PEKİŞTİRMELİ ÖĞRENME ALGORİTMASI İLE TAKIM OYUNU OYNAYABİLEN YAPAY ZEKA GELİŞTİRME

BİTİRME PROJESİ 3

Furkan Kaya

191216002

## 

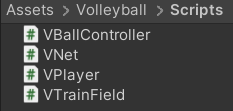
## DANIŞMAN Öğr. Gör. Ezgi Özer

## **ÖZET**

Daha önceki raporda bahsedilen ortam kurulumu bitirilmiş olup gerekli ödül/ceza fonksiyonlarının tanımlanması yapılmıştır. Ardından farklı parametreler ile eğitimler çalıştırılmış ve 1-1.5 haftalık eğitim sonucunda yaklaşık %80 başarı oranı sağlanmıştır. Oyuncular topu iyi bir şekilde kontrol etmeseler de topa vurmaları gerektiğini başarıyla öğrenmiş durumdalar.

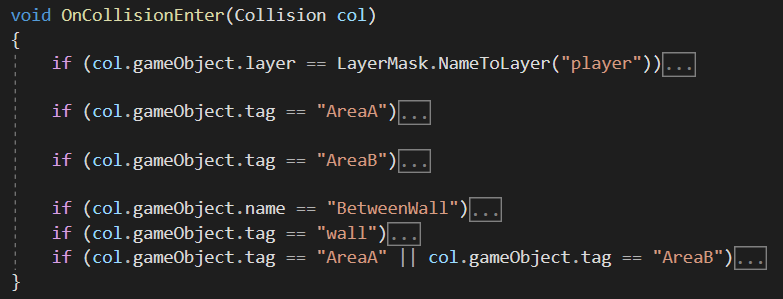
Projenin son haline ve dosyalarına <https://github.com/Wijt/graduation-project/> adresinden ulaşabilirsiniz.

# **Kodlar**

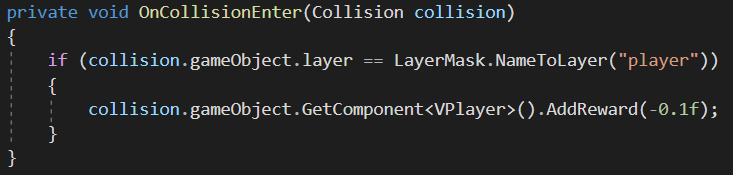


Eğitim sahasını oluşturabilmek için 4 sınıf kodlanıp oyun objelerine atanmıştır.

## Ball Controller:

Ödül fonksiyonunun büyük bir kısmının olduğu sınıftır. Topun hareketlerini, çarptığı oyuncuları ve atıldığı yönü kontrol ederek oyuncuya ya da takıma ödül/cezalar verilmiştir. 

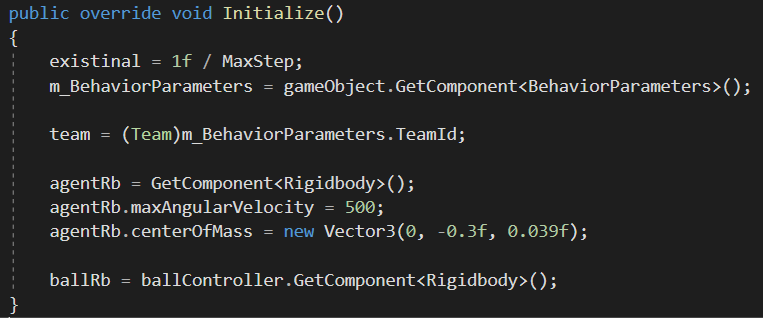
## Net:

