

**TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ**

**SAĞKALIM ANALİZİ İÇİN DERİN ÖĞRENME
YAKLAŞIMI**

İrem KAR

**BİYOİSTATİSTİK ANABİLİM DALI
DOKTORA TEZİ**

DANIŞMAN

Prof. Dr. Atilla Halil ELHAN

II. TEZ DANIŞMANI

Doç. Dr. Erdal COŞGUN

ANKARA

2021

ETİK BEYAN

Ankara Üniversitesi

Sağlık Bilimleri Enstitüsü Müdürlüğü'ne,

Doktora tezi olarak hazırlayıp sunduğum “Sağkalım Analizi için Derin Öğrenme Yaklaşımı” başlıklı tez; bilimsel ahlak ve değerlere uygun olarak tarafımdan yazılmıştır. Tezimin fikir/hipotezi tümüyle tez danışmanım ve bana aittir. Tezde yer alan araştırma tarafımdan yapılmış olup, tüm cümleler, yorumlar bana aittir.

Yukarıda belirtilen hususların doğruluğunu beyan ederim.

Öğrencinin Adı Soyadı: İrem KAR

Tarih:

İmza:

KABUL VE ONAY

Ankara Üniversitesi Sağlık Bilimleri Enstitüsü
Biyoistatistik Anabilim Dalında
İrem KAR tarafından hazırlanan
“Sağkalım Analizi İçin Derin Öğrenme Yaklaşımı” adlı tez çalışması
aşağıdaki jüri tarafından DOKTORA TEZİ olarak OY BİRLİĞİ ile kabul edilmiştir.

Tez Savunma Tarihi:21.06.2021

Prof. Dr. Erdem KARABULUT

Hacettepe Üniversitesi

Jüri Başkanı

Prof. Dr. Atilla Halil ELHAN

Ankara Üniversitesi

Doç. Dr. Serdal Kenan KÖSE

Ankara Üniversitesi

Doç. Dr. Beyza DOĞANAY ERDOĞAN

Ankara Üniversitesi

Dr. Öğr. Üyesi Osman DAĞ

Hacettepe Üniversitesi

Tez hakkında alınan jüri kararı, Ankara Üniversitesi Sağlık Bilimleri Enstitüsü
Yönetim Kurulu tarafından onaylanmıştır.

Prof. Dr. Fügen AKTAN

Sağlık Bilimleri Enstitüsü Müdürü

İÇİNDEKİLER

Etik Beyan	ii
Kabul ve Onay	iii
İçindekiler	iv
Önsöz	vi
Simgeler ve Kısaltmalar	vii
Şekiller	viii
Çizelgeler	ix

1. GİRİŞ	1
1.1. Sağkalım Analizi	2
1.1.1. Sağkalım Analizinde Kullanılan Temel Fonksiyonlar	3
1.2. Sağkalım Analizi Yöntemlerinin Sınıflandırılması	5
1.2.1. Geleneksel İstatistiksel Yöntemler	7
1.2.1.1. Cox Regresyon	8
1.2.2. Makine Öğrenmesi Yöntemleri	9
1.2.2.1. Random Survival Forest	11
1.2.2.2. Derin Öğrenme	12
1.2.2.2.1. Derin Öğrenme Mimarileri	13
1.2.2.2.2. Derin Öğrenme Kütüphaneleri	16
1.2.2.3. Sağkalım Analizinde Derin Öğrenme	18
1.3. Hiperparametreler	20
1.4. Değerlendirme Ölçütü	21

2. GEREÇ VE YÖNTEM	23
2.1. Sanal Makineler (Virtual Machines)	27

3. BULGULAR	29
3.1. Bağımsız Değişkenlere İlişkin Katsayılar $U(-1,1)$ iken Sonuçlar	31
3.1.1. Örneklem Büyüklüğü 100 iken DeepSurv Sonuçları ($U(-1,1)$)	31
3.1.2. Örneklem Büyüklüğü 1000 iken DeepSurv Sonuçları ($U(-1,1)$)	32
3.1.3. Örneklem Büyüklüğü 10000 iken DeepSurv Sonuçları ($U(-1,1)$)	34
3.1.4. Örneklem Büyüklüğü 100 iken RSF Sonuçları ($U(-1,1)$)	35
3.1.5. Örneklem Büyüklüğü 1000 iken RSF Sonuçları ($U(-1,1)$)	36
3.1.6. Örneklem Büyüklüğü 10000 iken RSF Sonuçları ($U(-1,1)$)	37
3.1.7. Örneklem Büyüklüğü 100 iken Cox Regresyon Sonuçları ($U(-1,1)$)	39
3.1.8. Örneklem Büyüklüğü 1000 iken Cox Regresyon Sonuçları ($U(-1,1)$)	41
3.1.9. Örneklem Büyüklüğü 10000 iken Cox Regresyon Sonuçları ($U(-1,1)$)	43
3.1.10. DeepSurv, RSF ve Cox Regresyon Yöntemlerinin Karşılaştırılması ($U(-1,1)$)	44
3.2. Bağımsız Değişkenlere İlişkin Katsayılar $N(0;0,1)$ iken Sonuçlar	48
3.2.1. Örneklem Büyüklüğü 100 iken DeepSurv Sonuçları ($N(0;0,1)$)	48
3.2.2. Örneklem Büyüklüğü 1000 iken DeepSurv Sonuçları ($N(0;0,1)$)	49
3.2.3. Örneklem Büyüklüğü 10000 iken DeepSurv Sonuçları ($N(0;0,1)$)	50
3.2.4. Örneklem Büyüklüğü 100 iken RSF Sonuçları ($N(0;0,1)$)	51
3.2.5. Örneklem Büyüklüğü 1000 iken RSF Sonuçları ($N(0;0,1)$)	52

3.2.6. Örneklem Büyüklüğü 10000 iken RSF Sonuçları ($N(0;0,1)$)	53
3.2.7. Örneklem Büyüklüğü 100 iken Cox Regresyon Sonuçları ($N(0;0,1)$)	54
3.2.8. Örneklem Büyüklüğü 1000 iken Cox Regresyon Sonuçları ($N(0;0,1)$)	56
3.2.9. Örneklem Büyüklüğü 10000 iken Cox Regresyon Sonuçları ($N(0;0,1)$)	57
3.2.10. DeepSurv, RSF ve Cox Regresyon Yöntemlerinin Karşılaştırılması ($N(0;0,1)$)	58
3.3. Evre I Akciğer Adenokarsinomu Veri Seti Üzerinde DeepSurv, RSF ve Cox Regresyon Yöntemlerinin Karşılaştırılması	61
4. TARTIŞMA	64
5. SONUÇ ve ÖNERİLER	70
ÖZET	71
SUMMARY	72
KAYNAKLAR	73
ÖZGEÇMİŞ	78



ÖNSÖZ

Sağlık alanında karmaşık ve yüksek boyutlu verilerin artmasıyla birlikte makine öğrenmesine olan ilgi artmıştır. Derin öğrenme, hiyerarşik yapıları kullanarak verideki yüksek boyutlu soyutlamaları öğrenmeye çalışan bir makine öğrenmesi yöntemidir. Derin öğrenmenin 2012 yılından bu yana hızla yaygınlaşmasında GPU mimarisinin kullanılması büyük bir rol oynamaktadır. Bu tez kapsamında, derin öğrenme yöntemi ile sağkalım analizi uygulanması ve sağkalım verisinde derin öğrenme ile Random Survival Forest ve Cox regresyon yöntemlerinin performanslarının karşılaştırılması amacıyla benzetim çalışması yapılmıştır.

Doktora tez sürecimin tüm aşamalarında bilgi ve tecrübelerinden yararlandığım, mesai saati gözetmeksizin hoşgörü ve sabırla desteğini hiçbir zaman esirgemeyen, kıymetli tez danışmanım Prof. Dr. Atilla Halil ELHAN'a ve ikinci tez danışmanım Doç. Dr. Erdal COŞGUN'a,

Tez izleme komitemde bulunan, öneri ve katkılarıyla tez sürecimde büyük destek olan Prof. Dr. Erdem KARABULUT'a ve Doç. Dr. Beyza DOĞANAY ERDOĞAN'a,

Doktora eğitimim boyunca yardım, bilgi ve tecrübeleriyle bana destek olan Anabilim Dalımız öğretim üyelerine,

Tüm eğitim hayatım boyunca benden maddi ve manevi desteklerini hiçbir zaman esirgemeyen her zaman yanımda olan başarılarımın en büyük kaynağı sevgili aileme sonsuz teşekkürlerimi sunarım.

Bu tez TÜBİTAK 2211-A Genel Yurtiçi Doktora Burs Programı kapsamında desteklenmiştir.

SİMGELER VE KISALTMALAR

AE	AutoEncoders
CNN	Evrişimli Sinir Ağları
DAE	Derin Oto-Kodlayıcılar
DBM	Derin Boltzmann Makineleri
DBN	Deep Belief Networks
DÖ	Derin Öğrenme
DSA	Derin Sinir Ağı
DVM	Destek Vektör Makineleri
EAKA	Eğri Altında Kalan Alan
GPU	Grafik İşlemci Birimleri
KM	Kaplan-Meier
LSTM	Uzun Kısa Süreli Hafıza Ağları
NA	Nelson Aalen
OOB	out-of-bag
RBM	Kısıtlı Boltzmann Makinesi
RF	Random Forest
RNN	Tekrarlayan Sinir Ağları
RSF	Random Survival Forest
VM	Sanal Makineler
YSA	Yapay Sinir Ağları
YT	Yaşam Tablosu

ŞEKİLLER

Şekil 1.1. Teorik Olarak ve Pratikte Sağkalım Eğrileri.	3
Şekil 1.2. Sağkalım Yöntemlerinin Sınıflandırılması.	6
Şekil 1.3. Yapay Zeka, Makine Öğrenmesi ve Derin Öğrenme.	12
Şekil 1.4. RNN Mimarisi.	14
Şekil 1.5. DAE Mimarisi.	15
Şekil 1.6. DeepSurv'ün Temel Bileşenleri.	19
Şekil 2.1. GitHub Ortamı.	27
Şekil 3.1. DeepSurv Çıktısı Örneği.	29
Şekil 3.2. Sağkalım Süresi 60 iken Yöntemlerin Karşılaştırılması ($U(-1,1)$).	45
Şekil 3.3. Sağkalım Süresi 120 iken Yöntemlerin Karşılaştırılması ($U(-1,1)$).	45
Şekil 3.4. Bağımsız Değişken sayısı 10 iken Yöntemlerin Karşılaştırılması ($U(-1,1)$).	46
Şekil 3.5. Bağımsız Değişken Sayısı 20 iken Yöntemlerin Karşılaştırılması ($U(-1,1)$).	48
Şekil 3.6. Sağkalım Süresi 60 iken Yöntemlerin Karşılaştırılması ($N(0;0,1)$).	58
Şekil 3.7. Sağkalım Süresi 120 iken Yöntemlerin Karşılaştırılması ($N(0;0,1)$).	59
Şekil 3.8. Bağımsız Değişken Sayısı 10 iken Yöntemlerin Karşılaştırılması ($N(0;0,1)$).	59

ÇİZELGELER

Çizelge 2.1. Verilerin Türetilmesine İlişkin Senaryolar.	24
Çizelge 2.2. Bağımsız Değişkenlere İlişkin Katsayılar.	25
Çizelge 2.3. Hiperparametreler.	26
Çizelge 2.4. NC Serisi Sanal Makineler.	28
Çizelge 3.1. Her Senaryo için Deepsurv Yönteminin Uygulanmasında Geçen Süre (sn).	30
Çizelge 3.2. Örneklem Büyüklüğü 100 iken Deepsurv Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).	32
Çizelge 3.3. Örneklem Büyüklüğü 1000 iken Deepsurv Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).	33
Çizelge 3.4. Örneklem Büyüklüğü 10000 iken Deepsurv Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).	34
Çizelge 3.5. Örneklem Büyüklüğü 100 iken RSF Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).	35
Çizelge 3.6. Örneklem Büyüklüğü 1000 iken RSF Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).	37
Çizelge 3.7. Örneklem Büyüklüğü 10000 iken RSF Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).	38
Çizelge 3.8. Örneklem Büyüklüğü 100 iken Cox Regresyon Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).	39
Çizelge 3.9. Örneklem Büyüklüğü 100 iken Cox Regresyon Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (NA=0) (U(-1,1)).	40
Çizelge 3.10. Örneklem Büyüklüğü 1000 iken Cox Regresyon Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).	42
Çizelge 3.11. Örneklem Büyüklüğü 10000 iken Cox Regresyon Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).	43
Çizelge 3.12. Örneklem Büyüklüğü 100 iken Deepsurv Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).	49
Çizelge 3.13. Örneklem Büyüklüğü 1000 iken Deepsurv Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).	50

Çizelge 3.14. Örneklem Büyüklüğü 10000 iken DeepSurv Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).	51
Çizelge 3.15. Örneklem Büyüklüğü 100 iken RSF Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).	52
Çizelge 3.16. Örneklem Büyüklüğü 1000 iken RSF Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).	53
Çizelge 3.17. Örneklem Büyüklüğü 10000 iken RSF Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).	54
Çizelge 3.18. Örneklem Büyüklüğü 1000 iken Cox Regresyon Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).	55
Çizelge 3.19. Örneklem Büyüklüğü 1000 iken Cox Regresyon Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).	56
Çizelge 3.20. Örneklem Büyüklüğü 10000 iken Cox Regresyon Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).	57
Çizelge 3.21. Değişkenlere İlişkin Tanımlayıcı İstatistikler.	62
Çizelge 3.22. Evre I Akciğer Adenokarsinomu Veri Seti ile Yapılan Analizler Sonucu DeepSurv, RSF Ve Cox Regresyon Yöntemlerinin Performansının Değerlendirilmesi.	62

1. GİRİŞ

Sağlık alanında, veri kümelerinin boyutunun artması ve derin öğrenmenin yaygınlaşması ile birlikte, makine öğreniminin kullanımına olan ilgi artmaktadır (Gensheimer ve Narasimhan, 2018). Literatürde, hastanın klinik durumundan hastalıkları tahmin etmeye kadar pek çok konuda derin öğrenmenin uygulandığı çalışma vardır. Derin öğrenmenin klinik verilere ilk uygulamaları, özellikle Alzheimer hastalığını tahmin etmek için beyin Manyetik Rezonans Görüntüleme (MRI) taramalarının analizinde görüntü işleme üzerinedir. Gulshan ve ark. (2016), retinal fundus fotoğraflarında diyabetik retinopati ve diyabetik maküler ödemin otomatik belirlenmesi için evrişimli sinir ağlarını kullanmışlar ve yüksek seçicilik ve duyarlılık elde etmişlerdir. Zhang ve ark. (2016) çalışmasında, SWE (shear wave elastography) verisinden elde edilen görüntülerden iyi-kötü huylu tümörlerin tespiti için otomatik özellik çıkarımı amacıyla bir derin öğrenme mimarisi kurmuştur. Kısıtlı Boltzmann Makinesi (RBM) ve point-wise gated Boltzmann Makinesi (PGBM) ile iki katmanlı oluşturulan model 227 SWE (135 benign tümör ve 92 malign tümör) verisiyle eğitilmiş, %93,4 doğruluk, %88,6 duyarlılık, %97,1 seçicilik ve 0,947 eğri altında kalan alan (EAKA) elde edilmiştir.

Son zamanlarda, derin öğrenme yöntemleri, çeşitli uygulama alanlarında sağkalım analizi problemlerini çözme konusunda da büyük ilgi görmüştür. Hastalık risklerini tahmin ederek ve kişiye özgü tedavi önerileri sunarak hekimlere hastalar hakkındaki klinik kararlarında yardımcı olmaları için birkaç derin sağkalım analizi yaklaşımı literatürde önerilmiştir. Bu kapsamda iki yaklaşım vardır. İlki, Cox Orantısal Hazard modelinin uyarlandığı yaklaşım altında Katzman ve ark. (2018) tarafından önerilen “DeepSurv” ve Ching ve ark. (2018) tarafından önerilen “Cox-nnet” yöntemleridir. İkinci yaklaşım ise parametrik sağkalım modellerini kullanan alternatif yöntemlerdir. Bunlar, Gensheimer ve Narasimhan (2018) tarafından önerilen “Nnet-survival”, Giunchiglia ve ark. (2018) tarafından önerilen “RNN-SURV” yöntemleridir.

1.1. Sağkalım Analizi

Tıp alanında sıklıkla kullanılan sağkalım analizi, ilgilenilen sonuç değişkeninin; olayın gerçekleştiği ana kadar geçen süre olduğu veriler için kullanılan istatistiksel yöntemlerin bütünüdür. Burada süre ile, bir bireyin takibinin başlangıcından olayın gerçekleştiği zamana kadar geçen günler, haftalar, aylar veya yıllar kastedilmektedir. Olay ise, ölüm, hastalık, remisyonadaki nöks, iyileşme veya bireyin yaşayabileceği herhangi bir deneyimin başlaması olabilir. Sağkalım analizinde, sağkalım süresi tam bilinmeyen olgular sansürlü (censored) olarak adlandırılır (Kleinbaum ve Klein, 2010).

Sansürün meydana gelmesinin genellikle üç nedeni vardır:

- kişi çalışma süresince olayı yaşamamıştır (administrative censoring).
- kişi çalışma süresinde takip dışı kalmıştır (lost to follow-up).
- kişi ölüm nedeniyle (ölüm ilgilenilen olay değilse) ya da başka bir nedenden (örn., ilacın yan etkisi ya da diğer riskler) dolayı çalışmadan çekilmiştir (withdrawing).

Sağdan sansürlü, soldan sansürlü ve aralık sansür olmak üzere temel olarak üç çeşit sansür vardır. Kişinin gerçek sağkalım süresi; çalışma bittiğinde, kişi takip dışı kaldığında ya da geri çekildiğinde (izleme süresinin sonunda) tamamlanıyorsa bu tür veri “sağdan sansürlü veri” olarak adlandırılır. Bu verilerde gerçek sağkalım süresi (T) gözlenen sağkalım süresine (t) eşit ya da daha uzundur ($T \geq t$). Eğer sağkalım süresi olayın gerçekleşmesi ile sonuçlanıyorsa ancak etkene maruz kalış zamanı bilinmiyorsa sansür soldan gerçekleşebilir. Diğer bir deyişle, kişinin gerçek sağkalım süresi gözlenene eşit ya da daha kısa ise bu durumda veri “soldan sansürlü veri” olarak adlandırılır ($T \leq t$). Eğer olay iki zaman noktası arasında gerçekleşmişse fakat kesin zaman bilinmiyorsa bu sansür “aralık sansür” adını alır ($a < T < b$) (Kleinbaum ve Klein, 2010).

1.1.1. Sağkalım Analizinde Kullanılan Temel Fonksiyonlar

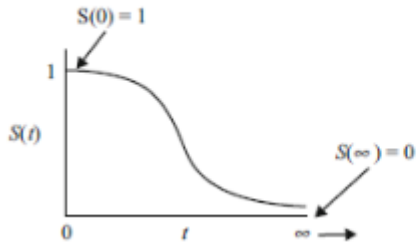
Sağkalım fonksiyonu:

Sağkalım fonksiyonu $S(t)$, bir kişinin belirli bir süreden (t) daha uzun süre hayatta kalma olasılığını verir: yani $S(t)$, rastgele değişken T 'nin t süresini aşma olasılığını verir (Hosmer ve ark., 2002; Kleinbaum ve Klein, 2010).

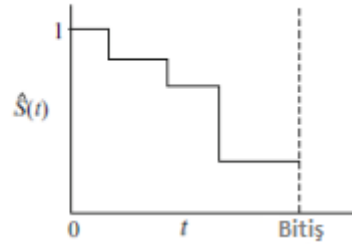
$$S(t) = P(T \geq t)$$

Sağkalım fonksiyonu sağkalım analizi için önemlidir, çünkü farklı t değerleri için sağkalım olasılıklarının elde edilmesi, sağkalım verilerinden önemli özet bilgiler çıkarılmasını sağlar. $t=0$ zamanında, $S(t)=S(0)=1$ 'dir. Yani, çalışmanın başında ($t=0$), henüz ilgilenilen olay kimsede gerçekleşmediğinden, sağ kalma olasılığı birdir. $t=\infty$ zamanında ise, $S(t)=S(\infty)=0$ 'dır. Yani, teorik olarak, çalışma süresi sonsuz olsaydı, çalışma sonunda kimse hayatta kalamazdı (Şekil 1.1) (Hosmer ve ark., 2002; Kleinbaum ve Klein, 2010).

Teorik olarak $S(t)$:



Pratikte $\hat{S}(t)$:



Şekil 1.1. Teorik Olarak ve Pratikte Sağkalım Eğrileri (Kleinbaum ve Klein, 2010).

Kümülatif dağılım fonksiyonu:

İlgilenilen olayın t 'den önce gerçekleşme olasılığını temsil eder ve $F(t)$ ile gösterilir.

$$F(t) = 1 - S(t) = P(T < t)$$

Olasılık yoğunluk fonksiyonu:

$$f(t) = \frac{d}{dt} F(t)$$

Hazard fonksiyonu:

Hazard fonksiyonu $h(t)$, t zamanına kadar sağ kalan bir kişi için birim zaman (Δt) başına ilgilenilen olayın anlık gerçekleşmesi potansiyelini vermektedir (Wang ve ark., 2019).

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t * S(t)} = \frac{f(t)}{S(t)}$$

$f(t)$ aynı zamanda

$$f(t) = -\frac{d}{dt} S(t)$$

olarak da ifade edilebilir. Bu durumda $h(t)$,

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} S(t) * \frac{1}{S(t)} = -\frac{d}{dt} [\ln S(t)]$$

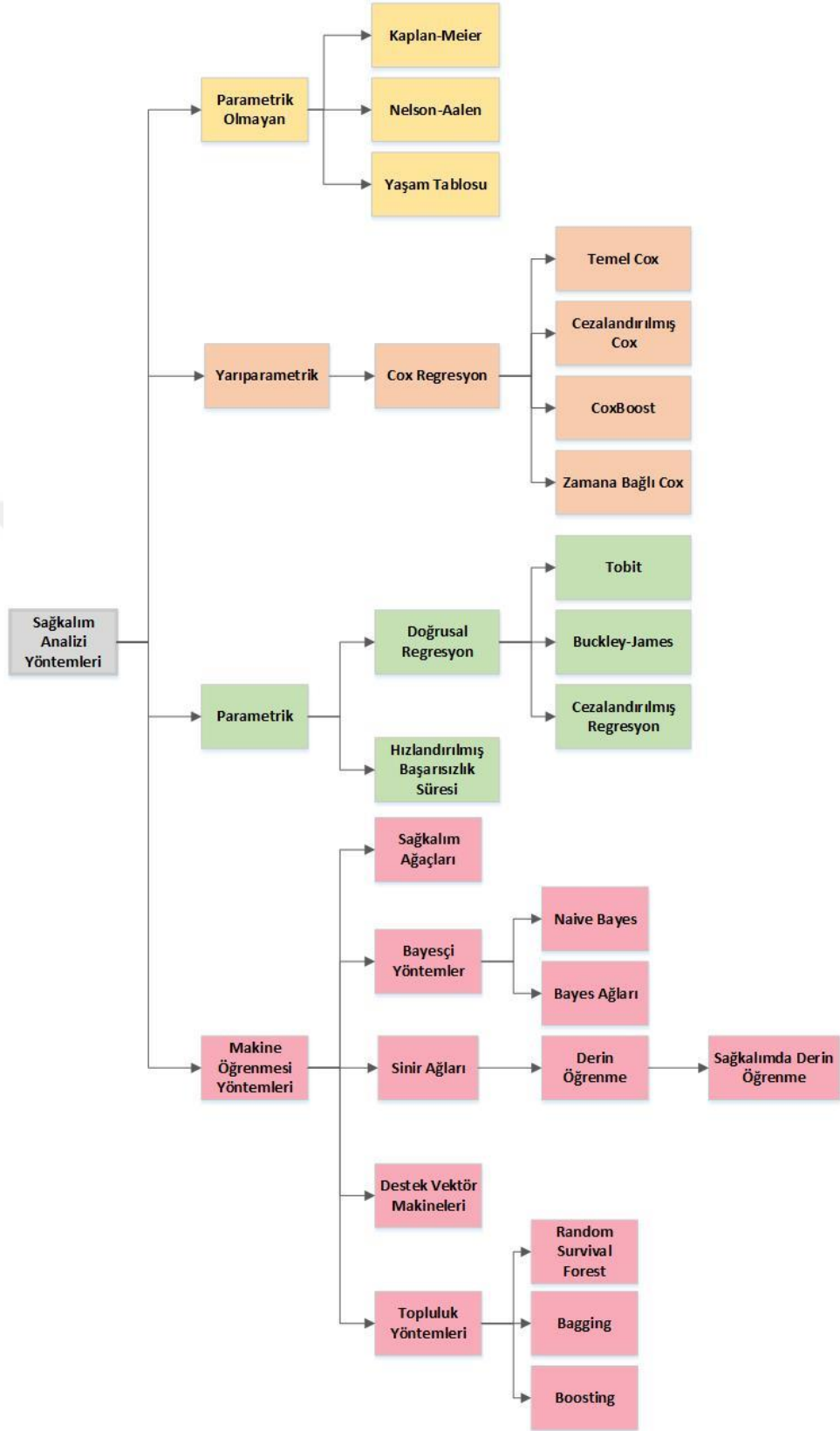
olarak yazılabilir. Ayrıca

$$S(t) = \exp(-H(t))$$

dir. Burada $H(t) = \int_0^t h(u) du$ olmak üzere kümülatif hazard fonksiyonunu göstermektedir (Wang ve ark., 2019).

