

**ERİTROSİT ENVANTER YÖNETİMİNDE DERİN PEKİŞTİRMELİ
ÖĞRENME**

İSTANBUL MEDİPOL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
SAĞLIK SİSTEMLERİ MÜHENDİSLİĞİ PROGRAMI
YÜKSEK LİSANS DERECEŚİNE İLİŐKİN BELİRLENEN ŐARTLARIN
YERİNE GETİRİLMESİ AMACIYLA SUNULAN
YÜKSEK LİSANS TEZİ

Ahmed Arif Őengil

Aralık, 2022

ERİTROSİT ENVANTER YÖNETİMİNDE DERİN PEKİŞTİRMELİ ÖĞRENME

Ahmed Arif Şengil

21 Aralık 2022

Jüri üyeleri olarak bu tezi okuduğumuzu ve ilgili programın yüksek lisans derecesi için yeterli kapsam ve kalitede olduğunu onaylıyoruz.

Prof. Dr. Hakan Tozan (Danışman)

Dr. Öğr. Üye. Mustafa Yağımlı

Dr. Öğr. Üye. Melis Almula Karadayı

Fen Bilimleri Enstitüsü tarafından onaylanmıştır:

Prof. Dr. Yasemin Yüksel Durmaz

Fen Bilimleri Enstitüsü Müdürü

İşbu belge ile bu tezdeki bütün bilgilerin akademik kurallara ve etik davranış ilkelerine uygun olarak toplanıp sunulduğunu beyan ederim. Bu kural ve ilkelerin gereği olarak, çalışmada bana ait olmayan tüm veri, malzeme ve sonuçları alıntılıdığımı ve kaynağını gösterdiğimi ayrıca beyan ederim.

İmza :

Adı Soyadı: AHMED ARİF ŞENGİL

TEŐEKKÜR

Bu alıőmanın baőından itibaren desteklerini eksik etmeyen Prof. Dr. Hakan Tozan, Dr. Öğr. Üye. Kevser Banu Köse, Dr. Öğr. Üye. Hülya Bilgen ve Dr. Öğr. Üye. Melis Almula Karadayı hocalarıma ve üniversitemizdeki diđer tüm saygıdeđer hocalarıma sonsuz desteklerinden dolayı teşekkür ederim.

Bu süreçte maddi manevi her açıdan yanımda olan amcama, aileme, tüm sevdiklerime ve üniversitedeki kardeşlerime de Őükranlarımı sunarım.

Ahmed Arif Őengil

Aralık, 2022



İÇİNDEKİLER

Sayfa

TEŞEKKÜR	iv
İÇİNDEKİLER	v
ŞEKİL LİSTESİ.....	vii
ÇİZELGE LİSTESİ.....	viii
SEMBOLLER	ix
KISALTMALAR	x
ÖZET	xi
ABSTRACT.....	xiii
1. GİRİŞ.....	1
1.1. Problem Tanımı	1
1.2. Vaka Tanımı	6
1.3. Araştırma Amacı	8
2. TEORİK KISIM.....	10
2.1. Pekiştirmeli Öğrenme	10
2.1.1. Pekiştirmeli öğrenmenin öğeleri	10
2.1.2. Markov karar süreci	11
2.1.3. Bellman denklemi	11
2.1.4. Yaklaşık dinamik programlama.....	12
2.1.5. Zamansal fark öğrenme	13
2.2. Derin Pekiştirmeli Öğrenme	17
2.2.1. Sinir ağları	17
2.2.2. Algoritmalar	19
2.3. Kan Ürünleri Tedarik Zinciri Yönetimi	22
2.3.1. Eritrosit envanter yönetimi	26
3. DENEYSEL KISIM	32
3.1. Vaka Tanımı	32
3.1.1. Ürünler.....	32
3.1.2. Talep	34
3.1.3. Amaç.....	37
3.1.4. Sistem	38
3.2. Durum Değişkeni	40
3.3. Eylem Değişkeni	40
3.4. Zaman Ufku	42
3.5. Ödül Fonksiyonu.....	43
3.6. Değer Fonksiyonu	43
3.7. Kıyaslama	44
3.8. Derin Pekiştirmeli Öğrenmenin Uygulanması.....	44
4. BULGULAR VE TARTIŞMA	46
4.1. Eğitim Eğrileri	46
4.2. Isı Haritaları	51
4.3. Performans	53
5. SONUÇ VE GELECEKTEKİ ÇALIŞMALAR	56
5.1. Sonuç	56
5.1.1. Akademik katkı	58
5.2. Öneriler	59

5.3. Gelecek Arařtırmalar	60
KAYNAKÇA	61
ÖZGEÇMİŐ	67



ŞEKİL LİSTESİ

Şekil 1.1: Kan tedarik zinciri [6].	3
Şekil 1.2: Pekiştirmeli öğrenme sisteminin öğeleri.	5
Şekil 2.1: Uçurum problemi Sutton ve Barto [8].	16
Şekil 2.2: Bir sinir ağının görselleştirilmesi.	18
Şekil 2.3: Sinir ağı içerisinde bir nöron.	19
Şekil 2.4: Q-tablosu ile Q-ağı arasındaki farkların görselleştirilmesi. Gemmink (2019)'dan [27] uyarlanmıştır	20
Şekil 2.5: Aktör-Kritik mimarisi Sutton ve Barto [8].	21
Şekil 4.1: A RH+ ve O RH+ kan grubu eğitim eğrileri.	48
Şekil 4.2: B RH+ kan grubu eğitim eğrisi.	48
Şekil 4.3: A RH-, AB RH+, O RH-, B RH- ve AB RH- kan grubu eğitim eğrileri.	50
Şekil 4.4: O RH+, B RH+ ve AB RH- kan grubu ısı haritaları.	53



ÇİZELGE LİSTESİ

Çizelge 2.1: Kan ürünlerinin özellikleri. American Red Cross (2017).....	24
Çizelge 2.2: Eritrosit envanter yönetimi.	27
Çizelge 3.1: Hastanenin Kızılay'dan sipariş ettiği ve teslim aldığı ünite sayıları.	35
Çizelge 3.2: Günlük Poisson ortalama talep yoğunlukları.	37
Çizelge 3.3: Farklı kan grupları için kullanılan eylem uzayları.	42
Çizelge 4.1: Farklı kan grupları için son 50 simülasyonda gerçekleşen ortalama israf ve kıtlık.....	54



SEMBOLLER

λ :Lamda



KISALTMALAR

MKS	: Markov Karar Süreci
DP	: Dinamik Programlama
YDP	: Yaklaşık Dinamik Programlama
PPO	: Proksimal Politika Optimizasyonu
LP	: Doğrusal Programlama
MILP	: Tamsayı Karışık Doğrusal Programlama
MINLP	: Tamsayı Karışık Doğrusal Olmayan Programlama
ILP	: Tamsayı Doğrusal Programlama
CMIP	: Kısıtlandırılmış Karışık Tamsayı Programlama
DES	: Kesikli Olay Simülasyonu
SDS	: Sistem Dinamikleri Simülasyonu
MCS	: Monte Carlo Simülasyonu
RS	: Yeniden Kullanılabilir Simülasyon
TDS	: İz-Sürümlü Simülasyon
PSO	: Parçacık Sürü Optimizasyonu
TA	: Eşik Kabulü
TS	: Tabu Araması

ERİTROSİT ENVANTER YÖNETİMİNDE DERİN PEKİŞTİRMELİ ÖĞRENME

ÖZET

Ahmed Arif Şengil

Sağlık Sistemleri Mühendisliği, Yüksek Lisans

Tez Danışmanı: Prof. Dr. Hakan Tozan

Eş-Danışman: Dr. Öğr. Üye. Kevser Banu Köse

Aralık, 2022

Dünya çapında en yaygın olarak uygulanan tedavi yöntemlerinden birisi kan transfüzyonudur. Sağlığın korunmasında ve hayatların kurtarılmasında kan ürünlerinin doğru bir şekilde temini çok önemli bir rol oynamaktadır. Her yıl, güvenli ve taze eritrosit erişimi sayesinde yüz binlerce hastanın hayatı kurtarılmaktadır. Dolayısıyla, kan kıtlığı sorununu çözebilmek ve kan talebinin ihtiyaçlarını karşılayabilmek için hastane düzeyinde de verimli bir kan ürünleri talep ve arz yönetimi sağlamak ve kan tedarik düzeyini yükseltmek ciddi bir ihtiyaçtır. Dünya çapında hastanelerin ortak hedefi, hizmet verilebilen hasta sayısını en üst düzeye çıkararak ve kan israfını azaltarak mevcut kan birimlerini daha iyi kullanmaktır. Kanı yönetmek zordur çünkü kan ürünleri bozulabilir ürünlerdir, arz stokastiktir ve talepler oldukça belirsiz bir şekilde gerçekleşmektedir. Ek olarak, eritrositler farklı gruplara ayrılmaktadırlar ve hasta uyumluluğu gerektirmektedirler. Pekıştirmeli öğrenme, dinamik ortamlarda sıralı karar verme problemlerini çözmek için geliştirilmiştir ve envanter yönetimi için ilgi çekici bir yöntemdir. Ölçeklendirilebilir, karmaşık problemlerde dahi optimal politikalara yakınsayabiliyor olmaları sebebiyle pekıştirmeli öğrenme metodları arasından yalnızca derin pekıştirmeli öğrenme yöntemleri dikkate alınmıştır. Son zamanlarda, derin pekıştirmeli öğrenme, birçok disiplinde sıralı karar verme problemleri için büyük potansiyel göstermiştir. Derin pekıştirmeli öğrenme, geleneksel yaklaşımlar kullanılarak elde edilmesi zor olan optimuma yakın politikalar geliştirmek için kullanılabilir. Ancak, Boute ve diğerlerinin (2021) kanıtladığı gibi, derin pekıştirmeli öğrenme algoritmaları, geleneksel yöntemlerle elde edilmesi zor, veya hatta bazı durumlarda imkansız olan optimuma yakın politikalar geliştirmek için kullanılabilir. Bu çalışma, hastane düzeyinde kan envanteri yönetimini iyileştirmek için pekıştirmeli öğrenme ve derin öğrenmedeki son gelişmelerden yararlanmaktadır. Kan bankası envanter yönetiminde derin pekıştirmeli öğrenme yöntemlerinden yararlanılıp yararlanılamayacağı, veya nasıl yararlanabileceği araştırılmıştır. Bu doğrultuda, İstanbul'da bulunan büyük ölçekli özel bir hastane örnek olarak ele alınarak eritrosit envanter yönetimi farklı kan grupları için derin pekıştirmeli öğrenme yöntemi ile modellenmiştir. Kullanılan Proksimal Politika Optimizasyonu [1] algoritmasının tüm kan gruplarında optimale yakın seviyelerde performans gösterebilecek sipariş politikaları

yakınsayabildiđi gözlemlenmiştir. Derin pekiştirmeli öğrenme yöntemi kullanılarak kan ürünleri envanter yönetiminden sorumlu uzmanlara sipariş kararları konusunda destek olabilecek modeller kurulabileceđi tespit edilmiştir.



Anahtar sözcükler: Envanter yönetimi, derin pekiştirmeli öğrenme, kan ürünleri.

**DEEP REINFORCEMENT LEARNING IN RED BLOOD CELL
INVENTORY MANAGEMENT
ABSTRACT**

Ahmed Arif Şengil

MSc in Healthcare Systems Engineering

Advisor: Prof. Dr. Hakan Tozan

Co-Advisor: Assist. Prof. Dr. Kevser Banu Köse

December, 2022

One of the most widely applied treatment methods worldwide is blood transfusion. Proper supply of blood products plays a very important role in maintaining health and saving lives. Lives of hundreds of thousands of patients' lives are saved each year via access to safe and fresh red blood cells. Therefore, it is a serious need to provide an efficient blood products demand and supply management at the hospital level and to increase the blood supply level in order to solve the problem of blood shortage and meet the needs of blood demand. The common goal of hospitals worldwide is to benefit from the acquired blood units as much as possible by maximizing the number of patients that can be served and minimizing the wastage of the units. It is hard to manage blood since blood products are perishable, the supply is stochastic, and demand occurs with high uncertainty. Additionally, red blood cells are divided into different groups and they require patient compatibility. Reinforcement learning has been developed to solve sequential decision making problems in dynamic environments and is an interesting method for inventory management. Among the reinforcement learning methods, only deep reinforcement learning methods have been taken into account, since they can converge to optimal policies even in scalable and complex problems. Deep reinforcement learning has shown great capability related to sequential decision-making problems lately, in many disciplines. Deep reinforcement learning is used to develop near-optimal policies which are hard to obtain using traditional methods. However, as Boute et. al. (2021) put forward recently, even though deep reinforcement learning methods can be used to develop near-optimal problems which are hard to obtain using traditional methods, even impossible time to time, they have not been applied at a satisfactory level in the inventory control field. This work makes use of the latest developments in the reinforcement learning field to be able to improve blood inventory management at a hospital level. Whether deep reinforcement learning can be benefited from in the field of blood bank inventory management, and if so, how would that be possible is the focus of the research. In this direction, red blood cell inventory management was modeled with a deep reinforcement learning method for different blood groups, using a private hospital in Istanbul as an example case. It has been observed that the Proximal Policy Optimization [1] algorithm used was able to converge to order policies that can perform at near-optimal levels for each blood type. It has been

confirmed that it is possible to develop models that can support the experts responsible from blood product inventory management with inventory decisions by using the deep reinforcement learning method.



Keywords: Inventory management, deep reinforcement learning, blood products.

BÖLÜM 1

1. GİRİŞ

Bu tez çalışması, sistem olarak büyük ölçekli bir özel hastanenin transfüzyon merkezi örnek alınarak ve bu merkezden temin edilen veriler kullanılarak hastane kan bankası eritrosit envanter yönetiminde derin pekiştirmeli öğrenme metodunun uygulanabilirliğini araştırmaktadır. Bölüm 1.1 problem tanımını içermektedir ve takip eden Bölüm 1.2’de çalışmada ele alınan vaka tanımlanmıştır. Son olarak, Bölüm 1.3’de çalışmanın amacı ve bu amacın yol açtığı araştırma soruları paylaşılmıştır.

1.1. Problem Tanımı

Kan; besinleri, oksijeni ve kimyasalları vücuttaki farklı hücre ve dokulara taşımak, enfeksiyonlarla mücadele etmek, vücut ısısını düzenlemek ve vücut asitliğini düzenlemek gibi bedende birçok gerekli işlevi yerine getirir [2]. Kan transfüzyonu, dünya çapında en önemli ve yaygın olarak uygulanan tedavi yöntemlerinden biridir. Kan talebini ve arzını yönetmenin daha doğru ve verimli yollarına duyulan ihtiyaç birçok ülkede artan bir endişe kaynağıdır. Kan sisteminin güvenliğini korurken, israfı ve kıtlığı azaltma hedeflerine ulaşabilecek, teknoloji tabanlı, sağlam bir kan ürünleri talep ve tedarik sistemi oluşturmak, modern sağlık sistemlerinde önemli bir ihtiyaçtır. Kan, sıkı bir kontrol mekanizması ile yönetilen ve raf ömrü olan bozulabilir bir üründür.

Kan bileşenlerinin bozulabilirliği, kan tedarik zincirlerinde talebin stokastik doğası gereği kullanılmayan fazla kan için israfa ve atıklara sebep olmaktadır. Bu, ek işlem ücretleri, envanter maliyetleri ve atık maliyetlerinin oluşmasına yol açmaktadır. Alternatif

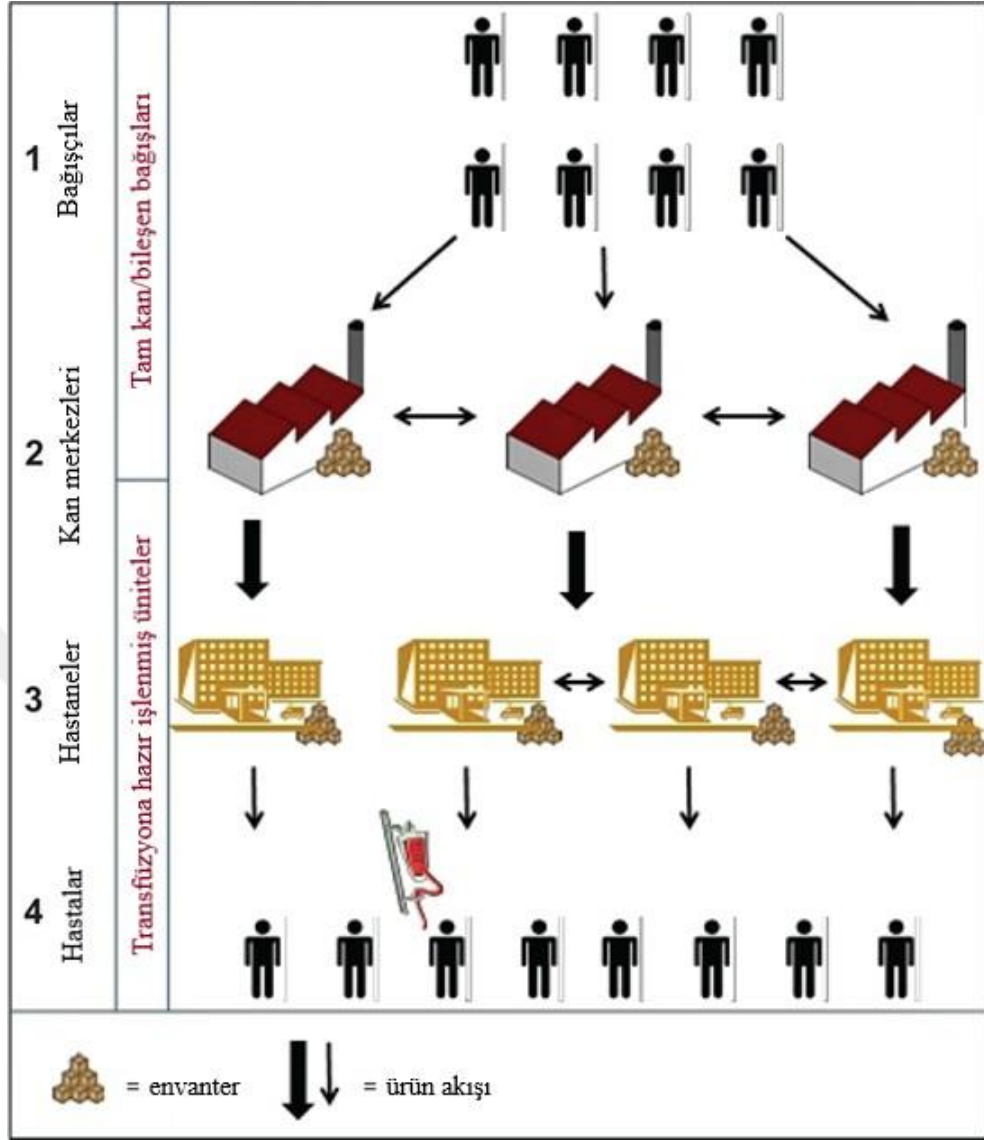
olarak, elde yetersiz miktarda kan bulunması, hastaların sađlıđı ve yařamları üzerinde somut olmayan sonuçlar dođuracaktır. Bu deđiř tokuř, özellikle arz ve talebin yuėsek belirsizliđi nedeniyle, karar vericilerin soruna net bir optimal ozuėm belirlemesini engellemektedir. Ařırı kan arzı veya kıtlıđının olumsuz etkilerini azaltmak iin, arz edilen kan miktarı, talebi mmkn olduđunca yakın dzeyde karřılayacak Őekilde seilmelidir. Kan rnnn mevcut olmamasının insan sađlıđı zerinde zararlı etkileri olabileceđi durumlarda kanın bozulabilirliđini hesaba katmak hayati nem tařımaktadır. Bu, aynı zamanda kanın raf mrnden daha uzun sre stokta kalmayarak eskimemesini ve atılmamasını sađlayacaktır.

Bozulabilir rnler iin iyi bir envanter ynetimi performansı, envanterde yeterli seviyelerde rn bulundurmayı garanti ederken aynı zamanda da raf mr dolan nite sayısını minimize etmeyi gerektirir [3]. Birok kuruluř iin bu dengeyi yanlıř ynetmenin sonuçları artan maliyetlere ve artan israfa yol aacaktır. Bununla birlikte, kan tedarik zinciri iinde iki hedefin karřılanamaması bir yandan mevcut olmayan grupta acil kan ihtiyaı olan bir hasta iin potansiyel olarak lmcl olabilirken, diđer yandan, fazla stok sađlık sistemi iin gereksiz maliyetlere yol aacaktır.

Trkiye’de her yıl 2,8 milyon niteden fazla kan bađıřlanmaktadır [4]. Bu leđe rađmen, kan bileřenleri kıt ve deđerli bir kaynak olmaya devam etmektedir. Trk Kızılay, Trkiye’deki tek kan rnleri tedarikisidir; bađıřıların temininden, bađıřların toplanmasından ve lke genelindeki hastanelere dađıtımından sorumludur [5]. Kan tedarik zincirinin basitleřtirilmiř bir versiyonu **Őekil 1.1**’de gsterilmektedir. Tedarik zincirinin drt ana unsuru vardır:

- 1 bađıřılar;
- 2 kan merkezleri;
- 3 hastaneler; ve
- 4 hastalar [6].

Kan; bađıř noktaları ađı, bađıř etkinlikleri ve mobil niteler aracılıđıyla bađıřılardan toplanır ve daha sonra kan merkezine nakledilir. Tedarik zincirinin bađıř tarafı ayrı bir birimdir ve kan merkezleri, “itme” ve talep odaklı tedarik zinciri arasında ayrıřtırma noktaları



Şekil 1.1: Kan tedarik zinciri [6].

olarak hareket eder [7]. Kan, genellikle tam kan üniteleri olarak toplanır ve daha sonra kan merkezlerinde işlenir. Bağışlanan kan bulaşıcı etkenler için test edilir ve kan grubu belirlenir. Tam kan daha sonra ana bileşenlerine ayrılır: eritrosit (kırmızı kan hücreleri), trombositler ve plazma. Tam kan ve bileşenlerinin farklı raf ömürleri, kullanımları vardır ve farklı saklama yöntemleri gerektirir. Bu çalışmanın analizinin yalnızca eritrosit bileşenine odaklanmıştır.

Gerekli işlemlerden sonra bileşenler, hastanelere gönderilmek üzere kan merkezlerinde depolanır. Her hastanedeki transfüzyon merkezi sorumlusu, daha sonra ilgili kan merkezinden teslim edilecek kan bileşenlerinin siparişinden sorumludur. Transfüzyon merkezi sorumlusu, hastanedeki müşteriler (tıbbi ve cerrahi departmanlar) ile üst

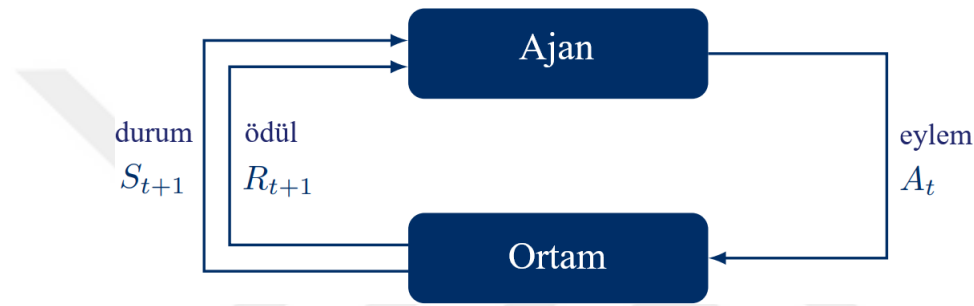
kademedeki tedarik zinciri arasındaki bağlantıdır ve hastanenin kan bileşenleri tedarikini yönetir. Hastanede kan ünitelerini boşa harcamanın önemli bedelleri vardır ve bu ünitelerin transfüzyon için kullanılmış olmak yerine israf edildiklerinden dolayı envanterdeki yerlerinin doldurulması gerektiğinden dolayı, tedarik zincirinin üst kademeleri üzerinde de zincirleme bir etkiye sahiptir. Tedarik zinciri gönüllü bağışlara dayanmaktadır ve bu nedenle hastanelerde boşa harcanan her ünite kan aynı zamanda bağışçıların kan vermek için harcadıkları zamanın ve emeğin de kaybıdır. Bu nedenle hastane talebinin kapsamı ve israfı, tedarik zincirinde yukarı kademelerdeki bağışçı sayısını etkiler.

Hastanelerdeki kan stokları, atanmış ve atanmamış olmak üzere iki çeşitten oluşmaktadır. Atanmış envanter, bireysel hastalar için ayrılan ünitelerden oluşur ve atanmamış envanter herhangi bir hasta tarafından kullanılabilir. Hastalara atanmadan önce, kan üniteleri, transfüzyon reaksiyonlarını önlemek ve uyumluluğu sağlamak için hastadan alınan kan numunesiyle çapraz eşleştirilir (cross-matching) [8]. Çapraz eşleştirme yapıldıktan sonra, atanan birimler, her bir hasta için 24 saat ile 72 saat arasında transfüzyona hazır olarak bekletilir. Bu işlem, ünitelere ihtiyaç duyulmazsa ve üniteler daha sonra genel atanmamış envantere iade edilirse değerli raf ömrü kaybına neden olur.

