

T.C.
SAKARYA ÜNİVERSİTESİ
BİLGİSAYAR VE BİLİŞİM BİLİMLERİ FAKÜLTESİ
BİLİŞİM SİSTEMLERİ MÜHENDİSLİĞİ

TİPTA YAPAY ZEKA PROJE TASARIM RAPORU

“HAFIZA TESTİ SONUÇLARINA DAYALI ANTİ-ANKSİYETE İLAÇ DOZU STRATEJİSİNİN REINFORCE ALGORİTMASI İLE ÖĞRENİLMESİ”

HAZIRLAYAN

Esra AYVAZ
B211200045

DANIŞMAN

Dr. Öğr. Üyesi Fatma AKALIN

ARALIK 2025

İÇİNDEKİLER

ÖZET	1
Veri Seti	1
Çalışmanın Kapsamı	1
YÖNTEM	1
Keşifsel Veri Analizi (EDA)	1
Veri Görselleştirme	2
Veri Önisleme	2
Modelleme ve Reinforcement Learning Süreci	2
Optimizasyon	3
Model Performans Analizi	3
DENEYSEL SONUÇLAR	4
Genel Veri Seti Bulguları	4
Veri Görselleri	4
Model Performansı ve Test Bulguları	6
DEĞERLENDİRME	8
Veri Setinin Yorumu	8
Model Performansının İncelenmesi	8
TARTIŞMA	9
Çalışmanın Çıkarımları	9
Sonuç	9

ÖZET

Veri Seti

Bu çalışmada anti-anksiyete ilaçlarının hafıza performansı üzerindeki etkilerini incelemek amacıyla “Memory Test on Drugged Islanders” veri seti kullanılmıştır. Veri setinde katılımcıların isim–soyisim, yaşı (age), uygulanan ilaç dozu (Dosage), ilaç türü (Drug) ve hafıza performansına ilişkin ön-test (memory_before) ile son-test (memory_after) değerleri bulunmaktadır. Katılımcılar ayrıca duygusal durumlarına göre Happy ve Sad olmak üzere iki deneysel gruba ayrılmıştır. Bu yapı, farklı yaş grupları ve duygusal durumların ilaç dozuna nasıl tepki verdığını; aynı zamanda ön-test ve son-test karşılaştırmaları üzerinden hafıza performansındaki değişimi gösteren Diff değişkeni aracılığıyla değerlendirmeyi mümkün kilmaktadır. Böylece çalışma, ilaç dozunun ve duygusal durumun hafıza performansı üzerindeki etkilerini bütüncül bir şekilde analiz etmek için uygun bir temel sunmaktadır.

Çalışmanın Kapsamı

Bu çalışmada, belirli bir hasta grubunda ilaç dozajı kararlarının optimize edilmesi amacıyla offline reinforcement learning tabanlı bir model geliştirilmiştir. Çalışmada öncelikle veri seti incelenmiş ve temel istatistiksel analizler ile veri ön işleme adımları uygulanmıştır. Daha sonra, ajan–çevre etkileşimini simüle edebilmek için özel bir ortam oluşturulmuş ve ajan, geçmiş veri üzerinden öğrenen bir politika ağı kullanılarak eğitilmiştir. Eğitim sürecinde toplam ödüller ve uzun vadeli getiriler takip edilerek modelin performansı değerlendirilmiş, sonuçlar ortalama ödül ve denge süresi gibi metriklerle raporlanmıştır. Genel olarak çalışma, veri temelli bir yaklaşım ile ilaç dozajı kararlarının iyileştirilmesini ve offline RL yöntemlerinin uygulanabilirliğini ortaya koymayı hedeflemektedir.

YÖNTEM

Keşifsel Veri Analizi (EDA)

Çalışmanın ilk adımımda veri seti `read_excel()` fonksiyonu ile Python ortamına aktarılmıştır. Veri yüklenikten sonra `df.shape` komutu kullanılarak satır–sütun bilgisi kontrol edilmiştir. Ardından `df.head()` ile ilk satırlar görüntülenmiş ve değişkenlerin genel yapısı incelenmiştir.

Veri setine ait tanımlayıcı istatistikleri görmek için `df.describe(include='all')` çalıştırılmış; sayısal ve kategorik değişkenlerin temel istatistikleri gözden geçirilmiştir. Son olarak `df.info()` çıktısı alınarak veri tipleri ve eksik değer durumu kontrol edilmiştir.

Veri setindeki eksik değerler incelenmiştir. Bunun için her sütundaki eksik gözlem sayıları ve eksik değer yüzdeleri hesaplanmıştır. Daha sonra sütunlardaki tekrarlayan veya düşük çeşitliliğe sahip kategorileri görmek amacıyla her değişkenin benzersiz değer sayısı (`nunique`) hesaplanmıştır.