1.2. Sağkalım Analizi Yöntemlerinin Sınıflandırılması

Genel olarak, sağkalım analizi yöntemleri iki ana kategoride sınıflandırılabilir: istatistiksel yöntemler ve makine öğrenmesine dayalı yöntemler (Şekil 1.2). İstatistiksel yöntemler ve makine öğrenmesine dayalı yöntemler, sağkalım süresini tahmin etmek ve kestirilen sağkalım süresinde sağkalım olasılığını kestirmek gibi ortak bir hedefi paylaşır. Bununla birlikte, istatistiksel yöntemler daha çok, sağkalım eğrilerini tahmin ederek hem sağkalım zamanlarının dağılımlarını hem de parametre tahmininin istatistiksel özelliklerini belirlemeye odaklanırken, makine öğrenmesine dayalı yöntemler, geleneksel sağkalım analizi yöntemlerinin gücünü çeşitli makine öğrenmesi teknikleriyle birleştirerek belirli bir zamanda olayın gerçekleşme olasılığının tahminine odaklanır. Makine öğrenmesi yöntemleri genellikle yüksek boyutlu verilere uygulanırken, istatistiksel yöntemler genellikle düşük boyutlu verilere uygulanır. Sağkalım analizi için makine öğrenmesi yöntemleri, hem istatistiksel yöntemleri hem de makine öğrenmesi yöntemlerini kullanarak sağkalım problemlerini analiz etme yetenekleri ile daha etkili algoritmalar sunar (Wang ve ark., 2019).



Şekil 1.2. Sağkalım Yöntemlerinin Sınıflandırılması.

1.2.1. Geleneksel İstatistiksel Yöntemler

Varsayımlara ve parametrelerin modelde kullanılma şekline bağlı olarak, geleneksel istatistiksel yöntemler üç kategoriye ayrılabilir:

- parametrik olmayan modeller
- yarı parametrik modeller
- parametrik modeller

Sağkalım süresi için herhangi bir dağılım belirlenmediyse veya orantısız hazard varsayımı sağlanmıyorsa parametrik olmayan yöntemler daha etkilidir. Parametrik olmayan yöntemlerde sağkalım fonksiyonunun ampirik bir tahmini, Kaplan-Meier (KM) yöntemi, Nelson-Aalen tahmincisi (NA) veya Yaşam Tablosu (YT) yöntemlerinden biri kullanılarak elde edilir.

Daha genel olarak, belirtilen sağkalım süresindeki sağkalım olasılığı için KM tahmincisi, bir önceki zamana kadar aynı tahmin ile verilen süre için gözlenen sağkalım oranının çarpımı olacaktır. NA yöntemi, modern sayma işlemi tekniklerine dayanan bir tahmin edicidir. YT yöntemi (Cutler ve Ederer, 1958), KM yönteminin aralıklı gruplanmış sağkalım verilerine uygulanmasıdır.

Yarı parametrik kategoride, Cox modeli sağkalım verileri için en sık kullanılan regresyon analizi yaklaşımıdır. Cox modeli, orantısız hazard varsayımı üzerine kurulduğundan ve parametre tahmini için kısmi olabilirlik kullandığından diğer yöntemlerden önemli ölçüde farklıdır. Cox regresyonu, bir parametrik regresyon modeline dayansa bile sonuç değişkeninin dağılımı bilinmediği için yarı parametrik bir yöntem olarak tanımlanmaktadır. Cezalandırılmış Cox modelleri (Zhang ve Lu, 2007), CoxBoost algoritması (Binder ve Schumacher, 2008) ve Zamana Bağlı Cox modeli (TD-Cox) gibi temel Cox modelinin birkaç çeşidi literatürde önerilmiştir (Kleinbaum ve Klein, 2010).

İlgilenilen olayın süresi belirli bir dağılıma sahip olduğunda, parametrik yöntemler tahmin için daha verimli ve doğrudur. Parametrik modellerle ilgilenilen

zamanı tahmin etmek nispeten kolaydır, ancak Cox modeliyle bunu yapmak olanaksızdır. Doğrusal regresyon ana parametrik sağkalım yöntemlerinden biridir, Tobit modeli, Buckley-James regresyon modeli ve cezalandırılmış regresyon, sağkalım analizi için en yaygın olarak kullanılan doğrusal modellerdir. Aynı zamanda değişkenlerin bir fonksiyonu olarak sağkalım süresini modelleyen hızlandırılmış başarısızlık süresi (Accelerated Failure Time - AFT) gibi diğer parametrik modeller de yaygın olarak kullanılmaktadır (Wang ve ark., 2019).

1.2.1.1. Cox Regresyon

Cox modeli (Cox, 1972), birden fazla değişkenin etkisinin eşanlı olarak modellenmesine izin veren bir sağkalım analizi regresyon modelidir (Bradburn ve ark., 2003; Schober ve Vetter, 2018). Sağlık alanında sağkalım verilerini analiz etmek için en yaygın kullanılan çok değişkenli yaklaşımdır. Cox modeli orantısal hazard olarak da ifade edilmektedir (Bradburn ve ark., 2003). Cox modelinin matematiksel olarak ifadesi

$$h(t) = h_0(t)\exp(b_1x_1 + b_2x_2 + \dots + b_px_p)$$

dir. Burada $h_0(t)$ temel hazard fonksiyonunu, $(x_1, x_2, \dots, x_p)$ bağımsız değişkenleri ve $(b_1, b_2, \dots, b_p)$ regresyon katsayılarını göstermektedir (Bradburn ve ark., 2003). Temel hazard fonksiyonu $h_0(t)$ 'nin tanımlanmış belirli bir fonksiyonu olmadığından Cox yarı parametrik bir modeldir (Kleinbaum ve Klein, 2010).

Cox modelinin yaygın olmasının önemli bir nedeni, temel hazard belirtilmese de, regresyon katsayılarının, ilgilenilen hazard oranlarının ve düzeltilmiş sağkalım eğrilerinin çok çeşitli veri yapıları için oldukça iyi tahmin edilebilmesidir. Yani, Cox modeli “sağlam (robust)” bir modeldir (Kleinbaum ve Klein, 2010).

Cox modelinin en önemli varsayımı hazardların orantısallığıdır. Bu varsayım, hazard oranlarının tüm zamanlarda sabit olduğunu ifade eder (Hosmer ve ark., 2002; Kleinbaum ve Klein, 2010).

1.2.2. Makine Öğrenmesi Yöntemleri

Son yıllarda makine öğrenmesi yöntemleri, doğrusal olmayan ilişkileri modelleme yetenekleri ve genel tahminlerinin kalitesi gibi avantajları nedeniyle, pek çok alanda önemli bir yere sahip olmuştur. Sağkalım analizinde, makine öğrenmesi yöntemlerinin karşı karşıya olduğu temel zorluk, sansürlenmiş bilgilerle ve sağkalım süresinin tahminiyle uygun şekilde başa çıkmaktır. Makine öğrenmesi yöntemleri, yüksek boyutlu veriler olduğunda etkilidir, ancak bu sağkalım analizindeki bazı problemler için geçerli değildir (Fadnavis, 2019; Wang ve ark., 2019). Bu bölümde, sağkalım analizinde en sık kullanılan makine öğrenmesi yöntemlerinden bahsedilecektir.

Sağkalım ağaçları (Survival Trees), sansürlü verileri işlemek için özel olarak uyarlanmış sınıflandırma ve regresyon ağaçlarıdır. İlk olarak Gordon ve Olshen tarafından 1985 yılında yayınlanan bir makale ile tanıtılmıştır (Gordon ve Olshen, 1985). Ağaç modellerinin arkasındaki temel, verilerin belirli bir ayırma ölçütüne göre yinelemeli olarak ayrılması ve ilgilenilen olaya dayalı olarak birbirine benzeyen nesnelerin aynı düğüme yerleştirilmesidir (Fadnavis, 2019; Wang ve ark., 2019). Bu yöntemin standart karar ağacına göre avantajı sansürlü verilerle başa çıkabilmesidir (Wang ve ark., 2019). Sansürlü sağkalım verileri için geliştirilen bu yöntem, her düğüm için, tüm özelliklerdeki olası her bölünmeyi inceler ve bir ağaçtaki düğümlerle temsil edilen gruplar arasındaki sağkalım farkını en üst düzeye çıkaran bölünmeyi gerçekleştirir (LeBlanc ve Crowley, 1993).

Makine öğrenmesinde iyi bilinen olasılıksal bir yöntem olan Naive Bayes, basit ama etkili tahmin algoritmalarından biridir. Bununla birlikte, Naive Bayes yönteminin bir dezavantajı, tüm değişkenler arasında bağımsızlık varsayımı

olmasıdır. Bu, sağkalım analizindeki birçok problem için sağlanması pek mümkün olmayan bir varsayımdır (Fadnavis, 2019; Wang ve ark., 2019).

Değişkenlerin çeşitli seviyelerde birbiriyle ilişkili olabileceği bir Bayesian ağı (Bayesian network), bir dizi değişken üzerinde teorik bir dağılımı, grafiksel olarak ifade edebilir. Bayesian ağları, değişkenler arasındaki tüm ilişkilerin görsel olarak yorumlanabilir olmasını sağlar (Fadnavis, 2019; Wang ve ark., 2019).

Denetimli öğrenme yaklaşımı olan Destek Vektör Makineleri (DVM), 1995 yılında Cortes ve Vapnik tarafından geliştirilen, istatistik öğrenme temeline dayalı bir algoritmadır (Cortes ve Vapnik, 1995). DVM, çoğunlukla sınıflandırma için kullanılırken regresyon problemleri için de kullanılabilir. Ayrıca, sağkalım analizi problemleri için de başarıyla adapte edilmiş bir yöntemdir (Fadnavis, 2019; Wang ve ark., 2019).

Biyolojik sinir sistemlerinden esinlenen Frank Rosenblatt, 1958'de yapay sinir ağları (YSA) üzerine ilk makalesini yayınladı (Rosenblatt, 1958). Burada, nöronlar olarak adlandırılan basit yapay düğümler, bir ağ oluşturmak için, biyolojik bir sinir ağını simüle ederek ağırlıklı olarak bir ağa bağlanır. Bu anlamda bir nöron, uyarlanabilir ağırlık kümelerinden oluşan ve belirli bir çeşit aktivasyon fonksiyonuna dayalı çıktı üreten bir hesaplama unsurudur. YSA, sağkalım analizinde yaygın olarak kullanılmaktadır (Fadnavis, 2019; Wang ve ark., 2019).

Derin sinir ağı (DSA - deep neural network), giriş ve çıkış katmanları arasında birden fazla katmana sahip yapay bir sinir ağıdır. DSA, ister doğrusal bir ilişki ister doğrusal olmayan bir ilişki olsun, girdiyi çıktıya dönüştürmek için doğru matematiksel işlemi bulur. Ağ, her çıktının olasılığını hesaplayan katmanlar arasında hareket eder. DSA'lar karmaşık doğrusal olmayan ilişkileri modelleyebilir (Fadnavis, 2019; Wang ve ark., 2019).

Son yıllarda, sansürlü verilere ilişkin tahminlerde bulunmak için birçok ileri düzey makine öğrenmesi yöntemi geliştirilmiştir. Bu yöntemler, diğer yöntemlere kıyasla sağkalım verilerinin analizi için pek çok avantaja sahiptir (Fadnavis, 2019; Wang ve ark., 2019).

Bagging, en eski ve en yaygın kullanılan topluluk yöntemlerinden biridir. Bagging sağkalım ağaçlarında (Bagging Survival Trees) toplanmış (aggregated) sağkalım fonksiyonu, çoğunluk oyu almak yerine tek bir sağkalım ağacının yaptığı tahminleri kullanarak hesaplanabilir (Fadnavis, 2019; Wang ve ark., 2019).

Random Forest (RF), ağaç yapılı modeller kullanarak tahmin yapmak için özel olarak geliştirilmiş bir topluluk yöntemidir (Breiman, 2001). Bagging'e benzer bir çerçeveye dayanmaktadır; aralarındaki temel fark, RF'nin tüm değişkenleri kullanmak yerine belli bir değişken alt kümesini kullanmasıdır. Bu rastgeleliğin ağaçlar arasındaki ilişkiyi azalttığı ve böylece tahmin performansını geliştirdiği gösterilmiştir (Fadnavis, 2019; Wang ve ark., 2019). Random Survival Forest (RSF), Breiman'ın RF yönteminin genişletilmişidir ve tahmin için sağkalım ağaçlarını kullanır (Ishwaran ve ark., 2008).

1.2.2.1. Random Survival Forest

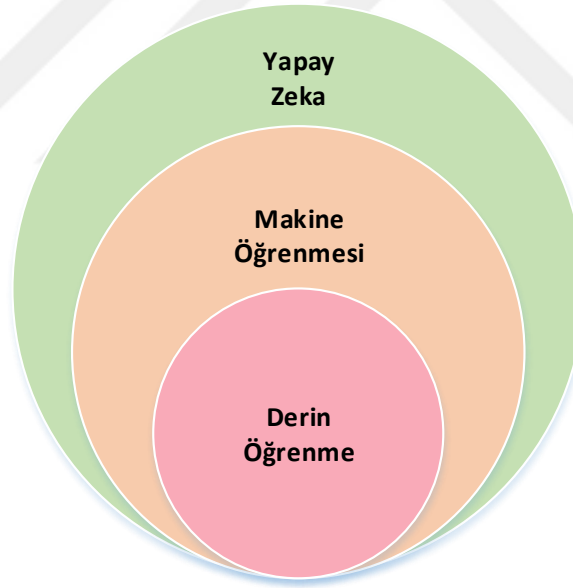
RSF, Breiman'ın RF yönteminin genişletilmişidir ve sağdan sansürlü sağkalım verilerinin analizinde kullanılan bir topluluk yöntemidir. Sağkalım verileri genellikle hazardların orantısallığı varsayımı gibi kısıtlayıcı varsayımları olan yöntemler ile analiz edilir. Ancak, RSF gibi yöntemlerde bu zorluklarla başa çıkılabilir. Bu nedenle sağdan sansürlü sağkalım verileri için genişletilen RF literatüre oldukça yüksek katkı sağlamıştır (Ishwaran ve ark., 2008).

RSF algoritması, orijinal veriden bootstrap ile örneklemeler seçerek başlar. Bu örneklemelerin %37'sini OOB (out-of-bag) verisi olarak ayırır. Her örneklem için sağkalım ağacı oluşturur ve ağacın her bir düğümünde (node) rastgele p tane aday

değişken seçer. Düğüm, alt düğümler arasındaki sağkalım farkını maksimum yapan değişken kullanılarak, her bir terminal düğümünün önceden belirlenmiş ölüm sayısından daha az olmaması koşuluyla, yeni düğümler oluşmayacak şekilde yinelemeli olarak bölünür. Ardından, her ağaç için kümülatif hazard fonksiyonu hesaplanır ve ortalaması alınır. Son olarak, OOB verileri kullanılarak tahmin hatası hesaplanır (Ishwaran ve ark., 2008).

1.2.2.2. Derin Öğrenme

Derin öğrenme (DÖ), nesne algılama, konuşma tanıma, dil çevirisi gibi konularda yüksek doğruluğu sağlamak için çok katmanlı yapay sinir ağlarını kullanan yapay zeka ve makine öğrenmesinin bir alt kümesidir (Şekil 1.3).



Şekil 1.3. Yapay Zeka, Makine Öğrenmesi ve Derin Öğrenme.

DÖ, özellik çıkarma ve dönüştürme için birçok doğrusal olmayan işlem birimi katmanını kullanır. Her ardışık katman, önceki katmandaki çıktıyı girdi olarak alır. Algoritmalar denetimli veya denetimsiz olabilir (Şeker ve ark., 2017).

DÖ'nün 2012 yılından bu yana hızla yaygınlaşmasında ve görüntü tespitindeki doğruluk oranının artmasında grafik işlemci birimi (Graphics Processing Unit – GPU) mimarisinin kullanılması büyük bir rol oynamaktadır (GPU mimarileri CPU mimarilerine göre daha fazla çekirdek barındırdığından paralel işlem bakımından üstünlük sağlamaktadır) (François, 2017).

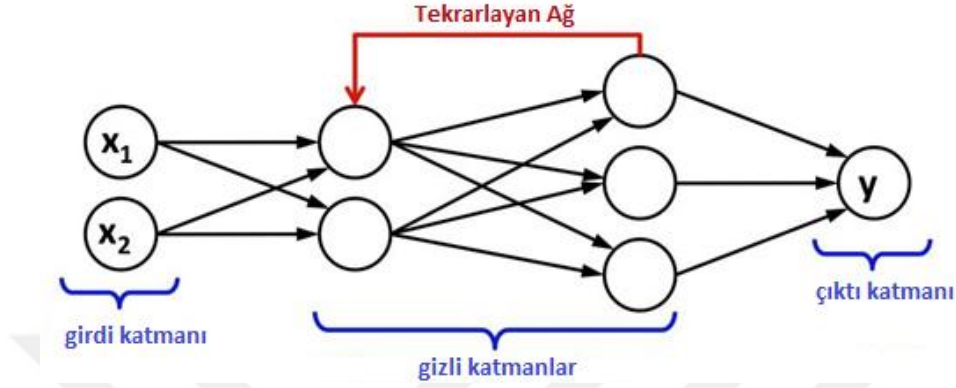
DÖ, yüksek boyutlu verilerdeki karmaşık yapıları keşfetmede çok başarılıdır ve pek çok alanda uygulanmaktadır. Görüntü tanıma ve konuşma tanımanın yanı sıra, potansiyel ilaç moleküllerinin aktivitesini tahmin etme, partikül hızlandırıcı verileri analiz etme, beyin devrelerini yeniden yapılandırma ve gen ekspresyonunda diğer makine öğrenmesi yöntemlerinden daha iyi performans göstermiştir (Miotto ve ark., 2018).

1.2.2.2.1. Derin Öğrenme Mimarileri

1988 yılında Yann LeCun ve ark. tarafından ortaya atılan Evrişimli Sinir Ağları (Convolutional Neural Network - CNN), çok sayıda gizli katman kullanarak görüntülerdeki farklı özelliklerin algılanmasını sağlamaktadır (LeCun ve ark., 1998). CNN ile görüntü sınıflama, nesne tanımlama, görüntü segmentasyonu gibi işlemler başarılı bir şekilde yapılmaktadır (Kayaalp ve Süzen, 2018; Şeker ve ark., 2017). Genellikle bir CNN, evrişimli katmanlar, ortaklama (pooling) katmanları ve tamamen bağlı katmanlar olmak üzere üç ana sinir katmanından oluşur. Her katman türünün farklı işlevi vardır.

Tekrarlayan Sinir Ağları (Recurrent Neural Network - RNN), ardışık bilgileri kullanan, gizli katman çıkışını tekrar aynı katmana giriş olarak gönderebilen bir DÖ algoritmasıdır (Şekil 1.4) (Kayaalp ve Süzen, 2018; Şeker ve ark., 2017). RNN'nin temelleri 1990 yılında Jeff Elman tarafından basit tekrarlayan sinir ağı oluşturularak atılmıştır (Elman, 1990). Konuşma tanıma, dil modelleme, çeviri gibi pek çok alanda başarılı bir şekilde uygulanmaktadır. RNN mimarisinin tekrarlayan olarak adlandırılmasının sebebi, bir dizinin her ögesi için aynı görevi önceki çıktılara bağlı

olarak yerine getirmesidir. Ancak, önceki bilgiye dayalı bu yaklaşım, yani uzun vadeli bağımlılıkları sağlamak teoride mümkün olsa da pratikte bazı sorunları beraberinde getirmiştir.

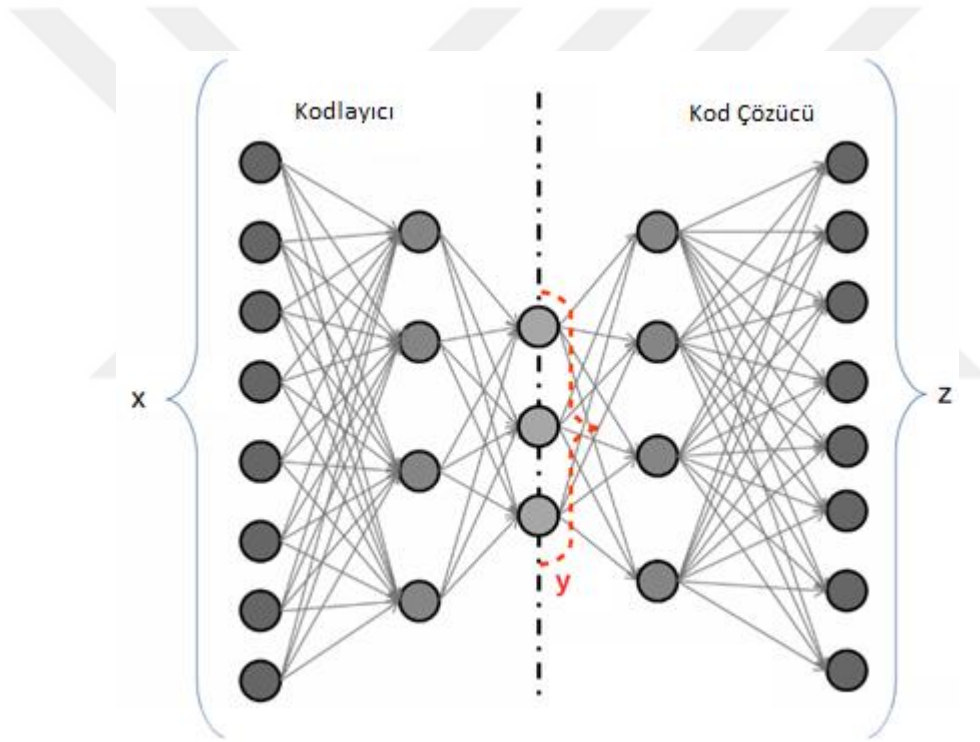


Şekil 1.4. RNN Mimarisi.

