TÜBİTAK–2209-B SANAYİYE YÖNELİK LİSANS ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI

ARAŞTIRMA ÖNERİSİ FORMU

2022 Yılı

2. Dönem Başvurusu

2209/B SANAYİYE YÖNELİK LİSANS ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI ARAŞTIRMA ÖNERİSİ FORMU

A. GENEL BİLGİLER

Araştırma Önerisinin Başlığı: Yapay Zekâ Tabanlı Sistemlerin Kalitesinin Ölçülmesi İçin Test Kütüphanesi Geliştirilmesi (Temiz Yapay Zekâ)
Başvuru Sahibinin Adı Soyadı: Osman Çağlar
Akademik Danışmanın Adı Soyadı: Dr. Öğr. Üyesi Uğur Yayan
Sanayi Danışmanının Adı Soyadı: Cem Bağlum
Araştırmanın Yürütüleceği Kurum/Kuruluşlar: ESKİŞEHİR OSMANGAZİ ÜNİVERSİTESİ & İNOVASYON MÜHENDİSLİK TEKNOLOJİ GELİŞTİRME DANIŞMANLIK SANAYİ VE TİCARET LİMİTED ŞİRKETİ

ÖZET

Özet

Günümüz dünyasında Yapay Zekâ (YZ) tabanlı sistemlerin kullanımı sağlıktan otomotiv sektörüne, savunma sanayiden eğitim gibi bir çok kritik alanda gün geçtikçe artmaktadır. Kritik birçok alanda kullanılması ile YZ tabanlı sistemlerin kalitesini ve güvenilirliğini ölçmek önemli hale gelmiştir. Mevcut YZ modellerinin geliştirme sürecinde Derin Sinir Ağı (DSA) modellerine oldukça sık rastlanmakta fakat DSA'nın yapısı, güvenliği ve kalite kontrolü gibi metriklerin analizini yapan araçlara aynı sıklıkla rastlanamamaktadır. Geliştireceğimiz projemizin baz aldığı motivasyon da budur. Projemizin amacı DSA modeli ile geliştirilen ürünler için modelin yapısal analizini, kalitesini ve güvenilirliğini ölçmek için kapsam parametrelerini ölçen beyaz kutu test kütüphanesi oluşturmaktır. Beyaz kutu test kütüphanesi ise geliştirilmiş modellerin iç yapısını inceleyerek belirlenen 9 adet kapsam test metotlarını uygulamaktadır. Böylece modelin gerekli analizleri sağlanacak ve geliştiricilere rapor halinde çıktı üretecektir. DSA modelinin analizi yapılırken geleneksel test metotlarından farklı olarak DSA modelinin kapsam kriterlerini ölçerek yeni test yaklaşımı ile çalışılmaktadır. Geliştirilen bu kütüphane ile çok çeşitli alanlarda kullanılan YZ tabanlı sistemlerde bulunan DSA modellerinin testinin kolaylığı sağlanacaktır. Proje kapsamında geliştirilecek kütüphane ile YZ tabanlı sistemlerin kalitesi artırılacak ve güvenlik / emniyet kritik modellerin testinin kolaylığı sayesinde firmalarda ve sanayide çalışan YZ geliştiricileri açısından zaman ve iş gücü maliyetlerinden tasarruf sağlanacaktır.

Proje sonucunda geliştirilecek Temiz Yapay Zekâ Test kütüphanesi sayesinde sanayi proje geliştiricilerine geliştirmiş oldukları modellerin test kolaylığı sağlanacaktır. Böylelikle zaman maliyeti, işgücü maliyeti vb. maliyetlerden tasarruf sağlanacaktır. Geliştirilen modellerin güvenliğinin ve kalitesinin sağlanması, bu modelleri geliştiren AR-GE firmalarına güven duyulmasına yol açacaktır. Bu sayede bu firmaların YZ tabanlı sistemlerin geliştirildiği pazarda daha yüksek pay almalarına sebep olacaktır.

Geliştirilen test metodunun geleneksel test metotlarından farklı olmasından dolayı bizden sonra yapılacak kapsamlı akademik çalışmalara yol gösterici olup makale, araştırma veya bildiri yayınlanmasını sağlayacaktır.

Anahtar Kelimeler: Güvenilir Yapay Zekâ, Derin Sinir Ağı Yapı Analizi, Açıklanabilir ve Hesap Verilebilir Yapay Zekâ, Yapay Zekâ Modellerinin Testi, Derin Sinir Ağlarında Kapsam Analizi

1. AMACI, YENİLİKÇİ YÖNÜ ve TEKNOLOJİK DEĞERİ

1.1. Projenin amacı

Yazılım testi, bir projenin geliştirilme aşamasında eklenen yeni özelliklerin proje başlangıcı esnasında belirtilen gereksinimler için yeterli olup olmadığını, yazılımın herhangi bir yanlış işlem yaparak hatalı davranıp davranmadığını, yazılımın beklenen performansı verip vermediğini belirtmeye yarayan ve yazılımın kusurlarını ortaya çıkarmaya imkân tanıyan önemli bir kavramdır. Bir yazılım geliştirme projesi sürecinde yazılım geliştirme yaşam döngüsünün herhangi bir aşamasında hataların meydana gelebileceği unutulmamalıdır. Son aşamada oluşturulan ürünün fonksiyonel olarak birtakım arızalara sahip olabileceği ve performans açısından da maksimum verimliliği veremeyebileceği bilinmektedir. Yazılım testindeki başlıca motivasyon, bu hataların yazılım geliştirilmesi sürecinde erken evrelerde farkına varılıp maliyet ve zaman kaybına uğramadan düzeltilmesini hedeflemektir.

YZ tabanlı sistemler günümüz teknolojisinde otomasyon sistemleri, tıbbi alanlar, karmaşık problemlerin çözümü ve hata analizi gerçekleştirme gibi çeşitli alanlarda kullanılmaktadır. YZ tabanlı sistem bulunduran yazılımlar, genellikle emniyet açısından kritik alanlarda kullanılması nedeniyle güvenlik açısından çok daha fazla önem

taşımaktadır. Bu sebeple oluşturulacak olan YZ modellerinin testinin yapılmasında kullanılacak olan araçlar önem arz etmektedir. (ELECTRONIC COMPONENTS AND SYSTEMS, 2022)

Yazılım testlerinin geleneksel yazılımın geliştirme sürecinde ciddi önem oluşturmasının yanı sıra YZ barındıran sistemlerin geliştirilmesi sürecinde de oldukça önemlidir. Geleneksel yazılım, yazılım geliştirici tarafından belirlenmiş kurallar çerçevesinde ortaya çıkmaktadır. Ancak YZ tabanlı bir yazılımın geliştirme süreci de aynı benzerliğe sahip olsa da bir modelin eğitim süreci ile gerçekleşen ve oluşan milyonlarca parametreyi de içermektedir. YZ tabanlı yazılımlar, yazılım geliştiricileri tarafından belirlenmiş kurallara göre değil, modelin veri setine göre eğitilmesi ile kuralları oluşturmaktadır. YZ tabanlı yazılım sistemlerinin geleneksel yazılımlara göre farklılık içermesi sebebiyle YZ'nin testi kavramı ortaya çıkmıştır. Bu kavram derin sinir ağlarının (DSA) içerdiği yapıların analizinin yapılması yoluyla gerçekleştirilmektedir. YZ yazılımının çıktılarının güvenilirliğini ve kalitesini denetlemek de oldukça önemli bir olgudur. Yaptığımız araştırmaların sonucunda YZ'nin güvensizliği konusunda ve bu konunun getirmiş olduğu etik kaygılar hakkında çeşitli yazılar bulunmaktadır. Ancak bu güvenilirlik kavramını da test edebilecek herhangi bir çalışma literatürde yer almamaktadır. Projemizin temelinde YZ'nin testini gerçekleştirerek YZ'nin güvenirliğini test etmede kullanılan en önemli parametrelerden olan adil karar verebilme, güvenirlik, açıklanabilirlik ve hesap verilebilirlik gibi unsurların kısacası YZ emniyetinin daha da iyileştirilmesi hedeflenmiştir. Bu projedeki temel amacımız YZ'yi test eden ve bu sayede de YZ sistemlerinin emniyetini sağlayan bir yazılım kütüphanesi oluşturmaktır.

Ekibimiz tarafından yapılan literatür çalışmaları sonucunda YZ gibi kritik sistemlerin oluşturmuş olduğu endişeleri azaltacak çözüm yöntemleriyle karşılaşılmıştır. Projemizin ana amacı YZ modelinin oluşturulma aşamasında veya oluşturulan YZ'nin test edilmesi esnasında DSA modelinin yapısının analizini gerçekleştirmektir. Bu analizin gerçekleştirilmesindeki temel amaç DSA yapısındaki katmanlarda bulunan nöronların, YZ'nin karar vermesinde ne kadar etkili olduğunu göstermektir. Ayrıca analizler sonucunda DSA modelinin yapısı tekrardan gözden geçirilerek anlaşılabilir, performans açısından daha etkili ve daha güvenilir bir DSA modelinin oluşturulması sağlanacaktır. Projemizin ana başlığında bulunan "Temiz Yapay Zekâ" kavramı ise bu doğrultuda hedeflenen motivasyon sonucunda ortaya çıkmıştır.

DSA'lar birbirine bağlı birtakım nöronlardan oluşan, karar verme ve öğrenme evrelerini gerçekleştiren insan beyninden esinlenerek oluşturulmuş yapılardır. Temiz Yapay Zekâ ise bu tür sistemlerin testini ve yapısal analizini gerçekleştirmek üzere geliştirilmesi hedeflenen bir projedir.

