

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

SİBER GÜVENLİK SİSTEMLERİ İÇİN DİNAMİK VE ARTIMLI
MAKİNE ÖĞRENMESİ YAKLAŞIMLARI

DOKTORA TEZİ

Engin BAYSAL

Bilgisayar ve Bilişim Mühendisliği Anabilim Dalı

HAZİRAN 2025

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

SİBER GÜVENLİK SİSTEMLERİ İÇİN DİNAMİK VE ARTIMLI
MAKİNE ÖĞRENMESİ YAKLAŞIMLARI

DOKTORA TEZİ

Engin BAYSAL

Bilgisayar ve Bilişim Mühendisliği Anabilim Dalı

Tez Danışmanı: Prof. Dr. Cüneyt BAYILMIŞ

HAZİRAN 2025

Engin BAYSAL tarafından hazırlanan “SİBER GÜVENLİK SİSTEMLERİ İÇİN DİNAMİK VE ARTIMLI MAKİNE ÖĞRENMESİ YAKLAŞIMLARI” adlı tez çalışması 20.06.2025 tarihinde aşağıdaki jüri tarafından oy birliği/oy çokluğu ile Sakarya Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar ve Bilişim Mühendisliği Anabilim Dalı’nda Doktora tezi olarak kabul edilmiştir.

Tez Jürisi

Jüri Başkanı : **Prof. Dr. Devrim AKGÜN**
Sakarya Üniversitesi

Jüri Üyesi : **Prof. Dr. Cüneyt BAYILMIŞ** (Danışman)
Sakarya Üniversitesi

Jüri Üyesi : **Prof. Dr. Ünal ÇAVUŞOĞLU**
Sakarya Üniversitesi

Jüri Üyesi : **Doç. Dr. Süleyman UZUN**
Sakarya Uygulamalı Bilimler Üniversitesi

Jüri Üyesi : **Doç. Dr. Selman HIZAL**
Sakarya Uygulamalı Bilimler Üniversitesi



ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Sakarya Üniversitesi Fen Bilimleri Enstitüsü Lisansüstü Eğitim-Öğretim Yönetmeliğine ve Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesine uygun olarak hazırlamış olduğum “SİBER GÜVENLİK SİSTEMLERİ İÇİN DİNAMİK VE ARTIMLI MAKİNE ÖĞRENMESİ YAKLAŞIMLARI” başlıklı tezin bana ait, özgün bir çalışma olduğunu; çalışmamın tüm aşamalarında yukarıda belirtilen yönetmelik ve yönergeye uygun davrandığımı, tezin içerdiği yenilik ve sonuçları başka bir yerden almadığımı, tezde kullandığım eserleri usulüne göre kaynak olarak gösterdiğimi, bu tezi başka bir bilim kuruluna akademik amaç ve unvan almak amacıyla vermediğimi ve 20.04.2016 tarihli Resmi Gazete’de yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince Sakarya Üniversitesi’nin abonesi olduğu intihal yazılım programı kullanılarak Enstitü tarafından belirlenmiş ölçütlere uygun rapor alındığını, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun ortaya çıkması halinde doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi beyan ederim.

(20/06/2025).

Engin BAYSAL





Arya...



TEŞEKKÜR

Bu zorlu ama bir o kadar da değerli akademik yolculuğumda bana destek olan herkese en içten teşekkürlerimi sunarım.

Öncelikle, danışmanım Sayın Prof. Dr. Cüneyt BAYILMIŞ'a sonsuz minnettarım. Bilgisi, sabrı, değerli yönlendirmeleri ve her aşamada gösterdiği özverili destek olmasaydı, bu tezin tamamlanması mümkün olmazdı. Kendisine yalnızca akademik anlamda değil, mesleki ve kişisel gelişimimde de bir rol model olduğu için ayrıca teşekkür ederim.

Tez İzleme Komitesi'nin değerli üyeleri Prof. Dr. Devrim AKGÜN ve Doç. Dr. Süleyman UZUN'a, araştırmamın kritik aşamalarında yaptıkları yapıcı eleştiriler ve yönlendirmeler için içten teşekkürlerimi iletirim. Ayrıca, tez savunma jürisinde yer alarak değerlendirmeleriyle çalışmamın olgunlaşmasına katkı sunan saygıdeğer jüri üyeleri Prof. Dr. Ünal ÇAVUŞOĞLU ve Doç. Dr. Selman HIZAL'a da teşekkür ederim.

Hayatımın en büyük motivasyon kaynağı olan sevgili kızım Arya'ya ve her koşulda yanımda duran, sabrı ve sevgisiyle bana güç veren eşim Derya'ya şükranlarımı sunarım. Bu uzun ve yorucu süreçte verdikleri destek, bana her zaman ilham oldu. Onların varlığı, bu yolculuğu anlamlı kılan en büyük güçtü.

Ayrıca, tez çalışmam süresince fikirleriyle, eleştirileriyle ve moral desteğiyle yanımda olan tüm akademisyenlere, meslektaşlarıma ve arkadaşlarıma teşekkürü bir borç bilirim.

Engin BAYSAL



İÇİNDEKİLER

Sayfa

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ	v
TEŞEKKÜR	ix
İÇİNDEKİLER	xi
KISALTMALAR	xiii
SİMGELER	xv
TABLO LİSTESİ	xvii
ŞEKİL LİSTESİ	xix
ÖZET	xxi
SUMMARY	xxiii
1. GİRİŞ	1
1.1. Problemin Tanımı	3
1.2. Tezin Amacı	6
1.3. Tez Çalışmasının Katkıları	7
1.4. Tezin Organizasyonu	10
2. ÖĞRENME PARADİGMALARI	13
2.1. Yaşam Boyu Öğrenme (Lifelong Learning, LL)	13
2.2. Sürekli Öğrenme (Continual Learning)	14
2.3. Artımlı Öğrenme	16
2.3.1. Artımlı öğrenme türleri	18
2.3.1.1. Örnek artımlı öğrenme (instance incremental learning)	18
2.3.1.2. Özellik artımlı öğrenme (feature incremental learning)	19
2.3.1.3. Sınıf artımlı öğrenme (class incremental learning)	19
2.4. Yaşam Boyu Öğrenmenin İlgili Öğrenme Paradigmalarıyla Karşılaştırılması	19
3. SINIF DENGESİZLİĞİ VE SINIF ARTIMLI ÖĞRENME	23
3.1. Sınıf Dengesizliği (Class Imbalance-CI)	23
3.2. Sınıf Artımlı Öğrenme (Class Incremental Learning)	25
3.2.1. Sınıf artımlı öğrenmede katastrofik unutma ve sınıf dengesizliği	25
3.2.2. Çözüm stratejileri	26
3.2.2.1. Bellek tekrarı (replay-based methods)	26
3.2.2.2. Düzenlileştirmeye dayalı yöntemler (regularization-based)	26
3.2.2.3. Yapı temelli yöntemler (architecture-based methods)	26
3.2.3. Sınıf artımlı öğrenmeyi destekleyen algoritmalar	27
4. ÖNERİLEN EWC DESTEKLİ ORTAK KODLAYICI VE CIL-DOE MİMARİLERİ	33
4.1. EWC Destekli Ortak Kodlayıcı Mimari	34
4.1.1. Model mimarisi	36
4.1.2. Bilginin saklanması için EWC	37
4.1.3. Eğitim Süreci	38
4.1.3.1. İlk Görev Eğitimi	38
4.1.3.2. EWC ceza terimi ile yeni görev eğitimi	38
4.1.3.3. EWC parametrelerinin güncellenmesi	39

4.1.3.4. EWC ceza terimi olmadan yeni görev eğitimi	39
4.1.3.5. Sınıf artımlı öğrenme ve sınıf dengesizliğinin giderilmesi	39
4.1.4. Veri setinin kullanımı	40
4.1.5. Eğitim ve test süreci	40
4.1.6. Değerlendirme metrikleri	41
4.2. Dinamik Çıktı Genişletme ile Hoeffding Tree	42
4.2.1. Model mimarisi	43
4.2.2. Hoeffding Tree	43
4.2.3. CIL-DOE öğrenme süreci	44
4.2.4. Veri setinin kullanımı	46
4.2.4.1. Aşamalı (kademeli) sınıf tanıtımı deneyi	47
4.2.4.2. Tam akış (end-to-end) deneyi	47
4.2.5. Değerlendirme metrikleri	48
4.3. Modellerin Karşılaştırılması	49
5. GELİŞTİRİLEN YÖNTEMLERİN BAŞARIM ANALİZİ	53
5.1. Uygulamalarda Kullanılan Veri Seti	53
5.2. EWC Destekli Ortak Kodlayıcılı Mimari Deneysel Sonuçlar	54
5.2.1. Sınıf çıkarmanın etkisi	57
5.2.2. Sınıf artımlı öğrenme başarımı	59
5.3. Dinamik Çıktı Genişletme ile Hoeffding Tree Deneysel Sonuçlar	59
5.3.1. Analiz sonuçları	60
5.4. Modellerin Karşılaştırmalı Analizi	65
6. SONUÇ ve TARTIŞMA	67
6.1. EWC Destekli Ortak Enkoder Mimarisinin Değerlendirmesi	67
6.2. CIL-DOE Yaklaşımının Genel Değerlendirmesi	67
6.3. Karşılaştırmalı Değerlendirme	68
KAYNAKLAR	69
ÖZGEÇMİŞ	75

KISALTMALAR

CCIoT2023	: Canadian Institute for Cybersecurity IoT 2023
CI	: Class Imbalance
CIL	: Class-Incremental Learning
CIL-DOE	: Class-Incremental Learning with Dynamic Output Expansion
CPU	: Central Processing Unit
EWC	: Elastic Weight Consolidation
GAN	: Generative Adversarial Networks
IoT	: Internet of Things
LL	: Lifelong Learning
MCC	: Matthews Correlation Coefficient
MTL	: Multi-Task Learning
OL	: Online Learning
RL	: Reinforcement Learning
SMOTE	: Synthetic Minority Over-sampling Technique



SİMGELER

Acc	: Doğruluk
f(x)	: Model çıktısı
F1	: F1 skoru
$\mathcal{L}$	: Kayıp fonksiyonu
MCC	: Matthews Korelasyon Katsayısı
Precision	: Kesinlik
Recall	: Duyarlılık
t	: Zaman adımı
x	: Girdi örneği
y	: Sınıf etiketi
$\hat{y}$	: Tahmin edilen etiket
θ	: Model parametreleri
κ	: Cohen's Kappa katsayısı
μ	: Ortalama
σ	: Standart sapma
Ω	: Parametre önem derecesi
$\mathcal{D}$	: Veri kümesi
$\mathcal{T}_i$	: i. görev



TABLO LİSTESİ

Sayfa

Tablo 2.1. Yaşam boyu öğrenmenin ilgili öğrenme paradigmalarıyla karşılaştırmalı değerlendirmesi.	22
Tablo 3.1. Önerilen modellerin ve literatürdeki yaklaşımların karşılaştırması.	29
Tablo 3.2. Benzer çalışmalar ile önerilen yöntemlerin karşılaştırması.	30
Tablo 4.1. Önerilen modellerin karşılaştırması.	50
Tablo 5.1. Genel performans metrikleri.	55
Tablo 5.2. Sınıf dengeli doğruluk.	56
Tablo 5.3. Bazı sınıflara ait özel performans metrikleri.	58
Tablo 5.4. Artımlı öğrenme performansı.	59
Tablo 5.5. Karşılaştırmalı performans özeti.	66



ŞEKİL LİSTESİ

Sayfa

Şekil 2.1. Artımlı öğrenme yaklaşımı.	18
Şekil 4.1. EWC destekli ortak kodlayıcı.	35
Şekil 4.2. Diyagonal fisher yakınsaması.	38
Şekil 4.3. Önerilen CIL-DOE çerçevesinin sistem mimarisi.	43
Şekil 4.4. CIL-DOE: dinamik çıktı genişletme ile sınıf artımlı öğrenme.	45
Şekil 4.5. Önerilen modellerin karşılaştırmalı grafiği.	51
Şekil 5.1. İlk ve son eğitim aşamalarında EWC ile ve EWC olmadan doğruluk ve makro F1 skoru karşılaştırması.	57
Şekil 5.2. Sınıf artımlı öğrenme kurulumundaki tüm eğitim aşamalarında zaman içindeki genel performans ölçümleri.	61
Şekil 5.3. Sınıf bazında performans.	62
Şekil 5.4. İstatistiksel güvenilirlik ölçümleri.	63
Şekil 5.5. Tam akış sırasında genel performans ölçümleri.	64
Şekil 5.6. Tam akış sırasında makro denge ölçümleri.	64
Şekil 5.7. İstatistiksel güvenilirlik ölçümleri.	65



SİBER GÜVENLİK SİSTEMLERİ İÇİN DİNAMİK VE ARTIMLI MAKİNE ÖĞRENMESİ YAKLAŞIMLARI

ÖZET

Anahtar kelimeler: Sınıf Artımlı Öğrenme, CIL-DOE, Katastrofik Unutma, Çevrimiçi Öğrenme, Siber Güvenlik, Sınıf Dengesizliği, Yaşam Boyu Öğrenme

Bu tez çalışması, sınıf artımlı öğrenme problemlerine yönelik olarak, bilgi kaybı (katastrofik unutma) ve sınıf dengesizliği gibi temel zorlukların üstesinden gelmeyi amaçlayan iki özgün yaklaşım geliştirmektedir. Çalışmanın temel amacı, farklı öğrenme senaryolarında hem önceki bilgiyi koruyabilen hem de yeni sınıfları etkili şekilde öğrenebilen sürdürülebilir ve uyarlanabilir sistemler ortaya koymaktır. Bu kapsamda, biri EWC (Elastic Weight Consolidation) destekli ortak kodlayıcıdan oluşan bir derin öğrenme temelli mimari, diğeri ise çevrim içi öğrenmeye uygun Sınıf Artımlı Dinamik Çıktı Genişletme (CIL-DOE-Class-Incremental Learning with Dynamic Output Expansion) mekanizmasıyla donatılmış Hoeffding Tree modeli çözüm olarak önerilmektedir.

Geliştirilen ilk model, ortak bir kodlayıcı (encoder) yapısı etrafında şekillendirilmekte ve bilgi koruma amacıyla EWC (Elastic Weight Consolidation) mekanizmasıyla desteklenen bir yapay sinir ağı mimarisi olarak tasarlanmaktadır. Söz konusu model, özellikle sınıf dengesizliği içeren artımlı öğrenme senaryolarında gözlemlenen performans düşüşlerinin önüne geçmeyi amaçlamaktadır. Çalışma kapsamında CICIoT2023 veri seti (Neto ve ark., 2023) kullanılarak, en sık ve en seyrek gözlemlenen sınıfların modelden çıkarıldığı çeşitli senaryolar test edilmektedir. Elde edilen ön bulgular, EWC mekanizmasının uygulanmadığı durumlarda makro ölçekte ciddi performans kayıplarının yaşandığını göstermektedir. Özellikle nadir sınıflarda F1 skorlarının %0 ila %9 arasında değiştiği gözlemlenmektedir. Buna karşın, EWC mekanizmasının uygulanmasıyla birlikte modelin geçmiş bilgilere ait parametrelerini koruyarak sınıf dengesizliğiyle daha etkili bir şekilde başa çıkabildiği görülmektedir. Makro F1 skorlarındaki %82'ye varan artışlar, EWC'nin hem genel başarıma hem de azınlık sınıfların öğrenilmesine anlamlı katkılar sunduğunu ortaya koymaktadır. Ayrıca, EWC'nin etkisiyle modelin zaman içindeki kararlılığının arttığı ve önceki görevlerde öğrenilen bilgilerin, yeni görevler öğrenilirken kaybolmasının büyük ölçüde engellendiği anlaşılmaktadır.

Bu tez çalışması kapsamında geliştirilen ikinci yaklaşım, çevrim içi öğrenmeye uygun biçimde tasarlanmakta olup, sınıf çıktılarının veri akışı sırasında dinamik olarak genişletilmesini sağlayan CIL-DOE stratejisini temel almaktadır. Hoeffding Tree algoritması kullanılarak geliştirilen bu model, gerçek zamanlı sistemler için düşük gecikmeli, maliyet açısından verimli ve sürekli öğrenebilen bir yapı sunmaktadır. Gerçekleştirilen deneysel analizler, mikro ortalama metriklerde (doğruluk, micro-F1, micro-recall) modelin erken aşamalarda %80'in üzerine çıkarak hızlı bir yakınsama sağladığını, son aşamalarda ise bu metriklerin %95 düzeyine ulaştığını göstermektedir. Bu bulgular, modelin yeni sınıfları öğrenme süreci sırasında daha önce öğrenilen sınıflardaki performansını büyük ölçüde koruyabildiğini ortaya koymaktadır.

Makro düzeyde ise, özellikle yeni sınıfların tanıtıldığı geçiş aşamalarında macro-F1 ve macro-recall metriklerinde geçici düşüşler gözlemlenmekte, ancak zamanla bu metriklerde düzenli bir iyileşme eğilimi ortaya çıkmaktadır. Bu durum, modelin nadir ve geç tanıtılan sınıflara yönelik genelleme kapasitesini zamanla geliştirebildiğini, ancak bu sürecin mikro metriklere kıyasla daha yavaş gerçekleşmekte olduğunu göstermektedir. Ayrıca, precision değerlerinin recall değerlerinin üzerinde seyretmesi, modelin tahmin sürecinde daha temkinli davrandığını ve özellikle azınlık sınıflarda yanlış pozitifleri önlemeye çalıştığını göstermektedir. Bu eğilim, sınıf dengesizliğinin etkisiyle modelin karar eşiğini daha sık gözlemlenen sınıflar lehine kaydırmakta olduğuna işaret etmektedir.

Modelin güvenilirliği, Cohen's Kappa ve Matthews Korelasyon Katsayısı (MCC - Matthews Correlation Coefficient) gibi istatistiksel doğruluk ölçütleriyle değerlendirilmektedir. Bu metrikler, hem sınıf artımlı öğrenme senaryolarında hem de veri setinin doğal zaman akışıyla işlendiği deneysel kurulumlarda %95'in üzerinde istikrar sergilemekte olup, modelin yalnızca doğruluk açısından değil, aynı zamanda güvenilirlik bakımından da güçlü bir performans ortaya koyduğunu göstermektedir.

Sonuç olarak, bu tez kapsamında geliştirilen iki farklı yaklaşım, farklı öğrenme gereksinimlerine sahip sistemler için çözüm üretmeyi amaçlamaktadır. Derin öğrenme mimarisine entegre edilen EWC mekanizması, bilgi koruma ve azınlık sınıfların öğrenimi açısından etkili bir yapı sunmaktadır. Öte yandan, CIL-DOE temelli Hoeffding Tree modeli ise çevrim içi, sürekli ve hafif yapıli sistemler için uygun bir sınıf artımlı öğrenme stratejisi olarak öne çıkmaktadır. Her iki yaklaşım da sınıf dengesizliği, unutmama ve yeni sınıflara adaptasyon gibi sınıf artımlı öğrenmenin temel zorluklarına karşı, yüksek genelleme kapasitesine sahip, esnek ve sürdürülebilir çözümler üretme potansiyeli taşımaktadır.

DYNAMIC AND INCREMENTAL MACHINE LEARNING APPROACHES FOR CYBER SECURITY SYSTEMS

SUMMARY

Keywords: Class-Incremental Learning, CIL-DOE, Catastrophic Forgetting, Online Learning, Cybersecurity, Class Imbalance, Lifelong Learning

This thesis explores strategies to effectively manage the intertwined challenges of catastrophic forgetting and class imbalance in Class-Incremental Learning (CIL) scenarios. The primary goal is to create adaptable and sustainable learning systems capable of efficiently retaining previously acquired knowledge while seamlessly acquiring new classes from dynamic, real-time data streams commonly encountered in practical applications.

To address these challenges, the study proposes and rigorously evaluates two innovative approaches:

1. **Elastic Weight Consolidation (EWC) Enhanced Deep Learning Architecture:** This approach employs a sophisticated deep learning model with a shared encoder structure. EWC helps to mitigate catastrophic forgetting by selectively preserving important weights representing previously learned tasks. This enables the model to maintain continuity in knowledge while allowing for the integration of new classes without significant loss of prior knowledge.
2. **Hoeffding Tree-based Online Learning Model with Dynamic Output Expansion (CIL-DOE):** This approach utilizes a Hoeffding Tree framework tailored for online learning, with a mechanism for dynamic output expansion specifically designed for real-time stream processing. This model adapts to the continuous influx of new classes, efficiently updating its classification capabilities in response to evolving data distributions while effectively managing class imbalance as new classes are introduced.

These two approaches, when integrated, establish a comprehensive framework that significantly enhances the effectiveness of machine learning systems in class incremental learning (CIL) scenarios. This framework ultimately contributes to creating more sophisticated and responsive artificial intelligence systems.

The first approach emphasizes the development of deep learning models and specifically targets the prevalent issues associated with CIL, particularly those arising from severe class imbalance. The framework utilizes a shared encoder neural network architecture to address these challenges. Within this architecture, EWC is deployed to protect previously learned parameters. EWC achieves this by implementing a mechanism that penalizes significant updates to critical weights for tasks learned earlier. This strategy enables the model to adapt to new distributions of classes while effectively minimizing the degradation of information retained from prior learning phases.

The CICIOT2023 cybersecurity dataset was selected to simulate various challenging learning scenarios in the experimental evaluation. The ten most frequently occurring or least frequently encountered classes were systematically excluded from the training sequences in this setup. This deliberate exclusion allowed a thorough assessment of the model's performance across differing class distributions. Extensive evaluations were conducted using macro-averaged metrics, which provide an overall performance measure, and class-specific metrics, which delve into the performance of individual classes.

The results from these experiments were telling. Without the application of EWC, the model exhibited significant declines in macro-level performance, particularly with regard to rare classes. In some instances, F1-scores for these classes plummeted to as low as 0%, indicating a failure to classify any samples from those classes accurately. Conversely, when EWC was implemented, the model demonstrated remarkable resilience, achieving macro F1-scores of up to 82%. This improvement underscores the model's enhanced capability to manage minority classes effectively while maintaining overall performance stability across the various learning phases. Such findings highlight the critical importance of incorporating advanced techniques like EWC in CIL frameworks to bolster the robustness and reliability of machine learning systems.

Class-wise analysis reveals that EWC significantly enhances classification accuracy, particularly for low-frequency classes such as Backdoor Malware and SQL Injection. These classes often suffer from inadequate representation in training datasets, making accurate classification challenging. EWC addresses this issue by effectively preserving critical model parameters related to these rarer classes, thereby reducing the risk of catastrophic forgetting, where the model forgets previously learned information when exposed to new data. This preservation of knowledge is essential for ensuring resilient and reliable performance in sequential learning tasks, where data is incrementally introduced over time.

The second approach incorporates a lightweight, real-time learning mechanism based on the Hoeffding Tree algorithm, augmented by a dynamic output expansion framework. This innovative design is specifically tailored for streaming data environments, where the influx of new classes can occur dynamically. Unlike traditional models that require extensive retraining or fixed output structures, this framework allows the system to adapt seamlessly as new classes emerge, thus facilitating continuous learning in real time.

A custom preprocessing pipeline has been developed to support this adaptive learning process. This pipeline ensures compatibility with the stream-learning paradigm by efficiently filtering out incomplete or irrelevant data, converting features to floating-point numbers without normalization, and dynamically assigning unique class identifiers as new classes are first encountered. This methodology streamlines the data processing and enables the model's output layer to expand fluidly in response to new information, allowing for the effective and flexible integration of newly identified classes.

Evaluations conducted across three distinct phases of incremental learning reveal remarkable insights into the model's performance. The micro-averaged metrics (acc, F1 score, and recall) demonstrate a rapid convergence, surpassing 80% in the initial stages and stabilizing at impressive levels above 95% in subsequent phases. This

notable performance underscores the model's capability to assimilate new class information effectively while maintaining its previously acquired knowledge.

In contrast, the macro averages—specifically the F1 score, precision, and recall—exhibit temporary downturns when new classes are integrated into the model. However, these metrics consistently show a trend of gradual improvement over time, reflecting the model's enhanced generalization abilities as it adapts to a more complex data landscape. Interestingly, macro precision reliably exceeds macro recall, suggesting a conservative classification strategy. This strategic approach prioritizes more prevalent classes, reducing the likelihood of false positives for rarer categories. This behavior illustrates the inherent challenges associated with learning from class-imbalanced data streams and underscores the pressing need for further refinement and optimization of the classification methodology.

To evaluate the statistical reliability of the model, we employ Cohen's Kappa and the Matthews Correlation Coefficient, both of which emerge with scores exceeding 0.95 during the final training phases. These robust metrics indicate the model's predictions are accurate and exhibit high statistical consistency across dynamically evolving data distributions. Furthermore, a complementary experiment conducted with the complete dataset, preserved in its natural chronological order, further validates the method's adaptability and efficiency in real-world streaming conditions, showcasing its capability to navigate the complexities of temporal data effectively.

This study presents two complementary strategies to tackle the unique challenges encountered in deep and online learning within class-incremental learning frameworks. The first approach, the EWC-enhanced shared encoder architecture, is designed to preserve and retain knowledge from previously learned tasks effectively. This architecture mitigates the risk of catastrophic forgetting and enhances the learning capabilities for minority classes in batch learning scenarios, ensuring that diverse data representations are maintained across various tasks.

In contrast, the second approach, the CIL-DOE framework, leverages the principles of Hoeffding Trees to provide a robust and efficient solution for processing streaming data. This innovative framework is optimized for low-latency performance, allowing it to swiftly adapt to evolving class distributions without declining accuracy or efficiency. This adaptability is crucial in real-world applications where data streams can change rapidly.

Together, these strategies foster the development of flexible, adaptable, and sustainable learning systems that excel in complex and dynamic data environments. By systematically addressing the dual challenges of catastrophic forgetting and class imbalance, this thesis significantly advances lifelong and incremental learning paradigms, paving the way for more effective machine learning applications that can continuously learn and adapt in real-time.



1. GİRİŞ

Her geçen gün hayatımızda daha çok yer edinen teknoloji, veri kümelerinin hem boyut hem de çeşitlilik olarak sürekli artışına neden olmaktadır. Özellikle Nesnelerin İnterneti (IoT), siber güvenlik, sağlık hizmetleri ve finansal teknolojiler gibi verinin yoğun olduğu sektörlerde, her geçen gün daha fazla sayıda cihazın çevrimiçi hale gelmesiyle birlikte büyük miktarda veri üretilmektedir (Li ve ark, 2023). Bu veriler, yalnızca büyük hacimli olmayıp, aynı zamanda yüksek çeşitlilik göstermekte, çok hızlı bir şekilde akmakta ve çoğunlukla sınıf dengesizliği içeren yapılar sergilemektedir (Gandomi ve Haider, 2015; Katal ve ark, 2013). Bu özellikler, verilerin işlenmesini ve anlamlı bilgiye dönüştürülmesini geleneksel yöntemler açısından son derece karmaşık hale getirmektedir. Özellikle geleneksel makine öğrenmesi algoritmaları, genellikle sabit, dengeli ve önceden tanımlanmış veri kümeleriyle eğitilecek şekilde tasarlanmaktadır (Read ve Žliobaitė 2022). Bu nedenle, sürekli değişen ve akış halinde gelen gerçek dünya verilerini işleme konusunda sınırlı bir yetkinliğe sahip olmaktadır. Statik veri varsayımına dayalı bu modeller, dinamik ortamlarda yeni örnekleri ya da sınıfları tanımada yetersiz kalmakta ve bu da öğrenmenin genellenebilirliğini ve doğruluğunu ciddi biçimde etkilemektedir (Zhang ve ark, 2022).

Büyük veri (big data) kavramı, hacim (volume), hız (velocity), çeşitlilik (variety), doğruluk (veracity) ve değer (value) olmak üzere literatürde genellikle “5V” olarak tanımlanmaktadır (Baysal ve Bayılmış, 2019). Bu beş temel özellik, modern veri kümelerinin sahip olduğu karmaşıklığı ve büyüklüğü ortaya koyarken, aynı zamanda bu verilerle çalışan öğrenme sistemlerinin karşılaştığı zorlukları da gözler önüne sermektedir. Söz konusu özelliklerin giderek artması, öğrenme sistemlerinin yalnızca hesaplama kaynakları bakımından değil, aynı zamanda öğrenme stratejileri açısından da daha esnek, çevik ve sürdürülebilir hale gelmesini zorunlu kılmaktadır. Bu gereksinimler doğrultusunda geliştirilen artımlı öğrenme (incremental learning) ve sürekli öğrenme (continual learning) yaklaşımları, modellerin gelen verileri sıralı şekilde işleyerek geçmiş bilgileri kaybetmeden güncellenmesine olanak tanımaktadır.