Uzun Kısa Süreli Hafıza Ağları (Long Short Term Memory – LSTM), Hochreiter ve Schmidhuber tarafından 1997 yılında tanıtılmıştır (Hochreiter ve Schmidhuber, 1997). RNN mimarisindeki uzun süreli bağımlılık sorununa çözüm olarak sunulan özel bir RNN türüdür. LSTM mimarisi bir girdi katmanı, bir tekrarlanan gizli katman ve bir çıktı katmanından oluşur. Geleneksel sinir ağından farklı olarak, gizli katmanın temel birimi bellek bloğudur. Bellek bloğu, zamansal durumu ezberleyen kendinden bağlantılı bellek hücreleri ve bloktaki bilgi akışını kontrol etmek için bir çift uyarlamalı, çarpımsal kapı ünitesi içerir. Giriş kapısı ve çıkış kapısı olarak adlandırılan iki ek kapı sırasıyla bloğa giriş ve çıkış aktivasyonlarını kontrol eder (Ma ve ark., 2015). Unutma kapısında ise LSTM kendisine gelen bilgilerden hangilerini sileceğine sigmoid fonksiyonu ile karar verir. Sigmoid fonksiyonu aldığı bilgileri 0 ile 1 arasındaki değerlere dönüştürür. Burada 0, bu bilgilerin tamamen unutulması, 1 ise bilgiler üzerinde herhangi bir değişiklik yapmadan bilgilerin iletilmesi anlamına gelmektedir (Phi, 2018).

Derin Oto-Kodlayıcılar (Deep AutoEncoders - DAE), 1980’li yıllarda Hinton ve PDP grup tarafından literatüre kazandırılan, denetimsiz öğrenme için kullanılan özel bir yapay sinir ağıdır (Kayaalp ve Süzen, 2018). AutoEncoders (AE), girdi

katmanındaki deęerleri çıktı katmanına kopyalayan ileri beslemeli bir sinir ağıdır. DAE ise, her bir katmandaki çıktıların ardışık katmanın girişlerine bağlandığı AE'lerin çok katmanlarından oluşan sinir ağıdır. DAE'de amaç, verilen girdiye en çok benzeyen çıktıyı üretecek olan veriye ait önemli yapısal bilgileri içeren fonksiyonun bulunmasıdır (Kayaalp ve Süzen, 2018; Şeker ve ark., 2017). Şekil 1.5'te DAE'nin yapısı verilmektedir. Burada girdi katmanı ile gizli katman arasında kodlayıcı (encoder) denir ve boyut indirgemeyi sağlar. Gizli katman ile çıktı katmanı arasında ise kod çözücü (decoder) denir ve gizli katmanlardan geçen kodun çözülerek çıktı olarak verilmesini sağlar. Yani kod çözücü, sıkıştırılmış gizli katmanın boyutunu arttırarak girdiyi yeniden inşa eder (Vachhani ve ark., 2017).



Şekil 1.5. DAE Mimarisi (Vachhani ve ark., 2017).

Deep Belief Networks (DBN) ve Derin Boltzmann Makineleri (Deep Boltzmann Machines – DBM), “Boltzmann ailesine” ait derin öğrenme modelleridir. Kısıtlı Boltzmann Makinesi'ni (Restricted Boltzmann Machine – RBM) öğrenme modülü olarak kullanırlar. RBM üretken stokastik bir sinir ağıdır. DBN'de katmanlara yönlendirilmiş bağlantılar var iken, DBM ağı tüm katmanları arasındaki yönlendirmeleri kaldırmıştır (Voulodimos ve ark., 2018).

1.2.2.2.2. Derin Öğrenme Kütüphaneleri

Makine öğrenmesi yöntemlerinin özellikle son yıllarda yaygın şekilde kullanılmasının nedeni artan veri büyüklüğünün yanı sıra bilgisayar donanımında yaşanan gelişmelerdir. Yani, paralel işlemeyi daha hızlı, daha ucuz ve daha güçlü hale getiren GPU'ların kullanılabilirliğinin artmasıdır. GPU'lar eğitim hızını önemli ölçüde artırabilir, çünkü sinir ağlarının eğitilme şekli mimarilerine iyi uyum sağlar. GPU'lardaki çok sayıda işlem birimi ve yüksek bellek bant genişliği sayesinde hızlı vektör ve matris çarpımları gerçekleştirilebilir (Zou ve ark., 2019). 2000'li yıllar boyunca NVIDIA ve AMD gibi şirketler, giderek daha gerçekçi hale gelen video oyunlarının grafiklerini güçlendirmek amacıyla, GPU geliştirmek için milyarlarca dolar yatırım yapmışlardır. Bu yatırımla, 2007 yılında NVIDIA, GPU'ları için bir programlama arayüzü olan CUDA'yı piyasaya sürdüğünde bilim dünyasına da fayda sağlamıştır (François, 2017).

DÖ kütüphaneleri, ileri düzey bir programlama arayüzü aracılığıyla derin sinir ağlarını tasarlamak, eğitmek ve geçerliğini doğrulamak için yapı taşlarıdır (Nvidia, Erişim Tarihi: 10/02/2021). MxNet, Pytorch, TensorFlow, Theano, CNTK, Caffe, Torch, Neon, Pylearn vb. gibi çeşitli DÖ kütüphaneleri vardır. Bu kütüphanelerin her birinin farklı özelliği ve performans karakteristiği vardır (Shatnawi ve ark., 2018). DÖ için en popüler olan üç kütüphane TensorFlow, CNTK ve Theano'dur. TensorFlow Google tarafından, CNTK ise Microsoft tarafından geliştirilmiştir. Theano, Université de Montréal'deki MILA laboratuvarı tarafından geliştirilmiştir. Bu kütüphaneler çok çekirdekli işlemcileri (CPU) ve birden çok GPU'yu destekler (François, 2017; Shatnawi ve ark., 2018).

TensorFlow, mükemmel performans elde etmek için tüm sayısal hesaplamaları ifade etmede tek bir veri akış grafiği kullanan, Google tarafından geliştirilen açık kaynaklı bir kütüphanedir (Abadi ve ark., 2016). TensorFlow'un esnek mimarisi, geliştiricilerin ve kullanıcıların çok çeşitli derin sinir ağı modellerini denemelerine ve eğitmelerine olanak sağlar. TensorFlow, grafik oluşturmak için Python, C++ ve Java gibi çeşitli dillerin farklı API'lerini kullanır (François, 2017; Shatnawi ve ark., 2018).

Keras, Python'da geliştirilen açık kaynaklı bir DÖ kütüphanesidir. CNTK, Theano veya TensorFlow kütüphanelerinin üstünde çalışır. Keras, Google mühendisi Chollet tarafından 2015 yılında ONEIROS (Açık uçlu Nöro-Elektronik Akıllı Robot İşletim Sistemi) araştırma projesinin bir parçası olarak kurulmuştur. Keras, derin sinir ağları ile hızlı ifade, kolay ve hızlı prototipleme (modülerlik ve genişletilebilirlik) sağlayacak şekilde tasarlanmıştır (François, 2017; Shatnawi ve ark., 2018).

Caffe, Berkeley Vision and Learning Center tarafından geliştirilen açık kaynaklı bir DÖ kütüphanesidir (Jia ve ark., 2014). Modüler ve hızlıdır, birden fazla GPU'yu destekler. Caffe, C++, Python ile yazılabilen ve Ubuntu, OS X, resmi olmayan Android portu ve diğer platformları destekleyen çapraz bir platformdur. Birçok uzantı aktif olarak eklenebilmektedir ve görüntü sınıflama için kullanılan en popüler araçtır. Caffe'nin avantajı, kodu yazmak yerine, modellerin ve optimizasyonların düz metin şemaları olarak tanımlanmasıdır (Parvat ve ark., 2017).

Microsoft Bilişsel Araç Seti (CNTK), Microsoft tarafından birden çok GPU veya sunucuda çeşitli tipteki sinir ağlarını eğitmek ve test etmek için geliştirilen açık kaynaklı bir DÖ kütüphanesidir (Agarwal ve ark., 2014). CNTK farklı DÖ mimarilerini destekler. CNTK arayüzü, hem GPU (CUDA) hem de CPU'da, Python, C++ ve C# gibi çeşitli dillerin farklı API'lerini destekler (François, 2017; Shatnawi ve ark., 2018).

Theano, Montreal Üniversitesi'ndeki MILA laboratuvarında, kullanıcıların ve geliştiricilerin NumPy'nin sözdizimini (syntax) (büyük ve çok boyutlu dizileri destekleyen bir Python kütüphanesi) kullanarak ifadelerini optimize etmelerini ve değerlendirmelerini sağlayan matematiksel ifadeler için bir derleyici olarak geliştirilen açık kaynaklı bir Python kütüphanesidir (Team ve ark., 2016). Theano, hesaplama seçimini optimize ederek otomatik olarak hesaplama yapmaya başlar, bunları C++ veya CUDA (GPU için) gibi diğer makine öğrenme dillerine çevirir ve daha sonra CPU'lar veya GPU'larda verimli bir şekilde Python modüllerine derler (François, 2017; Shatnawi ve ark., 2018).

1.2.2.3. Sağkalım Analizinde Derin Öğrenme

Sağkalım analizi problemi makine öğrenmesi alanında gerçekleşen RSF, sağkalım analizi için çoklu görev (Multi-task) öğrenme, sansürlü veriler için destek vektör makinelerinin kullanılması gibi yeniliklerle birlikte makine öğrenmesi alanında çalışan pek çok bilim adamının dikkatini çekmiştir. Ancak, bu yöntemlerin hiçbiri sinir ağı yaklaşımına dayanmamaktadır (Giunchiglia ve ark., 2018).

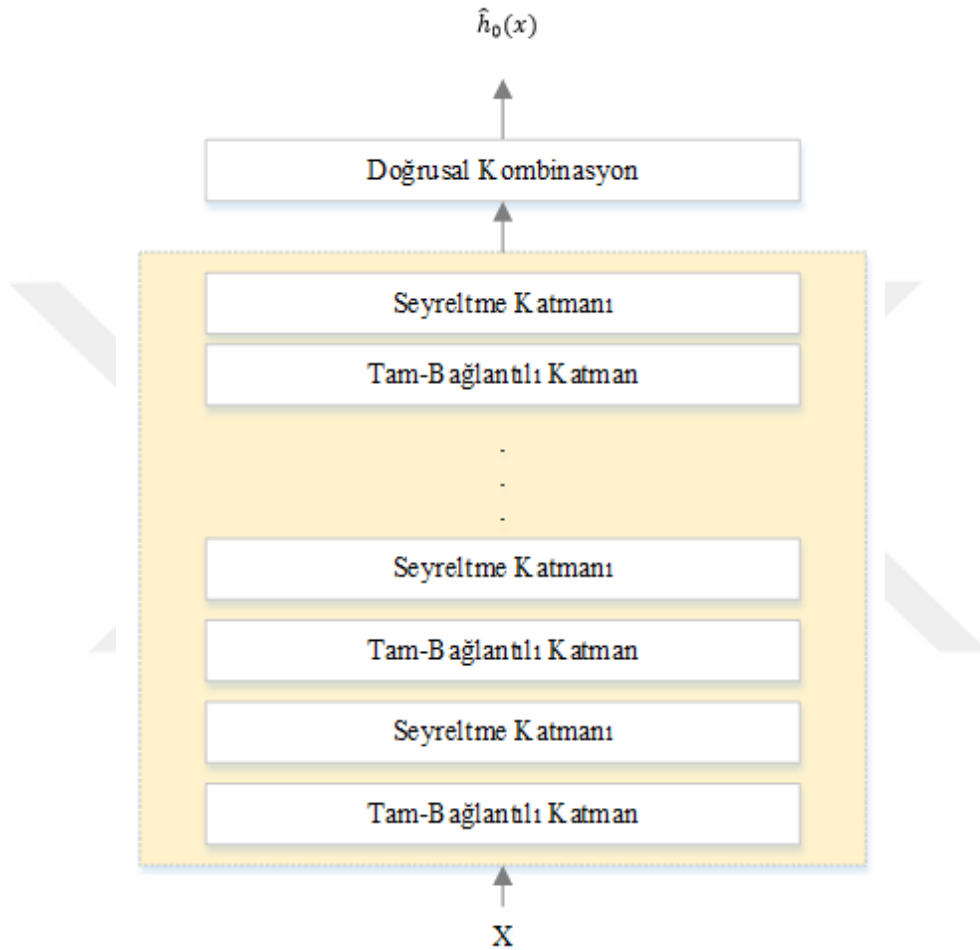
Sağkalım kestirimi için ilk yapay sinir ağı modeli 1995'te Faraggi ve Simon tarafından, dört girdi parametresine sahip prostat kanseri sağkalımını modellemek için geliştirilmiştir (Faraggi ve Simon, 1995). Ancak, bu model, girdi sayısının fazla olduğu veri setleri için uygun değildi (Ching ve ark., 2018).

1990'ların ortalarından itibaren ise sağkalım analizi için DÖ yöntemleri kullanılmaya başlanmıştır. DÖ'deki son gelişmeler de göz önüne alındığında, sağkalım analizi için DÖ'nün kullanılmasında oldukça başarılı sonuçlar elde edilmiştir. DÖ yöntemleri, ortak değişkenler (covariates) ve hazardlar arasında karmaşık bir ilişki olduğunda, diğer sağkalım modellerinden daha iyi performans göstermiştir. Ancak sağkalım analizi verisine derin öğrenme yaklaşımı hala çok yeni ve araştırılması gereken bir konudur.

Sağkalım analizi için derin öğrenme yaklaşımında iki kapsamlı yaklaşım söz konusudur. Bunlardan ilki Cox Orantısız Hazard modelinin uyarlanması, ikincisi ise tamamen parametrik sağkalım modellerini kullanan alternatif yöntemlerdir. Literatürde yer alan DeepSurv (Katzman ve ark., 2018) ve Cox-nnet (Ching ve ark., 2018), Cox Orantısız Hazard modelinin uyarlanması yaklaşımı altında; Nnet-survival (Gensheimer ve Narasimhan, 2018) ve RNN-SURV (Giunchiglia ve ark., 2018) ise ikinci yaklaşım altındadır.

İlk olarak 2018 yılında Katzman ve ark. tarafından ortak değişkenler ile risk fonksiyonu arasında doğrusal olmayan ilişkileri modellemek için “DeepSurv” olarak

adlandırılan derin tam bağlantılı ağ yapısı önerildi (Katzman ve ark., 2018). DeepSurv, ortak değişkenlerin, ağın ağırlıkları (θ) ile parametrelendirilen hazard oranları üzerindeki etkilerini tahmin eden derin ileri beslemeli bir sinir ağıdır. Şekil 1.6. DeepSurv'ün temel bileşenlerini göstermektedir.



Şekil 1.6. DeepSurv'ün Temel Bileşenleri (Katzman ve ark., 2018).

Transkriptomik verilerinden hasta prognozunu tahmin etmek için Ching ve ark. (2018) tarafından Cox-nnet adında bir sinir ağı geliştirilmiştir. Bu, Cox regresyon modelinin bir uzantısıdır (Ching ve ark., 2018). Cox-nnet ve DeepSurv birbirine çok benzer yöntemler olup, her ikisi için de hazardların orantısallığı varsayımı söz konusudur ve çıktı olarak sağkalım olasılığı yerine hazard oranı sunarlar (Zhao ve Feng, 2019).

Gensheimer ve Narasimhan (2018) tarafından önerilen Nnet-survival, sinir ağıları için sağkalım süresinin kesikli sayısal olduğu bir sağkalım modelidir. Nnet-survival, çok katmanlı algılayıcılar ve evrişimli sinir ağıları dahil olmak üzere çeşitli sinir ağı mimarilerine uygulanabilir (Gensheimer ve Narasimhan, 2018).

Giunchiglia ve ark. (2018) tarafından önerilen RNN-SURV ise, sansürlü verilerin varlığında her hasta için kişiselleştirilmiş bir risk skoru ve sağkalım olasılığı fonksiyonunu tahmin etmek için geliştirilen bir tekrarlayan sinir ağı modelidir (Ching ve ark., 2018).

1.3. Hiperparametreler

Çoğu makine öğrenmesi algoritmasında, modeli tasarlayan kişinin belirlemesi gereken bazı parametreler yer alır. Hiperparametre olarak adlandırılan bu parametrelerin değerinin ne olacağı genellikle başlangıçta net değildir, çünkü ilgilenilen probleme, veri setine vb. göre farklılık gösterir. Hiperparametrelerin seçimi genelde tasarımcının deneyimlerine, güncel eğilimlere vb. göre değişmektedir.

Sağkalım analizinde derin öğrenme için kullanılan hiperparametreler; epoch sayısı, L2 düzenlileştirmesi, seyreltme (dropout), gizli katman sayısı, öğrenme oranı, öğrenme oranı azalımı (learning rate decay), momentum ve aktivasyon fonksiyonudur.

Epoch sayısı, eğitim verisinin katmanlardan kaç kez geçeceği. Epoch sayısı arttıkça modelin başarısı artar.

L2 düzenlileştirmesi, aşırı öğrenmeyi önlemek için kullanılan bir parametredir. Bunun için cezalandırma fonksiyonları kullanılarak modelin ağırlıklarının sıfır ya da sıfıra yakın olması sağlanır. Cezalandırma işleminde sıklıkla L1 ve L2 fonksiyonları

kullanılır. L1 fonksiyonunda cezalandırma ağırlığın mutlak değeri, L2 fonksiyonunda ise ağırlığın karesi alınarak yapılır.

Seyreltme, katmanlardaki düğümlerin azaltılmasını ifade etmektedir. Aşırı öğrenmeyi önlemek için kullanılan bir düzenleştirme yöntemidir.

DÖ'yü diğer sinir ağı modellerinden ayıran en önemli özellik katman sayısıdır. Bu sayede DÖ karmaşık problemleri çözmede oldukça başarılıdır. Giriş ve çıkış katmanları arasındaki katman gizli katmandır.

Öğrenme oranı, bir ağı parametrelerini güncelleme hızıdır. Öğrenme oranı azalımı ise, ağırlık güncelleme kuralında planlanan farklı bir güncelleme yoksa ağırlıkların üssel olarak sifıra yakınsamasını sağlayan ek bir parametredir.

Momentum, önceki adımların bilgisi ile bir sonraki adımı tahmin etmeye yarar. Momentum değeri genellikle 0,8-0,9 olarak tercih edilir.

DÖ ile genellikle doğrusal olmayan problemler çözülmeye çalışılır. Aktivasyon fonksiyonu, bu doğrusal olmamayı DÖ modellerine tanıtmak için kullanılır. Sigmoid, tanh, ReLU (Rectified Linear Unit), PReLU (Parametric Rectified Linear Unit) gibi çeşitli aktivasyon fonksiyonları vardır. Bunlardan, en yaygın olarak kullanılanı ReLU aktivasyon fonksiyonudur.

1.4. Değerlendirme Ölçütü

Hata kareler ortalamasının karekökü (RMSE) ve R^2 gibi regresyonda kullanılan standart değerlendirme ölçütleri, sağkalım verilerindeki sansürlü gözlemler nedeniyle, sağkalım analizinin performansını ölçmek için uygun değildir. Bunun yerine, sağkalım analizindeki tahmin performansını değerlendirmede daha özel ölçütlerin kullanılması gerekir (Wang ve ark., 2019).

Sağkalım analizini değerlendirmede kullanılan temel ölçütlerden biri uyum indeksidir (C-indeksi - Concordance index). Bu çalışmada tahmin başarısı Harrell'in C-indeksi kullanılarak değerlendirildi (Harrell ve ark., 1982). C-indeks, sansürlü verileri dikkate alabilen, alıcı işlem karakteristiği eğrisi altında kalan alanın (EAKA) genelleştirilmesidir. C-indeks aşağıdaki adımlar kullanılarak hesaplanır (Ishwaran ve ark., 2008):

1. Veriler üzerinden olası tüm çiftler oluşturulur.
2. Oluşturulan çiftlerden sağkalım süresi daha kısa olan sansürlü veri ise o çift çıkarılır. Ayrıca, çiftlerden en az birinde ilgilenilen olay gerçekleşmemişse ve $T_i = T_j$ ise o çift çıkarılır. Böylece olası (izin verilebilir) çiftlerin sayısı belirlenir.
3. $T_i = T_j$ olan her olası çift için, çiftlerden sağkalım süresi daha kısa olanın tahmin edilen sağkalım süresi de daha kısaysa 1, tahmin edilen sonuçlar eşit ise 0,5 olarak hesaplanır. $T_i = T_j$ ve her ikisinde de olayın gerçekleştiği her olası çift için, tahmin edilen sonuçlar eşit ise 1, aksi halde 0,5 olarak hesaplanır. $T_i = T_j$ olan ve her ikisinde de olayın gerçekleşmediği her olası çift için, olayın gerçekleştiği kişide tahmin edilen sağkalım süresi daha kısa ise 1, aksi takdirde 0,5 olarak hesaplanır.
4. Sonuç olarak C-indeksi, $C\text{-indeks} = \text{Uyumlu Çiftler} / \text{Olası Çiftler}$ olarak hesaplanır.

EAKA'ya benzer şekilde, C-indeksi için 1 en iyi model tahminini, 0,5 ise rastgele bir tahmini temsil eder (Fotso, 2018).

Bu çalışmada, derin öğrenme yöntemi ile sağkalım analizi uygulanması ve sağkalım verisinde derin öğrenme ile makine öğrenmesi yöntemlerinden RSF ve klasik istatistiksel yöntemlerden Cox regresyonun performanslarının karşılaştırılması amaçlanmaktadır. Yöntemlerin başarısı C-indeks ile değerlendirilecek ve karşılaştırmalar bu ölçüte göre yapılacaktır.

2. GEREÇ VE YÖNTEM

Bu tez çalışması kapsamında derin öğrenme ile sağkalım analizinin uygulanması ve sağkalım verisinde derin öğrenme yaklaşımlarından DeepSurv ile makine öğrenmesi yöntemlerinden RSF ve klasik istatistiksel yöntemlerden Cox regresyon yöntemlerinin performanslarının karşılaştırılması amacıyla türetilmiş (simulated) veri setleri kullanılarak bir benzetim çalışması gerçekleştirilmiştir. Veri setlerinin türetilmesinde,

- örneklem büyüklüğü (n) 100, 1000 ve 10000;
- sağkalım süresi (T) 60 ve 120;
- bağımsız değişken sayısı (X) 10 ve 20;
- sansür oranı 0,10, 0,25, 0,50, 0,75 ve 0,90

olarak belirlenmiş ve 60 farklı senaryo oluşturulmuştur. Oluşturulan her senaryo ise 100 kez tekrarlanmıştır. Bu senaryolara ilişkin bilgiler Çizelge 2.1’de sunulmuştur.

Bağımsız değişkenler standart normal dağılımdan ($X_i \sim N(0,1)$), bu bağımsız değişkenlere ilişkin katsayılar ise Tekdüze dağılımdan ($\beta_i \sim U(-1,1)$) türetilmiştir.

Belirlenen 60 senaryoya ek olarak, bağımsız değişken katsayılarına ilişkin ön belirlemenin yapılmadığı, kullanılan R paketindeki varsayılan haliyle (default) bırakıldığı 60 ayrı senaryo daha oluşturulmuştur. Bu senaryolarda katsayılar, 0 ortalama ve 0,1 standart sapma ile normal dağılımdan türetilmektedir.

Veriler eğitim ve validasyon seti olmak üzere rastgele iki ayrı veri setine ayrılmıştır. DeepSurv, RSF ve Cox regresyon yöntemleri için aynı eğitim ile validasyon setleri kullanılmış olup, validasyon oranları örneklem büyüklüğü 100 iken 0,20; 1000 ve 10000 iken ise 0,10 olarak alınmıştır.

Çizelge 2.1. Verilerin Türetilmesine İlişkin Senaryolar.

