**İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

# MIL-STD 1553 TABANLI SİSTEMLER İÇİN YENİ BİR SALDIRI TESPİTİ YAKLAŞIMI

**YÜKSEK LİSANS TEZİ**

**Yunus Emre ÇİLOĞLU**

**Bilgisayar Mühendisliği Anabilim Dalı**

**Bilgisayar Mühendisliği Programı**

**OCAK 2024**

**İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

# MIL-STD 1553 TABANLI SİSTEMLER İÇİN YENİ BİR SALDIRI TESPİTİ YAKLAŞIMI

**YÜKSEK LİSANS TEZİ**

**Yunus Emre ÇİLOĞLU**
**(504191588)**

**Bilgisayar Mühendisliği Anabilim Dalı**

**Bilgisayar Mühendisliği Programı**

**Tez Danışmanı: Doç. Dr. Şerif BAHTİYAR**

**OCAK 2024**

**ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL**

# A NEW INTRUSION DETECTION APPROACH FOR MIL-STD 1553 BASED SYSTEMS

**M.Sc. THESIS**

**Yunus Emre ÇİLOĞLU**
**(504191588)**

**Department of Computer Engineering**

**Computer Engineering Programme**

**Thesis Advisor: Doç. Dr. Şerif BAHTİYAR**

**OCAK 2024**

İTÜ, Lisansüstü Eğitim Enstitüsü'nün 504191588 numaralı Yüksek Lisans Öğrencisi Yunus Emre ÇİLOĞLU, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı "MIL-STD 1553 TABANLI SİSTEMLER İÇİN YENİ BİR SALDIRI TESPİTİ YAKLAŞIMI" başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

| | | |
|---|---|---|
| **Tez Danışmanı :** | **Doç. Dr. Şerif BAHTİYAR**<br>İstanbul Teknik Üniversitesi | ..... |
| **Jüri Üyeleri :** | **Doç. Dr. Şerif BAHTİYAR**<br>Istanbul Teknik Üniversitesi | ..... |
| | **Doç. Dr. Yusuf YASLAN**<br>Istanbul Teknik Üniversitesi | ..... |
| | **Doç. Dr. Muhammed Ali AYDIN**<br>Istanbul Üniversitesi | ..... |

**Teslim Tarihi :** **4 OCAK 2024**
**Savunma Tarihi :** **26 OCAK 2024**

***Türk Havacılığına,***

## ÖNSÖZ

Bu çalışmada askeri havacılığın önemli sistemlerinden biri olan Mil-Std 1553 veriyolunun güvenliği ele alınmıştır.

Bu çalışmanın hayata geçirilmesi süresince bilgi ve tecrübelerinden bolca faydalandığım, her sorunuma büyük bir hassasiyetle yaklaşan, yardımcı olan ve yol gösteren değerli tez danışmanım Doç. Dr. Şerif BAHTİYAR'a sonsuz teşekkür ederim.

Bu çalışmanın yapılabilmesi için izin veren, bu çalışmada verilen bilgileri edinmemi sağlayan ve yüksek lisans eğitimim için destek veren ASELSAN'a teşekkür ederim.

Desteklerini her zaman hissettiğim kıymetli arkadaşlarım Melih, Osman ve Şenol'a teşekkür ederim.

Bütün eğitim hayatım boyunca benim yanımda duran, destekleyen, yol gösteren, cesaret veren, benden ümitlerini hiç kesmeyen sevgili anneme, babama, ablama ve abime sonsuz teşekkürü borç bilirim.

Bu tez süresince bana ilham veren, yolumu aydınlatan, devam etme gücü veren sevgili aşkım Damla'ya şükranlarımı sunarım.

OCAK 2024

Yunus Emre ÇİLOĞLU
(Bilgisayar Mühendisi)

**İÇİNDEKİLER**

## KISALTMALAR

**RT** : Remote Terminal
**BC** : Bus Controller
**BM** : Bus Monitor
**RF** : Random Forest
**DT** : Decision Tree
**KNN** : K-Nearest Neighbors
**GNB** : Gaussian Naive Bayes
**SGD** : Stochastic Gradient Descent
**LR** : Logistic Regresssion
**BERT** : Bidirectional Encoder Representations from Transformers
**NLP** : Natural Language Processing
**DoS** : Denial-of-Service
**TP** : True Positive
**TN** : True Negative
**FP** : False Positive
**FN** : False Negative
**VUHF** : Very-Ultra High Frequency
**EW** : Electronic Warfare
**IFF** : Identification Friend or Foe
**HF** : High Frequency
**RNN** : Recurrent Neural Networks
**CNN** : Convolutional Neural Networks

## SEMBOLLER

| | |
|---|---|
| **F1** | **:** F1 Skoru |
| $\mathbf{F1_{weighted}}$ | **:** Weighted Average F1 Skoru |
| $\mathbf{F1_{macro}}$ | **:** Macro Average F1 Skoru |
| $\mathbf{F1\ Skor_i}$ | **:** i'inci sınıfın F1 skoru |
| **Sınıf** $\mathbf{Ağırlığı_i}$ | **:** i'inci sınıfın ağırlığı |

## ÇİZELGE LİSTESİ

**Sayfa**

## ŞEKİL LİSTESİ

**Sayfa**

# MIL-STD 1553 TABANLI SİSTEMLER İÇİN YENİ BİR SALDIRI TESPİTİ YAKLAŞIMI

## ÖZET

Havacılık, insanlığın gelişimi ve teknolojik ilerlemesi açısından tarihi bir öneme sahiptir. İlk uçuş denemelerinden günümüze, havacılık endüstrisi büyük bir evrim geçirmiş, dünyayı daha yakın bir hale getirmiş ve bir dizi sektörde önemli yeniliklere öncülük etmiştir. Insanlığın sınırlarını zorlayan ve dünya genelinde birleştirici bir rol oynayan önemli bir sektördür.

Havacılığın insanları birleştirmesi gibi fonksiyonlarının yanıda askeri olarak da çok önemli bir yeri vardır. Havacılığın askeri açıdan önemi, savaş stratejilerini, keşif faaliyetlerini ve hatta lojistik operasyonları temelinden değiştirmiştir. Hava araçları, düşman hatlarını aşma, hedefleri hassas bir şekilde vurma ve genel olarak askeri güç projeksiyonu sağlama konusunda kritik bir rol oynamaktadır. Havacılık, özellikle askeri operasyonlarda, insan hayatının doğrudan etkilendiği bir alandır. Bu nedenle, uçakların, helikopterlerin ve diğer hava araçlarının güvenliği, sadece stratejik avantajlar sağlamakla kalmaz, aynı zamanda insan hayatını koruma açısından da hayati önem taşır. Bu önemli faktörler, havacılık güvenliği konusunda sürekli bir gelişimi ve yenilikçi çözümleri beraberinde getirmiştir.

Ancak, havacılık avantajlarına rağmen, siber tehditlerin ortaya çıkmasıyla birlikte, hava araçlarının güvenliği daha da kritik hale gelmiştir. Mil-Std 1553, 1975 yılında ABD Savunma Bakanlığı tarafından geliştirilen bir iletişim standardıdır ve hava araçlarında, kara araçlarında, deniz araçlarında ve uzay araçlarında kullanılmaktadır. Mil-Std 1553, birçok hava, kara ve deniz aracının iletişimini sağlayarak bu araçlar arasında entegrasyonu mümkün kılar. Bu standardın çift yedekli veri yoluna, yüksek güvenilirliğe ve düşük hata oranına sahip olması nedeniyle emniyet kritik sistemlerde kullanılmaktadır. Standart ilk geliştirildiği zamanlarda güvenli kabul edildiği halde, zaman içinde teknolojinin ve saldırı yöntemlerinin gelişmesiyle savunmasız hale gelmiş ve saldırganlar için kolay hedef haline gelmiştir. Siber saldırılar, bilgi sızdırma, veri manipülasyonu ve hatta sistemlerin tamamen devre dışı bırakılması gibi ciddi sonuçlara yol açabilir.

Önerdiğimiz sistem, Mil-Std 1553 sistemi üzerinde gerçekleşen saldırıları tespit etme ve bu saldırılara karşı etkili bir güvenlik çözümü sunma amacını taşımaktadır. Proje, makine öğrenmesi, doğal dil işleme teknikleri ve Mil-Std 1553 bus yapısının bağlantılarına dışarıdan eklenen bir switch kullanarak bus için genel bir saldırı tespiti ve korunma sistemi sağlamakatadır.

Mil-Std 1553 veri trafiğini analiz ederek anormal durumları belirlemeyi ve bus controller yapısını güçlendirerek bus üstündeki hakimiyeti kaybetmemeyi amaçlamaktadır. Bu, sistemin normal işleyişinden sapmaları tespit etmeye yardımcı olacaktır.

Bunu yaparken makine öğrenmesi algoritmalarından random forest,stochastic gradient descent, decision tree, k nearest neighbor, logistic regression ve gaussian naive bayes algoritmalarını kullanmaktadır. Bu 6 farklı algoritma ve farklı datasetler ile hangi algoritmanın nerelerde en iyi performans gösterdiği ölçülerek gösterilmek istenmiştir.

Bu yolla Mil-Std 1553 sistemine önerdiğimiz çözümü entegre etmek isteyen bir sistemin hangi algoritmayı tercih etmesi gerektiği netleştirmek istenmiştir. Doğal dil işleme algoritmalarından biri olan BERT algoritması da bir siber saldırı tespit yöntemi olarak iletişim trafiği üzerindeki anlamlı desenleri anlamak için kullanılacaktır. Bu yolla Mil-Std 1553 sistemine yapılabilecek saldırıların tespit mekanizması güçlendirilmiş olacaktır. Bus kontrolünü ele geçirmek için yapılabilecek herhangi bir atağa karşı savunmasız olan orijinal Mil-Std 1553 bus sistemi, önerdiğimiz bus controller yapısına bağlanan Bus controller switch ile bus kontrolünü kaybetmeyecektir ve saldırılara karşı koyacaktır. Bu yollarla hem insan hayatı için hem de görev için kritik olan Mil-Std 1553 sistemi korunacak ve aracın güvenle görevini tamamlanmasına yardımcı olunacaktır.

Çalışma sırasında ilk olarak Mil-Std 1553 bus yapısına yapılacabilecek saldırıların metodları düşünüldü. Bu saldırıların ilk olarak nereye yapılacabileceği ve Mil-Std 1553 bus yapısının en savunmasız yanlarının ne olduğu araştırıldı. Bu araştırmalara dayanarak makine öğrenimi algoritmaları ve doğal dil işleme algoritması seçildi. Çalışmamızın ikinci kısmında farkettiğimiz, bus controller yapılacak bir saldırının sonucunda Mil-Std 1553 sisteminin tüm kontrolünün kötü niyetli yazılımların eline geçebileceğidir. Bu durumda bunu önlemek için Mil-Std 1553 tasarımını ve protokolünü değiştirmeden yapılabilecek en ekonomik ve verimli yollardan birini yaparak bus yapısına dışardan bir switch yardımı ile bus üzerindeki cihazların fiziksel bağlantısının kontrolü bus controller'a verildi. Böylece bus controller kendisine saldırmaya çalışabilecek herhangi bir cihazın Mil-Std 1553 ile bağlantısını keserek hem kendisini hem de bus yapısını koruyabilecektir.

Çalışmalarımızı tamamladığımızda, makine öğrenmesi algoritmaları ve bert algoritmasıyla yapılan saldırı testpit sisteminde ortaya çıkan sonuçlar f1 skor, macro average f1 skor, weighted average f1 skor, accuracy ve confusion matrix gibi metriklerle test edilmiştir. Bunu yaparken makine öğrenmesi algoritmaları için üç farklı veri seti kullanıldı. BERT algoritmasının testi için de farklı bir dataset kullanıldı.

Sonuçlar farklı algoritmaların farklı eğitim setlerinde ve farklı yoğunluktaki verilerde iyi performanslar gösterdiğini ortaya koydu. Ayrıca Mil-Std 1553 için yapılacak bir saldırı tespit sisteminde tasarım ve cihazların bağlı olacağı yapıya göre hangi algoritmanın daha performanslı olacağı belirlendi. Test sonuçları, önerdiğimiz sistemin farklı veri boyutlarında ve farklı yoğunluktaki Mil-Std 1553 sistemlerinde bile iyi performanslar gösterdiğini ortaya çıkarttı. Hedeflenen f1 skor ve accuracy değerlerine yakın sonuçlar elde edildi.

Bu çalışmanın devamında yapay zeka ve makine öğrenmesi tekniklerinin daha etkin bir şekilde kullanılması, siber tehditlerin daha etkili bir şekilde değerlendirilmesi üzerinde odaklanabilir. Mil-Std 1553 sisteminin güvenlik açıklarını kapatmak için yeni nesil şifreleme teknolojilerinin entegrasyonu gibi çözümler üzerinde çalışmak da gelecek araştırmaların öncelikli konularından biri olabilir.

## A NEW INTRUSION DETECTION APPROACH FOR MIL-STD 1553 BASED SYSTEMS


### SUMMARY

Aviation holds historical significance in the development and technological progress of humanity. From the early attempts at flight to the present day, the aviation industry has undergone significant evolution, bringing the world closer together and pioneering important innovations across various sectors. It is a crucial sector that pushes the boundaries of humanity and plays a unifying role globally.

In addition to its functions in bringing people together, aviation also holds a crucial place in military contexts. The military importance of aviation has transformed war strategies, reconnaissance activities, and logistics operations at their core. Aircraft play a critical role in overcoming enemy lines, accurately striking targets, and providing overall military power projection. Aviation, especially in military operations, is an area where human lives are directly affected. Therefore, the safety of aircraft, helicopters, and other aerial vehicles is not only essential for providing strategic advantages but also crucial for preserving human life. These significant factors have led to continuous advancements and innovative solutions in aviation safety.

However, despite the advantages of aviation, the security of aircraft has become even more critical with the emergence of cyber threats. Mil-Std 1553, developed by the U.S. Department of Defense in 1975, is a communication standard used in military aircraft, ground vehicles, and spacecraft. Mil-Std 1553 enables the communication of many air, ground, and sea vehicles, facilitating integration among these platforms. Due to its dual redundant data bus, high reliability, and low error rate, this standard is employed in safety-critical systems. Over the course of time, advancements in technology and the evolution of attack methods have increased its susceptibility, thereby presenting adversaries with an exploitable target. Cyber attacks can result in serious consequences such as information leakage, data manipulation, and even complete system shutdown.

Securing and maintaining the integrity of critical systems within military aircraft stands as a paramount objective outlined by Mil-Std 1553. This protocol is designed to fortify the transmission of data among these systems by implementing robust security measures, such as encryption and authentication protocols. The incorporation of these security features not only safeguards sensitive information but also ensures the reliability of communication channels.

Mil-Std 1553 plays a pivotal role in the protection of critical systems, providing a comprehensive framework for encoding and verifying data exchanged between different components. By employing encryption, the protocol adds an extra layer of defense against unauthorized access and tampering, thereby enhancing the overall security posture of military aircraft.

Furthermore, the protocol's noteworthy attribute of maintaining low error rates contributes significantly to the overall reliability of military aircraft. The reduction in error rates minimizes the likelihood of data corruption during transmission, ensuring the accurate and consistent flow of information between critical systems. This reliability is particularly crucial in the context of military operations where precision and consistency are imperative for the successful execution of critical missions.

Unauthorized access and susceptibility to manipulation of data can heighten the risk of strategic information falling into the hands of adversaries. Additionally, cyber attacks may have a direct impact on flight safety. Aircraft utilizing the MIL-STD 1553 protocol become more vulnerable to cyber threats due to the weak security measures in this standard, potentially leading to accidents. From an operational perspective, the standard's poor cybersecurity can result in coordination issues in military operations and integration deficiencies among systems, negatively affecting operational efficiency. Lastly, the vulnerability of the standard can limit the defense capabilities of military platforms, preventing military forces from maintaining a strategically and operationally secure position. The protection of data and the security of military systems necessitate the strengthening of standards, such as MIL-STD 1553, with enhanced security measures. These measures are crucial for ensuring the protection of strategic information, increasing resilience against cyber attacks, and optimizing the effectiveness of military operations.

The proposed system aims to detect attacks on the Mil-Std 1553 system and provide an effective security solution against these attacks. The project involves machine learning, natural language processing techniques, and the use of an externally added switch to the Mil-Std 1553 bus structure to establish a general attack detection and protection system for the Mil-Std 1553 bus.

By analyzing Mil-Std 1553 data traffic, the system aims to identify abnormal situations and reinforce the bus controller structure to maintain dominance over the bus. This will help detect deviations from the normal operation of the system. In doing so, it utilizes machine learning algorithms such as random forest, stochastic gradient descent, decision tree, k nearest neighbor, logistic regression, and Gaussian naive Bayes. The goal is to demonstrate the performance of these six different algorithms and determine where each algorithm excels based on different datasets.

During the integration of the proposed solution into the Mil-Std 1553 system, the aim is to clarify the preferred algorithm for a system seeking to enhance its security. The BERT algorithm, one of the natural language processing algorithms, will be utilized as a method for detecting cyber attacks by understanding meaningful patterns in communication traffic. This approach will strengthen the detection mechanism for potential attacks on the Mil-Std 1553 system. The original Mil-Std 1553 bus system, vulnerable to any attack that aims to take control of bus operations, will not lose control of the bus when connected to the suggested Bus Controller Switch structure. It will effectively resist and defend against potential attacks. Through these means, the Mil-Std 1553 system, critical for both human safety and mission completion, will be safeguarded.

In the course of the study, considerations were given to potential attack methods on the Mil-Std 1553 bus structure. The initial focus was on identifying possible points of

vulnerability and researching the weakest aspects of the Mil-Std 1553 bus structure. Based on these investigations, machine learning algorithms and the BERT natural language processing algorithm were selected. In the second part of our study, a significant observation was made that a potential attack on the bus controller could lead to malicious software gaining control over the entire Mil-Std 1553 system. To prevent this, an economical and efficient approach was implemented without altering the Mil-Std 1553 design and protocol. This involved introducing an external switch to the bus structure, allowing the bus controller to control the physical connections of devices on the bus. Consequently, the bus controller can sever the connection of any device attempting to attack it, thereby protecting both itself and the bus structure.

Upon the conclusion of our research, the outcomes of the attack testing system, utilizing machine learning algorithms and the BERT algorithm, were scrutinized through metrics encompassing F1 score, macro-average F1 score, weighted-average F1 score, accuracy, and confusion matrix. In the course of conducting this study, three distinct datasets were utilized for the purpose of testing machine learning algorithms. Additionally, a separate dataset was utilized specifically for testing the BERT algorithm.

The findings underscored the nuanced performance of diverse algorithms across distinct training sets and datasets with varying densities. Furthermore, in the context of an intrusion detection system tailored for Mil-Std 1553, the selection of the algorithm hinged upon the intricacies of the system's design and the architecture to which connected devices would adhere. The test results substantiated the robust performance of the proposed system across Mil-Std 1553 systems featuring disparate data dimensions and densities, culminating in outcomes proximate to the designated F1 score and accuracy values.

Initially, our objective was to establish a comprehensive system aiding the defense of the Mil-Std 1553 system against sophisticated attacks. In contrast to prior studies, this research, which proved to be considerably extensive, aimed to uncover significant insights and outcomes, ultimately reaching the targeted conclusion.

As a prospect for future research endeavors, one might consider placing a heightened emphasis on the more effective utilization of artificial intelligence and machine learning techniques. This focus could be directed toward achieving a more nuanced assessment of cyber threats. Furthermore, exploring potential solutions, including the integration of state-of-the-art encryption technologies, to address security vulnerabilities in the Mil-Std 1553 system may merit prioritization in subsequent academic research pursuits.

# 1. GİRİŞ

Hava araçları, insanlığın gökyüzüne uzanan teknolojik yolculuğunda eşsiz bir başarı öyküsünü temsil eder. Her biri karmaşık sistemlerle donatılmış olan bu metal kuşlar, sadece uçmakla kalmayıp, aynı zamanda yüksek teknolojiyle entegre edilmiş bir dizi alt sistemle bir araya gelir. Bu alt sistemlerin etkili bir şekilde iletişim kurması, hava araçlarının güvenliği, performansı ve sürdürülebilirliği açısından kritik bir öneme sahiptir.

Hava aracı sistemlerinde veri iletimi, kara araçlarından farklı özelliklere ve gereksinimlere sahiptir. Hava araçları, kara araçlarına göre daha karmaşık bir topolojiye sahiptir. Bu durum, hava araçlarındaki veri iletim altyapısının, güvenli ve etkili iletişimi sağlayacak şekilde tasarlanmasını gerektirir. Ayrıca hava araçları genellikle askeri güvenlik uygulamalarında kullanılır, bu nedenle veri iletim altyapıları genellikle yüksek güvenlik standartlarına uymak zorundadır. Aviyonik, havacılık ve elektronik kelimelerinin birleşiminden türetilmiş bir terimdir ve genellikle hava araçlarının elektronik sistemlerini ifade eder. Bu sistemler, hava aracının kontrolünü, navigasyonun sistemlerini, haberleşme sistemlerini, silah sistemlerini ve diğer kritik görevleri desteklemek üzere tasarlanmış elektronik bileşenlerin bütünleşik bir ağıdır. Aviyonik sistemler, hava araçlarının güvenliğini, etkinliğini ve performansını artırmak için kullanılır. Aviyonik sistemlerde veri iletimi, farklı alt sistemler arasında bilgi paylaşımını ve koordinasyonu sağlamak amacıyla kullanılan bir dizi veri iletim protokolü bulunmaktadır. Mil-Std 1553 askeri hava araçlarında ve diğer havacılık uygulamalarında yaygın olarak kullanılan aviyonik veri iletim standartlarından biridir. Mil-Std 1553, Amerikan Savunma Bakanlığı tarafından geliştirilen ve 1973 yılında ortaya çıkan bu standart, askeri hava araçlarının karmaşık ve kritik sistemleri arasında güvenilir iletişimi sağlamak üzere tasarlanmıştır. İlk oluşturulduğunda hava araçları için tasarlanmış olsa da ilerleyen zamanlarda kara araçları, deniz araçları ve hatta uzay araçları da bu sistemi kullanmışlardı. F-16 Fighting Falcon, F-35 Lightning II, B-2

Spirit gibi savaş uçakları, AH-64 Apache, S70 Black Hawk gibi helikopterler, MQ-9 Reaper gibi insansız hava araçları, James Webb uzay teleskobu, M1 Abrams tankı ve diğer birçok askeri platform Mil-Std 1553 sistemini kullanmaktadır [3]. Birçok farklı platformda kullanılmasına rağmen bu sistemin hava araçlarında önemi çok daha büyüktür. Hava araçlarında, farklı sistemler arasında veri iletimini düzenleyerek, uçuş kontrolünden silah sistemlerine, navigasyondan sensör sistemlerine kadar birçok kritik bileşeni birbirine bağlar.

Askeri hava araçlarındaki kritik sistemler arasında güvenli ve bütünlük korunması, Mil-Std 1553'ün temel amaçlarından biridir. Protokol, bu sistemler arasındaki veri iletimini şifreleme ve doğrulama gibi güvenlik önlemleriyle destekler. Ayrıca düşük hata oranlarına sahip olması, hava aracının güvenilirliğini artırır ve kritik görevlerin sorunsuz bir şekilde gerçekleştirilmesini sağlar. Askeri hava araçlarındaki kritik sistemler arasında güvenli ve bütünlük korunması, Mil-Std 1553'ün temel amaçlarından biridir. Protokol, bu sistemler arasındaki veri iletimini şifreleme ve doğrulama gibi güvenlik önlemleriyle destekler. Ayrıca düşük hata oranlarına sahip olması, hava aracının güvenilirliğini artırır ve kritik görevlerin sorunsuz bir şekilde gerçekleştirilmesini sağlar [1].

Mil-Std 1553, zorlu hava koşullarına, elektromanyetik müdahalelere ve diğer dış etkenlere karşı dayanıklılık sağlamak üzere tasarlanmıştır. Bu özellik, askeri operasyonlarda güvenilir bir iletişim altyapısının oluşturulmasına katkı sağlar. Bu standart, uzun yıllardır askeri havacılık ve uzay uygulamalarında kullanılan bir standart olmuş, askeri platformlardaki elektronik sistemlerin koordinasyonu ve güvenli iletişimi için temel bir taşıyıcı olarak görev yapmıştır.

Yıllarca birçok farklı platformda kullanılan Mil-Std 1553 standardı, teknolojinin gelişmesiyle birlikte bazı zorluklarla karşılaşmıştır. Özellikle siber güvenlik açısından, bu protokol tabanlı sistemlerin, zamanla gelişen teknoloji ile siber saldırılara daha açık hale gelmesine neden olmuştur [4].

İlk tasarlandığı yıllarda, Mil-Std 1553'e yönelik siber saldırıların fazla mümkün olmaması, hava araçlarının bus yapısının dış sistemlere sınırlı bağlantıya sahip olması

gibi nedenlerle güvenli kabul edilmiştir. Ancak, teknolojik ilerlemelerle birlikte, özellikle 2010'lardan sonra, Mil-Std 1553 siber tehditlerine daha açık hale gelmiştir.

Siber tehditler, Mil-Std 1553 tabanlı sistemlere yönelik yeni saldırı vektörlerini ortaya çıkarmıştır. Bu tehditler, sistemlerin güvenliğini zayıflatabilir ve istenmeyen müdahalelere olanak tanıyabilir. Bu durum, mevcut güvenlik önlemlerinin yetersiz kaldığını ve Mil-Std 1553 standardının siber tehditlere karşı daha güçlü bir şekilde korunması gerektiğini göstermektedir.

Bu nedenle, MIL-STD 1553 standardının güvenlik açısından güçlendirilmesi ve güncellenmesi, modern siber tehditlere daha etkin bir şekilde karşı koymak için kritik bir öneme sahiptir. Bu araştırmada, Mil-Std 1553 tabanlı iletişim sistemleri için makine öğrenimi tabanlı bir sızma algılama öneriyoruz. Mil-Std 1553 iletişim otobüsündeki anormallikleri tespit etmek için makine öğrenimi ve doğal dil işleme tekniklerini kullanıyoruz. Önerdiğimiz sistem, Mil-Std 1553 tabanlı sistemlerde anormallik tespiti için Stokastik Gradyan İniş algoritmasını kullanan ilk araştırmadır. Bu çalışma, Mil-Std 1553 tabanlı sistemlerde anormallik tespiti için BERT algoritmasını kullanan ilk araştırmadır ve anormallikleri tespit etmede önemli bir katkı sağlar. Önerilen sistem, farklı türdeki saldırıları (örneğin, DOS, sahtecilik) tespit edebilir. Önerdiğimiz sistem, eğitim ve test amaçları için bir simülatörden elde edilen bir veri kümesi kullanılarak deneysel olarak analiz edilmiştir ve gerçek Mil-Std 1553 uygulamalarıyla çok benzer sonuçlar vermektedir.

