verilerinin sınırlı sayıda olması halinde, literatürde “Hughes etkisi” olarak bilinen yüksek boyutluluk sorunu (curse of dimensionality) ortaya çıkmaktadır [5]. Bunun sonucu olarak, sınıflandırma başarımları düşmektedir [6]. Bu sorunun üstesinden gelmek için, boyut azaltma yöntemleri ile, yüksek boyutlu veri, birtakım ölçütler ile, özü en iyi biçimde korunacak şekilde daha az boyutlu bir uzayda ifade edilmektedir. Boyut azaltma yöntemleri, literatürde öznitelik seçme ve öznitelik çıkarma yöntemleri olmak üzere iki grup altında incelenmektedirler. Bunlardan birincisinde, mevcut verinin spektral özniteliklerinden en iyi olanları belirlenirken, ikincisinde ise mevcut veri daha az boyutlu bir uzayda yeniden ifade edilmektedir [7, 8].

En çok bilinen öznitelik çıkarma yöntemlerinden biri olan temel bileşenler analizi, farklı disiplinlerdeki birçok problemin çözümünde kullanılmış ve halen kullanılmakta olan etkin doğrusal bir boyut azaltma yöntemidir. Ancak, performansları doğrusal yapıya sahip olmayan verilerde sınırlıdır. Bu tür durumlar için, doğrusal olmayan yapının modellendiği ve eniyileme süreçlerine dahil edildiği, manifold öğrenme (MÖ) olarak isimlendirilen çok sayıda doğrusal olmayan öznitelik çıkarma yöntemi geliştirilmiştir [9, 10]. MÖ yöntemleri, temelde, yüksek boyutlu uzayda yer alan verinin, saklı olduğu daha az boyutlu bir uzayın var olduğu varsayımını kullanarak, veriler arasındaki lokal ya da global konum ilişkisini çizgeler (graph) üzerinden modelleyerek bir eniyileme problemi çözmektedirler [11]. Bu problemin çözümü, genelleştirilmiş özdeğer denklemine karşılık gelmektedir. Yüksek boyutlu uzayda yer alan verilerin daha az boyutlu uzaydaki koordinatları ise, bu özdeğer denkleminin en büyük ya da en küçük özdeğerlerine karşılık gelen özvektörlerinden oluşmaktadır. MÖ yöntemleri özdeğer tabanlı yöntemler olduklarından spektral yöntemler olarak da isimlendirilmektedirler. “Laplacian eigenmaps” (LE) [12], “Locally linear embedding” (LLE) [9], ISOMAP [13], “Hessian eigenmaps” (HLE) [14] ve “Local tangent space alignment” (LTSA) [15] literatürde en yaygın olarak bilinen yöntemlerden bazılarıdır. MÖ yöntemleri, katışım (unmixing) problemleri [16], veri kaynaştırma (image fusion) [17], alan uyarlama (domain adaptation) [18] ve sınıflandırma [3] gibi farklı hiperspektral uygulama alanlarında da başarı ile uygulanmıştır. MÖ yöntemlerinin en büyük eksikliği, yüksek boyutlu uzaydan daha az boyutlu uzaya dönüşümü ifade eden bir dönüşüm matrisi veya gömüleme fonksiyonu (embedding function) üretmemeleridir. Bunun sonucu olarak, özellikle de

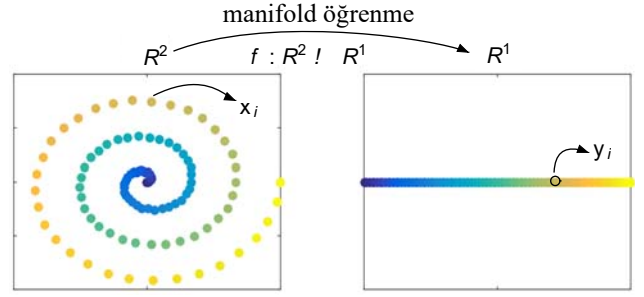
sınıflandırma amaçlı çalışmalarda, sonradan gelebilecek örneklem dışı verilerin (out-of-sample) az boyutlu uzaya dönüşümleri yapılamaz [19]. Bu gibi durumlarda, örneklem dışı veri mevcut eğitim verisi ile birleştirilerek, özdeğer problemi yeniden çözülür ve bu sorunun üstesinden gelinebilir. Ancak, sonradan gelebilecek her bir örneklem dışı veri için bu işlemin tekrar tekrar yapılması gerektiğinden, bu çözüm yolu hem pahalı hem de etkin olmayan bir yöntem olacaktır [20]. Literatürde örneklem dışı veri sorunu olarak bilinen bu durumun üstesinden gelmek için geliştirilmiş çok farklı yöntemler vardır. Bu yöntemler genel olarak iki farklı yaklaşımla incelenmiştir. Örneklem dışı uzantılar (out-of-sample extensions) temelli yaklaşımlarda, dönüşüm, MÖ yöntemlerinin eniyileme probleminde doğrusal bir fonksiyon ile modellenmekte ve bu sayede yüksek boyutlu uzaydaki verilerin daha az boyutlu uzaydaki koordinatları ve buna ek olarak da dönüşümü ifade eden dönüşüm matrisi elde edilmektedir. Bu yaklaşımlarda, yüksek boyutlu uzay ve daha az boyutlu uzay arasındaki ilişki doğrusal olarak kabul edilmektedir. Bu yöntemlere örnek yöntemler, LE yönteminin doğrusal uzantısı olan konum koruyan dönüşüm (locality preserving mapping) [21] ve LLE yönteminin doğrusal uzantısı olan komşuluk koruyan dönüşüm (neighborhood preserving embedding) [22] yöntemleridir. Bu yöntemlerdeki doğrusallık varsayımı, veri yapısının doğrusal olmayan bir yapıda olması halinde, manifold yapısının bozulmasına yol açabilir.

Yüksek boyutlu uzay ve daha az boyutlu uzay arasındaki ilişki, girdi-çıkı (input-output) problemi olarak düşünülebilir [23]. Eğitim verileri ve onların az boyutlu uzaydaki karşılık gelen koordinatları sırasıyla girdi ve çıktı verileri olacak şekilde herhangi bir regresyon yöntemi ile modellenilebilir. Bu durumda, az boyutlu uzayın boyut sayısı kadar çıktı veri seti olacağından, her bir boyut için ayrı ayrı regresyon modellerinin yapılması gerekmektedir. Bu çerçevede, MÖ yöntemlerinin regresyon yöntemleri ile nasıl bir arada kullanılabileceği [24] çalışmasında gösterilmiş ve “Spectral Regression” (SR) isimli bir yaklaşım önerilmiştir. Buradan esinlenerek, SR yöntemi, aşırı öğrenme (extreme learning) uzayında, dönüşüm fonksiyonu üretecek biçimde yeniden tasarlanmıştır [25]. Bir diğer çalışmada, veriler arasındaki korelasyon dikkate alınarak, Gauss süreç regresyonu (Gaussian process regression) yöntemi, ilk kez olarak örneklem dışı problemin çözümü için kullanılmış ve yöntemin Nyström yöntemi ilgili matematiksel ilişkisi gösterilmiştir [26]. Eğitim verilerinin sınırlı sayıda ve yüksek korelasyona sahip olması halinde, daha kararlı sonuçlar elde edebilmek için regresyon yöntemleri düzenlenilerek (regularization) daha kararlı sonuçlar elde edilmiştir. Hang ve diğerlerinin çalışmasında, uzaktan algılama verilerinin doğrusal olmayan yapıları çekirdek fonksiyonları ile ifade edilmiş ve örneklem dışı veri sorununun çözümü için yarı denetimli çizge tabanlı regresyon yöntemi kullanılmıştır [27]. Aynı motivasyonla, Tang ve diğerleri ise, çizge tabanlı bir boyut azaltma yöntemi geliştirmişlerdir [28]. Vural ve Guillemot ise denetimli manifold öğrenme yöntemlerinin sınıflandırma amaçlı