Bu yaklaşımlar, modelin yaşam boyu öğrenmesini mümkün kılmakla birlikte, dinamik veri ortamlarına etkin bir çözüm sağlamasını da amaçlamaktadır (Parisi ve ark, 2019).

Artımlı öğrenmenin karşılaştığı en temel zorluklardan biri, yeni görevlerin öğrenilmesi sürecinde daha önce edinilmiş bilgilerin unutulmasıyla ortaya çıkan "katastrofik unutma" (catastrophic forgetting) problemidir (McCloskey ve Cohen, 1989). Bu durum, özellikle modelin uzun vadeli bilgi birikimi sağlamasını engellediği için, yaşam boyu öğrenme hedefiyle çelişmektedir. Bir diğer önemli zorluk ise, sınıflar arasında gözlemlenen dengesiz örnek dağılımından kaynaklanan "sınıf dengesizliği" (class imbalance) problemidir (He ve Garcia, 2009). Gerçek dünya uygulamalarında, özellikle siber güvenlik gibi alanlarda az görülen bazı sınıflar kritik öneme sahip olmasına rağmen, veri setlerinde yeterince temsil edilmeyebilmektedir (Lopez-Ledezma ve Velarde, 2025). Bu durum, modelin azınlık sınıfları yeterince öğrenememesine ve sonuç olarak çoğunluk sınıflara karşı önyargılı tahminlerde bulunmasına yol açmaktadır.

Bunun yanı sıra, geleneksel sınıflandırma modelleri genellikle sabit bir çıktı katmanını varsayımına dayanmaktadır. Bu yapı, sınıf artımlı öğrenme (Class-Incremental learning, CIL) gerektiren durumlarda çeşitli yapısal sınırlamalara yol açmaktadır (Rebuffi ve ark, 2017). Yeni sınıfların öğrenilmesi gerektiğinde, bu tür modellerin tüm ağ yapısıyla birlikte baştan eğitilmesi gerekmektedir. Bu nedenle hem zaman hem de hesaplama kaynakları açısından önemli bir maliyet oluşturmaktadır (Masana ve ark, 2020). Sürekli veri akışının söz konusu olduğu gerçek dünya senaryolarında ise, modelin karşılaştığı yeni sınıfları dinamik olarak tanıyabilmesi ve çıktı katmanını yeniden eğitim gerektirmeksizin genişletebilmesi büyük önem taşımaktadır. Artımlı öğrenme yaklaşımları bu tür sistemlerin uzun vadeli ölçeklenebilirliği açısından bir gereklilik hâline gelmektedir (Zhou ve ark, 2023).

Bu tez çalışması, yukarıda belirtilen temel sorunlara çözüm getirmeyi amaçlayan iki özgün model önermektedir. İlk model, EWC (Elastic Weight Consolidation) destekli ortak bir encoder mimarisi kullanarak, geçmiş görevlerden edinilen bilgilerin korunmasını ve sınıf dengesizliğinin hâkim olduğu veri setlerinde öğrenmenin iyileştirilmesini hedeflemektedir. Aynı zamanda bu model, yeni sınıflarla karşılaştığında bu sınıfları da öğrenebilme kapasitesine sahiptir, böylece hem bilgi koruma hem de çıkışın genişleyebilmesi açısından güçlü bir yapı sunmaktadır. İkinci model ise, Hoeffding Tree algoritması üzerine inşa edilen dinamik çıkış genişlemesi

(Dynamic Output Expansion-DOE) yaklaşımını temel almakta olup, modelin sürekli akan veri ortamlarında karşılaştığı yeni sınıfları tanımasını ve bu sınıflara gerçek zamanlı olarak uyum sağlamasını mümkün kılmaktadır. Her iki yaklaşım da artımlı öğrenmenin gerektirdiği esneklik, bilgi sürekliliği ve yeni sınıflara adaptasyon yetkinliklerini bir arada ele alarak, geleneksel yöntemlerin aşamadığı sınırlamaların üstesinden gelmeyi amaçlamaktadır.

EWC destekli ortak kodlayıcı mimari, sıralı ve parçalanmış veri kümeleri üzerinde test edilmiş, geleneksel sınıf artımlı öğrenme senaryolarında bilgi korunumu ve sınıf dengesizliğiyle başa çıkma becerisi açısından değerlendirilmektedir. Diğer yandan, Hoeffding Tree temelli model, Apache Kafka altyapısı kullanılarak akan biçimde işlenen CICIOT2023 veri seti üzerinde test edilmektedir. Bu sayede, dağıtık sistem mimarileriyle entegre çalışabilen, kaynak kullanımı açısından verimli ve gerçek zamanlı uygulama ortamlarına uygun sürdürülebilir öğrenme modellerinin geliştirilmesi hedeflenmektedir. Özellikle gerçek zamanlı veri akışlarının yoğun olduğu uygulama senaryolarında, bu tür mimarilerin kullanımı modellerin çeviklik ve işlevsel verimlilik açısından önemli avantajlar sunmasını sağlamaktadır.

1.1. Problemin Tanımı

Verilerin hacmi, hızı ve çeşitliliği benzeri görülmemiş bir oranda artarken, geleneksel makine öğrenimi modelleri gerçek zamanlı veri akışlarını etkili bir şekilde yönetmede önemli zorluklarla karşılaşmaktadır. Özellikle sınıf dengesizliği, kavram kayması (concept drift) ve katastrofik unutma gibi problemler, bu modellerin doğruluk, genellenebilirlik ve adaptasyon yeteneklerini ciddi şekilde kısıtlamaktadır. Bu nedenle, sürekli öğrenme (continual learning) ve artımlı öğrenme (incremental learning) gibi yaklaşımlar, değişen veriler karşısında bilgi kaybını en aza indirerek daha esnek ve sürdürülebilir çözümler sunma potansiyeline sahiptir. Modern uygulamalarda bu tür yaklaşımlar, yalnızca daha doğru sonuçlar üretmekle kalmayıp, aynı zamanda sistemlerin uzun vadeli etkinliğini de artırmaktadır (Gama ve ark, 2014; Wang ve ark, 2024). Gerçek zamanlı veri akışlarının olduğu ortamlarda, modellerin tüm geçmiş veri kümesine erişimi olmadan gelişen verilerden sıralı olarak öğrenmesi gerekmektedir. Bu zorluk, artımlı veya sürekli öğrenme kavramları aracılığıyla ele alınmaktadır (Parisi ve ark, 2019).

Sürekli öğrenme sistemleri, sıklıkla “katastrofik unutma” olarak adlandırılan kritik bir sorunla karşı karşıya kalmaktadır. Bu problem, modelin yeni görevleri öğrenirken daha önceki görevlerden edindiği bilgileri kaybetmesi, yani üzerine yazması sonucunda ortaya çıkmaktadır (McCloskey ve Cohen, 1989). Bunun yanı sıra, gerçek dünya veri akışları çoğunlukla ciddi düzeyde sınıf dengesizliği içermektedir. Özellikle nadir görülen ancak büyük öneme sahip olan saldırı türleri gibi sınıflar, veri setlerinde genellikle yetersiz temsil edilmektedir. Bu dengesizlik, modellerin çoğunluk sınıflara karşı önyargılı hale gelmesine neden olmakta ve azınlık sınıflar üzerindeki performansın düşmesiyle sonuçlanmaktadır. Dolayısıyla, tüm sınıflar genelinde düşük genelleme başarımı gözlemlenmektedir (He ve Garcia, 2009).

Geleneksel sınıflandırma modelleri, genellikle önceden tanımlanmış sabit bir sınıf kümesine göre tasarlanmış statik çıktı katmanlarından oluşmaktadır. Bu mimari yapı, eğitim aşamasında belirlenmiş sınıflar dışındaki herhangi bir yeni sınıfın sisteme entegre edilmesini mümkün kılmadığından, esneklikten gösterememektedir. Bu durum, özellikle sınıf artımlı öğrenme gibi senaryolarda ciddi bir sınırlılık oluşturmaktadır. Bu tür senaryolarda, sistemin dağıtımına alındıktan sonra zamanla karşılaşabileceği yeni sınıfları tanıyabilmesi ve bu sınıflara uygun çıktılar üretebilmesi beklenmektedir (Rebuffi ve ark, 2017). Ancak statik çıktı katmanları bu ihtiyacı karşılayamamakta dolayısıyla modele yeni sınıfların eklenmesi genellikle tüm ağı sıfırdan yeniden eğitmeyi gerektirmektedir. Bu yeniden eğitim süreci yalnızca yüksek hesaplama kaynakları gerektirmekle kalmamakta, aynı zamanda gerçek zamanlı veri akışlarıyla çalışan sistemler için de pratik bir çözüm sunmamaktadır. Özellikle siber güvenlik, IoT ya da izleme gibi sistemlerde sınıfların zamanla ortaya çıktığı ya da öğrenilmesi gereken bilgilerin sürekli değiştiği uygulamalarda, yeniden eğitim süreçleri operasyonel verimliliği olumsuz etkilediği gibi kritik sorunlara da yol açabilmektedir (Masana ve ark, 2020). Dolayısıyla, bu sınırlamayı aşmak için dinamik çıktı genişletme mekanizmalarına veya yeniden eğitime ihtiyaç duymadan modele yeni sınıfları entegre edebilen yaklaşımlara duyulan ihtiyaç giderek artmaktadır.

Bu tür çözümlerin Apache Kafka gibi akış platformları ve Apache Spark gibi işleme motorlarıyla entegre edilmiş dağıtık büyük veri ekosistemlerine dahil edilmesi, beraberinde ölçeklenebilirlik, gecikme süresi ve kaynak verimliliği gibi konularda ek zorluklar doğurmaktadır (Montiel ve ark, 2018). Mevcut birçok yöntem, bu altyapısal engelleri yeterince ele alamamakta, çoğunlukla teorik çerçevelerle sınırlı kalmakta ya

da yalnızca küçük ölçekli deneysel uygulamalarla test edilmektedir. Bu durum, gerçek zamanlı ve yüksek hacimli veri ortamlarında bu yöntemlerin uygulanabilirliğini ciddi şekilde kısıtlamaktadır.

Bu tez çalışması, mevcut yöntemlerin pratik sınırlamaları ve günümüz veri ortamlarının gerektirdiği bazı kritik ihtiyaçlar doğrultusunda şekillenmektedir. Öncelikle, sürekli veri akışlarıyla çalışabilen ve bu akışlardan artımlı biçimde öğrenme yeteneğine sahip uyarlanabilir modellere duyulan ihtiyaç giderek artmaktadır. Buna ek olarak, tarihsel verilere yeniden erişim gerektirmeden önceki bilgileri koruyabilen mekanizmaların geliştirilmesi büyük önem taşımaktadır. Ayrıca, modellerin karşılaştıkları yeni sınıfları tanıyabilmesi ve çıktı katmanlarını bu sınıfları kapsayacak şekilde dinamik olarak genişletebilmesi, özellikle sınıf artımlı öğrenme bağlamında kritik bir gereklilik hâline gelmektedir. Yüksek hızlı ve gerçek dünya koşullarını yansıtan veri kümelerinde sıkça rastlanan sınıf dengesizliğinin etkin şekilde yönetilmesi de bu alandaki diğer önemli zorluklardan biri olmaktadır. Son olarak, tüm bu bileşenlerin, gerçek zamanlı ve dağıtık sistem mimarilerine entegre edilebilecek, kaynakları verimli kullanan uygulamalarla desteklenmesi gerekmektedir.

Bu tez çalışması, yukarıda belirtilen problemlerin çözümü için iki yenilikçi yaklaşım önermektedir;

1. Birinci yaklaşım, önceden öğrenilmiş temsilleri korurken aynı zamanda göreve özgü çıktı uyarlamasını kolaylaştırmayı hedefleyen, EWC destekli ortak kodlayıcı (encoder) mimarisine dayanmaktadır.
2. İkinci yaklaşım, akış verileri içerisinde zamanla ortaya çıkan yeni sınıflara verimli ve gerçek zamanlı bir şekilde uyum sağlayabilen, dinamik çıktı genişletme mekanizmasıyla güçlendirilmiş Hoeffding Tree temelli bir sınıf artımlı öğrenme sistemidir.

Her iki yöntem de sürekli öğrenmenin gerektirdiği bilgi koruma, yeni sınıflara adaptasyon ve veri akışı ortamlarına uyarlanabilirlik gibi temel gereksinimlere çözüm sunmaktadır. Birinci yaklaşımda EWC destekli ortak kodlayıcı mimari, sıralı ve parçalanmış veri kümeleri üzerinde test edilerek, geleneksel sınıf artımlı öğrenme senaryolarında bilgi kaybı ve sınıf dengesizliği gibi temel zorluklara çözüm sunması amaçlanmaktadır. Bu amaç doğrultusunda EWC mekanizması, paylaşılan bir encoder yapısı ve göreve özgü çıktı katmanlarıyla birleştirilerek, modelin geçmiş bilgileri

koruyarak yeni sınıfları öğrenmesi sağlanmaktadır. Ayrıca sınıf frekansına duyarlı ağırlıklandırma ve dengeli kayıp fonksiyonu gibi bileşenlerle desteklenen bu yapı, azınlık sınıfların daha etkili öğrenilmesini mümkün kılmaktadır.

İkinci yaklaşım, Hoeffding Tree temelli dinamik çıktı genişletme modeli ise, Apache Kafka destekli bir dağıtık akış mimarisi üzerine inşa edilmektedir. Bu model, gerçek zamanlı veri akışları sırasında sınıf artımlı öğrenmenin temel gereksinimlerinden biri olan yeni sınıflara otomatik ve kesintisiz adaptasyon yeteneğini hedeflemektedir. Kafka üzerinden akan CICIOT2023 veri setiyle test edilen bu sistem, sınıf etiketleri önceden bilinmeyen durumlarda çıktı katmanını dinamik olarak genişleterek yeniden eğitime gerek duymadan çalışabilmektedir. Bu yönüyle, büyük ölçekli, yüksek hızlı ve durağan olmayan veri ortamlarında hem çeviklik hem de kaynak verimliliği sağlayarak gerçek dünya uygulamalarına entegre edilebilecek sürdürülebilir bir öğrenme çözümü sunmaktadır.

1.2. Tezin Amacı

Bu tez çalışması, sınıf artımlı öğrenme bağlamında ortaya çıkan katastrofik unutma ve sınıf dengesizliği gibi temel sorunlara karşı, ölçeklenebilir, uyarlanabilir ve kaynak açısından verimli öğrenme mimarileri geliştirmeyi amaçlamaktadır. Bu doğrultuda, iki farklı yaklaşım önerilmektedir. İlk olarak, sıralı veri yapıları üzerinde çalışmak üzere yapılandırılan ve EWC destekli bir ortak kodlayıcı mimarisi geliştirilmektedir. Bu yapı, bilgi korunumu ve azınlık sınıfların öğrenilmesi konularına odaklanmaktadır. İkinci yaklaşım ise, Apache Kafka kullanılarak bir veri akışı altyapısı ile entegre çalışan, Hoeffding Tree algoritması üzerine kurulu ve dinamik çıktı genişletme mekanizmasıyla desteklenen bir model sunulmaktadır. Bu yaklaşım, durağan olmayan ve gerçek zamanlı veri akışları içerisinde yeni sınıfların otomatik olarak tanınması ve modellenmesine olanak sağlamaktadır. Her iki yöntem, sınıf artımlı öğrenme senaryolarının bilgi sürdürülebilirliği, yeni sınıflara adaptasyon ve dağıtık sistem entegrasyonu gibi temel gereksinimlere katkı sunmayı hedeflemektedir.

Bu tez çalışması, artımlı ve sürekli öğrenme sistemlerinin gerçek dünya koşullarında karşılaştığı temel zorluklara yenilikçi iki çözüm önermektedir. Birinci çözümde, EWC yöntemi temel alınarak bir ortak kodlayıcı mimarisi tasarlanması hedeflenmektedir. Bu mimari, paylaşılan özellik temsillerinin genelliğini ve sağlamlığını korurken, göreve özgü sınıflandırma katmanlarını içerecek şekilde yapılandırılmaktadır.

Modelin öğrenme sürecinde parametre konsolidasyonu ile birlikte dengeli bir eğitim stratejisinin entegre edilmesi yoluyla hem katastrofik unutmanın önlenmesi hem de sınıf dengesizliğinin olumsuz etkilerinin azaltılması amaçlanmaktadır.

İkinci çözümde ise Hoeffding Tree sınıflandırıcısının veri akışı sırasında karşılaştığı yeni sınıfları dinamik biçimde tanıyabilmesine olanak sağlayan bir çıktı genişletme mekanizması önerilmektedir. Bu çerçevede, sınıflandırıcının gerçek zamanlı olarak değişen etiket alanına uyum sağlamasını mümkün kılmaktadır. Böylece modelin tamamen yeniden eğitilmesine olan gereksinimi ve sınıf etiketlerine ilişkin ön bilgi ihtiyacını ortadan kaldırılması amaçlanmaktadır. Ayrıca, Hoeffding Tree temelli dinamik çıktı genişletme yaklaşımının, Apache Kafka destekli bir dağıtık mimariye entegre edilmesi ile önerilen çözüm gerçek sistem üzerinde çalışıldığı gösterilmektedir. Bu entegrasyonun, geliştirilen yöntemin ölçeklenebilirlik, hata toleransı ve yüksek hacimli veri işleme kapasitesi gibi sistem düzeyindeki yetkinliklerini değerlendirmeye olanak tanınması beklenmektedir. Böylece yalnızca teorik düzeyde değil, aynı zamanda pratik uygulamalarda da kullanılabilirliği yüksek, sürdürülebilir bir öğrenme sisteminin geliştirilmesi amaçlanmaktadır.

Geliştirilen yöntemlerin etkinliği, CICIOT2023 veri seti ile gerçekleştirilecek deneysel çalışmalarla değerlendirilmektedir. Bu değerlendirme kapsamında, aşamalı sınıf tanıtımları, sınıf frekans manipülasyonu ve tam kronolojik akış testleri gibi farklı senaryolar kullanılmaktadır. Model başarımı, doğruluk ve F1 skoru gibi geleneksel ölçütlerin yanı sıra, sınıfa özgü hatırlama, adaptasyon kabiliyeti ve dengesiz öğrenme davranışını yansıtan makro ortalamalı metriklerle analiz edilmektedir. Tüm bu hedefler doğrultusunda çalışma, özellikle siber güvenlik ve IoT sistemlerinde anomali tespiti gibi dinamik ortamlarda uygulanabilecek, artımlı ve sürekli öğrenme yeteneklerine sahip, esnek, yüksek performanslı ve dağıtıma hazır modeller geliştirilmesine katkı sunmayı amaçlamaktadır.

1.3. Tez Çalışmasının Katkıları

Bu tez çalışması, sınıf artımlı öğrenmeye yönelik özgün katkılardan oluşan bütüncül bir yaklaşım sunmaktadır. Çalışmanın temel hedefleri arasında, sürekli öğrenme süreçlerinde ortaya çıkan katastrofik unutmanın azaltılması, sınıf dengesizliğinin etkili biçimde yönetilmesi ve özellikle gerçek zamanlı uygulamalarda dinamik model adaptasyonunun sağlanması yer almaktadır. Sunulan katkılar, yalnızca model tasarımı

düzeyinde değil, aynı zamanda algoritmik yaklaşımlar ve sistem mimarisine yönelik entegrasyon çözümleri aracılığıyla hem teorik hem de pratik boyutta ele alınmaktadır. Bu kapsamda geliştirilen çerçeveler, artımlı öğrenme alanında sürdürülebilirlik, ölçeklenebilirlik ve uygulanabilirlik ilkeleri doğrultusunda anlamlı bir ilerleme sağlamayı amaçlamaktadır.

Bu araştırmanın temel katkıları aşağıdaki gibi sıralanabilir:

1. Sınıf dengesizliği altında sınıf artımlı öğrenme için yeni ve özgün bir EWC destekli ortak kodlayıcı çerçevesi önerilmesi;

Bu çalışma, yeni sınıfların verimli bir şekilde öğrenilmesini sağlarken önceki bilgilerin korunmasını da amaçlayan, EWC yöntemini paylaşılan bir kodlayıcı yapısıyla ve göreve özgü çıktı katmanlarıyla bütünleştiren hibrit bir mimari önermektedir. Önerilen model, sıralı öğrenme sürecinde sınıf dengesizliğinin olumsuz etkilerini en aza indirmek amacıyla dengelenmiş bir kayıp fonksiyonu ile sınıf frekansına duyarlı ağırlıklandırma stratejilerini birlikte kullanmaktadır. Bu yaklaşım, daha önce edinilen temsillerin üzerine yazılmasını engelleyerek bilgi sürekliliğini korumakta, aynı zamanda modüler ve ölçeklenebilir bir adaptasyon yapısıyla sürekli öğrenme literatürüne anlamlı bir katkı sunmaktadır.

2. Gerçek zamanlı akış senaryolarında Hoeffding Tree için dinamik bir çıktı genişletme mekanizması önerilmesi;

Bu tez çalışması, karar ağaçlarına dayalı bir sınıflandırıcı yapısı kullanarak, daha önce karşılaşılmamış sınıf etiketleriyle karşılaşıldığında çıktı alanını dinamik biçimde güncelleyebilen uyarlanabilir bir model önermektedir. Hoeffding Tree algoritmasından yararlanılarak geliştirilen bu yaklaşım, modelin yeniden eğitime gerek duymadan gerçek zamanlı veri akışı içerisinde yeni sınıflara yanıt verebilmesini mümkün kılmaktadır. Bu yönüyle, önceden tanımlanmış etiket kümelerine veya sabit görev sınırlarına olan bağımlılığa çözüm sunmaktadır. Söz konusu yöntem, geleneksel ağaç sınıflandırıcıların sınıf artımlı öğrenme bağlamında karşılaştığı temel kısıtlamaları aşmayı hedeflemektedir. Literatürde karar ağaçlarıyla çalışan modeller arasında, gerçek zamanlı akış ortamında otomatik çıktı genişletme yeteneğini sınıf artımlı öğrenme ile birlikte ele alan uygulamaların oldukça sınırlı olduğu görülmektedir. Bu bağlamda, önerilen yaklaşımın bu alandaki özgün örneklerden biri olduğu değerlendirilmektedir.

3. Hoeffding Tree temelli dinamik çıktı genişletme yaklaşımının Apache Kafka kullanılarak bir akış mimarisine entegrasyonu ile gerçek zamanlı sistem üzerinde testinin gerçekleştirilmesi;

Bu çalışma, Hoeffding Tree temelli dinamik çıktı genişletme yaklaşımını Apache Kafka temelli dağıtık bir akış işleme mimarisiyle bütünleştirerek, algoritma düzeyindeki yenilikleri sistem düzeyinde uygulamalarla buluşturmayı amaçlamaktadır. Bu entegrasyon, geliştirilen modelin yalnızca teorik çerçevede değil, aynı zamanda gerçek dünya koşullarında da etkin bir şekilde uygulanabileceğini göstermektedir. Özellikle izinsiz giriş tespiti ve diğer siber güvenlik uygulamaları gibi yüksek hacimli ve gerçek zamanlı veri akışı gerektiren alanlarda, önerilen sistemin ölçeklenebilirlik, hata toleransı ve pratik kullanılabilirlik açısından güçlü bir potansiyel sunduğu değerlendirilmektedir.

4. Kontrollü ve gerçek dünya akış koşulları altında kapsamlı değerlendirme için iki aşamalı deneysel çerçeve;

Önerilen yöntemlerin doğrulanması amacıyla, CICIOT2023 veri seti kullanılarak iki farklı deneysel çerçeve oluşturulmuştur. EWC kullanan model, sıralı ve parçalara ayrılmış veri kümeleriyle yapılandırılmış sınıf artımlı öğrenme senaryosu altında değerlendirilmektedir. Hoeffding Tree ve dinamik çıktı genişletme modeli ise, Apache Kafka üzerinden gerçek zamanlı olarak iletilen veri akışı ile hem aşamalı sınıf tanıtımlarını hem de durağan olmayan ortamlara karşılık gelen kronolojik tam akış senaryosunu içeren bir yapı içerisinde test edilmektedir. Deneysel değerlendirmede, makro ve mikro F1 puanlarının yanı sıra elde tutma (retention) ve adaptasyon kabiliyetini ölçen göstergeler kullanılmaktadır. Ayrıca, model güvenilirliğini daha derinlemesine analiz edebilmek amacıyla Kappa ve Matthews Korelasyon Katsayısı gibi ileri düzey metriklere de yer verilmektedir. Bu kapsamlı metrik seti, zaman içerisinde modellerin performans değişimlerini ortaya koyarak öğrenme davranışlarına ilişkin ayrıntılı iç görüler sağlamaktadır.

5. Akıllı sistemler için esnek, hafif modellerin tasarımına pratik katkılar;

Önerilen Hoeffding Tree temelli model, bellek kullanımı açısından verimli olacak şekilde tasarlanmaktadır. Ayrıca uç cihazlara dağıtımına uygun biçimde geliştirilmiş ve zaman içerisinde değişen sınıf dağılımlarına uyum sağlayabilecek esneklikle yapılandırılmaktadır. Bu özellikler, Nesnelerin İnterneti, siber güvenlik sistemleri ve

akıllı altyapılar gibi giderek önem kazanan uygulama alanlarında kritik bir rol oynamaktadır. Çalışmadan elde edilen bulgular, sürekli öğrenme yaklaşımlarının üretim ortamındaki veri akışlarına entegre edilebileceğini göstermekte ve çevrimiçi öğrenme, dengesiz veri yapıları ile sınıf evrimi gibi dinamik görevler üzerine yapılacak gelecekteki araştırmalar için deneysel açıdan güçlü bir temel sunmaktadır.

1.4. Tezin Organizasyonu

Bu tez çalışmasının organizasyonu aşağıda özetlenen 6 ana bölümden oluşmaktadır.

Bölüm 1 – Giriş: Bu bölümde çalışmanın problem alanı, amacı ve motivasyonu sunulmakta; sınıf artımlı öğrenme bağlamında karşılaşılan temel zorluklar olan katastrofik unutma, sınıf dengesizliği ve gerçek zamanlı adaptasyon gerekliliği tanımlanmaktadır. Ayrıca tezin özgün katkıları ve genel çerçevesi açıklanmaktadır.

Bölüm 2 – Öğrenme Paradigmaları: Bu bölümde yaşam boyu öğrenme, sürekli öğrenme, artımlı öğrenme ve sınıf artımlı öğrenme paradigmatları detaylı biçimde açıklanmaktadır. Kavram kayması, sınıf dengesizliği ve katastrofik unutma gibi temel problemler teorik çerçevede ele alınmaktadır. Ayrıca ilgili çözüm stratejileri (bellek, düzenleştirmeye ve mimariye dayalı yaklaşımlar) ile EWC, Hoeffding Tree ve Kafka destekli sistemler üzerine güncel literatür incelenmektedir..

Bölüm 3 – Sınıf Dengesizliği ve Sınıf Artımlı Öğrenme : Bu bölümde, önerilen iki modelin detaylı yapısı sunulmaktadır. Birinci model, EWC destekli ortak kodlayıcı mimarisiyle bilgi koruma ve sınıf dengesizliği sorunlarına odaklanmaktadır. İkinci model, Hoeffding Tree algoritması ile birlikte dinamik çıktı genişletme mekanizması içeren ve Apache Kafka ile bütünleşik çalışan bir sistem olarak tanıtılmaktadır.

Bölüm 4 – Önerilen EWC Destekli Ortak Kodlayıcı ve CIL-DOE Mimarileri: Bu bölümde, her iki model için oluşturulan deneysel senaryolar tanıtılmaktadır. Veri seti (CICIoT2023), veri işleme biçimleri (sıralı/parçalanmış, gerçek zamanlı akış), sınıf artımlı senaryolar ve sistem mimarisi sunulmaktadır. Ayrıca değerlendirme metrikleri (F1, MCC, kappa, retention/adaptation vb.) açıklanmakta ve deneylerin uygulanma adımları verilmektedir.

Bölüm 5 – Geliştirilen Yöntemlerin Başarım Analizi: Her iki modelin deneysel sonuçları kapsamlı biçimde analiz edilmektedir. Modellerin farklı senaryolar altındaki

performansları, katastrofik unutma ile başa çıkma yetenekleri ve sınıf dengesizliği karşısındaki davranışları karşılaştırmalı olarak değerlendirilmektedir.

Bölüm 6 – Sonuç ve Tartışma: Tezin genel bulguları özetlenmekte, önerilen yaklaşımların pratik faydaları ve literatüre katkıları tartışılmaktadır. Ayrıca, gelecekte yapılacak çalışmalar için öneriler sunulmaktadır.