Kan envanter yönetimi teorisi, daha geniş bir alan olan bozulabilir envanter teorisinin bir alt kümesidir. Ancak, hem atanmış hem de atanmamış envanterin varlığının da sebep olduğu ek karmaşıklık nedeniyle, bozulabilir envanter teorisindeki çoğu genel modelin uygulanması zordur veya imkansızdır. Özetle, kan envanterinde bugüne kadar yapılan araştırmalara, matematiksel modeller geliştiren ve bunları çeşitli politikalar türetmek için kullanan yöneylem araştırması uzmanları hakimdir [9]. Bir yandan da literatür, karmaşık envanter modellerinin uygulanmasının düşük israf seviyelerine ve iyi envanter performansına yol açtığı izlenimini yaratmaktadır [10]. Ancak eleştirmenler, bu modellerin çok karmaşık olduğunu ve yakınsamaların ve sezgisel yöntemlerin bozulabilir envanter sistemlerinde iyi performans için daha verimli bir yol olduğunu belirtmektedir.

Pekiştirmeli öğrenme, bir ödülü en üst düzeye çıkarmak için ne yapılacağını öğrenmekle ilgilidir [11]. Bir pekiştirmeli öğrenme modeli **Şekil 1.2**'deki gibi görselleştirilebilir. Bu modelde bir ajan, bir eylem aracılığıyla ortam ile etkileşime girer. Bu

işlem yapıldığında ortamın mevcut durumu güncellenir ve bir ödül verilir. Pekiştirmeli öğrenme, keşif (exploration) ve istifade (exploitation) arasında bir denge bulmaya odaklanır [12]. Keşif sırasında, pekiştirmeli öğrenme ajanı, ödülünü keşfetmek için rastgele eylemler seçer, istifade ise sahip olunan mevcut bilgiye dayalı yapılan eylemlerdir. Pekiştirmeli öğrenmenin bir diğer önemli yönü, gecikmeli ödüllerle başa çıkabilme yeteneğidir; bu, sistemin tanım gereği şu anda en yüksek ödüle sahip eyleme gitmeyeceği, ancak genel olarak toplamda en yüksek ödülü elde etmeye çalışacağı anlamına gelmektedir. Bu özellikler, pekiştirmeli öğrenmeyi belirsizlikler altında karar vermek için iyi bir yöntem haline getirmektedir.



Şekil 1.2: Pekiştirmeli öğrenme sisteminin öğeleri.

Pekiştirmeli öğrenmenin yaygın olarak bilinen bir uygulaması oyun alanındadır. Örneğin, Pac-Man oyununda pekiştirmeli öğrenmeyi kullansaydık, ajanımız Pac-Man'ın kendisi olurdu ve ortam da labirent olurdu. Ajanın alabileceği eylem belirli bir yönde hareket ediyor. Pac-Man bir yöne doğru hareket ettiğinde bu eyleme bağlı olarak ortam güncellenir. Daha sonra Pac-Man'e bir ödül verilir; bu, Pac-Man bir nokta yediğinde kazanılan olumlu bir ödül veya Pac-Man bir hayalet tarafından yenildiğinde karşılaşılan olumsuz bir ödül olabilir. Ancak, Pac-Man önce yanıp sönen büyük noktayı ve ardından hayaleti yediğinde, mümkün olan en büyük ödülü alır. Bu bir gecikmeli ödül örneğidir. Pekiştirmeli öğrenme sistemi oyunu oynayarak öğrenecektir. Keşif yoluyla, belirli durumlarda belirli eylemlere verilen ödülleri öğrenir. Pekiştirmeli öğrenme sistemi oyunu yeterince uzun süre oynadığında, Pac-Man oyununu oynamada üstün olan bir pekiştirmeli öğrenme sistemi elde edilmektedir.

Pekiştirmeli öğrenme, dinamik ortamlarda sıralı karar verme problemlerini çözmek için geliştirilmiştir [11]. Sıralı karar vermede, genel ödülü en üst düzeye çıkarmak için dinamik bir ortamla etkileşim halinde bir dizi karar verilmelidir [13]. Bu, pekiştirmeli

öğrenmeyi envanter yönetimi için ilgi çekici bir yöntem haline getirmektedir. Dinamik ortamları idare edebilmekle, sabit bir sipariş politikası yerine sistemin mevcut durumuna bağlı değişken bir envanter politikası oluşturmak mümkün olabilir. Sıralı karar verme ve gecikmeli ödülleri de envanter yönetimi için önemlidir. Envanter yönetiminde sürekli eylem alınması gerekmektedir, ancak bu kararların sonuçları her zaman doğrudan görünür değildir. Örneğin, bir şirket stokta hala ürünleri olduğu için belirli bir anda yenilememeyi seçtiğinde, doğrudan bir ceza uygulanmaz. Daha sonraki bir aşamada, ürünler stokta kalmadığında müşterilere hizmet verilemez. Bu durumda potansiyel satışlar kaybedilmiş olur ve şirketin, karını maksimize etmek için daha erken envanter yenilemeyi tercih etmiş olması gerektiği ortaya çıkar. Pekiştirmeli öğrenme, her durumdaki her karara belirli bir ödül atayabilir ve bu nedenle, envanterin ne zaman yenilenmesi gerektiği konusunda sağlam önerilerde bulunmak için umut verici bir yöntemdir. Bunun yanında pekiştirmeli öğrenme, daha fazla stokastiklik ve daha fazla değişkeni hesaba katma gibi daha karmaşık durumların çözülmesine yardımcı olabilmektedir.

Matematiksel modellerin karmaşıklığı ve hesaplama zamanının bir sonucu olarak, çok-ürünlü modeller şu anda literatürde çok az temsil edilmektedir [14]. Pekiştirmeli öğrenmenin kullanılmasıyla, modeli daha karmaşık durumlar için kullanmak daha kolay hale gelebilmektedir. Ayrıca, pekiştirmeli öğrenme talebin stokastikliğini hesaba katabilmektedir, oysa şu anda sadece birkaç model bunu dikkate almaktadır, çünkü talep stokastikliği ile başa çıkılması gerçekten zordur [14]. Pekiştirmeli öğrenmenin bir başka umut verici yönü, kapsamlı alan bilgisine veya kısıtlayıcı varsayımlara ihtiyaç duymadan çeşitli sorunları çözenin bir yolunu sunabilmesidir [15]. Bu nedenle, farklı durumlara uyum sağlamak için daha az çaba gerektiren genel bir yöntem olarak çalışabilir ve böylelikle gerçek dünya durumlarını yansıtmak daha kolay hale gelecektir.

1.2. Vaka Tanımı

Bu çalışmada İstanbul ilinde yer alan büyük ölçekli bir özel hastane ele alınmıştır. Hastanenin kan ürünleri yönetiminden tamamen hastanenin transfüzyon merkezi sorumludur ve bu birim yalnızca bu hastanenin kan ürünleri yönetimi ile ilgilenmektedir. Her gün, transfüzyon merkezi sorumlusu tarafından belirlenmiş olan rakamlar kullanılarak Kızılay'dan ihtiyaç duyulan kan ürünleri sipariş edilir. Hekimler, taleplerini genellikle

1 gün önceden transfüzyon merkezine bildirmektedirler. Envanter durumu her gün sorumlu hekim tarafından kontrol edilmektedir ve sonrasında; tarihi talep verileri, yakın tarihli talep verileri, hastanenin farklı departmanlarında bulunan hasta dağılımı, mevcut envanter gibi ilgili durumlar göz önünde bulundurularak sipariş miktarları belirlenmektedir. Merkez sorumlusu aynı zamanda gün, hafta, ay ve mevsimden yola çıkarak trend ve sezonsallık durumlarını da göz önünde bulundurarak karar vermektedir. Transfüzyon merkezi 24 saat operasyon halindedir.

Bozulabilir ürünler söz konusu olduğundan dolayı ünitelerin hastalara atanması işleminde önce raf ömrü dolmaya daha yakın olan ürünler tercih edilmektedir. Her ne kadar bazı tıbbi uygulamalarda kanın tazeliğine önem verilse de, genel olarak raf ömrünü doldurmuş olan tüm üniteler aynı şekilde kullanılabilir. Dolayısıyla, hekim taleplerinin karşılanması esnasında raf ömrü dolmaya yakın ünitelere öncelik verilerek israf edilen ünite sayısı minimize edilmektedir. Merkez, bu “ilk giren ilk çıkar” (FIFO) yöntemini katı bir şekilde takip etmemektedir ancak envanterdeki ünitelerin yaşı sistemden görüntülenebildiğinden ve düzenli olarak takip edildiğinden dolayı raf ömrü dolmaya yaklaşan ürünlere dikkat edilmektedir ve bu ürünler ilk fırsatta kullanıma sunulmaktadır. Transfüzyon merkezi yoğun tüketim seviyelerine sahip olduğundan ve envanter seviyeleri yakından takip edildiğinden dolayı eritrositlerde israf (wastage) ve kıtlık (shortage) ihmal edilebilecek seviyelerdedir. Tüm eritrosit siparişleri karşılanabilmektedir ve tüketilen binlerce ünitenin yanında yılda yalnızca 1-2 ünite israf gerçekleşebilmektedir. Kızılay’dan gelen ünitelerin yaşlarında rassallık söz konusudur. Teslim alınan eritrositlerin örneğin 5 günlük raf ömrü de kalmış olabilir, 30 günlük raf ömrü de kalmış olabilir. Eritrositlerin teslim alınırkenki ortalama yaşı 15’dir.

Merkez, Kızılay’a 2 farklı şekilde sipariş oluşturabilmektedir: standart sipariş ve acil sipariş. Standart siparişlerde de ürünler Kızılay tarafından hızlı ve sorunsuz bir şekilde teslim edilmektedir ancak kritik durumlarda acil siparişler tercih edilmektedir. Ayrıca, acil durumlarda da kullanılabilmesi adına envanterde her zaman bir miktar genel verici kan grubu (0 RH-) eritrosit bulundurulmaktadır. Kan toplama yetkisi yalnızca kendisinde bulunan Kızılay, envanterinde yeterli ürün bulunmadığı bazı durumlarda hastanelere geçici olarak kan toplama yetkisi verebilmektedir. Böyle durumlarda ele alınan hastaneye geçici olarak, Kızılay tarafından karşılanamayan ünite adedince kan toplama yetkisi verilebilmektedir.

Hastane, kullanılan ünitelerin ücretini Sosyal Güvenlik Kurumu'ndan alıp Kızılay'a ödemektedir. Yalnızca ünitelerin israf edilmesi durumunda ünitenin bedeli hastaneye yansıtılmaktadır. Tedarik edilen ünitelere ne olduğu (kullanıldı, israf edildi vb.) gün sonunda Kızılay ile paylaşılmaktadır.

1.3. Araştırma Amacı

Bu çalışma, hastane kan bankası eritrosit envanter yönetiminde derin pekiştirmeli öğrenme yönteminin uygulanabilirliğinin araştırılması üzerine odaklanmaktadır.

Kan temini, sağlığın korunmasında ve hayatların kurtarılmasında çok önemli bir rol oynamaktadır. Bu nedenle, kan kıtlığı sorununu çözmek ve kan talebinin ihtiyaçlarını karşılamak için kan tedarik düzeyini yükseltmek ve hastane düzeyinde sağlam bir kan talebi ve arzı yönetimi sağlamak büyük öneme sahiptir. Hastane kan bankaları düzeyinde, mevcut envanter yönetimi uygulamaları büyük ölçüde insan kararlarına dayanmaktadır. Kan ürünleri sipariş kararları insan deneyimine göre verilmektedir. Hastane talepleri, bireysel hasta teşhislerine dayalı olarak doktorlar tarafından yapılmaktadır ve aynı anda birden fazla ünite ve belirli doz gereksinimleri olan üniteler gibi farklı gereksinimlerle gerçekleşmektedir. Bu sınırlamalar, envanter yönetiminin karmaşıklığını artırmaktadır. Sonuç olarak, hastane kan bankaları, fazla envanter tutarak talepteki değişkenlikle başa çıkma eğilimi göstermektedirler. Gerekenden fazla envanter tutmak, hastane kan bankalarının kıtlık riskini azaltmasına yardımcı olurken, kan depolama maliyetlerini ve israf riskini artırmaktadır.

Ayrıca, süresi dolmak üzere olan birimlerin yeniden taşınması ekstra maliyetlere yol açmaktadır. Bununla birlikte, Heddle [16], eritrosit depolama süresinin, hasta sonuçlarını etkileyebilecek fonksiyonel bütünlüğü ve kalite standartlarını etkileyebileceğini göstermiştir. Ek olarak, hastane kan bankalarında büyük miktarda envanter bulunduğu Kızılay'ın gerçek talebi görmesi engellenmektedir ve Kızılay'ın talep değişkenliğine uyum sağlama yeteneği kısıtlanmaktadır. Sonuç olarak, hem hastane kan bankalarının, hem de Kızılay'ın, talep ve tedarik zincirinde verimli bir şekilde hareket etmeleri zorlaşmaktadır.

Kanın özgün fizyolojik özellikleri, envanter yönetimine büyük zorluklar getirmektedir. Kanın fizyolojik özellikleri takip eden şekildedir: 1- Kan ürünlerinin kıtlığı. Kanın tek

tedarik kaynağı insan vücududur, ki bu kolay bir tedarik kanalı değildir. 2- Kan ürünleri bozulabilir ürünlerdir. Kan ürünlerinin sabit bir raf ömrü vardır. Örneğin, farklı işleme yöntemlerine bağlı olarak kırmızı kan hücrelerinin raf ömrü genellikle 21 gün, 35 gün veya 42 gün; trombositlerin raf ömrü ise 5-7 gündür. 3- Kan ürünlerinin yeri doldurulamaz. Kan, başka malzemelerle değiştirilemez [17].

Pekiştirmeli öğrenme metodu envanter problemlerinin çözümünde başarılı bir yöntemdir ve özellikle de derin pekiştirmeli öğrenme yöntemlerinin daha karmaşık envanter problemlerinin çözümünde kullanılabildiği, bu alanda gelecek vaat ettiği literatürde belirtilmiştir [15]. Kan ürünleri envanter yönetiminin sahip olduğu yüksek karmaşıklık düzeyi göz önünde bulundurulduğunda derin pekiştirmeli öğrenme metodunun bu alanda da ümit verici bir teknik olduğu tahmin edilebilmektedir fakat literatürde böyle bir uygulama bulunamamıştır.

Bu araştırmanın amacı, ele alınan hastanede faaliyet göstermekte olan sistem örnek alınarak, hastane kan bankası envanter yönetiminde transfüzyon merkezi sorumlularına farklı envanter durumları için sipariş miktarı tavsiyelerinde bulunabilecek bir derin pekiştirmeli öğrenme yöntemi geliştirmektir. Envanter yönetiminin optimize edilip edilemeyeceğini öğrenmek için bir derin pekiştirmeli öğrenme yöntemi kurulmalıdır. Bu, aşağıdaki ana araştırma sorusuna yol açmaktadır:

Hastane kan bankası eritrosit envanter yönetiminde derin pekiştirmeli öğrenme yöntemi kullanılabilir mi?

BÖLÜM 2

2. TEORİK KISIM

Bu bölümde literatür taramasının bulguları tartışılmıştır. Bölüm 2.1 pekiştirmeli öğrenme ve öğelerinin açıklamasını, Bölüm 2.2 ise derin öğrenme ile pekiştirmeli öğrenmenin birleşimini kapsamaktadır. Bölüm 2.3’de önce kan ürünleri tedarik zinciri tartışılmıştır. Sonrasında, eritrosit envanter yönetimine giriş yapılmış ve bu alanda yapılmış olan çalışmalar özetlenmiştir. Son olarak, literatüre yapılan katkı belirtilmiştir.

2.1. Pekiştirmeli Öğrenme

Bölüm 1.1’de pekiştirmeli öğrenme kısaca tanıtılmıştı. Bu bölümde de pekiştirmeli öğrenmenin yapısı daha ayrıntılı olarak ele alınmıştır. Sonrasında ise yaklaşık dinamik programlamanın pekiştirmeli öğrenme ile nasıl ilişkili olduğu tartışılmış ve ardından da, en sık kullanılan iki algoritmayla birlikte, yaygın olarak kullanılan bir yöntem olan zamansal fark öğrenmesi tartışılmıştır.

2.1.1. Pekiştirmeli öğrenmenin öğeleri

Pekiştirmeli öğrenme, bir ödülü en üst düzeye çıkarmak için ne yapılacağını öğrenmekle ilgilidir. Pek çok makine öğrenmesi biçiminden farklı olarak, pekiştirmeli öğrenmede, öğreniciye her durumda hangi eylemi gerçekleştirmesi gerektiği açıkça belirtilmez [11]. Her zaman adımında, ajan mevcut durum s_t ’yi gözlemler ve bir eylem a_t seçer. Ajan eylemi deneme yanılma yoluyla (keşif) ve ortam hakkında halihazırda sahip olduğu bilgiye dayanarak (istifade) seçer [18]. Eylemi gerçekleştirdikten sonra ajana r_{t+1} ödülü verilir ve ajan yeni s_{t+1} durumunda gelişmiş olur. Ajan bu bilgileri kullanarak ortam hakkındaki bilgilerini günceller ve bir sonraki eylemi seçer. Ödülü en üst düzeye çıkarabilmek için keşif

ve istifade arasındaki denge çok önemlidir. Kazanılacak olan ödül değerlerini bulabilmek için ajan sürekli yeni durum-eylem çiftlerini keşfetmek zorundadır. Eksiksiz keşif tüm eylemlerin tüm durumlarda alınmış olmasını sağlar. İdeal olarak, keşif aşamasında öğrenilmiş olan değerlerle, her durumda en iyi eylemin hangisi olduğunu açıkça gösterecek olan bir eğilim bulunabilir. Her durumu, alınması gereken ilgili optimal eyleme eşleyen vektör öğrenilen optimal politikayı temsil eder [18]. Çoğu metod tam keşifle başlar ve keşif şansını yavaş yavaş azaltır, dolayısıyla istifade şansını artırır.

2.1.2. Markov karar süreci

Ajanın girdiği her durum, önceki durumun ve seçilen eylemin doğrudan bir sonucudur. Her adım ve bunların atılma sırası, mevcut durum hakkında bilgi içerir ve ajanın bir sonraki adımda hangi eylemi seçmesi gerektiğini etkilemektedir. Ancak, büyük problemlerde tüm bilgileri depolamak ve işlemek imkansız hale gelmektedir. Bunun üstesinden gelebilmek için tüm durumların Markov durumları olduğu varsayılır. Bu, herhangi bir durumun yalnızca kendisinden bir önceki duruma ve seçilen eyleme bağlı olduğu anlamına gelmektedir. Bir pekiştirmeli öğrenme sistemi bu Markov özelliğini sağladığında, buna Markov Karar Süreci (MKS) denir [11]. Hem tüm olası durumlar kümesinin hem de her durumdaki tüm olası eylemler kümesinin sonlu olduğunu varsaydığımızda, sonlu bir MKS'ye sahip oluruz. Sonlu MKS pekiştirmeli öğrenmede sık sık kullanılmaktadır çünkü bu şekilde tek-adımlı dinamikler sağlanmış olmaktadır. Sonlu bir MKS'nin olasılık tanımı şöyledir:

$$P(s', r | s, a) = P\{S_t = s', R_t = r | S_{t-1} = s, A_{t-1} = a\} \quad (2.1)$$

2.1.3. Bellman denklemi

Pekiştirmeli öğrenmeyi kullanırken her durum için en iyi eylemi seçmek önemlidir. Bellman denklemi hangi eylemin en iyi eylem olduğunu öğrenmek için kullanılmaktadır. Bellman denklemi, başlangıç noktanızın değerinin, o durumda olmaktan beklediğiniz ödül ile birlikte girecek olduğunuz durumun değerinin toplamı olduğu gerçeğine dayanan genel bir formüldür. Deterministik bir ortamın Bellman denklemi aşağıdaki gibi gösterilebilir [19]:

$$V(s) = \max_a (R(s, a) + \gamma V(s')) \quad (2.2)$$

Bu denklemde, $R(s, a)$ ifadesi s durumunda alınacak belirli bir eylem a 'nın ödül değerini temsil eder. $\gamma V(s')$ ise bir sonraki durumunun indirgenmiş değerini gösterir. Bellman denklemi ile bir durumda alınabilecek tüm olası eylemler dikkate alınır ve maksimum değere sahip olan eylem seçilir. Bellman denklemi pekiştirmeli öğrenmenin temel ilkelerinden biridir ve pekiştirmeli öğrenme metodlarında kullanılmaktadır, çoğu zaman uyarlanmış versiyonlar halinde.

2.1.4. Yaklaşık dinamik programlama

Basit MKS'ler olası tüm politikalar değerlendirilerek çözülebilmektedir. Tabii ki bu yaklaşım daha büyük ve daha karmaşık MKS'lerde pratik olarak hesaplanamaz hale gelecektir. MKS'leri çözenin bir başka, daha akıllıca yolu ise Dinamik Programlamadır (DP). Dinamik Programlama karmaşık MKS'leri daha küçük alt problemlere bölerek çözer [20]. MKS için optimal politika, MKS'nin tüm alt problemlerine optimal çözümü sağlayan politikadır [19]. DP, ardışık durumların değerlerinin tahminlerine dayalı olarak durumların değerlerinin tahminlerini günceller veya tahminleri geçmiş tahminlere dayalı olarak günceller. Dinamik Programlama ortamın mükemmel bir modelini gerektirir ve bu ortam bir MKS'ye eşdeğer olmalıdır. Bu nedenle ve yüksek hesaplama maliyetleri sebebiyle DP algoritmaları pekiştirmeli öğrenmede pek tercih edilmemektedir. Yine de, pekiştirmeli öğrenmenin teorik çerçevesini temsil ederler çünkü her iki yöntem de aynı amaca ulaşmaya çalışmaktadır. Pekiştirmeli öğrenmenin avantajı bunu daha düşük hesaplama maliyetleriyle ve mükemmel bir model varsayımı gerekmeksizin yapabilmesidir [11].

Yaklaşık Dinamik Programlama (YDP), MKS modeli temelli bir modelleme çerçevesidir. YDP, aynı pekiştirmeli öğrenme gibi, bir MKS'nin optimal çözümüne yaklaşmak için kullanılan geniş yöntem yelpazesi için kullanılan genel bir terimdir. Genellikle optimizasyonu simülasyonla birleştirir. Yaklaşık Dinamik Programlama ve pekiştirmeli öğrenme, sıralı karar-verme problemlerini çözmek için kullanılan birbiriyle yakından ilişkili iki paradigma olduğundan, bunlar bozulabilir envanter yönetiminde pekiştirmeli öğrenme bölümünde tartışılmıştır.

2.1.5. Zamansal fark öğrenme

Markov Karar Süreci ile, geçiş olasılıkları da dahil olmak üzere ortamın eksiksiz bir modeline sahibiz. Ancak, pekiştirmeli öğrenme böyle eksiksiz bir modele sahip olmadığımızda optimal bir politika elde etmek için de kullanılabilir [12]. Model ile etkileşim deneyimlerini optimal bir politika belirlemek için kullanılabilir bilgiye dönüştüren iki yaklaşım vardır. Bu yaklaşımlarda geçiş olasılıklarının bilinmesine gerek yoktur. Bir yaklaşım, bir model oluşturmadan doğrudan deneyimlerden bir denetleyici türetmektir ve bu yaklaşım modelsiz yaklaşım olarak adlandırılmaktadır. Model-tabanlı yaklaşım olarak adlandırılan diğer yaklaşımda ise deneyimlerden bir model öğrenilir ve bundan bir model türetilir [12].

Pekiştirmeli öğrenmede ajanın geçici krediyi nasıl atayacağını bilmesi önemlidir. Bir eylemin kısa vadede olumlu bir kredisi olsa da aynı eylem gelecekte büyük olumsuzluklara sebep olabilir [21]. Bununla başa çıkmanın bir yolu her iterasyonun sonunda, sonuca göre, her eyleme ödül atamaktır. Bu yöntem büyük miktarda bellek gerektirir ve bu nedenle her zaman uygun değildir. Bu durumda, anlık bir ödülün ve bir sonraki durumun beklenen ödülünden durum-eylem çiftlerinin değerini tahmin eden zamansal fark yöntemleri kullanılabilir [11]. Bu şekilde, hem tahminler yalnızca iterasyon sonlarında güncelleme yapılmasına kıyasla daha sık güncellenmiş olur, hem de gelecekteki ödüllerin hesaba dahil edilmesine devam edilmiş olur.

Zamansal fark öğrenmesi (Temporal Difference Learning - TDL) genellikle $TD(\lambda)$ veya $TD(0)$ olarak tanımlanır. İkinci tanımda, veya lambda'nın 0'a ayarlandığı tanımda, algoritma tek-adımlı bir algoritma olacaktır, bu nedenle tahmin sonraki bir durumun değerini hesaba katar. λ 1 olarak ayarlanırsa, formülasyon Monte Carlo yaklaşımını uygulamanın daha genel bir yolu olacaktır. 0 ile 1 arasındaki herhangi bir λ değeri, karmaşık bir n-adımlı dönüş algoritması oluşturacaktır [22]. Zamansal farkın sıklıkla kullanılan iki algoritması Q-Öğrenme ve SARSA'dır. Bu iki algoritma takip eden alt başlıklarda ele alınmıştır.

