CUKUROVA UNIVERSITY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

MSc THESIS

DDoS Attack Detection From Network Traffic Data Using Machine Learning Methods

Numan ERTİK

Computer Engineering Department

May, 2024

CUKUROVA UNIVERSITY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

MSc THESIS APPROVAL

DDoS Attack Detection from Network Traffic Data Using Machine Learning Methods

Numan ERTİK

Computer Engineering Department

This Master's Thesis was evaluated by the following Jury Members on 14/05/2024 and was approved by unanimity.

Jury : Assoc.Prof.Dr. Fatih ABUT (Advisor) ……

: Asst.Prof.Dr. Barış ATA ……

: Asst.Prof.Dr. Onur POLAT ……

This Thesis was written in the Department of Computer Engineering, Institute of Natural and Applied Sciences.

Thesis Number:

Prof. Dr. Sadık DİNÇER
Director
Institute of Natural and Applied Sciences

CONTENTS

CUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

MSc THESIS

DDoS Attack Detection From Network Traffic Data Using Machine Learning Methods

Numan ERTİK

Advisor: Assoc. Prof. Dr. Fatih ABUT

Department of Computer Engineering

ABSTRACT

This thesis aims to develop new DDoS attack prediction models by employing various machine learning and feature selection methods. Especially, Logistic Regression (LR), Multilayer Perceptron (MLP), Support Vector Machine (SVM), and Random Forest (RF) have been used to build different DDoS attack prediction models. Relief-F and minimum redundancy maximum relevance (mRMR) algorithms have been utilized as feature selectors. The dataset has been created with 10.000 rows of traffic data, considering an equal number of rows containing four different attack types and normal traffic. The dataset encompasses a diverse array of predictor variables, notably including packet size, transmission time, and indicators of attack occurrences, among numerous others. Accuracy, precision, recall, and F1-score metrics has been utilized to evaluate the performance of the models, whereas the generalization errors of the models have been assessed using 10-fold cross-validation. Results reveal that RF surpasses all other ML classifiers for predicting DDoS attack detection. Overall, mRMR performed better than Relief-F in predicting the DDoS attack detection. In both feature selection methods, it has been determined that packet size and temporal features are the most important predictors of DDoS attack prediction.

Key Words: DDoS Attack, Feature Selector, Machine Learning, Prediction Models, Cyber Security

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

Makine Öğrenmesi Yöntemleri Kullanılarak Ağ Trafiği Verilerinden DDoS Saldırısı Tespiti

Numan ERTİK

Danışman: Doç. Dr. Fatih ABUT

Bilgisayar Mühendisliği Anabilim Dalı

ÖZ

Bu tez, çeşitli makine öğrenimi ve öznitelik seçme yöntemlerini kullanarak yeni DDoS saldırısı tahmin modelleri geliştirmeyi amaçlamaktadır. Lojistik Regresyon (LR), Çok Katmanlı Algılayıcı (MLP), Destek Vektör Makinesi (SVM) ve Rastgele Orman (RF) yöntemleri kullanılarak çeşitli DDoS saldırısı tahmin modelleri oluşturulmuştur. Öznitelik seçiciler olarak Relief-F ve minimum gereksiz maksimum ilgi (mRMR) algoritmaları kullanılmıştır. Veri seti, dört farklı saldırı türünü ve normal trafik içeren eşit sayıda satır içeren 10.000 satırlık trafik verisinden oluşmaktadır. Veri seti, paket boyutu, iletim süresi ve saldırı olaylarının belirteçleri gibi dikkate değer pek çok değişkeni içeren geniş bir dizi tahmin değişkeni kapsamaktadır. Modellerin performansını değerlendirmek için doğruluk, hassasiyet, duyarlılık ve F1-skoru metrikleri kullanılmıştır, modellerin genelleme hataları ise 10 katlı çapraz doğrulama kullanılarak değerlendirilmiştir. Sonuçlar, RF'nin DDoS saldırısı tespitinde diğer tüm ML sınıflayıcılarının daha iyi bir performans sergilediğini göstermektedir. Genel olarak, mRMR, DDoS saldırısı tespiti tahmininde Relief-F'den daha iyi performans göstermiştir. Her iki öznitelik seçme yönteminde, paket boyutu ve zamansal özelliklerin DDoS saldırısı tahmininde en önemli belirleyiciler olduğu belirlenmiştir

Anahtar Kelimeler: DDoS Atak, Makine Öğrenmesi, Özellik Seçici, Tahmin Modelleri, Siber Güvenlik

EXTENDED ABSTRACT

The role of the Internet in our daily lives is increasing day by day. Internet service is widely used in almost every aspect of our lives, including work, school, the health sector, trade, and communication. With the development of Internet of Things (IoT) technology, the number of devices connected to the Internet has significantly increased. Now, users can access the internet not only with computers but also with household appliances such as televisions, refrigerators, and washing machines. While this increase provides great convenience, it also brings about risks. With the expansion of services through the Internet, there is an observed increase in the exposure of network infrastructure to attacks. Attackers can target almost any device connected to the internet. They can use devices connected to the internet as zombies, or they can attack different devices by taking over the system. The most common of these attacks are Distributed Denial of Service (DDoS) attacks.

DDoS attacks have been a persistent threat to the IT world for an extended period, prompting ongoing prevention studies. State-owned institutions, commercial companies, and all systems used by citizens are susceptible to these attacks. The struggle against attacks persists, with continuous development of new methods to counteract them. In world history, the first DDoS attacks occurred in the 1980s, with the initial large-scale attack organized at the University of Minnesota in 1999. The university's network was significantly affected, and services could not be provided for about two days. According to research, there has been a substantial annual increase in DDoS attacks. In 2015, DDoS attacks reached 500 Gbps, setting a record of 800 Gbps in 2016 with a 60% increase. In subsequent years, the attacks reached the terabyte level. Cisco conducted an analysis study on DDoS attacks, revealing that the number of DDoS attacks was 7.9 million in 2018, 9.5 million in 2019, 10.8 million in 2020, 12.1 million in 2021, and 15.4 million in 2023. This signifies an increase of approximately 100% over five years.

There are many features that distinguish DDoS attacks from other cyber attacks. The distinguishing features of DDoS attacks from other cyber attacks make DDoS attack detection studies more important than other attacks. DDoS attacks can be carried out in very large sizes. DDoS attack action can be done easily by attackers. It is not very easy for cyber security experts to completely block these attacks. Since the DDoS attack action pretends to be normal traffic, it cannot be completely detected by experts. Attacks have the potential to cause great material damage to systems and organizations. In addition to these material damages, it also causes loss of reputation. In order to minimize these damages, it becomes important to carry out detection and blocking activities in the fastest way.

The process of preventing the attack can terminate or delay the access of users seeking services. In this context, timely detection and swift mitigation of DDoS attacks are crucial. Numerous studies have been conducted for the detection of DDoS attacks. Despite the abundance of these studies, those carrying out DDoS attacks continuously devise new attack types over time. They

employ different methods to launch new attacks against the measures taken. The fact that executing DDoS attacks is not challenging gives rise to the need for new research, not only due to economic losses but also because it leads to reputational damage.

The purpose of this thesis is to detect DDoS attacks using machine learning methods. Alongside various machine learning techniques, two feature selectors have also been employed. Particularly, four different machine learning algorithms including Logistic Regression (LR), Multilayer Perceptron (MLP), Support Vector Machine (SVM), and Random Forest (RF) individually combined with Relief-F and mRMR feature selectors were employed.

The CICDDoS2019 dataset, shared by the Canadian Institute for Cybersecurity, is the most up-to-date dataset designed, taking into account the deficiencies in the institute's previous datasets. CICDDoS2019 includes benign and the latest DDoS attacks, resembling real data (PCAPs), along with timestamp, source and destination IPs, connection ports, protocols, and labeled flows based on CICFlowMeter-V3 used for network traffic analysis. This dataset encompasses various DDoS attack types such as PortMap, Network Basic Input/Output System (NetBIOS), Lightweight Directory Access Protocol (LDAP), Microsoft SQL Server (MSSQL), User Datagram Protocol (UDP), UDPLag, Simple Service Discovery Protocol (SSDP), WebDDoS, SYN, Network Time Protocol (NTP), Domain Name System (DNS), and Simple Network Management Protocol (SNMP). There are a total of 86 features and 50.063.112 examples in the dataset.

The dataset used in this study was obtained from the CICDDoS2019 dataset produced at the University of New Brunswick. The dataset comprises 86 features. Features with all values being zero and those with variance close to zero (considered to have no contribution to the study) were removed, resulting in the exclusion of 20 features. Consequently, a dataset with 66 potential prediction variables was obtained. There is one output variable, which contains information about attacks. If the data does not pertain to an attack, it includes information indicating normalcy. The utilized dataset encompasses a total of 10.000 entries, including four different types of attacks and normal traffic conditions.

Using the Relief-F and mRMR feature selection methods, the rankings of the predictive variables were determined. A total of 130 models were created, consisting of 65 Relief-F-based models and 65 mRMR-based models. The accuracy, precision, recall, and F1-Score metric values of all models were evaluated.

According to the Relief-F and mRMR feature selectors, features related to packet information and temporal features were more prominent among the best predictive variables. Furthermore, it was determined that the highest metric results were obtained with the RF method, independent of whether Relief-F or mRMR has been used as the feature selector. The lowest performance in all metrics was achieved by the LR method. Improvement in MLP, SVM, and RF methods was observed to be more significant in mRMR-based models compared to Relief-F-based models. There was no significant increase in performance obtained from models applied with feature

selection methods using MLP, SVM, and RF. In contrast, a considerable increase was observed in metric results obtained from models where the LR method was applied. The improvement in the LR method was found to be greater in Relief-F-based models compared to mRMR-based models.

GENİŞLETİLMİŞ ÖZET

İnternetin günlük hayatımızdaki rolü her geçen gün artmaktadır. İnternet hizmeti, çalışma hayatı, okul hayatı, sağlık sektörü, ticaret ve iletişim gibi hayatımızın hemen her alanında yaygın olarak kullanılmaktadır. İnternet kullanımı her geçen gün artmakta olup, Nesnelerin İnterneti (IoT) teknolojisinin gelişimi ile internete bağlı cihazların sayısı önemli ölçüde artmıştır. Artık kullanıcılar sadece bilgisayarlarla değil, televizyon, buzdolabı, çamaşır makinesi gibi ev aletleriyle de interneti kullanabilmektedirler. Bu artış, büyük kolaylıklar sağlarken aynı zamanda beraberinde riskleri de getirmektedir. İnternet aracılığıyla sunulan hizmetlerin sayısının artmasıyla birlikte ağ altyapısına yapılan saldırılar da artmaktadır. Saldırganlar, internete bağlanabilen neredeyse her türlü cihaza saldırı yapabilirler. İnternete bağlı cihazları zombi olarak kullanarak sistemi ele geçirebilir ve farklı cihazlara da saldırı gerçekleştirebilirler. Bu saldırıların en yaygını Dağıtılmış Hizmet Reddi (DDoS) saldırılarıdır.

DDoS saldırıları, bilişim dünyasını uzun süredir tehdit eden ve önleme çalışmalarının yapıldığı bir konu olmuştur. Devlete ait kurumlar, ticari şirketler ve vatandaşların kullandığı tüm sistemler bu tür saldırılara maruz kalmaktadır. Saldırılarla mücadele devam etmekte ve yeni yöntemlere karşı önlemler geliştirilmektedir.

Dünya tarihinde ilk DDoS saldırıları 1980'lerde gerçekleşmiştir. İlk büyük çaplı saldırı 1999'da Minnesota Üniversitesi'ne düzenlenmiş olup, üniversite ağı ciddi şekilde etkilenmiş ve 2 gün boyunca hizmet verilememiştir. Yapılan araştırmalara göre, her yıl DDoS saldırılarında ciddi bir artış gözlemlenmektedir. 2015 yılında DDoS saldırıları 500 Gbps boyutuna ulaşmış, 2016'da ise %60 artışla 800 Gbps boyutuna ulaşarak rekor kırmıştır (Jazi et al., 2017). Bu saldırılar ilerleyen yıllarda terabayt seviyelerine ulaşmıştır. CISCO'nun yaptığı bir analiz çalışmasına göre, 2018'de 7,9 milyon olan DDoS saldırı sayısı, 2019'da 9,5 milyon, 2020'de 10,8 milyon, 2021'de 12,1 milyon ve 2023'te 15,4 milyon adede ulaşmıştır. Bu, beş yılda yaklaşık %100'lük bir artışı göstermektedir.

DDoS saldırılarını diğer siber saldırılarından ayıran çok sayıda özellik bulunmaktadır. DDoS saldırılarının diğer siber saldırılardan ayıtr edici özellikler DDoS saldırı tespiti çalışmalarını diğer saldırılara göre önemli hale getirmektedir. DDoS saldırları çok büyük boyutlarda yapılabilir. DDoS saldırı eylemi saldrıganlar tarafından kolayca yapılabilmektedir. Siber güvenlik uzmanları bu saldırıları tamamen engellemeleri çok kolay olmamaktadır. DDoS saldırı eylemi normal trafikmiş gibi davrandığı için, uzmanlar tarafından tamamı tespit edilememektedir. Saldırıların sistemlere, kuruluşlara maddi açıdan çok büyük zarar verme potansiyeli bulunmaktadır. Bu maddi zararların yanı sıra itibar kaybına da sebep vermektedir. Bu zararları en aza indirmek için tespit, engelleme çalışmalarının en hızlı şekilde olması önemli hale gelmektedir.

Saldırıyı engelleme işlemi, hizmet almak isteyen kullanıcıların erişimini sonlandırabilir veya geciktirebilir. Bu bağlamda DDoS saldırılarının zamanında tespiti ve hızlı bir şekilde bertaraf edilmesi oldukça önemli bir durumdur. DDoS saldırılarının tespiti için çok sayıda çalışma

yapılmıştır. Çok sayıda çalışma yapılmış olmasına rağmen, DDoS saldırı yapanlar da her geçen zaman yeni saldırı çeşidi tasarlıyorlar. Alınan önlemlere karşı farklı yöntemlerle yeni saldırılar gerçekleştiriyorlar. DDoS saldırılarının gerçekleştirilmesinin zor olmaması, sadece ekonomik kayıp değil aynı zamanda itibar kaybına da sebep olması gibi sebeplerden dolayı yeni çalışmalara ihtiyaç var.

Bu tezin amacı, makine öğrenimi yöntemlerini kullanarak DDoS saldırılarını tespit etmektir. Çeşitli makine öğrenimi tekniklerinin yanı sıra, iki özellik seçicisi de kullanılmıştır. Özellikle, Lojistik Regresyon (LR), Çok Katmanlı Algılayıcı (MLP), Destek Vektör Makinesi (SVM) ve Rastgele Orman (RF) gibi dört farklı makine öğrenimi algoritması, Relief-F ve mRMR özellik seçicileri ile birlikte ayrı ayrı kullanılmıştır.

Kanada Siber Güvenlik Enstitüsü tarafından paylaşılan CICDDoS2019 veri seti, önceki enstitü veri setlerindeki eksiklikleri göz önüne alarak tasarlanmış en güncel veri setidir. CICDDoS2019, gerçek verilere (PCAP'ler) benzeyen iyi huylu ve en güncel DDoS saldırılarıyla birlikte, zaman damgası, kaynak ve hedef IP'ler, bağlantı noktaları, protokoller ve saldırılara (CSV dosyaları) dayalı etiketli akışlar içeren CICFlowMeter-V3 tarafından kullanılan ağ trafiği analizinin sonuçlarını içermektedir. Bu veri kümesinde PortMap, Ağ Temel Giriş/Çıkış Sistemi (NetBIOS), Hafif Dizin Erişim Protokolü (LDAP), Microsoft SQL Sunucusu (MSSQL), Kullanıcı Datagram Protokolü (UDP), UDPLag, Basit Hizmet Algılama Protokolü (SSDP), WebDDoS, SYN, Network Time Protocol (NTP), Alan Adı Sistemi (DNS) ve Basit Ağ Yönetim Protokolü (SNMP) gibi farklı DDoS saldırı çeşitleri bulunmaktadır. Veri setinde toplamda 86 öznitelik ve 50.063.112 örnek bulunmaktadır.

Bu çalışmada kullanılan veri seti, New Brunswick Üniversitesi'nde üretilen CICDDoS2019 veri setinden elde edilmiştir. Veri setinde 86 özellik bulunmaktadır. Tüm değerleri sıfır olan ve varyansı sıfıra yakın olan özellikler çalışmaya katkısı olmadığı gerekçesiyle 20 adet özellik çıkarılmıştır. Böylece, 66 adet tahmin değişkenli bir veri seti elde edilmiştir. Bir adet çıkış değişkeni bulunmaktadır. Bu değişken saldırı bilgisini içermektedir. Eğer saldırı verisi değilse normal olduğunu ifade eden bilgi bulunmaktadır. Kullanılan veri seti, dört farklı saldırı türünü ve normal trafik durumunu içeren toplam 10.000 veriyi kapsamaktadır.

Relief-F ve mRMR özellik seçme yöntemleri kullanılarak tahmin değişkenlerinin sıraları belirlendi. Relief-F tabanlı 65 adet model ve mRMR tabanlı 65 adet model olmak üzere toplam 130 model oluşturuldu. Tüm modellerin doğruluk, hassasiyet, duyarlılık ve F1-skoru metrik değerleri hesaplanmıştır.

Relief-F ve mRMR özellik seçicilerine göre, paket bilgileri ve zamansal özelliklerle ilgili özelliklerin en iyi tahmin edici değişkenler arasında olduğu belirlenmiştir. Ayrıca, Relief-F veya mRMR'nin özellik seçicisi olarak kullanılıp kullanılmadığından bağımsız olarak en yüksek metrik sonuçların RF yöntemiyle elde edildiği tespit edilmiştir. Tüm metriklerde en düşük performans LR yöntemiyle elde edilmiştir. MLP, SVM ve RF yöntemlerindeki iyileşmenin, Relief-F tabanlı

modellere kıyasla mRMR tabanlı modellerde daha belirgin olduğu gözlenmiştir. MLP, SVM ve RF kullanarak uygulanan modellerden elde edilen performansta anlamlı bir artış gözlenmemiştir. Bununla birlikte, LR yönteminin uygulandığı modellerden elde edilen metrik sonuçlarında dikkate değer bir artış gözlenmiştir. LR yöntemindeki iyileşmenin, mRMR tabanlı modellere kıyasla Relief-F tabanlı modellerde daha büyük olduğu tespit edilmiştir.

ACKNOWLEDGEMENTS

I wish to express my gratitude to my supervisor Assoc. Prof. Dr. Fatih ABUT for his guidance, patience, advice, encouragement, and insight throughout this study.

I would like to thank Asst. Prof. Dr. Onur POLAT and Asst. Prof. Dr. Barış ATA for their comments, suggestions, and serving on my committee.

I would like to thank Assoc. Prof. Dr. Erdal Başaran, Asst. Prof. Dr. Murat GENÇ, Hasan Tacettin ÜLBEĞİ, Cumali KAMURAN, Mehmet DOĞAN, Mehmet ÖZTÜRK, Vahdet Cemil ALTUN and Umut Ahmet ÇETİN for their suggestions, help, comments, advice, and support.

I would like to thank the Brunswick University for sharing the dataset with the public used in this thesis.

Last but not the least, I would like to thank my mother, my father, my wife, Ahmet Yahya, my brother and his family, my sister and her family, my relatives, and my friends for their never-ending support, encouragement, and patience throughout my life.

LIST OF TABLES

LIST OF FIGURES

SYMBOLS AND ABBREVIATIONS

DDoS : Distributed Daniel of Service
Gbps : Giga Bits Per Second
IoT : Internet of Things
LR : Logistic Regression
ML : Machine Learning
MLP : Multi-Layer Perceptron
mRMR : minimum Redundancy Maximum Relevance
RF : Random Forest
SDN : Software Defined Networking
SVM : Support Vector Machine

1. INTRODUCTION

1.1. Overview of DDoS Attack

Distributed Denial of Service (DDoS) attacks are cyber-attacks aimed at rendering internet-connected systems and services dysfunctional. DDoS attacks are carried out by numerous compromised devices. The compromised devices are called 'bots,' and the collective term for all bots that carry out DDoS attacks is referred to as a 'botnet.' DDoS attacks began in the 1980s, with the first large-scale attack occurring in 1999 at the University of Minnesota. The university's network was severely affected, and services couldn't be provided for about 2 days. According to research, there is a significant increase in DDoS attacks every year. DDoS attacks reached a size of 500 Gbps in 2015 and set a record by reaching 800 Gbps in 2016, with a 60% increase (Jazi et al., 2017). In the following years, they reached the terabyte level. An analysis of DDoS attacks was conducted by CISCO. According to this analysis, the number of DDoS attacks, which was 7.9 million in 2018, increased to 9.5 million in 2019, 10.8 million in 2020, 12.1 million in 2021, and 15.4 million in 2023. There is an approximately 100% increase in five years (Deniz, 2023). The key factors that differentiate this attack from others include the ease with which attacks of large scale can be executed and the difficulty in distinguishing and preventing them. The process of blocking the attack can lead to the termination or delay of access for users seeking services. In this context, timely detection and rapid elimination of DDoS attacks are crucial.

With the widespread use of the Internet, the diversity of DDoS attacks has increased. DDoS attacks on the Internet of Things, Cloud environments and Software-based networks are realised. Intelligent devices spread to all areas of our lives have caused a serious increase in internet usage. According to CISCO's study, the number of devices connected to the internet will reach 50 billion. Citation. Internet-connected devices are becoming widespread almost everywhere, from factories, private institutions, government agencies, schools and even home appliances. Cloud technology is used to store the data used in these environments. Data processing management centers manage network networks with software systems and devices connected to the internet. All these are the target of DDoS attacks. If precautions are not taken, it will cause serious material and moral losses.

With the proliferation of the internet, the variety of DDoS attacks has also increased. DDoS attacks targeting the Internet of Things, cloud environments, and software-based networks are occurring. Smart devices, spreading across every aspect of our lives, have led to a significant increase in internet usage. According to a study by CISCO, the number of connected devices to the internet is expected to reach 500 billion by 2030 (Ji et al., 2021). Connected devices to the internet are becoming widespread, from factories, private institutions, and government agencies to schools and even household appliances. Cloud technology is used for storing data in these environments. Information technology management centers handle network systems with software and internet-

connected devices. All of these are susceptible to DDoS attacks. Without preventative measures, they can lead to substantial financial and non-material losses.

1.2. Motivation for DDoS Attack Detection

With the rapid developments in today's digital age, there have been serious developments in cyber-attacks on internet-connected devices. DDoS attacks are at the forefront of these attacks. Many government agencies, commercial companies, and universities have been affected by these attacks, leading not only to financial losses but also to reputational damage. Every internet-connectable device faces the possibility of a DDoS attack.

