

**ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

DOKTORA TEZİ

**MAKİNE ÖĞRENMESİ TEKNİKLERİ İLE ÖRTEN ÇİFT YILDIZ
ALGILAMA VE SINIFLANDIRMA**

Bedri KESKİN

ASTRONOMİ VE UZAY BİLİMLERİ ANABİLİM DALI

**ANKARA
2026**

Her hakkı saklıdır

ÖZET

Doktora Tezi

MAKİNE ÖĞRENMESİ TEKNİKLERİ İLE ÖRTEN ÇİFT YILDIZ ALGILAMA VE SINIFLANDIRMA

Bedri KESKİN

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
Astronomi ve Uzay Bilimleri Anabilim Dalı

Danışman: Prof. Dr. Özgür BAŞTÜRK

Bu tez çalışmasında büyük astronomi arşivlerindeki örten çift yıldızların makine öğrenmesi teknikleriyle sınıflandırılması hedeflenmiştir.

Gaia DR3 arşivindeki örten çift yıldızların yaklaşık 48 bin kadarı görsel inceleme ile ışık eğrisi biçimine (EA, EB, EW) göre etiketlendi. 40 bin kadarı eğitim/doğrulama için, 8 bin kadarı da test için ayrıldı. Işık eğrilerinin hem görsellerini hem de geometrik model parametrelerini aynı anda işleyen çok modlu (multimodal) makine öğrenmesi modeli oluşturuldu. Etiketli veri ile çok modlu model eğitildi. Eğitilmiş model ile Gaia DR3 arşivindeki yaklaşık 2 milyon örten çift yıldız ışık eğrisi biçimine (EA, EB, EW) göre otomatik olarak sınıflandırıldı.

Kepler ve TESS arşivlerindeki örten çift yıldızlar morfoloji parametresine (Ayrık, Yarı-ayrık, Aşırı Değen, Elipsoidal) göre etiketlendi. Işık eğrilerinin görselleri ile görüntü sınıflandırma için özel olarak geliştirilmiş VGG-19 modeli eğitildi. Test verisinde %91 doğruluk elde edildi. Eğitilmiş model ile Kepler ve TESS arşivleri ile Gaia DR3 arşivinde ortak bulunan 2106 örten çift yıldız tahmin edildi, %64 doğruluk elde edildi.

Kepler, TESS ve Gaia DR3 arşivinde ortak bulunan 2106 örten çift yıldızın ışık eğrileri 11 kosinüs terimli Fourier serisi ile modellendi. Örten çift yıldızlar morfoloji parametresine (Ayrık, Yarı-ayrık, Aşırı Değen, Elipsoidal) göre etiketlendi. Kepler ve TESS örten çift yıldızlarının Fourier katsayıları ile makine öğrenmesi modeli eğitildi. Test verisinde %65 doğruluk elde edildi. Eğitilmiş model ile Kepler ve TESS arşivleri ile Gaia DR3 arşivinde ortak bulunan 2106 örten çift yıldız tahmin edildi, %33 doğruluk elde edildi.

Kepler arşivinde bulunan örten çift yıldızlar Fourier katsayılarına göre kümelendirildi. Yıldızlar 3 kümeye ayrıldı ve ışık eğrisi biçimine (EA, EB, EW) göre yapılan sınıflandırma ile ilişkisi araştırıldı.

Şubat 2026, 77 sayfa

Anahtar sözcükler: çift yıldızlar, örten çiftler, ışık eğrisi sınıflandırma, makine öğrenmesi

ABSTRACT

PhD Thesis

ECLIPSING BINARY DETECTION AND CLASSIFICATION WITH MACHINE LEARNING TECHNIQUES

Bedri KESKİN

Ankara University
Graduate School of Natural and Applied Science
Department of Astronomy and Space Sciences

Supervisor: Prof. Dr. Özgür BAŞTÜRK

This thesis aims to classify eclipsing binary stars in large astronomical archives using machine learning techniques.

Approximately 48,000 eclipsing binary stars in the Gaia DR3 archive were visually labeled according to their light curve shape (EA, EB, EW). About 40,000 were reserved for training/validation, and 8,000 for testing. A multimodal machine learning model was created that simultaneously processes both visualizations of the light curves and geometric model parameters. The multimodal model was trained with the labeled data. Using the trained model, approximately 2 million eclipsing binary stars in the Gaia DR3 archive were automatically classified according to their light curve shape (EA, EB, EW).

Eclipsing binary stars in the Kepler and TESS archives were labeled according to their morphology parameter (Detached, Semi-Detached, Overcontact, Ellipsoidal). The VGG-19 model, specifically developed for image classification, was trained using visualizations of light curves. 91% accuracy was achieved with the test data. The trained model predicted 2106 eclipsing binary stars common to the Kepler, TESS and Gaia DR3 archives, achieving 64% accuracy.

The light curves of 2106 eclipsing binary stars common to the Kepler, TESS and Gaia DR3 archives were modeled using an 11 terms cosine Fourier series. The eclipsing binary stars were labeled according to their morphology parameters (Detached, Semi-Detached, Overcontact, Ellipsoidal). The machine learning model was trained using the Fourier coefficients of the Kepler and TESS eclipsing binary stars. 65% accuracy rate was achieved with the test data. The trained model predicted 2106 eclipsing binary stars common to the Kepler, TESS and Gaia DR3 archives, achieving a 33% accuracy rate.

Eclipsing binary stars in the Kepler archive were clustered according to their Fourier coefficients. The stars were divided into 3 clusters and their relationship to classification based on light curve shape (EA, EB, EW) was investigated.

February 2026, 77 pages

Key words: binary stars, eclipsing binaries, light curve classification, machine learning

ÖNSÖZ ve TEŞEKKÜR

İlk olarak, 2018’de aramızdan ayrılan, yüksek lisans eğitimimde danışmanım sayın Prof. Dr. Berahitdin ALBAYRAK hocamı rahmetle anmak istiyorum. Yüksek lisans tezimi savunmam sonrasında “Bedri sen doktora devam et” sözleri hala kulağımdadır. Birkaç hafta sonrasında maalesef elim bir kazada kaybettik. Sanırım mezun ettiği son öğrencisi bendim. Nur içinde yat Hocam.

2019’da başladığım uzun doktora öğrenim hayatımın her aşamasında bilgi, öneri ve yardımlarını esirgemeyerek yetiştirme ve gelişmeye katkıda bulunan danışman hocam sayın Prof. Dr. Özgür BAŞTÜRK ve sayın hocam Prof. Dr. Selim Osman SELAM başta olmak üzere bütün hocalarıma, adlarını burada sayamadığım beraber ders aldığım bütün arkadaşlarıma en içten en samimi duygularıyla teşekkür ederim.

Işık eğrilerini görsel inceleme ile sınıflandırmasında emekleri için Alkım Özfıdan’a ayrıca teşekkürü borç bilirim.

Bu araştırmada yer alan kısmi nümerik hesaplamalar TÜBİTAK ULAKBİM, Yüksek Başarım ve Grid Hesaplama Merkezi’nde (TRUBA kaynaklarında) gerçekleştirilmiştir.

Bedri KESKİN
Ankara, Şubat 2026

İÇİNDEKİLER

TEZ ONAY SAYFASI

ETİK.....	i
ÖZET.....	ii
ABSTRACT.....	iii
ÖNSÖZ ve TEŞEKKÜR.....	iv
SİMGELER DİZİNİ.....	vi
ŞEKİLLER DİZİNİ.....	viii
ÇİZELGELER DİZİNİ.....	x
1. GİRİŞ.....	1
2. KAYNAK ÖZETLERİ.....	7
3. MATERYAL ve YÖNTEM.....	12
3.1 Örten Çift Yıldızların Sınıflandırılması.....	12
3.1.1 Roche geometrisine göre sınıflandırma.....	13
3.1.2 Işık eğrisinin biçimine göre sınıflandırma.....	17
3.2 Gözlemsel Veri.....	20
3.2.1 Gaia DR3 arşivi.....	20
3.2.2 Villanova Kepler ve TESS katalogları.....	20
3.3 Makine Öğrenmesi Modelleri.....	21
3.3.1 Çok modlu (multimodal) model.....	22
3.3.2 VGG-19 tabanlı model.....	24
3.3.3 Bir boyutlu CNN modeli.....	25
3.3.4 GMM modeli.....	26
3.4 Yazılım ve Donanım.....	27
4. BULGULAR.....	28
4.1 Çok Modlu (Multimodal) Sınıflandırma.....	28
4.2 Görsel Veri ile Sınıflandırılma.....	41
4.3 Fourier Katsayılarına Göre Sınıflandırma.....	43
4.4 Örten Çift Yıldızların Kümelendirilmesi.....	46
5. TARTIŞMA ve SONUÇ.....	51
KAYNAKLAR.....	58
EKLER.....	63
EK 1 Geliştirilen Makine Öğrenmesi Modelleri.....	64
EK 2 Gaia DR3 Örten Çift Yıldızlarının Makine Öğrenmesi ile Işık Eğrisinin Biçimine Göre Otomatik Olarak EA, EB, EW Sınıflandırılmasına Ait Örnekler.....	76
ÖZGEÇMİŞ.....	79

SİMGELER DİZİNİ

A	Genlik
c	Morfoloji parametresi
C	Geometrik modelin referans seviyesi
d	Minimum derinliği
EA	Algol türü ışık değişimi gösteren örten değişen sistemler
EB	Beta Lyrae türü ışık değişimi gösteren örten değişen sistemler
EW	W UMa türü ışık değişimi gösteren örten değişen sistemler
f	Değme oranı
k	Komşu sayısı
P	Dönem
q	Örten çift yıldızın bileşenlerinin kütle oranı
r	Düzenleştirme parametresi
μ	Gaussyen veya kosinüs bileşenin evresi
σ	Gaussyen bileşenin standart sapması
φ	Yörünge evresi

Kısaltmalar

AGN	Active Galactic Nuclei
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
BIC	Bayesian Information Criterion
BiLSTM	Bidirectional Long Short-Term Memory
C	Contact
CALEB	Catalog and Atlas of Eclipsing Binaries
CDT	Compound Decision Tree
CNN	Convolutional Neural Network
CSS	Catalina Sky Surveys
D	Detached
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DEBCat	Catalogue Of Well-Studied Detached Eclipsing Binaries
E	Ellipsoidal
ESA	European Space Agency
ESAC	European Space Astronomy Centre
Gaia	Gökyüzü tarama sistemi (http://sci.esa.int/gaia)
Gaia DPAC	Gaia Data Processing and Analysis Consortium
Gaia DR3	Gaia Data Release 3
GCVS	General Catalogue of Variable Stars
GMM	Gaussian Mixture Model
GRU	Gated Recurrent Unit
HIPPARCOS	High Precision Parallax Collecting Satellite
HJD	Heliocentric Julian Date

IBVS	Information Bulletin on Variable Stars
KIC	Kepler Input Catalog
k-NN	k-Nearest Neighbors
LSTM	Long Short-Term Memory
LSVC	Large Support Vector Classification
LDA	Linear Discriminant Analysis
mAP	mean Average Precision
ML	Machine Learning
MLP	Multi-Layer Perceptron
NN	Neural Network
OC	Overcontact
OGLE	Optical Gravitational Lensing Experiment
PCA	Principal Component Analysis
PNG	Portable Network Graphics
R-CNN	Region-based CNN
ResNet	Residual Networks
RF	Random Forest
SIMBAD	Astronomik Veri Tabanı (http://simbad.u-strasbg.fr/simbad/)
SD	Semi-detached
SOM	Self Organizing Map
SOS	Single Object Studies
SSD	Single Shot MultiBox Detector
TCE	Threshold Crossing Events
TESS	Transiting Exoplanet Survey Satellite
TRUBA	Türk Ulusal Bilim e-Altyapısı
TIC	TESS Input Catalog
VGG-19	Visual Geometry Group - 19
VVV	VISTA Variables of the Vía Láctea Survey
WUMaCat	W UMa Catalogue
XGBoost	eXtreme Gradient Boosting
YOLO	You Only Look Once
ZTF	Zwicky Transient Facility

ŞEKİLLER DİZİNİ

Şekil 3.1 Her iki bileşenin de Roche loblarının içinde olduğu ayrık sistem	13
Şekil 3.2 Bileşenlerden birinin Roche lobunu doldurduğu yarı-ayrık sistem	14
Şekil 3.3 Bileşenlerin ortak bir zarfla çevrildiği aşırı değen bir sistem	14
Şekil 3.4 Kepler arşivinde bulunan 1572 örten çift yıldızın bir eğri boyunca LLE izdüşümü.....	15
Şekil 3.5 Örnek 15 yıldızın polinom fitleri ve morfoloji parametresine göre sıralanışı .	16
Şekil 3.6 Örten değişen yıldızlara ait örnek ışık eğrileri; (a) Algol, (b) Beta Lyrae, (c) W	19
Şekil 4.1 Gaia DR3 arşivinde bulunan örten çift yıldızların katlanmış ışık eğrisi üzerine çizilen Mowlavi vd. (2023) geometrik modeline göre elde edilen EA, EB, EW sınıflarına ait örnek model grafikleri	31
Şekil 4.2 İki Gaussyen (TWOGAUSSIANS) ile modellenen yıldızların eğitim/doğrulama setinde yapılan eğitimin öğrenme eğrisi	33
Şekil 4.3 İki Gaussyen ve 1. Gaussyen üzerinde bir kosinüs (TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE1) ile modellenen yıldızların eğitim/doğrulama setinde yapılan eğitimin öğrenme eğrisi. Eğrilerin anlamları Şekil 4.2'deki gibidir.....	33
Şekil 4.4 İki Gaussyen ve 2. Gaussyen üzerinde bir kosinüs (TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE2) ile modellenen yıldızların eğitim/doğrulama setinde yapılan eğitimin öğrenme eğrisi. Eğrilerin anlamları Şekil 4.2'deki gibidir.....	34
Şekil 4.5 Bir Gaussyen ve bir kosinüs (ONEGAUSSIAN_WITH_ELLIPSOIDAL) ile modellenen yıldızların eğitim/doğrulama setinde yapılan eğitimin öğrenme eğrisi	34
Şekil 4.6 'TWOGAUSSIANS' ile modellenen örten çiftlerin test verisinin karışıklık matrisi	36
Şekil 4.7 'TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE1' ile modellenen örten çiftlerin test verisinin karışıklık matrisi.....	37
Şekil 4.8 'TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE2' ile modellenen örten çiftlerin test verisinin karışıklık matrisi.....	38
Şekil 4.9 'ONEGAUSSIAN_WITH_ELLIPSOIDAL' ile modellenen örten çiftlerin test verisinin karışıklık matrisi.....	39
Şekil 4.10 Kepler ve TESS arşivlerinden alınan ayrık (sol üst), yarı-ayrık (sağ üst), aşırı değen (sol alt) ve elipsoidal (sağ alt) örnek ışık eğrileri.....	41
Şekil 4.11 Kepler ve TESS arşivlerinde ortak bulunan Gaia DR3 örten çift yıldızların sınıflandırılması sonucu elde edilen karışıklık matrisi	42
Şekil 4.12 Villanova Kepler ışık eğrilerinin Fourier serileri ile modellenmesi	43

Şekil 4.13 Villanova TESS ışık eğrilerinin Fourier ile modellenmesi.....	44
Şekil 4.14 Gaia DR3 ışık eğrilerinin Fourier ile modellenmesi.....	44
Şekil 4.15 Gaia DR3 ışık eğrilerinin Fourier ile sınıflandırılması sonucu elde edilen karışıklık matrisi.....	45
Şekil 4.16 3 kez katlanmış Gaia DR3 ışık eğrilerinin Fourier ile sınıflandırılması sonucu elde edilen karışıklık matrisi.....	46
Şekil 4.17 Fourier a2 ve a4 katsayılarına göre ayırık çiftlerin yarı-ayırık ve değen çiftlerden ayrılması. EA türü çiftler kalın eğrinin üstünde, EB ve EW türü çiftler altında yer alır).....	47
Şekil 4.18 Fourier a1 ve a2 katsayılarına göre yarı-ayırık çiftlerin değen çiftlerden ayrılması. EB türü çiftler kalın çizginin üstünde, EW türü çiftler altında yer alır	47
Şekil 4.19 Kepler örten çift yıldızlarının GMM ile kümelendirilmesi	48
Şekil 4.20 Kepler örten çift yıldızlarının Fourier katsayılarına göre sınıflandırılması ile kümelendirilmesinin karşılaştırılmasını gösteren karışıklık matrisi	49
Şekil 4.21 Kepler örten çift yıldızlarında 0. kümenin dağılımı.....	49
Şekil 4.22 Kepler örten çift yıldızlarında 1. kümenin dağılımı.....	50
Şekil 4.23 Kepler örten çift yıldızlarında 2. kümenin dağılımı.....	50

ÇİZELGELER DİZİNİ

Çizelge 4.1 G-bant ışık eğrilerine fit edilen geometrik model türleri.....	29
Çizelge 4.2 Görsel inceleme ile sınıflandırılan örten çift yıldız sayıları.....	32
Çizelge 4.3 ‘TWOGAUSSIANS’ ile modellenen örten çiftlerin test verisinin sınıflandırma raporu	36
Çizelge 4.4 ‘TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE1’ ile modellenen örten çiftlerin test verisinin sınıflandırma raporu	37
Çizelge 4.5 ‘TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE2’ ile modellenen örten çiftlerin test verisinin sınıflandırma raporu	38
Çizelge 4.6 ‘ONEGAUSSIAN_WITH_ELLIPSOIDAL’ ile modellenen örten çiftlerin test verisinin sınıflandırma raporu	39
Çizelge 4.7 Gaia DR3 arşivinde çok modlu makine öğrenmesi modeli ile otomatik sınıflandırılan örten çift yıldızlarının sayıları ve oranları.....	40
Çizelge 5.1 Bu çalışma ile literatürdeki benzer çalışmaların karşılaştırması.....	53

1. GİRİŞ

Çift yıldızlar, iki bileşen yıldızın kütle çekimi nedeniyle ortak kütle merkezi etrafında Kepler yasalarına göre dolandığı yıldız sistemleridir. Çift yıldızlar gözlemsel olarak keşfedilme yöntemine göre görsel, astrometrik çiftler, tayfsal çiftler ve örten değişenler olmak üzere dört sınıfta incelenirler.

Örten çift yıldızlar bileşen yıldızların ortak kütle merkezi etrafındaki yörünge hareketi esnasında gözlemcinin bakış doğrultusuna göre dönemli olarak birbirlerini örtmeleri sonucu ışık değişimi gösteren yıldızlardır. Bir örten çift hakkında en temel bilgi kaynağı onun fotometrik ışık eğrisi verisidir (Çokina vd. 2021). Işık eğrisinin dönem, şekil, minimum derinliklerinin analizi ile örten çift yıldızın yarıçapları, yüzey sıcaklıkları, parlaklıkları, yörünge eğimleri ve yarı büyük eksen uzunlukları gibi çift sistemlerin temel parametreleri elde edilir. Radyal (dikine) hız verisinin ışık eğrisi ile beraber çözümü ile bileşen yıldızların ayrı ayrı kütle ve yoğunlukları belirlenebilir. Bu da örten çiftleri yıldız astrofizikinde yıldız evrimi teorilerini (Feiden ve Chaboyer 2012), uzaklıkları (Stassun ve Torres 2016), asterosismik teorileri (Gaulme vd. 2016) ve kütleçekim teorilerini (Baroch vd. 2021) test etmek için iyi laboratuvarlar haline getirir (Shan vd. 2025).

Çift yıldızların evrimi Roche geometrisi bağlamında ele alınır ve bu nedenle çift yıldız bileşenlerinin evrimlerinin hangi aşamasında olduklarını belirlemek için Roche şişimlerini doldurup doldurmadıkları önem kazanır. Çift yıldızların Roche potansiyel yüzeylerine göre fiziksel sınıflandırması ilk kez Kopal (1959) tarafından önerilmiştir. Çift yıldızlar bu bağlamda ayırık (ing. detached - D), yarı-ayırık (ing. semi-detached - SD), değen (ing. contact - C) ve aşırı değen (ing. overcontact - OC) olarak sınıflandırılır. Örten çift yıldızların fotometrik gözlemleri sonucu elde edilen ışık eğrilerinin morfolojik olarak, türün prototip yıldızının ışık eğrisi morfolojisinden hareketle, Algol (EA), Beta Lyrae (EB) ve W UMa (EW) sınıflandırılması, fiziksel özelliklere göre yapılan “Kopal sınıflandırmasıyla” benzerlikler taşır. Işık eğrisi Algol (EA) olarak sınıflandırılan çift yıldızlar genellikle ayırık sistemlerken, çoğu Beta Lyrae (EB) yarı-ayırık sistemdir. Değen ve aşırı değen sistemlerin ışık eğrileri ise W UMa (EW) türündedir.

Modern astronomik gökyüzü tarama projeleri örten çift yıldızların tespiti için benzeri görülmemiş bir kaynak sağlamıştır. Ancak bu devasa veri setlerinin manuel olarak analiz edilmesi imkansız hale gelmiş, bu sistemlerin algılanması, karakterize edilmesi ve sınıflandırılması için otomatik yöntemleri bir zorunluluk haline getirmiştir.

Gaia¹, Kepler², Transiting Exoplanet Survey Satellite (TESS)³ ve Optical Gravitational Lensing Experiment (OGLE)⁴ gibi büyük ölçekli gökyüzü tarama projelerinin sağladığı devasa veri setleri milyarlarca gök cisminin ışık eğrilerini ve diğer özelliklerini içeren bu veri hacmi, özellikle değişen yıldızların ve örten çift sistemlerin incelenmesinde yeni bir çağır açmıştır. Ancak bu veri zenginliği, geleneksel manuel analiz yöntemlerini yetersiz kılmakta ve bu sistemlerin tespiti, sınıflandırılması ve fiziksel parametrelerinin belirlenmesi için güçlü, otomatik ve ölçeklenebilir yöntemlere olan ihtiyacı zorunlu hale getirmektedir. Bu noktada, makine öğrenmesi (ing. Machine Learning – ML) ve istatistiksel analiz, bu büyük veri setlerinden anlamlı astrofiziksel bilgiler çıkarmanın anahtarı olarak öne çıkmaktadır.

Yer tabanlı ve uzay teleskoplarıyla gözlemlere dayalı fotometrik araştırmalar, yeni örten çiftlerin keşfini önemli ölçüde ilerletmiş ve bu sistemlere yeni bir pencere açmıştır. Örneğin Papageorgiou vd. (2018) Catalina Sky Survey (CSS)⁵ verilerini kullanarak kuzey gökyüzündeki 4680 örten çifti içeren bir katalog sunmuştur. Chen vd. (2020) Zwicky Transient Facility (ZTF)⁶ 2. Veri Yayını'ndaki 781 binden fazla periyodik değişen yıldızdan oluşan bir katalog sunmaktadır; bunların 350 bini örten çift olarak sınıflandırılmıştır. Uzay tabanlı Kepler görevi (Borucki 2016), 105 derece karelik bir gökyüzü alanında yaklaşık 2 milyon yıldız yaklaşık 4 yıl boyunca sürekli olarak gözlemlenmiş ve 2800'den fazla tutulmalı ve elipsoidal (ing. ellipsoidal – E) ikili sistem bulmuştur (Kirk vd. 2016). Kepler'in halefi olan TESS, Güneş Sistemi dışı gezegenleri (ötegezegenler) aramak için neredeyse tüm gökyüzünü çok daha geniş bir görüş alanıyla

¹ <https://gea.esac.esa.int/archive/>

² <https://archive.stsci.edu/missions-and-data/kepler>

³ <https://archive.stsci.edu/missions-and-data/tess>

⁴ <https://ogle.astrouw.edu.pl/>

⁵ <https://catalina.lpl.arizona.edu/>

⁶ <https://www.ztf.caltech.edu/>

taramaktadır (Ricker vd. 2014) ve verilerinin örten çift arařtırmaları için paha biçilmez olduđu kanıtlanmıřtır. Prša vd. (2022), ilk iki yıldaki 2 dakikalık ıřık eğrisi verilerini (Sektör 1-26) kullanarak 4584 EB tespit etmiřtir. Ayrıca IJspeert vd. (2021) aynı sektörlerdeki 30 dakikalık ıřık eğrisi verilerini analiz etmiř ve 3155 O-B-A tayf türünden örten çift adayı bulmuřtur. Bu çalıřmalar TESS verilerini kullanarak yeni örten çiftler bulmanın gücünü vurgulamaktadır (Shan vd. 2025).

Son yıllarda, yer tabanlı ve uzay tabanlı birçok arařtırma, neredeyse tüm gökyüzünü kapsayan büyük hacimli fotometrik gözlemler üretmiřtir. Bu gözlemler, astronomik arařtırmaların hassasiyetini ve zamansal kapsamını artırmıř ve astronomik veri analizinin iř akıřlarını önemli ölçüde dönüřtürmüřtür. Deđiřen yıldızların çok sayıda farklı ıřık eğrisinin analizinde yeni zorluklar ortaya çıkmıřtır. Büyük veri setlerini tek tek analiz etmek, parametrelerini (periyot, tutulma derinliđi vb.) çıkarmak ve sınıflandırmak için otomatik boru hatları (ing. pipelines) ve algoritmalar řarttır. Otomatik sınıflandırma, evrimin farklı ařamalarındaki sistemlerin istatistiksel dađılımını anlamamızı da sađlar. Deđiřen yıldızların sınıflara ayrılması fotometrik analizde önemli bir ařamadır (Parimucha vd. 2025). Örten çift yıldızların dođru sınıflandırılması yıldız evrimini anlamada kritik öneme sahiptir. Yıldız evrim modelleri bu sistemlerden elde edilen gerçek ölçümlerle test ve kalibre edilir. Örneđin Torres vd. (2010) ayrıık örten çift yıldızların, yıldızların kütle ve yarıçap gibi temel parametrelerini dođrudan ve yüksek hassasiyetle ölçebildiđimiz tek güvenilir astrofiziksel "laboratuvarlar" olduđunu vurgular.

Astronomik veri analizinde derin öğrenme gibi modern yapay zeka tekniklerinin kullanılması büyük veri setlerinin iřlenmesi sorununa yenilikçi ve etkili bir çözüml sunmaktadır. Literatürde makine öğrenimi tekniklerinin kullanıldıđı çok sayıda çalıřma vardır. Örneđin Papageorgiou vd. (2019), CSS tarafından tespit edilen 2200'den fazla örten çiftin fiziksel özelliklerini analiz etmek için Yapay Sinir Ađını (ing. Artificial Neural Network - ANN) kullanmıřtır. Shallue ve Vanderburg (2018) Kepler verilerinde yeni geçiř yapan ötegezegenleri tanımlamak için bir makine öğrenimi yaklařımı sunmuřtur. Osborn vd. (2020) TESS tarafından tespit edilen geçiř sinyallerini sınıflandırmak için Evriřimli Sinir Ađı (ing. Convolutional Neural Network – CNN) uygulayarak aday dođrulamanın verimliliđini artırmıřtır. Ofman vd. (2022) ötegezegen

geçiş sinyali araştırmasında eşik geçiş olaylarına (ing. Threshold Crossing Events - TCE) uygulanan yeni bir makine öğrenimi yöntemi geliştirmiş ve çalışma üç yeni ötegezegen adayının keşfiyle sonuçlanmıştır (Shan vd. 2025).

Çift yıldızlara özgü olarak, araştırmacılar çift sistemlerin ışık eğrilerini tespit etmek, modellemek ve sınıflandırmak için makine öğrenimi ve derin öğrenme tekniklerini kullanmışlardır. Wyrzykowski vd. (2003) OGLE verilerini kullanarak yapay sinir ağı yaklaşımı ile Büyük Macellan Bulutu'ndaki 2580 çift sistemi tanımlamışlardır. Prša vd. (2008) bazı fiziksel parametreleri belirlemek için 33235 ayrık örten çift sistemin ışık eğrisi verisini kullanarak eğitilmiş bir yapay sinir ağı sunmuşlardır. Kochoska vd. (2020) çeşitli modelleme tekniklerini ele almış ve çift sistemlerin başlangıç parametrelerinin makine öğrenimi teknikleri yardımıyla tahmin edilebileceği sonucuna varmışlardır. Makine öğrenimi yöntemleri kullanılarak değişen yıldızların görüntü tabanlı bir sınıflandırması Szklenár vd. (2020) tarafından gerçekleştirilmiştir. Čokina vd. (2021) tarafından yapılan iki sınıflı morfolojik sınıflandırma, sentetik ışık eğrisi verileri ile farklı derin öğrenme yöntemlerine odaklanmıştır. Bódi ve Hajdu (2021) OGLE çift yıldızlarının morfolojilerini sınıflandırmak için Locally Linear Embedding (LLE) yöntemini kullanan bir makine öğrenimi algoritması oluşturmuştur. Szklenár vd. (2022) OGLE-III verileriyle çok girişli (ing. multiple-input) bir sinir ağı kullanarak değişen yıldızları görsel özelliklerine göre sınıflandırmıştır. Healy vd. (2024) tarafından ZTF verilerine dayalı makine öğrenimi algoritmaları kullanılarak 60 binden fazla örten çift yıldızın sınıflandırması yapılmıştır (Ulaş vd. 2025).

