**KONU:** Kötü Amaçlı URL'ler Projesi  
**Proje Sahibi:** AISECLAB  
**Proje Danışmanı:** Elif PARLAK  
**HAZIRLAYAN:** Halil Berkay Şahin

**Kötü Amaçlı URL Tespiti Projesi**

Siber güvenlik, günümüz dijital dünyasında giderek önem kazanan bir konu haline geldi. Bu proje kapsamında, kötü amaçlı URL'lerin tespiti üzerine çalıştık ve bu konuda başarılı sonuçlar elde ettik. İşte proje detayları:

**Veri Seti ve Başlangıç**

Çalışmamızın temelini, "Malicious URLs Dataset" oluşturdu. Bu veri seti, toplamda 651.191 adet URL içeriyor ve şu türlere ayrılmış durumda:

* **Benign (Zararsız)**
* **Defacement (Tahrifat)**
* **Phishing (Kimlik Avı)**
* **Malware (Kötü Amaçlı Yazılım)**

Bu veri seti, kötü niyetli yazılımların ve tehditlerin izlenmesi ve sınıflandırılması için mükemmel bir kaynak sundu.

**Veri İşleme ve Temizlik**

Projemizde verilerin işlenmesi ve temizlenmesi önemli bir aşamayı oluşturdu. Şu adımları takip ettik:

1. **Veri Setini İnceleme:**  
   Pandas kütüphanesi ile veri setini inceledik ve head() komutu ile ilk 10 satıra göz attık. info() fonksiyonunu kullanarak sütunların yapısı ve eksik veri durumunu analiz ettik.
2. **Veri Temizleme:**
   * Tekrarlayan satırları kaldırdık.
   * Eksik verilerin analizini tamamladık ve veri setini temiz hale getirdik.

Bu aşamalar, modelin performansını artırmak için önemli bir temel oluşturdu.

**Modelleme ve Performans Değerlendirme**

Veri işleme tamamlandıktan sonra, verilerimizi eğitim ve test kümelerine ayırdık. Eğitim için Logistic Regression algoritmasını tercih ettik. Modeli şu şekilde uyguladık:

1. **TF-IDF Özellik Çıkarımı:**  
   Metin verilerinden özellik çıkarımı için TF-IDF vektörleştirme yöntemi kullanıldı.
2. **Eğitim ve Test:**  
   Eğitim verileri ile model eğitildi ve test verileri üzerinde doğruluğu ölçüldü.

**Elde Edilen Sonuçlar:**  
Modelimiz, %94.64 doğruluk oranı ile kötü amaçlı URL'lerin tespitinde başarılı bir performans sergiledi.

**Sonuç**

Bu proje ile kötü amaçlı URL tespiti konusunda değerli bilgiler elde ettik. Veri analizi, temizlik ve modelleme süreçlerini başarıyla uygulayarak, yüksek doğruluk oranına ulaştık.

Proje, siber güvenlik alanında etkili bir araç olabilecek önemli bir adım sunuyor. Dijital dünyada güvenliği artırmaya yönelik bu tür projeler, gelecekteki tehditlerin önlenmesi için kritik öneme sahiptir.