





International JOURNAL of SOCIAL and HUMANITIES SCIENCES RESEARCH (JSHSR)


Uluslararası Sosyal ve Beşeri Bilimler Araştırma Dergisi

Received/Makale Geliş 20.03.2023
Published /Yayınlanma 31.05.2023
Volume/Issue (Cilt/Sayı)-ss/pp 10(95),1146-1150

<http://dx.doi.org/10.26450/jshsr.3642>
Research Article
ISSN: 2459-1149

 Hüseyin YAŞAR

 <https://orcid.org/0000-0003-4081-6736>

 Dokuz Eylül Üniversitesi, Fen Bilimleri Enstitüsü, Coğrafi Bilgi Sistemleri, İzmir / TÜRKİYE

LANDSAT 8 ÖRNEĞİYLE MULTİSPEKTRAL GÖRÜNTÜ İŞLEMEDE YAPAY SİNİR AĞLARI, DESTEK VEKTÖR MAKİNELERİ VE KARAR AĞAÇLARI ALGORİTMALARININ DUYARLILIKLARININ KARŞILAŞTIRILMASI

COMPARISON OF SENSITIVITIES OF ARTIFICIAL NEURAL NETWORKS, SUPPORT VECTOR MACHINES, AND DECISION TREES ALGORITHMS IN MULTISPECTRAL IMAGE PROCESSING: A CASE STUDY WITH LANDSAT 8

ÖZET

Bu çalışma, sınıflandırma algoritmalarının duyarlılıklarını analiz etmeye yöneliktir. Çalışma alanı olarak İzmir'in Urla ilçesi seçilmiştir ve bölgenin aralık ve nisan LANDSAT 8 görüntüleri eş zamanlı olarak işlenmiştir. Sınıflandırmada Yapay Sinir Ağları (YSA), Destek Vektör Makineleri (DVM) ve Karar Ağaçları (KA) algoritmaları kullanılmıştır. Sınıflandırma için şehir, orman, sera, çayır, narenciye, ekili ve yanmış alanlardan oluşan dokular seçilmiştir. Bu dokuların seçimi, GPS cihazı ile yerinde gözlem (in-situ) araştırmaları yapılarak gerçekleştirilmiştir. Çalışmanın sonuçları, kompleks matris algoritması kullanılarak sınıflandırmaların duyarlılıklarının test edilmesiyle elde edilmiştir. Farklı mevsimlerden elde edilmiş görüntülere göre genel doğruluklar hesaba katıldığında, aralık verileri ile yapılan çalışma nisan verileri ile yapılan çalışmadan %8 daha fazla doğruluk sağlamıştır. Algoritmalara göre, en yüksek doğruluk oranı %85 ile YSA algoritması tarafından sağlanmıştır. KA algoritması, bilgisayarda harcanan zaman açısından avantajlı olmasına rağmen sınıflandırma konusunda en az doğruluk oranını sağlayan algoritma olarak tespit edilmiştir. DVM algoritmaları ise geniş alanlı sınıflandırma çalışmalarında hem bilgisayarda harcanan zaman hem de doğruluk analizleri göz önüne alındığında avantajlıdır. Bu çalışma, bölgesel doğal kaynakların korunması ve yönetimi konusunda faydalı olabilecek veriler sağlamaktadır. Sınıflandırma algoritmalarının duyarlılıklarının analiz edilmesi, benzer çalışmalar için bir temel sağlayabilir ve gelecekteki araştırmalar için bir referans olabilir.

Anahtar Kelimeler: Yapay sinir ağları, Karar Ağaçları, Zemin izleme, Destek Vektör Makineleri.

ABSTRACT

This study aims to analyze the sensitivity of classification algorithms. The study area was selected as the Urla district of Izmir, and the spring and autumn LANDSAT 8 images of the region were processed simultaneously. Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Decision Trees (DT) algorithms were used for classification. The tissues of the region consisting of urban, forest, greenhouse, meadow, citrus, cultivated, and burnt areas were selected for classification. The selection of these tissues was made by conducting in-situ observations using a GPS device. The results of the study were obtained by testing the sensitivity of classifications using the complex matrix algorithm. When considering the overall accuracies according to the seasons, the study conducted with December data provided 8% more accuracy than the one conducted with April data. According to the algorithms, the highest accuracy rate was achieved by the ANN algorithm with a rate of 85%. Although the DT algorithm was advantageous in terms of time spent on the computer, it was identified as the algorithm providing the least accuracy rate in classification. SVM algorithms are advantageous in large-scale classification studies in terms of both times spent on the computer and accuracy analyses. This study provides useful data for the conservation and management of regional natural resources. Analyzing the sensitivity of classification algorithms can provide a basis for similar studies and serve as a reference for future research.

