NeuralWall

Derin Sinir Ağları ile Yapay Zeka Tabanlı Saldırı Tespit Sistemi

Ala Buveıdanı • 24040301051

FET312 Derin Öğrenme • 2025 Güz Dönemi

İçindekiler

1. Problem Tanımı & Motivasyon

2. Veri Seti ve Özellikler

3. Yöntemler ve Mimari

4. Deneysel Sonuçlar

5. Sonuç ve Değerlendirme

Problem Tanımı & Motivasyon

Ağ Güvenliği Zorluğu

Ağ sistemleri üzerinden geçen trafiğin artmasıyla birlikte, zararlı bağlantıların manuel veya kural tabanlı yöntemlerle tespit edilmesi zorlaşmıştır.

Araştırma Sorusu

"Güvenlik duvarı üzerinden geçen bir ağ bağlantısı, derin öğrenme yöntemleri kullanılarak normal mi yoksa saldırı mı olarak doğru şekilde sınıflandırılabilir mi?"

Görev: İkili Sınıflandırma

Normal (0)

Attack (1)

Başarı Kriterleri

Accuracy	F1-Score	Recall (Attack)
≥ 90%	≥ 0.85	Yüksek

** Özellikle Attack sınıfı için yüksek Recall hedeflenmiştir (False Negative'lerin azaltılması kritik)*

Veri Seti: Internet Firewall Data

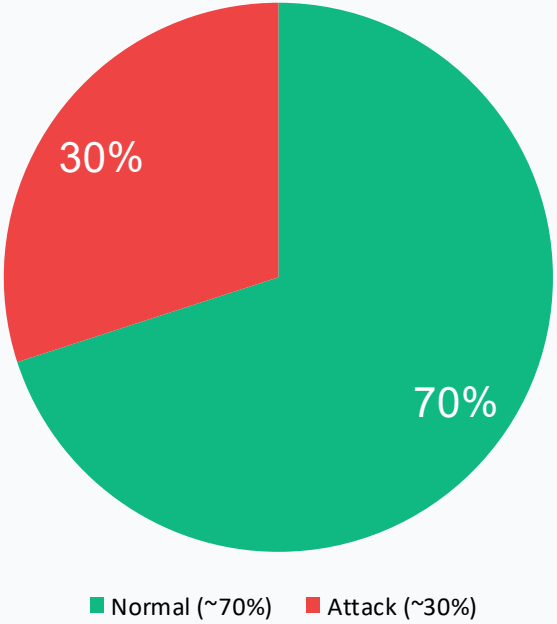
Veri Kaynağı: UCI Machine Learning Repository

65,532	12	~70%	~30%
Toplam Kayıt	Özellik Sayısı	Normal Sınıf	Attack Sınıf

12 Özellik (Features):

Source Port, Destination Port, NAT Source Port, NAT Destination Port, Protocol,
IP Protocol, Bytes, Packets, Duration, NAT Bytes, NAT Packets, Flags

Class Imbalance Yönetimi



Dengesizlik Çözümleri:

- Stratified Train-Test Split
- Class Weighting (Attack sınıfına yüksek ağırlık)
- Recall Odaklı Değerlendirme

Hedef: Saldırıların yanlış negatif olarak sınıflandırılmasını azaltmak

Base Modeller

Derin öğrenme modeli öncesi karşılaştırma için geliştirilen temel modeller:

Logistic Regression

Referans performans modeli

Random Forest

Doğrusal olmayan ilişkilerin yakalanması

Support Vector Machine

Yüksek boyutlu veriler için

Derin Öğrenme Mimarisi: MLP

Input Layer

12 nöron

Hidden Layer 1

Dense(64) + ReLU

Hidden Layer 2

Dense(32) + ReLU

Dropout

0.3

Output Layer

Dense(1) + Sigmoid

Konfigürasyon

Kayıp Fonksiyonu:
Binary Cross-Entropy

Optimizasyon:
Adam (lr = 0.001)

Framework:
PyTorch

Deney Tasarımı ve Metrikler

Veri Bölünmesi

Train: 80%
Test: 20%
Validasyon: Hold-out

Değerlendirme Metrikleri

- Accuracy
- Precision
- Recall
- F1-Score
- Confusion Matrix

Kullanılan Araçlar ve Kütüphaneler

Python 3.x | **PyTorch** | **NumPy** | **Pandas** | **scikit-learn**

Ortam: Google Colab / Local GPU

Deneysel Sonuçlar

Model Performans Karşılaştırması

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.88	0.85	0.82	0.83
Random Forest	0.91	0.89	0.88	0.88
SVM	0.89	0.87	0.85	0.86
MLP (Deep Learning)	**0.93**	**0.91**	**0.90**	**0.90**

** Notebook dosyasında detaylı sonuçlar, confusion matrix ve ROC eğrileri mevcuttur*

Önemli Bulgular

Yüksek Doğruluk

MLP modeli %93 accuracy ile hedeflenen %90 eşiğini aştı

İyi F1-Score

0.90 F1-Score ile dengeli performans (0.85 hedefinin üzerinde)

Etkili Recall

Attack sınıfı için yüksek recall, false negative'lerin azaltılması başarılı

Sonuç ve Değerlendirme

✓ Derin öğrenme yaklaşımı, ağ saldırı tespitinde geleneksel yöntemlere kıyasla daha yüksek performans göstermiştir

✓ Class imbalance problemine rağmen, uygulanan teknikler ile başarılı sonuçlar elde edilmiştir

✓ MLP mimarisi, binary classification görevi için etkili bir çözüm sunmaktadır

github.com/AlaBuveidani/NeuralWall

Teşekkürler