Bu projenin sonunda elde etmek istediğimiz 4 temel unsur bulunmaktadır:

1. **DSA Modelinin Yapısal Kapsam Kriterlerinin Ölçülmesi:** DSA içeren sistemlerde modelin güvenilirliğini ve kalitesini ölçmeye yardımcı olan kapsam kriterlerinin ölçülmesi birinci hedefimizdir. Temiz kapsam kriterlerinin ölçülmesi, farklı test girdi verileri ile girdi ve çıktı katmanları dışında kalan gizli katmanlar arasındaki her bir öznitelik çifti için 4 temel test kapsama yöntemini uygulamaktır. Bu testler ise bir öznitelik çifti olan $(\psi_{k,i}, \psi_{k+1,j})$, k ve k+1 komşu katmanlardaki; K toplam katman sayısı, k belirli katman olmak üzere, $1 \leq k < K$, ardışık iki katman nöronları olan $1 \leq i \leq t_k$ ve $1 \leq j \leq t_{k+1}$ olan her bir öznitelik çiftine uygulanacaktır. 4 temel kapsama yöntemleri (coverage methods) olan İşaret-İşaret Kapsamı (Sign-Sign Coverage) veya İİ Kapsamı (SS Coverage), Değer-İşaret Kapsamı (Value-Sign Coverage) veya Dİ Kapsamı (VS Coverage), İşaret-Değer Kapsamı (Sign-Value Coverage) veya İD Kapsamı (SV Coverage) ve Değer-Değer Kapsamı (Value-Value Coverage) veya DD Kapsamı (VV Coverage) ile her bir öznitelik çifti için bütün katmanlarda ve her bir nöron değerleri için kontroller sağlanacaktır. (Sun, Y., Huang, X., Kroening, D., Sharp, J., Hill, M., & Ashmore, R. (2018). Testing deep neural networks. arXiv preprint arXiv:1803.04792.)
2. **Zaman Maliyetini İndirgeme:** Geleneksel yazılım ürünleri, yazılım geliştirme sürecinin erken evrelerinde test edilerek ilerde oluşabilecek hatalar sonucunda zaman maliyetinden kaçınmak için önemlidir. Oluşturulacak Temiz Yapay Zekâ Kütüphanesi sayesinde YZ geliştiricileri için DSA modelinin yapısal analizini gerçekleştiren bir araç sunulacaktır ve geliştiriciler tarafından oluşturulan modellerin oluşturulan araca gerekli girdi verileri sağlanarak geliştirilen modelin girdi verilerinin yeterliliğine, ağırlık değerlerine ve modelin başarı ölçütünü etkileyebilecek bütün parameterlere iyileştirme yapılıp daha iyi sonuçlar almaları sağlanacaktır. Bu aracın temel amaçlarından birisi ise geliştiricinin kütüphane içindeki ilgili fonksiyonu kullanarak ve gerekli girdi değerlerini bu fonksiyona sağlayarak DSA'nın yapısal analiz sonuçlarını kolayca elde etmesidir. Elde edilen bu değerler ile YZ modelini geliştiren geliştiriciye, DSA yapısında yapılması gereken gerekli değişiklikler hakkındaki kriterler üzerinde ve kapsam kriterlerinin ölçümü üzerinde sonuçlar çıkarılmasına yardımcı olunacaktır. Kütüphane içinde oluşturulmuş fonksiyonlara ilgili girdi parametreleri verilip ortak kullanılacak olan DSA yapısal analiz metodu çalıştırılacaktır. Daha sonraki aşamada ise bu metodun çıktı değerleri oluşturulmuş her bir

kapsam metodlarına iletilecek ve işlemler tamamlanacaktır. Geliştiriciler sadece hazırlanmış olan DSA yapısal analiz metodunu kullanarak zaman maliyetinden tasarruf sağlayabileceklerdir. Ayrıca çıkarılan analizler sonucunda geliştiricinin DSA yapısında gerekli değişiklikleri uygulaması sonucu modelin performans kalitesinde iyileştirmelere gidilebilir ve bu durum da modelin kullanıldığı ortamda yerine getirmesi gereken görev için de zaman maliyetini aşağı çeker.

3. **Yeni Test Yöntemini YZ Tabanlı Sistemlere Uygulama:** Üçüncü hedefimiz ise, günümüzde çok az rastlanan bu tür model testlerini geleneksel test metotlarından farklı olarak uygulayacağımız yeni test yöntemi ile kritik güvenlik zafiyeti içeren sistemlerin performans kayıplarının önüne geçip aynı zamanda bu kayıplardan doğabilecek kazalara sebep olmasını engellemektir.
4. **Visual Studio Code Marketplace Eklentisi Oluşturma:** Oluşturulan YZ Test Kütüphanesi, kullanıcıların daha rahat erişim sağlayabilmesi adına Visual Studio Code uygulamasının Marketplace kısmında eklenti olarak da yayınlanacaktır. Bu doğrultudaki amacımız ürünün kullanılırlığının yaygınlaşması, kolaylaşması ve daha çok kitleye hitap edebilmesi yönündedir.

Sonuç olarak, oluşturulacak olan bu YZ sistemi test kütüphanesi sayesinde sistemde yapılacak olan değişiklikler ile sistemin emniyetli, beklenildiği gibi davranması, hatalarını bulma, farklı katmanlarda milyonlarca nöronların kullanımından oluşan zaman ve iş gücü maliyetleri kaybının önlenmesi ve buna buna bağlı olarak geliştirilen bütün AR-GE projelerinin geliştiricilerine, geliştirilen YZ modellerinin testlerine kolaylık sağlanması planlanmaktadır. Aynı zamanda YZ tabanlı sistemlerin test edilmesinde büyük ve yüksek hacimli verilerin elde edilmesinin oluşturduğu kayıpları, girdi verilerinin gerçek dünyadaki olayları temsil ettiği durumlarda, bu girdi verilerinin aslında gerçekte zamanla değişikliğe uğraması sonucu oluşan birtakım zorlukların üstesinden gelinmesi hedeflenmektedir.

Bu proje, daha önceden yapılmış olan literatür taramaları sonucunda elde edilen kısıtlı kazanımların ortaya koyulacağı ve ilerde YZ'nin testi hakkında yapılacak olan diğer çalışmalara yol haritası sağlayabileceği bir çalışma olacaktır. Temiz Yapay Zekâ Kütüphanesi, DSA yapısında bulunan nöronların analizini gerçekleştirmek amacıyla 9 adet farklı kapsam kriterlerini (İşaret-İşaret Kapsamı (Sign-Sign Coverage), Değer-İşaret Kapsamı (Value-Sign Coverage), İşaret-Değer Kapsamı (Sign-Value Coverage), Değer-Değer Kapsamı (Value-Value Coverage), Nöron Sınır Kapsamı (Neuron Boundary Coverage), En İyi Nöron Kapsamı (Top Neuron Coverage), Çok Bölmeli Nöron Kapsamı (Multisection Neuron Coverage), Nöron Kapsamı (Neuron Coverage) ve Emniyet Kapsamı (Safety Coverage)) içerisinde barındıracaktır. Bu kapsam kriterlerinin kullanılması sonucunda elde edilen analiz verileri DSA yapısının iyileştirilmesi ve gerekli performans artırımlarının yapılması amacıyla geliştiricilerin kullanımına sunulacaktır. Temiz Yapay Zekâ Kütüphanesi'nin oluşturulmasındaki amaç YZ barındıran sistemlere olan güvenirliği arttırmak, bu tarz sistemlerde oluşabilecek kazaların sayısını düşürmek, şeffaflık, yorumlanabilirlik ve açıklanabilirlik gibi YZ kalite gereksinimlerini karşılamaktır.

1.2. Yenilikçi Yönü ve Teknolojik Değeri

Grand View Research'da yayımlanan "Artificial Intelligence Market Size, Share & Trends Analysis Report By Solution, By Technology (Deep Learning, Machine Learning, Natural Language Processing, Machine Vision), By End Use, By Region, And Segment Forecasts, 2022-2030" isimli rapora göre YZ'nin küresel açıdan pazar boyutu 2021 yılında 93,5 Milyar ABD doları olarak ölçülmüştür. Bu küresel pazarın 2030 yılına kadar ise %38,1 oranında artış göstereceği öngörülmektedir. Şirketlerin daha iyi bir müşteri deneyimi sağlayabilmek amacıyla YZ teknolojisini bünyelerine katmak için önemli yatırımlarda bulunduğu da atlanılmaması gereken bir diğer önemli hususdur. Gartner isimli şirketin 2019 yılında yaptığı araştırmada ise gelecek 4 yıl içerisinde YZ kullanım oranının %270 oranında artış göstereceği tahmin edilmektedir (Grand View Research. Artificial Intelligence Market Size, Share & Trends Analysis Report By Solution, By Technology (Deep Learning, Machine Learning, Natural Language Processing, Machine Vision), By End Use, By Region, And Segment Forecasts, 2022 - 2030.). Bu nedenle pazar payı bu denli yüksek olan bir alanda ortaya çıkan ürünlerin test edilmesi gerekliliği barizdir ve bu durum da YZ'nin testini gerçekleştirmeye olanak sağlayacak bir araç ihtiyacı meydana getirir.