Sayısal değişkenlerde üç değer olup olmadığını belirlemek için IQR (Interquartile Range) yöntemi kullanılmıştır. Her sayısal sütun için birinci çeyrek (Q1), üçüncü çeyrek (Q3) ve IQR değeri hesaplanmış; alt ve üst sınırların dışında kalan değerler tespit edilerek sütun bazında üç değer sayıları elde edilmiştir.

Veri Görselleştirme

Bu adımda veri setindeki sayısal değişkenler arasındaki ilişkileri incelemek için bir korelasyon matrisi oluşturulmuştur. Sayısal sütunlardan oluşan matris görselleştirilmiş ve değişkenlerin birbiriyle pozitif veya negatif ilişki eğilimleri gözlemlenmiştir.

Ardından her sayısal değişkenin dağılımını görmek amacıyla histogram ve boxplot grafikleri oluşturulmuştur. Histogramlar verinin dağılım şeklini (simetrik, sağ/sol çarpık vb.) incelemek için; boxplot'lar ise olası uç değerleri görsel olarak tespit etmek için kullanılmıştır.

Veri Önisleme

Bu adımda veri setine çeşitli ön işleme işlemleri uygulanmıştır. İlk olarak her katılımcıyı benzersiz şekilde temsil etmek için bir ID sütunu oluşturulmuş ve analizde gereksiz olan first_name ve last_name sütunları veri setinden çıkarılmıştır.

Daha sonra veri yapısını düzenlemek amacıyla değişkenler kategorik ve sayısal olarak ayrılmıştır. Sayısal sütunlar belirlendikten sonra uç değerlerin etkisini azaltmak amacıyla tüm sayısal değişkenlere Winsorize işlemi uygulanmıştır. Bu işlem, aşırı uç değerleri belirli yüzdelik dilimlere sabitleyerek dağılımı daha dengeli hâle getirmiştir.

Uç değer azaltma sonrasında sayısal değişkenler RobustScaler ile ölçeklendirilmiştir. Bu yöntem, medyan ve IQR değerlerine dayandığı için uç değerlere karşı dayanıklıdır ve modelleme için daha tutarlı bir ölçekte veri elde edilmesini sağlar.

Son olarak kategorik değişkenlerin modellenebilir hale gelmesi için sayısal değerlere dönüştürme yapılmıştır. Bu kapsamında, katılımcıların duygusal durumunu ifade eden Happy_Sad_group değişkeninde "H" değeri 1, "S" değeri ise 0 olarak kodlanmıştır. Benzer şekilde, deneyde kullanılan ilaç türü gösteren Drug değişkeni de sayısal hale getirilmiş; "A" için 0, "S" için 1, "T" için ise 2 değeri atanmıştır. Bu dönüşüm, kategorik bilgilerin modeller tarafından doğrudan işlenebilmesini sağlamıştır.

Modelleme ve Reinforcement Learning Süreci

Bu çalışmada ilaç dozu kararlarının optimize edilmesi için offline reinforcement learning (offline RL) yaklaşımı kullanılmıştır. Ajan, gerçek bir ortamla etkileşime girmeden tamamen mevcut veri üzerinden öğrenmiş; dolayısıyla tüm öğrenme süreci “veriyle öğrenilmiş geri bildirim mekanizması” üzerinden yürütülmüştür. Ortam simülasyonu OpenAI Gym altyapısını takip eden özel bir ortam (MedicalEnv) ile sağlanmış, durum–eylem geçişleri bu yapı içinde modellenmiştir.

Öncelikle hastaların bir sonraki duruma geçişlerini temsil etmek için PyTorch ile bir dinamik model (DynamicsModel) tasarlanmıştır. Bu model Diff, mevcut doz, ilaç türü, Happy/Sad grup bilgisi ve ajan eylemini girdi olarak alıp bir sonraki adımda beklenen Diff değerini tahmin etmektedir. Böylece ajan, gerçek dünyanın yerine bu veriyle öğrenilmiş geri bildirim mekanizmasından aldığı tahminlerle hareket etmiş, tüm eğitim süreci offline RL çerçevesine uygun olarak yürütülmüştür.

Sistem içerisinde durum uzayı Diff, doz, ilaç türü ve grup bilgisi olmak üzere dört boyuttan oluşmaktadır; eylem uzayı ise üç seçenek içermektedir: dozu azalt (0), sabit tut (1) ve artır (2). Ortam her adımda dinamik model tarafından üretilen yeni Diff değeriyle güncellenmiştir. Ayrıca OpenAI Gym formatına uygun olarak her durum–eylem çifti için pozitif veya negatif ödül üreten bir ödül fonksiyonu tanımlanmış ve bu ödüller öğrenmenin temelini oluşturmuştur.

