



# Reinforcement Learning ile Yapay Zekaya Hayatta Kalmayı Öğretme

Hüseyin Buğra Taştan

Mart 2024



---

## Özet

”AI Learns To Survive” adlı proje, derin güçlendirme öğrenme (deep reinforcement learning) kullanarak yapay zeka ajanlarının hayatta kalma becerilerini öğrenmeye odaklanan bir araştırma veya uygulama projesidir. Projede amaç, yapay zeka ajanlarını belirli bir ortamda hayatta kalmaya yönlendirmek ve bu süreçte karşılaştıkları zorluklarla başa çıkabilmelerini sağlamaktır.

---

# 1 Giriş

Günümüzde, yapay zeka teknolojisinin hızla gelişmesiyle birlikte, yapay zeka ajanlarının gerçek dünya ortamlarında başarılı bir şekilde işlev görmesi ve hatta hayatta kalabilmesi giderek daha büyük bir önem kazanıyor. "AI Learns To Survive" (Yapay Zeka Hayatta Kalmayı Öğreniyor) projesi, bu bağlamda derin güçlendirme öğrenme tekniklerini kullanarak yapay zeka ajanlarının hayatta kalma becerilerini geliştirmeye odaklanan heyecan verici bir araştırma ve uygulama girişimidir. Projemizin temel amacı, yapay zeka ajanlarını belirli bir ortamda hayatta kalmaya teşvik etmek ve bu süreçte karşılaşacakları çeşitli zorluklarla başa çıkabilmelerini sağlamaktır. Bu zorluklar, gerçek dünya senaryolarından esinlenmiş olabileceği gibi sanal ortamlarda da simüle edilebilir. Örneğin, bir robotun doğal afetler, engeller veya diğer tehlikelerle karşılaştığı bir ortamda hayatta kalabilme yeteneği üzerine odaklanabiliriz. Derin güçlendirme öğrenme algoritmaları, projemizin temelini oluşturur. Bu algoritmalar, yapay zeka ajanlarının ortama uyum sağlamak ve belirlenmiş hedefleri başarmak için optimal eylemleri öğrenmelerini sağlar. Öğrenme süreci genellikle bir ödül sistemiyle desteklenir; ajanlar belirli görevleri başarıyla tamamladıklarında veya belirli zorlukları aştıklarında ödüllendirilirler. Bu ödül sistemi, ajanların istenilen davranışları öğrenmelerini teşvik eder. "AI Learns To Survive" projesi, yapay zeka alanında önemli bir boşluğu doldurmayı hedefliyor. Gerçek dünya uygulamalarında kullanılabilen yapay zeka ajanlarının, değişen ve bazen de tehlikeli ortamlarda başarılı bir şekilde işlev görebilmesi, teknolojinin ilerlemesinde kritik bir adımdır. Bu proje, bu hedefe ulaşmada bir adım daha atmaktadır.

## 2 Literatür Çalışması

---

Bu makalede, derin pekiştirmeli öğrenme yöntemlerini kullanarak, yüksek boyutlu duyuşal girdilerden (örneğin, görüntü veya konuşma gibi) doğrudan ajanları kontrol etmeyi öğrenmektir. Peşin başarılı pekiştirmeli öğrenme uygulamaları, bu tür alanlarda genellikle elle hazırlanmış özelliklerle birleştirilmiş doğrusal değer fonksiyonları veya politika temsillerine dayanmıştır[1].

Bu makalede, derin pekiştirmeli öğrenme modellerini daha hızlı ve daha verimli bir şekilde eğitmek ve genişletmek için asenkron yöntemlerin etkin bir şekilde kullanılmasıdır. Bu sayede, daha karmaşık ve gerçekçi problemler üzerinde daha iyi performans gösteren yapay zeka sistemleri geliştirmek mümkün olabilir[2].