Q-Öğrenme, uygulanması kolay olduğu ve gecikmeli ödülleri öğrenmede en etkili modelsiz algoritmalarından biri olduğu için, en çok kullanılan zamansal fark algoritmalarından biridir [12]. Q-Öğrenme politika-dışı bir algoritmadır. Bu da Q-Öğrenme fonksiyonunun politika dışındaki eylemlerden, rastgele eylemler gibi, öğrendiği anlamına gelmektedir, dolayısıyla bir politikaya ihtiyaç duyulmamaktadır. Durum-eylem çiftlerinin Q-değerleri Bellman denkleminde dayanmaktadır. Bu nedenle, aşağıdaki gibi tanımlanır:

$$Q^*(s, a) = R(s, a) + \gamma \max_a Q^*(s', a) \quad (2.3)$$

Dolayısıyla, s durumunda a eylemini gerçekleştirmenin değeri, bu s durumundan başlayarak a eylemini gerçekleştirmenin beklenen ödülü olarak tanımlanır. Q-Öğrenme tüm bilgileri bir tablo halinde saklar. Her durum-eylem çifti için, aşağıdaki değer fonksiyonunu [23] temel alarak optimal durum-eylem değeri $Q^*(s, a)$ 'nın bir $Q(s, a)$ tahminini öğrenir:

$$Q(s, a) \leftarrow Q(s, a) + \alpha [r + \gamma \max_a Q(s', a) - Q(s, a)] \quad (2.4)$$

Formül 3.4'teki γ gelecekteki ödüllerin önemiyle ilgili olan indirgeme faktörünü belirler. 0 faktörü, ajanın yalnızca mevcut ödülleri r göz önünde bulundurmasına sebep olarak miyop yapacaktır. Öte yandan, 1 değeri, uzun vadede yüksek ödüller için çaba sarf etmesini sağlayacaktır. α , öğrenme oranını belirtir ve 0 ile 1 arasında bir değere sahiptir. 0 değeri yalnızca istifade, 1 değeri ise yalnızca keşif anlamına gelmektedir. Öğrenme oranı uygun bir oranda azaltılırsa tahminlerin tüm olası çiftler için optimal değerlere yakınsadığı gösterilmiştir [24]:

$$Q(s, a) \rightarrow Q^*(s, a) \quad (2.5)$$

Her ne kadar bu keşif-istifadenin α kullanılarak ele alınması gerekse de, keşif stratejisinin detayları öğrenme algoritmasının yakınsamasını etkilemeyecektir. Bu nedenle, Q-Öğrenme keşif-duyarsızdır. Yani, Q değerleri veri toplanırken ajanın nasıl davrandığından

bağımsız olarak optimal değerlere yakınsar (tüm durum-eylem çiftleri yeterince sık denendiği sürece). Bu özellik Q-Öğrenme'yi çok popüler kılmaktadır [12]. Ancak, durum-eylem çiftlerinin sık sık ziyaret edilmesi gerektiğinden, bu durum Q-Öğrenme'nin aynı zamanda yavaş bir algoritma olması anlamına da gelmektedir. Bunun yanında, tüm bilgilerin tablo halinde saklanması sebebiyle tüm durumlar bir tabloda tutulduğundan dolayı bu algoritma aynı zamanda bellek-yoğun bir algoritmadır.

Diğer bir zamansal fark algoritması SARSA'dır. SARSA, "State-Action-Reward-State-Action" ($S_t, A_t, R_t, S_{t+1}, A_{t+1}$) ifadesinin kısaltmasıdır. Q-Öğrenme'nin aksine SARSA, politikaya dayalı bir algoritmadır, yani politikayı doğrudan eylemde bulunurken günceller. Hareket ederken bir kontrol politikası izler ve bunu, aşağıdaki formüle göre, Q-değerlerini güncellemek için kullanır [11]:

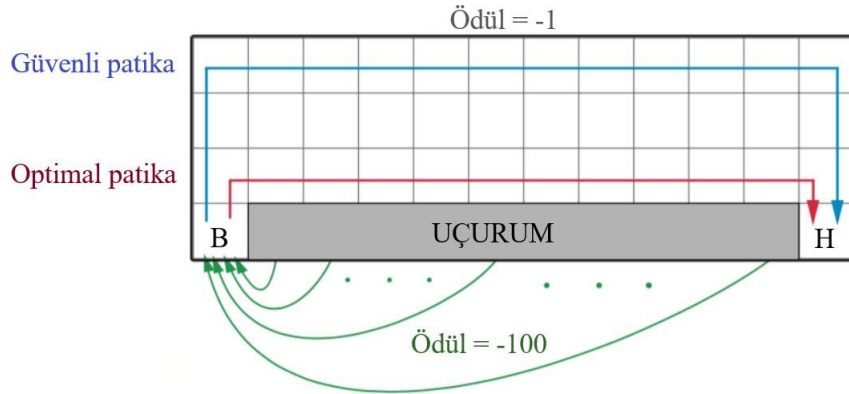
$$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma Q(s', a') - Q(s, a)] \quad (2.6)$$

SARSA ayrıca her durum-eylem çiftini saklamak için bir tablo kullanır ve davranış politikası π 'ye göre q_π 'yi tahmin eder ve aynı zamanda π 'yi q_π 'ye göre açgözlülük yönünde değiştirir. Başka bir deyişle, mevcut politikanın izlenmeye devam edildiğini ve bu arada daha az keşifçi ve daha istifadeci hale geldiğini varsayarak durum-eylem çiftlerinin dönüşünü tahmin etmektedir.

SARSA, Q-Öğrenme'nin daha genel bir versiyonu olarak düşünülebilir, algoritmalar çok benzerdir. Bir Q-Öğrenme uygulaması, pratikte, Q-değerleri için güncelleme yöntemini değiştirilerek ve her iterasyonda aynı politika kullanılarak mevcut eylem A seçilerek SARSA olarak değiştirilebilir. Bu durumda her adımda A yerine A' seçilirdi. Ancak performansta önemli bir fark gerçekleşebilir. Bunu göstermek için Sutton ve Barto'nun [11] ünlü bir örneği kullanılabilir: uçurumda yürüme problemi.

Bu uçurumda yürüme probleminde ajanın attığı her adım için -1, uçurumdan düşmek için ise -100 ceza vardır. Bu nedenle en optimal yol, tam olarak uçurumun kenarı boyunca

yürümek ve mümkün olan en az adımda ödüle ulaşmaktır. Ajan herhangi bir noktada uçurumdan düşmediği sürece bu strateji ödülü en üst düzeye çıkaracaktır.



Şekil 2.1: Uçurum problemi Sutton ve Barto [8].

Şekil 2.1 iki farklı yol göstermektedir. Q-öğrenme optimal yolu, SARSA ise güvenli yolu seçer. Sonuç olarak, keşif-temelli bir politika izlendiğinde, bir Q-öğrenme ajanının herhangi bir noktada keşif seçiminin bir sonucu olarak uçurumdan düşme riski vardır.

SARSA, ajanın bir sonraki adımda gerçekte ne yapacağını görmek için bir sonraki eyleme bakar ve mevcut durum-eylem çiftinin Q-değerini buna göre günceller. Bu nedenle, ajanın uçuruma düşebileceğini ve bunun büyük bir negatif ödüle yol açacağını öğrenir, bu nedenle bu durum-eylem çiftlerinin Q değerlerini buna göre düşürür. Sonuç olarak, Q-Öğrenme, ajanın mümkün olan en iyi politikayı (bu politikanın gerçekte ne olduğunu çözmeye çalışmaksızın) izlediğini varsayarken, SARSA, ajanın gerçek politikasını da hesaba katmaktadır.

Ajanın politikası sadece açgözlü politika iken Q-Öğrenme ve SARSA aynı sonuçları üretecektir. Pratikte çoğu uygulama basit bir açgözlü strateji kullanmaz ve bunun yerine bazı eylemlerin rastgele seçildiği e-açgözlü gibi bir strateji seçer.

SARSA birçok problemde Q-öğrenmeden daha iyi performans gösterecektir, özellikle de ajanın bir sonraki adımda rastgele bir optimal-altı eylem seçme şansının yüksek olduğu durumlarda, uçurum probleminde görüldüğü gibi. Bu durumda, Q-Öğrenme'nin ajanın optimal politikayı izlediğini varsayımı gerçek durumdan çok uzak olabilir ve sonuç olarak SARSA daha hızlı bir şekilde ve daha az hatayla yakınsayabilir [25].

2.2. Derin Pekiştirmeli Öğrenme

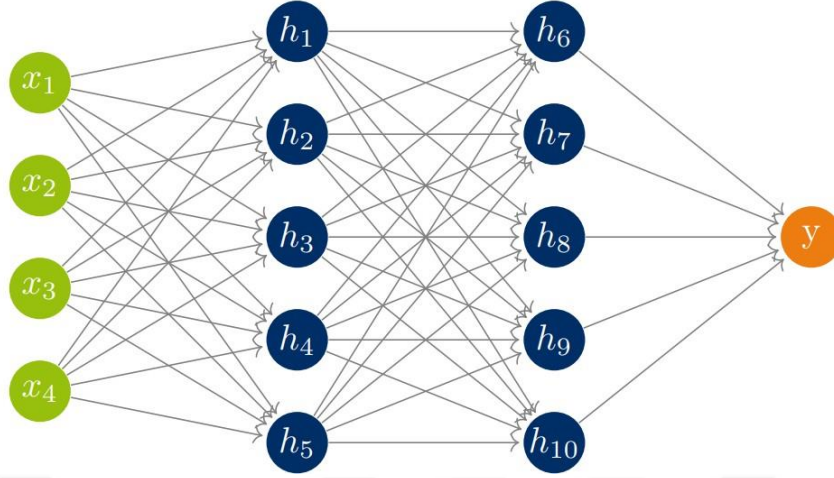
Şimdiye kadar tartışılan pekiştirmeli öğrenme yöntemlerinde tüm durumların keşfedilebileceği ve her bir durum hakkında öğrenilen her şeyin bir tabloda sunulabileceği varsayıldı. Ancak, büyük, gerçekçi problemleri çözmek için bu yöntemlerin kullanılması istendiğinde, durumların ve eylemlerin miktarı, tüm bilgileri bir tabloda depolamak için fazla büyük olacaktır. Yalnızca depolama sınırlayıcı bir faktör değildir, her bir durum-eylem çiftinin tahmin edilebilmesi için büyük miktarda zamana da ihtiyaç vardır aynı zamanda. Bu sınırlamalar bir genelleme problemi olarak tanımlanabilir. Bu sınırlamaların üstesinden gelmek için, daha büyük bir kümeye yakınsamak için durum-eylem çiftlerinin bir alt kümesi üzerinde eğitim alırken edindiği deneyimi genelleştirebilen bir yöntem kullanılması gerekmektedir. Değer fonksiyonu yakınsama böyle bir yöntemdir. Değer fonksiyonu yakınsama ile bir tablo yerine sadece parametrelerinin saklanması gereken bir fonksiyon kullanılarak Q-değerleri temsil edilebilir. Bu şekilde, daha az alana ihtiyaç duyulur ve ziyaret edilen durumlar ziyaret edilmeyen durumlara genellenebilir, böylece her olası durumun ziyaret edilmesi gerekmez.

Pekiştirmeli öğrenme için değer fonksiyonu yakınsama kullanmanın birçok yolu vardır ancak popülerlik kazanan bir yöntem derin öğrenmenin kullanılmasıdır. Derin öğrenme, öğrenme yönteminin değer fonksiyonunu tahmin etmek için bir veya birden fazla sinir ağı kullanması anlamına gelir. Yakın tarihli bir makalede Mnih [26] derin öğrenmeyi bir değer fonksiyonu yakınsayıcısı olarak uygulamış ve etkileyici sonuçlar elde etmiştir. Derin öğrenme pekiştirmeli öğrenme ile birleştirilmişti çünkü iki tekniğin de kendi dünyası ile etkileşime girebilen ve ondan öğrenebilen genel-amaçlı yapay zeka sistemleri oluşturmak gibi ortak bir hedefleri vardı [27]. Pekiştirmeli öğrenme karar verme için genel-amaçlı bir çerçeve sağlarken derin öğrenme ise temsili öğrenme için genel-amaçlı çerçeveyi sağlar. Belirli bir hedef ve minimal alan bilgisi ile doğrudan ham girdilerden en iyi temsili öğrenebilir [11].

2.2.1. Sinir ağları

Bir sinir ağı, birden çok doğrusal olmayan işlem birimi katmanından oluşur. Bu ağ girdiyi alır ve çıktılara dönüştürür. İki gizli katmana ve bir çıktı katmanına sahip bir sinir ağı örneği **Şekil 2.2**'te gösterilmektedir.

Girdi katmanı Gizli katman 1 Gizli katman 2 Çıktı katmanı



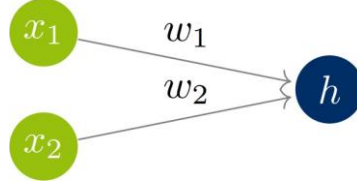
Şekil 2.2: Bir sinir ağının görselleştirilmesi.

Sinir ağının her katmanı bir dizi nöron içerir. Nöron, ağırlık parametresi olan bir hesaplama birimidir ve **Şekil 2.3**'te görselleştirilmiştir. Bu nöron önceki katmandan girdi alır ve bunu bağlantının ağırlığı (w) ile çarpar. Bu ağırlıklandırılmış girdiler bir eşik değeri (bias) ile birlikte toplanır. Son olarak, toplam bir aktivasyon fonksiyonundan geçirilir. **Şekil 2.3**'teki örnek ele alınırsa bu işlem aşağıdaki aktivasyon fonksiyonu ile sonuçlanacaktır:

$$y = f(x_1 * w_1 + x_2 * w_2 + b) \quad (2.7)$$

Aktivasyon fonksiyonu, sınırlanmamış bir girdiyi öngörülebilir biçimdeki bir çıktıya dönüştürmek için kullanılmaktadır ve genellikle doğrusal değildir. Sigmoid fonksiyonu gibi birden çok türde aktivasyon fonksiyonu vardır. Bu fonksiyonun 0 ile 1 aralığında bir çıktısı vardır, bu yüzden değerleri sıkıştırır. Bu aktivasyon fonksiyonunun sonucu daha sonra bir sonraki katmana iletilir. Çıktı katmanı gizli katmanla aynı şekilde hesaplanmaktadır fakat nihai değeri çıktı katmanı sağlar. Daha sonra, ihtiyaç duyulan doğru çıktı, çıktı katmanındaki bir kayıp fonksiyonu kullanılarak hesaplanır ve bu işlem için sıklıkla ortalama hata karesi veya log-olabilirlik fonksiyonu kullanılmaktadır [28]. Bu kayıp ne kadar düşükse tahmin de o kadar iyidir.

Sinir ağı eğitilirken, girdi, çıktısı bilinen bir ağa verilir. Bu girdi ile ağı ağırlığı ve



Şekil 2.3: Sinir ağı içerisinde bir nöron.

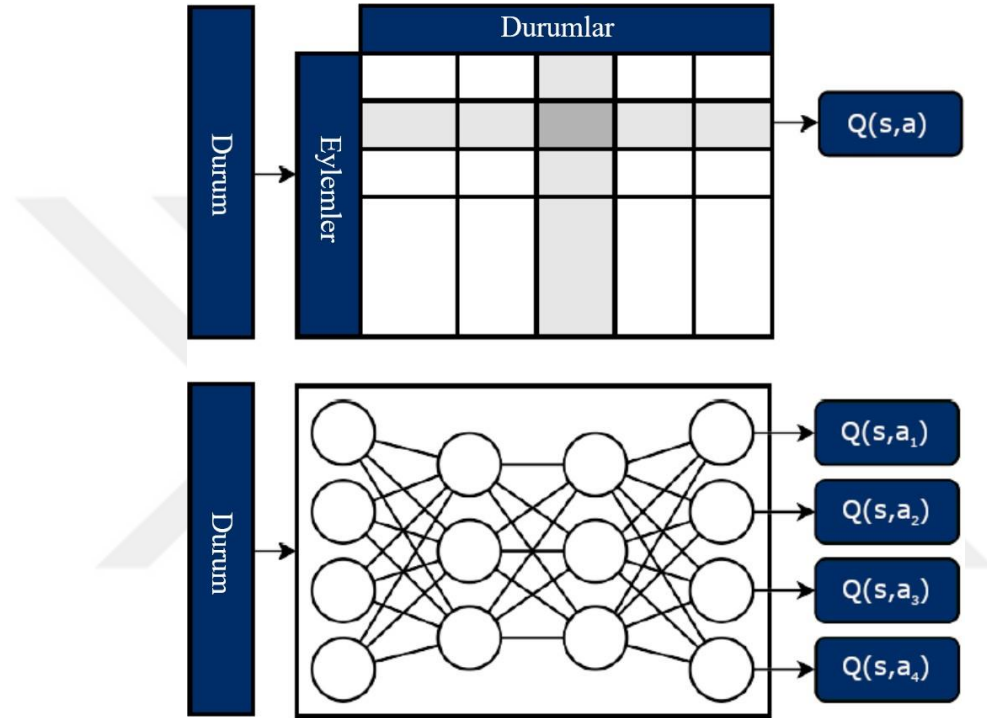
eşik değeri doğru çıktıya götürecektir şekilde ayarlanabilir. Bu genellikle, her bağlantı ve nöron için tam olarak ne kadar ağırlık ve eşik değeri gerektiğini hesaplayabilen stokastik gradyan iniş kullanılarak yapılmaktadır. Yöntem için belirli bir kayıp değeri veya doğruluk elde edildiğinde ağa yeni veriler eklenebilir. Sinir ağı daha sonra belirli bir girdinin tahmini çıktısını hesaplayabilir. Derin öğrenme ile pekiştirmeli öğrenme birleştirilirken kullanılan girdi, tanımlanan durumdur. Bu girdi katmanı ile ağ, hangi eylemin gerçekleştirilmesi gerektiğini tahmin edebilir. Çıktı katmanı eylem uzayına karşılık gelmektedir. Bu, her eylem için belirli bir tahmin ve standart sapma döndürerek veya tüm eylemler için birer olasılık döndürerek olabilir. Kazanılan ödüller veya cezalar ağı güncellemek için kullanılmaktadır. Bir ceza verilirse, ajan aynı alana düştüğünde başka bir eylem seçilecek şekilde ayarlanacaktır ağırlıklar.

2.2.2. Algoritmalar

Tıpkı pekiştirmeli öğrenmede olduğu gibi, derin pekiştirmeli öğrenme ile çalışmak için çeşitli algoritmalar geliştirilmiştir. Sonraki alt başlıklarda popüler derin pekiştirmeli öğrenme algoritmaları tartışılmıştır.

Mnih [26] yeni bir değer fonksiyonu tabanlı derin pekiştirmeli öğrenme yöntemi tanıtmıştır: Derin Q-Ağı. Bu makalede çeşitli sayıda Atari 2600 oyunu oynamayı öğrenmek için tek bir mimari kullanılmıştır. Derin Q-Ağı neredeyse tüm oyunlarda rakip yöntemlerden daha iyi performans göstermiştir. Oyun görüntülerini doğrudan girdi olarak kullanabilen ilk pekiştirmeli öğrenme algoritmasıdır ve birçok oyuna uygulanabilmiştir. Q-Öğrenme bir tablo kullanarak tüm durum-eylem kombinasyonları için Q değerlerini depolarken, Derin Q-Ağı bir eylemin performansını tahmin etmek için sinir ağları kullanmaktadır. Q-Öğrenmede, Q-değerleri bir durum-eylem çifti için hesaplanır. Derin Q-Ağı'nda

ise Q-değerleri spesifik bir durum için hesaplanır. Bu fark **Şekil 2.4**'te görselleştirilmiştir. Durum-eylem çiftleri yerine durumları kullanma kavramı yaklaşık dinamik programlamada da sıklıkla kullanılmaktadır [29]. YDP'de bu, bir karar verildikten hemen sonra ancak herhangi bir yeni bilgi gelmeden önce sistemin durumunu yakalayan karar-sonrası durum değişkeni olarak kullanılmaktadır. Karar-sonrası durum değişkenini kullanmanın avantajı, her eylem için tüm olası sonuçların değil de sadece farklı durumların değerlendirilmesinin gerekli olmasıdır [20].



Şekil 2.4: Q-tablosu ile Q-ağı arasındaki farkların görselleştirilmesi. Gemmink (2019)'dan [27] uyarlanmıştır.

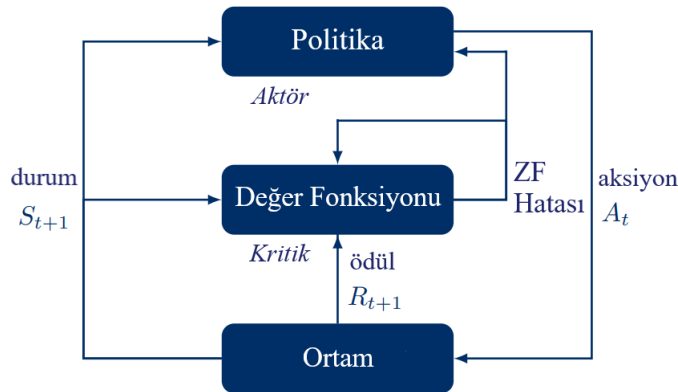
Derin Q-Ağı için değer fonksiyonu yakınsayıcısı **Denklemler 2.8**'de gösterilmiştir [26]. Görüldüğü üzere bu denklem hem durumu hem de eylemi göstermektedir çünkü hala hem durum hem de eylem kullanılarak Q değerlerinin yakınsanması istenmektedir.

$$Q^*(s, a) = \max_{\pi} E[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots | s_t = s, a_t = a, \pi] \quad (2.8)$$

Bu denklem, s durumunda a eylemini gerçekleştirdikten sonra, davranış politikası $\pi = P(a|s)$ ile başarılabilen, indirgeme faktörü γ ile zaman içinde indirgenen ödüllerin (r_t, r_{t+x}) en iyi toplamını tanımlamaktadır [26]. Pekiştirmeli öğrenmedeki diğer

değer fonksiyonu yakınsamalarının dengesizleştiği bilindiği halde Mnih [26] bu dengesizlik sorununu iki yeni fikirle çözmektedir: deneyim tekrarı ve hedef ağın iteratif güncellemeleri. Deneyim tekrarı ile ajanın her bir adımdaki deneyimleri bir veri setinde depolanır ve bir Q-Öğrenme güncelleme iterasyonu yürütüldüğünde bu deneyimlerin bir örneği rastgele çekilir. Bu yöntemle gözlem dizilerindeki korelasyonlar kaldırılır. Hedef ağ yalnızca birkaç iterasyondan sonra güncellenir. Deneyim tekrarı için bir gereklilik ortam kurallarının geçmiş deneyimlerin alakasız hale geleceği şekilde değiştirilmemesidir [30]. Bu fikirlerle Mnih [26] yakınsak değer fonksiyonunu başarılı bir şekilde parametrize edebilmektedir.

Aktör-kritik algoritması ilk olarak Lillicrap [31] tarafından tanıtılmıştır ve değer fonksiyonu tabanlı yöntemler (Q-Öğrenme ve Derin Q-Ağı gibi) ile politika tabanlı yöntemlerin bir kombinasyonudur. Aktör-kritik yöntemlerinde, aktör bir politika gibi hareket eder ve politika dağılımını kritiğin önerdiği doğrultuda günceller. Kritik değer fonksiyonunu tahminler. Tahmin, ortamın tüm eylemlerinin dönüşlerinin değer fonksiyonu yakınsayıcısına dayanır. Bu algoritma daha küçük gradyanlar verir, bu da daha düşük varyanslı daha dengeli güncellemeler sağlar. Aktör-Kritik yöntemlerinin mimarisi **Şekil 2.5**'de gösterilmektedir. Mnih'de [32], Aktör-Kritik algoritması Eşzamansız Avantaj Aktör-Kritik (Asynchronous Advantage Actor-Critic - A3C) algoritmasına genişletilmiştir. Eşzamansızdır, çünkü aynı derin sinir ağı ile ortamın birden çok örneğinde birden çok ajanı paralel olarak yürütmektedir. Deneyim tekrarı yerine farklı aktörlerin deneyimleri ile güncelleme yapılmaktadır. Bu şekilde, lokal bir optimumda takılı kalma şansı azaltılmış olmaktadır.



Şekil 2.5: Aktör-Kritik mimarisi Sutton ve Barto [8].

Şu anda, en başarılı derin pekiştirmeli öğrenme yöntemlerinden biri Aktör-Kritik mimarisine dayanan Proksimal Politika Optimizasyonu (Proximal Policy Optimization - PPO) yöntemidir ve Schulman [1] tarafından tanıtılmıştır. Politika arama yöntemleri genellikle doğal politika gradyanı tarafından ele alınan yakınsama sorunlarına sahiptir. Bu doğal politika gradyanı, bu yöntemi büyük problemler için ölçeklenemez hale getiren ikinci-dereceden bir türev matrisi içermektedir. PPO biraz farklı bir yaklaşım kullanmaktadır. Gradyan iniş yöntemini hedefi optimize etmek için kullanır. Bu şekilde, yeni politikanın eski politikadan çok uzaklaşmamasını sağlamak için hedef kademeli olarak güncellenmektedir. Bu, algoritmayı gözlemlerindeki aykırı değerlere karşı daha az duyarlı hale getirmektedir. Schulman [1], PPO yöntemini Mnih'in [26] Derin Q-Ağı ile karşılaştırmış ve birçok durumda bu yöntemi geride bırakmıştır. PPO verimli örnekler oluşturabilmektedir ve hesaplama açısından daha az yoğunur.