Keywords: Artificial Neural Network, Decision Tree, Ground Monitoring, Support Vector Machines.

Çalışmada seçilen Urla yarımadası İzmir'e bağlı bir ilçedir. $38^{\circ}12'53''$, $38^{\circ}27'14''$ kuzey ve $26^{\circ}38'50''$, $26^{\circ}54'40''$ doğu koordinatları arasında yer almaktadır. Toplam çalışma alanı 400 km^2 alanı kapsamaktadır. Görüntü işlemede spektral hatalardan kaçınmak için deniz kısmı maskelenmiştir.



Şekil 2 Uygulama Alanının Kuzey-Ege Coğrafyasındaki Yeri

Uygulamada Landsat 8 Nisan-2013 tarihli ve Aralık 2013 tarihli veriler kullanılmıştır. Çalışmanın amacına uygun olabilmesi için görüntüler seçilirken arazinin yeşillenme, ekim ve hasat zamanları göz önüne alınmıştır.

2.2. Metot

Kullanılan bu veriler sırasıyla destek vektör makineleri (DVM), ileri beslemeli yapay sinir ağları (IBYSA) ve karar ağaçları (KA) algoritmalarıyla işlenmiştir. Duyarlılıklarının analizi için kompleks matris kullanılmıştır.

Destek vektör makineleri algoritması sınıflandırma problemleri için, Vladimir Vapnik ve arkadaşları tarafından geliştirilen istatistiksel öğrenme teorisine dayalı parametrik olmayan bir sınıflandırma yöntemidir (Hong ve arkadaşları, 2008). Bu yöntem nesne tanıma, el yazısı tanımlama, metin sınıflandırmaları gibi birçok alanda kullanılmaktadır (Duan ve Jiang, 2012). DVM, eğitim verisi üzerinde örnekleme yaparak diğer veriler üzerinde tahmin etme ve genelleştirme temeline dayalı bir makine öğrenmesidir. İlk olarak iki sınıflı doğrusal verilerin sınıflandırılmasını hedefleyen DVM, bazı problemlerde yetersiz kalınca ilerleyen dönemlerde çok sınıflı ve doğrusal olmayan verilerin sınıflandırılması için geliştirilmiştir (Vapnik, 1995).

Yapay sinir ağları (YSA), insan beyninin özelliklerinden olan öğrenme yolu ile yeni bilgiler türetebilme, yeni bilgiler oluşturabilme ve keşfedebilme gibi yetenekleri, herhangi bir yardım almadan otomatik olarak gerçekleştirebilmek amacı ile geliştirilen bilgisayar sistemleridir. Bu çalışmada ileri beslemeli yapay sinir ağ yapısı kullanılmıştır. Giriş ve çıkış katmanından oluşan bu sistemde bir katmandaki hücrelerin çıkışları bir sonraki katmanda ağırlık olarak verilmekte; giriş katmanı, girdi vektörü yardımı ile elde ettiği bilgileri ağırlık katsayısı ile çarparak gizli katmandaki hücre yapılarına iletmektedir.

Karar ağaçları yapısal olarak üç temel bileşenden oluşur bunlar düğüm, dal ve yapraklardan oluşmaktadır. Bu ağaç yapısında her bir öznitelik (uydu görüntüsü için bant değeri) bir düğüm tarafından temsil edilir. Dallar ve yapraklar ağaç yapısının diğer elemanlarıdır. Ağaçta en son kısım yaprak en üst kısım ise kök olarak adlandırılır. Kök ve yapraklar arasında kalan kısımlar ise dal olarak ifade edilir. Karar ağaçlarının oluşturulmasındaki en önemli adım ağaçtaki dallanmanın hangi kriterlere veya kıstasa göre yapılacağı ya da hangi öznitelik değerlerine göre ağaç yapısının oluşturulacağıdır.

Değerlendirilmenin doğruluğu kompleks matris ile sağlanır. Doğrulama Eğitim alanı olarak ayrılan bölgeler dışında kalan test alanlarına ilişkin piksel değerlerinin, referans kabul edilen, haritalar ya da arazide eğer çok sayıda sınıflandırılmamış piksel varsa eğitim veri setlerinin gerçeği temsil etme olasılığı düşer [15]. Bu çalışmada doğruluk analizi yaparken kendi oluşturduğumuz $k \times k$ kompleks matris algoritması izlenmiş ve Python 2.7 Shell formatında programlanmıştır. İşlemin sonucunda değerin $0 < \chi <$

1 arasında olması beklenir. Sonuç 1'e yaklaştıkça doğruluk artar; 0'a yaklaştıkça doğruluk azalır. Yapılan çalışmada aynı sınıflar olmasına karşın mevsimlere göre algoritmalar farklılıklar gözlenmiştir. Buradaki en temel etken bitki örtüsünün mevsimsel değişkenliğidir.