Ajanın politika fonksiyonu PyTorch ile kurulan bir PolicyNetwork tarafından modellenmiş ve eğitim süreci REINFORCE (Policy Gradient) algoritmasıyla gerçekleştirılmıştır. Ajan, her epizotta duruma bağlı olarak olasılıksal eylemler seçmiş, aldığı ödüller kaydedilmiş ve epizot sonunda politika ağı geri yayılmış ile güncellenmiştir. Eğitim sırasında amaç, uzun vadeli getiriyi (return) maksimize etmektir. Bu kapsamda epizotların toplam ödülleri takip edilmiş ve eğitim ilerledikçe politikanın istikrarı gözlemlenmiştir.

Eğitim tamamlandıktan sonra ajan test aşamasına alınmıştır. Test sürecinde politika deterministik şekilde işletilerek her durumda en yüksek olasılıklı eylem (argmax) seçilmiştir. Bu aşamada 100 epizot boyunca ajan performansı ölçülmüş, özellikle iki ana metrik raporlanmıştır: ortalama ödül ve denge süresi. Denge süresi, pozitif ödül alınan adımların sayısını ifade eder ve ajanın doğru doz kararlarını ne kadar tutarlı verdienenğini gösterir.

Ek olarak tek bir epizot için accuracy değeri hesaplanmış; pozitif ödül üreten her durum–eylem seçimi “doğru karar” olarak değerlendirilmiştir. Bu ölçüm, modelin anlık karar doğruluğunu ortaya koymuştur. Son olarak dinamik modelin başarısını değerlendirmek amacıyla tahmin edilen Diff değerlerinin hataları incelenmiş, ortalama tahmin hatası ve standart sapma raporlanmıştır.

Optimizasyon

Bu sürümde önceki modele kıyasla üç temel iyileştirme yapılmıştır. İlk olarak, öğrenme oranı 0.01'den 0.005'e düşürülerek daha yumuşak, kararlı ve gradyan patlamalarından uzak bir öğrenme süreci hedeflenmiştir.

İkinci olarak, REINFORCE fonksiyonuna $\text{decay} = 0.999$ şeklinde ek bir stabilizasyon katsayısı eklenmiş, böylece politika güncellemelerinin her epizotta aşırı dalgalanmasının önüne geçilerek uzun vadeli stabilitenin artırılması amaçlanmıştır.

Üçüncü olarak ise eğitim süresi 800 epizottan 1150 epizoda çıkarılarak politikanın daha olgun, daha az değişken ve daha güvenilir bir seviyeye ulaşması sağlanmıştır. Test aşaması ise aynı bırakılmış, çoklu test yine 100 epizot olarak korunarak iki sürümün performansının adil ve karşılaştırılabilir olması garanti edilmiştir.

Model Performans Analizi

Test aşamasında modelin hangi satırlarda doğru ya da hatalı karar verdiği analiz edebilmek için her episode boyunca ajan tarafından alınan ödüller satır ID'leriyle eşleştirilmiştir. Test süreci 100 episode olarak yürütülmüş ve her adımda ajan, eğitimde öğrenilmiş politika ağı (policy_net) tarafından üretilen olasılık dağılımından en yüksek olasılıklı eylemi seçmiştir. Ortamda adım atıldığında, o ana karşılık gelen veri satırının ID'si alınmış ve elde edilen ödüle göre ilgili listeye kaydedilmiştir: pozitif ödül alan satırlar rewarded_ids, sıfır veya negatif ödül alan satırlar ise punished_ids listesinde toplanmıştır.

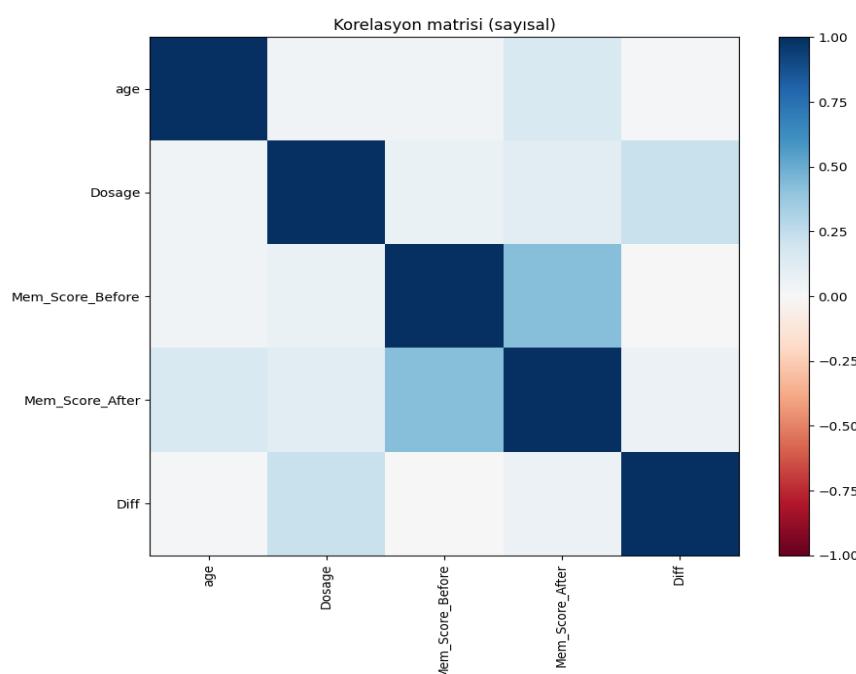
Çizgi grafik oluşturularak eğitim sürecindeki toplam ödül değişimleri görselleştirilmiştir. Eğitim çıktılarında yer alan epizot numaraları ile her epizot için hesaplanan total reward değerleri, düzenli ifadeler (regex) kullanılarak metinden ayrıstırılmış ve matplotlib kütüphanesi ile bir çizgi grafik halinde gösterilmiştir. Bu grafik, öğrenme süreci boyunca ödüllerde meydana gelen dalgalanmaları ortaya koymak modelin performansının epizotlar ilerledikçe nasıl değiştiğini göstermektedir.