2.3. Kan Ürünleri Tedarik Zinciri Yönetimi

Rutin tıbbi tedaviler veya cerrahi operasyonların bir parçası olarak ve bunlara ilave olarak acil durumlarda hastalara kan ürünleri transfüzyonu yapılmaktadır. Bu sebeple doğru kan ürünlerini elde bulundurmak hayati önem taşımaktadır çünkü envanter bulunmaması durumu can kaybıyla sonuçlanabilir. Bununla birlikte, kan ve kan ürünleri yalnızca insanlardan temin edilebilmektedir ancak dünya genelinde gönüllü bağışçı azlığı problemi yaşanmaktadır [33]. Kan bağışı (arz) ile kan nakli ihtiyacı (talep) arasındaki makas giderek açılmaktadır. Mevcut durumun yönetimi ve optimal çözümlerin üretilmesi konusunda aksiyonlar alınması gerekmektedir.

Kan ürünlerine artan talebin yanı sıra azalan bağışçı popülasyonu göz önüne alındığında tüm bunlar kan tedarik zinciri için karar vermenin zor olduğu anlamına gelmektedir. Diğer bir önemli faktör ise maliyettir. Özellikle gelişmiş ülkelerde, kanın kendisi gönüllü olarak bağışlansa da, süreç boyunca işçilik, test, fraksiyonasyon (tam kanın alt ürünlerine ayrılması), depolama ve dağıtım gibi birçok maliyet ortaya çıkmaktadır. Verimli bir kan tedarik zinciri talebi karşılarken aynı zamanda da israfı ve maliyetleri de en aza indirmekle yükümlüdür. Bununla birlikte, kan ürünlerinin sınırlı raf ömrü,

halihazırda son derece karmaşık olan bu tedarik zinciri üzerine katı kısıtlamalar getirmektedir ve bu durum sipariş karşılayamama ve raf ömrünü doldurma risklerini önemli ölçüde artırmaktadır.

İnsan hayatı söz konusu olduğundan kan tedarik zincirinin önemini kavramak kolaydır fakat bu tedarik zincirinin karmaşıklığını kavramak göreceli olarak belki de daha zor olacaktır. Katsaliaki ve Brailsford'a [34] göre, ürünler ve alt ürünler dahil olmak üzere kandan yüzden fazla farklı ürün elde edilebilir, ancak kırmızı kan hücreleri (eritrosit), plazma ve trombosit en önemlileri olarak kabul edilmektedir. Eritrosit toplam transfüze edilen ürünlerin %63.4'ünü temsil eder, bunu %17.8 ile plazma, %13.6 ile trombosit ve son olarak %5 ile kriyopresipitat takip eder [35]. Ayrıca bu ürünler, özel işlemler için işinlanılmış veya yıkanmış ürünler gibi alt ürünler elde etmek için veya rekombinant ürünler gibi diğer ürünler için hammadde olarak da işlenebilmektedir.

Kan bileşenleri farklı durumlarda kullanılmaktadır. Örneğin, anemi tedavilerinde eritrosit gerekirken kanser hastaları için trombosit ve yanıkları olan hastaları tedavi etmek için plazma gerekmektedir. Bu örnekler her bileşenin kullanımına ilişkin yalnızca birer örneği temsil etmektedir fakat her bir bileşenin sağlık hizmetlerinde farklı süreçlerde birçok kullanımı olabilir. **Çizelge 2.1**, ana kan bileşenlerinin başlıca kullanımlarını, raf ömrünü ve saklama koşullarını göstermektedir. Öte yandan, kan toplamak sürekli bir çaba gerektirmektedir; konfor, riskler, kolaylık ve erişilebilirlik gibi birçok faktör bağış kararını etkileyebilir. Ancak tabii ki talebi karşılayabilmek için her zaman yeterli kan toplanması gerekmektedir. Arz ile talebin eşleştirilebilmesi: kan ve ürünlerinin toplanması, işlenmesi ve dağıtılmasını sağlayan bir altyapı gerektirmektedir. Hastane dahili kan bankalarından, çeşitli talep noktaları sağlayan çoklu toplama ve işleme ve dağıtım merkezlerine kadar birçok farklı kan tedarik zinciri konfigürasyonu bulunabilir, hem gerçek dünyada hem de literatürde. İzlenen stratejiler ülkenin sağlık politikalarına göre değişiklik göstermektedir. Her halükarda amaç aynı kalmaktadır: kan ürünleri talebini minimum maliyetle ve minimum atıkla karşılamak.

Bunlarla birlikte, kan ve kan ürünlerinin kendine has özellikleri de dikkate alınmalıdır. Kan grupları, gruplar arası uyumluluklar ve ürünlerin farklı raf ömürlerine sahip olması gibi faktörler kan tedarik zinciri sistemlerine ve karar verme süreçlerine karmaşıklık katmaktadır. Sekiz ana kan grubu (her biri RH pozitif veya negatif olabilen; A, B,

Çizelge 2.1: Kan ürünlerinin özellikleri. American Red Cross (2017).

Kan ürünü	Ana kullanım alanları	Raf ömrü	Saklama koşulları
Tam kan	Travma ve cerrahi işlemler	21/35 gün	18–24 °C
Eritrosit	Travma, cerrahi işlemler, kansızlık ve kan kaybı	42 gün	2–10 °C
Trombosit	Kanser tedavisi ve organ nakli	5-7 gün	20–24 °C
Plazma	Yanık hastaları ve kanama bozuklukları	1 yıl	≤ -30 °C
Kriyopresipitat	Hemofili ve Von Willebrand hastalığı	1 yıl	≤ -30 °C

AB ve O) vardır ve her grubun popülasyonda farklı bir oranı vardır ve bu oranlar etnik kökenlere ve coğrafi bölgelere göre değişim göstermektedir. Örneğin Türkiye’de O-pozitif ve AB-negatif oranları sırasıyla %29,4 ve %0,7’dir [36]. Bazı çok nadir kan grupları olduğundan, sık sık ikame ürünlerin kullanılması gerekmektedir. Ancak, ikame ürün kullanmada belirli kısıtlamalar ve tercihler vardır. Genellikle, zorunlu durumlar hariç, aynı kan grubu tercih edilmektedir. Kan ürünlerinin raf ömrü de dikkate alınması gereken bir diğer önemli faktördür; trombosit, eritrosit ve plazma farklı raf ömürlerine sahiptirler. Trombosit sadece 5-7 günlük raf ömrü ile en kısa ömürlü bileşendir. Bunu, 42 gün raf ömrü ile eritrosit ve bir yıl raf ömrü ile plazma takip etmektedir. Raf ömrü sona ermeden önce transfüze edilememişse kan ürününün atılması gerektiği anlamına gelmektedir bu durum. Ayrıca, kan ürünlerinin bağımsız olarak üretilmediğini belirtmek de önemlidir. Çeşitli sayıda ürün üretebilen farklı fraksiyonlama alternatifleri ve kan toplama yöntemleri vardır.

Kan; bölgesel kan merkezi, kan bağıışı merkezi, hastane kan bankası veya mobil birim gibi toplama sitelerindeki birimlerde toplanır. Toplandıktan sonra bir dizi teste tabi tutulur ve tam kan şeklinde tutulur veya bileşenlere ayrılır. Daha sonra, transfüzyon taleplerini karşılamak üzere bir hastane kan bankasına gönderilir.

Hastane kan bankası, transfüzyon taleplerini karşılamak için uygun kan ünitelerini depolayan ve arz eden bir envanter noktası olarak çalışır. Kan Bankası gün içerisinde her bir kan grubu için, rastgele sayıda ünite için, rastgele sayıda transfüzyon talebi alır. Bir hasta için bir talep alındığında, bu kan grubundan uygun sayıda ünite serbest envanterden (atanmamış envanter) kaldırılır ve başarılı çapraz karşılaştırma (crossmatching) sonrasında belirli bir hasta için rezerve envantere (atanmış envanter) yerleştirilir. Çapraz karşılaştırma, bağışlanan kanın hastaya transfüzyon için uygun olup olmadığının

anlaşılması için yapılan bir testtir. Transfüze edilmeyen üniteler, gelecekte talebi karşılamak için kullanılacak atanmamış envantere döndürülür. Hasta ameliyatı ile kullanılmayan ünitelerin atanmamış envantere geri dönüşü arasında geçen süreye çapraz karşılaştırma serbest bırakma periyodu denir. Çapraz karşılaştırma/transfüzyon oranı, çapraz karşılaştırılan ünite sayısının transfüze edilen ünite sayısına oranı olarak tanımlanır.

Talep ve kullanım, sırasıyla istenen ünite sayısı ve transfüze edilen ünite sayısını ifade eder. Raf ömrü boyunca kullanılmayan üniteler, kullanılamaz kabul edilir ve envantere çıkarılır [37]. Bir talep için belirlenen süre içerisinde aynı grupta kan bulunamazsa, talep, başka bir uyumlu kan grubu tarafından karşılanır. Bu işleme uyumsuz karşılaştırma (mismatching), bu periyoda ise uyumsuz karşılaştırma aktivasyon periyodu adı verilir.

Kan bankacılığı sağlık sistemlerinin önemli bir parçasıdır ve uygulamalarının tıbbi tedavi prosedürlerinin başarısı üzerinde önemli bir etkisi vardır. Kan bankacılığı uygulamaları, özellikle 70'lerden itibaren, yöneylem araştırmacıları ve sağlık çalışanları tarafından her zaman yoğun ilgi görmüştür. Literatürde kan bankacılığını popüler yapan başlıca nedenler şu şekilde özetlenebilir: 1 - Birçok tıbbi tedavi prosedürünün başarısı üzerinde büyük etkisi vardır. 2 - Kan bağıışı ile kan nakli ihtiyacı arasındaki fark giderek açılmaktadır. 3 - Gönüllü kan bağıışçısı oranında azalma yaşanmaktadır. 4 - Kan tedarik zinciri planlaması; çoklu-ürün yapısı, bozulabilirlik, hem arz hem de talep tarafında belirsizlik gibi problemin özgün doğasından kaynaklanan sebeplerden ötürü karmaşık bir yapıya sahiptir. 5 - Güvenli kan tedarik maliyeti giderek artmaktadır.

Kan kaynaklarının yönetimi hiyerarşik olarak ilişkili üç düzeyde incelenebilir; bireysel hastane düzeyi, bölgesel düzey ve bölgeler arası düzey. Bu seviyelerin her birinde, karar verici çok çeşitli operasyonel, taktiksel ve stratejik yönetim sorunlarıyla karşı karşıya kalmaktadır [37]. Bu karar problemlerinin analitik olarak modellenmesi ve ardından da verimli ve etkili bir envanter yönetim sistemi olarak uygulanması karmaşık bir görevdir.

Bu çalışmada; karar verici olarak hastane kan bankası, taktiksel karar seviyesi olarak envanter yönetimi, kan ürünleri arasından ise eritrosit bileşeni ele alınmıştır. Takip eden bölümde, eritrosit envanter yönetimi literatürü özetlenmiş ve konu hakkında 2005 yılından itibaren günümüze dek yayınlanmış olan makaleler tablo halinde özetlenmiştir.

2.3.1. Eritrosit envanter yönetimi

Hastane kan bankasının temel görevleri; kan ürünlerini depolama, çapraz karşılaştırma, bulunduğu bölgenin kan merkezinden kan ürünleri sipariş verme ve bu ürünlerin teslim alınması, ürünlerin hastalara atanması ve ürünlerin envanter yönetimi olarak özetlenebilir. Kan ürünlerinin tedariği tek yetkili kurum olan Kızılay'dan sağlanmaktadır ve bu ürünlerin talebi rutin tıbbi işlemler, cerrahi operasyonlar, veya acil durumlarda kullanılmak üzere hekimler tarafından gerçekleştirilmektedir.

Hastaların tedavisinde kan çok önemli olduğu için yüksek bir hizmet seviyesi gereklidir. Hastaların gerekli kan transfüzyonlarını gecikmeden alması gerektiği ve bazı hastaların ihtiyaçları önceden bilinemediği için her daim yeterli miktarda kullanıma hazır kan ürünü bulundurulması önemlidir. Talep karşılayamama endişesi ile envantere fazla ürün bulundurulması ise uygunsuz envanter birikimine yol açabilir ve bu da israf düzeylerini artırabilir. Kanın, sınırlı üretimi olan ve yeri doldurulamaz bir ürün olduğu düşünüldüğünde, kan ünitelerinin israfı etik değildir. Ayrıca, israf sağlık sistemlerine gereksiz maliyetler de getirmektedir. Talep karşılayamama ve israf arasındaki bu denge kan tedarik zinciri boyunca kan envanterinin yönetilme biçiminden doğrudan etkilenir. Özellikle de hastane seviyesindeki envanter yönetiminden etkilenmektedir çünkü bu seviye tedarik zincirinin son aşamasıdır ve hastalarla doğrudan bağlantılıdır.

Bununla birlikte, arz ve talepteki belirsizlik, bozulabilirlik sebebiyle uzun süreler boyunca envanter biriktirmenin imkansız olması gibi sebepler dolayısıyla kan envanter yönetimi son derece karmaşıktır. Belirsiz ortam ve kısa raf ömrü, israfı ve ortaya çıkan maliyetleri olumsuz yönde etkileyen gereksiz stoklara yol açabilir [38]. Sekiz benzersiz ABO/Rhesus kan grubu (A+, A-, B+, B-, AB+, AB-, O+ ve O-) olduğundan, kan envanterini yönetmek daha da zorlaşmaktadır. Ek olarak, eritrositler kendi aralarında belirli bir ikame edilebilirlik düzeyine de sahiptir, ki bu konu literatürde pek ele alınmamıştır [39].

Eritrosit envanter yönetimi alanında 2005-2022 yılları arasında yayınlanan makalelerin özet bilgileri **Çizelge 2.2**'de paylaşılmıştır. Kan ürünleri, envanter yönetimi açısından farklı karakteristiklere sahip olduğundan dolayı eritrosit bileşenini incelemeye almayan makaleler kapsam dışı bırakılmıştır. Yalnızca doğrudan eritrosit, veya eritrositle beraber tam kan, trombosit veya plazma ürünlerini incelemeye alan çalışmalar göz önünde bulundurulmuştur. Kan tedarik zincirinin; transfüzyon, bileşen işleme, test etme, taşıma ve

bağış toplama gibi diğer ilgi alanlarına odaklanan makaleler ayıklanmış, geriye yalnızca envanter yönetimi problemine odaklanan çalışmalar literatür taramasına dahil edilmiştir. Tüm eleme adımlarından sonra geriye kalan 36 makale **Çizelge 2.2**'de özetlenmiştir.

Çizelge 2.2: Eritrosit envanter yönetimi.

Makale	Kademeler	Ağ topolojisi	Bileşenler	Yaklaşım
[40]	Kan merkezi – Talep noktası	Kan merkezi - Çoklu Talep noktası	Eritrosit	MINLP-MILP
[41]	Kan merkezi – Talep noktası	Kan merkezi - Çoklu Talep noktası	Eritrosit	LP, DES
[42]	Kan merkezi – Talep noktası	Çoklu Kan merkezi - Çoklu Talep noktası	Eritrosit	RS
[43]	Kan merkezi – Talep noktası	Kan merkezi - Çoklu Talep noktası	Eritrosit	MINLP-MILP
[33]	Kan merkezi – Talep noktası	Çoklu Kan merkezi - Çoklu Talep noktası	Eritrosit	DES
[44]	Kan merkezi – Talep noktası	Kan merkezi - Talep noktası	Eritrosit	MILP
[45]	Bağış noktası - Kan merkezi - Talep noktası	Çoklu Bağış noktası - Kan merkezi - Çoklu Talep noktası	Tam kan-Trombosit-Eritrosit-Plazma	DES
[46]	Bağış noktası - Kan merkezi - Talep noktası	Çoklu Bağış noktası - Çoklu Kan merkezi - Çoklu Talep noktası	Eritrosit-Trombosit	ILP
[47]	Talep noktası	Talep noktası	Eritrosit	MILP
[48]	Talep noktası	Talep noktası	Eritrosit	MILP
[49]	Kan merkezi – Talep noktası	Çoklu Kan merkezi - Çoklu Talep noktası	Eritrosit	RS
[50]	Kan merkezi – Talep noktası	Kan merkezi - Çoklu Talep noktası	Eritrosit	MILP

[51]	Bağış noktası - Kan merkezi - Talep noktası	Çoklu Bağış noktası - Kan merkezi - Çoklu Talep noktası	Eritrosit	MILP
[52]	Bağış noktası - Kan merkezi - Talep noktası	Çoklu Bağış noktası - Çoklu Kan merkezi - Çoklu Talep noktası	Eritrosit	CMIP
[53]	Kan merkezi – Talep noktası	Kan merkezi - Talep noktası	Eritrosit	SDS
[54]	Talep noktası	Talep noktası	Eritrosit	MILP
[55]	Kan merkezi	Kan merkezi	Eritrosit	DES
[56]	Kan merkezi – Talep noktası	Kan merkezi - Talep noktası	Eritrosit	MINLP-MILP
[38]	Kan merkezi – Talep noktası	Kan merkezi - Talep noktası	Tam kan-Eritrosit-Trombosit	DES
[57]	Kan merkezi – Talep noktası	Kan merkezi - Talep noktası	Eritrosit	MINLP-MILP
[58]	Talep noktası	Talep noktası	Eritrosit	TDS
[59]	Kan merkezi – Talep noktası	Kan merkezi - Talep noktası	Eritrosit-Trombosit	MINLP-MILP
[60]	Kan merkezi – Talep noktası	Kan merkezi - Çoklu Talep noktası	Eritrosit	DES
[61]	Kan merkezi	Kan merkezi	Eritrosit	DES
[62]	Talep noktası	Talep noktası	Eritrosit	DES
[63]	Kan merkezi	Kan merkezi	Eritrosit-Trombosit	DES
[64]	Kan merkezi – Talep noktası	Çoklu Kan merkezi - Talep noktası	Eritrosit	DES

[65]	Kan merkezi	Kan merkezi	Tam kan- Trombosit- Eritrosit- Plazma	DES
[66]	Kan merkezi – Talep noktası	Kan merkezi - Talep noktası	Eritrosit	MCS
[67]	Kan merkezi – Talep noktası	Çoklu Kan merkezi - Talep noktası	Eritrosit	DES
[68]	Talep noktası	Talep noktası	Eritrosit	DES
[69]	Kan merkezi	Kan merkezi	Eritrosit	DES
[10]	Kan merkezi	Kan merkezi	Eritrosit	DES
[70]	Kan merkezi – Talep noktası	Kan merkezi - Çoklu Talep noktası	Eritrosit- Trombosit	DES
[71]	Kan merkezi – Talep noktası	Kan merkezi - Talep noktası	Eritrosit- Plazma- Trombosit	DES
[72]	Kan merkezi – Talep noktası	Kan merkezi - Talep noktası	Eritrosit	DES
Bu tez	Talep noktası	Talep noktası	Eritrosit	DRL

İlk sınıflandırma elementi, kademeler, kan tedarik zincirinin ana kademeleri olan, sırasıyla: bağışçı, bağış noktası, kan merkezi, talep noktası ve hasta kademelerinden hangilerinin makalede ele alınan sisteme dahil edildiğini özetlemektedir. Bazı makaleler kan merkezlerini (hastanelere kan dağıtımından sorumlu olan merkezler) veya talep noktalarını (hastaneler, klinikler veya diğer transfüzyon noktaları) yalnız başına ele alsa da, çoğu makale, ana arz ve talep noktaları olan kan merkezleri ve talep noktalarını birlikte incelemeye almıştır. Makalelerden hiçbiri bağışçı veya hasta tarafına odaklanmamış, 4 makale ise bağış noktalarını da çalışmaya dahil etmiştir. Ele alınan kademelerin ağ topolojisi, hangi kademedan kaç sayıda noktanın ele alındığı, tekli veya çoklu, bir sonraki sütunda özetlenmiştir. 27 makale, bu çalışmadaki şekilde, yalnızca eritrosit ürününü araştırmaya dahil etmiş, geri kalan makaleler ise eritrosit bileşeni ile birlikte en az bir kan ürününü daha dikkate almıştır.

Literatürde en sık görülen amaçlar toplam maliyet minimizasyonu ve israf/kıtlık minimizasyonudur. Daha az sıklıkta karşılaşılan kriterler ise düzenli sipariş sayısı ve ABO- ikameli ünite sayısıdır. Toplam maliyetler genellikle şu terimleri dikkate almaktadır: envanter tutma maliyeti, sipariş maliyeti, israf maliyeti, kıtlık maliyeti. Eritrosit envanter yönetimi çalışmalarında tekli veya çoklu amaçlar göz önünde bulundurulmuştur. Kesin çözüm yöntemlerinin kullanıldığı makaleler genellikle tek bir amaç ele almıştır. Simülasyon modelleri kullanan çalışmaların neredeyse tamamı ise bir veya daha fazla değerlendirme kriterini, performans göstergelerini ele almıştır.

Deterministik bir karar verme ortamı için literatürde doğrusal programlama (Linear Programming - LP) modelleri ve tamsayı karışık doğrusal programlama (Mixed Integer Linear Programming - MILP) modelleri bulunmaktadır. Tamsayı karışık doğrusal olmayan programlama (Mixed Integer Nonlinear Programming - MINLP) modellerini dikkate alan çalışmalar bunları MILP modellerine dönüştürmek için doğrusallaştırma tekniklerinden yararlanmaktadırlar. Ayrıca, literatürdeki bir çalışmada [46] tamsayı doğrusal programlama (Integer Linear Programming – ILP), ve bir çalışmada da [52] kısıtlandırılmış karışık tamsayı programlama (Constrained Mixed Integer Programming – CMIP) modelleri kullanılmıştır.

Stokastik karar verme ortamında literatürde sıklıkla uygulanan bazı optimizasyon teknikleri bulunmaktadır. Stokastik programlama, en az bir parametrenin rastgele değişken olarak modellendiği problemler için en uygun kararı bulmak için kullanılır. Stokastik problemleri modellemenin bir yolu iki aşamalı stokastik programlamadır. İki aşamalı stokastik programlama problemlerinde veriler rastgele olarak kabul edilir. Birinci aşama kararları, rastgele veriler hakkında tam bilgi olmadan alınır. Devamında tam bilgi alınır ve ikinci aşama kararları veya düzeltici aksiyonlar alınır. Daha sonra iki aşamalı problemin doğrusal programlama eşdeğeri, ilgili olasılıklara sahip rastgele veriler üzerinde sonlu sayıda senaryo göz önüne alınarak geliştirilir [73].

Matematiksel modelleme geliştiren çalışmalar, çoğu durumda, bir çözüm elde etmek için CPLEX veya Gurobi gibi yazılımları kullanılmıştır. Kesin prosedürler, yalnızca küçük boyutlu örnekler için modelleri makul bir hesaplama süresi içinde çözebilir.

Bu nedenle sezgisel yöntemler, daha büyük örneklerde yüksek kaliteli çözümler bulmak için daha uygun alternatiflerdir. Metasezgisel yöntemler ise farklı stratejiler kullanarak arama uzaylarını keşfetmeye yönelik üst düzey kavramlardır. Belirsizlik altında optimizasyon problemlerini çözmek için genellikle başarılı alternatiflerdir. İncelenen makaleler arasında yalnızca Yadav [50] bir metasezgisel yöntem olan genetik algoritmayı kullanmıştır.

Yukarıda bahsedilen kesin çözüm yöntemleri ve sezgisel yaklaşımlara ek olarak, belirsizliği yönetmek için simülasyon da literatürde yaygın olarak kullanılmaktadır. Eritrosit envanter yönetimi literatüründe en yaygın olarak kullanılan simülasyon metodu kesikli olay simülasyonudur (Discrete Event Simulation - DES). Literatürde; sistem dinamikleri simülasyonu (System Dynamics Simulation - SDS) (Clay [53]), Monte Carlo simülasyonu (Monte Carlo Simulation - MCS) (Simonetti [66]), yeniden kullanılabilir simülasyon (Reusable Simulation - RS) (Ahmadimanesh [49]), iz-sürümlü simülasyon (Trace-Driven Simulation - TDS)(Sarhangian [58]) gibi yaklaşımlar da kullanılmıştır. Simülasyon modelleri, optimal bir çözüm bulma garantisine olmamasına rağmen, karmaşık stokastik sistemleri temsil etmek için idealdir. Simülasyon ve optimizasyon kombinasyonu, eritrosit envanter yönetiminde karmaşık kararları ele almanın bir yolunu sunmaktadır. Örneğin, Oluşma ve diğer [61] bir parçacık sürü optimizasyonu (Particle Swarm Optimization - PSO) tekniği içinde DES kullanmış, Duan ve Liao [62] ise eşik kabulü (Threshold Accepting - TA) ve tabu araması (Tabu Search - TS) gibi metasezgiselleri içeren bir simülasyon-optimizasyon çerçevesi önermiştir.

Çizelge 2.2'de de görüldüğü üzere çeşitli ağ topolojileri ve kan ürünleri ele alınırken birçok matematiksel model ve simülasyon yöntemi kullanılmış olsa dahi çalışmaların hiçbirinde pekiştirmeli öğrenme yöntemine başvurulmamıştır.

