

**Seydişehir Ahmet Cengiz Mühendislik Fakültesi**

**Bilgisayar Mühendisliği Bölümü**

**… Dersi Proje Formu**

|  |
| --- |
| **Proje Başlığı** |
| Makine Öğrenmesi Tabanlı Saldırı Tespit Sistemi Geliştirilmesi |

|  |
| --- |
| **Öğrenci Bilgileri** |
| **Öğrenci No** | … |
| **Öğrenci Ad Soyad** | … |

|  |
| --- |
| **Danışman Ünvan** **Danışman Ad Soyad** |

**Aralık 2021**

**Konya**

|  |
| --- |
| **Proje Raporu Hazırlama Kuralları** |
| Her öğrenci bitirme projesini aşağıda belirlenen kurallara göre hazırlayacaktır.**1.**       **Rapor Bölümleri**         Kapak         İçindekiler         Projenin Önemi         Projenin Amacı          Projenin Hedefi          Projenin iş-zaman çizelgesi         Projede kullanılan donanımlar ve yazılımlar ile ilgili bilgileri.         Projenin yapım aşamaları.         Kaynaklar**2.**       **Konu Anlatımı**         Etik kurallara uygun olarak, öğrenci konuyu kendi cümleleri ile sade bir şekilde anlatmalıdır. Yararlanılan kaynaklar belirtilmelidir.**3.**       **Sayfa Düzeni**         Kenar Boşlukları: 2.5 cm         Sayfa Numarası: Sağ Alt Köşede Projenin Önemi sayfasından itibaren başlamalıdır.**4.**       **Metin Özellikleri**         Paragraf Girintisi Yok         Metin iki yana yaslı         Tek satır aralıklı         12 Punto ve Normal         Paragraflar arasında 1 boşluk         Ana Başlıklar ve alt Başlıklar: 12 Punto ve Kalın         Tüm Metin Fontu: Times New Roman**5.**       **Şekiller ve Tablolar**         Şekiller sayfa içerisinde ortalı olmalıdır.         Şekil açıklama metni şekil altında tek satıra sığıyorsa ortalı, sığmıyorsa iki yana yaslı sol kenara yaslı olmalıdır.         Tablolar sayfa içerisinde sola yaslı olmalıdır         Tablo açıklama metni tablo üstünde ve sola yaslı olmalıdır.**6.**       **Kaynaklar**         Bitirme projesinde hazırlanmasında yararlanılan kaynaklar, metin içerisinde kullanım sırasına göre rapor sonunda ve “Kaynaklar” başlığı altında sıra numaraları verilerek listelenecektir. Proje çalışmalarınızda başarılar… |

**İçindekiler**

[1. Proje Konusunun Önemi ve Özgün Değeri 4](#_Toc51147498)

[2. Projenin Amacı ve Hedefi 4](#_Toc51147499)

[3. Projenin iş-zaman çizelgesi 5](#_Toc51147500)

[4. Projede kullanılan donanımlar ve yazılımlar ile ilgili bilgileri 6](#_Toc51147501)

[4.1. Donanımlar 6](#_Toc51147502)

[4.2. Yazılımlar 6](#_Toc51147503)

[5. Projenin yapım aşamaları 6](#_Toc51147504)

[Kaynaklar 6](#_Toc51147505)