The internet has become central to our lives. Internet access is now considered essential in almost every sector. With Industry 4.0, the necessity of internet usage has increased even further. For internet-dependent systems to continue functioning, technical requirements as well as security measures are essential. Having a well-functioning system in place is not sufficient to prevent service disruptions; measures must also be taken against potential attacks.

Not only computers and phones but also any device capable of connecting to the internet can be used in carrying out attacks. There has been a significant increase in the number of devices connecting to the internet every year, and substantial increases in internet-connectable devices are expected in the coming years. Alongside these developments, the increasing competition between companies has led to insufficient investments in device security to reduce costs. As a result, systems are at high risk. Therefore, all systems must be prepared for large-scale DDoS attacks. The fact that the number of internet-connected devices is in the billions underscores the importance of extensive efforts. In systems with large datasets, conducting studies using machine learning methods has become crucial. Failure to detect DDoS attacks and the problems faced by systems exposed to attacks can lead to serious losses. For example, if a DDoS attack is launched on a government's hospital systems, it could disrupt the treatment of millions of patients until the system is restored. Attacks on banks could result in significant financial losses due to the disruption of numerous financial transactions. DDoS attacks on universities could disrupt education and impede scientific research. Attacks on public institutions could disrupt public services, resulting in harm to the public and causing significant economic losses by halting production in factories.

The internet has become a necessity in almost every aspect of human life, leading to the widespread use of internet-connected devices. With this proliferation, we can provide examples of the negative consequences that may arise when DDoS attacks occur.

Numerous studies have been conducted to detect DDoS attacks. Researchers at different universities have created numerous DDoS datasets. Various machine learning methods have been used for DDoS attack detection, and their successes have been compared. Despite all these efforts, there is still a need for further research into DDoS attack detection. While DDoS attack detection

may be relatively easy compared to other cyber-attacks, multiple attacks can occur simultaneously. Completely preventing them is a highly challenging task.

All the difficulties mentioned, coupled with the ease of execution, enable attackers to devise new methods over time. Therefore, despite the numerous studies conducted, additional research is needed to keep pace with innovations in attacks and to access better solutions than the existing ones.

1.3. Purpose and Contributions of the Thesis

This thesis aims to develop new DDoS attack prediction models by employing various machine learning and feature selection methods. Relief-F and mRMR were employed as feature selectors. A dataset of 10.000 rows was created, ensuring equal numbers of four different attack data and normal traffic data. Four different machine learning algorithms, namely Logistic Regression (LR), Multilayer Perceptron (MLP), Support Vector Machine (SVM), and Random Forest (RF), were utilized. The dataset encompasses a diverse array of predictor variables, notably including packet size, transmission time, and indicators of attack occurrences, among numerous others. Performance metrics, including accuracy, precision, specificity, recall, and F1 score, were used to determine the models' performance in the study.

The contributions of this thesis to the scientific community can be summarized as follows.

- By using two feature selectors and four machine learning classifiers, this thesis allows the ranking of these applied methods for prediction of DDoS attack detection.
- To the best of our knowledge, there is no such comprehensive study in the literature on DDoS prediction using 2 attribute selectors and 4-5 machine learning methods.
- By using 130 models, a comprehensive study was carried out in terms of the number of models.

1.4. Dataset Generation

Numerous dataset studies have been conducted for the detection of DDoS attacks. The created datasets have not only been valuable but also formed the basis of many scientific articles. Despite all these positive developments, the emergence of new attacks has necessitated the need for new datasets. In this dataset, attacks are divided into two main groups: reflection-based and exploitation-based.

Reflection-based DDoS attacks are attacks where the attacker hides their identity. Many packets are sent from the attacker to the target device to overwhelm it. This process is carried out by redirecting to reflective servers. These attacks are carried out through TCP and UDP protocols. Reflection-based attacks are illustrated in Figure 1.1.

Exploitation-based DoS attacks are carried out while the identity of the attacker is concealed. They are conducted using TCP and UDP transfer protocols. UDP packets are sent to target devices

at very high speeds. As a result of the attack, the bandwidth of the target network is exhausted, rendering the system inoperable. UDP-Lag attack is a type of attack aimed at slowing down the connection of the target system. The types of attacks are shown in Figure 1.1.

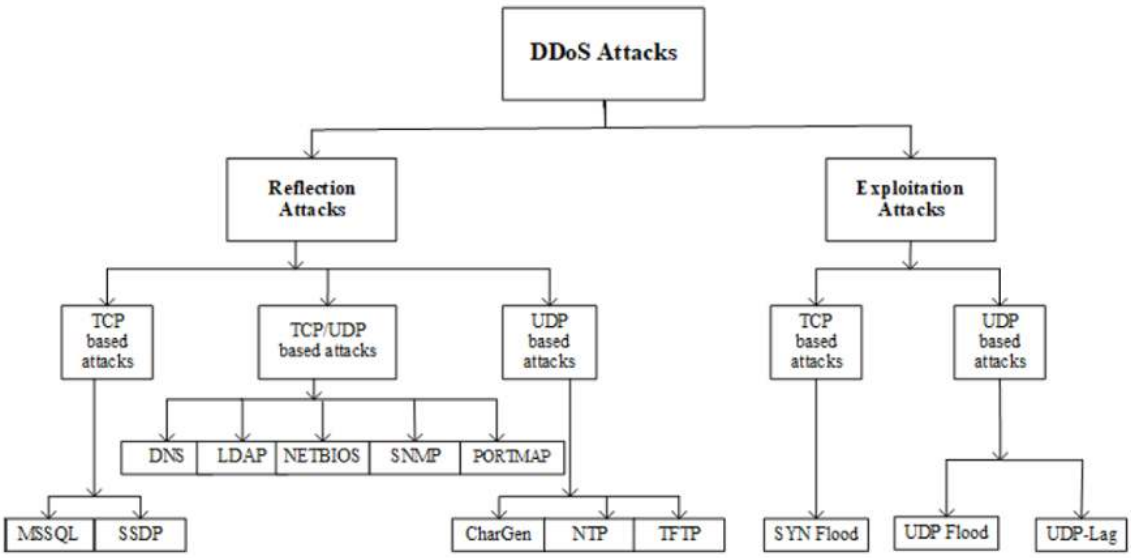

Figure 1.1. DDoS Attack Classification (Sharafaldin et al., 2018)

The CICDDoS2019 dataset was created in 2019 by Sharafaldin et al. It was downloaded from the database of the University of New Brunswick (UNB).

The data in CICDDoS2019 has 83 features. These features include many features information such as traffic occurrence time, prtcl information, ip information of the device where the packet occurs, target ip information. Along with these, there is also information about whether the traffic is harmless or whether it contains a DDoS attack. The data set was created in two days. On the first day, PortMap, NetBIOS, LDAP, MSSQL, UDP, UDP-Lag, SYN DDoS attacks were created, respectively. On the second day, NTP, DNS, LDAP, MSSQL, NetBIOS, SNMP, SSDP, UDP, UDP-Lag, WebDDoS, SYN and TFTP attacks were created, respectively. Files were created with pcap extension.

Data containing normal traffic with four different types of attacks were used: DNS, TFTP, SYN Flood, and NetBIOS attacks. Equal amounts were taken from all of them. There is a total of 10000 data. In this study, 66 features were utilized.

Variables generally contain information related to packets. Among this information are details such as the type of protocol used, packet sizes, header sizes, transmission times, standard deviation sizes of packets transmitted in both forward and reverse directions, and protocol information used in communication. Additionally, the target port number and the port information to which the packet is sent are provided. The maximum, minimum, and average sizes of packets transmitted in both forward and reverse directions, along with the standard deviation values of the packets, are specified. The maximum, minimum, and average values of packets sent per second and

transmission times are also among these features. Values with infinity values are denoted as -1 in this dataset.

All variables used in the study are detailed in Table 1.1 through Table 1.3. In addition to their descriptions, the data type and abbreviations are also provided.

Table 1.1. Descriptions and abbreviations of the features used in the study – Part 1

Feature Name	Abbreviation	Description	Unit
Destination Port	dport	Port number to which the packets was sent	Number
Protocol	prtcl	Types of communications internet protocols	Number
Flow Duration	fdrtn	Duration of the flow in Microsecond	Microsecond
Total Fwd Packets	tfpckts	Total packets in the forward direction	Number
Total Backward Packets	tbpckts	Total packets in the backward direction	Number
Total Length of Fwd Packets	tlfpckts	Total size of packet in forward direction	Number
Total Length of Bwd Packets	tlbpckts	Total size of packet in backward direction	Number
Fwd Packet Length Max	fplmax	Maximum size of packet in forward direction	Bytes
Fwd Packet Length Min	fplmin	Minimum size of packet in forward direction	Bytes
Fwd Packet Length Mean	fplmean	Mean size of packet in forward direction	Bytes
Fwd Packet Length Std	fplstd	Standard deviation size of packet in forward direction	Bytes
Bwd Packet Length Max	bplmax	Maximum size of packet in backward direction	Bytes
Bwd Packet Length Min	mplmin	Minimum size of packet in backward direction	Bytes
Bwd Packet Length Mean	bplmean	Mean size of packet in backward direction	Bytes

Table 1.2. Descriptions and abbreviations of the features used in the study – Part 2

Feature Name	Abbreviation	Description	Unit
Bwd Packet Length Std	bplstd	Standard deviation size of packet in backward direction	Bytes
Flow Bytes/s	fbytes	Number of flow bytes per second	Bytes/second
Flow Packets/s	fpckts	Number of flow packets per second	Bytes/second
Flow IAT Mean	fimean	Mean time between two packets sent in the flow	Microsecond
Flow IAT Std	fistd	Standard deviation time between two packets sent in the flow	Microsecond
Flow IAT Max	fimax	Maximum time between two packets sent in the flow	Microsecond
Flow IAT Min	fimin	Minimum time between two packets sent in the flow	Microsecond
Fwd IAT Total	fwditotal	Total time between two packets sent in the forward direction	Microsecond
Fwd IAT Mean	fwdimean	Mean time between two packets sent in the forward direction	Microsecond
Fwd IAT Std	fwdistd	Standart Deviation time between two packets sent in the forward direction	Microsecond
Fwd IAT Max	fwdimax	Maximum time between two packets sent in the forward direction	Microsecond
Fwd IAT Min	fwdimin	Minimum time between two packets sent in the forward direction	Microsecond
Bwd IAT Total	bitotal	Total time between two packets sent in the backward direction	Microsecond
Bwd IAT Mean	bimean	Mean time between two packets sent in the backward direction	Microsecond
Bwd IAT Std	bistd	Standart Deviation time between two packets sent in the backward direction	Microsecond
Bwd IAT Max	bimax	Maximum time between two packets sent in the backward direction	Microsecond
Bwd IAT Min	bimin	Minimum time between two packets sent in the backward direction	Microsecond
Fwd PSH Flags	fpflg	Number of times the PSH flag was set in packets travelling in the forward direction (0 for UDP)	Boolean
Fwd Header Length	fwdhl	Total bytes used for headers in the forward direction	Bytes
Bwd Header Length	bwdhl	Total bytes used for headers in the backward direction	Bytes
Fwd Packets/s	fwdpckts	Number of forward packets per second	Bytes/second
Bwd Packets/s	bwdpckts	Number of backward packets per second	Bytes/second
Min Packet Length	minpl	Minimum length of a packet	Bytes
Max Packet Length	maxpl	Maximum length of a packet	Bytes
Packet Length Mean	plmean	Mean length of a packet	Bytes
Packet Length Std	plstd	Standart Deviation length of a packet	Bytes
Packet Length Variance	plvrnc	Variance length of a packet	Bytes
RST Flag Count	rdtfc	Number of packets with RST	Boolean
ACK Flag Count	ackfc	Number of packets with ACK	Boolean

Table 1.3. Descriptions and abbreviations of the features used in the study – Part 3

Feature Name	Abbreviation	Description	Unit
URG Flag Count	urgfc	Number of packets with URG	Boolean
CWE Flag Count	cwefc	Number of packets with CWE	Boolean
Down/Up Ratio	duratio	Download and upload ratio	Number
Average Packet Size	avrpcks	Average size of packet	Bytes
Avg Fwd Segment Size	afwdss	Average size observed in the forward direction	Bytes
Avg Bwd Segment Size	abwdss	Average size observed in the backward direction	Bytes
Fwd Header Length.1	fwdhl1	Total bytes used for headers in the backward direction	Bytes
Subflow Fwd Packets	sbfwdpckt	The average number of packets in a sub flow in the backward direction	Number
Subflow Fwd Bytes	sbfwdbyt	The average number of bytes in a sub flow in the backward direction	Bytes
Subflow Bwd Packets	sbbwdpckt	The average number of packets in a sub flow in the forward direction	Number
Subflow Bwd Bytes	sbbwdbyt	The average number of bytes in a sub flow in the forward direction	Bytes
Init_Win_bytes_forward	iwbf	The total number of bytes sent in initial window in the forward direction	Bytes
Init_Win_bytes_backward	iwbb	The total number of bytes sent in initial window in the backward direction	Bytes
act_data_pkt_fwd	adpf	Count of packets with at least 1 byte of TCP data payload in the forward direction	Number
min_seg_size_forward	mssf	Minimum segment size observed in the forward direction	Bytes
Active Mean	actmean	Mean time a flow was active before becoming idle	Bytes
Active Std	actstd	Standart Deviation time a flow was active before becoming idle	Bytes
Active Max	actmax	Maximum time a flow was active before becoming idle	Microsecond
Active Min	actmin	Minimum time a flow was active before becoming idle	Microsecond
Idle Mean	imean	Mean time a flow was idle before becoming active	Microsecond
Idle Std	istd	Standart Deviation time a flow was idle before becoming active	Microsecond
Idle Max	imax	Maximum time a flow was idle before becoming active	Microsecond
Idle Min	imin	Minimum time a flow was idle before becoming active	Microsecond

Table 1.4. Descriptive statistics of the dataset – Part 1

Variables	Standard Deviation	Mean
Dport	7,70E+09	1,11E+09
Prtcl	829097,6	60187,24
Fdrtn	8,75E+12	9,29E+11
Tfpckts	3704707	447111,4
Tbpckts	4001897	420858,8
Tlfpckts	1,30E+09	69355053
Tlbpckts	3,32E+09	2,20E+08
Fplmax	1,12E+08	12479030
Fplmin	10637583	1727838
Fplmean	2,88E+09	57704628
Fplstd	48967708	6517393
Bplmax	1,96E+08	21603172
Mplmin	24726214	3730758
Bplmean	8,77E+09	1,33E+08
Bplstd	97939681	13206721
Fbytes	3,91E+13	4,23E+12
Fpckts	8,91E+10	2,41E+10
fimean	6,29E+12	3,29E+11
Fistd	1,84E+12	4,84E+11
Fimax	4,39E+12	4,41E+11
Fimin	1,43E+11	4,67E+09
Fwditotal	8,75E+12	9,09E+11
Fwdimean	1,23E+12	2,79E+11
Fwdistd	2,05E+12	5,21E+11
Fwdimax	4,31E+12	4,17E+11
Fwdimin	1,09E+11	3,11E+09
Bitotal	7,59E+12	6,92E+11
Bimean	8,95E+11	8,36E+10
Bistd	1,72E+12	1,87E+11
Bimax	3,94E+12	3,40E+11
Bimin	2658526	275375,4
Fpflg	0,147855	0,022359
Fwdhl	90375734	12260701
Bwdhl	101048526	12269761
Fwdpckts	1,16E+11	2,68E+10
Bwdpckts	3,96E+09	7,11E+08
Minpl	9786224	1665589
Maxpl	2,12E+08	25076314
Plmean	4,33E+08	14798665
Plstd	86512965	14596109
Plvrnc	6,84E+10	8,22E+09
Rdtfc	0,147855	0,022359
Ackfc	0,419991	0,228661
Urgfc	0,279076	0,085121
Cwefc	0,16591	0,028326
Duratio	215711,9	13766,4
Avrpcks	6,44E+09	1,10E+08
Afwdss	8,22E+09	1,81E+08
Abwdss	7,68E+09	1,41E+08
Fwdhl.1	100545723	13255566
Sbfwdpckt	3746988	435218,5

Table 1.5. Descriptive statistics of the dataset – Part 2

Variables	Standard Deviation	Mean
Sbfwdbyt	9,70E+08	64766748
Sbbwdpckt	4197890	419729,2
Sbbwdbyt	3,13E+09	2,11E+08
İwbf	4,16E+09	4,90E+08
İwbb	2,44E+09	2,02E+08
adpf	1973617	184522,2
mssf	22956249	1338176
Actmean	2,39E+11	9,68E+09
Actstd	5,80E+10	3,48E+09
Actmax	1,35E+11	8,28E+09
Actmin	1,35E+11	6,10E+09
İmean	4,36E+12	5,35E+11
İstd	1,10E+12	1,28E+11
İmax	4,31E+12	4,09E+11
İmin	3,82E+12	3,50E+11

1.5. Roadmap of the Thesis

In Chapter 1, the concept of DDoS attack detection is introduced, and the necessity of studies for detecting DDoS attacks is explained. Detailed information is provided about the dataset used in the study. The variables within the dataset are explained. Finally, a roadmap containing information about the content of the chapters in the thesis is presented.

In Chapter 2, academic studies related to DDoS attack detection are being thoroughly examined under different headings. These different headings have been determined according to the environments in which DDoS attacks occur: DDoS attacks conducted in SDN environment, DDoS attacks conducted in IoT environment, attacks conducted in cloud environment, and DDoS attacks conducted in other environments.

In Chapter 3, information about the machine learning methods used in the thesis has been provided. Definitions of machine learning methods have been made, their mathematical models have been expressed. In addition to the machine learning methods used, feature selectors have also been explained, and their mathematical formulas have been expressed.

In Chapter 4, information about feature selectors has been provided, and the importance ranks of variables have been determined using feature selector methods. The rankings of variables have been expressed in a table. Definitions of evaluation metrics used in the thesis have been made, and their mathematical formulas have been shown. The parameters of the machine learning methods used in the thesis have been explained, and the value ranges of the parameters have been indicated.

In Chapter 5, the results have been examined under two different headings based on the used feature selectors. The results obtained from all models have been listed. These results have been compared and conclusions have been drawn. Additionally, the results obtained from the experiments have been compared with the results obtained from similar studies conducted previously.

In Chapter 6, a synthesis of discussions conducted based on the results obtained from experiments in the thesis has been presented. The obtained results have been presented in brief summary form. Recommendations for future studies have been provided, and identified deficiencies related to the topic have been expressed to contribute to academic research.

2. RELATED WORKS

Numerous studies have been conducted on DDoS attack detection using machine learning methods. In this section, information about studies related to DDoS attack detection is provided. Literature is classified according to environment of DDoS attack. This chapter is divided in four sections: DDoS attack detection in SDN environments, DDoS attack detection in cloud environments, DDoS attack detection in IoT environments, and other environments.

2.1. DDoS Attack Detection in SDN Environments

(Tahirou et al., 2023) conducted a study to detect and mitigate DDoS attacks in SDN environments. Three different machine learning methods, namely KNN, SVM, and NB, were employed. The CICDDoS2019 dataset, published at the UNB, was utilized. From the dataset containing more than 80 features, 26 features were selected for the study. Accuracy, precision, recall, and F-measure metrics were measured. The best results were obtained with SVM, achieving accuracy, precision, recall, and F-measure scores of 98.9%, 99%, 98%, and 97.9%, respectively.

(Kohnehshahri et al., 2022) conducted a study to detect UDP-based Distributed Reflection Denial of Service (DRDoS) attacks in an SDN environment using machine learning algorithms in 2022. The study utilized the CICIDS2017 dataset created by the UNB in 2017. Five machine learning algorithms, namely Decision Tree (DT), RF, SVM, Gradient Boosting Classifier (GBC), and Adaptive Boosting algorithms (Ada), were employed. The aim was to detect DRDoS attacks using variables such as source IP address, destination IP address, the number of packets in a flow, and the volume of packets in a flow. Analyzing differences in the size and number of packets requested and transmitted in response is considered crucial for attack detection. Overall, the most successful results were achieved with SVM and GBC. The proposed method was compared with previous approaches, demonstrating its superior performance. A success rate of 100% has been achieved in all metrics, including accuracy, precision, recall, and F1-score.

(Makuvaza et al., 2021) conducted a study to detect DDoS attacks in the SDN environment using the DNN method. The study employed the CICIDS2017 dataset, created by Sharafaldin et al. at UNB, containing 3.1 million data points. DDoS attack detection in the SDN environment was compared with different studies, demonstrating superior effectiveness. The study identified estimation variables as backward packet length, fdrtn, avrpcks, and flow inter-arrival time. The obtained metrics, including accuracy, recall, precision, and F1-score, were 96.67%, 97.29%, 97.21%, and 97.25%, respectively.

(Wang & Liu, 2020) conducted a study about DDoS attack detection using machine learning algorithms within the SDN framework. The study employed four machine learning algorithms, namely Convolutional Neural Networks (CNN), DNN, SVM, and DT. CNN emerged as the most successful among them, showcasing its effectiveness in preventing DDoS attacks in the SDN

environment. The CICIDS2017 dataset, created by UNB in 2017, was utilized in the study, comprising 2,830,743 data points. The measured metric results for accuracy, precision, recall, and F1-score were 98.98%, 98.99%, 98.96%, and 98.97%, respectively.

(Polat et al., 2020) conducted a study to detect DDoS attacks using the SDN method. Estimation variables such as memory, processor, bandwidth, and CPU capacity were considered. The researchers generated the dataset used in the study. Classification models including SVM, Naive Bayes (NB), Artificial Neural Networks (ANN), and K-Nearest Neighbors (KNN) were employed. KNN exhibited the highest accuracy rate among them. The study revealed that SDN can effectively detect network scanning, inter-layer attacks, and malicious software. The dataset created for the study comprised 129,000 data points. The researchers performed the feature selection process using the Matlab platform. In their analysis, they compared the performance metrics of machine learning algorithms both before and after the feature selection process, observing that the metrics improved after the selection. The results for accuracy, precision, F1-score, specificity, and sensitivity obtained with KNN were determined as 98.30%, 97.72%, 97.70%, 99.45%, and 97.73%, respectively.

(Kyaw et al., 2020) proposed a study to detect vulnerabilities in SDN-designed systems, making them susceptible to DDoS attacks. The study proposed a solution utilizing machine learning methods for the detection and prevention of DDoS attacks. Specifically, SVM and Polynomial SVM methods were employed for detecting UDP flood attacks in the SDN environment, and their performance metrics were meticulously compared. The research was conducted using the KDD99 dataset, established in 1999, comprising 650 data points. Key forecast variables included no_packets, no_bytes, no_dst, duration, and Port_no. The accuracy and precision metric values were determined to be 95.38% and 95.05%, respectively.