Senaryo	n	X	T	Sansür Oranı
1	100	10	60	0,1
2				0,25
3				0,5
4				0,75
5				0,9
6		120	120	0,1
7				0,25
8				0,5
9				0,75
10				0,9
11	1000	20	60	0,1
12				0,25
13				0,5
14				0,75
15				0,9
16		120	120	0,1
17				0,25
18				0,5
19				0,75
20				0,9
21	10000	10	60	0,1
22				0,25
23				0,5
24				0,75
25				0,9
26		120	120	0,1
27				0,25
28				0,5
29				0,75
30				0,9
31	10000	20	60	0,1
32				0,25
33				0,5
34				0,75
35				0,9
36		120	120	0,1
37				0,25
38				0,5
39				0,75
40				0,9
41	10000	10	60	0,1
42				0,25
43				0,5
44				0,75
45				0,9
46		120	120	0,1
47				0,25
48				0,5
49				0,75
50				0,9
51	10000	20	60	0,1
52				0,25
53				0,5
54				0,75
55				0,9
56		120	120	0,1
57				0,25
58				0,5
59				0,75
60				0,9

Bağımsız değişken 10 ve 20 olduğu durumda kullanılan katsayılar Çizelge 2.2.'de verilmiştir.

Çizelge 2.2. Bağımsız Değişkenlere İlişkin Katsayılar.

X=10		X=20			
β_1	-0,94	β_1	-0,99	β_{11}	0,33
β_2	-0,82	β_2	-0,88	β_{12}	0,35
β_3	-0,62	β_3	-0,76	β_{13}	0,36
β_4	-0,47	β_4	-0,75	β_{14}	0,44
β_5	-0,17	β_5	-0,45	β_{15}	0,52
β_6	0,22	β_6	-0,42	β_{16}	0,55
β_7	0,36	β_7	-0,35	β_{17}	0,61
β_8	0,52	β_8	-0,23	β_{18}	0,68
β_9	0,88	β_9	0,17	β_{19}	0,82
β_{10}	0,99	β_{10}	0,24	β_{20}	0,84

Ayrıca, çalışmaya Ocak 2007 ve Aralık 2015 tarihleri arasında Ankara Üniversitesi Göğüs Cerrahisi Anabilim Dalında Evre I akciğer adenokarsinomu nedeniyle opere edilen 165 hasta verisi dahil edilmiştir. Neoadjuvan kemoterapi veya radyoterapi alan hastalar çalışma dışı bırakılmıştır. Eksik verisi olan 12 hasta verisi çıkarılmış ve toplam 153 hasta verisi çalışmada kullanılmıştır.

RSF'nin uygulanmasında ağaç sayısı 1000, düğüm (node) sayısı 15 ve bir düğümü bölmek için aday olarak rastgele seçilen değişkenlerin sayısı (mtry) ise bağımsız değişken sayısının üçte biri olarak belirlenmiştir.

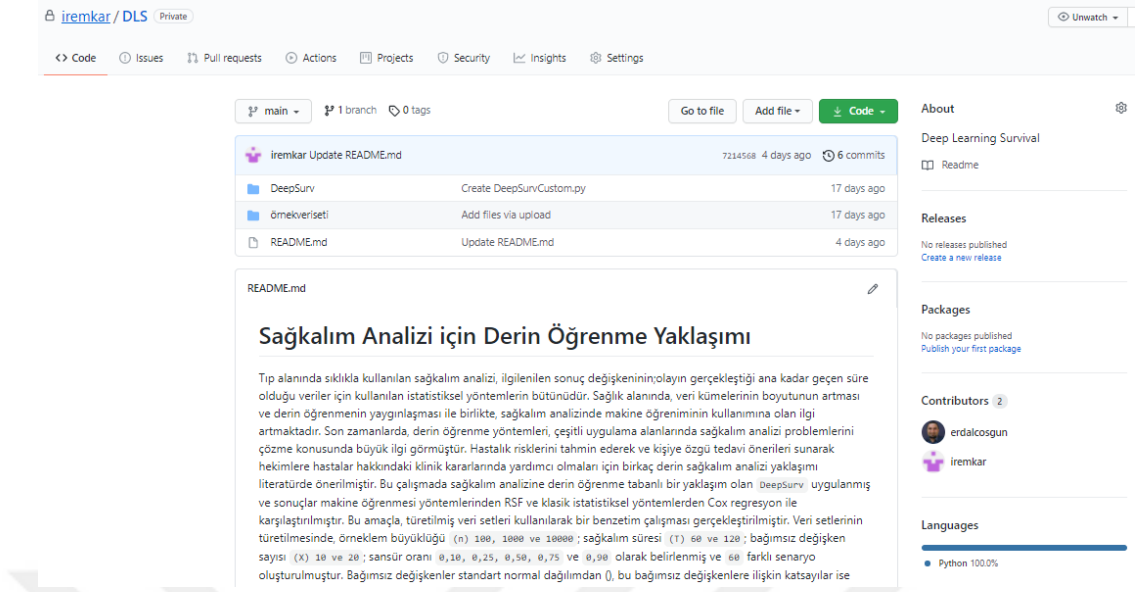
DeepSurv için kullanılan hiperparametreler; epoch sayısı, L2 düzenleştirmesi, seyreltme, gizli katman sayısı, nod sayısı, öğrenme oranı, öğrenme oranı azalımı, momentum ve aktivasyon fonksiyonudur. Bu hiperparametreler modele en uygun olacak, aşırı uyuma neden olmayacak şekilde rastgele arama (random search) yöntemi ile belirlenmiş ve Çizelge 2.3'te sunulmuştur.

Çizelge 2.3. Hiperparametreler.

Hiperparametre	
Epoch Sayısı	2000
L2 Düzenlestirmesi	0,01
Seyreltme	0,2
Gizli Katman Sayısı	100
Nod Sayısı	10
Öğrenme Oranı	0,0001
Öğrenme Oranı Azalımı	0,0001
Momentum	0,9
Aktivasyon Fonksiyonu	ReLU

DeepSurv, RSF ve Cox regresyon yöntemlerinin karşılaştırılmasında Harrell'in C-indeks değeri kullanılmıştır. Her senaryo için bu iki yönteme ait 100 eğitim ve validasyon C-indeks değeri kaydedilmiş ve bu değerlerin ortalaması alınarak yöntemlerinin tahmin başarısı değerlendirilmiştir. Yöntemlerin karşılaştırılmasına ilişkin ise ısı-harita grafiklerinden yararlanılmıştır. Isı-harita grafikleri validasyon seti sonuçları kullanılarak çizilmiştir.

Verilerin türetilmesi, RSF ve Cox regresyon yöntemlerinin uygulanması R programlama dili (ver. 3.6.3) ile gerçekleştirilmiştir. Verilerin türetilmesinde “coxed” paketi kullanılmıştır (Kropko ve Harden, 2019). Cox regresyon analizinde “BiocManager”, “survcomp” ve “survival” paketlerinden faydalanılmıştır (Morgan, 2019; Schröder ve ark., 2011; Therneau ve Grambsch, 2000). RSF yöntemi “randomForestSRC” paketi aracılığıyla uygulanmıştır (Ishwaran ve ark., 2021). Isı-harita grafikleri de R programlama dilinde “pheatmap” paketi yardımıyla çizilmiştir (Raivo, 2019). Derin öğrenmeye ilişkin analizler Python programlama dilinde gerçekleştirilmiştir. DeepSurv “lasagne” paketi ile Theano’da uygulanmıştır. Sonuç modellerinin diğer araştırmacılarla paylaşılabilmesi amacıyla GitHub ortamı kullanılmıştır (Şekil 2.1).



Şekil 2.1. GitHub Ortamı.

2.1. Sanal Makineler (Virtual Machines)

Genellikle VM olarak kısaltılan bir sanal makine, dizüstü bilgisayar, akıllı telefon veya sunucu gibi diğer fiziksel bilgisayardan farklı değildir. Dosyaları depolamak için bir CPU, bellek ve disklere sahiptir ve gerekirse internete bağlanabilir. Bilgisayarı oluşturan parçalar (donanım) fiziksel ve somut olduğundan, VM'ler genellikle fiziksel sunucular içinde yalnızca kod olarak bulunan sanal bilgisayarlar veya yazılım tanımlı bilgisayarlar olarak düşünülebilir (Microsoft Azure, Erişim tarihi: 23/04/2021).

Bu tez çalışmasında NVIDIA Tesla K80 Card ve Intel Xeon E5-2690 v3 (Haswell) işlemcisi tarafından desteklenen NC serisi sanal makinelerden, NC24 kullanılmıştır (Çizelge 2.4). Gerekli olan fayda maliyet göz önünde bulundurulduğunda, VM kullanılmasının nedeni, hız, ölçeklenebilirlik, azaltılmış kapalı kalma süresi ve güvenlik gibi avantajlarının olmasıdır.

Çizelge 2.4. NC Serisi Sanal Makineler.

Boyut	vCPU	Bellek: GiB	Geçici Depolama (SSD) GiB	GPU	GPU Belleği: GiB	Maksimum Veri Diskleri	Maksimum NIC
Standard_NC6	6	56	340	1	12	24	1
Standard_NC12	12	112	680	2	24	48	2
Standard_NC24	24	224	1440	4	48	64	4
Standard_NC24r*	24	224	1440	4	48	64	4

1 GPU: K80 kartın yarısı

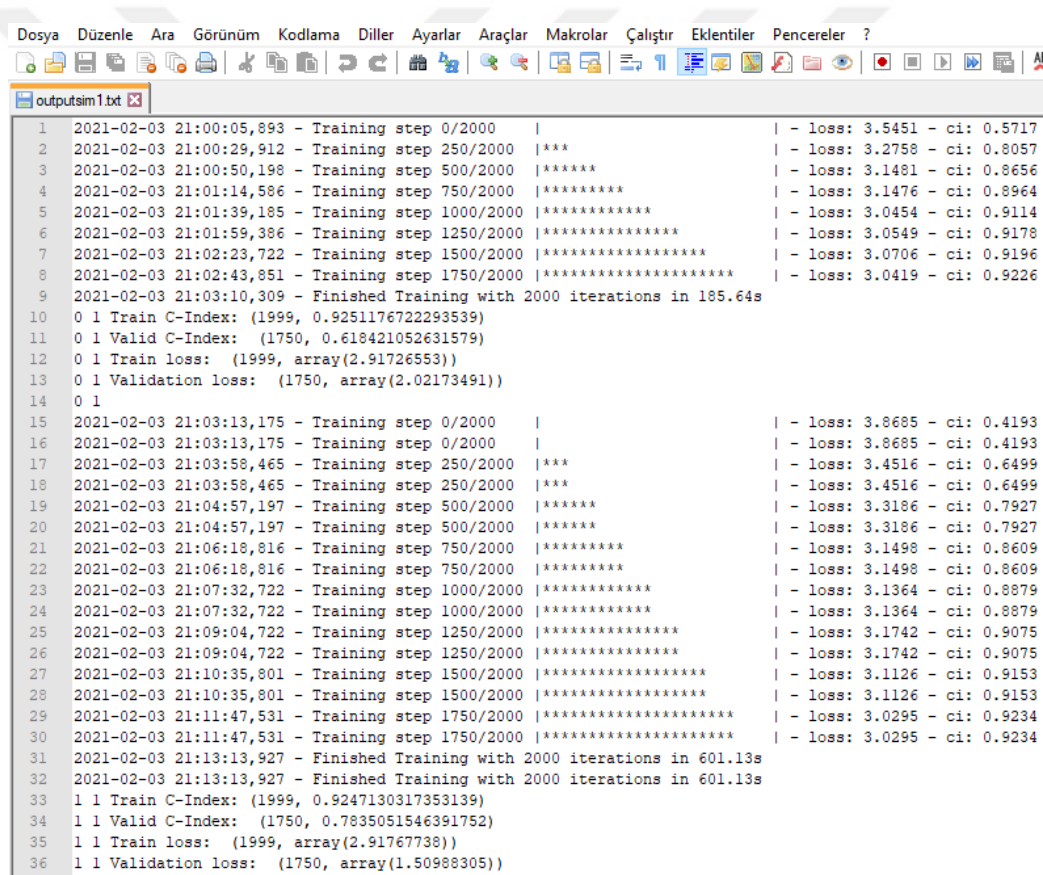
*RMDA özellikli

Çizelge 2.4'te yer alan bazı terimleri kısaca açıklayacak olursak, vCPU (sanal CPU), belirli bir sanal bir makineye (VM) atanan fiziksel CPU'nun bir kısmını veya bölümünü ifade eder ve sanal işlemci olarak da bilinir. GPU, esas olarak grafik oluşturmayı hızlandırmak için tasarlanmış özel bir işlemcidir. GPU'lar birçok veri parçasını aynı anda işleyebilir ve bu nedenle makine öğrenimi, video düzenleme gibi alanlar için kullanılırlar. Depolama kapasitesi GiB (1024³ bayt) cinsinden gösterilmiştir. NIC (Network Interface Card) bilgisayarların ağ ile iletişim kurmalarına izin veren kart olarak tanımlanabilir. RMDA (Uzaktan Doğrudan Bellek Erişimi) bir bilgisayarın belleğinden diğer bir bilgisayarın belleğine, her iki bilgisayarın da işletim sistemlerini etkilemeden veri aktarımını sağlar.

3. BULGULAR

Bu bölümde, oluşturulan senaryolarda DeepSurv, RSF ve Cox regresyon yöntemlerine ait çıktılar yorumlanmıştır. Senaryoların her birinde, eğitim ve validasyon setleri için 100 tekrar sonucu ortalama C-indeks değerleri elde edilmiş ve güven aralıkları ile birlikte sunulmuştur.

DeepSurv yönteminin uygulanması sonucu elde edilen örnek bir çıktı Şekil 3.1.'de gösterilmiştir.



```
Dosya Düzenle Ara Görünüm Kodlama Diller Ayarlar Araçlar Makrolar Çalıştır Eklentiler Pencerele ?
outputsim1.txt
1 2021-02-03 21:00:05,893 - Training step 0/2000 | - loss: 3.5451 - ci: 0.5717
2 2021-02-03 21:00:29,912 - Training step 250/2000 |*** - loss: 3.2758 - ci: 0.8057
3 2021-02-03 21:00:50,198 - Training step 500/2000 |***** - loss: 3.1481 - ci: 0.8656
4 2021-02-03 21:01:14,586 - Training step 750/2000 |***** - loss: 3.1476 - ci: 0.8964
5 2021-02-03 21:01:39,185 - Training step 1000/2000 |***** - loss: 3.0454 - ci: 0.9114
6 2021-02-03 21:01:59,386 - Training step 1250/2000 |***** - loss: 3.0549 - ci: 0.9178
7 2021-02-03 21:02:23,722 - Training step 1500/2000 |***** - loss: 3.0706 - ci: 0.9196
8 2021-02-03 21:02:43,851 - Training step 1750/2000 |***** - loss: 3.0419 - ci: 0.9226
9 2021-02-03 21:03:10,309 - Finished Training with 2000 iterations in 185.64s
10 0 1 Train C-Index: (1999, 0.9251176722293539)
11 0 1 Valid C-Index: (1750, 0.618421052631579)
12 0 1 Train loss: (1999, array(2.91726553))
13 0 1 Validation loss: (1750, array(2.02173491))
14 0 1
15 2021-02-03 21:03:13,175 - Training step 0/2000 | - loss: 3.8685 - ci: 0.4193
16 2021-02-03 21:03:13,175 - Training step 0/2000 | - loss: 3.8685 - ci: 0.4193
17 2021-02-03 21:03:58,465 - Training step 250/2000 |*** - loss: 3.4516 - ci: 0.6499
18 2021-02-03 21:03:58,465 - Training step 250/2000 |*** - loss: 3.4516 - ci: 0.6499
19 2021-02-03 21:04:57,197 - Training step 500/2000 |***** - loss: 3.3186 - ci: 0.7927
20 2021-02-03 21:04:57,197 - Training step 500/2000 |***** - loss: 3.3186 - ci: 0.7927
21 2021-02-03 21:06:18,816 - Training step 750/2000 |***** - loss: 3.1498 - ci: 0.8609
22 2021-02-03 21:06:18,816 - Training step 750/2000 |***** - loss: 3.1498 - ci: 0.8609
23 2021-02-03 21:07:32,722 - Training step 1000/2000 |***** - loss: 3.1364 - ci: 0.8879
24 2021-02-03 21:07:32,722 - Training step 1000/2000 |***** - loss: 3.1364 - ci: 0.8879
25 2021-02-03 21:09:04,722 - Training step 1250/2000 |***** - loss: 3.1742 - ci: 0.9075
26 2021-02-03 21:09:04,722 - Training step 1250/2000 |***** - loss: 3.1742 - ci: 0.9075
27 2021-02-03 21:10:35,801 - Training step 1500/2000 |***** - loss: 3.1126 - ci: 0.9153
28 2021-02-03 21:10:35,801 - Training step 1500/2000 |***** - loss: 3.1126 - ci: 0.9153
29 2021-02-03 21:11:47,531 - Training step 1750/2000 |***** - loss: 3.0295 - ci: 0.9234
30 2021-02-03 21:11:47,531 - Training step 1750/2000 |***** - loss: 3.0295 - ci: 0.9234
31 2021-02-03 21:13:13,927 - Finished Training with 2000 iterations in 601.13s
32 2021-02-03 21:13:13,927 - Finished Training with 2000 iterations in 601.13s
33 1 1 Train C-Index: (1999, 0.9247130317353139)
34 1 1 Valid C-Index: (1750, 0.7835051546391752)
35 1 1 Train loss: (1999, array(2.91767738))
36 1 1 Validation loss: (1750, array(1.50988305))
```

Şekil 3.1. DeepSurv Çıktısı Örneği.

Bu tez çalışması kapsamında oluşturulan 60 senaryonun her biri için DeepSurv yönteminin uygulanmasında geçen sürelerle ilişkin bilgiler Çizelge 3.1.'de sunulmuştur. Burada, 100 tekrar sonucu ortalama süreler her bir senaryo için standart sapma değerleri ile birlikte verilmiştir. Süreler saniye cinsinden kaydedilmiştir.

Çizelge 3.1. Her Senaryo için Deepsurv Yönteminin Uygulanmasında Geçen Süre (sn).

Senaryo	n	X	T	Sansür Oranı	Süre (sn)
1	100	10	60	0,1	313,05±36,91
2				0,25	312,33±38,32
3				0,5	316,38±38,14
4				0,75	319,53±36,60
5				0,9	-
6			120	0,1	223,59±21,11
7				0,25	320,32±13,23
8				0,5	318,94±15,22
9				0,75	236,79±3,58
10				0,9	-
11		20	60	0,1	421,19±93,45
12				0,25	250,58±21,80
13				0,5	250,00±20,83
14				0,75	610,98±189,28
15				0,9	-
16			120	0,1	237,25±3,39
17				0,25	251,21±20,13
18				0,5	252,61±18,87
19				0,75	247,39±15,92
20				0,9	-
21	1000	10	60	0,1	743,00±261,33
22				0,25	782,25±49,53
23				0,5	779,96±40,62
24				0,75	772,26±30,36
25				0,9	738,94±9,13
26			120	0,1	942,03±177,67
27				0,25	933,15±148,92
28				0,5	925,71±135,87
29				0,75	916,04±127,18
30				0,9	908,90±123,35
31		20	60	0,1	1121,01±94,57
32				0,25	1115,07±65,67
33				0,5	1104,28±41,28
34				0,75	1091,55±24,44
35				0,9	1079,21±21,57
36			120	0,1	1120,60±101,45
37				0,25	1115,96±74,03
38				0,5	1084,76±64,23
39				0,75	1070,15±28,93
40				0,9	1054,15±13,03
41	10000	10	60	0,1	1914,22±852,59
42				0,25	1890,90±815,41
43				0,5	2365,98±200,45
44				0,75	2274,52±92,14
45				0,9	2170,31±18,50
46			120	0,1	2463,51±468,17
47				0,25	2433,20±334,80
48				0,5	1364,86±77,82
49				0,75	1308,43±35,68
50				0,9	1265,83±25,89
51		20	60	0,1	1433,88±184,46
52				0,25	1404,76±122,34
53				0,5	1383,65±132,54
54				0,75	1320,12±83,97
55				0,9	1284,40±67,28
56			120	0,1	1467,74±201,09
57				0,25	1266,64±121,63
58				0,5	1231,99±67,53
59				0,75	1196,59±38,08
60				0,9	1164,61±29,62

$U(-1,1)$ ve $N(0;0,1)$ dağılımlarından türetilen bağımsız değişken katsayılarına ilişkin farklı örneklem büyüklüklerinde yapılan analiz sonuçları sırasıyla DeepSurv, RSF ve Cox regresyon yöntemleri için alt başlıklarda verilmiştir. Ancak, örneklem büyüklüğü 100 iken sansür oranının 0,90 olduğu senaryolarda validasyon setlerinin 100 tekrarının yarısından fazlasında ilgilenilen olayın gerçekleşmediği görülmüştür. Bu nedenle, bu senaryolara ilişkin sonuçların sunulmamasına karar verilmiştir. Örneklem büyüklüğünün 100 olduğu başlıklarda yorumlar 16 senaryo üzerinden yapılmıştır. Bu bölümde son olarak Başlık 3.3'te gerçek veri setine ilişkin açıklamalar ve analiz sonuçları sunulmuştur.

3.1. Bağımsız Değişkenlere İlişkin Katsayılar $U(-1,1)$ iken Sonuçlar

3.1.1. Örneklem Büyüklüğü 100 iken DeepSurv Sonuçları ($U(-1,1)$)

Örneklem büyüklüğü 100 iken DeepSurv sonuçları Çizelge 3.2'de gösterildiği gibidir. Çizelge 3.2 incelendiğinde sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,895 ve 0,992; validasyon setinde ise 0,648 ve 0,790 olarak kaydedilmiştir. Validasyon setindeki en yüksek C-indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 20 ve sansür oranının 0,10 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 120, bağımsız değişken sayısının 20 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir.

Çizelge 3.2. Örneklem Büyüklüğü 100 iken DeepSurv Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	T=60				
	0,10	0,902 [0,898;0,906]	0,789 [0,776;0,802]	0,978 [0,975;0,982]	0,790 [0,775;0,805]
	0,25	0,908 [0,903;0,912]	0,768 [0,751;0,786]	0,929 [0,925;0,934]	0,738 [0,721;0,756]
	0,50	0,930 [0,926;0,935]	0,744 [0,724;0,764]	0,958 [0,955;0,961]	0,685 [0,660;0,710]
	0,75	0,975 [0,972;0,978]	0,690 [0,654;0,726]	0,992 [0,990;0,994]	0,682 [0,649;0,715]
	0,90	-	-	-	-
	T=120				
Sansür Oranı	0,10	0,895 [0,891;0,899]	0,769 [0,756;0,783]	0,919 [0,914;0,923]	0,743 [0,725;0,760]
	0,25	0,906 [0,902;0,910]	0,766 [0,750;0,783]	0,929 [0,925;0,933]	0,735 [0,716;0,754]
	0,50	0,931 [0,927;0,936]	0,723 [0,698;0,748]	0,955 [0,951;0,958]	0,703 [0,678;0,727]
	0,75	0,973 [0,969;0,977]	0,709 [0,676;0,743]	0,984 [0,981;0,987]	0,648 [0,615;0,681]
	0,90	-	-	-	-

Tüm senaryolarda validasyon setlerinde sansür oranı arttıkça C-indeks değerinin azaldığı gözlemlenmiştir. Sağkalım süresinin 60 ve 120 olduğu her iki durumda da bağımsız değişken sayısı 10 iken 20 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir. Her iki bağımsız değişken sayısında da sağkalım süresi 60 iken 120 olduğu duruma göre sonuçlar daha yüksek bulunmuştur. Ancak, bağımsız değişken sayısı 20 ve sansür oranı 0,50 iken ve bağımsız değişken sayısı 10 ve sansür oranı 0,75 iken sağkalım süresi 120 olduğu durumda, 60 olduğu duruma göre sonuçların daha yüksek olduğu görülmüştür.