kullanımı için radyal tabanlı interpolasyon fonksiyonları ile genelleştirilmiş bir öğrenme modeli önermişlerdir [20]. MÖ yöntemlerindeki örneklem dışı veri sorunu, çok çıktılı (multi-output) bir girdi-çıkı problemi olması nedeni ile, çok çıktılı çekirdek tabanlı Tikhonov düzenleme regresyonu (kernel ridge regression) yöntemi bu amaçla kullanılmış ve performansı sınıflandırma problemi üzerinde detaylıca incelenmiştir [29]. Bir diğer çalışmada, alt uzaydaki verilerin kendi aralarındaki korelasyonu da modellenerek çok çıktılı genel bir regresyon modeli geliştirilmiş ve ilk kez olarak destek vektör makineleri ile uygulanmıştır [30]. Literatüre genel olarak bakıldığı zaman, örneklem dışı veri probleminin regresyon yöntemleri ile çözümlenmesinin birçok fayda getirdiği görülmektedir. Özellikle de manifold öğrenme yönetiminden bağımsız genel bir model önerebilme esnekliği, birçok uygulama alanında önemli faydalar sağlamaktadır. Literatürde, hiperspektral verilerin sınıflandırılması özelinde, MÖ yöntemlerinin ileri seviye regresyon yöntemleri ile genelleştirilmesini konu alan herhangi bir çalışma bulunmamaktadır. Bu çalışmanın amacı, MÖ yöntemlerinin önemli bir sorunu olan örneklem dışı veri probleminin regresyon yöntemleri ile modellenmesini konu almaktadır. Bu çalışmada, ileri seviye regresyon yöntemleri olarak, 5 farklı regresyon yöntemi kullanılmıştır. Bunlar, çekirdek tabanlı Tikhonov düzenleme regresyonu (kernel ridge regression) [31], destek vektör regresyonu (support vector regression) [32], Gauss süreç regresyonu (gaussian process regression) [33], ilişkili vektör makineleri (relevance vector machines) regresyonu [34] ve aşırı öğrenme regresyonu (extreme learning regression) [35] yöntemleridir. Hiperspektral veriler olarak, Botswana ve Indian Pines verileri kullanılmıştır. MÖ yöntemleri olarak, "Laplacian Eigenmaps (LE)" ve LLE manifold öğrenme yöntemleri için analizler yapılmıştır [36]. Regresyon yöntemlerinin performansı, sınıflandırma başarımları ve öğrenme süreleri açısından, standart yöntem olan Nyström yöntemi ile kıyaslanmıştır. Elde edilen sonuçlara göre, her bir veri setinde farklı regresyon yöntemleri başarılı olurken, genelde ELM yönteminin en iyi performansı verdiği gözlemlenmiştir. Bu çalışmada kullanılan tüm regresyon yöntemlerinin, Nyström yönteminden daha başarılı olduğu da gözlemlenmiştir.

2. ÖNERİLEN YAKLAŞIM (PROPOSED METHODS)

Yüksek boyutlu uzaydaki eğitim verileri $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ veri kümesi ile ifade edilsin. Burada m eğitim örneklerinin sayısına karşılık gelmektedir. Her bir eğitim verisinin $\mathbf{x}_i \in R^n$ boyutlu uzayda yer aldığını varsayalım. MÖ yöntemleri, eğitim verilerinin birbirleri ile olan konumsal ilişkisini bir çizge matrisi yardımı ile oluşturarak eniyileme probleminin dönüşür. Problemin analitik olarak çözülmesinin ardından genelleştirilmiş özdeğer problemi elde edilir. Bu özdeğer probleminin en büyük ya da en küçük özdeğerlerine karşılık gelen özvektörler kümesini $\mathbf{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_m\}$ ile ifade edelim. Bu durumda \mathbf{y}_i vektörü, \mathbf{x}_i örneğinin dönüştürülmüş alt uzaydaki koordinatlarına karşılık gelmektedir. Burada $\mathbf{y}_i \in R^{n'}$ ve $n' \ll n$ olmaktadır (Şekil 1).



Şekil 1. Manifold öğrenme yöntemi ile yapılan izdüşümün gösterilimi (Illustration of manifold learning embedding)

Önerilen yaklaşım ile \mathbf{x}_i ve \mathbf{y}_i örnekleri arasındaki ilişki, regresyon yöntemleri ile modellenerek n boyutlu uzaydan n' boyutlu uzaya dönüşümü sağlayan, gömüleme fonksiyonu (embedding function) diye de isimlendirilen $f: R^n \rightarrow R^{n'}$ fonksiyonunun belirlenmesi amaçlanmaktadır. Eğitim verisinin az boyutlu uzaydaki koordinatlarının belirlenebilmesi için herhangi bir MÖ yönteminin bir kez çalıştırılması gerekmektedir. Ardından, eğitim verilerinin her iki uzaydaki \mathbf{x}_i ve \mathbf{y}_i koordinatları ilgili regresyon yöntemine girdi ve çıktı verileri olarak verilir ve gömüleme fonksiyonu elde edilir. Burada dikkat edilmelidir ki, alt uzayın boyutu n' sayısı kadar gömüleme fonksiyonunun modellenmesi gerekmektedir. Ardından, eğitim veri seti içinde bulunmayan örneklem dışı veri ya da test verisi, ilgili gömüleme fonksiyonlarına verilerek, verilerin alt uzaydaki koordinatları belirlenir.

Regresyon yöntemleri ile belirlenen gömüleme fonksiyonlarının az boyutlu uzaya yaptıkları dönüşümlerin performanslarının belirlenmesi için k-nn sınıflandırma yöntemi ölçüt olarak kullanılmıştır. k-nn yöntemi, MÖ yöntemlerinin çizge matrislerinin oluşturulmasında kullanıldığı için özellikle tercih edilmiştir. Önerilen yaklaşımın işlem adımları Şekil 2'de gösterilmektedir.

Şekil 2'de belirtildiği üzere, hiperspektral görüntünün yer gerçekliği (ground truth) verisinden rastgele olarak eğitim ve test verisi oluşturulmaktadır. Eğitim verilerinin (girdi) ilgili manifold öğrenme yöntemi ile daha az boyutlu uzaya olan dönüşümleri yapılır. Bu sayede yüksek boyutlu uzaydaki her bir verinin daha az boyutlu uzaydaki koordinatları (çıkı) elde edilir. Elde edilen her bir girdi-çıkı çiftinin arasındaki ilişki regresyon yöntemleri ile modellenerek, örneklem dışı verilerin daha az boyutlu uzaya indirgeyecek gömüleme modelleri elde edilir. Regresyon yöntemleri tarafından üretilen gömüleme modellerinin performansları, alt uzaya izdüşürülen verilerin k-nn yöntemi ile sınıflandırma başarımları üzerinden değerlendirilir. Önerilen yaklaşıma ait sözde kod aşağıda verilmektedir:

Algoritma: Örneklem dışı verilerin regresyon yöntemleri ile alt uzaya indirgenmesine ait sözde kod.

Girdi: $\mathbf{x}_i \in R^n$, eğitim verileri.

1) Herhangi bir manifold öğrenme yöntemi ile eğitim verilerini daha az boyutlu uzaya izdüşür ve

her bir örneğin $y_i \in R^{n'}$ koordinatlarını belirle.