DENEYSEL SONUÇLAR

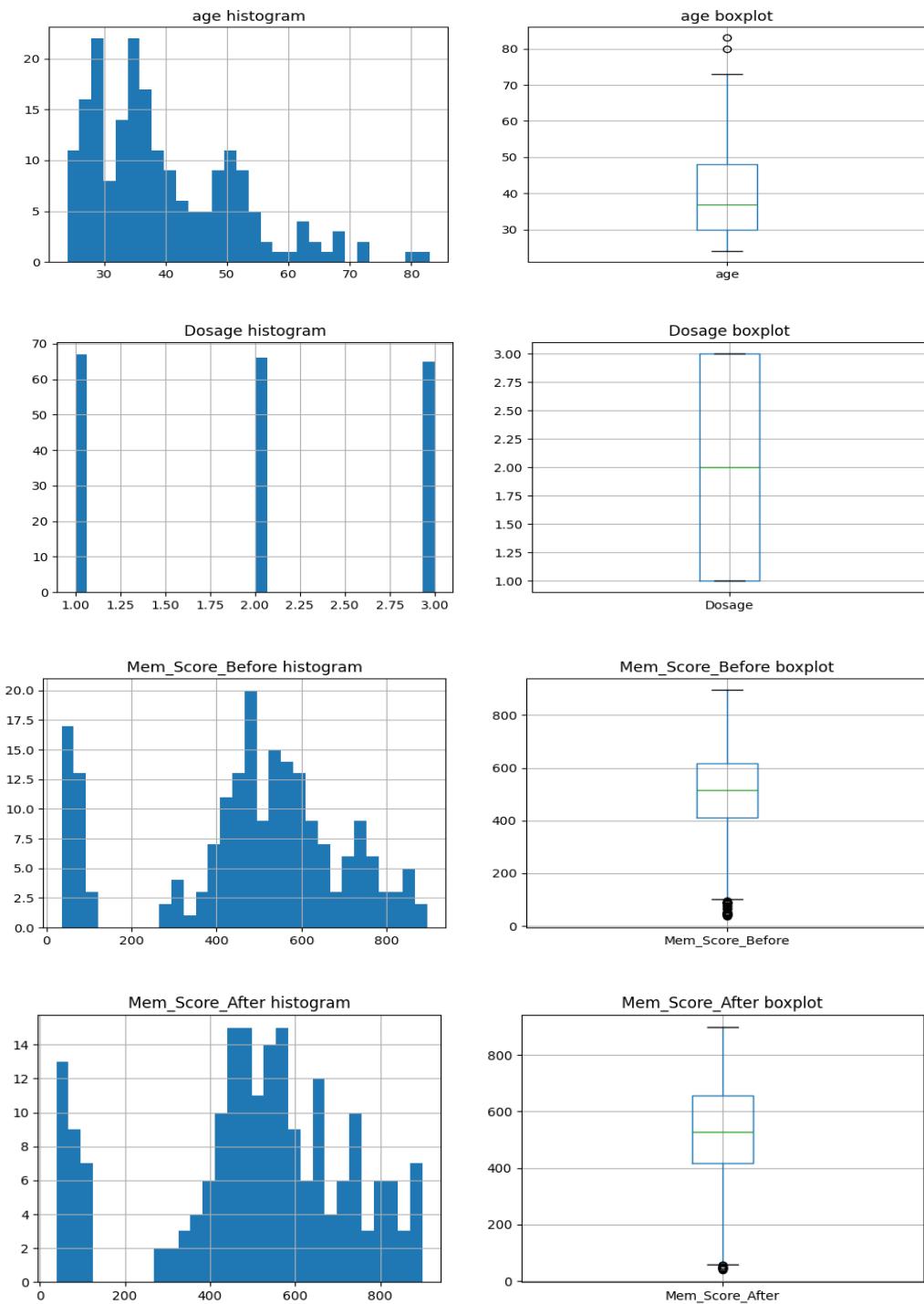
Genel Veri Seti Bulguları

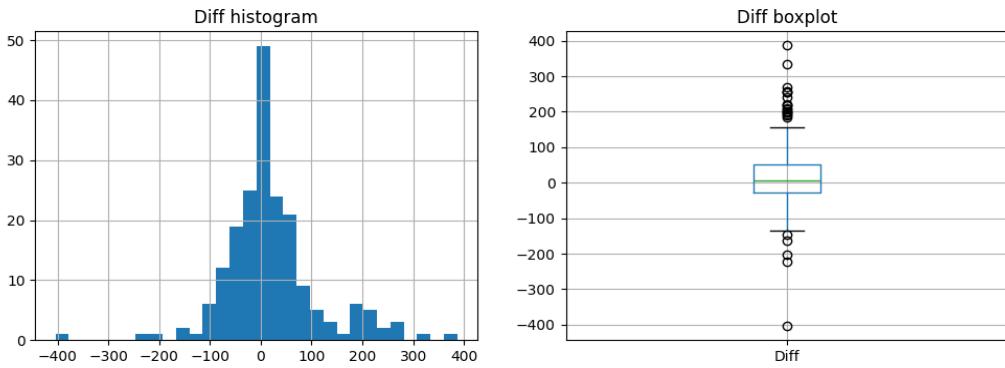
- Katılımcıların yaş ortalaması 39,5 yıl olup, yaş dağılımı 24–83 yıl arasında değişmektedir.
- Kullanılan ilaç türleri üç farklı kategoridedir ve dozlar 1–3 arasında değişmektedir; ortalama doz yaklaşık 1,99'dur.
- Hafıza skorları incelendiğinde, ortalama skor 481'den 507'ye yükselmiş ve Diff ortalaması +20 olarak hesaplanmıştır.
- Hafıza skorlarındaki değişimler geniş bir aralıkta görülmüş (-404 ile +387), yani bazı katılımcılarda ciddi artış veya düşüşler olmuştur.
- Veri setinde eksik değer bulunmamaktadır; tüm sütunlarda missing_count = 0 ve missing_pct = 0.0 olarak gözlenmiştir.
- last_name sütununda 18 farklı soyadı bulunmaktadır en sık rastlananlar: Durand (44), Carrasco (43), Lopez (23), Takahashi (20) ve Summers (12) şeklindedir.
- Happy_Sad_group iki kategoriye sahiptir ve her biri eşit sayıda gözlem içermektedir: H = 99, S = 99.
- Dosage üç farklı değere sahiptir ve dağılımı neredeyse eşittir: 1 = 67, 2 = 66, 3 = 65.
- Drug değişkeninde de üç farklı ilaç türü bulunmaktadır ve dağılımları şu şekildedir: A = 67, S = 66, T = 65.
