



GAIA DR3 ÖRTEN ÇİFT YILDIZ IŞIK EĞRİLERİNİN SINIFLANDIRILMASI: BİR MAKİNE ÖĞRENMESİ YAKLAŞIMI

Bedri Keskin

Ankara Üniversitesi, Fen Fakültesi, Astronomi ve Uzay Bilimleri Bölümü, 06100, Tandoğan, Ankara, Türkiye
bedri.keskin@gmail.com

UAK
2026

Özet

Gaia DR3 yaklaşık 2.1 milyon örten çift yıldız adayı içeren benzersiz bir veri seti sunmaktadır. Bu kadar büyük bir veri hacminin manuel olarak sınıflandırılmasının sürdürülemezliği, güvenilir ve ölçeklenebilir otomatik tekniklerin geliştirilmesini zorunlu kılmıştır. Bu çalışmada, Gaia DR3 arşivindeki örten çift yıldızların ışık eğrisi morfolojisine (EA, EB, EW) göre otomatik sınıflandırılması için çok modlu makine öğrenmesi (ML) modeli geliştirilmiştir. Geliştirilen model, ışık eğrisi görüntülerinden görsel özellikleri çıkaran bir Evrişimsel Sinir Ağı (CNN) ve geometrik model parametrelerini işleyen bir Çok Katmanlı Algılayıcı (MLP) kullanmaktadır. Modelin geometrik şekillere odaklanmasını sağlamak için eğitim sürecinde gürültüsüz sentetik ışık eğrileri kullanılmıştır. Testler, modelin tüm sınıflar için %95'in üzerinde doğruluk oranına ulaştığını ve özellikle EA tipi sistemlerde mükemmel ayırım performansı sergilediğini göstermiştir. Eğitilen modelle gerçekleştirilen otomatik sınıflandırma sonucunda, Gaia DR3 örten çift sistemlerinin %40'ı EA, %30'u EB ve %30'u EW olarak sınıflandırılmıştır. Bu çalışma, gelecekteki büyük ölçekli gökyüzü araştırmaları için oldukça doğru bir sınıflandırma çerçevesi sunmaktadır.

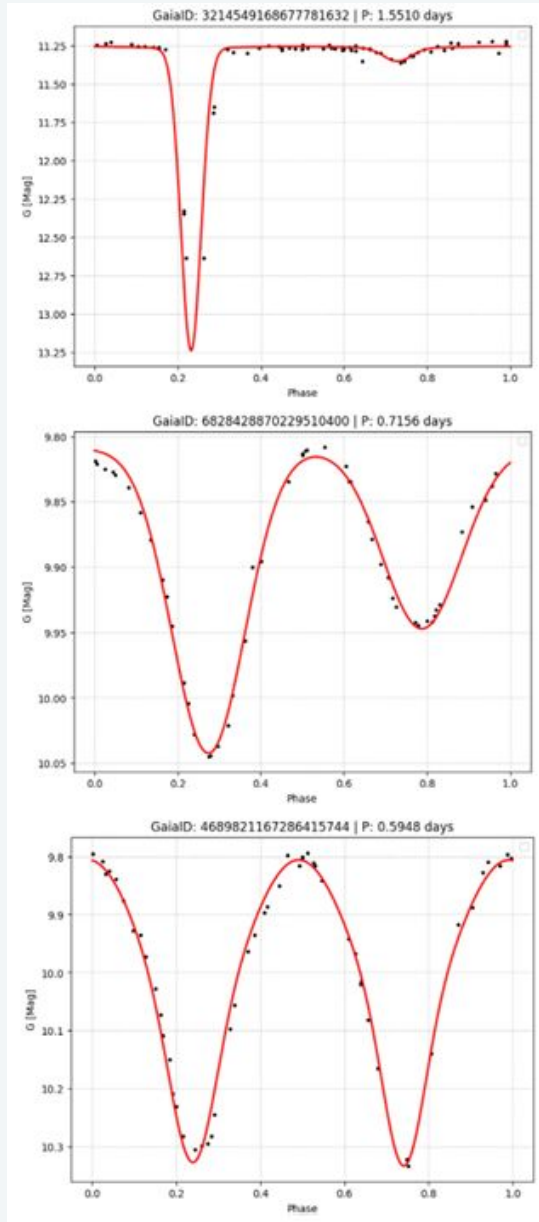
Veri Hazırlama

Mowlavi vd. (2023), Gaia DR3 örten çift yıldızlarını bir veya iki Gauss ve/veya bir kosinüs fit ederek geometrik 6 farklı şekilde modellemiştir (Çizelge 1).