3.1.2. Örneklem Büyüklüğü 1000 iken DeepSurv Sonuçları (U(-1,1))

Örneklem büyüklüğü 1000 iken DeepSurv sonuçları Çizelge 3.3'te gösterildiği gibidir. Burada, sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,819 ve 0,903; validasyon setinde 0,679 ve 0,823 olarak kaydedilmiştir. Validasyon setindeki en yüksek C-indeks değeri, sağkalım süresinin 120, bağımsız değişken sayısının 10 ve sansür oranının 0,10 olduğu senaryoda gerçekleşirken, en düşük C-

indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 10 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir.

Örneklem büyüklüğü 1000 iken de, 100 olduğu duruma benzer biçimde tüm senaryolarda validasyon setlerinde sansür oranı arttıkça C-indeks değerinin azaldığı gözlemlenmiştir.

Çizelge 3.3. Örneklem Büyüklüğü 1000 iken Deepsurv Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	T=60				
	0,10	0,835 [0,833;0,836]	0,819 [0,815;0,824]	0,832 [0,829;0,836]	0,806 [0,799;0,813]
	0,25	0,826 [0,824;0,828]	0,807 [0,801;0,813]	0,825 [0,822;0,829]	0,793 [0,787;0,799]
	0,50	0,819 [0,816;0,822]	0,790 [0,782; 0,798]	0,820 [0,816; 0,824]	0,768 [0,760;0,776]
	0,75	0,820 [0,816; 0,823]	0,765 [0,754;0,776]	0,835 [0,831;0,840]	0,733 [0,721;0,745]
	0,90	0,867 [0,862;0,873]	0,743 [0,723;0,763]	0,902 [0,897;0,906]	0,679 [0,657;0,702]
	T=120				
Sansür Oranı	0,10	0,833 [0,832;0,835]	0,823 [0,818;0,828]	0,831 [0,828;0,834]	0,799 [0,792;0,805]
	0,25	0,826 [0,824;0,828]	0,805 [0,799;0,811]	0,827 [0,824;0,830]	0,791 [0,785;0,798]
	0,50	0,821 [0,818; 0,823]	0,792 [0,785;0,799]	0,824 [0,821;0,828]	0,777 [0,769;0,785]
	0,75	0,828 [0,824;0,832]	0,770 [0,758;0,781]	0,841 [0,837;0,845]	0,751 [0,740;0,762]
	0,90	0,870 [0,866;0,874]	0,753 [0,733;0,772]	0,903 [0,898;0,908]	0,729 [0,711;0,748]

Her iki sağkalım süresinde de bağımsız değişken sayısı 10 iken 20 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir. Bağımsız değişken sayısı 20 ve sansür oranı $\geq 0,50$ iken sağkalım süresi 120 olduğu durumda, 60 olduğu duruma göre daha yüksek C-indeks değerleri bulunmuştur. Bağımsız değişken sayısı 10 iken ise sansür oranının 0,25 olduğu dışında sağkalım süresi 120 olduğu durumda daha yüksek sonuçlar elde edilmekle birlikte, her iki sağkalım süresinde de sonuçların benzer olduğu söylenebilir.

3.1.3. Örneklem Büyüklüğü 10000 iken DeepSurv Sonuçları (U(-1,1))

Örneklem büyüklüğü 10000 iken DeepSurv sonuçları incelendiğinde sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,781 ve 0,827; validasyon setinde ise 0,752 ve 0,823 olarak kaydedilmiştir. Burada, güven aralıklarının oldukça dar olduğu görülmüştür. Validasyon setindeki en yüksek C-indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 10 ve sansür oranının 0,10 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 20 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir (Çizelge 3.4).

Çizelge 3.4. Örneklem Büyüklüğü 10000 iken DeepSurv Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	T=60				
	0,10	0,827 [0,826;0,828]	0,823 [0,821;0,825]	0,820 [0,818;0,823]	0,810 [0,807;0,814]
	0,25	0,818 [0,816;0,819]	0,813 [0,810;0,815]	0,807 [0,804;0,810]	0,796 [0,791;0,800]
	0,50	0,802 [0,799;0,805]	0,796 [0,792;0,799]	0,794 [0,790;0,797]	0,781 [0,777;0,785]
	0,75	0,789 [0,786;0,793]	0,782 [0,778;0,787]	0,781 [0,776;0,785]	0,762 [0,756;0,768]
	0,90	0,789 [0,784;0,794]	0,771 [0,764;0,778]	0,783 [0,778;0,788]	0,752 [0,746;0,759]
Sansür Oranı	T=120				
	0,10	0,824 [0,823;0,825]	0,819 [0,817;0,821]	0,822 [0,820;0,825]	0,813 [0,810;0,816]
	0,25	0,818 [0,817;0,819]	0,812 [0,810;0,815]	0,811 [0,808;0,814]	0,801 [0,797;0,805]
	0,50	0,806 [0,804;0,808]	0,801 [0,798;0,804]	0,800 [0,797;0,803]	0,787 [0,783;0,791]
	0,75	0,799 [0,796;0,802]	0,790 [0,786;0,794]	0,788 [0,784;0,792]	0,771 [0,766;0,776]
	0,90	0,798 [0,795;0,802]	0,778 [0,772;0,783]	0,790 [0,785;0,794]	0,757 [0,750;0,764]

Örneklem büyüklüğünün 100 ve 1000 olduğu durumdakine benzer biçimde burada da tüm senaryolarda validasyon setlerinde sansür oranı arttıkça C-indeks değerinin azaldığı gözlemlenmiştir. Sağkalım süresinin 60 ve 120 olduğu her iki durumda da bağımsız değişken sayısı 10 iken 20 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir. Bağımsız değişken sayısı 20 iken, sağkalım süresinin 120

olduğu durumda 60 olduğu duruma göre daha yüksek C-indeks sonuçları bulunmuştur. Ancak, bağımsız değişken sayısı 10 iken sağkalım oranının çok düşük (0,10 ve 0,25) ve çok yüksek (0,90) olduğu durumlarda sağkalım süresi 60 iken 120 olduğu duruma göre sonuçların daha yüksek olduğu görülmüştür.

3.1.4. Örneklem Büyüklüğü 100 iken RSF Sonuçları (U(-1,1))

Örneklem büyüklüğü 100 iken RSF sonuçları Çizelge 3.5'te gösterildiği gibidir. Burada, sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,548 ve 0,690; validasyon setinde 0,610 ve 0,711 olarak kaydedilmiştir. Ayrıca, validasyon ve eğitim setindeki en büyük ve en küçük C-indeks değerlerinin elde edildiği senaryoların aynı olduğu görülmüştür. Buna göre, en yüksek C-indeks değeri, sağkalım süresinin 120, bağımsız değişken sayısının 10 ve sansür oranının 0,10 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 120, bağımsız değişken sayısının 20 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir.

Çizelge 3.5. Örneklem Büyüklüğü 100 iken RSF Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	<u>T=60</u>				
	0,10	0,687 [0,680;0,694]	0,701 [0,686;0,715]	0,658 [0,649;0,666]	0,685 [0,670;0,699]
	0,25	0,675 [0,667;0,683]	0,699 [0,681;0,717]	0,640 [0,631;0,649]	0,683 [0,666;0,701]
	0,50	0,644 [0,633;0,655]	0,696 [0,675;0,717]	0,613 [0,601;0,626]	0,625 [0,602;0,649]
	0,75	0,614 [0,595;0,632]	0,651 [0,621;0,681]	0,558 [0,540;0,576]	0,632 [0,597;0,667]
	0,90	-	-	-	-
Sansür Oranı	<u>T=120</u>				
	0,10	0,690 [0,684;0,697]	0,711 [0,695;0,727]	0,658 [0,650;0,665]	0,690 [0,676;0,704]
	0,25	0,686 [0,678;0,693]	0,691 [0,674;0,709]	0,647 [0,639;0,655]	0,683 [0,665;0,702]
	0,50	0,648 [0,636;0,659]	0,668 [0,641;0,695]	0,611 [0,598;0,624]	0,656 [0,634;0,678]
	0,75	0,612 [0,594;0,629]	0,666 [0,631;0,700]	0,548 [0,528;0,568]	0,610 [0,575;0,645]
	0,90	-	-	-	-

Genel olarak, sansür oranı arttıkça C-indeks değerlerinin azaldığından söz edilebilir. Validasyon setlerinde, sağkalım süresinin 60 ve 120 olduğu her iki durumda da bağımsız değişken sayısı 10 iken 20 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir. Bağımsız değişken sayısı 20 iken, sansür oranının 0,75 olduğu senaryo hariç; sağkalım süresinin 120 olduğu durumda 60 olduğu duruma göre daha yüksek C-indeks sonuçları bulunmuştur. Ancak, bağımsız değişken sayısı 10 iken sağkalım oranının orta büyüklükte (0,25 ve 0,50) olduğu durumlarda sağkalım süresi 60 iken 120 olduğu duruma göre sonuçların daha yüksek olduğu görülmüştür.

3.1.5. Örneklem Büyüklüğü 1000 iken RSF Sonuçları (U(-1,1))

Örneklem büyüklüğü 1000 iken RSF sonuçları incelendiğinde sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,678 ve 0,772; validasyon setinde ise 0,669 ve 0,777 olarak kaydedilmiştir. Hem eğitim hem de validasyon setindeki en yüksek C-indeks değeri sağkalım süresinin 120, bağımsız değişken sayısının 10 ve sansür oranının 0,10 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 20 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir (Çizelge 3.6).

Çizelge 3.6. Örneklem Büyüklüğü 1000 iken RSF Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	<u>T=60</u>				
	0,10	0,769 [0,767;0,771]	0,770 [0,765;0,776]	0,761 [0,759;0,763]	0,766 [0,760;0,771]
	0,25	0,760 [0,757;0,763]	0,763 [0,757;0,769]	0,757 [0,754;0,759]	0,759 [0,753;0,765]
	0,50	0,755 [0,752;0,758]	0,756 [0,747;0,765]	0,741 [0,737;0,744]	0,742 [0,735;0,749]
	0,75	0,734 [0,728;0,739]	0,732 [0,720;0,745]	0,718 [0,714;0,723]	0,712 [0,700;0,723]
	0,90	0,706 [0,698;0,714]	0,709 [0,688;0,730]	0,678 [0,671;0,685]	0,669 [0,648;0,690]
Sansür Oranı	<u>T=120</u>				
	0,10	0,772 [0,770;0,773]	0,777 [0,772;0,781]	0,765 [0,763;0,768]	0,764 [0,758;0,769]
	0,25	0,769 [0,767;0,771]	0,768 [0,763;0,774]	0,760 [0,757;0,762]	0,761 [0,755;0,767]
	0,50	0,761 [0,758;0,764]	0,758 [0,750;0,765]	0,749 [0,745;0,752]	0,758 [0,752;0,765]
	0,75	0,741 [0,737;0,746]	0,744 [0,732;0,756]	0,724 [0,720;0,729]	0,730 [0,720;0,741]
	0,90	0,715 [0,708;0,721]	0,714 [0,694;0,733]	0,685 [0,679;0,692]	0,690 [0,672;0,709]

Tüm senaryolarda sansür oranı arttıkça C-indeks değerinin azaldığı gözlemlenmiştir. Validasyon setlerinde, sağkalım süresinin 60 ve 120 olduğu her iki durumda da bağımsız değişken sayısı 10 iken 20 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir. Bağımsız değişken sayısı 10 iken, sağkalım süresinin 120 olduğu durumda 60 olduğu duruma göre daha yüksek C-indeks sonuçları bulunmuştur. Bağımsız değişken sayısı 10 iken ise sansür oranı $\geq 0,25$ olduğu durumlarda sağkalım süresi 120 iken elde edilen sonuçların daha yüksek olduğu görülmüştür.

3.1.6. Örneklem Büyüklüğü 10000 iken RSF Sonuçları (U(-1,1))

Örneklem büyüklüğü 10000 iken RSF sonuçları Çizelge 3.7’de gösterildiği gibidir.

Çizelge 3.7. Örneklem Büyüklüğü 10000 iken RSF Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	<u>T=60</u>				
	0,10	0,787 [0,785;0,789]	0,788 [0,785;0,790]	0,781 [0,779;0,783]	0,782 [0,780;0,785]
	0,25	0,783 [0,781;0,786]	0,783 [0,780;0,786]	0,777 [0,774;0,779]	0,777 [0,774;0,780]
	0,50	0,776 [0,773;0,780]	0,776 [0,772;0,780]	0,770 [0,767;0,773]	0,772 [0,768;0,775]
	0,75	0,767 [0,764;0,770]	0,769 [0,765;0,773]	0,762 [0,758;0,766]	0,761 [0,757;0,766]
	0,90	0,753 [0,748;0,757]	0,750 [0,742;0,757]	0,744 [0,740;0,748]	0,748 [0,741;0,754]
Sansür Oranı	<u>T=120</u>				
	0,10	0,790 [0,789;0,792]	0,790 [0,788;0,792]	0,784 [0,782;0,786]	0,785 [0,783;0,787]
	0,25	0,789 [0,787;0,791]	0,789 [0,786;0,792]	0,783 [0,781;0,785]	0,785 [0,782;0,787]
	0,50	0,782 [0,780;0,785]	0,784 [0,780;0,787]	0,779 [0,777;0,782]	0,780 [0,776;0,783]
	0,75	0,776 [0,773;0,779]	0,775 [0,770;0,779]	0,767 [0,764;0,771]	0,769 [0,765;0,774]
	0,90	0,762 [0,759;0,766]	0,763 [0,757;0,768]	0,750 [0,746;0,754]	0,751 [0,745;0,756]

Burada, sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,744 ve 0,790; validasyon setinde 0,748 ve 0,790 olarak kaydedilmiştir. Eğitim ve validasyon setlerinin her ikisinde de, en yüksek C-indeks değeri, sağkalım süresinin 120, bağımsız değişken sayısının 10 ve sansür oranının 0,10 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 20 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir.

Örneklem büyüklüğü 100 ve 1000 olduğu duruma benzer şekilde burada da, tüm senaryolarda sansür oranı arttıkça C-indeks değerinin azaldığı gözlemlenmiştir. Validasyon setlerinde, sağkalım süresinin 60 ve 120 olduğu her iki durumda da bağımsız değişken sayısı 10 iken 20 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir. Bağımsız değişken sayısının 10 ve 20 olduğu her iki koşulda da, sağkalım süresi 120 olduğu durumda 60 olduğu duruma göre elde edilen C-indeks değerleri daha yüksektir.

3.1.7. Örneklem Büyüklüğü 100 iken Cox Regresyon Sonuçları (U(-1,1))

Bölüm 3'te senaryoların her birinde, eğitim ve validasyon setleri için 100 tekrar sonucu ortalama C-indeks değeri elde edildiğinden bahsedilmiştir. Ancak, Cox regresyon analizinde örneklem büyüklüğü 100 ve 1000 iken bazı senaryoların validasyon setlerinde 100 tekrarın tümünde sonuç alınamamıştır. Bu nedenle analiz sonuçları iki ayrı tabloda sunulmuştur ve örneklem büyüklüğü 100 iken Cox regresyon sonuçları Çizelge 3.8 ve Çizelge 3.9'da özetlenmiştir.

Çizelge 3.8'de, i tekrar sayısını göstermek üzere; Cox regresyon analizi sonucu elde edilen ortalama C-indeks değerleri ve güven aralıkları sunulmuştur.

Çizelge 3.9'da ise Cox regresyon analizi sonunda C-indeks değeri elde edilemeyenler başarısız kabul edilip sonuçları "0" olarak alınarak 100 tekrar sonucu ortalama C-indeks değerleri ve güven aralıkları sunulmuştur.

Çizelge 3.8. Örneklem Büyüklüğü 100 iken Cox Regresyon Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).

		X=10			X=20		
		Eğitim	i	Validasyon	Eğitim	i	Validasyon
<u>Sansür Oranı</u>	<u>T=60</u>						
	0,10	0,857 [0,852;0,862]	100	0,837 [0,826;0,848]	0,892 [0,888;0,896]	100	0,847 [0,835;0,859]
	0,25	0,848 [0,843;0,853]	98	0,821 [0,807;0,836]	0,878 [0,873;0,883]	100	0,843 [0,829;0,857]
	0,50	0,832 [0,825;0,839]	95	0,813 [0,796;0,831]	0,876 [0,870;0,882]	97	0,795 [0,773;0,817]
	0,75	0,841 [0,832;0,851]	81	0,730 [0,700;0,761]	0,900 [0,893;0,906]	79	0,728 [0,694;0,762]
	0,90	-	-	-	-	-	-
<u>Sansür Oranı</u>	<u>T=120</u>						
	0,10	0,852 [0,848;0,857]	99	0,824 [0,811;0,838]	0,885 [0,880;0,889]	99	0,842 [0,830;0,854]
	0,25	0,848 [0,843;0,853]	99	0,817 [0,803;0,832]	0,883 [0,879;0,887]	98	0,831 [0,818;0,845]
	0,50	0,841 [0,834;0,849]	96	0,790 [0,769;0,811]	0,874 [0,868;0,879]	99	0,798 [0,777;0,820]
	0,75	0,852 [0,843;0,860]	80	0,754 [0,724;0,784]	0,897 [0,889;0,904]	87	0,706 [0,671;0,742]
	0,90	-	-	-	-	-	-

i:tekrar sayısı

Çizelge 3.8 incelendiğinde, sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,832 ve 0,900; validasyon setinde 0,706 ve 0,847 olarak kaydedilmiştir. Validasyon setindeki en yüksek C-indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 20 ve sansür oranının 0,10 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 120, bağımsız değişken sayısının 20 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir.

Validasyon setlerinde sansür oranı arttıkça C-indeks değerinin azaldığı gözlemlenmiştir. Sağkalım süresi 60 ve sağkalım oranı $<0,50$ iken bağımsız değişken sayısının 20 olduğu durumda 10 olduğu duruma göre sonuçların daha yüksek olduğu görülmüştür. Sağkalım süresi 120 ve sağkalım oranı $\leq 0,75$ iken bağımsız değişken sayısının 20 olduğu durumda 10 olduğu duruma göre sonuçların daha yüksek olduğu görülmüştür. Bağımsız değişken sayısı 10 ve sağkalım oranı $\leq 0,75$ iken sağkalım süresinin 60 olduğu durumda 120 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir.

Çizelge 3.9. Örneklem Büyüklüğü 100 iken Cox Regresyon Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (NA=0) (U(-1,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	<u>T=60</u>				
	0,10	0,857 [0,852;0,862]	0,837 [0,826;0,848]	0,892 [0,888;0,896]	0,847 [0,835;0,859]
	0,25	0,848 [0,843;0,853]	0,805 [0,778;0,832]	0,878 [0,873;0,883]	0,843 [0,829;0,857]
	0,50	0,832 [0,825;0,839]	0,773 [0,734;0,812]	0,876 [0,870;0,882]	0,771 [0,736;0,805]
	0,75	0,841 [0,832;0,851]	0,592 [0,529;0,654]	0,900 [0,893;0,906]	0,575 [0,510;0,640]
	0,90	-	-	-	-
	<u>T=120</u>				
Sansür Oranı	0,10	0,852 [0,848;0,857]	0,816 [0,795;0,821]	0,885 [0,880;0,889]	0,834 [0,813;0,854]
	0,25	0,848 [0,843;0,853]	0,809 [0,787;0,831]	0,883 [0,879;0,887]	0,815 [0,788;0,841]
	0,50	0,841 [0,834;0,849]	0,758 [0,721;0,795]	0,874 [0,868;0,879]	0,790 [0,764;0,817]
	0,75	0,852 [0,843;0,860]	0,603 [0,538;0,668]	0,897 [0,889;0,904]	0,615 [0,558;0,671]
	0,90	-	-	-	-

Çizelge 3.9 incelendiğinde, sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,832 ve 0,900; validasyon setinde 0,575 ve 0,847 olarak

kaydedilmiştir. Validasyon setindeki en yüksek C-indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 20 ve sansür oranının 0,10 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 20 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir.

Validasyon setlerinde sansür oranı arttıkça C-indeks değerinin azaldığı gözlemlenmiştir. Bağımsız değişken sayısı 20 ve sağkalım oranı $<0,50$ iken sağkalım süresinin 60 olduğu durumda 120 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir. Sağkalım süresi 120 iken 60 olduğu duruma göre bağımsız değişken sayısının 10 ve 20 olduğu her iki durumda da daha yüksek sonuçlar bulunmuştur. Sağkalım süresi 60 ve sağkalım oranı $<0,50$ iken bağımsız değişken sayısının 20 olduğu durumda 10 olduğu duruma göre sonuçların daha yüksek olduğu görülmüştür.

3.1.8. Örneklem Büyüklüğü 1000 iken Cox Regresyon Sonuçları (U(-1,1))

Örneklem büyüklüğü 1000 iken Cox regresyon sonuçları Çizelge 3.10'da gösterildiği gibidir. Burada, sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,804 ve 0,872; validasyon setinde 0,824 ve 0,882 olarak kaydedilmiştir. Validasyon setindeki en yüksek C-indeks değeri, sağkalım süresinin 60, bağımsız değişken sayısının 20 ve sansür oranının 0,10 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 20 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir.

Çizelge 3.10. Örneklem Büyüklüğü 1000 iken Cox Regresyon Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).

		X=10			X=20		
		Eğitim	i	Validasyon	Eğitim	i	Validasyon
<u>T=60</u>							
Sansür Oranı	0,10	0,845 [0,844;0,847]	100	0,848 [0,843;0,852]	0,872 [0,871;0,874]	100	0,882 [0,877;0,887]
	0,25	0,833 [0,831;0,835]	100	0,851 [0,845;0,856]	0,860 [0,859;0,862]	100	0,877 [0,873;0,881]
	0,50	0,820 [0,817;0,822]	100	0,846 [0,839;0,853]	0,837 [0,833;0,840]	100	0,874 [0,869;0,880]
	0,75	0,801 [0,797;0,806]	100	0,852 [0,843;0,861]	0,824 [0,820;0,828]	100	0,869 [0,860;0,877]
	0,90	0,804 [0,798;0,810]	100	0,833 [0,817;0,850]	0,823 [0,818;0,829]	98	0,824 [0,805;0,843]
<u>T=120</u>							
Sansür Oranı	0,10	0,843 [0,841;0,844]	100	0,848 [0,844;0,853]	0,871 [0,869;0,872]	100	0,875 [0,871;0,880]
	0,25	0,834 [0,833;0,836]	100	0,840 [0,835;0,846]	0,859 [0,858;0,861]	100	0,873 [0,868;0,878]
	0,50	0,823 [0,820;0,825]	100	0,839 [0,833;0,846]	0,845 [0,842;0,847]	100	0,870 [0,865;0,876]
	0,75	0,812 [0,808;0,816]	100	0,837 [0,827;0,848]	0,832 [0,828;0,836]	100	0,869 [0,861;0,877]
	0,90	0,810 [0,805;0,814]	100	0,833 [0,816;0,850]	0,832 [0,826;0,837]	99	0,846 [0,830;0,862]

i:tekrar sayısı

Bağımsız değişken sayısı 10 ve sansür oranı 0,10 ile 0,90 iken sonuçlar çok benzer olmakla birlikte, diğer sansür oranlarında sağkalım süresi 60 olduğu durumda 120 olduğu duruma göre sonuçların daha yüksek olduğu görülmüştür. Bağımsız değişken sayısı 20 ve sansür oranı $\leq 0,50$ iken, sağkalım süresi 60 olduğu durumda 120 olduğu duruma göre sonuçlar bir miktar daha yüksek bulunmuştur. Sağkalım süresi 60 ve sansür oranının 0,90 olduğu senaryo dışında, bağımsız değişken sayısı 20 iken 10 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir.

Örneklem büyüklüğü 1000 iken yalnızca iki senaryoda 100 tekrardan daha az sonuç alınmıştır. Bu nedenle ikinci bir çizelge yapılmasına gerek duyulmamış olup, yalnızca bu senaryolara ilişkin sonuç alınamayan tekrarlarda C-indeks değeri “0” kabul edilip hesaplanan sonuçlar sunulmuştur. Buna göre sonuçlar, sağkalım süresi 60, bağımsız değişken sayısı 20 ve sansür oranının 0,90 iken 0,808 [0,778;0,837]; sağkalım süresi 120, bağımsız değişken sayısı 20 ve sansür oranının 0,90 iken 0,838 [0,814;0,861] dir.