2) **for** her bir $y_i \in R^{n'}$ **do**
Alt uzayın her bir boyutu için regresyon modelini kur ve $f(x)$ dönüşüm fonksiyonunu belirle.

end

3) Örneklem dışı verileri, \bar{x} , $f(x)$ ile alt uzaya indirge ve y koordinatlarını hesapla

4) Alt uzaya indirgenen örneklem dışı verileri k-nn yöntemi ile sınıflandır ve sınıflandırma başarımlar oranını hesapla.

Çıktı : $f(x) : R^n \rightarrow R^{n'}$, dönüşüm fonksiyonu.

3. REGRESYON YÖNTEMLERİ (REGRESSION METHODS)

3.1. Çekirdek Tabanlı Tikhonov Düzenleştirilmesi Regresyonu (Kernel Ridge Regression)

Çekirdek Tabanlı Tikhonov düzenleştirilmesi regresyonu, doğrusal regresyon yönteminin çekirdek fonksiyonları yardımı ile genelleştiren bir yöntemdir. Yöntemin üreteceği model fonksiyonu aşağıdaki amaç fonksiyonun en küçüklenmesi ile elde edilmektedir:

$$J(w) = \sum (y_i - w^T x_i)^2 + \lambda \|w\|^2 \quad (1)$$

Burada λ düzenleme terimine, w vektörü ise x_i örneklerinin bilinmeyen katsayılarına karşılık gelir. Eş. 1'in parçalı türevlerinin alınması ile elde edilen denklemin sıfıra eşitlenmesi sonucu aşağıda eşitliğe ulaşılır.

$$w = (X^T X + \lambda I_n)^{-1} X^T y \quad (2)$$

Burada X ve I_n matrisleri sırasıyla yüksek boyutlu uzaydaki tüm eğitim verilerini tutan matris ile $n \times n$ boyutlu birim matrise karşılık gelmektedir. Bilinmeyen katsayıların Eş. 2 ile bulunması halinde, örneklem dışı verinin (\bar{x}) az boyutlu uzaydaki koordinatları Eş. 3 ile belirlenebilir.

$$f(\bar{x}) = w^T \bar{x} \quad (3)$$

Eş. 1'de verilen amaç fonksiyonun birincil yol ile (primal) çözümü yerine, eşiz en iyi çözümünün (dual) yapılması halinde Eş. 4'deki gömüleme fonksiyonu elde edilir.

$$f(x, \bar{x}) = y^T (K + \lambda I_m)^{-1} x^T \bar{x} \quad (4)$$

Burada K çekirdek matrisi olarak ifade edilir ve Hilbert uzayında doğrusal olmayan Φ fonksiyonlarının iç çarpımlardan oluşmaktadır. Çekirdek fonksiyonu ise $K(x, \bar{x}) = \Phi(x)^T \Phi(\bar{x})$ ile ifade edilmektedir. Burada Φ fonksiyonu, radyal tabanlı fonksiyon ya da d . mertebeden bir polinom olarak seçilebilir.

Destek Vektör Regresyonu (Support vector regression)

Vapnik tarafından geliştirilmiş, Eş. 5'de verilen ϵ -insensitive yitim fonksiyonunu (loss function) en küçüklemeyi amaçlayan bir yöntemdir.

$$J_\epsilon(y_i, f(x_i, w)) = \begin{cases} 0 & \text{if } |y_i - f(x_i, w)| \leq \epsilon \\ |y_i - f(x_i, w)| & \text{otherwise} \end{cases} \quad (5)$$

Burada ϵ , yitim fonksiyonu parametresine karşılık gelmektedir. Gömüleme fonksiyonu aşağıdaki biçimde seçilmiştir.

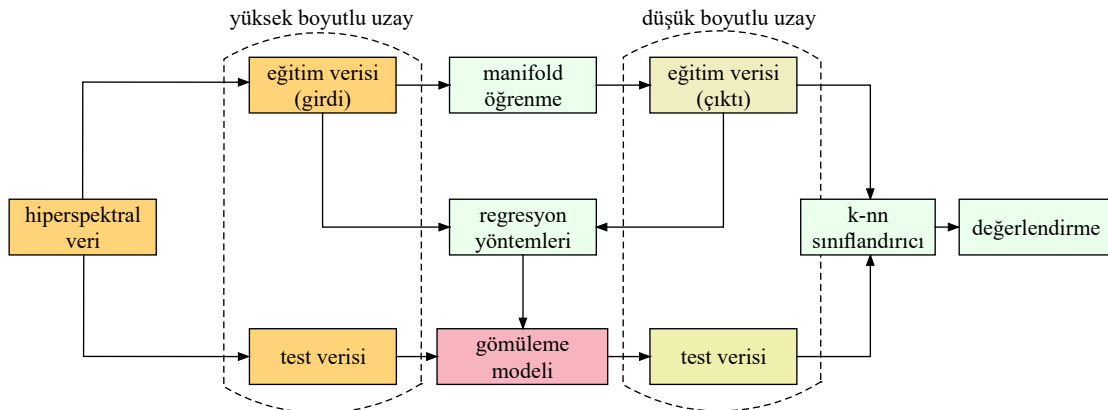
$$f(x, w) = w^T \Phi(x) + b \quad (6)$$

Eş. 6'nın bilinmeyen parametrelerinin belirlenmesi için, Eş. 7'de verilen eniyileme probleminin çözülmesi gerekmektedir.

$$\sum_{i=1}^n \max(|y_i - w^T \Phi(x) + b| - \epsilon, 0) \quad (7)$$

Eş. 7'de verilen kısıtsız eniyileme problemine düzenleştirme parametresi de dahil edilerek, çekirdek hesaplamalarına olanak tanınabilir. Böylece Eş. 7 denkleminin Eş. 8'de belirtilen kısıtlı eniyileme problemine dönüştürülebilir.

$$\begin{aligned} \min_{w, b, \xi} \quad & \frac{1}{2} w^T w + C \sum_{i=1}^n (\xi_i, \xi'_i) \\ \text{subject to} \quad & y_i - w^T \Phi(x_i) - b \leq \epsilon + \xi_i \\ & w^T \Phi(x_i) + b - y_i \leq \epsilon + \xi'_i \\ & \xi_i > 0, \xi'_i > 0, \quad i = 1 \end{aligned} \quad (8)$$



Şekil 2. Önerilen Yöntemin İşlem Adımları (Flow chart of the proposed method)

Lagrange yöntemi ile Eş. 8'in eşiz en iyi çözümü (dual solution) yapılarak gömüleme fonksiyonunun analitik yapısı aşağıdaki biçimde elde edilir:

$$f(x) = \sum_{i=1}^n (\alpha_i, \alpha_i^*) K(x, x_i) + b \quad (9)$$

Destek vektörleri sıfırdan farklı $\alpha_i, \alpha_i^* \in [0, C]$ katsayılarına karşılık gelen vektörlerdir. Eş. 9'daki $K(x, x_i)$ ise doğrusal olmayan fonksiyonlardan oluşturulmuş çekirdek matrisine karşılık gelir.

3.2. Gauss Süreç Regresyonu (Gaussian Process Regression)

Gauss bileşik (joint) dağılımına sahip sonlu sayıda rastgele değişkenlerin toplamına Gauss süreci denilmektedir. Eğitim verilerinden oluşan küme, $D = \{(x_i, f_i), i = 1, \dots, n\}$ olmak üzere, $f_i = f(x_i)$, x_i değerinin ilgili fonksiyondaki aldığı değere karşılık gelmektedir. Bu durumda, Gauss süreci

$$f(x) \sim GP(m(x_i), k(x_i, x_j)) \quad (10)$$

Eş. 10 ifadesi ile tanımlanabilir. Burada x ve $k(x_i, x_j)$ sırasıyla ortalamaya ve kovaryans fonksiyonuna karşılık gelmektedir. Kovaryans matrisinin elemanları, rastgele değişkenler arasındaki benzerliği ifade eder ve bu benzerlikler Eş. 11'deki üstel fonksiyon aracılığı ile belirlenebilmektedir.