3. SONUÇLAR

2013 nisan ve aralık ayına ait 7 sınıftan oluşan görüntü işleme yapılmış, algoritmalara ve mevsimlere göre farklılıklar gözlenmiştir.

LANDSAT-8 08.04.2013	IBYS		DVM		KA	
	SINIF	Doğruluk%	SINIF	Doğruluk%	SINIF	Doğruluk%
KOMPLEKS MATRİS	ŞEHİR	94,36	ŞEHİR	91,80	ŞEHİR	78,10
	ÇAYIR	90,87	ÇAYIR	87,58	ÇAYIR	43,55
	SERA	85,82	SERA	63,91	SERA	81,81
	NARENCİYE	81,60	NARENCİYE	79,82	NARENCİYE	88,82
	TARIM ALANLARI	78,57	TARIM ALANLARI	69,57	TARIM ALANLARI	20,27
	YANMIŞ ARAZI	98,79	YANMIŞ ARAZI	77,83	YANMIŞ ARAZI	83,57
	LANDSAT-8 04.12.2013	IBYS		DVM		KA
KOMPLEKS MATRİS	SINIF	Doğruluk%	SINIF	Doğruluk%	SINIF	Doğruluk%
	ŞEHİR	92,75	ŞEHİR	91,80	ŞEHİR	73,75
	ÇAYIR	94,80	ÇAYIR	82,38	ÇAYIR	82,38
	SERA	90,81	SERA	63,91	SERA	81,80
	NARENCİYE	81,73	NARENCİYE	73,75	NARENCİYE	82,80
	TARIM ALANLARI	65,22	TARIM ALANLARI	69,57	TARIM ALANLARI	64,77
	YANMIŞ ARAZI	96,82	YANMIŞ ARAZI	77,83	YANMIŞ ARAZI	91,47

Şekil 3 Sınıflandırmaların Doğruluk Oranları

Gözlemlenen veriler, Şekil 3'te sunulmuştur. Buna göre, nisan ve kasım aylarına ait analizlerde, şehir alanlarının sınıflandırılmasıyla ilgili algoritmaların genel doğruluk değerleri arasında bir fark gözlenmemiştir ve ortalama doğruluk değeri %87 olarak belirlenmiştir. Çayır alanlarının sınıflandırılmasında, algoritmalar arasında önemli farklılıklar gözlemlenmiştir. Nisan ayında algoritmaların doğruluk ortalaması %43,55 iken aralık ayında %85'tir. Sera alanlarının sınıflandırılmasında ise, DVM sonuçları aylara göre farklılık göstermeden %63 doğrulukla sınıflandırma yaparken, IBYSA aralık verilerini %96 ile sınıflandırabilmişken, nisan verilerinde %69 doğrulukla sınıflandırmıştır. KA algoritması ise nisan verilerini %81 doğrulukla sınıflandırırken, aralık verilerinde %62 doğrulukla sınıflandırma yapabildiği görülmüştür. Narenciye alanlarının sınıflandırılmasında, DVM analizi sonuçları aylara göre ciddi farklılık göstermemiştir ve ortalama doğruluk %74'tür. IBYSA nisan verilerini %82 doğrulukla sınıflandırabilmişken, aralık verilerinde %62 doğrulukla sınıflandırma yapabildiği görülmüştür. KA algoritması ise nisan verilerini %50 doğrulukla, aralık verilerini ise %82 doğrulukla sınıflandırma yapabildiği görülmüştür. Ekili tarım alanlarının sınıflandırılmasında, DVM sınıflayıcısı aylara göre farklılık göstermeden %70 doğrulukla sınıflandırma yaparken, IBYSA nisan verilerinde %79 doğruluk, aralık verilerinde %65 doğrulukla sınıflandırma yapmıştır. KA algoritması ise nisan ayı sınıflandırmasında %20, aralık ayında %70 doğrulukla sınıflandırma yapabildiği görülmüştür. Son olarak, yanmış orman arazilerinin sınıflandırılmasında DVM ve KA algoritmaları nisan ve aralık verilerinde benzer doğruluk değerleri göstermiştir. IBYSA ise nisan verileriyle %91, aralık verileriyle %97 doğrulukla sınıflandırma yapabildiği görülmüştür.