Bu araştırmanın amacı, bir hastane kan bankası envanter sistemi için pekiştirmeli öğrenme yöntemi geliştirmektir. Mevcut literatürde, pekiştirmeli öğrenme, bozulabilir envanter yönetim sistemlerine çeşitli şekillerde uygulanmıştır ancak henüz bir kan ürünleri envanter yönetim sistemine uygulanmamıştır. Dolayısıyla bu tezin mevcut literatüre katkısı şu şekilde olacaktır: derin pekiştirmeli öğrenme yönteminin bir hastane kan bankası envanter sistemine uygulanması.

BÖLÜM 3

3. DENEYSEL KISIM

Bu bölümde, ele alınan hastanenin kan bankası ayrıntılı olarak ele alınmıştır. Önce, Bölüm 3.1’de bu vaka tekrardan tanıtılmıştır ve sonrasında da bu problem için pekiştirmeli öğrenme sisteminin farklı elementleri sıra sıra tanımlanmıştır. Bölüm 3.2 ve Bölüm 3.3’de sırasıyla durum değişkeni ve eylem değişkeni tanıtılmıştır. Sonrasında, Bölüm 3.4’de zaman ufkunun nasıl planlandığı açıklanmıştır. Takip eden Bölüm 3.5’de ödül fonksiyonu ele alınmıştır ve Bölüm 3.6’de ilgili değer fonksiyonuna değinilmiştir. Derin pekiştirmeli öğrenme metodunun çıktılarının doğru değerlendirilebilmesi adına Bölüm 3.7’de kıyaslama (benchmark) tartışılmıştır ve son olarak Bölüm 3.8’de geriye kalan uygulama detayları paylaşılmıştır.

3.1. Vaka Tanımı

Bölüm 1’de, hastanenin gerçek yaşam örneğini tanıtılmıştır. Sistemle ilgili tüm bilgiler toplanmış ve bu bölümde sistem yakından analiz edilmiştir. Sistemin tüm elementleri ve özellikleri yakından incelenerek hangilerinin derin pekiştirmeli öğrenme metodunda yer alıp, hangilerinin yer almayacağı belirlenmiştir.

3.1.1. Ürünler

Hastanenin transfüzyon merkezi eritrosit süspansiyonları ile birlikte plazma, trombosit gibi farklı kan ürünleri, ve hatta bu kan ürünlerinin de farklı tıbbi ihtiyaçlara yönelik olarak çeşitli şekillerde işlenmiş olarak sunulan muhtelif türlerini sunmaktadır.

İhtiyaç duyduğu ürünleri Kızılay'dan talep etmektedir ve gelen ürünlerin hastalara atanmasını sağlamaktadır. Bu çalışma dahilinde yalnızca eritrosit envanter yönetimi ile ilgilenilmiştir. Kan grubuna bağlı olarak farklılık gösteren 8 farklı eritrosit ürünü vardır (0+, 0-, A+, A-, ...) ve her biri incelenmiştir fakat, modelde bir seferde yalnızca bir ürünün ele alınmasına karar verilmiştir. Simülasyon birden fazla ürünü içerecek olsaydı, sipariş politikası her ürün için belirlendiğinden, bu durum eylem sayısını büyük ölçüde etkilerdi. Dolayısıyla bu, derin pekiştirmeli öğrenme yönteminin çalışma süresini de etkilerdi. Birden çok ürün için sipariş politikasını belirlemek için yöntem birden çok kez çalıştırılabilir.

Gerçek hayatta bir kan grubundaki ürün bulunmadığı zaman, kan grubuna bağlı olarak, bir veya birçok kan grubu ile ikame (substitution) söz konusudur. Örneğin, AB Rh Pozitif kan grubuna sahip olan bir hastaya diğer tüm kan gruplarından eritrosit transfüzyonu mümkündür. Gerçek hayatta olduğu gibi, kan grupları arası ikame ihtimali incelemeye dahil edilecek olsaydı 8 ürünün de aynı anda ele alınması bir zorunluluk olurdu fakat ürünler arası ikame kapsam dışında tutulmuştur. Gerçek uygulamalarda da kan grupları arası ikame pek tercih edilen bir uygulama değildir, aynı grupta kan transfüzyonuna önem verilmektedir ve o anki envanterle karşılanamayan siparişler genellikle Kızılay'a bildirilecek olan acil siparişlerle karşılanmaktadır. Ayrıca, özellikle kritik durumlarda kullanılmak üzere envantere her zaman bir miktar genel verici 0 Rh Negatif grubu kan tutulmaktadır. Kurulan modelde, her ürün ayrı ayrı ele alınacak olduğundan dolayı, düşük envanter seviyesinden dolayı o günün talebi karşılanamazsa, karşılanamayan ürünler kayıp satış olarak kaydedilecektir.

Bunlarla birlikte, envanterdeki ürünlerin beklemede olan, henüz atanmamış ünite ve atanmış ünite olarak 2 ayrı şekilde ele alınması da mümkündür. Hekimler tarafından herhangi bir talep alındığında belirli bir sayıda ünite rezerve edilmesi ve bu ünitelerin transfüzyon anına kadar farklı işlemlerde kullanılmadan bekletilmesi gerekmektedir. Bu ürünler atanmış ürünlerdir fakat her durumda rezerve edilen ürünlerin tamamı hekim tarafından kullanılmamaktadır. Örneğin bir cerrah talep etmiş olduğu 6 üniteden, ihtiyaca bağlı olarak, yalnızca 4 tanesini kullanabilir operasyon esnasında, dolayısıyla kalan 2 ürüne ihtiyacı kalmayacaktır. Bu gibi senaryolarda, artan ürünlerin birkaç günlüğüne, farklı talepleri karşılayamayacak şekilde "envanter dışı kalma" durumu ortaya çıkmaktadır. Bahsedilen durum çalışma kapsamı dışında tutulmuştur. Amerika, Kanada gibi ülkelerde aynı kademede bulunan birimlerin, kan bankalarının kendi aralarında envanter

alışverişi de söz konusudur: yanal sevkiyat. Kan bankaları bilgi alışverişinde bulunarak ihtiyaç duydukları ürünleri birbirlerinden temin edebilmektedirler. Örneğin, ömrü dolmakta olan eski üniteler, tüketimin daha yoğun olduğu büyük hastanelere sevk edilebilir. Türkiye’de ise kan bankaları birbirlerinin envanter bilgilerini görememektedirler ve yanal sevkiyat gerçekleştirememektedirler. Dolayısıyla, bu araştırmada yanal sevkiyat dikkate alınmamıştır.

Bir diğer ünite transferi ise raf ömrü dolmakta olan ürünler ile ilgilidir: geri dönüşüm. Kan bankaları, ihtiyaç duymadıkları veya raf ömrü yaklaşmakta olan ve zamanında kullanamayacaklarını düşündükleri ürünleri Kızılay’a iade edebilirler. Sonrasında ise Kızılay bu ürünleri tekrardan farklı kan bankalarına dağıtmaktadır. Bu durum özellikle nadir kan gruplarında gözlemlenmektedir. Bu çalışmada geri dönüşüme izin verilmemektedir, kullanılmayan ürünler israf olarak kaydedilmektedir.

3.1.2. Talep

Hastanenin Ağustos 2016 – Ağustos 2021 tarihleri arasındaki 5 yıllık tüketim verisi hastanenin Bilgi Teknolojileri departmanından temin edilmiştir. Hastane tarafından 2 veri seti paylaşılmıştır. İlk veri seti, her bir kan grubu için ayrı ayrı, hastanenin Kızılay’dan talep etmiş olduğu toplam ünite sayısını içermektedir. İkinci veri seti ise, Kızılay’ın bu siparişlere karşılık olarak her bir kan grubu için toplamda hastaneye arz etmiş olduğu ünite sayısını içermektedir.

Çizelge 3.1’de görülebileceği üzere hastane tarafından talep edilen tarafından rakamlar ile Kızılay tarafından temin edilen rakamlar arasında önemli bir fark bulunmaktadır. Bunun sebebi kan bankaları tarafından bir sipariş alındığında Kızılay’ın bu talebi karşılayamıyor olması değildir. Talep edilen rakamlar ile karşılanan rakamlar arasında ciddi farklar açılıyor olmasının sebebi kan bankalarının arz tarafında belirsizlik hissettikleri durumlarda siparişlerinin karşılanamayabileceğini düşünerek gerçekte ihtiyaç duydukları rakamlardan fazlasını Kızılay’a yansıtılmalarından kaynaklanmaktadır. Buradaki mantık, tek yetkili tedarikçi Kızılay olduğu için buradan gerekenden fazla sayıda ünite talep ederek, siparişin kısmen karşılanması durumunda dahi gerçekte ihtiyaç duyulan rakamları yakalamaya çalışmaktır. Bu uygulamaya literatürde de değinilmiştir [74]. İncelenen verilerin Covid-19 pandemisi dönemini içermesi ve bu sebeple kan tedarik zincirinin de bu sebeple daha önceden tecrübe edilmemiş olan belirsizliklerle karşı karşıya kaldığı da

hesaba katılırsa kan bankalarının gerekenden çok daha fazla ünite sipariş vermiş olması daha anlaşılabilir hale gelecektir. Aksi takdirde, hastane ihtiyaç duyduğu tüm ürünleri istediği miktarda, acil siparişler dahil, Kızılay'dan temin edebilmektedir.

Çizelge 3.1: Hastanenin Kızılay'dan sipariş ettiği ve teslim aldığı ünite sayıları.

Kan grubu	Sipariş edilen ünite sayısı	Karşılanan ünite sayısı
0 +	82026	46013
0 -	10927	5878
A +	99808	55654
A -	14860	8722
B +	38571	20957
B -	6966	3876
AB +	17401	9356
AB -	2312	1273
Toplam	272871	151729

Bu çalışmada da sipariş edilen miktarın tamamının deterministik bir şekilde karşılandığı varsayılacaktır. Ayrıca, hastanedeki israf oranı ihmal edilebilecek seviyelerdedir ve hasta talepleri her zaman karşılanabilmektedir. Dolayısıyla, karşılanan ürün değerlerini içeren ikinci veri seti, hem hastanenin gerçek ihtiyacını, hem de aynı zamanda hastaların hastaneden gerçekte talep ettiği miktarı yansıtmaktadır. Dolayısıyla, her bir kan grubu için 5 yıl içerisinde karşılanan toplam eritrosit sayısı kullanılarak hastaneye yansiyacak olan günlük talep hesaplanabilir.

Hastane, siparişlerini günlük olarak Kızılay'a bildirmektedir ve ele alınan sistem de siparişler günlük olarak karar verilecek şekilde kurgulanmıştır. Hastaneden gün bazında veri temin edilememiş, yalnızca 5 yıl sonundaki kümülatif toplam rakamlar öğrenilebilmiştir. Dolayısıyla, 5 yıllık toplam ünite sayılarının makul yöntemlerle günlük talep yoğunluklarına dönüştürülmesi ihtiyacı bulunmaktadır. Bu dönüşümler planlanırken trend, sezonsallık ve günden güne yaşanan talep farklılığına dikkat edilmiştir.

Öncelikle, her ne kadar kan ürünleri talebinin zaman içerisinde değişkenlik gösterdiği bilinse de, sistemin performansı 1 yıllık simülasyon üzerinden değerlendirileceği için trend faktörü modelde göz ardı edilmiştir. Yıl içerisinde aydan aya, mevsimden mevsime kan ürünleri talebinde sezonsallık olduğu hem hastanedeki uzmanlar tarafından ifade edilmiştir, hem de literatürde yer almaktadır [75] Sezonsallığın bu çalışmada ele alınmamasının iki temel sebebi sahip olunan veri setinin bu bilgiyi içermemesi ve bununla

birlikte literatürde de modeli kurarken örnek olarak ele alınabilecek, aynı zamanda okuyucularla da paylaşılmış olan, bir veri seti ile karşılaşılammış olmaktan kaynaklanmaktadır. Literatürde, örnek olarak incelenebilme potansiyeline sahip olan yalnızca bir makale ile karşılaşmıştır [76]. Bu makalede ise, her ne kadar 5 yıllık bir eritrosit talep veri seti okuyucuyla paylaşılmış olsa dahi, veri incelendiğinde sezonsallık tespit edilememiştir. Dolayısıyla bu çalışmada örnek olarak ele alınabilecek bir yayın bulunamamıştır ve sezonsallık dikkate alınmamıştır.

Sahip olunan verinin kullanıma uygun bir günlük talep yoğunluğuna dönüştürülmesi sürecinde son olarak taleplerde günden güne yaşanan farklılıklar incelenmiştir. Kan ürünleri kullanımının haftanın günlerine bağlı olarak değişim gösterdiği bilinmektedir [62]. Hafta sonları talep genellikle minimum seviyelere düşerken, Pazartesi ve Cuma günü gibi cerrahi işlemler için ağırlıklı olarak tercih edilebilen günlerde tüketim seviyeleri en üst düzeylere çıkabilmektedir. Örneğin hastanede eritrosit tüketimi Salı günleri en üst seviyelere ulaşmaktadır. Sahip olunan veri seti 5 yıllık toplam tüketimi temsil ettiğinden dolayı hastanenin günden güne değişen tüketim yoğunlukları bilinmemektedir. Dolayısıyla, kan bankalarının gerçek taleplerini daha gerçekçi bir şekilde temsil edilebilecek bir model kurabilmek adına, sezonsallık faktörü ile ilgili denendiği şekilde, literatürden bir örnek bulunarak bu örneğin bu çalışmaya uyarlanması denenmiştir.

Tüm kan gruplarını, haftanın tüm günleri için ayrı ayrı ele alan örnek bir talep yoğunlu paylaşan Duan'daki [62] yoğunluklar kurulan sistemde denenmiştir. Derin pekiştirmeli öğrenme modeli, her gün için, o güne özel bir lamda (λ) değerine sahip Poisson dağılımıyla talep oluşacak şekilde, ve durum vektörüne de, ajana o günün hangi gün olduğu bilgisini aktaracak ek bir değişken eklenecek şekilde düzenlenmiştir. Ortaya çıkan modelin ise, günden güne değişen talep dağılımlarının getirdiği ek karmaşıklık dolayısıyla, bir türlü ne şekilde aksiyon alması gerektiğine dair fikir edinemediği, kullanışlı herhangi bir politika geliştiremediği gözlemlenmiştir. Günden güne talep yoğunluğunda farklılığı dikkate alan bir model geliştirilemediği için bu durum kapsam dışında tutulmuştur.

Her gün aynı yoğunlukta talep dikkate alan bir sistem kurabilmek için hastaneden temin edilen karşılanan ürün verisi transfüzyon merkezinin haftanın her günü operasyon halinde olduğu dikkate alınarak günlük ortalama talep yoğunluklarına dönüştürülmüştür. Her bir kan grubu için günlük talep yoğunlukları **Çizelge 3.2**'den görülebilir. Duan'da

da [62] belirtildiği üzere hastane kan bankalarının karşılaştıkları talepler, her bir kan grubu için ayrı yoğunlukta olacak şekilde, Poisson varsayılabilir. Dolayısıyla, nihai modelde, her bir kan grubunun günlük talebi, **Çizelge 3.2**'de ilgili yoğunlukta lamdaya sahip Poisson varsayılmıştır.

Çizelge 3.2: Günlük Poisson ortalama talep yoğunlukları.

Kan grubu	Günlük talep yoğunluğu (λ)
0 +	25
0 -	3
A +	30
A -	5
B +	11
B -	2
AB +	5
AB -	1

3.1.3. Amaç

Bölüm 1'de hastanenin amacının kıtlık ve israfı eş zamanlı olarak minimize etmek olduğundan bahsetmiştik. Tıbbi işlemlerde ihtiyaç duyulduğundan ve yerini tutacak herhangi bir alternatif henüz geliştirilemediğinden dolayı kan ürünleri envanterinin yeterli seviyede tutulması kritiktir. Ayrıca, ünitelerin israfı da, hem maliyetli olduğundan hem de kan transfüzyon hizmetlerinde olumsuz etkiye yol açabileceğinden dolayı [77] istenmeyen bir durumdur. Yalnızca insan vücudundan gönüllü bağışlar ile elde edilebilen üretimi kısıtlı ve değerli kan ürünleri, hem kıtlık hem de ürün israfı durumunda parasal açıdan ölçümü son derece zor olan maliyetlere yol açtığı için, ciddi vakalarda can kaybı gibi, bu çalışmada ekonomik maliyetler dikkate alınmamış, yalnızca hizmet seviyesine ve israf miktarına odaklanılmıştır. Literatürde, örneğin, izin verilen maksimum kıtlık miktarını belirten bir hizmet seviyesi kısıtlamasının, kıtlık maliyetlerinin bir temsilcisi olarak varsayılması gibi uygulamalar yaygındır. Can kaybının genel olarak kabul edilemez olması, tedarik zincirinde tüm maliyetlerin üzerinde güvenliği ön planda tutmaktadır. Çoğu makale, gerçek uygulamalarda, takip edilmesi beklenen politika sonuçlarının yumuşatılabileceği konusunda uyarıda bulunarak bunu kabul etmektedir [78].

3.1.4. Sistem

Hastane, kan ürünlerini Kızılay'dan tedarik edip, hekim taleplerine göre hastalara atamaktadır. Amaç, kıtlık ve israfı minimize etmektir. Sistem, incelenmekte olan kan grubuna uygun bir şekilde parametreler güncellendikten sonra, bir seferde yalnızca bir kan grubu için çalıştırılmaktadır.

Karar değişkeni, her bir zaman adımında Kızılay'a iletilmek üzere belirlenmesi gereken sipariş miktarıdır. Modelde, iki ortamsal belirsizlik bulunmaktadır. Birincisi hekim talepleri ile ilgilidir. Hekim talepleri hastaneye rassallık içerecek şekilde Poisson dağılımıyla yansımaktadır. Kızılay tarafında sipariş karşılama performansında bir belirsizlik bulunmamaktadır, sipariş edilen tüm ürünler Kızılay tarafından karşılanabilmektedir. Teslim süresinde de bir belirsizlik bulunmamaktadır, siparişler bir adım sonraki talepten önce teslim alınmaktadır.

İkinci belirsizlik ise hastaneye gelen ünitelerin kalan raf ömürleri ile ilgilidir. Kızılay'dan gelen üniteler belirli bir yaşta değil de rassallık içerecek şekilde gelmektedir. Gelen ünite 30 günlük de olabilir, 12 günlük raf ömrü kalmış olur; 3 günlük de olabilir, 39 günlük raf ömrü kalmış olur. Hastane gelen ürünlerin ortalama 15 günlük üniteler olduğu transfüzyon merkezi sorumlusu tarafından ifade edilmiştir. Dolayısıyla, kolaylık amacıyla, gelen ünitelerin $\lambda=15$ Poisson yaşında teslim alındıkları varsayılmaktadır.

Herhangi bir zaman adımında, ortaya çıkan talebin envantere bulunan ürün sayısından büyük olması durumunda, envantere bulunan ünite adedince talep karşılanacak, talebin kalan kısmı miktarınca karşılanamayan sipariş kaydedilecektir. Gerçek hayattan farklı olarak, bu sistemde ek bir kaynak olarak acil siparişlere izin verilmemektedir. Takip eden adımda karşılaşılabilecek talep bilgisi ajan tarafından sipariş miktarı belirlenmeden önce bilinmemektedir. Günlük tüketimden sonra, her zaman adımının sonunda, bir sonraki adımın talep bilgisi edinilmeden önce Kızılay'a sipariş miktarı bildirilmektedir. Talep bilgisi ancak sipariş miktarı ajan tarafından belirlendikten sonra meydana çıkmaktadır.

Her zaman adımında, simülasyonun başında envanter başlangıç seviyesine getirildikten sonra, 5 işlem gerçekleşir: sipariş miktarı belirlenir, belirlenen miktarda ünite envantere eklenir, ortaya çıkacak talep üretilir, hekim siparişleri envanterden karşılanır, envanter seviyesi güncellenir.

Envanter başlangıç seviyesi minimum olacak şekilde, 0 ürün mevcut halde, simülasyon başlatılır. Ajan, geçmişte deneme yanılma yoluyla edindiği tecrübelerle dayanarak ve yalnızca mevcut envanter durumundan yola çıkarak sipariş miktarını belirlemektedir ve bu siparişe gelen üniteler bir önceki günden kalan envantere eklenir (ilk zaman adımında kalan envanter 0'a eşit olmaktadır). Eritrosit raf ömrü 42 gündür ve üniteler, toplandıkları andan itibaren en erken 2 gün sonra kullanıma hazır hale gelmektedir (taşıma, test edilme, fraksinasyon, dağıtım gibi işlemlerden dolayı). Dolayısıyla, ünitelerin 40 zaman adımından sonra envanterden atılması sağlanarak bu durum dikkate alınmıştır. Ayrıca, yukarıda da bahsedildiği üzere, gelen her üniteye ayrı ayrı $\lambda=15$ Poisson dağılımından bir yaş atanarak ürünler envantere eklenmektedirler.

Daha sonra, ortaya çıkacak olan günlük talebin Poisson dağılımına göre belirlenmesi gerekmektedir. Bu adımda her kan grubu için ayrı bir talep yoğunluğu kullanılmalıdır. Örneğin, **Çizelge 3.2**'den görülebileceği üzere, A Rh Pozitif kan grubunun günlük talep miktarı $\lambda=30$ Poisson dağılımından üretilmelidir. Bir sonraki adımda hekim siparişleri karşılanmaktadır. Örnek vermek gerekirse: talep envanter seviyesinden az ise, envanterden talep miktarınca ünite eksiltilir ve kalan envanter bir sonraki güne devreder. Aksi durumda, talebin envanter seviyesinden fazla olması durumunda ise, envanterin tamamı talebin kısmi olarak karşılanmasında kullanılır, karşılanamayan miktarda ürün karşılanamayan sipariş olarak kaydedilir ve ertesi güne boş bir envanter devretmiş olur. Ek olarak, hekim siparişlerinin karşılanması esnasında sırasıyla en eski ünitelerden en yeni ünitelere doğru bir öncelik tanımlandığı da belirtilmelidir, önce raf ömrü dolmaya yaklaşan üniteler kullanılmaktadır her zaman.

Bir sonraki zaman adımına geçiş yapılmadan önce son olarak envanter durumunun güncellenmesi gerekmektedir. Önce, kalan tüm ürünler bir gün yaşlandırılır ve sonrasında da, eğer varsa, raf ömrü dolan ünite sayısına israf kaydedilir ve bu ürünler envanterden çıkartılır. Böylelikle, ajanın bir sonraki zaman adımında sipariş miktarını belirlerken dikkate alacağı envanter bilgisi güncel halini almış olur ve döngü yeniden başlamış olur. Bu adımlar, aynı sırayla, planlama ufku olan 360 gün boyunca (simülasyon koşum tekrarında sağladığı kolaylıktan dolayı 1 yıl 360 gün olarak ele alınmıştır) tekrar etmektedir.

3.2. Durum Değişkeni

Bölüm 2.1’de belirtildiği gibi, eğer bir sistem durumu Markov özelliğine sahipse, durum yalnızca ondan önce gelen duruma ve seçilen eyleme bağlıdır. Bu şekilde, mevcut durum ve eylem göz önüne alındığında bir sonraki durumu ve beklenen ödülü tahmin etmek mümkündür. Durum sinyali kesinlikle Markov olmasa bile, ek bir durum değişkeni ekleyerek ona istenilen doğrulukta yaklaşabiliriz. Bu şekilde problemi pekiştirmeli öğrenme problemi olarak formüle etmek mümkündür. Pekiştirmeli öğrenme modellerinde son karar sistem durumuna göre verildiğinden, sistem durumunun ajanların karar verme süreci için uygun bilgileri sağlaması önem arz etmektedir. Sağlıklı bir politika geliştirme süreci için sinir ağına gerekli tüm bilgilerin, ve yalnızca gerekli bilgilerin iletilmesi gerekmektedir.

Sonrasında, bu bilgiler ağı giriş katmanıyla paylaşılmaktadır. Durum vektörümüzdeki her değişken için sinir ağıımızda ayrı bir girdi düğümü oluşturulmaktadır. Birçok denemenin ardından durum vektörü sadece iki değişkeni dikkate alacak şekilde sadeleştirilebilmiştir: envanterdeki ünite sayısı ve ünitelerin medyan yaşı. Bu iki değişken arasından, envanterdeki ünite sayısını temsil eden değişken 0 ile 1 aralığına yakın değerlere normalize edilmiştir. Alt sınır 0’ın altına inilemez fakat üst sınır 1’in üstünde kalınabilir. Envanterde birikebilecek ürün sayısında bir kısıtlamaya sahip olunmadığından dolayı normalizasyon üst sınırı serbest bırakılmıştır. Ünitelerin medyan yaşını temsil eden değişken ise $[0, 1]$ aralığına normalize edilmiştir çünkü ünite yaşları sınırlı bir aralıkta, 0 ile 40 arasında, seyretmek zorundadır.