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (11)$$

Burada σ parametresi çekirdek genişliği parametresidir ve eniyilenmesi gerekmektedir. Örneklem dışı X_* verisi için Gauss Süreç Regresyonu (GPR) yönteminin vereceği gömüleme fonksiyonu

$$p(f_* | X_*, X, f) = N(f_* | \mu_*, \sigma_*) \quad (12)$$

şeklinde olacaktır. Eş. 12'deki μ_* ve σ_*^2 değerleri, aşağıdaki Eş. 13'de verilen bağıntılar ile belirlenebilir.

$$\begin{aligned} \mu_* &= \underline{k}_*^T [K(X, X) + \sigma^2 I]^{-1} \underline{y} \\ \sigma_*^2 &= k(X_*, X_*) - k^T [K(X, X) + \sigma_n^2]^{-1} \underline{k}_* \end{aligned} \quad (13)$$

Burada k_* değişkeni, X eğitim verileri ve X^* örneklem dışı verisi arasındaki kovaryans değerlerini içeren vektöre karşılık gelmektedir. σ_n^2 ise gürültü varyansına karşılık gelmektedir. σ_x^2 değişkeni de girdi ve çıktı verileri arasında oluşturulan modelin güven (confidence) ölçütüne karşılık gelmektedir.

3.3. İlişkili Vektör Makineleri Regresyonu (Relevance Vector Machines Regression)

Destek vektör makinelerinin (DVM) birtakım kısıtlarını ortadan kaldırmak için önerilmiş, öğrenme sürecine Bayesçi yaklaşım uyarlanmış çekirdek tabanlı olasılıksal bir yöntemdir. İlişkili vektör makinelerinde (RVM), çekirdek

fonksiyonu herhangi bir fonksiyon olarak seçilebilmektedir. Ayrıca, DVM yöntemlerinde olduğu gibi herhangi bir parametre optimizasyonu gerektirmemektedir. RVM yönteminde, $\{x_i, y_i\}_{i=1}^m$ girdi-çıkı verilerinin normal dağılımda olduğu varsayılmaktadır. Buna göre, gömüleme fonksiyonunun olasılıksal gösterilimi Eş. 14'de verilmektedir.

$$p(y|\alpha, \sigma^2) = (2\pi)^{-\frac{n+1}{2}} |\mathbf{B}^{-1} + \Phi \mathbf{A}^{-1} + \Phi^T|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \mathbf{y}^T (\mathbf{B}^{-1} + \Phi \mathbf{A}^{-1} + \Phi^T)^{-1} \mathbf{y}\right\} \quad (14)$$

Burada $\mathbf{A} = \text{diag}(\alpha_0, \dots, \alpha_n)$ ve $\mathbf{B} = \sigma^{-2} \mathbf{I}_n$ şeklindedir. σ parametresi, veriden hesaplanabilen bir değer, α parametresi ise en iyilenmesi gereken parametre olmak üzere, bu değerler birtakım formüller yardımı ile analitik olarak belirlenebilmektedir. Φ matrisi "design" matrisi olarak isimlendirilmekte olup, $\Phi = K(x, x_i)$ çekirdek fonksiyonlarından oluşmaktadır.

3.4. Aşırı Öğrenmeli Makineler (Extreme Learning Machines)

Aşırı öğrenmeli makineler (ELM), ileri beslemeli yapay sinir ağları temelli regresyon yöntemidir. ELM yönteminin üreteceği gömüleme fonksiyonunun analitik yapısı Eş. 15'de verilmektedir.

$$f_L(x) = \sum_{i=1}^L \beta_i h_i(x) = \mathbf{h}(x) \boldsymbol{\beta} \quad (15)$$

Bu eşitlikte, $\boldsymbol{\beta} = [\beta_1, \dots, \beta_L]$ vektörü, L sayıdaki nöron ve çıktı düğümleri arasındaki ağırlık katsayılarına karşılık gelmektedir. $\mathbf{h}(x) = [h_1(x), \dots, h_L(x)]$ vektörü ise h_i sigmoit fonksiyonlarından (doğrusal olmayan öznitelik dönüşümleri) oluşmaktadır (Eş. 16).

$$h_i(x) = G(\mathbf{a}_i, b_i, x) = \frac{1}{1 + \exp(-\mathbf{a}_i x + b_i)} \quad (16)$$

Bilinmeyen ağırlık katsayıları $\boldsymbol{\beta}$, aşağıdaki doğrusal denklem takımının çözülmesi ile elde edilebilir.

$$\boldsymbol{\beta} = \left(\frac{L}{c} + \mathbf{H}^T \mathbf{H}\right) \mathbf{H}^T \mathbf{Y} \quad (17)$$

Eş. 17'deki \mathbf{I} matrisi, L boyutlu bir birim matristir. Saklı katmanlardaki düğüm noktalarının parametreleri (\mathbf{a}_i vektörü ve b_i) rastgele olarak üretilir. \mathbf{Y} matrisi ise çıktı verilerinin oluşturduğu matrise karşılık gelmektedir. Eş. 17'deki \mathbf{H} matrisinin yapısı aşağıdaki biçimdedir:

$$\mathbf{H} = \begin{bmatrix} h_1(x_1) & \dots & h_1(x_1) \\ \vdots & \ddots & \vdots \\ h_1(x_m) & \dots & h_1(x_m) \end{bmatrix} \quad (18)$$

Düğüm noktalarının sayısı L parametresi, genellikle öznitelik uzayının boyutu ile aynı seçilmektedir [35].

4. NYSTRÖM YAKLAŞIMI (NYSTRÖM APPROACH)

Nyström yaklaşımı, manifold öğrenme yöntemlerinin örneklem dışı veri problemi için özel olarak geliştirilmiş bir

yöntemdir [19]. Girdi ve çıktı verileri arasındaki ilişkiyi modellemek yerine, doğrudan ilgili MÖ yönteminin dönüşümünü esas alarak, \bar{x} örneklem dışı verisinin, daha az boyutlu uzaydaki koordinatlarını, hesaplanan özdeğer ve özvektörlerin ağırlıklı toplamlarını alarak belirler (Eş. 19).

$$y^T = \sum_{i=1}^n K(\bar{x}, x_i) y_i^T \Lambda^{-1} \quad (19)$$

Burada, $K(\cdot, \cdot)$ eğitim ve test verileri ile belirlenen çekirdek matrisine, Λ matrisi ise elemanları çekirdek matrisinin en büyük özdeğerlerinden oluşan $n' \times n'$ boyutlu bir matrise karşılık gelmektedir. Her bir manifold öğrenme yöntemi için farklı bir yaklaşım mevcuttur. Bunlara ilişkin ayrıntılar [19] çalışmasında verilmektedir.

5. DENEYSEL SONUÇLAR (EXPERIMENTAL RESULTS)

Regresyon yöntemlerinin örneklem dışı veri problemindeki performansları, hiperspektral verilerin sınıflandırma problemi üzerinden test edilmiştir. Buna göre, literatürde sıkça kullanılan “benchmark” verisi olarak kabul görmüş iki farklı hiperspektral veri kullanılmıştır. Bu veriler, Indian Pines ve Botswana verileridir.

