



KÜTAHYA
SAĞLIK BİLİMLERİ
ÜNİVERSİTESİ

REINFORCEMENT LEARNING TEKNIĞINI KULLANARAK YAPAY ZEKA TARAFINDAN YÖNETİLEN BİR ARABAYA PISTİ TAMAMLATMAK

Yazar: Mert Koca
March 14, 2024



Özet

Bu proje, yapay zeka kullanarak bir arabanın otomatik olarak bir pisti tamamlamasını sağlamayı hedefler. Yapay zeka, görsel girdi olarak pistin harmasını alacak ve aracı uygun şekilde yönlendirerek pisti en kısa sürede tamamlamaya çalışacak. Bu projede reinforcement learning (teşvik öğrenimi) algoritması kullanılarak gerçekleştirilecek. Proje, yapay zeka temelli oyun geliştirme, teşvik öğrenimi ve simülasyon gibi alanlara odaklanacak.

1 Giriş

Reinforcement learning algoritması, bir ajanın çevresiyle etkileşime geçerek deneyimlerinden öğrenmesini sağlar ve bu öğrenme sürecini en iyi şekilde nasıl yönlendireceğini belirler. Araba yarışı, özellikle karmaşık çevrelerde hızlı ve doğru kararlar gerektiren bir projedir. Bu proje, reinforcement learning'in gücünü araç yarışı simülasyonlarıyla bireleştirerek, bir aracın yarış pistinde optimum performansı elde etmesi için nasıl öğrenebileceğini araştırmayı amaçlamaktadır. Bu projenin başarısı, gerçek dünya uygulamalarında otonom araçların geliştirilmesi gibi birçok alana ilham verebilir.

2 Literatür Çalışması

Ibrahim Sobh Ibrahim'in yürüttüğü araştırma[1], Reinforcement Learning (RL) algoritmalarını kullanarak ajanların dinamik ortamlarda hızlı bir şekilde öğrenmesini ve adapte olmasını vurgulamaktadır. Derin RL'yi kullanarak, çalışma, özellikle müzakere gerektiren ve dinamik etkileşimler gerektiren senaryolarda otonom ajanların karar verme yeteneklerini artırmayı amaçlamaktadır.

Ayrıca, çalışma, RL ajanlarının öğrenme sürecini yönlendirmek için etkili ödül fonksiyonlarının tasarlanmasıının önemini vurgulamaktadır. Literatürde tartışıldığı gibi, ödül şekillendirme, istenen davranışlara dayalı olarak ajanlara ek rehberlik sağlayarak öğrenme sürecini hızlandırmada kritik bir rol oynamaktadır. Bu yaklaşım, ajanların ortamla etkileşimler yoluyla kümülatif ödülleri maksimize etmeye çalıştığı RL'nin temel prensibiyle uyumludur.

Son yıllarda derin öğrenme ve özellikle de Derin Pekiştirmeli Öğrenme (DPO) alanındaki ilerlemeler[2], otonom sürüs sistemlerinin geliştirilmesinde önemli bir rol oynamaktadır. El Sallab ve diğerleri (2017), otonom sürüs için bir Derin Pekiştirmeli Öğrenme çerçevesi önermişlerdir. Bu çerçeve, Reküran Sinir Ağları (RSA) ve dikkat modellerini içерerek karmaşık sürüs senaryolarıyla başa çıkmakta ve hesaplama karmaşıklığını azaltmaktadır. Araştırmacılar, bu çerçevenin TORCS simülatöründe test edildiğini ve başarılı öğrenme sonuçları elde edildiğini belirtmektedirler. Ayrıca, önerilen çerçeve, RNN'leri kısmen gözlemlenebilir senaryoları ele almak için entegre etmektedir. Bu çalışma, otonom sürüs sistemlerinin geliştirilmesinde derin öğrenme tekniklerinin etkili bir şekilde kullanılmasını vurgulamaktadır.

flyyufelix adlı github kullanıcısının yaptığı "donkeyrl" projesinde[3] Donkey Car adı verilen bir arabanın bir pist etrafında RL kullanarak kendi kendine dönmesini sağlamıştır. Bu projede Double Deep Q Learning (DDQN) kullanılmıştır. DDQN, bir ajanın çevresiyle etkileşime girerek belirli bir durumda alabileceği eylemler arasında en uygun olanını seçmeyi öğrenen bir güçlendirme öğrenme algoritmasıdır. Bunu, bir Q-değer fonksiyonu kullanılarak gerçekleştirir. Q-değer fonksiyonu, bir durum ve bir eylem çifti için beklenen toplam ödülu temsil eder.