Örten çift yıldızlar astrofizikte yıldızların temel parametrelerini yüksek doğrulukla belirlemek, yıldız yapısı ve evrim modellerini test etmek ve hatta yakın galaksilerin uzaklıklarını belirlemek gibi önemli bir özelliğe sahiptir (Sarro vd. 2006). Örneğin yıldızların kütlelerini model bağımsız olarak belirlemenin tek yolu çift çizgili örten çift yıldızların ışık ve dikine hız eğrilerinin analizidir. Kepler, TESS, Gaia gibi gökyüzü tarama sistemlerinden muazzam miktarda ışık eğrisi verisi toplanmaktadır. Bu devasa veri akışı zorunlu olarak otomatikleştirilmiş ve akıllı işleme yöntemlerini gerektirmektedir; zira geleneksel manuel prosedürlerle bu veri hacminin üstesinden gelmek imkansızdır. Bu nedenle, makine öğrenimi ve yapay zeka teknikleri, desen

tanıma, sınıflandırma ve bilgi çıkarımı (ing. inference) gibi görevlerdeki üstün performansları sayesinde, gelecekteki örten çift yıldız araştırmaları yoluyla yıldız astrofiziğine yapılacak büyük katkılar için de kritik araçlar haline gelmiştir.

Geniş ölçekli gökyüzü tarama projeleri tarafından üretilen yüksek hacimli veri setleri, manuel analiz kapasitesinin çok ötesinde bir boyutta olduğundan, astrofiziksel açıdan kritik öneme sahip olan örten çift yıldız sistemlerinin tespiti ve karakterizasyonu için otomatize edilmiş metodolojilere ihtiyaç duyulmaktadır. Bu tez çalışmasında makine öğrenmesi teknikleri ile örten çift yıldızların ışık eğrilerinin morfolojik sınıflandırması hedeflenmiştir.

Bu hedef doğrultusunda Gaia DR3, Kepler ve TESS arşivlerinden alınan örten çift yıldızların akı ölçümlerinin hem denetimli öğrenme ile sınıflandırılması hem de denetimsiz öğrenme ile kümelendirilmesi yapıldı. Gaia DR3 arşivinden alınan örten çift yıldızlar, ışık eğrilerinin görselleri ve sayısal verileri olmak üzere iki şekilde ifade edildi. Bu görev için iki farklı veri türünü (görsel ve sayısal veri) aynı anda işleyebilen çok modlu (multimodal) bir derin öğrenme modeli kullanıldı. Model, görüntü verilerinden özellik çıkarmak için bir CNN ve sayısal verilerden anlamlı desenler öğrenmek için çok katmanlı bir algılayıcı (ing. Multi-Layer Perceptron - MLP) içerir. Bu iki yapıdan elde edilen bilgiler birleştirilerek nihai sınıflandırma yapıldı. Kepler ve TESS arşivinden örten çift yıldızların görselleri ile Visual Geometry Group - 19 (VGG-19) (Simonyan ve Zisserman 2014) modeli eğitildi ve eğitilen model ile Gaia DR3 örten çiftleri sınıflandırıldı. Gaia DR3, Kepler ve TESS arşivlerinden alınan örten çift yıldızların ışık eğrileri Fourier serisi ile modellendi ve Fourier model katsayıları kullanılarak sınıflandırma yapıldı. Kepler örten çiftlerinin denetimsiz öğrenme ile dağılımı araştırıldı.

2. Bölüm’de literatürdeki benzer çalışmalardan örnekler verilmiştir. 3. Bölüm’de örten çift yıldızlarını sınıflandırma yöntemleri, bu çalışmada kullanılan Gaia DR3, Kepler ve TESS arşivlerinden alınan ışık eğrisi verilerinin detayları, denetimli-denetimsiz makine öğrenmesi modellerinin detayları ve geliştirilen yazılım ile kullanılan donanım hakkında bilgi verilmiştir. 4. Bölüm’de hem görsel veri hem de sayısal veri ile yapılan çok modlu sınıflandırma, sadece görsel veri ile hazır eğitilmiş VGG-19 modeli ile yapılan

sınıflandırma, Fourier katsayılarına göre yapılan sınıflandırma ve Fourier katsayılarına göre Kepler arşivindeki örten çiftlerin dağılımının araştırıldığı kümelendirme anlatılmıştır. 5. Bölüm’de elde edilen sonuçlar değerlendirilmiştir.

2. KAYNAK ÖZETLERİ

Bu bölümde örten çift yıldızların ve diğer değişen cisimlerin sınıflandırılmasına yönelik makine öğrenmesi uygulamalarındaki son gelişmeleri özetleyen kapsamlı bir literatür taraması sunulmaktadır. Bu bağlamda, Principal Component Analysis (PCA) ve Random Forest (RF) gibi klasik yöntemlerden, CNN, Transformatörler ve You Only Look Once (YOLO) gibi modern derin öğrenme ve bilgisayarla görü mimarilerine kadar geniş bir metodolojik yelpaze incelenmektedir.

Süveges vd. (2017) CALEB, HIPPARCOS ve Kepler veri setlerini kullanmıştır. Hem ışık eğrisi morfolojisine göre (EA, EB, EW) hem de sistem morfolojisine göre (Detached, Semi Detached, Overcontact) sınıflandırma yapmıştır. PCA, Linear Discriminant Analysis (LDA), RF, Self Organizing Map (SOM) yöntemlerini kullanmıştır. Gürültüyü gidermek ve interpolate etmek için ışık eğrilerine ilk önce çift Gaussyen ve spline fit yapmıştır. PCA ile değişenliği belirten en önemli elemanları seçmiştir. HIPPARCOS veri setinde ışık eğrisi morfolojisine göre yapılan sınıflandırmada %5 gibi düşük bir genel hata oranı elde etmiştir. CALEB veri setinde sistem morfolojisine göre sınıflandırmada %10.5 genel hata oranı elde etmiştir. Sistem morfolojisine göre yapılan sınıflandırmada hata oranının yüksek çıkmasını fiziksel olarak farklı sistemlerin çok benzer ışık eğrileri üretebilmesi şeklinde açıklamıştır. CALEB veri setinde eğitilen modelleri Kepler çift yıldızlarına uygulamıştır. Buradaki temel amaç, bir veri setiyle (CALEB) eğitilmiş bir modelin başka bir veri setinde (Kepler) ne kadar iyi performans gösterdiğini test etmektir. Çalışmada Kepler veri setine uygulanan model için spesifik bir sayısal hata oranı belirtilmemiştir. Ancak CALEB verisiyle eğitilen RF modelinin Kepler verilerine uygulandığında elde ettiği tahminlerin, SOM tabanlı sınıflandırıcıya göre daha isabetli olduğu ifade edilmiştir.

Bódi ve Hajdu (2021) OGLE veri setini kullanmıştır. Işık eğrilerindeki aykırı değerler Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algoritması ile temizlenmiştir. Temizlenmiş ışık eğrilerine polinom fitleri yapılmıştır. Her bir yıldızın yüksek boyutlu (1000 noktadan oluşan) ışık eğrisi modeli, iki aşamalı bir Locally Linear Embedding (LLE) (Roweis ve Saul 2000) süreciyle önce 3 boyutlu, sonra da 2 boyutlu

bir parametre uzayına indirgenmiştir. Boyut indirgeme benzer ışık eğrisine sahip yıldızların bu yeni uzayda birbirine yakın kalmasını sağlar. Ardından her bir yıldıza 0 ile 1 arasında sistemin ayrıklığını belirten “morfoloji parametresi (c)” parametresi atanmıştır. Sistem morfolojisine göre sınıflandırma yapılmıştır. LLE denetimsiz öğrenme yöntemi kullanılmıştır. LLE algoritmasının komşu sayısı (k) ve düzenleme parametresi (τ) deneysel olarak ve sonuçların görsel olarak karşılaştırılmasıyla seçilmiştir. Bütün yıldızlar için hesaplanan morfoloji parametrelerini içeren bir katalog ve tüm bu süreci otomatikleştiren açık kaynaklı bir yazılım paketi sunulmuştur.

Čokina vd. (2021) ML modeli eğitimi için yaklaşık 500 bin adet sentetik veri üretmiştir. Sistem morfolojisine göre sınıflandırma yapılmıştır. Örtlen çift yıldızları "ayrık" ve "değen" olmak üzere iki ana sınıfa ayırmayı hedefleyen bir sınıflandırıcı geliştirilmiştir. Yarı-ayrık sistemler bu sınıflandırmada ayrık sistemler olarak değerlendirilmiştir. Doğrulama için literatürden 100 tane ışık eğrisi kullanılmıştır. CNN, Long Short-Term Memory (LSTM) ve Kapılı Tekrarlayan Birim (ing. Gated Repeating Unit - GRU) gibi farklı derin öğrenme mimarileri test edilmiştir. En yüksek performansı çift yönlü LSTM (ing. Bidirectional LSTM - BiLSTM) ile bir boyutlu CNN'nin bir kombinasyonunu kullanan "Model 6" adlı mimari göstermiştir. Yarı-ayrık sistemler ayrık olarak kabul edildiğinde değerlendirme setinde %98'lik bir doğruluk (ing. accuracy) oranına ulaşılmıştır. Yarı-ayrık sistemler veri setinden çıkarıldığında ise %100 doğruluk elde edilmiştir. Çok modlu (multimodal) makine öğrenmesi modellerinin örtlen çiftlerin sınıflandırılmasında yüksek doğruluğa ulaşabildiği gösterilmiştir.

Daza-Perilla vd. (2023) VISTA Variables of the Vía Láctea Survey (VVV) veri setindeki örtlen değişenleri ayrık, yarı-ayrık ve değen alt sınıflara ayırmıştır. 100 adet örtlen çift yıldız (50 ayrık, 13 yarı-ayrık, 37 değen) kullanmıştır. Bu yıldızların ışık eğrilerinden FEETS adlı bir yazılım paketi kullanılarak ortalama, eğim, periyot gibi istatistiksel ve şekilsel özellikleri içeren 69 adet öznitelik çıkarılmıştır. En önemli öznitelikleri belirlemek amacıyla varyans analizi (ing. analysis of variance - ANOVA), karşılıklı bilgi (ing. mutual information) ve ki-kare istatistiksel yöntemleri kullanılmış ve 35 öz niteliğin yeterli olduğu tespit edilmiştir. Bileşik karar ağacı (ing. Compound Decision Tree - CDT) olarak adlandırılan yeni bir model geliştirilmiştir. Bu model karar ağacı (ing. Decision

Tree - DT), RF, k-en yakın komşu (ing. k-nearest neighbors - k-NN) ve destek vektör makineleri (ing. Support Vector Machine - SVM) algoritmalarını bir araya getiren ve oylama sistemiyle çalışan bir modeldir. CDT modeli diğer tekil modellere (DT, RF gibi) göre belirgin şekilde daha iyi performans gösterdiği ve yaklaşık %80'lik bir görsel doğrulama isabet oranı elde ettiği sonucuna varılmıştır.

Eyer vd. (2023) tarafından fotometrik, astrometrik ve tayfsal veriler kullanılarak bütün gökyüzünü kapsayan en büyük değişen analizi yapılmıştır. Gaia DR3'de yayınlanan 1,8 milyar cismin değişenlik analizi ile fotometrik ve tayfsal zaman serilerinin analizi yapılmıştır. Değişen cisimlerin karakterize edilmesi ve sınıflandırılması için istatistiksel ve makine öğrenmesi metotları kullanılmıştır. Gaia DR3'de 10.5 milyon cisim değişen olarak sınıflandırılmıştır. Bunlardan 9.5 milyonu değişen yıldız, 1 milyonu aktif galaksi çekirdeği (AGN)/kuazar olarak sınıflandırılmıştır.

Gavras vd. (2023) literatürde bilinen cisimlerden kapsamlı bir set oluşturmuş ve Gaia DR3 ile çapraz eşleştirmesini yapmıştır. Literatürdeki cisimlerin Gaia eşlerini bulmak için astrometrik ve fotometrik yöntemler kullanmıştır. Literatürdeki 152 farklı kaynağın Gaia DR3 ile çapraz eşleşmesi yapılmıştır. Sonuçta 112 tür sabit, galaksi ve değişen toplam 7.8 milyon cisim bulunan katalog elde edilmiştir. Bunların 1.2 milyonu sabit, 1.7 milyonu galaksi, 4.9 milyonu değişen cisimdir. Yüzden fazla değişen sınıf/alt sınıf içermektedir.

Rimoldini vd. (2023) tarafından Gaia DR3 arşivindeki değişen cisimler makine öğrenmesi yöntemiyle sınıflandırılmıştır. eXtreme Gradient Boosting (XGBoost) ve RF yöntemleri kullanılmıştır. 22 türde yaklaşık 9 milyon değişen yıldız, binlerce süpernova, 1 milyon aktif galaksi çekirdeği ve 2.5 milyon galaksi olmak üzere toplam 12.4 milyon cisim sınıflandırılmıştır. 11 değişen tür SOS (Single Object Studies) adı altında ayrıca çalışılmıştır. Örneğin Mowlavi vd. (2023) bu çalışma sonucu örten değişen olarak sınıflandırılan yıldızları ayrıca çalışmıştır.

Mowlavi vd. (2023) Gaia DR3 arşivindeki 2.1 milyon örten değişen olarak sınıflandırılan yıldızları ele almıştır. G-band ışık eğrilerine örten değişen kriterlerine göre filtre

uygulanmıştır. Yörünge dönemini (P) bulmak için 3 tane dönem arama metodu temizlenmiş G ışık eğrilerine uygulamıştır. Her bir örten değişen adayının G-band ışık eğrisine bir veya iki adet Gaussyen fit ve bir adet kosinüs fiti yapmıştır. Gaussyen fit minimum zamanları, kosinüs fiti ise ışık eğrisindeki elipsoidal benzeri değişimleri modellemek içindir. Yörünge dönemi (P) ve geometrik modelin en iyi kombinasyonu Bayesian model karşılaştırma ile seçilmiştir. Işık eğrilerine seçilen modelin kalitesini belirten bir sıralama değeri (ing. global rank) verilmiştir. Katalog, dönemi 0.2 günden uzun sistemlerle sınırlandırılmıştır. Literatür ile yapılan 600 bin çapraz eşleşmenin 530 bin tanesi yine örten değişen olarak sınıflandırılmış ve bunların %93'ünün dönemi Gaia dönemleri ile uyumludur.

Parimucha vd. (2025) örten çift yıldızların ışık eğrilerini sınıflandırmak için bilgisayarla görü yöntemlerini kullanmıştır. Sınıflandırmayı yapmak için evrişimli sinir ağlarına (ResNet50) ve görüş transformatörlerine (vision transformers) dayalı önceden eğitilmiş derin öğrenme modelleri kullanmıştır. Modeli eğtmek, değerlendirme ve test etmek için fiziksel modellere dayalı sentetik veriler üretmiştir. Modelin gerçek gözlem koşullarına uyum sağlaması için sentetik verilere gürültü (ing. noise) ve aykırı değerler (ing. outliers) eklemiştir. Yıldızları lekeli ve lekesiz olarak sınıflandırabilmek için bazılarını leke de eklemiştir. Geliştirilen modelin nihai başarımı ise literatürde iyi tanımlanmış kataloglardan (OGLE, DEBCat, WUMaCat) alınan gerçek gözlem verileri üzerinde test edilmiştir. Çalışmanın temel yeniliklerinden biri, yıldızların zamanla parlaklık değişimini gösteren ışık eğrilerinin, modelin genelleme yeteneğini artırmak ve ezberlemeyi (overfitting) azaltmak amacıyla kutupsal koordinatlar ve altıgen gruplama (hexbin) tekniğiyle özel iki boyutlu görüntülere dönüştürülmesidir. Sınıflandırma hiyerarşik bir yaklaşımla iki aşamada gerçekleştirilmiştir: ilk aşamada sistemler "ayrık" ve "aşırı degen" olarak iki ana gruba ayrılmış, ikinci aşamada ise bu gruplar içinde yıldız lekelerinin olup olmadığı tespit edilmeye çalışılmıştır. Çalışmanın sonuçları ana sınıflandırma görevinin oldukça başarılı olduğunu göstermiştir. Geliştirilen modeller sentetik verilerle yapılan testlerde %96'nın üzerinde, OGLE, DEBCat ve WUMaCat gibi büyük astronomi kataloglarından elde edilen gerçek gözlemsel veriler üzerinde %94 ile %100 arasında yüksek bir doğruluk oranıyla ayrık ve aşırı degen sistemleri ayırt edebilmiştir. Ancak yıldız lekelerini tespit etmeyi amaçlayan ikincil görevde modellerin performansı oldukça

düşük kalmış ve güvenilir sonuçlar üretememiştir. Bu durum, geliştirilen yöntemin çift yıldızların temel morfolojik sınıflandırması için büyük potansiyele sahip olduğunu, ancak yıldız lekelerinin tespiti konusunda başarısız olduğu ve bu alanda daha fazla araştırmaya ihtiyaç duyulduğunu ortaya koymuştur.

Ulaş vd. (2025) örten çift yıldızların ışık eğrilerinde salınım benzeri yapıları tespit etmek için CNN tabanlı yöntemleri kullanmıştır. Araştırmanın nihai hedefi örten çift yıldız sistemlerindeki zonklayan (ing. pulsating) bileşenlerin otomatik olarak saptanmasını sağlayacak sağlam bir çerçeve oluşturmaktır. Single Shot MultiBox Detector (SSD), Faster Region-based CNN (Faster R-CNN), You Only Look Once (YOLO) ve EfficientDet gibi önceden eğitilmiş (ing. pre-trained) modellerin yanı sıra sıfırdan oluşturulmuş özel bir model de kullanmıştır. ML modellerini eğitmek için hem sentetik veri üretilmiş hem de TESS arşivinde bilinen zonklayan bileşenli örten çift yıldız verileri kullanılmıştır. Işık eğrilerinde salınım ve minimum noktalarının yerleri el ile etiketlenmiştir. Eğitilen modellerin performansı Kepler verileri üzerinde test edilmiştir. Önceden eğitilmiş modellerin hedeflenen desenleri tespit etmede yüksek doğruluk ve güvenilirlik sergilediği gösterilmiştir. Faster R-CNN ve YOLO modelleri %99 'u aşan ortalama kesinlik (ing. mean average precision - mAP) değeriyle üstün bir performans göstermiştir. SSD modeli ise biraz daha düşük bir performansa (%97 mAP) sahip olmasına rağmen en hızlı model olmuştur. Sıfırdan geliştirilen özel model ise önceden eğitilmiş modellere kıyasla oldukça zayıf bir performans sergilemiştir.

Gerçekleştirilen literatür taraması ML tekniklerinin örten çift yıldızların sınıflandırılmasında kullanıldığını göstermektedir. Ancak çalışmaların bir kısmı örten çift yıldızları ayırık/değen şeklinde sınıflandırdığı, fiziksel olarak geçiş evresini temsil eden yarı-ayırık sistemleri diğer sınıflara dahil ederek genelleştirdiği görülmüştür. Ayrıca, çalışmaların bir kaç bin yıldızdan oluşan veri setleri üzerinde yapıldığı, sınıflandırma sonucu bulunan istatistiksel dağılımlar arasında fikir birliğinin oluşmadığı görülmüştür. Bu tez çalışmasında çift sistemlerin hem fiziksel (ayırık, yarı-ayırık, değen, elipsoidal) hem de ışık eğrisi biçimine göre (Algol, Beta Lyrae, W UMa) hassas bir şekilde sınıflandırılması ve çok sayıda veri içeren arşivler (Gaia DR3, TESS, Kepler) kullanılarak bulunan sonuçların literatür için referans teşkil etmesi amaçlanmıştır.

3. MATERYAL ve YÖNTEM

Bu tez çalışmasının temel amacı örten çift yıldızların makine öğrenmesi teknikleri ile sınıflandırılmasıdır. Örten çift yıldızların sınıflandırılması;

- Geleneksel olarak ışık eğrisinin biçimine göre EA, EB, EW
- Morfoloji parametresine göre Ayrık, Yarı-ayrık, Değen, Aşırı Değen, Elipsoidal

olmak üzere iki şekilde yapılabilir. Bu tez çalışmasında her ikisi de ele alındı. ESA'nın Gaia DR3 arşivi ile NASA'nın Kepler ve TESS arşivlerindeki ışık eğrisi verileri kullanıldı. Verilerin bir kısmı görsel inceleme ile etiketlenerek denetimli öğrenme ile büyük miktarda verinin otomatik sınıflandırılması yapıldı. Ayrıca Kepler verisinde denetimsiz öğrenme ile örten çiftlerin dağılımı incelendi. Bu incelemeler ve verinin hazırlanması için gereken tüm bilgisayar kodları Python⁷ dilinde yazıldı. Yazılan kodlar ihtiyaca göre Türk Ulusal Bilim e-Altyapısı (TRUBA)⁸ sisteminde veya kişisel bilgisayarlarda çalıştırıldı.

3.1 Örten Çift Yıldızların Sınıflandırılması

Örten çift yıldızları sınıflandırmak çift yıldızların evrimsel yollarını anlamak için bir başlangıç noktasıdır. Işık eğrisi analizleri dikine hız ölçümleri ile birleştirilerek bileşen yıldızların ayrı ayrı kütle, yarıçap, sıcaklık gibi fiziksel parametreleri belirlenmektedir.

Makine öğrenmesi ile ışık eğrilerini sınıflandırmada ışık eğrisi verisinin nasıl temsil edileceği temel bir karardır. Işık eğrisi iki şekilde temsil edilebilir: zaman serisi vektörleri ve ışık eğrisi görselleri. Işık eğrisi verisini zaman serisi vektörleri olarak işlemek yaygındır ancak Gaia gibi seyrek ve düzensiz örneklenmiş veriye sahip veri setlerinde önemli zorluklar çıkarır (Parimucha 2025).

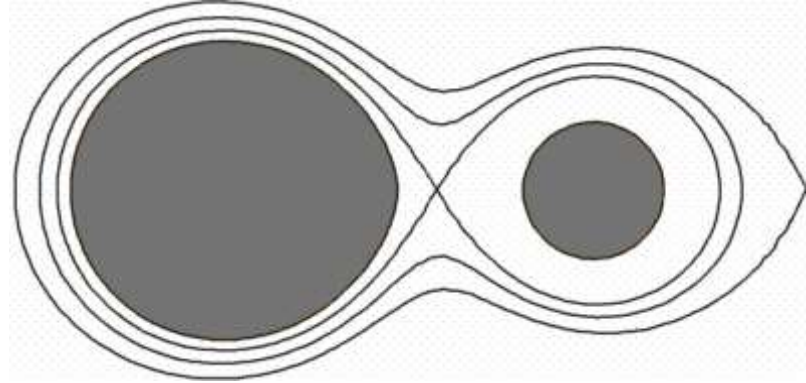
⁷ <https://www.python.org/>

⁸ <https://www.truba.gov.tr/>

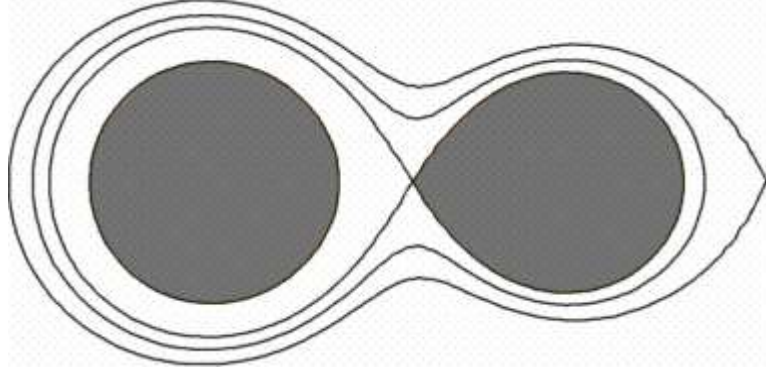
3.1.1 Roche geometrisine göre sınıflandırma

Kopal (1959) Roche modeline göre bir sınıflandırma geliştirmiştir. Roche modeli çift sistemdeki her bir yıldızın etrafında, kütle çekim kuvvetlerinin baskın olduğu ve maddenin o yıldıza bağlı olduğu bir uzay bölgesi tanımlar. Wilson (1979) genelleştirilmiş Roche modelini geliştirmiş ve bileşenlerin eliptik yörüngesini ve eşzamanlı olmayan dönüşünü dikkate almıştır. Dairesel yörünge ve senkron dönme için 3 olasılık vardır (Parimucha 2025):

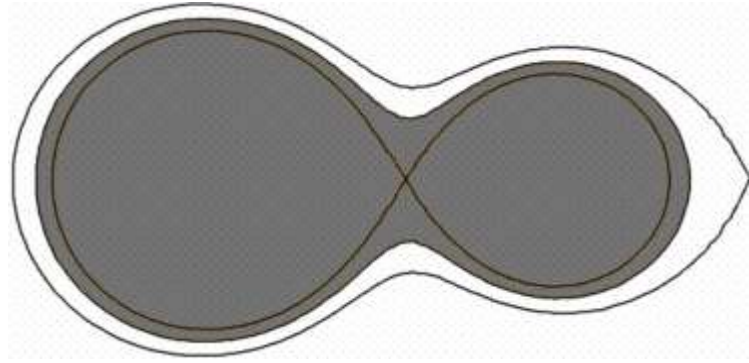
- Ayrık: Her iki bileşen de Roche loblarının içindedir. Bileşenler Roche loblarına kıyasla küçük olsalar bile, şekilleri kürelere çok yakındır (Şekil 3.1).
- Yarı-ayrık: Bileşenlerden biri lobunu doldurur ve kütle iç Lagrangian noktası L_1 aracılığıyla diğer bileşene aktarılabilir (Şekil 3.2).
- Değen veya Aşırı Değen: Her iki bileşen Roche şişimlerini doldurmuş ise "değen", bileşenler bu sınırları aşarak eşpotansiyel yüzeye sahip bir dış zarf (ing. common envelope) içinde ise "aşırı değen" sistemler olarak adlandırılır (Şekil 3.3).



Şekil 3.1 Her iki bileşenin de Roche loblarının içinde olduğu ayrık sistem (Özuyar 2007)



Şekil 3.2 Bileşenlerden birinin Roche lobunu doldurduğu yarı-ayrık sistem (Özuyar 2007)

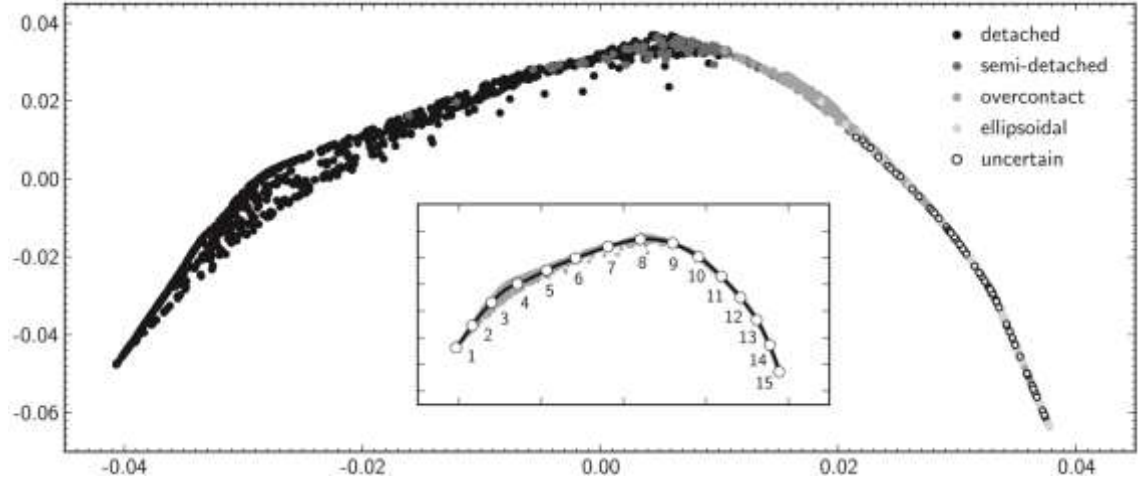


Şekil 3.3 Bileşenlerin ortak bir zarfla çevrildiği aşırı değen bir sistem (Özuyar 2007)

Bir sistemi ayrık, yarı-ayrık veya değen olarak sınıflandırmak, o sistemin evrimsel durumu (kütle transferi olup olmadığı, ortak bir zarf paylaşp paylaşmadıkları) hakkında doğrudan bilgi verir. Fiziksel modellemede en kritik ayrım, çift sistemde bileşen yıldızların ayrık olup iki ayrı potansiyelle (Ω_1 ve Ω_2) veya ortak zarf içinde bulunup tek bir potansiyelle ($\Omega_1 = \Omega_2 = \Omega$) ifade edilebilmeleridir.