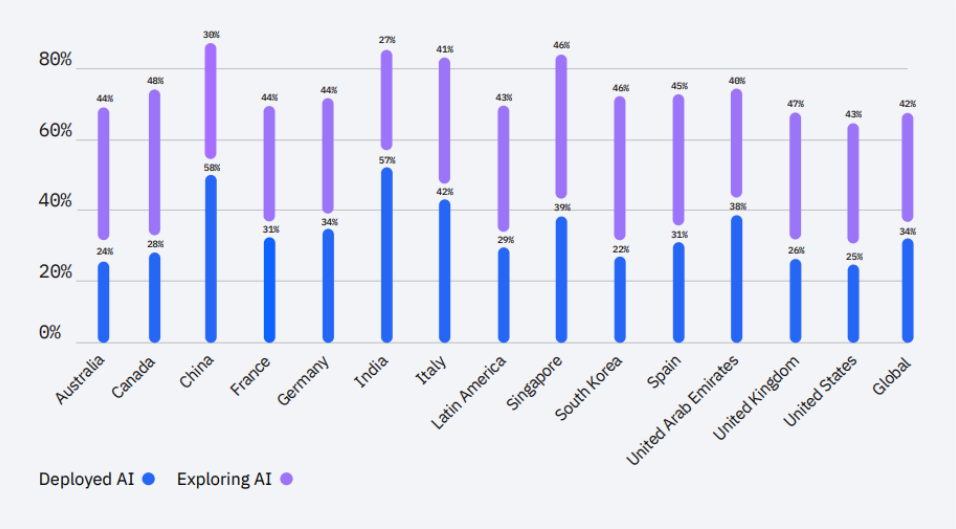
Yapay Zeka Pazar Raporu Kapsamı

Rapor Özelliği	Detaylar
2022'de pazar büyüklüğü değeri	136.6 milyar dolar
2030'da gelir tahmini	1,811,8 milyar ABD doları
Büyüme oranı	2022'den 2030'a kadar %38,1'lik CAGR
Tahmin için temel yıl	2021
Tarihsel veri	2017 - 2020
Tahmin dönemi	2022 - 2030

Tablo 1: YZ Pazar Raporu Kapsamı (Grand View Research)

(Grand View Research. Artificial Intelligence Market Size, Share & Trends Analysis Report By Solution, By Technology (Deep Learning, Machine Learning, Natural Language Processing, Machine Vision), By End Use, By Region, And Segment Forecasts, 2022 - 2030.)

YZ'nin benimsenmesi 2022'de hızlı bir şekilde devam etti. Şirketlerin %35'i işlerinde YZ kullandığını raporladı. (IBM Global AI Adoption Index 2022, 2022) YZ, özellikle iş dünyasında kimin işe alınacağından kendi kendini süren arabalar gibi güvenlik açısından kritik uygulamalara kadar çok çeşitli problemlerde insanlardan daha hızlı, doğru, güvenilir ve tarafsız bir şekilde çalışabilir. Ancak YZ'ler de çok fazla sayıda, bazı ölümcül derecede başarısızlıklar yaşadı ve YZ'nin artan kullanım sıklığı bize gösterir ki, başarısızlıkların sadece bireyleri değil, milyonlarca insanı da etkileyebileceği anlamına geliyor. (Choi, C. Q. (2021). 7 Revealing Ways AIs Fail: Neural Networks can be Disastrously Brittle, Forgetful, and Surprisingly Bad at Math. IEEE Spectrum, 58(10), 42-47.) Amazon şirketinin işe alım süreçleri için geliştirmiş olduğu YZ'nin işe alım süreçlerinde cinsiyetçi bir yaklaşım sürdüğü görüldü. (Dastin Jeffrey. 2018. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. Ulaşılabilir: https://www.reuters.com.) Tesla şirketinin pazara sürmüş olduğu Autopilot özellikli araçlar geçen yıldan bu yana 273 kazaya karışmış bulunmakta (Siddiqui, F., Lerman, R., Merrill, J.B. (2022). Teslas running Autopilot involved in 273 crashes reported since last year. The Washington Post. Ulaşılabilir: https://www.washingtonpost.com.). Bizim bu projeyi gerçekleme motivasyonumuz ise bunlar gibi kritik durumların önüne geçmektir.



Şekil 1: Ülkelere göre 2022 yılında YZ'nin kullanım ve yayılım oranları

(IBM, (2022). "IBM Global AI Adoption Index 2022" https://www.ibm.com/watson/resources/ai-adoption)

İçerisinde kritik sistemler barındıran uygulamaların DSA teknolojisini kullanması sebebiyle yazılımın ilgili güvenlik standartlarının gerekliliklerini karşılamak için nasıl test edilmesi, doğrulanması ve nihai olarak da sertifikalandırılması soruları ortaya çıkmaktadır. Endüstri tarafında paydaşlara yazılım ürünü veya hizmetinin

kalitesi hakkında bilgi sağlamak açısından birincil araç olarak teste güvenilmektedir. Nisan 2021'de duyurulan Avrupa Yapay Zekâ Yasası ile geliştiricilerin YZ testi uygulamalarındaki stratejilerini yeniden gözden geçirmeleri gerekli kılınmıştır. (Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, 2022) Bu proje ile YZ, test ve YZ'nin testi alanları için söz konusu gelişmelere katkı sağlamakla beraber yeni metot ve yöntemlerin araştırılması ile bu araştırmalardan çıktılar üretilmesi amaçlanmaktadır.

YZ'nin birçok sektöre ve teknolojiye hızlı bir şekilde entegre olması beraberinde YZ'nin güvenilirliği ve performansı gibi önemli alt başlıkları da ortaya çıkarmıştır. Bu ilişki de doğrudan YZ'nin testi ihtiyacını meydana getirmiştir. Günümüz tarihi baz alındığında ise YZ'nin testi alanında hazırlanan akademik makaleler haricinde somut anlamda DSA'nın performansını ölçen test araçlarına çok fazla rastlanılmamaktadır. Üzerinde hazırlık yürüttüğümüz çalışma ise daha önceleri kullanılan klasik metotlardan olan YZ'nin ürettiği çıktı değerlerinin doğruluğunu incelemekten ziyade DSA'da bulunan nöronlar arasındaki bağlantıları incelemekte ve ağda bulunan bu nöronların yüzde kaçının karar vermede etkin rol aldığı, sinir ağının performansını önemli ölçüde etkilemeyen nöronların var olup olmadığı gibi bazı daha önceden kullanılmamış istatiksel değerleri gösterecektir.

Yazılım testindeki araştırmalar, farklı yazılım kritiklik düzeylerinin değerlendirilmesi için farklı yaklaşımlar sunmaktadır.

DSA'lar beyin benzeri işlevselliği hedefler ve beyin gibi DSA'lar da nöronlardan oluşur. Bu sözde nöronlar katmanlar halinde toplanır ve bir katmanın çıktıları bir sonraki katmanın girdileri olur. DSA'lar son zamanlarda görüntü sınıflandırma, metin veya konuşma tanıma gibi karmaşık işlerde etkileyici performanslar sergilemiştir. Farklı tiplerde DSA'lar bulunmaktadır, bunlardan bazıları: Çok Katmanlı Algılayıcılar, Evrişimli Sinir Ağları, Tekrarlayan Sinir Ağları.

Çok Katmanlı Algılayıcılar: Bir dizi tamamen bağlantılı katmandan oluşan en temel DSA'dır. Her yeni katman, bir önceki katmandan gelen tüm çıktıların ağırlıklı toplamının doğrusal olmayan bir işlevidir. Günümüzde, modern derin öğrenme mimarilerinin gerektirdiği yüksek bilgi işlem gücü gereksiniminin üstesinden gelmek için Çok Katmanlı Algılayıcılar makine öğrenme yöntemleri ile kullanılabilir.

Evrişimli Sinir Ağları: Görüntü işleme ve sınıflandırma için en popüler ve güçlü araçtır. Bir evrişimli sinir ağı bir girdi görüntüsünü alabilen, görüntüdeki çeşitli nesnelere önem (weight ve bias) atayan ve nesnelerin birbirinden ayırt edilebilmesini sağlayabilen, aynı zamanda nesnelerin birbirleriyle olan ilişkilerini çıkarabilen bir derin öğrenme algoritmasıdır.

Yinelemeli Sinir Ağları: Yinelenen sinir ağları, ileri beslemeli bir sinir ağının aksine, nöronlar arasındaki bağlantıların yönlendirilmiş bir döngü oluşturduğu yinelemeli yapay sinir ağının bir çeşididir. Yinelenen sinir ağlarında çıktı sadece mevcut girdilere göre değil, aynı zamanda önceki adımın nöronunun durumuna da bağlıdır. Yinelemeli Sinir Ağları, kullanıcıların el yazısını tanıma veya konuşma tanıma gibi doğal dil işleme sorunlarını çözmede kullanılır. (Boesch, G. (2021). "Deep Neural Network: The 3 Popular Types (MLP, CNN and RNN)" https://viso.ai/deep-learning/deep-neural-network-three-popular-types/)

Yukarıda verilen örneklerin yanında DSA'lar sağlık sektöründe hastalık teşhisinde, savunma sanayisinde uydudan gelen görüntüler ile tehdit algılamada ve de otonom araçlarda görüntü tanıma görevlerinde kullanılır. Özellikle Türkiye'de Togg şirketinde üretilen elektrikli araçlarda ileriki zamanlarda ortaya çıkabilecek bir otonom araç teknolojisinde bu tip sinir ağları da kullanılacaktır.