## 1.1 Tezin Amacı

MIL-STD 1553 standardının savunmasızlığı, öncelikle hassas verilerin korunmasını tehlikeye atar. Verilerin yetkisiz erişim ve manipülasyona açık olması, stratejik bilgilerin düşmanların eline geçme riskini artırabilir. Ayrıca, siber saldırıların uçuş güvenliği üzerinde doğrudan etkisi olabilir. MIL-STD 1553'ü kullanan hava araçları, bu standardın zayıf güvenlik önlemleri nedeniyle siber saldırılara daha açık hale gelir, bu da potansiyel kazalara yol açabilir. Operasyonel açıdan, standardın siber güvenliğinin zayıf olması, askeri operasyonlarda koordinasyon sorunlarına ve sistemler arasındaki entegrasyon eksikliklerine neden olabilir. Bu da operasyonel

etkinliği olumsuz yönde etkileyebilir. Son olarak, standardın savunmasızlığı, askeri platformların savunma yeteneklerini sınırlayabilir. Düşman saldırılarına karşı savunmasızlık, askeri operasyonlarda dezavantaj yaratır ve stratejik hedeflere ulaşmayı zorlaştırır [5].

Bu nedenle Mil-Std 1553 sistemlerini saldırılara karşı korumak ve güçlendirmek büyük bir önem arz etmektedir. Bu araştırmada, Mil-Std 1553 tabanlı iletişim sistemleri için hibrit bir saldırı tespiti sistemi öneriyoruz. Bu hibrit saldırı sistemi ile sadece tek yönlü bir saldırı tespit sistemi değil çok daha geniş bir güvenlik sistemi oluşturmayı hedefliyoruz. Bu saldırı tespit sistemi birkaç alt sistemden oluşmaktadır. Bu alt sistemlerden her biri kendi başlarına bir savunma sistemleridir. Bizim önerdiğimiz bu sistem bu farklı savunma sistemlerini bir araya getirmektedir.

İlk olarak Mil-Std 1553 sisteminde RT-RT arasındaki mesajlaşmalarda olabilecek manüpülasyon ve saldırılara karşı iki farklı sistemi aynı anda planalamayı öneriyoruz. Bu durum çift taraflı bir kontrol mekanizması oluşmasını sağlayacak ve herhangi bir şekilde hatalı bir tespit yapılmasını veya kötü niyetli mesajarın tespit edilememesini önlemeyi sağlamak için yapılmaktadır.

Bu amaçla ilk olarak kullandığımız yöntem makine öğrenmesi algoritmalarıyla Mil-Std 1553 sisteminde anomali tabanlı saldırı tespit sistemi çalışmamızdır. Bu çalışmayı yaparken 6 farklı makine öğrenmesi algoritması kullanıldı. Bu algoritmalar stochastic gradient descent algoritması, random forest algoritması, decision tree algoritması, k-nearest neighbors algoritması, logistic regression algoritması ve gaussian naive bayes algoritmasıdır. Bu algoritmaların her birinin farklı şekillerde pozitif ve ya negatif yanları vardır. Bazı şartlarda ve gereksinimlerde bir makine öğrenmesini kullanmak diğerlerine göre daha avantajlı olabilirken başka bir şartta veya gereksinimde bu ihtiyaç değişmektedir. Bu durumu derinlemesine incelemek için farklı boyutlarda ve farklı dağılım özelliğinde veri setleri kullanılmıştır. Bu çalışmanın sonucunda, hem makine öğrenmesi ile Mil-Std 1553 1553 sisteminde yapılacak saldırı tespit sisteminde göstereceği performansın yüksekliği ve kullanılabilirliği hem de makine öğrenmesi algoritmalarının hangi koşullarda daha yüksek performans göstereceği gösterilmek istenmiştir.

RT-RT mesajlaşmasında saldırı tespitine yönelik ikinci kısmında bir doğal dil işleme algoritması olan BERT algoritmasını kullanıyoruz. Doğal dil işleme alanında önemli bir konsept olan ve büyük başarı elde eden bir öğrenme modelidir. Google tarafından geliştirilmektedir. Normalde Mil-Std 1553 sistemini için çok uygun olmayan BERT sistemi verisetinin düzenlenmesiyle çalışabilir duruma gelmiştir. Bu sistem ile makine öğrenmesi algoritmlarından farklı bir altyapı kullanan bir tespit sistemi elde edilmek istenmiştir. 2 farklı tespit yapısı ile herhangi birinin farkedemeyeceği bir saldırı olursa diğerinin farketmesi amaçlanmıştır.

Hibrit savunma sisteminin son ayağını ise BC savunma sistemi oluşturmaktadır. Mil-Std 1553 sisteminin üstündeki RT veya BM cihazlardan biri kötü niyetli yazılımlar tarafından ele geçirmesi ve bus üstünde BC olmaya çalışması durumunda güvenliği sağlaması için eklenmiştir. Mil-Std 1553 sisteminin bu durumda kötü niyetli cihazın engellemesi için neredeyse hiçbir güvenlik sistemi yoktur. Bu durumda önerdiğimiz sistem, kötü niyetli cihazı devreden çıkarıp, bus controller cihazı, tekrar Mil-Std 1553 busını yönetebilir pozisyona getirecek etkili bir yöntem olarak görünmektedir.

## 1.2 Katkı

Mil-Std 1553 sistemlerinde günümüz teknolojisinde ortaya çıkan güvenlik açıkları büyük tehlike oluşturmaktadır. Bu nedenle Mil-Std 1553 sisteminin güvenliği için oldukça kapsamlı ve her türlü saldırıyı tespit edebilecek bir sistem ortaya çıkarmayı odaklandık. Bu çalışmalar sırasında ilk olarak Mil-Std 1553 sistemine hangi tip saldırıların yapılabileceğinin üstünde durduk. Bu araştırmalardan sonra ilk olarak öne yapımına başladığımız makine öğrenmesi ile Mil-Std 1553 sisteminde saldırı tespiti metodu ortaya çıktı. Bu sistemi yaparken stochastic gradient descent algoritması, random forest algoritması, decision tree algoritması, k-nearest neighbors algoritması, logistic regression algoritması ve gaussian naive bayes algoritması kullandık. Önerdiğimiz sistem, Mil-Std 1553 tabanlı sistemlerde anormallik tespiti için stokastik gradyan descent algoritmasını kullanan ilk araştırmadır. Bunun yanında algoritmaları farklı veri setleriyle test ederek en iyi yolun hangi koşullarda hangi makine öğrenmesi algoritmasının daha verimli olabileceğini gösteren bir çalışma yaptık.

Bu çalışmalardan sonra günümüzde oldukça kullanılan ve gelişmiş bir deep learning tabanlı doğal dil işleme BERT algritması üzerine yoğunlaştık. Bert algoritması ile Mil-Std 1553 sisteminde saldırı tespiti ilk bakışta birbirlerine çok uyabilen sistemler değildir. Ancak bu sistemleri birbirlerine uyarlamaya odaklandık ve bu konuda başarılı olduk. Önerdiğimiz sistemde kullanılan BERT algoritması, ilk defa Mil-Std 1553 busında anomalikleri tespit etmek kullanılmış oldu.

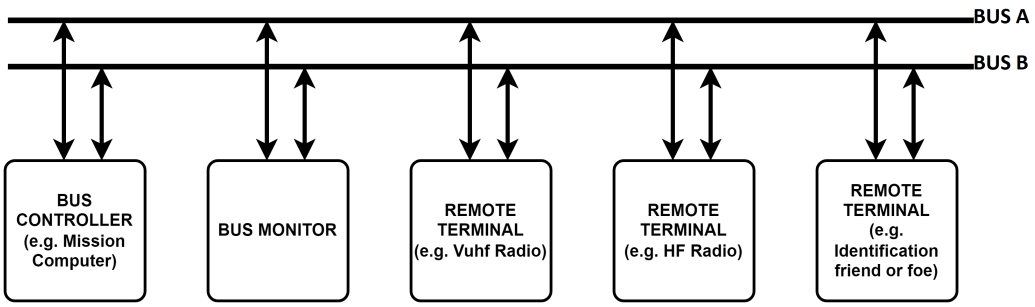
Daha sonra Mil-Std bus protokolündeki en önemli zayıflıklardan birini ortadan kaldırmayı hedefledik. Bu zayıflık bir şekilde bus sistemi üzerinde bulunan bir RT ya da BM cihazın bus kontrolünü ele geçirmeye çalışması durumunda öncelikle bus'ı sonra da Mil-Std 1553 bus'ının üzerinde bulunduğu aracı korumaktadır. Böyle bir saldırı olması durumunda olabilecek aşamalarını ilk defa bu araçtırma ile detaylı olarak açıkladık ve bütün sistemi koruyabilecek çözüm önerimizi ortaya koyduk.

En başta hedeflediğimiz daha önce benzeri olmayan, Mil-Std 1553 sisteminin, saldırılara karşı savunulmasına yardımcı olacak tüm bir sistemdi. Daha önce yapılan çalışmalara göre oldukça geniş kapsamlı olan çalışma ile önemli bilgileri ve sonuçlara ortaya koyarak istediğimi sonuca ulaştık.

# 2. MIL-STD 1553 VE SALDIRI TESPİT YÖNTEMLERİ

## 2.1 Mil-Std 1553

MMil-Std 1553, askeri uygulamalarda kullanılan bir data bus standartıdır. Amerika Birleşik Devletleri Savunma Bakanlığı (USDoD) tarafından 1975 oluşturulmuştur. İlk olarak savaş uçakları için tasarlanmış olmasına rağmen daha sonraları diğer hava araçları, deniz araçları, kara araçları ve hatta uzay araçlarında da kullanılmıştır. Mil-Std 1553 silah sistemleri, telsiz sistemleri, navigasyon sistemleri gibi farklı sistemlerinde entegrasyonunda kullanılan temel araçlardan biri haline gelmiştir. Bu standart sayesinde bus'a bağlı alt cihazlar birbirleriyle iletişim kurabilirler. [4].




**Şekil 2.1 :** 1 BC,1 BM ve 3 RT cihazdan oluşan Mil-Std 1553 bus.

Mil-Std 1553, Manchester kodlama tekniğini kullanarak veriyi çift yönlü bir veri yolu üzerinden iletmek için tasarlanmıştır. Manchester kodlama ise bir veri iletim tekniğidir. Bu teknik, veriyi temsil etmek için sinyali iki seviyeli bir sinyale dönüştürerek gerçekleşir. Her bitin durumu, belirli bir zamansal periyotta değişir, bu da sinyalin yüksek ve düşük seviyelerinin belirli bir deseni oluşturmasını sağlar. Manchester kodlama, veri iletiminde güvenilirlik sağlamak ve veri senkronizasyonunu yönetmek amacıyla kullanılır. Bu özellikler, askeri ve uzay uygulamalarındaki karmaşık sistemlerde önem taşır [1].

Mil-Std 1553 bus sisteminde genellikle 2'li bir bus yapısı kullanılır. Şekil 2.1'de görüldüğü üzere bu buslar Bus A ve Bus B olarak ifade edilir. Bu şekilde çift bus

kullanılmasının asıl sebebi yedeklilik ve güvenliktir. Askeri sistemlerde güvenilirlik kritik bir öneme sahiptir. Bus A ve Bus B'nin kullanılması, sistemin daha yüksek bir seviyede güvenilirlik sağlamasına yardımcı olabilir. Eğer bus arızalanırsa, diğer bus devreye girer ve iletişim sürekli olarak devam eder. Bu, sistemin arızalara karşı daha dayanıklı olmasına olanak tanır. Benzer şekilde eğer bir bus düşman müdahalesi, arıza veya başka bir nedenle etkilenirse, diğer bus işine devam edebilir. Bu, sistemin güvenlik açısından daha dirençli olmasına yardımcı olabilir [1].

### 2.1.1 Mil-std 1553 bileşenleri

Mil-Std 1553 bus ağı genelde 1 adet aktif bus controller, bir veya daha fazla bus monitor ve bir veya daha fazla remote terminalden oluşur. Mil-Std 1553 standardında bus controller, askeri ve havacılık uygulamalarındaki bus sistemlerinde kritik bir rol üstlenen ve karmaşık görevleri yöneten bir ana yönetici cihazdır. BC'nin temel görevi, bus üzerindeki iletişimi kontrol etmek, senkronize etmek ve düzenlemektir. Bu çerçevede, BC, belirli görevlere yönelik komutları tanımlayarak ve bu komutları diğer bus bileşenlerine ileterek, data transfer süreçlerini başlatır, bitirir veya durdurur. BC, zamanlama ve senkronizasyonun sağlanmasıyla birlikte, bus üzerindeki hata yönetimini de üstlenir. Hata durumlarını izler ve gerekli önlemleri alarak iletişim güvenilirliğini artırır. Adresleme yetenekleri sayesinde, BC, veri yolu üzerinde bulunan diğer remote terminal cihazlarına spesifik komutları yönlendirebilir. Bu, belirli bir RT'nin özel bir görevi yerine getirmesine olanak tanır ve genel sistem performansını optimize eder [3] [4].

Ayrıca, BC'nin rolü genellikle bir uçağın veya aracın merkezi kontrol sistemini temsil etmekle sınırlı değildir; aynı zamanda bu cihaz, karmaşık alt sistemlerin etkileşimini yöneterek ve bus üzerindeki iletişimi düzenleyerek genel sistem entegrasyonunu koordine eder. Bu şekilde, BC, Mil-Std 1553 standardının gereksinimlerini karşılamak üzere tasarlanan sistemlerde, güvenilir ve etkili bir iletişim altyapısının anahtar bileşenidir.

Remote Terminal (RT), Mil-Std 1553 bus ağındaki bir cihaz türünü temsil eder. Mil-Std 1553 sistemine göre bus kontrolü yapmayan bir cihazı tespit eder. Yani, normal olarak çalışan bir RT, veri iletimini başlatma yeteneğine sahip değildir ve aynı

zamanda iletimi pasif bir şekilde izleme görevine de sahip değildir. Ancak, belirli komutları alıp yanıt verebilir ve belirli görevleri yerine getirebilir [3]. RT, ağ içindeki diğer cihazlarla iletişim kurmak için belirli bir adrese sahiptir. Bu adres, BC veya diğer RT'ler tarafından kullanılarak RT'ye yönlendirilecek iletilerin hedefini belirler. RT, belirli bir komut aldığında veya belirli bir durum oluştuğunda veri iletimi yapabilir. İletişim genellikle BC tarafından başlatılır ve RT, bu komutlara yanıt olarak belirli görevleri yerine getirir. RT'ler, genellikle askeri uygulamalarda sensörler, cihazlar veya diğer alt sistemlerin bir parçası olarak kullanılır. Örneğin, bir savaş uçağındaki bir sensör, RT olarak hareket ederek belirli veri görevlerini yerine getirebilir. RT'nin içerdiği kontrol mantığı ve algoritmalar, bu görevleri başarıyla yerine getirmesini sağlar [3].

Mil-Std 1553 bir diğer önemli parçası da Bus Monitor(BM)'dür. BM'nin temel görevi, veri yolu üzerinde iletilen mesajları pasif bir şekilde izlemek, kaydetmek, analiz etmek ve bu mesajlardan belirli bilgileri çıkarmaktır. BM, iletişimi başlatma veya yanıt verme yeteneğine sahip değildir. Tamamen pasif bir rol oynar. Ancak, BM'nin içerdiği özellikler, ağ üzerindeki iletişimi dikkatlice izleyerek, hataları tespit etmek, veriyi analiz etmek ve gerektiğinde bu veriyi saklamak gibi işlevleri gerçekleştirmesine imkan tanır.

BM'nin birincil işlevi, ağ üzerindeki veri transferini kaydetmek ve bu kayıtları daha sonra analiz için kullanılmak üzere tutmaktır. Bu, ağ üzerindeki iletişimde oluşan hataları tespit etmek ve gerekirse bu hataları düzeltmek için önemlidir. Ayrıca, BM, belirli bir cihazın durumunu izleyebilir ve bu cihazın sağlığını değerlendirebilir. BM, ayrıca bir yedek BC (Bus Controller) olarak hareket edebilir; bu durumda, başlıca BC'nin bir hata durumunda yerine geçebilmesi için gereken bilgileri sağlar.

BM'nin kullanılabilirliği, ağın güvenilirliğini ve performansını artırabilir. Ayrıca, BM'nin sağladığı analiz verileri, ağdaki genel durumu değerlendirmek ve gerektiğinde müdahale etmek için kullanılabilir. Bu özellikler, Mil-Std 1553 standardı altındaki sistemlerde, özellikle askeri ve havacılık uygulamalarında, kritik öneme sahip veri iletişimi altyapısının güvenilirliğini artırmak için tasarlanmıştır.

Aviyonik, uçaklarda kullanılan elektronik sistemleri içeren bir terimdir, bu sistemler iletişim, navigasyon ve ekran sistemlerini içerir. Bu sistemler, uçuş operasyonlarının güvenliği ve güvenilirliği konusunda kilit bir rol oynadıkları için güvenlik açısından kritik olarak değerlendirilir. Aviyonik gibi güvenlik açısından kritik sistemler, hata tolere edilebilirlik, gerçek zamanlı performans ve güvenlik sertifikasyonu gibi katı gereksinimlere tabidir [6].

### 2.1.2 Mil-std 1553 mesaj yapısı

Mil-Std 1553 standardı, Şekil 2.2'de görüldüğü iletişim protokolünde kullanılan üç temel kelime türü tanımlar. Bunlar veri kelimesi(data word), durum kelimesi(status word) ve komut kelimesi(command word) şeklindedir. Bu kelimeler, Mil-Std 1553 veri yol mimarisi içinde bilgi ve komut alışverişini kolaylaştırmada kritik roller oynar.

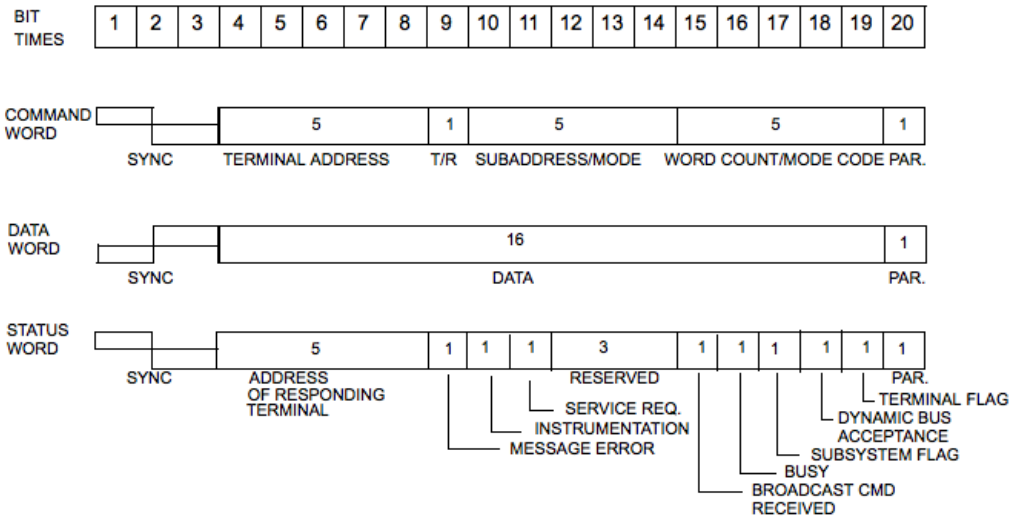
- Command Word: Eylemleri başlatma ve RT'yi belirli görevleri gerçekleştirmesi için yönlendirme sorumluluğundadır. Bu komut, BC tarafından RT'lere komutlar, talepler ve direktifler vermek için kullanılır. Command word, Mil-Std 1553 ağı içinde bağlı cihazların iletişim akışını düzenleme ve aktivitelerini koordine etme konusunda kilit bir rol oynar [7].

- Status Word: RT'den BC'ye durum ve kontrol bilgisi iletmek için kullanılır. Bu, RT'lerin durumu hakkında kritik geri bildirim sağlar. Alınan komutları doğrulama, hataları rapor etme ve verinin uygun olup olmadığını belirtme gibi durumları raporlar. Status word, Mil-Std 1553 mimarisi içinde iletişim sürecinin bütünlüğünü ve güvenilirliğini sağlamada etkili bir rol oynar [5].

- Data Word: Mil-Std 1553'deki veri kelimesi, BC ile RT arasında gerçek veri iletimi için kullanılır. Sensör okumaları, kontrol komutları veya başka ilgili yük bilgilerini taşır. Veri kelimesi, Mil-Std 1553 veriyolunu kullanan aviyonik sistemlerde gerçek zamanlı bilgi ve komut alışverişi için temel bir rol oynar [8].

## 2.2 Makine Öğrenmesi

İnsanlar tarihin ben başından bu yana yapılması gereken bazı işleri daha kolay, daha hızlı ya da az kişiyle yapmak için bu bazı araç ve gereçleri kullanmışlardır.




**Şekil 2.2 :** Mil-Std 1553 Word Yapısı [1].

Bu araçlar ilk başta çok daha basit iken ve günümüzde çok daha gelişmişlerdir. İlk araçlar sadece fiziksel iş gücünü azaltma ihtiyacını görüyorlardı. Ancak insalığın gelişimiyle beraber yaratıcılığı, kabiliyeti arttı ve kendisine çok daha gelişmiş yardımcılar oluşturmaya başladı. Günümüz teknolojisiyle birlikte artık sadece insanın fiziksel işgücünü azaltan araçlar değil, insan yerine ona benzer şekilde düşünen, karar veren yapılar var. Bu yapıları oluşturmamızı sağlayan çalışma alanına makine öğrenimi diyoruz. Makine öğrenimi bilgisayar sistemlerine belirli bir görevleri otomatik olarak öğrenmeyi ve uygulama yeteneği kazandırmayı amaçlar. Bu uygulama yeteneği karar verme, yönetim, sorun çözme, veri tespiti, oyun oynama, istatistiksel analiz, optimizasyon, algoritmik modelleme gibi çok çeşitli alanlarda olabilir [9].

Makine öğrenmesinin Şekil 2.3'de görüldüğü gibi 3 tamel kategorisi olduğu kabul edilir:

- Denetimli Öğrenme (Supervised Learning):

  Denetimli öğrenme, makine öğrenimi alanında yaygın olarak kullanılan bir paradigmadır. Bu öğrenme yaklaşımı, belirli bir çıkışa ulaşmak için gerekli olan model parametrelerini, etiketlenmiş bir eğitim veri kümesi üzerinden öğrenen bir algoritma içerir. Bu süreç, genellikle iki temel bileşeni içerir: giriş özellikleri veya değişkenleri ve hedef çıkış değerleri veya etiketleri.

Öğrenme süreci, bir modelin eğitim veri setindeki giriş özellikleri ve hedef çıkış değerleri arasındaki ilişkiyi anlamasını içerir. Örneğin, bir e-ticaret platformundaki aktif kullanıcı sayısını tahmin etmek için, satılan ürün sayısı ve kullanıcı değerlendirmeleri gibi giriş özellikleri kullanılabilir. Bu özellikler, modelin, hedef olan aktif kullanıcı sayısını (y değişkeni) tahmin etmek için kullanılacak olan parametreleri öğrenmesine yardımcı olur [10].

Denetimli öğrenme, genellikle iki temel problem türünü içerir: regresyon ve sınıflandırma. Regresyon problemlerinde, model, sürekli bir sayısal çıkışı tahmin etmeye çalışırken, sınıflandırma problemlerinde model, giriş özelliklerine dayanarak belirli bir kategoriye ait olma olasılığını tahmin etmeye çalışır.

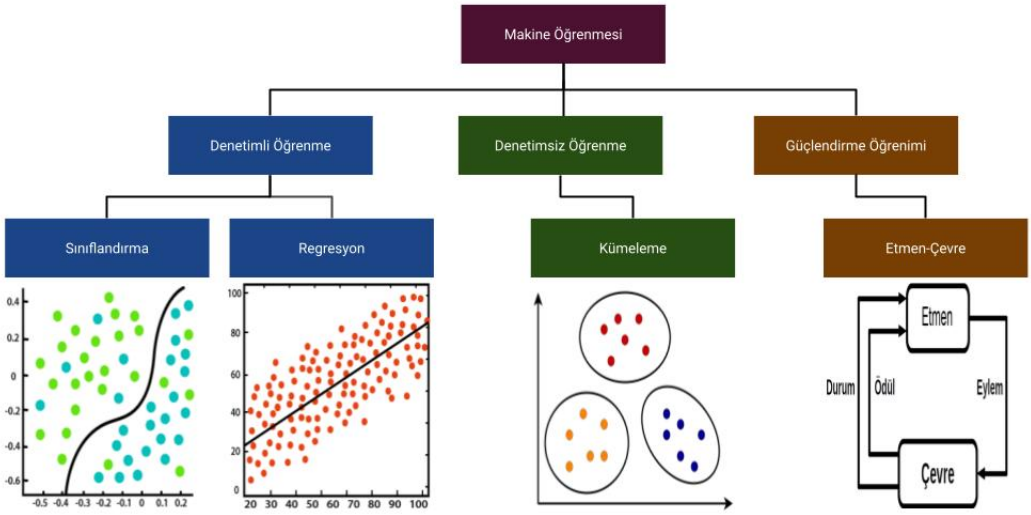
Eğitim sürecinden sonra, model yeni veya görülmemiş veri noktaları için tahminlerde bulunabilir. Denetimli öğrenme, tahmin doğruluğunu artırmak için modelin performansını değerlendirmek ve gerektiğinde ayarlamak için kullanıcıya olanak tanır.

Bu bağlamda, denetimli öğrenme, gerçek dünya uygulamalarında geniş bir kullanım alanına sahiptir. Örneğin, pazarlama platformlarında abone tahminleri, tıbbi teşhisler, finansal tahminler ve daha pek çok alanda başarıyla kullanılmaktadır. Bu teknik, bilgisayar sistemlerinin karmaşık veri setlerini anlamalarını ve öngörmelerini sağlamak için güçlü bir araç sunar.

- Denetimsiz Öğrenme (Unsupervised Learning):

  Unsupervised learning, öğrenme sisteminin herhangi bir önceden var olan etiket veya spesifikasyon olmadan desenleri tespit etmeye odaklandığı bir öğrenme paradigmadır. Bu öğrenme yaklaşımında, eğitim verileri sadece değişkenlerin kombinasyonlarından oluşur ve temel amaç, ilgi çekici yapısal bilgileri bulmaktır. Bu yapısal bilgiler, örneğin ortak özelliklere sahip öğelerin grupları (kümeleme) veya yüksek boyutlu uzaydan daha düşük boyutlu bir uzaya yansıyan veri temsillerini içerebilir [10].

  Bu öğrenme paradigması altında iki temel görev bulunur: kümeleme (clustering) ve boyut azaltma (dimensionality reduction). Kümeleme, veri setinde benzer özelliklere sahip veri noktalarını gruplandırmayı amaçlar.




**Şekil 2.3 :** Makine Öğrenmesi Tipleri.

Boyut azaltma ise yüksek boyutlu veri setlerini daha anlamlı ve işlenebilir bir formata dönüştürmeyi hedefler.