- Mem_Score_Before sütununda 31 uc değer tespit edilmiştir, bu da başlangıç bellek skorlarında aşırı düşük veya yüksek değerler bulunduğunu göstermektedir.
- Diff sütununda 23 uc değer saptanmıştır, yani bellek skoru farklarında önemli aşırılıklar mevcuttur.
- Mem_Score_After sütununda 9 uc değer vardır, son bellek skorlarında daha az aşırı değer bulunmuştur.
- age sütununda 2 uc değer tespit edilmiştir, yaş dağılımında nadir istisnalar mevcuttur.
- Dosage sütununda uc değer bulunmamaktadır, dozaj verisi tutarlı ve dağılımı sınırlı aralıkta kalmıştır.

Veri Görselleri



- Mem_Score_Before ve Mem_Score_After: Yaklaşık 0.4 civarında pozitif korelasyon var; yani ilaçı almadan önce hafızası iyi olan kişiler, ilaçı aldıktan sonra da göreceli olarak iyi performans göstermektedir.
- Dosage ve Diff: Yaklaşık 0.2–0.25 civarında pozitif korelasyon var; doz arttıkça hafıza puanındaki değişim biraz yükseliyor, ilacın etkisi dozla artıyor olabilir.
- age ile Mem_Score_After veya Diff: Korelasyonlar oldukça düşük (0 civarı), yaş hafıza değişimi üzerinde anlamlı bir etkiye sahip görünmemektedir.
- Dosage ve diğer hafıza skorları: Hafif pozitif ilişki mevcut, ancak güçlü değil; doz hafıza puanındaki değişimini sınırlı şekilde etkileyebilir.