Şekil 2: Çekişmeli örnek

(Sun, Y., Huang, X., Kroening, D., Sharp, J., Hill, M., & Ashmore, R. (2018). Testing deep neural networks. arXiv preprint arXiv:1803.04792.)

Geleneksel yazılım için kullanılan testler (white-box, black-box testleri vs.) YZ tabanlı sistemlerin kendi kendine öğrenmesi, gelişmesi ve farklılaşması nedeniyle doğrudan DSA'lara uygulanamamaktadır. DSA'ların kod akışı, eğitim aşamasında öğrenilen bilgileri temsil etmek için yeterli değildir. Bu sebeple DSA'lar için yapısal kapsama kriterlerinin nasıl tanımlanacağı açık değildir. Ayrıca, DSA'lar geleneksel yazılımlardan farklı hatalar sergilemektedir. Özellikle de insan gözüyle ayırt edilemeyen iki girdi (input) verisinin çelişkili kararlara neden olduğu çekişmeli örnekler (adversarial samples), DSA'lardaki en belirgin güvenlik endişelerinden birisidir. Şekil 1'de çekişmeli örnekler için bir görsel yer almaktadır. (Sun, Y., Huang, X., Kroening, D., Sharp, J., Hill, M., & Ashmore, R. (2018). Testing deep neural networks. arXiv preprint arXiv:1803.04792.)

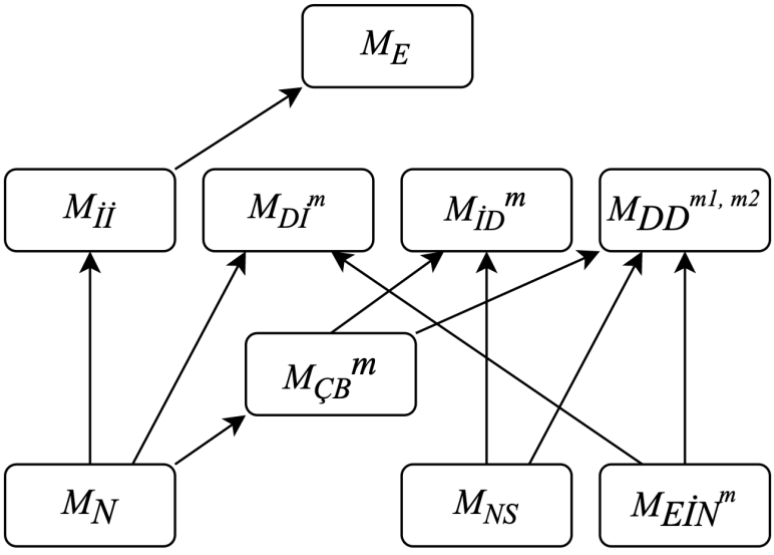
Uygun kapsama kriterleri tarafından yönlendirilmiş DSA'ların test edilmesi ile geliştiricilerin hatalarını bulması, modelin ağ sağlamlığının ölçülmesi ve iç yapısının analiz edilmesi sağlanacaktır. Ayrıca geliştiriciler, model ağını yeniden eğitmek ve geliştirmek için oluşturulan bu çekişmeli örnekleri kullanabilir. Bunun sonucunda geliştiriciler güvenlikle ilgili herhangi bir argüman için farklı ağları anlayabilir ve karşılaştırmasını sağlayabilir.

Sinir ağları için tanımlanmış kapsama kriterlerinin test edilmesiyle dört açıdan fayda sağlanabilir:

1. Hata (çekişmeli örnek) bulma
2. DSA güvenlik istatistikleri
3. Test verimliliği
4. DSA yapı analizi

Bizim üzerinde duracağımız asıl fayda ise dördüncü maddede belirtilen "DSA yapı analizi" olacaktır.

Hızlı uyum sağlayan sistemlerde her bir değişiklikten sonra yeni testlerin manuel olarak yürütülmesi otomatize edilmiş hale göre daha çok zaman ve maliyete sebep olmaktadır. Sistemin kendini her değiştirdiği durumda otomatik olarak çalışabilecek yeni testler yazmak gerekliliği var olabilir. Bazı durumlarda sistemin normalde çalışması gereken kısıtlamaları bilinebilir fakat sistemin kendisinin meydana getirebileceği değişiklikler hakkında bilgi sahibi olmak zordur. Normal şartlar altında belirlenen gereksinimlere karşı sistemi test etmek kolaydır ancak sistemin yenilikçi bir gerçekleme ürettiği durumlarda üretilen bu gerçeklemeleri test etmek çok da kolay değildir. [7] Geliştirilecek bu kütüphane ile DSA'nın yapısının analizi ile bu testlerin gerekliliği ortadan kaldırılacaktır. Uygulanacak işlemlerin detayları Yöntem başlığı altında detaylı bir şekilde anlatılmıştır.



Şekil 3: DSA Test kriterleri arasındaki ilişki

(Sun, Y., Huang, X., Kroening, D., Sharp, J., Hill, M., & Ashmore, R. (2018). Testing deep neural networks. arXiv preprint arXiv:1803.04792.)

Yaptığımız araştırmalar sonucunda uygulayacağımız yönteme benzer bir yöntem kullanan araç olan DeepXplore ve TensorFuzz ile karşılaştık. Bu araçlar ile olan benzerlik ve farklılıklarımızı ise aşağıda belirttik.

- **DeepXplore:** Bu araç beyaz-kutu (white-box) testlerini gerçek zamanlı derin öğrenme sistemlerine uygulamaya yarayan bir araç sunmaktadır. DeepXplore, DSA'nın yapısının analizini yapmaktan ziyade bir YZ modelinin hatalı davranışlar gösterebileceği köşe durumları belirlemeye çalışır. Bu araç bizim de gerçekleştirmeyi hedeflediğimiz amaçlardan en önemlisi olan DSA sistemlerinin güvenilirliğini ve performansını test ederken nöron kapsamının hesaplanması mantığını da sade bir şekilde barındırmaktadır [https://github.com/peikexin9/deepxplore]. Bizim bu araçtan farklı olarak sunduğumuz hizmetler ise İşaret-İşaret Kapsamı (Sign-Sign Coverage), Değer-İşaret Kapsamı (Value-Sign Coverage), İşaret-Değer Kapsamı (Sign-Value Coverage), Değer-Değer Kapsamı (Value-Value Coverage), Nöron Sınır Kapsamı (Neuron Boundary Coverage), En İyi Nöron Kapsamı (Top Neuron Coverage), Çok Bölmeli Nöron Kapsamı (Multisection Neuron Coverage), Nöron Kapsamı (Neuron Coverage) ve Emniyet Kapsamı (Safety Coverage) olarak dokuz başlığa ayrılmış kapsama yöntemlerinin verilen model üzerinde hesaplanması ve üretilen analizlerin kullanıcıya sunulması olacaktır (Pei, K., Cao, Y., Yang, J., & Jana, S. (2017, October). Deepxplore: Automated whitebox testing of deep learning systems. In proceedings of the 26th Symposium on Operating Systems Principles (s. 1-18).).
- **TensorFuzz:** TensorFuzz, nadir şekilde meydana gelen girdi verileri için sistemin ürettiği hataları saptamaya yarayan DSA'lara yönelik oluşturulmuş yazılım testi tekniğidir. Bu teknikte, bir sinir ağına girdilerin rastgele mutasyonları verilerek sinir ağları için kapsama değerlerine ulaşılması hedeflenir. DeepXplore aracında belirttiğimiz gibi bizim bu araçtan farklı olarak sunmuş olduğumuz dokuz maddeden oluşan kapsama yöntemi mevcuttur. Temiz Yapay Zekâ Projesi de daha önce ele alınmamış bu dokuz kapsama yöntemini kullanarak DSA yapısının analizini ve testini gerçekleştirecektir.

2. YÖNTEM

1. Projede Kullanılacak Teknolojiler ve Araçlar

1.1 Kullanılacak Teknolojiler

1.1.1 Python Programlama Dili

Python; nesne yönelimli, yorumlamalı, modüler ve etkileşimli yüksek seviyeli bir programlama dilidir. Python ile uzaktan kontrol veya görüntü işleme, veri analizi ve veri kontrolü yapılabilir. Kullanım alanındaki çeşitliliği ve YZ uygulamalarında sıkça kullanılması sebebi ile önerilen projemizi Python programlama dilinde kodlamaya karar verdik.

1.1.2 Python Kütüphaneleri

TensorFlow, PyTorch ve JAX gibi kütüphaneler ile derin öğrenme uygulamaları yapılmaktadır. Oluşturulacak olan YZ testi kütüphanesinin yapımında da Python programlama dilinde sıkça kullanılan PyTorch ve TensorFlow kütüphaneleri kullanılacaktır. Bu kütüphanelerin kullanım sebebi olarak derin öğrenme uygulaması geliştiricileri tarafından oluşturulan modellerin bu kütüphaneler kullanılarak gerçekleştirilmesidir. Oluşturulan modelleri, kodlanmış olduğu kütüphaneleri kullanarak analiz edebilmektir.

1.2 Kullanılacak Araçlar

1.2.1 Visual Studio Code

Visual Studio Code; Microsoft tarafından Windows, Linux ve MacOS için geliştirilen bir kaynak kod düzenleyicisidir. Hata ayıklama, Git kontrolü, sözdizimi vurgulama, akıllı kod tamamlama, kod parçacıkları ve kod yeniden yapılandırma desteği içermektedir. Aynı zamanda özelleştirilebilir yapısı sayesinde çeşitli eklentiler ile bu editörün kullanımı kolaylaşabilir. Bu esnekliği sebebiyle Visual Studio Code kullanılacaktır.