Örneğin, unsupervised learning algoritmaları kullanılarak bir müşteri veri seti üzerinde kümeleme gerçekleştirilebilir. Bu, benzer alışveriş alışkanlıklarına sahip müşterileri aynı grupta bir araya getirerek pazarlama stratejilerini daha iyi özelleştirmeyi sağlar. Aynı şekilde, boyut azaltma teknikleri, yüksek boyutlu bir veri setindeki karmaşık ilişkileri anlamak ve temsil etmek için kullanılabilir.

Unsupervised learning, veri analitiği ve keşifsel veri madenciliği uygulamalarında yaygın olarak kullanılmaktadır. Bu yöntem, büyük veri setlerinden anlamlı bilgiler çıkarmak ve karmaşık yapıları anlamak için güçlü bir araçtır.

– Güçlendirme Öğrenimi (Reinforcement Learning):

Güçlendirme Öğrenimi, bir sistemin belirli bir durumda en uygun eylemi seçerek toplam ödülü maksimize etmeye yönelik bir makine öğrenimi paradigmasıdır. Bu öğrenme yaklaşımı, öğrenme algoritmasına optimal çıkış örnekleri sağlanmadığı için deneme-yanılma sürecine dayanır [11].

Öğrenme süreci genellikle sistem ve çevresi arasında gerçekleşen bir dizi durum ve eylemi içerir. Bu durum ve eylemler arasındaki karmaşık ilişkilerin çözülmesi, öğrenme algoritmasının ödülü en etkili şekilde artıran faktörleri belirlemesini gerektirir. Örneğin, bir sinir ağı, tavla oyununu öğrenmek

amacıyla kullanılabilir, bu noktada ağ, tahta durumu ve zar atış sonuçları gibi giriş özelliklerini alarak optimal bir hamle yapmak için çıkış üretir [11].

Bu öğrenme sürecindeki temel zorluklardan biri, bir tavla oyununun onlarca hamleyi içerebileceği ve ödülün sadece oyunun sonunda ortaya çıkacağıdır. Ödül, oyunun tamamındaki her hamleye uygun bir şekilde atandığında, bu durum bir kredi atama problemi olarak adlandırılır.

Güçlendirme öğreniminin önemli bir özelliği, sistemin keşif ve sömürü arasında bir denge kurmaya çalışmasıdır. Sistem, yeni eylem türlerini deneyerek etkilerinin etkinliğini anlamaya çalışırken, bilinen etkili eylemleri kullanarak elde ettiği ödülü artırmaya çalışır. Bu denge, sistemin optimize edilmiş performansını sağlamak için kritiktir ve aşırı odaklanma durumunda suboptimal sonuçlara yol açabilir. Güçlendirme öğrenimi, halen yoğun bir araştırma konusu olup, uygulama alanları oldukça geniş bir yelpazede bulunmaktadır [12].

### 2.2.1 Makine öğrenmesi algoritmaları

Bu çalışmada 6 farklı makine öğrenmesi algoritması kullanılmıştır.

- Random Forest (RF) Algoritması:

  Random Forest olarak adlandırılan makine öğrenimi algoritması temelde birden fazla karar ağacını içerir. Her bir ağaç, veri setinin farklı alt kümelerini kullanarak eğitilir ve genellikle farklı özelliklere odaklanır. Random Forest, bu ağaçlardan gelen farklı sınıflandırmaları birleştirerek daha güvenilir ve genelleştirilebilir bir sonuç elde etmeyi amaçlar. Random Forest algoritmasının temel formülü Denklem 2.1'de görüldüğü şekilde ifade edilebilir.

$$\hat{Y}_{\mathrm{RF}} = \mathrm{argmax}_j \left( \sum_{i=1}^{N_{\mathrm{trees}}} \mathbb{1}(\hat{Y}_i = j) \right) \tag{2.1}$$

Burada:

$\hat{Y}_{\text{RF}}$ : Random Forest için. Tahmin edilen sınıf

$N_{\text{trees}}$ : Toplam karar ağacı sayısı.

$\hat{y}_i$ : İki rastgele özellik seti ile eğitilmiş $i$-inci karar ağacının tahmini.

$\mathbb{1}(\cdot)$ : İndikatör fonksiyonu; içindeki ifade doğruysa 1, aksi takdirde 0 döndürür.

$\text{argmax}_j$ : Toplamı en büyük yapan $j$ değerini bulma işlemi.

Algoritmanın temel özelliği, her bir karar ağacının bağımsız olarak eğitilmesi ve tahminlerde bulunmasıdır. Her ağacın kendi benzersiz görüş açısı ve öğrenilen desenleri vardır. Random Forest, bu farklılıkları kullanarak daha güçlü bir model oluşturur. [13]

**Algorithm 1** Random Forest Algoritması

```
function RandomForest(S, F, B)
    H ← {}
    for i ← 1 to B do
        S_i ← BootstrapSample(S)
        h_i ← RandomizedTreeLearn(S_i, F)
        H ← H ∪ {h_i}
    end for
    return H
end function
function RandomizedTreeLearn(S, F)
    f ← SmallSubsetOfFeatures(F)
    split ← BestFeatureSplit(S, f)
    return LearnedTree
end function
```

Sınıflandırma ve regresyon görevlerinde kullanılabilen Random Forest, sınıflandırma durumunda veri noktalarını belirli sınıflara atar, regresyon durumunda ise bir tahmin yapar. Her bir karar ağacının bağımsız tahminleri algoritmanın genel tahminini oluşturur. Yeni bir örnek verildiğinde, her ağaç tahminde bulunur ve ardından çoğunluk kararı alınarak nihai tahmin yapılır.

Bu güçlü ve esnek algoritma, özellikle büyük ve karmaşık veri setleri üzerinde etkili bir şekilde çalışabilir. Leo Breiman tarafından geliştirilen Random Forest, geniş bir uygulama alanına sahiptir ve özellikle sınıflandırma ve regresyon

problemlerine yönelik çözümlerde tercih edilmektedir [14]. Random Forest sınıflandırmanın genel algoritması Algorithm 1'de gösterilmiştir.

– Decision Tree (DT) Algoritması: Karar ağacı sınıflandırma algoritması, yorumlanabilirliği, hesaplama verimliliği ve doğruluğu nedeniyle çeşitli alanlarda yaygın olarak kullanılan bir yöntemdir. Karar ağaçları, giriş özelliklerinin değerine dayanarak veriyi alt kümeler halinde bölerek ve her alt kümeye bir etiket atayarak oluşturulur. Bu algoritma, resim sınıflandırması, uzaktan algılama, tıp ve bilgisayar bilimleri gibi farklı alanlarda kullanılmıştır [15] [16] [17]. Belirli senaryolarda diğer sınıflandırma algoritmalarını geride bıraktığı tespit edilmiştir [18]. Karar ağacı algoritması, basitliği ve yorumlanabilirliği ile bilinir, bu nedenle sınıflandırma görevleri için popüler bir tercihtir. Ayrıca, karar ağaçları, sınıflandırma yeteneklerini artırmak için evrimsel algoritmalar kullanılarak geliştirilmiştir [19].

**Algorithm 2** Decision Tree Training

**function** TRAINDECISIONTREE($S, F,$ depth)
  **if** depth $= 0$ **then**
    **return** LeafNode(MostFrequentClass($S$))
  **end if**
  $BestFeature, BestScore \leftarrow$ null, $\infty$
  **for** $f \in F$ **do**
    $score \leftarrow$ ComputeScore$(S, f)$
    **if** $score < BestScore$ **then**
      $BestFeature, BestScore \leftarrow f, score$
    **end if**
  **end for**
  $left\_S, right\_S \leftarrow$ SplitData$(S, BestFeature)$
  **return** Node(BestFeature, TrainDecisionTree($left\_S, F,$ depth $-$ 1), TrainDecisionTree($right\_S, F,$ depth $-$ 1))
**end function**

Bununla birlikte, karar ağacı algoritması performansını ve verimliliğini artırmaya yönelik araştırmalara da konu olmuştur. Örneğin, mobil kullanıcı sınıflandırması için genetik algoritmaya dayalı modifiye bir karar ağacı algoritması önerilmiş ve karar ağacı algoritmasının sonuçlarını optimize etmeyi amaçlamıştır [20]. Ancak, karar ağaçlarının performansının eğitim veri setinin boyutu gibi faktörlerden etkilenebileceği unutulmamalıdır.

Karar ağaçlarının yeterli eğitim verisi olduğunda sınıflandırma doğruluğunu artırabildiği, ancak eğitim örneği yetersizse performanslarının düşebileceği gözlemlenmiştir [21]. Decision Tree, bir ağaç yapısı oluşturarak sınıflandırma veya regresyon problemlerini çözmeye çalışan bir algoritmadır. Bu nedenle, algoritmanın genel yapı ve işleyişini ifade eden bir matematiksel denklemle tam olarak ifade edilemez. Ancak genel işleyişini temsil eden bir formül, Denklem 2.2'de görüldüğü gibi ifade edilebilir. Decision Tree sınıflandırmanın genel algoritması Algorithm 2'de gösterilmiştir.

$$f(x) = \sum_{i=1}^{M} c_i \cdot I(x \in R_i) \tag{2.2}$$

Özetle, karar ağacı sınıflandırma algoritması, yorumlanabilirliği, hesaplama verimliliği ve doğruluğu nedeniyle çeşitli alanlarda yaygın olarak kullanılmaktadır. Resim sınıflandırması, uzaktan algılama, tıbbi teşhis ve diğer alanlarda başarı göstermiştir. Ayrıca, genetik algoritmalar ve paralel formulasyonlar gibi yöntemlerle karar ağaçlarının performansını ve verimliliğini artırmaya yönelik araştırmalar yapılmıştır.

– K-Nearest Neighbors (KNN) Algoritması: K nearest neighbor sınıflandırma algoritması, basitliği ve desen tanıma ile sınıflandırma görevlerindeki etkinliği nedeniyle çeşitli alanlarda yaygın olarak kullanılan bir yöntemdir. KNN, yeni veri noktalarını eğitim verilerine olan benzerlik ölçülerine dayanarak sınıflandıran, parametrik olmayan bir örnek tabanlı öğrenme algoritmasıdır [22] [23] [24]. Tıp, bilgisayar bilimi, uzaktan algılama ve desen tanıma gibi çeşitli alanlarda kullanılmaktadır.

  Araştırmalar, KNN algoritmasının doğruluğunu ve verimliliğini artırmaya odaklanmıştır. Bu doğrultuda, ağırlıklı KNN, bulanık KNN ve sezgisel denetimli Öklidyen veri farkı boyut indirgeme gibi yöntemler kullanılarak algoritmanın performansı iyileştirilmeye çalışılmıştır. Ayrıca, algoritma, performansını artırmak için minimum genişleme ağacı ve komşuluk sınıflandırıcı

**Algorithm 3** K-Nearest Neighbors Classification

```
function KNN(X, y, k, new_data)
    distances ← {}
    for all (x_i, y_i) ∈ (X, y) do
        dist ← ComputeDistance(x_i, new_data)
        distances ← distances ∪ {(dist, y_i)}
    end for
    sorted_distances ← SortByDistance(distances)
    k_neighbors ← SelectTopK(sorted_distances, k)
    predicted_class ← MajorityVote(k_neighbors)
    return predicted_class
end function
```

gibi diğer tekniklerle entegre edilmiştir [25]. KNN algoritmasının temel formülü Denklem 2.3'de görüldüğü şekilde ifade edilebilir.

$$\hat{y} = \arg\max_{c_i} \left( \sum_{i=1}^{k} I(y^{(i)} = c_i) \right) \tag{2.3}$$

Burada:

$\hat{y}$ : Örnek için tahmin edilen sınıf.

$y^{(i)}$ : Eğitim veri setindeki $i$-inci örneğin sınıfı.

$c_i$ : KNN algoritması tarafından seçilen $i$-inci sınıf.

$k$ : Komşu sayısı (K).

– Gaussian Naive Bayes (GNB) Algoritması: Gaussian Naive Bayes sınıflandırma algoritması, desen tanıma ve sınıflandırma görevlerindeki basitliği ve etkinliği nedeniyle çeşitli alanlarda yaygın olarak kullanılan bir yöntemdir. Algoritma, sürekli değerli özellikleri modellemek için Gaussian dağılımını kullanır ve özelliklerin sınıfa göre koşullu bağımsız olduğunu varsayar, bu da veri noktalarını etkili ve verimli bir şekilde sınıflandırmayı mümkün kılar [26].

$$P(y|x) = \frac{P(x|y) \cdot P(y)}{P(x)} \tag{2.4}$$

Burada:

| | |
|---|---|
| $P(y\|x)$ : | Belirli bir sınıf $y$ için veri $x$ olduğunda olasılık. |
| $P(x\|y)$ : | Sınıf $y$ için veri $x$ olduğunda olasılık. |
| $P(y)$ : | Belirli bir sınıfa ait olasılık. |
| $P(x)$ : | Veri $x$ olduğunda gözlemlenen olasılık. |

Gaussian Naive Bayes algoritmasının önemli avantajlarından biri, basitliği ve uygulama kolaylığıdır, bu da onu çeşitli sınıflandırma görevleri için popüler bir tercih haline getirir. Ayrıca, algoritmanın hesaplama verimliliği ve yüksek boyutlu verilerle başa çıkma yeteneği, pratik uygulamalarda yaygın bir şekilde kullanılmasına katkıda bulunur [27].

**Algorithm 4** Gaussian Naive Bayes Classification

```
function GaussianNaiveBayes(X, y, new_data)
    class_probs ← {}
    for all class ∈ UniqueClasses(y) do
        class_data ← ExtractClassData(X, y, class)
        class_mean, class_std ← ComputeMeanAndStd(class_data)
        class_prob ← ComputeClassProbability(new_data, class_mean, class_std)
        class_probs ← class_probs ∪ {(class, class_prob)}
    end for
    predicted_class ← SelectClassWithMaxProb(class_probs)
    return predicted_class
end function
```

Bu algoritma metin sınıflandırma, duygu analizi ve spam tespiti gibi alanlarda kullanılmıştır. Bu da gaussian naive bayes algoritmasının doğal dil işleme ve bilgi çekme konularındaki önemini vurgulamaktadır. Ayrıca, algoritma genomik veri analizi ve protein fonksiyon tahmini gibi biyoteknolojide kullanılmış, karmaşık biyolojik olayları açığa çıkarmada potansiyelini ortaya koymuştur [27]. Gaussian Naive Bayes sınıflandırmanın genel algoritması Algorithm 4'de gösterilmiştir.

– Stochastic Gradient Descent (SGD) Algoritması: Makine öğrenimi ve derin öğrenmede yaygın olarak kullanılan bir optimizasyon algoritmasıdır. Özellikle büyük ölçekli modelleri eğitirken hesaplama verimliliği ve büyük

veri setleriyle başa çıkma yeteneği nedeniyle popülerdir. Algoritma, model parametrelerini eğitim verilerine göre kayıp fonksiyonunun negatif gradyanı yönünde iteratif olarak güncelleyerek çalışır. Bu iteratif süreç, modelin kayıp fonksiyonunu en aza indiren optimal parametre setine doğru yaklaşmasına olanak tanır. Stochastic Gradient Descent algoritmasının temel formülü Denklem 2.5'de görüldüğü şekilde ifade edilebilir.

$$\theta_{t+1} = \theta_t - \eta \nabla J(\theta_t; x^{(i)}, y^{(i)}) \tag{2.5}$$

Burada:

$\theta_{t+1}$ : Algoritmanın $(t+1)$ iterasyonunda model parametreleri.

$\theta_t$ : Algoritmanın $t$ iterasyonundaki model parametreleri.

$\eta$ : Öğrenme oranı (learning rate).

$\nabla J(\theta_t; x^{(i)}, y^{(i)})$ : Kayıp fonksiyonunun $\theta_t$ parametrelerine göre gradyanı, ve $x^{(i)}, y^{(i)}$ veri setindeki $i$-inci örneği temsil eder.

Stochastic gradient descent algoritmasının önemli avantajlarından biri, büyük ölçekli veri setleri ve yüksek boyutlu parametre uzaylarıyla başa çıkma konusundaki etkililiğidir. Her iterasyonda eğitim verisinin bir alt kümesine dayanarak model parametrelerini günceller. SGD, özellikle büyük veri setleri üzerinde karmaşık modelleri eğitmek için yüksek hesaplama yükünü önemli ölçüde azaltır ve bu da özellikle geniş veri setlerinde karmaşık modelleri eğitmek için uygundur [28]. Stochastic gradient descent sınıflandırmanın genel algoritması Algorithm 5'de gösterilmiştir.

Ayrıca, geleneksel gradyan azalma yöntemlerine kıyasla daha hızlı yakınsama avantajı sunar. Sürekli yapılan güncellemeler, optimizasyon sürecine gürültü ekler, bu da algoritmanın yerel minumumlardan kaçınmasına ve özellikle yüksek boyutlu uzaylarda daha optimal bir çözüme daha kısa sürede

**Algorithm 5** SGD Algoritması

```
function SGD(X, y, α, E)
    w ← InitializeWeights()
    for epoch ← 1 to E do
        for i ← 1 to len(X) do
            xi, yi ← RandomSample(X, y)
            w ← w − α · xi · (Sigmoid(DotProduct(xi, w)) − yi)
        end for
    end for
    return w
end function
function RANDOMSAMPLE(X, y)
    return X[RandomIndex(len(X))], y[RandomIndex(len(X))]
end function
```

ulaşmasına yardımcı olabilir [29]. Ayrıca, algoritmanın non-convex ve gürültülü amaç fonksiyonlarıyla başa çıkma yeteneği, derin sinir ağları ve karmaşık makine öğrenimi modellerini eğitmek için özellikle uygun kılar. Güncellemeler algoritmanın karmaşık kayıp yüzeylerinde gezinmesine ve zorlu optimizasyon manzaralarında iyi çözümler bulmasına olanak tanır [30].

– Logistic Regresssion (LR) Algoritması:Logistic regression algoritması, bir veya daha fazla bağımsız değişkenin bir sonucu belirlediği bir veri setini analiz etmek için kullanılan istatistiksel bir yöntemdir. Genellikle, sonucun ikili olduğu ikili sınıflandırma problemleri için kullanılır. Algoritma, belirli bir girişin belirli bir kategoriye ait olma olasılığını tahmin ederek çalışır. Logistic regression algoritmasının temel avantajlarından biri, basitliği ve yorumlanabilirliğidir. Algoritma, her bir giriş özelliğinin önemini ve bir sonucun olasılığını nasıl etkilediğini anlamamıza yardımcı olur [31]. Ayrıca, gürültüye karşı dirençlidir ve kategorik ve sürekli giriş özelliklerini işleyebilir, bu da çeşitli uygulamalar için çok yönlü olmasını sağlar [32]. Logistic Regresssion sınıflandırmanın genel algoritması Algorithm 6'de gösterilmiştir. Ek olarak, logistik regresyon hesaplama açısından etkili ve yüksek hesaplama kaynaklarına ihtiyaç duymayan bir özellik taşır; bu da gerçek zamanlı uygulamalar ve sınırlı hesaplama kapasitesi olan senaryolar için uygundur [33]. Aynı zamanda, özellikle eğitim örneklerinin sayısı sınırlı olduğunda aşırı uydurmaya karşı daha az duyarlıdır ve giriş özellikleri arasındaki çoklu

**Algorithm 6** Logistic Regression Classification

```
function LogisticRegression(X, y, α, epochs)
    W, b ← InitializeWeights(X)
    for epoch ← 1 to epochs do
        W, b ← W − α · Gradient(X, Sigmoid(XW + b) − y, X)
    end for
    return W, b
end function
function Sigmoid(Z)
    return 1/(1 + exp(−Z))
end function
function InitializeWeights(X)
    W ← Array of zeros with shape(shape of X[1], 1)
    b ← 0
    return W, b
end function
function Gradient(X, error, input)
    return X^T · error/len(X), Sum(error)/len(X)
end function
```

doğrusallıkla başa çıkabilir [34]. Logistic Regression'un temel formülü Denklem 2.6'de görüldüğü gibi ifade edilebilir.

$$P(y=1|\mathbf{x}) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n)}} \tag{2.6}$$

Burada:

$P(y=1|\mathbf{x})$ : Veri $\mathbf{x}$ için sınıf $y=1$ olasılığı.

$\beta_0, \beta_1, \beta_2, \ldots, \beta_n$ : Model parametreleri.

$x_1, x_2, \ldots, x_n$ : Veri özellikleri.

## 2.3 Doğal Dil İşleme ve BERT

Doğal Dil İşleme (NLP), Yapay Zeka ve Dilbilim'in kesişim noktasında özel bir branşı temsil eder. Temel misyonu, bilgisayarların insan dillerini anlamalarını kolaylaştırmak ve onlara insan dilini yansıtan bir şekilde yazılı ifadeleri ve kelimeleri anlamalarını sağlamaktır. NLP'nin ortaya çıkmasındaki ana hedef, bilgisayarlarla kullanıcı etkileşimini basitleştirmek ve insan-bilgisayar iletişimini doğal bir dilbilimsel tarzda gerçekleştirmekti.

Bu alandaki gelişmeler, özellikle bilgisayar kullanıcılarının makineye özgü dillerde uzmanlık sahibi olma ihtiyacını ortadan kaldırmak ve herkesin bu özel dillerde yetenek kazanmak için büyük zaman ve çaba harcamasına gerek olmaksızın, bilgisayarlarla daha doğal bir şekilde iletişim kurmalarını sağlama amacını taşımıştır. NLP, bilgisayarların metinleri analiz etmesini, anlamasını ve hatta cevaplamasını sağlayan karmaşık algoritmalar, dil modelleri ve makine öğrenimi teknikleri kullanır.

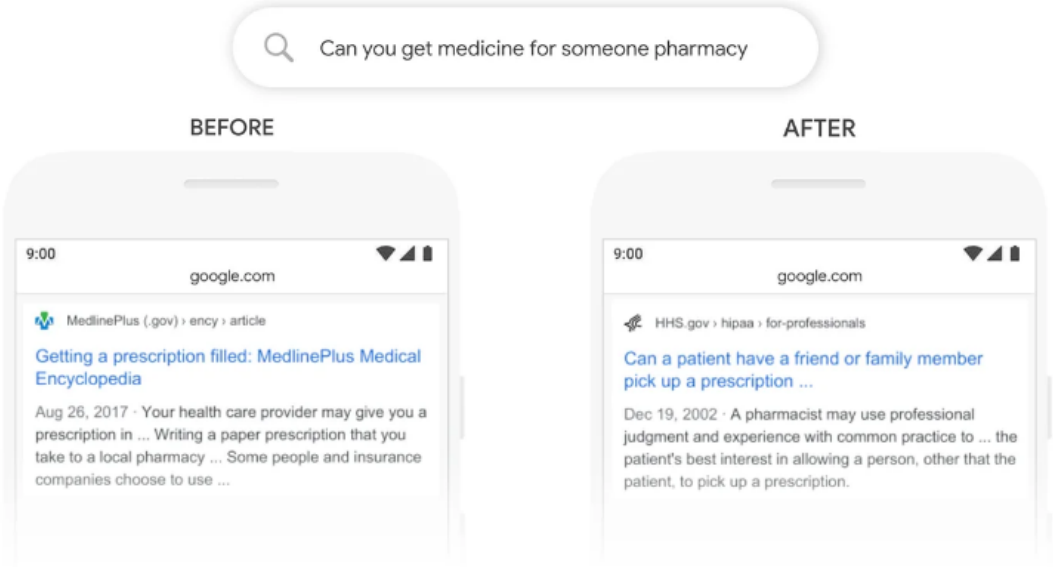
Doğal Dil İşleme, metin madenciliği, dil anlama, dil üretme, konuşma tanıma ve çeviri gibi çeşitli alt alanlarda uygulanabilir. Metin verilerini anlamak ve çıkarılan bilgileri kullanmak, NLP'nin temel görevlerindendir. Bu, büyük veri setlerinden anlam çıkarmak, duygu analizi yapmak ve özetleme gibi çeşitli uygulamalara olanak tanır.

NLP'nin evrimi, kullanıcı deneyimini geliştirmek ve bilgisayarların daha etkili bir şekilde insan dilini anlamalarını sağlamak için devam etmektedir. Bu alandaki araştırmalar, dil modellerinin, kelime gömme tekniklerinin ve derin öğrenme yöntemlerinin kullanımını içerir. Bu sayede, bilgisayarlar metinleri daha iyi anlayabilir, kullanıcı taleplerine daha doğru yanıt verebilir ve karmaşık dil yapılarını çözümleyebilir hale gelir [35].

BERT (Bidirectional Encoder Representations from Transformers), doğal dil işleme (NLP) alanında önemli bir derin öğrenme modelidir. Google tarafından 2018 yılında tanıtılan BERT, metni hem sağdan sola hem de soldan sağa doğru anlama yeteneğiyle öne çıkar. Bu model, tek yönlülük kısıtlamasını aşarak dil anlama görevlerinde çığır açıcı bir gelişme sağlamıştır [36].

BERT tanıtılmadan önce doğal dil işleme işlemleri için genellikle RNN ve CNN kullanılıyordu. Bu modeller çok kötü olmasa da BERT bu modellerin aksine verilerin sırayla işlenmesini gerektirmez. BERT ile veriler herhangi bir sırayla işlenebildiği için diğer modellerin çok daha fazla miktarda veri üzerinde train yapmak çok daha kolaylaşır [37].

BERT'in temel amacı, metinsel verilerin kapsamlı anlayışı, çözümü ve analizi için derin öğrenme tekniklerini kullanmaktır. Bu, metinsel içerikteki dilbilimsel yapıları




**Şekil 2.4 :** BERT Algoritmasının Cümleyi Algılama Farkı [2].

ve anlamsal nüansları anlamayı mümkün kılar. BERT'in bazı önemli uygulamaları şunlardır:

- Metin Sınıflandırma: BERT, metinsel verileri kapsamlı bir analizle belirli sınıflara veya kategorilere ayırma konusunda yeteneklidir. Bu uygulamalar arasında spam e-postaların tespiti ve duygu analizi bulunmaktadır. [37]
- Soru Cevaplama: BERT, bir metin belgesi veya kaynağa dayalı sorulara cevap verme konusunda uzmanlık gösterir. Kullanıcıların belirli bir metin belgesinin içeriği hakkında sorduğu sorulara yanıt verebilme kapasitesine sahiptir.
- Dil Çıkarımı: Şekil 2.4'de görüldüğü gibi BERT, bağlamın anlaşılmasını, çıkarımların formülasyonunu ve metin pasajları arasındaki anlamsal benzerliklerin ölçümünü gerektiren çeşitli NLP görevlerinde etkili bir araç olarak ortaya çıkar. Bu sayede aranılanı çok daha iyi anlar ve aranılan şeylere uygun sonuçlar ortaya koyabilir
- Metin Oluşturma: BERT, özel metinsel içerik oluşturmak için kullanılır ve bu esneklik, makaleler, denemeler veya metin tabanlı oyunlar gibi görevlerde kullanılır.
- Belge Sıralama: BERT, belirli kriterlere yanıt olarak metin belgelerinin sıralanması veya derecelendirilmesi konusunda değerli bir araçtır ve belirli

bir sorguya karşılık gelen en uygun metinleri belirleme konusunda yardımcı olabilir [37].