# Proje Konusunun Önemi ve Özgün Değeri

**Saldırı Tespit Sistemi Nedir?**

Günümüzde bilgisayar ağların karmaşık bir yapıda olması, başta internet üzerinden olmak üzere bilgisayar ağlarının birbirleriyle erişim noktaları üzerinden bağlantıların olması, siber tehditlerin/saldırıların gün geçtikçe çeşitliliğin artması ve bunun yanında karmaşık yapıda olan bilgisayar ağlarının güvenlik duvarı veya veri şifreleme gibi klasik güvenlik teknikleri ile korunamayacağının ortaya çıkması bilgisayar ağlarındaki trafiğin sürekli olarak izlenip siber tehditlerin/saldırıların gerçek zamanlı olarak belirlenmesini kaçınılmaz kılmıştır. Saldırı Tespit Sistemi (STS) gelişen teknoloji ile birlikte önemli bir rol oynayarak günümüz sistem ve ağ güvenliğini kolaylaştıran yeni bir savunma mekanizmasıdır. Bir STS, bilgisayar ağlarını sık sık izlemek, olası siber tehditleri/saldırıları tanımlamak ve bunlarla ilgili olayların kayıtlarını tutmak, tehditleri/saldırıları durdurmak ve siber güvenlik ile ilgili çalışanlara raporlamak gibi işlevlere sahiptir. Bir STS aynı zamanda ağ keşfi yapan saldırganları algılayarak saldırının erken aşamada tespitini gerçekleştirebilir.

**Özgün Değer**

Siber güvenlik bilgisayar sistemlerinin, bilgisayar ağlarının ve yazılımların dijital saldırılara karşı korunmasıdır. Bu siber saldırılar genellikle hassas bilgilere erişmeyi, değiştirmeyi veya imha etmeyi amaçlar; kullanıcılardan para çekmek, kullanıcıların mahrem bilgilerine ulaşmak veya normal iş süreçlerini kesintiye uğratmak siber saldırılara örnek olarak verilebilir. Gelişen teknoloji, internet ortamında kullanılan cihaz sayısının gün geçtikçe artması ve saldırganların sürekli kendilerini geliştirerek taktik değiştirmelerinden dolayı günümüzde etkili bir siber güvenlik önleminin uygulanması oldukça zordur. Güvenlik duvarları, güvenlik tarayıcıları, veri şifreleme ve kullanıcı kimlik doğrulaması gibi geleneksel güvenlik teknikleri siber güvenliğin sağlanmasında yetersiz kalmaktadır. Geleneksel sistemlerin yanında virüsler, servis reddi saldırıları, port taramaları, tampon taşmaları ve CGI atakları gibi güvenlik ihlallerini otomatik olarak tespit eden akıllı STS’lerin kullanılması siber güvenliğin daha etkili ve efektif olarak uygulanmasını sağlamaktadır. Makine öğrenmesi ve veri madenciliği algoritmaları günümüzde sağlık, bankacılık, ticaret, askeriye olmak üzere hemen hemen bütün alanlarda uygulanmaktadır. Sonra yıllarda, siber güvenlik ile alakalı akıllı STS modeli geliştirilmesinde yapılan birçok çalışma bulunmaktadır. Gerçekleştirilecek araştırmanın özgün yönü veri madenciliği veri önişleme tekniklerini de kullanarak makine öğrenmesi tabanlı çevrimiçi çalışabilen bir akıllı STS modeli geliştirmektir.

# Projenin Amacı ve Hedefi

Araştırma kapsamında literatürde bulunan siber güvenlik veri kümeleri ve veri madenciliği veri önişleme teknikleri kullanılarak makine öğrenmesi tabanlı bir STS modeli geliştirilecektir. Geliştirilecek olan STS modelinin literatürde önerilen STS modellerine göre daha yüksek doğruluk oranına sahip, hızlı, güçlü ve çevrimiçi olarak çalışabilen olması hedeflenmektedir. Araştırma sonrasında Türkiye’ye dönülmesine mutabık geliştirilen STS modeli milli teknolojimize önemli bir katkı sağlayacak olan milli bir ürün haline dönüştürülmesi ve ülkemize uluslararası pazarda önemli bir nitelik kazandırması için gerekli proje başvurularının yapılması amaçlanmaktadır.

# Projenin iş-zaman çizelgesi

Projede yer alacak başlıca iş paketleri (İP) ve hedefleri, her bir iş paketinin hangi sürede gerçekleştirileceği, başarı ölçütü ve projenin başarısına katkısı “İş-Zaman Çizelgesi” doldurularak verilir.