Durum değişkenlerinin değerleri, değerleri birbirine yakın tutmak için normalleştirilmiştir çünkü sinir ağıları küçük değerlerle başlatılır ve küçük gradyan adımlarıyla güncellenir. Eğer girdi değerleri sinir ağına yansıtılmadan önce normalleştirilmemişse, sinir ağının ağırlıklarını doğru değerlere getirmek uzun zaman almaktadır. Gözlemlerde yöntemin sınırlandırılmadığından emin olunabilmesi adına durum değişkenlerinin normalleştirilmeden önceki değerleri sınırlandırılmamıştır.

3.3. Eylem Değişkeni

Pekiştirmeli öğrenme problemleri, ayrık veya sürekli bir eylem alanına sahip olabilir. Her çıktı düğümü için, aktörün sinir ağı ya bu eylemin gerçekleştirilme şansını (ayrık

eylemler) ya da eylem-değerini (sürekli eylemler) çıktı verir. Ayrık bir eylem uzayında, ajan, sonlu bir eylem kümesi içerisinde hangi eylemi seçmesi gerektiğine karar verir. Bu eylem setindeki her farklı eylem için sinir ağının ayrı bir çıktı düğümü vardır. Sürekli bir eylem uzayında, sinir ağının çıktı düğümlerinin sayısı, eylem kümesinin uzunluğuna eşittir. Sinir ağının çıktısı, her düğüm için bir ortalama ve standart sapmadan oluşmaktadır. PPO algoritması hem ayrık, hem de sürekli eylem uzaylarıyla çalışabilmektedir.

Sipariş edilen ünite sayısı tamsayı olmak zorunda olduğu için ayrık eylem alanı tercih edilmiştir fakat bu modelde günlük izin verilen sipariş miktarı aralığı ne çok dar ne de çok geniş olduğu için yukarıda bahsedilenden farklı bir çıktı düğümü tercih edilmiştir. En yoğun talebe sahip olan kan grubu A Rh Pozitifdir ve bu kan grubu için günlük talep $\lambda=30$ Poisson dağılımından üretilmektedir. Çeşitli denemeler sonucu ajana en fazla 30'un iki katı, yani 60 ünite, sipariş verme imkanı tanınmanın uygun olduğu gözlemlenmiştir. Bu şekilde, hem ajanın hareketlerinin kısıtlanmadığı, hala optimal politikalara yakınsanabilmektedir çünkü, hem de öğrenme sürecinin, ihtiyaç duyulmayan eylem seçeneklerinin kaldırılmış olması sebebiyle, daha hızlı gerçekleştiği gözlemlenmiştir. Dolayısıyla, eylem uzayının en geniş olduğu A Rh Pozitif modelinde bile 0 ile 60 arasında bir seçimde bulunmak yeterli olacaktır. Çıktı düğümü bu sebeple tanımlanmış olan sipariş aralığı ile aynı uzunlukta bir vektör çıktı vermektedir. Vektördeki her değer $[0, 1]$ aralığındadır ve listedeki değerlerin toplamı 1'dir. Sonrasında bu vektör bir kategorik dağılıma dönüştürülmektedir. Böylelikle, bir eylem seçilmesi gerektiğinde her sipariş miktarı farklı olasılıklarla seçilme ihtimaline sahip olacaktır. Sinir ağı, durumu gözlemleyerek her seferinde farklı bir vektör oluşturmaktadır.

Sipariş esnasında, envantere bulunan ünite sayısı ve ünitelerin medyan yaşı bilinmemektedir fakat talep bilinmemektedir. Derin pekiştirmeli öğrenme yöntemi her zaman adımında sipariş miktarını belirlemelidir. Belirlenmesi gereken tek bir değer vardır, sipariş miktarı, dolayısıyla sipariş miktarı eylem vektörünün tek değişkenidir. Kan grupları farklı talep yoğunluklarına sahip olduklarından dolayı her kan grubu için çalıştırılacak olan modelin eylem uzayı talep yoğunluğu dikkate alınarak belirlenmelidir. Eylem uzayını fazla geniş tutmak öğrenme sorunlarına, fazla dar tutmak ise optimalite problemlerine yol açacaktır. **Çizelge 3.3**'de her kan grubunun günlük talep yoğunluğu ve karşılığında modelde kullanılan eylem uzayı paylaşılmıştır.

Çizelge 3.3: Farklı kan grupları için kullanılan eylem uzayları.

Kan grubu	Günlük talep yoğunluğu (λ)	Eylem uzayı
0 +	25	[0, 50]
0 -	3	[0, 9]
A +	30	[0, 60]
A -	5	[0, 15]
B +	11	[0, 27]
B -	2	[0, 7]
AB +	5	[0, 15]
AB -	1	[0, 5]

3.4. Zaman Ufku

Derin pekiştirmeli öğrenmenin önemli bir yönü de zaman ufkuudur. Sistemin performansı 1 yıl üzerinden değerlendirilecektir ancak her koşul, gerçek durumu daha iyi temsil edebilmek adına 10 yıl olarak planlanmıştır. Böyle bir strateji takip edilmesinin ana sebebi, zaman ufkunun sonlarına doğru gerçekleştirilen eylemlerin etkilerinin sisteme yeterince yansımamasından kaynaklanmaktadır çünkü bu eylemlerin etkileri ufku bitiminden sonra sisteme yansiyabilmektedir. Bu durumda, ajan her ne kadar doğru eylem tercihlerinde bulunduğunu düşünse de, bu eylemlerin olumsuz etkileri sonradan sisteme yansiyacaktır.

Herhangi bir adımda alınan bir kararın etkileri onlarca adım sonra kendisini gösterebilmektedir ve simülasyon 10 yıllık çalıştırılarak bu etki azaltılmıştır. Zaman ufkunun 1 yıl olarak planlandığı uygulamaların performanslarının, zaman ufkunun 10 yıl olarak planlandığı çalışmalardan çok daha düşük olduğu da ayrıca farklı denemeler esnasında gözlemlenmiştir. Zaman ufku planlanırken PPO'da iki önemli parametre vardır: arabellek boyutu ve parti boyutu. Arabellek boyutu, bir iterasyonun uzunluğudur ve genellikle ufuk olarak da belirtilir. Parti boyutu, modelin ağırlıklarını güncellemek için arabelleği bölmek için kullanılan boyuttur. Ayrıca, yakın eylemlerin uzun vadedeki eylemlere kıyasla daha ağırlıklı olmalarını sağlayan bir indirgeme faktörü de sıklıkla kullanılmaktadır. Bu sonlu ufukta, arabelleğin uzunluğu 3600'e, parti boyutu 300'e ve indirim faktörü 0.99'a ayarlanmıştır.

3.5. Ödül Fonksiyonu

PPO, ödülleri maksimize etmeye çalışır. Bölüm 2.2.2’de belirtildiği gibi, PPO algoritması bir Aktör-Kritik mimarisi kullanır. Aktör, eylemlerin değerlerini çıkarırken, Kritik değer fonksiyonunu tahmin eder. Bu değer fonksiyonunu tahmin etmek için sinir ağı, gözlemlenen durum-eylem çiftleri ve gelecekteki değerler üzerinde bir regresyon gerçekleştirir. Ağ ağırlıkları daha sonra ağ çıktıları ve eğitim veri çıktıları arasında minimum kayıp elde etmek için optimize edilir. Ancak, bu gelecekteki değerler, özellikle ağ henüz eğitilmemişken, yani başlangıçta çok büyük olabilir. Ağ çıktılarının ölçeği, girdi özelliklerinden ağırlık ve eşik değerinden önemli ölçüde farklıysa, sinir ağı, doğru öğrenemeyen bir ağ ile sonuçlanabilecek dengesiz ağırlık ve eşik değerleri dağılımlarını öğrenmeye zorlanacaktır. Bu durumla mücadele etmek için çıktı değerlerinin ölçeklendirilmesi önerilmektedir.

Amaç kıtlık ve israfı minimize etmek olduğu için ödül fonksiyonu doğrudan bu iki değer üzerinden tanımlanmıştır. Günlük talebin eksiksiz karşılanabildiği ve hiçbir ünitenin israf edilmediği zaman adımlarında ödül 1, diğer tüm zaman adımlarında ödül 0 olarak tanımlanmıştır. Her zaman adımında, yalnızca günlük tüm siparişleri karşılayabildiğinde ve hiçbir ünite israf edilmediğinde ödül 1, diğer tüm durumlarda ise ödül 0 olarak tanımlanmıştır. Böylelikle, hem yukarıda bahsedilen sebeplerden dolayı değerler dar bir aralıkta tutulabilmiş, hem de yalnızca en başarılı adımlar ödüllendirilerek Kritik doğru yönlendirilebilmiştir.

3.6. Değer Fonksiyonu

Değer fonksiyonu ödül fonksiyonu ile ilişkilidir, zaman ufku boyunca beklenen maksimum toplam ödülü temsil etmek için kullanılır. Değer fonksiyonu, **Denklem 2.2**’de tanımlanan Bellman’ın değer fonksiyonuna dayanmaktadır:

$$V(s) = \max_a (r(t) + \gamma E[V(s'|s, a)]) \quad (3.1)$$

Burada $r(t)$, bir önceki başlık altında ele alınan ödül fonksiyonudur. Gama 0.99 olarak ayarlanmıştır ve değer fonksiyonu, PPO algoritması tarafından tahmin edilecektir.

3.7. Kıyaslama

Derin pekiştirmeli öğrenme yöntemimizin sonucunu karşılaştırmak için uygun bir kıyaslama (benchmark) bulunması gerekmektedir. Hastanenin gerçek hayattaki durumu potansiyel bir kıyaslama olarak düşünülebilir. Hastanede, Bölüm 3.1’de de belirtildiği üzere israf oranı ihmal edilebilecek seviyelerdedir. Tüm kan gruplarından yılda ortalama toplam 30.346 ünite eritrosit alınıp kullanılmaktadır ancak yılda toplam maksimum 1 veya 2 ünite israf edilmektedir. Hasta talepleri de her zaman karşılanabilmektedir. Derin pekiştirmeli öğrenme yönteminin hastanedeki uzmanlara destek olabilecek şekilde sipariş miktarları üretip üretemeyeceği araştırma sorusu kapsamında olduğundan dolayı bu kıyaslama, ele alınan model için makul bir kıyaslama olarak kabul edilebilir. Ayrıca, her ne kadar gerçek hayatta hastanenin, ele alınan modelden farklı olarak, oluşan talebi acil siparişler ile veya sürekli olarak envantere bulundurduğu 0 RH- üniteler ile karşılama imkanı bulunduğundan dolayı bu kıyaslama fazla zorlayıcı gözükse dahi kurulan modelde de, özellikle talep tarafında, bazı varsayımlardan yararlanılmıştır.

3.8. Derin Pekiştirmeli Öğrenmenin Uygulanması

Bölüm 3.1.4’de de belirtildiği üzere durum değişkeni $[0, 0]$ iken, yani envantere hiç ürün bulunmazken simülasyon başlatılmaktadır. Sonrasında, Bölüm 3.1.4’de tanımlanmış olan adımlar takip edilerek ve Bölüm 3.4’de ifade edildiği üzere her bir bölüm 10 yıl sürecek şekilde öğrenme gerçekleşene kadar simülasyon yüzlerce bölüm tekrar etmektedir. Simülasyonun yüzlerce bölüm tekrar etmesi zaman alan bir işlem olduğundan dolayı (her 1000 iterasyon yaklaşık 10 saat sürmektedir) her koşulda kaç bölüm tekrar edilmesi gerektiğine karar verilmelidir. Genellikle yüksek seviyelerde günlük talebe sahip olan kan grupları için 1500 iterasyonun yeterli olduğu tespit edilmiştir. Bu kan gruplarında 1500 iterasyondan sonra, halihazırda optimale yakın politikalar öğrenilmiş olduğundan dolayı, kayda değer ilerlemeler gerçekleşmediği gözlemlenmiştir.

Nispeten düşük talep yoğunluğuna sahip olan kan gruplarında ise eğitim diğer kan gruplarına kıyasla daha uzun sürdüğü için 1500 yerine 2000 iterasyon tercih edilmiştir. Her bir koşul 15 ile 20 saat arası sürdüğünden dolayı eğitimin tamamlandığı tecrübe edilene kadar her kan grubu için model önce 1500’er iterasyon ile, 1500 iterasyonda

eğitimin tamamlanamadığı kan grupları için ise model ikinci kez 2000'er iterasyon ile çalıştırılmıştır. Her kan grubu için yalnızca son koşullar dikkate alınmıştır. Başlarda öğrenmede ciddi istikrarsızlıklar gözlemlenebilse dahi, özellikle de nadir kan gruplarında, ajan her zaman ne yapması gerektiğini kavrayıp doğru stratejilere yönelebilmektedir ve aynı kan grubunun modeli istikrarlı bir şekilde aynı sonuçlara yakınsamaktadır. Başlarda gözlemlenen istikrarsızlıklar ise modelin nasıl bir başlangıç yaptığıyla doğrudan ilişkilidir, dolayısıyla bu durum normal karşılanmaktadır.

Simülasyon OpenAI Gym yapısı kullanarak kurulmuştur. Bu yapı, OpenAI ve yine OpenAI yapısını temel alan diğer paketlerin algoritmalarının kullanılmasını mümkün kılmaktadır. PPO algoritmasının uygulanabilmesi için Python Derin Öğrenme Kütüphanesi PyTorch'tan yararlanılmaktadır. PyTorch, sinir ağları oluşturmak için kullanılmaktadır. PPO algoritması uygulanırken, Schulman ve diğerlerinin [1] çalışmasından alınan hiperparametreler kullanılmıştır, öğrenme oranı hariç. Öğrenme oranı 10^{-5} 'e düşürülerek, daha uzun bir eğitim süreci ile karşılaşılsa dahi, daha optimal politikalara yakınsanabilmesi sağlanmıştır.

Bir sinir ağı uygulanırken, verilmesi gereken birkaç tasarım kararı vardır. Bölüm 2.2.1'de sinir ağları konsepti açıklanmıştır. Modelde, Aktör ve Kritik ayrı sinir ağları kullanılmaktadır, dolayısıyla toplamda 2 sinir ağı tanımlanmıştır. Her bir sinir ağının; girdi düğümü sayısı, gizli katman sayısı ve bu gizli katmanlardaki nöron sayısı, çıktı düğümü sayısı ve aktivasyon fonksiyonu tanımlanmalıdır. İki sinir ağı için de girdi düğümü sayısı, durum vektörü uzunluğuna eşittir, yani 2'dir. Kullanılan gizli katman sayısı ve bu katmanlardaki nöron sayısı da eşittir. Her iki sinir ağında da, 256 nörondan oluşan 2 gizli katman tercih edilmiştir. Kullanılan aktivasyon fonksiyonu da yine aynıdır: ReLU (Rectified Linear Unit). Tek fark çıktı düğümü sayısında gözlemlenmektedir. Aktör kritik ağındaki çıktı düğümü sayısı, izin verilen maksimum sipariş miktarına eşittir (örneğin A Rh Pozitif kan grubunun modeli için 60). Kritik sinir ağındaki çıktı düğümü sayısı ise 1'dir çünkü yalnızca durumun değerinin tahmin edilmesi gerekmektedir.

Model, makine öğrenmesi alanındaki popülerliğinden ve kullanım kolaylığından dolayı Python programlama dilinde, Jupyter Notebook'da programlanmıştır. Deneyler, Intel(R) Pentium(R) CPU G4600 @ 3.60GHz ve 8GB RAM'e sahip olan bir bilgisayarda çalıştırılmıştır.

BÖLÜM 4

4. BULGULAR VE TARTIŞMA

Bu bölümde, PPO algoritması kullanılarak kurulan derin pekiştirmeli öğrenme modelinin çıktıları ele alınmıştır. İlk olarak, kurulan modelin nasıl bir eğitim sürecine tabi tutulduğu ve tarif edilmiş olan eğitim sürecinde her kan grubunun nasıl bir performans sergilediği eğitim eğrileri üzerinden yorumlanmıştır. Bölümün devamında, bu eğitim sonrasında modelin nasıl bir politika öğrenmiş olduğu, hangi durumlarda nasıl hareket etmeyi tercih ettiği durum-sipariş kararı çiftlerine göre oluşturulmuş olan ısı haritaları aracılığıyla ele alınmıştır. Son olarak, modelin, karşılaşmış olduğu durumlarda verdiği tepkilerin mantıklı olup olmadığı da, veya ne kadar mantıklı oldukları, her kan grubunun son 50 simülasyonda sergilemiş olduğu performans incelenerek değerlendirilmiştir. Elde edilen sonuçlar araştırma soruları ve amacı çerçevesinde kıyaslama ile ilişkilendirilmiştir.

4.1. Eğitim Eğrileri

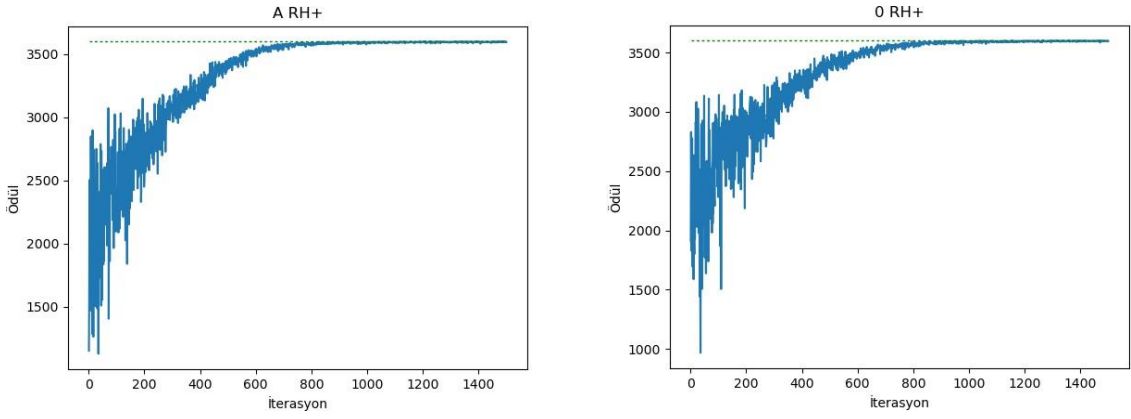
Verimli bir eğitim gerçekleşebilmesi adına ödül fonksiyonu şu şekilde tanımlanmıştır: herhangi bir ünitenin israf edilmediği ve tüm siparişlerin karşılanabildiği her gün ödül 1'dir, diğer günler ödül 0'dır. Böylelikle, ajana, iki amacın da aynı anda gerçekleştirilmesi gerektiğinin öğretilmesi hedeflenmiştir. Sipariş miktarlarının yüksek tutularak tüm siparişlerin karşılanması bir anlam ifade etmemektedir, çünkü bu durumda israf da yüksek seviyelerde seyredecektir. Bu politikanın tersini izlemek de bize istenen sonuçları sağlamayacaktır. Sipariş miktarlarını düşük tutmak her ne kadar israfı minimize edecek olsa da, bu durumda da ihtiyaç duyulan üniteler karşılanamayacaktır. Önemli olan dengeyi yakalayabilmek, envanter seviyesini düzenli olarak optimal seviyelerde tutmaya çalışarak hem israfı, hem de kıtlığı minimize edebilmektir.

Bölüm 3.4'de de belirtildiği üzere her iterasyon 3600 zaman adımı boyunca devam etmektedir. Dolayısıyla, yalnızca tüm siparişlerin karşılanabildiği ve hiçbir ünitenin israf edilmediği zaman adımlarında ödül değerinin 1 olduğu da göz önüne alındığında, ulaşılabilecek en yüksek toplam ödül her iterasyon için 3600'dür. Eğer 10 yıl boyunca hiçbir ünite israf edilmez ve tüm siparişler karşılanabilirse, yani ideal senaryoda, toplam ödül 3600 olacaktır. Aşağıdaki eğitim eğrilerinden her bir kan grubu için nasıl bir eğitim gerçekleştiği görülebilmektedir.

Örneğin A RH+ kan grubu için model ilk iterasyonlarda çok istikrarsız bir performans sergilemektedir ve zaman adımları arası toplam ödül arasındaki fark 2000'lere kadar çıkabilmektedir. Buna rağmen, ortalama toplam ödül değeri ise istikrarlı bir şekilde iyileşmektedir. Ortalama toplam ödül değerinin 2000'lerden hızlı bir şekilde 2500-3000 seviyelerine, ve sonrasında bir yavaşlama gerçekleşse dahi 3500, ve nihai olarak 3600 seviyelerine ulaştığı görülebilmektedir. Grafikteki yeşil kesikli çizgiler optimal seviyeyi temsil etmektedir, 3600.

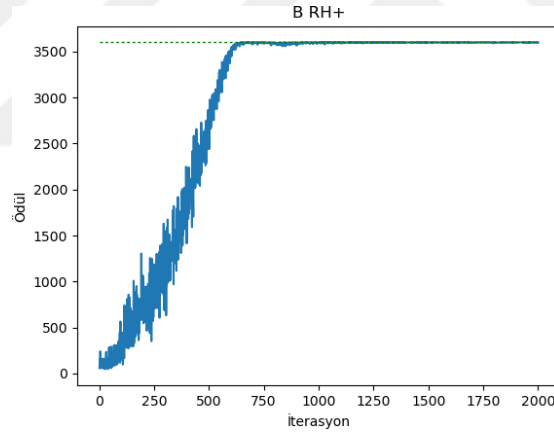
Kıyaslama olarak hastanenin gerçek hayattaki durumunun ele alındığı ve hastanenin %100 sipariş karşılama performansına ve yaklaşık %0 israf oranına sahip olduğu Bölüm 3.7'de ifade edilmiştir. Dolayısıyla, hastanenin performansı 10 yılda toplam 3600 ödülü yakalayacak şekilde, yani kesikli yeşil çizgilerin seviyesinde düşünülebilir. A RH+ için yöntem optimal seviyeleri 800 iterasyondan sonra yakalayabilmektedir ve 1000 iterasyona ulaştıktan sonra istikrarlı bir şekilde optimal seviyelerde performans gösterebilmektedir. 8 kan grubunun eğitim eğrileri: A+ ve 0+ yan yana, B+ yalnız başına, ve kalan gruplar bir arada olacak şekilde 3 grup halinde paylaşılmıştır. Buradaki amaç eğitim eğrilerinin günlük talep yoğunluğuna bağlı olarak nasıl farklılık gösterdiğini daha net bir şekilde gösterebilmektir. **Çizelge 3.2**'de de paylaşıldığı üzere A+ ve 0+ için günlük talep sırasıyla 30 ve 25 ünite seviyelerindedir, B+ için günlük talep ortalama seviyelerde, 11 ünite civarındadır ve kalan kan grupları için de günlük talep minimal seviyelerde, 1-5 ünite aralığındadır.

Şekil 4.1'den görülebileceği üzere talebin 25-30 ünite civarında olduğu durumlarda yöntem başlangıçtan itibaren hızla ortalama toplam ödülü iyileştirebilmektedir istikrarlı bir şekilde ve 800 iterasyon civarlarında optimal seviyelere yaklaşabilmektedir. 1000 iterasyona ulaşıldıktan sonra yöntem stabil bir şekilde optimal performans gösterebilmektedir.



Şekil 4.1: A RH+ ve 0 RH+ kan grubu eğitim eğrileri.

Talebin günlük 11 ünite olduğu B RH+'de durumun biraz farklı olduğu ise **Şekil 4.2**'den görülebilmektedir. Yöntem yavaş bir öğrenme ile başladıktan sonra takip eden iterasyonlarda optimal seviyelere yaklaşılan kadar öğrenme hızı artarak devam etmektedir. 1000 iterasyondan sonra da A+ ve 0+ gruplarına benzer şekilde optimal performans düzenli olarak başarılabilir.



Şekil 4.2: B RH+ kan grubu eğitim eğrisi.