2. ÖĞRENME PARADİGMALARI

Klasik makine öğrenmesi, bir modelin belirli bir görev için sabit bir veri kümesi üzerinde eğitilmesini ve sonrasında bu modelin uygulama aşamasında kullanılarak performansının değerlendirilmesini ilkesine dayanmaktadır. Ancak bu paradigma, her öğrenme sürecini izole biçimde ele almaktadır (Hoi ve ark, 2021). Geçmiş görevlerden öğrenilen bilgiler yeni görevlerde tekrar kullanılmamaktadır. Bu yaklaşımın birtakım sınırlılıkları bulunmaktadır. İlk olarak, model eğitimi sonrasında edinilen bilgi sistem tarafından tutulmamaktadır. Bu nedenle her yeni görev için modelin sıfırdan eğitilmesini gerekmektedir. Ayrıca klasik sistemler genellikle statik ve kapalı veri ortamlarında çalışmak üzere tasarlanmaktadır (Suárez-Cetrulo ve ark, 2023). Ancak gerçek dünya verisi dinamik, sürekli değişmektedir. Tüm bu nedenlerle, klasik makine öğrenmesi paradigmaları günümüzün karmaşık ve değişken veri ortamlarında sınırlı etkiye sahip olmaktadır (Zhu ve ark, 2025).

Makine öğrenmesi sistemlerinin gerçek dünyadaki değişken, sürekli ve yüksek hacimli veri akışlarıyla başa çıkabilmesi, klasik öğrenme yöntemlerinin ötesinde yeni yaklaşımlar gerektirmektedir (Baysal ve Bayılmış, 2025). Bu kapsamda artımlı öğrenme (incremental learning), sürekli öğrenme (continual learning) ve yaşam boyu öğrenme (lifelong learning) kavramları, sistemlerin zaman içerisinde gelen verilerle modelini güncelleyebilmesi ve daha önce öğrenilmiş bilgileri unutmadan yeni bilgileri öğrenebilmesi amacıyla geliştirilen yöntemleri ifade etmektedir (Parisi ve ark, 2019; Hoi ve ark, 2021; Delange ve ark, 2021).

2.1. Yaşam Boyu Öğrenme (Lifelong Learning, LL)

Yaşam boyu öğrenme, bir makine öğrenmesi sisteminin zaman içinde karşılaştığı çok sayıda farklı görevden edindiği bilgileri biriktirerek, bu bilgileri gelecekteki görevlerde yeniden kullanabilmesini amaçlayan gelişmiş bir öğrenme paradigmaları bütünüdür. Bu yaklaşım, yalnızca bir görevi öğrenip sonrasında pasifleşen klasik sistemlerin aksine, bilginin sürekli edinimi, korunması ve yeniden kullanımı gibi insan benzeri öğrenme yeteneklerini yapay sistemlere kazandırmayı hedeflemektedir (Chen ve Liu, 2018).

Yaşam boyu öğrenmenin temel özellikler aşağıda gösterildiği şekilde sıralanabilmektedir:

1. Süreklilik: LL, klasik makine öğrenmesinden farklı olarak, öğrenme sürecini tek seferlik bir işlem olarak değil, zaman içinde kesintisiz devam eden bir süreç olarak ele almaktadır (Chen ve Liu, 2018).
2. Bilgi Tabanı Kullanımı: Öğrenilen bilgiler, sistem belleğinde bir bilgi tabanı aracılığıyla saklanmakta ve gelecekteki görevlerde yeniden kullanılmak üzere korunmaktadır (Thrun ve Mitchell, 1995).
3. Çeşitli Bilgi Türlerinin Saklanması: Bu bilgi yalnızca modelleri değil, öznitelikleri, temsilleri ve çıkarımları da kapsamaktadır (Hess ve ark, 2024).
4. Bilgi Transferi: Sistem, önceki görevlerden elde edilen bilgileri yeni görevlerde aktif biçimde kullanarak öğrenme verimliliğini artırmakla birlikte veri ihtiyacını azaltmaktadır (Zhu ve ark, 2020).
5. Görev Keşfi: LL sistemleri, yalnızca verilen görevlerle sınırlı kalmayıp aynı zamanda çevredeki yeni öğrenme fırsatlarını da kendiliğinden keşfetme yeteneğine sahip olmaktadır (Yang ve ark, 2024).
6. İş Başında Öğrenme: Öğrenme süreci, yalnızca eğitim aşamasıyla sınırlı kalmayıp aynı zamanda sistem çalışırken, uygulama sırasında da öğrenme devam edebilmektedir (Losing ve ark, 2018).
7. Uyarlanabilir ve Bilgi Temelli Yapılar: Bu bütüncül yaklaşım sayesinde LL sistemleri, çevresel değişkenlere dinamik biçimde uyum sağlayabilen esnek ve öğrenen yapılara dönüşmektedir (Parisi ve ark, 2019).

Yaşam boyu öğrenme, çoklu görevler boyunca edinilen bilginin kalıcı olarak biriktirilmesini ve yeniden kullanılmasını hedefleyen geniş kapsamlı bir yaklaşımdır. Bu bağlamda, ardışık görevler üzerinde unutmayı önleyerek öğrenmeyi sürdüren sürekli öğrenme paradigmasını içermektedir (Chen ve Liu, 2018).

2.2. Sürekli Öğrenme (Continual Learning)

Sürekli öğrenme, yapay zekâ sistemlerinin zaman içinde karşılaştığı yeni görevleri öğrenirken daha önce öğrendiği görevleri unutmadan bilgi edinmesini amaçlayan bir öğrenme paradigmasıdır. Bu yaklaşım, özellikle yaşam boyu öğrenme çatısı altında önemli bir yer tutmakta olup, insan benzeri bilgi birikimi ve adaptasyon yeteneğini yapay sistemlere kazandırmayı hedeflemektedir (Chen ve Liu, 2018).

Sürekli öğrenme, bir modelin zaman içinde ardışık olarak sunulan görevleri öğrenirken, önceki görevlerden elde ettiği bilgileri koruyabilmesini ifade etmektedir. Bu öğrenme yaklaşımı, geleneksel makine öğrenimi yöntemlerinden farklı olarak, tüm veriye aynı anda erişim gerektirmeksizin, bilgi birikimini sürdürülebilir şekilde güncelleyebilme yetisine dayanmaktadır (Parisi ve ark, 2019; Gama ve ark, 2014). Siber güvenlik, otonom sistemler ve IoT altyapıları gibi sürekli değişen ortamlar veri akışının zamanla değişime uğradığı durumları içermektedir. Bu süreçte karşılaşılan temel zorluklardan biri, modelin yeni görevleri öğrenirken önceki görevlerde edindiği bilgileri unutmasıdır. Bu problem literatürde katastrofik unutma (catastrophic forgetting) olarak bilinmektedir (Parisi ve ark, 2018).

Sürekli öğrenme yaklaşımları, literatürde genellikle üç temel kategori altında sınıflandırılmaktadır, bunlar düzenlileştirmeye dayalı yaklaşımlar, bellek temelli yaklaşımlar ve model temelli yaklaşımlardır (Parisi ve ark, 2019; Delange ve ark, 2021). İlk kategori olan düzenlileştirmeye dayalı yöntemler (regularization-based methods), modelin daha önce öğrendiği görevleri unutmadan yeni görevleri öğrenmesini sağlamak amacıyla, belirli parametrelerin değişimini sınırlandıran düzenleyici terimler kullanmaktadır. Bu yöntemler, özellikle ağırlıkların geçmiş görevler için önem derecesine göre korunmasını hedeflemektedir. Bu yaklaşıma en bilinen örneklerden biri, sinir ağı parametrelerinin önemini ölçerek güncellemeleri denetleyen EWC yöntemidir (Kirkpatrick ve ark, 2017). İkinci kategori olan bellek temelli yaklaşımlar (memory-based methods), geçmiş görevlerden elde edilen verilerin veya temsil örneklerinin bellekte saklanması ve yeni görevler öğrenilirken bu örneklerin yeniden kullanılmasını öngörmektedir. Bu yöntemler, geçmiş bilgiye doğrudan erişimi mümkün kılarak unutmayı azaltmaktadır. Bu gruba örnek olarak Experience Replay ve Rehearsal stratejileri verilebilmektedir (Lopez-Paz ve Ranzato, 2022). Üçüncü kategori olan model temelli ya da dinamik mimari yaklaşımları (model-based or dynamic architectures) ise, yeni görevlerle karşılaşıldığında modelin yapısının dinamik olarak genişletilmesini veya modifiye edilmesini içermektedir. Bu sayede sistem, görevler arası etkileşimleri daha esnek şekilde ele alabilmektedir. Örneğin, Progressive Neural Networks (Rusu ve ark, 2022), her yeni görev için ayrı bir ağ bileşeni ekleyerek bilgi birikimini ve transferini sağlamaktadır. Bu üç yaklaşım, katastrofik unutma (catastrophic forgetting) problemini azaltmak için farklı stratejiler geliştirerek sürekli öğrenmenin uygulanabilirliğini güçlendirmektedir.

Sürekli öğrenme yaklaşımları, zaman içinde görevler arası bilgi aktarımını ve unutmayı önleyerek öğrenmeyi sürdürmeyi amaçlayan geniş bir paradigma sunarken, bu yapının altında yer alan daha dar kapsamlı bir öğrenme biçimi olarak artımlı öğrenme (incremental learning) öne çıkmaktadır. Artımlı öğrenme, çoğunlukla tek bir görev veya alan içerisinde, veri örneklerinin zamanla, küçük partiler veya tek tek sisteme sunulması durumunda kullanılan bir yaklaşımdır (Maltoni ve Lomonacco, 2018). Bu yaklaşım, sürekli öğrenmenin temel ilkeleriyle uyumlu olmakla birlikte, görev değişiminden ziyade veri akışı üzerinde güncel modelin artımlı olarak güncellenmesine odaklanmaktadır. Artımlı öğrenme, özellikle büyük veri kümeleriyle çalışan sistemlerde bellek ve hesaplama verimliliği açısından önemli avantajlar sunmaktadır.

2.3. Artımlı Öğrenme

İnsan öğrenmesi incelendiğinde, bu sürecin genellikle kademeli ve artımlı olduğu görülmektedir. İnsanlar, kavramlara ilişkin açıklamaları somut gerçeklerden yola çıkarak öğrenmekte ve zamanla karşılaştıkları yeni bilgiler ve gözlemler doğrultusunda bu açıklamaları aşamalı olarak geliştirmektedirler (van de Ven ve ark, 2022). Edinilen yeni bilgiler, mevcut bilgi yapısını ve zihinsel modelleri zenginleştirmek amacıyla kullanılmaktadır. Bu süreçte birey, ilgili konu hakkında sahip olduğu tüm bilgileri baştan inşa etmek zorunda kalmadan mevcut yapıyı güncelleyebilmektedir. İnsan öğrenmesinin doğası gereği artımlı olmasının iki temel nedeni olduğu kabul edilmektedir. Bunlardan ilki, bilginin zamana yayılan sürekli ve ardışık bir biçimde edinilmesi ve diğeri ise insan zihninin sınırlı bellek kapasitesi ile işlem gücüne sahip olmasıdır (Baddeley ve Hitch, 1974).

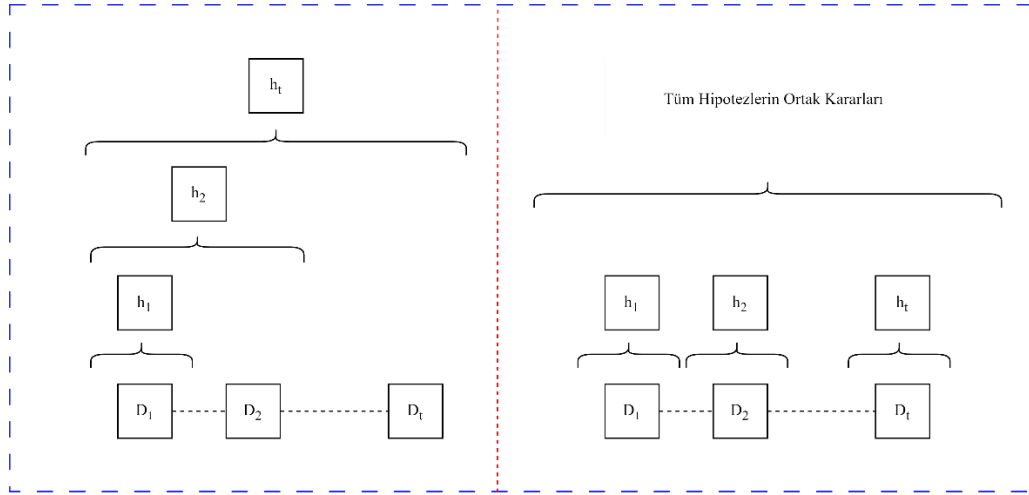
Biyolojik sistemler, yaşamları boyunca sürekli öğrenme yeteneğine sahip oldukları ve zamanla bilgi biriktirebildikleri için, artımlı öğrenme beyin benzeri zekâ sistemleri açısından temel bir yetenek olarak kabul edilmektedir (Pisupati ve Niv, 2022). Geleneksel makine öğrenmesinde temel amaç, geçmiş verilerden öğrenerek genelleme yapmak ve bu genellemeleri karar verme süreçlerinde kullanmaktır. Buna karşılık, artımlı makine öğrenmesinde akıllı sistem, içinde bulunduğu çevreden gelen verileri sürekli ve aşamalı bir biçimde işleyebilmektedir. Bu durum, öğrenmenin yalnızca belirli bir eğitim aşamasında değil, sistemin tüm yaşam döngüsü boyunca devam

etmesini gerektirmektedir (Ade ve Deshmukh, 2013). Dolayısıyla öğrenme modelinin sürekli güncellenebilir olması, artımlı öğrenmenin en belirleyici özelliği olmaktadır.

Makine öğrenmesi yöntemleri, veri setlerinden anlamlı çıkarımlar yapabilmek ve etkili modeller geliştirebilmek için gelişmiş teknolojilere ve sofistike yaklaşımlara ihtiyaç duymaktadır. Ancak geleneksel uygulamaların çoğu, modelin eğitilmesinden önce tüm verinin hazır olduğu ve veri kümesinin zaman içinde değişmediği varsayımına dayanmaktadır. Bu çerçevede, model seçimi genellikle mevcut veri kümesine özgü olarak gerçekleştirilmektedir. Buna karşın, artımlı öğrenme yaklaşımı, verilerin zaman içinde sürekli olarak geldiği bir senaryoyu esas almakta ve modelin bu akışa paralel şekilde sürekli güncellenmesini gerektirmektedir (Parisi ve ark, 2019). Bu öğrenme biçimi, özellikle otonom sürüş sistemleri ve otonom robotlar gibi gerçek zamanlı karar verme gerektiren uygulamalarda kullanılmaktadır (Vapnik, 2000). Öte yandan, bazı veri kümeleri her ne kadar statik olsa da, boyutlarının aşırı büyük olması ya da işlenmesinin uzun zaman alması nedeniyle fiilen bir veri akışı şeklinde ele alınabilmektedir. Bu tür durumlarda da artımlı öğrenme, tüm veri seti ile kademeli bir biçimde uygulanabilmektedir (Schlimmer ve Fisher, 1986).

Artımlı öğrenme, sınırlı bellek ve işlem kaynaklarıyla çalışan sistemlerde kullanılan öğrenme stratejilerini ifade edmektedir. Bu yaklaşımda, makine öğrenmesi modeli yeni veriler geldikçe mevcut bilgisine dayanarak bir hipotezi ya da hipotezler kümesini aşamalı olarak geliştirmektedir (Syavasya ve Muddana, 2020). Temel amacı, veri akışı boyunca modeli sürekli olarak güncellemek ve öğrenme sürecini kesintisiz hâle getirmektir (Junsawang ve ark, 2019). Artımlı makine öğrenmesi, özellikle verinin tümüne aynı anda erişimin mümkün olmadığı durumlarda, mevcut model bilgisini yeni örneklerle zenginleştirerek daha güçlü ve güncel bir temsil oluşturmayı hedeflemektedir.

Bu yaklaşım, şekil 2.1’de gösterildiği gibi bellekte veri depolamadan akan verilerden öğrenerek model oluşturma esasına dayanmaktadır. Buna göre t_i anında gelen D_{t_i} verisi kullanılarak bir hipotez (h_{t_i}) geliştirilmektedir. Her yeni bir veri geldiğinde bu hipotez güncellenir ya da yeni bir hipotez üretilmektedir. Herhangi bir t anına kadar gelen verileri biriktirip doğrudan kullanmak yerine, model güncellemesi verilerin özetlenmiş istatistikleri, model parametrelerinin ortak bir temsili ya da gözlemlenen verilerin sıkıştırılmış biçimdeki temsilleri üzerinden gerçekleştirilmektedir.



Şekil 2.1. Artımlı öğrenme yaklaşımı.

Kavram kayması (concept drift), artımlı öğrenme sürecinde modelin sağlamlığı ve uyarlanabilirliği açısından kritik bir öneme sahip olmaktadır. Zamanla veri dağılımında meydana gelen değişiklikler, modelin performansını olumsuz etkileyebileceğinden, modelin yeni veri yapılarıyla karşılaştığında bu değişikliklere uyum sağlayabilmesi gerekmektedir. Artımlı öğrenme yaklaşımı, yalnızca modeli sürekli olarak güncellemekle kalmaz, aynı zamanda veri kümesinin farklı alanlarından bilgi edinerek veri boyutundaki değişikliklere uyum sağlama yeteneği de sunmaktadır. Bu sayede, veri setinde doğrudan gözlemlenmemiş olsa bile, tüm özellik uzayı hakkında bilgi edinme imkânı doğmaktadır. Artımlı öğrenme ile sistem, yeni veriler geldikçe bilgi tabanını güncelleyebilir ve mevcut yapıya yeni bilgileri sorunsuz biçimde entegre edebilmektedir.

2.3.1. Artımlı öğrenme türleri

Artımlı öğrenme, yalnızca veri miktarındaki artışa değil, aynı zamanda sınıfların, örnek dağılımlarının ve özellik uzayının zamanla değişmesine karşı da duyarlı olmak mecburiyetindedir. Literatürde artımlı öğrenme, örnek artımlı öğrenme, sınıf artımlı öğrenme ve özellik artımlı öğrenme olmak üzere üç temel başlık altında sınıflandırılmaktadır (He ve ark, 2011). Bu ayrım, öğrenme sürecinde zamanla değişen ya da artan bileşenlere göre yapılmaktadır.

2.3.1.1. Örnek artımlı öğrenme (instance incremental learning)

Bu öğrenme yaklaşımında, başlangıç eğitim veri setiyle oluşturulan sınıflandırma modeli, zamanla gelen yeni veri örnekleriyle güncellenmektedir. Ancak yeni veriler, sınıflar arasındaki dengesizlik, kavramsal kayma veya örnek dağılımındaki

değişiklikler gibi nedenlerle mevcut modelin performansını olumsuz etkileyebilmektedir. Bu nedenle, modelin hem eski bilgiyi koruması hem de yeni gelen veri örneklerinden etkili biçimde öğrenmesi gerekmektedir. Örnek artımlı öğrenme, özellikle büyük veri ortamlarında ve verinin sürekli olarak aktığı durumlarda tercih edilmektedir (He ve ark, 2011).

2.3.1.2. Özellik artımlı öğrenme (feature incremental learning)

Özellik artımlı öğrenme, öğrenme sürecinde yeni özniteliklerin modele eklenmesi gerektiği durumlarda uygulanmaktadır. Başlangıç eğitim verisiyle oluşturulan özellik uzayı, zamanla ortaya çıkan yeni özniteliklerle genişletilmesi gerekmektedir. Yeni öznitelikler, sınıflandırma yeteneğini ciddi biçimde artırabilmektedir. Bu nedenle modelin, mevcut özellik uzayını yeniden yapılandırarak daha doğru bir temsil sağlamasını gerektirmektedir (Baysal ve Bayılmış, 2022). Özellikle IoT, kullanıcı profillemesi ve dinamik sensör verileri gibi alanlarda işlevsel bir yaklaşım olmaktadır.

2.3.1.3. Sınıf artımlı öğrenme (class incremental learning)

Sınıf artımlı öğrenmede, başlangıçta sınırlı sayıda sınıfa göre eğitilmiş bir modelin, zamanla karşılaştığı yeni sınıfları tanıyıp bu sınıflara uygun biçimde genişletilmesi amaçlanmaktadır. Yeni sınıflar genellikle daha önceki sınıflardan belirgin biçimde farklıdır ve bu durum sınıflandırma doğruluğunu olumsuz etkileyebilmektedir. Bu nedenle, sınıf artımlı öğrenme algoritmaları, yeni sınıfların tespiti, temsil edilmesi ve mevcut sınıf yapısına uyumlu biçimde entegrasyonu üzerine odaklanmaktadır (Rebuffi ve ark, 2017).

2.4. Yaşam Boyu Öğrenmenin İlgili Öğrenme Paradigmalarıyla Karşılaştırılması

Yaşam boyu öğrenme, yalnızca ardışık görevleri öğrenmekle kalmayıp, aynı zamanda önceki görevlerden edinilen bilgileri kalıcı şekilde bellekte tutarak yeni görevlerin öğrenilmesinde bu bilgileri etkin biçimde kullanmayı hedefleyen bir üst öğrenme paradigmasıdır (Chen ve Liu, 2018). LL, çevresel değişimlere uyum sağlayan ve bilgi birikimini sürdürebilen sistemler geliştirmeyi amaçlarken, diğer birçok makine öğrenmesi paradigması ya tek seferlik bilgi aktarımına ya da sınırlı görev kapsamına odaklanmaktadır.

Bu doğrultuda, yaşam boyu öğrenmenin kapsamını, temel bileşenlerini ve özgün yönlerini daha net biçimde ortaya koymak amacıyla, sıklıkla ilişkilendirilen ya da

kavramsal olarak benzeşen diğer öğrenme yaklaşımlarının karşılaştırmalı olarak ele alınması önem arz etmektedir. Özellikle transfer öğrenme, çok görevli öğrenme, çevrimiçi öğrenme, pekiştirmeli öğrenme ve meta-öğrenme gibi paradigmalara, yaşam boyu öğrenme ile sınırlı benzerlikler taşısalar da, öğrenmenin sürekliliği, edinilen bilginin kalıcılığı, bilgi transferinin sürekliliği ve çevresel değişkenlere uyum kapasitesi gibi temel açılardan önemli ölçüde farklılaştıkları gözlemlenmektedir.

Transfer öğrenme, bir modelin bir kaynak görevde veya etki alanında edindiği bilgileri, benzer veya farklı bir hedef görevde yeniden kullanmasını amaçlayan bir öğrenme yöntemidir. Bu yaklaşım, genellikle etiketli verinin çokça bulunduğu bir kaynak alanda eğitim almış modelin, az etiketli veya etiketsiz veriye sahip bir hedef alana uygulanması şeklinde işlemektedir. Ancak transfer öğrenme süreci çoğunlukla tek yönlü ve tek seferlik olmaktadır. Model, elde ettiği bilgileri gelecekteki öğrenme süreçlerinde yeniden kullanmak üzere bellekte saklamamaktadır. Bu bağlamda transfer öğrenme, bilgiyi yalnızca belirli bir geçiş süreci için kullanmakta ve bilgi birikimi ya da görevler arası kalıcı bellek mekanizması barındırmamaktadır. Buna karşın yaşam boyu öğrenme, önceki görevlerden edinilen bilgilerin sistem belleğinde tutulmasını ve bu bilgilerin gelecekteki görevlerde öğrenmeyi kolaylaştırmak üzere yeniden kullanılmasını hedeflemektedir (Pan ve Yang, 2010; Chen ve Liu, 2018).

Çok Görevli Öğrenme (Multi-Task Learning-MTL), aynı anda birden fazla görev üzerinde öğrenme gerçekleştiren bir yaklaşımdır. Bu paradigmada, görevler arasında bilgi paylaşımı yapılarak her bir görevin genel performansı artırılmaya çalışılmaktadır. Ancak bu yaklaşım, toplu (batch) öğrenme paradigması kapsamında çalışır ve zaman içinde ardışık olarak gelen yeni görevlerle karşılaşıldığında bilgi biriktirme veya yeniden kullanım yeteneği göstermemektedir. Bu nedenle çok görevli öğrenme, sürekli ve evrimsel bir bilgi edinim süreci sunmamaktadır. Buna karşılık yaşam boyu öğrenme, görevler arasında zamana yayılan bir bilgi birikimi sağlayarak, her yeni görevin öğrenilmesinde geçmiş görevlerde edinilen deneyimlerden yararlanmaktadır (Caruana, 1997).

Çevrimiçi Öğrenme (Online Learning-OL), verilerin zaman içinde sıralı biçimde sunulduğu senaryolarda kullanılan bir yöntemdir. Model, her yeni veri noktasını aldığı anda parametrelerini günceller ve böylece eğitimi artımlı şekilde sürdürülmektedir. Ancak bu öğrenme biçimi, genellikle tek bir görev bağlamında çalışır ve sistemin geçmişte öğrendiği bilgileri uzun vadede saklamasını ya da farklı

görevlerde yeniden kullanmasını amaçlamamaktadır. Öte yandan yaşam boyu öğrenme yalnızca artımlı güncellemeyi değil, aynı zamanda geçmiş görevlerde edinilen bilgilerin bellekte kalıcı biçimde tutulmasını ve bu bilginin farklı görevlerde yeniden kullanılmasını da içeren bütüncül bir öğrenme stratejisi sunmaktadır (Bottou, 1998; Chen ve Liu, 2018).

Pekiştirmeli Öğrenme (Reinforcement Learning-RL), bir ajanın bir çevre ile etkileşim kurarak deneme-yanılma yoluyla optimal eylem politikasını öğrenmesini sağlamaktadır. Bu süreçte ajan, ödül ve ceza geri bildirimlerine dayanarak uzun vadeli kazancı maksimize etmeye çalışmaktadır. Pekiştirmeli öğrenme güçlü bir karar verme stratejisi sunmasına rağmen, öğrenme süreci genellikle tek bir görev ve çevre ile sınırlı kalmaktadır. Ajana kazandırılan bilgi, ileride farklı görevlerde kullanılmak üzere bir bellekte saklanmaz ve bu nedenle bilgi birikimi gerçekleştirilmemektedir. Buna karşın yaşam boyu öğrenme, bilgi birikimini teşvik eden ve geçmiş deneyimleri gelecekteki karar süreçlerine entegre edebilen bir yapıya sahip olmaktadır (Sutton ve Barto, 2018).

Meta-öğrenme, bir sistemin az sayıda örnekle yeni bir görevi hızlı şekilde öğrenebilmesini sağlamak amacıyla çok sayıda benzer görevden öğrenilen deneyimlere dayanan bir paradigmadır. Bu yaklaşımda genellikle iki düzeyde öğrenici vardır. Biri hızlı öğrenici (temel öğrenici), diğeri ise yavaş öğrenicidir (meta öğrenici). Meta öğrenici, önceki görevlerden genel öğrenme stratejileri çıkarmakta ve bu stratejileri yeni görevlere uygulayarak hızlı adaptasyonu sağlamaktadır. Ancak meta-öğrenme sistemleri genellikle eğitim ve test görevlerinin aynı dağılımdan geldiğini varsaymaktadır. Yaşam boyu öğrenme ise görevler arası dağılımsal çeşitliliği gözeterik, geçmiş bilgilerin yalnızca uygun olanlarını seçip yeniden kullanabilen bir esneklik sunmaktadır (Finn ve ark, 2017).

Tablo 2.1., yaşam boyu öğrenme paradigması ile transfer öğrenme, çok görevli öğrenme, çevrim.içi öğrenme, pekiştirmeli öğrenme ve meta-öğrenme gibi yaklaşımlar arasındaki temel farkları sunmaktadır. Bu karşılaştırma bilgi birikiminin kalıcılığı, görevler arası transfer yeteneği, öğrenmenin sürekliliği, belleğe dayalı yapıların kullanımı ve görev dağılımına ilişkin varsayımlar gibi ölçütler üzerinden yapılmaktadır. Yapılan karşılaştırma sonucunda, yaşam boyu öğrenmenin bütüncül ve esnek yapısıyla diğer yaklaşımlardan anlamlı biçimde ayrıştığı ortaya konmaktadır.

Tablo 2.1. Yaşam boyu öğrenmenin ilgili öğrenme paradigmalarıyla karşılaştırmalı değerlendirmesi.