İş-zaman çizelgesi

|  |  |  |  |
| --- | --- | --- | --- |
| İP.No. | İş Paketlerinin Adı ve Hedefleri | Zaman Aralığı (Haftalık) | Başarı Ölçütü  |
| 1 | **Siber Güvenlik Veri Setlerinin Hazırlanması:** Sınıflandırma başarısı yüksek bir STS geliştirmek için literatürde sıklıkla kullanılan 3 farklı veri seti kullanılacaktır. Bu iş paketinde, geliştirilecek olan STS yapısında veri setlerinin direk olarak veri önişleme aşamasına giriş olarak alınabilmesi için hepsinin tek bir standarda çevrilmesi hedeflenmektedir.  | 1. hafta | Veri setleri için ortak bir format belirlenmesi ve araştırma aşamasında kullanılacak olan veri setlerinin belirlenen ortak formata dönüştürülmesi veri önişleme aşaması ve daha sonrasında gerçekleştirilecek makine öğrenmesi algoritmalarının eğitimi için önem arz etmektedir. Siber güvenlik veri setlerinin hazırlanması iş paketinin araştırmanın başarısına katkısı %10 oranındadır. |
| 2 | **Veri Normalizasyonu:** Kullanılacak olan veri kümelerinde kesikli nitelikler olabileceği gibi sürekli niteliklerde bulunmaktadır. Makine öğrenmesi aşamasına geçmeden önce ortak formata dönüştürülmüş olan veri setlerinin sürekli değerleri literatürde sıklıkla kullanılan normalizasyon yöntemleri ile belirli aralıklarda normalize edilecektir.  | 2. ve 3. haftalar | Siber güvenlik veri kümelerinde farklı tiplerde ve faklı değer aralıklarda nitelikler bulunmaktadır.. Sağlam ve başarılı bir STS sistemi geliştirebilmek için eğitim ve test aşamasında kullanılacak olan verilerin istenen ve uygun formatta gelmesi büyük önem arz etmektedir. Bu sebeplerden dolayı veri normalizasyonu iş paketinin araştırmanın başarısına katkısı %10 oranındadır. |
| 3 | **Veri Ayrıklaştırma:** Veri ayrıklaştırma aynı sınıfa ait sürekli değerleri belirli aralık değerlerine dönüştürerek makine öğrenmesi algoritmalarının sınıflandırma başarısını arttırır. Literatürde sıklıkla kullanılan veri ayrıklaştırma yöntemleri kullanılarak veri setleri daha anlamlı hale çevrilmesi hedeflenmektedir. | 4., 5. ve 6. haftalar | Siber güvenlik veri kümelerinde nitelikler genellikle sürekli değerlere sahiptir. Bu sürekli nitelikler makine öğrenmesi algoritmalarının başarısını etkilemektedir. Bundan dolayı veri ayrıklaştırma iş paketinin araştırmanın başarısına katkısı %10 oranındadır. |
| 4 | **Özellik seçimi:** Kullanılacak olan veri kümelerinde faklı atak tipleri bulunmaktadır. Her atak tipinin sınıflandırılmasında etkili olan öznitelikler birbirinden farklı olabilmektedir. Bu öznitelikler hem makine öğrenmesi algoritmasının çalışma zamanını yükseltmekte hem de makine öğrenmesi algoritmasının gereksiz veriler öğrenmesine ve başarısının düşmesine sebep olabilmektedir. Bu sebeple literatürde kullanılan özellik seçme algoritmalarının kullanılarak geliştirilecek olan STS sisteminin daha hızlı ve başarılı eğitilebilmesi için uygun veri seti alt kümelerinin tespit edilmesi ve elde edilen alt kümelere göre öznitelik vektörlerinin oluşturulması hedeflenmektedir. | 7., 8. ve 9. haftalar | Araştırma esnasında geliştirilecek olan STS sisteminin başarısı veri önişleme işlemlerinin