- Yaş dağılımı çoğunlukla 25–50 yaş arasında yoğunlaşmıştır ve birkaç yaşı katılımcı (70–80 yaşı) outlier olarak gözlemlenmiştir. Dağılım sağa çarpık olduğundan medyan, ortalamadan daha düşük olabilir; genel olarak katılımcılar genç ve orta yaşıdır.
- İlaç dozu üç seviyeli bir değişken olarak görülmektedir (1: düşük, 2: orta, 3: yüksek) ve her doz seviyesinde yaklaşık eşit sayıda katılımcı bulunmaktadır. Boxplot medyanı 2’yi göstermektedir, yani doz dağılımı dengelidir.
- Hafıza skorları ilaç almadan önce oldukça dağınık olup bazı katılımcılarda düşük performans gözlemlenmiştir. Boxplot ve histogram, alt ucta outlier’lar olduğunu göstermektedir. İlacı aldıktan sonra hafıza skorları genel olarak yükselmiş, medyan artmış ve düşük skorlar toparlanmıştır.
- Hafıza skorlarındaki değişim (Diff) pozitif yönde yoğunlaşmıştır. Coğu katılımcının hafıza puanı artmış, medyan artış pozitif ve yaygın olmakla birlikte bazı bireylerde oldukça yüksek iyileşmeler gözlenmiştir. Genel olarak ilaç uygulamasının hafıza skorlarını artırıcı etkisi olduğu söylenebilir.

Model Performansı ve Test Bulguları

Modelin eğitim ve test sonuçları, ajanın performansını açık şekilde göstermektedir. Toplam ödüller 16 ile 72 arasında değişmiş, 100 epizotluk test sürecinde ortalama ödül 67.96 ve ortalama denge süresi 132.98 olarak hesaplanmıştır. Tek-run sonuçlarında ise toplam ödül 68 olarak elde edilmiş ve accuracy %67.17 olarak ölçülmüştür. Dinamik modelin tahmin hatası incelendiğinde, ortalama hata -0.3578 ve standart sapma 1.0004 olarak bulunmuştur.

```

Episode 50/800, total reward: 20.00
Episode 100/800, total reward: 70.00
Episode 150/800, total reward: 54.00
Episode 200/800, total reward: 68.00
Episode 250/800, total reward: 68.00
Episode 300/800, total reward: 68.00
Episode 350/800, total reward: 68.00
Episode 400/800, total reward: 68.00
Episode 450/800, total reward: 68.00
Episode 500/800, total reward: 68.00
Episode 550/800, total reward: 68.00
Episode 600/800, total reward: 68.00
Episode 650/800, total reward: 16.00
Episode 700/800, total reward: 68.00
Episode 750/800, total reward: 72.00
Episode 800/800, total reward: 68.00
100 epizot için ortalama test ödül: 67.96
100 epizot için ortalama denge süresi: 132.98

--- Tek-run sonuçları ---
Toplam ödül: 68.00
Accuracy = 67.17%

--- Dynamics Model Tahmin Hatası ---
Ortalama hata: -0.3578
Standart sapma: 1.0004