Paradigma	Bilgi Birikimi ve Yeniden Kullanım	Görevler Arası Öğrenme Transferi	Sürekli Öğrenme Desteği	Bellek Kullanımı	Görev Dağılımı Varsayımı
Transfer Öğrenme	Bilgi tek seferlik olarak aktarılır, kalıcı birikim yoktur	Kaynak ve hedef görev arasında bilgi transferi yapılır	Süreklilik yoktur, aktarım genellikle tek aşamalıdır	Bellek yapısı kullanılmaz	Genellikle benzer dağılım varsayılır
Çok Görevli Öğrenme	Bilgi eşzamanlı görevler arasında paylaşılır, kalıcılık yoktur	Görevler arası ortak öğrenme yapılır	Öğrenme toplu biçimde gerçekleşir, sürekli değildir	Bellek yapısı bulunmaz	Tüm görevlerin aynı anda öğrenileceği varsayılır
Çevrimiçi Öğrenme	Her yeni veriye göre güncelleme yapılır, bilgi uzun vadeli saklanmaz	Genellikle tek görev üzerine odaklanır	Veri akışıyla birlikte artımlı öğrenme sağlanır	Kalıcı bellek kullanılmaz	Tek görev bağlamı kabul edilir
Pekiştirmeli Öğrenme	Öğrenilen stratejiler çevreye özgüdür, bilgi genellenmez	Genellikle farklı görevler arası transfer yapılmaz	Ajan çevreyle sürekli etkileşim halindedir	Kalıcı bellek mekanizması yoktur	Belirli bir çevreye bağlı tek görev dikkate alınır
Meta-Öğrenme	Önceki görevlerden öğrenilen stratejiler yeni görevlere aktarılır	Görevler arası genelleme yapılır	Görevlere hızlı adaptasyon sağlanır, ancak uzun vadeli birikim sınırlıdır	Sınırlı düzeyde bellek kullanımı olabilir	Eğitim ve test görevlerinin benzer dağılımdan geldiği varsayılır
Yaşam Boyu Öğrenme	Bilgi uzun vadede saklanır ve yeni görevlerde kullanılır	Görevler arası bilgi transferi temel hedeftir	Zamanla gelen görevler üzerinden sürekli öğrenme mümkündür	Kalıcı bellek mekanizması kullanılır	Görevler farklı dağılımlardan gelebilir

3. SINIF DENGESİZLİĞİ VE SINIF ARTIMLI ÖĞRENME

3.1. Sınıf Dengesizliği (Class Imbalance-CI)

Sınıf dengesizliği, makine öğrenimi ve veri analizinde belirli sınıfların diğerlerine kıyasla önemli ölçüde yetersiz temsil edildiği durumları ifade etmektedir. Gerçek hayattaki makine öğrenimi uygulamalarında oldukça yaygın bir problem olan sınıf dengesizliği, sınıflandırma modellerinin performansını olumsuz yönde etkileyerek, çoğunluk sınıfını desteklerken azınlık sınıfını ihmal eden önyargılı tahminlere yol açmaktadır.

Siber güvenlik, sağlık, finans gibi alanlarda bazı azınlık sınıflar nadir gözlemlenmekle birlikte, kritik öneme sahip olabilmektedirler. Geleneksel öğrenme ortamlarında bu durum, modelin nadir fakat önemli kategorileri hatalı sınıflandırması ve çoğunluk sınıflara karşı önyargı geliştirmesiyle sonuçlanabilmektedir (He ve Garcia, 2009). Ancak bu zorluk, sınıfların sıralı olarak tanıtıldığı ve geçmiş veriye erişimin kısıtlı olduğu sınıf artımlı öğrenme senaryolarında daha da karmaşık hale gelmektedir (Buda ve ark, 2018).

Sınıf dengesiz veri kümeleri üzerinde eğitilen modeller, azınlık sınıflarının nüanslarını göz ardı eden önyargılı karar sınırlarını öğrenme eğilimindedirler. Bu genellikle düşük doğruluk, yüksek yanlış negatif oranları ve görülmeyen verilerde azalan genelleme yetenekleriyle kendini göstermektedir (Akintola ve ark, 2022).

Sınıf dengesizliği sorununu gidermeye yönelik literatürde çeşitli yöntemler geliştirilmektedir. Bu kapsamda en çok başvurulanan yaklaşımlardan biri, azınlık sınıfına ait örneklerin sayısını artırmayı hedefleyen aşırı örnekleme ile çoğunluk sınıfına ait örneklerin sayısını azaltmayı amaçlayan yetersiz örnekleme tekniklerinin uygulanmasıdır. Sentetik Azınlık Aşırı Örnekleme Tekniği (Synthetic Minority Over-sampling Technique-SMOTE) ve türevleri, azınlık sınıfına ait sentetik örnekler üreterek veri setindeki dengesizliği azaltmayı hedeflemektedir (Chawla ve ark, 2002). Öte yandan, yetersiz örnekleme yöntemleri, çoğunluk sınıfına ait örneklerin sayısını düşürmek suretiyle sınıf dağılımını dengelemeye çalışmaktadır. Ancak bu durum, bazı değerli bilgilerin kaybına neden olabilmektedir (Seol ve ark., 2023). Nitekim,

örnekleme tekniklerinin özellikle topluluk (ensemble) yöntemleri ile birlikte kullanılması durumunda sınıflandırma performansında anlamlı iyileşmeler sağlanabildiğini ortaya koymaktadır (Akintola ve ark, 2022).

Yapay zeka ve makine öğrenimi, maliyet duyarlı öğrenme yaklaşımları ile sınıf dengesizliğini ele almada ilerleme kaydetmektedir. Bu yaklaşımlar, öğrenme algoritmasını azınlık sınıfı örneklerine daha yüksek yanlış sınıflandırma maliyetleri atayacak şekilde değiştirmekte ve böylece modeli bu durumları doğru bir şekilde tahmin etmeye daha fazla odaklanmaya zorlamaktadır (Firdausanti ve ark, 2024). Odak kaybı gibi değiştirilmiş kayıp işlevlerini uygulama gibi teknikler, modelin çoğunluk sınıfından aşırı etkilenmesini önleyerek öğrenme sürecini önemli ölçüde iyileştirebilmektedir (Goorbergh ve ark, 2022).

Bunlarla birlikte Üretken Çekişmeli Ağlar'a (Generative Adversarial Networks-GAN'lar) dayalı veri artırma stratejilerini kullanmak, azınlık sınıflarından yeni örnekleri sentezlemek için güçlü bir çerçeve sunarak gelişmiş sınıflandırıcı doğruluğuna sağlamaktadır (Zola ve ark, 2022). Son zamanlarda yapılan çalışmalar, bu tür tekniklerin yalnızca sınıf dengesizliği sorununu ele almakla kalmayıp aynı zamanda eğitim veri setine değişkenlik getirerek model sağlamlığını da artırdığını vurgulamaktadır (Rebuffi ve ark, 2021).

Sınıf dengesizliğinin giderilmesine yönelik kayda değer ilerlemeler sağlanmış olsa da, özellikle sürekli öğrenme bağlamında hâlâ çözüm bekleyen bazı temel zorluklar mevcuttur. Bu zorluklardan en önemlilerinden biri, veri akışlarının dinamik yapısından kaynaklanmaktadır. Gerçek dünya veri kümelerinde sıklıkla gözlenen, zaman içinde sınıf dağılımlarında meydana gelen değişimler, literatürde genellikle kavram kayması (concept drift) olarak adlandırılmaktadır (Lu ve ark, 2020). Kavram kayması, modellerin statik varsayımlar altında eğitilmesi durumunda sınıflandırma performansının zamanla bozulmasına neden olabilmektedir. Bu nedenle, değişen sınıf dağılımlarına etkin biçimde uyum sağlayabilecek, uyarlanabilir algoritmaların geliştirilmesi ve modellerin sürekli güncellenmesi büyük önem taşımaktadır (Lyu ve ark, 2024).

Sınıf dengesizliği, verinin bağlamına ve yapısal özelliklerine bağlı olarak farklı şekillerde ortaya çıkabilen, çok yönlü ve karmaşık bir sorun olmaktadır. Bu nedenle, etkili bir şekilde ele alınabilmesi, genelleştirilmiş yaklaşımlardan ziyade, verinin

özgün doğasına uyarlanmış stratejilerin geliştirilmesini gerektirmektedir. Güncel araştırmalar, yalnızca algoritmik iyileştirmelere odaklanmakla kalmayıp, aynı zamanda veri setlerinin altında yatan yapısal karmaşıklıkların anlaşılmasını da içeren bütüncül yaklaşımların önemine dikkat çekmektedir. Bu doğrultuda, sınıf dengesizliğinin olumsuz etkilerini en aza indirmeyi amaçlayan yeni metodolojiler, öğrenme çerçeveleri ve stratejiler geliştirilmeye devam etmektedir.

3.2. Sınıf Artımlı Öğrenme (Class Incremental Learning)

Makine öğrenimi modelleri çoğunlukla sabit veri kümeleri ve görevler üzerinde eğitilmekte olup, bu durum onları dinamik, değişken ve sürekli genişleyen veri ortamlarında yetersiz kılmaktadır. Gerçek dünya uygulamalarında, modellerin zaman içerisinde yeni sınıfları öğrenirken önceki bilgilere dair performanslarını sürdürebilmesi büyük önem taşımaktadır. Bu bağlamda Sınıf Artımlı Öğrenme, modelin sürekli olarak yeni sınıfları öğrenmesini sağlarken daha önce öğrendiği bilgileri de unutmamasını amaçlayan bir sürekli öğrenme paradigmasıdır (Delange ve ark, 2021).

Sürekli veri akışının söz konusu olduğu dinamik ortamlar için CIL kritik öneme sahip olmaktadır. CIL'in temel amacı, daha önce öğrenilmiş sınıflardaki performansı koruyarak yeni sınıfları mevcut öğrenme yapısına entegre edebilmektir. Ancak bu süreç çoğunlukla katastrofik unutma (catastrophic forgetting) olarak adlandırılan modelin önceki öğrenmeleri unutma eğilimine girmesi ve veri dağılımının zamanla değişmesinin neden olduğu sınıf dengesizliği gibi sorunlarla daha da karmaşık bir hale gelmektedir (Zhou ve ark, 2023; van de Ven ve ark, 2022).

3.2.1. Sınıf artımlı öğrenmede katastrofik unutma ve sınıf dengesizliği

CIL senaryolarında karşılaşılan başlıca zorluklardan biri, modelin yeni sınıflara odaklandıkça önceki bilgileri kaybetmesine yol açan katastrofik unutma olgusudur (Kirkpatrick ve ark, 2017). Bu sorun, özellikle yeni sınıfların veri setinde düşük temsil oranına sahip olduğu, yani sınıf dengesizliği durumlarında daha da şiddetlenmektedir (Wu ve ark, 2019; Hou ve ark, 2019). Modelin sık karşılaşılan sınıflara önyargı göstermesi, azınlık sınıfların yeterince temsil edilememesi sonucunu doğurmaktadır.

3.2.2. Çözüm stratejileri

Sınıf artımlı öğrenmede karşılaşılan katastrofik unutma ve sınıf dengesizliği gibi temel sorunlara yönelik çözüm stratejileri; bellek tekrarına dayalı yaklaşımlar, düzenlileştirmeye dayalı yöntemler ve mimariye müdahale eden yapısal çözümler olmak üzere üç ana grupta ele alınmaktadır (Delange ve ark, 2021; Lesort ve ark, 2021).

3.2.2.1. Bellek tekrarı (replay-based methods)

Bellek tekrar yöntemleri, önceki sınıflara ait bazı örneklerin modelin belleğinde tutulmasını ve bu örneklerin yeni sınıfların eğitimi sırasında yeniden kullanılmasını esas almaktadır. En yaygın teknikler arasında Experience Replay (deneyim tekrarı) ve Generative Replay (üretici modeller aracılığıyla sentetik örneklerin üretilmesi) yer almaktadır.

- Experience Replay, önceki sınıflardan rastgele seçilen gerçek örneklerin belirli bir tampon bellekte saklanmakta ve modelin eğitimi sırasında bu örneklerin yeniden kullanılmaktadır (Rebuffi ve ark, 2017; Castro ve ark, 2018).
- Generative Replay, geçmiş sınıflara ait sentetik örneklerin üretici modeller kullanılarak yeniden üretilmesini içermektedir (Shin ve ark, 2017). Bu sayede geçmiş verilerin fiziksel olarak saklanmasına gerek kalmadan tekrar yapılabilir.

3.2.2.2. Düzenlileştirmeye dayalı yöntemler (regularization-based)

Düzenleme temelli yöntemlerin asıl hedefi, modelin parametrelerini korumaya yönelik ceza terimleri (regularization loss) ekleyerek unutmayı azaltmaktır. Özellikle EWC yöntemi, geçmiş görevlerde önemli olan parametrelerin yeni öğrenme süreçlerinde değişmesini kısıtlayarak bilgi kaybını azaltmaktadır (Kirkpatrick ve ark, 2017).

Benzer şekilde, Li ve Hoiem (2017) Learning Without Forgetting (LwF) yaklaşımı da önceki çıktıları hedefleyen bir kayıp fonksiyonu üzerinden modelin çıktılarının korunmasını amaçlamaktadırlar.

3.2.2.3. Yapı temelli yöntemler (architecture-based methods)

Yapı temelli stratejiler, modelin mimarisinin yeni bilgilerle birlikte esnek şekilde genişletilmesini sağlamaktadır. Bu sayede önceki bilgiler korunurken yeni sınıfların öğrenimi için uygun bir alan oluşturulmaktadır. Özellikle Dinamik Omurga Mimarileri

(Dynamic Backbone Architectures), her yeni sınıf için yeni bir yapı modülü veya alt ağ tahsis ederek hem önceki hem de yeni bilgilerin paralel olarak temsil edilmesini mümkün kılmaktadır (Douillard ve ark, 2022; Yan ve ark, 2021).

3.2.3. Sınıf artımlı öğrenmeyi destekleyen algoritmalar

Sınıf artımlı öğrenme, modelin zamanla yeni sınıfları öğrenmesini ve önceki bilgileri unutmadan bu süreci sürdürebilmesini gerektirmektedir. Bu çerçevede, bazı geleneksel makine öğrenmesi sınıflandırma algoritmaları, doğal yapıları gereği artımlı öğrenme gerçekleştirebilmektedirler. Bu algoritmalar, verinin zamanla artarak veya değişerek gelmesine rağmen, performans kaybı yaşamadan öğrenme sürecini sürdürebilmektedir. Sınıf artımlı öğrenmede sıklıkla tercih edilen beş temel algoritma bulunmaktadır.

- Naive Bayes: Naive Bayes algoritması, olasılık temelli bir sınıflandırma yaklaşımı olup, Bayes teoremi aracılığıyla koşullu olasılıkların hesaplanmasına dayanmaktadır. Artımlı öğrenme bağlamında öne çıkan en önemli özelliği, yeni verilerin istatistiksel olarak mevcut öğrenme sürecine dâhil edilmesine olanak tanımasıdır. Bu yapı sayesinde, her yeni veri örneği sisteme eklendiğinde, modelin parametreleri önceki olasılıkların yeni gözlemlerle güncellenmesi yoluyla yeniden ayarlanmakta ve böylece öğrenme süreci devam ettirilebilmektedir. Bu durum, modelin geçmiş bilgileri koruyarak yeni sınıfları öğrenmesini ve bilgi bütünlüğünü muhafaza etmesini mümkün kılmaktadır (Domingos ve Hulten, 2000).
- Destek Vektör Makineleri (SVM): Veri örneklerini yüksek boyutlu bir uzayda en iyi ayıran karar sınırlarını, yani hiper düzlemleri, öğrenmeye çalışan bir sınıflandırma algoritmasıdır. Modelin karar mekanizması, yalnızca sınıfları birbirinden ayırmada kritik rol oynayan veri noktaları olan destek vektörlerine dayanmaktadır. Bu özellik, SVM algoritmasının artımlı öğrenmeyi doğal olarak desteklemesini sağlamaktadır. Çünkü yeni gelen verilerle elde edilen destek vektörleri, mevcut destek vektörleriyle birlikte değerlendirilerek karar sınırı güncellenmekte ve böylece model, geçmiş bilgilere erişmeden güncellenebilir hale gelmektedir (Cauwenberghs ve Poggio, 2000).
- Karar Ağaçları: Karar ağacı algoritmaları, örnekleri özelliklerine göre dallandırarak hiyerarşik yapılar kuran ve sınıflandırma işlemini yaprak düğümler üzerinden gerçekleştiren yöntemlerdir. Geleneksel karar ağaçları genellikle tüm veri kümesi mevcutken eğitilirken, artımlı öğrenmeye uygun sürümleri, yalnızca

yeni gelen örnekleri dikkate alarak mevcut ağaç yapısını güncelleyebilmektedir. Bu sayede, sistem belleğinde önceki verilere yeniden ihtiyaç duymaksızın karar ağacı evrim geçirebilmekte ve yeni sınıfların öğrenilmesi mümkün hâle gelmektedir (Utgoff, 1989; Schlimmer ve Fisher, 1986).

- Rastgele Ormanlar: Rastgele Orman algoritması, birden fazla karar ağacının toplu öğrenmesini ifade eden bir topluluk (ensemble) yöntemidir. Her bir ağaç, eğitim verisinin rastgele alt kümeleri üzerinden bağımsız olarak oluşturulmakta ve sınıflandırma işlemi tüm ağaçların oy çokluğu prensibine göre yapılmaktadır. Bu yapısı sayesinde, artımlı öğrenme bağlamında yeni sınıfların öğrenilmesi için yeni karar ağaçlarının eğitilip mevcut ormana entegre edilmesi yeterli olmakta; böylece önceki bilgiler korunarak modelin kapasitesi artırılabilir (Rahman ve Islam, 2022).
- Yapay Sinir Ağları: Yapay Sinir Ağları (YSA), çok katmanlı ve bağlantılı nöron yapıları sayesinde karmaşık örüntüleri öğrenebilen doğrusal olmayan modellerdir. Bu yapılar, her yeni veriye göre ağırlıkların sürekli olarak güncellenmesi yoluyla öğrenmeyi sürdürmektedir. Bu nedenle, artımlı öğrenme bağlamında YSA'lar, yeni gelen veri örneklerini mevcut ağırlıklar üzerinden değerlendirmekte ve ağı yeniden yapılandırarak öğrenmeye devam edebilmektedir. Ancak YSA'lar, artımlı öğrenme sürecinde “katastrofik unutma” sorunuyla karşı karşıya kalmaktadır. Bu nedenle, genellikle bellek tekrar, düzenlileştirme ve modüler yapı gibi ek stratejilerle desteklenmeleri gerekmektedir (Goodfellow ve ark, 2015).

Bu tez çalışması, sınıf artımlı öğrenmenin (CIL) çeşitli senaryolarda karşılaştığı temel zorluklara odaklanmaktadır. Sürekli veri akışı içeren uygulamalarda geçmiş veriye erişimin sınırlı olduğu, yeni sınıfların zamanla ortaya çıktığı ve sınıf dengesizliğinin yüksek olduğu durumlar ele alınarak, bu problemlere yönelik iki özgün model önerilmektedir.

Birinci model, EWC yöntemini modüler çıktı katmanlarıyla bütünleştirerek, düzenlileştirmeye dayalı bir bilgi koruma mekanizması sunmaktadır. Bu yapı yalnızca önceki görevlerde öğrenilen bilgileri muhafaza etmekle kalmayıp, aynı zamanda yeni gelen sınıfların etkili biçimde öğrenilmesini de mümkün kılmaktadır. Böylece hem sınıf dengesizliğine karşı direnç sağlamakta hem de görevler arası bilgi sürekliliğini korumaktadır.

İkinci model ise, Hoeffding Tree temelli dinamik çıktı genişletme (DOE) mekanizması ile gerçek zamanlı veri akışında yeni sınıfları otomatik olarak tanıyarak çıktı katmanını yeniden yapılandırmaktadır. Böylece esnek ve ölçeklenebilir bir öğrenme çerçevesi sunmaktadır. Bu yapı, özellikle siber güvenlik gibi kritik alanlarda pratik avantajlar sağlamaktadır.

Bu tez çalışmasında, yapılandırılmış veriler üzerinde bellek erişimi ve üretici tabanlı tekrar stratejilerinin uygulanmasında karşılaşılan pratik zorluklar nedeniyle tekrar temelli yöntemlere yer verilmemiştir. Bunun yerine, bu yöntemlerin sunduğu avantajlar dikkate alınarak, tekrar mekanizmasına ihtiyaç duymadan tatmin edici performans sergileyebilen modeller geliştirilmesi amaçlanmaktadır.

Bu bağlamda önerilen her iki model, tablo 3.1.'de görüldüğü üzere literatürdeki mevcut sınıf artımlı öğrenme yaklaşımları ile karşılaştırmalı olarak değerlendirildiğinde, hem yapısal esneklik hem de uygulama bağlamında önemli avantajlar sunmaktadır. Özellikle CIL-DOE (Class-Incremental Learning with Dynamic Output Expansion) modeli, dinamik çıktı katmanı genişletme yeteneği, gerçek zamanlı çalışabilirliği ve Apache Kafka ile sağladığı tam uyum sayesinde literatürdeki sabit yapılı yöntemlerden ayrılmaktadır. Bu yönüyle, yeniden eğitime gerek duymadan yeni sınıflara adaptasyon sağlayabilmesi, modeli dağıtık büyük veri ortamlarında uygulanabilir kılmaktadır. Ayrıca belleğe dayalı olmayan yapısı sayesinde, kaynak kısıtlı sistemlerde de sürdürülebilir bir öğrenme süreci sunmaktadır.

Tablo 3.1. Önerilen modellerin ve literatürdeki yaklaşımların karşılaştırması.

Yöntem	Katastrofik Sınıf Unutma	Sınıf Dengesizliği	Çıktı Katmanı Uyarlanabilirliği	Bellek Kullanımı	Gerçek Zamanlılık	Dağıtık Sistem Uyumluluğu	Yöntem Türü
Experience Replay	Saklama ile azaltılmış	Dengesizlik hâlâ mevcut	Ön tanımlı sınıflar	Bellek gerektirir	Gerçek zamanlı değil	Offline senaryolarda kullanılır	Bellek tabanlı
Generative Replay	Sentetik tekrar ile çözüm	Azınlık temsilde zayıf	Sabit çıktı katmanı	Üretici model kullanımı	Gerçek zamanlı zayıf	Dağıtık senaryo yok	Bellek + Üretici model tabanlı
Dynamic Architectures	Modül genişletme ile çözüm	Spesifik strateji yok	Mimari düzeyde genişleme	Genellikle hafıza dostu	Gerçek zamanlılığı değişken	Genellikle tekil senaryolar	Mimari tabanlı
EWC Destekli Ortak Encoder	EWC ile azaltılmış	Ağırlıklandırma ve dengeli kayıpla ele alınır	(Statik yapı, görev bazlı çıkış katmanı)	Harici belleğe ihtiyaç olabilir	Genellikle çevrimdışı	Kafka ile test edilmedi	Düzenleme + Mimari tabanlı
CIL-DOE	Parametre tutarlılığı korunur	Azınlık sınıflarda metrik farkı	Dinamik olarak genişleyebilir	Belleğe ihtiyaç duymaz	Gerçek zamanlı (Kafka destekli)	Apache Kafka ile entegre	Mimari + Akış tabanlı

EWC destekli ortak kodlayıcı mimarisi ise özellikle katastrofik unutma ve sınıf dengesizliği problemlerine karşı etkili bir çözüm sağlamaktadır. Ağırlıklandırma temelli dengeleme stratejileri ve dengeli kayıp fonksiyonları ile desteklenen bu yapı, geçmiş görevlerden öğrenilen bilgilerin korunmasına olanak tanımakta ve azınlık sınıfların temsil kabiliyetini artırmaktadır.

Tablo 3.2., bu tez kapsamında önerilen iki yöntem ile literatürde yer alan yaygın sınıf artımlı öğrenme yaklaşımlarını, yapısal ve işlevsel yönleriyle karşılaştırmalı olarak sunmaktadır. Söz konusu karşılaştırma, yöntemlerin katastrofik unutma, yeni sınıflara uyum sağlama, çıktı katmanı esnekliği, bellek kullanımı, gerçek zamanlı veri akışına uygunluk ve sistem düzeyinde uygulanabilirlik gibi temel ölçütler bakımından değerlendirilmesini kapsamaktadır.

Tablo 3.2. Benzer çalışmalar ile önerilen yöntemlerin karşılaştırması.

Yöntem	Model Türü	Katastrofik Unutma	Yeni Sınıflara Uyum	Çıktı Katmanı Genişletme	Bellek Kullanımı	Gerçek Zamanlılık	Özgün Katkı
Rebuffi ve ark, (2017) – iCaRL	Replay + Nearest Class Mean	×	✓	Sabit	Gerektirir	Çevrimdışı	Bellek gerektirir, çıktı katmanı statik
Hou ve ark, (2019) – LUCIR	Cosine Similarity + Regularization	✓	✓	×	×	×	Çıkış genişletmez, derin ağlarla çalışır
Douillard ve ark, (2022) – DyTox	Transformer + Dynamic Architecture	✓	✓	Modüler	×	×	Yapı genişletilebilir, ancak ağır ve çevrimdışı
Yan ve ark, (2021) – DER	Replay + Regularization	✓	✓	×	✓	×	Bellek destekli klasik strateji
Shin ve ark, (2017) – Generative Replay (GR)	Üretici Ağlar	✓	✓	×	Fiziksel değil	×	GAN tabanlı, üretim hatasına duyarlı
Batra ve Clark (2024) – Elastic Variational Continual Learning	Derin ağ	✓	✓	×	Orta	×	Varyasyonel posteri tahmin ve EWC ile hibrit model
Skiers ve Deja (2024)-Joint Diffusion Continual Learning	Üretken tekrar oynatma	✓	✓	×	Düşük	×	Üretken tekrar ve yayılma modellerle entegrasyon
Elsayed ve Mahmood (2023)-Utility-based Perturbed Gradient Descent	Derin ağ	✓	✓	×	Minimal	×	Plastisiteyi arttırma, unutmayı engelleme

Tablo 3.2.(Devamı) Benzer çalışmalar ile önerilen yöntemlerin karşılaştırması.

CIL-DOE	Karar Ağacı + $\checkmark$ Dinamik Çıktı	Otomatik	Dinamik	Gerektirmez	Kafka ile uyumlu	Gerçek zamanlı, yeniden eğitim gerektirmeyen, ilklerinden biri
EWC + Ortak Encoder	Derin Ağ + $\checkmark$ Regularization	Görev Bazlı	Statik	Ağırlıklı	Çevrimdışı	Katastrofik unutmayı EWC ile azaltan ve sınıf dengesizliğine odaklı

Bu tez çalışmasında önerilen CIL-DOE modeli, Hoeffding Tree ve dinamik çıktı genişletme mekanizmasıyla, önceden tanımlı sınıflara ihtiyaç duymadan yeni sınıflara gerçek zamanlı uyum sağlayabilmektedir. Bu model Apache Kafka ile entegre çalışabilmekte, yeniden eğitim ve bellek gerektirmemektedir. Bu yönüyle büyük veri ve siber güvenlik gibi dinamik alanlarda yüksek uygulanabilirliğe sahip olmaktadır. EWC destekli ortak kodlayıcı mimarisi ise, çevrimdışı ve statik yapısına rağmen, katastrofik unutmayı azaltması ve sınıf dengesizliğine duyarlı dengeli kayıplar kullanmasıyla, düzenlileştirmeye dayalı yaklaşımlar arasında öne çıkan güçlü bir alternatif olmaktadır.



4. ÖNERİLEN EWC DESTEKLİ ORTAK KODLAYICI VE CIL-DOE MİMARİLERİ

Bu tez çalışmasında, Sınıf Artımlı Öğrenme problemine yönelik iki özgün yaklaşım geliştirilmelidir. Çalışmanın ilk modeli, görev ayrımlarının önceden bilindiği çevrimdışı öğrenme senaryolarında, katastrofik unutma sorununu azaltmak ve sınıf dengesizliği altında bilgi sürekliliğini korumak amacıyla tasarlanmaktadır. Bu doğrultuda, EWC yöntemi, ortak bir kodlayıcı mimari ile entegre edilerek, görevler arası parametre stabilitesini düzenlemeye dayalı bir çerçevede sağlamaktadır. Modelin modüler çıktı katmanları aracılığıyla yeni sınıflar etkili biçimde öğrenilmekte ve daha önce edinilen bilgiler korunmaktadır.

İkinci model, sınıf artımlı öğrenmenin gerçek zamanlı ve akış tabanlı uygulama gereksinimlerine yönelik olarak tasarlanmaktadır. Özellikle sınıfların önceden tanımlanmadığı ve zamanla ortaya çıktığı senaryolarda etkinlik gösterecek şekilde yapılandırılan bu yaklaşımda, karar ağaçları ailesinden gelen Hoeffding Tree algoritması temel alınmakta ve buna entegre edilen dinamik çıktı genişletme mekanizması geliştirilmektedir. Önerilen yöntem, veri akışı sırasında yeni sınıf etiketlerini otomatik olarak tanımlayarak, sınıflandırıcının çıktı katmanını eşzamanlı biçimde yeniden yapılandırmakta ve bu sayede yapısal düzeyde esneklik ve ölçeklenebilirlik sağlamaktadır. Özellikle sınıf sayısının önceden belirli olmadığı büyük ölçekli IoT veri kümelerinde (örneğin CIIoT2023), bu mekanizma, öğrenme sürecine dinamik uyum kabiliyeti kazandırmakta ve sınıflandırma doğruluğunu sürdürülebilir kılmaktadır.