(Tonkal et al., 2021) conducted a study to determine DDoS attacks in an SDN environment. SDN offers various advantages, including usability and manageability, making it a preferred choice for next-generation networks. However, these conveniences come with inherent risks, particularly in the context of DDoS attacks, which pose a significant threat to SDN environments. The study utilized machine learning algorithms enhanced with Neighborhood Component Analysis (NCA) to classify SDN traffic into hacked and normal categories. Estimation variables included byte_count, duration_sec, packet rate, and packets per flow. The 'DDoS attack SDN Dataset' was employed, consisting of 104.345 data points. The study highlighted the dataset's superiority over previous ones and emphasized its up-to-dateness. Machine learning algorithms such as ANN, KNN, DT, and SVM were applied. Despite DT's superiority in precision, the overall effectiveness was demonstrated by the ANN algorithm. The study compared accuracy, precision, sensitivity, F1-score, and specificity metric values, with the ANN method showing the highest success rates: 97.35%, 97.78%, 95.55%, 96.65%, and 98.55%, respectively.

(Tan et al., 2020) conducted a study with two machine learning algorithms to address the vulnerability of SDN technology, which has recently gained widespread adoption, against DDoS

attacks. The study utilized the NSL-KDD dataset, originally developed by Tavallaee et al. in 2009, comprising a total of 1,152,281 data points (Tavallaee et al., 2009). Additionally, a novel machine learning approach was introduced, combining K-Means and KNN, resulting in more successful outcomes compared to other methods. Estimation variables included average bytes per second at the time, average durations per flow, the percentage of pair-flow, variation rate of single flows, and the percentage of flows with small packets. The accuracy and recall metric results were 98.85% and 98.47%, respectively.

Table 2.1. Overview of DDoS attack detection studies in SDN environment – Part 1

Study	**Number of Data**	**Method**	**Predictor Variables**	**Performance Metrics and Values (%)**
(Tahirou et al., 2023)	n.a.	KNN, SVM, NB	Fwd Packet Length, Min PacketLength	Accuracy = 98.90, Precision = 99.00, Recall = 98.00, F1-score = 97.90
(Kohnehshahri et al., 2022)	n.a.	SVM, GBC, RF, DT, Ada	Total requested packets	Accuracy = 100.00, Precision = 100.00, Recall = 100.00, F1-score = 100.00
(Makuvaza et al., 2021)	3.100.000	DNN	BPL, FD, APS, FIAT	Accuracy = 96.67, Precision = 97.21, Recall = 97.29, F1-score = 97.25
(Wang & Liu, 2020)	2.830.743	CNN, DNN, SVM, DT	SA_IP, DA_IP, SP, DP, PL, Prtcl	Accuracy = 98.98, Precision = 98,99, Recall = 98,96, F1-score = 98.97
(Kyaw et al., 2020)	650	P- SVM, L-SVM	No_packets, no_bytes, no_dst, duration, Port_no	Accuracy = 95.38, Precision = 95.05
(Tonkal et al., 2021)	104.345	ANN, DT, KNN, SVM	byte_count, duration_sec, packet rate, and packet per flow	Accuracy = 97.35, Precision = 97.78, Sensitivity = 95.55, F1-score = 96.65
(Polat et al., 2020)	129.000	KNN, SVM, NB, ANN	Memory, processor, bandwith, CPU capacity	Accuracy = 98.30, Precision = 97.72, F1-score = 97.70, Sensitivity = 97.73
(Tan et al., 2020)	1.152.281	K-Means, KNN	ABST, ADF, PPF, VRSF, PFSP	Accuracy = 98.85, recall = 98.47

Abbreviations: n.a.: Not Available, KNN: K-Nearest Neighbors, FIAT: Flow Inter Arrival Time, BPL: Backward Packet Length, FD: Flow Duration, P-SVM: Polynomial SVM, L-SVM: Linear SVM, GBC: Gradient Boosting Classifier, DT: Decision Tree, RF: Random Forest, CNN: Convolutional Neural Networks, RT: Random Tree, REPT: Reduced Error Pruning Tree, DNN: Deep Neural Network, KPCA: Kernel Principal Component Analysis, GA: Genetic Algorithm, ANN: Artificial Neural Networks, ELM: Extreme Learning Mach, SA_IP: Source IP address, DA_IP: Destination IP Address, APS: Avrpcks, SP: Source Port, DP: Destination Port, PL: Packet Length, ABST: Average bytes per second at a time, ADF: Average durations, per flow, PPF: percentage of pair-flow, VRSF: variation rate of single flows, PFSP: percentage of flows with small packets.

2.2. DDoS Attack Detection in Cloud Environments

(Pandithurai et al., 2024) conducted a study on the detection and prevention of DDoS attacks in the cloud environment using the Bi-LSTM-based porcupine optimization method. In this study, comparisons were made with LSTM, DNN, ANN, and DBN algorithms. Accuracy, sensitivity, specificity, and precision metric values were determined, yielding success rates of 97%, 95%, 90%, and 94%, respectively. The dataset used in the study consists of 12.794.627 data points and includes 84 features.

(Islam et al., 2023) proposed a study on the detection and prevention of DOS attacks in the cloud environment. A hybrid model was created by combining SVM, KNN, and LR methods. The study employed the Relief-F feature selection method. This study utilized a dataset gathered from diverse sources, which underwent preprocessing to extract pertinent features essential for the detection of attacks. The best predictive variables were identified as std_cpu_util, avg_memory_util, and max_memory_util. Accuracy, precision, recall, and F1 score values were measured as 95%, 96%, 93%, and 94%, respectively.

(Kiruthiga et al., 2022) conducted a study to detect DDoS attacks in the cloud using the Deep Generative Radial Neural Network (DGRNN) method. Comparisons were conducted with the Iterative Dichotomiser 3 Maximum Multifactor Dimensionality Posteriori Method (ID3-MMDP), Ada Shuffled Leaping Optimal Selection (ASLOS), and DNN methods. It was determined that the DGRNN method was more successful with a 90% accuracy rate. The NSL-KDD dataset was utilized in the study, with estimation variables including traffic, time response, src-bytes, dst-bytes, and wrong fragment count. NSL-KDD was developed by Tavallaee et al. in 2009. The study observed that the success rate of the proposed method is better than the detection results achieved with other methods.

(Gumaste et al., 2021) proposed a study on an intrusion detection method in the cloud environment with OpenStack software. Three machine learning methods were employed, including RF, DT, and LR. It was observed that the results obtained with RF were better than the others. In this study, the KDD CUP dataset and real-time cloud traffic were utilized. The dataset was created in 1999, and only ICMP flood attacks were considered. Predictor variables such as prtcl type, service, src bytes, dest bytes, count, srv count, and same srv rate were determined. High results were achieved in the experiments with the KDD Cup dataset, with accuracy and precision values measured as 99.21% and 99.91%, respectively.

(Kushwah & Ranga, 2021) developed a new approach to detect DDoS attacks on cloud services on the internet. Performance metrics that emerged as a result of experiments on four different datasets with a large number of machine learning methods were compared. They developed the Self-Adaptive Evolutionary Extreme Learning Machine (SaE-ELM) System and adapted it to the intrusion detection system. NSL-KDD, ISCX IDS 2012, UNSW-NB15, and CICIDS 2017. After testing on four different datasets, comparisons were made with commonly used machine learning

methods such as ANN, DT, and SVM. NSL-KDD was developed (Tavallaee et al., 2009). ISCX IDS was developed in 2012, and CICIDS 2017 was developed at UNB in 2017. The UNSW-NB15 dataset was developed at UNSW Sydney University. The most successful result among these datasets was obtained with the CICIDS 2017 dataset. Forecast variables Flags, fdrtn, flow inter-arrival time, number of bytes from destination to source, number of bytes from source to destination, number of packets from destination to source, number of packets from source to destination, Prtcl name, Service name has been determined. Performance metrics accuracy, sensitivity, specificity, precision, and F-score were measured as 99.99%, 100%, 100%, 100%, 100%, respectively.

(Sambangi & Gondi, 2020) presented an approach to contribute the prevention systems of DDoS attacks in the cloud environment with a machine learning method. The Multiple Linear Regression (MLR) method was employed. The CICIDS2017 dataset, produced by researchers at UNB (Sharafaldin, Lashkari, et al., 2018) , was utilized in this study. The dataset consists of 225,746 data points, and the experiments were concluded successfully. In the study, the accuracy rate was determined as 97.86%. Estimation variables included destination port, fwd packet length max, bwd packet length mean, avrpcks, afwdss, and sbfwdbyt.

(Virupakshar et al., 2020) developed a system to mitigate DDoS attacks with the OpenStack platform. Different machine learning methods were used in this study. The dataset was created in the cloud environment for classification, training, and testing. Three types of DDoS attack traffic were generated. Four machine learning methods were used: DT, KNN, NB, and DNN. Among these machine learning methods, the recall and f1-score metric results in the cloud dataset were found to be the best with 99% and 98% values, respectively. Prtcl type, Service, Src bytes, Dest bytes, Count, Srv count, Same srv rate were determined as predictor variables.

(Rivas et al., 2019) conducted a study to detect DDoS attack with machine learning methods which are SVM and CNN methods. The dataset was created by the same researchers. Experiments were carried out on datasets obtained from honey pots on the cloud network. It was visualized using a three-axis hive plot graph using source ip address, source country, and attack time information. High success was achieved with both algorithms. A more successful result was achieved with CNN. Accuracy, precision, and recall metric values were determined as 100%, 99.5%, and 100%, respectively, with the CNN method. In the dataset used, 16,000 data were processed. Forecast variables are set as Source IP, country of origin, and time of the attack.

(Alsirhani et al., 2019) conducted a study to develop a new DDoS attack detection algorithm. First, the classification algorithm, then a distributed system, and finally, a fuzzy logic system. A system that can detect different DDoS models is aimed. Many machine learning algorithms have been used to perform these operations. Four different machine learning methods were used: NB, DT (Gini), DT (Entropy), and RF. Some predicted variables are frame.len, ip.src, ip.dest, ip.proto, ip.ttl. The best results were obtained with DT (Gini). Accuracy, precision, recall, and F1-score were

measured at 98%, 98%, 98%, and 97.9%, respectively. Two million data were used in the study. The dataset produced by the University of California in 2015 was used.

(Rukavitsyn et al., 2017) developed a two-step learning method to detect DDos attack in cloud environment. They created the dataset, and their proposed methods were tested by increasing the number of servers in the cloud system. 12, 21, 30, 39, and 48 clients were used for this purpose. While collecting the data, the researchers grouped the attack and normal traffics and tagged these traffics. Subsequently, the groundwork for re-learning studies was established. Accuracy, precision, and recall are all measured at 100%. As the number of clients increases, the success rates decrease. Predicted variables are CachePool, AttackPool, TrafficPool.

(He et al., 2017) conducted a study to develop a machine learning-based method for detecting DDoS attacks on the cloud environment. LR, SVM Linear Kernel, SVM RBF Kernel, SVM Poly Kernel, DT, NB, RF, K-Means, Gaussian Expectation-Maximization (GEM) machine learning methods are proposed. Four different DDoS attacks were tried to be detected: SSH brute-force, DNS reflection, ICMP flooding, and TCP SYN attack. The best result was achieved with the SVM Linear Kernel among the nine different machine learning methods. The attack was carried out in a real cloud environment. Prediction variables were set as IP Address, DNS Record, and Prtcl Information. The metric variables accuracy, precision, recall, and F1-score were 97.77%, 99.68%, 96.18%, 97.9%, respectively.

Table 2.2. Overview of DDoS attack detection studies in Cloud environment – Part 1

Study	**Number of Data**	**Method**	**Predictor Variables**	**Performance Metrics and Results (%)**
(Pandithurai et al., 2024)	12.794.627	LSTM, DNN, ANN, DBN	dst_host_count, srv_diff_host_rate	Accuracy = 97.00, Precision = 95.00, Recall = 90.00, F1-score = 94.00
(Islam et al., 2023)	12.000	SVM, KNN, and LR	std_cpu_util, avg_memory_util, and max_memory_util	Accuracy = 95.00, Precision = 96.00, Recall =93.00, F1-score = 94.00
(Kiruthiga et al., 2022)	10.000	DGRNN, ID3-MMDP, ASLOS	traffic, time response, src-bytes, DST-bytes, WFC.	Accuracy = 90.00
(Gumaste et al., 2021)	n.a.	RF, DT, LR.	Prtcl type, Service, Src bytes, Dest bytes, Count, Srv count, Same srv rate	Accuracy = 99.21, Precision = 99.91,

Abbreviations: DNN: Deep Neural Network, dst_host_count: duration destination bytes, LR: Logistic Regression, SVM: Support Vector Machine, DGRNN: Deep Generative Radial Neural Network, ID3-MMDP: Iterative Dichotomiser 3 Maximum Multifactor Dimensionality Posteriori Method, ASLOS: Ada Shuffled Leaping Optimal Selection, WFC: Wrong fragment count.

Table 2.3. Overview of DDoS attack detection studies in Cloud environment – Part 2

Study	Number of Data	Method	Predictor Variables	Performance Metrics
(Kushwah & Ranga, 2021)	2.491.6894	SaE-ELM-Ca	Flags, fdrtn, FIAT, NBFDS, NBFSD, NPFDS, NPFSD, Prtcl, SN	Accuracy = 99.99, recall = 100.00, precision = 100.00, and F-score = 100.00
(Virupakshar et al., 2020)	n.a.	DNN, DT, NB, KNN	Prtcl type, Service, Src bytes, Dest bytes, Count, Srv count, Same srv rate.	Recall = 99.00, F1-score = 98.00
(Sambangi & Gondi, 2020)	225.746	MLR	DP, FPLM, BPLM, APSA, AFSS, SFB.	Accuracy = 97.86
(Rivas et al., 2019)	16.000	CNN, SVM	Source IP, country of origin, time of attack	Accuracy = 100.00, Precision = 99.50, Recall = 100.00
(Alsirhani et al., 2019)	2.106	DT (Gini), DT (Entropy), RF, NB	Source, length, destination of ip, tcp, udp	Accuracy = 98.00, Precision = 98.00, Recall = 98.00, F1-score = 97.90
(Rukavitsyn et al., 2017)	n.a.	Self-Learning	CachePool, AttackPool, TrafficPool	Accuracy = 100.00, Precision = 100.00, F1-score = 100.00
(He et al., 2017)	n.a.	LR, SVM, DT, NB, RF, K-Means, GEM	IP address, DNS record, prtcl	Accuracy = 97.77, precision = 99.68, Recall = 96.18, F1-score = 97.90

Abbreviations: RF: Random Forest, CNN: Convolutional Neural Networks, DT: Decision Tree, LR: Logistic Regression, NB: Naive Bayes, DP: Destination Port, APSA: Avrpcks, AFSS: Afwdss, SFB: Sbfwdbyt, KNN: K Nearest Neighbor, MLR: Multiple Linear Regression, SVM: Support Vector Machine, FPLM: Fwd Packet Length Max, BPLM: Bwd Packet Length Mean, NBFSD: Number of bytes from source to destination, FIAT: flow inter-arrival time, NPFSD: Number of packets from source to destination, SN: Service name, n.a.: Not Available, GEM: Gaussian Expectation-Maximization.

2.3. DDoS Attack Detection in IoT Environments

(Fenanir & Semchedine, 2023) developed a Federation Learning-Based Intrusion Detection system to protect internet-connected devices from DDoS attacks using CNN, LSTM and DNN deep learning methods. IoT-23, N-BaIoT and IoTID20 datasets were used in the study. There are 325.307.990, 7.062.606, 625.783 samples in the datasets, respectively. In this study, various attack types were used as prediction variables. LSTM was found to be more successful than other deep learning methods. Accuracy value was found to reach 99.00%.

(Alimi et al., 2022) developed a new model to detect DDoS attacks in the IoT environment, using Refined long short-term memory (RLSTM), RF, MLP, DT, K-Nearest Neighbors (k-NN), NB, LSTM, and CNN. The results were also evaluated in a hybrid form of these methods. It was stated that the performance metrics were good with the RLSTM model. NSL-KDD containing 125.973 data and CICIDS 2017 datasets containing 2.830.743 data were used. The result obtained with CICIDS 2017 was found to be more successful. Accuracy, Precision, Recall, F-Score metric results obtained

with RLSTM were 99.22%, 99.23%, 99.22%, 99.22%, respectively. Timestamp, source IP, destination IP, source ports, destination ports, prtcls, and attacks were determined as predictor variables.

(Aydin & Bahtiyar, 2021) conducted a study to detect DDoS attack in IoT environment. They used CNN method. USTC-TFC 2016 dataset was used. Compared with the Mirai-RGU dataset. It has been observed that more successful results have been obtained with USTC-TFC 2016. It was created by Wang et al. between 2011 and 2015. It contains 752,040 pieces of data. It was stated that it performed successfully not only in terms of performance but also in terms of time. Estimation variables are determined as source IP, destination IP, port, and prtcl. Performance metrics Accuracy, Precision, Recall, F1-score were determined as 99.22%, 99.23%, 99.22%, 99.22%, respectively.

(Roopak et al., 2020) conducted a study to develop a new approach to protect against DDoS attacks in the IoT environment with LSTM and CNN deep learning models using the CICIDS2017 dataset in 2020. The CICIDS2017 dataset was produced by researchers in 2017 within UNB (Sharafaldin et al., 2018). The success of the methods proposed by the researchers was demonstrated by comparing different datasets (KDD Cup99, NSL-KDD, UNB ISCX 2012) and machine learning methods (MLP, SVM, NB, RF). Performance metrics Accuracy, precision, recall, F1-score were 99.03%, 99.26%, 99.35%, 99.36%, respectively.

(Ma et al., 2020) conducted a study to prevent DDoS attacks in the IoT environment using an approach with the CNN learning method. The NSL-KDD dataset was used. The dataset used in the study was created by UNB in 2007. In this study, 17,171 pieces of data were used. It has been shown that deep learning algorithms are more successful than many machine learning algorithms. In addition, the results obtained with SVM, KNN, DT, NB, and Recurrent Neural Network (RNN) methods were compared in the study. The accuracy value of the proposed method was determined as 92.99%. Estimation variables were determined as IP, port, network prtcl, transmission flow, and network connection frequency.

(Roopak et al., 2019) conducted a study to protect IoT networks from cyber threats including DDoS attacks using deep learning methods. In this study, using the CICIDS2017 dataset created by UNB in 2017, a comparison was made with CNN – LSTM, MLP, SVM, NB, RF, and DCNN methods. The best accuracy value was obtained with the CNN and LSTM hybrid models. Accuracy, precision, and recall metric results were measured as 97.16%, 98.44%, and 99.1%, respectively.

(Soe et al., 2019) presented a study to detect DDoS attacks using ANN and (Synthetic Minority Over-Sampling Technique) SMOTE technique with the Bot-IoT dataset. There are more than 1.9 million data in the data set. The dataset was developed at UNSW Sydney University in 2018. Success has been achieved in detecting DDoS attacks. The number of benign traffic in the data set used is quite low compared to the total. SMOTE technique was used to overcome this situation. Prediction Variables were determined as saddr, sport, daddr, dport, proto. Performance metrics (accuracy, precision, recall, F1-score) were all measured at 100%.

(Li1 et al., 2019) proposed a new method with the Extreme Learning Method (ELM) to prevent DDoS attacks in the IoT environment. The entropy method was used in the feature extraction stage. The dataset used in the experiment was created by the researchers in 2019. Although the number of datasets used in the experiments varies between 1000-20000, 15,000 data were used in the experiment with ELM. DDoS attacks were carried out over ICMP, UDP, and TCP prtcls. IP, destination IP, source port, destination port, packets, bytes, prtcl, and TCPFlag were specified as predictor variables. The accuracy value was determined as 99.4%.

(Su et al., 2018) developed a new method called Light-Weight Malware Detection to detect DDoS attacks in IoT environments. IoT DDoS dataset with 500 attack content obtained through IoTPOT. The black and white images of the malware were extracted, and the CNN success was revealed by making analyzes through image processing techniques. 94% accuracy result was achieved. Malware images were determined as the prediction variable.

(Doshi et al., 2018) conducted a study to detect DDoS attacks in the IoT environment using the dataset they created with five different machine learning methods, i.e., RF, KNN, LSVM, DT, and Neural Network (NN). In the dataset, a total of 491.855 data were studied. After the results were obtained, the success of machine learning algorithms in detecting DDoS attacks was demonstrated. Accuracy, precision, recall, and F1-score values were determined as 99%. The predictor variables are set to source IP address, source port, and destination IP address.

Table 2.4. DDOS attack detection studies in IoT environment – Part 1

Study	Number of Data	Method	Predictor Variables	Performance Metrics and Results (%)
(Fenanir & Semchedine, 2023)	n.a.	CNN, LSTM, DNN	Attack types	Accuracy = 99.00
(Alimi et al., 2022)	2.830.743	RLSTM, RF, MLP, DT, k-NN, NB, LSTM, CNN	Timestamp, sourceIP, destinationIP, sourceports, destination ports,prtcls,attacks.	Accuracy = 99.22, Precision = 99.23, Recall = 99.22, F1-score = 99.22
(Aydin & Bahtiyar, 2021)	752.040	CNN	source IP, destination IP, port, prtcl	Accuracy = 99.97, precision = 99.95, recall = 99.99, F1-score = 99.97
(Roopak et al., 2020)	225.742	CNN – LSTM, MLP, SVM, NB, RF	DP, FPLM, BPLM, APSA, AFSS, SFB.	Accuracy = 99.03, precision = 99.26, Recall = 99.35, F1-score = 99.36
(Ma et al., 2020)	17.171	CNN, SVM, KNN, DT, NB, RNN	IP, port, network prtcl, transmission flow, network connection frequency	Accuracy = 92.99
(Roopak et al., 2019)	n.a.	CNN – LSTM, MLP, SVM, NB, RF, DCNN	DP, FPLM, BPLM, APSA, AFSS, SFB.	Accuracy = 97.16, Precision = 98.44, Recall = 99.10
(Soe et al., 2019)	1.900.000	ANN	saddr, sport, daddr, dport, proto	Accuracy = 100.00, Precision = 100.00, Recall = 100.00, F1-score = 100.00
(Li1 et al., 2019)	15.000	ELM	IP, destinationIP, sourceport, destination port, packets, bytes, prtcl, and TCPFlag	Accuracy = 99.40
(Su et al., 2018)	n.a.	CNN	malware images	Accuracy = 94.00
(Doshi et al., 2018)	n.a.	RF, KNN, LSVM, DT, NN	Source IP address, source port, destination IP address	Accuracy = 99.90, Precision = 99.90, Recall = 99.90, F1-score = 99.90

Abbreviations: n.a.: Not Available, CNN: Convolutional Neural Networks, LSTM: Long Short-Term Memory, ANN: Artificial Neural Networks, ELM: Extreme Learning Mach, RF: Random Forest, RLSTM: Refined Long Short-Term Memory, DT: Decision Tree, RF: Random Forest, NB: Naive Bayes, MLP: Multi-Layer Perceptron, RNN: Recurrent Neural Network, DP: Destination Port, FPLM: Fwd Packet Length Max, BPLM: Bwd Packet Length Mean, APSA: Average Packet Size, SFB: Subflow Fwd Bytes, AFSS: Average Packet Size

2.4. DDoS Attack Detection in Other Environments

(Said, 2023) conducted a study to detect to DDoS attack. They proposed Quantum vector machine (QSVM), a variant of SVM, to detect DDoS attacks targeting smart microgrids (SMG). With QSVM recall, accuracy and precision are 99.94%, 99.93% and 99.93%, respectively. CICDDoS2019 dataset was used in this study. There are 38 features and 2950 samples in dataset.