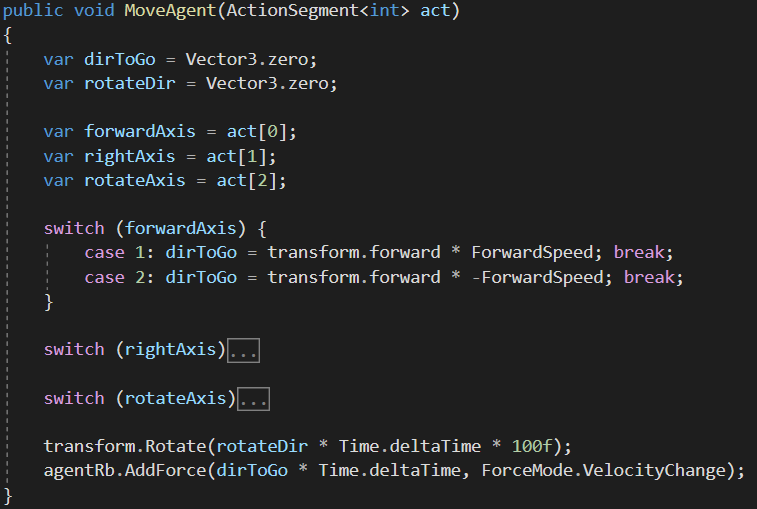
Bu sınıfın MonoBehaviour adlı Unity oyun objesi sınıfından türetilmiş bir sınıftır. Herhangi bir objenin çarpması durumunda çalışan OnCollisionEnter fonksiyonunun içine dokunan oyuncuya 0.1 ceza puanı verecek kodlar yazılmıştır.

## VPlayer:

ML-Agents kütüphanesinin içinde bulunan Agent sınıfından türetilmiş ve yapay zekanın giriş çıkışlarından sorumlu olan sınıftır.

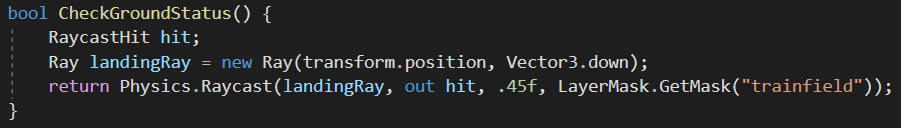
**Initialize():** Eğitim ilk başladığında çalışıp daha sonra çalışmayacak olan script için gerekli tanımlamaların yapıldığı fonksiyon.



**MoveAgent():** Verilen axis değerlerine göre hareket sağlayan fonksiyon.

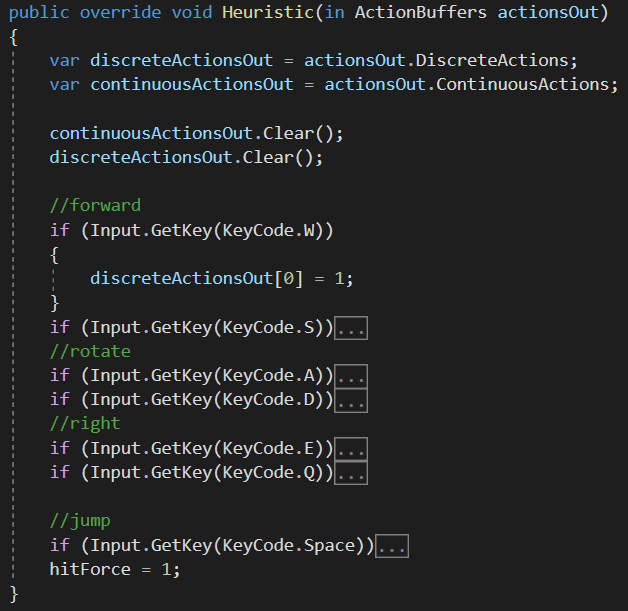
**CheckGroundStatus():** Oyuncunun altından gönderilen ışın zemine

değerse true değmez ise false değer döndüren, oyuncunun yerde mi havada mı olduğunu kontrol eden fonksiyon.

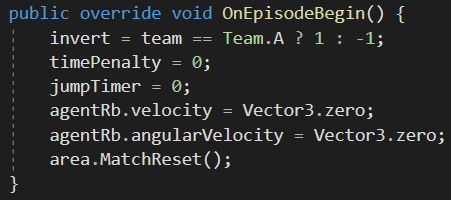


**OnActionReceived():** Agent sınıfından override edilmiştir. Fonksiyonun değişkeni olan actionBuffers içine yapay zekadan aldığımız çıktılar atanmaktadır.