Tüm deneysel çalışmalar, makine öğrenmesi modellerinin geliştirilmesi ve değerlendirilmesi amacıyla bulut tabanlı bir platform olan Google Colab Pro üzerinde yürütülmektedir. Bu platform, GPU ve yüksek RAM gibi gelişmiş donanım seçeneklerine erişim imkânı sunarak, büyük veri kümeleri ve karmaşık modelleme süreçleri için uygun bir çalışma ortamı sağlamaktadır. Her iki deney de aynı platformda gerçekleştirilmiş olmakla birlikte, kullanılan donanım kaynakları açısından farklılık göstermektedir.

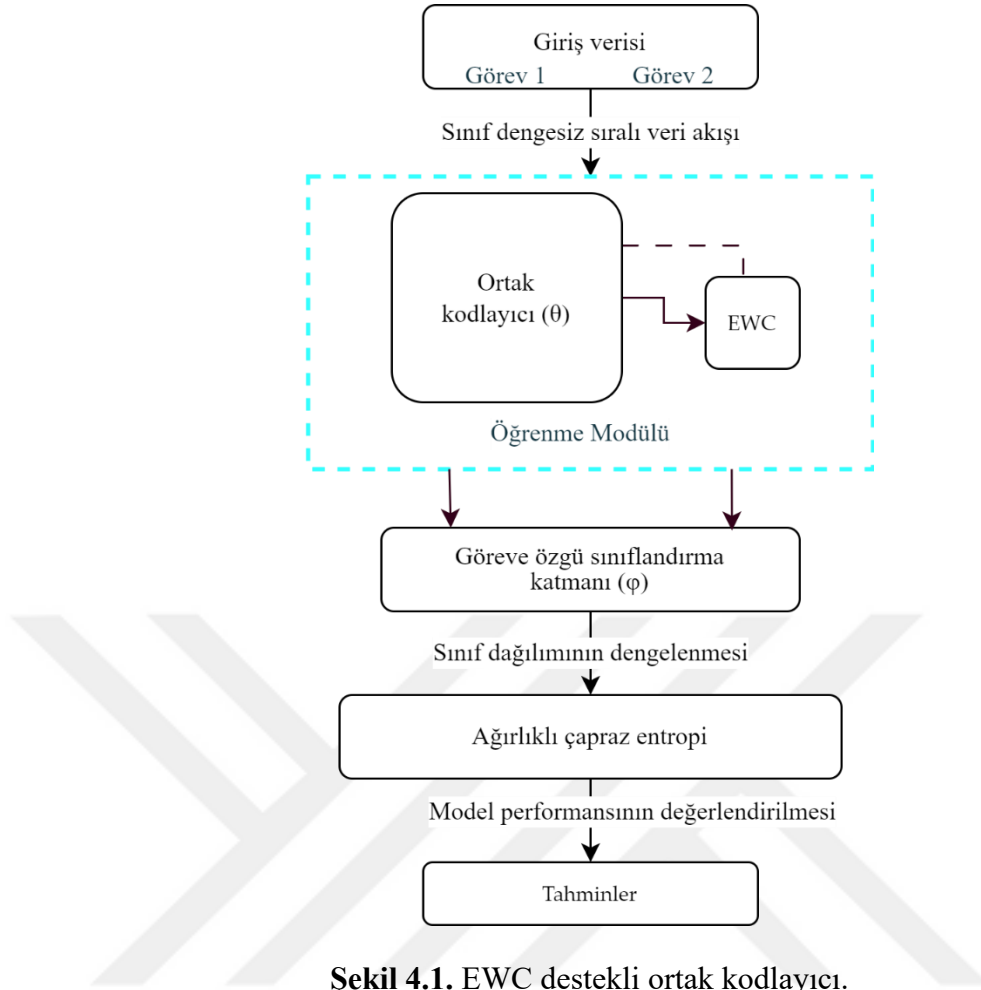
İlk model olan EWC destekli ortak kodlayıcı mimarisi, yüksek hesaplama gücü ve bellek kapasitesi gerektirdiğinden, deneyler Google Colab Pro platformunda NVIDIA A100 Tensor Core GPU ve yüksek bellek (High-RAM) seçeneği ile gerçekleştirilmektedir. Bu donanım seçeneği, modelin derin öğrenme tabanlı yapısını verimli şekilde eğitmeye ve hiperparametre ayarlamalarını hızlıca uygulamaya olanak sağlamaktadır. Böylece modelin hem parametre konsolidasyonu hem de farklı sınıf yapılarında performansını karşılaştırmak üzere çoklu deney senaryoları kısa sürede tamamlanabilmektedir. Buna karşılık, ikinci model olan Hoeffding Tree tabanlı CIL-DOE yaklaşımı, çevrim içi ve hafif olması nedeniyle GPU'ya ihtiyaç olmadan yalnızca CPU ve standart bellek konfigürasyonu kullanılarak aynı platform üzerinde uygulanmaktadır.

Her iki yaklaşım, sınıf artımlı öğrenmenin temel zorluklarından olan katastrofik unutma ve sınıf dengesizliği problemlerine farklı yönlerden çözüm getirmeyi amaçlamaktadır. Düzenlileştirmeye dayalı parametre koruma ve yapısal çıktı genişletme stratejileri ile sürekli öğrenme süreçlerinde bilgi sürekliliği ve model uyarlanabilirliği sağlamayı hedeflemektedir.

4.1. EWC Destekli Ortak Kodlayıcı Mimari

Bu tez çalışmasında önerilen ilk model, sınıf artımlı öğrenme bağlamında katastrofik unutmayı azaltmak ve sınıf dengesizliğiyle başa çıkmak amacıyla, ortak bir kodlayıcı mimarisini EWC yöntemi ile birleştiren düzenlileştirmeye dayalı bir yaklaşımı temel almaktadır.

Şekil 4.1, önerilen yöntem olan EWC destekli ortak kodlayıcı yaklaşımını sınıf dengesizliği içeren sınıf artımlı öğrenme bağlamında kavramsal olarak özetlemektedir. Giriş verisi, ardışık olarak gelen ve sınıflar arasında dengesizlik içeren görevler (Görev 1 ve Görev 2) şeklinde modele sunulmaktadır.



Şekil 4.1. EWC destekli ortak kodlayıcı.

Öğrenme modülü, parametreleri θ ile gösterilen bir ortak kodlayıcı (shared encoder) ve önceden öğrenilen bilginin korunmasını sağlamak amacıyla EWC mekanizmasından oluşmaktadır. Bu modül, görevler arası bilgi paylaşımını sağlarken, önceki görevlerde öğrenilen bilgilerin unutulmasını en aza indirmeyi hedeflemektedir. Kodlayıcıdan elde edilen çıktılar, her görev için ayrı ayrı tanımlanan parametreleri ϕ ile gösterilen göreve özgü sınıflandırıcı başlıklara (task-specific classification head) iletilmektedir.

Bu başlıklar, sınıf dengesizliğini ters frekans tabanlı yeniden ağırlıklandırma stratejisi ile dengelemekte ve sınıflandırma kaybı, ağırlıklı çapraz entropi kaybı (weighted cross-entropy loss) ile hesaplanmaktadır. Model performansı ise doğruluk (accuracy), makro F1 skoru, bilgi koruma (retention) ve uyarlanabilirlik (adaptability) gibi çeşitli ölçütler üzerinden değerlendirilmektedir. Bu yapı, hem sürekli öğrenme sürecinde bilgi kaybını önlemeyi hem de sınıf dengesizliğinin olumsuz etkilerini azaltmayı amaçlamaktadır.

4.1.1. Model mimarisi

Önerilen model mimarisi, sınıf artımlı öğrenme bağlamında esnek ve modüler bir yapı sunarak bilgi korunumu ve yeni görevlerin öğrenimi arasındaki dengeyi sağlamayı amaçlamaktadır. Bu yapı üç temel bileşenden oluşmaktadır:

- **Girdi Katmanı:** Her bir görev için gelen verilerin, modelin kodlayıcı bileşeni tarafından işlenebilecek uygun formata getirilmesini sağlamaktadır. Bu katman, görev özelindeki veri yapılarının standart hale getirilmesi sürecini yürütmektedir.
- **Ortak Kodlayıcı (f_θ):** Ortak kodlayıcı, θ parametresiyle yapılandırılmıştır ve farklı görevler arasında göreve bağımlı olmayan (task-agnostic) özellik temsillerini öğrenmekle sorumludur. Bu katman, görevden bağımsız olarak genel özellikleri çıkartmakta ve böylece önceki görevlerde öğrenilen temsillerin yeni görevlere aktarımını mümkün kılmaktadır. Bu yapı, görev özelinde model çoğaltımına gerek kalmadan belleğin verimli kullanılmasını sağlar ve öğrenilen temsillerin yeniden kullanılmasına olanak tanımaktadır.
- **Göreve Özel Çıkış Katmanları (h_{ϕ_t}):** Her bir görev t için ayrı bir sınıflayıcı başlığı h_{ϕ_t} tanımlanmaktadır. Bu başlıklar, kodlayıcıdan gelen ortak temsilleri görev sınıflarına eşleyen, hafif ve göreve özel çıktılar üreten katmanlardır. Yeni bir görev tanıtıldığında, modelin önceki görevleri etkilenmeden öğrenmeye devam edebilmesi için yalnızca yeni bir çıktı katmanı eklenmektedir. Bu yapı sayesinde, kodlayıcı sabit tutulur ve bilgi kaybı en aza indirilmektedir.

Bu mimari çerçevede, herhangi bir x girdisi için görev t 'ye ait model çıktısı y_t aşağıdaki şekilde hesaplanmaktadır:

$$y_t = h_{\phi_t}(f_\theta(x)) \quad (4.1)$$

Burada, $f_\theta(x)$ ifadesi, ortak kodlayıcı tarafından üretilen özellik vektörünü, $h_{\phi_t}(f_\theta(x))$, ise bu temsili görev t 'ye özgü sınıf tahminlerine dönüştüren sınıflayıcı başlığı ifade etmektedir.

4.1.2. Bilginin saklanması için EWC

EWC, ardışık görevler arasında öğrenilen bilgilerin korunmasını amaçlayan bir düzenleştirme yaklaşımıdır. Katastrofik unutmayı azaltmak için, model parametrelerinin geçmiş görevlerde öğrenilen değerlerinden fazla sapmasını engelleyen bir ceza terimini kayıp fonksiyonuna eklemektedir. Bu terim, her bir parametrenin önceki görevler için ne kadar önemli olduğunu gösteren fisher bilgi matrisi (Fisher Information Matrix - FIM) ile ağırlıklandırılmaktadır. Bu sayede, modelin önceki görevlerde kritik olan parametreler yeni görevler öğrenilirken korunmaktadır (Kirkpatrick ve ark, 2017).

t görevi için, kodlayıcıya ait θ parametresine uygulanan EWC ceza terimi (penalty), aşağıdaki şekilde hesaplanmaktadır:

$$L_{EWC} = \frac{\lambda}{2} \sum_i F_i (\theta_i - \theta_i^*)^2 \quad (4.2)$$

Burada, λ terimi, EWC cezasının etkisini belirleyen bir hiperparametredir ve yeni görevleri öğrenme ile önceki görevlerde edinilen bilgiyi koruma arasındaki dengeyi ayarlamak için kullanılmaktadır. θ_i^* , i'nin önceki görevler için öğrenilmiş en uygun değerini temsil ederken, F_i ise ilgili parametrenin önceki görevler açısından önemini yansıtan fisher bilgi değerini ifade etmektedir. Bu bağlamda, fisher bilgi matrisi, her bir parametrenin önceki görevlerde model performansı üzerindeki katkısını niceliksel olarak değerlendirerek, öğrenme sürecine göreceli bir katsayısı sunmaktadır.

FIM, model parametrelerinin gözlemlenen verilerin olasılığı üzerindeki etkisini değerlendirerek, her bir parametrenin öğrenme süreci açısından önemini belirlemektedir. Ancak FIM'in tam olarak hesaplanması, özellikle yüksek boyutlu derin öğrenme modellerinde yüksek maliyetli ve pratikte uygulanması zor bir işlem olmaktadır. Bu nedenle, yaygın olarak çapraz (diagonal) bir yaklaşım tercih edilmektedir. Bu basitleştirilmiş yaklaşımda, FIM'in yalnızca köşegen elemanları dikkate alınarak yaklaşık bir çözüm elde edilmektedir (Kirkpatrick ve ark, 2017). Şekilde 4.2'de sunulan sözde kod, bu yaklaşım doğrultusunda fisher bilgi matrisinin yaklaşık olarak nasıl hesaplandığını göstermektedir.

Adım	Açıklama
1	Model parametrelerini başlat: $\theta \leftarrow \theta_0$
2	Eğitim verisini sırayla dolaş
3	Her örnek (x, y) için:
4	Tahmini hesapla: $\hat{y} = f(x; \theta)$
5	Kayıp fonksiyonunu hesapla: $L(y, \hat{y})$
6	Kayıp fonksiyonuna göre gradyanı hesapla: $g = \nabla_{\theta} L$
7	Fisher katkısı: $F \leftarrow F + g^2$
8	Fisher matrisini ortalama: $F \leftarrow \frac{F}{N}$, N : örnek sayısı)

Şekil 4.2. Diyagonal fisher yakınsaması.

EWC cezası terimi, modeli kritik parametreleri korumaya teşvik etmekte, böylece yenilerini öğrenirken önceki görevlerden gelen bilgilere müdahaleyi en aza indirmektedir.

4.1.3. Eğitim Süreci

4.1.3.1. İlk Görev Eğitimi

İlk görev için model, önceki görevlerden korunması gereken herhangi bir bilgi bulunmadığından yalnızca görev kaybı (L) kullanılarak eğitilmektedir. Görev kaybı, çapraz entropi kaybı olarak hesaplanmaktadır:

$$L_{görev} = \sum_j y_j \log(\hat{y}_j) \quad (4.3)$$

Burada, y_j ve $\hat{y}_j$ gerçek ve tahmin edilen etiketleri temsil etmektedir. İlk görev eğitimi tamamlandıktan sonra, her bir parametre için fisher bilgi matrisi (F) hesaplanır ve bu görevde elde edilen en uygun parametre değerleri θ^* sonraki aşamalarda kullanılmak üzere EWC ceza terimi için saklanmaktadır.

4.1.3.2. EWC ceza terimi ile yeni görev eğitimi

Yeni bir görev tanıtıldığında, modele yeni bir görev-bağımlı çıkış katmanı $h_{\phi_{t+1}}$ eklenmektedir. Daha sonra model, yeni görev verisi ile eğitilmektedir. Bu aşamada kullanılan toplam kayıp fonksiyonu, görev kaybı L_{task} ile EWC ceza terimini L_{EWC} birleştirmektedir.

$$L_{toplam} = L_{görev} + L_{EWC} \quad (4.4)$$

Burada ki toplam kayıp, modelin yeni görevi öğrenirken önceki görevlerde kritik olan bilgileri de korumasını sağlamaktadır. Eğitim sırasında yalnızca o göreve ait çıkış başlığı optimize edilirken, EWC ceza terimi ortak kodlayıcıya ait parametreler üzerinde uygulanmaktadır.

4.1.3.3. EWC parametrelerinin güncellenmesi

Her görev öğrenildikten sonra, kodlayıcının parametreleri θ güncellenmektedir. Sonrasında yeni fisher bilgi matrisi (F) ve en uygun parametre değerleri θ^* yeniden hesaplanarak, gelecekteki görevlerde kullanılmak üzere kaydedilmektedir. Bu sayede modelin görevler arasında edinilen bilgileri kademeli olarak koruması sağlanmaktadır.

4.1.3.4. EWC ceza terimi olmadan yeni görev eğitimi

EWC mekanizmasının etkisini değerlendirebilmek amacıyla, yeni görevler için EWC ceza teriminin uygulanmadığı ayrı bir eğitim adımı gerçekleştirilmektedir. Bu adım, EWC mekanizmasının bilgi koruma ve uyum sağlama açısından ne ölçüde katkı sunduğunu gözlemlemeye yönelik bir karşılaştırma (baseline) görevi görmektedir.

4.1.3.5. Sınıf artımlı öğrenme ve sınıf dengesizliğinin giderilmesi

Sınıf dengesizliği, özellikle sınıf artımlı öğrenme süreçlerinde modelin azınlık sınıfları öğrenmede yetersiz kalmasına ve tahmin performansında ciddi sapmalara yol açabilmektedir. Bu çalışmada, sınıf dengesizliğini azaltmak amacıyla üç temel strateji uygulanmaktadır:

1. **Ağırlıklı Kayıp Fonksiyonu:** Model eğitimi sırasında, her sınıfa ait örneklerin veri setindeki frekanslarına ters orantılı olarak ağırlıklandırıldığı dengelenmiş bir çapraz entropi kaybı fonksiyonu kullanılmaktadır. Bu sayede, sık görülen sınıfların toplam kayıptaki etkisi azaltılırken, az görülen sınıfların modele katkısı korunmaktadır.
2. **Ortak Özellik Temsili:** Model mimarisinde yer alan ortak kodlayıcı katman (common encoder), farklı görevler ve sınıflar arasında genellenebilir temsil öğrenimini desteklemektedir. Böylece, sınıf bazlı önyargıların etkisini azaltarak nadir sınıfların da etkili bir biçimde temsil edilmesi sağlanmaktadır.

Ortak kodlayıcı, yalnızca görev-bağımlı çıkış katmanlarından bağımsız olarak güncellenmekte ve bilgi koruma mekanizması (EWC) ile desteklenmektedir.

3. Ayrıştırılmış Çıkış Katmanları: Her yeni görev tanıtıldığında modele göreve özgü bir çıkış katmanı (task-specific head) eklenmektedir. Bu ayrıştırılmış yapı, önceki görevlerden kaynaklanabilecek önyargıların yeni görevlere taşınmasını engellerken, her görev için daha esnek ve özelleştirilmiş bir öğrenme süreci sağlamaktadır. Bu sayede modelin yeni sınıfları öğrenme kabiliyeti artırılırken, sınıf dengesizliğinden doğabilecek negatif transfer etkileri en aza indirgenmektedir.

4.1.4. Veri setinin kullanımı

Bu tez çalışmasında kullanılan CICIoT2023 veri seti, toplam 34 farklı sınıfı temsil eden örneklerden oluşmaktadır. Modelin artımlı öğrenme yetisini değerlendirebilmek ve farklı sınıf dağılımları altında sergilediği davranışı gözlemleyebilmek amacıyla veri seti iki aşamalı bir yapıya dönüştürülmüştür.

1. Başlangıç Eğitim Verisi: İlk deney kapsamında, veri seti içerisinde en sık görülen 10 sınıf çıkarılmış ve model bu sınıflar olmadan eğitilmektedir. Bu yöntem, veri setine bilinçli bir şekilde dengesizlik kazandırarak modelin nadir sınıflara odaklanmasını sağlamayı ve ilk aşamada azınlık sınıfların temsiliyetini artırmayı hedeflemektedir.
2. Artımlı Eğitim Verisi: İkinci aşamada, başlangıçta dışlanan 10 yüksek frekanslı sınıfı içeren veri seti modele artımlı olarak sunulmaktadır. Bu adımda amaç, modelin hem yeni sınıflara adapte olma kapasitesini hem de önceki görevlerde öğrenilen bilgileri ne ölçüde koruyabildiğini değerlendirmektir.

4.1.5. Eğitim ve test süreci

Her iki deneysel senaryo için de eğitim ve değerlendirme süreci üç aşamalı olarak yürütülmektedir.

1. İlk olarak, veri setinin belirlenen ilk bölümü kullanılarak model eğitilmektedir. Bu aşamada ortak kodlayıcı ve ilk göreve ait görev-bağımlı çıkış katmanı başlatılmaktadır. Elde edilen model başlangıç referansı olarak kaydedilmektedir.
2. İkinci aşamada, başlangıçta dışlanan sınıfları içeren veri setiyle artımlı eğitim gerçekleştirilmektedir. Bu süreçte, ortak kodlayıcı, EWC mekanizması ile

düzenlenerek ve yeni sınıflara uyum sağlayabilmesi amacıyla modele yeni bir görev-bağımlı çıkış katmanı eklenmektedir. Artımlı eğitim tamamlandığında model yeniden kaydedilmektedir.

3. Son olarak, değerlendirme sürecinde model önce yalnızca ilk görevdeki sınıflar kullanılarak test edilmekte, ardından tüm sınıfları içeren bir test veri seti ile genel performansı ölçülmektedir. Bu değerlendirme, modelin önceki bilgileri ne ölçüde koruyabildiğini (bilgi tutma), yeni sınıflara uyum sağlama kapasitesini (uyarlanabilirlik) ve tüm sınıflar üzerinde dengeli bir performans sergileyip sergilemediğini (genel doğruluk) ortaya koymayı amaçlamaktadır.

4.1.6. Değerlendirme metrikleri

Test süreci boyunca toplanan metrikler, modelin hem önceki görevlerdeki bilgileri koruma düzeyini hem de yeni sınıflara ne ölçüde başarılı bir şekilde uyum sağlayabildiğini çok boyutlu olarak analiz etmek amacıyla kullanılmaktadır. Bu tez çalışmasında kullanılan metrikler aşağıda sıralanmaktadır.

Doğruluk (Accuracy): Modelin tüm görevler üzerindeki genel sınıflandırma başarımını ölçmektedir.

$$Doğruluk = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.5)$$

Sınıf Dengeli Doğruluk (Class-Balanced Accuracy): Sınıf dengesizliğini dikkate alarak her sınıf için doğruluk oranının ortalamasını almaktadır. Böylece azınlık sınıfların performansı da hesaba katılmaktadır.

$$Dengeli Doğruluk = \frac{1}{C} \sum_{i=1}^C \frac{TP_i}{TP_i + FN_i} \quad (4.6)$$

C , toplam sınıf sayısını göstermektedir.

Bilgi Tutma Skoru (Retention Score): Modelin artımlı eğitim sonrası önceki görevlerde öğrendiği bilgileri ne ölçüde koruyabildiğini değerlendirmektedir.

$$\text{Hatırlama Skoru} = \frac{\text{Yeni Görev Eğitimi Sonrası Eski Sınıflardaki Performans}}{\text{Yeni Görev Eğitimi Öncesi Eski Sınıflardaki Performans}} \quad (4.7)$$

F1-Skoru: Hassasiyet (precision) ve geri çağırma (recall) metriklerinin dengeli bir birleşimi olarak, özellikle dengesiz sınıf dağılımlarında etkili bir performans ölçütü olarak hesaplanmaktadır.

$$F1 = 2 \frac{\text{Kesinlik} * \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (4.8)$$

Makro ve Mikro Metrikler: Makro metrikler, her sınıf için performans ölçütlerini ayrı ayrı hesaplayarak ortalamasını almaktadır. Mikro metrikler ise tüm sınıflardaki doğru ve yanlış tahminleri birleştirerek tek bir genel skor üretmektedir.

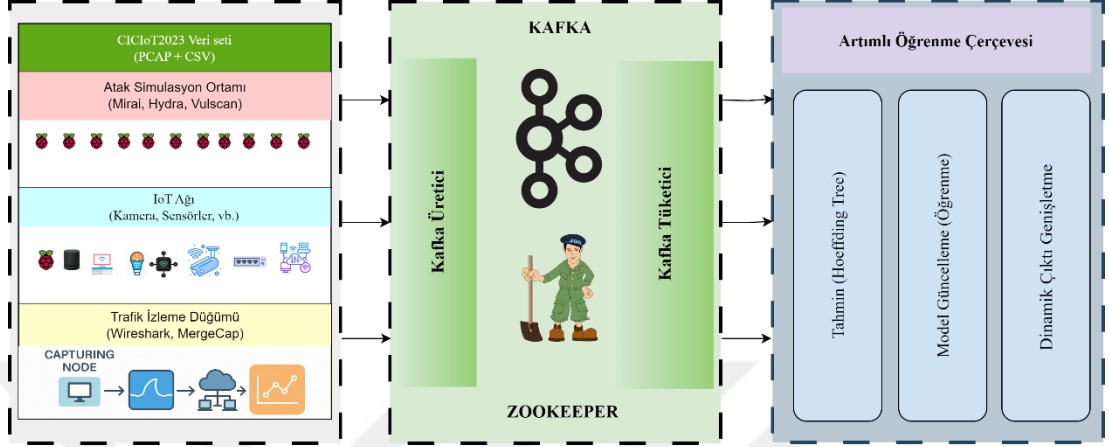
4.2. Dinamik Çıktı Genişletme ile Hoeffding Tree

Bu tez kapsamında önerilen ikinci model olan CIL-DOE çerçevesi, akan veri ortamlarında sınıf artımlı öğrenmeyi mümkün kılmak amacıyla Hoeffding Tree algoritması timeline yapılandırılmaktadır. Geleneksel Hoeffding Tree modelleri, sabit bir çıktı sınıf uzayı varsayımıyla çalışmaktadır. Öğrenme süreci sırasında yeni sınıf etiketleriyle karşılaşıldığında modelin mimarisinde kapsamlı değişiklikler ya da yeniden eğitim gerektirmektedir. Bu yapısal sınırlılığı aşmak üzere geliştirilen dinamik çıktı genişletme mekanizması, modelin çevrim içi eğitim sürecinde ilk kez karşılaştığı sınıf etiketlerini otomatik olarak tanımasını ve çıktı uzayını gerçek zamanlı olarak genişletmesini sağlamaktadır. Böylece, model yeni sınıfları tanıma ve öğrenme sürecini kesintisiz şekilde sürdürebilmektedir. Ayrıca daha önce edinilmiş bilgileri koruyarak artımlı öğrenmeye olanak tanımaktadır.

Hoeffding Tree algoritması, sınırlı bellekle çalışabilen ve yalnızca istatistiksel olarak anlamlı özellik ayrımlarını takip eden bir çevrim içi karar ağacı yapısıdır. Bu özellikleri nedeniyle özellikle akan veri ortamlarında gerçekleştirilen uygulamalarda tercih edilmektedir. Bu tez çalışması kapsamında önerilen model, her gelen örnek için önce tahmin yapar, ardından doğru etiketle öğrenme işlemini gerçekleştirmektedir (prediction-before-update). Bu yaklaşım, modelin değerlendirme metriklerinin eğitim sırasında dinamik olarak izlenmesine imkân tanımaktadır.

4.2.1. Model mimarisi

Önerilen CIL-DOE çerçevesi, Şekil 4.3.'de sunulduğu üzere, akan veri ortamlarında sınıf artımlı öğrenmeyi desteklemek amacıyla bütüncül biçimde kurgulanmış üç temel bileşenden oluşmaktadır.



Şekil 4.3. Önerilen CIL-DOE çerçevesinin sistem mimarisi.

İlk bileşen, CICIoT2023 veri seti kullanılarak oluşturulan trafik üretim ve izleme altyapısıdır. Bu yapı, gerçekçi IoT ağ ortamlarını ve güncel siber saldırı senaryolarını simüle ederek, modelin dinamik koşullar altında gerçek dünyaya yakın bir biçimde değerlendirilmesine olanak tanımaktadır.

İkinci bileşen, Apache Kafka altyapısını kullanan bir veri akış hattıdır. Bu bileşen, veri örneklerinin kaynak sistemden modele zaman açısından tutarlı ve kesintisiz biçimde iletilmesini sağlayarak, çevrim içi öğrenme senaryolarının güvenilir ve yeniden üretilebilir biçimde gerçekleştirilmesine imkân sunmaktadır.

Üçüncü ve son bileşen ise, Hoeffding Tree algoritması üzerine inşa edilen dinamik öğrenme modülüdür. Bu modül, modelin çevrim içi olarak sürekli güncellenmesini mümkün kılmaktadır. Ayrıca eğitim sürecinde yeni sınıf etiketleri ile karşılaşıldığında çıktı uzayının anlık olarak genişletilmesini sağlamaktadır. Böylece sistem, önceden tanımlı sınıf bilgisine ihtiyaç duymaksızın yeni sınıfları tanıma ve öğrenme kapasitesi kazanmakta, yeniden eğitim gereksinimi olmaksızın evrimsel sınıf dağılımlarına uyum sağlayabilmektedir.

4.2.2. Hoeffding Tree

Hoeffding Tree, sınırlı sayıda örnekle yüksek doğrulukta kararlar verebilmesini sağlayan Hoeffding sınırına dayalı istatistiksel güvence mekanizması sayesinde, akan

veri ortamlarında artımlı öğrenme için sıklıkla tercih edilen bir yöntemdir. Bu algoritma, karar noktalarında bölünme işlemini gerçekleştirmek üzere entropi veya Gini indeksi gibi bilgi kazancı ölçütlerini kullanmaktadır. Yalnızca istatistiksel olarak anlamlı farklar tespit edildiğinde düğüm bölünmesine izin vermektedir. Bu sayede, düşük hesaplama ve bellek maliyetiyle birlikte gerçek zamanlı öğrenmeyi mümkün kılmaktadır. Özellikle kaynak kısıtlı senaryolarda önemli bir avantaj sunmaktadır. Söz konusu özellikleri nedeniyle Hoeffding Tree, çevrim içi karar verme sistemleri, veri akış madenciliği ve siber güvenlik gibi çeşitli alanlarda yaygın şekilde kullanılmaktadır (Domingos ve Hulten, 2000).