(Kumar et al., 2022) proposed a hybrid method for detecting low-level DDoS attacks on 5G networks, which has recently become widespread. The dataset was created by the same researchers in 2021. 794.919 pieces of data were used in the dataset. The accuracy value was measured as 98.397% using the CNN-RF hybrid method. CNN-RF, LSTM-RF, and Light Gradient Boosting Machine (LightGBM) hybrid models were compared. The predictor variables are birthplace IP address, purpose IP address, port location, destiny port, and procedure.

(Cil et al., 2021) introduced a study the DDoS attack detection with a dataset containing a small amount of data. The CICDoS2019 dataset produced by Sharafaldin et al. was used. 395.977 pieces of data were processed. An attempt was made to create a deep learning model with the best level of success. A good service is provided for use in DDoS attack detection systems. Feature extraction and classification processes were performed. The DNN model was used for the classification study in the dataset. Flow ID, SourceIP, SourcePort, DestinationIP, DestinationPort, Prtcl, Timestamp, SimillarHTTP were determined as predictor variables. Performance metrics are Accuracy, precision, recall, F1-score 99.97%, 99.99%, 99.98%, 99.98%, respectively.

(Adhikary et al., 2020) proposed a study to protect the private network system from DDoS attacks for vehicles created by researchers. A hybrid system was created with NN and DT algorithms, and an experiment was carried out on the dataset of 1000 vehicles. According to the results obtained with NN and DT separately on the same dataset, it was seen that the hybrid system was superior. Estimation variables were determined as as collision, delay, jitter, packet drop, ratioinout, and throughput. The accuracy metric value was measured as 96.4%.

(Aytaç, 2020) conducted a study to detect DDoS attacks. The dataset named CICDDoS2019 was used. In the related study, the success of LG, Gaussian Naive Bayes (GNN), KNN, Multinomial Naive Bayes (MNB), Bernoulli Naive Bayes (BNB), DT (entropy), DT (gini), RF, SVM machine learning algorithms in detecting DDoS attacks were tested. In this study, the related dataset was divided into five different datasets as a quartet, hexadecimal, octal, decimal, and twelfth, and the 5th dataset with 12 features, which was the most successful, was determined. Each dataset contains 500,000 pieces of data. The dataset used in the study was created by UNB in 2019. Since the estimation variables of the five different datasets differ, the estimation variable of the most successful group has been added to the table. The estimation variables are set to Fwd_Seg_Size_Min, Bwd_Pkt_Len_Min. Performance metrics (Accuracy, precision, recall, F1-score) were measured as 99%.

(Özekes & KARAKOÇ, 2019) proposed a study to detect DDoS attack. CICIDS2017 dataset was used. The dataset was created by UNB in 2017. 1.042.557 data were used in the dataset. DDoS attack detection study was performed with DT, and RF methods and performance metrics were compared. RF was more successful and performance metric results were measured as 99.966%, 99.948%, 99.968%, and 99.976% for accuracy, precision, recall, and F1-score, respectively.

(Sahingoz et al., 2019) presented a study to improve DDoS attack detection systems using seven machine learning methods. The successes of these algorithms were compared. NSL-KDD dataset, an improved version of KDD CUP99, was used as the dataset. NSL-KDD was developed by UNB University in 2009. The best result was obtained with the Ada method. Ada, KNN, DT, ANN, RF, GB, and Linear Discriminant Analysis (LDA) machine learning methods were used, and metric results were compared. The predictor variables are set to prtcl_type, service, and flag. The accuracy value obtained with Ada was measured as 99.95%.

(Sharafaldin et al., 2019) introduced a study to detect DDoS attacks with machine learning algorithms. The CICDDoS2019 dataset was used in the study. The dataset was created in 2019. There are 50,063,112 data in the CICDDoS2019 dataset. As a result of this study, it has been shown that the ID3 machine learning algorithm is successful in detecting DDoS attacks. Distributions of packet sizes of a prtcl, the number of packets per flow, certain patterns in the payload, the size of the payload, and request time distribution of prtcls were determined as estimation variables. Performance metrics precision, recall, F1-score were measured as 78.0%, 65.0%, and 69.0%, respectively.

(Hasan et al., 2018) conducted a study to show that the Deep Convolution Neural Network (DCNN) method is more successful than traditional machine learning methods in DDoS attack detection. An attempt was made to detect the Burst Header Packet Flood (BHPF) attack. In this study, performance metrics of DCNN and NB, SVM, and KNN methods were compared. 1075 pieces of data were used in the study. The dataset used was published in 2017. As a result of the classification estimation processes, it was determined that the performance metrics obtained with DCNN were better than other algorithms, and the targeted result was achieved. Performance metrics Accuracy, Precision, F1-Score measured 99.0%, 99.0%, and 99.0%, respectively. Variables are Utilized Bandwidth Rate, Used Bandwidth

(Sharafaldin, Lashkari, et al., 2018) proposed a study to prevent DDoS and different attackers with machine learning algorithms. CICIDS-2017 dataset was used in this study. The success of the RF method in DDoS attack detection has been demonstrated. Quadratic Discriminant Analysis (QDA), RF, KNN, ID3, Ada, MLP, and NB methods were used. Performance metrics precision, recall, F1-score were measured as 98.0%, 97.0%, and 97.0%, respectively.

Table 2.5. DDOS attack detection studies in other environments – Part 1

Study	Number of Data	Method	Predictor Variables	Performance Metrics and Results (%)
(Said, 2023)	2950	QSVM	n.a.	Accuracy = 99.9, Precision = 99.93, Recall = 99.94
(Kumar et al., 2022)	794.919	CNN-RF, LSTM-RF, LSTM-LGBM	Birthplace IP address, purpose IP address, port location, destiny port, procedure	Accuracy = 98.40
(Cil et al., 2021)	365.977	DNN	FlowID, SourceIP, SourcePort,DIP, DP, Prtcl, Timestamp, SimillarHTTP	Accuracy = 99.97, Precision = 99.99, Recall = 99.98, F1-score = 99.98
(Adhikary et al., 2020)	1000	NN-DT, NN, DT	Collision, delay, jitter, packetdrop, RT	Accuracy = 96.40
(Aytaç, 2020)	500.000	KNN, LR, GNN, BNB, DT, RF, SVM	FSSM, BPLM	Accuracy = 99.00, Precision = 99.00, Recall = 99.00, F1-score = 99.00
(Kushwah & Ranga, 2021)	2.491.6894	SaE-ELM-Ca	Flags, fdrtn, FIAT, NBFDS, Prtcl, SN	Accuracy = 99.99, Recall = 100, Precision = 100, F-score = 100
(Özekes et al, 2019)	1.042.557	RF, DT	DP, FPLM, BPLM, APSA, AFSS, SFB.	Accuracy = 99.97, Precision = 99.95, Recall = 99.97, F1-score = 99.98
(Sahingoz et al., 2019)	149.470	ADA, KNN, DT, ANN, RF, GB, LDA	Prtcl_type, service, flag	Accuracy = 99.95
(Sharafaldin et al., 2019)	50.063.112	ID3, RF, NB, LR	DPSP, NPPF, CPP, SP, RTDP	Precision = 78.00, Recall = 65.00, F1-score = 69.00
(Hasan et al., 2018)	1075	DCNN	Utilized Bandwidth Rate, Used Bandwidth	Accuracy = 99.00, Precision = 99.00, F1-Score = 99.00
(Sharafaldin, Lashkari, et al., 2018)	2.830.743	RF, KNN, ID3, Ada,MLP, NB, QDA	DPSP, NPPF, CPP, RTDP, SP	Precision = 98.00, Recall = 97.00, F1-score = 97.00

Abbreviations: QSVM: Quantum SVM, LGBM: Light Gradient Boosting Machine, n.a.: Not Available, SaE-ELM-Ca: Self-adaptive evolutionary extreme learning machine, DNN: Deep Neural Network, CNN: Convolutional Neural Networks, NN: Neural Network, RT: Ratioinout and throughput BNB: Binomial Naive Bayes, BPLM: Bwd Packet Length Mean, FSSM: Fwd_Seg_Size_Min, LSTM: Long Short-Term Memory, RF: Random Forest, DT: Decision Tree, KNN: KNearest Neighbors, ADA: Ada Boost, LDA: Linear Discriminant Analysis, DCNN: Deep Convolutional Neural Networks, RTDP: request time distribution of prtcls, CPP: certain patterns in the payload, DP: Destination Port, FPLM: Fwd Packet Length Max, BPLM: Bwd Packet Length Mean, APSA: Avrpcks, FSSM: Fwd_Seg_Size_Min, AFSS: Afwdss, SFB: Sbfwdbyt, SFB: Sbfwdbyt, NPPF: number of packets per flow, APSA: Avrpcks, AFSS: Afwdss, SFB: Sbfwdbyt, SFB: Sbfwdbyt, DPSP: distributions of packet sizes of a prtcl, SP: size of the payload, RTDP: request time distribution of packet, DPSP: distributions of packet sizes of a protocol.

3. OVERVIEW OF METHODS

In this section, the ML classifiers and feature selection algorithms used in the study are examined under two sections. The first part describes the ML methods, while the second part explains the feature selectors. algorithms are introduced.

In this section, four different machine learning methods have been addressed. Definitions of LR, MLP, SVM, and RF machine learning methods were provided, and their mathematical formulas and parameters were discussed.

There are many features in the dataset. Since using all of them would increase the time cost, it needs to be reduced. This reduction process is performed using feature selection algorithms. Thus, the features in the models that achieve the best results are determined. In this study, Relief-F and mRMR feature selectors were used. After defining the feature selectors, their mathematical formulas were expressed.

3.1. Machine Learning Classifiers

LR, MLP, SVM, and RF machine learning methods were utilized in the thesis. This section will define the machine learning algorithms employed in the thesis.

3.1.1. Logistic Regression

The LR machine learning method is a preferred statistical technique in binary and multiclass classification studies. It is highly effective in binary classification. This study is associated with probability concepts, and the Logistic function plays a significant role in the LR machine learning method. (Kolukisa, 2024).

The LR machine learning method is easier to understand compared to other ML methods. It is more effective in terms of time cost because it works faster than other methods.

In classical regression analysis, the response is a continuous random variable. The use of classical regression methods is not appropriate when the response is a discrete variable. In this case, logistic regression is proposed as an alternative to classical regression methods. Logistic regression is a powerful supervised ML algorithm. Logistic regression involves modelling the probabilities of occurrences of the categories of the response given the features. Logistic regression fundamentally uses the logistic function to predict the probability of a response (Tolles and Meurer, 2016; Nick and Campell, 2007).

The logistic regression is model is of the form

$$\log \frac{p(X)}{1-p(X)} = X\beta. \tag{3.1.}$$

where X is the design matrix, $p(X)$ is the probability that an observation is in a certain category of the response variable and β is the parameter vector to be estimated. In Equation (3.1), $\text{logit}(p) = \log\frac{p}{1-p}$ is called the logit transformation. According to Equation (3.1), the class probabilities are given in terms of logistic function $h(x) = \frac{1}{1+e^{-x}}$ as

$$p(y = 1 \mid X = x) = \frac{1}{1 + e^{-X\beta}} \tag{3.2.}$$

The main method for the parameter estimation of the model in Equation (3.2) is the minimizing the maximum likelihood function of the parameter vector β by using the iteratively weighted least squares method (Hosmer et al, 2013). Logistic regression can also be used when the response variable has more than two classes. In this case, the

$$p(y = k \mid X = x) = \frac{e^{X\beta_k}}{\sum_{\ell=1}^{K}(1 + e^{X\beta_\ell})} \tag{3.3.}$$

symmetric log-linear representation is used to estimate the class probabilities, where K represents the number of classes and k represents the class probability (Hastie et al, 2015).

3.1.2. Multi-Layer Perceptron

MLP (Multi-Layer Perceptron), a type of artificial neural network function, consists of three layers: input layer, output layer, and hidden layer. Inspired by the human nervous system, it is composed of nodes forming layers. Each node is referred to as a neuron. All layers are connected through neurons. In the model, each connection has a weight value to adjust the obtained results. These weights are calculated using the backpropagation algorithm (Ramchoun et al., 2016).

Activation functions are the key elements of artificial neural networks. Each neuron in MLP is activated by an activation function. Among the activation functions, ReLU is the most commonly used one because of its simplicity and its consistent performance (Misra, 2019). Also, softmax and sigmoid activation functions are widely used in classification problems. This study uses ReLU activation function for hidden layers and softmax activation function for output layer. In this study, the ReLU activation function defined by

$$f_{ReLU}(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0' \end{cases} \tag{3.4.}$$

is used for hidden layers and the softmax activation function defined by

$$f_{softmax}(x) = \begin{cases} \alpha(e^x - 1), & x < 0 \\ x, & x \geq 0 \end{cases}. \quad (3.5.)$$

is used for output layer.

To get the network output, the output of every unit in each layer has to be computed. Suppose that an input layer with n_0 neurons is ($X=(x_0,x_1,\ldots,x_{n0})$) and a set of hidden layers are $(h_1, h_2, \ldots\ h_N)$. Also, assume that each hidden layer h_i has n_i neurons. Then, the expression for the first hidden layer is

$$h_1^j = f\left(\sum_{k=1}^{n_{i-1}} w_{k,j}^0\, x_k\right) \quad j = 1, \ldots, n_1 \quad (3.6.)$$

and the expression for the i-th hidden layer h_i^j is

$$h_i^j = f\left(\sum_{k=1}^{n_{i-1}} w_{k,j}^{i-1}\, h_{i-1}^k\right) \quad i = 2, \ldots, N \text{ and } j = 1, \ldots, n_i \quad (3.7.)$$

where $w_{k,j}^i$ represents the weight in hidden layer i and hidden layer i+1 between neurons k and j, respectively. Also, the number of the neurons in the i^{th} hidden layer is given by n_i. The notation for the output corresponding to the i^{th} hidden layer is as follows:

$$h_i = {}^t(h_i^1, h_i^2, \ldots, h_i^{n_i}) \quad (3.8.)$$

The network output is calculated by using Equation (3.4.) as follows:

$$y_i = f\left(\sum_{k=1}^{n_N} w_{k,j}^N\ h_N^k\right) \quad Y = \left(y_1, \ldots, y_j, \ldots, y_{N+1}\right) = F(W, X) \quad (3.9.)$$

where $w_{k,j}^N$ is the weight between neuron k and neuron j in the N^{th} hidden layer and output layer, respectively, n_N is the number of neurons in the N^{th} hidden layer, Y is the vector of the hidden layer, F is the transfer function, and W is the matrix of the weights, as shown in (3.6.):

$$W = \left[W^0, \ldots, W^j, \ldots, W^N\right] \quad (3.10.)$$

$$W^i = (w_{j,k}^i)_{\substack{0 \leq i \leq N \\ 1 \leq j \leq n_{i+1} \\ 1 \leq k \leq n_i}} \quad where\ w_{j,k}^i \in \mathbb{R} \quad (3.11.)$$

To make things simpler, let it be said that the hidden layers of the network are defined as $n=n_i \; \forall i=1,\ldots,N$. Then, X stands for the input of the neural network in Equation (3.5), Also, f donates the activation function, and W^i donates the matrix of weights between the i^{th} hidden layer and the $(i+1)^{th}$ hidden layer for i = 1,...,N-1. Then, according to Ramchoun et al. (2016), the weight matrix between the input layer and the first hidden layer is represented by W^0, while the weight matrix between the Nth hidden layer and the output layer is represented by W^N in Equations (3.6) and (3.7).

3.1.3. Support Vector Machine

SVM is a robust classification algorithm with strong regularization properties. SVM simultaneously maximizes the geometric margin and reduces the empirical classification error. SVM is based on the reduction of structural risk. The input vector is mapped by SVM to a higher-dimensional space where a maximal separation hyperplane is constructed. To divide the data, two hyperplanes are constructed on either side of the hyperplane. The distance between two parallel hyperplanes is maximized by the separating hyperplane. It can be assumed that the generalization error of the classifier will be better if the margin or separation between these parallel hyperplanes gets greater. Let the data points be assumed as $\{(x_1, y_1), \ldots \; (x_n, y_n)\}$, where n is the number of samples and y_n= 1/-1 is a constant indicating the class to which x_n belongs. Each x_n represents a real vector in p dimensions. In SVM, scaling is necessary to protect against over-variation. In order to analyze training data, the determination of the hyperplane is needed.$w \cdot x + b = 0$ that seperates the classes. In this equation, b is a scalar, w, a p-dimensional vector, is perpendicular to the separating hyperplane. The offset parameter b can be added to increase the margin. The hyperplane must pass through the origin if b is zero. In order to maximize the maximum margin, parallel hyperplanes are important. The parallel hyperplanes are formulated as:

$$\begin{aligned} w \cdot x + b &= 1 \\ w \cdot x + b &= -1 \end{aligned} \tag{3.12.}$$

When training data can be separated linearly, the hyperplanes can be chosen so that there are no points between them and hence their distance nned to to be maximized. The distance between the hyperplanes can be computed as $2/|w|$. As a result, $|w|$ needs to be minimized. Consequently, to highlight the data points:

$$w \cdot x_i - b \geq 1 \text{ or } w \cdot x_i - b \leq -1 \tag{3.13.}$$

It can also be expressed like this:

$$y_i(w \cdot x_i - b) \geq 1\,, 1 \leq i \leq n \tag{3.14.}$$

Since this study employs non-linear SVM, a kernel function is required. The RBF kernel is used in this study. One way to formulate an RBF kernel is as (Bhavsar & Panchal, 2012)

$$K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right), \gamma > 0 \tag{3.15.}$$

3.1.4. Random Forest

The first step performed when implementing the Random Forest algorithm is the selection of random data. Once the random data is selected, a separate decision tree is created for each data point. Predictions are obtained through these decision trees. The final results to be obtained in the Random Forest algorithm are determined based on the frequency of predictions obtained through decision trees. Random forest makes adjustments to overcome the overfitting problem. The number of decision trees and the parameters for generating random numbers are used to achieve a balanced and effective outcome in the algorithm (Hamid Arif et al., 2024).

Random Forest is a tree-based machine learning method. It was proposed by L. Breiman in 2001. It is a classification and regression method that can be used for large data sets. It can also be used for feature selection. In each partitioning process, a number of features close to the square root of the number of features is selected. The RF algorithm was found to be the most successful in the data sets obtained with both feature selections.

3.2. Feature Selection Algorithms

In this section, the Relief-F and mRMR feature selectors used in the thesis have been examined. Relief-F and mRMR feature selectors play an important role in improving the performance of machine learning methods. They significantly benefit by reducing the size of the database, thus reducing the time cost.

3.2.1. Relief-F

The Relief-F algorithm was developed by Kira and Rendell for binary classification studies (Kira & Rendell, 1992). When initially proposed by Kira and Rendell, it was limited compared to its current usage. It lacked a solution for handling missing data. Relief-F randomly divides the problem into two classes (Stief et al., 2018) The targeted feature is scored positively if the condition is met and negatively if not. Based on this scoring, the success of features is determined and ranked.

Figure 3.1 illustrates the step-by-step algorithm of the Relief-F feature selection algorithm. The first two lines provide definitions. The third line represents Ri, a randomly selected example. The algorithm searches for its nearest neighbors. If correct, it increases the score positively and

denotes it as H. If incorrect, it decreases the score and denotes it as M. The predicted value W[A] is altered based on the value of variable A. The update processes are expressed in lines 7 and 8 (Robnik-Sikonja & Kononenko, 2003).

Input: a vector containing the class value and attribute values for each training instance

Output: The vector W representing the qualities of attribute estimations.

1. Set all weights W[A] := 0.0;
2. **for i** :=1 **to** m **do begin**
3. randomly select an instance R_i;
4. find k nearest hits H_j;
5. **for** each class C ≠ class(R_i) **do**
6. from class C find k nearest misses M_j(C);
7. **for** A := 1 **to** a **do**
8. $W[A] := W[A] - \sum_{j=1}^{k} diff(A, R_i, H_j)/m.k +$
9. $\sum_{C \neq class(R_i)} [\frac{P(C)}{1-P(class(R_i))} \sum_{j=1}^{k} diff(A, R_i, M_j(C))]/(m.k.)$;

end;

Figure 3.1. Relief-F algorithm's pseudo code. (Robnik-Sikonja & Kononenko, 2003)

3.2.2. Minimum Redundancy Maximum Relevance

In the mRMR feature selection method, the first of the two main objectives is to avoid using prediction variables that are highly similar to each other. The aim is to preserve data that is moderately dissimilar to each other to prevent overfitting (Acid et al., 2011). The degree of dissimilarity between variables is formulated in the formula (3.16).

$$W_I(S) = \frac{1}{|S|^2} \sum_{X_i, X_j \in S} MI(X_i, X_j) \tag{3.16.}$$

where $MI(X_i, X_j)$ represents the measure of mutual information between variables *Xi* and *Xj*. The method's other aim is to minimize redundancy. It achieves this by relying on the principle of maximizing the relevance of predictor variables to the class variable. This aim is explained in the formula (3.17).

The most straightforward approach to amalgamating relevance and redundancy for acquiring a satisfactory subset of features is outlined below:

$$S^* argmax_{S \subseteq U}(V_I(S) - W_I(S)) \tag{3.18.}$$

Looking for the best subset through a comprehensive search is meaningless. Sm-1, the feature that the selected subset maxima, is added to the current subset progressively. This process starts with the feature that has a maximum value of $MI(C; X_i)$ $(S_0 = \{X_{i_0}\})$ (Acid et al., 2011).

$$\left(\underset{X_j \in U \setminus S_{m-1}}{max} \left(MI(C, X_j) - \frac{1}{m-1} \sum_{X_i \in S_{m-1}} MI(X_j, X_i) \right) \right) \tag{3.19.}$$

4. DEVELOPMENT OF PREDICTION MODELS

This chapter includes two sections. In the first section, the methods of Relief-F and mRMR techniques, along with the analyses of machine learning methods based on obtained metrics, have been explained. The second section provides details about the machine learning methods used in the study.