Matijević vd. (2012) örten çift yıldızların sınıflandırılması için yeni bir yöntem sunmuştur. Kepler Uzay Teleskobu tarafından gözlenen 2165 örten çift yıldızı “polyfit” adı verilen bir polinom zinciri ile modelledi. Aykırı değerler (gürültülü, bünyesel değişkenlik gösteren veya kötü işlenmiş ışık eğrileri) çıkarıldıktan sonra kalan 1572 ışık eğrisi modelini 1000 noktada örnekledi ve Locally Linear Embedding (LLE) yöntemi ile boyut indirgedi. LLE yöntemi yüksek boyutlu veri setlerindeki noktalar arasındaki yerel geometrik ilişkileri koruyarak veriyi daha düşük boyutlu bir uzaya yansıtır. LLE analizi

sonucunda ışık eğrilerinin iki boyutlu bir uzayda belirgin tek boyutlu bir eğri boyunca dağıldığı görülmüştür (Şekil 3.4).

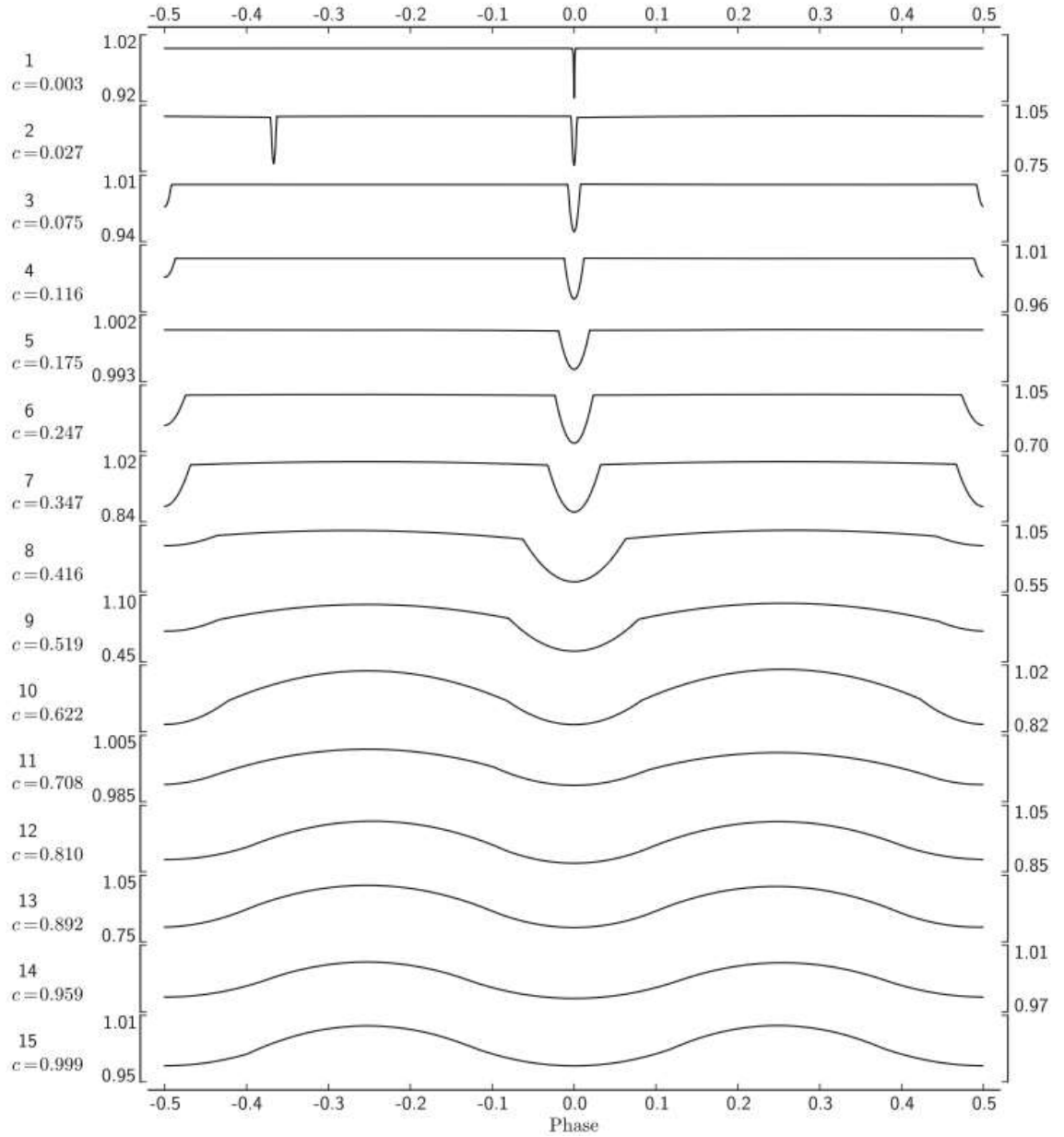


Şekil 3.4 Kepler arşivinde bulunan 1572 örten çift yıldızın bir eğri boyunca LLE izdüşümü (projeksiyon). İzdüşümler global ötelemelere, döndürmelere ve ölçeklemelere bağlı olmadığından, grafiklerin x ve y eksenlerindeki değerler rasgeledir. İçteki grafikte örnek olarak gösterilen 15 adet yıldızın polinom fitleri Şekil 3.5'te gösterilmiştir (Matijević vd. (2012)'den değiştirilerek alınmıştır)

Bu eğri boyunca ilerledikçe çift sistemlerin morfolojilerinin sürekli bir şekilde değiştiği gözlenmiştir: çok dar tutulmalı ayırık sistemlerden başlayarak daha geniş tutulmalı ayırık sistemlere, yarı-ayırık sistemlere, değen sistemlere ve son olarak elipsoidal/belirsiz sistemlere doğru bir geçiş mevcuttur.

Elipsoidal değişenler çift yıldızlardır ancak tutulma koşulunu sağlamazlar, örten değişen değildirler. Parlaklık değişimleri yıldızlar dolanırken elipsoidal şekillerinden dolayı değişen yüzey alanlarından kaynaklanır. Dolayısıyla ışık eğrileri sinüzoidale çok yakındır.

Bu tek boyutlu yapıdan yola çıkarak, her bir ışık eğrisini temsil eden ve sistemin "ayrıklık derecesini" ölçen, 0 ile 1 arasında değişen tek bir sınıflandırma parametresi "c" tanımlanmıştır. Örnek olarak alınan 15 yıldızın "c" parametresine göre sıralanışı Şekil 3.5'te gösterilmiştir.



Şekil 3.5 Örnek 15 yıldızın polinom fitleri ve morfoloji parametresine göre sıralanışı (Matijević vd. (2012)’den değiştirilerek alınmıştır)

Parametre 0’dan 1’e doğru gittikçe örten çiftlerin ayrıktan elipsoidallere doğru sıralandığı görülmüştür. Bu parametre “morfoloji parametresi” olarak Kepler örten çift kataloğuna bir sütun olarak eklenmiştir⁹. Morfoloji parametresi TESS örten çift yıldızları için de Prša vd. (2022) tarafından hesaplanıp TESS kataloğuna eklenmiştir¹⁰.

⁹ <https://keplerebs.villanova.edu/>

¹⁰ <https://tessebs.villanova.edu/>

Matijević vd. (2012) morfoloji parametresine (morph) göre örten çift yıldızları aşağıdaki şekilde sınıflandırmıştır:

$\text{morph} < 0.5$ detached (D)

$0.5 < \text{morph} < 0.7$ semi-detached (SD)

$0.7 < \text{morph} < 0.8$ overcontact (OC)

$0.8 < \text{morph}$ ellipsoidal (E)

3.1.2 Işık eğrisinin biçimine göre sınıflandırma

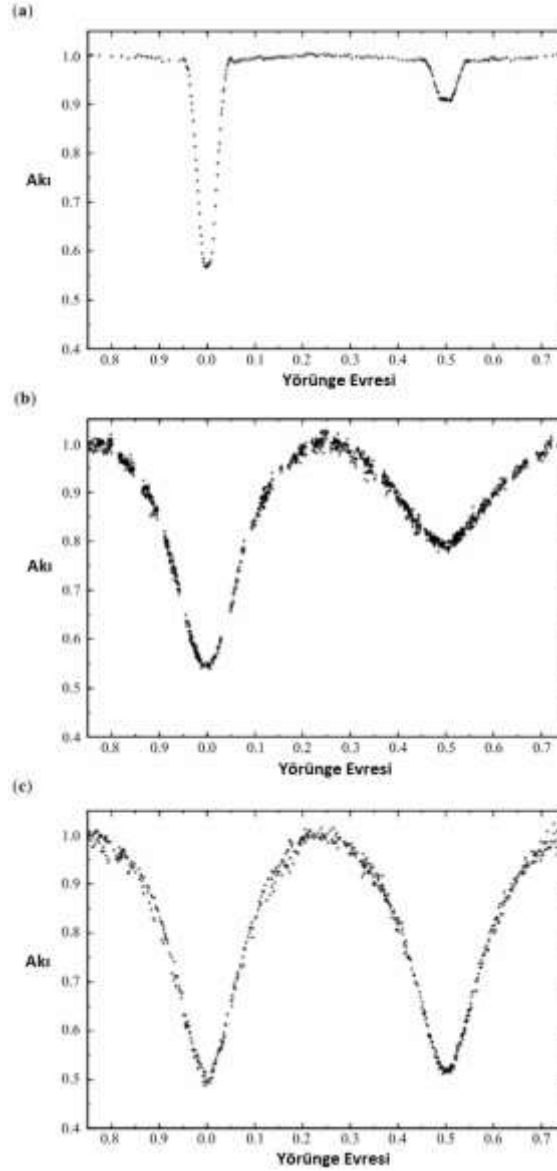
Örten çift yıldız sistemleri, General Catalogue of Variable Stars (GCVS) standartları çerçevesinde, ışık eğrilerinin morfolojik karakteristiklerine göre temel olarak EA (Algol), EB (Beta Lyrae) ve EW (W Ursae Majoris) olmak üzere üç sınıfa ayrılır (Samus vd. 2017). Bu sınıflandırma ışık eğrilerinin şekillerini yansıtmakta olup, örten çift yıldızların evrimini ve fiziksel özelliklerini dikkate almamaktadır.

EA (Algol): Işık eğrisinde minimumlar belirgindir. Birinci minimum daha parlak yıldızın sönük yıldız tarafından örtülmesiyle oluşur ve oldukça derindir. İkinci minimum daha sönük yıldızın ışınım gücü daha yüksek yıldız tarafından örtülmesiyle meydana gelir ve birinci minimuma nazaran daha sığdır. Minimumların başlangıç ve bitiş noktaları oldukça nettir. Tutulmalar haricinde ışık eğrisi, yüzey parlaklık dağılımının homojenliğini bozan leke ve plaj bölgeleri ile yıldız zonklamaları gibi olguların neden olduğu değişkenlikler dışında, neredeyse düzdür. Işık eğrisinin tutulmalar dışında büyük ölçüde sabit kalması yıldızların küresel şekle sahip olduğu anlamına gelir. Dolayısıyla bu yıldızlar birbirlerinden yeterince uzaktır ve birbirlerinin şeklini gelgit etkileriyle önemli ölçüde bozmazlar. Ayrıca bileşenlerin uzaklıkları nedeniyle yansıma etkileri de minimal düzeydedir. Bu tür sistemler genellikle ayrı sistemlerdir (Şekil 3.6.a). Sıcak alt-dev ya da beyaz cüce barındıran kısa dönemli sistemler Roche geometrisine göre ayrı sistemler oldukları halde ışık eğrilerinde tutulmalar dışında yüksek yansıma etkisi nedeniyle sürekli değişimler gösterirler.

EB (Beta Lyrae): EB türü sistemlerde parlaklık yörünge dönemi boyunca sürekli olarak değişir. Minimumların başlangıç ve bitişleri net değildir. Birinci ve ikinci minimum derinlikleri birbirinden oldukça farklıdır. Işık eğrisindeki sürekli değişim, yıldızların şekillerinin önemli ölçüde bozulmuş olduğunu gösterir. Bileşenlerden biri genellikle Roche şişimini doldurmuş ve yoldaşının kütleçekimi kaynaklı kuvvetli gelgit etkileri nedeniyle şekli “armutlaşmıştır”. Buna bağlı olarak yörünge hareketi sırasında görünen yüzey alanı sürekli değişir (Şekil 3.6.b) ve ışık eğrilerinde tutulmalar dışı sürekli parlaklık değişimleri gözlenir. Ancak, ayırık sistemler olduğu halde ışık eğrileri EB olarak sınıflandırılan sistemler de bulunmaktadır. Bu sistemlerin bileşenleri Roche şişimlerini doldurmamış olsa dahi yakınlık etkileri nedeniyle elipsoid bozulmalar ve yüksek yansıma etkileri gösterirler. EE Pegasi sistemi bu durumun iyi bir örneğidir. Wilson-Devinney (Wilson ve Devinney 1971) yöntemi ile yapılan ışık ve dikine hız eğrisi analizleri sistem bileşenlerinin Roche şişimlerinin içinde bulunduğunu, dolayısı ile sistemin ayırık bir sistem olduğunu ortaya koymuştur. Ancak sistemin ışık değişimi EB türü ile uyumludur. Tutulmalar dışı ışık değişimi elipsoidal bozulma ve yansıma etkileri nedeniyle sürekli ve kontakt noktaları belirgin değildir (Kharchenko vd. 2007). Başka bir örnek olarak OO Aql gerçekte değen olup EB türü ışık eğrisine sahiptir. Bileşen yıldızlar ortak zarf içerisindedir ancak termal dengeye ulaşmadıklarından ışık eğrisinde minimumlar arasında derinlik farkı vardır. Bu da ışık eğrisinde EB profili yaratır (Hrivnak 1989).

EW (W Ursae Majoris): Birinci ve ikinci minimum derinlikleri birbirine çok yakındır. Işık eğrisi sürekli değişken yapıda olup, kontakt noktalarını belirlemek mümkün değildir. EB türü gibi sürekli parlaklık değişimi gösterirler ancak daha simetrik yapıdadırlar ve bu nedenle sinüs eğrilerini andırırlar. Minimum derinliklerinin birbirine çok yakın olması çift yıldızın yüzey sıcaklıklarının neredeyse aynı olduğunu gösterir. Bu tür ışık eğrileri genellikle her iki yıldızın da kendi Roche loblarını doldurduğu değen ve aşırı değen sistemlerde gözlenir. Bu yıldızlar birbirlerine fiziksel olarak temas ederler ve ortak bir zarfın içinde bulunurlar. Bu nedenle de etkin sıcaklıkları birbirine çok yakındır ve yörünge dönemleri de genellikle bir günden kısadır (Şekil 3.6.c). Ancak bu durumun da aykırı örnekleri mevcuttur. Örneğin CN And örten çiftin ışık eğrisi neredeyse sinüs eğrisine benzer, maksimumlarında düzlük yoktur ve minimum derinlikleri neredeyse eşit derinliktedir. Fotometrik olarak EW türü gibi gözükse de tayfsal çalışmalar bileşenlerin

birbirine değmediğini, bileşenlerden birinin Roche şişimini doldurduğunu, dolayısıyla sistemin yarı-ayrık olduğunu göstermiştir. Yakınlık etkisinden dolayı yıldızların aldıkları elipsoidal şekil tutulma dışındaki düzlükleri yok ederek ışık eğrisine sinüs eğrisi biçimi vermiştir (Lee ve Lee 2013).



Şekil 3.6 Örtlen değişen yıldızlara ait örnek ışık eğrileri; (a) Algol, (b) Beta Lyrae, (c) W UMa (Skelton ve Smits (2009)’dan değiştirilerek alınmıştır)

Geleneksel sınıflandırma ile Kopal sınıflandırması arasında benzerlik göze çarpar; EA ışık eğrileri ayırık, EB ışık eğrileri yarı-ayrık ve EW ışık eğrileri değen sistemlerle ilişkilendirilir. Ama gerçekte durum çok daha karmaşıktır. Bir sistemin ışık eğrisi türü her zaman onun fiziksel durumuyla örtüşmeyebilir. Ayırık sistem olup EB veya EW türü

ışık eğrisine (örn. EE Peg ve CN And), yarı-ayrık olup EA türü ışık eğrisine (örn. Algol (Beta Per)'in kendisi), değen olup EB türü ışık eğrisine (örn. OO Aql) sahip olabilir.

3.2 Gözlemsel Veri

Bu çalışmada Gaia DR3 (Gaia Data Release 3), TESS ve Kepler arşivlerinde bulunan örten çift yıldızların ışık eğrileri kullanılmıştır.

3.2.1 Gaia DR3 arşivi

Gaia, Avrupa Uzay Ajansı'nın (ESA) 2013'te uzaya gönderilen ve en temelde Samanyolu'ndaki yıldızların konumlarını ve uzaklıklarını hassas olarak ölçmeyi amaçlayan bir astrometri uydusudur. Gaia Uydu Teleskobu'ndan elde edilen veriler ESAC (European Space Astronomy Centre) veri merkezi tarafından işlenir. Bu arşivin verileri herkesin kullanımına açıktır. Gaia DR3, bu arşivde yayımlanan üçüncü büyük veri setidir. 1,8 milyardan fazla yıldız için konum, parlaklık, renk, paralaks bilgisi sağlamaktadır. Gaia DR3 arşivinde 2,1 milyon örten çift yıldız bulunmaktadır¹¹. Gaia Data Processing and Analysis Consortium (Gaia DPAC)'un değişen olarak sınıflandırdığı yıldızlardır.

3.2.2 Villanova Kepler ve TESS katalogları

Villanova Üniversitesi, Kepler ve TESS uzay teleskopları tarafından gözlenen örten çift yıldızların ışık eğrilerini ve bu ışık eğrilerinin çözümlerini barındıran birer veritabanı sunmaktadır.

Kepler, NASA tarafından geçiş yöntemiyle ötegezegen keşfi amaçlı olarak geliştirilmiş bir uzay teleskobudur. Kuğu ve Çalgı takımyıldızları civarında 115 kare derecelik bir bölgeyi 2009 – 2018 yılları arasındaki yaklaşık 4 yıllık sürede kesintisiz olarak gözlemiştir. Villanova Kepler örten çift yıldız kataloğunda¹² Kepler uzay teleskobunun

¹¹ https://cdn.gea.esac.esa.int/Gaia/gdr3/Variability/vari_eclipsing_binary/

¹² <https://keplerebs.villanova.edu/>

birincil görevi sırasında (2009-2013) gözlemlenen 2907 adet örten çift yıldızın ışık eğrisi verisi bu tez çalışmasında kullanılmak üzere indirilmiştir.

TESS de NASA tarafından parlak yıldızlar etrafında öte gezegen keşfi amaçlı geliştirilmiş dört küçük uzay teleskobundan oluşan bir düzendir. Kepler'den farklı olarak bütün gökyüzünü gözlemler ve yörüngesinde geçiş (ing. transit) yapan öte gezegen bulunma olasılığı olan parlak yıldızları hedefler. TESS 2018 yılından beri çalışmaktadır. Bu tez çalışmasında kullanılmak üzere Villanova TESS örten çift yıldız kataloğundan¹³ 4349 adet örten çift yıldızın ışık eğrisi verisi indirilmiştir.

Bu kataloglarda örten çift yıldızların her biri için sistemin Roche geometrisine göre ayrıklığını belirten morfoloji (morph) parametresi verilmiştir. Morfoloji parametresi 0 ile 1 arasında değer alır. Morfoloji parametresi bu çalışmada, bu kataloglardaki örten çift yıldızların ayırık, yarı-ayırık, değen ve elipsoidal şeklinde sınıflandırılmasında kullanılmıştır.

3.3 Makine Öğrenmesi Modelleri

Bu tez çalışmasında örten çift yıldızların 3 farklı denetimli makine öğrenmesi yöntemi ile sınıflandırılması yapılmıştır:

- Hem ışık eğrisi görselleri hem de sayısal ışık eğrisi verisi kullanılarak çok modlu (multimodal) model ile Gaia DR3 örten çift yıldızları sınıflandırılmıştır.
- Kepler ve TESS örten çift yıldızlarının ışık eğrisi görselleri kullanılarak VGG-19 tabanlı model eğitilmiş ve bu model ile Gaia DR3 örten çift yıldızları sınıflandırılmıştır.
- Kepler ve TESS örten çift yıldızlarının Fourier katsayıları kullanılarak bir boyutlu CNN modeli eğitilmiş ve bu model ile Gaia DR3 örten çift yıldızları sınıflandırılmıştır.

¹³ <https://tessebs.villanova.edu/>

Ayrıca Kepler örten çift yıldızlarının ışık eğrilerinin Fourier dönüşümleri sonucu elde edilen fonksiyonların katsayıları da, GMM denetimsiz makine öğrenmesi yöntemiyle kümelendirme amacıyla kullanılmıştır.

3.3.1 Çok modlu (multimodal) model

Gaia DR3 örten çift yıldızlarını sınıflandırmak için hem ışık eğrisi görsel verileri hem de sayısal verileri aynı anda işleyen çok modlu (multimodal) bir makine öğrenmesi modeli geliştirilmiştir (Bölüm 4.1). Bu model hem görsel (resim) hem de sayısal veriyi aynı anda kullanarak veri kümelerinin birleşiminden daha zengin bir özellik çıkarımı yapabilir. Her bir örten çift yıldızın Mowlavi vd. (2023)'te verilen geometrik model uyumlaması (fit) parametreleri ve bu parametrelere karşılık gelen PNG formatında (640×480 piksel) ışık eğrisi görselleri elde edildi. Görseller 128×96 piksel olarak yeniden boyutlandırıldı. Ardından sayısal verilere dönüştürüldü ve $[0, 1]$ aralığına normalize edildi. Etiketler (EA, EB, EW sınıfları) 'Label Encoder' kullanılarak EA için 1, EB için 2, EW için 3 şeklinde sayısal değerlere dönüştürüldü. Veriler %80 eğitim (ing. training) ve %20 doğrulama (ing. validation) olarak bölündü. Sayısal veri ortalaması 0, standart sapması 1 olan bir normal dağılım 'Standart Scaler' kullanılarak ölçeklendi. Bu işlem makine öğrenmesi modelinin daha doğru öğrenmesini sağlamaktadır. Standart Scaler veri setini dengeli hale getiren bir araçtır.

Görsel veri CNN katmanları ile işlendi. İlk olarak Conv2D katmanı ile görsel özellikler çıkarıldı. Ardından MaxPooling2D katmanı ile daha düşük boyutlu ve anlamlı özellikler elde edildi. Flatten katmanı ile özellikler düzleştirildi ve sonrasında yoğun (dense) katmanında öğrenilen özellikler birleştirildi.

Sayısal veri MLP ile işlendi. Görüntü ve sayısal veriden elde edilen özellikler 'concatenate' fonksiyonu ile birleştirildi. Böylece her iki veri türünden gelen bilgilerin birleşerek daha zengin ve anlamlı temsili sağlandı. Son olarak birleşen veriler üzerinde sınıflandırma yapıldı. Son katman olan çıktı katmanında 'Softmax' aktivasyon fonksiyonu her bir sınıf için bir olasılık değeri üretti.

Makine öğrenmesi modelinin “ezberlemesini” önlemek ve performansı artırmak için bazı mekanizmalar eklendi;

- **MacroF1 Callback:** Her çevrim¹⁴ (ing. epoch) sonunda modelin ‘macro F1’ skorunu hesaplar ve eğitim sırasında doğrulama seti üzerinde bu skoru takip eder. Bu metrik, özellikle dengesiz veri setlerinde modelin başarısını daha iyi değerlendirmek için kullanılmaktadır. F1 skoru bir modelin Precision (Kesinlik) ve Recall (Duyarlılık) metriklerinin harmonik ortalamasıdır. Macro F1 skoru ise her bir sınıf için ayrı ayrı hesaplanan F1 skorlarının aritmetik ortalamasıdır.
- **Model Checkpoint:** Doğrulama seti üzerindeki ‘macro F1’ skoru en iyi olan modeli kaydeder. Modelin aşırı uyum yapmaması için erken durdurma stratejisinin de önüne geçmektedir.
- **Early Stopping:** Modelin doğrulama seti üzerindeki performansı uzun süre iyileşmediyse eğitim erken durdurulmaktadır. Gereksiz yere harcanan hesaplama gücünden ve zamandan tasarruf sağlamaktadır.
- **Dropout Katmanları:** Eğitim sırasında rasgele nöronları kapatarak modelin ezberlemesini engelleyip kavramaya zorlar.
- **L2 Regularization:** Yoğun (ing. dense) katmanlara eklenen parametredir. Modelin ağırlıklarının aşırı büyümesini önler.
- **ReduceLROnPlateau:** Eğitim sırasında modelin bir lokal minimumda (ing. local minima) kalmasını engellemek için doğrulama kaybı (ing. validation loss) değeri 4 çevrim boyunca iyileşmezse öğrenme hızını yarıya indirir.
- **Data Augmentation:** Işık eğrisi görselleri birbirine çok benzediği için modelin ezberlemesi kolaydır. Işık eğrisi görsellerini farklılaştırmak amacıyla rasgele yatay çevirme (ing. horizontal flip), sağa/sola veya yukarı/aşağı kaydırma (ing.

¹⁴ Bu çalışmada "çevrim" terimi iki farklı anlamda kullanılmıştır. Makine öğrenmesi bağlamında yapay sinir ağı modelinin tüm eğitim veri setini bir kez tam olarak işlediği iterasyon adımını, astrofiziksel bağlamda ise örten çift yıldızın ortak kütle merkezi etrafındaki tam yörünge dolanımını ifade etmektedir.

translation), yakınlaştırma (ing. zoom) işlemleri uygulanmıştır. Dikey çevirme (ing. vertical flip) ışıık eğrilerinde asla yapılmamalıdır.

Eğitim tamamlandığında, modelin kayıp (loss) ve doğruluk (accuracy) değerleri ile macro F1 skorunu gösteren öğrenme eğrisi (learning curve) çizdirilmektedir. Bu grafik modelin eğitim sürecindeki performansını görsel olarak göstermektedir.

Makine öğrenmesi modelini test etmek için veri setinin bir kısmı test veri seti olarak ayrıldı. Test verisi de eğitim ve doğrulama veri seti ile aynı şekilde işlendi, etiketlendi ve ölçeklendi. Modelin test verisi üzerinde tahminleri sonrasında modelin başarısı “doğruluk” (ing. accuracy), “kesinlik” (ing. precision), “duyarlılık” (ing. recall) ve F1 skoru metrikleri ile değerlendirildi. Modelin her sınıf için doğru ve yanlış tahminlerini görselleştiren bir karışıklık matrisi (confusion matrix) oluşturuldu. Bu matris modelin hangi sınıflarda hata yaptığını görmeyi sağlamaktadır. Bu makine öğrenmesi modelinin detayları Ek 1A’da verilmiştir.

3.3.2 VGG-19 tabanlı model

Örten çiftleri makine öğrenmesi teknikleri ile sınıflandırmada alternatif bir yaklaşımı değerlendirmek açısından, kendi modelimizi oluşturmak yerine, makine öğrenmesi literatüründe iyi bilinen hazır eğitilmiş model olan VGG-19 ile transfer öğrenme (ing. transfer learning) yapılmıştır (Bölüm 4.2).

VGG-19, Oxford Üniversitesi Visual Geometry Group (VGG) departmanı tarafından geliştirilmiş ve önceden eğitilmiş (ing. pretrained) bir CNN modeldir. 19 katmandan oluşan model, ağ genelinde tek tip 3×3 evrişimsel filtreler kullanması sayesinde sade bir mimariye sahiptir. VGG-19’un ince ayarı, düzenleme tekniklerinin kullanımıyla birleştirildiğinde, “ezberleme” (overfitting) olmadan sonuçlar vermektedir. Modelin ezberlemesini önlemek için "platoda LR'yi azaltma (reduce LR on plateau)" ve "erken durdurma (early stopping)" yöntemleri kullanıldı. Ayrıca, değişken hiperparametrelerle deneyler gerçekleştirildi ve aşırı eğitim (overtraining) meydana gelmediği doğrulandı (Keskin ve Baştürk 2025).

Kepler ve TESS arşivindeki örten çift yıldız görselleri morfoloji parametresine göre etiketlenerek VGG-19 modeli eğitildi. Eğitilmiş model ile Gaia DR3 arşivindeki 2106 örten çift yıldız sınıflandırıldı.