Bu projede, derin sinir ağları ve ağaç araması gibi ileri öğrenme tekniklerinin kullanılmasıyla, Go oyununu oynamak için daha etkili bir yaklaşım geliştirilmiştir. Derin sinir ağları, oyun tahtasındaki durumu analiz etmek ve olası hamleleri tahmin etmek için kullanılırken, ağaç araması algoritması, bu tahminleri daha ileri seviyeye taşımak ve olası hamlelerin sonuçlarını değerlendirmek için kullanılır[3].

Bu proje, derin takviyeli öğrenme tekniklerini kullanarak genel video oyunları için yapay zeka geliştirmeyi amaçlamaktadır. Derin takviyeli öğrenme, yapay zeka ajanlarının belirli bir görevi gerçekleştirmek için çevreleriyle etkileşimde bulunarak deneyimlerinden öğrenmelerini sağlayan bir makine öğrenimi yaklaşımıdır.[4]

## 3 Metodoloji

---

### 3.1 Ortam ve Görev Belirleme

İlk adım olarak, yapay zeka ajanlarının etkileşime gireceği bir ortam belirlenir. Bu ortam, gerçek dünya senaryolarını veya sanal ortamları simüle edebilir. Örneğin, bir doğal afet simülasyonu veya engellerle dolu bir alan seçilebilir. Ardından, ajanların karşılaştıkları belirli görevler ve hedefler tanımlanır. Bu görevler, ajanların hayatta kalma becerilerini geliştirmek için temel oluşturur.

### 3.2 Derin Güçlendirme Öğrenme Algoritmalarının Uygulanması

Projede genellikle derin güçlendirme öğrenme algoritmaları kullanılır. Örnek olarak, Deep Q-Networks (DQN) veya Proximal Policy Optimization (PPO) gibi algoritmalar tercih edilebilir. Bu algoritmalar, ajanların ortama uyum sağlamak ve belirlenmiş hedefleri başarmak için optimal eylemleri öğrenmelerini sağlar. Algoritmalar, ajanların mevcut durumu değerlendirerek en uygun eylemi seçmelerine yardımcı olur.

### 3.3 Ödül Sistemi Tasarımı

Projede bir ödül sistemi tasarlanır. Ajanlar, belirli görevleri başarıyla tamamladıklarında veya belirli zorlukları aştıklarında ödüllendirilirler. Ödüller, ajanların istenilen davranışları öğrenmelerini teşvik eder. Örneğin, hedefe ulaşma, engelleri aşma veya riskli durumlardan kaçınma gibi davranışlar ödüllendirilebilir.

### 3.4 Eğitim ve Deneyler

Ajanlar, belirlenen ortamda ve görevlerde eğitilir. Eğitim süreci, derin güçlendirme öğrenme algoritmaları kullanılarak gerçekleştirilir. Ajanlar, deneme yanılma yöntemiyle ortama uyum sağlamayı ve belirlenmiş hedefleri başarmayı öğrenirler. Ardından, eğitilen ajanların performanslarını değerlendirmek için çeşitli deneyler yapılır. Bu deneyler, ajanların gerçek dünya uygulamalarında başarılı olup olmadığını belirlemek için kullanılır.

### 3.5 Sonuçların Değerlendirilmesi ve İyileştirme

Projenin sonunda, elde edilen sonuçlar değerlendirilir. Ajanların performansı analiz edilir ve iyileştirme noktaları belirlenir. Bu noktalar, ajanların hayatta kalma becerilerini daha da geliştirmek için kullanılır. Gerekirse, algoritmalarda veya ödül sistemlerinde değişiklikler yapılabilir.

## 4 Sistem Tasarımı

---

### 4.1 Problem Belirleme

Projenin amacı ve hedefi netleştirilmelidir. AI'nın hangi ortamda hayatta kalmayı öğrenmesi gerektiği belirlenmelidir.

#### 4.1.1 Ortamın Belirlenmesi

AI'nın etkileşime gireceği ortamın belirlenmesi gerekmektedir. Bu ortam, bir simülasyon ortamı olabilir veya gerçek dünyada bir robotun deneyimleyeceği bir ortam olabilir.