4.1. Methodology

This section was divided into two subsections. In the first subsection, the list was presented in the form of a prioritized ranking of prediction variables obtained through feature selection. The second subsection provided information about metric variables obtained through machine learning methods.

4.1.1. Model Creation

Datasets used in DDoS attack detection typically contain a large amount of data. Packet features contain a lot of information related to packets, as they are important in attack detection. The abundance of row count and packet features increases the time cost. It is necessary to sort the features of the data set according to their importance. Starting from the least important features and removing them one by one, the most successful model should be determined. In this study, the features of the dataset used are sorted according to their importance using Relief-F and mRMR feature selection methods.

The rank of each variable has been determined and sorted in decreasing order by their rankings using the mRMR and Relief-F feature selectors, as shown in Table 4.1 and Table 4.2. The variables with the lowest rank were then deleted one by one. As a result, 65 mRMR-based and 65 Relief-F-based prediction models were developed, as shown in the Appendix (in Table A.1. through Table A.2).

The total number of models obtained using Relief-F and mRMR feature selectors is 130. Metric values were measured using LR, MLP, SVM, and RF machine learning algorithms. Accuracy, precision, recall, and F1 score metric values were examined separately in tables. The models yielding the best results were identified and evaluated.

Table 4.1. List of ranks of predictor variables assigned by Relief-F and mRMR – Part 1

Rank	Predictor Variables Ranked by Relief-F	Predictor Variables Ranked by mRMR
1	ackfc	fplmean
2	prtcl	cwefc
3	fplmin	actmin
4	minpl	ackfc
5	fbytes	bimin
6	dport	mssf
7	fplmax	urgfc
8	maxpl	fbytes
9	urgfc	minpl
10	fpckts	iwbf
11	iwbf	plmean
12	fwdpckts	fplmin
13	cwefc	sbbwdpckt
14	fpflg	maxpl
15	rdtfc	prtcl
16	fdrtn	avrpcks
17	fwditotal	fimin
18	duratio	afwdss
19	fimax	fwdistd
20	fwdimean	fplmax
21	fwdimax	iwbb
22	Bplmax	fwdhl.1
23	imax	sbfwdbyt
24	fwdistd	bimean
25	plstd	fpflg
26	sbfwdbyt	bwdpckts
27	imean	adpf
28	imin	tbpckts
29	plmean	fwdpckts
30	iwbb	rdtfc
31	fistd	tlfpckts
32	bplstd	fwdhl
33	tlfpckts	bimax
34	bitotal	plstd
35	mplmin	fdrtn
36	mssf	bitotal
37	bimin	bplmean
38	sbfwdpckt	plvrnc
39	tfpckts	fwdimin
40	adpf	dport
41	bimax	istd
42	fwdhl.1	fistd
43	plvrnc	abwdss
44	fwdhl	fpckts
45	fplstd	bplmax
46	istd	fimax
47	bistd	tlbpckts
48	tbpckts	fplstd
49	sbbwdpckt	fimean
50	bwdpckts	sbbwdbyt
51	bwdhl	imean
52	fplmean	mplmin

Table 4.2. List of ranks of predictor variables assigned by Relief-F and mRMR – Part 1

Rank	Predictor Variables Ranked by Relief-F	Predictor Variables Ranked by mRMR
53	avrpcks	bplstd
54	fimin	imax
55	fwdimin	bwdhl
56	afwdss	sbfwdpckt
57	actmax	actmax
58	actmin	tfpckts
59	actmean	actmean
60	sbbwdbyt	fwdimax
61	tlbpckts	bistd
62	fimean	duratio
63	bimean	fwdimean
64	abwdss	actstd
65	bplmean	fwditotal
66	actstd	imin

4.1.2. Evaluation Metrics

Prediction errors have been calculated using accuracy, precision, recall, and F1-score metrics, whose equations are given in (4.1.), (4.2.), (4.3), and (4.4), respectively (Shurman et al., 2020).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{4.1.}$$

$$Precision = \frac{TP}{TP+FP} \tag{4.2.}$$

$$Recall = \frac{TP}{TP+FN} \tag{4.3.}$$

$$F1 - \text{Score} = 2 \times \frac{Precision \times Recall}{Precision+Recall} \tag{4.4.}$$

True Positive (TP): The attack record was correctly classified as an attack.

True Negative (TN): The benign record is correctly classified as benign.

False Positive (FP): Benign record misclassified as an attack.

False Negative (FN): Attack record incorrectly classified as benign.

Cross-validation serves as an alternative to hold-out validation and is implemented to achieve higher accuracy compared to hold-out estimates while maintaining the dataset's training examples intact. Among the various forms of cross-validation, the most widely recognized is k-fold cross-validation. In this method, the dataset is divided into k equally sized folds. Each fold functions akin to a hold-out set, where the model is trained on the remaining k-1 folds and evaluated on the held-

out fold to assess its quality and performance. This process iterates for each fold, and the overall performance is computed as the average across all k-folds. Within the realms of data mining and machine learning, tenfold cross-validation (k=10) emerges as the prevailing approach. (Abut & Akay, 2015).

4.2. Details of Prediction Models

This section explains the preferred parameter values of five different ML classifiers.

4.2.1. LR-Based Prediction Models

Since the LR does not have any hyper parameter, the parameter and range table for the LR is not reported. The LR function is based on mathematical functions. The foundation of LR machine learning is formed by the logistic function. It can be used in multiple classification tasks. It is particularly effective in problems requiring binary classification. The Logistic Function equation (4.5.) is shown.

$$h(x) = \frac{1}{1 + e^{-x}} \tag{4.5.}$$

4.2.2. MLP-Based Prediction Models

Optimizing the MLP-based models required six different parameters, which are given in Table 4.3. These parameters are the number of the hidden layers, hidden layer activation functions, output layer activation functions, number of neurons in the hidden layer, epochs, and batch size.

The hidden layer is the layer located between the input and output layers in the MLP algorithm. The functions responsible for processing the input data from the input layer and producing results from the output layer are located in this layer. It plays an important role in solving complex problems. There are many nodes in this layer, and they are all interconnected.

Table 4.3. The intervals for values of the utilized parameters for MLP-based DDoS attack prediction models

Parameter	Range
Number of the hidden layers	2
Hidden layer activation function	ReLu
Output layer activation function	Softmax
Number of neurons in the hidden layer	30
epochs	30
Batch size	30

The activation functions generate non-linear outputs for the neurons by computing the weighted sum and applying bias to it. Two hidden layers are used for the network in this study. The ReLu activation function given by Equation (3.4.) is chosen for hidden layers while the softmax

activation function given by Equation (3.5.) is chosen for the output layer. The parameters are explained in section 3.1.2.

Since neurons employ activation functions to generate output, the number of neurons in the hidden layers is critical for the model (Sherrod, 2016). The number of neurons in each hidden layer in the network is determined as 30. Also, epoch and batch size are set to 30.

4.2.3. SVM-Based Prediction Models

Optimizing the SVM-based models required two parameters, which are given in Table 4.4. These parameters are cost and gamma.

Table 4.4. The intervals for values of the utilized parameters for SVM-based DDoS attack prediction models

Parameter	Range
Cost (C)	[0.1-100]
Gamma (γ)	[0.5-3]

In SVM, the cost and gamma parameters are calculated by grid search. In this study, the RBF kernel is used for SVM. In SVM, the cost parameter is a regularization parameter and is used to optimize the margin. It specifies how soft or hard the margin should be. High C values can lead to overfitting and low C values can lead to underfitting. Therefore, an attempt was made to prefer the C value at an ideal level. When the experiments conducted were examined, an increase in the success value was observed as the number of features in the models decreased. It is understood from this examination that overfitting was not observed.

Furthermore, the gamma parameter is used to optimize non-linear SVM models. If the gamma parameter is higher or lower than it should be, it negatively affects the performance of the models (Sherrod, 2016). The gamma parameter is effective in modeling non-linear data. As the gamma value decreases, it approaches underfitting. As the gamma value increases, it approaches overfitting. Therefore, the gamma parameter should neither be too high nor too low. Based on the results obtained from the models, it is understood that the preferred gamma value falls within a good range.

4.2.4. RF-Based Prediction Models

In this thesis, the two parameters of the RF are given in Table 4.5. If the number of trees and the number of candidate samples at each split in RF are overly high, the model can become complex, and overfitting can occur. In this case, the model may lose its ability to learn and generalize from the noise in the training data and adapt worse to new data. Hence, the parameters are chosen in line with the literature (Breiman, 2001; James et al., 2023).

Table 4.5. The intervals for values of the utilized parameters for RF-based DDoS attack prediction models

Parameter	Range
Number of trees to grow.	500
Number of variables randomly sampled as candidates at each split.	9

5. RESULTS AND DISCUSSIONS

In this section, the accuracy, precision, recall, F1 score metric values obtained with ML prediction models applied to the models obtained using Relief-F and mRMR feature selectors are listed. The metric values are discussed separately and the success of LR, MLP, SVM and RF machine learning methods are compared.

5.1. Results

5.1.1. Relief-F Results

Table 5.1 through Table 5.8 show the accuracy, precision, recall, and F1-score metric values for Relief-F-based DDoS attack prediction models.

Figure 5.1. through Figure 5.4. illustrates the percentage increase rates in accuracy, precision, recall, and F1-score values for DDoS attack prediction models.

Table 5.1. Accuracy values for Relief-F-based DDoS attack detection prediction models using different four machine learning methods – Part 1

Models	Accuracy (%)			
	LR	MLP	SVM	RF
Model 1	71.13	98.37	97.77	99.24
Model 2	72.67	98.40	97.77	99.25
Model 3	71.20	98.40	97.77	99.25
Model 4	78.30	98.33	97.77	99.25
Model 5	76.00	98.43	97.77	99.24
Model 6	70.83	98.20	97.77	99.25
Model 7	70.13	98.40	97.77	99.25
Model 8	69.03	98.53	97.77	99.26
Model 9	70.27	98.47	97.77	99.25
Model 10	69.93	98.37	97.77	99.25
Model 11	70.87	98.33	97.90	99.22
Model 12	71.53	98.40	97.90	99.26
Model 13	71.87	98.23	97.90	99.26
Model 14	70.27	98.23	97.90	99.25
Model 15	72.97	98.27	97.90	99.26
Model 16	70.90	98.17	97.90	99.26
Model 17	70.10	98.23	97.90	99.28
Model 18	78.83	98.37	97.90	99.26
Model 19	78.73	98.50	97.90	99.28
Model 20	78.73	98.43	97.90	99.22
Model 21	79.40	98.47	97.90	99.25
Model 22	78.53	98.37	97.80	99.26
Model 23	79.57	98.50	97.80	99.28
Model 24	79.20	98.37	97.80	99.24
Model 25	77.93	98.47	97.80	99.22
Model 26	77.93	98.47	97.80	99.22
Model 27	86.47	98.37	97.90	99.19
Model 28	86.60	98.33	97.97	99.21
Model 29	84.93	98.40	97.97	99.22
Model 30	86.23	98.33	98.00	99.21

Table 5.2. Accuracy values for Relief-F-based DDoS attack detection prediction models using different four machine learning methods – Part 2

Models	Accuracy (%)			
	LR	**MLP**	**SVM**	**RF**
Model 31	86.50	98.40	98.10	99.22
Model 32	86.57	98.30	98.30	99.16
Model 33	88.03	98.33	98.30	99.15
Model 34	87.47	98.33	98.40	99.19
Model 35	87.13	98.47	98.37	99.15
Model 36	84.80	98.33	98.37	99.18
Model 37	75.80	98.40	98.37	99.18
Model 38	80.87	98.40	98.37	99.19
Model 39	75.03	98.30	98.27	99.21
Model 40	81.17	98.30	98.30	99.18
Model 41	85.70	98.33	98.30	99.16
Model 42	86.23	98.43	98.37	99.19
Model 43	88.40	98.37	98.37	99.13
Model 44	79.27	98.43	98.37	99.13
Model 45	87.20	98.40	98.37	99.16
Model 46	85.77	98.37	98.37	99.22
Model 47	89.23	98.27	98.37	99.18
Model 48	86.10	98.50	98.37	99.16
Model 49	85.77	98.40	98.40	99.21
Model 50	93.03	98.47	98.43	99.18
Model 51	90.70	98.50	98.43	99.18
Model 52	90.97	98.30	98.50	99.16
Model 53	88.13	98.37	98.50	99.19
Model 54	88.40	98.27	98.50	99.19
Model 55	89.40	98.40	98.50	99.19
Model 56	88.80	98.47	98.53	99.22
Model 57	89.90	98.30	98.63	99.12
Model 58	93.70	98.47	98.53	99.07
Model 59	97.03	98.47	98.40	98.97
Model 60	96.80	98.33	98.40	99.01
Model 61	89.80	98.37	98.33	99.03
Model 62	96.97	98.57	96.43	98.22
Model 63	95.27	98.40	96.33	96.09
Model 64	95.27	98.33	96.00	95.51
Model 65	49.90	98.53	49.83	50.40

Table 5.3. Precision values for Relief-F-based DDoS attack detection prediction models using different four machine learning methods – Part 1

Models	Precision (%)			
	LR	MLP	SVM	RF
Model 1	71.26	98.37	97.76	99.23
Model 2	72.80	98.40	97.76	99.24
Model 3	71.32	98.40	97.76	99.24
Model 4	78.45	98.34	97.76	99.24
Model 5	76.15	98.44	97.76	99.23
Model 6	70.96	98.20	97.76	99.24
Model 7	70.25	98.40	97.76	99.24
Model 8	69.14	98.54	97.76	99.26
Model 9	70.38	98.47	97.76	99.24
Model 10	70.04	98.37	97.76	99.24
Model 11	70.99	98.34	97.90	99.21
Model 12	71.66	98.40	97.90	99.26
Model 13	72.00	98.23	97.90	99.26
Model 14	70.39	98.24	97.90	99.24
Model 15	73.11	98.27	97.90	99.26
Model 16	71.03	98.17	97.90	99.26
Model 17	70.22	98.24	97.90	99.27
Model 18	78.98	98.37	97.90	99.26
Model 19	78.88	98.50	97.90	99.27
Model 20	78.88	98.44	97.90	99.21
Model 21	79.55	98.47	97.90	99.24
Model 22	78.66	98.37	97.80	99.26
Model 23	79.72	98.50	97.80	99.27
Model 24	79.34	98.37	97.80	99.23
Model 25	78.05	98.47	97.80	99.21
Model 26	78.08	98.47	97.80	99.21
Model 27	86.47	98.37	97.90	99.18
Model 28	86.61	98.34	97.97	99.20
Model 29	84.92	98.40	97.97	99.21
Model 30	86.23	98.34	98.00	99.20
Model 31	86.48	98.40	98.10	99.21
Model 32	86.57	98.30	98.30	99.15
Model 33	88.06	98.34	98.30	99.14
Model 34	87.47	98.34	98.40	99.18
Model 35	87.13	98.47	98.37	99.14
Model 36	84.82	98.34	98.37	99.17
Model 37	75.94	98.40	98.37	99.17
Model 38	80.84	98.40	98.37	99.18
Model 39	75.13	98.30	98.27	99.20
Model 40	81.18	98.30	98.30	99.17

Table 5.4. Precision values for Relief-F-based DDoS attack detection prediction models using different four machine learning methods – Part 2

Models	Precision (%)			
	LR	MLP	SVM	RF
Model 41	85.71	98.34	98.30	99.15
Model 42	86.24	98.44	98.37	99.18
Model 43	88.39	98.37	98.37	99.12
Model 44	79.23	98.44	98.37	99.12
Model 45	87.24	98.40	98.37	99.15
Model 46	85.76	98.37	98.37	99.21
Model 47	89.22	98.27	98.37	99.17
Model 48	86.04	98.50	98.37	99.15
Model 49	85.70	98.40	98.40	99.20
Model 50	93.02	98.47	98.44	99.17
Model 51	90.70	98.50	98.44	99.17
Model 52	90.95	98.30	98.50	99.15
Model 53	88.11	98.37	98.50	99.18
Model 54	88.40	98.27	98.50	99.18
Model 55	89.36	98.40	98.50	99.18
Model 56	88.76	98.47	98.54	99.21
Model 57	89.87	98.30	98.64	99.11
Model 58	93.70	98.47	98.54	99.07
Model 59	97.04	98.47	98.40	98.97
Model 60	96.81	98.34	98.40	99.01
Model 61	89.76	98.37	98.33	99.02
Model 62	96.97	98.57	96.43	98.24
Model 63	95.27	98.40	96.33	96.16
Model 64	95.27	98.34	96.00	95.57
Model 65	50.35	98.54	50.28	50.37

Table 5.5. Recall values for Relief-F-based DDoS attack detection prediction models using different four machine learning methods – Part 1

Models	Recall (%)			
	LR	MLP	SVM	RF
Model 1	75.34	98.38	97.79	99.23
Model 2	75.22	98.42	97.79	99.25
Model 3	74.57	98.41	97.79	99.25
Model 4	81.01	98.35	97.79	99.25
Model 5	79.57	98.45	97.79	99.23
Model 6	71.73	98.22	97.79	99.25
Model 7	72.06	98.41	97.79	99.25
Model 8	69.69	98.54	97.79	99.26
Model 9	72.33	98.47	97.79	99.25
Model 10	72.07	98.38	97.79	99.25
Model 11	77.16	98.34	97.91	99.22
Model 12	76.77	98.41	97.91	99.26
Model 13	78.19	98.24	97.91	99.26
Model 14	75.57	98.26	97.91	99.25
Model 15	76.07	98.28	97.91	99.26
Model 16	77.67	98.19	97.91	99.26
Model 17	75.97	98.25	97.91	99.28
Model 18	82.71	98.39	97.91	99.26
Model 19	82.63	98.52	97.91	99.28
Model 20	82.62	98.45	97.91	99.22
Model 21	84.32	98.47	97.91	99.25
Model 22	82.52	98.38	97.82	99.26
Model 23	84.37	98.51	97.82	99.28
Model 24	81.86	98.38	97.82	99.23
Model 25	80.29	98.48	97.82	99.22
Model 26	80.57	98.47	97.82	99.22
Model 27	87.42	98.38	97.92	99.19
Model 28	87.85	98.35	97.98	99.20
Model 29	86.37	98.42	97.98	99.22
Model 30	88.09	98.35	98.02	99.20
Model 31	87.31	98.40	98.11	99.22
Model 32	87.68	98.32	98.30	99.16
Model 33	89.40	98.34	98.30	99.14
Model 34	89.56	98.35	98.40	99.19
Model 35	89.21	98.47	98.37	99.14
Model 36	87.10	98.35	98.37	99.17
Model 37	80.60	98.41	98.37	99.17
Model 38	81.93	98.40	98.37	99.19
Model 39	79.74	98.32	98.27	99.20
Model 40	84.73	98.32	98.30	99.17
Model 41	87.03	98.35	98.30	99.16
Model 42	87.67	98.44	98.37	99.19

Table 5.6. Recall values for Relief-F-based DDoS attack detection prediction models using different four machine learning methods – Part 2

Models	Recall (%)			
	LR	MLP	SVM	RF
Model 43	89.01	98.37	98.37	99.13
Model 44	79.97	98.46	98.37	99.13
Model 45	88.47	98.42	98.37	99.16
Model 46	87.01	98.38	98.37	99.22
Model 47	90.18	98.27	98.37	99.17
Model 48	86.71	98.51	98.37	99.16
Model 49	86.00	98.41	98.41	99.20
Model 50	93.05	98.47	98.44	99.17
Model 51	91.25	98.51	98.44	99.17
Model 52	91.33	98.32	98.51	99.16
Model 53	88.70	98.38	98.51	99.19
Model 54	88.74	98.28	98.51	99.19
Model 55	90.28	98.41	98.51	99.19
Model 56	89.11	98.48	98.54	99.22
Model 57	90.27	98.31	98.64	99.12
Model 58	94.30	98.48	98.53	99.07
Model 59	97.15	98.48	98.40	98.97
Model 60	96.93	98.35	98.40	99.02
Model 61	89.94	98.38	98.33	99.03
Model 62	97.07	98.58	96.70	98.28
Model 63	95.58	98.40	96.58	96.49
Model 64	95.58	98.35	96.25	95.94
Model 65	95.57	98.54	96.25	95.94

Table 5.7. F1-Score values for Relief-F-based DDoS attack detection prediction models using different four machine learning methods - Part 1

Models	F1-Score (%)			
	LR	MLP	SVM	RF
Model 1	73.25	98.37	97.77	99.23
Model 2	73.99	98.41	97.77	99.24
Model 3	72.91	98.41	97.77	99.24
Model 4	79.71	98.34	97.77	99.24
Model 5	77.82	98.44	97.77	99.23
Model 6	71.34	98.21	97.77	99.24
Model 7	71.14	98.41	97.77	99.24
Model 8	69.41	98.54	97.77	99.26
Model 9	71.34	98.47	97.77	99.24
Model 10	71.04	98.38	97.77	99.24
Model 11	73.95	98.34	97.91	99.21
Model 12	74.12	98.41	97.91	99.26
Model 13	74.97	98.24	97.91	99.26
Model 14	72.89	98.25	97.91	99.24
Model 15	74.56	98.28	97.91	99.26
Model 16	74.20	98.18	97.91	99.26
Model 17	72.98	98.25	97.91	99.28
Model 18	80.80	98.38	97.91	99.26
Model 19	80.71	98.51	97.91	99.28
Model 20	80.71	98.44	97.91	99.22
Model 21	81.86	98.47	97.91	99.24
Model 22	80.54	98.37	97.81	99.26

Table 5.8. F1-Score values for Relief-F-based DDoS attack detection prediction models using different four machine learning methods - Part 2

Models	F1-Score (%)			
	LR	MLP	SVM	RF
Model 23	81.98	98.51	97.81	99.28
Model 24	80.58	98.38	97.81	99.23
Model 25	79.15	98.48	97.81	99.22
Model 26	79.30	98.47	97.81	99.22
Model 27	86.94	98.38	97.91	99.19
Model 28	87.23	98.35	97.98	99.20
Model 29	85.63	98.41	97.98	99.22
Model 30	87.15	98.34	98.01	99.20
Model 31	86.90	98.40	98.11	99.21
Model 32	87.12	98.31	98.30	99.16
Model 33	88.73	98.34	98.30	99.14
Model 34	88.51	98.34	98.40	99.19
Model 35	88.16	98.47	98.37	99.14
Model 36	85.95	98.34	98.37	99.17
Model 37	78.20	98.40	98.37	99.17
Model 38	81.38	98.40	98.37	99.18
Model 39	77.37	98.31	98.27	99.20
Model 40	82.92	98.31	98.30	99.17
Model 41	86.37	98.34	98.30	99.15
Model 42	86.95	98.44	98.37	99.19
Model 43	88.70	98.37	98.37	99.13
Model 44	79.60	98.45	98.37	99.13
Model 45	87.85	98.41	98.37	99.16
Model 46	86.38	98.38	98.37	99.22
Model 47	89.70	98.27	98.37	99.17
Model 48	86.38	98.51	98.37	99.15
Model 49	85.85	98.41	98.41	99.20
Model 50	93.03	98.47	98.44	99.17
Model 51	90.97	98.51	98.44	99.17
Model 52	91.14	98.31	98.51	99.16
Model 53	88.40	98.38	98.51	99.19
Model 54	88.57	98.28	98.51	99.19
Model 55	89.82	98.41	98.51	99.19
Model 56	88.93	98.47	98.54	99.21
Model 57	90.07	98.30	98.64	99.11
Model 58	94.00	98.48	98.53	99.07
Model 59	97.09	98.48	98.40	98.97
Model 60	96.87	98.34	98.40	99.01
Model 61	89.85	98.37	98.33	99.03
Model 62	97.02	98.57	96.57	98.26
Model 63	95.42	98.40	96.45	96.32
Model 64	95.42	98.34	96.12	95.76
Model 65	95.41	98.54	96.11	95.75

5.1.2. mRMR Results

Table 5.9 through Table 5.16 show the accuracy, precision, recall, and F1-score values for mRMR-based DDoS attack prediction models. Figure 5.2 illustrates the percentage increase rates in

accuracy, precision, recall, and F1-score values for DDoS attack prediction models with mRMR method.