başarısı ile doğru orantılıdır. Özellik seçimi işlemi hem veri setlerini basitleştirirken hem de makine öğrenmesi algoritmasının hızlı ve başarılı çalışmasını sağlamaktadır. Özellik seçme işlemi sonrasında yapılacak makine öğrenmesi eğitimleri ile hangi saldırı tipleri için hangi öznitelik altkümelerinin kullanılabileceği belirlenecek ve böylelikle her bir saldırı için aynı öznitelikleri kullanmak yerine faklı saldırı tipleri için farklı öznitelikleri kullanan esnek ve modüler bir STS sistemi geliştirilmesi araştırmanın başarısını arttıracaktır. Bu sebeplerden dolayı özellik seçimi iş paketinin araştırmanın başarısına katkısı %20 oranındadır. |
| 5 | **Makine Öğrenmesi:** Veri işleme aşamasından sonra elde edilen öznitelik vektörleri makine öğrenmesi algoritmalarını eğitmek için kullanılacak ve çapraz doğrulama yöntemi ile öğrenme ve doğrulama başarımlarının maksimize edildiği yöntem ve parametreleri tayin edilmesi hedeflenmektedir.  | 10., 11. ve 12. haftalar | Belirli bir dünya probleminin makine öğrenmesi algoritmaları ile çözümü için uygun algoritmanın belirlenmesi kadar algoritmanın eğitim kümesinin hazırlanması ve algoritma parametrelerinin de optimize edilmesi büyük önem arz edilmektedir. Makine öğrenmesi iş paketinde STS modelinin oluşturulmasında kullanılacak olan makine öğrenmesi algoritmaları ile benchmark testlerinin yapılması araştırmanın başarıya ulaşması için gereklidir. Bu sebeplerden dolayı makine öğrenmesi iş paketinin araştırmanın başarısına katkısı %35 oranındadır. |
| 6 | **Doğrulama ve Test:** Eğitim sürecinde dışarıda tutulan ve eğitim sürecine dâhil edilmeyen veriler sistemin doğrulanması için kullanılmalı ve doğruluk başarımı gün mertebesine yaklaşılacak şekilde artırılması hedeflenmektedir. Ayrıca araştırma yapılacak kuruluşun laboratuvarlarında kurulacak küçük bir bilgisayar ağında gerçek zamanlı deneme yapılması planlanmaktadır. | 13. ve 14. haftalar | Geliştirilecek olan STS’in bu iş paketinde eğitim aşamasında kullanılmamış olan veriler ile testi yapılacaktır. Test işlemi sonucunda elde edilen doğruluk oranları araştırma projesinin başarısını belirleyecektir. Bu sebeplerden dolayı doğrulama ve test iş paketinin araştırmanın başarısına katkısı %15 oranındadır. |

# Projede kullanılacak yöntem, donanımlar ve yazılımlar ile ilgili bilgiler

Proje yapımı esnasında kullanılacak yöntemler, algoritma ve donanımlar burada yazılacaktır.

# Projenin yapım aşamaları

Haftalık yapılanlar burada yazılacaktır. Vize sınavına ve final sınavına gelmeden önce geçmiş haftalar doldurulmuş şekilde getirilecektir.

1. **Hafta:** İlgili haftaya ait projede gerçekleştirilenler buraya yazılacaktır. Bu hafta için proje konusunun belirlenmesi ve iş planının çıkarılması yazılabilir.
2. **Hafta:** İlgili haftaya ait projede gerçekleştirilenler buraya yazılacaktır.
3. **Hafta:** İlgili haftaya ait projede gerçekleştirilenler buraya yazılacaktır.
4. **Hafta**: İlgili haftaya ait projede gerçekleştirilenler buraya yazılacaktır.

….

….

…

**14. Hafta:** İlgili haftaya ait projede gerçekleştirilenler buraya yazılacaktır.

# Kaynaklar