Villanova arşivinden indirilen 2907 adet Kepler ve 4349 adet TESS ışık eğrisi verilerinden PNG formatında görselleri elde edildi. Görseller morfoloji parametresine göre “detached”, “semi-detached”, “over-contact” ve “ellipsoidal” şeklinde etiketlendi (Matijević vd. 2012). Kepler ve TESS veri kümeleri daha sonra tek bir liste halinde birleştirildi. Ardından %67 eğitim, %25 doğrulama ve %8 test olmak üzere ayrıldı.

Doğrulama kaybı (val_loss) değeri 10 çevrim boyunca gözlemlendi ve iyileşme olmazsa early stopping mekanizması ile eğitim durduruldu. Eğitimin durdurulduğu andaki değil doğrulama kaybının en düşük olduğu çevrimdeki model ağırlıklarının kullanılması sağlandı. Modelin başarısı ‘accuracy’ metriği ile değerlendirildi ve karışıklık matrisi oluşturuldu. Bu makine öğrenmesi modelinin detayları Ek 1B’de verilmiştir.

3.3.3 Bir boyutlu CNN modeli

Örten çift yıldızları sınıflandırmada sayısal veri kullanarak sınıflandırma yapıldı (Bölüm 4.3). Yalnızca sayısal veri kullanarak sınıflandırma yapmak yaygın bir yaklaşımdır. Sayısal verilerde her sütun bir özelliği (ing. feature), her satır ise bir örneği temsil eder. Makine öğrenmesi modeli sayısal özellikler ile o özelliklere ait etiket arasındaki matematiksel ilişkiyi öğrenir.

Işık eğrileri zaman serisi verisidir. Bu veriden türetilen Fourier katsayıları sayısal yapıdadır. Bu katsayılar yıldızın ayırık, yarı-ayırık, aşırı degen ve ellipsoidal türüne ait bilgi içerir (Rucinski 1993 ve Selam 2004).

Rucinski (1993) W UMa türü degen çift yıldız sistemlerinin ışık eğrilerini analiz etmek için basit ve pratik bir yöntem sunmuş ve degen çift yıldızların temel parametrelerini karmaşık ışık eğrisi sentezleme modellerine gerek kalmadan tahmin etmiştir. Bu tür sistemlerin ışık eğrileri büyük ölçüde geometrik faktörlere (kütle oranı 'q', yörünge eğimi

'i' ve değme derecesi 'f') bağlıdır. Rucinski (1993) bu üç parametrenin farklı değerleri için ($0.05 \leq q < 1$; $30^\circ < i < 90^\circ$; $f = 0, 0.5, 1$) sentetik ışık eğrileri oluşturmuştur. Sentetik ışık eğrileri kütle, çekim kararması ve albedo etkilerini içermektedir. Daha sonra sentetik ışık eğrilerini Fourier kosinüs serileri ile modellemiştir. Model fonksiyonu Eşitlik 3.1'de verilmiştir (Selam 2004).

$$l(\theta) = \sum_{i=0}^{10} a_i \cos(2\pi i \theta) \quad (3.1)$$

Eşitlik 3.1'de a_i Fourier katsayısını, θ evre açısını, $l(\theta)$ normalize ışık akısını ifade etmektedir. Rucinski (1993) ilk altı terimin katsayılarını ışık eğrisinin şeklini tanımlamakta kullanmıştır. Bu katsayılardan a_2 ve a_4 'ün sistemin değme derecesini (f) tahmin etmek için kullanılabileceğini göstermiştir.

2907 adet Kepler, 4349 adet TESS ve 2106 adet Gaia DR3 örten çift yıldızının ışık eğrileri Eşitlik 3.1'e göre Fourier serisi ile modellendi. Modelleme, kosinüslü terimler için a_0 'dan a_{10} 'a kadar 11 tane parametre ve bir tane de temel frekansı belirten ω (omega) olmak üzere 12 parametre üretmektedir. Bu fit sonucunda elde edilen Fourier katsayıları ışık eğrisinin biçimi hakkında (minimumların derinliği ve genişliği, maksimumların şekli vb.) temel bilgileri sayısal olarak içerir. Bu 12 parametre makine öğrenmesi modeli için öznitelik olarak kullanıldı.

Kepler ve TESS yıldızlarının Fourier katsayıları ile bir boyutlu CNN modeli eğitildi ve eğitilen model ile Gaia DR3 yıldızları tahmin edildi. Bu makine öğrenmesi modelinin detayları Ek 1C'de verilmiştir.

3.3.4 GMM modeli

Örten çift yıldızların ışık eğrisi verisi içinde doğal gruplar araştırılmıştır (Bölüm 4.4). Bunun için Kepler veri setindeki örten çift yıldızların ışık eğrilerine yapılan Fourier fit

katsayıları kullanılmıştır. Buradaki amaç Fourier katsayılarına göre yapılan sınıflandırma sonucu elde edilen sınıflar ile Fourier katsayılarının dağılımının uyumlu olup olmadığının araştırılmasıdır. Fourier a_1 , a_2 , a_4 katsayıları “Standart Scaler” ile standart hale getirildi ve PCA ile 2 boyuta indirildi. Gaussian Mixture Model (GMM) ile 3 kümeye ayrıldı. Kepler örten çift yıldızlarının Fourier katsayılarına göre EA, EB, EW sınıflarına ayrılması ile Fourier katsayılarının 3 kümeye ayrılması arasındaki ilişki araştırılmıştır. Bu makine öğrenmesi modelinin detayları Ek 1Ç’de verilmiştir.

3.4 Yazılım ve Donanım

Bu araştırma için Apple MacBook Pro, M3 işlemci, 16 GB RAM, 512 GB SSD, macOS Sequoia işletim sistemi olan kişisel bilgisayar üzerine kurulu PyCharm¹⁵ geliştirme ortamında Python dilinde kodlar yazılmıştır. Kod yazımında ve makine öğrenmesi modellerinin geliştirilmesinde ChatGPT¹⁶ ve Google Gemini¹⁷ yapay zeka araçlarından faydalanılmıştır. Yazılan kodların bir kısmı söz konusu kişisel bilgisayarda çalıştırılmıştır. Makine öğrenmesi modelinin eğitilmesi, milyonlarca gözlem verisinin internetten indirilmesi, ışık eğrilerinin çizilmesi gibi çok uzun süren işler için TRUBA sistemi kullanılmıştır. Bu tez kapsamında yazılan kodlar, kullanılan bütün Gaia DR3, Kepler ve TESS arşiv verileri, elde edilen sonuçlar, grafikler vd. bütün dijital materyal Github¹⁸ sayfasında paylaşılmıştır.

¹⁵ <https://www.jetbrains.com/pycharm/>

¹⁶ <https://chatgpt.com/>

¹⁷ <https://gemini.google.com/app>

¹⁸ <https://github.com/BedriKeskin/EclipsingBinaryClassification>

4. BULGULAR

4.1 Çok Modlu (Multimodal) Sınıflandırma

Bu bölümde Gaia DR3 arşivinde bulunan örten çift yıldızların çok modlu (multimodal) makine öğrenmesi modeli ile sınıflandırılması yapıldı. Gaia DR3 arşivinde bulunan örten çift yıldızlar iki şekilde temsil edilebilir:

- Mowlavi vd. (2023) geometrik model parametreleri ile sayısal veri olarak,
- Mowlavi vd. (2023) geometrik model parametreleriyle çizilen ışık eğrisi grafiği ile PNG (Portable Network Graphics) formatında görsel olarak.

Bu iki veri tipinin de kullanılabileceği çok modlu denetimli makine öğrenmesi modeli oluşturuldu (Bölüm 3.3.1). Gaia DR3 veri setinde G-bandında 4 kadirde 20 kadire kadar parlaklıklara sahip ve tüm gökyüzünü kapsayan 2.1 milyon kaynak içeren ilk Gaia örten çift yıldız adayları kataloğu da yer almaktadır. Örten çift yıldız adayları, Rimoldini vd. (2023) tarafından, Gaia DR3 veri setindeki değişen yıldızların XGBoost ve RF yöntemleriyle sınıflandırılması sonucunda elde edilmiştir. Mowlavi vd. (2023) örten çift yıldız adaylarını daha sonra G-bandındaki ışık eğrilerine dayalı örten çift yıldızlara uyarlanmış kriterlere göre filtrelemiş, yıldızların yörünge dönemlerini bulmak için 3 farklı periyot arama yöntemi kullanmıştır. Daha sonra her bir örten çift yıldız için, temizlenmiş G-bandı zaman serisine bir veya iki Gauss ve/veya bir kosinüs uyumlanarak (fit edilerek) G-bant ışık eğrisinin geometrik bir modeli oluşturulmuştur¹⁹. Son olarak, yörünge periyodu ve geometrik modelin en iyi kombinasyonu, BIC (Bayesian Information Criterion)'a dayalı olasılıkçı (Bayesian) bir model karşılaştırması kullanılarak seçilmiştir.

Geometrik modellerin Gaussyen bileşenleri tutulmaların ve ışık eğrisinin genel şeklinin, kosinüs bileşeni ise elipsoidal ya da yansıma etkisi benzeri bir değişkenliğin temsili için

¹⁹ Buradaki "model" kavramı, Wilson-Devinney gibi bir ışık eğrisi modellemesini değil, gözlem verisine matematiksel bir eğri uydurma (ing. curve fitting) işlemini ifade etmektedir.

kullanılmaktadır. Bileşen sayısına göre geometrik model türlerinin listesi, her model türüne ait parametre sayısı, model açıklaması ve modellenen örten çift sayısı Çizelge 4.1’de verilmiştir.

Çizelge 4.1 G-bant ışık eğrilerine fit edilen geometrik model türleri (Mowlavi vd. (2023)’den değiştirilerek alınmıştır)

Model türü	Parametre sayısı	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

Çizelge 4.1’deki 'ELLIPSOIDAL' ifadeleri, bir değişen yıldız türü olan 'Elipsoidal Değişenler' sınıfını değil, uygulanan geometrik modeli temsil etmektedir. Işık eğrisinde elipsoidal değişim gösteren, bir kosinüs ile fit edilen modeli tanımlar.

Gaussyen bileşenin fonksiyonu Eşitlik 4.1’de verilmiştir.

$$f_{Gauss,k}(\varphi) = d_k \exp\left(-\frac{(\varphi - \mu_k)^2}{2\sigma_k^2}\right) \quad (4.1)$$

Eşitlik 4.1’de k indisi 1. veya 2. Gaussyen bileşeni, φ yörünge evresini, d minimum derinliğini, μ Gaussyen bileşenin evresini, σ Gaussyen bileşenin standart sapmasını ifade etmektedir. Kosinüs bileşenin fonksiyonu Eşitlik 4.2’de verilmiştir.

$$f_{cos}(\varphi) = A_{ell} \cos[4\pi(\varphi - \mu_{ell})] \quad (4.2)$$

Eşitlik 4.2’de φ yörünge evresini (Eşitlik 4.1’deki ile aynı), A_{ell} kosinüs bileşenin genliğini, μ kosinüs bileşenin evresini ifade etmektedir. Eğer geometrik modelde bütün bileşenler varsa model fonksiyonu Eşitlik 4.3’teki şekilde olur.

$$f(\varphi) = C + f_{Gauss,1}(\varphi) + f_{Gauss,2}(\varphi) + f_{cos}(\varphi) \quad (4.3)$$

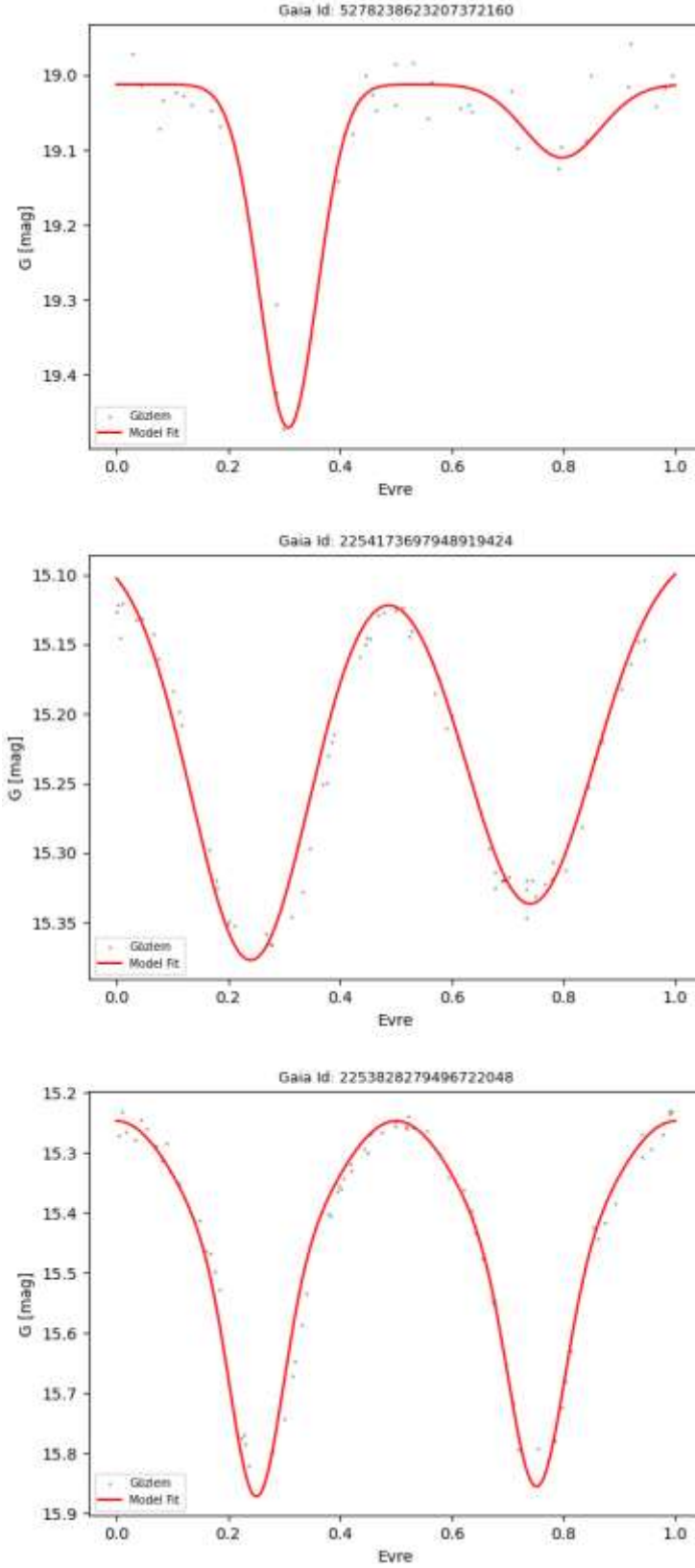
Eşitlik 4.3’de C geometrik modelin referans seviyesini ifade etmektedir.

Bütün örten çift yıldızların ışık eğrisi verisi ve geometrik model parametreleri internetten indirildi²⁰. Yıldızların geometrik model parametreleri ve fonksiyonları kullanılarak ışık eğrisi grafikleri PNG formatında elde edildi.

Işık eğrisi ile modelin ne kadar uyumlu olduğunu göstermek amacıyla, Şekil 4.1’de katlanmış ışık eğrisi üzerine çizilmiş model grafikleri örnek olarak verilmiştir. Daha fazla örnek için Github sayfasına bakılabilir²¹. Makine öğrenmesinde sadece kırmızı ile gösterilen model grafikleri kullanılmıştır.

²⁰ https://cdn.gea.esac.esa.int/Gaia/gdr3/Variability/vari_eclipsing_binary/

²¹ https://github.com/BedriKeskin/EclipsingBinaryClassification/tree/main/Gaia/Combined_Plots



Şekil 4.1 Gaia DR3 arşivinde bulunan örten çift yıldızların katlanmış ışık eğrisi üzerine çizilen Mowlavi vd. (2023) geometrik modeline göre elde edilen EA, EB, EW sınıflarına ait örnek model grafikleri

Gaia DR3 örten çift yıldızlarının sınıflandırılması, Mowlavi vd. (2023)'in ışık eğrilerine uyguladığı 6 ayrı model türüne göre ayrı ayrı değerlendirildi. Bir Guassien (ONEGAUSSIAN) ile modellenen yıldızların tamamının EA (Algol) türü ışık eğrisi şeklinde, elipsoidal (ELLIPSOIDAL) ile modellenen yıldızların tamamının kosinüs eğrisi şeklinde olduğu görüldü. Dolayısıyla bu iki tür modellenen yıldızlarda sınıflandırma yapılmadı. Diğer yıldızlarda her bir model türünde ayrı ayrı EA, EB, EW şeklinde sınıflandırmaya gidildi.

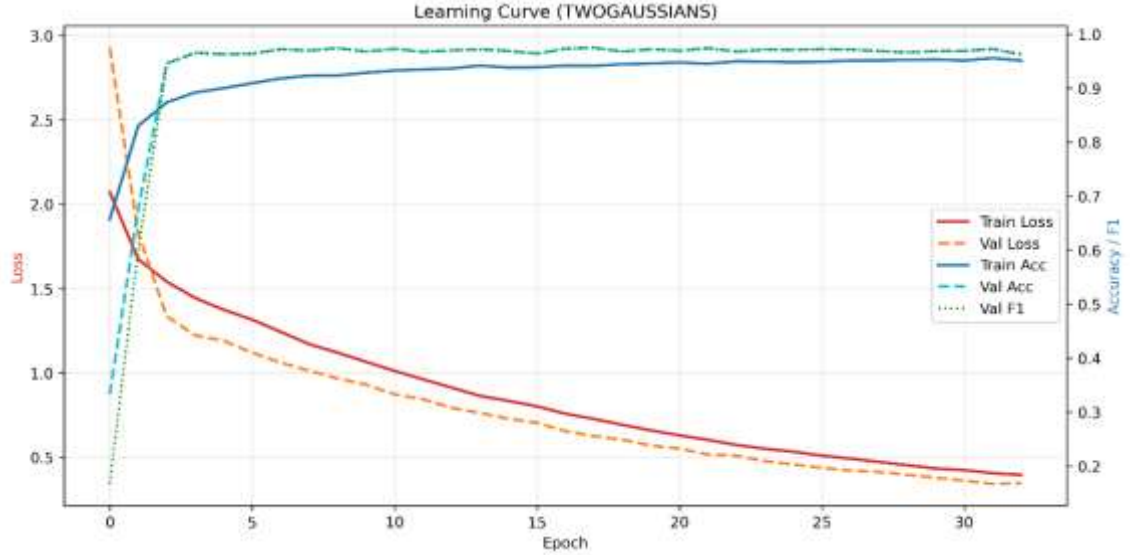
Yıldızların bir kısmı makine öğrenmesi modelini eğitmek üzere eğitim/doğrulama için, bir kısmı da eğitilmiş makine öğrenmesi modelini test etmek için ışık eğrisinin biçimine göre görsel inceleme ile EA, EB ve EW şeklinde sınıflandırıldı (etiketlendi). Makine öğrenimi uygulamalarında milyonlarca veriden oluşan bir popülasyondan dikkatlice seçilmiş ve doğru etiketlenmiş binlerce örneğin kullanılması, genelleme yeteneği yüksek bir model oluşturmak için yaygın bir yaklaşımdır. Bütün veri seti içerisinde ne kadar verinin etiketlenmesi gerektiğine dair bir standart yoktur. Makine öğrenmesi modelinin öğrenme sürecindeki metriklerine ve test veri setindeki sonuçlara bakılarak, modelin öğrendiğine kanaat getirilene kadar etiketli veriye ihtiyaç vardır. Görsel inceleme ile sınıflandırılan örten çift yıldızların sayısı Çizelge 4.2'de verilmiştir.

Çizelge 4.2 Görsel inceleme ile sınıflandırılan örten çift yıldız sayıları

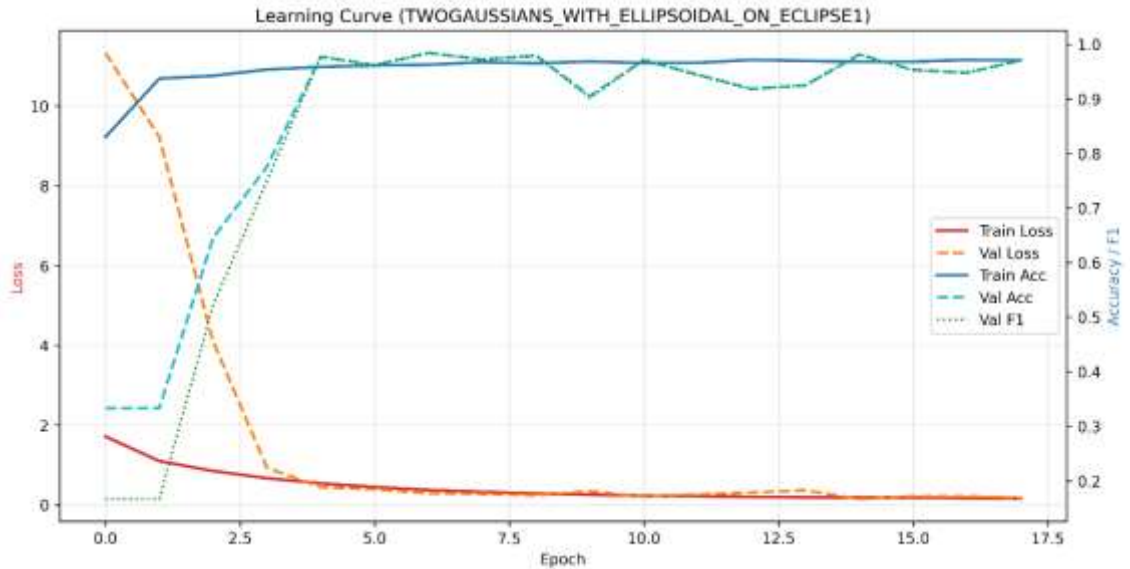
Model türü	Eğitim/ Doğrulama	Test
TWOGAUSSIANS	15 094	3 031
TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE1	10 032	1 983
TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE2	7 981	1 608
ONEGAUSSIAN_WITH_ELLIPSOIDAL	7 537	1 558
Toplam	40 644	8 180

Her bir model türü için aynı çok modlu makine öğrenmesi mimarisi kullanılmış olup, sadece sayısal veri girişinde Çizelge 4.1'de verilen parametre sayısına bağlı olarak değişiklik yapılmıştır. Çok modlu denetimli makine öğrenmesi modeli, eğitim/doğrulama için etiketlenen yıldızların sayısal ve görsel verisi ile eğitildi. Eğitim sürecinde yıldızların

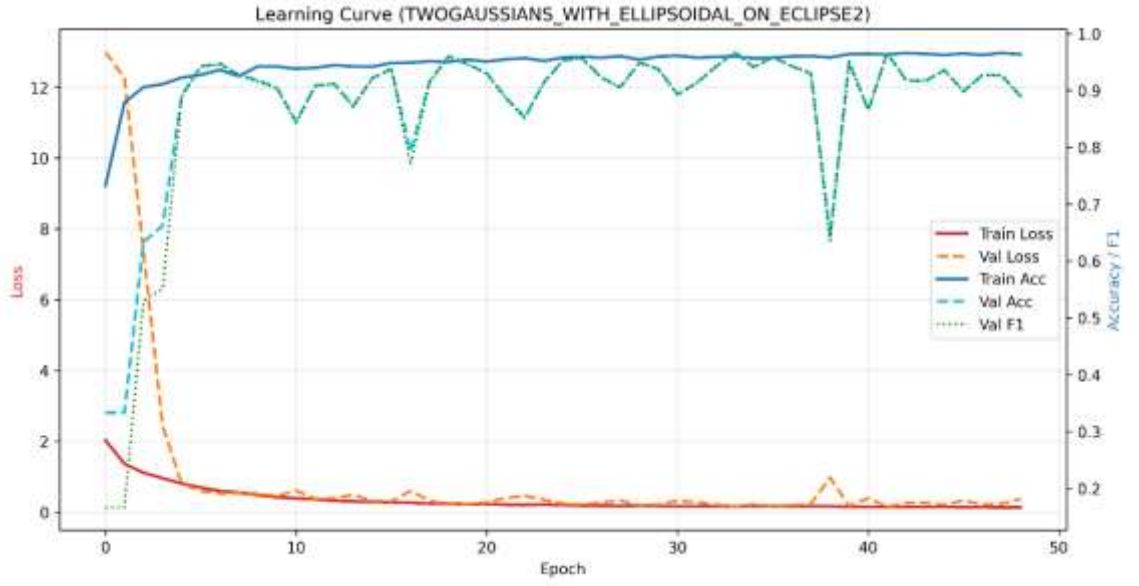
%20'si doğrulama (ing. validation) için kullanıldı. Her bir model türü için yapılan ayrı ayrı eğitimlerin öğrenme eğrileri Şekil 4.2 ile Şekil 4.5 arasında verilmiştir.



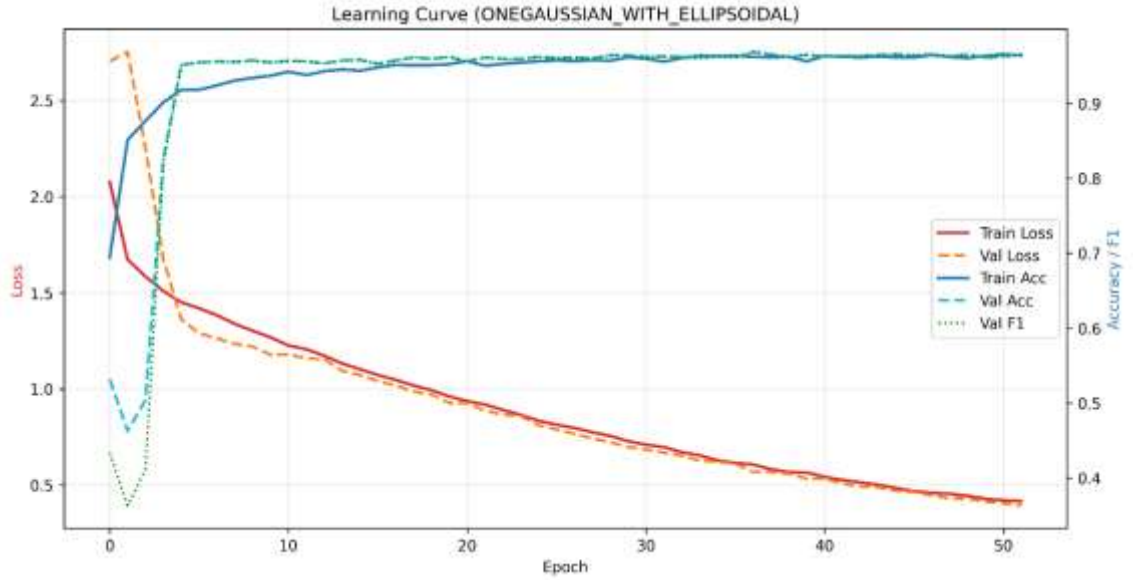
Şekil 4.2 İki Gaussyen (TWOGAUSSIANS) ile modellenen yıldızların eğitim/doğrulama setinde yapılan eğitimin öğrenme eğrisi. Yatay eksen çevrimi, sol düşey eksende kırmızı düz çizgi eğitim kaybını (ing. train loss) ve turuncu kesikli çizgi doğrulama kaybını (ing. val loss), sağ düşey eksende mavi düz çizgi eğitim doğruluğunu (ing. train accuracy), açık mavi kesikli çizgi doğrulama doğruluğunu (ing. validation accuracy) ve yeşil noktalı çizgi doğrulama (ing. val.) F1 skorunu ifade etmektedir.



Şekil 4.3 İki Gaussyen ve 1. Gaussyen üzerinde bir kosinüs (TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE1) ile modellenen yıldızların eğitim/doğrulama setinde yapılan eğitimin öğrenme eğrisi. Eğrilerin anlamları Şekil 4.2'deki gibidir.



Şekil 4.4 İki Gaussyen ve 2. Gaussyen üzerinde bir kosinüs (TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE2) ile modellenen yıldızların eğitim/doğrulama setinde yapılan eğitimin öğrenme eğrisi. Eğrilerin anlamları Şekil 4.2’deki gibidir.



Şekil 4.5 Bir Gaussyen ve bir kosinüs (ONEGAUSSIAN_WITH_ELLIPSOIDAL) ile modellenen yıldızların eğitim/doğrulama setinde yapılan eğitimin öğrenme eğrisi. Eğrilerin anlamları Şekil 4.2’deki gibidir.