BERT, özellikle etiketsiz metinden türetilmiş derin çift yönlü temsillerin ön eğitimi için tasarlanmıştır. Bu ön eğitim süreci, modelin tüm katmanlarında hem sol hem de sağ bağlam bilgisinin eşzamanlı olarak dikkate alınmasını içerir. Bu nedenle, ön eğitimli BERT modeli, basit bir ek çıkış katmanının eklenmesi ile daha da optimize edilebilir. Bu optimize edilmiş fine-tuning prosedürü, geniş bir dizi görevde (soru cevaplama, dil çıkarsama gibi) uzmanlaşmış, ancak önemli ölçüde görev özel mimari ayarlamalarını gerektirmeyen modeller üretir [36].

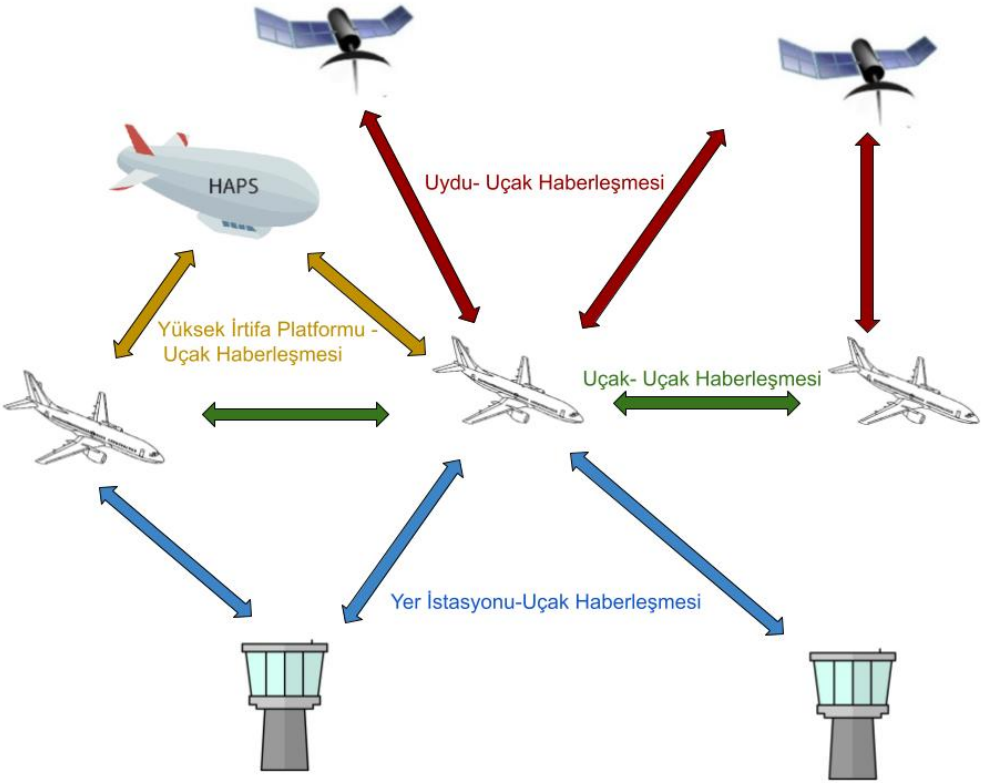
BERT'in Temel Prensipleri:

- BERT'in temel konsepti, kelime anlamını modelin tüm katmanlarında çift yönlü bir bağlam içinde anlama yeteneğidir. Bu çift yönlü yaklaşım, BERT'i önceki tek yönlü modellerden ayırarak dilin daha derinlemesine anlaşılmasını sağlar [37].
- BERT'in gücü, geniş ve çeşitli bir metin veri koleksiyonunda gerçekleştirilen kapsamlı ön eğitiminde yatmaktadır. Bu aşamada BERT, cümle içindeki eksik kelimeleri tahmin etme ve sağlanan metindeki karmaşık kelime ilişkilerini anlama sürecini öğrenir. Bu ön eğitim, BERT'e dilin derinlemesine bir anlayışını kazandırarak downstream NLP görevlerinde çok yönlü uygulamalara olanak tanır [37].
- Ön eğitimi takiben BERT, görev özel doğal dil işleme görevleri için bir görev özel çıkış katmanı eklenerek verimli bir şekilde fine-tune edilebilir. BERT, bu optimize edilmiş fine-tuning süreci aracılığıyla metin sınıflandırma, soru cevaplama ve dil çıkarsama gibi geniş bir NLP görev yelpazesinde başarılı olabilir.
- BERT, çeşitli dillerde yetenek göstererek çapraz-dilsel uygulamalar için çok yönlü bir seçenek haline gelir. Ayrıca, BERT'in mimarisi çoklu görev öğrenimini destekler, böylece tek bir ön eğitimli modelin farklı NLP görevlerini ayrı, görev özel modellere ihtiyaç duymadan karşılamasına olanak

tanır. Bu çok yönlülük, NLP çözümlerinin geliştirilmesini kolaylaştırır ve BERT'i doğal dil işleme alanında temel bir unsur olarak konumlandırır [37].

# 3. MIL-STD 1553 SİSTEMİNE YAPILAN SİBER SALDIRILAR

Mil-Std 1553 sistemi ilk olarak 1975 yılında geliştirildi. Mil-Std 1553 sistemi askeri bir veri iletim yolu olarak oldukça stabil, hızlı, yüksek dayanıma sahip düşük hata oranlı bir sistemdir. Geliştirildiği yıllar için oldukça güvenli ve sorunsuz bir yapısı vardı. O yıllarda teknoloji herhangi bir sisteme uzaktan erişip bozmaya, manipüle etmeye elverişli değildi. Bu nedenle uzun yıllar boyunca hava, deniz, kara ve hatta uzay araçlarında da kullanıldı.




**Şekil 3.1 :** Aviyonik Sistemlerde İletişim Mimarisi

Mil-Std 1553, 1970'lerde ortaya çıkan bir standart olduğu için, orijinal tasarımı siber güvenlik tehditlerinin günümüzdeki karmaşıklığına cevap verecek şekilde tasarlanmamıştır. Bu dönemde siber güvenlik, günümüzdeki kadar büyük bir endişe kaynağı değildi ve bus yapısını fiziksel erişim olmadan bu sistemi bozmak ya da manipüle etmek mümkün değildi. Fiziksel bir erişim olsa bile kolaylıkla

farkedilip sızma girişimi engellenebilirdi. Dolayısıyla dışardan gelebilecek bir saldırı öngürülmediği için herhangi bir savunma önlemi alınmadı ve tehditlere karşı güçlü bir savunma mekanizması eklenmedi. Ancak Aradan geçen uzun yıllar sonucu günümüzde gelişen teknoloji ile birlikte artık Mil-Std 1553 sistemi geçmişte olduğunun aksine siber tehditlere karşı oldukça savunmasız bir hale geldi.

Aradan geçen uzun yıllarda teknoloji oldukça gelişti ancak Mil-Std 1553 sisteminde kapsamlı bir güvenlik geliştirmesi yapılmadı. Bunun nedeni, Mil-Std 1553, geniş bir kullanım alanına sahiptir ve hava,kara,deniz,uzay araçları dahil birçok sistemde kullanılmaktadır. Bu sistemlerde radikal değişiklikler yapmak ve standartı güncellemek oldukça zorlu bir süreçtir. Mevcut sistemlerin bu tür güncellemelere uygun olması için maliyetli bir revizyon sürecine gerek duyulur. Bu durumun yanında kapsamlı bir değişiklik yaparak siber güvenlik önlemlerini entegre etmek, maliyetli bir dönüşüm sürecini ve geriye uyum sorunlarını gündeme getirebilir.

Mil-Std 1553 standardının askeri ve havacılık uygulamalarında yaygın olarak kullanılması, bu sistemlerin güvenliği üzerinde stratejik bir etki yaratır. Çünkü bu sistemler, ülkelerin savunma ve güvenlik stratejilerinin temelini oluşturan kritik bileşenlerdir. Bu nedenle, bu sistemlerin güvenliğinin sağlanması, ulusal savunma kapasitelerini güçlendirmek, stratejik avantajları korumak ve dış tehditlere karşı direnç sağlamak açısından hayati bir önem taşır.

Ancak siber saldırılara karşı olan savunmasızlığı, bu sistemi oldukça kötü etkilemektedir. MIL-STD 1553 bus yapısı üzerinde Şekil 3.1'de görüldüğü gibi konum ,yükselik, hız gibi son derece önemli olan navigasyon sistemleri, havadaki ve yerdeki diğer sistemlerle iletişimi sağlayan telsiz sistemleri, hava aracının olduğu bölgedeki trafiği izlemesini sağlayan gözetleme sistemleri ve eğer askeri bir hava aracı ise elektronik harp gibi güvenlik sistemleri bulunabilir. Bu sistemlerden herhangi birinin aksaması oldukça maliyetli sonuçlar doğuracaktır.

Şekil 3.2'de Mil-Std 1553 bus sistemine yapılabilecek siber saldırı tipleri gösterilmektedir. Bu saldırı tiplerini daha detaylı olarak inceleyelim:

A. Veri Manipülasyonu: Veri manipülasyonu, saldırganın sistem üzerinden geçen veriyi değiştirme veya bozma sürecidir. Saldırgan, veri terminaline erişerek iletilen


**Şekil 3.2 :** Mil-Std 1553 Güvenlik Tehditleri.

veya alınan veriyi manipüle edebilir. Bu, veri bütünlüğünü ve güvenilirliğini tehlikeye atabilir. Saldırgan, veri paketlerini ele geçirerek içeriklerini değiştirebilir veya yanlış bilgi gönderebilir. Örneğin, bir askeri uçakta Mil-Std 1553 protokolü ile birleşik bir silah sistemine ateş emri veriliyor olabilir. Saldırgan, bu veri paketini ele geçirerek içeriğini değiştirip yanlış hedeflere ateş emri verebilir veya silah sistemini devre dışı bırakabilir [3].

- Veri Paketi Değiştirme: Saldırgan, veri otobüsündeki iletişim trafiğini izleyerek, normalde beklenen veri paketlerini değiştirir. Saldırgan, veri otobüsündeki veri paketlerini yakalar, içeriğini değiştirir ve ardından değiştirilmiş veriyi hedef cihazlara gönderir. Bu, hedef sistemde yanıltıcı veri işlemlerine neden olabilir.
- Veri Paketi Enjeksiyonu: Saldırgan, veri otobüsüne sahte veri paketleri göndererek, hedef sistemde istenmeyen etkiler oluşturabilir. Saldırgan, veri otobüsüne sahte bir veri paketi ekleyerek, hedef cihazlara zararlı bilgiler gönderir. Bu, sistemi yanıltabilir veya normal işleyişi bozabilir. Veri paketi değiştirme ile çok benzerdir sadece sonuç açısından farklılık gösterir.

B. Komut Yönlendirme: Komut yönlendirme saldırıları, saldırganın komutları yanlış bir hedefe yönlendirmesini içerir. Mil-Std 1553 protokolü, veri terminali

arasında komut iletimi için kullanılır. Saldırgan, iletilen komutları ele geçirerek hedef terminali değiştirebilir veya yanlış bir cihaza yönlendirebilir. Bu durumda, istenmeyen eylemler gerçekleşebilir veya sistem işlevsiz hale gelebilir. Örneğin, bir askeri araçta Mil-Std 1553 protokolü kullanılan bir iletişim sistemi düşünelim. Saldırgan, bu sistem üzerinden yanlış bir komut göndererek hedeflenmeyen bir cihaza ateş emri verebilir veya sistemi yanlış çalıştırabilir. Bu durum, stratejik bir hedefin yanlış şekilde etkilenmesine veya askeri operasyonlarda istenmeyen sonuçlara yol açabilir [5].

- Komut İletim Paketlerini Manipüle Etme: Saldırgan, busdaki komut paketlerini değiştirerek, hedef cihazlara yanıltıcı veya zararlı komutlar gönderebilir. Saldırgan, normalde beklenmeyen veya hedef cihazın normal işlevselliğini bozan sahte komutları iletebilir.
- Güvenlik Zafiyetlerini Kullanma: Sistemdeki güvenlik açıklarını sömürerek, saldırganlar hedef sistemi etkileyebilir. Saldırgan, sistemdeki bir güvenlik açığından yararlanarak, hedef cihazlara zararlı komutlar gönderebilir veya sistemdeki komutları etkileyebilir.

C. Hizmet Reddi Saldırısı (DoS): Hizmet reddi saldırıları, bir sistemin normal işlevselliğini engelleyerek hedefin hizmet vermeme durumuna getirilmesini içerir. Mil-Std 1553 protokolünün bu tür bir saldırıya maruz kalması durumunda, sistemdeki veri iletimi ve kontrol işlemleri durabilir. Saldırgan, aşırı yükleme veya kaynak tüketimi gibi yöntemlerle sistem kaynaklarını tüketebilir veya sistem üzerindeki trafiği engelleyebilir. Bu, askeri operasyonlarda ciddi sonuçlara yol açabilir ve savunma sistemlerinin etkinliğini azaltabilir. Örneğin, bir askeri savaş gemisinde Mil-Std 1553 protokolü kullanılan bir veri haberleşme sistemi düşünelim. Saldırgan, sistem üzerinde aşırı miktarda veri trafiği oluşturarak sistem kaynaklarını tüketebilir. Bu durumda, veri iletimi kesintiye uğrayabilir ve geminin diğer sistemlerine veri aktarımı engellenebilir. Bu, savaş gemisinin etkinliğini azaltabilir ve kritik operasyonlarda sorunlara yol açabilir [4].

- Trafiği Aşırı Yükleme: Saldırganlar, sistemdeki veri otobüsünü aşırı yükleyerek, normal veri trafiğini bloke edebilirler. Saldırgan, sürekli olarak

sahte veri paketleri göndererek veri otobüsünü aşırı yükleyebilir. Bu, normal veri iletimini engeller ve sistemde hizmet kesintisine neden olabilir.

- İletim Kuyruklarını Doldurma: Sistemdeki veri iletim kuyruklarını doldurarak, diğer cihazların veri göndermesini engelleyebilir. Saldırgan, sürekli olarak sahte talepler göndererek veya iletim kuyruklarını manipüle ederek, bus üzerindeki iletişimi engelleyebilir.
- Sistemi Hatalı Konfigürasyonla Kilitleme: Sistem konfigürasyonunda hatalı bir şekilde yapılan değişikliklerle, normal işleyişi bozarak hizmet kesintisine neden olabilir. Saldırgan, sistem ayarlarını hatalı bir şekilde değiştirerek veya yanlış yapılandırarak, normal işleyişi etkileyebilir ve hizmetin kapanmasına neden olabilir.

D. İzleme/Casusluk: Mil-Std 1553 protokolü, askeri ve endüstriyel sistemler arasında veri paylaşımı ve iletişim için kullanılan bir standarttır. Bu protokol, veri paketlerini belirli bir format ve yapıda ileterek güvenli bir iletişim sağlar. Ancak, bu iletişim süreci, saldırganlar için bir hedef olabilir. Saldırganlar, Mil-Std 1553 protokolünü izleyerek veya casusluk yaparak, sistem üzerinden geçen veri paketlerini ele geçirebilir. Bu veri paketlerini analiz ederek, içerisindeki bilgileri okuyabilir ve hassas bilgilere erişebilirler. Bu durum, askeri operasyonlarda veya endüstriyel sistemlerde ciddi güvenlik açıklarına neden olabilir. Örneğin, bir askeri sistemde Mil-Std 1553 protokolü kullanılan bir haberleşme ağı düşünelim. Bu ağ aracılığıyla askeri birlikler arasında stratejik bilgiler paylaşılıyor olabilir. Saldırgan, bu ağı izleyerek iletilen veri paketlerini ele geçirebilir ve içerisinde yer alan askeri stratejik bilgilere erişebilir. Bu durum, askeri operasyonların gizliliği ve güvenliği açısından büyük bir risk oluşturur. Bu nedenle, Mil-Std 1553 protokolünü kullanan sistemlerde izleme ve casusluk saldırılarına karşı güçlü güvenlik önlemleri alınması önemlidir. Veri şifreleme, yetkilendirme ve erişim kontrolleri gibi güvenlik mekanizmaları, bu tür saldırıları önlemek ve sistemlerin güvenliğini sağlamak için kullanılabilir [4].

- Dinleme: Saldırgan, veri otobüsüne fiziksel veya mantıksal bir noktada erişim sağlar.

- Veri Paketi Yakalama: Saldırgan, iletişim trafiğini kaydetmek için casus bir cihaz veya yazılım kullanır. Bu sayede, veri otobüsünde iletilen paketlerin içeriğini görebilir.
- Analiz ve Keşif: Elde edilen verileri analiz ederek, sistemdeki cihazların kimlik bilgilerini, iletişim protokollerini ve kullanılan komutları belirler. Bu adım, hedef sistemin zayıf noktalarını tespit etmeye yardımcı olur.
- Gizli Bilgi Ele Geçirme: Saldırgan, ele geçirilen bilgileri kullanarak sistemin güvenlik önlemlerini atlamaya çalışır. Hassas verilere ulaşabilir ve gelecekteki saldırılar için planlama yapabilir.

Mil-Std 1553 standardının kullanıldığı askeri ve havacılık sistemlerindeki güvenlik ihlalleri veya başarısızlıklar, potansiyel olarak insan hayatını riske atabilir. Bu sistemler genellikle taktik savaş uçakları, helikopterler veya insansız hava araçları gibi kritik görevlerde kullanıldığından, güvenlik açıkları veya hatalı çalışmalar ciddi sonuçlara yol açabilir. Güvenilir bir veri iletim sistemine sahip olmak, bu tür araçların personeli ve operatörleri için hayati öneme sahiptir. Ayrıca sisteminin güvenliği, operasyonel görevlerin başarıyla tamamlanması açısından kritiktir. Bu sistemler, askeri keşif, gözetleme, ve saldırı operasyonları gibi önemli görevlerde kullanılabilmektedir. Güvenli olmayan veya güvenilir olmayan bir veri iletim sistemi, iletişim hatası veya veri bütünlüğü sorunlarına yol açarak, operasyonların başarısız olmasına veya hedeflere ulaşılamamasına neden olabilir. Bu da stratejik hedeflere ulaşmada ve ulusal savunma yeteneklerini güçlendirmede sorunlara yol açabilir. Mil-Std 1553 sisteminin güvenliği, içsel ve dışsal tehditlere karşı etkili bir koruma mekanizması sağlamayı gerektirir. Bu sistemlerin ilettiği veriler genellikle hassas askeri bilgileri içerdiğinden, yetkisiz erişimlere karşı önlem alınmalı ve veri bütünlüğü sağlanmalıdır.

## 3.1 Değerlendirme Metrikleri

Makine öğrenimi ve doğal dil işleme modellerinin performansını değerlendirmek için çeşitli metrikler kullanılmaktadır. Bu metrikler, modelin sınıflandırma yeteneğini ve tahmin doğruluğunu değerlendirmek amacıyla kullanılır. Bu

çalışmada, model performansının ölçülmesi için binart ortalama F1 skoru, makro ortalama F1 skoru, ağırlıklı ortalama F1 skoru, doğruluk ve karışıklık matrisi (confusion matrix) gibi metrikler kullanılmıştır.

F1 skoru, bir sınıflandırma modelinin başarısını ölçen bir metriktir ve genellikle hassasiyet (precision) ve duyarlılık (recall) değerlerinin birleştirilmiş bir değerini ifade eder. F1 skoru, özellikle dengesiz sınıflar veya yanlış pozitif ve yanlış negatif hataların eşit öneme sahip olduğu durumlarda kullanışlıdır. Precision, pozitif olarak tahmin edilen durumların gerçekten pozitif olma olasılığını ölçer ve denklem 3.1'de gösterildiği şekilde hesaplanır. Recall, gerçekten pozitif olan durumların ne kadarını doğru bir şekilde tahmin ettiğimizi ölçer. Denklem 3.2'de gösterildiği şekilde hesaplanır. Precision ve recall değerleri dengeli olduğunda F1 skoru yüksek olur. F1 skoru Denklem 3.3'de belirtildiği gibi hesaplanır.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3.1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3.2}$$

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3.3}$$

Makro ortalama (Macro Average) F1 skoru, her sınıfın ayrı ayrı hesaplanan F1 skorlarının ortalamasıdır. Bu metrik, her sınıfın performansını eşit şekilde değerlendirir, bu nedenle dengesiz sınıf dağılımlarına karşı daha dirençlidir.

$$\text{Macro F1 Score} = \frac{\sum_i \text{F1 Score}_i}{\text{Toplam Sınıf Sayısı}} \tag{3.4}$$

Makro ortalama F1 skoru, her sınıfın ayrı ayrı hesaplanan F1 skorlarının ortalamasıdır. Bu metrik, her sınıfın performansını eşit olarak değerlendirir. Makro ortalama F1 skoru denklem 3.4'de gösterildiğişekilde hesaplanır.

Weighted average F1 skor(ağırlıklı ortalama F1 skoru), her sınıfın ağırlığıyla çarpılmış F1 skorlarının toplamının normalizasyonu ile hesaplanır. Bu metrik,

dengesiz sınıf dağılımlarını göz önünde bulundurur. Ağırlıklı ortalama F1 skoru formülü denklem 3.5'de gösterildiği gibidir.

$$\text{Weighted F1 Score} = \frac{\sum_i (\text{F1 Score}_i \cdot \text{Sınıf Ağırlığı}_i)}{\sum_i \text{Sınıf Ağırlığı}_i} \tag{3.5}$$

Doğruluk (Accuracy) Doğruluk, doğru tahmin edilen örneklerin toplam örnek sayısına oranını ifade eder. Bu metrik, genel model doğruluğunu değerlendirmek için kullanılır. Accuracy hesaplamının yolu denklem 3.6'de gösterildiği şekildedir.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{3.6}$$

Karışıklık matrisi (confusion matrix) bir sınıflandırma modelinin performansını değerlendirmek için kullanılan bir metrik matristir. her bir sınıfın tahminlerini gerçek sınıf etiketleriyle karşılaştırır. Genellikle dört temel değeri içerir: TP (doğru pozitif- true positive (TP), TN (true negative-doğru negatif), FP (false positive-yanlış pozitif), FN (false negative-yanlış negatif).

$$\begin{bmatrix} \text{TP} & \text{FP} \\ \text{FN} & \text{TN} \end{bmatrix} \tag{3.7}$$

Karışıklık matrisindeki gösterimi 3.7'de belirtilmiştir. Her bir değerin açıklaması ise şu şekildedir:

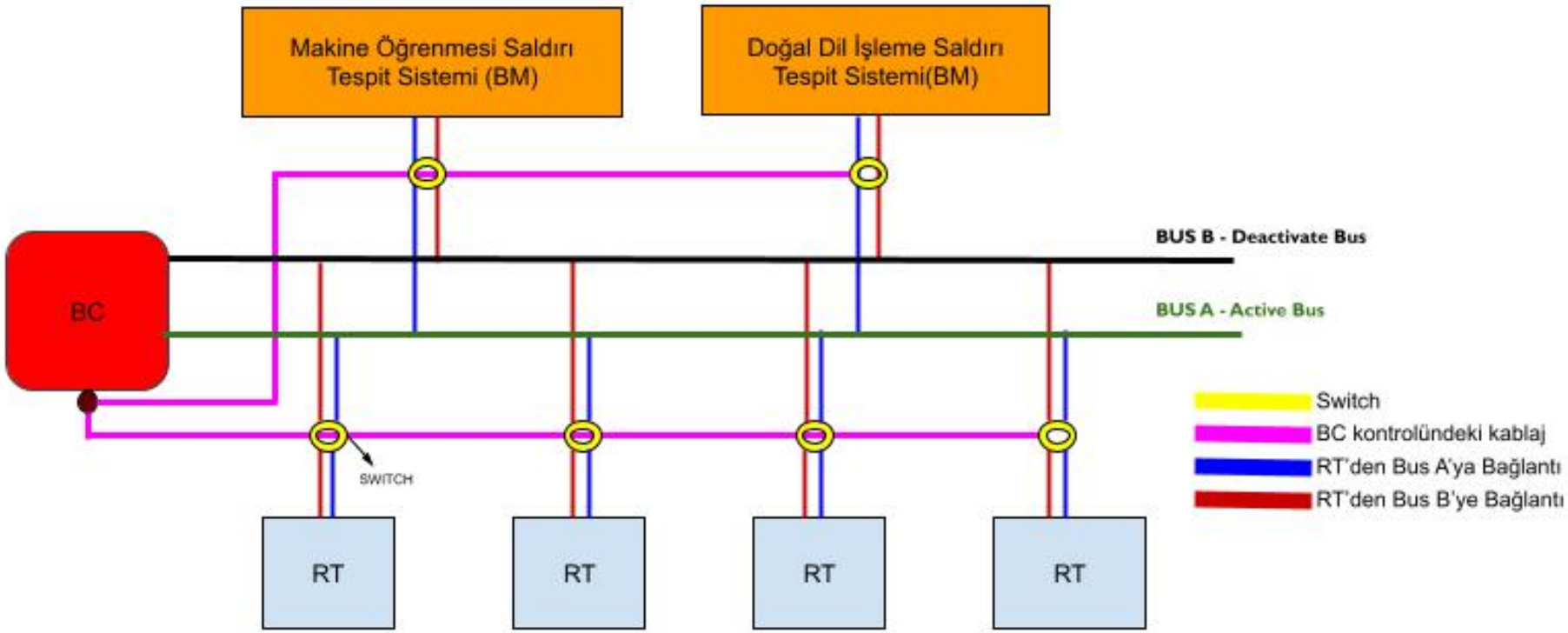
- True Positive (TP): Gerçekte pozitif olan ve model tarafından doğru bir şekilde pozitif olarak sınıflandırılan örnek sayısı. Yani, modelin doğru bir şekilde pozitif tahmin yaptığı durum.
- True Negative (TN): Gerçekte negatif olan ve model tarafından doğru bir şekilde negatif olarak sınıflandırılan örnek sayısı. Yani, modelin doğru bir şekilde negatif tahmin yaptığı durum.
- False Positive (FP): Gerçekte negatif olan ancak model tarafından pozitif olarak yanlış bir şekilde sınıflandırılan örnek sayısı. Yani, modelin yanlış bir şekilde pozitif tahmin yaptığı durum.

- False Negative (FN): Gerçekte pozitif olan ancak model tarafından negatif olarak yanlış bir şekilde sınıflandırılan örnek sayısı. Yani, modelin yanlış bir şekilde negatif tahmin yaptığı durum.

Bu değerlendirme metrikleri, modelin genel performansını anlamak için kapsamlı bir görünüm sunar ve çeşitli metriklerin kombinasyonu, modelin belirli güçlü ve zayıf yönlerini anlamak için kullanılır.