3.1.9. Örneklem Büyüklüğü 10000 iken Cox Regresyon Sonuçları (U(-1,1))

Örneklem büyüklüğü 10000 iken Cox regresyon sonuçları Çizelge 3.11’de gösterildiği gibidir.

Çizelge 3.11. Örneklem Büyüklüğü 10000 iken Cox Regresyon Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (U(-1,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	T=60				
	0,10	0,844 [0,844;0,845]	0,852 [0,850;0,854]	0,872 [0,871;0,873]	0,882 [0,880;0,883]
	0,25	0,833 [0,832;0,835]	0,852 [0,849;0,854]	0,858 [0,856;0,859]	0,881 [0,879;0,883]
	0,50	0,817 [0,814;0,819]	0,851 [0,848;0,854]	0,836 [0,833;0,839]	0,882 [0,880;0,884]
	0,75	0,803 [0,800;0,806]	0,853 [0,850;0,857]	0,820 [0,816;0,824]	0,879 [0,877;0,882]
	0,90	0,795 [0,790;0,800]	0,848 [0,843;0,854]	0,811 [0,807;0,816]	0,879 [0,875;0,884]
	T=120				
Sansür Oranı	0,10	0,841 [0,840;0,842]	0,846 [0,844;0,848]	0,870 [0,869;0,871]	0,877 [0,875;0,879]
	0,25	0,834 [0,833;0,835]	0,846 [0,844;0,848]	0,859 [0,858;0,860]	0,878 [0,876;0,879]
	0,50	0,821 [0,819;0,823]	0,849 [0,846;0,851]	0,843 [0,841;0,846]	0,876 [0,874;0,877]
	0,75	0,810 [0,807;0,813]	0,846 [0,842;0,849]	0,826 [0,822;0,830]	0,876 [0,873;0,879]
	0,90	0,805 [0,802;0,808]	0,844 [0,839;0,849]	0,818 [0,814;0,822]	0,876 [0,872;0,880]

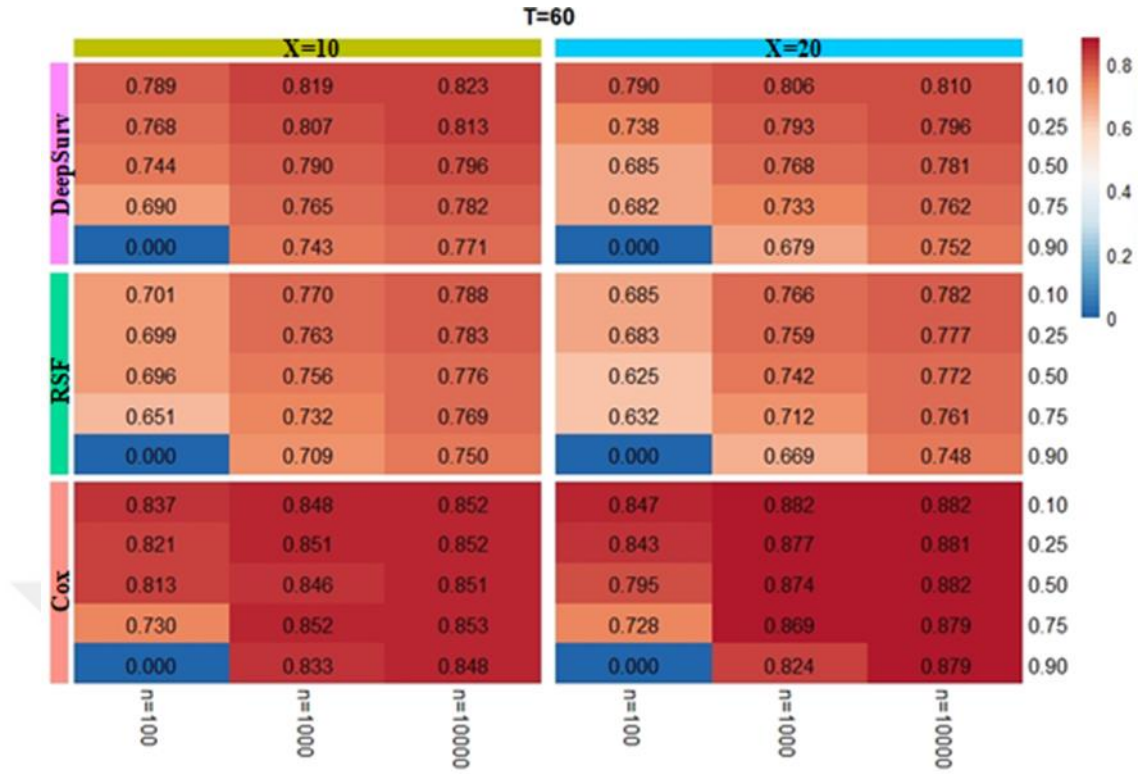
Örneklem büyüklüğü 10000 iken Cox regresyon analizi sonuçları incelendiğinde sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,795 ve 0,872; validasyon setinde ise 0,844 ve 0,882 olarak kaydedilmiştir. Validasyon setindeki en yüksek C-indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 20 ve sansür oranının 0,10 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 120, bağımsız değişken sayısının 10 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir.

Burada sansür oranı artması/azalması ile C-indeks değeri arasında bir ilişki gözlemlenememiştir. Her iki sağkalım süresinde de bağımsız değişken sayısı 20 iken 10 olduğu duruma göre sonuçlar daha yüksek bulunmuştur. Bağımsız değişken

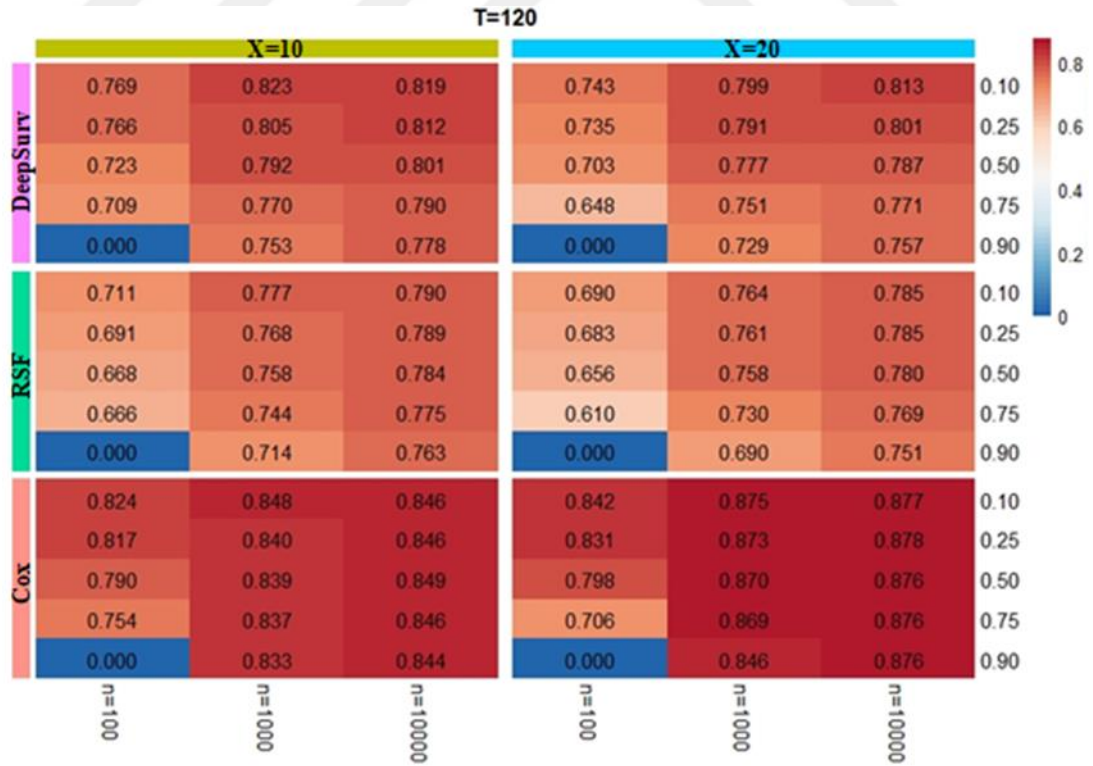
sayısının 10 ve 20 olduđu her iki durumda da sađkalım süresi 60 iken 120 olduđu duruma göre sonuçlar bir miktar daha yüksek olsa da, aslında sonuçların çok benzer olduđu görölmüştür. Özetle, sađkalım süresinin örneklem sayısı yüksek iken bir etkisinin olmadığı saptanmıştır.

3.1.10. DeepSurv, RSF ve Cox Regresyon Yöntemlerinin Karşılaştırılması (U(-1,1))

Sađkalım süresi 60 iken DeepSurv ve Cox regresyon yöntemlerinin karşılaştırılmasına ilişkin ısı-harita grafiđi Şekil 3.2’de verilmiştir. Buna göre, DeepSurv ve RSF yönteminde bağımsız deđişken sayısı 10 iken 20 olduđu duruma göre daha yüksek sonuçlar elde edilmiştir. Ancak, Cox regresyon yönteminde tam tersi gözlemlenmiş olup bağımsız deđişken sayısı 20 olduđu durumda 10 olduđu duruma göre sonuçların daha yüksek olduđu görölmüştür. DeepSurv ve RSF yönteminde tüm senaryolarda sansür oranı arttıkça C-indeks deđerlerinin azaldığı, Cox regresyon yönteminde ise sansür oranındaki deđişimin (artış/azalış) C-indeks deđerini etkilemediđi sonucuna varılmıştır. Her üç yöntemde de örneklem büyüklüğü arttıkça C-indeks deđerinin arttığı ve Cox regresyon sonuçlarının diđer iki yöntemle göre daha yüksek olduđu saptanmıştır. DeepSurv yönteminden elde edilen sonuçların ise RSF yönteminden elde edilen sonuçlara göre daha yüksek olduđu gözlemlenmiştir.



Şekil 3.2. Sağkalım Süresi 60 iken Yöntemlerin Karşılaştırılması ($U(-1,1)$).

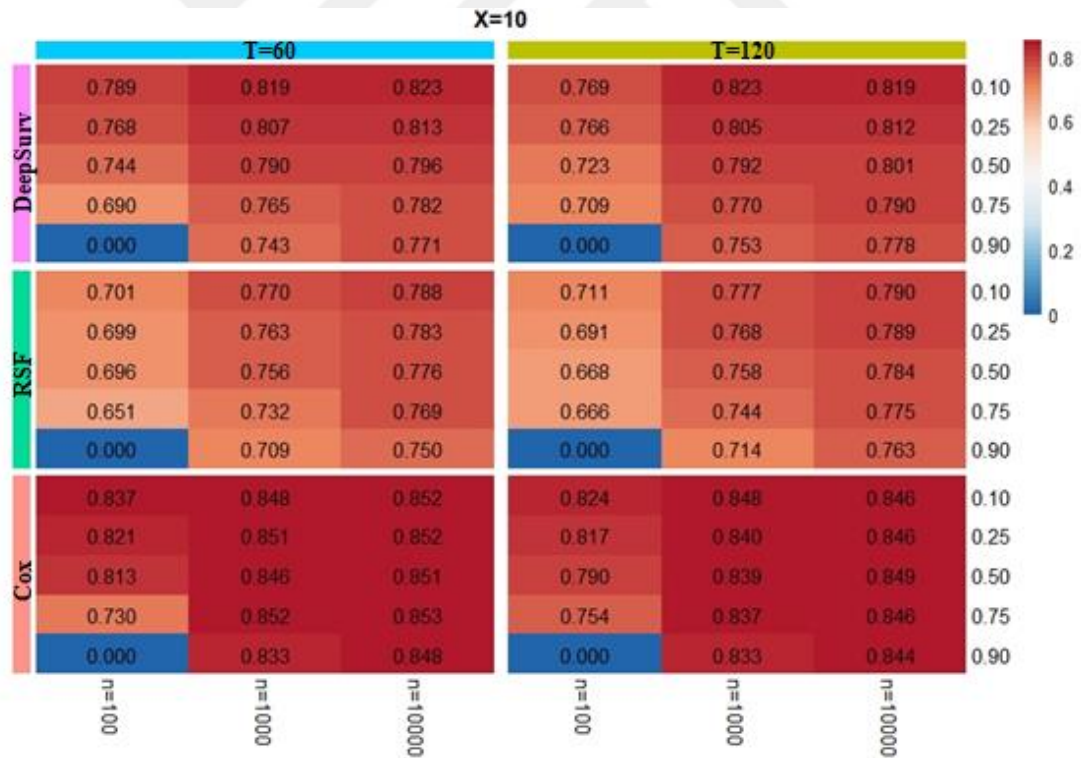


Şekil 3.3. Sağkalım Süresi 120 iken Yöntemlerin Karşılaştırılması ($U(-1,1)$).

Sağkalım süresi 120 iken bu üç yöntemin karşılaştırılmasına ilişkin yorumlar sağkalım süresinin 60 olduğu durumdakine oldukça benzerdir (Şekil 3.3).

Şekil 3.2 ve Şekil 3.3 birlikte incelendiğinde, Cox regresyon sonuçları sağkalım süresi 120 iken sağkalım süresi 60 olduğu duruma göre bir miktar daha düşük bulunmuştur.

DeepSurv yönteminde sansür oranı $\geq 0,50$ iken sağkalım süresi 120 olduğu durumda 60 olduğu duruma göre bir miktar daha yüksek sonuçlar elde edilmiştir. RSF yönteminde örneklem büyüklüğü 1000 ve 10000 iken sağkalım süresi 120 olduğu durumda 60 olduğu duruma göre sonuçlar bir miktar daha yüksektir. Ancak, Cox regresyon yönteminde tüm sansür oranlarında sağkalım süresi 60 iken 120 olduğu duruma göre sonuçların daha yüksek olduğu görülmüştür (Şekil 3.4).



Şekil 3.4. Bağımsız Değişken sayısı 10 iken Yöntemlerin Karşılaştırılması (U(-1,1)).

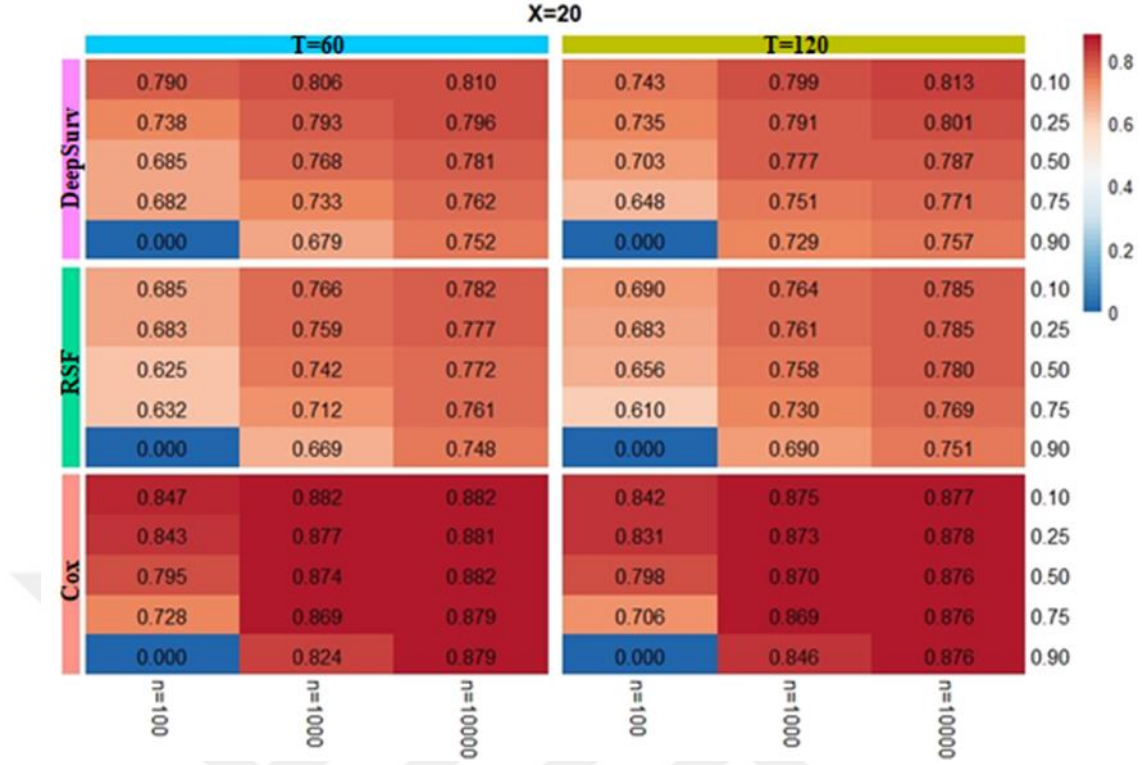
DeepSurv yönteminde örneklem büyüklüğü 100 ve sansür oranı 0,25 iken, örneklem büyüklüğü 1000 ve sansür oranı $\geq 0,50$ iken ve örneklem büyüklüğü 10000

iken tüm sansür oranlarında sağkalım süresi 120 iken 60 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir. RSF yönteminde örneklem büyüklüğü 100 ve sansür oranı $\leq 0,50$ iken, örneklem büyüklüğü 1000 ve sansür oranı $\geq 0,25$ iken ve örneklem büyüklüğü 10000 iken tüm sansür oranlarında sağkalım süresi 120 iken 60 olduğu duruma göre daha yüksek sonuçlar bulunmuştur. Cox regresyon yönteminde ise tüm örneklem büyüklüklerinde sağkalım süresi 60 olduğu durumda 120 olduğu duruma göre daha yüksek sonuçlar saptanmıştır (Şekil 3.5).

Şekil 3.2-3.5 birlikte incelendiğinde, her üç yöntemde de oldukça yüksek C-indeks değerleri kaydedilmekle birlikte, Cox regresyon yönteminde DeepSurv ve RSF yöntemine göre; DeepSurv yönteminde ise RSF yöntemine göre daha yüksek C-indeks değerleri elde edilmiştir.

Cox regresyon yönteminde özellikle örneklem büyüklüğü arttıkça tüm senaryolarda çok benzer sonuçlar olduğu görülmüş olup, senaryolardaki değişimin etkisi bu yöntemde saptanamamıştır. RSF yönteminin de küçük örneklem büyüklüğünde oldukça düşük sonuçlar verdiği görülmüştür.

Ayrıca, sansür oranındaki artışla birlikte C-indeks değerlerinde bir azalma DeepSurv ve RSF yöntemlerinde görülürken, Cox regresyon yönteminde bu durum gözlemlenememiştir ve Cox regresyon yönteminde sansür oranının etkisi olmadığı gibi bir sonuç ortaya çıkmıştır.



Şekil 3.5. Bağımsız Değişken Sayısı 20 iken Yöntemlerin Karşılaştırılması (U(-1,1)).

3.2. Bağımsız Değişkenlere İlişkin Katsayılar $N(0;0,1)$ iken Sonuçlar

3.2.1. Örneklem Büyüklüğü 100 iken DeepSurv Sonuçları ($N(0;0,1)$)

Örneklem büyüklüğü 100 iken DeepSurv sonuçları Çizelge 3.12’de gösterildiği gibidir. Burada da başlık 3’te belirtilen nedenden ötürü, sansür oranının 0,90 olduğu senaryolarda sonuçlar sunulmamıştır.

Çizelge 3.12 incelendiğinde, sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,796 ve 0,986; validasyon setinde 0,513 ve 0,558 olarak kaydedilmiştir. Validasyon setindeki en yüksek C-indeks değeri, sağkalım süresinin 60, bağımsız değişken sayısının 20 ve sansür oranının 0,10 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 20 ve sansür oranının 0,75 olduğu senaryoda gerçekleşmiştir.

Çizelge 3.12. Örneklem Büyüklüğü 100 iken DeepSurv Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	<i>T=60</i>				
	0,10	0,807 [0,800;0,813]	0,534 [0,513;0,556]	0,861 [0,854;0,867]	0,558 [0,542;0,574]
	0,25	0,831 [0,824;0,838]	0,541 [0,519;0,562]	0,880 [0,874;0,886]	0,544 [0,522;0,567]
	0,50	0,886 [0,880;0,892]	0,534 [0,507;0,560]	0,933 [0,928;0,938]	0,513 [0,486;0,539]
	0,75	0,968 [0,963;0,976]	0,531 [0,490;0,572]	0,983 [0,980;0,986]	0,547 [0,510;0,584]
	0,90	-	-	-	-
Sansür Oranı	<i>T=120</i>				
	0,10	0,796 [0,790;0,801]	0,549 [0,528;0,570]	0,854 [0,848;0,860]	0,547 [0,528;0,567]
	0,25	0,826 [0,820;0,833]	0,541 [0,520;0,562]	0,883 [0,877;0,889]	0,553 [0,530;0,577]
	0,50	0,887 [0,881;0,893]	0,543 [0,511;0,575]	0,934 [0,929;0,939]	0,547 [0,523;0,570]
	0,75	0,965 [0,960;0,970]	0,553 [0,510;0,596]	0,986 [0,983;0,989]	0,538 [0,503;0,572]
	0,90	-	-	-	-

Sonuçlar tüm senaryolarda çok benzer olmakla birlikte oldukça düşük bulunmuştur. Çizelge 3.12 incelendiğinde, bağımsız değişken sayısı 10 iken sağkalım süresi 120 olduğu durumda 60 olduğu duruma göre sonuçların bir miktar daha yüksek olduğu görülmüştür. Sansür oranındaki değişim ile C-indeks artış ya da azalışına ilişkin bir yargıya varılamamıştır.

3.2.2. Örneklem Büyüklüğü 1000 iken DeepSurv Sonuçları (N(0;0,1))

Örneklem büyüklüğü 1000 iken DeepSurv yönteminde sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,597 ve 0,831; validasyon setinde 0,525 ve 0,573 olarak kaydedilmiştir. Validasyon setindeki en yüksek C-indeks değeri, sağkalım süresinin 120, bağımsız değişken sayısının 20 ve sansür oranının 0,10 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 120, bağımsız değişken sayısının 10 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir (Çizelge 3.13).

Çizelge 3.13. Örneklem Büyüklüğü 1000 iken DeepSurv Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	<u>T=60</u>				
	0,10	0,598 [0,594;0,602]	0,553 [0,544;0,562]	0,621 [0,617;0,624]	0,570 [0,563;0,578]
	0,25	0,605 [0,601;0,609]	0,551 [0,540;0,561]	0,623 [0,618;0,627]	0,565 [0,557;0,574]
	0,50	0,617 [0,612;0,622]	0,552 [0,542;0,563]	0,644 [0,639;0,649]	0,563 [0,553;0,573]
	0,75	0,661 [0,656;0,665]	0,541 [0,525;0,557]	0,700 [0,695;0,705]	0,554 [0,540;0,569]
	0,90	0,769 [0,762;0,776]	0,547 [0,519;0,575]	0,831 [0,823;0,838]	0,563 [0,540;0,587]
Sansür Oranı	<u>T=120</u>				
	0,10	0,597 [0,593;0,601]	0,551 [0,542;0,560]	0,624 [0,619;0,628]	0,573 [0,564;0,581]
	0,25	0,605 [0,601;0,609]	0,566 [0,555;0,576]	0,626 [0,621;0,631]	0,571 [0,561;0,581]
	0,50	0,619 [0,615;0,623]	0,552 [0,541;0,562]	0,647 [0,642;0,651]	0,561 [0,551;0,571]
	0,75	0,664 [0,659;0,669]	0,547 [0,533;0,561]	0,703 [0,698;0,708]	0,564 [0,548;0,579]
	0,90	0,769 [0,762;0,775]	0,525 [0,499;0,550]	0,824 [0,818;0,831]	0,529 [0,505;0,553]

Sağkalım süresinin 60 ve 120 olduğu her iki durumda da bağımsız değişken sayısı 20 iken 10 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir. Sağkalım süresi 120 ve bağımsız değişken sayısı 20 iken sansür oranı arttıkça C-indeks değerlerinin azaldığı gözlemlenmiştir.