### 4.2 Gözlem ve Aksiyon Uzayının Tanımlanması

Gözlemler, AI'nın ortamı algıladığı verilerdir. Bu, bir görüntü, sensör verisi veya başka bir tür veri olabilir.

Aksiyonlar, AI'nın ortama müdahale ettiği eylemlerdir. Bu, bir robotun hareketi, bir oyun karakterinin hareketi veya başka bir tür aksiyon olabilir.

### 4.3 Mevcut Verinin Toplanması veya Oluşturulması

Eğer bir simülasyon ortamı kullanılıyorsa, bu ortamda AI'nın eğitilmesi için yeterli miktarda veri toplanmalı veya oluşturulmalıdır.

### 4.4 Derin Öğrenme Modelinin Tasarımı

Gözlem verilerini girdi olarak alacak bir derin öğrenme modeli tasarlanmalıdır. Bu, genellikle evrişimli sinir ağları (CNN) veya rekürrent sinir ağları (RNN) gibi mimariler içerebilir.

Model, veriyi işleyip çıktı olarak aksiyonları veya aksiyon değerlerini üretmelidir.

### 4.5 Pekiştirmeli Öğrenme Algoritmasının Seçimi

Projenin gereksinimlerine uygun bir pekiştirmeli öğrenme algoritması seçilmelidir. Bu, Q-learning, Deep Q-Networks (DQN), Advantage Actor-Critic (A2C), veya Proximal Policy Optimization (PPO) gibi olabilir.

### 4.6 Model Eğitimi

Seçilen pekiştirmeli öğrenme algoritması kullanılarak derin öğrenme modeli eğitilmelidir. Bu süreçte, model ortamla etkileşime girer, gözlem ve ödül alır, ve bu veri üzerinden güncellenir.

#### **4.7 Modelin Değerlendirilmesi**

Eğitilen modelin performansı test edilmeli ve değerlendirilmelidir. Bu, modelin ortamda ne kadar iyi performans gösterdiğinin analizini içerir.

#### **4.8 Modelin Ayarlanması ve İyileştirilmesi**

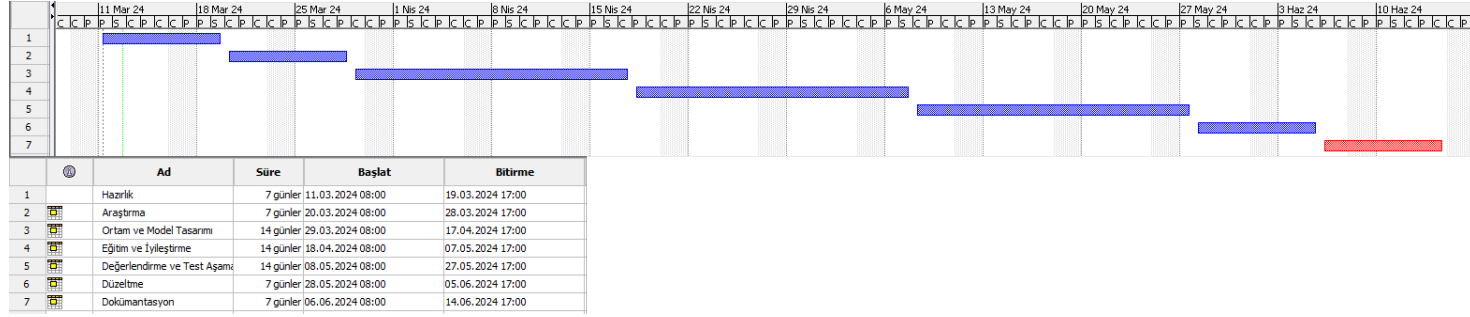
Modelin performansını artırmak için gerektiğinde ayarlamalar yapılmalı ve eğitim süreci tekrarlanabilir.