Botswana veri seti, Mayıs 2001'de, NASA EO-1 uydusu tarafından Botswana'da bulunan Okavango Delta'sı üzerinde kaydedilmiştir. Verinin elde edilmesinde kullanılan sensör, 30 m. uzamsal çözünürlük ve 400-2500 nm. arasında 10 nm.'lik geçişlerde 242 bantlık bir hiperspektral çözünürlük sunmaktadır. Botswana üzerinde bu çözünürlüklerle elde edilen veri, Teksas Üniversitesi Uzay Araştırmaları Merkezi tarafından kalibre edilmiş ve su soğurmasını içeren bantlar gibi gürültülü bantların kaldırılmasıyla, analizlere uygun hale getirilmiştir. Yapılan bu işlemlerin ardından bant sayısı 145'e düşürülmüştür (silinen bantlar: 10-55, 82-97, 102-119, 134-164, 187-220). Botswana verisi içerisinde 14 farklı arazi sınıfına ait gözlemler etiketlenerek analizlere hazır hale getirilmiştir. Bu arazi örtülerinden, spektral yansıma değerleri birbirlerine en yakın olanları “riparian” ve “Acacia woodlands” isimli sınıflardır ve sınıflandırma açısından ayırt edilmesi en zor sınıflar olarak raporlanmıştır.

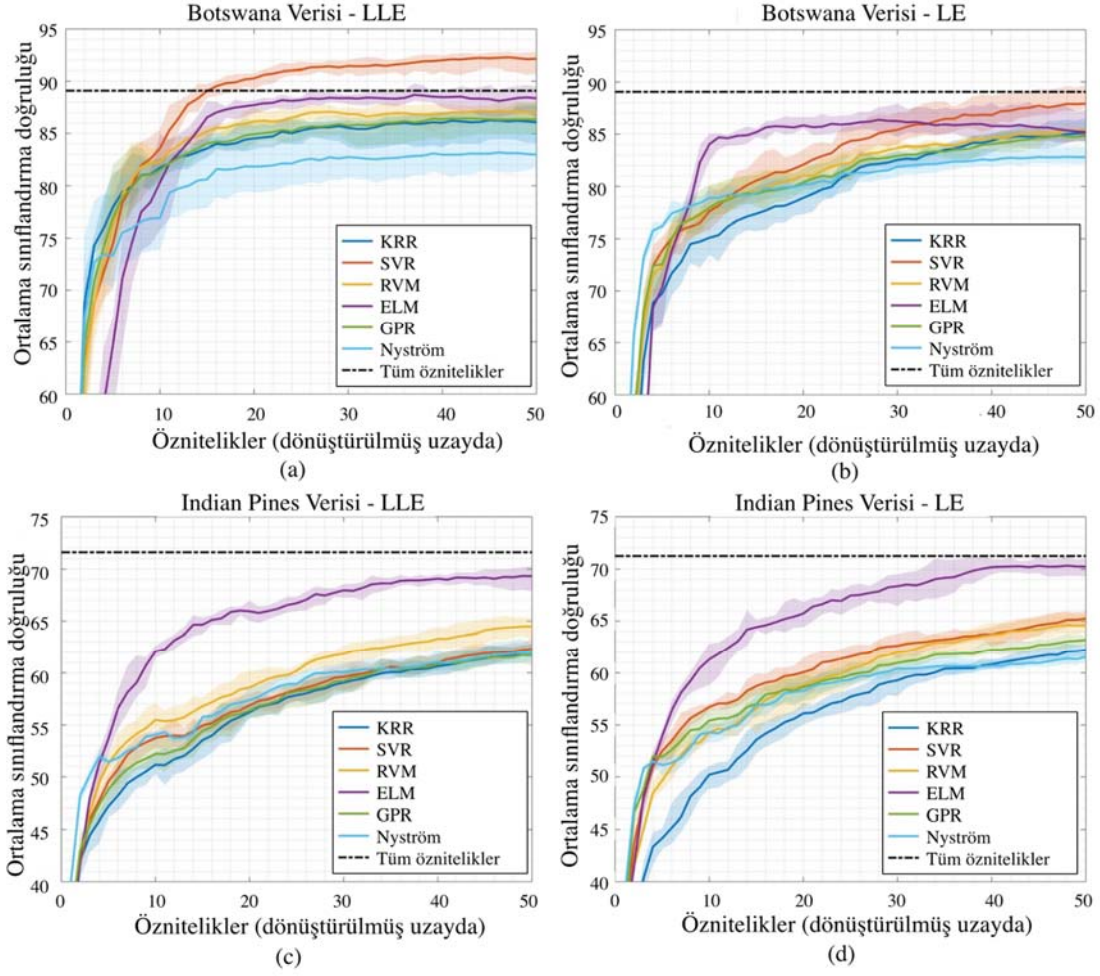
Indian Pines verisi, 1992 yılında AVIRIS sensörü tarafından elde edilmiş 20 m. uzamsal çözünürlüğe sahip bir veridir. Görüntünün üçte ikisi tarımsal alan, geriye kalan kısımları orman veya çeşitli bitki türlerinden oluşmaktadır. Görüntü, 145×145 sayıda pikselden ve 224 spektral banttan oluşmaktadır. Su emilim etkisi kaldırıldıktan sonra bant sayısı 200'e düşürülmüştür. Bu veri, toplamda 16 farklı arazi sınıfından oluşmaktadır. “Corn” ve “Soybean” bitkilerinin aynı zamanlarda ekimleri yapıldığı için spektral yansımaları birbirlerine çok benzemektedir. Bu nedenle bu iki sınıf, Indian Pines verisi için sınıflandırma ile ayırt edilmesi en zor olan sınıflardandır.

Bu çalışmada, regresyon yöntemlerinin performansları iki açıdan incelenmiştir: sınıflandırma başarımları ve hesaplama zamanı. Sınıflandırma performansının ölçülmesinde k-NN yöntemi özellikle tercih edilmiştir. Bunun nedeni, manifold öğrenme yöntemlerinin esas olarak

komşuluk tabanlı ölçütler üzerinden eniyileme yapmasından kaynaklanmaktadır. Hesaplama zamanlarının belirlenmesinde ise regresyon yöntemlerinin sadece öğrenme süreleri dikkate alınmıştır. Hiperspektral verilerin alt uzaya izdüşümleri için iki farklı manifold öğrenme yöntemi kullanılarak regresyon yöntemlerinin genelleştirme kapasiteleri de ayrıca incelenmiştir. MÖ yöntemi olarak “Laplacian Eigenmaps (LE)” ve “Linear Locality Embedding (LLE)” yöntemleri kullanılmıştır.

Deneysel analizlerde, LE ve LLE manifold öğrenme yöntemleri ile eğitim verisi alt uzaya izdüşürülmüştür. Her bir MÖ yöntemi için, yüksek boyutlu uzaydan daha az boyutlu uzaya olan dönüşüm, regresyon yöntemleri ile öğrenilmiş ve rastgele olarak oluşturulan test verisi (örneklem dışı veri), ilgili gömüleme fonksiyonları ile daha az boyutlu uzaya izdüşürülmüştür. Bu işlem rastgele olarak oluşturulan on farklı eğitim ve test verisi için tekrar edilmiştir. Verilerin % 20'si eğitim, geri kalanı test verisi olacak şekilde bölümlenmiştir. Manifold öğrenme yöntemlerinin daha az eğitim verisinde nasıl bir performans vereceğinin görülebilmesi için eğitim verisinin oranı özellikle daha az seçilmiştir. Her bir regresyon modeli ile yapılan izdüşümlerin doğruluğu, izdüşürülen test verilerinin k-en yakın komşuluk sınıflandırıcısının performansı üzerinden karşılaştırılmıştır. Test verileri üzerinde elde edilen sınıflandırma başarımlarının ortalama değerleri sunulmuştur. Elde edilen sonuçlar Şekil 3'de verilmektedir.