Çizelge 1. Geometrik model türleri (#: parametre sayısı)

Model türü	#	Açıklama	Örten çift sayısı
TWOGAUSSIANS	8	İki Gaussyen	1 587 926
TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE1	9	İki Gaussyen ve 1. Gaussyen üzerinde bir kosinüs	389 725
TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE2	9	İki Gaussyen ve 2. Gaussyen üzerinde bir kosinüs	85 400
ONEGAUSSIAN	5	Bir Gaussyen	36 984
ONEGAUSSIAN_WITH_ELLIPSOIDAL	6	Bir Gaussyen ve bir kosinüs	48 215
ELLIPSOIDAL	4	Bir kosinüs	36 227
Toplam Örten Çift Sayısı			2 184 477

Söz konusu yıldızların gözlem verileri ve model parametreleri internetten indirildi¹. Model parametreleri ve denklemleri kullanılarak PNG formatında sentetik ışık eğrileri oluşturuldu (Şekil 1).



Şekil 1. EA, EB, EW sınıflarına ait örnek model grafikleri

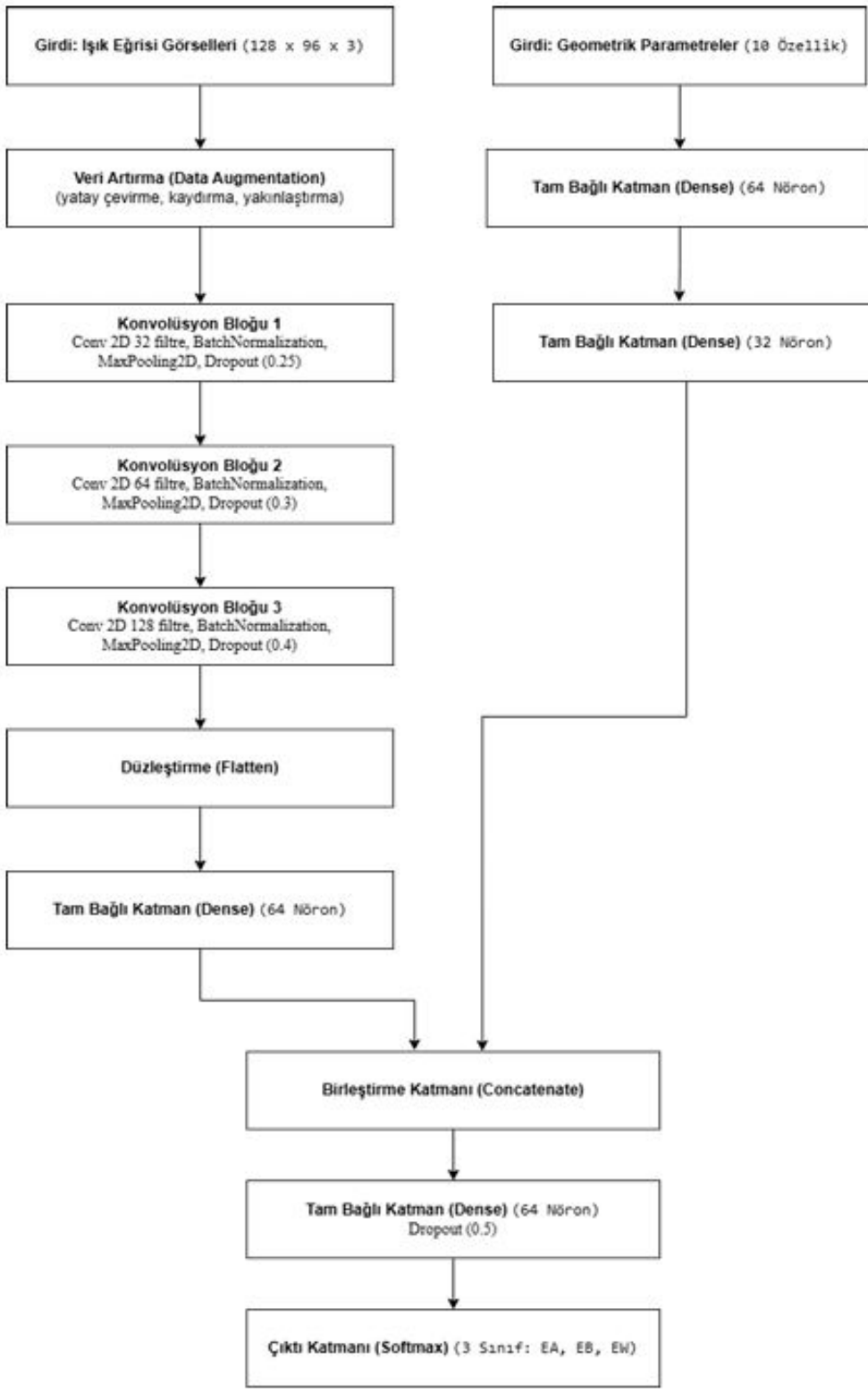
Yıldızların bir kısmı ML modelinin eğitilmesi için manuel olarak EA, EB, EW şeklinde sınıflandırıldı (Çizelge 2).