3 Yöntemler

Bu projede, reinforcement learning algoritmalarını kullanarak bir aracın yarış pistinde optimal performansı öğrenmesini sağlamak için bir metodoloji izlenecektir. İlk olarak, yarış simülasyonu için uygun bir ortam oluşturulacak ve aracın bu ortamda hareket etmesi sağlanacaktır. Daha sonra, aracın davranışlarını kontrol etmek için bir reinforcement learning algoritması seçilecek ve bu algoritma modeli, aracın çevresini algılamasını ve çevresel faktörlere tepki vermesini sağlayacak şekilde ayarlanacaktır. Ardından, aracın deneyimlerinden öğrenerek, yarış pistinde en iyi performansı elde etmesini sağlamak için algoritma sürekli olarak güncellenecek ve iyileştirilecektir. Son olarak, algoritmanın eğitimi tamamlandıktan sonra, aracın gerçek zamanlı olarak yarış pistinde nasıl davranışlığını test etmek için simülasyonlar üzerinde çeşitli denemeler yapılacak ve sonuçlar analiz edilecektir. Bu metodoloji, reinforcement learning'in araç yarışı gibi karmaşık görevlerde nasıl başarılı bir şekilde uygulanabileceğini anlamak için bir çerçeve sağlayacaktır.

Bu projeyi uygulamaya geçireceğim platform olarak Unity oyun motorunu seçtim. Arayüz kullanılabilirliği, uygulama anlaşılmabilirliği açısından en iyi performansı elde eteceğimi düşünüyorum. Kodlama açısından C Sharp dilini kullanacağım için Unity içerisinde Visual Studio Code uygulamasını kullanacağım. Arabanın kontorları, zamanlayıcı gibi kodlama gerektiren kısımları burada kodlayacağım.

Bu projede kullanılacak pist ve checkpoint noktaları aşağıda verilmiştir. Checkpointler tur zamanı tutmayı kolaylaştırırken ilerde ödül sisteminin bir parçası olacak



Figure 1: Harita

Arabanın etrafında döneceği pist



Figure 2: Kontrol Noktaları

Ödül sistemi olarak da kullanıldıkları checkpointler

4 Veri Odaklı Sistem Tasarımı

4.1 Veri Toplama ve İşleme Bileşeni

Sensörler aracılığıyla pistin gerçek zamanlı haritasını oluşturacak verileri toplama. Kullanılacak kamera veya diğer algılayıcılarla çevrenin görüntülerini toplama.

4.2 Yapay Zeka ve Öğrenme Bileşeni

Veri odaklı algoritma seçimi, Reinforcement Learning (teşvik öğrenimi) algoritması kullanılacak. Modelin eğitilmesi: Toplanan veriler üzerinde modelin eğitilmesi, pisti tamamlama görevini öğrenmesi. Modelin güncellenmesi: Sürekli olarak yeni verilerle modelin güncellenmesi ve iyileştirilmesi.

4.3 Karar ve Yönlendirme Bileşeni:

Model tarafından verilen kararların alınması ve işlenmesi. Arabanın mevcut konumu ve pist haritası dikkate alınarak bir sonraki adımın belirlenmesi. Arabanın hareketini yönlendirme: Hız, yönlendirme açısı vb. ayarlama.

4.4 Simülasyon ve Test Bileşeni

Geliştirilen sistemin simülasyon ortamında test edilmesi. Farklı senaryolarda ve gevresel koşullarda sistemin performansının değerlendirilmesi. Modelin güvenilirliğinin ve dayanıklığının test edilmesi.

4.5 Görselleştirme ve Raporlama Bileşeni

Sistemin çalışmasıyla ilgili verilerin görselleştirilmesi. Performans metriklerinin ve sonuçların raporlanması. Karar vericilere veya paydaşlara sunumlar için raporlar hazırlama.

Veri odaklı sistem tasarımı, bir sistemin geliştirilmesi veya iyileştirilmesi sürecinde veriye dayalı bir yaklaşımı benimseyen bir metodolojidir. Bu yaklaşım, karar alma süreçlerinin verilere dayandığı, performansın sürekli olarak izlendiği ve geri bildirimlerin değerlendirildiği bir çerçeve sunar.

5 Proje Aşamaları

5.1 Hazırlık ve Araştırma

Tesvik öğrenimi ve reinforcement learning konseptleri incelenecek. Gerekli yazılım araçlarını ve kütüphaneleri (Unity, C Sharp, vs.) yüklenip yapılandırılacak. Basit bir araba simülasyonu oluşturulacak ve Unity'de çalışır hale getirilecek..