Buna göre, Şekil 3a'da SVR yöntemi en iyi performansı verirken, diğer durumlarda ELM yöntemi SVR yönteminden daha iyi bir performans vermektedir. Genel olarak bakıldığında, ELM yönteminin her iki veri setinde ve her iki MÖ yöntemi için en iyi sınıflandırma performansını ürettiği söylenebilir. Botswana veri setinde, tüm regresyon yöntemlerinin Nyström yönteminden daha iyi bir performans verdiği gözlemlenmiştir. Indian Pines verisinde, LLE dönüşümünde RVM yönteminin ve LE dönüşümünde ise SVR yönteminin ikinci en iyi performans veren yöntemler oldukları görülmektedir. ELM, RVM ve SVR hariç olmak üzere, Nyström yaklaşımının diğer yöntemlerden daha yüksek performans verdiği görülmektedir. Bu sonuçlara göre, regresyon yöntemlerinden özellikle ELM, SVR ve RVM yöntemlerinin, örneklem dışı veri problemi için Nyström yöntemine alternatif olabilecek yöntemler olduğu söylenebilir. Şekil 3a hariç olmak üzere, regresyon yöntemlerinin sınıflandırma performansının, tüm özneliklerin sınıflandırmada kullanılması durumundaki sınıflandırma performansını geçemediği görülmektedir. Bunun nedeni, MÖ yöntemlerinin dönüşümdeki başarımlarına ya da eğitim veri sayısının gerçek manifoldun çıkarılması için yetersiz olmasına bağlı olabilir. Bu çalışmada, regresyon yöntemlerinin örneklem dışı veri performanslarının karşılaştırılması hedeflendiğinden, MÖ yöntemleri varsayılan parametre değerleri ile uygulanmış, MÖ yöntemlerinin performanslarını en iyilenmesi amaçlanmamıştır [37]. Tablo 1-Tablo 4'de, her bir veri seti için, alt uzaydaki ilk 50 öznelik ile sınıflandırma yapılması halinde, her bir arazi sınıfı için elde edilen sınıflandırma



Şekil 3. Her bir regresyon yöntemi ile örneklem dışı verilerin dönüştürülmüş uzaydaki k-nn sınıflandırma doğruluklarına ait öğrenme eğrileri

(Learning curves of classification accuracies for the out-of-samples, achieved by k-nn in the embedding space obtained by each regression method)

Tablo 1. LLE yönteminin Botswana verisindeki sınıflandırma doğrulukları ve Kappa değerleri
(Classification accuracies and Kappa values of LLE for Botswana Dataset)

Sınıflar	Eğitim	Test	KRR	SVR	RVM	ELM	GPR	Nyström
Water	54	216	99,9	100,0	99,8	99,6	99,8	99,8
Hippo grass	21	80	96,6	96,8	96,1	95,4	97,4	94,7
Floodplain grasses 1	51	200	96,9	99,7	97,4	94,9	97,3	89,2
Floodplain grasses 2	43	172	86,2	95,3	88,9	91,3	87,0	79,1
Reed 1	54	215	73,6	82,5	74,5	85,1	73,7	73,7
Riparian	54	215	60,4	69,3	60,1	71,8	60,1	59,2
Firescar2	52	207	99,6	99,8	99,4	99,5	99,6	99,6
Island interior	41	162	92,4	97,9	93,1	97,1	92,6	90,2
Acacia woodlands	63	251	76,6	87,3	78,4	80,9	77,2	71,8
Acacia shrublands	50	198	77,5	91,1	81,2	82,6	78,0	71,3
Acacia grasslands	61	244	91,8	95,8	92,5	94,0	92,4	88,0
Short mopane	37	144	91,3	93,4	91,3	83,6	91,7	90,7
Mixed mopane	54	214	85,9	93,6	86,8	78,3	85,3	81,7
Exposed soils	19	76	99,2	100,0	98,9	100,0	98,9	99,0
Toplam	654	2594						
		Ortalama doğruluk	86,1	92,1	87,0	88,4	86,4	83,0
		Kappa	84,7	91,2	85,3	86,3	84,1	81,2

doğrulukları, eğitim ve test verilerinin sayıları ile birlikte verilmektedir. Her bir tabloda ayırt edilmesi en zor olan

sınıflara ait sınıflandırma başarımları koyu italik yazı tipi ile ifade edilmektedir.

Tablo 2. LE yönteminin Botswana verisindeki sınıflandırma doğrulukları ve Kappa değerleri
(Classification accuracies and Kappa values of LE for Botswana Dataset)

Sınıflar	Eğitim	Test	KRR	SVR	RVM	ELM	GPR	Nyström
Water	54	216	93,9	99,8	99,4	99,6	99,7	99,4
Hippo grass	21	80	92,8	95,2	92,8	94,9	93,2	86,4
Floodplain grasses 1	51	200	94,0	98,1	94,7	90,8	94,8	92,5
Floodplain grasses 2	43	172	87,5	88,5	87,0	86,0	86,5	85,0
Reed 1	54	215	78,1	81,2	77,1	84,8	78,3	73,9
Riparian	54	215	60,9	61,8	60,2	67,9	58,2	52,3
Firescar2	52	207	98,3	99,7	99,0	99,3	98,5	98,5
Island interior	41	162	93,1	95,3	93,0	92,6	91,2	85,5
Acacia woodlands	63	251	72,6	75,2	73,3	75,9	71,6	71,5
Acacia shrublands shrublands	50	198	75,0	84,0	80,6	71,2	80,8	75,7
Acacia grasslands	61	244	88,5	92,9	85,2	91,9	87,3	91,2
Short mopane	37	144	87,4	90,1	89,1	80,8	88,6	90,6
Mixed mopane	54	214	85,3	91,2	85,3	78,4	82,4	78,0
Exposed soils	19	76	93,9	99,8	99,4	99,6	99,7	99,4
Toplam	654	2594						
		Ortalama doğruluk	85,2	87,9	85,4	85,2	84,9	82,8
		Kappa	83,3	86,2	82,3	82,2	81,7	80,4

Tablo 3. LLE yönteminin Indian Pines verisindeki sınıflandırma doğrulukları ve Kappa değerleri
(Classification accuracies and Kappa values of LLE for Indian Pines Dataset)

Sınıflar	Eğitim	Test	KRR	SVR	RVM	ELM	GPR	Nyström
Alfalfa	10	36	31,7	32,7	34,9	57,7	31,1	24,7
Corn-no till	286	1142	48,1	51,5	50,3	58,1	48,8	47,1
Corn-min till	166	664	51,4	54,7	52,4	58,3	51,3	51,9
Bldgs-Grass-TreeDrives	48	189	28,6	35,6	31,7	45,9	28,5	29,8
Grass-pasture	97	386	77,9	81,7	81,1	87,4	79,8	77,1
Grass-trees	146	584	74,9	78,2	79,2	83,6	76,1	74,3
Grass-pasture-mowed	6	22	50,0	52,1	57,4	80,6	54,7	29,6
Corn	96	382	87,4	90,2	89,3	92,9	88,4	86,0
Oats	4	16	11,7	20,6	16,7	60,7	12,5	21,7
Soybeans-no till	195	777	54,3	57,2	57,9	60,1	56,1	53,4
Soybeans-no till	491	1964	63,8	66,2	66,5	66,7	65,5	63,9
Soybeans-clean till	119	474	34,5	39,1	39,3	55,6	35,6	33,5
Wheat	41	164	78,1	80,9	81,0	90,0	79,0	78,3
Woods	253	1012	88,8	89,4	89,8	91,2	89,3	88,9
Hay-windrowed	78	308	32,3	39,3	36,6	55,8	33,8	30,8
Stone-stell-towers	19	74	95,0	97,0	98,0	97,6	95,1	95,5
Toplam	2055	8194						
		Ortalama doğruluk	62,2	65,2	64,6	70,2	63,2	61,5
		Kappa	60,4	62,1	62,3	68,1	61,1	59,8