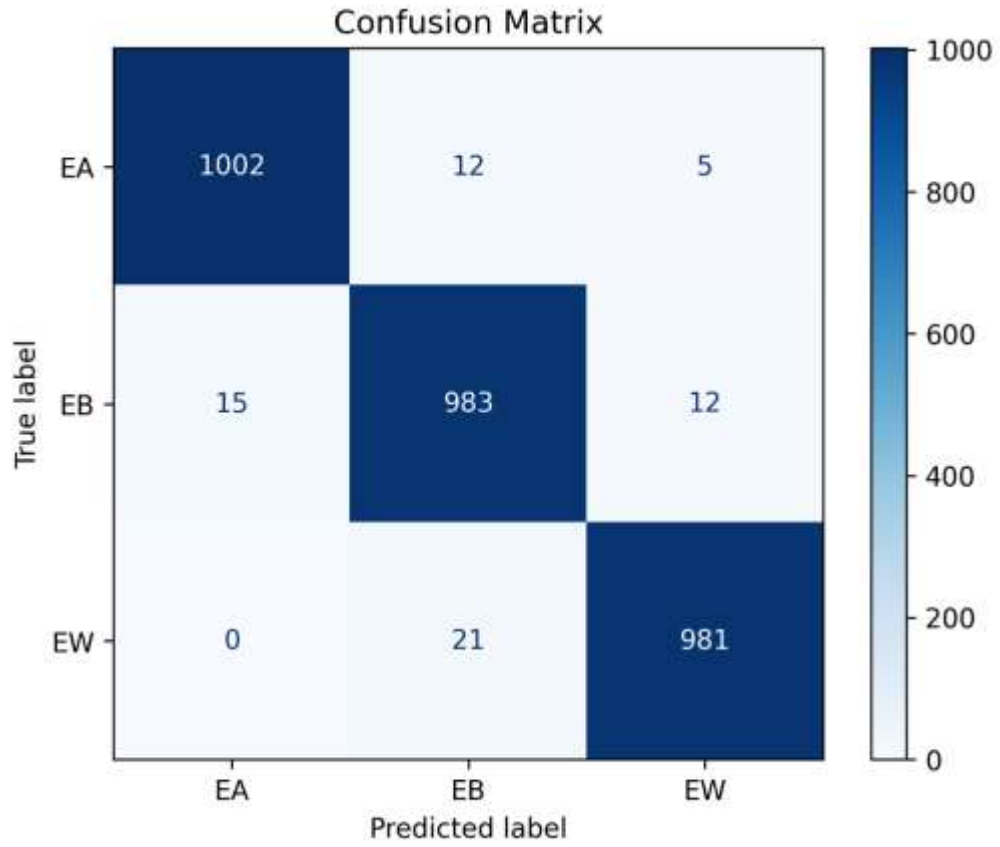
Eğitimlerin sürecinde seyreltme (ing. dropout) regülarizasyonu kullanıldığı için, eğitimlerin erken aşamalarında doğrulama kaybı (ing. validation loss) değeri ile eğitim kaybı (ing. training loss) değeri arasında büyük fark gözlemlenmiştir. Ancak eğitim ilerledikçe (yakınsadıkça), beklenen teorik davranışa uygun olarak doğrulama kaybı, eğitim kaybı ile dengelenmiş ve genelleme aralığı (ing. generalization gap) stabilize olmuştur. Bu durum modelin ezberlemediğini gösterir. Eğitimin başlangıcında görülebilen geçici kayıp artışı modelin veri çeşitliliğine (ing. augmentation, ışık eğrisi görsellerinin yatayda çevrilmesi, sağa/sola veya yukarı/aşağı kaydırılması, yakınlaştırılması) adaptasyon sürecini göstermektedir. Eğitim ilerledikçe doğrulama kaybının eğitim kaybının hemen altında seyretmesi, eğitim sırasında bazı nöronların kapatılması ve veri çeşitliliği uygulanması nedeniyle tahmin (ing. prediction) etmenin zorlaşmasıyla eğitim kaybının daha yüksek, doğrulama aşamasında ise bütün nöronların açık olması ve veri çeşitliliğinin uygulanmaması nedeniyle tahmin etmenin kolaylaşmasıyla doğrulama kaybının daha düşük çıkmasından kaynaklanmaktadır. Bazı nöronların kapatıldığı ve veri çeşitliliğinin uygulandığı makine öğrenmesi modellerinde beklenen ideal durumdur. Doğrulama doğruluğu (ing. validation accuracy) ve doğrulama F1 (ing. validation F1) skorunun zikzaklı yapısı, veri seti her çevrimde (ing. epoch) dinamik olarak değiştiği için modelin performansının anlık olarak dalgalanmasından kaynaklanmaktadır. Bu durum modelin ezberlemediğini, her seferinde yeni bir durumla mücadele ettiğini gösterir. Önemli olan doğruluk çizgilerinin yukarı doğru seyretmesidir.

Makine öğrenmesi modelinin eğitimi test veri setlerinde test edildi ve doğruluk (ing. accuracy), kesinlik (ing. precision), duyarlılık (ing. recall), F1-skor (F1-score), makro (ing. macro) ve ağırlıklı (ing. weighted) ortalama metrikleri hesaplandı.

‘TWOGAUSSIANS’ ile modellenen örten çiftlerin test veri setinde yapılan sınıflandırmanın sınıf bazlı raporu Çizelge 4.3’te, karışıklık matrisi Şekil 4.6’da verilmiştir. Bu test verisinde %98 doğruluk (ing. accuracy) elde edildi.

Çizelge 4.3 ‘TWOGAUSSIANS’ ile modellenen örten çiftlerin test verisinin sınıflandırma raporu

	Precision	Recall	F1-score	Örten çift sayısı
EA	0.99	0.98	0.98	1019
EB	0.97	0.97	0.97	1010
EW	0.98	0.98	0.98	1002
macro avg	0.98	0.98	0.98	3031
weighted avg	0.98	0.98	0.98	3031

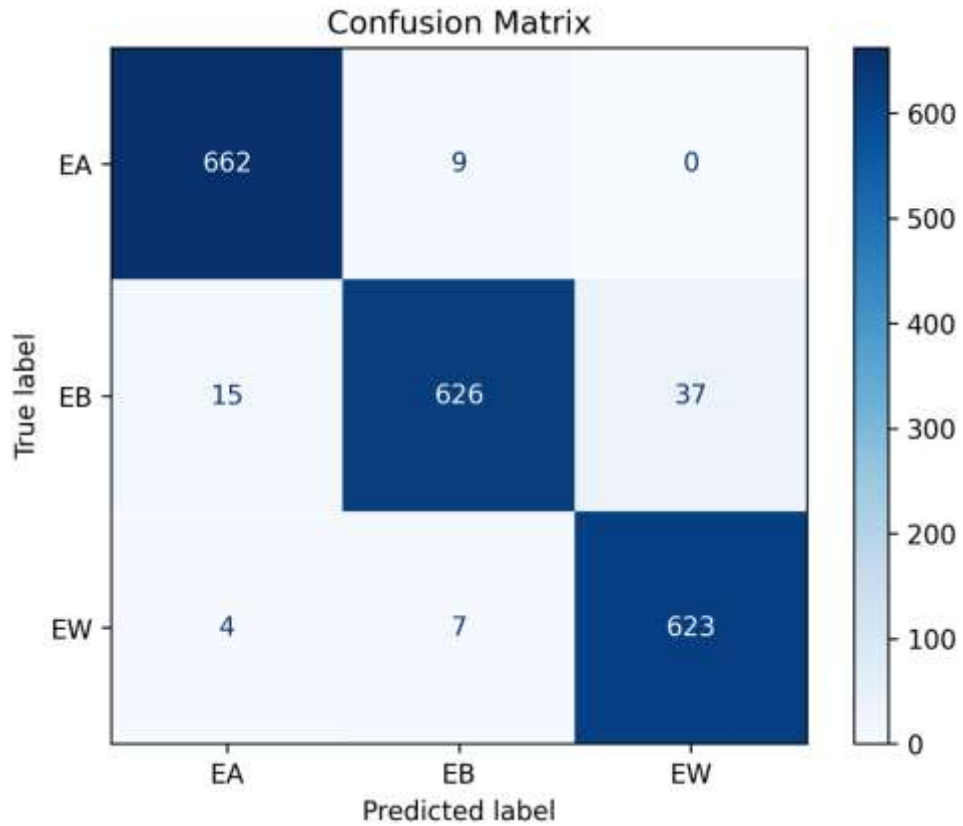


Şekil 4.6 ‘TWOGAUSSIANS’ ile modellenen örten çiftlerin test verisinin karışıklık matrisi

‘TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE1’ ile modellenen örten çiftlerin test veri setinde yapılan sınıflandırmanın sınıf bazlı raporu Çizelge 4.4’te, karışıklık matrisi Şekil 4.7’te verilmiştir. Bu test verisinde %96 doğruluk elde edildi.

Çizelge 4.4 ‘TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE1’ ile modellenen örten çiftlerin test verisinin sınıflandırma raporu

	Precision	Recall	F1-score	Örten çift sayısı
EA	0.97	0.99	0.98	671
EB	0.98	0.92	0.95	678
EW	0.94	0.98	0.96	634
macro avg	0.96	0.96	0.96	1983
weighted avg	0.96	0.96	0.96	1983

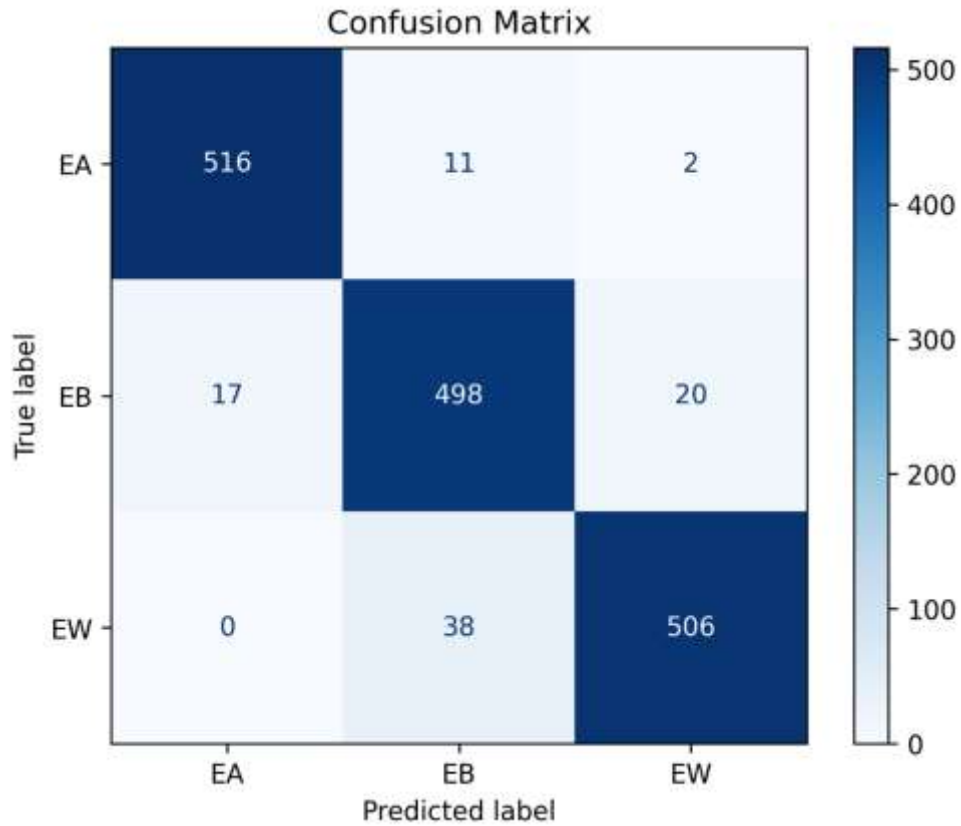


Şekil 4.7 ‘TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE1’ ile modellenen örten çiftlerin test verisinin karışıklık matrisi

‘TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE2’ ile modellenen örten çiftlerin test veri setinde yapılan sınıflandırmanın sınıf bazlı raporu Çizelge 4.5’da, karışıklık matrisi Şekil 4.8’te verilmiştir. Bu test verisinde %95 doğruluk elde edildi.

Çizelge 4.5 ‘TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE2’ ile modellenen örten çiftlerin test verisinin sınıflandırma raporu

	Precision	Recall	F1-score	Örten çift sayısı
EA	0.97	0.98	0.97	529
EB	0.91	0.93	0.92	535
EW	0.96	0.93	0.94	544
macro avg	0.95	0.95	0.95	1608
weighted avg	0.95	0.95	0.95	1608

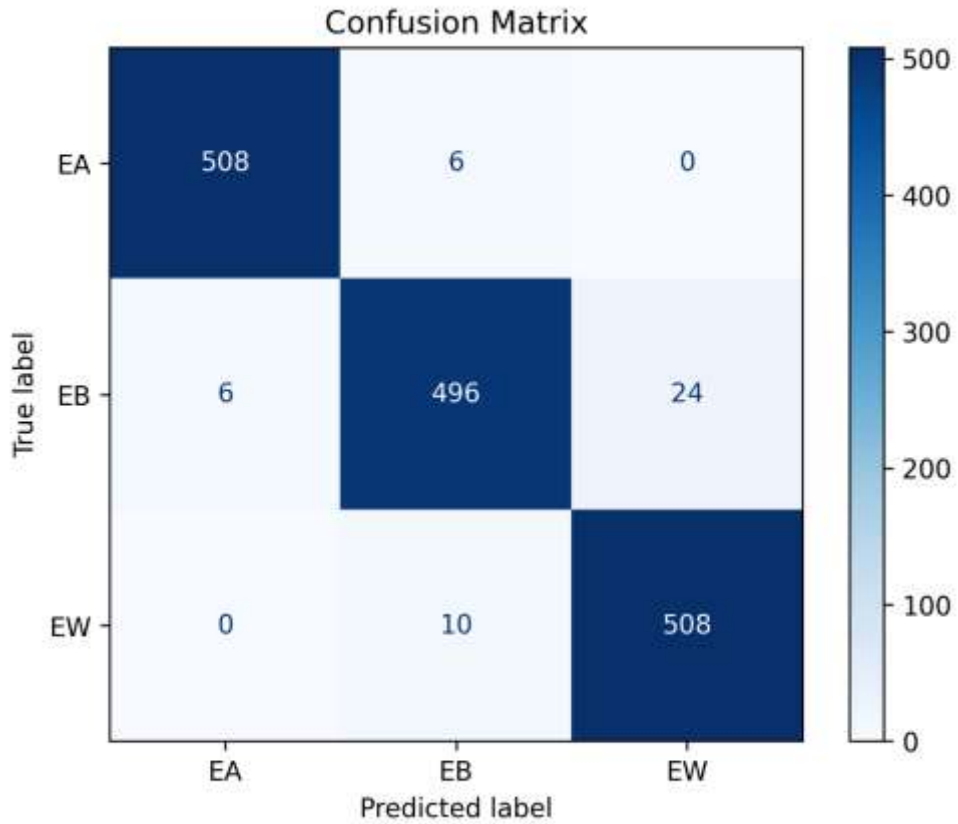


Şekil 4.8 ‘TWOGAUSSIANS_WITH_ELLIPSOIDAL_ON_ECLIPSE2’ ile modellenen örten çiftlerin test verisinin karışıklık matrisi

‘ONEGAUSSIAN_WITH_ELLIPSOIDAL’ ile modellenen örten çiftlerin test veri setinde yapılan sınıflandırmanın sınıf bazlı raporu Çizelge 4.6’da, karışıklık matrisi Şekil 4.9’da verilmiştir. Bu test verisinde %97 doğruluk elde edildi.

Çizelge 4.6 ‘ONEGAUSSIAN_WITH_ELLIPSOIDAL’ ile modellenen örten çiftlerin test verisinin sınıflandırma raporu

	Precision	Recall	F1-score	Örten çift sayısı
EA	0.99	0.99	0.99	514
EB	0.97	0.94	0.96	526
EW	0.95	0.98	0.97	518
macro avg	0.97	0.97	0.97	1558
weighted avg	0.97	0.97	0.97	1558



Şekil 4.9 ‘ONEGAUSSIAN_WITH_ELLIPSOIDAL’ ile modellenen örten çiftlerin test verisinin karışıklık matrisi

Sınıf bazlı raporlarda (Çizelge 4.3 - Çizelge 4.6) performans metriklerinin (Precision, Recall, F1-Score) birbirine yakınsaması ve makro/ağırlıklı ortalamaların eşdeğer çıkması oluşturulan test veri setinin sınıflar arası dengeli (ing. balanced) olduğunu ve modelin hiçbir sınıfa karşı yanlılık (ing. bias) göstermediğini kanıtlamaktadır.

Gaia DR3 arşivinde bulunan toplam 2 184 477 örten çift yıldızdan sınıflandırılmayan ('ONEGAUSSIAN' ve 'ELLIPSOIDAL' ile modellenenler) 73 211 yıldız hariç 2 111 266 örten çift yıldız içeren veri seti içerisinde görsel inceleme ile sınıflandırılan (40 644 eğitim/doğrulama ve 8 180 yıldız test için) yıldızlar hariç geri kalan 2 062 442 örten çift yıldızın tamamı eğitilen makine öğrenmesi modelleri ile ışık eğrisinin biçimine göre EA, EB, EW otomatik olarak sınıflandırıldı. Sınıflandırılan yıldızların sayıları ve oranları herbir model türü için ayrı ayrı Çizelge 4.7'de verilmiştir.

Çizelge 4.7 Gaia DR3 arşivinde 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

Otomatik sınıflandırılan örten çift yıldızların bazı örnekleri Ek 2'de, tamamı Github sayfasında verilmiştir²².

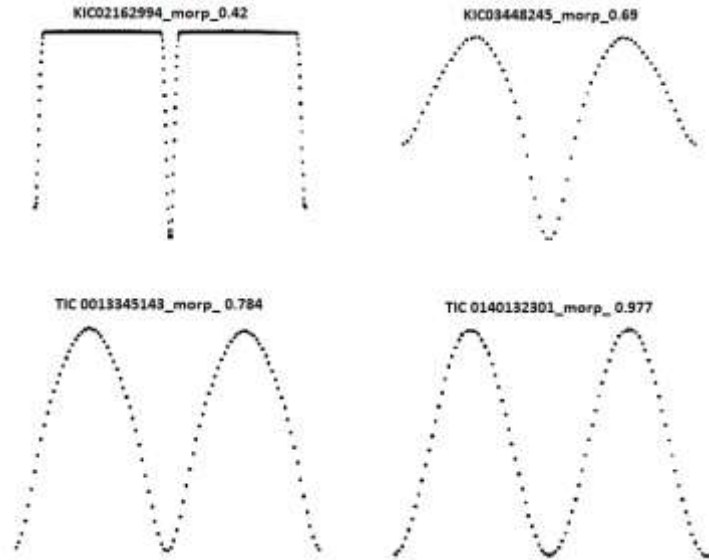
²² <https://github.dev/BedriKeskin/EclipsingBinaryClassification/tree/main/Gaia>

4.2 Görsel Veri ile Sınıflandırılma

Bu bölümde Villanova Kepler ve TESS kataloğunda bulunan örten çift yıldızların ışık eğrisi görselleri ile VGG-19 tabanlı denetimli makine öğrenmesi modeli eğitildi ve eğitilen model ile Gaia DR3 yıldızlarının bir kısmı sınıflandırıldı (Bölüm 3.3.2).

Villanova Kepler kataloğundan 2907 adet, TESS kataloğundan 4349 adet örten değişen çift yıldız ışık eğrisi verisi indirildi. Villanova kataloğunda her bir örten değişen yıldız için morfoloji (morph) parametresi atanmıştır. Morfoloji parametresi 0 ile 1 arasında değer alır.

Bölüm 3.1.1’de belirtildiği şekilde Matijević vd. (2012) morfoloji parametresi kriterlerine göre Kepler ve TESS örten çift yıldızları etiketlendi. Işık eğrileri Villanova kataloğundan alınan ışık elemanları ile tek çevrim olarak katlandı (phase folded) ve veri setindeki tutarlılığı sağlamak için üstüste bindirildi (binlendi). Örnek ışık eğrileri Şekil 4.10’da gösterilmiştir.



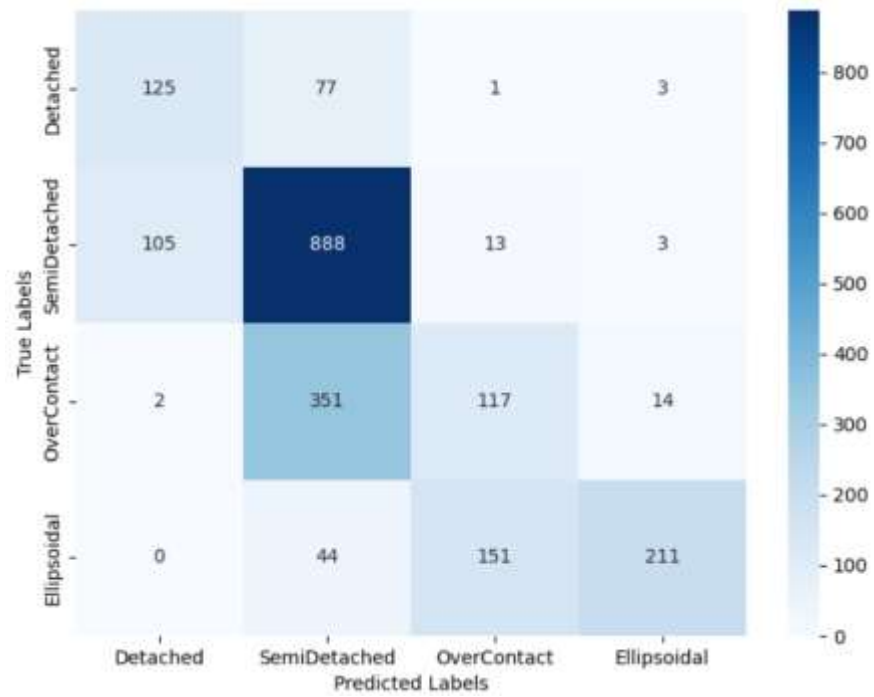
Şekil 4.10 Kepler ve TESS arşivlerinden alınan ayırık (sol üst), yarı-ayırık (sağ üst), aşırı degen (sol alt) ve elipsoidal (sağ alt) örnek ışık eğrileri

Sınıflandırma için Gaia DR3 arşivinden 2106 adet örten çift yıldız seçildi. Seçilen yıldızlar Gaia DR3 arşivi ile Kepler ve TESS arşivlerinde ortak bulunan yıldızlardır. Bu

eşleşme SIMBAD²³ veri tabanı kullanılarak yapıldı. Gaia DR3 arşivindeki 2106 örten çiftin 789 adeti Kepler arşivinde, 1317 adeti TESS arşivinde bulunmaktadır. Dolayısıyla 2106 yıldızın hangi etikete sahip olduğu bilinmekte ve böylece etiketlenen sınıf ile tahmin edilen sınıf arasında ilişki kurulabilmektedir.

Villanova kataloğundan alınan 2907 adet Kepler ve 4349 adet TESS örten değişen çift yıldızın PNG formatında ışık eğrisi çizildi. Sonrasında veri seti %67 eğitim, %25 doğrulama ve %8 test olmak üzere 3 gruba ayrıldı.

Kepler ve TESS veri seti ile VGG-19 tabanlı model eğitildi. Test verisinde %91 doğruluk (ing. accuracy), 2106 adet Gaia DR3 örten çift yıldızlarının tahmininde ise %64 doğruluk elde edildi. Karışıklık matrisi Şekil 4.11’te verilmiştir.

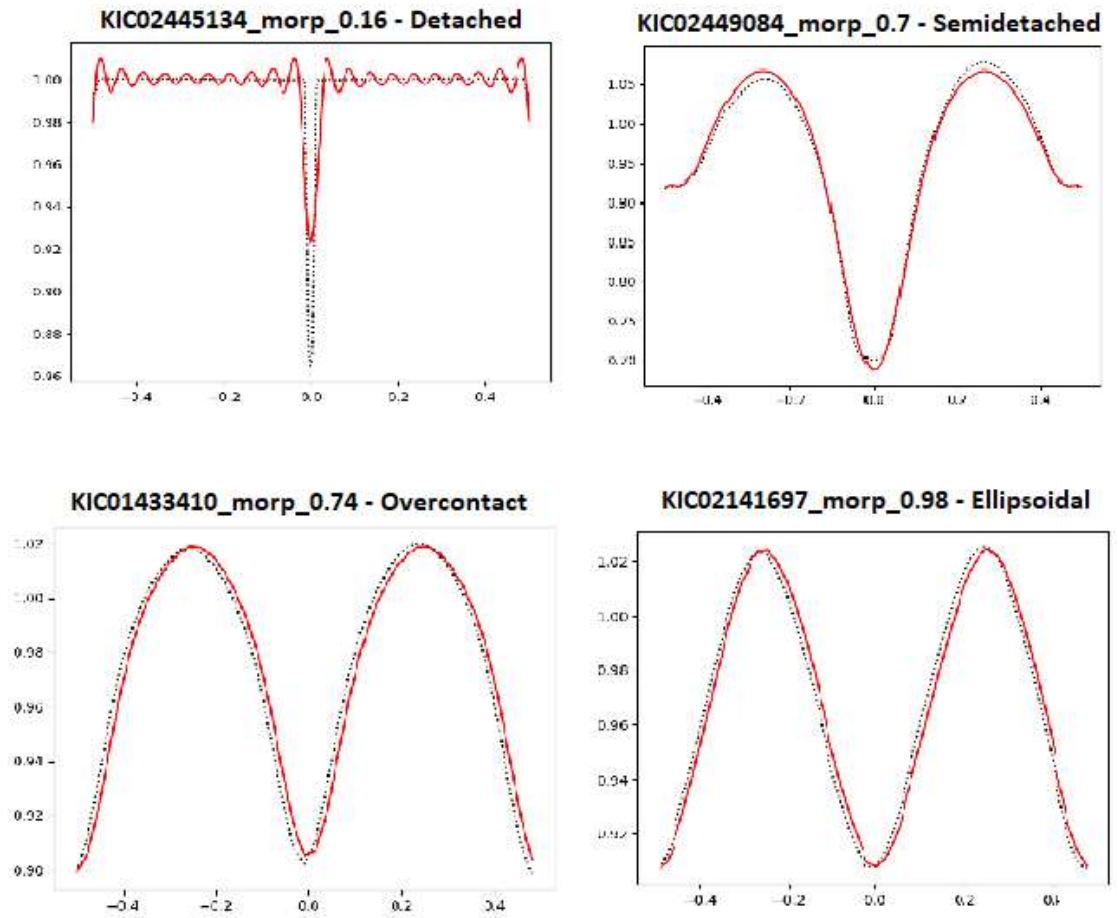


Şekil 4.11 Kepler ve TESS arşivlerinde ortak bulunan Gaia DR3 örten çift yıldızların sınıflandırılması sonucu elde edilen karışıklık matrisi

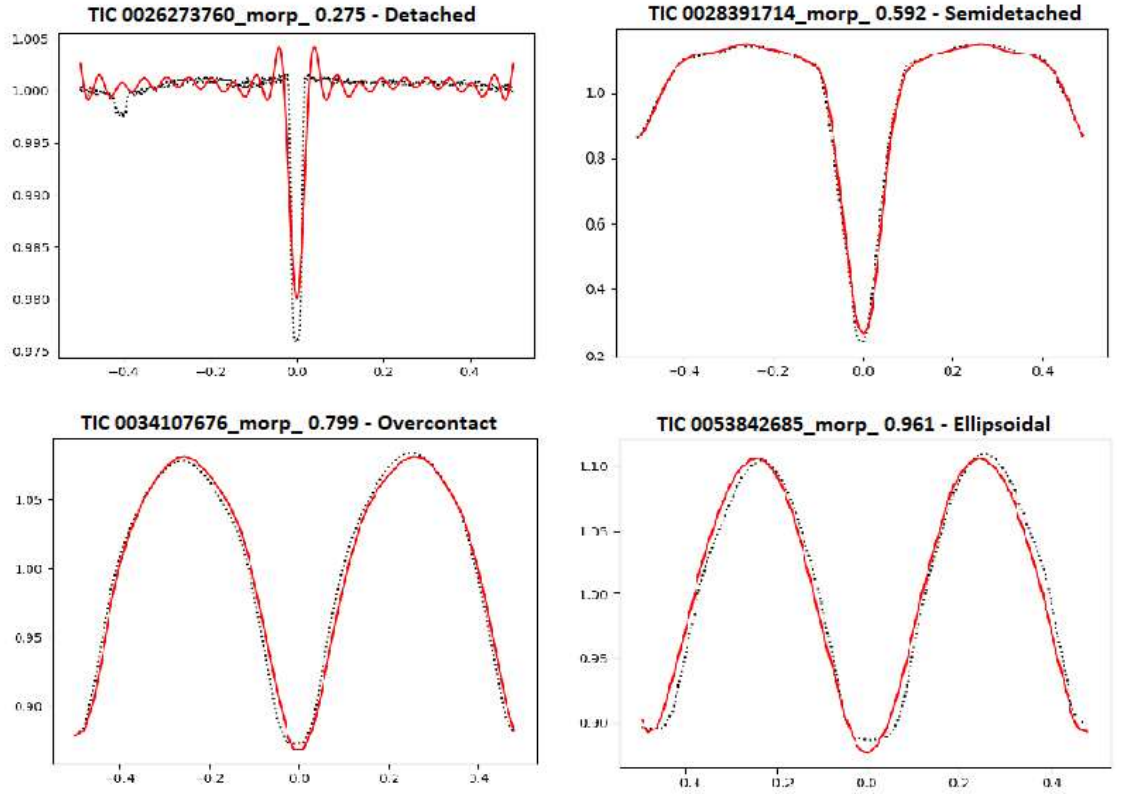
²³ <http://simbad.u-strasbg.fr/simbad/>

4.3 Fourier Katsayılarına Göre Sınıflandırma

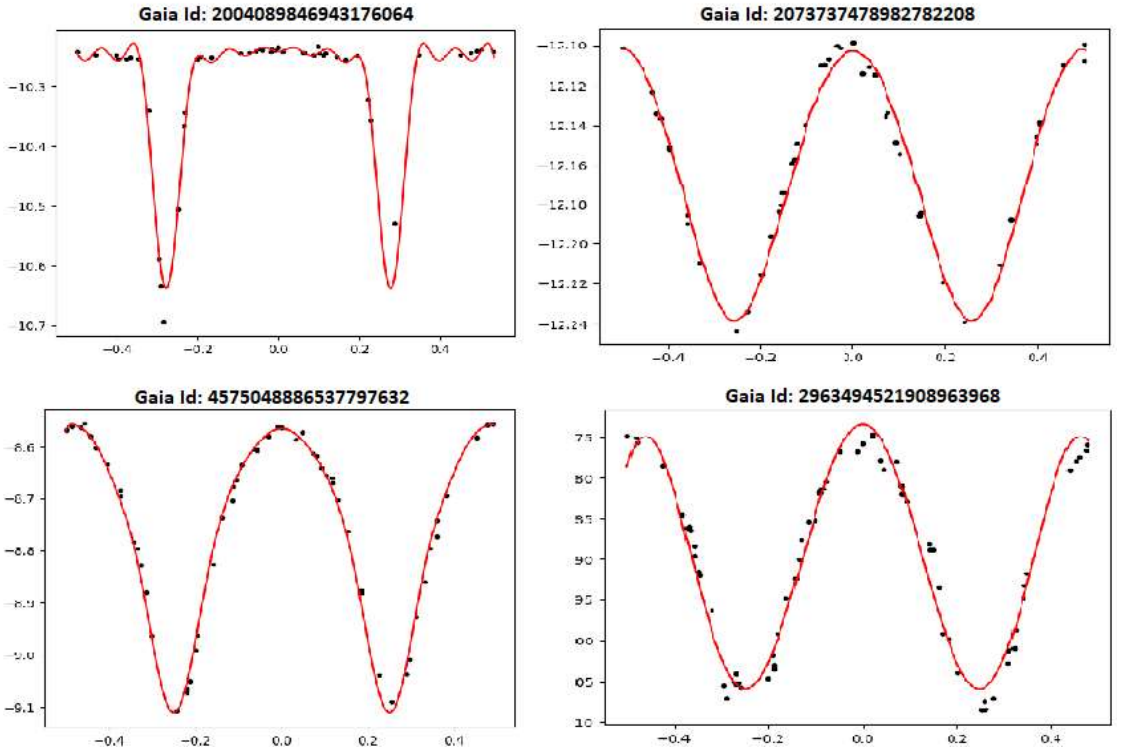
Bu tez çalışmasında Fourier katsayıları kullanılarak da sınıflandırma yapıldı. Villanova Kepler (Şekil 4.12), TESS (Şekil 4.13) ve Gaia DR3 (Şekil 4.14) birer çevrim katlanmış ışık eğrileri 11 kosinüs terimli Fourier serisi ile Eşitlik 3.1'e göre modellendi. Kepler ve TESS örten çift yıldızlarının Fourier katsayıları ile bir boyutlu CNN makine öğrenmesi modeli eğitildi ve eğitilen model ile Gaia DR3 örten çift yıldızları tahmin edildi (Bölüm 3.3.3).