Table 5.9. Accuracy values for mRMR-based DDoS attack detection prediction models using different four machine learning methods – Part 1

Models	Accuracy (%)			
	LR	MLP	SVM	RF
Model 66	73.23	98.30	97.77	99.28
Model 67	70.73	98.53	97.77	99.24
Model 68	70.60	98.27	97.77	99.24
Model 69	70.80	98.47	97.77	99.24
Model 70	71.00	98.43	97.80	99.25
Model 71	71.93	98.43	97.80	99.26
Model 72	71.90	98.27	97.80	99.25
Model 73	70.70	98.40	97.80	99.25
Model 74	73.87	98.40	97.80	99.26
Model 75	73.37	98.33	97.67	99.24
Model 76	72.00	98.30	97.93	99.25
Model 77	73.27	98.37	97.77	99.26
Model 78	71.47	98.47	97.77	99.25
Model 79	71.40	98.47	97.80	99.21
Model 80	71.40	98.37	97.80	99.25
Model 81	71.60	98.53	97.83	99.21
Model 82	71.43	98.47	97.80	99.24
Model 83	70.33	98.40	97.80	99.26
Model 84	71.57	98.37	97.80	99.25
Model 85	70.07	98.43	97.83	99.24
Model 86	70.83	98.23	97.83	99.25
Model 87	71.07	98.27	98.00	99.24
Model 88	72.77	98.40	98.00	99.25
Model 89	70.20	98.53	98.03	99.25
Model 90	73.97	98.33	98.03	99.25
Model 91	74.87	98.27	98.03	99.25
Model 92	77.37	98.37	98.20	99.25
Model 93	82.90	98.30	98.47	99.24
Model 94	78.80	98.30	98.47	99.24
Model 95	74.20	98.37	98.47	99.25
Model 96	74.07	98.33	98.47	99.28
Model 97	78.23	98.37	98.47	99.24
Model 98	86.13	98.47	98.50	99.24
Model 99	82.27	98.40	98.50	99.29
Model 100	81.43	98.37	98.50	99.25
Model 101	81.53	98.27	98.50	99.25
Model 102	83.70	98.23	98.50	99.25
Model 103	80.50	98.37	98.50	99.25
Model 104	74.47	98.37	98.50	99.19
Model 105	75.30	98.50	98.50	99.26
Model 106	76.30	98.33	98.50	99.24
Model 107	73.03	98.37	98.50	99.25
Model 108	73.43	98.37	98.50	99.25
Model 109	81.43	98.27	98.77	99.26
Model 110	72.27	98.37	98.60	99.22

Table 5.10. Accuracy values for mRMR-based DDoS attack detection prediction models using different four machine learning methods – Part 2

Models	Accuracy (%)			
	LR	MLP	SVM	RF
Model 111	79.33	98.47	98.60	99.25
Model 112	72.17	98.27	98.60	99.24
Model 113	75.60	98.33	98.53	99.26
Model 114	73.53	98.33	98.97	99.22
Model 115	73.53	98.43	98.97	99.24
Model 116	81.87	98.37	98.93	99.21
Model 117	83.97	98.43	98.90	99.22
Model 118	53.87	98.33	98.90	99.26
Model 119	63.37	98.37	98.87	99.21
Model 120	55.40	98.50	98.87	99.26
Model 121	60.43	98.23	98.80	99.22
Model 122	54.17	98.20	98.50	99.15
Model 123	59.50	98.60	98.10	99.00
Model 124	51.50	98.50	97.77	98.76
Model 125	90.17	98.43	97.30	98.78
Model 126	94.03	98.43	97.23	98.81
Model 127	94.07	98.40	97.27	98.37
Model 128	84.50	98.40	97.27	98.15
Model 129	82.77	98.40	93.83	95.51
Model 130	82.13	98.33	93.10	95.79

Table 5.11. Precision values for mRMR-based DDoS attack detection prediction models using different four machine learning methods – Part 1

Models	Precision (%)			
	LR	MLP	SVM	RF
Model 66	73.38	98.30	97.76	99.27
Model 67	70.86	98.54	97.76	99.23
Model 68	70.72	98.27	97.76	99.23
Model 69	70.91	98.47	97.76	99.23
Model 70	71.12	98.44	97.80	99.24
Model 71	72.06	98.44	97.80	99.26
Model 72	72.03	98.27	97.80	99.24
Model 73	70.82	98.40	97.80	99.24
Model 74	74.01	98.40	97.80	99.26
Model 75	73.51	98.34	97.66	99.23
Model 76	72.14	98.30	97.93	99.24
Model 77	73.41	98.37	97.76	99.26
Model 78	71.60	98.47	97.76	99.24
Model 79	71.53	98.47	97.80	99.20
Model 80	71.52	98.37	97.80	99.24
Model 81	71.73	98.54	97.83	99.20
Model 82	71.57	98.47	97.80	99.23
Model 83	70.46	98.40	97.80	99.26
Model 84	71.71	98.37	97.80	99.24
Model 85	70.18	98.44	97.83	99.23
Model 86	70.94	98.24	97.83	99.24
Model 87	71.21	98.27	98.00	99.23
Model 88	72.91	98.40	98.00	99.24
Model 89	70.32	98.54	98.03	99.24
Model 90	74.09	98.34	98.03	99.24

Table 5.12. Precision values for mRMR-based DDoS attack detection prediction models using different four machine learning methods – Part 2

Models	Precision (%)			
	LR	MLP	SVM	RF
Model 91	75.02	98.27	98.03	99.24
Model 92	77.49	98.37	98.20	99.24
Model 93	83.07	98.30	98.47	99.23
Model 94	78.94	98.30	98.47	99.23
Model 95	74.33	98.37	98.47	99.24
Model 96	74.18	98.34	98.47	99.27
Model 97	78.34	98.37	98.47	99.23
Model 98	86.09	98.47	98.50	99.23
Model 99	82.23	98.40	98.50	99.29
Model 100	81.41	98.37	98.50	99.24
Model 101	81.51	98.27	98.50	99.24
Model 102	83.69	98.23	98.50	99.24
Model 103	80.45	98.37	98.50	99.24
Model 104	74.62	98.37	98.50	99.18
Model 105	75.46	98.50	98.50	99.26
Model 106	76.45	98.34	98.50	99.23
Model 107	73.20	98.37	98.50	99.24
Model 108	73.59	98.37	98.50	99.24
Model 109	81.46	98.27	98.77	99.26
Model 110	72.39	98.37	98.60	99.21
Model 111	79.50	98.47	98.60	99.24
Model 112	72.30	98.27	98.60	99.23
Model 113	75.74	98.34	98.54	99.26
Model 114	73.64	98.34	98.97	99.21
Model 115	73.63	98.44	98.97	99.23
Model 116	81.86	98.37	98.93	99.20
Model 117	83.95	98.44	98.90	99.21
Model 118	54.05	98.34	98.90	99.26
Model 119	63.34	98.37	98.87	99.20
Model 120	55.36	98.50	98.87	99.26
Model 121	60.40	98.24	98.80	99.21
Model 122	54.12	98.20	98.50	99.14
Model 123	59.46	98.60	98.10	99.00
Model 124	51.68	98.50	97.77	98.76
Model 125	90.17	98.44	97.31	98.77
Model 126	94.03	98.44	97.24	98.81
Model 127	94.07	98.40	97.27	98.38
Model 128	84.48	98.40	97.27	98.15
Model 129	82.75	98.40	93.83	95.56
Model 130	82.11	98.34	93.10	95.86

Table 5.13. Recall values for mRMR-based DDoS attack detection prediction models using different four machine learning methods - Part 1

Models	Recall (%)			
	LR	MLP	SVM	RF
Model 66	77.49	98.31	97.79	99.28
Model 67	75.17	98.54	97.79	99.23
Model 68	73.95	98.28	97.79	99.23
Model 69	73.67	98.48	97.79	99.23
Model 70	72.98	98.44	97.82	99.25
Model 71	73.96	98.45	97.82	99.26
Model 72	74.88	98.28	97.82	99.25
Model 73	75.35	98.41	97.82	99.25
Model 74	79.69	98.41	97.82	99.26
Model 75	78.62	98.35	97.69	99.23
Model 76	77.95	98.32	97.94	99.25
Model 77	79.31	98.38	97.79	99.26
Model 78	77.65	98.47	97.79	99.25
Model 79	77.84	98.47	97.82	99.20
Model 80	78.14	98.38	97.82	99.25
Model 81	78.30	98.55	97.85	99.20
Model 82	74.96	98.48	97.82	99.23
Model 83	72.35	98.42	97.82	99.26
Model 84	77.87	98.38	97.82	99.25
Model 85	73.53	98.45	97.85	99.23
Model 86	75.47	98.25	97.85	99.25
Model 87	76.53	98.28	98.01	99.23
Model 88	78.31	98.40	98.01	99.25
Model 89	76.49	98.55	98.04	99.25
Model 90	79.31	98.35	98.04	99.25
Model 91	81.15	98.28	98.04	99.25
Model 92	79.28	98.38	98.20	99.25
Model 93	86.17	98.32	98.47	99.23
Model 94	82.32	98.31	98.47	99.23
Model 95	80.45	98.39	98.47	99.25
Model 96	77.70	98.34	98.47	99.28
Model 97	82.18	98.38	98.47	99.23
Model 98	85.92	98.47	98.50	99.23
Model 99	82.12	98.41	98.50	99.29
Model 100	83.02	98.38	98.50	99.25
Model 101	84.56	98.28	98.50	99.25
Model 102	86.58	98.24	98.50	99.25
Model 103	81.10	98.39	98.50	99.25
Model 104	80.84	98.39	98.50	99.19
Model 105	81.71	98.51	98.50	99.26
Model 106	81.11	98.35	98.50	99.23
Model 107	81.53	98.37	98.50	99.25
Model 108	80.73	98.38	98.50	99.25
Model 109	86.45	98.28	98.76	99.26
Model 110	76.80	98.38	98.60	99.22
Model 111	85.24	98.48	98.60	99.25
Model 112	77.61	98.28	98.60	99.23
Model 113	79.01	98.35	98.54	99.26
Model 114	77.88	98.35	98.96	99.22
Model 115	78.25	98.44	98.96	99.23

Table 5.14. Recall values for mRMR-based DDoS attack detection prediction models using different four machine learning methods - Part 2

Models	Recall (%)			
	LR	MLP	SVM	RF
Model 116	86.99	98.39	98.93	99.20
Model 117	87.14	98.45	98.90	99.22
Model 118	66.49	98.35	98.90	99.26
Model 119	63.46	98.38	98.87	99.20
Model 120	58.08	98.51	98.87	99.26
Model 121	59.63	98.25	98.80	99.22
Model 122	58.70	98.22	98.50	99.14
Model 123	61.75	98.61	98.11	99.01
Model 124	44.50	98.51	97.80	98.77
Model 125	91.43	98.44	97.36	98.78
Model 126	94.67	98.44	97.30	98.81
Model 127	94.75	98.41	97.33	98.40
Model 128	89.41	98.41	97.34	98.17
Model 129	87.93	98.41	94.35	95.72
Model 130	88.12	98.34	93.74	96.04

Table 5.15. F1-Score values for mRMR-based DDoS attack detection prediction models using different four machine learning methods – Part 1

Models	F1-Score (%)			
	LR	MLP	SVM	RF
Model 66	75.38	98.31	97.77	99.27
Model 67	72.95	98.54	97.77	99.23
Model 68	72.30	98.28	97.77	99.23
Model 69	72.26	98.47	97.77	99.23
Model 70	72.04	98.44	97.81	99.24
Model 71	73.00	98.44	97.81	99.26
Model 72	73.43	98.27	97.81	99.25
Model 73	73.02	98.41	97.81	99.24
Model 74	76.75	98.41	97.81	99.26
Model 75	75.97	98.34	97.68	99.23
Model 76	74.93	98.31	97.94	99.25
Model 77	76.25	98.37	97.78	99.26
Model 78	74.50	98.47	97.78	99.25
Model 79	74.55	98.47	97.81	99.20
Model 80	74.69	98.38	97.81	99.24
Model 81	74.87	98.54	97.84	99.20
Model 82	73.22	98.48	97.81	99.23
Model 83	71.39	98.41	97.81	99.26
Model 84	74.66	98.37	97.81	99.25
Model 85	71.82	98.44	97.84	99.23
Model 86	73.13	98.25	97.84	99.25
Model 87	73.77	98.28	98.00	99.23
Model 88	75.51	98.40	98.00	99.25
Model 89	73.27	98.54	98.04	99.25
Model 90	76.61	98.34	98.04	99.24
Model 91	77.97	98.28	98.04	99.24
Model 92	78.38	98.38	98.20	99.24
Model 93	84.59	98.31	98.47	99.23
Model 94	80.59	98.31	98.47	99.23
Model 95	77.27	98.38	98.47	99.24

Table 5.16. F1-Score values for mRMR-based DDoS attack detection prediction models using different four machine learning methods – Part 2

Models	F1-Score (%)			
	LR	MLP	SVM	RF
Model 96	75.90	98.34	98.47	99.27
Model 97	80.22	98.37	98.47	99.23
Model 98	86.00	98.47	98.50	99.23
Model 99	82.18	98.41	98.50	99.29
Model 100	82.21	98.37	98.50	99.24
Model 101	83.01	98.28	98.50	99.24
Model 102	85.11	98.24	98.50	99.24
Model 103	80.78	98.38	98.50	99.24
Model 104	77.61	98.38	98.50	99.18
Model 105	78.46	98.51	98.50	99.26
Model 106	78.71	98.34	98.50	99.23
Model 107	77.14	98.37	98.50	99.24
Model 108	76.99	98.37	98.50	99.24
Model 109	83.88	98.28	98.77	99.26
Model 110	74.53	98.38	98.60	99.21
Model 111	82.27	98.48	98.60	99.24
Model 112	74.86	98.27	98.60	99.23
Model 113	77.34	98.34	98.54	99.26
Model 114	75.70	98.34	98.97	99.21
Model 115	75.87	98.44	98.97	99.23
Model 116	84.35	98.38	98.93	99.20
Model 117	85.51	98.44	98.90	99.22
Model 118	59.63	98.34	98.90	99.26
Model 119	63.40	98.37	98.87	99.20
Model 120	56.69	98.51	98.87	99.26
Model 121	60.02	98.24	98.80	99.22
Model 122	56.32	98.21	98.50	99.14
Model 123	60.58	98.60	98.11	99.00
Model 124	47.82	98.51	97.79	98.77
Model 125	90.80	98.44	97.33	98.78
Model 126	94.35	98.44	97.27	98.81
Model 127	94.41	98.41	97.30	98.39
Model 128	86.88	98.41	97.30	98.16
Model 129	85.26	98.41	94.09	95.64
Model 130	85.01	98.34	93.42	95.95

5.2. Discussion

In this section, metric results are analyzed in detail. Along with comparing the success of ML methods, successful models have also been compared. The importance of Relief-F and mRMR feature selectors is particularly understood. Models obtained with Relief-F and mRMR feature selectors have been evaluated under different headings. In the general discussion section, all models have been evaluated and completed.

5.2.1. Discussion for Relief-F Method

Relief-F feature selection was found to be more successful than the mRMR method in the predictions made by the LR method. In addition, the improvement in the LR method was found to

be the highest not only in the models obtained with the Relief-F method but also in the models obtained with the mRMR method. The evaluations on the models obtained with Relief-F are listed below.

When models created with Relief-F feature selection algorithm are compared, the best ML methods in sequence have been RF, SVM, MLP, LR.

- Among the LR-based models, Model 60 had the lowest value in all metrics. Model 59 was the most successful model. When the Actmin feature was removed from Model 64, a significant decrease was observed in Model 65. It was observed that Model 59 was the most successful model. In the best model, 97.03%, 97.04%, 97.15% and 97.09% values were obtained for accuracy, precision, recall and F1-score, respectively. This is the best improvement in predictions with Relief-F feature selection. According to Relief-F feature selection, the most successful features in LR models are Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes. 25.90% increase compared to the prediction without feature selection.
- Among the MLP-based models, Model 16 was found to be the least successful for all metric values. The most successful model is Model 62. In the best model, the accuracy, precision, recall and F1-score values were 98.57%, 98.57%, 98.58% and 98.57%, respectively. It was observed that the maximum decrease was realized by removing the Avrpcks feature. According to Relief-F feature selection, Fplmean, Cwefc, Actmin, Ackfc, Bimin were found to be the most successful features in MLP models. There was a 0.20% increase compared to the prediction without feature selection. The increase between the worst model and the best model was 0.40%.
- Among the SVM-based models, Model 65 was found to be the least successful for all metric values. The most successful model is Model 57. In the best model, the accuracy, precision, recall and F1-score values were 98.63%, 98.64%, 98.64% and 98.64%, respectively. The lowest values were obtained in all metrics by removing the Actmin feature. There was a 46.17% decrease in accuracy. According to Relief-F feature selection, the most successful features in MLP models are Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf. There was a 0.89% increase compared to the prediction without feature selection. The increase between the worst model and the best model was 48.80%.
- Among the RF-based models, Model 65 was found to be the least successful for all metric values. The most successful models were observed to be Model 17, Model 19 and Model 23. A value of 99.28% was obtained in the best models. By removing the Actmin feature, the lowest values were obtained in all metrics. There was a 45.11% decrease in accuracy. According to Relief-F feature selection, the features found in the most successful RF models are Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc,

Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl. 1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts. There was a 0.04% increase in accuracy compared to the prediction without feature selection. The increase between the worst model and the best model was 48.88%.

5.2.2. Discussion for mRMR Method

In the predictions made with MLP, SVM and RF ML methods, mRMR feature selection was found to be more successful than Relief-F. The evaluations about the models obtained with mRMR are listed below.

When models created with mRMR feature selection algorithm are compared, the best ML methods in sequence have been RF, SVM, MLP, LR.

- Among the LR-based models, Model 124 had the lowest value in all metrics. The most successful model was observed to be Model 127. When the Fbytes feature was removed from Model 124, it was found that there was a high improvement in Model 60. Accuracy value of 94.07% was obtained in the best model. This value is the best improvement in predictions made with mRMR feature selection. The most successful features in LR models according to mRMR feature selection are Fplmean, Cwefc, Actmin, Ackfc, Bimin. There was a 20.84% increase compared to the prediction without feature selection.
- Among the MLP-based models, Model 122 was found to be the least successful for all metric values. The most successful model was observed to be Model 123. A value of 98.60% was obtained in the best model. The maximum value was obtained by subtracting the İwbf feature from the minimum value. According to mRMR feature selection, the most successful features in MLP models were Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl. There was a 0.3% increase compared to the prediction without feature selection. The increase between the worst model and the best model was 0.4%.
- Among the SVM-based models, Model 130 was found to be the least successful for all metric values. The most successful models are Model 114 and Model 115. A value of 98.97% was obtained in the best models. By removing the Cwefc feature, the lowest values were obtained in all metrics. There was a decrease of 0.70% in the accuracy value. According to mRMR feature selection, the most successful features in MLP models are Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss. There was a 1.20%

increase compared to the prediction without feature selection. The increase between the worst model and the best model was 5.77%.

- Among the RF-based models, Model 129 was found to be the least successful for all metric values. The most successful model was observed to be Model 99. A value of 99.29% was obtained in the best model. By removing the Ackfc feature, the lowest values were obtained in all metrics. There was a 2.64% decrease in the accuracy value. According to mRMR feature selection, the features found in the most successful RF models are Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl. 1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax. There was a 0.10% increase compared to the prediction without feature selection. The increase between the worst model and the best model was 3.78%.

5.3. General Discussion

The best method in all models was obtained with RF machine learning applied to the dataset prepared with the mRMR feature selection method. SVM, MLP followed the RF model. The best results obtained with SVM and MLP were obtained from models created with mRMR. The worst model was LR.

When models created with both feature selection algorithms are compared, the best ML methods in sequence have been RF, SVM, MLP, LR

In RF, SVM, MLP prediction methods, mRMR feature selection performed better than Relief-F. However, Relief-F was found to be more successful than the mRMR method in the predictions made in the LR method.

The LR model showed the highest improvement among the models. SVM, MLP followed the LR model. The least improved model was RF.

In general, the best results were obtained mRMR between Model 99 and Model 127. The features used in these models are Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss.

In general, the best results were obtained Relief-F between Model 23 and Model 62. The features used in these models are ackfc, prtcl, fplmin, minpl, fbytes, dport, fplmax, maxpl, urgfc, fpckts, iwbf, fwdpckts, cwefc, fpflg, rdtfc, fdrtn, fwditotal, duratio, fimax, fwdimean, fwdimax, bplmax, imax, fwdistd, plstd, sbfwdbyt, imean, imin, plmean, iwbb, fistd, bplstd, tlfpckts, bitotal, mplmin, mssf, bimin, sbfwdpckt, tfpckts, adpf, bimax, fwdhl1, plvrnc, fwdhl.