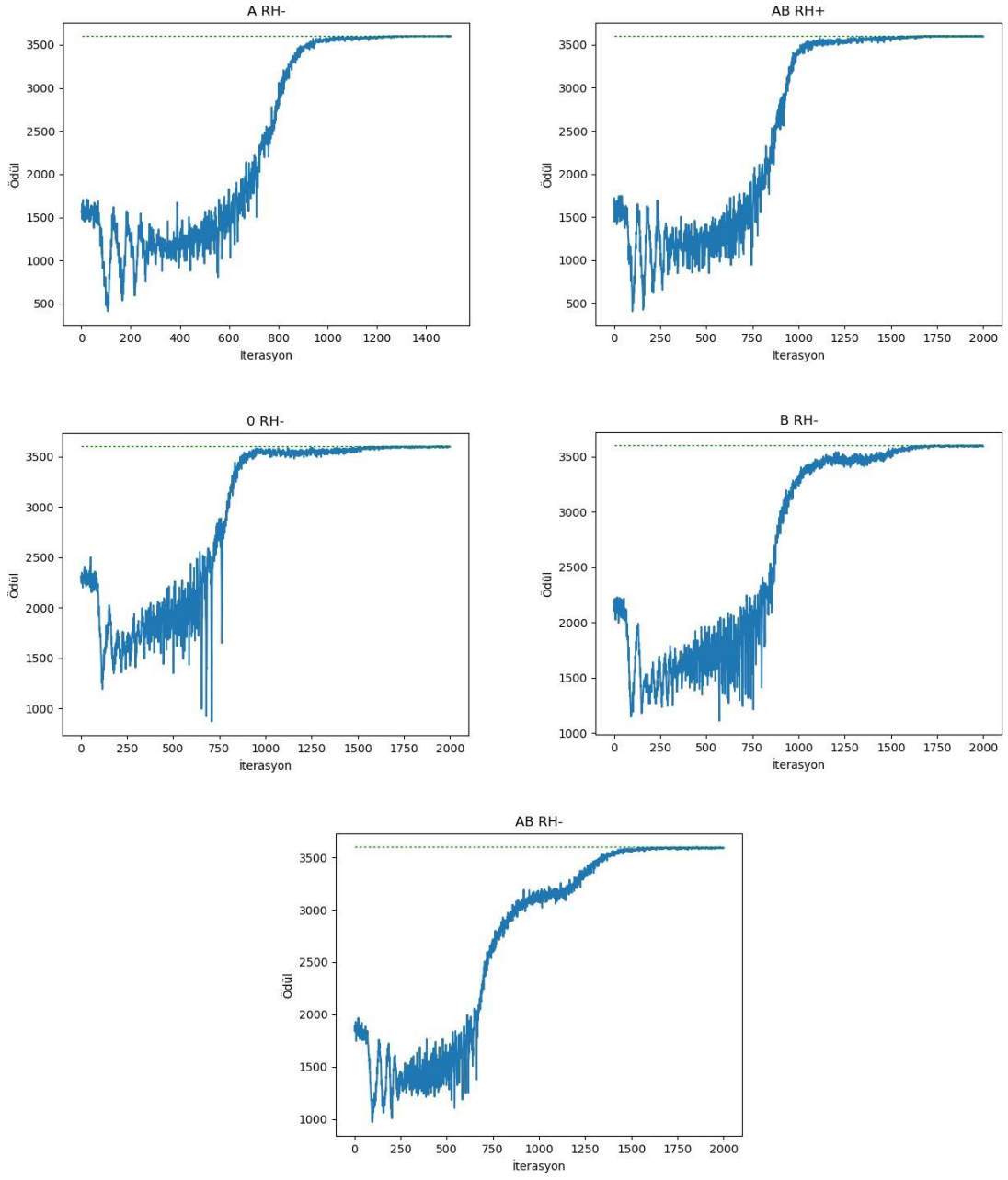
Son olarak, **Şekil 4.3**'e bakıldığında, günlük talebin 1-5 ünite aralığında olduğu kan gruplarında eğitim süreçlerinin birbirine benzer şekilde ilerlediği görülebilmektedir. Başlangıçta toplam ödüller 1500-2000 aralığında artarak-azalarak başlayıp bir süre bu aralıkta seyrettikten sonra ve yukarı aşağı dalgalandıktan sonra yavaş yavaş ortalama toplam ödül artmaktadır. Devamında da, henüz istikrarlı bir şekilde performans gösteremese dahi, ortalama toplam ödül hızlı bir şekilde yükselerek optimal seviyelere yaklaşabilmektedir. İlerleyen iterasyonlarda ortalama toplam ödüldeki dalgalanmalar azalmaktadır ve 1800-1900 iterasyondan sonra istikrarlı bir şekilde optimal seviyelerde performans sergilenebilmektedir. Tüketimin diğer kan gruplarına kıyasla düşük olduğu

kan gruplarının envanter yönetiminin daha zor olması sebebiyle optimal seviyelerde performans çok daha geç öğrenilebilmektedir ve bu sebepten dolayı Bölüm 3.8’de bahsedildiği şekilde, eğitim eğrilerinden de görülebileceği üzere, 5 kan grubu 2000 iterasyon, 3 kan grubu 1500 iterasyon boyunca çalıştırılmıştır.

Eğitim eğrileri kullanılarak model hakkında farklı çıkarımlarda bulunulmuştur. Yöntemin; farklı zaman dilimlerinde, farklı kan grupları için ne seviyelerde performans gösterebildiği, ne kadar istikrarlı bir performans sergileyebildiği, hangi kan gruplarında eğitimin daha zorlayıcı olduğu gibi bilgiler bu eğriler incelenerek keşfedilebilmiştir.

Eğitim eğrileri, her ne kadar eğitim süreci ve genel performans konularında aydınlatıcı olsalar dahi, tam performans hakkında kesin bilgi sağlayamamaktadırlar. Eğitim eğrilerinde ödül değeri kullanıldığı için bu durum ödül fonksiyonlarının nasıl tanımlandığıyla alakalıdır. Bölüm 3.5’de de ifade edildiği üzere israfsız ve karşılanamayan siparişin olmadığı günlerde ödül 1’dir, diğer günlerde ödül 0’dır. Bu sebepten dolayı, karşılanamayan sipariş durumunda kaç ünite karşılanamadığı, veya, israf durumunda kaç ünite israf edildiği bilgisi bu eğrilere yansımamaktadır. Gerçek hayatta ise bu şekilde ikili bir durum söz konusu değildir. Yalnızca kıtlık veya israf olup olmaması değil, aynı zamanda, kıtlık durumu varsa toplamda kaç ünite karşılanamadığı, israf durumu varsa toplamda kaç ünite israf edildiği de önemlidir.

Dolayısıyla, önce ısı haritaları kullanılarak yöntemin hangi durumlarda nasıl hareket etmeyi tercih ettiği incelenmiş, devamında da takip edilen bu politikaların performansları simülasyon sonuçları incelenerek daha hassas bir şekilde değerlendirilmiştir.



Şekil 4.3: A RH-, AB RH+, O RH-, B RH- ve AB RH- kan grubu eğitim eğrileri.

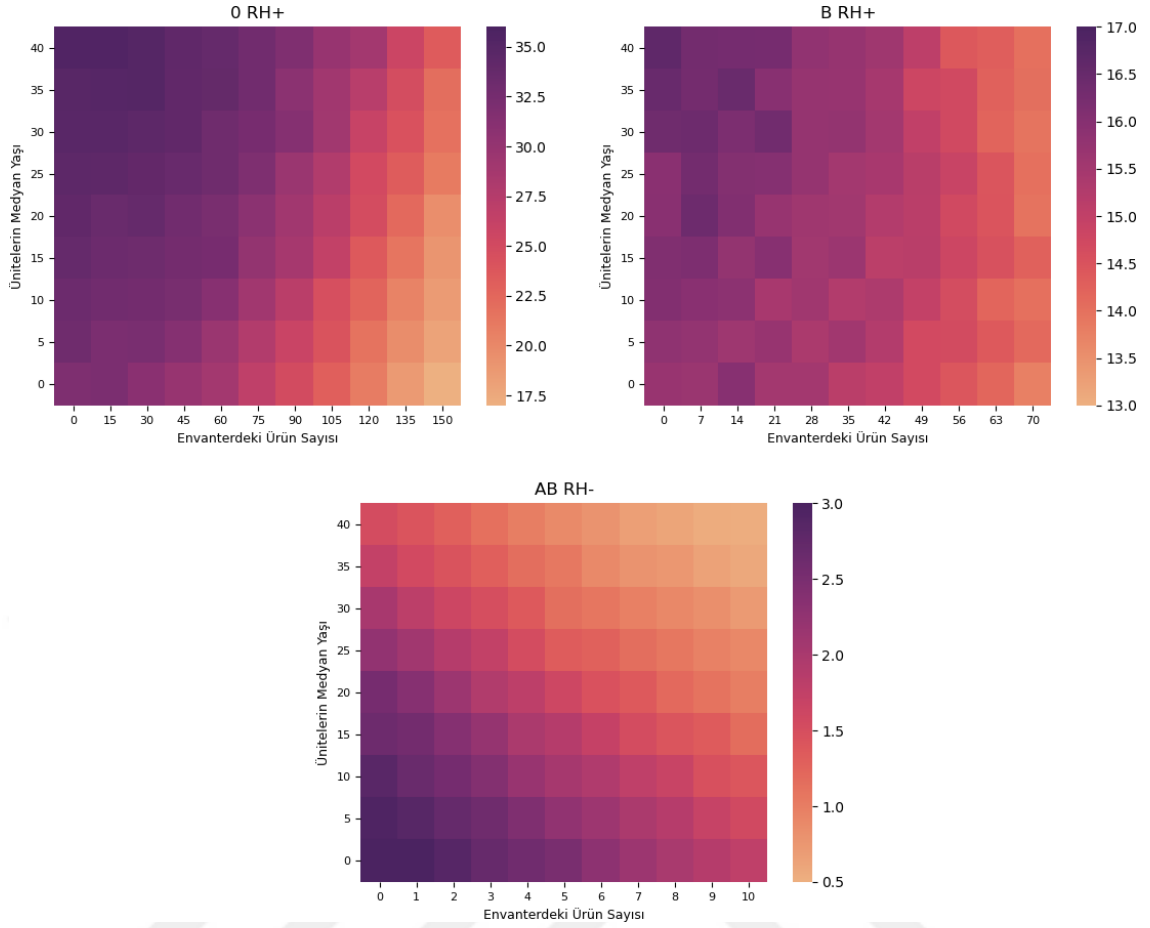
4.2. Isı Haritaları

Isı haritaları incelenerek yöntemin çeşitli durumlara nasıl tepki verdiği görülebilmektedir. Bu analiz için yüksek, orta ve düşük seviyelerde farklı günlük talep yoğunluklarına sahip olan 3 kan grubunun ısı haritası incelenmiştir örnek olarak; 0 RH+, B RH+ ve AB RH-. Durum vektörü 2 durum değişkeninden oluştuğu için 2 boyutlu bir ısı haritası yeterli olmaktadır, karar vermek için gerekli olan tüm durum bilgisi bu şekilde paylaşılabilmektedir. Isı haritaları **Şekil 4.4**'de paylaşılmıştır. Yatay ekseninde durum vektörünün ilk değişkeni olan envanterdeki ünite sayısı, dikey ekseninde ise vektörün ikinci değişkeni olan ünitelerin medyan yaşı yer almaktadır. Dikey ekseninde dikkate alınan aralık tüm kan grupları için aynıdır: 0-40 aralığı. Üniteler, Bölüm 3.1'de ifade edildiği üzere, taşıma ve laboratuvar testleri gibi gerekli işlemlerden geçirildikten sonra en erken 2 gün içerisinde kan bankasında kullanıma hazır hale gelebildiklerinden dolayı bu aralık gerçekte 2-42 gün arasındadır fakat modelde 0-40 aralığında ele alınmaktadır. Her kan grubu için raf ömrü 42 gündür, dolayısıyla 0-40 aralığı tüm ısı haritalarına uygulanmıştır. **Şekil 4.4**'den de görülebileceği üzere yatay eksenlerde ise farklı aralıklara başvurulmuştur. Kan gruplarının günlük talep yoğunluklarının farklı olması sebebiyle her kan grubunun envanteri farklı seviyelerde seyretmektedir. Örneğin 0 RH+ kan grubu için günlük ünite talebi 25 olduğundan dolayı bu durumda envanter seviyesi 0 seviyelerine düşebilir, günlük talebin 5-6 katı seviyelere, 140-150 ünitelere de ulaşabilir. Aynı durum AB RH- gibi düşük seviyede günlük talep ihtiyacı olan bir kan grubu için, eğer sağlıklı bir politika uygulanıyorsa tabii ki, söz konusu olmayacaktır çünkü bu kan grubunda 140-150 ünite envanter seviyelerine ulaşıldığında ürünlerin büyük bir miktarının israf olacağı aşikardır. Böyle bir kan grubu için envanterdeki ünite sayısının 0-10 aralığında seyretmesi doğal olacaktır. Bu sebeplerden dolayı, yatay eksenlerdeki aralıklar kan grubundan kan grubuna değişiklik göstermekte iken dikey eksenlerdeki aralıklar hep 0-40 aralığında tutulmuştur.

Ünite sayılarının ve medyan yaşların kesişimlerine denk gelen renk tonları da yöntemin bu durumda sipariş etmeyi tercih ettiği miktarı temsil etmektedir. Açık tonlar düşük sipariş miktarlarını, koyu tonlar yüksek sipariş miktarlarını yansıtmaktadır. Renk tonlarına karşılık gelen sipariş miktarları şekillerin sağ tarafında yer alan renk skalalarından görülebilmektedir.

Sağlıklı bir politika geliştirebilen her modelden bekleneceği üzere 3 kan grubunda da envanter seviyesinin düşük olduğu durumlarda daha koyu renk tonları, envanter seviyesinin yüksek olduğu durumlarda ise daha açık renk tonları gözlemlenebilmektedir. Bu da demek oluyor ki, 3 modelde de, envanter seviyesinin düşük olduğu durumlarda yüksek sipariş miktarlarına, envanter seviyesinin yüksek olduğu durumlarda ise düşük sipariş miktarlarına karar verilmektedir. İki durumda da optimal envanter seviyelerine dönmek hedeflenmektedir.

Kan gruplarının, envanterdeki ünitelerin medyan yaşına verdiği tepkilerde ise modeller arasında farklılık görülmektedir. Günlük talebin yoğun olduğu 0 RH+ kan grubunda medyan yaşın artmasıyla birlikte sipariş miktarlarının da arttığı görülebilmektedir. B RH+ kan grubunda de benzer bir tablo görülmekle birlikte medyan yaşın sipariş miktarını 0 RH+'deki kadar etkilediği söylenemez. AB RH- kan grubunda ise, medyan yaşın sipariş miktarlarına etkisi söz konusu olduğunda, tam tersi bir durum gözlenmektedir. Sipariş miktarları medyan yaş ile ters orantılı bir şekilde hareket etmektedir. Envanterdeki ürün sayısına bağlı olmaksızın, medyan yaş arttıkça sipariş miktarlarının azaldığı gözlemlenebilmektedir. Derin pekiştirmeli öğrenme tekniği kullanıldığından dolayı yöntemin hangi sebeplerle bu şekilde hareket ettiğinin tam olarak anlaşılabilmesi zordur. Her halükarda, eğitim tamamlandıktan sonra yöntemin sergilemiş olduğu performansa bakılarak mantıklı politikalar öğrenilip öğrenilemediği değerlendirilebilir.



Şekil 4.4: 0 RH+, B RH+ ve AB RH- kan grubu ısı haritaları.

4.3. Performans

Yukarıda paylaşılmış olan eğitim eğrilerinden görülebileceği üzere tüm kan grupları son iterasyonlarda stabil bir performans sergileyebilmektedir. **Çizelge 4.1**'de, her kan grubu için son 50 simülasyonda gerçekleşen israf ve karşılanamayan ünite ortalamaları paylaşılmıştır. Her simülasyon 3600 zaman adımında tamamlandığı, dolayısıyla 10 yıl çalıştırıldığı için tabloda paylaşılan rakamlar 10 yıllık verilerin 1 yıla dönüştürülmüş hallerini içermektedirler. 0 RH- kan grubunu ele alacak olursak israf edilen ünite sayısının 0.2, karşılanamayan ünite sayısının 0.7 olduğu görülmektedir. Bu da demek oluyor ki, son 50 iterasyonun ortalamasına bakıldığında, 10 yıl süresince ortalama 2 ünite israf edilmiş, 7 ünite eritrosit karşılanamamıştır.

Çizelge 4.1: Farklı kan grupları için son 50 simülasyonda gerçekleşen ortalama israf ve kıtlık.

Kan grubu	İsraf edilen ünite sayısı	Karşılanamayan ünite sayısı
0+	0.0	1.2
0-	0.2	0.7
A+	0.0	1.1
A-	0.0	0.3
B+	0.0	0.4
B-	0.5	0.2
AB+	0.0	0.6
AB-	0.9	0.2
Genel ortalama	0.2	0.6

Tüm kan gruplarının benzer seviyelerde performans sergilediği görülebilmektedir. Her ne kadar bazı kan gruplarında diğer kan gruplarına kıyasla daha uzun eğitim süreleri gerekmiş olsa dahi, her kan grubu optimal seviyelere, yani hastanedeki 0 israf ve %100 sipariş karşılama oranına, yakın bir performansa ulaşabilmiştir. İsraf ve karşılanamayan sipariş oranları 0'a indirilememiş olsa dahi ortalama israf edilen ünite sayısı yılda 0.2 üniteye, ortalama karşılanamayan ünite sayısı yılda 0.6 üniteye kadar düşürülebilmektedir.

Kıyaslama olarak hastane seçildiğinden dolayı yöntemin sergilemiş olduğu performansı değerlendirirken hastanenin hangi koşullar altında faaliyet gösterdiğini modelin hangi koşullar altında çalıştığıyla karşılaştırmak, israf ve kıtlık oranlarını doğrudan veya dolaylı yoldan etkileyebilecek varsayımları detaylandırmak isabetli olacaktır. Örneğin, hastanenin, ihtiyaca bağlı olarak, gün içerisinde acil sipariş aracılığıyla talep karşılama imkanı ve daha acil durumlarda da genel verici kan grubu 0 RH- ile talep karşılama imkanı bulunmaktadır. Acil siparişler ne kadar maliyetli olursa olsun, veya 0 RH- ile uyumsuz eşleştirme yapmak ne kadar istenmeyen bir durum olursa olsun hastane, karşılaştığı talepleri bu iki yöntem aracılığıyla da karşılayabilmektedir ancak kurulan modelde bu 2 yönteme de izin verilmemektedir. Aynı zamanda, hastanenin, tüketemeyeceğini düşündüğü ürünleri Kızılay'a geri verme imkanı da bulunmaktadır. Raf ömrü dolma ihtimali artan ürünler gerçek hayatta Kızılay'a iade edilebilmektedir ve bu şekilde israf oranı düşürülebilmektedir fakat ele alınan sistemde bu işlem de kapsam dışı tutulmuştur.

Hastane, kurulan derin pekiştirmeli öğrenme modeline kıyasla sahip olduğu bu “kolaylıkların” yanında, envanter yönetimini zorlaştıracak birçok etken ile de baş etmelidir. Çalışma kapsamı dışında tutulan günden güne değişiklik gösteren talep seviyeleri, gerçek hayatta yaşanabilecek afet, kaza vb. durumlar, hekimlere atandığı halde sonrasında ihtiyaç duyulmayarak envantere döndürülen ünitelerin yaşlanması, talebe yansıtılacak olan trend ve sezonsallık, veya ünitelerin kan bankasına teslimi esnasında rastgele yaşlarda gelmesi durumu bu etkenlerden birkaçı olarak sıralanabilir. Bu etkenler arasından yalnızca rastgele yaşlarda gelen ürün durumu derin pekiştirmeli öğrenme modeline yansıtılmıştır. Hem hastanede, hem de kurulan modelde envanter yönetiminin kolaylaşmasını sağlayan bir durum ise talep edilen ünitelerin Kızılay tarafından tamamen karşılanabiliyor olmasından kaynaklanmaktadır.

Ayrıca hastane, çalışılan sistemden farklı olarak, hekimlerin ertesi gün için taleplerini sipariş vermeden önce toplamaktadır ve farklı ünitelerde bulunan hasta sayıları, rutin kan ihtiyacı olan hasta bilgisi gibi yardımcı bilgilere de sahiptir. Dolayısıyla, hastanenin yüzleştiği bazı zorluklar çalışma kapsamında tutulmuş veya modellenememiş olsa dahi kurulan model de hastanenin sahip olduğu bazı kolaylıklardan mahrum bırakılmıştır. Eritrosit envanter yönetiminin kendine özgün karakteristiklerinden kaynaklanan karmaşıklığı göz önüne alındığında derin pekiştirmeli öğrenme yönteminin her kan grubunda optimal seviyelerde performans gösterebilecek politikalar geliştirebilmesi etkileyicidir.

BÖLÜM 5

5. SONUÇ VE GELECEKTEKİ ÇALIŞMALAR

Bu bölümde, bu araştırmanın en önemli bulguları ve sonuçları tartışılmıştır. Bölüm 5.1’de en önemli bulgular, literatüre yapılan katkılar ve mevcut sınırlamalar tartışılmıştır. Bölüm 5.2, bu araştırmanın uygulanması konusundaki önerileri ve Bölüm 5.3 de gelecek araştırmalar için sunulan önerileri kapsamaktadır.

5.1. Sonuç

Hastane kan bankası envanter yönetiminde geleneksel çözüm metodları yaygın bir şekilde kullanılmaktadır ancak bu çalışmada derin pekiştirmeli öğrenme yönteminin bu alandaki uygulanabilirliğine odaklanılmıştır. Bu araştırma için uygun bir örnek olarak İstanbul’daki büyük ölçekli bir özel hastane seçilmiştir. Hastanenin transfüzyon merkezi hastane seviyesinde faaliyet göstermektedir ve envanter yönetimi şu anda insan tecrübesine, uzman tecrübesine dayalı olarak yürütülmektedir. Dolayısıyla, aşağıdaki araştırma sorusu formüle edilmiştir:

Hastane kan bankası eritrosit envanter yönetiminde derin pekiştirmeli öğrenme yöntemi kullanılabilir mi?

Ölçeklendirilebilir, karmaşık problemlerde dahi optimal politikalara yakınsayabiliyor olmaları sebebiyle pekiştirmeli öğrenme metodları arasından yalnızca derin pekiştirmeli öğrenme yöntemleri dikkate alınmıştır. Bu metodlar arasından da yakın zamanda başarılı sonuçlar sergileyerek popülerleşen Schulman’ın [1] Proksimal Politika Optimizasyonu tercih edilmiştir. Bu yöntem, değer fonksiyonunu yakınsayabilmek için sinir ağlarını kullanmaktadır, dolayısıyla büyük problemlerde kullanılabilir.

Modelde herhangi bir düzenleme, parametre değişikliği yapılmadan derin pekiştirmeli

öğrenme ile farklı kan grupları için birçok deney gerçekleştirilmiştir ve yöntemin sağlıklı bir öğrenme gerçekleştirebildiği, her deneyde optimale yakın performanslar yakalanabildiğinden dolayı, kanıtlanmıştır. Denemeler esnasında durum ve eylem vektörlerinin tanımlarının algoritma performansında etkili olduğu da gözlemlenmiştir. Elde edilen sonuçların kıyaslanabilmesi adına hastanenin gerçek hayattaki performansı kıyaslama olarak ele alınmıştır. Kan gruplarının son 50 simülasyondaki israf ve kıtlık ortalamaları hastanenin israf ve kıtlık ortalamaları ile kıyaslanarak yöntemin performansı ölçülmüştür. Her bir kan grubu için uzun süren simülasyonlar çalıştırılarak derin pekiştirmeli öğrenme metodunun her seferinde optimal politikalara yakınsayabildiği gözlemlenmiştir. Başlangıçta her kan grubu için model 1500 iterasyon çalıştırılmıştır. 8 kan grubundan 5 tanesinin 1500 iterasyonda optimal politikalara yakınsayamadığı gözlemlendiğinden dolayı bu kan grupları birer defa daha 2000'er iterasyon çalıştırılmıştır ve yeterli iterasyon imkanı sağlandığında bu 5 kan grubunun da optimal performansı stabil bir şekilde sergileyebilecek seviyeye ulaşabildiği keşfedilmiştir.

Sinir ağlarının öğrenme biçimlerinin ve karar verme mekanizmalarının daha iyi anlaşılabilmesi adına farklı yöntemler kullanılmıştır. Eğitim eğrileri üzerinden yöntemin farklı kan grupları için tam olarak nasıl bir öğrenme sürecinden geçtiği ve yakınsanan politikaların ne kadar stabil performanslar sergileyebildiği tartışılmıştır. Bununla birlikte, yöntemin hangi durumda nasıl karar vermeyi tercih ettiğinin daha iyi anlaşılabilmesi için iki durum değişkenine, envanterdeki ürün sayısı ve ünitelerin medyan yaşı, bağlı olarak karar verilen sipariş miktarlarının gösterildiği ısı haritaları oluşturulmuştur. Bu şekilde karar verme mekanizması yakından incelenebilmiş ve verilen kararların mantıklı kararlar olup olmadığı daha kolay yorumlanabilmiştir.

Kurulan derin pekiştirmeli öğrenme yöntemi, eğitilmiş bir sinir ağı kullanılarak hastanede uygulanabilir. Hastane her sipariş vermek istediğinde, envanter sisteminin mevcut durumu sinir ağına girdi olarak iletilebilir ve çıktı olarak da eğitim esnasında belirlenen en iyi sipariş miktarları alınabilir. Ancak envanter sisteminin mevcut durumunu belirlemek zor olabilir. Bu nedenle, sinir ağının mevcut durum hakkında gerekli bilgileri otomatik olarak çıkarabilmesi için sinir ağını hastanenin envanter sistemine bağlamak anlamlı bulgular elde edilmesine yardımcı olabilir. Derin pekiştirmeli öğrenme yönteminin farklı hastanelere veya farklı kan ürünlerine uygulanabilmesi için iki parametre dikkate alınmalıdır. Dikkat edilmesi gereken ilk husus envanter sistemi ile ilgili tüm bilgilerin toplanmasıdır.

Talep yoğunluğu ve takip edilen atama politikası gibi. İkinci husus ise her özel sistem için durum ve eylem değişkenlerinin doğru bir şekilde ayarlanması olacaktır. En uygun durum ve eylem değişkenleri bulunduğu anda yöntem çalıştırılabilir ve ağ eğitilebilir. Bu ağ daha sonra her durum için uygun sipariş miktarlarını elde etmek için kullanılabilir.

Her ne kadar sonuçlar umut verici olsa da, bu araştırma ve yöntem bazı sınırlamalar altında gerçekleştirilmiştir. Bu çalışmada kullanılan verilerin ve yöntemin sınırlılıkları aşağıda tekrardan özetlenmiştir.