#### **4.9 Uygulama ve Dağıtım**

Eğitilen model, gerçek dünyada veya istenilen bir uygulama ortamında kullanılmak üzere dağıtılmalıdır.

#### **4.10 Sürekli Öğrenme ve İyileştirme**

Model, kullanıldıkça ve yeni veriyle beslendikçe sürekli olarak güncellenmelidir. Yeni veri ve deneyimlerle modelin performansı artırılmalıdır.



Resim 1: Gantt Chart

## 5 Beklenen Sonular

---

### 5.1 Hayatta Kalma Becerilerinin Gelişmesi

Projenin en temel beklenen sonuçlarından biri, yapay zekanın belirlenen ortamda hayatta kalmayı öğrenmesidir. Derin öğrenme algoritmaları, AI'nın ortama uyum sağlamasını ve tehlikeleri önceden tahmin etmesini sağlayarak hayatta kalma becerilerini geliştirecektir.

### 5.2 Esneklik ve Adaptasyon Yeteneđi

AI'nın öğrendiđi hayatta kalma stratejileri, deđişen ortam koşullarına uyum sağlamak için esnek olacaktır. Bu, AI'nın çeşitli senaryolara ve farklı koşullara hızlı bir şekilde uyum sağlayabilme yeteneđini içerecektir.

### 5.3 Optimize Edilmiş Karar Alma Yeteneđi

Pekiştirmeli öğrenme algoritmaları, AI'nın belirlenen hedeflere ulaşmak için en uygun kararları almasını sağlayacaktır. Bu, kaynakları verimli kullanma, riskleri minimize etme ve belirlenen hedeflere ulaşma konularında optimize edilmiş karar alma yeteneđi anlamına gelir.

### 5.4 Belirsizlik ve Karmaşıklıkla Başa Çıkma Yeteneđi

Proje, AI'nın belirsizlik ve karmaşıklıkla başa çıkma yeteneđini geliştirecek ve öngörülemeyen durumlarla nasıl başa çıkacağını öğretecektir. Bu, AI'nın deđişen ortam koşullarında tutarlı bir şekilde performans göstermesini sağlayacaktır.

### 5.5 Transfer Öğrenme ve Genelleme Yeteneđi

AI, öğrendiđi hayatta kalma stratejilerini farklı ortamlara ve senaryolara transfer edebilecek ve bu stratejileri genelleştirebilecektir. Bu, AI'nın bir ortamda öğrendiklerini başka bir ortama veya görev üzerine uygulayabilme yeteneđi anlamına gelir.

### 5.6 Sürdürülebilir Performans ve Güvenilirlik

Eğitilen modelin sürdürülebilir performansı ve güvenilirliđi, projenin önemli sonuçlarından biridir. AI'nın belirlenen görevleri başarıyla yerine getirebilmesi ve sürekli olarak istikrarlı bir şekilde performans gösterebilmesi beklenmektedir.



## 5.7 Endüstriyel ve Ticari Uygulamalar

Elde edilen sonuçlar, çeşitli endüstriyel ve ticari uygulamalarda kullanılabilir. Örneğin, robotik sistemlerde, otonom araçlarda, oyun geliştirme alanında ve daha birçok alanda AI'nın hayatta kalma becerilerini geliştirmek için kullanılabilir.

## Kaynakça

- [1] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, “Playing atari with deep reinforcement learning,” *arXiv preprint arXiv:1312.5602*, 2013.
- [2] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, “Asynchronous methods for deep reinforcement learning,” in *International conference on machine learning*, pp. 1928–1937, PMLR, 2016.
- [3] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, *et al.*, “Mastering the game of go with deep neural networks and tree search,” *nature*, vol. 529, no. 7587, pp. 484–489, 2016.
- [4] R. R. Torrado, P. Bontrager, J. Togelius, J. Liu, and D. Perez-Liebana, “Deep reinforcement learning for general video game ai,” in *2018 IEEE Conference on Computational Intelligence and Games (CIG)*, pp. 1–8, IEEE, 2018.