```

Optimizasyon sonrası modelin performans sonuçları, eğitim süreci boyunca ve test aşamasında önemli bir gelişim gösterdiğini ortaya koymaktadır. Toplam ödüller 46 ile 140 arasında değişmiş, eğitim ilerledikçe özellikle 550. epizottan itibaren yüksek ödüller gözlemlenmiştir. 100 epizotluk test sürecinde ortalama ödül 134.00 ve ortalama denge süresi 166.00 olarak ölçülmüştür. Tek-run sonuçlarında toplam ödül 134.00 olarak elde edilmiş ve modelin doğruluk oranı %83.84 olarak kaydedilmiştir. Dinamik modelin tahmin hatası incelendiğinde, ortalama hata -0.1922 ve standart sapma 1.0011 olarak bulunmuştur.

```

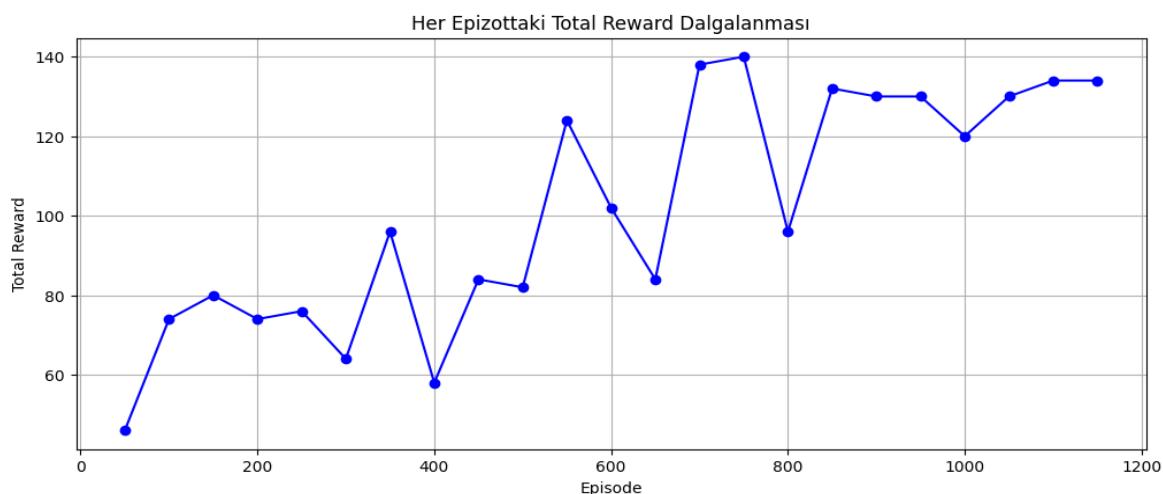
Episode 50/1150, total reward: 46.00
Episode 100/1150, total reward: 74.00
Episode 150/1150, total reward: 80.00
Episode 200/1150, total reward: 74.00
Episode 250/1150, total reward: 76.00
Episode 300/1150, total reward: 64.00
Episode 350/1150, total reward: 96.00
Episode 400/1150, total reward: 58.00
Episode 450/1150, total reward: 84.00
Episode 500/1150, total reward: 82.00
Episode 550/1150, total reward: 124.00
Episode 600/1150, total reward: 102.00
Episode 650/1150, total reward: 84.00
Episode 700/1150, total reward: 138.00
Episode 750/1150, total reward: 140.00
Episode 800/1150, total reward: 96.00
Episode 850/1150, total reward: 132.00
Episode 900/1150, total reward: 130.00
Episode 950/1150, total reward: 130.00
Episode 1000/1150, total reward: 120.00
Episode 1050/1150, total reward: 130.00
Episode 1100/1150, total reward: 134.00
Episode 1150/1150, total reward: 134.00
100 epizot için ortalama test ödül: 134.00
100 epizot için ortalama denge süresi: 166.00

--- Tek-run sonuçları ---
Toplam ödül: 134.00
Accuracy = 83.84%

--- Dynamics Model Tahmin Hatası ---
Ortalama hata: -0.1922
Standart sapma: 1.0011

```

Test aşamasında modelin doğru ve hatalı karar verdiği örnekler de raporlanmıştır. Pozitif ödül alan ilk 5 satır ID'si [2, 37, 158, 48, 9] olarak kaydedilmiştir; yani model bu satırlarda doğru dozaj kararları vermiştir. Öte yandan negatif veya sıfır ödül alan ilk 5 ID [111, 57, 56, 126, 189] olup, bu satırlarda modelin kararları beklenen sonucu üretmemiştir.



Başlangıçta (0–200. epizot) ödüller 45–80 arasında dalgalandırmaktır ve ajan çevreyi keşfetmektedir. Bu dönemde düşük ve değişken ödüller, keşif aşamasının tipik bir göstergesidir. Orta bölümde (200–600. epizot) dalgalandırmalar belirginleşmiş, bazı epizotlarda ödüller 60'a kadar düşerken bazı epizotlarda 120'ye çıkmıştır. Özellikle 500–600 arası, ödüllerde sert bir sıçrama ve kısa bir geri düşüş gözlemlenmiştir. 600 sonrası epizotlarda ise ödüller belirgin bir yükseliş trendi göstermiş ve daha istikrarlı bir hal almıştır. 700–1200 arası ödüller genellikle 110–140 bandında seyretmiş, 900 civarında hafif bir düşüş olsa da kısa sürmüştür. Son 5–6 epizotta ödüller 130 seviyesinin üzerinde sabitlenmiştir.

DEĞERLENDİRME

Veri Setinin Yorumu

Bu sonuçlar, çalışmada kullanılan ilaç dozlarının hafiza performansını genel olarak artırmada etkili olduğunu, ancak bireyler arasında farklılıklar gözlemediğini göstermektedir. Yaşın etkisinin sınırlı olması, ilaçın genç ve orta yaşı katılımcılarda benzer şekilde istediğini ortaya koymaktadır. Dengeli doz dağılımı, modelin her doz seviyesini yeterince öğrenmesini ve karşılaştırmayı mümkün kılar. Hafiza skorlarındaki yaygın pozitif değişim, tedavinin genel olarak faydalı olduğunu desteklerken, üç değerlerdeki yüksek iyileşmeler bireysel varyasyonların önemini vurgulamaktadır. Korelasyon analizleri, hafiza performansının büyük ölçüde bireylerin başlangıç seviyelerine bağlı olduğunu ve ilaçın etkisinin dozla sınırlı bir artış sağladığını göstermektedir. Eksik veri bulunmaması, analizlerin güvenilirliğini artırırken, üç değerler bireyler arasındaki farklılıkları yansıtmaktadır.