Çizelge 2. Görsel inceleme ile sınıflandırılan yıldız sayıları

Model türü	Eğitim ve Doğrulama	Test
TWOGAUSSIANS	15 094	3 031
TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE1	10 032	1 983
TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE2	7 981	1 608
ONEGAUSSIAN_WITH_ELLIPSOIDAL	7 537	1 558
Toplam	40 644	8 180

Çok Modlu ML Modeli

İki farklı veri türünü eş zamanlı olarak işleyen çok modlu bir derin öğrenme modeli kurulmuştur: görsel ışık eğrisi görüntüleri ve sayısal model parametreleri. Mimari, iki paralel giriş dalından oluşur; görsel özellik çıkarımı için CNN ve sayısal veriler için MLP. Bu dallar daha sonra nihai sınıflandırma için birleştirilir (Şekil 2).



Şekil 2. Çok modlu (multimodal) makine öğrenmesi modeli

Görsel Dal: Işık eğrisi görüntüleri 128x96 piksele küçültülüp [0, 1] aralığında normalize edildikten sonra CNN (Conv2D, MaxPooling2D, Flatten ve Dense) katmanlarıyla işlenerek özellik çıkarımı yapılır.

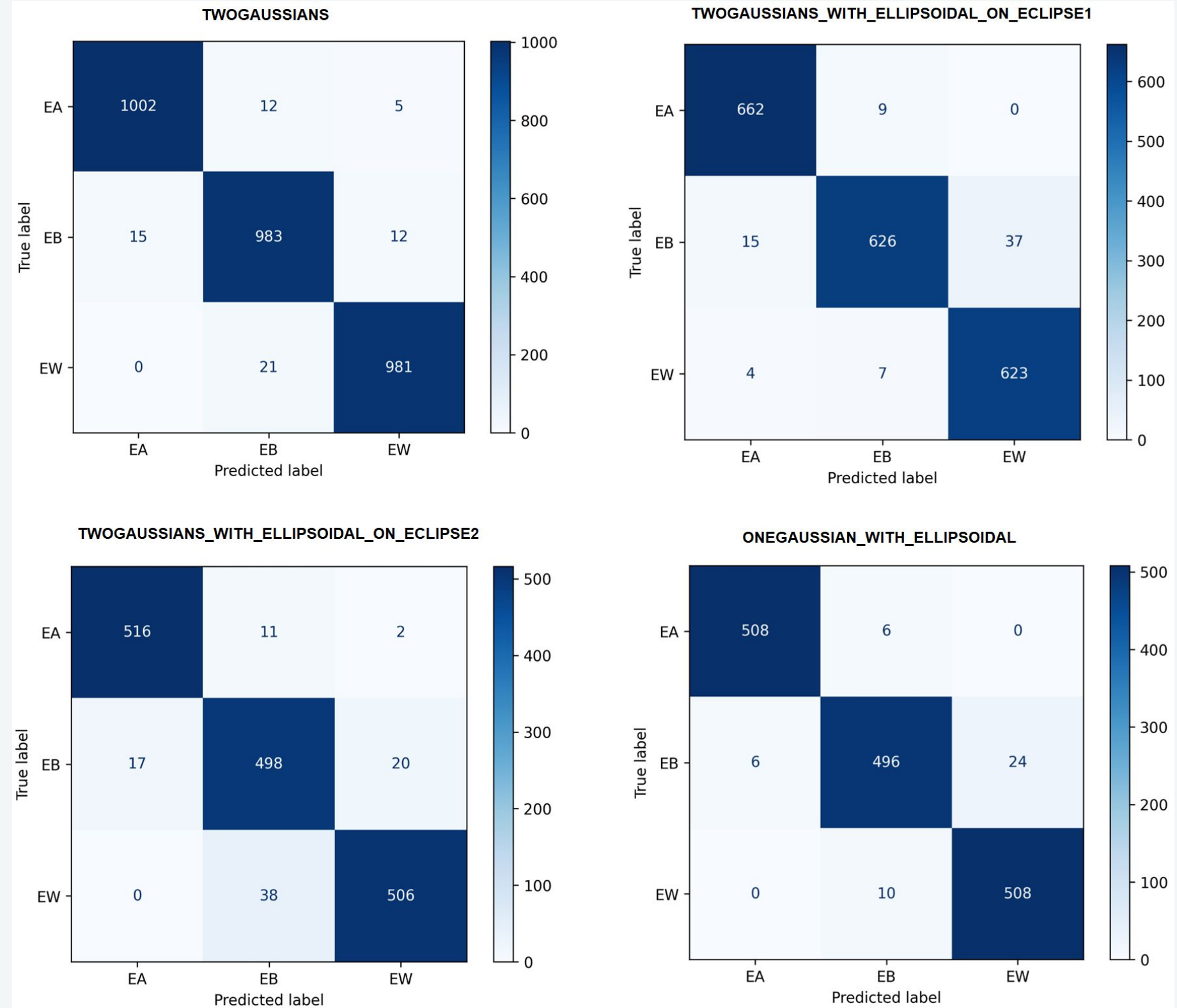
Sayısal Dal: İlgili parametreler standartlaştırılır (ortalama 0, standart sapma 1) ve ReLU aktivasyonlu iki katmanlı (sırasıyla 32 ve 16 nöronlu) çok katmanlı algılayıcıdan (MLP) geçirilir.