4.2.3. CIL-DOE öğrenme süreci

Bu tez çalışmasında önerilen CIL-DOE çerçevesinde, model tüm veri kümesini belleğe önceden yüklemek yerine, akan veri senaryolarına uygun olarak kayıtları bir döngü içerisinde sırasıyla ve tek tek işlemektedir. Bu yapı sayesinde, gerçek zamanlı veri akışının doğal yapısı korunmakta ve model, her yeni örnek için anlık tahmin ve güncelleme döngüsünü sürdürebilmektedir.

Çerçevenin öğrenme mekanizması üç temel bileşenden oluşmaktadır:

1. Tahmin öncesi değerlendirme (prediction-before-update)
2. Yeni sınıf etiketleriyle karşılaşıldığında dinamik çıktı katmanının genişletilmesi
3. Anlık olarak hesaplanan değerlendirme metriklerinin kaydedilmesi.

Bu süreçlerin bütüncül şekilde tanımlandığı öğrenme algoritması, Şekil 4.4.'te ayrıntılı olarak sunulmaktadır.

```

Input:
  D: Streaming dataset (ordered instances:  $x_1, x_2, \dots, x_n$ )
  M: Hoeffding Tree model
  C_seen: Set of known class labels (initially empty)
  Kafka_Consumer: stream consumer delivering instances one by one
Initialize:
  M  $\leftarrow$  new HoeffdingTree()
  C_seen  $\leftarrow \emptyset$ 
  metrics_log  $\leftarrow \emptyset$ 
For each incoming instance (x, y) from Kafka_Consumer do:
  If y  $\notin$  C_seen then
    Add y to C_seen
    Dynamically expand the output layer of M to include class y
    Log (step, y) to class_log
  y_pred  $\leftarrow$  M.predict(x)
  If y_pred is not None, then
    Update evaluation metrics:
      Accuracy.update(y, y_pred)
      MacroFl.update(y, y_pred)
      MicroFl.update(y, y_pred)
      Precision_Macro.update(y, y_pred)
      Recall_Macro.update(y, y_pred)
      Kappa.update(y, y_pred)
      MCC.update(y, y_pred)
  M.learn(x, y)
  If step % k == 0 then
    Save metrics_log[step]  $\leftarrow$  current metric values
End For

```

Şekil 4.4. CIL-DOE: dinamik çıktı genişletme ile sınıf artımlı öğrenme.

Bu tez çalışmasında model geliştirme sürecinde hafif ve akış odaklı bir ön işleme stratejisi benimsenmektedir. Özellikle sızma tespiti gibi zaman duyarlı alanlarda, kurulan sistemin gerçek dünya konuşlandırma koşullarına yakın kalmasını sağlanmaktadır.

Önerilen sistem temel olarak, veri akışı sırasında daha önce hiç karşılaşılmamış sınıf etiketleriyle karşılaşıldığında çıktı uzayını dinamik olarak genişletebilmektedir. Bu sayede model, önceden tanımlanmış sabit bir sınıf kümesine bağlı kalmaksızın, akış sürecinde yeni sınıf etiketleri ile karşılaştıkça içsel sınıf haritalama yapısını artımlı olarak güncelleyebilmektedir.

Gerçek zamanlı öğrenme altyapısında Apache Kafka kullanılmaktadır. Veri üreticisi tarafından gönderilen örnekler sırayla modele aktarılmaktadır. Her bir veri örneği model tarafından önce tahmin edilip, ardından doğru sınıf etiketiyle güncellenmektedir. Eğer bu etiket mevcut sınıf kümesinde yer almıyorsa, modelin çıktı yapısı genişletilmekte ve yeni sınıf sürece dâhil edilmektedir. Burada y_t , t adında gelen sınıfı gösterirken, Y_{mevcut} modelin öğrenmiş olduğu sınıf kümesini göstermektedir.

$$y_t \notin Y_{mevcut} \leftarrow Y_{mevcut} \cup \{y_t\} \quad (4.9)$$

Akış sırasında yeni bir sınıf etiketi içeren veri örneği gözlemlendiğinde, bu etiket modelin iç sınıf listesine eklenmektedir. Dinamik olarak yönetilen bir eşleme yapısı aracılığıyla söz konusu etikete karşılık gelen sayısal bir kimlik (ID) atanmaktadır. Bu sayede modelin, akan veri ortamında sınıf artımlı öğrenme sürecini nasıl gerçekleştirdiği net bir biçimde ortaya konmaktadır. Ayrıca sistemin çevrim içi koşullarda yeni bilgilere uyum sağlama kapasitesi somut biçimde gösterilmektedir.

CIL-DOE çerçevesinin öğrenme mekanizması, akış verisinin düzenli biçimde işlenmesini, çıktı alanının dinamik olarak genişletilmesini ve Hoeffding sınırına dayalı güncelleme kuralını içeren yapılandırılmış bir yaklaşıma dayanmaktadır. Modelin akan veriden artımlı biçimde öğrenmesini sağlarken önceki bilgileri etkin şekilde koruyabilmesinin temelini oluşturmaktadır.

Her t anında mevcut model hipotezi h_t , gelen veri örneği x_t ve bununla birlikte doğru sınıf etiketi y_t gelmektedir. Modelin t anına kadar öğrendiği sınıf kümesi $Y_{current}$ gelen yeni sınıf etiketi ile güncellenmektedir. Bu işlem, modelin çıktı yapısını genişletmesini ve yeni sınıf bilgilerini kesintisiz biçimde sistemine entegre etmesini sağlayan CIL_DOE fonksiyonu aracılığıyla gerçekleştirilmektedir.

$$h_{t+1} = CIL_DOE(h_t, x_t, y_t, Y_t) \quad (4.10)$$

Bu yaklaşım sayesinde model, önceki öğrenmeleri muhafaza edebilmektedir. Bununla birlikte yeni gelen sınıf etiketlerine uyum sağlayarak evrimleşen veri dağılımına adapte olabilmektedir. Dinamik sınıf genişletme mekanizması, çevrim içi (online) olarak uygulanmakta ve akış döngüsü içerisinde kesintisiz biçimde yürütülmektedir. Bu sayede zamanla yeni tehdit türlerinin ortaya çıkabildiği siber güvenlik gibi durağan olmayan (non-stationary) ortamlarda modelin adaptasyon kapasitesini artırmaktadır. Özellikle sınıf dağılımının önceden belirli olmadığı ve sürekli değiştiği saldırı tespiti sistemleri gibi uygulamalar için kritik önem taşımaktadır.

4.2.4. Veri setinin kullanımı

Bu tez çalışmasında CICIOT2023 veri seti, çevrim içi ve akış tabanlı öğrenme yaklaşımlarına uygun olarak, verilerin modele tek tek ve sıralı biçimde işlendiği bir

yapıda değerlendirilmektedir. Her yeni örnek geldiğinde model önce tahmin yapmakta, ardından gerçek etiketle güncellenmektedir. Bu yapı, modelin kararsız veri dağılımlarına uyum yeteneğini test etmektedir. Ayrıca akış sırasında karşılaşılan yeni sınıf etiketleri çıktıya dinamik olarak eklenmekte, böylece sabit sınıf uzayına bağlı kalmadan öğrenme sürdürülebilmektedir.

Bu tez çalışmasında önerilen ikinci modelde, CICIOT2023 veri seti, çevrim içi ve akış tabanlı öğrenme senaryolarına uygun şekilde yapılandırılmakta ve artımlı öğrenme bağlamında iki temel deneysel düzende değerlendirilmektedir.

4.2.4.1. Aşamalı (kademeli) sınıf tanıtımı deneyi

İlk aşamada veri seti üç ayrı faza ayrılarak işlenmektedir. Her faz, modelin hem sık hem de seyrek görülen sınıflara karşı gösterdiği uyumu ve bilgi koruma yetisini değerlendirmeyi amaçlamaktadır.

1. Aşama 1: Veri setinden en az sıklıkta gözlemlenen 10 sınıf çıkarılarak ve model yalnızca sık sınıflarla eğitilmektedir. Böylece modelin az görülen sınıflarla karşılaşmadan genelleme kapasitesi gözlemlenmektedir.
2. Aşama 2: Eğitim süreci Faz 1'in sonunda elde edilen model üzerinden devam ettirilerek en sık görülen 10 sınıf veri setinden çıkarılarak model az görülen sınıflarla eğitilmektedir. Bu aşamada amaç, önceki bilginin ne ölçüde korunabildiğini ve yeni sınıflara ne düzeyde adapte olunabildiğini değerlendirmektir.
3. Aşama 3: Bu fazda tüm sınıflar modele tanıtılarak, yani CICIOT2023 veri seti hiçbir sınıf dışlanmadan akışa aktarılmaktadır. Bu aşamanın temel amacı, modelin tüm sınıf etiketleriyle eş zamanlı olarak karşılaştığı gerçekçi bir dağıtım senaryosu altında gösterdiği genel öğrenme performansını değerlendirmektir. Böylece modelin, hem daha önceki sınıfları unutma eğilimi hem de yeni sınıflara adaptasyon kapasitesi bütünsel olarak gözlemlenebilmekte ve önceki fazlarla karşılaştırmalı analiz yapılabilmektedir.

4.2.4.2. Tam akış (end-to-end) deneyi

İkinci deneysel uygulamada veri seti herhangi bir sınıf ayrımı yapılmaksızın, zaman sırasına göre tek seferde modele aktarılmaktadır. Bu uygulama, modelin sınıfları önceden görmeden, akış sırasında karşılaştığı sırayla gördüğü ve böylece sınıf bilgisini

zaman içinde edinilirken öğrenmenin sürdürüldüğü bir tam akışlı öğrenme senaryosunu temsil etmektedir. Bu uygulama, modelin gerçek zamanlı koşullarda önceden tanımlı bir sınıf bilgisine sahip olmadan öğrenme yapabilmesini değerlendirmeyi amaçlamaktadır. Ayrıca bu uygulama kademeli tanıtım stratejileri ile karşılaştırılabilir bir referans (baseline) oluşturmaktadır.

4.2.5. Değerlendirme metrikleri

Önerilen modelin sınıf artımlı öğrenme sürecindeki başarımını çok yönlü olarak değerlendirebilmek amacıyla, yalnızca genel doğruluk değil, aynı zamanda sınıf dengesizliği, bilgi koruma ve yeni sınıflara uyum gibi kritik açılara odaklanan çeşitli performans metrikleri kullanılmaktadır. Aşağıda bu çalışmada kullanılan temel değerlendirme ölçütleri özetlenmektedir.

Doğruluk (Accuracy): Modelin tüm örnekler üzerindeki doğru sınıflandırma oranını vermektedir. Ancak sınıf dengesizliğinin olduğu durumlarda yanıltıcı olabileceği için, diğer tamamlayıcı metriklerle desteklenmektedir.

$$Doğruluk = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.11)$$

Makro Ortalama Metrikler: Her sınıf için ayrı ayrı hesaplanan değerlere dayalı olarak ağırlıksız ortalama alınmaktadır. Bu yaklaşım sayesinde az görülen sınıfların performansı değerlendirilmektedir. En önemli makro ortalama metriklerinden biri olan Makro F1 skoru, her sınıfın F1 değerinin eşit ağırlıkla ortalamasını alarak sınıflar arası dengeyi korunup korunmadığını ölçmektedir. Azınlık sınıfların öğrenilme düzeyini değerlendirmektedir.

$$Makro - F1 = \frac{1}{|Y|} \sum_{c \in Y} F1_c \quad (4.12)$$

Burada Y sınıf kümesini, $F1_c$ ise her bir sınıf için hesaplanan F1 değerini ifade etmektedir. Her sınıfa eşit ağırlık verilerek modelin genel başarımının değerlendirilmesini sağlamaktadır.

Mikro Ortalama Metrikler: Tüm sınıflardan gelen katkılar birleştirilerek genel bir performans ölçümü yapılmaktadır. Bu ölçüm özellikle sınıf dengesizliğinin baskın olduğu durumlarda genel sınıflandırma davranışını yansıtmaktadır.

Balanced Accuracy (Dengeli Doğruluk): Tüm sınıflar için ortalama duyarlılık (recall) değeri alınarak hesaplanmaktadır. Sınıf dağılımının zaman içinde değiştiği akış ortamlarında daha sağlıklı sonuçlar sunmaktadır.

Cohen's Kappa (κ): Modelin tahminleri ile gerçek etiketler arasındaki rastlantısal uyumu düzelten bir ölçüdür. Doğruluğa alternatif bir güvenilirlik göstergesi sağlamaktadır.

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (4.13)$$

Burada p_o gözlenen doğruluk, p_e ise rastgele tahminle beklenen doğruluktur.

Matthews Korelasyon Katsayısı (MCC): Karmaşıklığı yüksek çok sınıflı ve dengesiz sınıf yapılarına uygun bir metriktir. Karışıklık matrisinin tüm bileşenlerini dikkate alarak daha dengeli bir değerlendirme yapmaktadır.

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (4.14)$$

Bunlara ek olarak, sınıf düzeyinde duyarlılık (recall), kesinlik (precision) ve F1 skorları ayrı ayrı hesaplanarak ve özellikle düşük temsil oranına sahip sınıflar üzerindeki model performansı detaylı biçimde incelenmektedir. Bu yaklaşımlar, sınıflar arası ayırım gücünü ve temsil adaletini ortaya koymaktadır. Tüm bu değerlendirme kriterleri, her deney süresi boyunca ayrı ayrı hesaplanmaktadır. Böylece zaman içerisinde modelin öğrenme eğilimi, yeni sınıflara adaptasyon yeteneği ve önceki bilginin korunma düzeyi bütüncül olarak analiz edilmektedir.

4.3. Modellerin Karşılaştırılması

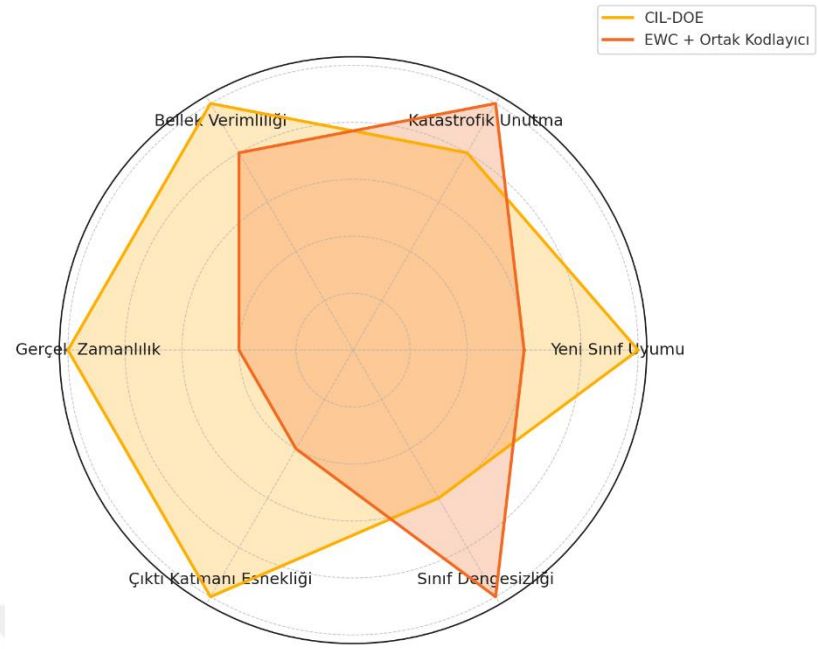
Tablo 4.1., önerilen CIL-DOE ve EWC tabanlı modellerin sınıf artımlı öğrenme bağlamındaki temel özelliklerini karşılaştırmalı olarak sunmaktadır. CIL-DOE modeli, özellikle yeni sınıflara gerçek zamanlı adaptasyon, çıktı katmanının dinamik genişletilebilmesi ve belleğe ihtiyaç duymadan çalışabilmesi gibi özellikleriyle öne çıkmaktadır. Bu yönleriyle, yüksek frekansta akan veri içeren ve sınıfların zamanla değiştiği senaryolarda pratik bir çözüm sunmaktadır. EWC destekli model ise katastrofik unutmanın önlenmesi, sınıf dengesizliğine duyarlı kayıplar ile öğrenmenin dengelenmesi ve düşük bellek kullanımı ile dikkat çekmektedir. Görev bazlı yapısı ve

bilgi sürekliliğini koruma kapasitesi, özellikle geçmiş bilginin uzun vadede korunmasının önemli olduğu görevlerde avantaj sağlamaktadır. Her iki model, farklı sistem ihtiyaçlarına uygun olarak yapılandırılmıştır ve tamamlayıcı güçlü yönler sergilemektedir.

Tablo 4.1. Önerilen modellerin karşılaştırması.

Kriter	CIL-DOE	EWC + Ortak Kodlayıcı
Model Türü	Karar ağacı + Dinamik çıktı genişletme	Derin ağ + EWC tabanlı düzenleme
Yeni Sınıf Uyumu	Otomatik, dinamik çıktı katmanı	Görev bazlı, ön tanımlı sınıf gerektirir
Katastrofik Unutma	Parametre kararlılığı ile azaltılır	EWC ile bilgi korunumu sağlanır
Bellek Kullanımı	Gerektirmez	Düşük (önemli parametrelerin izlenmesi için)
Gerçek Zamanlılık	Kafka entegrasyonu ile tam uyumlu	Çevrimdışı senaryolara uygun
Çıktı Katmanı	Dinamik	Statik
Sınıf Dengesizliği	Azınlık sınıflarda performans farkı gözlemlenebilir	Ağırlıklı ve dengeli kayıpla ele alınır
Uygulama Alanı	Büyük veri, siber güvenlik, akış verisi	Görev bazlı sınıflandırma, azınlık sınıflar
Özgünlük	Yeniden eğitim gerektirmeyen, bellek içermeyen CIL	Bilgi sürekliliği ve denge odaklı encoder yaklaşımı

Şekil 4.5., önerilen iki modelin altı temel değerlendirme ölçütü kapsamında nicel olarak karşılaştırılmasını sağlamaktadır. Bu ölçütler; yeni sınıflara uyum sağlama kapasitesi, katastrofik unutma direnci, bellek verimliliği, gerçek zamanlı işlem kabiliyeti, çıktı katmanı esnekliği ve sınıf dengesizliğiyle başa çıkma başarımı olarak belirlenmiştir. Şekil, CIL-DOE modelinin, özellikle yüksek adaptasyon yeteneği, gerçek zamanlı işleme uygunluğu ve dinamik çıktı yapısı ile dikkat çektiğini ortaya koymaktadır. Öte yandan, EWC destekli model, katastrofik unutmanın azaltılması, azınlık sınıflardaki başarımın korunması ve dengeli öğrenme performansı açısından daha güçlü bir görünüm sergilemektedir. Her iki modelin de bellek kullanımı açısından verimli olduğu gözlemlenmektedir. Bu karşılaştırma, uygulama bağlamına göre model tercihi yönlendirecek bütüncül ve çok boyutlu bir değerlendirme sunmaktadır.



Şekil 4.5. Önerilen modellerin karşılaştırmalı grafiği.



5. GELİŞTİRİLEN YÖNTEMLERİN BAŞARIM ANALİZİ

Bu bölümde, çalışmada önerilen iki farklı sınıf artımlı öğrenme yaklaşımının deneysel bulguları sunulmakta ve kapsamlı bir başarı analizi gerçekleştirilmektedir. İlk olarak, sınıf dengesizliği ve katastrofik unutma sorunlarını hedef alan EWC destekli ortak kodlayıcı mimari ele alınmaktadır. Bu yaklaşım, paylaşımlı bir özellik çıkarım katmanı ile görev-özgü çıktı başlıklarını birleştirerek, önceki bilgi birikiminin korunmasını amaçlamaktadır. İkinci olarak, CIL-DOE adlı ve Hoeffding Tree modeli değerlendirilmektedir. Bu yöntem, gerçek zamanlı veri akışı sırasında yeni sınıflarla karşılaşıldığında, çıktı katmanını dinamik biçimde genişletme yeteneğine sahiptir. Her iki yöntem de CICIoT2023 veri seti kullanılarak, farklı sınıf bölünmeleri ve artımlı senaryolar çerçevesinde test edilmektedir.

Elde edilen sonuçlar, her iki modelin başarımlarını ayrı ayrı değerlendirmek üzere analiz edilmektedir. EWC destekli ortak kodlayıcı yaklaşım, özellikle sınıf dengesizliği altındaki senaryolarda azınlık sınıflara karşı yüksek duyarlılık göstermektedir. Bu yöntem bilgi kaybını azaltmada etkili olmaktadır. CIL-DOE yaklaşımı ise, gerçek zamanlı veri akışında yeni sınıfların dinamik olarak entegre edilmesini başarıyla gerçekleştirmektedir. Yüksek mikro düzeyde doğruluk ve kararlılık sergileyerek model dayanıklılığını ön plana çıkarmaktadır.

5.1. Uygulamalarda Kullanılan Veri Seti

Bu çalışmada, sınıf artımlı öğrenme yaklaşımlarının değerlendirilmesi amacıyla CICIoT2023 veri seti kullanılmaktadır. Canadian Institute for Cybersecurity (CIC) araştırmacıları tarafından oluşturulan bu veri seti, IoT ortamlarında gerçekleştirilen çeşitli siber saldırı türlerini ve normal trafik örneklerini içeren, gerçekçi bir ağ trafiği yapısını sunmaktadır (Pinto Neto ve ark, 2023)..

Veri seti DDoS-RSTFINFlood, DoS-TCP_Flood, DDoS-ICMP_Flood, MITM, DoS-UDP_Flood, DoS-SYN_Flood, DDoS-UDP_Flood, Port Scanning, DDoS-SYN_Flood, DNS_Spoofing, SQL_Injection, OS_Fingerprint, Backdoor ve Benign

(normal trafik) gibi 1 normal (benning) trafik sınıfı ve 33 farklı saldırı sınıfı olmak üzere toplam 34 sınıf içermektedir.

Veri seti, CSV formatında yapılandırılmış olup, CICFlowMeter gibi araçlar tarafından türetilmiş özniteliklerle zenginleştirilmiş ağ akışı kayıtlarını içermektedir.

Veri seti yaklaşık 16 milyon örnek içermektedir. Her bir örnek 46 sayısal öznitelikten oluşmaktadır. Bu öznitelikler, ağ trafiğine ilişkin akış bazlı detaylı istatistikleri yansıtmaktadır. Öne çıkan bazı öznitelikler şu şekildedir:

- Flow Duration
- Total Forwarded Packets
- Mean Forwarded Packet Length
- Flow Bytes per Second
- Mean Active Time
- Maximum Idle Time
- Vb. diğer trafik davranışlarını yansıtan istatistiksel özellikler

5.2. EWC Destekli Ortak Kodlayıcı Mimari Deneysel Sonuçlar

Bu uygulamada, bir sınıflandırma modelinin farklı eğitim senaryoları altındaki başarımı değerlendirilmektedir. Deneysel yapılandırmalarda, EWC destekli öğrenme, EWC uygulanmadan gerçekleştirilen öğrenme ve veri setinden en sık ve en seyrek görülen 10 sınıfın çıkarılması şeklinde senaryolar ele alınmaktadır. Yapılan değerlendirmeler sonucunda, modelin davranış özellikleri, sınıf dengesizliğine karşı gösterdiği tepki ve artımlı öğrenme stratejilerinin başarıma etkisi kapsamlı biçimde ortaya konmaktadır.

Önerilen yöntemin genel doğruluk düzeyi ile birlikte makro ölçekteki metrikler (precision, recall ve F1-score), modelin sınıf dengesizliği içeren artımlı öğrenme senaryolarıyla etkili biçimde başa çıkabildiğini göstermektedir. Tablo 5.1.'de sunulan sonuçlara göre, doğruluk oranı hem başlangıç hem de son eğitim aşamalarında yüksek düzeyde istikrar sergilemektedir. Başlangıç aşamasında doğruluk %90,63 ile %98,38 arasında değişirken, son aşamada bu oran %86,07 ile %97,52 arasında seyretmektedir.

Bununla birlikte, makro precision, recall ve F1 değerlerinde özellikle şiddetli sınıf dengesizliği durumlarında belirgin dalgalanmalar gözlemlenmektedir. Azınlık sınıflardaki başarıyı değerlendirmede kritik bir gösterge olan makro F1 skoru, modelin

nadir görülen sınıfları öğrenmedeki zorluklarını ortaya koymaktadır. EWC uygulanmadığı durumlarda, makro F1 değerlerinde kayda değer düşüşler meydana gelirken, EWC mekanizmasının devreye alınmasıyla bu değerler önemli ölçüde iyileşmiş, bazı senaryolarda %82,01 düzeyine kadar ulaşmaktadır.

Tablo 5.1. Genel performans metrikleri.

Metrik	EWC	Doğruluk	Makro Kesinlik	Makro Duyarlılık	Makro F1-Skor
İlk Eğitim (En Sık 10 Sınıf Hariç)	X	0.9063	0.7077	0.5442	0.5633
Son Eğitim (En Sık 10 Sınıf Dahil)	X	0.9685	0.6301	0.5143	0.5143
İlk Eğitim (En Az 10 Sınıf Hariç)	X	0.9838	0.8714	0.8346	0.8457
Son Eğitim (En Az 10 Sınıf Dahil)	X	0.9752	0.6232	0.5546	0.5609
İlk Eğitim (En Sık 10 Sınıf Hariç)	✓	0.9099	0.6866	0.5507	0.5713
Son Eğitim (En Sık 10 Sınıf Dahil)	✓	0.8607	0.665	0.5787	0.593
İlk Eğitim (En Az 10 Sınıf Hariç)	✓	0.9809	0.871	0.8028	0.8201
Son Eğitim (En Az 10 Sınıf Dahil)	✓	0.9809	0.871	0.8028	0.8201

EWC mekanizmasının uygulanması, genel olarak modelin kararlılığını ve başarımını artırmış, özellikle de azınlık sınıflar açısından olumlu etkiler sağlamaktadır. Makro F1 skorlarında çoğu senaryoda gözlemlenen artış, modelin hem nadir hem de orta sıklıkta gözlemlenen sınıflarla başa çıkma yeteneğini geliştirdiğini göstermektedir. Örneğin, "en az görülen 10 sınıfın çıkarıldığı" senaryoda, son eğitim aşamasında makro F1 skoru, EWC uygulanmadan önce %56,09 iken, EWC uygulandığında %59,29'a yükselmektedir. Bu iyileşme, EWC'nin kritik parametreleri koruyarak bilgi kaybını en aza indirme ve aynı anda yeni görevleri öğrenmeye uyum sağlama konusundaki etkinliğini ortaya koymaktadır.

Tablo 5.2.'de ayrıntılı biçimde sunulan bulgular, EWC'nin hem yüksek hem de düşük frekanslı sınıflar üzerinde sınıf özelinde doğruluk oranlarını artırdığını göstermektedir. BenignTraffic ve DDoS-ICMP_Flood gibi sık görülen sınıflarda doğruluk oranı sırasıyla %95'ten %96'ya, %98'den %99'a çıkmıştır. Ancak EWC'nin etkisi, seyrek sınıflar üzerinde çok daha belirgin olmaktadır. Örneğin, Backdoor_Malware sınıfında

doğruluk oranı EWC olmadan %2 iken, EWC ile bu oran %5'e yükselmektedir. SqlInjection sınıfı için ise bu değer %1'den %3'e çıkmaktadır.