Şekil 4.12 Villanova Kepler ışık eğrilerinin Fourier serileri ile modellenmesi

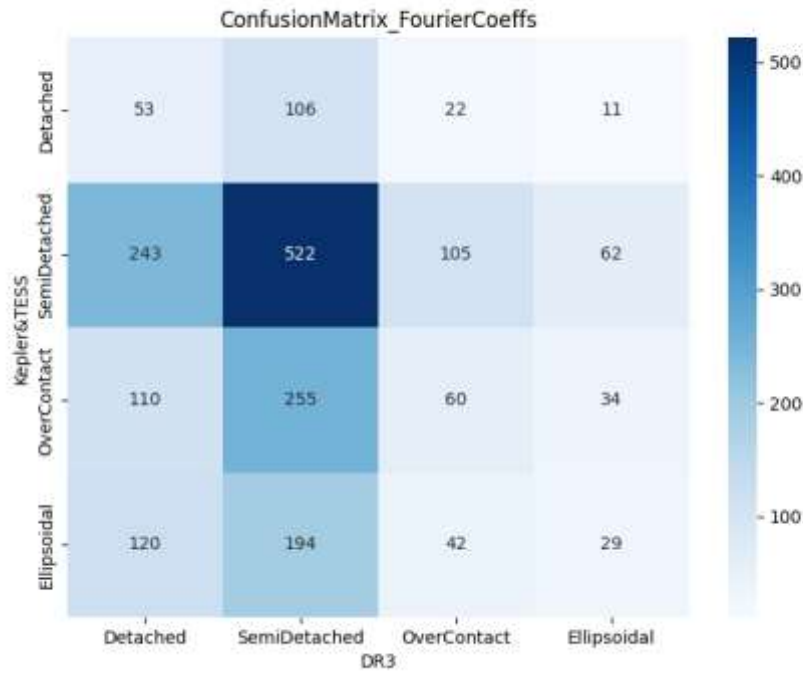


Şekil 4.13 Villanova TESS ışık eğrilerinin Fourier ile modellenmesi



Şekil 4.14 Gaia DR3 ışık eğrilerinin Fourier ile modellenmesi

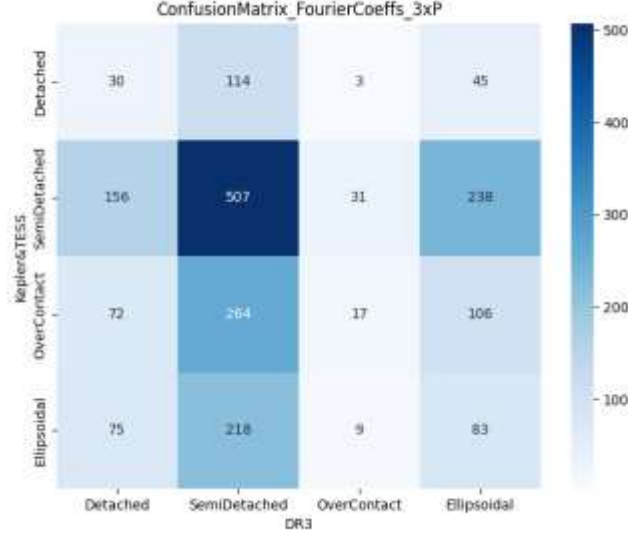
Kepler ve TESS ışık eğrilerinin Fourier katsayıları ile eğitilen makine öğrenmesi modeli test verisinde %65 doğruluk elde edildi. Gaia DR3 ışık eğrilerinde yapılan sınıflandırmada ise %33 doğruluk elde edildi (Şekil 4.15). Aslında bu sonucun çok değerli bir negatif sonuç olduğu düşünülmektedir. Rucinski (1993)'nin Fourier katsayıları ile çift yıldızların ışık eğrisi çözümlerinin sadece W UMa türü için geçerli olması bu negatif sonuç ile desteklendiği değerlendirilmektedir.



Şekil 4.15 Gaia DR3 ışık eğrilerinin Fourier ile sınıflandırılması sonucu elde edilen karışıklık matrisi

11 kosinüs terimli Fourier serileri EW türü örten çiftleri yüksek başarımla modellerken EA ve EB türü örten çift yıldızları benzer başarımla modelleyemediği görüldü. EW türündeki ışık eğrilerinin minimum derinliklerinin hemen hemen aynı derinlikte olması nedeniyle Fourier modelleri yörünge döneminin yarısı bir dönemlilik belirlerken diğer türlerde minimum derinliklerinin farklı olması nedeniyle yörünge dönemi doğru belirleniyor olabilir. Bu da EW ışık eğrilerinin diğerlerinden ayrılmasını mümkün kılıp, yanlış bir dönemliliğe (dönemin yarısına) dayanmasına rağmen doğru sınıflandırılmasını sağlıyor olmalıdır.

Aynı süreç Kepler, TESS ve Gaia DR3 örten çift yıldızlarının 3 kez katlanmış ışık eğrilerine de uygulandı ve Şekil 4.16’da görüldüğü gibi benzer sonuçlar elde edildi.



Şekil 4.16 3 kez katlanmış Gaia DR3 ışık eğrilerinin Fourier ile sınıflandırılması sonucu elde edilen karışıklık matrisi

4.4 Örten Çift Yıldızların Kümelendirilmesi

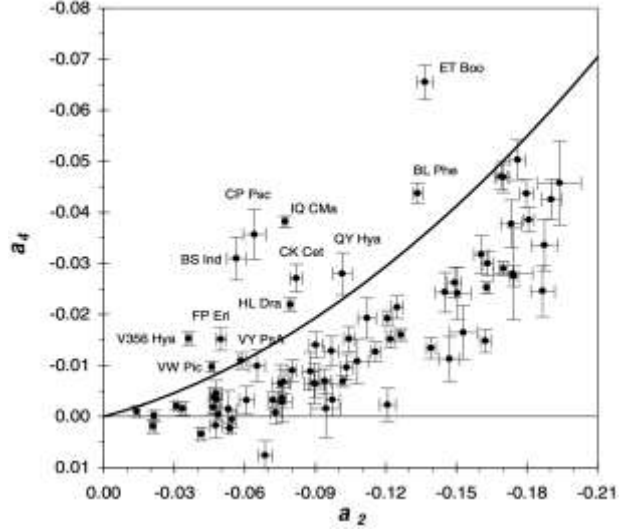
Bu tez çalışmasında kümelendirme yapılarak örten çift yıldızların Fourier katsayılarına göre dağılımı da incelendi. Kümelendirme, veri setindeki doğal grupları ve gizli yapıları ortaya çıkarır. Bir veri setindeki veri noktalarını aralarındaki benzerliklere göre gruplara ayırır. 2748 adet Kepler örten çift yıldızının Fourier a_1 , a_2 , a_4 katsayıları, GMM yöntemiyle kümelendirildi (Bölüm 3.3.4).

Fourier a_1 , a_2 , a_4 katsayılarına göre (Selam 2004);

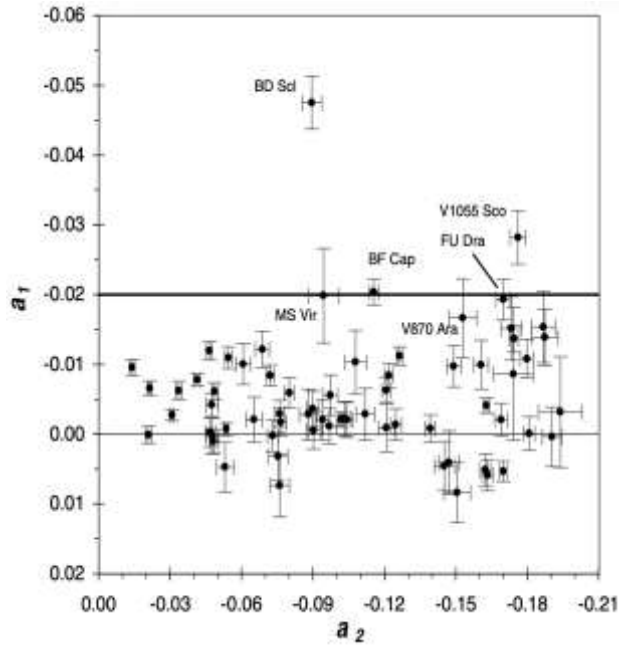
- $a_4 < a_2$ ($0.125 - a_2$) ilişkisinden EA'lar EB ve EW'lerden ayıklanır.
- $a_1 < -0.02$ ilişkisinden EB'ler EW'lerden ayıklanır.

Fourier a_2 ve a_4 katsayılarına göre ayırık çiftler, yarı-ayırık ve değen çiftlerden ayrılır (Şekil 4.17). Fourier a_1 ve a_2 katsayılarına göre de yarı-ayırık çiftler değen çiftlerden ayrılır (Şekil 4.18). Dolayısıyla örten çift yıldızları, Fourier a_2 ve a_4 katsayılarına göre EA'lar ve

diğerleri (EB ve EW'lar), Fourier a_1 ve a_2 katsayılarına göre de EB'ler ve EW'lar şeklinde iki aşamada türlerine ayrılabilir. Buna kural tabanlı sınıflandırma denilir.

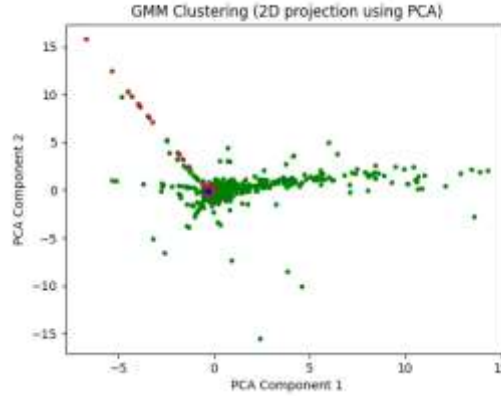


Şekil 4.17 Fourier a_2 ve a_4 katsayılarına göre yarı-ayrık çiftlerin yarı-ayrık ve değen çiftlerden ayrılması. EA türü çiftler kalın eğrinin üstünde, EB ve EW türü çiftler altında yer alır (Selam 2004).



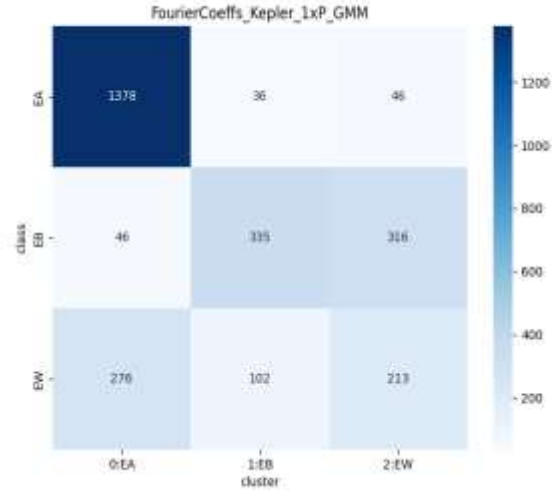
Şekil 4.18 Fourier a_1 ve a_2 katsayılarına göre yarı-ayrık çiftlerin değen çiftlerden ayrılması. EB türü çiftler kalın çizginin üstünde, EW türü çiftler altında yer alır (Selam 2004).

Yıldızları, Selam (2004)'de belirtildiği şekilde, a_1 , a_2 ve a_4 Fourier katsayılarına göre EA, EB, EW şeklinde 3 adete sınıfa ayırdığımız için kümelendirmede de GMM yöntemi yıldızları 3 kümeye ayıracak şekilde ayarlandı. Kepler örten çiftlerinin a_1 , a_2 , a_4 Fourier katsayıları “Standart Scaler” ile standart hale getirildi. Ardından PCA ile 2 boyuta indirildi. Elde edilen kümeleme grafiği Şekil 4.19’te gösterilmiştir.



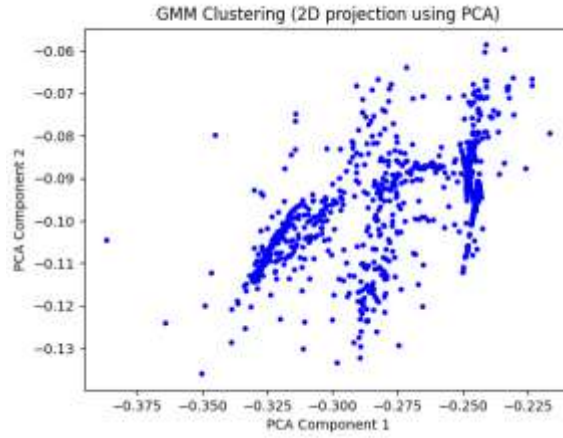
Şekil 4.19 Kepler örten çift yıldızlarının GMM ile kümelendirilmesi

Şekil 4.19’da mavi ile gösterilen 0. kümede 1700 yıldız bulunmakta ve EA sınıfına karşılık gelmektedir, dar bir alanda toplandıkları için nokta gibi gözükmektedir. Kırmızı ile gösterilen 1. kümede 473 yıldız bulunmakta ve EB sınıfına karşılık gelmektedir. Yeşil ile gösterilen 2. kümede 575 yıldız bulunmakta ve EW sınıfına karşılık gelmektedir. EA kümesi belli bir alanda toplanmışken EB ve EW kümelerinde saçılma gözlenmiştir. Buna göre çizilen karışıklık matrisi Şekil 4.20’de gösterilmiştir. Bu matris GMM kümelemesi ile kural tabanlı sınıflandırmanın ne kadar uyumlu olduğunu gösterir. Matrisin köşegeni üzerindeki yüksek sayılar uyumu, dışındaki sayılar ise GMM'nin farklı bir şekilde grupladığını gösterir.

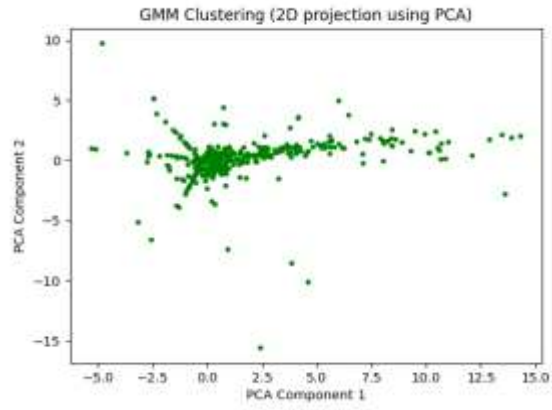


Şekil 4.20 Kepler örten çift yıldızlarının Fourier katsayılarına göre sınıflandırılması ile kümelendirilmesinin karşılaştırılmasını gösteren karışıklık matrisi

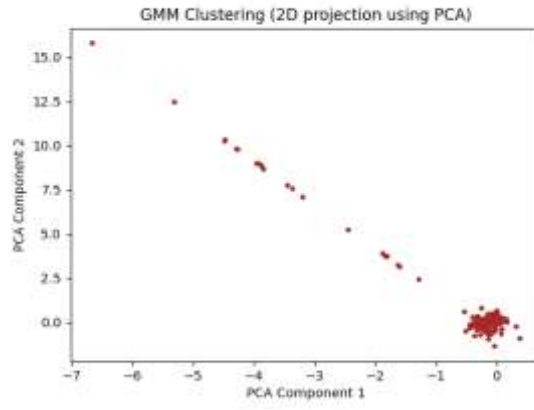
Şekil 4.20’de beraber gösterilen 3 kümenin grafiği daha anlaşılır olması için ayrı ayrı grafikleri Şekil 4.21, Şekil 4.22 ve Şekil 4.23’te gösterilmiştir.



Şekil 4.21 Kepler örten çift yıldızlarında 0. kümenin dağılımı



Şekil 4.22 Kepler örten çift yıldızlarında 1. kümenin dağılımı



Şekil 4.23 Kepler örten çift yıldızlarında 2. kümenin dağılımı

5. TARTIŞMA ve SONUÇ

Bu tez çalışmasında, Gaia DR3, Kepler ve TESS gibi büyük hacimli astronomi veri arşivlerindeki örten çift yıldızların, makine öğrenmesi teknikleri kullanılarak otomatik olarak algılanması ve sınıflandırılması hedeflenmiştir. Bu amaç doğrultusunda, hem denetimli öğrenme (çok modlu model, VGG-19 ile transfer öğrenme, Fourier katsayıları) hem de denetimsiz öğrenme (GMM ile kümeleme) yöntemleri uygulanmış ve elde edilen sonuçlar karşılaştırmalı olarak analiz edilmiştir.

Çalışmada en başarılı sonuç çok modlu (multimodal) derin öğrenme modelinde elde edilmiştir. Gaia DR3 örten çift yıldızlarının sınıflandırılması için geliştirilen bu model, Mowlavi vd. (2023) geometrik model parametrelerini (sayısal veri) ve bu parametrelerden üretilen ışık eğrisi görsellerini (görsel veri) aynı anda işlemiştir. İki Gaussyen ile modellenen yıldızlarda %98, iki Gaussyen ve 1. Gaussyen üzerinde bir kosinüs ile modellenen yıldızlarda %96, iki Gaussyen ve 2. Gaussyen üzerinde bir kosinüs ile modellenen yıldızlarda %95, bir Gaussyen ve bir kosinüs ile modellenen yıldızlarda %97 doğruluk elde edildi. Bu sonuç farklı veri türlerini (sayısal ve görsel) birleştiren çok modlu yaklaşımların, tekil veri türlerine dayalı yöntemlere kıyasla üstünlüğe sahip olduğunu kanıtlamıştır.

Mowlavi vd. (2023)'nin model parametrelerinden oluşan sayısal veri, geleneksel ışık eğrisi biçimine göre sınıfları (EA, EB, EW) tanımlayan temel öznitelikleri (tutulma derinliği ve süresi gibi) ihtiva etmektedir. Buna ek olarak, CNN'in ışık eğrisi görsellerinden öznitelik çıkarımıyla veri setinin zenginleşmesiyle çok modlu model yüksek başarıma ulaşmıştır. Bu sonuç EA, EB, EW sınıflandırmasının örten çift sistemin temel fiziksel bilgilerini içerdiğini göstermektedir.

Bu yüksek başarı sayesinde, Gaia DR3 arşivindeki yaklaşık 2 milyon örten çift yıldız, eğitilen çok modlu makine öğrenmesi modeli kullanılarak otomatik olarak sınıflandırılmıştır. Bu tez çalışması, Gaia DR3 örten çift yıldızlarının ışık eğrilerini morfolojilerine göre alt sınıflarına ayıran literatürdeki ilk çalışmadır.

Roche geometrisine dayalı sistem morfolojisi parametresine göre (Ayrık, Yarı-ayrık, Aşırı Değen, Elipsoidal) sınıflandırma amacıyla, önceden eğitilmiş VGG-19 modeli ile transfer öğrenme yaklaşımı test edilmiştir. Sadece ışık eğrisi görselleri kullanılarak yapılan bu sınıflandırmada, ışık eğrisinin biçimi ile sistemin Roche geometrisi arasında ilişki kurulmaktadır. Sistem morfoloji parametresine göre (Ayrık, Yarı-ayrık, Aşırı Değen, Elipsoidal) sınıflandırma, ışık eğrisinin biçimine dayalı sınıflandırmadan (EA, EB, EW) daha karmaşıktır. VGG-19 modeli, Kepler ve TESS arşivlerinden alınan ışık eğrisi verileri ile oluşturulan PNG görsellerle eğitilmiş ve test verisinde %91 doğruluk ile yüksek bir başarı elde etmiştir. Ancak, bu model Kepler ve TESS ile ortak olan 2106 adet Gaia DR3 yıldızına uygulandığında, doğruluk oranı %64'e gerilemiştir. Bu performans düşüşünün temel nedeni, modelin eğitiminde kullanılan Kepler ve TESS verileri ile hedeflenen Gaia DR3 verilerinin farklı veri işleme ve modelleme yöntemlerinden gelmesidir. Kepler ve TESS ışık eğrisi görselleri zaman ve akı değerlerinden oluşan ışık eğrisi verisi kullanılarak elde edilmişken, Gaia DR3 ışık eğrisi görselleri ise Mowlavi vd. (2023)'nin geometrik model parametrelerinden türetilmiştir. Veri temsili ve etiketin kaynağı arasındaki bu temel fark transfer öğrenmenin önündeki büyük bir engeldir. Transfer öğrenme için veri homojenliği şarttır. Test verisinde yüksek, ancak Gaia DR3 yıldızlarında düşük doğruluk elde edilmesi makine öğrenmesi modelinin değil, veri setlerinin doğasından kaynaklanan bir sorundur. Bu durum, farklı arşivlerden elde edilen verilerle transfer öğrenme yapılırken veri homojenliğinin ve model tutarlılığının kritik önem taşıdığını göstermiştir.

Fourier katsayılarına dayalı sınıflandırma yöntemi çalışmadaki en düşük performansı sergilemiştir. Kepler ve TESS verileriyle eğitilen model test verisinde %65 gibi düşük bir doğruluk sağlarken, Gaia DR3 verilerinde sadece %33 doğruluk elde edebilmiştir. Aslında bu negatif sonuç beklenen sonuçtur ve astrofizik açıdan değerlidir. Rucinski (1993) 11 kosinüs terimli Fourier serisinin simetrik yapıdaki EW türü yıldızları modellemede başarılı olduğunu, ancak farklı minimum derinliklerine sahip EA ve EB türü yıldızların ışık eğrisi şekillerini temsil etmekte yetersiz kaldığı gösterilmiştir.

Bu çalışma ile literatürdeki denetimli öğrenme ile sınıflandırma yapan benzer çalışmaların karşılaştırması Çizelge 5.1'de verilmiştir. Bu tez çalışmasında sentetik veri

kullanılmamıştır. Čokina vd. (2021) eğitim, doğrulama ve test için sentetik veri kullanmıştır. Gözlemsel verilerin yetersiz olduğu durumlarda yazılımla sentetik veri üretilmesi model eğitimi için gerekli olan büyük veri setlerinin oluşturulmasını sağlar. Bu veri azlığı sorununa pratik bir çözümdür. Ancak sentetik veride yüksek doğruluk beklenen bir durumdur ve tek başına bir başarı kriteri değildir. Sentetik veri setleri üzerinde elde edilen yüksek doğruluk oranları modelin gerçek veri üzerindeki performansını garanti etmez.

Çizelge 5.1 Bu çalışma ile literatürdeki benzer çalışmaların karşılaştırması

Çalışma	Veri Arşivi	Sınıflandırma Türü	Yöntem	Doğruluk
Bu Çalışma	Gaia DR3	Işık Eğrisi Biçimi (EA, EB, EW)	Çok Modlu CNN + MLP	%95 - %98 (Test Verisinde)
Bu Çalışma	Kepler, TESS, Gaia DR3	Sistem Morfolojisi (D, SD, OC, E)	VGG-19 Transfer Öğrenme	%91 (Test Verisinde), %64 (Gaia Tahmini)
Bu Çalışma	Kepler, TESS, Gaia DR3	Sistem Morfolojisi (D, SD, OC, E)	1D CNN (Fourier Katsayıları)	%65 (Test Verisinde), %33 (Gaia Tahmini)
Čokina vd. (2021)	Sentetik Veri	Sistem Morfolojisi (Detached, Contact)	BiLSTM + 1D CNN	%98 (Yarı-ayrık dahil), %100 (Yarı-ayrık çıkarılınca)
Parimucha vd. (2025)	Sentetik Veri, OGLE, DEBCat, WUMaCat	Sistem Morfolojisi (Detached, Overcontact)	ResNet50	%94 - %100 (Gerçek Veri)
Süveges vd. (2017)	HIPPARCOS, CALEB, Kepler	Işık Eğrisi (EA, EB, EW) & Sistem (D, SD, OC)	PCA, LDA, RF, SOM	%95 (EA, EB, EW-HIPPARCOS), %89,5 (D, SD, OC-CALEB)

Parimucha vd. (2025) eğitim ve doğrulamada sentetik, testte ise gerçek veri kullanmıştır. Gerçek test veri seti üzerinde sınıflandırmada yüksek doğruluğa ulaşmıştır ancak leke tespiti gibi daha ince fotometrik detayların analizinde sentetik veri ile eğitimin yetersiz kaldığı görülmüştür.