5.2 Ortam ve Veri Toplama

Pistin çevresi ve düzeni gibi ortamı oluşturulacak. Veri toplama için bir veri toplama arayüzü oluşturulacak ve araba simülasyonunu bu arayüz üzerinden kullanarak veri toplanacak. Toplanan verileri düzenlenip işlenecek, uygun formatlara dönüştürülecek.

5.3 Teşvik Öğrenme Algoritmasının Uygulanması

Seçilen bir teşvik öğrenme algoritması (örneğin, Q-learning, Deep Q-Networks, vs.) uygulanacak. Algoritma, veri kümesi üzerinde eğitilecek ve başlangıçta basit bir ortamda test edilecek. Algoritmanın performansını değerlendirilecek ve gerektiğinde ayarlamalar yapılacak.

5.4 Gelişmiş Simülasyon ve Ayarlamalar

Daha karmaşık pistler, engeller ve dinamik ortamlar eklenecek. Teşvik öğrenme algoritması bu geliştirilmiş simülasyon üzerinde yeniden eğitilecek. Algoritmanın performansını iyileştirmek için parametreler ayarlanacak.

5.5 Son Dokunuşlar ve Değerlendirme

Kullanıcı arayüzü eklenecek, gerektiğinde grafiksel iyileştirmeler yapılacak. Proje test edilecek ve hata ayıklama yapılacak. Sonuçları ve elde edilen son performans değerlendirilecek. Projenin dokümantasyonunu hazırlanacak.

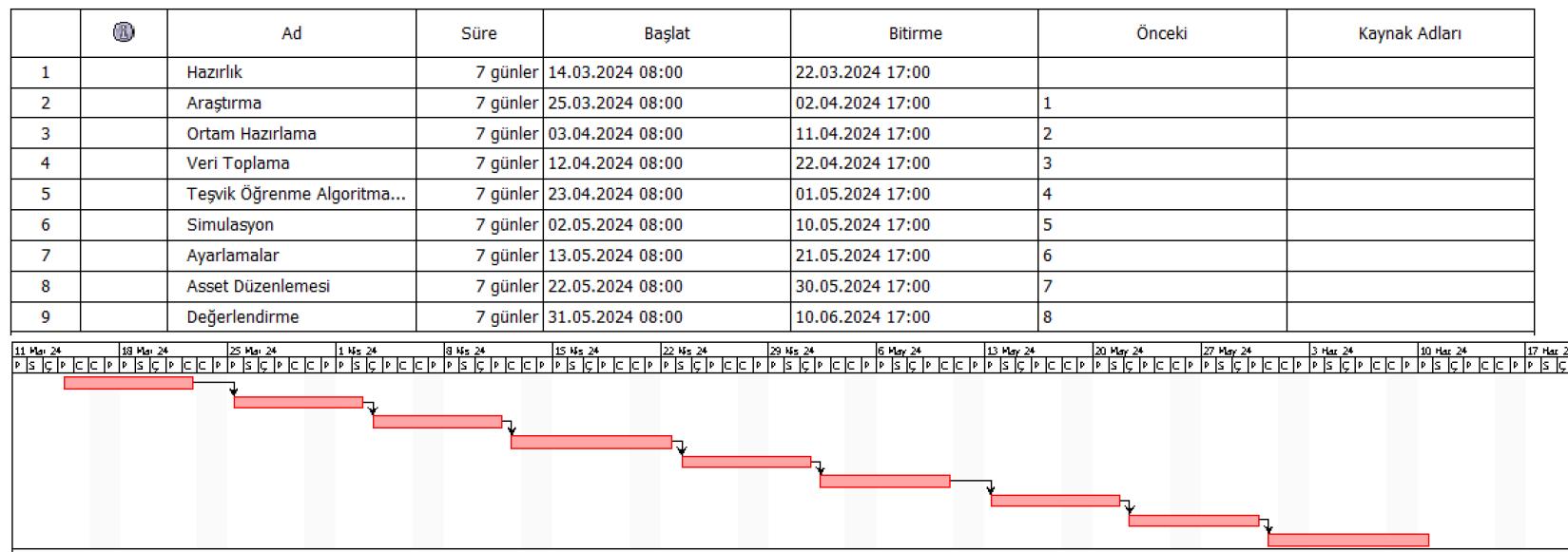


Figure 3: Gantt Chart

Kaynakça

- [1] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. Al Sallab, S. Yogamani, and P. Pérez, “Deep reinforcement learning for autonomous driving: A survey,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 4909–4926, 2021.
- [2] A. E. Sallab, M. Abdou, E. Perot, and S. Yogamani, “Deep reinforcement learning framework for autonomous driving,” *arXiv preprint arXiv:1704.02532*, 2017.
- [3] flyyufelix, “Train donkey car in unity simulator with reinforcement learnin,” *github*, 2018.