# 4. MIL-STD 1553 SİSTEMİNDE HİBRİT SALDIRI TESPİT SİSTEMİ

Bu çalışmada önerilen sistem birden fazla saldırı tespit ve koruma sisteminin birleştirilmesiyle elde edilen birkaç alt parçadan oluşan bir sistemdir. Önerdiğimiz sistem, Mil-Std 1553 sistemini hedef alan potansiyel saldırıları saptama ve bu saldırılara karşı etkin bir güvenlik çözümü sunma amacını taşımaktadır. Projemiz, Şekil 4.1'de gösterildiği gibi makine öğrenimi ve doğal dil işleme tekniklerini içermekle birlikte, aynı zamanda Mil-Std 1553 veri yolu yapısına dışarıdan entegre edilen ve bus controller kontrolüne verilen anahtar (switch) aracılığıyla genel bir saldırı tespiti ve koruma sistemi sağlamayı hedeflemektedir. Sistemin temel odak noktalarından biri, Mil-Std 1553 veri trafiğini anlamak ve analiz etmek suretiyle olası anormallikleri belirlemektir. Bu, makine öğrenimi algoritmalarını ve BERT algoritmasını kullanarak sistemin normal işleyişinden sapmaları tespit etmeyi amaçlar. Ayrıca, dışarıdan eklenen switch aracılığıyla bus yapısına güçlü bir savunma katmanı ekleyerek, bus kontrolünün dış etkenlere karşı güvenliğini sağlamayı hedefler.




**Şekil 4.1 :** Mil-Std 1553 Saldırı Önleme Sistemii

## 4.1 Makine Öğrenmesi ile Anomali Tabanlı Intrusion Tespiti

Mil-Std 1553, askeri ve havacılık sistemlerinde temel bir rol oynayan bir seri veri bağlantı standardıdır. Bu standart, çeşitli askeri platformlarda, uçaklarda, helikopterlerde, askeri araçlarda ve uzay araçlarında kullanılan veri iletişim sistemlerini standartlaştırmayı amaçlayarak güvenli ve güvenilir bir veri iletişim altyapısı sağlamaktadır. Mil-Std 1553, veri yolu üzerindeki cihazlar arasındaki koordinasyonu ve iletişimi düzenlemek için kullanılan bir protokol setini içermektedir. Bu standardın kullanıldığı sistemler, büyük ve karmaşık veri setleri üretir, bu da veri iletişimini yönetmeyi ve anlamlandırmayı zorlaştırır.

Bu bağlamda, Mil-Std 1553 sistemi üzerinde makine öğrenimi tabanlı bir anomaly detection yaklaşımı geliştirmenin önemi ortaya çıkmaktadır. Veri yolu üzerindeki karmaşık yapı, potansiyel anormalliklerin tespiti ve bu anormalliklere etkili bir şekilde müdahale etme gerekliliğini doğurur. Mil-Std 1553 veri yolu üzerinde iletilen mesajlar, farklı tiplerde ve içeriklerde olabilir. Mesajların içerdikleri bilgiler, iletim frekansları ve veri yapısı, sürekli olarak değişebilir. Bu durum, veri yolu üzerindeki dinamik yapının ana nedenlerinden biridir. Farklı cihazlar arasındaki iletişimde kullanılan mesaj tipleri ve içerikleri, standart protokollere rağmen geniş bir çeşitlilik gösterebilir. Ayrıca Mil-Std 1553 yüksek hızlı iletişim ihtiyacını karşılamak üzere tasarlanmıştır. Veri yolu üzerindeki mesajların hızlı bir şekilde iletilmesi, veri setinin hızla büyümesine neden olabilir. Hızlı iletişim, veri yolunu takip etmeyi ve hızla değişen veri setlerini izlemeyi zorlaştırır. İnsan gözüyle bu hızlı değişimleri takip etmek ve anlamak genellikle mümkün değildir. Hızlı iletişimden dolayı veri seti oldukça hızlı büyür. Ayrıca bu büyük veri setleri, çeşitli cihazlardan gelen farklı veri türlerini içerebilir. Büyük veri setleri içindeki bu çeşitlilik, geleneksel yöntemlerle manuel olarak veriyi analiz etmeyi ve anlamlandırmayı zorlaştırır.

Bunlarla birlikte Mil-Std 1553 protokolünün yapısından dolayı ortaya çıkan ve anomaly tespitini zorlaştıran bazı etmenler vardır. Mil-Std 1553 sistemi, kendi veri yolu protokollerii takip eder. Bu protokoller, belirli bir formatta veri iletilmesini ve

belirli iletişim kurallarının takip edilmesini sağlar. Ancak, protokollerin karmaşıklığı, geleneksel yöntemlerin bu protokoller altında çalışmasını zorlaştırabilir.

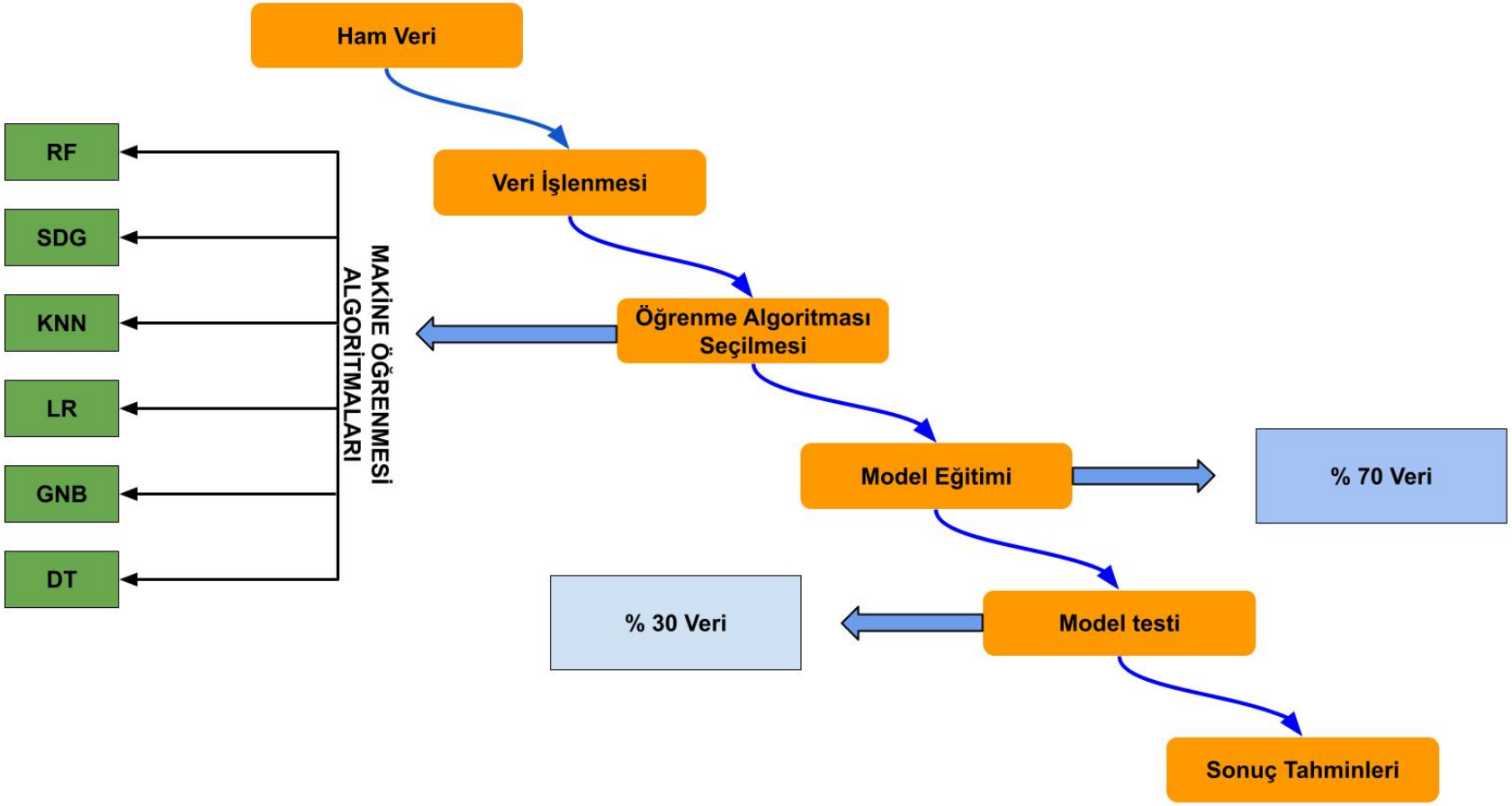
Mil-Std 1553 protokolündeki mesaj uzunluğu ve mesajlaşmanın karışık yapısı, anormallik tespiti için belirli bir model oluşturmayı zorlaştırabilir. Bu çalışma, Mil-Std 1553 sistemi içindeki veri yolu üzerindeki anormallikleri tespit etmek için makine öğrenimi algoritmalarının kullanılabilirliğini ve etkinliğini değerlendirmeyi amaçlamaktadır.

Makine öğrenimi ile anomali tabanlı saldırı tespiti önermemizin en temel sebepleri şu şekildedir:

- Mesaj Uzunluğu: Mil-Std 1553 veri yolu üzerinde iletilen mesajların uzunluğu, iletilen bilgi miktarını belirler. Mesaj uzunluğundaki bir anormallik, belirli bir mesajın normalden sapmasını ifade edebilir. Bu durum, bir veri setinin beklenmeyen bir şekilde büyüyüp küçülmesi veya bir mesajın sıradışı bir uzunluğa sahip olması şeklinde ortaya çıkabilir. Makine öğrenimi modelleri, bu tür anormallikleri tespit edebilir ve modelin normal işleyişe ait bir referans noktası oluşturarak anormal durumları belirleyebilir.
- Mesaj Yapısı: Mil-Std 1553 veri yolu mesajları belirli bir formatta iletilir ve bu formatta belirli alanları içerir. Mesaj yapısındaki anormallikler, örneğin bir mesajın beklenen sıraya göre gelmemesi veya belirli bir alanın hatalı olması durumlarında ortaya çıkabilir. Bu sapmalar, bir cihazın veya sistemdeki bir değişikliğin habercisi olabilir. Makine öğrenimi, mesaj yapısındaki bu tip anormallikleri belirleyebilir ve öğrenme sürecinde bu yapıları dikkate alarak normal işleyişle ilgili bir model oluşturabilir.
- Mesaj İçeriği: Mesajların içeriği, iletilen verinin türünü ve içeriğini belirler. Bu içerik, tipik olarak belirli bir formatta ve veri tipinde gelmelidir. Mesaj içeriğindeki anormallikler, verinin beklenen değerlerden sapması veya anormal durumlarla ilgili belirli bir deseni içermesi şeklinde ortaya çıkabilir. Makine öğrenimi modelleri, bu içerik bazlı anormallikleri belirleme konusunda etkili olabilir.

– Zamanlamaya Bağlı Anormallikler: Mil-Std 1553 sistemi, belirli bir zamanlama ve frekansta çalışan bir bus standardıdır. Mesajların belirli aralıklarla ve belirli bir zaman diliminde iletilmesi beklenir. Zamanlamaya bağlı anormallikler, bu belirlenen zaman aralıklarından sapma durumlarını ifade eder. Bu sapmalar, bir sistemin normal işleyişinden ayrıldığını ve potansiyel bir güvenlik tehdidi oluşturabileceğini gösterebilir.

  Zamanlamaya bağlı anormallikleri daha detaylı anlamak için, bir Mil-Std 1553 veri yolu sistemi ele alalım. Bu sistemde, belirli bir cihazdan belirli bir süre içinde belirli bir frekansta mesajlar beklenir. Ancak, bir saldırı veya sistemin içsel bir hatası nedeniyle bu zamanlamaya uymayan mesajlar gönderilirse, bu durum anormal bir durumu işaret eder. Örneğin, bir mesajın belirlenen zaman aralıklarından çok önce veya sonra gönderilmesi, zamanlamaya bağlı bir anormallik oluşturabilir.




**Şekil 4.2 :** Makine Öğrenmesi ile Saldırı Tespit Süreçleri

Makine öğrenimi modelleri, bu zamanlamaya bağlı anormallikleri tespit edebilir. Model, belirlenen zaman aralıkları içinde beklenen mesajları ve aralıkları analiz ederek sistemin normal işleyişini öğrenir. Ardından, bu öğrenilen modele dayanarak, gerçek zamanlı mesajları değerlendirir ve beklenmeyen zamanlamalara sahip olanları belirler. Bu sayede, sistemdeki olası güvenlik açıkları veya hatalar hakkında uyarılar oluşturabilir.

– Toplu Anormallikler: Toplu anormallikler, mesaj uzunluğu, yapısı ve içeriği üzerinde birbirleriyle ilişkili olarak ortaya çıkan anormal durumları ifade eder. Bu durumlar genellikle izole anormalliklerin bir kombinasyonu olarak ortaya çıkar ve birlikte değerlendirildiğinde sistematik bir sorunu gösterebilir.

  Örneğin, bir saldırganın sisteme sızma girişimi sonucunda mesaj uzunluğunda bir değişiklik olabilir ve aynı zamanda mesaj yapısında tutarsızlıklar görülebilir. Bu durum, tek bir anormallikten ziyade bir dizi birbirine bağlı anormalliği içerebilir. Makine öğrenimi modelleri, bu tür toplu anormallikleri belirleyebilir, izole durumlardan daha kapsamlı bir perspektif sunabilir ve sistematik sorunları tespit ederek daha güçlü bir güvenlik değerlendirmesi yapabilir. Model, mesaj uzunluğu, yapısı ve içeriği üzerindeki anormallikleri bir araya getirerek, sistemin bütünlüğünü ve güvenliğini daha etkili bir şekilde değerlendirebilir.

  Bu sebeplerden dolayı, hibrit Mil-Std 1553 güvenlik sisteminin bir parçası olarak makine öğrenimi ile anomali tabanlı saldırı tespit sistemi oluşturmayı öneriyoruz. Makine öğrenmesi için 6 farklı algoritmayı kullanarak bir sistem oluşturacağız. Bu algoritmalar bu algoritmalar Random Forest, Decision tree, K-Nearest Neighbors, Gaussian Naive Bayes, Stochastic Gradient Descent, Linear Regresssion'dır. Farklı algoritmaların kullanılması, farklı boyuttaki veri setlerine uyum sağlamak için oldukça önemlidir. Veri setinin boyutu birçok durumda algoritmanın performansı için en belirleyici girdidir. Büyük boyutlu veri setlerinde bazı algoritmalar çok daha iyi performans gösterirken daha küçük boyutlu veri setlerinde aynı performansı göstermeyebiliyor. Ya da eğitim kısmında yeterince örnek olmaması bazı algoritmaları çok daha kötü etkilerken diğerlerini nispeten daha az kötü etkiliyor. Bu durumların ve en

verimli performans sonuçlarının hangi şartlar altında oluştuğunu göstermek ve kullanılacak algoritmayı seçmektir.

Makine öğrenmesi tabanlı anomali tespiti için ilk olarak Şekil 4.2'de görüldüğü gibi üç farklı ham veri seti oluşturuldu. Bu veri setleri 6547, 15363 ve 32262 satır veriden oluşan bir Mil-Std 1553 bus haberleşmesinin benzerleridir. Bu veri seti bir simülatör verisi olarak düşünülebilir ve gerçek Mil-Std 1553 bus haberleşmesine oldukça yakındır. Her bir mesaj, data wordler, parity bitleri, kaynak adresleri, hedef adresleri, mod komutları gibi çok sayıda veriden oluştuşur. Bu kadar çok sayıda veriden oluşması makine öğrenimi için hem eğitim hem test kısımlarında gereken süreyi attırmaktadır. Bu nedenle veri setleri saldırı tespiti için önemli olan hiçbir veriye dokunulmadan sadeleştirilmiştir. Daha sonra makine öğrenmesi ile eğitilmesi ve test edilmesini kolay hale getirilmek için yapıları değiştirilmiştir. Yeni yapıda bütün mesaj yapısı sayısal değerlere dönüştürülmüştür. Bu işlem veri setleri ile makine öğrenmesi algoritmalarına daha uyumlu hale getirilmiş,tespit oranları artırılmış ve eğitim-test için gerekli süre azaltılmıştır.

Bu noktadan sonra makine öğrenmesi algoritması seçme noktasına gelinmektedir. Buradaki her bir algoritmanın diğerlerine göre avantajlı olduğu ya da dezavantajlı olduğu yerler vardır. Random forest yüksek doğruluk ve overfitting'e karşı dirençlidir. Ancak büyük veri setlerinde eğitim süresi oldukça uzundur ve model yapısı daha karmaşıktır. Decision tree açık ve anlaşılır bir model oluşturur. Karar ağacının dalları ve düğümleri, sistemdeki anomali tespitinde hangi özelliklerin ve değerlerin belirleyici olduğunu anlamayı kolaylaştırır. Ancak overfitting'e çok yatkındır. K-nearest neighbors basittir yüksek performanslıdır ve öğrenme süresi yoktur ancak veri setindeki değişkenlerin sayısı arttıkça ve veri seti büyüdükçe performansı düşer. Gaussian naive bayes hızlıdır ve küçük veri setlerinde etkili bir performans gösterir ancak değişkenler arasındaki bağımlılıkları hesaba katamaz, karmaşık ilişkileri modelleme konusunda sınırlıdır. Bu nedenle değişken sayısı arttıkça yani Mil-Std 1553 içinde RT sayısı ve onların esajları çoğaldıkça performansı düşecektir. Stochastic gradient descent büyük ver setlerinde ve

train datasının az olduğu durumlarda etkilidir. Ancak iyi performans almak için modelin parametrelerini çok hassas şekilde ayarlamak gerekir. Linear Regression ise yeterli sayıda eğitim ile iyi performanslar gösterir ancak çok karmaşık bağlantılar olursa yani yine bus üzerindeki cihaz sayısı çok artarsa performansı düşmeye başlar. Bu kısımda 6 farklı algoritmamız ihtiyacımız olan bir çok soruna cevap veriyor. Veri setinizin büyüklüğü mesaj sayısı Mil-Std 1553 sistemi üzerinde yer alan remote terminal sayısı, her lru'nun iletşimde kullandığı mesajlar, bus monitor olup olmadığı gibi etmenler hangi algoritmanın seçileceği konusunda önemlidir.

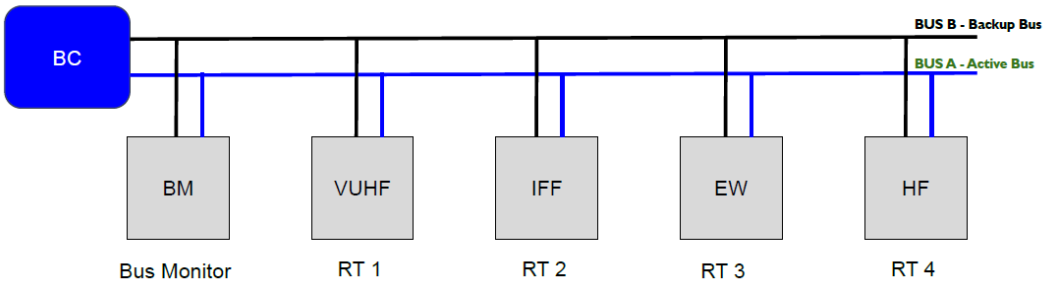
Algoritma seçimi yapıldıktan sonra modelin eğitimi başlatılacaktır. Burada Şekil 4.2'de belirtildiği gibi verisetinin % 70 oranında eğitim seti olarak ayarlanmıştır. Seçilen machine learning algoritması eğitimi tamamladıktan sonra modelin test süreci başlayacaktır. Modelin test sürecin veri setinin % 30 luk kısmı ile yapılacaktır. Test sürecinin de tamamlandıktan sonra ilgili verilerin doğru tahin oranları confusion matrix verileri elde edilecektir.

## 4.2 Doğal Dil İşleme ile Anomali Tabanlı Intrusion Tespiti

Mil-Std 1553 sistemlerinde şifreleme, kimlik doğrulama ve ek güvenlik protokollerinin eksikliği, ayrıca sınırlı güncelleme yetenekleri, sistemi çağdaş teknoloji ortamında önemli güvenlik zafiyetlerine karşı savunmasız kılmaktadır. Bu temel zayıflık, Mil-Std 1553'ü siber saldırılara karşı hassas hale getirir.

Mil-Std 1553 iletileri, komut kelimeleri, veri kelimeleri, durum kelimeleri ve teklik bitleri gibi çeşitli ileti segmentlerini içerir. Sonuç olarak, tek bir ileti için mesaj boyutu oldukça geniş olabilir ve iki ileti arasındaki zaman çok kısa milisaniye cinsindendir. Bu mesajların karmaşık, yoğun ve uzun yapısı, Mil-Std 1553 veri yolundaki sınıflandırma ve sızma tespiti göreviyle görevlendirilmiş makine öğrenme algoritmaları için önemli bir zorluk oluşturur.

Bu zorluğa çözüm olarak, başlangıçta karmaşık cümleleri anlama ve kategorize etme amacıyla geliştirilen BERT yöntemi uygulanmıştır. Mil-Std




**Şekil 4.3 :** Normal Çalışan Mil-Std 1553 Bus Yapısı.

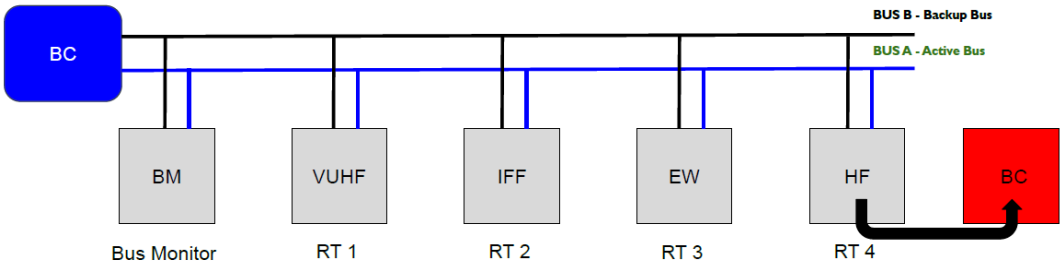
1553 mesajlarına benzeyen karmaşık cümle yapısı analiz edilerek, veri kümesinde güncellemeler yapılmıştır. Mesajlardaki sütunlar kaldırılmış ve her bir ileti, tek bir sıra ve tek bir sütun formatına dönüştürülmüştür. Bu yapılandırma, mesajları günlük konuşma cümlelerine daha benzer hale getirmeyi amaçlamıştır. Sonuç olarak, BERT algoritması, bu karmaşık mesajları etkili bir şekilde sınıflandırmak için adapte edilmiştir.

Eğitim süreci, BERT algoritmasını eğitmek için siber tehditleri içeren iletileri kullanmayı içeriyordu. Ardından, eğitilen BERT modelinin etkinliği, geriye kalan iletiler üzerinde test edilerek değerlendirildi ve Mil-Std 1553 iletişim ortamındaki karmaşık tehditleri başarıyla sınıflandırma ve tespit etme yeteneğini sergiledi.

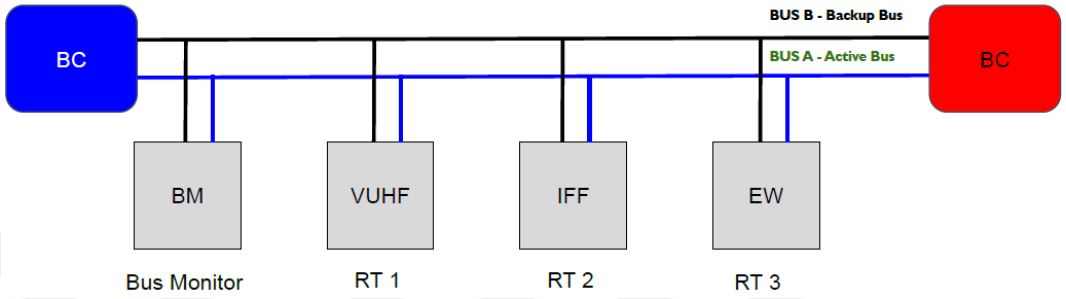
## 4.3 BC Çakışması ile Saldırı Tespiti

Standart bir Mil-Std 1553 busında bir tane bus controller bulunur. Yedekli sistemlerde ise 1 adet aktif Bus controller 1 adet te pasif durumda bus controller bulunabilir. Pasif durumdaki bus controller, sistemdeki aktif BC çalışmayı sürdürdükce, bus üzerinde bus monitor veya herhangi bir remote terminal olarak görev yapabilir [38].

Şekil 4.3 ile belirtilen standart Mil-Std 1553 bus haberleşmesinde 1 adet bus controller, 1 adet bus monitor ve 4 adet remote terminal bulunmaktadır. Mil-Std 1553 protokolüne uygun olarak 2 tane bus hattı vardır: Bus A ve Bus B. Bus iletişiminde herhangi bir sorun olmadığı için Bus A aktif olarak kullanılmaktadır ve Bus B backup olarak beklemektedir.




**Şekil 4.4 :** Kötü Niyetli RT'nin BC Olma Çabası.

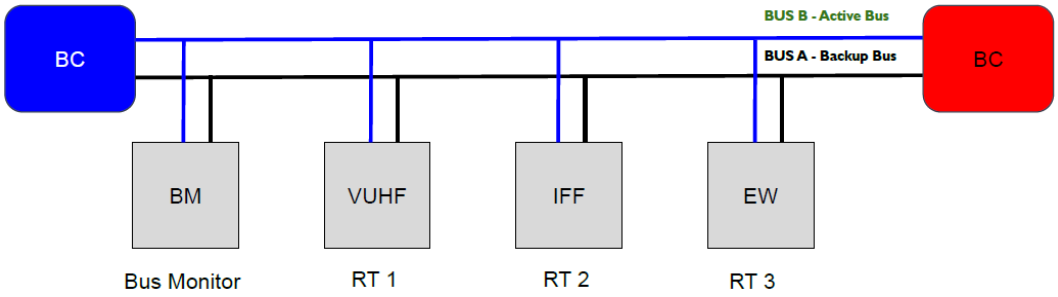



**Şekil 4.5 :** BC'ye dönen Kötü Niyetli RT.

Ancak normalde RT olarak çalışan cihazlardan bir tanesi kötü niyetli bir yazılıma sahipse veya kötü niyetli bir yazılım tarafından ele geçirilmişse kendini RT yerine BC olarak gösterip busı yönetmeye çalışabilir.

Şekil 4.4 ile belirtilen Mil-Std 1553 bus haberleşmesinde Şekil 4.3 'de RT olarak çalışan HF cihazı kötü niyetli bir yazılıma sahip olduğu için bus çalışırken ve BC varken RT yerine BC olarak çalışmaya niyetleniyor.