RF is the most successful machine learning method and the other methods are respectively Accuracy, precision, recall and F1-Score metric values are analyzed from Figure 5.1. to Figure 5.4.

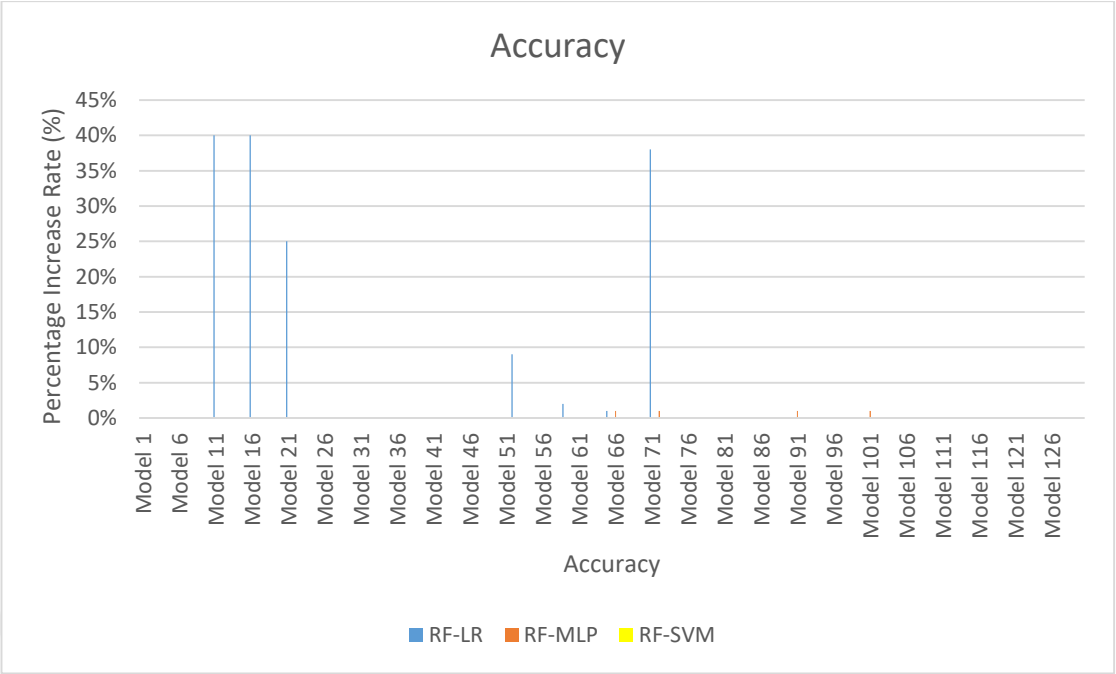

Figure 5.1. Percentage increase rates in accuracies of DDoS attack detection for RF-based models compared to accuracies obtained by the rest of the machine learning methods

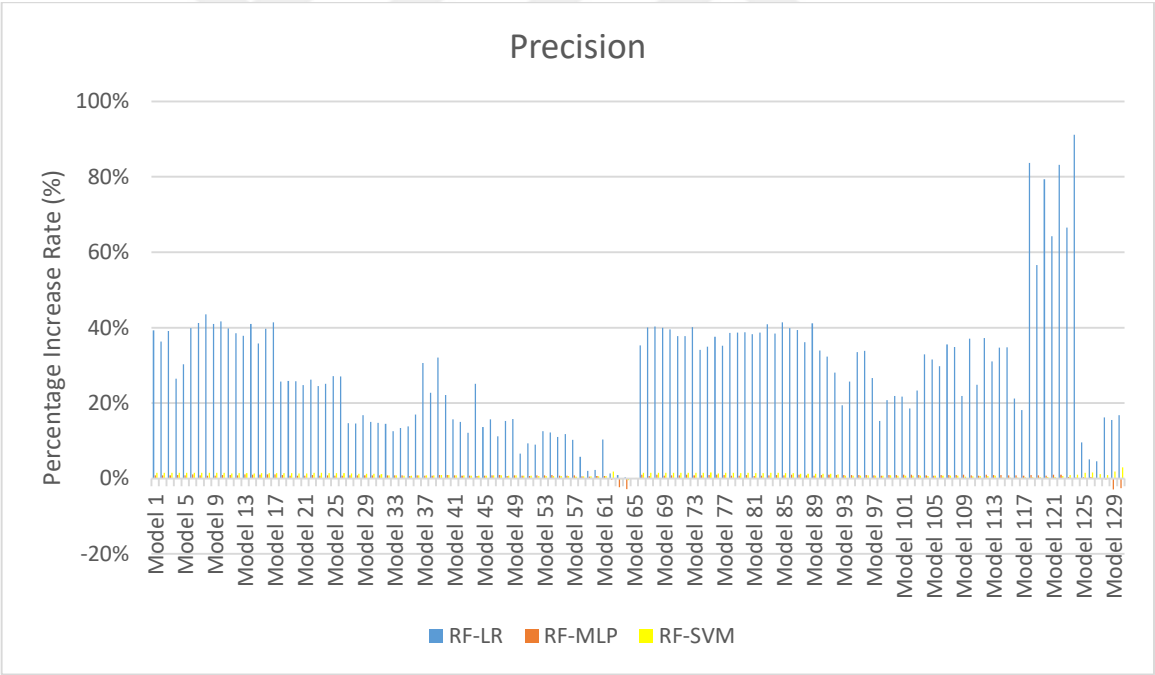

Figure 5.2. Percentage increase rates in precision of DDoS attack detection for RF-based models compared to precision obtained by the rest of the machine learning methods.

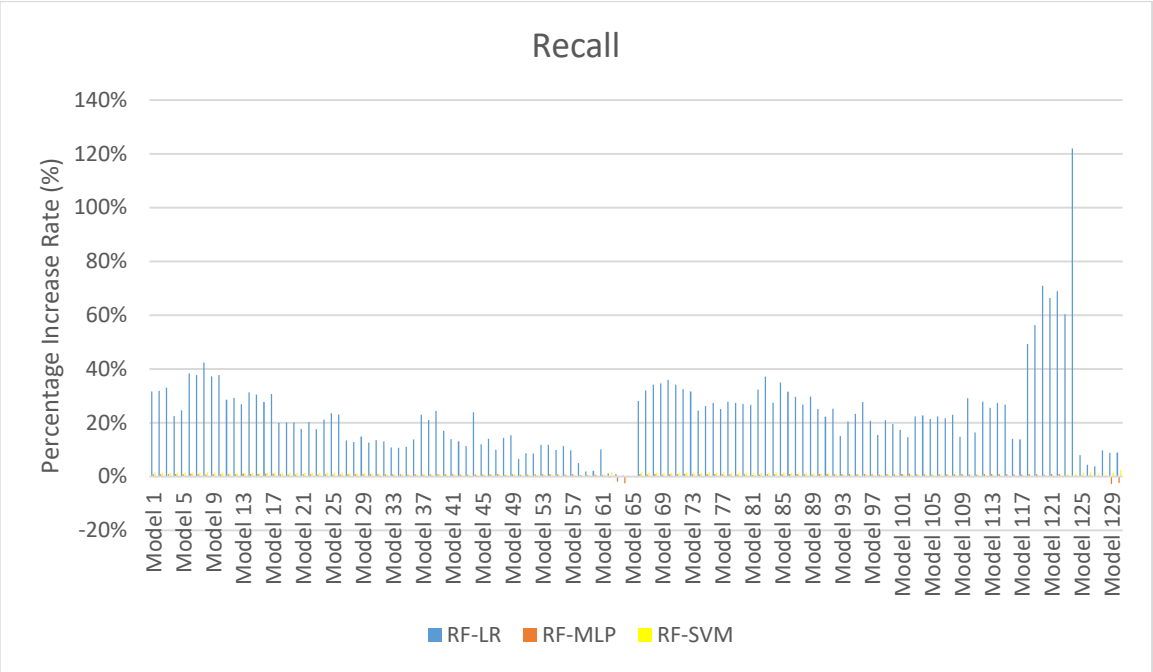

Figure 5.3. Percentage increase rates in recall of DDoS attack detection for RF-based models compared to recall obtained by the rest of the machine learning methods.

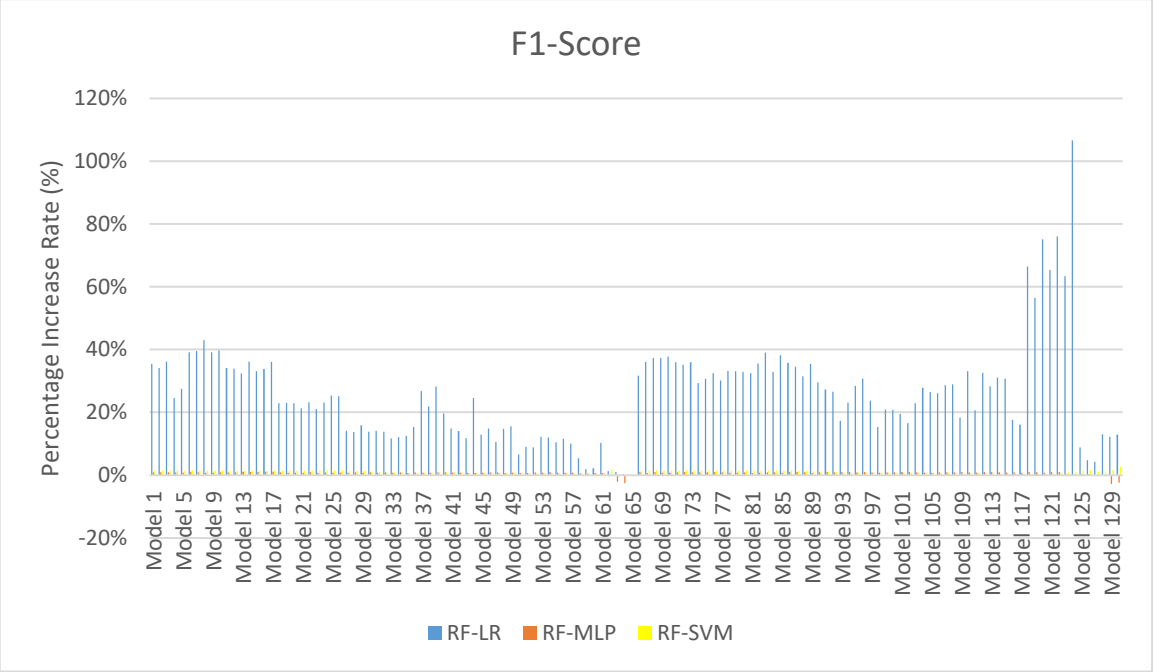

Figure 5.4. Percentage increase rates in F1-Score of DDoS attack detection for RF-based models compared to F1-Score obtained by the rest of the machine learning methods.

5.4. Comparing Results with Other Studies in the Literature

This section compares the accuracy, precision, recall and F1-Score metrics of several studies with the DDoS dataset. The best methods of the compared studies are added to the table. Among the included studies, k-NN was found to be the best model.

The models achieving the highest accuracy performance is Model 99 in this study, all attaining a value of 99.29%.

(Tahirou et al., 2023) have employed KNN and NB machine learning methods, different from this study. In both studies, SVM machine learning method and the CICDDoS2019 dataset have been utilized. They conducted a study on DDoS attack detection in SDN environments. In this study, fewer features were used. While the best model in this study was the traditional method RF, in Tahirou and his colleagues' study, it was SVM. Despite the use of two different feature selections in this study, Tahirou and his colleagues did not use any feature selectors. Sequentially, accuracy, precision, recall, and F-measure metrics yielded results of 98.9%, 99%, 98%, and 97.9%, whereas in Model 99, a result of 99.29% was obtained across all metrics.

(Almiani et al., 2021), conducted a study to detect DDoS attacks with ANN. Similar to this study, Almiani employed neural network methods in his own research to detect DDoS attacks. Unlike this study, Almiani and his colleagues employed the Kalman propagation method with ANN. In this study, RF, SVM, MLP, and LR machine learning methods were utilized. The dataset for both studies was selected as CICDDoS2019. While the accuracy, precision, and F1-score metrics were determined as 97.71, 96.13, and 97.75, respectively, in sequence, in this study, Model 99 achieved a result of 99.29% across all metrics.

(Lawall et al., 2021) conducted a study to detect DDoS attacks with k-NN. (Lawall et al., 2021) and the purpose of this study is DDoS attack detection. In both studies, the CICDDoS2019 dataset has been utilized. (Lawall et al., 2021) and their colleagues, in contrast to using k-NN, DT, and NB machine learning methods, employed RF, SVM, MLP, and LR machine learning methods in this study. The accuracy value has been calculated as 99.99%, and precision, recall, and F1-score metric values have been determined as 100.00%. In this study, Model 99 achieved a result of 99.29% across all metrics. While (Lawall et al., 2021) and their colleagues did not prefer feature selectors, in this study, two feature selection methods were used.

The objectives of the studies conducted with the CICDDoS2019 dataset are generally DDoS attack detection. This is also the case in the studies discussed here. Not only ML methods but also the dataset affects the success. The type of data set is not enough, and the number of samples taken from the data set and the number of features also affect the results. Therefore, a one-to-one comparison would not be very accurate.

In general, the following conclusion can be reached. The success rate in DDoS attack detection studies is notably high. Machine learning methods achieve remarkable efficiency in detecting DDoS attacks. Given that billions of attacks occur, and large-sized packets are sent, even a 1% failure implies that the damage cannot be completely prevented. Therefore, these elevated success rates are attributed to the simplicity of DDoS attacks, the increasing frequency of such attacks each year, and the substantial damages incurred. Consequently, there is a need for further current studies in this field.

Table 5.20. Comparison of the results with previous works

Study	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
(Lawall et al., 2021)	k-NN	99.99	100	100	100
(Tahirou et al., 2023)	SVM	98.90	99.00	98.00	97.90
(Almiani et al., 2021)	ANN	97.71	96.13	n.a.	97.75
This Study (Model 99)	RF	99.29	99.29	99.29	99.29
This Study (Model 115)	SVM	98.97	98.97	98.96	98.97
This Study (Model 123)	MLP	98.60	98.60	98.61	98.60
This Study (Model 59)	LR	97.03	97.04	97.15	97.09

6. CONCLUSION AND FUTURE WORK

6.1. Conclusion

DDoS attacks are a type of cyber-attack that threatens all internet-connected systems and software and is extremely difficult to prevent. Recently, various studies have been conducted on DDoS attacks. However, considering the difficulty in preventing the attack and the negative impact on the systems, it is obvious that more work needs to be carried out. Filtering efforts to prevent DDoS attacks are insufficient. It not only disrupts the work and functioning of the institutions it affects, but also causes damage on material and prestige. It is vital to carry out up-to-date studies to overcome such negative situations.

In this study, DDoS attacks have been detected with dataset which is CICDDOS2019. Relief-F and mRMR feature selectors were used to rank the features in the CICDDOS2019 dataset based on their scores. 130 models were created in accordance with the determined feature order. Metric values for these 130 models were calculated using LR, MLP, SVM, and RF machine learning classifiers. In this context, accuracy, precision, recall, and F1-Score metric values have been calculated. According to the two feature selectors considered in the study, the RF and LR algorithms are the most successful and the least successful methods, respectively. However, in terms of the improvement in the accuracy obtained from the methods based on selected features, the LR method was found to be superior to the other methods.

The features that come to the prominence through the mRMR and Relief-F feature selectors vary. In general, it is observed that the best features are provided in the first 17 features. According to mRMR, most important features are: Fplmean, Actmin, Bimin, mssf, Plmean, Sbbwdpckt, Fwditotal, Duratio. According to Relief-F, most important features are: Dport, Fplmax, Fpckts, Fwdpckts, Fpflg, Rdtfc, Fimin and Afwdss. Among the first 17 features, the prominent features according to both feature selectors can be listed as follows: Cwefc, Ackfc, Urgfc, Fbytes, Minpl, İwbf, Fplmin, Maxpl, Prtcl.

The four different machine learning methods have been applied on the built 130 models. The best results for each machine learning method have been obtained in different models. In experiments conducted with the LR-based machine learning method, the best result was achieved with the model obtained through Relief-F feature selection at Model 59. In experiments conducted with the MLP machine learning method, the best result was obtained with the model obtained through mRMR feature selection at Model 123. In experiments conducted with the SVM machine learning method, the best result was obtained with the model obtained through mRMR feature selection at Model 115. In experiments conducted with the RF machine learning method, the best result was obtained with the model obtained through mRMR feature selection at Model 99.

The LR machine learning experiments revealed that features such as ackfc, prtcl, fplmin, minpl, fbytes, dport, fplmax, and maxpl were incorporated into the optimal model. In the MLP

machine learning experiments, key features like dport, prtcl, fdrtn, tfpckts, tbpckts, tlfpckts, tlbpckts, fplmax, and fplmin emerged as integral components of the top-performing model. As for the SVM machine learning experiments, essential features including dport, prtcl, fdrtn, tfpckts, tbpckts, tlfpckts, tlbpckts, fplmax, fplmin, fplmean, fplstd, bplmax, mplmin, bplmean, bplstd, fbytes, and fpckts were identified in the optimal model. In the RF machine learning experiments, a comprehensive set of features, such as dport, prtcl, fdrtn, tfpckts, tbpckts, tlfpckts, tlbpckts, fplmax, fplmin, fplmean, fplstd, bplmax, mplmin, bplmean, bplstd, fbytes, fpckts, fimean, fistd, fimax, fimin, fwditotal, fwdimean, fwdistd, fwdimax, fwdimin, bitotal, bimean, bistd, bimax, bimin, fpflg, and fwdhl, were deemed essential for optimal model performance.

In both feature selection methods, it has been determined that packet size and temporal features are important. Relief-F feature selection prioritized packet sizes, while mRMR feature selection prioritized temporal features. According to the obtained results, the most successful model in MLP, SVM, and RF machine learning methods was achieved with temporal features. On the other hand, according to LR machine learning method, the most successful model was obtained with packet size features.

6.2. Future Work

Researchers have undertaken extensive investigations into DDoS attack detection, exploring diverse datasets, traditional and modern machine learning techniques, as well as hybrid methodologies. This comprehensive exploration has yielded numerous effective approaches, resulting in notable success rates in DDoS attack detection. Both quantitatively and qualitatively, these studies have demonstrated a high level of proficiency. Nonetheless, the pervasive threat of DDoS attacks across various domains highlights the ongoing relevance and significance of continued research in DDoS attack detection.

In this context, it is crucial to focus not only on relying on old methods but also on improving existing methods and emphasizing new machine learning detection studies. Additionally, using different datasets can significantly contribute to DDoS attack detection studies. For instance, creating new datasets from IoT devices and virtual machines is feasible. Investment should be made for the creation of new datasets. Devices should be procured and laboratories should be established specifically for the purpose of detecting DDoS attacks. Despite ongoing efforts in conducting up-to-date studies against DDoS attacks, attackers also update their software.

Developing hybrid scenarios in machine learning methods and conducting new analysis studies are of great importance. Newly created datasets, when combined with old datasets, can provide different approaches. Alongside all these new approaches, when feature selection methods are applied, an increase in success rates is anticipated. Developing original scenarios by using new machine learning methods and up-to-date feature selection techniques on new datasets will contribute significantly to the studies of DDoS attack detection.

Additionally, it is crucial to create datasets that encompass a wide variety of attacks. As the diversity of attacks increases, the success rates of models tend to decrease. Improvement efforts conducted on challenging datasets will also make a significant contribution to DDoS attack detection studies. In machine learning studies, the issue of overfitting hinders the progress of research. Features causing overfitting in data set creation studies should be identified and clearly expressed in articles.

Studies in the literature universally report success in attack detection. Attackers engage in meticulous preparation prior to launching attacks. Hence, it is imperative to underscore preventive measures to thwart these preparations. DDoS attacks leverage a multitude of internet-connected devices, including computers, tablets, and phones, making any such device a potential tool for attackers. Ensuring the security of devices implicated in DDoS attacks is paramount. Often, users of these devices are unaware of their devices' involvement. Consequently, there is a pressing need to develop software aimed at preventing internet-connected devices from being commandeered as bots.

REFERENCES

Abut, F., & Akay, F. (2015). Machine learning and statistical methods for the prediction of maximal oxygen uptake: recent advances. *Medical Devices: Evidence and Research*, 369. https://doi.org/10.2147/MDER.S57281

Acid, S., de Campos, L. M., & Fernandez, M. (2011). Minimum redundancy maximum relevancy versus score-based methods for learning Markov boundaries. *2011 11th International Conference on Intelligent Systems Design and Applications*, 619–623. https://doi.org/10.1109/ISDA.2011.6121724

Adhikary, K., Bhushan, S., Kumar, S., & Dutta, K. (2020). Hybrid Algorithm to Detect DDoS Attacks in VANETs. *Wireless Personal Communications*, *114*(4), 3613–3634. https://doi.org/10.1007/s11277-020-07549-y

Alimi, K. O. A., Ouahada, K., Abu-Mahfouz, A. M., Rimer, S., & Alimi, O. A. (2022). Refined LSTM Based Intrusion Detection for Denial-of-Service Attack in Internet of Things. *Journal of Sensor and Actuator Networks*, *11*(3), 32. https://doi.org/10.3390/jsan11030032

Almiani, M., AbuGhazleh, A., Jararweh, Y., & Razaque, A. (2021). DDoS detection in 5G-enabled IoT networks using deep Kalman backpropagation neural network. *International Journal of Machine Learning and Cybernetics*. https://doi.org/10.1007/s13042-021-01323-7

Alsirhani, A., Sampalli, S., & Bodorik, P. (2019). DDoS Detection System: Using a Set of Classification Algorithms Controlled by Fuzzy Logic System in Apache Spark. *IEEE Transactions on Network and Service Management*, *16*(3), 936–949. https://doi.org/10.1109/TNSM.2019.2929425

Aydin, E., & Bahtiyar, S. (2021). OCIDS: An Online CNN-Based Network Intrusion Detection System for DDoS Attacks with IoT Botnets. *Proceedings - 2021 14th International Conference on Security of Information and Networks, SIN 2021*. https://doi.org/10.1109/SIN54109.2021.9699288

Aytaç, T. (2020). *DDoS Saldırılarının Tespit Edilmesinde Makine Öğrenimi Yöntemlerinin Uygulanması [Yüksek Lisans Tezi, İstanbul Ticaret Üniversitesi Fen Bilimleri Enstitüsü]*.

Bhavsar, H., & Panchal, M. H. (2012). A Review on Support Vector Machine for Data Classification. In *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)* (Vol. 1, Issue 10).

Breiman, L. (2001). *Random Forests* (Vol. 45).