Tablo 4. LE yönteminin Indian Pines verisindeki sınıflandırma doğrulukları ve Kappa değerleri
(Classification accuracies and Kappa values of LE for Indian Pines Dataset)

Sınıflar	Eğitim	Test	KRR	SVR	RVM	ELM	GPR	Nyström
Alfalfa	10	36	23,0	21,2	24,7	61,3	21,2	23,3
Corn-no till	286	1142	47,1	47,7	50,2	58,4	47,2	48,5
Corn-min till	166	664	48,3	49,7	53,0	56,3	47,8	50,8
Bldgs-Grass-TreeDrives	48	189	33,4	34,9	33,1	45,4	33,0	35,7
Grass-pasture	97	386	80,8	81,8	81,2	89,0	81,0	77,3
Grass-trees	146	584	75,7	76,7	77,6	84,3	76,3	77,1
Grass-pasture-mowed	6	22	83,2	85,6	82,2	87,1	83,9	80,6
Corn	96	382	90,8	90,8	91,2	93,1	90,3	90,7
Oats	4	16	32,7	31,9	33,4	47,0	28,3	15,7
Soybeans-no till	195	777	51,1	50,9	56,3	56,1	51,3	47,5
Soybeans-no till	491	1964	63,3	63,5	65,7	65,4	62,7	62,9
Soybeans-clean till	119	474	31,8	32,7	38,3	52,8	32,2	31,0
Wheat	41	164	83,9	85,9	84,4	93,2	84,4	83,5
Woods	253	1012	89,2	89,6	89,9	91,3	89,2	89,2
Hay-windrowed	78	308	34,5	36,4	37,5	55,0	35,1	36,2
Stone-stell-towers	19	74	97,5	98,2	97,1	98,4	97,4	95,8
Toplam	2055	8194						
		Ortalama doğruluk	61,8	62,3	64,5	69,3	61,7	61,9
		Kappa	59,6	60,2	62,4	67,2	59,6	60,1

Tablo 5. Regresyon yöntemlerinin öğrenme süreleri (Elapsed computing time for Training of regression methods)

Veri	MÖ Yöntemi	KRR (sn.)	SVR (sn.)	RVM (saat)	ELM(dk.)	GPR (dk.)	Nyström(sn.)
Botswana	LE	3,8	20,6	2,5	1,8	3,1	7,5
	LLE	2,9	18,9	2,0	1,4	2,3	1,7
Indian Pines	LE	19,9	84,9	6,0	39,9	33,8	36,2
	LLE	32,5	209,3	10,3	51,0	41,8	10,4

Tablo 1'e göre, ortalama sınıflandırma doğruluğu açısından, SVR'in en yüksek performansı vermesine rağmen, "Riparian" sınıfı için ELM yöntemi en yüksek sınıflandırma doğruluğunu vermektedir (%71.8). Tablo 2'ye göre, aynı hiperspektral verinin LE yöntemi ile izdüşürülmesi halinde, yine ELM yöntemi "Riparian" sınıfı için en yüksek performansı vermektedir (%67.9). "Acacia woodlands" sınıfı için ise, LLE ile SVR yöntemi en yüksek performansı verirken, LE ile SVR ve ELM yöntemleri birbirlerine çok yakın performanslar vermektedir.

Tablo 3'de, "Corn" ve "Soybean" sınıfları için ELM en yüksek performansı vermektedir. "Oat" sınıfına ait başarı oranının çok düşük olmasının sebebi, bu sınıf için eğitim veri sayısının çok az olmasından ötürüdür. Tablo 4'de ise, "Corn" sınıfı için en yüksek performansı ELM yöntemi vermektedir. "Soybean" sınıfları için ("Soybeans-no till" ve "Soybeans-no till") ELM ve RVM yöntemleri birbirlerine yakın performanslar üretirken, "Soybeans-clean till" için ELM yönteminin en yüksek performansı verdiği gözlemlenmektedir. Ayrıca, eğitim örneklerinin dağılımına bakıldığında, "Alfalfa", "Grass-pasture-mowed" ve "Oat" isimli sınıfların örnek sayılarının, diğer sınıflara göre çok daha az olduğu görülmektedir. Bu sınıflar için, ELM yönteminin genelleştirme performansının diğer regresyon yöntemlerinininkine oranla daha iyi olduğu görülmektedir.

Elde edilen tüm sonuçlara göre, regresyon yöntemlerinin sınıflandırma performansları incelendiğinde, ELM, SVR ve RVM yöntemlerinin öne çıkan yöntemler olduğu görülmektedir. Tüm yöntemlerin, yüksek boyutlu uzay ile az boyutlu uzay arasındaki dönüşümü öğrenme süreleri ise Tablo 5'de verilmektedir. Eğitim süreleri kullanılan veriye göre değişim göstermekle birlikte, genel olarak en yavaş yöntem RVM, en hızlı yöntem de Nyström ve KRR yöntemleridir. ELM yönteminin SVR yöntemine göre daha yavaş bir yöntem olduğu görülmektedir. Eğitim süreleri ve sınıflandırma performansları birlikte düşünüldüğünde, SVR ve ELM yöntemlerinin örneklem dışı veri sorununun çözümü için en etkin yöntemler olduğu sonucu çıkarılabilir.

6. SONUÇLAR (CONCLUSIONS)

Manifold öğrenme (MÖ) yöntemleri, çizge tabanlı yaklaşımlar ile doğrusal olmayan verilerin yüksek boyutlu uzaydan daha az boyutlu uzaya izdüşümlerini konu almaktadırlar. MÖ yöntemleri, eğitim verilerini daha az boyutlu uzaya indirgerler, ancak dönüşüme ait bir dönüşüm matrisi ya da gömüleme fonksiyonu üretmezler. Bunun sonucu olarak da test verisi şeklinde sonradan gelebilecek örneklem dışı veriler az boyutlu uzaya indirgenemezler. Bu çalışmada, ileri seviye regresyon yöntemleri ile MÖ yöntemlerinin öğrenme süreçleri modellenerek, gömüleme

fonksiyonlarının üretilmesi amaçlanmaktadır. Buna göre, beş farklı regresyon yöntemi kullanılmıştır: Çekirdek tabanlı Tikhonov düzenleme regresyonu (KRR), destek vektör regresyonu (SVR), Gauss süreç regresyonu (GPR), ilişkili vektör makineleri (RVM) regresyonu ve aşırı öğrenmeli makineler (ELM) regresyonu. Bu yöntemlerin performansları, MÖ yöntemlerinin örneklem dışı veri sorunu için sıkça kullanılan bir yöntem olan Nyström yöntemi ile hiperspektral veriler üzerinde hem sınıflandırma performansı hem de öğrenme süreleri açısından karşılaştırılmıştır. Hiperspektral veriler olarak Botswana ve Indian Pines verileri, sınıflandırma yöntemi olarak ise k-NN yöntemi kullanılmıştır. Elde edilen sonuçlara göre, çalışmada kullanılan tüm regresyon yöntemlerinin Nyström yönteminden genelde daha iyi performans verdiği, ELM yönteminin ise regresyon yöntemleri içinde en iyi sınıflandırma performansı ürettiği gözlemlenmiştir. Öğrenme süreleri açısından ELM yönteminin, KRR, SVR ve Nyström yöntemlerine göre daha yavaş olmasına rağmen, kabul edilebilir sürelerde sonuç ürettiği gözlemlenmiştir. ELM yönteminden sonra SVR yönteminin de hem öğrenme süresi hem de performans açısından etkin bir yöntem olduğu sonucuna varılmıştır.