Şekil 4.4 de daha önce RT iken BC çalışma kararı veren cihaz Şekil 4.5'de BC olarak Mil-Std 1553 bus haberleşmesine katılıyor. Bu anda Mil-Std 1553 bus üzerinde iki tane BC olduğu anlaşılıyor. Mil-Std 1553 sisteminde iki tane aktif bus controller aynı anda bus'ı kontrol etmeye çalıştığında, bus mücadelesi (bus contention) meydana gelir. Bu durumda, bus'ın kontrolünün kim tarafından ele geçirileceği belirsiz hale gelir ve bus'ın düzgün çalışması riske girer. Mil-Std 1553, hangi bus controller'ın aktif olarak kalacağını belirleyen spesifik bir otomatizasyon stratejisi veya standart bir protokol sunmaz. Bunu sağlamak için sistemin genel tasarımına ve uygulamanın ihtiyaçlarına bağlı değişebilecek şekilde bazı adımlar atılabilir. Ancak bu adımlar zorunluluk değildir ve bütün Mil-Std 1553 sistemlerinde olmak

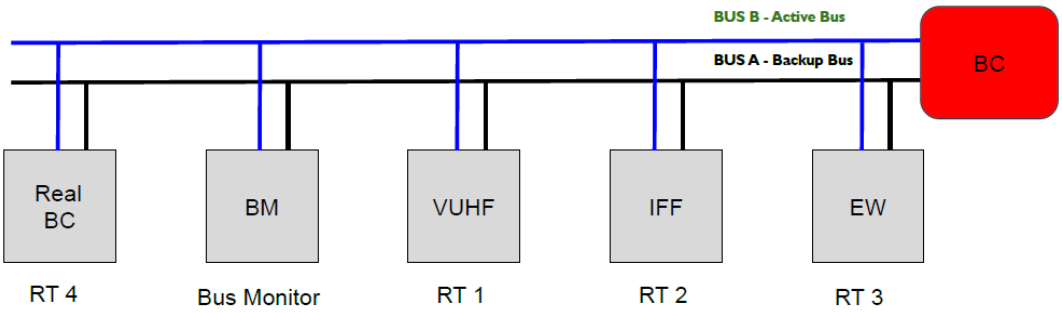



**Şekil 4.6 :** Bus A'da 2 Aktif Bus Olmasıyla Bus B'ye Geçilmesi.

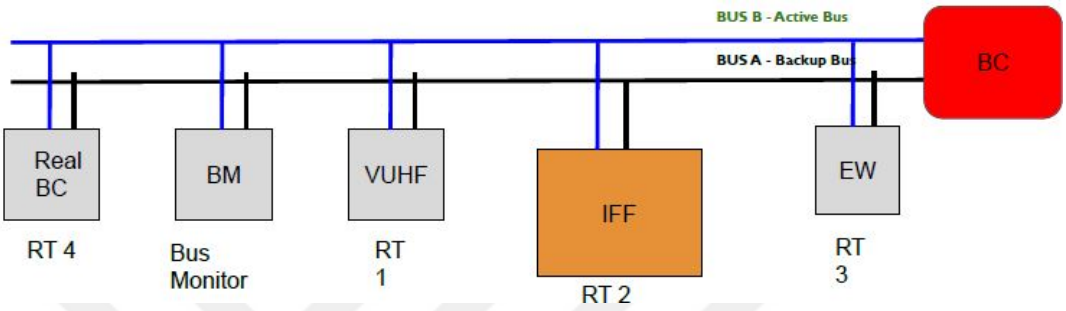
zorunda değildir. Mil-Std 1553 protokolünü kullanan kullanıcının kendi istekleri doğrultusunda şekillenir. Bu nedenle buradan sonra anlatılacak çözüm önerileri, kullanıcının tercihen Mil-Std 1553 tasarımına eklemiş olduğunu düşündüğümüz kısımlardır. Bu durum ortaya çıktığında Mil-Std 1553 tasarımında şu şekillerde sorun çözülmeye çalışılabilir:

* Öncelik Sıralaması: Mil-Std 1553 yazılımı ile, her BC ve RT için belirli bir öncelik seviyesi belirlenebilir. Öncelik seviyesi yüksek olan cihazlar, bus kontrolünü kazanma hakkına sahiptir. Öncelik seviyeleri, sistem tasarımı sırasında atanır ve belirli bir protokol kurallarına göre düzenlenir.
* Arabuluculuk: Bus mücadelesi yaşandığında, bus kontrolünün kim tarafından ele geçirileceğini belirlemek için arabuluculuk süreci başlar. BC'ler ve RT'ler, bus'ı kontrol etmek için arabuluculuk bitlerini kullanırlar.

Arabuluculuk süreci sonunda, bus kontrolünü kazanan cihaz, bus üzerinde komutlar ve veri transferleri gerçekleştirmek için yetkilendirilir. Diğer cihazlar ise bus kontrolünü kazanamadıkları için beklemek zorundadır ve bus'ın kullanılmasına izin verirler. Bu önlemlerden hepsi bus controllerların kötü niyetli olmayan cihazlar olduğu varsayımıyla yapılmaktadır. Ama bus yapısında kötü niyetli bir BC ya da RT var ise bu önlemler işe yaramayacaktır. Çünkü bu önlemlerin hepsi bir noktada BC'lerden birinin kendi isteğiyle BC olmayı bırakması prensibine dayanır. Ama kötü niyetli BC bunu yapmayacaktır. İki aktif BC aynı anda bus kontrolünü ele geçirmeye

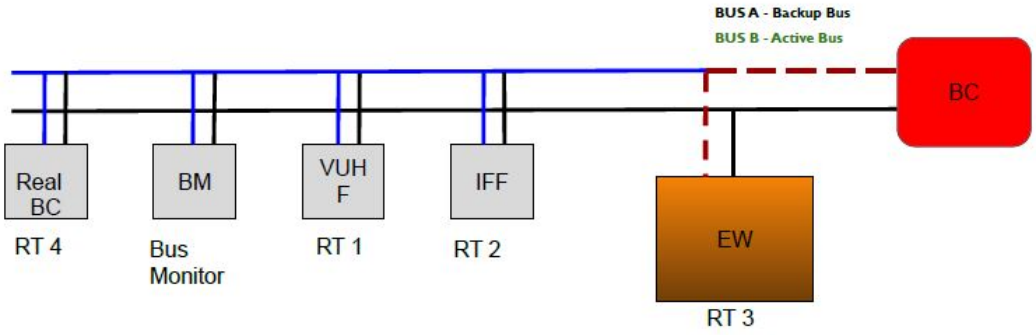



**Şekil 4.7 :** Kötü Niyetli BC'nin Kontrolüne Giren Bus.

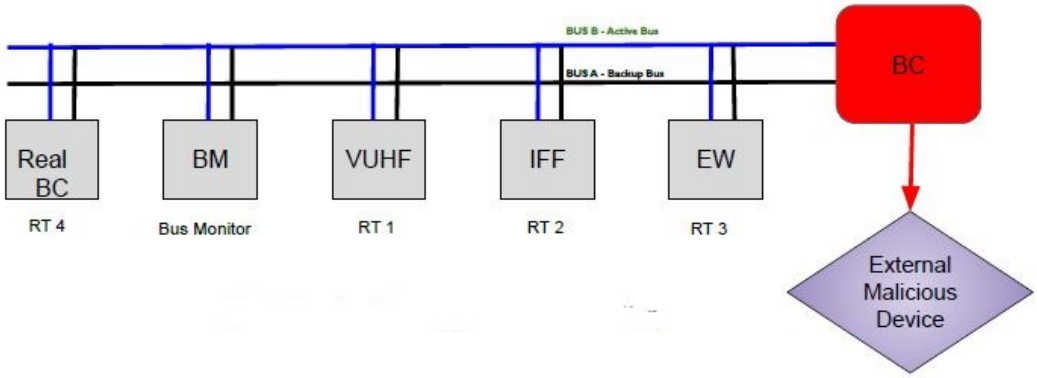



**Şekil 4.8 :** Kötü Niyetli BC'nin Kontrolünde Elekronik Harp Sisteminin Hedef Alınması.

çalıştığında, bir BC diğerine izin vermezse bir bus mücadelesi olabilir. Bu, bus kontrolünün atanamamasına ve bus'ın kullanılmamasına neden olabilir. Mil-Std 1553 üzerinde 2 tane BC var ise ilk olarak Şekil 4.6'de görüldüğü gibi aktif bus hattı değiştirilir. Ancak kötü niyetli BC geri çekilmediği için çözüm bulunamaz. Aktif BC'lerden bir tanesi kötü niyetli ise bu durumun olası sonuçlarından biri gerçek BC'nin çekilip kötü niyetli BC'ye bus kontrolünü bırakmasıdır. Bunun nedeni, gerçek BC karşı taraftaki BC'nin kötü niyetli olup olmadığını anlayamaz. Burada iki farklı durum oluşabilir. Bunlardan biri gerçek BC'nin geri çekilmesi ve yönetimi bırakmasıdır. Mil-Std 1553 kullanıcısı, bus'ı konfigüre ederken aktif bus controller ve backup bus controller'ı belirlemez ve bunlardan başka bir cihazın bus controller olamayacağına yönelik bir tasarım yapmazsa, gerçek BC bit arabulucuğu işlemleri sonucunda karşı taraf BC'likten çekilmeyeceği için bir süre sonra BC'likten çekilir ve RT veya BM olarak çalışmaya başlar. Bu durumda Şekil 4.7'de görüldüğü gibi kötü niyetli BC bus üzerinden tam hakimiyet kurar. Bir kötü niyetli BC, bir Mil-Std 1553 busını kontrol altına aldığında, Şekil 4.8'de



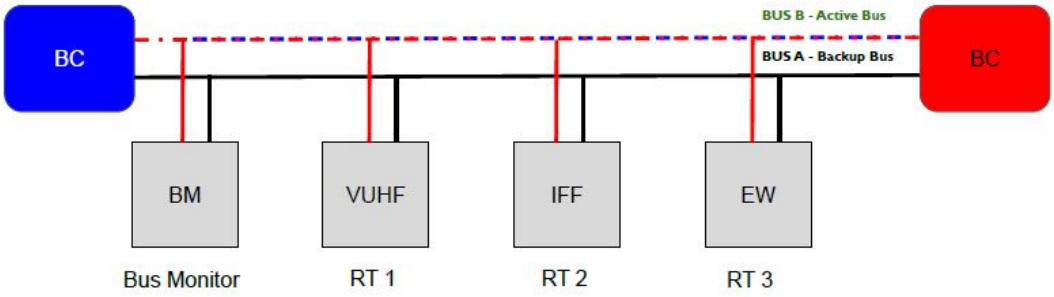

**Şekil 4.9 :** Elektronik Harp Sisteminin Ele Geçirilmesi.




**Şekil 4.10 :** Elektronik Harp Verilerinin Sızdırılması.

olduğu gibi, askeri ve havacılık gibi kritik sistemlerde ciddi sonuçlara yol açabilecek iletişimi bozma veya manipüle etme potansiyeline sahip olabilir. Bir kötü niyetli BC'nin yapabileceği bazı olası eylemler şunlar olabilir:

* Sahtecilik (Spoofing): Bir kötü niyetli BC, sahte adresler veya kimlik bilgileri ile mesajlar göndererek meşru bir BC gibi davranarak sistemi izinsiz olarak kontrol edebilir. Şekil 4.9'de kötü niyetli BC IFF ile erişimini keserek hava aracına büyük zararlar verdirebilir.
* Veri Manipülasyonu: Gönderilen komutlar ve verileri yakalayabilir ve değiştirebilir, bu da sistemin yanlış bilgi almasına veya işlemesine yol açabilir. Burada kötü niyetli BC, Şekil 4.9'de görülebileceği gibi elektronik harp sistemine çarpıtılmış mesajlar göndererek savunma sistemini çökertebilir. Sistemi yanlış çalıştırarak müttefiklerine zarar verilmesine sebep olabilir ya da düşmanlarına yarar sağlayabilir.




**Şekil 4.11 :** Gerçek BC'nin Bus Kontrolünü Bırakmaması .

* Dinleme: Bir kötü niyetli BC, tüm otobüs iletişimini pasif olarak dinleyebilir, potansiyel olarak algılanmadan hassas bilgileri toplayabilir. Normal bir Mil-Std 1553 sisteminde dış dünya ile net bir veri aktarımı mekanizması bulunmaz. Burada kötü niyetli BC tüm sistemi dinliyor ve istediği bilgileri kendinde topluyor. Bu bilgiler IFF gibi dost ve düşman uçakları tanıyan sistem olabilir. Bu bilgileri içinde bulunan external bir cihaz ile Şekil 4.10'de görüldüğü gibi herhangi bir düşman unsuruna gönderebilir. Bu yöntem ile herhangi bir kritik öneme sahip askeri bilgi çalınmış olur.

Gerçek BC'nin çekilmeyi reddettiğini durumda, Mil-Std 1553 otobüs sisteminde iki aktif BC aynı anda otobüs kontrolünü ele geçirmeye çalışır. Ancak hiçbiri geri çekilmez. Bunun sonucunda Şekil 4.11'da görüldüğü gibi bir bus mücadelesi sonucu ortaya çıkar. Bus mücadelesi, birçok ciddi soruna yol açabilir:

* Veri Çakışmaları: İki BC, farklı komutlar ve verileri aynı anda otobüse göndermeye çalışırlarsa, bu çakışmalara ve çakışan verilere neden olabilir. Bu durum, bus üzerindeki iletişimin tutarlılığını bozar.
* Belirsizlik: Hangi BC'nin otobüsü kontrol ettiği belirsiz hale gelir. Bu, verilerin doğru bir şekilde iletilmediği veya işlenmediği anlamına gelir.
* Veri Kaybı: Veriler eksik veya hatalı bir şekilde iletilirse, bu veri kaybına ve iletişim hatalarına neden olabilir.

* Hizmet Reddi (DoS): Otobüs mücadelesi durumunda, otobüs sürekli olarak meşgul olabilir, bu da diğer BC'lerin işleyişini engelleyebilir ve otobüsü kullanılamaz hale getirebilir.
* İletişimin Bozulması: Bir kötü niyetli BC, otobüse geçersiz veya çelişen komutlar ve veriler gönderebilir, böylece BC ve RT'ler arasındaki iletişimi karıştırabilir ve bozabilir.

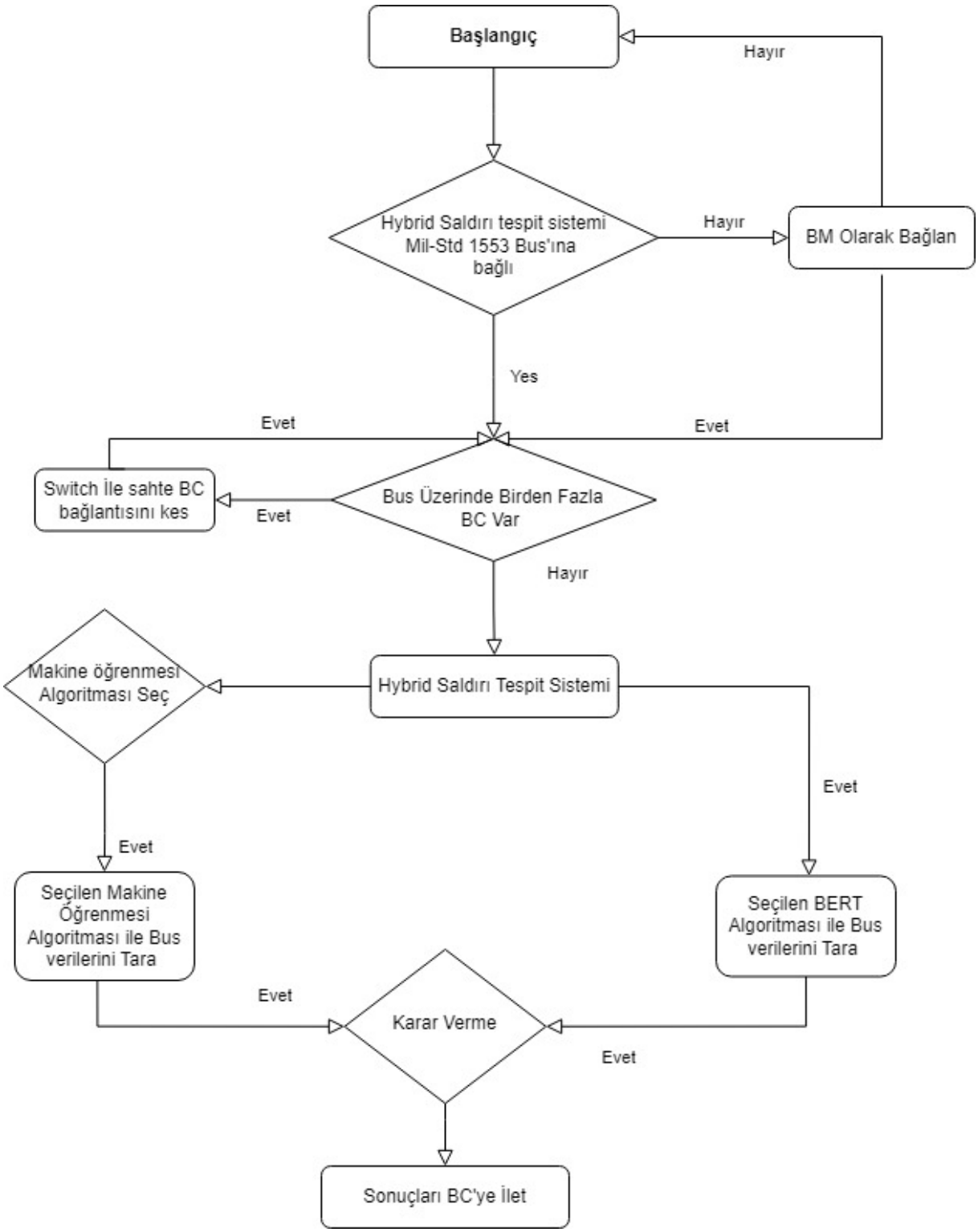
# 5. ÖLÇME VE DEĞERLENDİRME

Mil-Std 1553 sisteminin siber saldırılara karşı tespit etmede nasıl sonuçlar elde ettiğini test etmek için bazı deneyler yapılmıştır. Bu deneyler Mil-Std 1553 sisteminin doğruluk değerleri, zaman performansı ile ilgili çok yönlü bilgiler sunacaktır. Şekil 4.1'de gösterilen Mil-Std 1553 savunma sistemi makine öğrenmesi ile anomali tabanlı tespiti tespiti, Bert ile saldırı tespiti ve Mil-Std 1553 bus BC manipülasyonu şeklinde farklı alt savunma parçalarının entegrasyonu ile hazırlanmıştır. Bu nedenle bu alt savunma sistemlerinin performanslarının ölçülmesi ve değerlendirilmesi ayrı olarak yapılacaktır.

Makine öğrenmesi ve BERT algoritmalarıyla yapılan Mil-Std 1553 testleri için toplamda 4 farklı veriseti kullanılmıştır. 3 farklı veriseti makine öğrenmesi algoritmalarının testleri için, 1 tanesi BERT algoritmasının testleri için kullanılmıştır. Bu 4 farklı datasetin boyutları birbirinden farklıdır. Verisetlerinin oluşturulmasında taban olarak [39]'de belirtilen çalışma taban alınmıştır. Taban olarak alınan ham veriler işlenmiş makine öğrenmesi ve BERT algoritmalarıyla çalışacak 4 farklı dataset oluşturulmuştur. Sistemin genel karar verme ve işleyiş mekanizması Şekil 5.1'da gösterilmiştir.

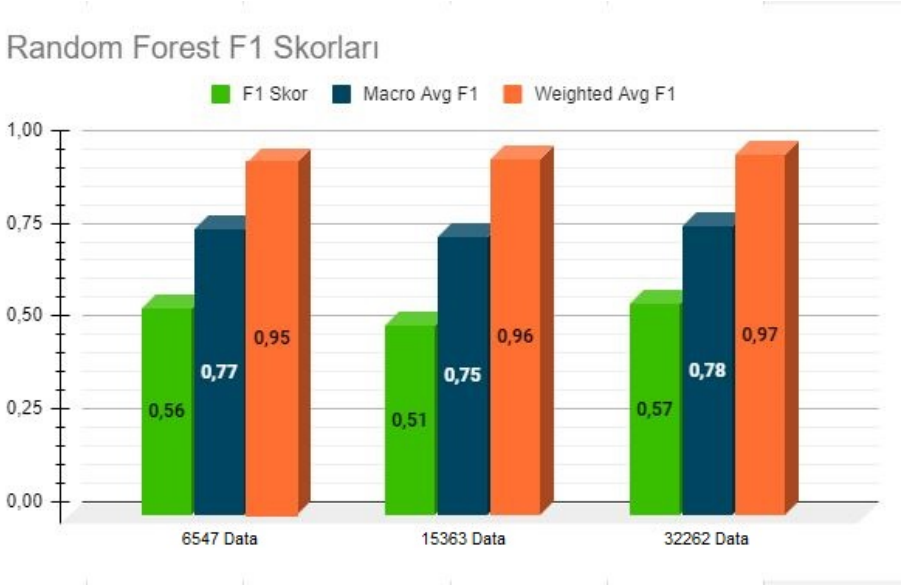
## 5.1 Makine Öğrenmesi ile Anomali Tabanlı Saldırı Tespiti Sonuçları

Bu çalışmadan makine öğrenmesi metoduyla Mil-Std 1553 sisteminde anomaly tabanlı saldırı tespit sistemi yapılması amaçlanmıştır. Bu hedef doğrultusunda Mil-Std 1553 datasetleri üzerinde atochastic gradient descent, decision tree, logistic regression, k-nearest neighbors, gaussian naive bayes ve random forest olmak üzere altı farklı makine öğrenmesi algoritması kullanılmıştır. Bu bölümde bu algoritmaların farklı veri boyutlarındaki performansını değerlendirerek, Mil-Std 1553 sisteminde güvenlik uygulamalarında kullanılmak üzere etkili bir model seçimini amaçlamıştır. Değerlendirme




**Şekil 5.1 :** Mil-Std 1553 Savunma Sistemi Karar Mekanizması

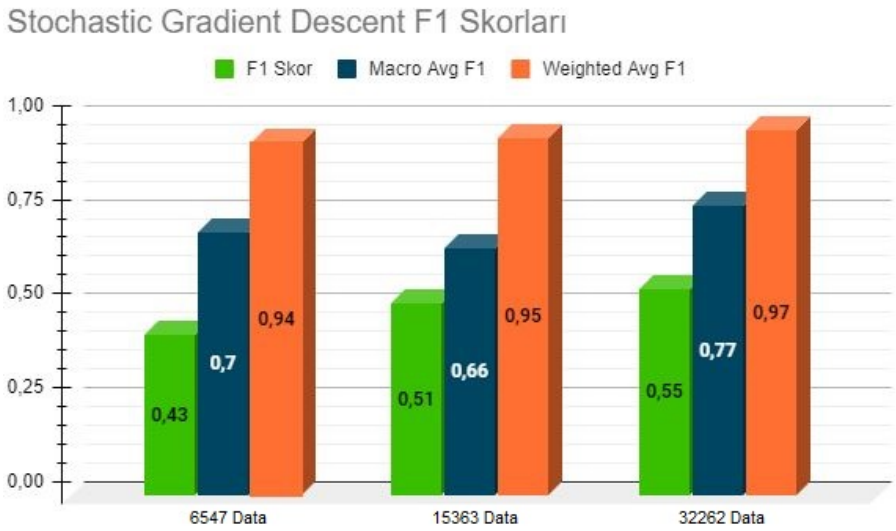
sürecinde f1 score, macro average ve weighted average f1 score, accuracy ve confusion matrix metrikleri kullanılmıştır.




**Şekil 5.2 :** Random Forest Algoritması F1 Skorları.

Mil-Std 1553 saldırı tespit sistemi makine öğrenmesi algoritmalarını test etmek için farklı boyutta üç veri seti kullanılmıştır. Veri setlerindenden bir tanesi 6547 veriden, diğeri 15363 veriden sonuncusu ise 32262 veriden oluşmaktadır. Farklı veri setleri hangi makine öğrenmesi algoritmasının hangi durumlarda daha verimli çalıştığı ve daha doğru sonuç verdiğini görmek için kullanılmıştır. Veri setleri oluşturulurken makine öğrenmesi algoritmalarının farklı veri setlerinde göstereceği performans ölçülmek istenmiştir. Bu duruma bağlı olarak 15363 adet veriden oluşan veri seti oluşturulurken daha dengesiz bir dağılım yapılmıştır. 15363 veriden oluşan setinin diğer veri setlerinde göre daha dengesiz dağılımlı bir yapıya sahip olmasıyla, dengeli dağılım ve dengesiz dağılım arasındaki verimlilik farklarının ölçülmesi amaçlanmıştır. Bunun yanında sadece eğitim setinin artırılmasıyla yani örnek sayısının çoğaltılmasıyla makine öğrenmesi algoritmalarının performans değişimleri de gözlenmek istenmiştir. Bu isteğe uygun olarak 6547 ve 32262 verilik setlerin dağılımı birbirlerine yakın yapılmıştır.

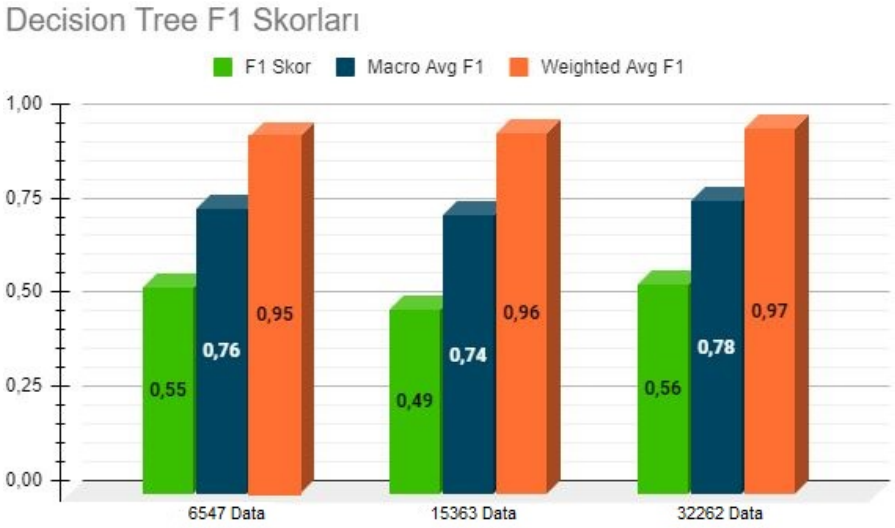
Random forest algoritması genel olarak veri boyutu yüksek eğitim setlerinde daha verimli çalışır ve performansı artar. Bu duruma uygun olarak Şekil 5.2'de görüldüğü gibi random forest algoritması F1 skor,macro f1 skor ve weighted f1 skora göre en iyi performansını 32262 boyutundaki veri setinde




**Şekil 5.3 :** Stochastic Gradient Descent Algoritması F1 Skorları.

gösteriyor. 6547 veriden 15363 veriye geçerken F1 ve macro averaj f1 skor bir miktar azalmaktadır. 15363 boyutundaki veri setinin daha düzensiz olması bu azalmanın temel sebebidir. Weighted f1 skora göre, eğitim seti boyutu yani örnek veri miktarı arttıkça random forest algoritmasının verimini artırdığı açıkça görülmektedir. Genel olarak bütün veri setlerinde random forest algoritması weighted f1 skora göre oldukça yüksek performans göstermiştir. Diğer f1 skor metriklerine göre de ortalamanın üzerinde performans göstermiştir.