1.2.2 Google Colaboratory

Colaboratory (ya da kısaca Colab), Google Research tarafından sunulan bir üründür. Özellikle veri analizi makine

öğrenimi ve derin öğrenme model eğitimini gerçekleştirmek için uygundur. Colab sayesinde isteyen herkes web tarayıcısı üzerinden Python kodunu yazabilir ve çalıştırabilir. Daha teknik açıdan tanımlanacak olursa Colab, kullanım için kurulum gerektirmeden GPU gibi bilgi işleme kaynaklarına ücretsiz erişim imkânı sunan bir Jupyter not defteri hizmetidir.

2. Yapım Aşamaları

Bu proje kendi içerisinde 8 iş paketine bölünmüştür. Bu paketlerin gerçeklenmesi ile projemizin Haziran 2023 itibari ile ortaya çıkması öngörülmektedir.

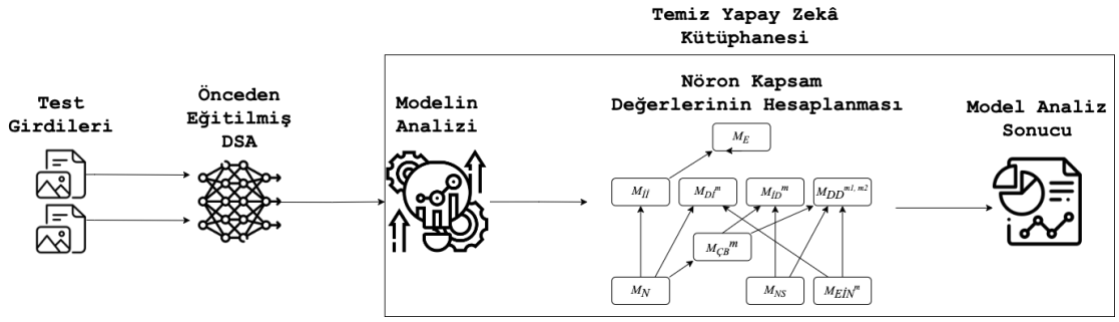
2.1 Literatür ve Kullanılacak Teknolojilerin Araştırılması

Projemizin ilk iş paketinde yapılacak olanlar, halihazırda yapmış olduğumuz literatür araştırmalarını daha geniş kapsamlı araştırmalar ile desteklemektir. Google'ın YZ için oluşturmuş olduğu prensiplerden biri olan "Be built and tested for safety." (Google "Artificial Intelligence at Google: Our Principles" https://ai.google/principles/) ilkesine bağlı kalarak YZ'lerin test edilebilir hale getirilmesi amaçlanmaktadır. Literatür araştırmaları kapsamında; YZ testi için önemli olan adalet, güvenilirlik, açıklanabilirlik ve hesap verilebilirlik [9] kavramlarının YZ bağlamında incelenecek, test etme kavramının YZ üzerine nasıl uygulanabileceğinin araştırılması gerçekleştirilecek ve eğer var ise bu tür projeler kapsamlı olarak incelenecektir.

2.2 Kullanılacak YZ Test Metodunun Belirlenmesi

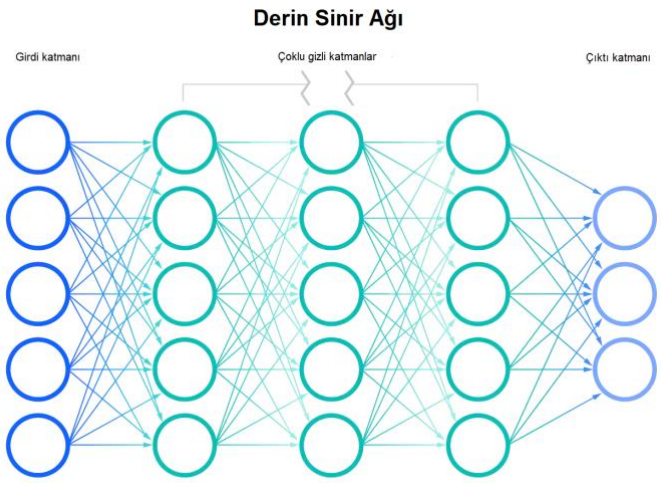
Projemizin ikinci iş paketi ise, projemizin ayağa kaldırılması için gerekli olan teknolojiler belirlenecek ve bu teknolojilerin projemiz özelinde nasıl kullanılabileceği araştırılacaktır.

2.3 Kütüphane Gereksinimlerinin Belirlenmesi ve Tasarımının Yapılması



Şekil 4: Oluşturulacak olan Kütüphanenin Mimarisi

DSA'lar; katmanlar, katmanlar arasındaki bağlantı kümeleri (ağırlıklar) ve nöronlardan oluşmaktadır. DSA yapısında bulunan bu nöronların arasındaki bağlantılar YZ'nin nihai kararlarını oluşturmaktadır. YZ'nin nihai kararının oluşması sırasında katmanlarda bulunan her bir nöronun üretilen sonucu oluştururken ne derece etkili olduğunu analiz edilmektedir. (Sun, Y., Huang, X., Kroening, D., Sharp, J., Hill, M., & Ashmore, R. (2018). Testing deep neural networks. arXiv preprint arXiv:1803.04792.)



Şekil 5: Basit bir derin sinir ağı yapısı

(IBM, "Neural Networks" https://www.ibm.com/cloud/learn/neural-networks)

Bir DSA, L katman kümesi, T katmanlar arasındaki bağlantı kümesi ve Φ aktivasyon fonksiyonu olmak üzere N = (L, T, Φ) kümesi olarak ifade edilebilmektedir. Buna göre bir DSA'nın yapısı:

- L_1 girdi katmanı (input layer) ve L_k çıktı katmanı (output layer) arada kalan katmanlar ise gizli katmanlar (hidden layers) olarak ifade edilmektedir.
- Her L_k katmanı, s_k nöronlarından oluşmaktadır. *K*'nın *l*. düğümü $n_{k,l}$ ile gösterilmektedir.
- Katman $1 < k < K$ ve nöron $1 \leq l \leq s_k$ ifadesi için her $n_{k,l}$ nöronu sırasıyla aktivasyon fonksiyonunun önceki ve sonraki değerlerini kaydetmek için $u_{k,l}$ (öncesi) ve $v_{k,l}$ (sonrası) iki değişkeni ile ilişkilendirilmektedir.

Şekil 4. te gerçekleyeceğimiz kütüphanenin sistem mimarisi tasarımı bulunmaktadır. Bu tasarıma göre kullanıcılardan alınacak test girdileri ve eğitilmiş modeli bir arada kullanarak bir sonraki katmana gerekli olan girdilerin sağlanması yapılacaktır. Bir sonraki aşamada ise Model Analizi katmanı kullanılarak modelin analizi yapılacaktır. Analiz sonucunda bir sonraki katman için kullanılan aktivasyon değerlerinin isimleri, nöron değerleri, vb. bazı gerekli çıktı değerleri bir sonraki katmana (Temiz Yapay Zekâ Kütüphanesi) iletilecektir. Bir sonraki katmanın da ise kullanıcılardan test girdileri ile birlikte bir önceki katmanda sağlanan çıktılar alınacaktır. Bu katmanda aşağıda belirteceğimiz 9 kapsam yöntemi uygulanacaktır.

1. **İşaret-İşaret Kapsamı (Sign-Sign Coverage) veya İİ Kapsamı (SS Coverage)**: Bir öznitelik çifti $a = (\psi_{k,i}, \psi_{k+1,j})$ için aşağıdaki koşullar DSA N[x1] ve N[x2] tarafından karşılanıyorsa, İİ(a, x1, x2) ile gösterilen iki test durumu x1 ve x2 tarafından İİ kapsamındadır. Özet olarak; iki bitişik katmandaki nöronlardan ilk katmanda işaret değişiyor ve ardından tekrar bitişik katmanda işaret değişiyorsa bu kapsam içerisindedir. sc işaret değişikliği ve nsc işaret değişikliği bulunmama durumunu ifade etmek üzere:
   - $sc(\psi_{k,i}, x1, x2)$ ve $nsc(P_k \setminus \psi_{k,i}, x1, x2)$. ($P_k$, k katmanındaki düğümler kümesidir.)
   - $sc(\psi_{k+1,j}, x1, x2)$.
2. **Değer-İşaret Kapsamı (Value-Sign Coverage) veya Dİ Kapsamı (VS Coverage)**: Bir g uzaklık değer fonksiyonu verildiğinde, bir öznitelik çifti $a = (\psi_{k,i}, \psi_{k+1,j})$ için aşağıdaki koşullar DSA N[x1] ve N[x2] tarafından karşılanıyorsa, $VS^g(a, x1, x2)$ ile gösterilen iki test durumu x1 ve x2 tarafından Dİ kapsamındadır. Özet olarak; iki bitişik katmandaki nöronlardan ilk katmanda değer değişiyor ve ardından bitişik katmanda işaret değişiyorsa bu kapsam içerisindedir. sc işaret değişikliği, nsc işaret değişikliği bulunmama durumunu ve vc değer değişikliğini ifade etmek üzere:
   - $vc(g1, \psi_{k,i}, x1, x2)$ ve $nsc(P_k, x1, x2)$.
   - $sc(\psi_{k+1,j}, x1, x2)$.
3. **İşaret-Değer Kapsamı (Sign-Value Coverage) veya İD Kapsamı (SV Coverage)**: Bir g uzaklık değer fonksiyonu verildiğinde, bir öznitelik çifti $a = (\psi_{k,i}, \psi_{k+1,j})$ için aşağıdaki koşullar DSA N[x1] ve N[x2] tarafından karşılanıyorsa, $İD^g(a, x1, x2)$ ile gösterilen iki test durumu x1 ve x2 tarafından İD kapsamındadır. Özet olarak; iki bitişik katmandaki nöronlardan ilk katmanda işaret değişiyor ve ardından bitişik katmanda değer değişiyorsa bu kapsam içerisindedir. sc işaret değişikliği, nsc işaret değişikliği bulunmama durumunu ve vc değer değişikliğini ifade etmek üzere:
   - $sc(\psi_{k,i}, x1, x2)$ ve $nsc(P_k \setminus \psi_{k,i}, x1, x2)$.
   - $vc(g, \psi_{k+1,j}, x1, x2)$ ve $nsc(\psi_{k+1,j}, x1, x2)$.