3.2.3. Örneklem Büyüklüğü 10000 iken DeepSurv Sonuçları (N(0;0,1))

Örneklem büyüklüğü 10000 iken DeepSurv yönteminde sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,564 ve 0,612; validasyon setinde 0,548 ve 0,577 olarak kaydedilmiştir. Validasyon setindeki en yüksek C-indeks değeri, sağkalım süresinin 120, bağımsız değişken sayısının 20 ve sansür oranının 0,25 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 120, bağımsız değişken sayısının 10 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir (Çizelge 3.14).

Çizelge 3.14. Örneklem Büyüklüğü 10000 iken Deepsurv Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	<u>T=60</u>				
	0,10	0,565 [0,561;0,569]	0,558 [0,553;0,563]	0,581 [0,577;0,584]	0,574 [0,570;0,578]
	0,25	0,565 [0,562;0,569]	0,557 [0,552;0,561]	0,584 [0,580;0,587]	0,575 [0,570;0,579]
	0,50	0,566 [0,563;0,570]	0,555 [0,550;0,560]	0,582 [0,579;0,585]	0,570 [0,566;0,574]
	0,75	0,572 [0,569;0,575]	0,555 [0,549;0,560]	0,590 [0,587;0,593]	0,571 [0,566;0,577]
	0,90	0,589 [0,585;0,592]	0,559 [0,551;0,568]	0,608 [0,604;0,612]	0,567 [0,559;0,575]
Sansür Oranı	<u>T=120</u>				
	0,10	0,567 [0,564;0,571]	0,561 [0,556;0,565]	0,581 [0,577;0,584]	0,571 [0,567;0,576]
	0,25	0,564 [0,560;0,568]	0,559 [0,554;0,564]	0,585 [0,581;0,589]	0,577 [0,573;0,581]
	0,50	0,568 [0,565;0,572]	0,560 [0,555;0,565]	0,582 [0,578;0,586]	0,570 [0,565;0,575]
	0,75	0,571 [0,567;0,574]	0,554 [0,548;0,559]	0,592 [0,589;0,596]	0,569 [0,563;0,574]
	0,90	0,588 [0,585;0,592]	0,548 [0,540;0,556]	0,612 [0,608;0,616]	0,569 [0,561;0,577]

Sağkalım süresinin 60 ve 120 olduğu her iki durumda da bağımsız değişken sayısı 20 olduğu durumda 10 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir.

3.2.4. Örneklem Büyüklüğü 100 iken RSF Sonuçları (N(0;0,1))

Örneklem büyüklüğü 100 iken RSF sonuçları Çizelge 3.15'te gösterildiği gibidir. Buna göre, sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,453 ve 0,523; validasyon setinde 0,505 ve 0,558 olarak kaydedilmiştir. Validasyon setindeki en yüksek C-indeks değeri, sağkalım süresinin 120, bağımsız değişken sayısının 20 ve sansür oranının 0,25 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 120, bağımsız değişken sayısının 20 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir (Çizelge 3.15).

Çizelge 3.15. Örneklem Büyüklüğü 100 iken RSF Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	<i>T=60</i>				
	0,10	0,506 [0,494;0,517]	0,534 [0,515;0,553]	0,509 [0,496;0,522]	0,556 [0,540;0,573]
	0,25	0,509 [0,495;0,522]	0,534 [0,512;0,555]	0,504 [0,491;0,518]	0,555 [0,536;0,574]
	0,50	0,486 [0,471;0,502]	0,551 [0,524;0,577]	0,506 [0,490;0,521]	0,542 [0,518;0,567]
	0,75	0,476 [0,456;0,496]	0,526 [0,489;0,563]	0,453 [0,432;0,473]	0,552 [0,516;0,589]
	0,90	-	-	-	-
	<i>T=120</i>				
Sansür Oranı	0,10	0,523 [0,511;0,535]	0,543 [0,524;0,563]	0,514 [0,502;0,525]	0,544 [0,524;0,564]
	0,25	0,496 [0,483;0,508]	0,529 [0,508;0,550]	0,508 [0,496;0,520]	0,558 [0,537;0,579]
	0,50	0,483 [0,469;0,497]	0,531 [0,506;0,556]	0,497 [0,480;0,513]	0,523 [0,498;0,548]
	0,75	0,466 [0,445;0,487]	0,514 [0,472;0,556]	0,455 [0,435;0,474]	0,505 [0,471;0,540]
	0,90	-	-	-	-

Örneklem büyüklüğü 100 iken RSF yönteminden elde edilen sonuçların tüm senaryolarda oldukça düşük olmakla birlikte genel olarak benzer olduğu söylenebilir.

3.2.5. Örneklem Büyüklüğü 1000 iken RSF Sonuçları (N(0;0,1))

Örneklem büyüklüğü 1000 iken RSF yönteminde, eğitim ve validasyon setinin her ikisi için de sırasıyla en küçük ve en büyük C-indeks değerleri aynı senaryolarda kaydedilmiş olup bu değerler 0,515 ve 0,593 tür. Buna göre, en yüksek C-indeks değeri, sağkalım süresinin 120, bağımsız değişken sayısının 20 ve sansür oranının 0,10 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 10 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir (Çizelge 3.16).

Çizelge 3.16. Örneklem Büyüklüğü 1000 iken RSF Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	<u>T=60</u>				
	0,10	0,562 [0,557;0,567]	0,559 [0,551;0,567]	0,587 [0,584;0,591]	0,589 [0,581;0,596]
	0,25	0,561 [0,556;0,565]	0,561 [0,552;0,570]	0,582 [0,578;0,586]	0,584 [0,576;0,591]
	0,50	0,550 [0,544;0,556]	0,546 [0,535;0,556]	0,572 [0,567;0,578]	0,576 [0,567;0,586]
	0,75	0,532 [0,527;0,538]	0,543 [0,528;0,559]	0,553 [0,546;0,560]	0,563 [0,548;0,577]
	0,90	0,515 [0,506;0,524]	0,515 [0,489;0,542]	0,535 [0,526;0,544]	0,541 [0,519;0,563]
	<u>T=120</u>				
Sansür Oranı	0,10	0,562 [0,557;0,566]	0,559 [0,550;0,567]	0,593 [0,589;0,598]	0,593 [0,585;0,601]
	0,25	0,564 [0,559;0,569]	0,568 [0,560;0,577]	0,586 [0,581;0,591]	0,592 [0,583;0,600]
	0,50	0,554 [0,549;0,560]	0,552 [0,543;0,562]	0,576 [0,571;0,582]	0,578 [0,567;0,589]
	0,75	0,539 [0,532;0,545]	0,550 [0,536;0,564]	0,554 [0,549;0,560]	0,575 [0,559;0,591]
	0,90	0,529 [0,520;0,538]	0,533 [0,509;0,558]	0,530 [0,520;0,540]	0,534 [0,509;0,559]

Bağımsız değişken sayısının 20, sağkalım süresinin 60 ve sansür oranının 0,90 olduğu senaryo dışındaki tüm senaryolarda sağkalım süresi 120 iken elde edilen sonuçlar sağkalım süresi 60 iken elde edilen sonuçlara göre bir miktar daha yüksek bulunmuştur. Her iki sağkalım süresinde de, bağımsız değişken sayısı 20 olduğu durumda 10 olduğu duruma göre sonuçların daha yüksek olduğu görülmüştür.

3.2.6. Örneklem Büyüklüğü 10000 iken RSF Sonuçları (N(0;0,1))

Örneklem büyüklüğü 10000 iken RSF sonuçları Çizelge 3.17’de gösterildiği gibidir. Buna göre, sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,544 ve 0,605; validasyon setinde 0,550 ve 0,608 olarak kaydedilmiştir. Validasyon setindeki en yüksek C-indeks değeri, sağkalım süresinin 120, bağımsız değişken sayısının 20 ve sansür oranının 0,25 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 120, bağımsız değişken sayısının 10 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir (Çizelge 3.17).

Çizelge 3.17. Örneklem Büyüklüğü 10000 iken RSF Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	<u>T=60</u>				
	0,10	0,571 [0,567;0,575]	0,572 [0,568;0,577]	0,601 [0,598;0,605]	0,604 [0,600;0,607]
	0,25	0,570 [0,566;0,573]	0,570 [0,566;0,574]	0,602 [0,598;0,605]	0,604 [0,600;0,608]
	0,50	0,565 [0,562;0,569]	0,565 [0,560;0,570]	0,594 [0,591;0,597]	0,595 [0,591;0,599]
	0,75	0,557 [0,553;0,561]	0,562 [0,557;0,568]	0,589 [0,585;0,592]	0,594 [0,589;0,600]
	0,90	0,544 [0,540;0,548]	0,551 [0,544;0,559]	0,568 [0,563;0,572]	0,580 [0,572;0,588]
	<u>T=120</u>				
Sansür Oranı	0,10	0,574 [0,571;0,578]	0,577 [0,573;0,581]	0,601 [0,598;0,604]	0,602 [0,598;0,606]
	0,25	0,570 [0,566;0,574]	0,572 [0,568;0,577]	0,605 [0,602;0,609]	0,608 [0,604;0,612]
	0,50	0,568 [0,564;0,571]	0,571 [0,566;0,575]	0,598 [0,594;0,601]	0,598 [0,593;0,603]
	0,75	0,558 [0,554;0,562]	0,557 [0,552;0,563]	0,590 [0,586;0,594]	0,593 [0,588;0,599]
	0,90	0,547 [0,543;0,551]	0,550 [0,541;0,558]	0,575 [0,570;0,580]	0,581 [0,573;0,589]

Örneklem büyüklüğü 10000 iken sonuçlar çok benzer olmakla birlikte, tüm senaryolarda sansür oranı arttıkça C-indeks değerinin azaldığı gözlemlenmiştir. Her iki sağkalım süresinde de, bağımsız değişken sayısı 20 olduğu durumda 10 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir.

3.2.7. Örneklem Büyüklüğü 100 iken Cox Regresyon Sonuçları (N(0;0,1))

Örneklem büyüklüğü 100 iken Cox regresyon sonuçları Çizelge 3.18’de gösterildiği gibidir. Sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,642 ve 0,828; validasyon setinde 0,516 ve 0,571 olarak kaydedilmiştir. Validasyon setindeki en yüksek C-indeks değeri, sağkalım süresinin 120, bağımsız değişken sayısının 10 ve sansür oranının 0,10 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 10 ve sansür oranının 0,75 olduğu senaryoda gerçekleşmiştir.

Çizelge 3.18. Örneklem Büyüklüğü 1000 iken Cox Regresyon Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).

		X=10			X=20		
		Eğitim	i	Validasyon	Eğitim	i	Validasyon
<u>T=60</u>							
Sansür Oranı	0,10	0,642 [0,635;0,649]	100	0,553 [0,531;0,574]	0,704 [0,697;0,712]	100	0,564 [0,546;0,582]
	0,25	0,657 [0,648;0,665]	100	0,544 [0,523;0,565]	0,709 [0,701;0,718]	100	0,570 [0,549;0,590]
	0,50	0,662 [0,653;0,671]	100	0,547 [0,519;0,575]	0,748 [0,740;0,757]	100	0,544 [0,515;0,573]
	0,75	0,721 [0,710;0,733]	90	0,516 [0,480;0,552]	0,828 [0,815;0,840]	90	0,536 [0,505;0,566]
	0,90	-	-	-	-	-	-
<u>T=120</u>							
Sansür Oranı	0,10	0,650 [0,643;0,657]	100	0,571 [0,551;0,591]	0,704 [0,698;0,711]	100	0,560 [0,538;0,581]
	0,25	0,642 [0,635;0,650]	100	0,560 [0,537;0,583]	0,711 [0,704;0,718]	100	0,560 [0,537;0,584]
	0,50	0,664 [0,656;0,672]	100	0,544 [0,514;0,574]	0,742 [0,734;0,750]	100	0,542 [0,518;0,567]
	0,75	0,719 [0,707;0,731]	89	0,524 [0,484;0,564]	0,827 [0,816;0,839]	99	0,533 [0,497;0,568]
	0,90	-	-	-	-	-	-
		i:tekrar sayısı					

Bağımsız değişken sayısı 10 ve sansür oranı 0,10; 0,25 ile 0,75 iken sağkalım süresi 60 olduğu durumda 120 olduğu duruma göre sonuçların daha yüksek olduğu görülmüştür. Bağımsız değişken sayısı 20 ve sansür oranı $\leq 0,50$ iken, sağkalım süresi 60 olduğu durumda 120 olduğu duruma göre sonuçlar bir miktar daha yüksek bulunmuştur.

Örneklem büyüklüğü 100 iken yalnızca sansür oranının 0,75 olduğu dört senaryoda 100 tekrardan daha az sonuç alınmıştır. Bu nedenle ikinci bir çizelge yapılmasına gerek duyulmamış olup, yalnızca bu senaryolara ilişkin sonuç alınamayan tekrarlarda C-indeks değeri “0” kabul edilip hesaplanan sonuçlar sunulmuştur. Buna göre sansür oranı 0,75 olmak üzere sonuçlar, sağkalım süresi 60 ve bağımsız değişken sayısı 10 iken 0,465 [0,420;0,509]; sağkalım süresi 60 ve bağımsız değişken sayısı 20 iken 0,487 [0,445;0,529]; sağkalım süresi 120 ve bağımsız değişken sayısı 10 iken 0,466 [0,418;0,514]; sağkalım süresi 120 ve bağımsız değişken sayısı 20 iken 0,527 [0,491;0,564] olarak bulunmuştur.

3.2.8. Örneklem Büyüklüğü 1000 iken Cox Regresyon Sonuçları (N(0;0,1))

Örneklem büyüklüğü 1000 iken sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,591 ve 0,670; validasyon setinde 0,533 ve 0,616 olarak kaydedilmiştir. Validasyon setindeki en yüksek C-indeks değeri, sağkalım süresinin 120, bağımsız değişken sayısının 20 ve sansür oranının 0,25 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 60, bağımsız değişken sayısının 10 ve sansür oranının 0,90 olduğu senaryoda gerçekleşmiştir (Çizelge 3.19).

Çizelge 3.19. Örneklem Büyüklüğü 1000 iken Cox Regresyon Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
<u>T=60</u>					
Sansür Oranı	0,10	0,594 [0,590;0,598]	0,577 [0,569;0,585]	0,628 [0,625;0,632]	0,612 [0,603;0,620]
	0,25	0,596 [0,592;0,600]	0,584 [0,574;0,594]	0,626 [0,622;0,630]	0,608 [0,599;0,616]
	0,50	0,593 [0,588;0,598]	0,570 [0,559;0,582]	0,628 [0,623;0,632]	0,605 [0,595;0,615]
	0,75	0,595 [0,591;0,600]	0,565 [0,548;0,582]	0,635 [0,631;0,640]	0,591 [0,574;0,608]
	0,90	0,616 [0,611;0,622]	0,533 [0,508;0,559]	0,670 [0,665;0,675]	0,583 [0,559;0,606]
<u>T=120</u>					
Sansür Oranı	0,10	0,591 [0,587;0,595]	0,575 [0,566;0,584]	0,632 [0,628;0,636]	0,614 [0,605;0,622]
	0,25	0,595 [0,590;0,599]	0,591 [0,582;0,600]	0,630 [0,626;0,634]	0,616 [0,607;0,625]
	0,50	0,594 [0,589;0,598]	0,575 [0,565;0,585]	0,631 [0,626;0,635]	0,607 [0,598;0,617]
	0,75	0,598 [0,594;0,602]	0,567 [0,552;0,582]	0,637 [0,633;0,642]	0,596 [0,582;0,609]
	0,90	0,627 [0,620;0,634]	0,551 [0,526;0,576]	0,666 [0,660;0,673]	0,556 [0,532;0,581]

Çizelge 3.19 incelendiğinde, sağkalım süresinin 60 ve 120 olduğu her iki durumda da bağımsız değişken sayısı 20 olduğu durumda 10 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir. Bağımsız değişken sayısı 20 iken sağkalım süresi 120 iken 60 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir. Bağımsız değişken sayısı 10 ve sansür oranı $>0,25$ olduğu senaryolarda sağkalım süresi 120 iken 60 olduğu duruma göre sonuçların bir miktar daha yüksek olduğu görülmüştür. Bağımsız değişken sayısı 20 iken, sansür oranı arttıkça C-indeks

değerlerinin azaldığı sonucuna varılmıştır. Bağımsız değişken sayısı 10 iken ise sansür oranı 0,50'nin üzerine çıktıkça C-indeks değerinin azaldığı gözlemlenmiştir.

3.2.9. Örneklem Büyüklüğü 10000 iken Cox Regresyon Sonuçları (N(0;0,1))

Örneklem büyüklüğü 10000 iken sırasıyla en küçük ve en büyük C-indeks değerleri eğitim setinde 0,583 ve 0,622; validasyon setinde 0,581 ve 0,625 olarak kaydedilmiştir. Validasyon setindeki en yüksek C-indeks değeri, sağkalım süresinin 120, bağımsız değişken sayısının 20 ve sansür oranının 0,25 olduğu senaryoda gerçekleşirken, en düşük C-indeks değeri sağkalım süresinin 120, bağımsız değişken sayısının 10 ve sansür oranının 0,75 olduğu senaryoda gerçekleşmiştir (Çizelge 3.20).

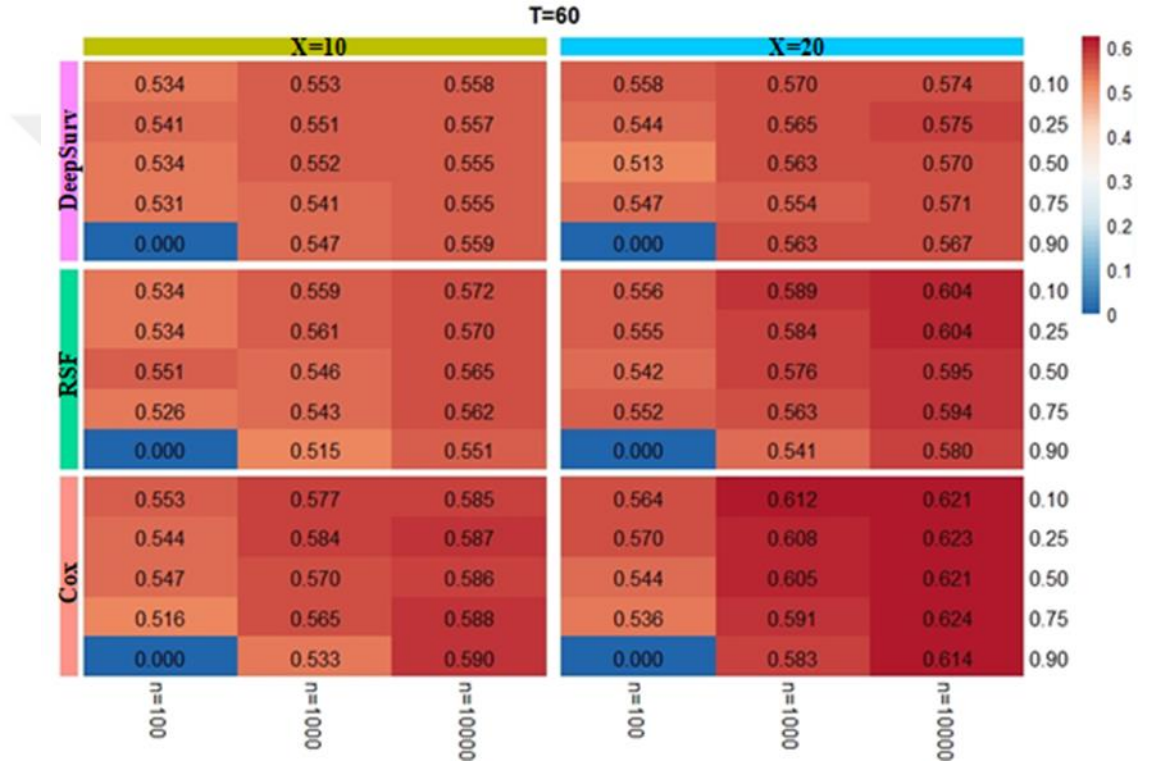
Çizelge 3.20. Örneklem Büyüklüğü 10000 iken Cox Regresyon Yöntemine ait C-indeks [%95 Güven Aralığı] Sonuçları (N(0;0,1)).

		X=10		X=20	
		Eğitim	Validasyon	Eğitim	Validasyon
Sansür Oranı	<u>T=60</u>				
	0,10	0,585 [0,581;0,589]	0,585 [0,580;0,590]	0,620 [0,617;0,623]	0,621 [0,616;0,625]
	0,25	0,585 [0,581;0,588]	0,587 [0,583;0,592]	0,620 [0,616;0,624]	0,623 [0,619;0,628]
	0,50	0,583 [0,579;0,587]	0,586 [0,581;0,591]	0,615 [0,612;0,618]	0,621 [0,617;0,625]
	0,75	0,584 [0,580;0,587]	0,588 [0,582;0,593]	0,617 [0,614;0,621]	0,624 [0,618;0,629]
	0,90	0,584 [0,580;0,587]	0,590 [0,582;0,599]	0,616 [0,612;0,619]	0,614 [0,606;0,622]
	<u>T=120</u>				
Sansür Oranı	0,10	0,587 [0,584;0,591]	0,588 [0,584;0,592]	0,618 [0,614;0,621]	0,617 [0,612;0,621]
	0,25	0,583 [0,580;0,587]	0,585 [0,581;0,590]	0,622 [0,618;0,625]	0,625 [0,620;0,629]
	0,50	0,585 [0,582;0,589]	0,587 [0,582;0,592]	0,616 [0,612;0,620]	0,618 [0,613;0,622]
	0,75	0,583 [0,580;0,586]	0,581 [0,575;0,587]	0,618 [0,614;0,621]	0,621 [0,616;0,627]
	0,90	0,585 [0,581;0,589]	0,581 [0,572;0,590]	0,621 [0,617;0,625]	0,613 [0,605;0,621]

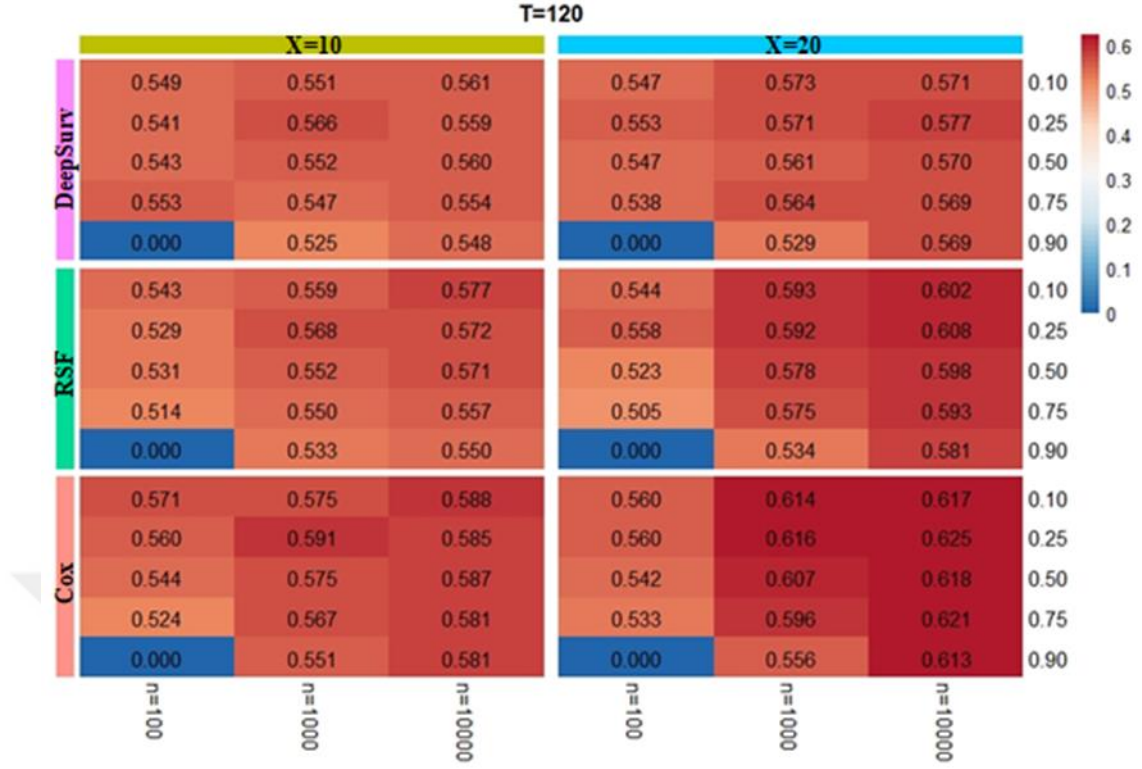
Bağımsız değişken sayısı 10 ve 20 olduğu her iki durumda da sağkalım süresi 60 ve 120 iken benzer sonuçlar elde edilmiştir.