Süveges vd. (2017) hem model eğitimi hem de test için fiziksel parametreleri ve sınıfları iyi bilinen CALEB (306 çift yıldız) ve HIPPARCOS (521 çift yıldız) kataloglarını kullanmıştır. Ancak bunlar küçük veri setleridir. Sadece birkaç yüz nesneden oluşan eğitim setleri milyonlarca nesnenin sınıflandırılması için yetersiz kalabilir. Kepler veri setini ise test için kullanmıştır. Kepler veri setinde aşırı değen çiftlerde %91, ayırık çiftlerde %85, yarı-ayırık çiftlerde %55 doğruluk elde etmiştir. Yarı-ayırık çiftlerde %55 doğruluk düşük bir başarıdır. Bunun nedeni CALEB ve Kepler veri setleri arasındaki uyumsuzluktur. CALEB kataloğu belirgin tutulmalara sahip yıldızlardan oluşmakta iken Kepler sık ve dar tutulmalara sahip yıldızları da içermektedir. CALEB ile eğitilen model, Kepler'deki ayırık sistemleri yarı-ayırık olarak sınıflandırmaktadır.

Parimucha vd. (2025) sentetik veriden gerçeğe geçerken leke tespitinde (%50 doğruluk) sorun yaşamıştı. Süveges vd. (2017) ise gerçek veriden (CALEB) başka bir gerçek veriye (Kepler) geçerken yarı-ayırık sistemlerde (%55 doğruluk) sorun yaşamıştır. Bu durum eğitim verisinin sadece gerçek olmasının yetmediğini, aynı zamanda hedefte sınıflandırılacak veri setini iyi temsil etmesi gerektiğini göstermiştir.

Denetimsiz öğrenme ile Kepler örten çift yıldızlarının Fourier katsayılarına (a_1 , a_2 , a_4) göre doğal dağılımı da incelenmiştir. GMM yöntemiyle elde edilen 3 küme, Selam (2004) kriterlerine göre yapılan EA, EB, EW sınıflandırmasıyla önemli ölçüde örtüşmüştür. Özellikle EA sınıfına (0. küme) ait yıldızların PCA uzayında yoğun ve belirgin bir küme oluşturması bu yıldızların Fourier katsayıları temelinde diğerlerinden net bir şekilde ayrıldığını göstermiştir. EB ve EW sınıflarının ise daha dağınık bir yapı sergilemesi bu iki sınıf arasındaki ayrımın daha karmaşık olduğunu ve sadece bu katsayılara dayalı bir ayrımın zorluklar içerdiğini göstermiştir.

GMM kümelemesi EA türü yıldızları diğerlerinden ayırmada çok başarılı olmuştur (1378 doğru eşleşme). PCA düzleminde bu yıldızların dar bir alanda toplanması EA sistemlerinin ışık eğrisi morfolojisinin (düz maksimumlar, minimumlar arasındaki derinlik farkı, tutulmaların başlangıç ve bitişinin belirginliği) Fourier katsayılarında ayırt edici bir etki bıraktığını göstermektedir. Bu sonuç Fourier katsayıları kullanılarak

denetimsiz öğrenme yöntemlerinin EA türü sistemleri ayırt etmede yüksek güvenilirlikle kullanılabileceğine işaret etmektedir.

EB ve EW sınıfları PCA düzleminde daha dağınık ve birbirine girmiş durumdadır. EB sınıfının önemli bir kısmı (316 yıldız) GMM tarafından EW kümesine (2. küme) dahil edilmiştir. Yarı-ayrık (EB) ile değen (EW) sistemler arasındaki fiziksel sınırların belirgin olmadığını göstermektedir. EB ve EW sistemlerinin birbirine benzer ıřık eğrileri bu iki sınıfı PCA düzleminde birbirine yaklařtırmıştır.

Rucinski (1993)'nin Fourier serisi ile ıřık eğrisini temsil etme yöntemi W UMa türü sistemler için geçerlidir. 11 terimli Fourier kosinüs serisi özellikle Algol (EA) türü sistemlerin karakteristięi olan minimumlardaki düzlükleri ve tutulmaların başlangıç ve bitiş belirginliğini modellemede yetersizdir. GMM yönteminin EA türünü belirgin kümelemesi ancak EB ve EW türlerini daęınık ve birbirine girmiş kümelemesi Rucinski (1993)'nin yönteminin sadece W UMa türü sistemler için geçerli olmasını destekledięi düşünölmektedir.

Fourier serisinin sadece W UMa türünü iyi temsil etmesi ile Fourier katsayılarına göre kümelemenin sadece Algol türünü iyi kümelemesi tezat gibi görünmektedir. Dięer bir ifadeyle, Fourier serisi W UMa türünü iyi temsil ediyorsa Fourier katsayılarının daęılımında W UMa türünün belirgin bir küme oluřturması beklenebilir ancak Algol türü belirgin bir küme oluřturmuřtur. Bu çeliřki gibi görönen durum, Fourier serisinin modelleme (temsil) başarısı ile Fourier katsayılarının sınıfları ayırt edebilirlięi farklı kavramlardır řeklinde açıklanabilir. Yani Fourier serisi W UMa türünü matematiksel olarak iyi modelleyebilir, ancak bu durum W UMa yıldızlarının Fourier katsayılarının belirgin bir küme oluřturacaęı anlamına gelmez.

Algollerin belirgin bir küme oluřturmasının sebebi hepsinde Fourier a_2 katsayısı sıfıra yakındır. Fourier a_2 katsayısı ıřık eğrisinde sürekli deęiřimi temsil eder. Beta Lyrae ve W UMa ıřık eğrilerinde sürekli deęiřim gözlenir ancak Algollerde gözlenmez. Bu Algollerin ayrık, dolayısıyla bileřen yıldızların küresel řeklini korumasından kaynaklanır. Rucinski (1993) Fourier a_2 ve a_4 katsayılarının deęme derecesini (f) tahmin etmek için en kullanışlı

parametreler olduğunu belirtmiştir. Değme derecesi fiziksel olarak yıldızların Roche şişimlerini (loblarını) ne kadar doldurduğunu, yani küresellikten ne kadar uzaklaşıp elipsoid bir şekil aldığını ifade eder.

Fourier katsayıları ile yapılan kümeleme performansı çoklu model ile yapılan sınıflandırmaya nazaran düşük kalmıştır. Bu durum Fourier katsayılarının örten çift yıldızların ışık eğrilerini temsil kabiliyetinin, ışık eğrisi görselleri ve geometrik model parametrelerinin ışık eğrilerini temsil kabiliyetine göre daha sınırlı olduğunu gösterir. Bu durum EB ve EW sistemlerinin sınıflandırılmasında, bu tezde geliştirilen farklı veri türlerini aynı anda işleyen çok modlu model gibi karmaşık derin öğrenme mimarilerine duyulan ihtiyacı astrofiziksel olarak desteklemektedir.

İnsan müdahalesi olmadan (denetimsiz öğrenme) verinin doğal yapısını analiz eden GMM kümelemesi, kural tabanlı sınıflandırma (Selam 2004) ile özellikle EA sınıfında yüksek uyum göstermiştir. Ancak EB ve EW sınıflarındaki karışıklık örten çiftlerin ayrımında sadece Fourier katsayılarının kullanımı yetersiz kalmıştır. Denetimsiz öğrenme veri setindeki genel yapıyı anlamak güçlü bir araçtır; ancak yüksek doğruluklu bir sınıflandırma için denetimli derin öğrenme modellerinin kaçınılmaz olduğu sonucuna varılmıştır.

Bu tez çalışmasının öncelikli amacı yöntemlerin işlerliğini kanıtlamak (kavram kanıtı) ve bu tür sınıflandırma görevlerindeki potansiyel zorlukları tespit etmektir. Özellikle VGG-19 ve Fourier yaklaşımlarında karşılaşılan model uyumsuzluğu ve düşük performans sorunları gelecekteki çalışmalar için önemli araştırma alanları açmıştır.

Yüksek başarımla elde edilen çok modlu öğrenmede yanlış tahmin edilen (karışıklık matrisindeki köşegen haricindeki hücreler) yıldızlar gelecekteki çalışmalarda ayrıca ele alınabilir. Bu yıldızlar, denetimli öğrenmede, etiketlenen ile tahmin edilen sınıfın farklı olduğu yıldızlar olup, evrimsel olarak farklı aşamalarındaki yıldızlara örnek olabilir. Makine öğrenmesi modelinin çıkış katmanındaki sınıflandırma güven skor değerleri (softmax olasılıkları) bu yıldızları tanımlamada kullanılabilir.

Gelecekteki çalışmalarda daha tutarlı sonuçlar elde etmek amacıyla eğitim veri setlerinin farklı gözlem arşivlerinden verilerle zenginleştirilmesi ve ışık eğrilerinin eş baza getirilmesi için ışık eğrisi modelleme yöntemlerinin araştırılması hedeflenmektedir. Ayrıca yansıma, elipsoidal bozulma ve yıldız lekeleri gibi fiziksel etkilerin modellere entegre edilmesi örten çiftlerin sınıflandırılmasının fiziksel temellerini güçlendirme potansiyeline sahiptir.

Sonuç olarak bu tez çalışması, örten çift yıldızların sınıflandırılmasında çok modlu derin öğrenme modellerinin tekil veri türlerine (ışık eğrisi görselleri veya Fourier katsayıları gibi) dayalı yaklaşımlara kıyasla çok daha üstün ve güvenilir olduğunu olarak ortaya koymuştur. Geliştirilen çok modlu model Gaia DR3 gibi devasa arşivlerdeki on binlerce yıldızın otomatik, hızlı ve yüksek doğrulukla sınıflandırılmasına olanak tanıyarak, bu alandaki istatistiksel çalışmalara ve yıldız evrimi araştırmalarına önemli bir katkı sunmaktadır.

KAYNAKLAR

- Baroch, D., Giménez, A., Ribas, I., Morales, J. C., Anglada-Escudé, G. and Claret, A. 2021. Analysis of apsidal motion in eclipsing binaries using TESS data - I. A test of gravitational theories. *Astronomy and Astrophysics*, 649, 14.
- Bódi, A. and Hajdu, T. 2021. Classification of OGLE Eclipsing Binary Stars Based on Their Morphology Type with Locally Linear Embedding. *The Astrophysical Journal Supplement Series*, 255, 1.
- Borucki, J. W. 2016. KEPLER Mission: development and overview. *Reports on Progress in Physics*, Volume 79, Issue 3, article id. 036901.
- Chen, X., Wang, S., Deng, L., de Grijs, R., Yang, M. and Tian, H. 2020. The Zwicky Transient Facility Catalog of Periodic Variable Stars. *The Astrophysical Journal Supplement Series*, 249, 18.
- Čokina, M., Maslej-Krešňáková, V., Butka, P. and Parimucha, Š. 2021. Automatic classification of eclipsing binary stars using deep learning methods. *Astronomy and Computing*, 36, 100488.
- Daza-Perilla, I. V., Gramajo, L. V., Lares, M., Palma, T., Ferreira Lopes, C. E., Minniti, D. and Clariá, J. J. 2023. Automated classification of eclipsing binary systems in the VVV Survey. *Monthly Notices of the Royal Astronomical Society*, 520, 828–838.
- Eyer, L., Audard, M., Holl, B., Rimoldini, L., Carnerero, M. I., Clementini, G., De Ridder, J., Distefano, E., Evans, D. W., Gavras, P., Gommel, R., Lebzelter, T., Marton, G., Mowlavi, N., Panahi, A., Ripepi, V., Wyrzykowski, Ł., Nienartowicz, K., Jevardat de Fombelle, G., Lecoœur-Taibi, I., Rohrbasser, L., Riello, M., García-Lario, P., Lanzafame, A. C., Mazeh, T., Raiteri, C. M., Zucker, S., Abraham, P., Aerts, C., Aguado, J. J., Anderson, R. I., Bashi, D., Binnemfeld, A., Faigler, S., Garofalo, A., Karbevská, L., Kóspál, Á and Kruszyńska, K., Kun, M., Lanza, A. F., Leccia, S., Marconi, M., Messina, S., Molinaro, R., Molnár, L., Muraveva, T., Musella, I., Nagy, Z., Pagano, I., Palaversa, L., Plachy, E., Prša, A., Rybicki, K. A., Shahaf, S., Szabados, L., Szegedi-Elek, E., Trabucchi, M., Barblan, F., Grenon, M., Roelens, M. and Süveges, M. 2023. GaiaData Release 3: Summary of the variability processing and analysis. *Astronomy and Astrophysics*, 674, A13.
- Feiden, G. A. and Chaboyer, B. 2012. Reevaluating the Mass-Radius Relation for Low-Mass, Main Sequence Stars. *The Astrophysical Journal*, 757, 42.
- Gaulme, P., McKeever, J., Jackiewicz, J., Rawls, M. L., Corsaro, E., Mosser, B., Southworth, J., Mahadevan, S., Bender, C. and Deshpande, R. 2016. Testing the Asteroseismic Scaling Relations for Red Giants with Eclipsing Binaries Observed by Kepler. *The Astrophysical Journal*, 832, 19.

- Gavras, P., Rimoldini, L., Nienartowicz, K., de Fombelle, G. J., Holl, B., Ábrahám, P., Audard, M., Carnerero, M. I., Clementini, G., De Ridder, J., Distefano, E., Garcia-Lario, P., Garofalo, A., Kóspál, Á., Kruszyńska, K., Kun, M., Lecoeur-Taïbi, I., Marton, G., Mazeh, T., Mowlavi, N., Raiteri, C. M., Ripepi, V., Szabados, L., Zucker, S. and Eyer, L. 2023. *Astronomy and Astrophysics*, 674, A22.
- Healy, B. F., Coughlin, M. W., Mahabal, A. A., Jegou du Laz, T., Drake, A., Graham, M. J., Hillenbrand, L. A., van Roestel, J., Szkody, P., Zielske, L. A., Guiga, M., Hassan, M. Y., Hughes, J. L., Nir, G., Parikh, S., Park, S., Purohit, P., Rebbapragada, U., Reed, D., Warshofsky, D., Wold, A., Bloom, J. S., Masci, F. J., Riddle, R. and Smith, R. The ZTF Source Classification Project. III. A Catalog of Variable Sources. *The Astrophysical Journal Supplement Series*, 272, 14.
- Hrivnak, B. J. 1989. Radial velocity studies and absolute parameters of contact binaries. II - OO Aquilae. *Astrophysical Journal*, 340, 458-467.
- Ijspeert, L. W., Tkachenko, A., Johnston, C., Garcia, S., De Ridder, J., Van Reeth, T. and Aerts, C. An all-sky sample of intermediate- to high-mass OBA-type eclipsing binaries observed by TESS. 2021. *Astronomy and Astrophysics*, 652, A120.
- Keskin, B. and Baştürk, Ö. 2025. Eclipsing binary classification with machine learning techniques. *Contributions of the Astronomical Observatory Skalnaté Pleso*, 55, 357-361.
- Kharchenko, N. V., Scholz, R.-D., Piskunov, A. E., Röser, S. and Schilbach, E. 2007. Astrophysical supplements to the ASCC-2.5: Ia. Radial velocities of ~55000 stars and mean radial velocities of 516 Galactic open clusters and associations. *Astronomische Nachrichten*, 328, 889.
- Kirk, B., Conroy, K., Prša, A., Abdul-Masih, M., Kochoska, A., Matijević, G., Hambleton, K., Barclay, T., Bloemen, S., Boyajian, T., Doyle, L. R., Fulton, B. J., Hoekstra, A. J., Jek, K., Kane, S. R., Kostov, V., Latham, D., Mazeh, T., Orosz, J. A., Pepper, J., Quarles, B., Ragozzine, D., Shporer, A., Southworth, J., Stassun, K., Thompson, S. E., Welsh, W. F., Agol, E., Derekas, A., Devor, J., Fischer, D., Green, G., Gropp, J., Jacobs, T., Johnston, C., LaCourse, D. M., Sætre, K., Schwengeler, H., Toczyski, J., Werner, G., Garrett, M., Gore, J., Martinez, A. O., Spitzer, I., Stevick, J., Thomadis, P. C., Vrijmoet, E. H., Yenawine, M., Batalha, N. and Borucki, W. 2016. Kepler Eclipsing Binary Stars. VII. The Catalog of Eclipsing Binaries Found in the Entire Kepler Data-Set. *The Astronomical Journal*, 151, 68.
- Kochoska, A., Conroy, K., Hambleton, K. and Prša, A. 2020. Beyond DC and MCMC: alternative algorithms and approaches to fitting light curves. *Contributions of the Astronomical Observatory Skalnaté Pleso*, 50, 539-545.
- Kopal, Z. 1959. Close binary systems. The International Astrophysics Series, London: Chapman & Hall.

- Lee, C.-U. and Lee, J.-W. 2006. Photometric Study of the Near-Contact Binary CN Andromedae. *Journal of The Korean Astronomical Society*, 39, 25-30.
- Matijević, G., Prša, A., Orosz, J. A., Welsh, W. F., Bloemen, S. and Barclay, T. 2012. Kepler Eclipsing Binary Stars. III. Classification of Kepler Eclipsing Binary Light Curves with Locally Linear Embedding. *The Astronomical Journal*, 143, Issue 5, article id. 123, p 6.
- Mowlavi, N., Holl, B., Lecoœur-Taïbi, I., Barblan, F., Kochoska, A., Prša, A., Mazeh, T., Rimoldini, L., Gavras, P., Audard, M., Jevardat de Fombelle, G., Nienartowicz, K., García-Lario, P. and Eyer, L. 2023. Gaia Data Release 3: The first Gaia catalogue of eclipsing-binary candidates. *Astronomy and Astrophysics*, 674, A16.
- Ofman, L., Averbuch, A., Shlisselberg, A., Benaun, I., Segev, D. and Rissman, A. 2022. *New Astronomy*, 91, 101693.
- Osborn, H. P., Ansdell, M., Ioannou, Y., Sasdelli, M., Angerhausen, D., Caldwell, D., Jenkins, J. M., Räissi, C. and Smith, J. C. 2020. Rapid classification of TESS planet candidates with convolutional neural networks. *Astronomy and Astrophysics*, 633, A53.
- Özuyar, D. 2007. W UMa türü değen çift U pegasi'nin ışık eğrisi ve dönem analizi. Yüksek Lisans Tezi. Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.
- Papageorgiou, A., Catelan, M., Christopoulou, P.E., Drake, A. J. and Djorgovski, S. G. 2018. An Updated Catalog of 4680 Northern Eclipsing Binaries with Algol-type Light-curve Morphology in the Catalina Sky Surveys. *The Astrophysical Journal Supplement Series*, 238, 4.
- Papageorgiou, A., Catelan, M., Christopoulou, P. E., Drake, A. J. And Djorgovski, S. G. 2019. Physical Parameters of Northern Eclipsing Binaries in the Catalina Sky Survey. *The Astrophysical Journal Supplement Series*, Volume 242, Issue 1, article id. 6, 10 pp.
- Parimucha, Š., Gabdeev, M., Markus, Y., Vaňko, M. and Gajdoš, P. 2025. Morphological classification of eclipsing binary stars using computer vision methods. eprint arXiv:2508.12802.
- Prša, A., Kochoska, A., Conroy, K. E., Eisner, N., Hey, D. R. and IJspeert, L. 2022. TESS Eclipsing Binary Stars. I. Short-cadence Observations of 4584 Eclipsing Binaries in Sectors 1-26. *The Astrophysical Journal Supplement Series*, 258, Issue 1, id.16, p 22.
- Prša, A., Guinan, E. F., Devinney, E. J., DeGeorge, M., Bradstreet, D. H., Giammarco, J. M., Alcock, C. R. and Engle, S. G. 2008. Artificial Intelligence Approach to the Determination of Physical Properties of Eclipsing Binaries. I. The EBAI Project. 687, 542-565.
- Ricker, G. R., Winn, J. N., Vanderspek, R., Latham, D. W., Bakos, G., Bean, J. L., Berta-Thompson, Z. K., Brown, T. M., Buchhave, L., Butler, N., Butler, R. P., Chaplin,

- W. J., Charbonneau, D., Christensen-Dalsgaard, J., Clampin, M., Deming, D., Doty, J., De Lee, N., Dressing, C., Villaseñor, J. 2014. The transiting exoplanet survey satellite. *Proceedings of SPIE - The International Society for Optical Engineering*; Vol. 9143.
- Rimoldini, L., Holl, B., Gavras, P., Audard, M., De Ridder, J., Mowlavi, N., Nienartowicz, K., Jevardat de Fombelle, G., Lecoœur-Taïbi, I., Karbevská, L., Evans, D. W., Ábrahám, P., Carnerero, M. I., Clementini, G., Distefano, E., Garofalo, A., García-Lario, P., Gomel, R., Klioner, S. A., Kruszyńska, K., Lanzafame, A. C., Lebzelter, T., Marton, G., Mazeh, T., Molinaro, R., Panahi, A., Raiteri, C. M., Ripepi, V., Szabados, L., Teyssier, D., Trabucchi, M., Wyrzykowski, Ł., Zucker, S. and Eyer, L. 2023. GaiaData Release 3: All-sky classification of 12.4 million variable sources into 25 classes. *Astronomy and Astrophysics*, 674, A14.
- Roweis, S. T. and Saul, L. K. 2000. Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science*, 290, 2323-2326.
- Rucinski, S. M. 1993. A Simple Description of Light Curves of W UMa Systems. *Publications of the Astronomical Society of the Pacific*, 105, 1433.
- Samus, N. N., Kazarovets, E. V., Durlevich, O. V., Kireeva, N. N., and Pastukhova, E. N. 2017. General catalogue of variable stars: Version GCVS 5.1. *Astronomy Reports*, 61, 80–88.
- Sarro, L. M., Sánchez-Fernández, C. and Giménez, Á. 2006. Automatic classification of eclipsing binaries light curves using neural networks. *Astronomy and Astrophysics*, 446, 395–402.
- Selam, S. O. 2004. Key parameters of W UMa-type contact binaries discovered by HIPPARCOS. *Astronomy and Astrophysics*, 416, 1097-1105.
- Shallue, C. J. and Vanderburg, A. 2018. Identifying Exoplanets with Deep Learning: A Five-planet Resonant Chain around Kepler-80 and an Eighth Planet around Kepler-90. *The Astronomical Journal*, 155, 94.
- Shan, Y., Chen, J., Zhang, Z., Wang, L., Zou, Z. and Li, M. 2025. Identifying eclipsing binary stars with TESS data based on a new hybrid deep learning model. *Publications of the Astronomical Society of the Pacific*, 137(4), 044503.
- Simonyan, K. and Zisserman, A. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. eprint arXiv:1409.1556.
- Skelton, P. L. and Smits, D. P. 2009. Modelling of W UMa-type variable stars. *South African Journal of Science*, 105, 120-126.
- Stassun, K. G. and Torres, G. 2016. Eclipsing Binary Stars as Benchmarks for Trigonometric Parallaxes in the Gaia Era. *The Astronomical Journal*, 152, 180.

- Süveges, M., Barblan, F., Lecoœur-Taïbi, I., Prša, A., Holl, B., Eyer, L., Kochoska, A., Mowlavi, N. and Rimoldini, L. 2017. Gaia eclipsing binary and multiple systems. Supervised classification and self-organizing maps. *Astronomy and Astrophysics*, 603, A117.
- Szklenár, T., Bódi, A., Tarczay-Nehéz, D., Vida, K., Marton, G., Mező, Gy., Forró, A. and Szabó, R. 2020. Image-based Classification of Variable Stars: First Results from Optical Gravitational Lensing Experiment Data. *The Astrophysical Journal Letters*, 897, L12.
- Szklenár, T., Bódi, A., Tarczay-Nehéz, D., Vida, K., Mező, Gy. and Szabó, R. 2022. Variable Star Classification with a Multiple-input Neural Network. *The Astrophysical Journal*, 938, 37.
- Torres, G., Andersen, J., Giménez, A. 2010. Accurate masses and radii of normal stars: modern results and applications. *Astronomy and Astrophysics Review*, 18, 67-126.
- Ulaş, B., Szklenár, T. and Szabó, R. 2025. Detection of oscillation-like patterns in eclipsing binary light curves using neural network-based object detection algorithms. *Astronomy and Astrophysics*, 695, A81.
- Wilson, R. E. 1979. Eccentric orbit generalization and simultaneous solution of binary star light and velocity curves. *The Astrophysical Journal*, 234, 1054-1066.
- Wilson, R. E., and Devinney, E. J. 1971. Realization of Accurate Close-Binary Light Curves: Application to MR Cygni. *The Astrophysical Journal*, 166, 605.
- Wyrzykowski, L., Udalski, A., Kubiak, M., Szymanski, M., Zebrun, K., Soszynski, I., Wozniak, P. R., Pietrzynski, G. and Szewczyk, O. 2003. The Optical Gravitational Lensing Experiment. Eclipsing Binary Stars in the Large Magellanic Cloud. *Acta Astronomica*, 53, 1-25.

EKLER

EK 1 Geliştirilen Makine Öğrenmesi Modelleri

**EK 2 Gaia DR3 Örtün Çift Yıldızlarının Makine Öğrenmesi ile Işık Eğrisinin
Biçimine Göre EA, EB, EW Sınıflandırmasına Ait Bazı Örnekler**

EK 1 Geliştirilen Makine Öğrenmesi Modelleri

A. Çok Modlu (Multimodal) Sınıflandırma Modeli:

Bu model derin öğrenme mimarisi olan Evrişimli Sinir Ağlarını (CNN) ve Çok Katmanlı Algılayıcıları (MLP) birleştirir. Model, değişen yıldızları EA, EB, EW sınıflandırmak için iki farklı veri türünü (modu) aynı anda kullanır: ışık eğrisi görselleri ve yıldızın Mowlavi vd. (2023) tarafından elde edilen geometrik model parametrelerini içeren sayısal veriler. Dolayısıyla model iki bağımsız girdi yoluna sahiptir.

Bu modelde önceden eğitilmiş bir ağ kullanmadığı için uçtan uca öğrenme prensibiyle çalışır. Eğitim sürecinde görsel moddaki evrişim filtreleri (Conv2D) ışık eğrisinin morfolojik desenlerini çıkarırken, eş zamanlı olarak sayısal moddaki tam bağlı katmanlar (Dense) geometrik parametrelerin ilişkilerini öğrenir. Birleştirme (Concatenate) katmanını takip eden yoğun katmanlarda ise her iki veri kaynağından elde edilen özniteliklerin sınıflandırma kararındaki ağırlıklarını optimize eder.

1. Görsel Mod (Işık Eğrisi Görselleri)

Evrişimli Sinir Ağı (CNN) bileşenidir.

- **Girdi Katmanı:** image_input (128x96x3 boyutunda, ışık eğrilerinin görselleştirilmiş bir görüntüsü).
- **Veri Artırma (Data Augmentation):** Modelin genelleme yeteneğini artırmak için girdi aşamasında rastgele yatay çevirme (flip), kaydırma (translation) ve yakınlaştırma (zoom) uygulanır.
- **Konvolüsyon Blokları:** Üç aşamalı (32, 64 ve 128 filtreli) derin mimari kullanılır. Her blokta:
 - **Evrişim (Conv2D):** Evrişim katmanı (32 filtre, 3x3 çekirdek) ile görüntüdeki yerel desenler (örneğin ışık eğrisinin şekli) çıkarılır.

- **Normalizasyon (BatchNormalization):** Eğitim sürecini stabilize eder ve hızlandırır.
- **Maksimum Havuzlama (MaxPooling2D):** Boyut azaltılarak temel özelliklere odaklanılır.
- **Seyreltme (Dropout):** Ezberlemeyi (overfitting) engellemek için %25 ile %40 arasında değişen oranlarda nöronlar devre dışı bırakılır.
- **Tam Bağlı Katman:** Flatten işleminden sonra 64 nöronlu, L2 düzenlileştirmeli bir yoğun katman bulunur.
- **Çıktı:** Görüntü özelliklerinin temsil edildiği 64 boyutlu vektör (cnn_output).

2. Sayısal Mod (Mowlavi vd. (2023) Geometrik Model Parametreleri)

Çok Katmanlı Algılayıcı (MLP) bileşenidir.

- **Girdi Katmanı:** tabular_input (yıldızın model parametreleri; frekans, genlik, evre, sigma vd.).
- **Yoğun Katmanlar:** İki aşamalı MLP yapısı kullanılır:
 - **Birinci Katman:** 64 nöron, BatchNormalization, Dropout (%40) ve L2 düzenlileştirme ile verideki karmaşık ilişkiler öğrenilir.
 - **İkinci Katman:** 32 nöronlu yoğun katman ile özellikler rafine edilir.
- **Çıktı:** Sayısal verideki özelliklerin temsil edildiği 32 boyutlu vektör (mlp_output).