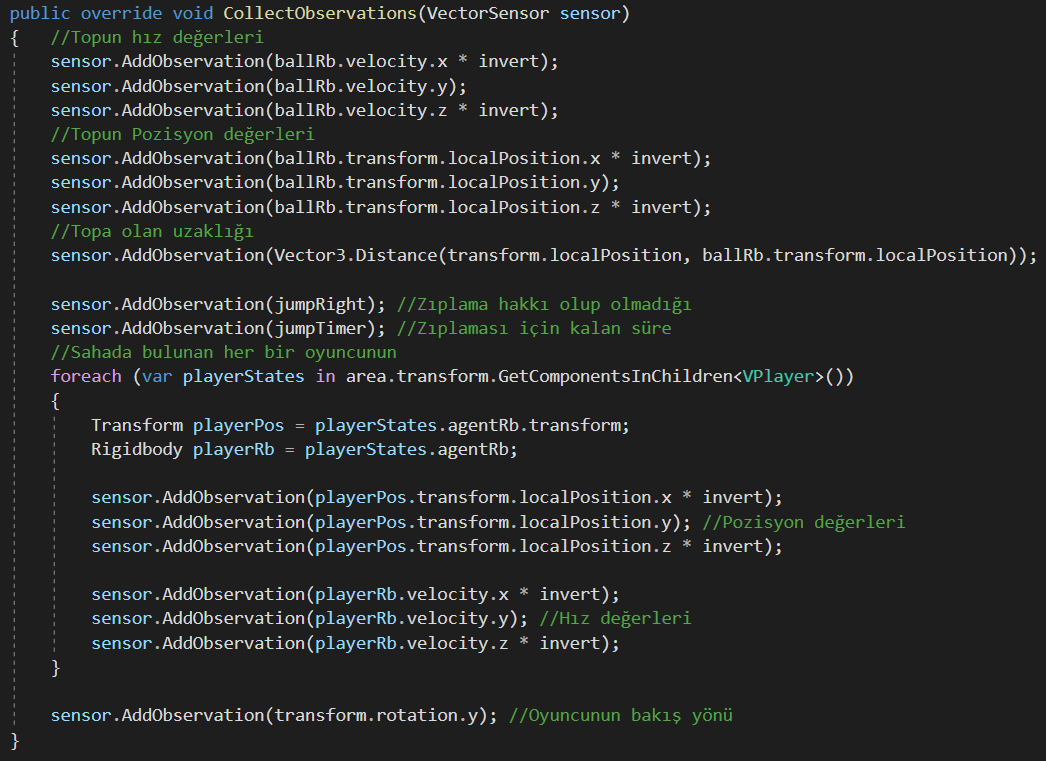


**Heuristic():** Yapay zekanın ürettiği sonuçları kendimiz vererek yazdığımız kodları test edip yapay zekayı kontrol edebileceğimiz bir fonksiyon.

**OnEpsidoBegin():** Her bölüm başında çalışacak fonksiyonlar oyuncunun değerlerini resetlemek için kullanıyoruz.



**CollectObservation():** Bu fonksiyon yapay zekanın giriş işlemlerinden sorumlu olan fonksiyondur. Agent sınıfındaki CollectObservation fonksiyonu override edilerek bir önceki raporda belirtilen giriş

değerlerimiz yapay zekaya iletilmektedir.