4. **Değer-Değer Kapsamı (Value-Value Coverage) veya DD Kapsamı (VV Coverage)**: İki adet g1 ve g2 uzaklık değer fonksiyonları verildiğinde, bir öznitelik çifti $a = (\psi_{k,i}, \psi_{k+1,j})$ için aşağıdaki koşullar DSA N[x1] ve N[x2] tarafından karşılanıyorsa, $DD^{g1, g2}(a, x1, x2)$ ile gösterilen iki test durumu x1 ve x2 tarafından DD kapsamındadır. Özet olarak; iki bitişik katmandaki nöronlardan ilk katmanda değer değişiyor ve ardından tekrar bitişik katmanda değer değişiyorsa bu kapsam içerisindedir. nsc işaret değişikliği bulunmama durumunu ve vc değer değişikliğini ifade etmek üzere:
   - $vc(g1, \psi_{k,i}, x1, x2)$ ve $nsc(P_k, x1, x2)$.
   - $vc(g2, \psi_{k+1,j}, x1, x2)$ ve $nsc(\psi_{k+1,j}, x1, x2)$.
5. **Nöron Kapsamı (Neuron Coverage):** Bir $n_{k,i}$ nöronu işaret($n_{k,i}$, x) = +1 ise, DSA N x test durumu tarafından kapsanmaktadır. Bu tanımda yalnızca işaret aktivasyon fonksiyonu için tanımlama yapılmıştır. Ancak bu tanımlamaya göre diğer aktivasyon fonksiyonları düzenlenebilir.
6. **Nöron Sınır Kapsamı (Neuron Boundary Coverage)**: Bu kapsama göre nöron için aktivasyon fonksiyonu çıkış değerine göre üst sınır bulunmaktadır. Bir nöron $n_{k,i}$ ve X eğitim veri seti olmak üzere:
   - $n_{k,i}$ nöronu için üst sınır eğitim verisindeki o nöron için aktivasyon fonksiyonu çıkışındaki üst değer olarak belirlenmektedir. Örneğin, eğitim veri setinde nöron için üst sınır 0.6 ise o nöronun test verisindeki üst sınırı da budur. Test verisinin bu nöron için geçmesi gerekir.
7. **En İyi Nöron Kapsamı (Top Neuron Coverage)**: Bu kapsama göre üst sınırı geçmiş nöronların sıralanmasıyla (büyükten küçüğe) bir sıralama listesi oluşturulmaktadır. Burada kullanıcı tarafından belirlenmiş olan m değerine göre bu sayının altında nöron kapsanmış ise bu kapsama dahildir.

8. **Çok Bölmeli Nöron Kapsamı (Multisection Neuron Coverage)**: Bu kapsama göre 6. Maddede bahsedilmiş olan Nöron Sınır Kapsamı gibi hem üst sınır hem de alt sınır aralığı bulunmaktadır. Bu aralığın içerisinde kapsamaya dahil ise nöronlar, bu kapsam dahilinde olduğunu söylenebilir.
9. **Emniyet Kapsamı (Safety Coverage):** Bu kapsam, DSA modelinin test kapsama kriterleri geçtiği anlamına gelmektedir. Emniyet Kapsamının sağlanabilmesi için ilişkinin bağıntılarındaki tüm kapsama kriterleri de sağlanmak zorundadır. Bu kapsama kriterini sağlayan nöron sinir ağları, çekişmeli örnek (adversarial example) olarak bahsettiğimiz modelin "hatalarına" karşı duyarlı olacaktır. DSA modelini kullanan kritik güvenlikli sistemler, daha güvenli bir YZ yazılımı haline getirilecektir. (Sun, Y., Huang, X., Kroening, D., Sharp, J., Hill, M., & Ashmore, R. (2018). Testing deep neural networks. arXiv preprint arXiv:1803.04792.)

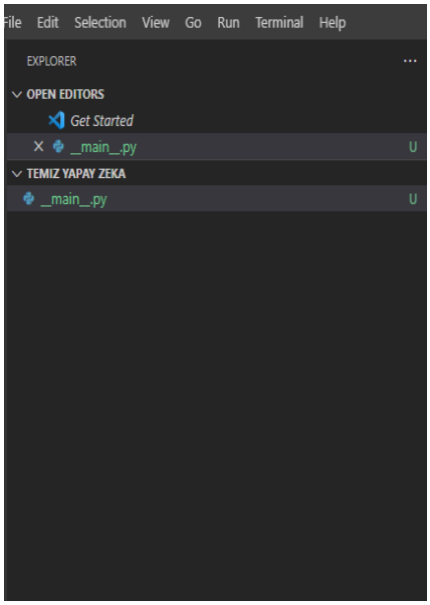
Projemizin üçüncü iş paketinde ise bu yöntemler kapsamında kullanıcıdan alınan modellerin analiz sonucunda her bir öznitelik çifti için ilk ve son katmanlar haricinde bu bahsi geçen metotlar uygulanacaktır.

2.4 Kütüphane Yazılımı İçin Gerekli Yazılımların Araştırılması ve Öğrenilmesi

Projemizin dördüncü iş paketinde ise, oluşturulacak olan bu kütüphane yazılımı için gerekli olan Python yazılım dili, versiyon kontrol teknolojisi (Github) öğrenilecek ve bu şekilde de projenin açık kaynak hale getirilmesi sağlanacaktır.

2.5 Test Kütüphanesinin Oluşturulması

Projemizin beşinci iş paketi ise, dördüncü adımda elde edilen teorik bilgilerin pratiğe dökülmesi ile de beşinci adım boyunca proje yapımına başlanılacak ve tamamlanacaktır. Kütüphane oluşturulurken aşağıdaki adımlar izlenilecektir.



Şekil 6: Visual Studio Code Dizin Yapısı

A. Kütüphanenin yer alacağı dizin belirlenecektir
B. Oluşturulan klasör için sanal ortamın oluşturulması
C. Klasör yapısının oluşturulması
D. Kütüphane geliştirilmesi için gerekli Python dosyaların oluşturulması
E. Kütüphane geliştirilmesi için bağımlılıkların kurulması
F. Test Kütüphanesinin oluşturulması

2.6 Geliştirilen Kütüphanenin Yazılım Testi

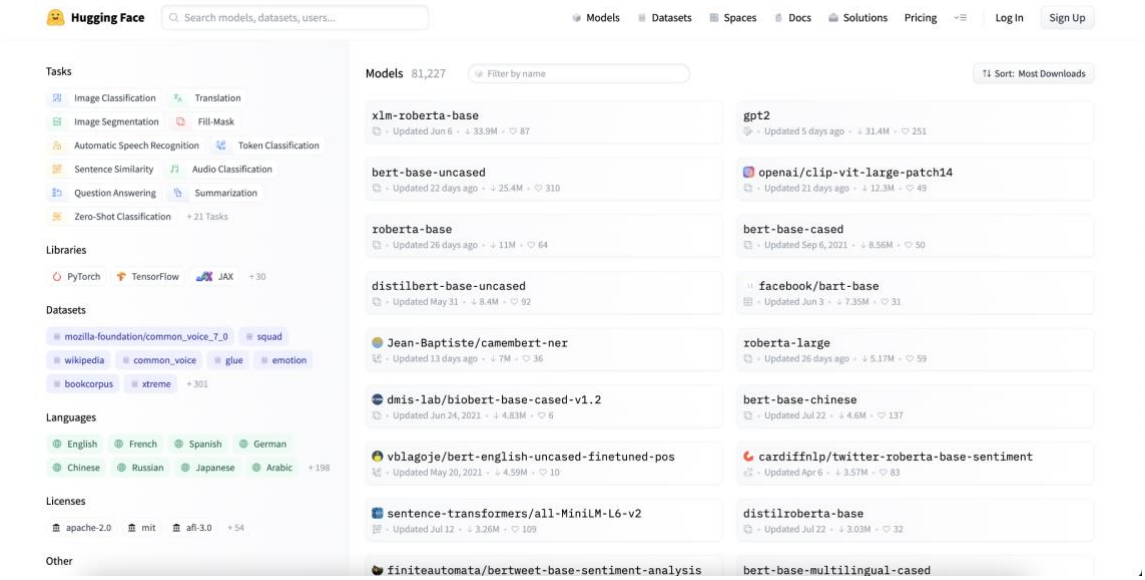
Projemizin altıncı iş paketinde ise oluşturduğumuz kütüphanenin birim, entegrasyon, sistem ve kabul testleri yapılacaktır. Oluşan hatalar sonucunda meydana gelen hatalar giderilecek ve bu sayede kütüphanenin hatasız çalışması sağlanacaktır.