3. Birleştirme ve Sınıflandırma

İki yolun çıktıları birleştirilerek (concatenate) ortak bir alanda temsil edilir ve son sınıflandırma yapılır:

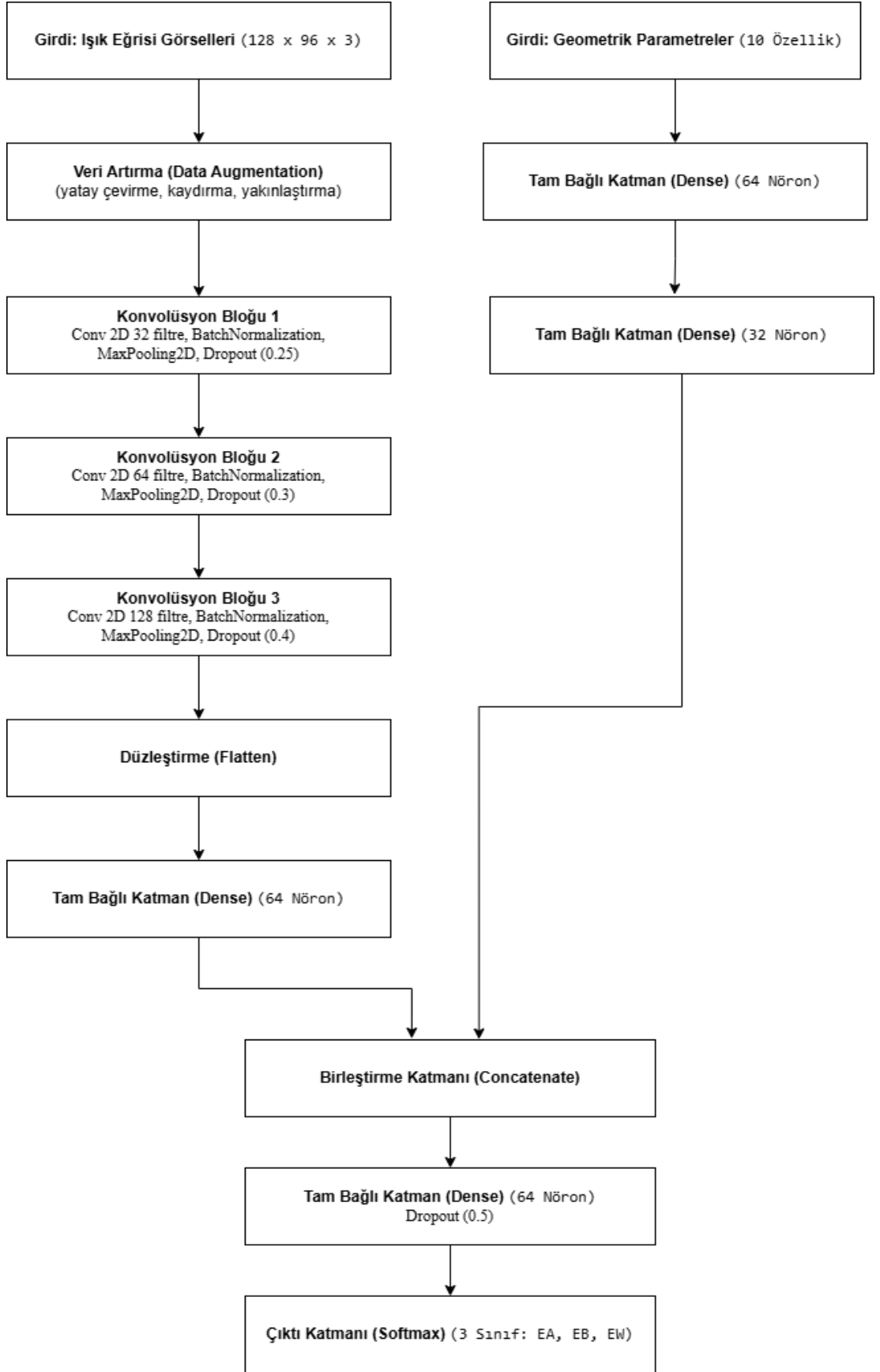
- **Birleştirme (Concatenate):** cnn_output (64) ve mlp_output (32) vektörleri birleştirilerek 96 boyutlu bir öznitelik uzayı oluşturulur.

- **Karar Katmanı:** 64 nöronlu, Dropout (%50) ve L2 düzenlileştirmeli bir yoğun katmandan geçirilerek sınıflar arası ayırım keskinleştirilir.
- **Çıkış Katmanı:** EA, EB, EW sınıfları için 3 nöronlu Softmax aktivasyonu kullanılır. Model, her sınıf için %0 ile %100 arasında bir olasılık değeri üretir.

4. Temel İş Akışı

- Işık eğrisi görselleri “TrainValidation” klasöründen, sayısal veriler “MowlaviModelParams.csv” dosyasından yüklenir.
- Görüntüler boyuta uygun hale getirilir ve [0, 1] aralığına normalize edilir (her bir değer 255’e bölünür).
- Sayısal veriler, eğitim setine göre StandardScaler ile ölçeklenir (ortalama 0, standart sapma 1). Ölçekleyici, test aşamasında kullanılmak üzere kaydedilir (tabular_scaler.joblib).
- Etiketler (EA, EB, EW), LabelEncoder ve one-hot encoding (to_categorical) ile sayısal hale getirilir.
- Veri seti, eğitim ve doğrulama setlerine stratify (sınıf oranlarını koruyarak) ile %80 ve %20 oranında bölünür.
- Model, Adam optimize edici ve categorical_crossentropy kaybı ile derlenir.
- MacroF1Callback geri çağrısı (callback) doğrulama kümesi üzerindeki Macro F1-skoru metriğini hesaplar.
- ReduceLROnPlateau geri çağrısı doğrulama kaybı (val_loss) iyileşmediğinde öğrenme hızını otomatik olarak düşürür.
- ModelCheckpoint ve EarlyStopping geri çağrıları, en iyi modeli Macro F1-skoruna göre kaydetmek ve eğitimdeki ilerlemeye göre durdurmak için kullanılır.
- Ayrı bir Test dizinindeki veriler üzerinde tahminler yapılır.

- Accuracy, Precision, Recall ve F1-score (macro ve weighted) gibi metrikler hesaplanır.
- Sınıflandırma Raporu ve Confusion Matrix görselleştirilir ve kaydedilir.
- Eğitim/Doğrulama Kayıp, Doğruluk ve Macro F1-skorunu gösteren bir Öğrenme Eğrisi grafiği oluşturulur.
- Test verilerindeki görüntü dosyaları, Confusion Matrix'teki gerçek ve tahmin edilen sınıfa göre klasörlere kopyalanır (örneğin, EA-EB klasörü, gerçekte EA iken EB olarak tahmin edilen görüntüleri içerir).



Çok modlu model mimarisi

B. Görsel Veri ile Sınıflandırma Modeli:

VGG-19 mimarisine dayalı bir CNN modeli tanımlayan transfer öğrenme (ing. transfer learning) metodu geliştirildi. ImageNet²⁴ veritabanının önceden eğitilmiş ağırlıkları kullanıldı. ImageNet, bilgisayarlı görü (ing. computer vision) alanında devrim yaratan, büyük ölçekli bir görsel veritabanıdır. 20.000'den fazla farklı kategoriye ayrılmış 14 milyondan fazla elle etiketlenmiş yüksek çözünürlüklü görüntü içerir. ImageNet modern derin öğrenme uygulamalarında transfer öğrenmenin temelini oluşturur. Bir modeli sıfırdan eğitmek yerine, ImageNet ağırlıklarıyla başlamak, daha az veriyle daha hızlı bir şekilde yüksek doğrulukta sonuçlar elde edilmesini sağlar.

VGG-19 modeli ImageNet'in devasa veri kümesi üzerinde önceden eğitilmiştir. Bu süreçte, model evrişim katmanlarında genel görüntü özelliklerini (kenarlar, dokular, temel şekiller vb.) öğrenir. Bu önceden eğitilmiş ağırlıklar VGG-19 mimarisine dayalı CNN modelini eğitmede kullanılır.

VGG-19 modelinde Transfer Öğrenme yaklaşımı kullanıldığı için, özellik çıkarıcı (feature extractor) kısmının büyük çoğunluğu dondurulmuş, sadece üst düzey özelliklerin çıkarıldığı son blok ve sınıflandırma katmanları (classifier head) eğitilmiştir.

1. Temel Model

- **VGG-19 Mimarisi:** Keras/TensorFlow kütüphanesinden önceden eğitilmiş VGG-19 modeli yüklenir.
- **Önceden Eğitilmiş Ağırlıklar:** weights='imagenet' kullanılarak modelin ImageNet üzerinde zaten öğrendiği ağırlıkları devralması sağlanarak transfer öğrenme yapıldı.
- **Tam Bağlı Katmanların Hariç Tutulması:** include_top=False ayarı ile VGG-19'un orijinal üst sınıflandırma katmanları dahil edilmeyerek yeni üst katmanlar eklendi.

²⁴ <https://www.image-net.org/>

2. Özellik Çıkarma (Feature Extraction)

- VGG-19'un evrişimli katmanlarının son 4 katmanı hariç diğerleri donduruldu. Böylece önceden öğrenilmiş ImageNet özelliklerinin korunması ve eğitim sırasında sadece modelin üst kısımlarının (VGG-19'un son bloklarının) güncellenmesi sağlandı.

3. Yeni Katmanlar

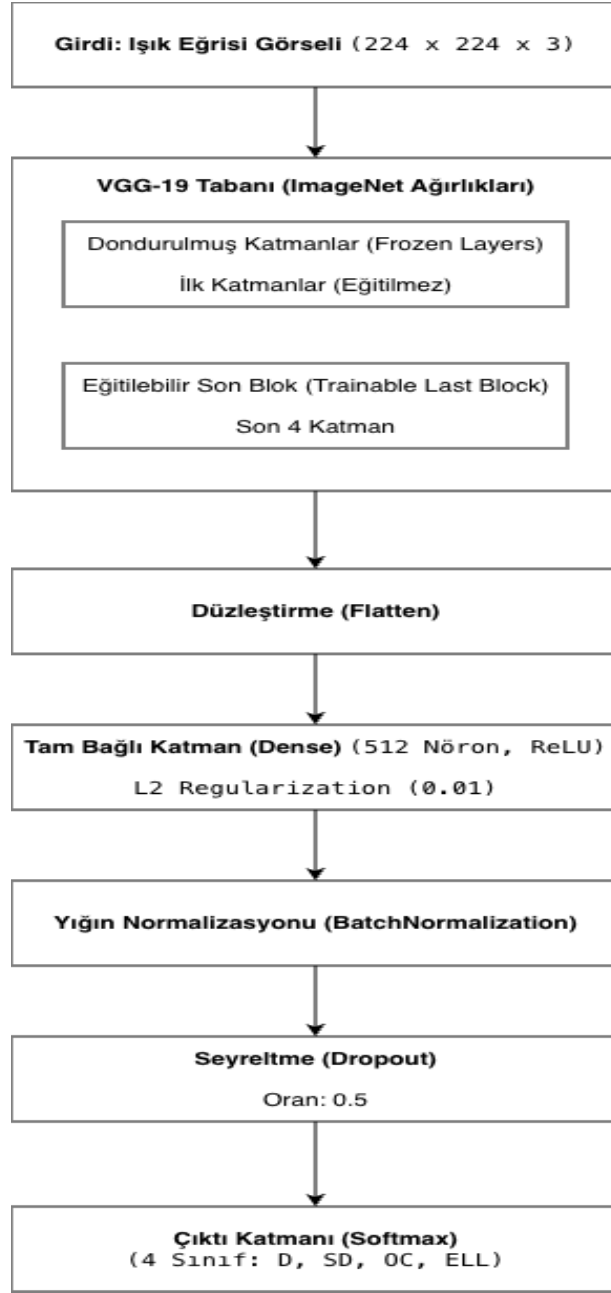
Dondurulmuş VGG-19 tabanının üzerine yeni katmanlar eklendi:

- **Düzleştirme (Flatten):** VGG-19'dan çıkan 3 boyutlu özellik haritası bir boyutlu vektöre dönüştürüldü.
- **Yoğun Katman (Dense):** 512 nöronlu gizli bir katman eklendi.
- **Çıkış Katmanı (Dense):** Detached, semi-detached, overcontact ve ellipsoidal sınıflarının sayısı olan 4 adet nöron içerir ve softmax aktivasyon fonksiyonu ile her sınıf için olasılık tahmini yapar.

4. Temel İş Akışı

- Işık eğrisi görselleri “Kepler” ve “TESS” klasörlerinden yüklenir.
- Kepler ve TESS verileri birleştirilip diziye dönüştürülür.
- 4 sınıfı temsil eden etiketler one-hot encoding kullanılarak sayısal hale getirilir.
- Veri seti 3 parçaya bölünür: %67 eğitim, %25 doğrulama, %8 test.
- Model 100 çevrim boyunca eğitilir. Eğitim sürecini kontrol etmek ve optimize etmek için üç geri çağrı (ing. call back) kullanılır:
 - **Model Kontrol Noktası (ModelCheckpoint):** val_accuracy metrik değeri en yüksek olan model ağırlıkları diske kaydedilir.

- **Erken Durdurma (EarlyStopping):** Aşırı öğrenmeyi önlemek için val_loss değeri 10 çevrim boyunca iyileşmezse eğitim durdurulur ve en iyi ağırlıklar geri yüklenir.
- **Öğrenme Oranı Azaltıcı (ReduceLROnPlateau):** val_loss 5 epoch boyunca iyileşmezse, öğrenme oranı 0.1 faktörle (yani 10 kat) azaltılır. İzlenen bir metriktaki iyileşme belirli bir süre durduğunda öğrenme hızı düşürülerek modelin daha iyi bir çözüme ulaşması hedeflenir.
- Eğitim tamamlandıktan sonra model test seti üzerinde değerlendirilir ve kayıp (loss) ve doğruluk (accuracy) değerleri hesaplanır.



Transfer öğrenme için özelleştirilmiş VGG-19 Mimarisi

C. Sayısal Veri ile Sınıflandırma Modeli:

Kepler, TESS ve Gaia DR3 örten çift yıldız ışık eğrileri Fourier serisi ile modellendi. Kepler ve TESS yıldızlarının Fourier katsayılarını kullanarak 1 boyutlu CNN (1D CNN) modeli eğitildi ve ardından Gaia DR3 örten çift yıldızları tahmin edildi.

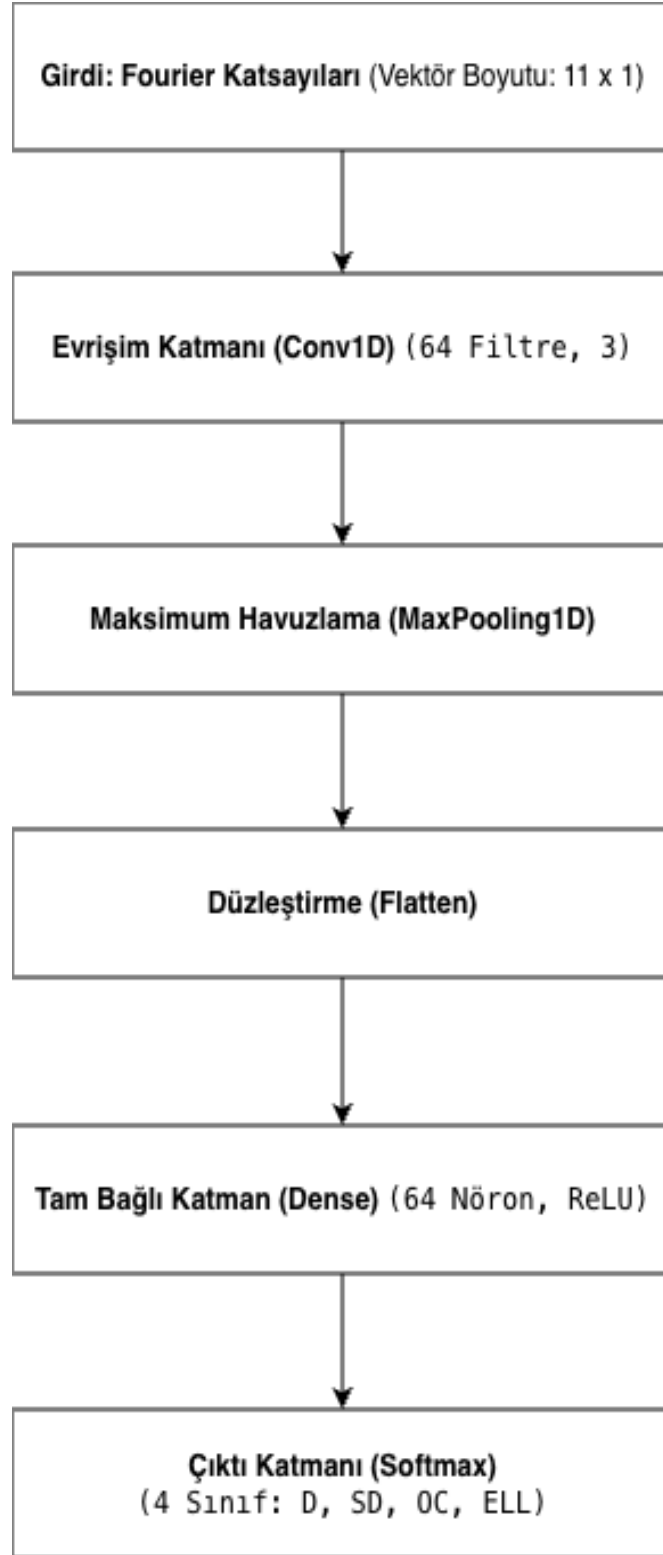
Bu modelde ağ sıfırdan başlatıldığı için tüm ağırlıklar rassal (random initialization) olarak başlar ve tüm katmanlarda öğrenme gerçekleşir.

1. Model

- **Conv1D:** 1D Evrişim Katmanı. Fourier katsayıları içindeki yerel kalıpları ve ilişkileri öğrenir.
- **MaxPooling1D:** Boyutu küçülterek hesaplama yükünü azaltır ve önemli özellikleri korur.
- **Düzleştirme (Flatten):** Evrişim katmanlarının çıktısını düzleştirerek yoğun katmanları besler.
- **Yoğun Katman (Dense):** 64 nöronlu gizli katman.
- **Çıkış Katmanı (Dense):** Detached, semi-detached, overcontact ve ellipsoidal sınıflarının sayısı olan 4 adet nöron içerir ve softmax aktivasyon fonksiyonu ile her sınıf için olasılık tahmini yapar.

2. Temel İş Akışı

- Kepler, TESS ve Gaia DR3 örten çift yıldızlarının ışık eğrisi zaman serisi verilerine fit yapılarak Fourier katsayıları CSV dosyalarına kaydedilir.
- Kepler ve TESS yıldızlarına ait Fourier katsayıları ilgili CSV dosyasından okunur ve tek bir dizi haline getirilir.
- 4 sınıfı temsil eden etiketler one-hot encoding kullanılarak sayısal hale getirilir.
- Veri seti 2 parçaya bölünür: %80 eğitim, %20 doğrulama/test.
- Model 20 çevrim boyunca eğitilir.
- Eğitim tamamlandıktan sonra model test seti üzerinde değerlendirilir ve kayıp (loss) ve doğruluk (accuracy) değerleri hesaplanır.



1 Boyutlu CNN Mimarisi

Ç. GMM ile Kümelenendirme Modeli:

Kepler örten çift yıldızlarının Fourier katsayıları GMM yöntemiyle kümelendirildi. Bölüm 4.3'te yıldızlar Selam (2004)'e göre EA, EB, EW şeklinde 3 adet sınıfa ayrıldığı için GMM'nin de veri setini 3 kümeye ayırması sağlandı ($n_components=3$). Böylece sınıflandırma sonucu ile kümelendirme sonucunun karşılaştırılması yapıldı. Kümeleme için sadece sınıflandırmada kullanılan a_1 , a_2 ve a_4 Fourier katsayıları seçildi.

1. Model

- **$n_components=3$:** Veri 3 farklı kümeye ayrılsın.

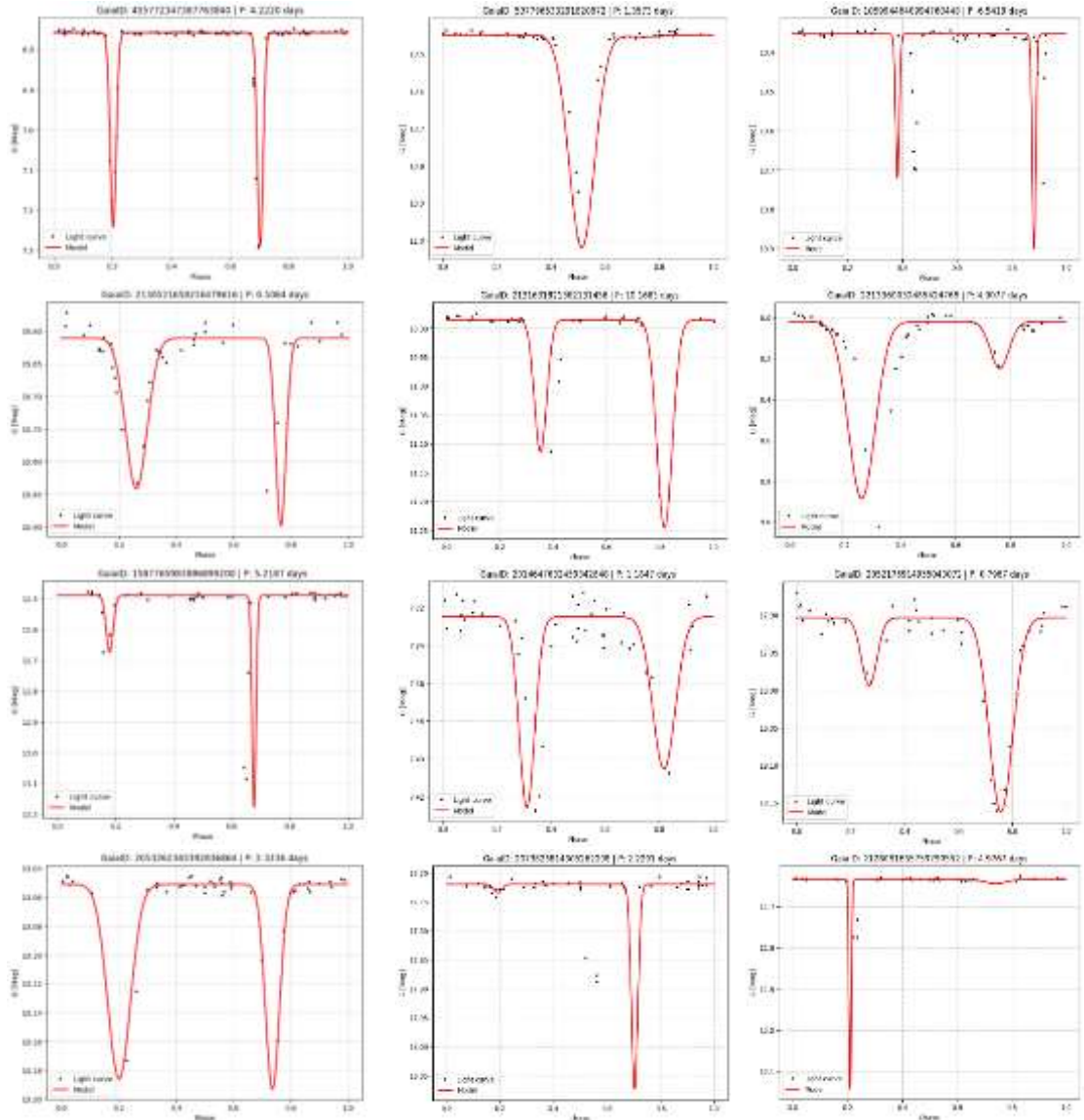
2. Temel İş Akışı

- Kepler yıldızlarına ait Fourier katsayıları ilgili CSV dosyasından okunur.
- Selam (2004) kriterlerine göre yıldızlar EA, EB, EW şeklinde etiketlenir.
- Yıldızların a_1 , a_2 ve a_4 Fourier katsayıları özellik (feature) olarak alınır.
- 'StandardScaler' kullanılarak özellikler standartlaştırılır. Standartlaştırma her özelliğin ortalamasını 0 ve standart sapmasını 1 yapar. Böylece farklı aralıklara sahip özelliklerin GMM algoritmasını eşit derecede etkilemesi sağlanır.
- Veri PCA ile 2 boyuta indirgenir.
- Veri GMM ile 3 kümeye ayrılır.
- Kümeleme sonuçları saçılım grafiğinde görselleştirilir.
- Sınıflandırma ile kümelemenin karşılaştırıldığı karışıklık matrisi çizilir.

EK 2 Gaia Dr3 Örtün Çift Yıldızlarının Makine Öğrenmesi İle Işık Eğrisinin Biçimine Göre Otomatik Olarak Ea, Eb, Ew Sınıflandırılmasına Ait Örnekler

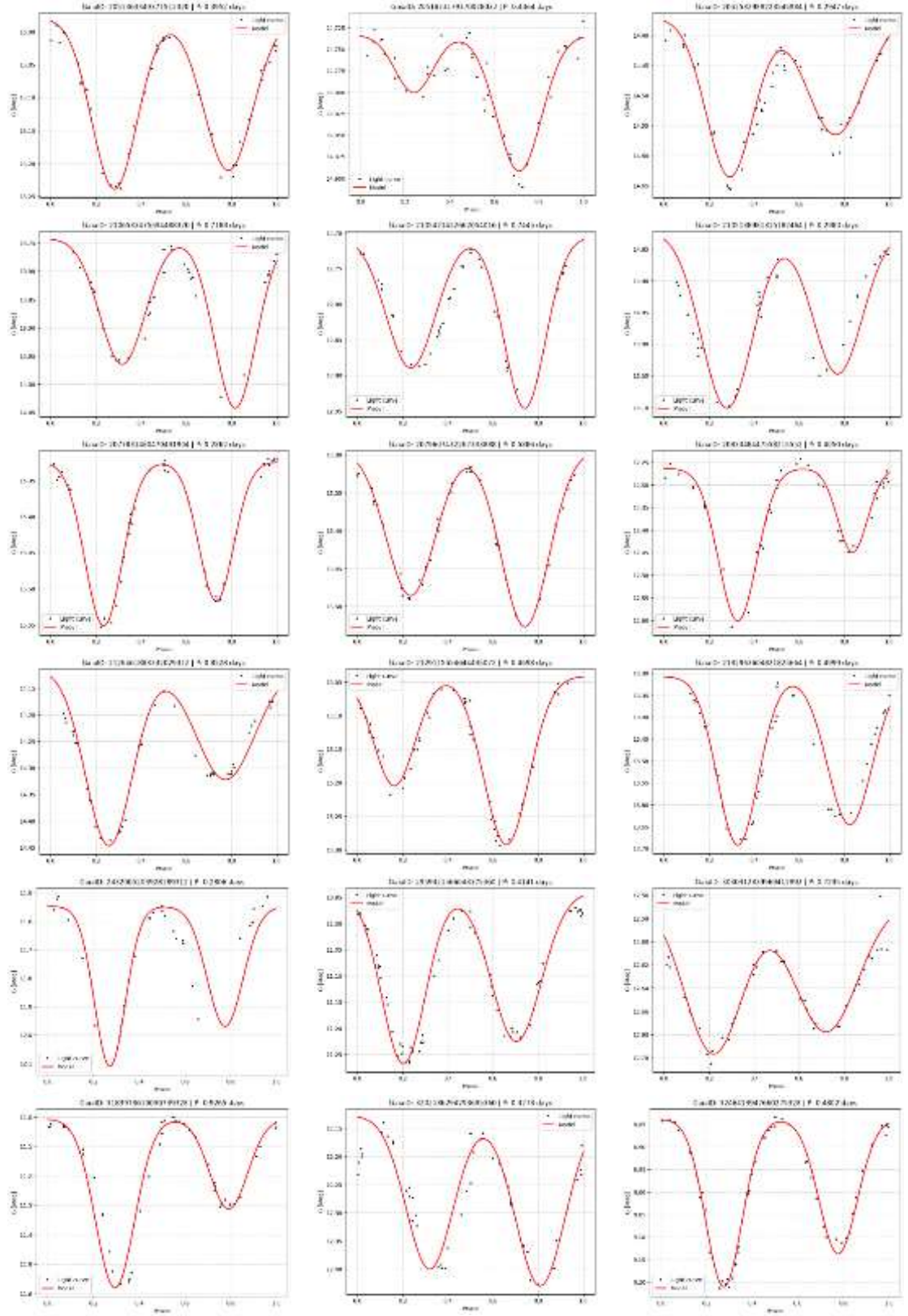
Örtün çift yıldızın Gaia ID'si görselinin altında verilmiştir. Otomatik olarak sınıflandırılan yaklaşık 2 milyon yıldızın tamamı GitHub sayfasında verilmiştir²⁵.

A. EA Olarak Sınıflandırılan Örtün Çift Yıldız Örnekleri



²⁵ <https://github.com/BedriKeskin/EclipsingBinaryClassification/tree/main/Gaia/>

B. EB Olarak Sınıflandırılan Örtün Çift Yıldız Örnekleri



C. EW Olarak Sınıflandırılan Örtün Çift Yıldız Örnekleri