## Trainfield:

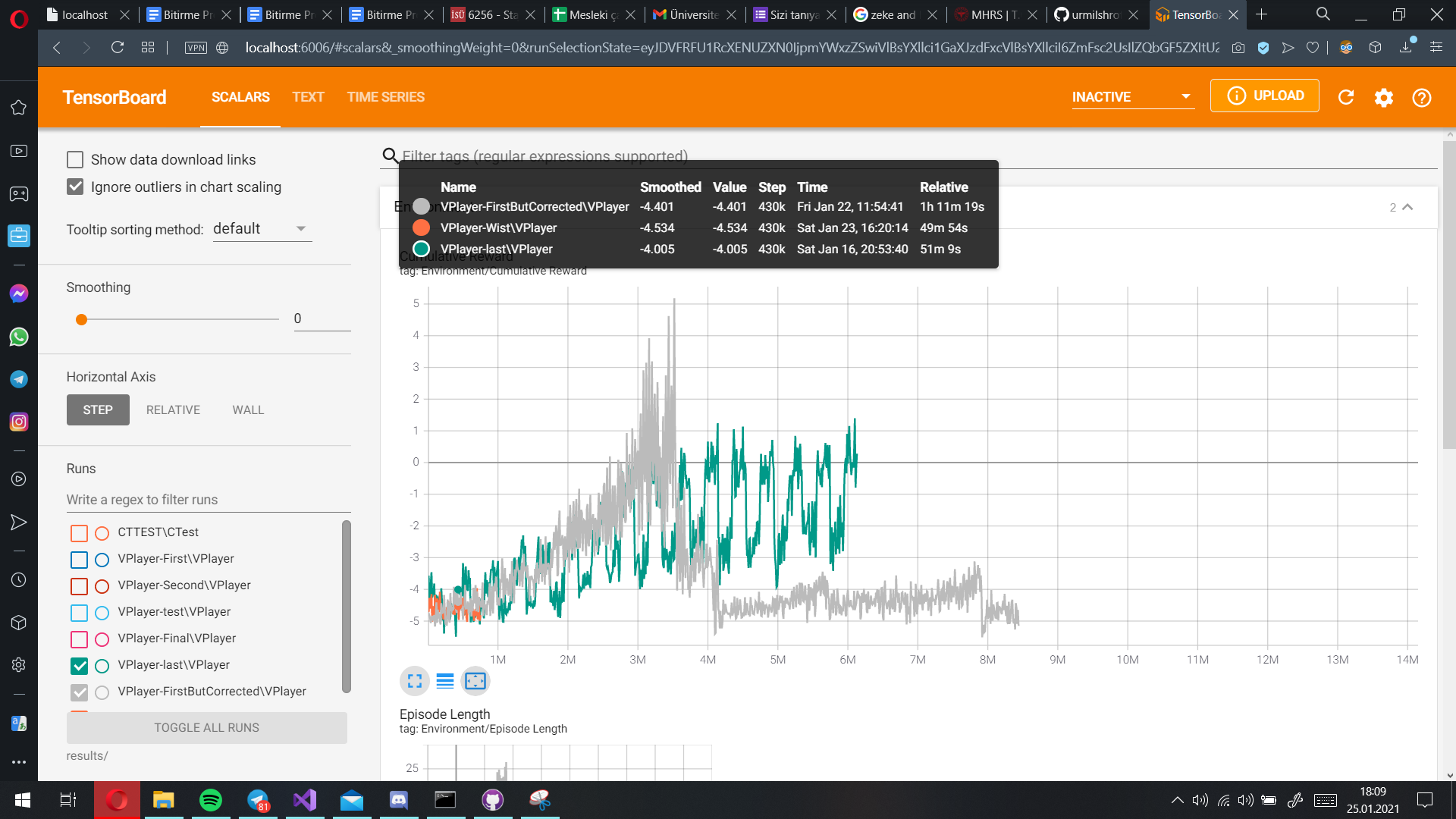
Oyuncuların bilgisini tutan, bölümü bitirip yeniden başlatmakla sorumlu olan sınıftır. MonoBehaviour sınıfından türetilmiştir.