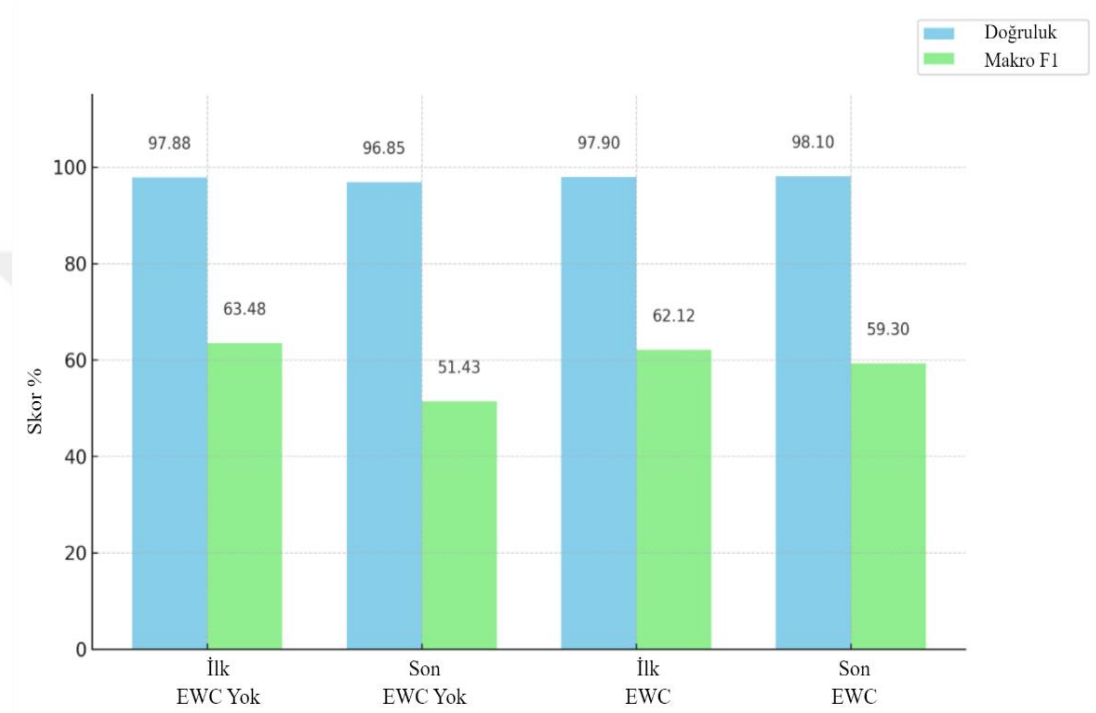
Tablo 5.2. Sınıf dengeli doğruluk.

Sınıf	Doğruluk (EWC olmadan)	Doğruluk (EWC ile)
BenignTraffic	0.95	0.96
DDoS-ICMP_Flood	0.98	0.99
Backdoor_Malware	0.02	0.05
SqlInjection	0.01	0.03
XSS	0	0
Recon-OSScan	0.45	0.5
MITM-ArpSpoofing	0.7	0.75
DDoS-UDP_Flood	0.98	0.99

Buna karşın, bazı zorlu sınıflarda herhangi bir gelişme gözlemlenmemektedir. Örneğin, XSS sınıfı için doğruluk oranı hem EWC'li hem de EWC'siz senaryolarda %0 olarak kalmaktadır. Buna karşılık, orta sıklıktaki sınıflar arasında yer alan Recon-OSScan ve MITM-ArpSpoofing sınıflarında doğruluk oranı EWC etkisiyle sırasıyla %45'ten %50'ye ve %70'ten %75'e yükselmektedir. Bu sonuçlar, EWC'nin farklı sınıf dağılımları arasında performans istikrarını artırmadaki etkinliğini göstermektedir. Ancak aynı zamanda daha fazla optimizasyon gerektiren sınıfları da işaret etmektedir.

XSS gibi nadir sınıflarda gözlemlenen düşük model başarımı, birden fazla etkenin birleşimine dayanmaktadır. Bu sınıfların eğitim sürecinde yeterince öğrenilememesinin başlıca nedenleri arasında, veri setinde çok az sayıda örneğe sahip olmaları, sınıf içi örneklerin oldukça değişkenlik göstermesi ve bazı öz niteliklerinin daha yaygın sınıflarla benzerlik taşıması yer almaktadır. Bunun sonucunda model, bu sınıfları diğerlerinden ayırt etmekte zorlanmaktadır. Öte yandan, EWC mekanizmasının önceki görevlerde edinilen bilgiyi korumak amacıyla model parametrelerine uyguladığı güncelleme kısıtlamaları, öğrenme sürecinin ilerleyen aşamalarında karşılaşılan nadir sınıflar için yeni ve özgün temsillerin oluşturulmasını dolaylı olarak engelleyebilmektedir. Bu durum, özellikle sonradan tanıtılan düşük frekanslı sınıflar için modelin uyum sağlama kapasitesini sınırlandırmaktadır.

Şekil 2, önerilen modelin EWC mekanizmasıyla ve EWC olmadan elde ettiği performansın karşılaştırmalı analizini sunmaktadır. Değerlendirme, başlangıç ve son eğitim fazları boyunca doğruluk ve makro F1 skoru üzerinden gerçekleştirilmektedir. Elde edilen bulgular, EWC uygulamasının özellikle son eğitim aşamasında, sınıf dengesizliği mevcut olsa dahi, makro düzeydeki başarıyı belirgin şekilde iyileştirdiğini ortaya koymaktadır.



Şekil 5.1. İlk ve son eğitim aşamalarında EWC ile ve EWC olmadan doğruluk ve makro F1 skoru karşılaştırması.

5.2.1. Sınıf çıkarmanın etkisi

Veri setinden en sık görülen 10 sınıfın çıkarılması, modelin odağını daha küçük ve az temsil edilen sınıflara yönlendirmiş ve bunun sonucunda makro metriklerde önemli bir düşüş meydana gelmektedir. EWC uygulanmadan gerçekleştirilen ilk eğitim aşamasında makro recall değeri yalnızca %54,42 olarak gerçekleşmektedir. Bu durum, özellikle Backdoor_Malware, SqlInjection ve XSS gibi nadir sınıflarda yetersiz performansı ortaya koymaktadır. EWC mekanizması, bu olumsuz etkileri sınırlı da olsa azaltmaktadır. İlk eğitim aşamasında, makro F1 skoru EWC olmadan %56,33 iken, EWC uygulandığında %57,13'e yükselmektedir.

En az temsil edilen 10 sınıfın veri setinden çıkarılması, sınıflandırma görevini nispeten basitleştirmekte ve kalan sınıflar için modelin performansını artırmaktadır. EWC

uygulanmadan yapılan ilk eğitimde, makro F1 skoru %84,57'ye ulaşmaktadır. En sık görülen sınıfların çıkarıldığı senaryoda bu oran yalnızca %56,33 olarak gözlemlenmektedir. Bu gözlem, yüksek frekanslı sınıfların varlığının, artımlı öğrenme görevlerine önemli ölçüde karmaşıklık kattığını göstermektedir. Bu yapılandırmada EWC'nin dahil edilmesi model başarımını artırmaktadır. Son eğitim aşamasında EWC sayesinde kritik parametrelerin korunması sağlanmakta ve model %82,01 makro F1 skoruna ulaşmaktadır. Bu sonuç, sınıf dağılımı oldukça dengesiz olsa dahi, EWC'nin artımlı öğrenme sürecinde bilgi koruma açısından önemli bir iyileştirme sağladığını göstermektedir.

Tablo 5.2., farklı sınıflar için elde edilen precision, recall ve F1 skoru değerlerini içermekte olup, en iyi ve en kötü durum senaryolarını sınıf bazında ortaya koymaktadır. BenignTraffic ve DDoS-ICMP_Flood gibi yüksek frekanslı sınıflar için precision ve recall değerleri istikrarlı biçimde yüksek kalmakta, en kötü senaryolarda dahi F1 skoru %85'in üzerinde gerçekleşmektedir. Örneğin, DDoS-ICMP_Flood sınıfı, en iyi senaryoda 0.9994 gibi neredeyse mükemmel bir F1 skoruna ulaşmaktadır.

Tablo 5.3. Bazı sınıflara ait özel performans metrikleri.

Sınıf	Kesinlik (En iyi)	Duyarlılık (En iyi)	F1-Skor (En iyi)	Kesinlik (Enkötü)	Duyarlılık (Enkötü)	F1-Skor (Enkötü)
BenignTraffic	0.8009	0.975	0.8777	0.5648	0.9279	0.851
DDoS- ICMP_Flood	1	0.999	0.9994	0.9715	0.9736	0.9866
Backdoor_Malware	0.4679	0.049	0.089	0	0	0
SqlInjection	0.4301	0.024	0.0458	0	0	0
XSS	0	0	0	0	0	0

Buna karşılık, Backdoor_Malware ve SqlInjection gibi azınlık sınıflar oldukça düşük performans sergilemektedir. En iyi senaryolarda bile bu sınıflar için F1 skorları sırasıyla %4,58 ve %8,9 seviyelerinde kalmaktadır. Özellikle XSS gibi zorluk derecesi yüksek sınıflarda, precision, recall ve F1 skorları tüm senaryolarda 0 olarak kalmaktadır. Bu durum, söz konusu sınıflar için hedefe yönelik özel stratejilerin geliştirilmesi gerektiğini ortaya koymaktadır.

5.2.2. Sınıf artımlı öğrenme başarımı

Elde edilen sonuçlar, artımlı öğrenmenin doğasında var olan zorlukları, özellikle azınlık sınıfların işlenmesi ve eğitim fazları boyunca başarımın korunması bakımından açıkça ortaya koymaktadır. EWC uygulanmadığında, modelin makro recall ve F1 skorlarında, özellikle son eğitim fazlarında kayda değer düşüşler gözlemlenmektedir. Tablo 5.4.'te görüldüğü gibi en sık görülen sınıfların çıkarıldığı senaryoda, başlangıç aşamasındaki makro F1 skoru %56,33 iken, son fazda %51,43'e gerilemektedir. Bu azalma, sınıf dengesizliği içeren artımlı öğrenme senaryolarında, ek mekanizmalar olmaksızın bilginin korunmasının zor olduğunu göstermektedir. Bu zorluklara karşı EWC mekanizmasının dahil edilmesi, daha önce öğrenilen görevlere ait kritik parametrelerin korunmasını sağlayarak başarımdaki düşüşleri hafifletmektedir. Aynı senaryoda, EWC ile makro F1 skoru başlangıç fazında %57,13'ten, son fazda %59,30'a yükselmektedir. Bazı durumlarda %3'e varan bu iyileşmeler, EWC'nin katastrofik unutmayı azaltma ve sınıf frekansından bağımsız olarak genel performansı dengeleme konusundaki etkinliğini ortaya koymaktadır.

Tablo 5.4. Artımlı öğrenme performansı.

Senaryo	EWC	Makro Kesinlik	Makro Duyarlılık	Makro F1-Skor
İlk Eğitim (En Sık 10 Sınıf Hariç)	X	0.7077	0.5442	0.5633
Son Eğitim (En Sık 10 Sınıf Dahil)	X	0.6301	0.5143	0.5143
İlk Eğitim (En Sık 10 Sınıf Hariç)	✓	0.6866	0.5507	0.5713
Son Eğitim (En Sık 10 Sınıf Dahil)	✓	0.665	0.5787	0.593

5.3. Dinamik Çıktı Genişletme ile Hoeffding Tree DeneySEL Sonuçlar

Bu uygulamada kullanılan CICIOT2023 veri seti, çevrim içi öğrenme yapısına uygunluğu sağlamak ve gerçek zamanlı veri akışında dinamik sınıf keşfine olanak tanımak amacıyla özel bir ön işleme sürecinden geçirilmektedir. Uygulanan işlemler, geleneksel toplu öğrenme senaryolarından farklı olarak, akış tabanlı öğrenme çerçevesine entegre olacak şekilde tasarlanmaktadır.

1. Eksik veya hatalı değerlere sahip örnekler, veri akışı sırasında doğrudan filtrelenerek yalnızca eksiksiz örnekler modele sunulmaktadır. Bu yaklaşım, öğrenme döngüsünde kesintileri önleyerek veri bütünlüğünü korumaktadır.

2. Tüm öznitelik değerleri modele aktarılmadan önce açık biçimde ondalıklı sayı türüne dönüştürülmektedir. Bu işlem, kullanılan Hoeffding Tree algoritmasının (River kütüphanesi üzerinden) sayısal veri girişine olan gereksinimini karşılamak üzere gerçekleştirilmektedir. Ayrıca, bu süreçte herhangi bir Min-Max normalizasyonu veya ölçekleme uygulanmamaktadır. Böylece, verinin ham hâli korunmakta ve gerçek zamanlı veri akışının yapısına uygunluk sağlanmaktadır.
3. Her örneğe ait sınıf etiketi, doğrudan veri setindeki label alanından alınmaktadır. Deneyisel aşamalarda belirli sınıfların akıştan geçici olarak dışlanması amacıyla önceden tanımlanmış bir sınıf listesi kullanılmaktadır. Bu liste üzerinden sınıf görünürlüğü zamana bağlı olarak kurgulanmaktadır. Bu yöntem, sınıf artımlı öğrenme senaryosunun gerçekçi bir biçimde benzetim yapılmasını sağlamaktadır.
4. Sabit bir etiket kodlaması işlemi uygulanmamaktadır. Sınıf etiketleri öğrenme süreci boyunca dinamik olarak yönetilmektedir. Model, daha önce karşılaşmadığı yeni bir sınıf etiketi ile ilk defa karşılaştığında, dahili bir eşleme yapısı üzerinden bu sınıfa yeni bir sayısal kimlik atanmaktadır. Böylece çıktı katmanı dinamik biçimde genişleyebilmektedir.

Bu ön işleme yaklaşımı, modelin yalnızca statik veri yapılarıyla değil, aynı zamanda zaman içinde değişen sınıf uzaylarıyla da etkili biçimde başa çıkabilmesini mümkün kılmaktadır. Böylece, gerçek dünya koşullarına yakın, sürdürülebilir ve sürekli öğrenen bir system amaçlanmaktadır.

5.3.1. Analiz sonuçları

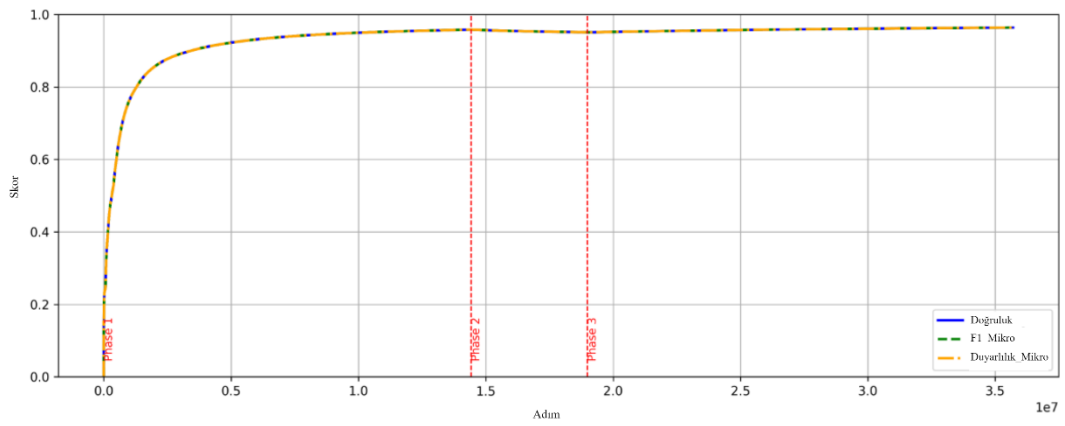
Bu tez çalışması kapsamında önerilen Dinamik Çıktı Genişletme (CIL-DOE) mekanizması sayesinde model, yeni sınıfları kademeli olarak öğrenme sürecine entegre ederken, daha önce öğrenilmiş sınıflara ilişkin performansını büyük ölçüde korumayı başarmaktadır. Bu durum, özellikle mikro ortalama metrikler (micro-averaged) açısından modelin başarılı olduğunu ortaya koymaktadır.

Birinci aşamada, modelin mevcut sınıf alt kümesine yönelik öğrenme süreci hızlı bir şekilde yakınsamaktadır. Doğruluk, mikro-F1 ve mikro-hatırlama (micro-recall) gibi mikro düzey göstergeler kısa sürede 0.8 seviyelerine ulaşmaktadır. Bu artış, modelin başlangıçta sunulan sınıfları etkin biçimde öğrenebildiğini göstermektedir.

İkinci ve üçüncü aşamalarda, veri akışına yeni sınıf grupları dahil edildikçe, modelin mikro düzeydeki performansı istikrarlı bir biçimde devam etmektedir. Her bir geçişin

ardından ilgili metrikler hızla dengeye oturmuş ve hatta zamanla iyileştiği gözlenmektedir. Böylece yeni sınıf bilgisinin başarıyla entegre edildiğini, aynı zamanda önceki bilgilerin önemli ölçüde korunabildiğini söylemek mümkündür.

Şekil 5.1.'de gösterildiği üzere, sınıfların kademeli olarak tanıtıldığı bu deneysel senaryoda model, genel anlamda güçlü ve tutarlı bir performans sergilemektedir. Bu bulgular, önerilen CIL-DOE yaklaşımının sınıf artımlı öğrenme bağlamında hem güncel hem de geçmiş bilgi arasında dengeli bir öğrenme sağlayabildiğini ortaya koymaktadır.



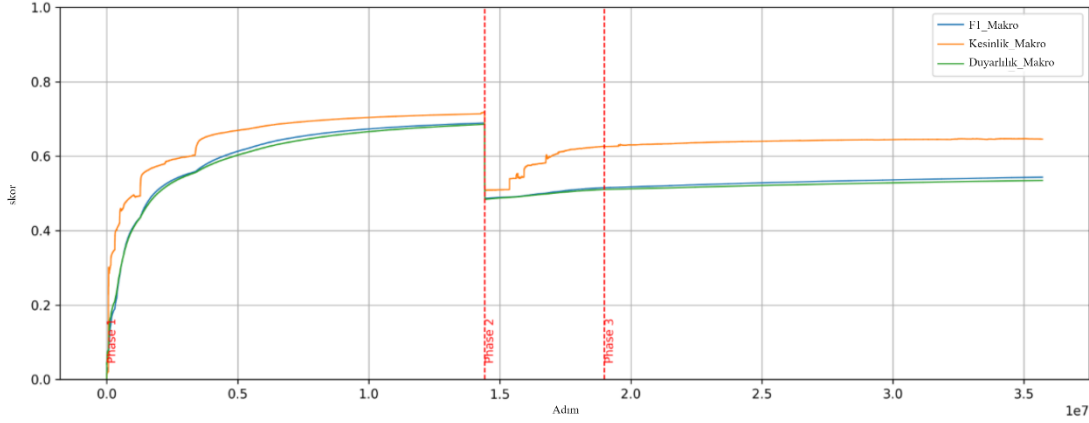
Şekil 5.2. Sınıf artımlı öğrenme kurulumundaki tüm eğitim aşamalarında zaman içindeki genel performans ölçümleri.

Makro ortalama performans göstergeleri, macro-F1, macro-precision ve macro-recall gibi metrikler aracılığıyla, modelin tüm sınıflar üzerindeki öğrenme etkinliğini dengeli bir şekilde değerlendirme olanağı sunmaktadır. Mikro metriklerde gözlemlenen ani dalgalanmalara kıyasla, makro metrikler daha yavaş ancak kararlı bir şekilde ilerleme göstermektedir.

İkinci ve üçüncü aşamalara geçiş sırasında, modelin henüz öğrenmediği yeni sınıflarla karşılaşması sonucunda, özellikle macro-F1 ve macro-recall değerlerinde belirgin düşüşler meydana gelmektedir. Bu azalmalar, modelin söz konusu sınıflara ilişkin bilgi eksikliğinden kaynaklanan düşük hatırlama oranları ve yüksek yanlış sınıflandırma oranları ile ilişkilidir.

Bununla birlikte, ilgili metriklerde zamanla gözlemlenen düzenli iyileşme, modelin yeni sınıfları genelleme becerisini aşamalı olarak geliştirdiğini göstermektedir. Bu eğilim, önerilen CIL-DOE stratejisinin sınıf dengesizliklerine sistematik şekilde yanıt verme yeteneğini ortaya koymaktadır.

Şekil 5.2.'de görülebileceği üzere, makro düzeydeki değerler zaman içerisinde artış göstermektedir. Nadir görülen sınıflar üzerindeki öğrenmenin daha zorlu olduğu ve modelin genel başarımını artırmak için daha uzun süreli eğitime ihtiyaç duyması nedeniyle, bu değerlerin hâlâ mikro metriklerin gerisinde kaldığı gözlemlenmektedir.



Şekil 5.3. Sınıf bazında performans.

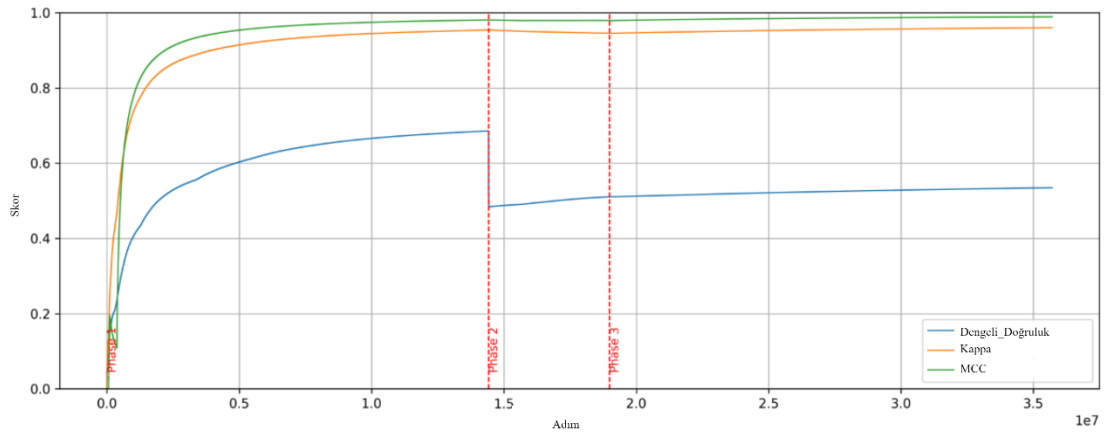
Eğitim süreci tamamlandığında, mikro ve makro metrikler arasındaki belirgin fark, daha geniş kapsamlı sistemsel bir soruna işaret etmektedir. Model genel doğruluk açısından neredeyse kusursuz bir performans sergilese de, bu başarı tüm sınıflar üzerinde homojen bir genelleme sağlamamaktadır. Mikro metrikler, daha sık karşılaşılan sınıflardaki başarıya ağırlık verirken, makro metrikler, tüm sınıfları eşit derecede değerlendirerek daha dengeli bir analiz sunmaktadır.

Bu doğrultuda, modelin azınlık sınıflar için yanlış pozitifleri minimize etme amacıyla daha temkinli tahminlerde bulunması, geri çağırma değerlerinin düşmesine neden olmaktadır. Şekil 5.2.'de görüldüğü üzere, makro düzeyde kesinlik, makro duyarlılık değerlerinin üzerinde seyretmektedir. Bu da modelin tahmin yaparken az görülen bir sınıfa atamak yerine sık görülen bir sınıfı tercih ettiğini, yani daha temkinli tahmin eğilimini ortaya koymaktadır. Aynı zamanda bu durum modelin öğrenme sürecinde kalıcı bir dengesizliğin varlığını da göstermektedir.

Bu bağlamda, CIL-DOE çerçevesi altında modelin sınıf dengesizliğine karşı dayanıklılığı, Cohen's Kappa ve Matthews Korelasyon Katsayısı (MCC) gibi istatistiksel güvenilirlik ölçütleri ile değerlendirilmektedir. Her iki metrik de, özellikle birinci fazda anlamlı bir yükseliş göstererek, sonraki fazlarda da yüksek seviyelerde istikrarını korumaktadır. Yeni sınıfların tanıtıldığı geçişlerde gözlemlenen küçük

düşüşler ise kısa sürede telafi edilmektedir. Bu durum, modelin yeni bilgilere hızlı adapte olabilme ve öğrenilmiş bilgiyi koruyabilme yetkinliğini ortaya koymaktadır.

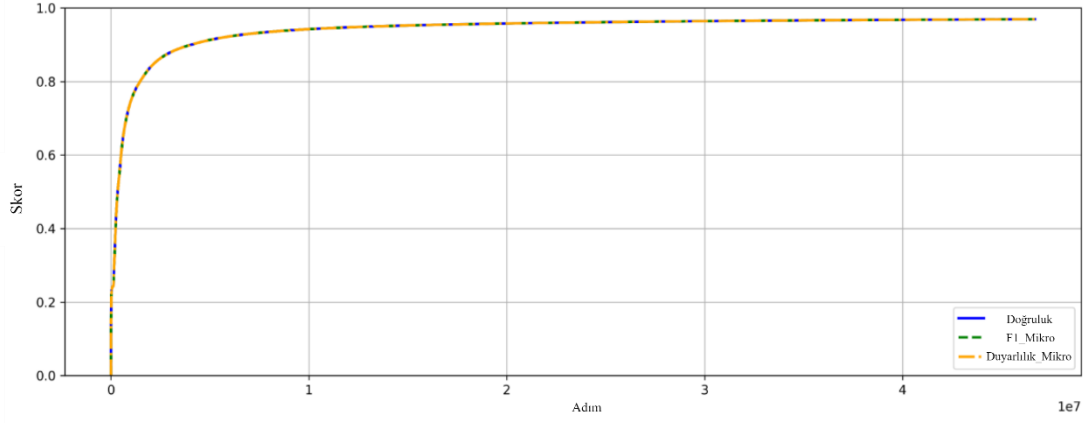
Üçüncü aşamanın sonunda, hem Kappa hem de MCC değerleri 0.95'in üzerine çıkarak, modelin tahminleri ile gerçek etiketler arasında güçlü bir korelasyon sağlandığını göstermektedir. Buna karşın, dengeli doğruluk değeri nispeten daha düşük seyretmekle birlikte, modelin tüm sınıflardaki ortalama hatırlama düzeyine dair tamamlayıcı değerlendirme sunmaktadır. Şekil 5.3., bu metriklerin zaman içerisindeki eğilimini ortaya koyarak, modelin artımlı öğrenme senaryolarındaki istatistiksel güvenilirliğini doğrulamaktadır.



Şekil 5.4. İstatistiksel güvenilirlik ölçümleri.

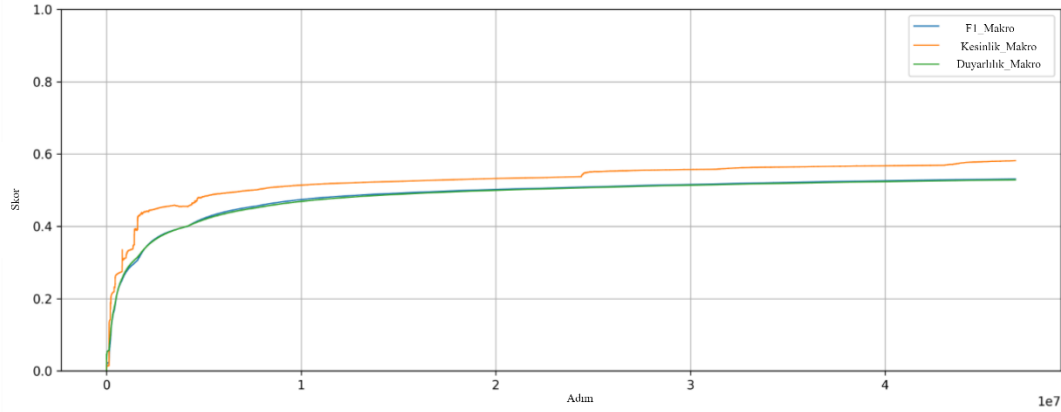
Aşamalı eğitim stratejisine ek olarak, veri setinin özgün zaman sırasını bozmadan uygulanan tamamlayıcı bir deney gerçekleştirilmektedir. Bu deneysel senaryoda, herhangi bir sınıf dışlama veya veri manipülasyonu yapılmaksızın, veri seti olduğu gibi modellenmektedir. Bu yaklaşım, sınıfların gerçek dünyada karşılaşılma sıklığına göre zamanla tanıtıldığı doğal bir veri akışı ortamını simüle etmektedir.

Bu aşamada, önerilen CIL-DOE çerçevesi, yeni sınıflarla karşılaştığında bile yüksek düzeyde genelleme yeteneği sergilemektedir. Şekil 5.4'te gösterildiği gibi, doğruluk ve mikro-F1 gibi mikro ortalamalı metrikler, kısa sürede 0.95'in üzerine çıkmış ve bu seviyelerde kararlılığını korumaktadır. Bu sonuçlar, modelin farklı sınıf dağılımlarına etkili biçimde uyum sağlama kapasitesini ve doğal veri akışına uygun biçimde çalışabildiğini göstermektedir.



Şekil 5.5. Tam akış sırasında genel performans ölçümleri.