Birleştirme ve Sınıflandırma: CNN ve MLP dallarından gelen 16'şar boyutlu vektörler birleştirilerek 32 nöronlu (ReLU) ek bir yoğun katmana aktarılır. Çıktı, EA, EB ve EW sınıflarını belirlemek üzere 3 nöronlu softmax katmanıyla sınıflandırılır.

Eğitim Detayları: Model; 'Adam' optimizasyonu, 'kategorik çapraz entropi' kayıp fonksiyonu, 64 batch boyutu ve maksimum 60 epoch ile eğitilmiş, aşırı uyumu (overfitting) önlemek için çeşitli geri çağırma (callback) fonksiyonları kullanılmıştır.

Sonuç

Görsel CNN ve sayısal MLP özelliklerini birleştiren çok modlu model, örten çift yıldızları sınıflandırmada %95'in üzerinde doğruluk (accuracy) oranına ulaşarak geleneksel tekli yöntemlere kıyasla oldukça başarılı bir alternatif olduğunu kanıtlamıştır (Şekil 3). Model, belirgin özellikler taşıyan EA türü yıldızlarda neredeyse kusursuz sonuçlar verirken; fiziksel ve morfolojik olarak birbirine daha benzeyen EB ve EW türlerinde hafif bir performans düşüşü gösterse de yüksek hassasiyetini korumuştur. Bu sonuçlar, modelin ışık eğrisi morfolojilerindeki ince geçişlerde bile herhangi bir sınıf yanlılığı olmaksızın, örten çift yıldız türlerini yüksek doğrulukla ayırt edebildiğini göstermektedir.



Şekil 3. Hata matrisleri (Confusion matrices)

211266 örten çift yıldızdan 40644'ü eğitim/doğrulama için, 8180'i ise test için manuel olarak sınıflandırılırken, kalan 2062442 yıldızın tamamı otomatik olarak %40 EA, %30 EB ve %30 EW olarak sınıflandırıldı (Çizelge 3).

Çizelge 3. Çok modlu makine öğrenmesi modeli ile otomatik sınıflandırılan örten çift yıldızlarının sayıları ve oranları

Model türü	EA	EB	EW	Toplam
TWOGAUSSIANS	703 865 (%45)	446 653 (%28)	419 283 (%27)	1 569 801
TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE1	88 235 (%23)	120 673 (%32)	168 802 (%45)	377 710
TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE2	22 784 (%30)	28 834 (%38)	24 193 (%32)	75 811
ONEGAUSSIAN_WITH_ELLIPSOIDAL	8 294 (%21)	25 014 (%64)	5 812 (%15)	39 120
Otomatik sınıflandırılan toplam örten çift sayısı	823 178 (%40)	621 174 (%30)	618 090 (%30)	2 062 442

¹ https://cdn.gea.esac.esa.int/Gaia/gdr3/Variability/vari_eclipsing_binary/

Kaynak: Mowlavi vd. (2023) Gaia Data Release 3: The first Gaia catalogue of eclipsing-binary candidates. Astronomy and Astrophysics, 674, A16.

Bu çalışma New Astronomy dergisinde yayınlanmıştır: Keskin, B. and Baştürk, Ö. Classification of eclipsing binary light curves in Gaia DR3: A machine learning approach. New Astronomy, Volume 127, 2026, 102585, ISSN 1384-1076.