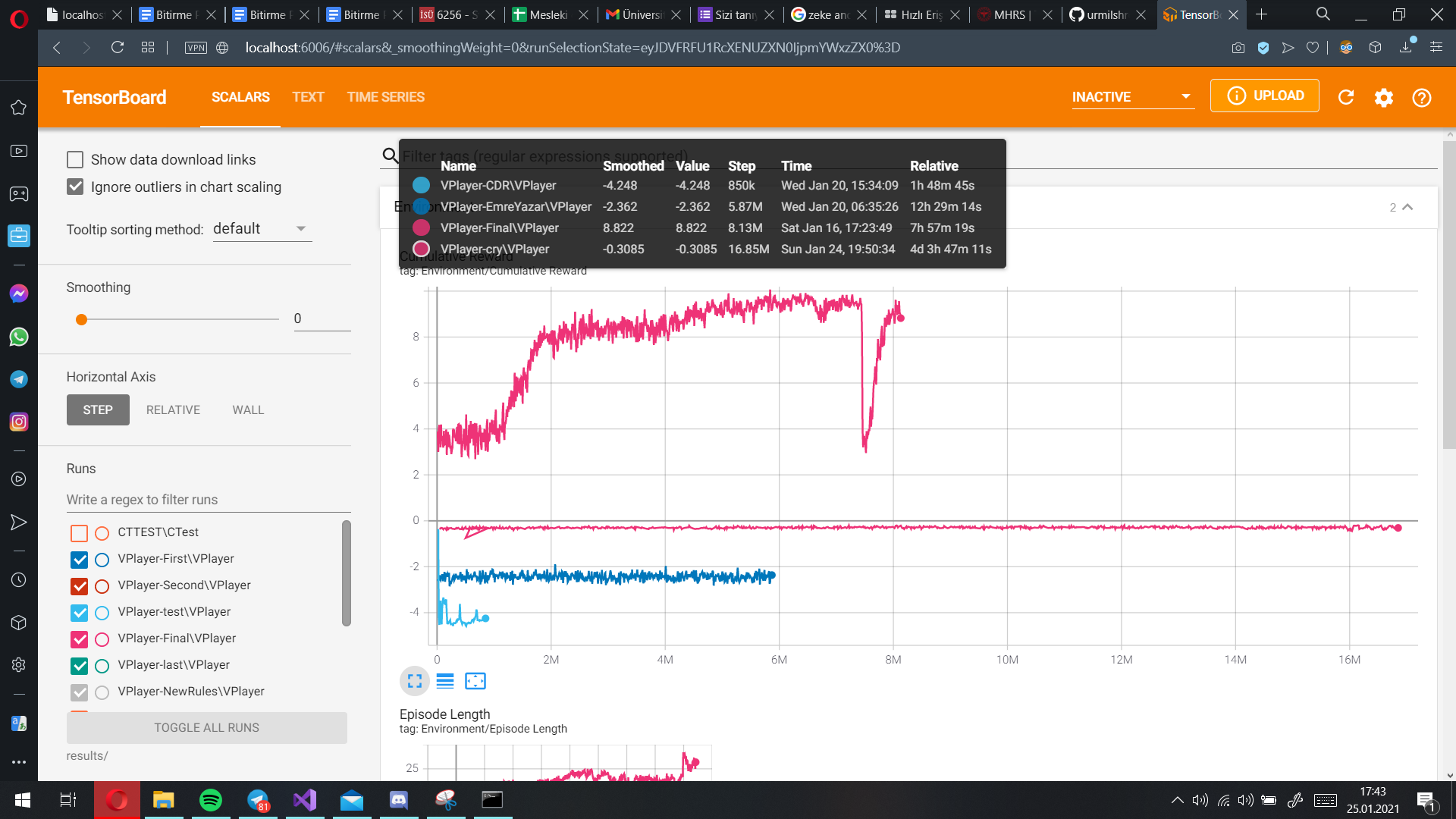


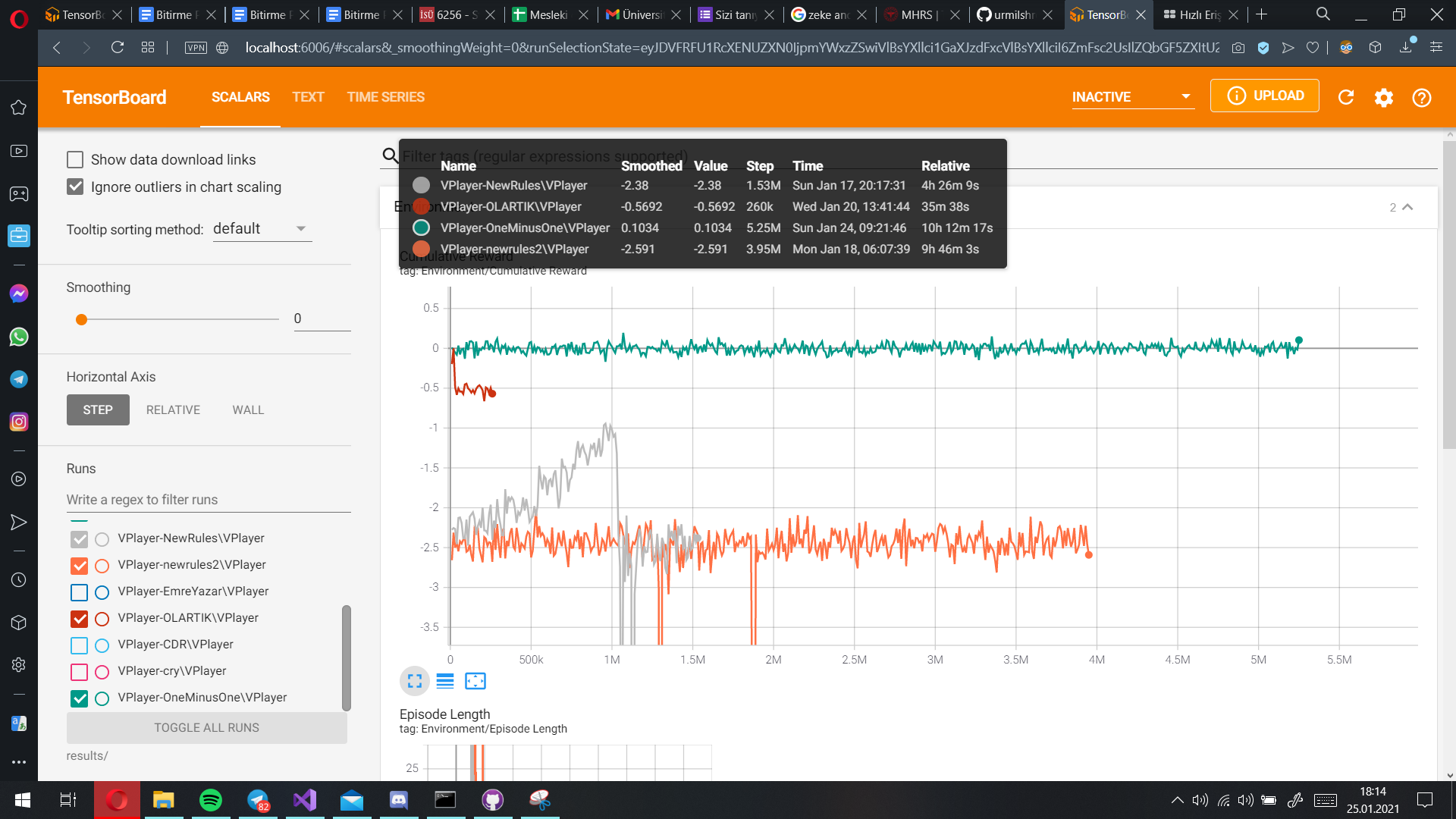
MonoBehaviour sınıfında; Awake fonksiyonu, Start fonksiyonundan önce çalışmaktadır. Start fonksiyonumuz maçı başlangıç pozisyonuna getirmekten sorumludur. Bu sebeple de start fonksiyonunda ihtiyacımız olan tanımlamaları awake fonksiyonunda yazıyoruz.