Cil, A. E., Yildiz, K., & Buldu, A. (2021). Detection of DDoS attacks with feed forward based deep neural network model. *Expert Systems with Applications*, *169*, 114520. https://doi.org/10.1016/j.eswa.2020.114520

Deniz, E. (2023). *Bulut Ağ Ortamında Dağıtılmış Hizmet Reddi (DDoS) Saldırılarını Algılamak İçin Derin Öğrenme Tabanlı Bir Sistem [Yüksek Lisans Tezi, Kütahya Fen Bilimleri Enstitüsü]*.

Doshi, R., Apthorpe, N., & Feamster, N. (2018). Machine learning DDoS detection for consumer internet of things devices. *Proceedings - 2018 IEEE Symposium on Security and Privacy Workshops, SPW 2018*, 29–35. https://doi.org/10.1109/SPW.2018.00013

Fenanir, S., & Semchedine, F. (2023). Smart Intrusion Detection in IoT Edge Computing Using Federated Learning. *Revue d'Intelligence Artificielle*, *37*(5), 1133–1145. https://doi.org/10.18280/ria.370505

Gumaste, S., D. G., N., Shinde, S., & K, A. (2021). Detection of DDoS Attacks in OpenStack-based Private Cloud Using Apache Spark. *Journal of Telecommunications and Information Technology*, *4*(2020), 62–71. https://doi.org/10.26636/jtit.2020.146120

Hamid Arif, N., Reza Faisal, M., Farmadi, A., Turianto Nugrahadi, D., Abadi, F., & Ali Ahmad, U. (2024). ShareAlike 4.0 International License (CC BY-SA 4.0). How to cite: Predicting the Need for Cardiovascular Surgery: A Comparative Study of Machine Learning Models An Approach to ECG-based Gender Recognition Using Random Forest Algorithm. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, *6*(2), 107–115. https://doi.org/10.35882/jeeemi.v6i2.363

Hasan, Md. Z., Hasan, K. M. Z., & Sattar, A. (2018). Burst Header Packet Flood Detection in Optical Burst Switching Network Using Deep Learning Model. *Procedia Computer Science*, *143*, 970–977. https://doi.org/10.1016/j.procs.2018.10.337

He, Z., Zhang, T., & Lee, R. B. (2017). Machine Learning Based DDoS Attack Detection from Source Side in Cloud. *2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud)*, 114–120. https://doi.org/10.1109/CSCloud.2017.58

Islam, U., Al-Atawi, A., Alwageed, H. S., Ahsan, M., Awwad, F. A., & Abonazel, M. R. (2023). Real-Time Detection Schemes for Memory DoS (M-DoS) Attacks on Cloud Computing Applications. *IEEE Access*, *11*, 74641–74656. https://doi.org/10.1109/ACCESS.2023.3290910

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2023). *An Introduction to Statistical Learning with Applications in R Second Edition*.

Jazi, H. H., Gonzalez, H., Stakhanova, N., & Ghorbani, A. A. (2017). Detecting HTTP-based application layer DoS attacks on web servers in the presence of sampling. *Computer Networks*, *121*, 25–36. https://doi.org/10.1016/j.comnet.2017.03.018

Ji, W., Ebrahimi, T., Li, Z., Yuan, J., Wu, D. O., & Xin, Y. (2021). Guest Editorial: Emerging Visual IoT Technologies for Future Communications and Networks. *IEEE Wireless Communications*, *28*(4), 10–11. https://doi.org/10.1109/MWC.2021.9535459

Karan, B. V., D. G., N., & Hiremath, P. S. (2018). Detection of DDoS Attacks in Software Defined Networks. *2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS)*, 265–270. https://doi.org/10.1109/CSITSS.2018.8768551

Kira, K., & Rendell, L. A. (1992). The Feature Selection Probşem: Traditional Methods and a New Algorithm. *AAAI*, *2*, 129–134.

Kiruthiga, G., Saraswathi, P., Rajkumar, S., Suresh, S., Dhiyanesh, B., & Radha, R. (2022). Effective DDoS Attack Detection using Deep Generative Radial Neural Network in the Cloud Environment. *2022 7th International Conference on Communication and Electronics Systems (ICCES)*, 675–681. https://doi.org/10.1109/ICCES54183.2022.9835916

Kohnehshahri, M. A., Mohammadi, R., Abdoli, H., & Nassiri, M. (2022). An Efficient Method for Online Detection of DRDoS Attacks on UDP-Based Services in SDN Using Machine Learning Algorithms. *Mobile Information Systems*, *2022*, 1–13. https://doi.org/10.1155/2022/1169035

Kolukisa, B. (2024). *MACHINE LEARNING APPROACHES FOR INTERNET OF THINGS BASED VEHICLE TYPE CLASSIFICATION AND NETWORK ANOMALY DETECTION*.

Kumar, H., Aoudni, Y., Ortiz, G. G. R., Jindal, L., Miah, S., & Tripathi, R. (2022). Light Weighted CNN Model to Detect DDoS Attack over Distributed Scenario. *Security and Communication Networks*, *2022*, 1–10. https://doi.org/10.1155/2022/7585457

Kushwah, G. S., & Ranga, V. (2021). Optimized extreme learning machine for detecting DDoS attacks in cloud computing. *Computers & Security*, *105*, 102260. https://doi.org/10.1016/j.cose.2021.102260

Kyaw, A. T., Zin Oo, M., & Khin, C. S. (2020). Machine-Learning Based DDOS Attack Classifier in Software Defined Network. *2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 431–434. https://doi.org/10.1109/ECTI-CON49241.2020.9158230

Li1, Z., Wei2, L., Li2, W., Wei2, L., Chen2, M., Lv2, M., Zhi3, X., Wang3, C., & Gao2, N. (2019). *Research on DDoS Attack Detection Based on ELM in IoT Environment; Research on DDoS Attack Detection Based on ELM in IoT Environment*.

Ma, L., Chai, Y., Cui, L., Ma, D., Fu, Y., & Xiao, A. (2020). A Deep Learning-Based DDoS Detection Framework for Internet of Things. *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 1–6. https://doi.org/10.1109/ICC40277.2020.9148944

Makuvaza, A., Jat, D. S., & Gamundani, A. M. (2021). Deep Neural Network (DNN) Solution for Real-time Detection of Distributed Denial of Service (DDoS) Attacks in Software Defined Networks (SDNs). *SN Computer Science*, *2*(2). https://doi.org/10.1007/s42979-021-00467-1

Muhammad Aminu, L., Shaikh, R. A., & Hassan, S. R. (2021). A DDoS Attack Mitigation Framework for IoT Networks using Fog Computing. *Procedia Computer Science*, *182*, 13–20. https://doi.org/10.1016/j.procs.2021.02.003

Özekes, S., & KARAKOÇ, E. N. (2019). Makine Öğrenmesi Yöntemleriyle Anormal Ağ Trafiğinin Tespit Edilmesi. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, *7*(1), 566–576. https://doi.org/10.29130/dubited.498358

Pandithurai, O., Venkataiah, C., Tiwari, S., & Ramanjaneyulu, N. (2024). DDoS attack prediction using a honey badger optimization algorithm based feature selection and Bi-LSTM in cloud environment. *Expert Systems with Applications*, *241*, 122544. https://doi.org/10.1016/j.eswa.2023.122544

Polat, H., Polat, O., & Cetin, A. (2020). Detecting DDoS Attacks in Software-Defined Networks Through Feature Selection Methods and Machine Learning Models. *Sustainability*, *12*(3), 1035. https://doi.org/10.3390/su12031035

Ramchoun, H., Amine, M., Idrissi, J., Ghanou, Y., & Ettaouil, M. (2016). Multilayer Perceptron: Architecture Optimization and Training. *International Journal of Interactive Multimedia and Artificial Intelligence*, *4*(1), 26. https://doi.org/10.9781/ijimai.2016.415

Rivas, P., DeCusatis, C., Oakley, M., Antaki, A., Blaskey, N., LaFalce, S., & Stone, S. (2019). Machine Learning for DDoS Attack Classification Using Hive Plots. *2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, 0401–0407. https://doi.org/10.1109/UEMCON47517.2019.8993021

Robnik-ˇ Sikonja, M., & Kononenko, I. (2003). *Theoretical and Empirical Analysis of ReliefF and RReliefF* (Vol. 53).

Roopak, M., Tian, G. Y., & Chambers, J. (2020). An Intrusion Detection System Against DDoS Attacks in IoT Networks. *2020 10th Annual Computing and Communication Workshop and Conference, CCWC 2020*, 562–567. https://doi.org/10.1109/CCWC47524.2020.9031206

Roopak, M., Yun Tian, G., & Chambers, J. (2019). Deep Learning Models for Cyber Security in IoT Networks. *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, 0452–0457. https://doi.org/10.1109/CCWC.2019.8666588

Rukavitsyn, A., Borisenko, K., & Shorov, A. (2017). Self-learning method for DDoS detection model in cloud computing. *2017 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, 544–547. https://doi.org/10.1109/EIConRus.2017.7910612

Sahingoz, O. K., Çebi, C. B., Bulut, F. S., Fırat, H., & Karataş, G. (2019). Saldırı Tespit Sistemlerinde Makine Öğrenmesi Modellerinin Karşılaştırılması. *Erzincan Üniversitesi Fen Bilimleri Enstitüsü Dergisi*. https://doi.org/10.18185/erzifbed.573648

Said, D. (2023). Quantum Computing and Machine Learning for Cybersecurity: Distributed Denial of Service (DDoS) Attack Detection on Smart Micro-Grid. *Energies*, *16*(8), 3572. https://doi.org/10.3390/en16083572

Sambangi, S., & Gondi, L. (2020). A Machine Learning Approach for DDoS (Distributed Denial of Service) Attack Detection Using Multiple Linear Regression. *The 14th International*

Conference on Interdisciplinarity in Engineering—INTER-ENG 2020, 51. https://doi.org/10.3390/proceedings2020063051

Sharafaldin, I., Habibi Lashkari, A., & Ghorbani, A. A. (2018). Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. *Proceedings of the 4th International Conference on Information Systems Security and Privacy*, 108–116. https://doi.org/10.5220/0006639801080116

Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. *ICISSP 2018 - Proceedings of the 4th International Conference on Information Systems Security and Privacy*, *2018-January*, 108–116. https://doi.org/10.5220/0006639801080116

Sharafaldin, I., Lashkari, A. H., Hakak, S., & Ghorbani, A. A. (2019). Developing Realistic Distributed Denial of Service (DDoS) Attack Dataset and Taxonomy. *2019 International Carnahan Conference on Security Technology (ICCST)*, 1–8. https://doi.org/10.1109/CCST.2019.8888419

Shurman, M., Khrais, R., & Yateem, A. (2020). DoS and DDoS attack detection using deep learning and IDS. *International Arab Journal of Information Technology*, *17*(4A Special Issue), 655–661. https://doi.org/10.34028/iajit/17/4A/10

Soe, Y. N., Santosa, P. I., & Hartanto, R. (2019). DDoS Attack Detection Based on Simple ANN with SMOTE for IoT Environment. *2019 Fourth International Conference on Informatics and Computing (ICIC)*, 1–5. https://doi.org/10.1109/ICIC47613.2019.8985853

Stief, A., Ottewill, J. R., & Baranowski, J. (2018). Relief F-Based Feature Ranking and Feature Selection for Monitoring Induction Motors. *2018 23rd International Conference on Methods & Models in Automation & Robotics (MMAR)*, 171–176. https://doi.org/10.1109/MMAR.2018.8486097

Su, J., Danilo Vasconcellos, V., Prasad, S., Daniele, S., Feng, Y., & Sakurai, K. (2018). Lightweight Classification of IoT Malware Based on Image Recognition. *Proceedings - International Computer Software and Applications Conference*, *2*, 664–669. https://doi.org/10.1109/COMPSAC.2018.10315

Tahirou, A. K., Konate, K., & Soidridine, M. M. (2023). Detection and Mitigation of DDoS Attacks in SDN Using Machine Learning (ML). *2023 International Conference on Digital Age & Technological Advances for Sustainable Development (ICDATA)*, 52–59. https://doi.org/10.1109/ICDATA58816.2023.00019

Tan, L., Pan, Y., Wu, J., Zhou, J., Jiang, H., & Deng, Y. (2020). A New Framework for DDoS Attack Detection and Defense in SDN Environment. *IEEE Access*, *8*, 161908–161919. https://doi.org/10.1109/ACCESS.2020.3021435

Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009). A detailed analysis of the KDD CUP 99 data set. *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, 1–6. https://doi.org/10.1109/CISDA.2009.5356528

Tonkal, Ö., Polat, H., Başaran, E., Cömert, Z., & Kocaoğlu, R. (2021). Machine Learning Approach Equipped with Neighbourhood Component Analysis for DDoS Attack Detection in Software-Defined Networking. *Electronics*, *10*(11), 1227. https://doi.org/10.3390/electronics10111227

Virupakshar, K. B., Asundi, M., Channal, K., Shettar, P., Patil, S., & Narayan, D. G. (2020). Distributed Denial of Service (DDoS) Attacks Detection System for OpenStack-based Private Cloud. *Procedia Computer Science*, *167*, 2297–2307. https://doi.org/10.1016/j.procs.2020.03.282

Wang, L., & Liu, Y. (2020). A DDoS Attack Detection Method Based on Information Entropy and Deep Learning in SDN. *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 1084–1088. https://doi.org/10.1109/ITNEC48623.2020.9085007

CURRICULUM VITAE

Numan ERTİK graduated from Ahmet Kurttepeli High School. He completed his studies at Çukurova University, earning a degree in Computer Engineering. Presently, he serves as a research assistant in the Department of Computer Engineering at Ağrı İbrahim Çeçen University.

APPENDICES

Appendices A.1. Relief-F Based Prediction Models

The tables between Table A.1.1 and Table A.1.6 provide an overview of Relief-F based prediction models, showing each one along with its unique prediction variables.

Table A.1.1. Overview of Relief-F-based prediction models - Part 1

Models	**Predictor Variables**
Model 1	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean, Avrpcks, Fimin, Fwdimin, Afwdss, Actmax, Actmin, Actmean, Sbbwdbyt, Tlbpckts, Fpckts, Bimean, Abwdss, Bplmean, Actstd
Model 2	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean, Avrpcks, Fimin, Fwdimin, Afwdss, Actmax, Actmin, Actmean, Sbbwdbyt, Tlbpckts, Fpckts, Bimean, Abwdss, Bplmean
Model 3	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean, Avrpcks, Fimin, Fwdimin, Afwdss, Actmax, Actmin, Actmean, Sbbwdbyt, Tlbpckts, Fpckts, Bimean, Abwdss
Model 4	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean, Avrpcks, Fimin, Fwdimin, Afwdss, Actmax, Actmin, Actmean, Sbbwdbyt, Tlbpckts, Fpckts, Bimean
Model 5	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean, Avrpcks, Fimin, Fwdimin, Afwdss, Actmax, Actmin, Actmean, Sbbwdbyt, Tlbpckts, Fpckts

Table A.1.2. Overview of Relief-F-based prediction models -Part 2

Models	**Predictor Variables**
Model 6	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean, Avrpcks, Fimin, Fwdimin, Afwdss, Actmax, Actmin, Actmean, Sbbwdbyt, Tlbpckts
Model 7	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean, Avrpcks, Fimin, Fwdimin, Afwdss, Actmax, Actmin, Actmean, Sbbwdbyt
Model 8	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean, Avrpcks, Fimin, Fwdimin, Afwdss, Actmax, Actmin, Actmean
Model 9	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean, Avrpcks, Fimin, Fwdimin, Afwdss, Actmax, Actmin
Model 10	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean, Avrpcks, Fimin, Fwdimin, Afwdss, Actmax
Model 11	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean, Avrpcks, Fimin, Fwdimin, Afwdss
Model 12	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean, Avrpcks, Fimin, Fwdimin

Table A.1.3. Overview of Relief-F-based prediction models - Part 3

Models	**Predictor Variables**
Model 13	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean, Avrpcks, Fimin
Model 14	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean, Avrpcks
Model 15	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl, Fplmean
Model 16	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts, Bwdhl
Model 17	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt, Bwdpckts
Model 18	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts, Sbbwdpckt
Model 19	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd, Tbpckts
Model 20	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd, Bistd

Table A.1.4. Overview of Relief-F-based prediction models - Part 4

Models	Predictor Variables
Model 21	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd, İstd
Model 22	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl, Fplstd
Model 23	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc, Fwdhl
Model 24	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1, Plvrnc
Model 25	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax, Fwdhl.1
Model 26	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf, Bimax
Model 27	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts, adpf
Model 28	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt, Tfpckts
Model 29	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin, Sbfwdpckt
Model 30	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf, Bimin

Table A.1.5. Overview of Relief-F-based prediction models - Part 5

Models	Predictor Variables
Model 31	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin, mssf
Model 32	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal, Mplmin
Model 33	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts, Bitotal
Model 34	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd, Tlfpckts
Model 35	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd, Bplstd
Model 36	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb, Fistd
Model 37	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean, İwbb
Model 38	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin, Plmean
Model 39	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean, İmin
Model 40	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt, İmean
Model 41	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd, Sbfwdbyt
Model 42	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd, Plstd
Model 43	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax, Fwdistd
Model 44	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax, İmax

Table A.1.6. Overview of Relief-F-based prediction models -Part 6

Models	**Predictor Variables**
Model 45	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax, Bplmax
Model 46	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean, Fwdimax
Model 47	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax, Fwdimean
Model 48	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio, Fimax
Model 49	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal, Duratio
Model 50	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn, Fwditotal
Model 51	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc, Fdrtn
Model 52	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg, Rdtfc
Model 53	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc, Fpflg
Model 54	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts, Cwefc
Model 55	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf, Fwdpckts
Model 56	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts, İwbf
Model 57	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc, Fpckts
Model 58	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl, Urgfc
Model 59	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax, Maxpl
Model 60	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport, Fplmax
Model 61	Ackfc, Prtcl, Fplmin, Minpl, Fbytes, Dport
Model 62	Ackfc, Prtcl, Fplmin, Minpl, Fbytes
Model 63	Ackfc, Prtcl, Fplmin, Minpl
Model 64	Ackfc, Prtcl, Fplmin
Model 65	Ackfc, Prtcl

Appendices A.2. mRMR Based Prediction Models

Each model has unique features. The tables between Table A.1.1 and Table A.1.6 provide an overview of mRMR based prediction models, showing each one along with its unique prediction variables.

Table A.2.1. Overview of mRMR-based prediction models - Part 1

Models	**Predictor Variables**
Model 66	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin, Bplstd, İmax, Bwdhl, Sbfwdpckt, Actmax, Tfpckts, Actmean, Fwdimax, Bistd, Duratio, Fwdimean, Actstd, Fwditotal, İmin
Model 67	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin, Bplstd, İmax, Bwdhl, Sbfwdpckt, Actmax, Tfpckts, Actmean, Fwdimax, Bistd, Duratio, Fwdimean, Actstd, Fwditotal
Model 68	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin, Bplstd, İmax, Bwdhl, Sbfwdpckt, Actmax, Tfpckts, Actmean, Fwdimax, Bistd, Duratio, Fwdimean, Actstd
Model 69	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin, Bplstd, İmax, Bwdhl, Sbfwdpckt, Actmax, Tfpckts, Actmean, Fwdimax, Bistd, Duratio, Fwdimean
Model 70	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin, Bplstd, İmax, Bwdhl, Sbfwdpckt, Actmax, Tfpckts, Actmean, Fwdimax, Bistd, Duratio

Table A.2.2. Overview of mRMR-based prediction models - Part 2

Models	Predictor Variables
Model 71	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin, Bplstd, İmax, Bwdhl, Sbfwdpckt, Actmax, Tfpckts, Actmean, Fwdimax, Bistd
Model 72	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin, Bplstd, İmax, Bwdhl, Sbfwdpckt, Actmax, Tfpckts, Actmean, Fwdimax
Model 73	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin, Bplstd, İmax, Bwdhl, Sbfwdpckt, Actmax, Tfpckts, Actmean
Model 74	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin, Bplstd, İmax, Bwdhl, Sbfwdpckt, Actmax, Tfpckts
Model 75	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin, Bplstd, İmax, Bwdhl, Sbfwdpckt, Actmax
Model 76	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin, Bplstd, İmax, Bwdhl, Sbfwdpckt
Model 77	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin, Bplstd, İmax, Bwdhl

Table A.2.3. Overview of mRMR-based prediction models - Part 3

Models	Predictor Variables
Model 78	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin, Bplstd, İmax
Model 79	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin, Bplstd
Model 80	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean, Mplmin
Model 81	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt, İmean
Model 82	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts, Sbbwdbyt
Model 83	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd, Fpckts
Model 84	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts, Fplstd
Model 85	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax, Tlbpckts
Model 86	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax, Fimax

Table A.2.4. Overview of mRMR-based prediction models - Part 4

Models	Predictor Variables
Model 87	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts, Bplmax
Model 88	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss, Fpckts
Model 89	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd, Abwdss
Model 90	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd, Fistd
Model 91	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport, İstd
Model 92	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin, Dport
Model 93	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc, Fwdimin
Model 94	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean, Plvrnc
Model 95	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal, Bplmean
Model 96	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn, Bitotal
Model 97	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd, Fdrtn

Table A.2.5. Overview of mRMR-based prediction models - Part 5

Models	Predictor Variables
Model 98	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax, Plstd
Model 99	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl, Bimax
Model 100	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts, Fwdhl
Model 101	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc, Tlfpckts
Model 102	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts, Rdtfc
Model 103	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts, Fwdpckts
Model 104	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf, Tbpckts
Model 105	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts, adpf
Model 106	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg, Bwdpckts
Model 107	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean, Fpflg
Model 108	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt, Bimean
Model 109	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1, Sbfwdbyt
Model 110	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb, Fwdhl.1
Model 111	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax, İwbb
Model 112	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax

Table A.2.6. Overview of mRMR-based prediction models -Part 6

Models	**Predictor Variables**
Model 112	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd, Fplmax
Model 113	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss, Fwdistd
Model 114	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin, Afwdss
Model 115	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks, Fimin
Model 116	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl, Avrpcks
Model 117	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl, Prtcl
Model 118	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt, Maxpl
Model 119	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin, Sbbwdpckt
Model 120	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean, Fplmin
Model 121	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf, Plmean
Model 122	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl, İwbf
Model 123	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes, Minpl
Model 124	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc, Fbytes
Model 125	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf, Urgfc
Model 126	Fplmean, Cwefc, Actmin, Ackfc, Bimin, mssf
Model 127	Fplmean, Cwefc, Actmin, Ackfc, Bimin
Model 128	Fplmean, Cwefc, Actmin, Ackfc
Model 129	Fplmean, Cwefc, Actmin
Model 130	Fplmean, Cwefc