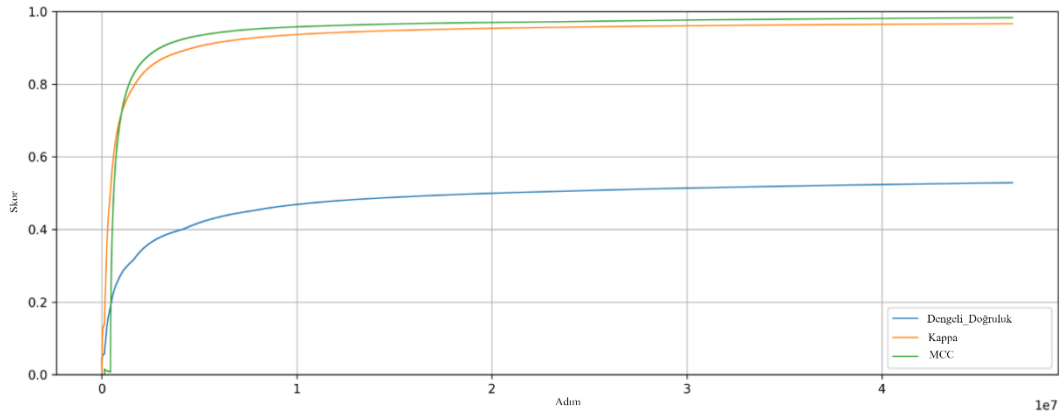
Buna karşılık, Şekil 5.5, makro ortalama metrikler olan makro-F1, makro-kesinlik ve makro-duyarlılık değerlerinin zaman içerisinde kademeli olarak iyileştiğini göstermektedir. Ancak bu metrikler ile mikro düzey göstergeler arasında belirgin bir fark gözlemlenmeye devam etmektedir. Bu durum, özellikle veri akışında geç tanıtılan veya düşük temsile sahip sınıflar için modelin tutarlı bir öğrenme gerçekleştirmekte zorlandığını ortaya koymaktadır. Bu durum, önerilen yöntemin genel başarımına rağmen, tüm sınıflarda eşit düzeyde genelleme sağlamanın hâlen önemli bir zorluk olarak varlığını sürdürdüğünü vurgulamaktadır.



Şekil 5.6. Tam akış sırasında makro denge ölçümleri.

Ayrıca, Şekil 5.6., modelin bu süreç boyunca gösterdiği istatistiksel güvenilirliği vurgulamaktadır. Dengeli doğruluk, Cohen's Kappa ve Matthews Korelasyon Katsayısı gibi ölçütler, akışın tamamı boyunca yüksek seviyelerde istikrarını korumaktadır. Bu değerler, önerilen CIL-DOE yaklaşımının, gerçekçi veri akış

senaryoları altında da dayanıklı, güvenilir ve uyarlanabilir bir performans sergilediğini göstermektedir.



Şekil 5.7. İstatistiksel güvenilirlik ölçümleri.

5.4. Modellerin Karşılaştırmalı Analizi

EWC destekli paylaşımlı kodlayıcı ile dinamik çıktı genişletmeli Hoeffding Tree (CIL-DOE) modelleri, yapılandırılmış sınıf artımlı öğrenme senaryoları ve gerçek zamanlı tam akış koşulları altında çok boyutlu performans metrikleri üzerinden karşılaştırılmaktadır. Analizlere göre doğruluk, makro-F1 skoru, elde tutma oranı (RS) gibi temel ölçütlere dayanmaktadır.

EWC modeli, özellikle azınlık sınıfların yer aldığı ve fazlar arası bilgi geçişinin gerekli olduğu durumlarda sınıf dengesizliğine karşı daha dayanıklı bir yapı sergilemekte ve yüksek makro-F1 puanlarıyla öne çıkmaktadır. Öte yandan CIL-DOE modeli, karar ağacı tabanlı hızlı bölünme yapısı sayesinde yeni sınıflara adaptasyonda daha başarılı olmaktadır. Düşük çıkarım gecikmesi ve hafif kaynak gereksinimleri sayesinde gerçek zamanlı uygulamalarda avantaj sunmaktadır.

Elde tutma skorları incelendiğinde, EWC modelinin tüm fazlar boyunca daha yüksek değerler sunduğu ve önceki bilgileri daha etkin koruduğu gözlemlenmektedir. Buna karşın CIL-DOE modeli, ilk örneklerde hızlı öğrenme eğilimiyle daha yüksek adaptasyon skorlarına ulaşmakta, ancak bu yüksek plastisite düzeyi eski bilgilerin daha hızlı unutulmasına ve özellikle nadir sınıflarda performans kaybına yol açmaktadır.

Kaynak kullanımı açısından CIL-DOE, daha düşük bellek ve işlemci tüketimiyle öne çıkarken, EWC modeli daha fazla hesaplama gücü gerektirmektedir. Ancak bu durum, EWC'nin zengin temsil kapasitesi ve daha stabil uzun vadeli performans sağlamasına

imkân tanımaktadır. Kavramsal sapma analizlerinde EWC modeli daha sabit bir öğrenme eğrisi sergilerken, CIL-DOE modeli bu sapmalara daha hızlı uyum sağlamaktadır. Ancak yerel güncellemelerin etkisiyle eski sınıflarda yanlış sınıflandırma eğiliminde bulunmaktadır.

Elde edilen bulgular, her iki modelin sınıf artımlı öğrenme bağlamında farklı güçlü yönleri sahip olduğunu ortaya koymaktadır. EWC modeli, yüksek doğruluk ve makro-F1 skorları ile özellikle dengesiz ve çok aşamalı ortamlarda üstün performans sergilerken; CIL-DOE modeli, yeni sınıflara hızlı adaptasyonu, düşük çıkarım gecikmesi ve daha az kaynak kullanımı ile öne çıkmaktadır. Bu farklılıklar, modellerin dağıtım ortamlarına göre tercih edilmesinde belirleyici rol oynayabilmektedir.

Karşılaştırmaya ilişkin özet performans metrikleri Tablo 5.5’te sunulmuştur. Bu tablo, her iki modelin doğruluk, F1 puanı, öğrenilen bilgileri saklama yetisi, yeni sınıflara adaptasyon hızı, sistem kaynaklarına duyarlılık ve azınlık sınıflardaki başarımlar gibi temel ölçütler açısından nasıl farklılaştığını açıkça göstermektedir.

Tablo 5.5. Karşılaştırmalı performans özeti.

Ölçüt	EWC	DOE
Ortalama Doğruluk	%78,6	%76,2
Ortalama Makro-F1	0.738	0.702
Hatırlama (RS)	%75,9	%66,8
Adaptasyon (AS)	0.69	0.83
Bellek Kullanımı	Yüksek	Düşük
CPU Kullanımı	Yüksek	Düşük
Azınlık Sınıf Başarısı	Yüksek	Düşük-Orta

6. SONUÇ VE TARTIŞMA

6.1. EWC Destekli Ortak Enkoder Mimarisinin Değerlendirmesi

Bu çalışma, sınıf dengesizliği ve katastrofik unutma gibi artımlı öğrenmeye özgü iç içe geçmiş zorlukları ele alarak EWC yönteminin bu problemlerin azaltılmasındaki rolünü incelemektedir. Elde edilen bulgular, EWC'nin farklı sınıf dağılımları altında modelin istikrarını ve bilgi koruma yeteneğini önemli ölçüde artırdığını ortaya koymaktadır. En az görülen 10 sınıfın çıkarıldığı senaryoda, makro F1 skoru EWC uygulanmadan önce %56,09 iken, EWC ile %59,30'a ulaşmış, bu da artımlı öğrenme sürecinde yaşanan performans kayıplarının azaltılmasında EWC'nin etkinliğini açıkça göstermektedir.

Paylaşımlı bir kodlayıcı ile göreve özgü çıktı katmanlarının birleştirildiği hibrit mimari, öznitelik temsilleri üretme ve göreve uyum sağlama açısından başarılı sonuçlar vermektedir. Bu yaklaşım, özellikle IoT sistemleri gibi dinamik ve dengesiz veri dağılımlarının yaygın olduğu gerçek dünya uygulamalarında ölçeklenebilir ve pratik bir çözüm sunmaktadır. Ancak tüm bu gelişmelere rağmen, nadir sınıflarda anlamlı performans artışı sağlama ve orta sıklıkta görülen sınıflarda model kararlılığını sürdürebilme gibi bazı temel zorluklar hâlâ varlığını sürdürmekte olup, şiddetli sınıf dengesizliğinin karmaşıklığını göstermektedir.

6.2. CIL-DOE Yaklaşımının Genel Değerlendirmesi

Bu çalışma kapsamında önerilen CIL-DOE çerçevesi, Hoeffding Tree sınıflandırıcılarının akıntı hâlindeki verilerde karşılaşılan yeni sınıf etiketlerine dinamik biçimde uyum sağlayabilmesini mümkün kılmaktadır. Model, önceden tanımlanmış sınıf kümesine ihtiyaç duymadan ve yeniden eğitim süreci gerektirmeksizin çıktı katmanını gerçek zamanlı olarak genişletebilmektedir. Bu sayede sürekli öğrenme ortamlarında yüksek düzeyde esneklik ve uyarlanabilirlik sunmaktadır.

Yürütülen deneysel çalışmalar, CIL-DOE yaklaşımının mikro düzey doğruluk ve F1 skorlarında yüksek performans sergilemektedir. Yeni sınıflar tanıtıldıkça önceden

edinilen bilgileri koruyarak öğrenme sürecini sürdürebildiğini göstermektedir. Model, katastrofik unutmayı büyük ölçüde engellemekte ve aynı zamanda sınıf dengesizliğine karşı da yüksek direnç göstermektedir. Bu durum, özellikle Cohen's Kappa ve Matthews Korelasyon Katsayısı (MCC) gibi istatistiksel güvenilirlik ölçütlerinin tüm aşamalar boyunca yüksek seviyelerde seyretmesiyle desteklenmektedir. Elde edilen bulgular, CIL-DOE çerçevesinin siber güvenlik, anomali tespiti ve benzeri alanlarda, değişken veri dağılımlarına sahip gerçek zamanlı uygulamalar için uygulanabilir ve pratik bir çözüm sunduğunu ortaya koymaktadır.

CIL-DOE'nin en önemli avantajlarından biri, mimarının sade ve düşük maliyetli olmasıdır. Bu yapı, dış bellek (external memory) ya da tekrar örnekleme (rehearsal) gibi ek destek mekanizmalarına ihtiyaç duymadan, çıktı katmanını dinamik şekilde genişletebilme yeteneğine sahiptir. Bu durum, özellikle çevrim içi ve kaynak kısıtlı ortamlarda uygulanabilirliğini artırmaktadır. Ek olarak, modelin daha önce karşılaşılmamış sınıflar karşısında genelleme yapabilme ve yeniden tanıtılan sınıflarda performansını hızlı şekilde geri kazanabilme yetkinliği, sistemin kararlılık ile esneklik arasında başarılı bir denge kurduğunu göstermektedir.

6.3. Karşılaştırmalı Değerlendirme

Yapılandırılmış sınıf artımlı öğrenme senaryoları ve gerçek zamanlı tam akış koşulları altında yapılan karşılaştırmalar, EWC ve CIL-DOE modellerinin farklı güçlü yönleri sahip olduğunu ortaya koymaktadır. EWC destekli model, özellikle dengesiz veri yapılarında ve önceki bilgilerin korunmasının kritik olduğu senaryolarda, elde ettiği yüksek doğruluk ve makro-F1 skorları ile öne çıkmaktadır. Aşamalar arası bilgi tutarlılığı ve azınlık sınıflardaki başarıyla, öğrenilen bilgilerin uzun süre korunması açısından avantaj sağlamaktadır.

Öte yandan, CIL-DOE modeli, yeni sınıflara yönelik hızlı adaptasyon yeteneği, düşük çıkarım gecikmesi ve yüksek kaynak verimliliği ile gerçek zamanlı öğrenme bağlamında dikkat çekmektedir. Yüksek uyum kabiliyetine rağmen, zaman içinde önceden öğrenilen bilgilerin zayıflaması ve nadir sınıflardaki performans kaybı, yöntemin bazı sınırlı yönlerini oluşturmaktadır. Bu sonuçlar, her iki modelin de farklı uygulama senaryolarına göre optimize edilebilecek özgün katkılar sunduğunu göstermektedir.

KAYNAKLAR

- Ade, R. R., ve Deshmukh, P. R. (2013). Methods for incremental learning: A survey. *International Journal of Data Mining & Knowledge Management Process (IJDKP)*, 3(4), 119–132. <https://doi.org/10.5121/ijdkp.2013.3408>.
- Akintola, A. G., Balogun, A. O., Mojeed, H. A., Usman-Hamza, F. E., Salihu, S. A., Adewole, K. S., Balogun, G. B., ve Sadiku, P. O. (2022). Performance analysis of machine learning methods with class imbalance problem in Android malware detection. *International Journal of Interactive Mobile Technologies (iJIM)*, 16(10), 140–154. <https://doi.org/10.3991/ijim.v16i10.29687>.
- Baddeley, A. D., ve Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory (Vol. 8, ss. 47–89)*. [https://doi.org/10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1).
- Batra, H., & Clark, R. (2024). EVCL: Elastic variational continual learning with weight consolidation. <https://doi.org/10.48550/arXiv.2406.15972>.
- Baysal, E., & Bayılmış, C. (2025). Overcoming Class Imbalance in Incremental Learning Using an Elastic Weight Consolidation-Assisted Common Encoder Approach. *Mathematics*, 13(11), 1887. <https://doi.org/10.3390/math13111887>.
- Baysal, E., ve Bayılmış, C. (2019). Machine learning with distributed data management and process architecture. In 4th International Conference on Computer Science and Engineering (UBMK) (ss. 53–58). IEEE. <https://doi.org/10.1109/UBMK.2019.8907073>.
- Baysal, E., ve Bayılmış, C. (2022). Incremental machine learning: Incremental classification. In 7th International Conference on Computer Science and Engineering (UBMK) (ss. 268–273). IEEE. <https://doi.org/10.1109/UBMK55850.2022.9919487>.
- Bottou, L. (1998). Online learning and stochastic approximations. In D. Saad (Ed.), *Online Learning in Neural Networks* (ss. 9–42). Cambridge University Press.
- Buda, M., Maki, A., ve Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks*, 106, 249–259. <https://doi.org/10.1016/j.neunet.2018.07.011>.
- Caruana, R. (1997). Multitask learning. *Machine Learning*, 28, 41–75. <https://doi.org/10.1023/A:1007379606734>.
- Castro, F. M., Marín, J., Guil, N., Moreno, G., ve Martinez, P. (2018). End-to-end incremental learning. In *Proceedings of the European Conference on Computer Vision (ECCV)* (ss. 233–248). https://doi.org/10.1007/978-3-030-01258-8_15.
- Cauwenberghs, G., ve Poggio, T. (2000). Incremental and decremental support vector machine learning. In *Advances in Neural Information Processing Systems (NIPS)* (ss. 409–415).

- Chawla, N. V., Bowyer, K. W., Hall, L. O., ve Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/jair.953>.
- Chen, Z., ve Liu, B. (2018). Lifelong machine learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 12(3), 1–207. <https://doi.org/10.1007/978-3-031-01581-6>.
- De Lange, M., Aljundi, R., Masana, M., Parisot, S., Jia, X., Leonardis, A., Slabaugh, G., ve Tuytelaars, T. (2021). A continual learning survey: Defying forgetting in classification tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7), 3366–3385. <https://doi.org/10.1109/TPAMI.2021.3057446>.
- Domingos, P., ve Hulten, G. (2000). Mining high-speed data streams. In Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (ss. 71–80). <https://doi.org/10.1145/347090.347107>.
- Douillard, A., Ramé, A., Couairon, G., ve Cord, M. (2022). DyTox: Transformers for continual learning with dynamic token expansion. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (ss. 9285–9295). <https://doi.org/10.1109/CVPR52688.2022.00907>.
- Elsayed, M., ve Mahmood, A. R. (2023). Utility-based perturbed gradient descent: An optimizer for continual learning. <https://doi.org/10.48550/arXiv.2302.03281>
- Finn, C., Abbeel, P., ve Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. In Proceedings of the 34th International Conference on Machine Learning (ICML) (ss. 1126–1135).
- Firdausanti, N. A., Mendonça, I., ve Aritsugi, M. (2024). Noise-free sampling with majority framework for an imbalanced classification problem. *Knowledge and Information Systems*, 66(7), 4011–4042. <https://doi.org/10.1007/s10115-024-02079-6>.
- Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., ve Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys (CSUR)*, 46(4), 44. <https://doi.org/10.1145/2523813>.
- Gandomi, A., ve Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>.
- Goodfellow, I. J., Mirza, M., Xiao, D., Courville, A., ve Bengio, Y. (2015). An empirical investigation of catastrophic forgetting in gradient-based neural networks. <https://arxiv.org/abs/1312.6211>.
- He, H., Chen, S., Li, K., ve Xu, X. (2011). Incremental learning from stream data. *IEEE Transactions on Neural Networks*, 22(12), 1901–1914. <https://doi.org/10.1109/TNN.2011.2171713>.
- He, H., ve Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263–1284. <https://doi.org/10.1109/TKDE.2008.239>.

- Hess, T., Verwimp, E., van de Ven, G. M., ve Tuytelaars, T. (2024). Knowledge accumulation in continually learned representations and the issue of feature forgetting. <https://arxiv.org/abs/2304.00933>.
- Hoi, S. C. H., Sahoo, D., Lu, J., ve Zhao, P. (2021). Online learning: A comprehensive survey. *Neurocomputing*, 459, 249–289. <https://doi.org/10.1016/j.neucom.2021.04.112>.
- Hou, R., Chen, C., Guo, H., ve Change Loy, C. (2019). Learning lightweight lane detection CNNs by self attention distillation. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 1013–1021. <https://doi.org/10.1109/ICCV.2019.00110>.
- Junsawang, P., Phimoltare, S., ve Lursinsap, C. (2019). Streaming chunk incremental learning for class-wise data stream classification with fast learning speed and low structural complexity. *PLoS ONE*, 14(9). <https://doi.org/10.1371/journal.pone.0220624>.
- Katal, A., Wazid, M., ve Goudar, R. H. (2013). Big data: Issues, challenges, tools and good practices. 2013 Sixth International Conference on Contemporary Computing (IC3), 404–409. <https://doi.org/10.1109/IC3.2013.6612229>.
- Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwińska, A., Hassabis, D., Clopath, C., Kumaran, D., ve Hadsell, R. (2017). Overcoming catastrophic forgetting in neural networks. Proceedings of the National Academy of Sciences, 114(13), 3521–3526. <https://doi.org/10.1073/pnas.1611835114>.
- Lesort, T., Lomonaco, V., Stoian, A., Maltoni, D., Filliat, D., ve Díaz-Rodríguez, N. (2020). Continual learning for robotics: Definition, framework, learning strategies, opportunities and challenges. *Information Fusion*, 58, 52–68. <https://doi.org/10.1016/j.inffus.2019.12.004>.
- Li, Z., Li, Z., ve Ren, Y. (2023). The impact of industrial internet and the digital economy on the management and development of manufacturing information systems triggering digitization as IoT and artificial intelligence. *Journal of Information Systems Engineering & Management*. <https://doi.org/10.55267/iadt.07.14038>.
- Li, Z., ve Hoiem, D. (2017). Learning without forgetting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(12), 2935–2947. <https://doi.org/10.1109/TPAMI.2017.2773081>.
- Lopez-Ledezma, M., ve Velarde, G. (2025). Cyber security data science: Machine learning methods and their performance on imbalanced datasets. In *Digital Management and Artificial Intelligence* (ss. 512–527). Springer. https://doi.org/10.1007/978-3-031-88052-0_45.
- Lopez-Paz, D., ve Ranzato, M. (2022). Gradient episodic memory for continual learning. *Neural Information Processing Systems (NeurIPS)*, 31, 6467–6476. <https://arxiv.org/abs/1706.08840>.
- Losing, V., Hammer, B., ve Wersing, H. (2018). Incremental on-line learning: A review and comparison of state of the art algorithms. *Neurocomputing*, 275, 1261–1274. <https://doi.org/10.1016/j.neucom.2017.06.084>.

- Lu, J., Liu, A., Dong, F., Gu, F., Gama, J., ve Zhang, G. (2020). Learning under concept drift: A review. *IEEE Transactions on Knowledge and Data Engineering*, 31(12), 2346–2363. <https://doi.org/10.1109/TKDE.2018.2876857>.
- Lyu, F., Liu, D., Zhao, L., Zhang, Z., Shang, F., Hu, F., Feng, W., ve Wang, L. (2024). Overcoming domain drift in online continual learning. <https://doi.org/10.48550/arXiv.2405.09133>.
- Maltoni, D., ve Lomonaco, V. (2019). Continuous learning in single-incremental-task scenarios. *Neural Networks*, 116, 56–73. <https://doi.org/10.1016/j.neunet.2019.03.010>.
- Masana, M., Liu, X., Twardowski, B., Menta, M., Bagdanov, A. D., ve van de Weijer, J. (2022). Class-incremental learning: Survey and performance evaluation on image classification. <https://arxiv.org/abs/2010.15277>.
- McCloskey, M., ve Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. *Psychology of Learning and Motivation*, 24, 109–165. [https://doi.org/10.1016/S0079-7421\(08\)60536-8](https://doi.org/10.1016/S0079-7421(08)60536-8).
- Montiel, J., Read, J., Bifet, A., ve Abdesslem, T. (2018). Scikit-multiflow: A multi-output streaming framework. *Journal of Machine Learning Research*, 19, 1–5. <http://jmlr.org/papers/v19/18-251.html>.
- Neto, E. C. P., Dadkhah, S., Ferreira, R., Zohourian, A., Lu, R., ve Ghorbani, A. A. (2023). CICIOT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment. *Sensors*, 23(13), 5941. <https://doi.org/10.3390/s23135941>.
- Pan, S. J., ve Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359. <https://doi.org/10.1109/TKDE.2009.191>.
- Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., ve Wermter, S. (2019). Continual lifelong learning with neural networks: A review. *Neural Networks*, 113, 54–71. <https://doi.org/10.1016/j.neunet.2019.01.012>.
- Pisupati, S., ve Niv, Y. (2022). The challenges of lifelong learning in biological and artificial systems. *Trends in Cognitive Sciences*, 26(12), 1051–1053. <https://doi.org/10.1016/j.tics.2022.09.022>.
- Rahman, M. G., ve Islam, M. Z. (2022). Adaptive decision forest: An incremental machine learning framework. *Pattern Recognition*, 122, 108345. <https://doi.org/10.1016/j.patcog.2021.108345>.
- Read, J., ve Žliobaitė, I. (2022). Learning from data streams: An overview and update. <https://arxiv.org/abs/2212.14720>.
- Rebuffi, S.-A., Gowal, S., Calian, D. A., Stimberg, F., Wiles, O., ve Mann, T. (2021). Fixing data augmentation to improve adversarial robustness. <https://doi.org/10.48550/arXiv.2103.01946>.
- Rebuffi, S.-A., Kolesnikov, A., Sperl, G., ve Lampert, C. H. (2017). iCaRL: Incremental classifier and representation learning. <https://arxiv.org/abs/1611.07725>.

- Rusu, A. A., Rabinowitz, N. C., Desjardins, G., Soyer, H., Kirkpatrick, J., Kavukcuoglu, K., Pascanu, R., ve Hadsell, R. (2022). Progressive neural networks. <https://arxiv.org/abs/1606.04671>.
- Schlimmer, J. C., ve Fisher, D. (1986). A case study of incremental concept induction. In *Proceedings of the Fifth National Conference on Artificial Intelligence* (ss. 496–501).
- Seol, D. H., Choi, J. E., Kim, C. Y., ve Hong, S. J. (2023). Alleviating class-imbalance data of semiconductor equipment anomaly detection study. *Electronics*, 12(3), Article 585. <https://doi.org/10.3390/electronics12030585>.
- Shin, H., Lee, J. K., Kim, J., ve Kim, J. (2017). Continual learning with deep generative replay. In *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)* (ss. 2990–2999).
- Skierś, P., ve Deja, K. (2024). Joint diffusion models in continual learning. <https://doi.org/10.48550/arXiv.2411.08224>.
- Suárez-Cetrulo, A. L., Quintana, D., ve Cervantes, A. (2023). A survey on machine learning for recurring concept drifting data streams. *Expert Systems with Applications*, 213, 118934. <https://doi.org/10.1016/j.eswa.2022.118934>.
- Sutton, R. S., ve Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press. <http://incompleteideas.net/book/the-book-2nd.html>.
- Syavasya, C. V. S. R., ve Muddana, A. L. (2020). A review on incremental machine learning methods, applications and open challenges. *International Journal of Advanced Research in Engineering and Technology (IJARET)*, 11(10), 919–928. <https://doi.org/10.34218/IJARET.11.10.2020.091>.
- Thrun, S. (1995). Lifelong learning: A case study. Carnegie Mellon University, School of Computer Science.
- Utgoff, P. E. (1989). Incremental induction of decision trees. *Machine Learning*, 4, 161–186. <https://doi.org/10.1023/A:1022699900025>.
- van de Ven, G. M., Tuytelaars, T., ve Tolias, A. S. (2022). Three types of incremental learning. *Nature Machine Intelligence*, 4, 1185–1197. <https://doi.org/10.1038/s42256-022-00568-3>.
- van den Goorbergh, R., van Smeden, M., Timmerman, D., ve Van Calster, B. (2022). The harm of class imbalance corrections for risk prediction models: Illustration and simulation using logistic regression. *Journal of the American Medical Informatics Association*, 29(9), 1525–1534. <https://doi.org/10.1093/jamia/ocac093>.
- Vapnik, V. N. (2000). The nature of statistical learning theory. Springer.
- Wang, L., Zhang, X., Su, H., ve Zhu, J. (2024). A comprehensive survey of continual learning: Theory, method and application. <https://doi.org/10.48550/arXiv.2302.00487>.
- Wu, Y., Chen, Y., Wang, L., Ye, Y., Liu, Z., Guo, Y., ve Fu, Y. (2019). Large scale incremental learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (ss. 253–261). <https://doi.org/10.1109/ICCV.2019.00032>.

- Yan, S., Xie, J., ve He, X. (2021). DER: Dynamically Expandable Representation for Class Incremental Learning. Dalam Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (hlm. 10419-10428). <https://doi.org/10.48550/arXiv.2103.16788>.
- Yang, P., Wang, X., Zhang, R., Wang, C., Oliehoek, F. A., ve Kober, J. (2025). Task-free lifelong robot learning with retrieval-based weighted local adaptation. <https://arxiv.org/abs/2410.02995>.
- Zhang, S.-S., Liu, J.-W., ve Zuo, X. (2021). Adaptive online incremental learning for evolving data streams. *Applied Soft Computing*, 105, 107255. <https://doi.org/10.1016/j.asoc.2021.107255>.
- Zhou, D.-W., Wang, Q.-W., Qi, Z.-H., Ye, H.-J., Zhan, D.-C., ve Liu, Z. (2024). Class-incremental learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12), 9851–9878. <https://doi.org/10.1109/TPAMI.2024.3429383>.
- Zhu, F., Ma, S., Cheng, Z., Zhang, X.-Y., Zhang, Z., Tao, D., ve Liu, C.-L. (2025). Open-world machine learning: A systematic review and future directions. <https://doi.org/10.48550/arXiv.2403.01759>.
- Zhu, Z., Lin, K., Jain, A. K., ve Zhou, J. (2023). Transfer learning in deep reinforcement learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(11), 13344–13362. <https://doi.org/10.1109/TPAMI.2023.3292075>.
- Zola, F., Seguro-Gil, L., Bruse, J. L., Galar, M., ve Orduna-Urrutia, R. (2022). Attacking Bitcoin anonymity: Generative adversarial networks for improving Bitcoin entity classification. *Applied Intelligence*, 52(15), 17289–17314. <https://doi.org/10.1007/s10489-022-03378-7>.

ÖZGEÇMİŞ

Ad-Soyad : Engin BAYSAL

ÖĞRENİM DURUMU:

- **Yüksek Lisans** : 2019, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar ve Bilişim Mühendisliği
- **Lisans** : 2016, Sakarya Üniversitesi, Bilgisayar ve Bilişim Bilimleri Fakültesi, Bilgisayar Mühendisliği
- **Lisans** : 2008, Marmara Üniversitesi, Teknik Eğitim Fakültesi, Bilgisayar ve Kontrol Öğretmenliği

MESLEKİ DENEYİM VE ÖDÜLLER:

- 2023 yılından bu yana İstanbul Teknik Üniversitesinde Öğretim görevlisi olarak çalışmaya devam etmekte
- 2021-2023 yılları arasında Ardahan Üniversitesinde Öğretim Görevlisi olarak çalıştı
- 2019-2021 yılları arasında Erzincan Binali Yıldırım Üniversitesinde Öğretim Görevlisi olarak çalıştı
- 2011-2019 yılları arasında İstanbul Gedik Üniversitesinde Öğretim Görevlisi olarak çalıştı.

TEZDEN TÜRETİLEN ESERLER:

- Baysal, E.; Bayılmış, C. Overcoming Class Imbalance in Incremental Learning Using an Elastic Weight Consolidation-Assisted Common Encoder Approach. *Mathematics* 2025, 13, 1887. <https://doi.org/10.3390/math13111887>.
- Baysal E.; Bayılmış C. (2022, 14-16, Eylül). Incremental Machine Learning: Incremental Classification. *7th International Conference on Computer Science and Engineering (UBMK)*, Diyarbakır, Turkey.