# **Eğitim Süreci**

Olası en iyi yapay zekayı bulmak üretebilmek için defalarca kez farklı ödül fonksiyonu ve farklı parametreler ile eğitim verilmiştir. Aşağıda kayıt altına alınan eğitim süreçlerinin grafikleri verilmiştir. Eğitim sürecinde 15 farklı çalışma ID’si oluşturulmuş ve toplam 90 saatten fazla eğitim verilmiştir. Aşağıdaki grafiklerde yapılan eğitim denemelerini inceleyebilirsiniz.







Ne yapılırsa yapılsın yapay zeka local bir minimuma inip oradan çıkmayı başaramadı.

Gelecek çalışmalarda genetik algoritma kullanılarak ve eğitim süresi arttırılarak geliştirme yapılması planlanmaktadır.

# 

# 

# Sonuç:

Ulaşılan maksimum başarıda oyuncumuz 10 servis kullanımının 8’inde karşı sahaya topu atıp puan almayı başarmıştır. 10 servisten 1 inde ise rakipten gelen topu karşılamayı başarmıştır. Simülasyonunun karmaşıklığı nedeniyle eğitimin yeterli gelmediği düşünülmektedir. Eğitim süresi artırılırsa daha da güçlü bir yapay zekaya ulaşılabilir.

## 

## Başarı sağlanan parametreler:

### Config.yaml:

behaviors:

VPlayer:

trainer\_type: ppo

hyperparameters:

batch\_size: 2048

buffer\_size: 20480

learning\_rate: 0.0003

beta: 0.005

epsilon: 0.2

lambd: 0.95

num\_epoch: 3

learning\_rate\_schedule: linear

network\_settings:

normalize: false

hidden\_units: 256

num\_layers: 8

vis\_encode\_type: simple

reward\_signals:

extrinsic:

gamma: 0.99

strength: 1.0

keep\_checkpoints: 50

max\_steps: 50000000

time\_horizon: 1000

summary\_freq: 10000

threaded: false

self\_play:

save\_steps: 15000

team\_change: 100000

swap\_steps: 2000

window: 10

play\_against\_latest\_model\_ratio: 0.5

initial\_elo: 1200.0

### Ödül fonksiyonu:

* Top takımın sahasına düşerse -0.75
* Top rakip takımın sahasına düşerse ve takım top üzerinde vuruş sahibiyse +1
* Topa dokunuş +0.1
* Topa çift dokunuş -0.1
* Topu fırlatmak +0.2
  + Doğru yöne fırlatmak +0.1
  + Yanlış yöne fırlatmak -0.1
* Oyuncunun fileye değmesi -0.1
* Topun rakip sahaya geçmesi +0.4
* Dokunulan topun duvara değmesi -0.02

Eğitim süresi ~36 saat.