4. BULGULAR

Araştırma sonuçlarına göre, Landsat 8 ve ETM verileriyle yapılan sınıflandırmalar incelendiğinde, Çok Katmanlı Yapay Sinir Ağı (YSA) yöntemi, genel doğruluk oranlarının yüksek olduğu ve mevsimlere göre ve eğitim sınıflarına göre en başarılı sonuçları veren yöntem olarak belirlenmiştir. Ancak, YSA algoritması bilgisayar zamanı işlem sürelerinin uzunluğu nedeniyle pratik uygulamalarda dezavantajlı olabilir.

Buna karşılık, DVM lineer Kernel fonksiyonu, geniş yer yüzeyi sınıflandırmalarında avantaj sağlayan bir yöntem olarak belirlenmiştir. Ayrıca, K-En Yakın Komşu (KA) yöntemi, hızlı ve etkin sonuçlar vermesi nedeniyle tercih edilebilir bir seçenek olarak değerlendirilmiştir.

Çalışmada, Landsat 8 görüntüleri için sınıflandırmalarda maksimum 1000m² homojen eğitim setleri oluşturulması önerilmiştir. Tek gizli katmanlı YSA yöntemi, çoğu durumda yeterli sonuçlar verirken, daha hassas çalışmalarda piksel içindeki farklı her bir yapı için bir gizli katman eklenmesi gerekebilir.

Öte yandan, Landsat ETM verileriyle yapılan çalışmada, İğne yapraklı, geniş yapraklı ağaçlar, bozkır ve yerleşim sınıflarından oluşan bir sınıflandırma yapılmıştır. Bu çalışmada, KA yöntemi oldukça etkin ve

hızlı sonuçlar vermiştir. Ancak, bozkır alanlarının sınıflandırılmasında, KA yöntemi aylara göre farklılıklar göstermiştir ve nisan verileriyle düşük doğruluk oranları elde edilmiştir.

Sonuç olarak, yapılan araştırmaların bulguları, Landsat 8 ve ETM verileriyle yapılan sınıflandırmalarda farklı yöntemlerin kullanılabilmesini ve seçilen yöntemin veri seti, mevsim ve eğitim sınıfları gibi faktörlere bağlı olarak değişebileceğini göstermektedir. Bu nedenle, sınıflandırma yöntemi seçerken bu faktörlerin dikkate alınması ve önceden yapılmış benzer çalışmaların incelenmesi önerilmektedir.

KAYNAKLAR

- Ayhan, E., Karşlı, F. ve Tunç, E. (2009). *Uzaktan Algılanmış Görüntülerde Sınıflandırma ve Analiz*. Trabzon: ResearchGate.
- Campbell, J. B. ve James, B. (1996). *Introduction to Remote Sensing*. New York: Guilford Press.
- Hong, J., Min, J., Cho, U. & Cho, S. (2008). Fingerprint classification using one-vs-all support vector machines dynamically ordered with naïve Bayes classifiers. *Pattern Recognition*, 41(2), 662–671.
- Kavzaoğlu, T. ve Çölkesen, E. (2010). Destek Vektör Makineleriyle Uydu Görüntülerinin Sınıflandırılmasında Kernel Fonksiyonlarının Etkilerinin İncelenmesi. *Harita Dergisi*, 144, 73-82
- Maktav, K. ve Kalkan, D. (2010). *Nesne Tabanlı ve Piksel Tabanlı Sınıflandırma Yöntemlerinin Karşılaştırılması*. Kocaeli: III. Uzaktan Algılama ve Coğrafi Bilgi Sistemleri Sempozyumu 11-13 Ekim, 2010.
- Karakis, S., Marangoz, A. M. ve Büyüksalih, G. (2006). *Analyses of segmentation parameters in eCognition software using high-resolution QuickBird MS imagery*. Ankara: Türkiye Ulusal Fotogrametri ve Uzaktan Algılama Birliği.
- Sabins, F. ve Floyd, R. (1997). *Remote sensing: principles and interpretation*. New York: Trove.
- Song, X., Duan, Z. ve Jiang, X. (2012). Comparison of artificial neural networks and support vector machine classifiers for land cover classification in Northern China using a SPOT-5 HRG image. *International Journal of Remote Sensing*, 33(10), 3301-3320.
- Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York: Springer-Verlag.
- Yaşar, H. (2016). *İleri Sınıflandırma Algoritmalarının Kıyaslanması*. Yayımlanmış yüksek lisans tezi. Dokuz Eylül Üniversitesi, İzmir, Türkiye.