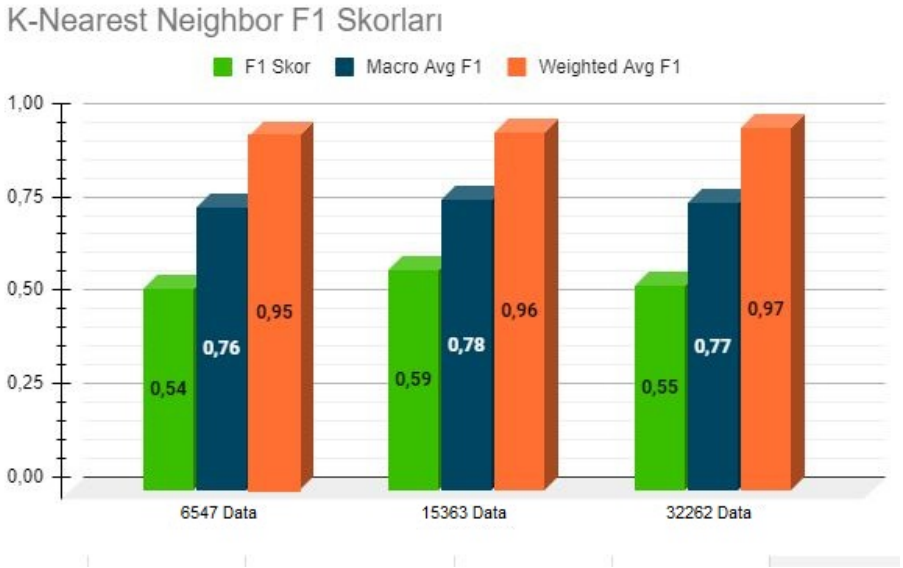
Şekil 5.3'de stochastic gradient descent algoritması ile veri setinin boyutu yükseldikçe genel olarak f1 skorların arttığı görülmektedir. F1 skor ve weighted f1 skora göre en düşük performansını 6547 boyutlu veri setinde göstermektedir. Veri sayısı artarken performansın artmadığını görülen tek yer, 6547 veriden 15363 veriye geçerken macro f1 skorun bir miktar azalmasıdır.Bu durumun oluşmasında 15363 boyutundaki veri setinin dengesiz bir dağılıma sahip olması büyük bir etkendir. Weighted f1 skorun veri boyutu arttıkça verimliliğinin yükselmektedir ve veri seti dağılım bunu etkilememektedir. Bütün veri setleri ve sonuçlar incelendiğinde f1 skorların genel olarak kabul edilebilir seviyede olduğu özellikle weighted average f1 skora göre ise oldukça iyi ve umut verici olduğu görülmektedir.




**Şekil 5.4 :** Decision Tree Algoritması F1 Skorları.

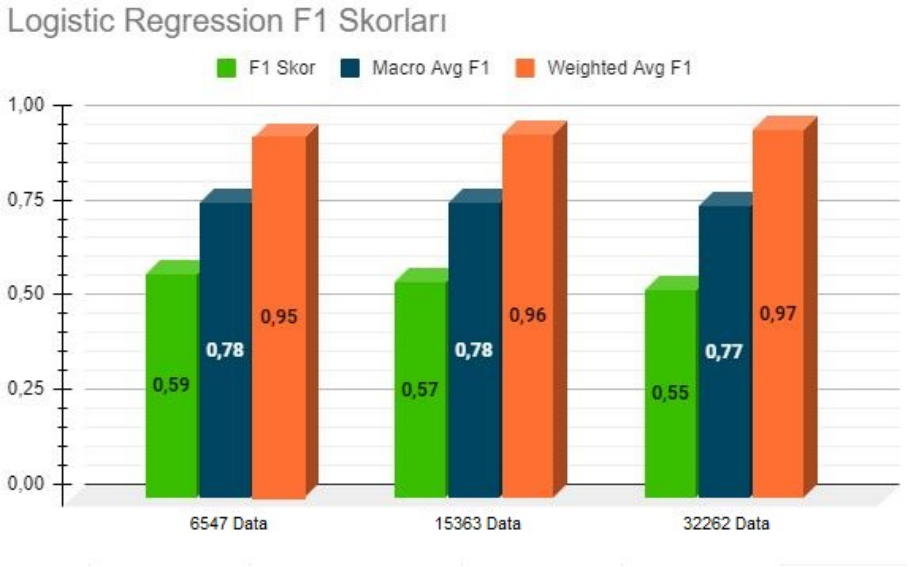
Decision tree algoritmasında Şekil 5.4'de görüldüğü gibi veri seti boyu arttıkça f1 skor performansı artmaktadır. F1 skor ve macro average f1 skorlarına göre en düşük performansı 15363 boyutundaki veri setinde göstermiştir. 15363 verilik setteki diğerlerine göre daha az f1 skor ve macro f1 skor çıkmasının sebebi bu veri setinde dengesiz dağılım yapılmasıdır. Dengesiz dağılım decision tree algoritmasında görüldüğü gibi performans düşmesine sebep olmaktadır. Özellikle f1 score'da 6547 lik veri setinden 1500'lik veri setinde geçildiğinde kayda değer bir düşüşe neden olmuştur. Bu durumun ortaya çıkmasında decision tree algoritmasının overfitting'e karşı dayanıksız bir yapıda olmasının payı vardır. Diğer taraftan weighted average f1 skora göre veri boyutu arttıkça sürekli bir artış söz konusudur. Tüm şekile bakıldığında decision tree weighted average ve macro average f1 skorlarına göre decision tree algoritması Mil-Std 1553 veri setlerinde saldırı tespit etmek de oldukça başarılı olduğu gözlenmektedir.

K-nearest neighbor algoritması decision tree, stochastic gradient descent ve random forest algoritmalarından farklı sonuçlar elde etmiştir. Şekil 5.5'de bu durumu kolay bir şekilde görmek mümkündür. KNN algoritması veri setinin boyutu arttıkça, macro average F1 skorları ve f1 skorları sürekli olarak yükselmiyor. 15363 boyutundaki veri setinden 32262 boyutundaki veri setine geçerken macro average ve f1 skorları düşüyor. Bunun en büyük

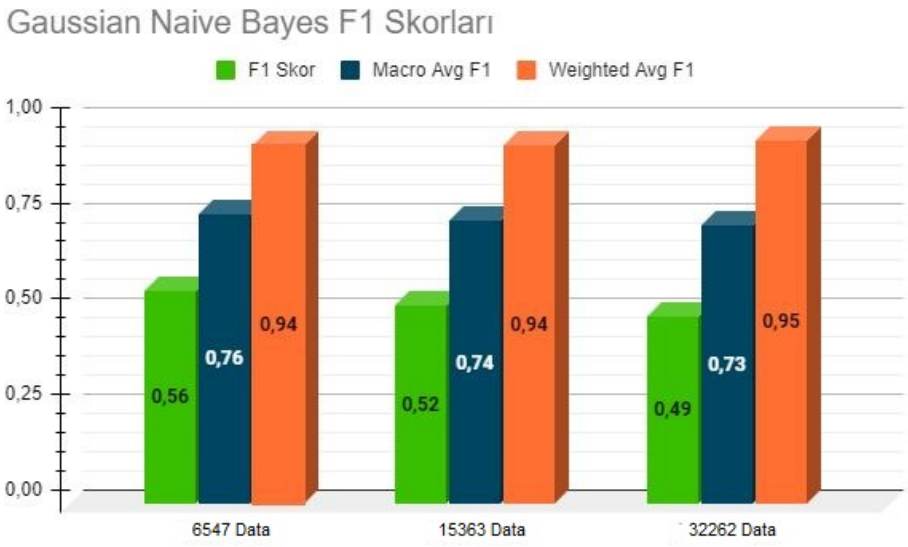



**Şekil 5.5 :** K-Nearest Neighbor Algoritması F1 Skorları.

nedeni KNN algoritmasının overfitting'e karşı dayanıksız bir algoritması olmasıdır. Diğer taraftan daha düzensiz bir veri setinde(15363 boyutundaki veri seti) f1 skorlarını yükseltmiş. Bu sonuçlar da KNN algoritmasının Mil-Std 1553 veri setinde düzensiz bir dağılım var ise KNN algoritmasının kullanılmasının oldukça uygun olabileceğini göstermektedir. Tüm tabloya bakıldığında ise KNN algoritmasının Mil-Std 1553 saldırı tespit sisteminde oldukça iyi sonuçlar verdiği görülmektedir. Genel olarak bu sistemde saldırı tespit etmede tercihlerden biri olabilecek bir sistemdir.




**Şekil 5.6 :** Logistic Regression Algoritması F1 Skorları.




**Şekil 5.7 :** Gaussian Naive Bayes Algoritması F1 Skorları.

Logictic regression algoritması diğer makine öğrenmesi algoritmlarından oldukça farklı sonuçlar vermektedir. Şekil 5.6'de ilk farkedilen durum logistic regression algoritması veri seti büyüdükçe macro average F1 ve F1 skorlarına göre performansı düşmektedir. Weighted average f1 skorlarına göre veri seti boyutu arttıkça performansı artmaktadır. Weighted average f1 skor veri seti boyutuna göre ağırlıklı bir ortalama almakatadır. Bu durum logistic regression'ın veri seti boyutu arttıkça saldırı tespit oranının azaldığını ancak yanlış tespit oranının artmadğını gibi ihtimali öne çıkartmaktadır. Ayrıca logistic regression algoritmasının Mil-Std 1553 sisteminde saldırı tespiti yaparken overfitting'e oldukça hassas olduğu göstermektedir. Tekil olarak bakıldığında 6547 verilik sette tüm machine learning algoritmları arasında en yüksek sonucu elde etse de veri boyutunun artması logistic regression algoritmasının kullanımı için uygun olmayabilmektedir. Ancak veri miktarının artması durumunda performans düşmesi yaşaması logistic regression'ın 32262'lik veri setinde bile oldukça iyi sonuçlar ortaya çıkardığını gölgelememelidir.

Gaussian Naive Bayes algoritması 6547 boyutu olan veri setinde iyi bir performans gösterirken daha yüksek veri sayısı olan veri setlerine geçtiğinde macro average f1 skor ve f1 skora göre düşüş yaşamaktadır. Şekil 5.7'de

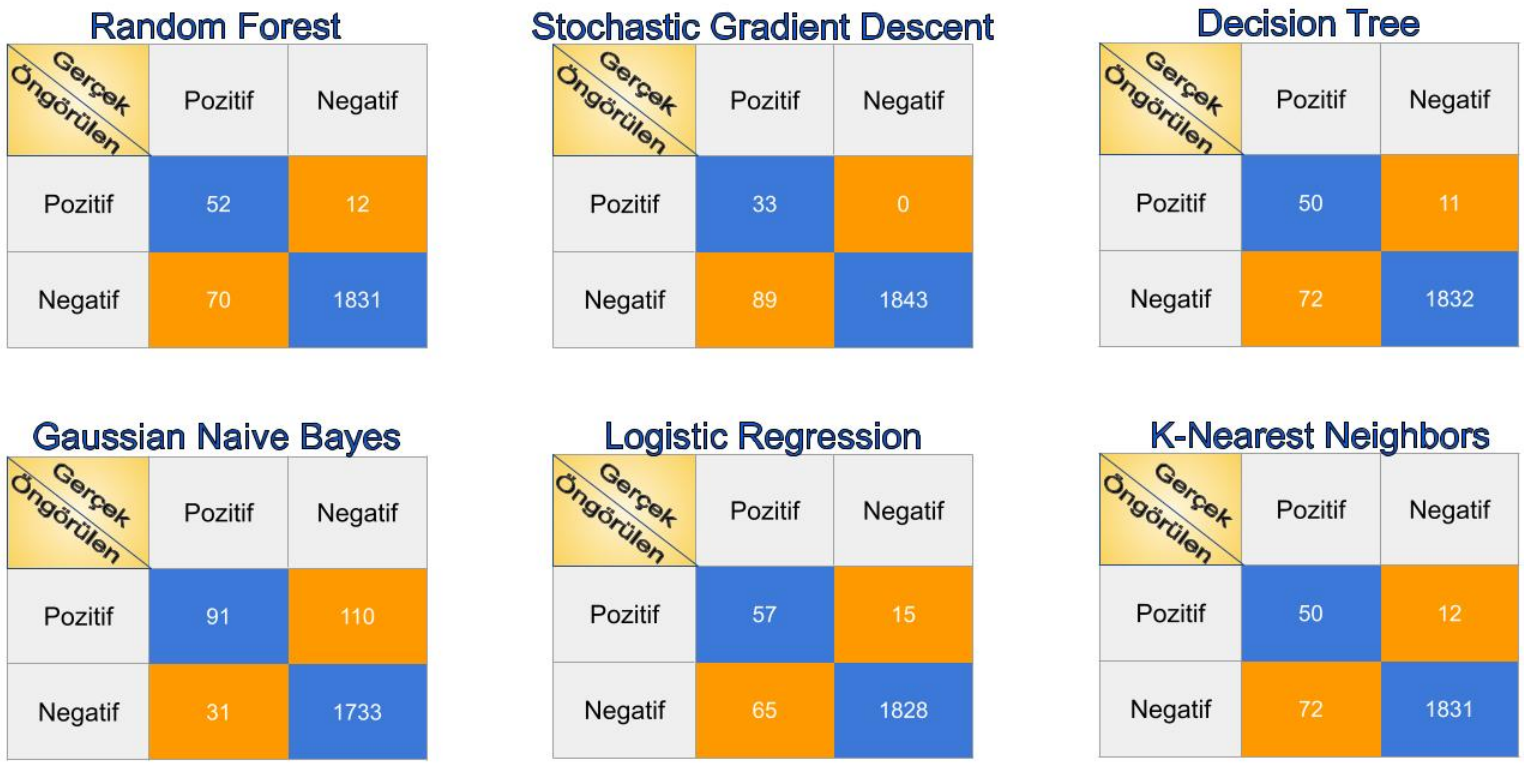
görüldüğü gibi 6547 verilik set ile 32262 verilik set arasında f1 skor ve macro average f1 skorları açısında kayde değer bir fark vardır. Bunun nedeni naive bayes algoritmasının basit bir algoritma yapısı olması ve karmaşık ve uzun modeller için efektif çalışmamasıdır. 6547 verilik sette performansı iyi olsa da veri seti boyutunun ve karmaşıklığın artması kötü etkilemiştir. Bu durum Mil-Std 1553 sisteminde küçük boyutlu eğitim ve train setleriyle çalışmasının uygun olabileceğini ancak büyük boyutlu veri setleriyle çalışmasının uygun olmayabileceğini göstermektedir.

**Çizelge 5.1 :** Makine Öğrenmesi Algoritmaları Accuracy Değerleri.

| Algoritmalar | 6547 Veri | 15363 Veri | 32262 Veri |
|---|---|---|---|
| Random Forest | %95.8 | %96.6 | %97.3 |
| Stochastic Gradient Descent | %95.8 | %96.9 | %97.0 |
| Decision Tree | %95.7 | %96.5 | %97.3 |
| K-Nearest Neighbor | %95.7 | %95.6 | %97.0 |
| Logictic Regression | %96.0 | %96.7 | %97.0 |
| Gaussian Naive Bayes | %92.9 | %93.4 | %93.6 |

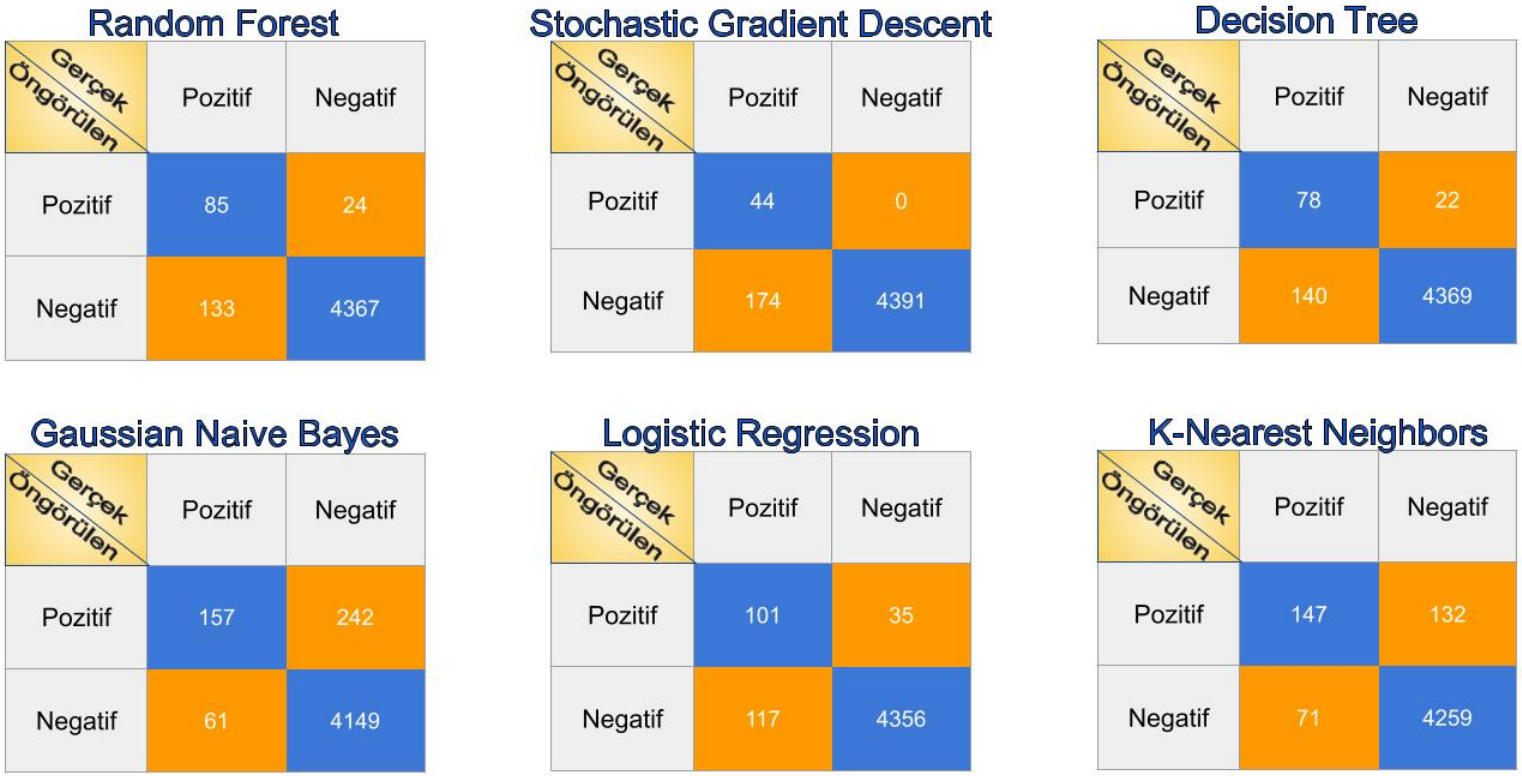
Çizelge 5.1'de görüldüğü üzere accuracy değerleri genel olarak birbirine yakın sonuçlar vermişlerdir. Accuracy sonuçlarına göre bütün algoritmalar %90 ve üzeri bir sonuç çıkarmıştır. Bu sonuçlar accuracy metriğine göre kullanılan bütün makine öğrenmesi sistemlerinin iyi sonuçlar verdiğini gösterir. 6547 verilik sette en iyi sonucu az bir farkla Logistic Regression algoritması vermiştir. En az başarılı sonucu ise Gaussian Naive Bayes verdiği görülmektedir. Gaussian Naive Bayes algoritması genel olarak bütün veri setlerinde en düşük performans verdiği görülebilir. Bunun en büyük nedeni, confusion matrix tablolarında da görülebileceği gibi, çok fazla veriyi kötü niyetli olmayan veriyi kötü niyetli olarak etiketlemesidir. 15363 verilik sette en yüksek başarıyı Stochastic Gradient Descent algoritması minik bir farkla göstermiştir. 32262 verilik sette ise en yüksek başarı Decision Tree ve Random Forest algoritmalarının olmuştur. Çizelgeden çıkarılabilecek başka bir sonuç ise kullanılan bütün algoritmalar, veri sayısı arttıkça daha başarılı accuracy sonuçları elde etmektedir.

Makine öğrenmesi algoritmalarının 6547 verilik, 15363 verilik ve 32262 verilik veri setlerine göre oluşturulmuş confusion matrix tabloları 5.8, 5.9 ve 5.10 'de gösterilmiştir. Bu tablolarda veriler doğru pozitif, yanlış pozitif, yanlış negatif, doğru negatif olarak belirtilmiştir.




**Şekil 5.8 :** Makine Algoritmalarının 6547'lik Veri Setine Göre Confusion Matrix Sonuçları.

6547 veriden oluşan veri setinin confusion matrix değerleri Şekil 5.8'de gösterilmiştir. Bu tablolarda Mil-Std 1553 sistemi için makine öğrenmesi ile saldırı tespiti yaparken incelenmesi gereken kısımlar sistemi kullanacak yapının istekleri doğrultusunda belirlenmelidir. Şekil 5.8'de en yüksek TP oranı Gaussian Naive Bayes algoritması ile yakalanmıştır. Bir Mil-Std 1553 sisteminde saldırı tespit sistemi yapılacaksa ve en ön önemli görülen şey gelen saldırıları yüksek oranda tespit etmekse bu algoritma kullanılabilir. Ancak bu algoritma kullanıldığında yüksek oranda saldırı olmayan Mil-Std 1553 mesajını saldırı olarak etiketleyecektir. Yine aynı şekle bakıldığında Stochastic Gradient Descent algoritması saldırı olmayan hiçbir mesajı saldırı olarak etiketlememiştir. Bir saldırı tespit sistemi eğer güvenli Mil-Std 1553 mesajlarının saldırı olarak etiketlenip engellenmesini istemiyorsa bu algoritmayı kullanabilir. Ancak bu Stochastic Gradient Descent algoritmasının dezavantajı da saldırı olan mesajları yanlış tespit oranı(FN) diğer bütün algoritlarından daha yüksek orandadır. Stochastic Gradient Descent sistemin bu durumu kabul ediyor olması gerekir. Random Forest , K-Nearest Neighbors ve Decision Tree algoritmaları 6547 verilik sette benzer sonuçlar vermişlerdir. FP oranları düşüktür ve bu durum onları tercih edilebilir bir algoritma yapabilir ancak FN oranları Stochastic Gradient Descent algoritması kadar olmasa da yüksektir. Yani bu algoritmarla saldırı olarak etiketlenip gerçekte saldırı olmayan mesaj sayısı çok az olacaktır. Ancak güvenli olarak etiketlenen mesajlar kötü niyetli mesajlar olabilecektir. Bu durumda kullanılması gereken algoritma Mil-Std 1553 sisteminin tasarım gereksinilere göre şekillencektir. 15363 veriden oluşan veri setinin confusion matrix değerleri Şekil 5.9'de gösterilmiştir. Stochasic Gradient Descent algoritmasının FP değeri burada da 0 görünmektedir. Bu bazı Mil-Std 1553 saldırı tespit sistemleri için önemli bir metriktir. 15363 veriden oluşan verisetinin geneline bakıldığında FN değelerinin arttığı görülmektedir.



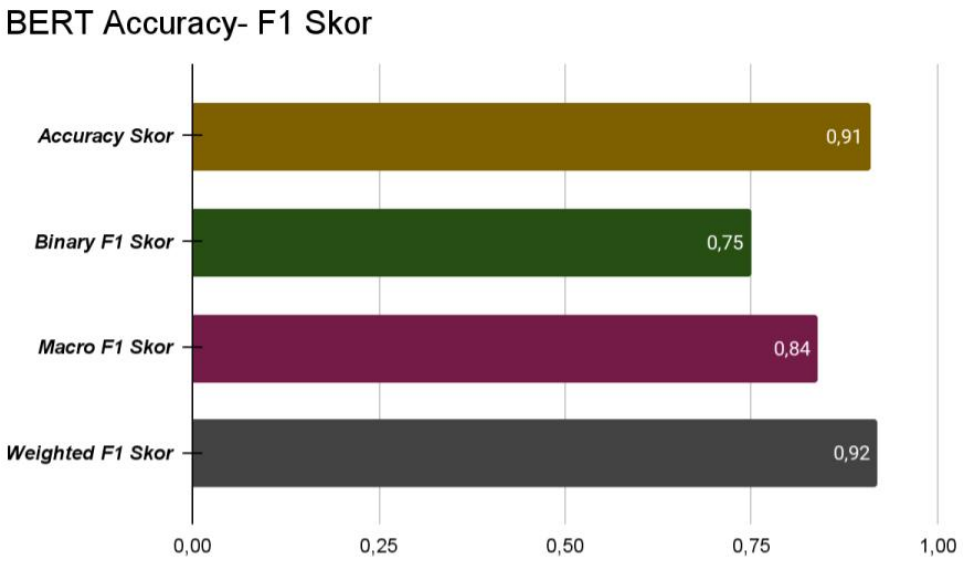

**Şekil 5.9 :** Makine Algoritmalarının 15363'lik Veri Setine Göre Confusion Matrix Sonuçları.

Bu durumun oluşmasında veri setleri arasındaki kötü niyetli mesaj sayısı oranının farkı etkili olmuştur. 6547 verilik sette normal Mil-Std 1553 mesaj sayısı yaklaşık % 93.8 oranında iken 15363 veriden oluşan sette bu oran %95.3 miktarına yükselmiştir. Şekil 5.10'da görülebilecek 32262 veriden oluşan sette ise bu oran % 95.7 seviyesine gelmiştir. 15363 verilik sette veri dağılımı dengesiziği de olduğu için genel olarak TP oranları düşmüştür. Ancak bu veri setinde K-Nearest Neighbors algoritması TP oranını yükseltmiş ve TN oranını da kayda değer oranda düşürmüştür. Her ne kadar FP oranı yükselse de bu veri seti için K-Nearest Neighbors algoritması oldukça iyi bir oran yakalamıştır. Stochastic gradient descent bu veri setinde de % 0 FP değerini korumuştur. Gaussian Naive Bayes 6547 verilik sette olduğu gibi 15363 ve 32262 verilik setlerde de en yüksek TP oranını korumuştur. Ancak veri sayısı yükseldikçe TP/FN oranı azalmıştır. Her ne kadar veri sayısının artışına paralel olarak FP oranı düşmüş olsa da Gaussian Naive Bayes algoritmasının yüksek veri miktarında daha düşük bir performans gösterdiği net bir şekilde görülebilmektedir. Bu şekillerin tamamı göz önüne alındığında Mil-Std 1553 sisteminde saldırı tespiti yaparken, daha önce de bahsedildiği gibi, gereksinimler ve veri durumu göz önünde bulundurularak algoritmalar seçilmesi gerekir. Normal bir mesajın kötü niyetli olarak tespit edilmesi ihtimalini düşürmek isteyen bir tasarımın stochastic gradient descent algoritmasını kullanması mantıklı olabilecekken, kötü niyetli bir mesajın bulunması ihtimalini yükseltmek isteyen bunun yanında normal mesajların da kötü niyetli mesajlar olarak etiketlenebileceği kabul eden bir sistem Gaussian Naive Bayes kullanabilir. Ya da dengesiz bir eğitim seti ile eğitim yapıldıktan sonra tespit yapılmak isteniyorsa K-Nearest Neighbors algoritması kullanılabilir.