Model Performansının İncelenmesi

Başlangıçta model, sınırlı bir performans göstermiş ve doğruluk oranı orta seviyedeydi. Ödüller ve denge süresi, ajanın çevreyi öğrenme ve keşfetme sürecinde dalgalandırmalar yaşadığını gösteriyor. Optimizasyon ve eğitim sürecinin ilerlemesiyle modelin performansı belirgin şekilde artmıştır. Artan ödüller ve denge süresi, ajanın çevreye uyum sağlama ve etkili karar alma yeteneğinin geliştiğini gösterir. Doğruluk oranının yükselmesi, modelin daha güvenilir ve tutarlı bir politika oluşturduğunu doğrulamaktadır. Dinamik modelin tahmin hatasındaki azalma, çevresel davranışların daha doğru öngörildiğini ve modelin artık öğrenme sürecini daha etkin yönettiğini gösterir.

Optimizasyon sonrası modelin başlangıç döneminde ajanın ödüllerinin düşük ve değişken olması, keşif ağırlıklı bir öğrenme sürecinde olduğunu gösterir; ajan çevreyi anlamaya ve olası aksiyonların sonuçlarını keşfetmeye çalışmaktadır. Orta dönemde ödüllerdeki dalgalandırmalar, ajanın politikasını optimize ederken hâlâ keşif ve yararlanma arasında bir denge kurmaya çalıştığını ortaya koyar. Bu, RL süreçlerinde tipik bir “öğrenme kaosu” olarak yorumlanabilir ve modelin deneme-yanılma yoluyla öğrenme sürecinde olduğunu doğrular.

Son dönemde ödüllerdeki istikrarlı artış ve yüksek seviyelere ulaşılması, ajanın artık çevreyi iyi kavrayıp kararlı bir politika geliştirdiğini gösterir. Ödüllerin uzun süreli sabitlenmesi, modelin öğrenmenin doruk noktasına ulaştığını ve etkili bir stratejiyle istikrarlı performans sergilediğini işaret eder.

TARTIŞMA

Çalışmanın Çıkarımları

Hafıza skorlarındaki iyileşme ve doz ilişkisi, ilaçın genel olarak etkin olduğunu göstermektedir. Bu bulgu, yalnızca bireyler arasındaki performans farklılıklarını ortaya koymakla kalmayıp, aynı zamanda tedavi stratejilerinin optimize edilebileceğini de işaret etmektedir. Benzer şekilde, RL modelinin eğitim ve test sürecindeki performans artışı, algoritmanın çevreyi öğrenme ve etkili bir politika geliştirme yeteneğini doğrulamaktadır. Bu durum, yapay zeka tabanlı modellemelerin gelecekte karar destek sistemlerinde uygulanabilirliğini desteklemektedir.

Yaşın hafıza skorları üzerindeki minimal etkisi, ilaçın farklı yaş gruplarında benzer şekilde istediğini ve modelin doz-yanıt ilişkisini yaş faktöründen bağımsız olarak öğrenebileceğini göstermektedir. Ayrıca, dengeli doz dağılımı ve veri setindeki eksiksizlik, analizlerin güvenilirliğini artırmış ve gelecekte yapılacak çalışmalar için sağlam bir temel oluşturmuştur.

Bununla birlikte, çalışmanın bazı sınırlılıkları bulunmaktadır. Katılımcı sayısı ve yaş aralığı sınırlı olduğundan, elde edilen bulgular tüm popülasyona genellenemez. RL modelinin yalnızca belirli senaryolarda test edilmiş olması, modelin farklı durumlarda performansının sınırlı olabileceğini göstermektedir. Ayrıca, üç değerler ve bireysel varyasyonlar, bazı sonuçların tüm katılımcılar için geçerli olmayabileceğini ortaya koymaktadır.

Gelecekteki çalışmalar, daha geniş yaş ve demografik gruplarda gerçekleştirilebilir. Doz optimizasyonu, bireysel bazda daha ayrıntılı incelenerek kişiye özel tedavi stratejileri geliştirilebilir. RL modeli ise farklı çevreler ve senaryolar kullanılarak test edilerek, öğrenme sürecinin genellenebilirliği ve modelin güvenilirliği değerlendirilebilir.

Sonuç

Bu çalışma, farklı dozlarda uygulanan ilaçların hafıza performansını artırdığını ve bireyler arası farklılıkların önemli olduğunu göstermiştir. Yaşın etkisi minimal olup, doz-yanıt ilişkisi daha belirleyicidir. Hafıza skorlarındaki genel iyileşme, tedavinin etkinliğini ortaya koymaktadır. RL modeli, eğitim ve optimizasyon sürecinde performansını belirgin şekilde artırarak çevreyi öğrenme ve etkili politika geliştirme yeteneğini kanıtlamış; model, tutarlı ve güvenilir kararlar üretmektedir. Sonuç olarak, bu proje hem ilaç etkilerini hem de yapay zekâ tabanlı öğrenme süreçlerini başarılı şekilde değerlendirmiştir.