TEŞEKKÜRLER (ACKNOWLEDGEMENT)

Bu çalışma, İstanbul Teknik Üniversitesi 41152 numaralı BAP Genel Araştırma Projesi kapsamında desteklenmiştir. Bu nedenle, İstanbul Teknik Üniversitesi'ne desteklerinden ötürü teşekkür ederim.

KAYNAKLAR (REFERENCES)

1. Celik C., Bilge H., Feature selection with weighted conditional mutual information, Journal of the Faculty of Engineering and Architecture of Gazi University, 30 (4), 585-596, 2015.
2. Hanbay K., Hyperspectral image classification using convolutional neural network and two-dimensional complex Gabor transform, Journal of the Faculty of Engineering and Architecture of Gazi University, 35 (1), 443-456, 2020.
3. Lungu, D., Prasad, S., Crawford, M. M., Ersoy, O., and Data, H., A review of advances in manifold learning, IEEE Signal Processing Magazine, 2014, 55-66, 2014.
4. Camps-Valls, G., Tuia, D., Gomez-Chova, L., Jimenez, S., and Malo, J., Remote Sensing Image Processing, Morgan & Claypool Publishers series, 2012.
5. Bellman, R., Adaptive Control Processes: A Guided Tour. Princeton, NJ: Princeton Univ. Press, 1961.
6. Bishop, C. M., Neural networks for pattern recognition, Oxford University Press, 1995.

7. D. Xiao and J. Zhang, Feature selection: Evaluation, application, and small sample performance, *Pattern Analysis and Machine Intelligence*, 19 (2), 197–201, 2009.
8. Jia, X., Kuo, B.-C., and Crawford, M. M., Feature Mining for Hyperspectral Image Classification, *Proc. IEEE*, 101 (3), 2013.
9. Roweis, S. T., and Saul, L. K., Nonlinear dimensionality reduction by locally linear embedding, *Science*, 290 (5500), 2323–2326, 2000.
10. Saul L. K. and Roweis, S. T., Think globally, fit locally: Unsupervised learning of low dimensional manifold, *J. Mach. Learn. Res.*, 4, 119–155, 2003.
11. Yan, S., Xu, D., Zhang, B., Zhang, H. J., Yang, Q., and S. Lin, Graph embedding and extensions: A general framework for dimensionality reduction, *IEEE Trans. Pattern Anal. Mach. Intell.*, 29 (1), 40–51, 2007.
12. Belkin M, Niyogi P., Laplacian eigenmaps and spectral techniques for embedding and clustering. MIT Press, In: *Advances in neural information processing systems*, 14 (14), 585–591, 2001.
13. Tenenbaum, J., Silva, V. de., and Langford, L., A global geometric framework for nonlinear dimensionality reduction, *Science*, 290 (5500), 2319–2323, 2000.
14. Donoho D. and Grimes C., Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data, *Proc Nat Acad Sci USA*, 100 (10), 5591–5596, 2003.
15. Zhang Z. and Zha H., Principal manifolds and nonlinear dimension reduction via local tangent space alignment, *SIAM J Scientif Comput.*, 26 (1), 313–338, 2005.
16. Zhu, F., Wang, Y., Xiang, S., Fan, B., and Pan, C., Structured Sparse Method for Hyperspectral Unmixing, *ISPRS Journal of Photogrammetry and Remote Sensing*, 88, 101-118, 2014.
17. Danping, L., Yuntao, Q. Jun Z. et. al., A Manifold Alignment Approach for Hyperspectral Image Visualization with Natural Color, *IEEE Transactions on Geoscience and Remote Sensing*, 54 (6), 3151-3162, 2016.
18. Hsiuhan L. Y. and Crawford, M. M., Domain Adaptation with Preservation of Manifold for Hyperspectral Image Classification, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9 (2), 543-555, 2016.
19. Bengio, Y., Paiement, J., Vincent, Dellalieu, O., Roux, L., and Quimet, M., Out-of sample extensions for LLE, Isomap, MDS, eigenmaps, and spectral clustering, in *Advances in Neural Information Processing System*, vol. 16. Cambridge, MA, MIT Press, 2004.
20. Vural, E. and Guillemot, C., Out-of-Sample Generalizations for Supervised Manifold Learning for Classification, *IEEE Transactions on Image Processing*, 25 (3), 1410-1424, 2016.
21. He, X. and Niyogi, P., Locality preserving projections, in *Advances in Neural Information Processing System*, Cambridge, MA, MIT Press, 16, 2004.
22. He, X., Cai, D., Yan, S., and Zhang, H., Neighborhood preserving embedding, in *Proc. IEEE Int. Conf. Comput. Vis.*, 1208–1213, 2005.
23. Taşkın, G. and Crawford M. M., An out-of-sample extension to manifold learning via meta-modelling, *IEEE Transactions on Image Processing*, 28 (10), 5227-5237, 2019.
24. Cai, D., He, X., and Han, J., Spectral Regression for Dimensionality Reduction Spectral Regression for Dimensionality Reduction, University of Illinois at Urbana-Champaign, Tech. Rep. May, 2007.
25. Liu, B., Xia, S. X., Meng, F. R., and Zhou, Y., Extreme spectral regression for efficient regularized subspace learning, *Neurocomputing*, 149 (Part A), 171–179, 2015.
26. Barkan, O., Weill J., and Averbuch, A., Gaussian process regression for out-of-sample extension, in *IEEE Int. Workshop Mach. Learn. Signal Process. (MLSP)*, 1-6, 2016.
27. Hang, R., Liu, Q., Song, H., Sun, Y., Zhu, F., and Pei, H., Graph regularized nonlinear ridge regression for remote sensing data analysis, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10 (1), 277–285, 2017.
28. Tang, M., Nie, F., and Jain, R., A graph regularized dimension reduction method for out-of-sample data, *Neurocomputing*, 225, 58-63, 2017.
29. Orsenigo and C. Vercellis, Kernel ridge regression for out-of-sample mapping in supervised manifold learning, *Expert Systems with Applications*, 39 (9), 7757–7762, 2012.
30. G. Liu, Z. Lin, and Y. Yu, Multi-output regression on the output manifold, *Pattern Recognition*, 42 (11), 2737–2743, 2009.
31. Saunders, C., Gammerman, A., and Vovk, V., Ridge regression learning algorithm in dual variables. In *Proceedings of the 15th Int. Con. on Mac. Learn.*, Morgan Kaufmann, 515–521, 1998.
32. Smola, A. J. and Schölkopf, B., A tutorial on support vector regression, *Statistics and Computing*, 14 (3), 199–222, 2004.
33. Williams, C. K., and Rasmussen, C. E., *Gaussian processes for machine learning*, MIT Press, 2006.
34. Tipping, M. E., *The Relevance Vector Machine*, *Advances in Neural Information Processing Systems 12*, Cambridge, Mass: MIT Press, 2000.
35. Huang, G.-B., and Chen, L., Enhanced random search based incremental extreme learning machine. *Neurocomputing*, 71 (16–18), 3460–3468, 2008.
36. Van der Maaten, L.J.P., *An Introduction to Dimensionality Reduction Using MatLab*, Report MICC 07-07, July, 2007.