Bu çalışmanın en önemli sınırlaması, derin pekiştirmeli öğrenme yönteminin performansının hastanenin gerçek performansı ile karşılaştırılmasının zorluğudur. Ele alınan hastane örneğinin çeşitli varsayımlar altında ele alınmasına karar verildiğinden dolayı bu sınırlama, yapılan değişikliklerin, kullanılan varsayımların veya basitleştirmelerin bir ürünüdür. Bu yaklaşımlar, örnek sistemin simülasyona aktarılmasını kolaylaştırmıştır. Bu nedenle, bu araştırma, derin pekiştirmeli öğrenme yönteminin mümkün kıldığı iyileştirmeleri yaklaşımsal olarak gösterebilmektedir.

Bu sadeleştirmelerden biri de taleple ilgilidir. Talep verilerinin toplam adet değil de günlük veri olarak elde edilememesinden dolayı talep verileri için bazı sadeleştirmeler yapılarak Poisson dağılımları kullanılmıştır. Bu nedenle, bu varsayımın yöntem performansı üzerinde büyük bir etki sahibi olma ihtimali vardır. Son olarak, hastanenin mevcut performansı bir kıyaslama olarak kullanılmıştır. Bu nedenle, derin pekiştirmeli öğrenme yönteminin diğer simülasyon veya matematiksel modelleme gibi tekniklerle karşılaştırıldığında ne kadar iyi performans gösterdiği bilinmemektedir.

5.1.1. Akademik katkı

Bu araştırma, derin pekiştirmeli öğrenme ve hastane kan bankası envanter yönetimi alanlarının her birine özgün katkılarda bulunmaktadır. Bilindiği kadarıyla, bu araştırma sinir ağlarını bir hastane kan bankası envanter sistemine uygulayan ilk örnektir. Proksimal Politika Optimizasyonu algoritması kullanılarak her kan grubu için kıyaslama performansı yakalanabilmiştir. Bunu başarabilmek için, çeşitli talep yoğunluklarına sahip eritrositlerin modellenip çözülebildiği bir bozulabilir ürün ortamı oluşturulmuştur. Bu ortam, gelecekteki araştırmalarda da kullanılabilir ve kan bankası envanter yönetiminin daha da önemli yönlerini simüle etmek için genişletilebilir. Bu çalışma ile, inovatif bir katkı ile ilk kez

derin pekiştirmeli öğrenme yöntemi bir hastane kan bankası envanter sisteminde uygulanmıştır.

5.2. Öneriler

Bu araştırma, derin pekiştirmeli öğrenmenin kan bankası envanter sorunlarını çözmek için uygun bir teknik olabileceğini göstermiştir. Bu çalışma, elde ettiği bulgular eşliğinde yeni araştırma konularına da girdi temin etmiştir. Devam niteliğindeki araştırmalara yön gösterici olarak şu maddeler önerilebilir:

- Henüz ana çözüm olarak derin pekiştirmeli öğrenme kullanılmaya başlanmamalı. Bunun yerine, yöntemin uzmanlara kıyasla nasıl performans gösterdiğinin doğrulanabilmesi için bu yöntem destekleyici bir yöntem olarak kullanılabilir.
- Çözüm hastanelere sunulmadan önce yöntemin açıklanabilirliğine odaklanılmalı. Yaygın olarak bilinmektedir ki, derin pekiştirmeli öğrenme bazen bir “kara kutu” sistemi olarak kabul edilebilmektedir çünkü yöntemin neden belirli bir eylemi seçtiği belirsiz olabilmektedir. Ancak, durum böyle olmak zorunda değildir. Çeşitli araştırmacılar sinir ağlarının açıklanabilirliği ve derin pekiştirmeli öğrenme konusunu incelemektedir ve bu çalışmada da uygulanan yöntemin verdiği kararların sebepleri anlaşılmaya çalışılmıştır. Araştırmacıların yöntemin açıklanabilirliğine de odaklanmaları önerilmektedir.
- Ortamın ve derin pekiştirmeli öğrenme yönteminin karmaşıklığını azaltmanın yolları aranmalıdır. Bu, eylem alanını daha da küçültebilir ve çeşitli durumlara daha az bağımlı hale getirebilir. Bu durum, ajanın eğitim süresini azaltacaktır ve daha karmaşık ortamlar için uygun çözümler bulma olasılığını da artıracaktır.
- Derin pekiştirmeli öğrenme alanındaki gelişmeler yakından takip edilmelidir. Derin pekiştirmeli öğrenme için açık kaynak kodlu algoritmalar mevcuttur. Üstelik bu algoritmalar genellikle birkaç Python paketinde halihazırda uygulanmıştır. Hangi tür problem için en iyi algoritmanın ne olduğunu araştırmak, derin pekiştirmeli öğrenme yönteminin performansını artırabilir.

5.3. Gelecek Arařtırmalar

Bu tez, hastane kan bankası envanter sistemlerinde iyi performans gösterebilen bir derin pekiřtirmeli öğrenme yöntemini tanıtmaktadır. Bununla birlikte, bu yöntemin kullanımı, farklı arařtırmalar için önerilerle sonuçlanan yeni soruları da gündeme getirmiřtir.

- Durum ve eylem vektörlerinin performans üzerinde büyük bir etkisi olduđu fark edilmiřtir. Bu vektörlerin tanımlanması önemsiz deđildir ve duruma özgüdür. Bu nedenle, bu süreci otomatikleřtirmenin ve optimize etmenin yolları aranabilir.

- Bu çalışmada ele alınan hastane örneğinin basitleřtirilmiř bir versiyonuna derin pekiřtirmeli öğrenme yöntemi uygulanmıřtır. Bu nedenle gelecekteki arařtırmalar için simülasyonun hastanenin gerçek hayat durumunu yansıtacak şekilde daha da genişletilmesi önerilebilir. Bu şekilde yöntem daha karmařık farklı envanter sistemlerinde de kullanılabilir. Bu, sezgisel yöntemler ve matematiksel modeller gibi daha geleneksel yöntemlerle çözülmesi zor olan envanter sistemleri için özellikle anlamlı bilgiler sağlayacaktır.

- Bunun yanında, hastanenin orijinal günlük talep verilerinin kullanılabilmesi de mümkündür. Derin pekiřtirmeli öğrenmenin, talebin tahmin edilmesi zorlařtıđında ve ortamın stokastikliđi arttıđında hala sipariř miktarlarını belirleyip belirleyemediđi gözlemlenebilirdi.

- Bu tezde, hiperparametrelere daha az duyarlı olduđu için ve bu nedenle farklı vakalara uygulanması daha kolay olan mevcut son teknoloji algoritma olduđu için Proksimal Politika Optimizasyonu algoritması uygulanmıřtır. Bununla birlikte, belirli durumlar için daha uygun olabileceğinden, diđer derin pekiřtirmeli öğrenme algoritmalarının incelenmesi de mümkündür. Bu algoritmaların çeřitli oyunlardaki karşılařtırılması üzerine arařtırmalar yapılmıřtır ancak henüz bir hastane kan bankası envanter ortamı üzerinde kıyaslama yapılmamıřtır. Bu nedenle, hastane kan bankası envanter ortamında çeřitli algoritmaların uygulanması, bu alandaki performanslarının karşılařtırılması önerilebilir.

- Talep tahmini ve envanter optimizasyonunu birleřtiren veriye dayalı sipariř yönetimi stratejileri sađlık sistemi sađlayıcı aktörler ve kan tedarikçileri için önemli maliyet tasarrufları sađlayabilir ve sistem iyileřtirici aksiyonlar alınmasına yardımcı olabilir.

KAYNAKÇA

- [1] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, ve O. Klimov, “Proximal policy optimization algorithms,” *CoRR*, c. abs/1707.06347, 2017.
- [2] H. Lowalekar ve R. Narasimhan, “A combined age-and-stock-based policy for ordering blood units in hospital blood banks,” *International Transactions in Operational Research*, c. 24, 07 2015.
- [3] S. Nahmias, “Perishable inventory theory: A review,” *Operations research*, c. 30, s. 680–708, 08 1982.
- [4] L. Sağdur, “The role of kızılay and the red cross in the blood supply the turkish red crescent blood banking,” *Global Journal of Transfusion Medicine AATM*, c. 13, s. 63–67, 2020.
- [5] S. H. G. M. K. ve Kan Ürünleri Dairesi Başkanlığı, “Kan ve kan Ürünleri yönetmeliği,” 2020. URL: <https://shgmkanhizmetleridb.saglik.gov.tr/TR-63680/kan-ve-kan-urunleri-yonetmeliği.html>
- [6] J. Chapman, “Unlocking the essentials of effective blood inventory management,” *Transfusion*, c. 47, 2007.
- [7] M. Christopher, *Logistics and Supply Chain Management: Creating Value-adding Networks*. FT Prentice Hall, 2005. URL: <https://books.google.com.tr/books?id=7Uq4swEACAAJ>
- [8] T. . U. I. Programı, “Ulusal kan ve kan bileşenleri hazırlama, kullanım ve kalite güvencesi rehberi 2016,” 2016. URL: <https://shgmkanhizmetleridb.saglik.gov.tr/Eklenti/36313/0/kan-ve-kan-bilesen-rehb-tr-9-mayis-2016-revpdf.pdf>
- [9] N. Van Dijk, R. Haijema, J. Van Der Wal, ve C. S. Sibinga, “Blood platelet production: a novel approach for practical optimization,” *Transfusion*, c. 49, sayı. 3, s. 411–420, 2009.
- [10] R. Kopach, B. Balcioglu, ve M. Carter, “Tutorial on constructing a red blood cell inventory management system with two demand rates,” *European Journal of Operational Research*, c. 185, s. 1051–1059, 02 2008.
- [11] R. S. Sutton ve A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
- [12] L. P. Kaelbling, M. L. Littman, ve A. W. Moore, “Reinforcement learning: A survey,” *CoRR*, c. cs.AI/9605103, 1996.
- [13] J. Shin ve J. Lee, “Multi-timescale, multi-period decision-making model development by combining reinforcement learning and mathematical programming,” *Computers Chemical Engineering*, c. 121, 11 2018.
- [14] V. Chaudhary, R. Kulshrestha, ve S. Routroy, “State-of-the-art literature review on inventory models for perishable products,” *Journal of Advances in Management Research*, c. 15, 03 2018.
- [15] J. Gijsbrechts, R. Boute, D. Zhang, ve J. Van Mieghem, “Can deep reinforcement learning improve inventory management? performance on dual sourcing, lost sales and multi-echelon problems,” *SSRN Electronic Journal*, 07 2019.
- [16] N. Heddle, R. Cook, D. Arnold, Y. Liu, R. Barty, M. Crowther, P. Devereaux, J. Hirsh, T. Warkentin, K. Webert, D. Roxby, M. Sobieraj-Teague, A. Kurz,

- D. Sessler, P. Figueroa, M. Ellis, ve J. Eikelboom, “Effect of short-term vs. long-term blood storage on mortality after transfusion,” *The New England journal of medicine*, c. 375, 10 2016.
- [17] B. Zahiri, S. Torabi, M. Mousazadeh, ve A. Mansouri, “Blood collection management: Methodology and application,” *Applied Mathematical Modelling*, 05 2015.
- [18] I. Giannoccaro and P. Pontrandolfo, “Inventory management in supply chains: a reinforcement learning approach,” *International Journal of Production Economics*, c. 78, sayı. 2, s. 153–161, 2002.
- [19] R. Bellman, *Dynamic Programming*. Dover Publications, 1957.
- [20] M. R. K. Mes ve A. P. Rivera, “Approximate dynamic programming by practical examples,” *Operations Research and Management Science*, s. 63–101, 2017.
- [21] R. S. Sutton, *Temporal credit assignment in reinforcement learning*. University of Massachusetts Amherst, 1984.
- [22] M. Çimen ve C. Kirkbride, “Approximate dynamic programming algorithms for multidimensional inventory optimization problems,” in *MIM*, 2013.
- [23] C. J. C. H. Watkins, “Learning from delayed rewards,” Ph.D. dissertation, King’s College, Cambridge, UK, May 1989.
- [24] T. Jaakkola, M. I. Jordan, and S. Singh, “On the convergence of stochastic iterative dynamic programming algorithms,” *Neural Computation*, c. 6, s. 1185–1201, 1994.
- [25] N. Habib, *Hands-On Q-Learning with Python: Practical Q-learning with OpenAI Gym, Keras, and TensorFlow*. Packt Publishing, 2019.
- [26] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, ve D. Hassabis, “Human-level control through deep reinforcement learning,” *Nature*, c. 518, sayı. 7540, s. 529–533, Feb. 2015.
- [27] K. Arulkumaran, M. P. Deisenroth, M. Brundage, ve A. A. Bharath, “A brief survey of deep reinforcement learning,” *CoRR*, c. abs/1708.05866, 2017.
- [28] S. J. Russell ve P. Norvig, *Artificial Intelligence*. Pearson Education, 2009.
- [29] W. Powell, “Approximate dynamic programming: Solving the curses of dimensionality,” 08 2011.
- [30] L. Lin, “Self-improvement based on reinforcement learning, planning and teaching,” in *ML*, 1991.
- [31] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. M. O. Heess, T. Erez, Y. Tassa, D. Silver, ve D. Wierstra, “Continuous control with deep reinforcement learning,” *CoRR*, c. abs/1509.02971, 2016.
- [32] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. P. Lillicrap, T. Harley, D. Silver, ve K. Kavukcuoglu, “Asynchronous methods for deep reinforcement learning,” *CoRR*, c. abs/1602.01783, 2016.
- [33] Y. Zhou, T. Zou, C. Liu, H. Yu, L. Chen, ve J. Su, “Blood supply chain operation considering lifetime and transshipment under uncertain environment,”

Applied Soft Computing, c. 106, s. 107364, 03 2021.

- [34] K. Katsaliaki ve S. C. Brailsford, “Using simulation to improve the blood supply chain,” *Journal of the operational research society*, c. 58, sayı. 2, s. 219–227, 2007.
- [35] DHHS, “The 2011 national blood collection and utilization survey report,” 2013.
- [36] C. Eren, “İstanbul İlinde abo ve rh kan grupları dağılımının analizi,” *Dicle Tıp Dergisi*, c. 46, sayı. 2, s. 241 – 246, 2019.
- [37] G. P. Prastacos, “Blood inventory management: An overview of theory and practice,” *Management Science*, c. 30, sayı. 7, s. 777–800, 1984.
- [38] H. Lowalekar ve N. Ravichandran, “A combined age-and-stock-based policy for ordering blood units in hospital blood banks,” *International Transactions in Operational Research*, c. 24, sayı. 6, s. 1561–1586, 2017.
- [39] A. Pirabán Ramírez, W. Guerrero, and N. Labadie, “Survey on blood supply chain management: Models and methods,” *Computers Operations Research*, c. 112, 07 2019.
- [40] N. Li, D. M. Arnold, D. G. Down, R. Barty, J. Blake, F. Chiang, T. Courtney, M. Waito, R. Trifunov, ve N. M. Heddle, “From demand forecasting to inventory ordering decisions for red blood cells through integrating machine learning, statistical modeling, and inventory optimization,” *Transfusion*, c. 62, sayı. 1, s. 87–99, 2022.
- [41] J. van Sambeeck, S. van Brummelen, N. van Dijk, ve M. Janssen, “Optimal blood issuing by comprehensive matching,” *European Journal of Operational Research*, c. 296, sayı. 1, s. 240–253, 2022.
- [42] M. Ahmadimanesh, A. Pooya, H. Safabakhsh, ev S. Sadeghi, “Designing an optimal model of blood logistics management with the possibility of return in the three-level blood transfusion network,” 12 2021.
- [43] N. Li, F. Chiang, D. G. Down, ve N. M. Heddle, “A decision integration strategy for short-term demand forecasting and ordering for red blood cell components,” *Operations Research for Health Care*, c. 29, s. 100290, 2021.
- [44] M. Meneses, I. Marques, ve A. Barbosa-Póvoa, “Blood inventory management: Ordering policies for hospital blood banks under uncertainty,” *International Transactions in Operational Research*, c. 30, sayı. 1, s. 273 – 301, 2023, cited by: 4. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85135518778>
- [45] M. Arani, X. Liu, ve S. Abdolmaleki, “Scenario-based simulation approach for an integrated inventory blood supply chain system,” in *2020 Winter Simulation Conference (WSC)*, 2020, s. 1348–1359.
- [46] A. Araújo, D. Santos, I. Marques, ve A. Barbosa-Povoá, “Blood supply chain: a two-stage approach for tactical and operational planning,” *OR Spectrum*, c. 42, 12 2020.
- [47] X. Guo, A. Liu, X. Li, ve T. Liu, “A two-stage stochastic model for daily reserve in inventory management of rh-negative red blood cells,” *Journal of Intelligent and Fuzzy Systems*, c. 39, sayı. 5, s. 6919 – 6933, 2020, cited by: 0.

- [48] A. Bozorgi ve M. Najafi, “Improving blood bank inventory management using double cross-match and hybrid issuance policy,” in *2020 IEEE 7th International Conference on Industrial Engineering and Applications (ICIEA)*, 2020, s. 819–826.
- [49] M. Ahmadimanesh, A. Tavakoli, A. Pooya, ve F. Dehghanian, “Designing an optimal inventory management model for the blood supply chain: Synthesis of reusable simulation and neural network,” *Medicine*, c. 99, s. e21208, 07 2020.
- [50] G. K. Ajay Singh Yadav, A. Swami, “Blood bank supply chain inventory model for blood collection sites and hospital using genetic algorithm,” *Selforganizology*, c. 6, 12 2019.
- [51] B. Hamdan ve A. Diabat, “A two-stage multi-echelon stochastic blood supply chain problem,” *Computers Operations Research*, c. 101, s. 130–143, 2019.
- [52] S. Ghorashi, M. Hamed, ve R. Sadeghian, “Modeling and optimization of a reliable blood supply chain network in crisis considering blood compatibility using mogwo,” *Neural Computing and Applications*, c. 32, 08 2020.
- [53] N. M. Clay, B. Abbasi, A. Eberhard, ve J. Hearne, “On the volatility of blood inventories,” *International Transactions in Operational Research*, c. 25, sayı. 1, s. 215–242, 2018.
- [54] J. Duan, Q. Su, Y. Zhu, ve Y. Lu, “Study on the centralization strategy of the blood allocation among different departments within a hospital,” in *2017 International Conference on Service Systems and Service Management*, 2017, s. 1–5.
- [55] V. Sarhangian, H. Abouee-Mehrzi, O. Baron, ve O. Berman, “Threshold-based allocation policies for inventory management of red blood cells,” *Manufacturing & Service Operations Management*, c. 20, sayı. 2, s. 347–362, 2018.
- [56] M. Dillon, F. Oliveira, ve B. Abbasi, “A two-stage stochastic programming model for inventory management in the blood supply chain,” *International Journal of Production Economics*, c. 187, s. 27–41, 2017.
- [57] M. Najafi, A. Ahmadi, ve H. Zolfagharinia, “Blood inventory management in hospitals: Considering supply and demand uncertainty and blood transshipment possibility,” *Operations Research for Health Care*, c. 15, s. 43–56, 2017.
- [58] V. Sarhangian, H. Abouee-Mehrzi, O. Baron, O. Berman, N. M. Heddle, ve R. Barty, “Reducing the age of transfused red blood cells in hospitals: ordering and allocation policies,” *Vox Sanguinis*, c. 110, sayı. 4, s. 385–392, 2016.
- [59] S. Gunpınar ve G. Centeno, “Stochastic integer programming models for reducing wastages and shortages of blood products at hospitals,” *Computers Operations Research*, c. 54, s. 129–141, 02 2015.
- [60] A. Grasas, A. Pereira, M.-A. Bosch, P. Ortiz, ve L. Puig, “Feasibility of reducing the maximum shelf life of red blood cells stored in additive solution: a dynamic simulation study involving a large regional blood system,” *Vox Sanguinis*, 10 2014.
- [61] M. Olusanya, M. Arasomwan, ve A. Adewumi, “Particle swarm optimization algorithm for optimizing assignment of blood in blood banking system,” *Computational and Mathematical Methods in Medicine*, c. 2015, s. 1–12, 03

2015.

- [62] Q. Duan ve T. Liao, “Optimization of blood supply chain with shortened shelf lives and abo compatibility,” *International Journal of Production Economics*, c. 153, s. 113–129, 07 2014.
- [63] B. Abbasi ve Z. HosseiniFard, “On the issuing policies for perishable items such as red blood cells and platelets in blood service,” *Decision Sciences*, c. 45, 10 2014.
- [64] J. Blake ve M. Hardy, “A generic modelling framework to evaluate network blood management policies: The canadian blood services experience,” *Operations Research for Health Care*, c. 3, 09 2014.
- [65] F. Baesler, M. Nemeth, C. Martínez, ve A. Bastias, “Analysis of inventory strategies for blood components in a regional blood center using process simulation,” *Transfusion*, c. 54, 06 2013.
- [66] A. Simonetti, R. Forshee, S. Anderson, ve M. Walderhaug, “A stock-and-flow simulation model of the us blood supply,” *Transfusion*, c. 54, 08 2013.
- [67] J. T. Blake, M. Hardy, G. Delage, ve G. Myhal, “Déjà-vu all over again: using simulation to evaluate the impact of shorter shelf life for red blood cells at héma-québec,” *Transfusion*, c. 53, sayı. 7, s. 1544–1558, 2013.
- [68] M. Atkinson, M. Fontaine, L. Goodnough, ve L. Wein, “A novel allocation strategy for blood transfusions: Investigating the tradeoff between the age and availability of transfused blood,” *Transfusion*, c. 52, s. 108–17, 07 2011.
- [69] M. Fontaine, Y. Chung, F. Erhun, ve L. Goodnough, “Age of blood as a limitation for transfusion: Potential impact on blood inventory and availability,” *Transfusion*, c. 51, s. 662–3, 03 2011.
- [70] K. Katsaliaki, “Cost-effective practices in the blood service sector,” *Health policy (Amsterdam, Netherlands)*, c. 86, s. 276–87, 06 2008.
- [71] J. Ryttilä ve K. Spens, “Using simulation to increase efficiency in blood supply chains,” *Management Research News*, c. 29, s. 801–819, 12 2006.
- [72] A. Pereira, “Blood inventory management in the type and screen era,” *Vox Sanguinis*, c. 89, sayı. 4, s. 245–250, 2005.
- [73] A. Shapiro, D. Dentcheva, ve A. Ruszczyński, *Lectures on stochastic programming. Modeling and theory*, 01 2009.
- [74] A. Paul, T. Rajapakshe, ve S. Mallik, “Socially optimal contracting between a regional blood bank and hospitals,” *Production and Operations Management*, c. 28, 10 2018.
- [75] A. Ekici, O. Örsan Özener, ve E. Çoban, “Blood Supply Chain Management and Future Research Opportunities,” in *Operations Research Applications in Health Care Management*, ser. International Series in Operations Research & Management Science, C. Kahraman and Y. I. Topcu, Eds. Springer, June 2018, c. 0, s. 241–266.
- [76] K. Guo, S. Song, L. Qiu, X. Wang, ve S. Ma, “Prediction of red blood cell demand for pediatric patients using a time-series model: A single-center study in china,” *Frontiers in Medicine*, c. 9, 2022.
- [77] K. Simon, M. Ambroise, ve A. Ramdas, “Analysis of blood and blood com-

ponents wastage in a tertiary care hospital in south india,” *Journal of Current Research in Scientific Medicine*, c. 6, s. 39, 01 2020.

- [78] J. Beliën ve H. C. Force, “Supply chain management of blood products: A literature review,” *Eur. J. Oper. Res.*, c. 217, s. 1–16, 2012.



ÖZGEÇMİŞ

Adı Soyadı :

E-posta :

EĞİTİM BİLGİLERİ:

Lisans : 2017, İstanbul Teknik Üniversitesi, İstanbul, Türkiye ve Southern Illinois University Edwardsville, Illinois, Amerika Birleşik Devletleri (Çift Diploma), Endüstri Mühendisliği

PROFESYONEL TECRÜBE VE ÖDÜLLER:

- Mayıs 2017, Magna Cum Laude
- Nisan 2017, Onur Sertifikası – Güz 2016 Dekan Listesi
- Ağustos 2016, International Undergraduate Scholarship (\$7,500)
- Nisan 2015, Onur Sertifikası – Güz 2014 Dekan Listesi

ERİTROSİT ENVANTER YÖNETİMİNDE DERİN PEKİŞTİRMELİ ÖĞRENME

ORIGINALITY REPORT

2%

SIMILARITY INDEX

2%

INTERNET SOURCES

1%

PUBLICATIONS

1%

STUDENT PAPERS

PRIMARY SOURCES

1	acikerisim.medipol.edu.tr Internet Source	<1%
2	sens.medipol.edu.tr Internet Source	<1%
3	RANDA, Ali Cem, CÖMERT, alican, ADIGÜZEL, Berker, BALIKÇIOĞLU, Cem, ÖRNEKOL, Ceren, BAYINDIR, Pelin and BAKAL, İsmail Serdar. "Türk Kızılayı Orta Anadolu Bölgesi kan ürünleri tedarik zinciri yönetimi projesi", TUBITAK, 2011. Publication	<1%
4	pure.tue.nl Internet Source	<1%
5	acikbilim.yok.gov.tr Internet Source	<1%
6	Submitted to TechKnowledge Student Paper	<1%
7	baadalsg.inflibnet.ac.in Internet Source	<1%