2.7 Farklı YZ Kütüphanelerinden Hazır YZ Modellerinin Belirlenmesi ve Çalıştırılması

Projemizin yedinci iş paketi ise kullanıcılardan ve platformlardan alınan farklı YZ modellerini oluşturduğumuz

kütüphane mimarisi yapısına göre analizleri yapılacaktır. Bu analizlerde ise kullanıcıların modellerinin hangi hazır YZ modelleri (TensorFlow, PyTorch, vb.) ile geliştirdikleri, modellerde hangi aktivasyon fonksiyonlarının uygulandığı, üst sınır neuron değerleri vb. bilgiler elde edilecektir. Alınan modeller ve bilgiler doğrultusunda geliştirdiğimiz kütüphanenin testleri gerçekleştirilecektir.

Hugging Face platformunda hazır halde bulunan çeşitli YZ modellerini belirlenip kendi kurulumları yapılacak ve çalıştırılacaktır.



Şekil 7: Hugging Face platformunda bulunan hazır modeller

2.8 Visual Studio Code Entegrasyonu ve Marketplace'e Eklenmesi

Son iş paketimiz olan sekizinci iş paketinde ise, çalışabilir ve testleri tamamlanmış durumdaki ürünün geliştiricilere daha kolay ulaşması adına çeşitli geliştirici pazar yerlerine (VSC Marketplace vb.) konulunacaktır. Ve bu şekilde de ürünün tanınabilirliği artırılacak ve gelecek olan bağışlar ile sürdürülebilirlik sağlanacaktır.

3 PROJE YÖNETİMİ

3.1 İş- Zaman Çizelgesi

İŞ-ZAMAN ÇİZELGESİ (*)

No	İş Paketlerinin Adı ve Hedefleri	Kim(ler) Tarafından Gerçekleştirileceği	Zaman Aralığı (..-.. Ay)	Başarı Ölçütü ve Projenin Başarısına Katkısı
1	Literatür ve Kullanılacak Teknolojilerin Araştırılması	Abdul Hannan Ayubi Furkan Taşkın Osman Çağlar	1 Ekim - 1 Kasım	En uygun test metodunun belirlenmesi amacıyla YZ'nin testi hakkında son 5 yılda oluşturulmuş en az 20 adet akademik literatür çalışmalarının araştırılması. (%10)
2	Kullanılacak YZ Test Metodunun Belirlenmesi	Abdul Hannan Ayubi Furkan Taşkın Osman Çağlar	15 Ekim - 15 Kasım	YZ testi için gerçekleştirilen literatür ve teknoloji araştırmaları sonucunda elde edilen veriler ışığında geliştirilen test metodlarından YZ testi için kullanılabilecek en uygun 5 test için tüm test metotlarının değerlendirilmesi ve bu kapsamda en uygun 5 adet test metodunun belirlenmesi amaçlanmaktadır. (%15)
3	Kütüphane Gereksinimlerinin Belirlenmesi ve Tasarımının Yapılması	Abdul Hannan Ayubi Furkan Taşkın Osman Çağlar	15 Kasım - 15 Aralık	YZ tabanlı sistemlerde kullanılan modellerin eğitimi GPU ağırlıklı yapıldığı için projemize uygun olan ve GPU barındıran sunucuların hazır bir hale getirilmesi amaçlanmaktadır. (%10)
4	Kütüphane Yazılımı İçin Gerekli Yazılımların Araştırılması ve Öğrenilmesi	Abdul Hannan Ayubi Furkan Taşkın Osman Çağlar	15 Aralık - 15 Ocak	YZ tabanlı sistemlerin model eğitimlerinde en çok kullanılan programlama dili olan Python dilinin orta düzeyde öğrenilmesi ve model eğitiminde gerekli olacak kütüphanelere (PyTorch, TensorFlow) hakim olunması. (%10)

5	Test Kütüphanesinin Oluşturulması	Abdul Hannan Ayubi Furkan Taşkın Osman Çağlar	1 Ocak - 1 Nisan	Elde edilen bilgiler ışığında ürünün oluşturulması ve oluşturulan ürünün test yapabilmesi için en düşük kriterleri karşılaması. Minimum olarak uygulanabilir bir ürünün (MVP) meydana getirilmesi. (%25)
6	Geliştirilen Kütüphanenin Yazılım Testi	Abdul Hannan Ayubi Furkan Taşkın Osman Çağlar	15 Mart - 1 Mayıs	Oluşturulan kütüphane üzerinde birim, entegrasyon, sistem ve kabul testleri uygulanarak meydana gelebilecek hataların saptanması ve giderilmesi hedeflenecektir. (%15)
7	Farklı YZ Kütüphanelerinden Hazır YZ Modellerinin Belirlenmesi ve Çalıştırılması	Abdul Hannan Ayubi Furkan Taşkın Osman Çağlar	1 Mart - 1 Nisan	En az 2 adet YZ kütüphanesinde (PyTorch, TensorFlow vb.) oluşturulan modellerin analizinin gerçekleştirilmesi. Geliştirilen kütüphanenin Hugging Face platformu üzerinde sunulan en az 5 adet hazır model ile çalıştırılmasının sağlanması gerçekleştirilecektir. (%10)
8	Geliştirici Pazar Yerine Entegrasyon	Abdul Hannan Ayubi Furkan Taşkın Osman Çağlar	15 Nisan - 1 Haziran	Ortaya çıkan ürünün geliştiricilere daha kolay ulaşması adına en az 1 adet geliştirici pazar yerine (VSC Marketplace vb.) konulması. (%5)

(*) Çizelgedeki satırlar ve sütunlar gerektiği kadar genişletilebilir ve çoğaltılabilir.

3.2 **Risk Yönetimi**

RİSK YÖNETİMİ TABLOSU*

No	En Büyük Riskler	Risk Yönetimi (B Planı)
1	Geliştirilecek olan kütüphanenin doğru çıktılar üretememesi.	Bir önceki sürümden edinilen deneyimler ile yeni bir sürümün oluşturulması ve böylelikle de önceden yapılan hataların önüne geçilmesi sağlanacaktır.
2	Oluşturulacak kütüphanenin farklı hazır YZ'ler ile entegre bir biçimde çalışmaması.	Kütüphane farklı YZ sistem model gerekliliklerine göre yeniden düzenlenecektir.
3	Gerçek zamanlı ve kendi kendine öğrenen sistemlerde sistemin ilerleyen zamanlardaki değişikliğine kütüphanenin ayak uyduramaması.	Kütüphanenin versiyon kontrol sistemini belli zaman aralıklarla güncellenmesinin otomatikleştirilmesi sağlanacaktır.
4	Farklı aktivasyon nöronlarına sahip DSA yapıları için nöron değerlerine erişim esnasında sıkıntı yaşanması.	PyTorch ve TensorFlow kütüphaneleri yardımıyla DSA yapısında bulunan her bir katmana erişim sağlanabilir ve bu katmanlarda hangi aktivasyon fonksiyonlarının kullanıldığı görülebilir.
5	Farklı YZ kütüphaneleri (TensorFlow, PyTorch vs.) ile oluşturulmuş olan YZ modellerinin nöron değerlerine erişim esnasında kısıtlama yaşanması.	Kütüphane spesifik yaklaşım yöntemi uygulanabilir.
6	Sinir ağları için tanımlanmış kapsama kriterlerinin test edilmesi esnasında kullanıcıdan örnek girdi verileri beklenmektedir. Farklı modeller farklı türde girdi verilerini (bazı modeller resim alırken bazı modeller metin türünde veri almaktadır) desteklemektedir. Her modelin aynı tipte girdi verisi desteklememesinin problem oluşturması.	Modelin yüklenmesi esnasında kullanıcıdan gireceği girdi verisinin tipinin belirtilmesi istenebilir.
7	Modelin test edilmesi için modelin tekrar yüklenmesi aşamasında model sınıfının belirtilmemesinin problem oluşturması.	Kullanıcıdan model sınıfına ait dosya türünün de yüklenmesi talep edilerek problemin üstesinden gelinebilir. Ya da bir diğer yol olarak ise PyTorch kütüphanesi kullanılarak modelin TorchScript türünde kaydedilmesi istenerek model sınıfının tanımlanması ihtiyacı ortadan kaldırılabilir.
8	Oluşturulan kütüphanenin kod editörlerine eklenti olarak sunulması esnasında problem ortaya çıkması.	Oluşturulan kütüphaneye ait bir API kurularak bu problemin üstesinden gelinebilir.

(*) Tablodaki satırlar gerektiği kadar genişletilebilir ve çoğaltılabilir.

3.3. **Araştırma Olanakları**

ARAŞTIRMA OLANAKLARI TABLOSU (*)

Altyapı/Ekipmanın Bulunduğu Kuruluş	**Kuruluşta Bulunan Altyapı/Ekipman Türü, Modeli** (Laboratuvar, Araç, Makine-Teçhizat vb.)	**Projede Kullanım Amacı**
ESOGU SRLAB	**SUNUCU Sistem Özellikleri:** İşlemci: Intel(R) Xeon(R) Silver 4110 CPU @ 2.10GHz Ekran Kartı: Matrox G200eW3 (Nuvoton) WDDM 1.2 RAM: 16GB	Veri setlerinin toplanması, işlenmesi ve modele uygun hazırlanması amaçlarını gerçekleştirmek için kullanılmaktadır.