**Şekil 5.10 :** Makine Algoritmalarının 32262'lik Veri Setine Göre Confusion Matrix Sonuçları.

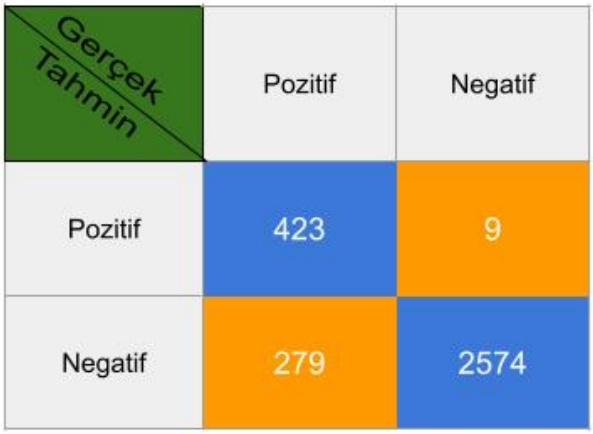
## 5.2 Bert ile Mil-Std 1553 Saldırı Tespiti Sonuçları

Bert algoritması bir cümlede kelimeler arasında bağlantı kurarak bir cümleyi anlamlandırmaya çalışır. Mil-Std 1553 sisteminin mesajlaşmasındaki veriler bir cümle haline getirildi ve BERT ile çalışacak şekilde düzenlendi. Bert Sistemi ile anomali tabanlı saldırı tespit sistemi yapılırken 16425 veriden oluşan bir veri seti kullanıldı.




**Şekil 5.11 :** BERT Algoritması Accuracy ve F1 Skorları.

Bert algoritması ile yapılan eğitim ve testler sonucunda şekil 5.11'de görülen sonuçlar elde edilmiştir. %91 oranında accuracy skoru elde edilmiştir. Bu accuracy skoru genel olarak başarılı bir performans sergiledini göstermektedir. F1 skoru ise %75 oranındadır. Macro average f1 skoru % 84, weighted average F1 skoru ise %92 oranındadır. Bu oranlar oldukça umut verici oranlardır. Daha önce bahsedilen makine öğrenmesi algoritmaları ile karşılaştırıldığında yeni bir sistem için oldukça iyi sonuçlar verdiği daha iyi anlaşılabilir.



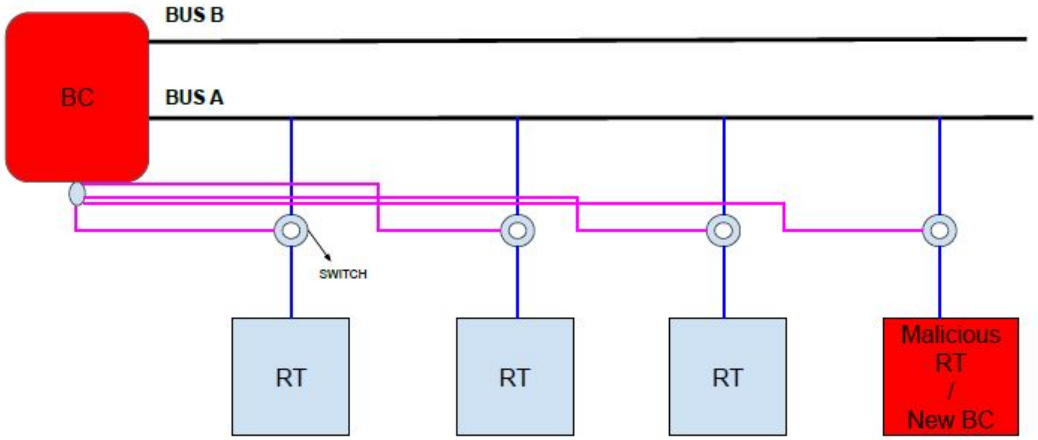

**Şekil 5.12 :** BERT Algoritması Confusion Matrix Sonuçları.

Bert algoritması ile yapılan Mil-Std 1553 saldırı tespit sisteminin confusion matrix sonuçları şekil 5.12'de gösterilmiştir. Confusion matrix sonuçlarına göre BERT sistemi kötü niyetli Mil-Std 1553 mesajlarını tespit ederken ortalamanın üstünde bir performans göstermiştir. TP değeri kötü niyetli mesajların tespit ederken % 60 üzerinde bir oran yakalamıştır. Ancak FN değeri oldukça yüksek çıktıştır ve bu durum performansının daha iyi çıkmasını engellemiştir. Bu durum bütün kötü niyetli mesajları tespit etmek isteyen bir sistem için BERT algoritmasının en iyi yöntem olmayabileceğini göstermektedir. Ancak BERT algoritmasının FP değerinin oldukça düşük olması normal mesajları kötü niyetli olarak etiketlememe ve doğru yorumlama konusunda oldukça iyi performansı olduğunu gösterir. Bir Mil-Std 1553 saldırı tespit sistemi normal mesajların kötü niyetli olarak yorumlaması tolere etmiyorsa BERT algoritması ile Mil-Std 1553 sistemini kullanması çok yerinde olabilir. BERT algoritması Mil-Std 1553 sisteminde saldırı tespiti ilk defa yapılan bir sistem olduğu halde oldukça umut verici sonuçlar elde etmiştir. Bu sistem daha da geliştirilerek Mil-Std saldırı tespit yöntemleri arasında en ön sıralarda yer alabilir.

## 5.3 Bus Controller'a Yapılan Saldırıları Engelleme

Mil-Std 1553 sistemlerinde BC saldıralarına bir savunma önlemi bulunmamaktadır. Bus yapısının tasarlandığı yıllarda bus controller cihazına herhangi

bir saldırı olabileceği öngürülmemiştir. Ancak daha önce bahsedildiği gibi Mil-Std 1553 fiziksel bus yapısına bağlı remote terminal cihazlardan bir tanesi kötü niyetli bir cihaz olabilir. Bu remote terminal Mil-Std 1553 sisteminin çalışma zamanının bir noktasında BC olmaya çalışabilir. Bu durumda daha önce yukarıda da bahsedilen bus üstünde birden fazla aktif BC oluşması yaşanır. Bu durumu çözmek Mil-Std 1553 sisteminin arabuluculuk protokolleri devreye girer. Ancak burada bus tasarımı yapılırken iki bus controller'ın da iyi niyetli olacağı düşünülmüştür. Yani bus üzerinde birden fazla aktif bus controller varsa tekrar remote terminal olmayı cihazların kendisinin seçmesi gerekir. Buna zorlayan bir durum olmadığı için kötü niyetli bir bus controller geri çekilmeyerek bütün yapının kontrolünü ele geçirebilir, bus üzerindeki iletişimi mesaj bombardımanıyla engelleyebilir ya da mesajlaşmayı manipüle eder görev kaybı ya da insan kaybıyla sonuçlandırabilir. Ayrıca bunlara ek olarak gizli bilgilerin sızdırılmasına sebep olabilir. Buna karşı önerdiğimiz çözüm Şekil 5.13'da görüldüğü gibi Mil-Std 1553 sistemine fiziksel olarak eklenen bir yapı öneriyoruz. Bus üzerinde birden fazla aktif cihaz olması durumda Mil-Std 1553 sisteminin tasarımdan dolayı fiziksel engel haricinde bus controller durumuna gelmiş bir cihazı engellemenin bir yolu bulunmamaktadır. Yazılımsal bir çözüm ancak bus yazılımsal mimarisinin değişmesiyle mümkün olabilir. Ancak Mil-Std 1553 sistemi gibi pahalı ve çok kullanılan bir sistemde büyük bir tasarımsal değişiklik yüksek ekonomik maliyete ve belirli bir süre için görev kaybına sebep olacaktır. Ancak önerilen sistem ile bu ihtimaller minimuma indirilmiştir. Önerilen bu sistemde Bus controller cihazına bağlı ancak Mil-Std 1553 sistemine bağlı olmayan bir fiziksel erişim yapısını öneriyor. Bus üzerindeki bütün remote terminal ve bus monitor cihazlarının Mil-Std 1553 bağlantılarına fiziksel bir kablaj ile bağlı bir yapı amaçlanıyor. Bu yapı bus üzerendeki bus controller harici bütün cihazların Mil-Std 1553 sistemine bağlandıkları kablaja bir switch yapısı ekliyor. Bus controllerdan çıkan kablaj her cihazın switch yapısına bağlanıyor. Bu switchin tüm kontrolü




**Şekil 5.13 :** Bus Controller Switch Sistemi.

fiziksel olarak bus controller cihazında olması amaçlanmıştır. Bu şekilde BC cihazınının bağlı olduğu görev bilgisayarı üzerinden, Mil-Std 1553 sisteminin fiziksel kontrolü sağlanmış olur. Herhangi bir şekilde cihazda aktif birden fazla bus controller olursa, gerçek bus controller kötü niyetli sistemi ele geçirmeye çalışan bus controller'ın switchine bir komut göndererek o cihazı devre dışı bırakır. Mil-Std 1553 sistemine erişimini de engellemiş olur. Bu sistem ile bus controller kaybı önlenmiş olması ve bu Mil-Std 1553 sistemini kullanan araçların bu zayıf noktasının önüne geçilmesi hedeflenmiştir.

# 6. SONUÇLAR

Mil-Std 1553 sistemlerinin güvenliği özellikle havacılık için önemli oldukça faktörlerden birisidir. Bu yolla hem insan hayatının korunması ve güvenliği sağlanmaktadır hem askeri ve sivil görevlerin yerine getirilebilmektedir hem de bir ülkenin kritik ve gizli bilgilerinin başka ülkelerce ya da tehdit unsurları tarafından ele geçirilmesi engellenmektedir. Bu nedenle Mil-Std 1553 sistemlerinin korunması aviyonik sistemlerde öncelikli konulardan birisidir. Makine öğrenmesi, doğal dil işleme ve bus controller sisteminin fiziksel kablaj ile cihaz kapatma ile kendini koruması yollarının birleştirerek oluşturduğumuz savunma sistemi bu konuda yardımcı olabilecek bir yapıdır. Bu tez süresince, saldırganların saldırması en muhtemel olan Mil-Std 1553 noktaları bulundu ve bu noktalara nasıl saldırılar yapılabileceği belirtildi. Bu noktalara yapılacak saldırıların olası sonuçları ele alındı ve bu saldırı stratejilerinin odak noktaları incelendi. Bu saldırı stratejilerini tespit etmek ve sistemi korumak için neler yapılabileceği açıklandı. Son olarak ise bu tez boyunca anlattığımız makine öğrenimi , doğal dil işleme ve fiziksel kablaj ile cihaz kapatma gibi yöntemler ile olabilecek tespit ve koruma yöntemleri test edildi. Bu testler ile önerdiğimiz sistemin sonuçları ortaya kondu. Sistemin makine öğrenme algoritmaları ile iyi sonuçlar elde ettiğini ve saldırıları tespit etme konusunda başarılı olduğu vurgulandı. Benzer şekilde BERT algoritması ile yapılan testlerde makine öğrenmesi kadar yüksek sonuçlar elde etmese de Mil-Std 1553 sistemi ile oldukça uyumlu şekilde çalışabileceği gösterildi.

## 6.1 Çalışmanın Uygulama Alanı

Mil-Std 1553 sistemi oldukça eski bir sistemdir. 1975 yılında tasarlanan bu sistem oldukça uzun yıllar boyunca bir çok hava aracında deniz aracında kara aracında ve hatta uzay aracında kullanılmıştır. Ancak özellikle havacılık ve askeri havacılık için vazgeçilmez olmuştur. Yoğun miktarda kullanımının

oldukça mantıklı nedenleri vardır ve bu tez boyunca bu nedenler detaylı olarak incelenmiştir. Tasarlandığı yıllar için saldırı tehdit söz konusu olmaması nedeniyle ciddi bir güvenlik sistemine sahip değildir. Ancak günümüzde teknolojinin geldiği noktayla beraber artık tehditlere karşı savunmasız bir durumu düşmüştür [4]. Bu durum insan güvenliğini, askeri ve sivil görevlerin yerine getirilmesini ve kritik ve gizli bilgilerinin korunmasını tehdit etmektedir. Bu nedenle Mil-Std 1553 sistemlerinin korunması büyük önem arz etmektedir. Bu sebeplerle oluşturduğumuz ve buz tezde önerdiğimiz savunma sistemi Mil-Std 1553 yapısının korunmasında test sonuçlarına göre başarılı çıktılar elde etmiştir. Bu tezde önerdiğimiz sistemi tasarlarken ilk odak noktalarımızdan birisi de gerçek Mil-Std 1553 bus yapısına uyarlanabilir olmasıydı. Burada önemli noktalardan biri Mil-Std 1553 sistemini kullanan araçların çok uzun süreden beri bu sistemi kullanmasıdır. Ayrıca çok fazla hava aracında aktif olarak kullanılmaktadır. Bu nedenle Mil-Std 1553 üstünde yapılabilecek tasarımsal değişiklik ile yapılacak bir güvenlik sistemi, özellikle yüksek maliyeti ve görev kaybına uğraması gibi sebeplerden dolayı çok uygulanabilir değildir. Ancak bu tezde önerdiğimiz sistem tasarımsal bir değişikliğe sebep olmayacak ve sadece bus üzerinde cihaz bağlantı noktalarına eklenecek bir switch mekanizması ve görev bilgisayarlarına eklenecek yazılımlar ile sistem aktif olarak çalışabilecektir. Bu nedenle aktif olarak çalışan bir hava aracına kolayla eklenebilecek bir sistemdir. Tezde önerdiğimiz sistem Mil-Std 1553 bus protokolünü kullanan herhangi bir araca eklenip büyük bir kullanıma sahip olabilir. Ayrıca hava araçlarının yanında kara araçları deniz araçları ve uzay araçlarında da kullanılablir.

## 6.2 Tartışma ve Gelecek Çalışmalar

Bu tezde, Mil-Std 1553 güvenliği ile ilgili çalışmalar yapılırken bu durumla ilgilenen ve çözüm bulmaya çalışan çeşitli araştırmalar incelendi. Genel olarak hepsinde Mil-Std 1553 sisteminin siber saldırılara karşı savunmsız olduğunun vurgulandığı görüldü. Tasarlandığı dönemde, saldırı tehdidi olmaması nedeniyle Mil-Std 1553 sistemleri ciddi bir güvenlik altyapısına

ihtiyaç duymuyordu. Ancak günümüzde, teknolojinin ilerlemesiyle birlikte, bu sistemler artık çeşitli tehditlere karşı savunmasız hale gelmiştir. Bu durum, insan güvenliğini, askeri ve sivil görevlerin etkili bir şekilde yerine getirilmesini ve kritik, gizli bilgilerin korunmasını ciddi şekilde tehlikeye atmaktadır. Bu sebeplerle Mil-Std 1553 sistemlerinin güvenliğinin sağlanması son derece kritik bir öneme sahiptir. ve Mil-Std 1553 sistemlerinin korunması, aviyonik sistemlerde öncelikli konulardan biridir. Bu tezde önerdiğimiz sistem, Mil-Std 1553 sistemi üzerinde meydana gelen saldırıları tespit etmeyi ve bu saldırılara karşı etkili bir güvenlik çözümü sunmayı amaçlamaktadır. Proje, makine öğrenmesi, doğal dil işleme teknikleri ve Mil-Std 1553 bus yapısına dışarıdan eklenen bir switch aracılığıyla, genel bir saldırı tespiti ve korunma sistemi sağlamayı hedeflemektedir. Çalışma sürecinde öncelikle Mil-Std 1553 bus yapısına yönelik potansiyel saldırı yöntemleri üzerinde düşünüldü. Bu saldırıların nereye yapılacağı ve Mil-Std 1553 bus yapısının en savunmasız noktalarının neler olduğu araştırıldı. Bu analizlere dayanarak, makine öğrenimi ve doğal dil işleme algoritmaları seçildi. Çalışmanın ikinci aşamasında fark edilen bir durum, bus controller üzerine gerçekleştirilecek bir saldırının sonucunda Mil-Std 1553 sisteminin tüm kontrolünün kötü niyetli yazılımların eline geçebileceğiydi. Bu riski önlemek için Mil-Std 1553 tasarımını ve protokolünü değiştirmeden, ekonomik ve etkili bir çözüm olarak bus yapısına dışarıdan bir switch ekleyerek, bus üzerindeki cihazların fiziksel bağlantısının kontrolünü bus controller'a vermek oldu. Böylece, bus controller, kendisine yönelik olası saldırıları engelleyerek hem kendi güvenliğini hem de bus yapısını koruyabilme imkanına sahip oldu. Bu tez, araştırılan bütün saldırı senaryolarını kapsayacak ve çözüm üretecek bir yapı sunmakadır. Tezin başlangıcında hedeflenen savunma sistemi oluşturulmuştur. Bu savunma sistemi oldukça iyi performans göstermektedir. Ancak her geçen gün saldırı methodları ve teknikleri değişmekte ve gelişmektedir. Ayrıca bahsedilen makine öğrenmesi algoritmaları ve BERT algoritması her ne kadar başarılı performans gösterse de, hatasız bir performans göstermemiştir. Mil-Std 1553 gibi son derece

kritik olan bir sistemde hata oranını en düşük miktarda tutmak önemlidir. Saldırı tespit sisteminde yapılabilecek küçük bir performans artırımının bile bir insanın hayatını kaybetmesini önleyebileceği unutulmamalıdır. Bu nedenle bu sistemi tamamlanmış bir sistem olarak değil üzerinde çalışılması gereken bir yapı olarak görüyoruz. Gelecek çalışmalarda ilk önceliğimiz kullanılan algoritmaların performanslarını yükseltmek olacaktır. Bunun yanında bu çalışmada BERT sistemine entegre ettiğimiz gibi farklı saldırı tespit mekanizmalarını ekleyerek çok yönlü ve saldırıyı mutlak olarak tespit eden bir sistem oluşturmak istiyoruz. Ayrıca bu tezde önerdiğimiz sistemi gerçek zamanlı çalışan bir Mil-Std 1553 sistemi üzerine entegre etmek de en önemli hedeflerimizden biridir.

## KAYNAKLAR


[1] **DDC** (1998). *MIL-STD-1553 DESIGNER'S GUIDE SIXTH EDITION*, ILC Data Device Corporation.

[2] **Nayak, P.** *Understanding searches better than ever before*, `https://blog.google/products/search/search-language-understanding-bert/`, (Accessed on 01/02/2024).

[3] (2017). Protecting Military Avionics Platforms from Attacks on MIL-STD-1553 Communication Bus.

[4] **Santo, D.**, **Malavenda, C.S. ve Romano, S.P.** (2021). Exploiting the MIL-STD-1553 avionic data bus with an active cyber device, *Computers & Security*, *100*, 102097.

[5] **Levy** (2022). AnoMili: Spoofing Prevention and Explainable Anomaly Detection for the 1553 Military Avionic Bus.

[6] **Schoofs, T.**, **Santos, S.L.d.**, **Tatibana, C. ve Anjos, J.M.S.** (2009). An integrated modular avionics development environment, *2009 IEEE/AIAA 28th Digital Avionics Systems Conference*.

[7] **Zhang, J.D.**, **Liu, M.Y.**, **Shi, G.Q. ve Pan, W.** (2011). A mil-std-1553b bus command optimization algorithm based on load balance, *Applied Mechanics and Materials*, *130-134*, 3839–3842.

[8] **Duren, R. ve Thompson, M.** (2008). Application of data compression to the mil-std-1553 data bus, *2008 IEEE Aerospace Conference*.

[9] **Mahesh, B.** (2023). Machine Learning Algorithms - A Review, *ISSN: 2319-7064*.

[10] **Janiesch, C.**, **Zschech, P. ve Heinrich, K.** (2021). Machine learning and deep learning, *Electronic Markets*, *31*, 685–695.

[11] **, R.S. ve Barto, A.G.** (1998). *Reinforcement learning: An introduction*, MIT press.

[12] **Bishop, C.M.** (2006). *Pattern Recognition and Machine Learning*, Springer.

[13] **Nelson, M.R.** (1988). *C. Crisci, B. Ghattas, and G. Perera, "A review of supervised machine learning algorithms and their applications to ecological data," Ecological Modelling, vol. 240, pp. 113–122, 2012.*, (Doctoral dissertation). Retrieved from http://edt.missouri.edu/.

[14] **Nelson, M.R.** (1988). *D. R. Cutler, T. C. Edwards Jr, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler, "Random Forests for Classification in Ecology," Ecology, vol. 88, pp.*

*2783–92, Dec. 2007*, (Doctoral dissertation). Retrieved from http://edt.missouri.edu/.

[15] **Yin, Q.**, **Cheng, J.**, **Zhang, F.**, **Zhou, Y.**, **Shao, L. ve Hong, W.** (2020). Interpretable polsar image classification based on adaptive-dimension feature space decision tree, *IEEE Access*, *8*, 173826–173837.

[16] **Zhai, L.**, **Sun, J.**, **Sang, H.**, **Yang, G. ve Jia, Y.** (2012). Large area land cover classification with landsat etm+ images based on decision tree, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XXXIX-B7*, 421–426.

[17] **Zhang, Y.**, **Zhang, J.**, **Zhang, X.**, **Wu, H.R. ve Guo, M.** (2015). Land cover classification from polarimetric sar data based on image segmentation and decision trees, *Canadian Journal of Remote Sensing*, *41*, 40–50.

[18] **Sinlae, F.**, **Yudhasti, A.S. ve Wibowo, A.** (2022). Comparative analysis of naïve bayes and decision tree algorithms in data mining classification to predict weckerle machine productivity, *Journal of Systems Engineering and Information Technology (JOSEIT)*, *1*, 47–51.

[19] **Mabu, S.**, **Obayashi, M. ve Kuremoto, T.** (2016). An evolutionary algorithm for making decision graphs for classification problems, *Journal of Robotics, Networking and Artificial Life*, *3*, 45.

[20] **Li, D. ve Fan, S.** (2014). A modified decision tree algorithm based on genetic algorithm for mobile user classification problem, *The Scientific World Journal*, *2014*, 1–11.

[21] **Wang, C.**, **Du, Z.**, **Liu, Z. ve Liu, Y.** (2008). Study on decision tree land cover classification based on modis data, *2008 International Workshop on Earth Observation and Remote Sensing Applications*.

[22] **Qasem, M. ve Nour, M.A.** (2015). Improving accuracy for classifying selected medical datasets with weighted nearest neighbors and fuzzy nearest neighbors algorithms, *2015 International Conference on Cloud Computing (ICCC)*.

[23] **Astuti, N.K.M.**, **Utami, N.W. ve Juliharta, I.G.P.K.** (2022). Classification of blood donor data using c4.5 and k-nearest neighbor methods (case study: utd pmi bali province), *Jurnal Pilar Nusa Mandiri*, *18*, 9–16.

[24] **Sun, Y.** (2022). Multispectral remote sensing data analysis based on knnlc algorithm and multimedia image, *Journal of Sensors*, *2022*, 1–8.

[25] **Zhou, C. ve Wan, L.** (2010). A hybrid algorithm of minimum spanning tree and nearest neighbor for classifying human cancers, *2010 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE)*.

[26] **Xu, S.** (2016). Bayesian naïve bayes classifiers to text classification, *Journal of Information Science*, *44*, 48–59.

[27] **Fathima, M.D.**, **Samuel, S.J. ve Raja, S.P.** (2021). Regression imputation and optimized gaussian naïve bayes algorithm for an enhanced diabetes mellitus prediction model, *Brazilian Archives of Biology and Technology*, *64*.

[28] **Bottou, L.** (2011). Large-scale machine learning with stochastic gradient descent, *Chapman Amp; Hall/CRC Computer Science Amp; Data Analysis*, 17–25.

[29] **Song, C.**, **Pons, A.P. ve Yen, K.K.** (2021). Ag-sgd: angle-based stochastic gradient descent, *IEEE Access*, *9*, 23007–23024.

[30] **Hu, W.**, **Li, C.J.**, **Li, L. ve Liu, J.** (2019). On the diffusion approximation of nonconvex stochastic gradient descent, *Annals of Mathematical Sciences and Applications*, *4*, 3–32.

[31] **Peng, C.J.**, **Lee, K.L. ve Ingersoll, G.M.** (2002). An introduction to logistic regression analysis and reporting, *The Journal of Educational Research*, *96*, 3–14.

[32] **Kirişçi, M.** (2019). Comparison of artificial neural network and logistic regression model for factors affecting birth weight, *SN Applied Sciences*, *1*.

[33] **Landwehr, N.**, **Hall, M. ve Frank, E.** (2005). Logistic model trees, *Machine Learning*, *59*, 161–205.

[34] **Wiest, M.M. ve Lee, K.J.** (2015). Statistics for clinicians: an introduction to logistic regression, *Journal of Paediatrics and Child Health*, *51*, 670–673.

[35] **Fox, E.** (2020). natural language processing advancements by deep learning: a survey.

[36] **Devlin, J.** (2018). bert: pre-training of deep bidirectional transformers for language understanding.

[37] *What is BERT (Language Model) and How Does It Work? — techtarget.com*, `https://www.techtarget.com/searchenterpriseai/definition/BERT-language-model`, [Accessed 04-01-2024].

[38] (2007). The MIL-STD-1553B data bus: What does the future hold?, *The Aeronautical Journal*, *111*(1118), 231–246.

[39] **Yahalom, R.**, **Barishev, D.**, **Steren, A.**, **Nameri, Y.**, **Roytman, M.**, **Porgador, A. ve Elovici, Y.** (2019). Datasets of rt spoofing attacks on mil-std-1553 communication traffic, *Data in Brief*, *23*, 103863.

## ÖZGEÇMİŞ

**Adı SOYADI: Yunus Emre ÇİLOĞLU**

**ÖĞRENİM DURUMU:**

* **Lisans:** 2018, Ankara Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği
* **Y. Lisans:** Devam Ediyor, İstanbul Teknik Üniversitesi, Bilgisayar Mühendisliği Anabilim Dalı, Bilgisayar Mühendisliği Programı.

**MESLEKİ DENEYİMLER VE ÖDÜLLER:**

* 2016-2017 yılları arasında Pixel Bilişim'de back-end yazılım üzerine çalıştı.
* 2017-2018 yılları arasında BiSoft'da database administrator olarak çalıştı.
* 2018 yılından itibaren Aselsan A.Ş'de gömülü yazılım mühendisi olarak çalışıyor.

**MASTER TEZİNDEN TÜRETİLEN YAYINLAR:**

* **Çiloğlu Y.E.**, Bahtiyar Ş., (2023). A New Anomaly-Based Intrusion Detection System for Mil-Std-1553, *2023 10th International Conference on Recent Advances in Air and Space Technologies (RAST)* 10.1109/rast57548.2023.10197927.
* **Çiloğlu Y.E.**, Bahtiyar Ş., (SUBMITTED). A Hybrid Machine Learning Based Intrusion Detection System for MIL-STD-1553, *Journal of Aeronautics and Space Technologies (JAST)*