	Disk: Dell PERC H730P 1TB	
ESOGU SRLAB	İşlemci: Intel(R) Core(TM) i5-10400 CPU @ 2.90GHz Ekran Kartı: Intel UHD Graphics 630 RAM: 8GB Dİsk 1: Sandisk Disk 240GB Disk 2: TOSHIBA DT01ACA1 1TB	Veri setlerinin toplanması, işlenmesi ve modele uygun hazırlanması amaçlarını gerçekleştirmek için kullanılmaktadır.
İnovasyon	DELL EMC R540 Sunucu	Veri setlerinin toplanması, işlenmesi ve modele uygun hazırlanması amaçlarını gerçekleştirmek için kullanılmaktadır.

(*) Tablodaki satırlar gerektiği kadar genişletilebilir ve çoğaltılabilir.

4. SANAYİ ODAKLI ÇIKTILARI ve YAYGIN ETKİ

Yaygın Etki Türleri	Önerilen Araştırmadan Beklenen Çıktı, Sonuç ve Etkiler
Bilimsel/Akademik (Makale, Bildiri, Kitap Bölümü, Kitap)	Ortaya çıkacak proje sonucunda yapay zekâ testi konusu hakkında bir adet makale yayınlanacak ve INISTA2023 konferans için bir adet bildiri ile literatür çalışmalarına katkıda bulunulacaktır.
Ekonomik/Ticari/Sosyal (Ürün, Prototip, Patent, Faydalı Model, Üretim İzni, Çeşit Tescili, Spin-off/Start- up Şirket, Görsel/İşitsel Arşiv, Envanter/Veri Tabanı/Belgeleme Üretimi, Telife Konu Olan Eser, Medyada Yer Alma, Fuar, Proje Pazarı, Çalıştay, Eğitim vb. Bilimsel Etkinlik, Proje Sonuçlarını Kullanacak Kurum/Kuruluş, vb. diğer yaygın etkiler)	● Otomotiv sektöründen eğitim sistemine, hukuk sistemlerinden sağlık hizmetlerine gibi birçok sektörde kullanılan YZ' modellerinin testini kolay bir biçimde gerçekleştirerek YZ'ye olan güvenin artırılması ve bunun sonucunda YZ teknolojisinin kullanımının arttırılması ile ülke ve dünya ekonomisine olumlu bir katkıda bulunacaktır. ● Visual Studio Code Marketplace üzerinde yayınlanması hedeflenen eklenti sayesinde üretilen ürünün kullanımının yaygınlaştırılması hedeflenmektedir. Ürün tanıtımının gerçekleştirilmesinin ve ürünün stabil çalıştığından emin olunmasının ardından ürün için ekonomik anlamda gelir beklentisi oluşacaktır. Bunlara ek olarak eklenti üzerindeki kullanımın sınırlandırılması ve ek kullanımların standart bir tarife üzerinden ücretlendirilmesi sağlanılarak da gelir modeli oluşturulabilinir. Bunların gerçekleştirilmesi sırasında B2C (Business to Customer) ve B2B (Business to Business) iş modelleri esas alınacaktır. B2B iş modeli ile elde edilen proje çıktısının, başta yazılım firmaları olmak üzere yapay zekâ kullanılan bütün işyerlerine ulaşmayı hedeflemektedir. Ayrıca kişisel olarak yapay zekâ projeleri geliştiren kişilere de B2C modeli ile ulaşılması amaçlanmaktadır.
Araştırmacı Yetiştirilmesi ve Yeni Proje(ler) Oluşturma (Yüksek Lisans/Doktora Tezi, Ulusal/Uluslararası Yeni Proje)	Bu proje kapsamında YZ modellerine uygulanacak olan yeni test metodları ile ilgili ulusal ve uluslararası projelere katılım gerçekleştirilecektir.

5. BÜTÇE TALEP ÇİZELGESİ

Bütçe Türü	Talep Edilen Bütçe Miktarı (TL)	Talep Gerekçesi
Sarf Malzeme	0	–
Makina/Teçhizat (Demirbaş)	0	–
Hizmet Alımı	7500	Toplanan veriseti ile birlikte hazır modellerin hızlıca çalıştırılması ve nöron değerlerine hızlıca erişilmesi için GPU barındıran sunucu hizmeti alınarak hızlı prototipleme üzerine çalışmalar yürütülecektir. Kapsam değerleri ölçülürken yapılacak olan testlerde de kullanılacaktır.
Ulaşım	0	–
TOPLAM	7500	

6. BELİRTMEK İSTEDİĞİNİZ DİĞER KONULAR

Proje Ekibi Hakkında

Osman Çağlar (Proje Yöneticisi): Eskişehir Osmangazi Üniversitesi, Bilgisayar Mühendisliği Bölümü son sınıf öğrencisidir. Halihazırda 2247 - C Stajyer Araştırmacı Burs Programı (STAR) kapsamında TÜBİTAK'da çalışmaktadır. Arka-uç ve ön-uç yazılım geliştirme, veritabanı yönetim sistemleri, konteynerleştirme teknolojileri ve bulut bilişim teknolojilerinin yanı sıra yazılım testi üzerinde çalışmalar yapmaktadır. Proje süreci boyunca tüm iş paketlerinde aktif bir şekilde rol alacaktır.

Abdul Hannan Ayubi (Proje Ekip Üyesi): Eskişehir Osmangazi Üniversitesi, Bilgisayar Mühendisliği Bölümü son sınıf öğrencisidir. Halihazırda Yıldız Teknik Üniversitesi Teknoparkı'nda bulunan Tripenia Bilişim Turizm Seyahat Acentası ve Tic. A.Ş Şirketi'nde yazılım mühendisi olarak yarı zamanlı çalışmaktadır. Bunun yanı sıra Vilnius Üniversitesinde altı aylık staj dönemi boyunca yapay zekâ projesinde araştırma görevlisi olarak aktif rol almıştır. Proje süreci boyunca tüm iş paketlerinde aktif bir şekilde rol alacaktır.

Furkan Taşkın (Proje Ekip Üyesi): Eskişehir Osmangazi Üniversitesi, Bilgisayar Mühendisliği Bölümü son sınıf öğrencisidir. Halihazırda 2247 - C Stajyer Araştırmacı Burs Programı (STAR) kapsamında TÜBİTAK'da çalışmaktadır. Arka-uç, ön-uç ve mobil yazılım geliştirme, konteynerleştirme teknolojileri ve bulut bilişim teknolojileri yanı sıra yazılım testi üzerinde çalışmalar yapmaktadır. Proje süreci boyunca tüm iş paketlerinde aktif bir şekilde rol alacaktır.

7. EKLER

EK-1: Kaynaklar

- ELECTRONIC COMPONENTS AND SYSTEMS, 2022
- Sun, Youcheng, et al. "Testing deep neural networks." arXiv preprint arXiv:1803.04792 (2018).
- Grand View Research. Artificial Intelligence Market Size, Share & Trends Analysis Report By Solution, By Technology (Deep Learning, Machine Learning, Natural Language Processing, Machine Vision), By End Use, By Region, And Segment Forecasts, 2022 - 2030.
- IBM Global AI Adoption Index 2022, 2022
- Choi, C. Q. (2021). 7 Revealing Ways AIs Fail: Neural Networks can be Disastrously Brittle, Forgetful, and Surprisingly Bad at Math. IEEE Spectrum, 58(10), 42-47.
- Dastin Jeffrey. 2018. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. Ulaşılabilir: https://www.reuters.com.
- Siddiqui, F., Lerman, R., Merrill, J.B. (2022). Teslas running Autopilot involved in 273 crashes reported since last year. The Washington Post. Ulaşılabilir: https://www.washingtonpost.com.
- IBM, (2022). "IBM Global AI Adoption Index 2022" https://www.ibm.com/watson/resources/ai-adoption
- Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING

DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, 2022

- Boesch, G. (2021). "Deep Neural Network: The 3 Popular Types (MLP, CNN and RNN)" https://viso.ai/deep-learning/deep-neural-network-three-popular-types
- Sun, Y., Huang, X., Kroening, D., Sharp, J., Hill, M., & Ashmore, R. (2018). Testing deep neural networks. arXiv preprint arXiv:1803.04792.
- Pei, K., Cao, Y., Yang, J., & Jana, S. (2017, October). Deepxplore: Automated whitebox testing of deep learning systems. In proceedings of the 26th Symposium on Operating Systems Principles (s. 1-18).
- Google "Artificial Intelligence at Google: Our Principles" https://ai.google/principles/
- IBM, "Neural Networks" https://www.ibm.com/cloud/learn/neural-networks
- ISTQB® Certified Tester AI Testing (CT-AI) Syllabus, Sürüm 2021 V1.0 https://www.istqb.org/downloads/category/2-foundation-level-documents.html (erişildi: Kas 2022).
- Tian, Y., Pei, K., Jana, S., & Ray, B. (2018, May). Deeptest: Automated testing of deep-neural-network-driven autonomous cars. In Proceedings of the 40th international conference on software engineering (s. 303-314).
- Pei, K., Cao, Y., Yang, J., & Jana, S. (2017, Ekim). Deepxplore: Automated whitebox testing of deep learning systems. In proceedings of the 26th Symposium on Operating Systems Principles (s. 1-18).
- Odena, A., Olsson, C., Andersen, D., & Goodfellow, I. (2019, May). Tensorfuzz: Debugging neural networks with coverage-guided fuzzing. Uluslararası Makine Öğrenmesi Konferansında (s. 4901-4911). PMLR.