3.2.10. DeepSurv, RSF ve Cox Regresyon Yöntemlerinin Karşılaştırılması (N(0;0,1))

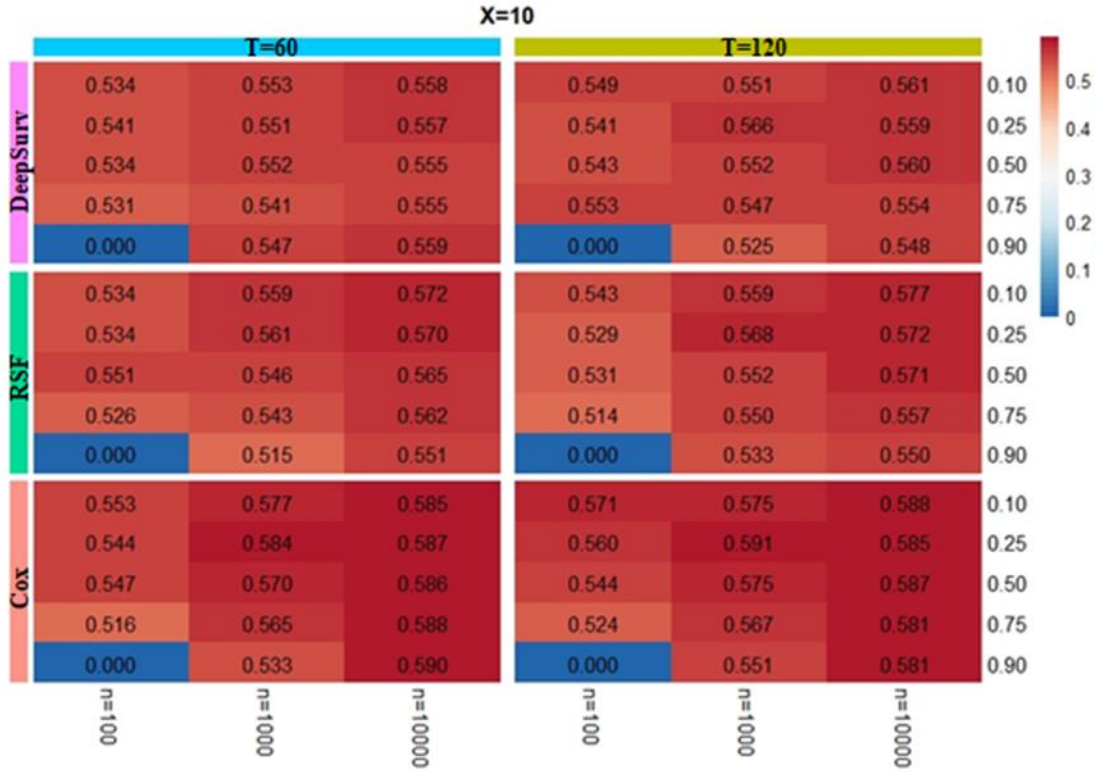
Sağkalım süresinin 60 ve 120 olduğu her iki durumda da DeepSurv, RSF ve Cox regresyon yöntemlerinde tüm örneklem büyüklüklerinde bağımsız değişken sayısı 20 olduğu durumda 10 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir (Şekil 3.6 ve Şekil 3.7).



Şekil 3.6. Sağkalım Süresi 60 iken Yöntemlerin Karşılaştırılması (N(0;0,1)).



Şekil 3.7. Sağkalım Süresi 120 iken Yöntemlerin Karşılaştırılması ($N(0;0,1)$).

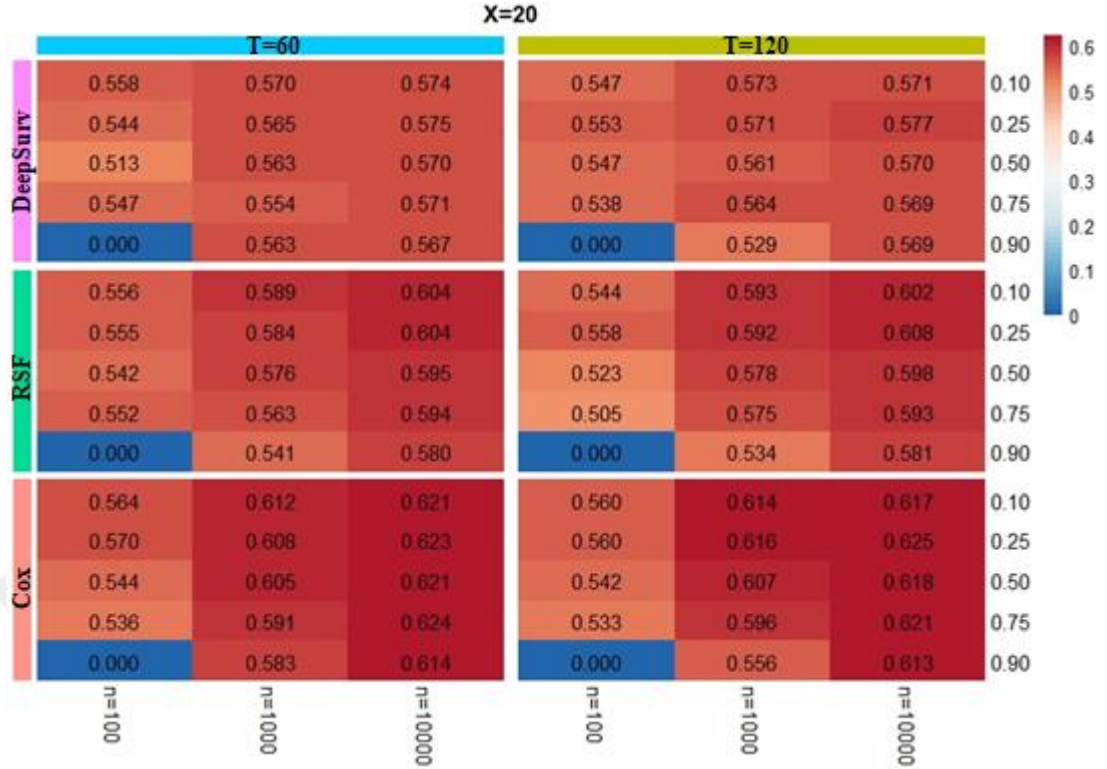


Şekil 3.8. Bağımsız Değişken Sayısı 10 iken Yöntemlerin Karşılaştırılması ($N(0;0,1)$).

Bağımsız değişken sayısı 10 iken, DeepSurv ve Cox regresyon yöntemlerinde genel olarak sansür oranı arttıkça C-indeks değerlerinin azaldığı söylenebilir. RSF yönteminde ise sansür oranındaki artışla birlikte C-indeks değerlerindeki azalma bağımsız değişken sayısının 10 ve 20 olduğu her iki durumda da net olarak gözlemlenmektedir (Şekil 3.8).

Bağımsız değişken sayısı 20 iken sonuçlar incelendiğinde, Cox regresyon yönteminde tüm örneklem büyüklüklerinde sağkalım süresi 60 olduğu durumda 120 olduğu duruma göre daha yüksek sonuçlar elde edilmiştir. RSF yönteminde örneklem büyüklüğü yüksek (1000 ve 10000) iken sağkalım süresi 120 olduğu durumda genel olarak sonuçlar sağkalım süresi 60 olduğu duruma göre bir miktar daha yüksektir. DeepSurv yöntemi için ise özet bir yargıya varılamamış olup, bazı senaryolarda sağkalım süresi 60 iken bazılarında ise sağkalım süresi 120 iken sonuçların daha yüksek olduğu görülmüştür (Şekil 3.9).

Şekil 3.6-3.9 birlikte incelendiğinde ise, Cox regresyon yönteminde elde edilen sonuçların DeepSurv ve RSF yöntemlerine göre bir miktar yüksek olduğu görülmüştür. Ancak, elde edilen C-indeks değerleri 0,50-0,60 civarında olup, her üç yöntemin de performansının oldukça düşük olduğu sonucuna varılmıştır. Çünkü C-indeksinde 0,50 rastgele bir tahmini temsil etmektedir.



Şekil 3.9. Bağımsız Değişken Sayısı 20 iken Yöntemlerin Karşılaştırılması ($N(0;0,1)$).

3.3. Evre I Akciğer Adenokarsinomu Veri Seti Üzerinde DeepSurv, RSF ve Cox Regresyon Yöntemlerinin Karşılaştırılması

Bu çalışmanın temel amacı, nükse etkisi olan risk faktörlerini tespit etmektir. Burada nükse etkisi olduğu düşünülen risk faktörleri; sigara içme durumu, nodülün yapısı, tümör kaybolma oranı, patolojik alt tip, adenokarsinoma, tümör boyutu, patolojik tümör evresi, lenfovasküler ve perinöral invazyon, cerrahi sınır, visseral plevra invazyonu, örneklenen lenf nodu istasyon sayısı, rezeksiyon tipi, SUVmax değeri ve adjuvan tedavi alma durumudur.

Çalışmaya dahil edilen hastaların 107(%69,9)'si erkek, 46(%30,1)'si ise kadındır ve yaşları $64,33 \pm 9,46$ 'dır. Veri ön işleme sürecinde, birbiriyle yüksek ilişkili olan ve kategorilerdeki hasta sayılarında dengesiz dağılım olan değişkenler çalışmadan çıkarılmıştır. Sonuç olarak ise nükse etkisi araştırılan risk faktörleri nodülün yapısı, tümör boyutu, patolojik tümör evresi, örneklenen lenf nodu istasyon

sayısı ve rezeksiyon tipidir. Değişkenlere ilişkin tanımlayıcı istatistikler Çizelge 3.21’de sunulmuştur.

Çizelge 3.21. Değişkenlere İlişkin Tanımlayıcı İstatistikler.

Değişkenler	
Nüks, n(%)	
Var	36(23,5)
Yok	117(76,5)
Takip Süresi, ort±ss	38,76±27,21
ortanca(min-maks)	32,00(1,30-108,00)
Nodülün Yapısı, n(%)	
Solid	102(66,7)
Subsolid	51(33,3)
Tümör Boyutu, ort±ss	19,88±7,57
ortanca(min-maks)	20,00(6,00-45,00)
Patolojik Tümör Evresi, n(%)	
1A	112(73,2)
1B	41(26,8)
Lenf Nodu İstasyon Sayısı, ort±ss	3,91±2,54
ortanca(min-maks)	4,00(0,00-10,00)
Rezeksiyon Tipi, n(%)	
Lobektomi	104(68,0)
Wedge+Segmentektomi	49(32,0)

ort:ortalama, ss:standart sapma, min:minimum, maks:maksimum

Burada, bu tez çalışması kapsamında kullanılan DeepSurv, RSF ve Cox regresyon yöntemleri ile analizler gerçekleştirilmiştir. Veri seti %80 eğitim ve %20 validasyon seti olacak şekilde rastgele ayrılmıştır. Bu işlem, 100 kez tekrarlanmıştır ve bu üç yönteme ilişkin 100 tekrar sonucu ortalama C-indeks değerleri güven aralıkları ile birlikte eğitim ve validasyon setleri için ayrı ayrı Çizelge 3.22’de sunulmuştur.

Çizelge 3.22. Evre I Akciğer Adenokarsinomu Veri Seti ile Yapılan Analizler Sonucu DeepSurv, RSF Ve Cox Regresyon Yöntemlerinin Performansının Değerlendirilmesi.

	Eğitim	Validasyon
DeepSurv	0,803 [0,798;0,807]	0,541 [0,511;0,571]
RSF	0,524 [0,518;0,530]	0,546 [0,534;0,559]
Cox Regresyon	0,632 [0,627;0,637]	0,535 [0,518;0,552]

Çizelge 3.22 incelendiğinde, eğitim setinde en yüksek C-indeks değerine sahip olan yöntem DeepSurv olup, onu takip eden yöntemler sırasıyla Cox regresyon ve RSF yöntemleridir. Validasyon seti sonucuna göre ise sırasıyla en yüksek C-indeks değerine sahip yöntemler RSF, DeepSurv ve Cox regresyon yöntemleridir. Ancak her üç yöntemin performansı da 0,54 civarında olup, validasyon setlerinde yöntemlerin performanslarının benzer olduğu görülmüştür.

Eğitim setinde DeepSurv yönteminin başarısı Cox regresyona göre %27, RSF yöntemine göre ise %53 daha yüksek bulunmuştur. DeepSurv yönteminde eğitim setindeki performansının yüksek olması, bu yöntemin veriyi eğitmede çok başarılı olduğunu göstermektedir. Burada, validasyon setinde “nüks olan” hastaların sayısı daha yüksek olsaydı, validasyon setinden elde edilen C-indeks değerinin daha yüksek olacağı düşünülmektedir.

Sonuç olarak gerçek veri üzerinde yapılan bu çalışma sonucunda, DeepSurv, RSF ve Cox regresyon yöntemlerinin her üçünün de uygulanabileceği sonucuna varılmıştır.

4. TARTIŞMA

Sağlık alanında yaygın olarak kullanılan sağkalım analizinde, klasik istatistiksel yöntemler ve makine öğrenmesine dayalı yöntemler kullanılmaktadır. Makine öğrenmesi yöntemleri genellikle yüksek boyutlu verilere uygulanırken, istatistiksel yöntemler düşük boyutlu verilere uygulanır. Klasik istatistiksel yöntemler ile sağkalım verilerinin analiz edilmesinde hazardların orantısallığı gibi kısıtlayıcı varsayımlar söz konusu iken, makine öğrenmesi yöntemleri bunlarla başa çıkabilirler.

Sağlık alanında, veri kümelerinin boyutunun artması ve makine öğreniminin kullanımına olan ilginin çoğalmasıyla birlikte sinir ağı yaklaşımına dayanan derin öğrenme yöntemleri, sağkalım analizi problemlerini çözme konusunda büyük ilgi görmüştür. Bu tez çalışmasında sağkalım analizine derin öğrenme yaklaşımının uygulanması birincil hedef olmakla birlikte, elde edilen sonuçların makine öğrenmesi yöntemlerinden RSF ve klasik istatistiksel yöntemlerden Cox regresyon ile karşılaştırılması da yapılmıştır. Yöntemlerin performansları Harrell'in C-indeks değeri ile değerlendirilmiştir. Bu kapsamda örneklem büyüklüğü, sağkalım süresi, bağımsız değişken sayısı, sansür oranı ve bağımsız değişkenlere ilişkin katsayıların çeşitli düzeylerinin ele alındığı çok sayıda senaryo oluşturulmuştur. Buna ek olarak, gerçek veri seti kullanılarak da yöntemlerin performansı karşılaştırılmıştır.

Bu çalışmada yer alan senaryoların oluşturulmasında gerçek veri setlerinde karşılaşılabilecek durumlar göz önünde bulundurulmuştur. Burada amaç, hangi senaryoda yaklaşık olarak ne gibi bir sonuçla karşılaşılabilecek sorusuna yanıt vermek ve araştırmacılara, beklentileri doğrultusunda analizlere yön verme konusunda yardımcı olmaktır.

Literatürde derin öğrenmenin sağkalım analizinde kullanılmasına ilişkin çeşitli yöntemler vardır. Bunlardan ilki Katzman ve ark. (2018) tarafından “DeepSurv” olarak adlandırılan yaklaşımdır. Katzman ve ark. yayınladıkları makalede, çeşitli

benzetim ve gerçek sađkalım verileri üzerinde DeepSurv uygulamaları yapmışlar ve sađkalım analizinde sıklıkla kullanılan RSF ve Cox regresyon yöntemleri ile karşılaştırmışlardır. Çalışmanın sonunda, DeepSurv'ün en az diğerk sađkalım modelleri kadar iyi performans gösterdiğini ve DeepSurv'ün ortak değışkenler ile risk fonksiyonu arasındaki karmaşık ilişkileri başarılı bir şekilde modellediğini ortaya koymuşlardır.

İkincisi, Ching ve ark. (2018) tarafından transkriptomik verilerden hasta prognozunu tahmin etmek için geliştirilen Cox-nnet adındaki bir sinir ağı yaklaşımıdır. The Cancer Genome Atlas'tan aldıkları 10 veri setinde Cox-nnet, Cox regresyon, CoxBoost ve RFS yöntemlerini karşılaştırmışlar. Her yöntemin ortalama dağılımını değıerlendirmek için uygulamaları 10 kez tekrarlamışlardır ve Cox-nnet ile başarılı sonuçlar elde etmişlerdir.

Gensheimer ve Narasimhan (2018) tarafından önerilen parametrik bir yaklaşım olan Nnet-survival, sinir ağıları için sađkalım süresinin kesikli sayısal olduğı bir sađkalım modelidir. Yaptıkları çalışmada benzetim ve gerçek veri setleri üzerinde Nnet-survival yönteminin performansını Cox-nnet ve DeepSurv ile karşılaştırmışlardır. Bu karşılaştırma sonucunda yöntemlerinin performansının benzer olduğı ancak hesaplama süresi göz önünde bulundurulduğunda Nnet-survival yönteminin Cox-nnet'e göre daha iyi, DeepSurv ile ise benzer olduğı sonucuna varmışlardır.

Giunchiglia ve ark. (2018) tarafından önerilen ise sansürlü verilerin varlığında her hasta için kişiselleştirilmiş bir risk skoru ve sađkalım olasılığı fonksiyonunu tahmin etmek için geliştirilen RNN-SURV adındaki tekrarlayan sinir ağı modelidir. Giunchiglia ve ark. yaptıkları çalışmada RNN-SURV ile aralarında Cox regresyon, RSF ve DeepSurv yöntemlerinin de bulunduğı çeşitli klasik sađkalım yöntemleri ile makine öğrenmesi yöntemlerini karşılaştırmışlardır. Uygulamaları ikisi büyük, üçü açık kaynaklı küçük veriseti olmak üzere toplam beş gerçek veriseti üzerinde yapmışlar ve RNN-SURV'ün diğerk yöntemlere göre bir miktar daha iyi sonuç verdiğini göstermişlerdir.

Literatürde sağkalım analizinde derin öğrenmeye ilişkin yukarıda bahsedilen dört yöntem olmakla birlikte, bunlardan en yaygın kullanılan yöntemin DeepSurv olduğu görülmüştür. Kim ve ark. (2019), oral skuamöz hücreli karsinom (SCC) hastalarında derin öğrenme tabanlı bir sağkalım tahmin algoritması olan DeepSurv kullanarak yaptıkları bir çalışmada, DeepSurv yönteminin performansını RSF ve Cox regresyon yöntemleri ile karşılaştırdılar. DeepSurv, üç yöntem arasında en iyi performansı göstermiştir (Kim ve ark., 2019).

Hung ve ark. (2019), robotik cerrahi ile radikal prostatektomi sonrası üriner kontinans iyileşmesini tahmin etmek için DeepSurv modeli kullanarak bir çalışma yapmışlardır. DeepSurv, RSF ve Cox regresyon sonuçlarını karşılaştırdıkları bu çalışmada, RSF ve DeepSurv yöntemleri Cox regresyona göre daha başarılı sonuçlar vermiştir (Hung ve ark., 2019).

Bice ve ark. (2020) tarafından ulusal kanser merkezi veritabanından alınan 1673 hastaya ait 27 değişkenin verisi ile beyin metastazına ilişkin derin öğrenmeye dayalı bir sağkalım analizi yapılmıştır. Burada, DeepSurv yaklaşımı kullanılmıştır. Ek olarak, bu yöntemin performansı RSF ve Cox regresyon yöntemleri ile karşılaştırılmıştır. DeepSurv yönteminde modelin başarısını en iyi yapan hiperparametreler belirlenmiş ve 7000 epoch ile çalıştırılmıştır. C-indeks sonucuna göre performanslar değerlendirildiğinde en iyi performansa sahip yöntemin DeepSurv olduğu görülmüştür. Ancak, RSF yöntemine göre sonuçlar bir miktar yüksek olmakla birlikte, benzerdir (Bice ve ark., 2020).

Huang ve ark. (2020) tarafından RNA-Seq verilerinden kanser sağkalım prognozuna ilişkin, derin öğrenmeye dayalı yöntemlerin performanslarının incelendiği bir çalışma yapılmıştır. Bu amaçla, The Cancer Genome Atlas'tan aldıkları 12 veri setinde, Cox-nnet, DeepSurv dahil olmak üzere çeşitli DÖ tabanlı modeller kullanılarak geniş bir analiz gerçekleştirilmiştir. DÖ tabanlı modelleri, RSF, DVM ve Cox regresyon yöntemleri ile kıyaslamışlardır. Sonuçta, DÖ tabanlı algoritmaların, geleneksel makine öğrenmesi tabanlı modellerden daha üstün performanslar sergilediğini ortaya koymuşlardır (Huang ve ark., 2020).

Byun ve ark. (2021), renal hücreli karsinomlarda prognoz tahmini için derin öğrenme yaklaşımından yararlanmışlardır. DeepSurv kullandıkları bu çalışmada yöntemin başarısını RSF ve Cox regresyon ile karşılaştırmışlardır. En yüksek tahmin başarısını gösteren yöntem DeepSurv olmuştur (Byun ve ark., 2021).

Trinh ve ark. (2021), yaptıkları çalışmada robotik cerrahi ile radikal prostatektomiye takiben üriner kontinans iyileşmesini öngören modelleme için DeepSurv ve Cox regresyon yöntemlerini kullanmışlardır. Bu çalışmada, DeepSurv yönteminin Cox regresyona göre daha iyi modelleme başarısına sahip olduğu görülmüştür (Trinh ve ark., 2021).

Morelli ve ark. (2021), önümüzdeki 10 yıl içerisinde ilk kez depresyon veya anksiyete tanısı alma riskini modellemeyi amaçladıkları bir çalışma yapmışlardır. Bunun için DeepSurv ve Cox regresyon yöntemlerini kullanmışlardır. Her iki yöntemin de performansı oldukça yüksek bulunmuştur (Morelli ve ark., 2021).

Bu tez çalışması kapsamında gerçekleştirilen benzetim çalışmasındaki kadar parametrelerin çeşitli seviyelerinin yer aldığı detaylı bir çalışmaya literatürde rastlanmamıştır. Başlık 2’de bağımsız değişken katsayılarının iki ayrı dağılımdan türetildiği ve başarılarının ayrı ayrı incelendiğinden söz edilmişti. Katsayıların, Kropko and Harden (2019) tarafından geliştirilen “coxed” paketindeki mevcut haliyle kullanıldığı durumda, tüm senaryo ve yöntemlerde sonuçların 0,50 civarında olduğu gözlemlenmiştir. Bu durum yöntemlerin performanslarını karşılaştırmaya ve senaryolardaki değişimin etkisini gözlemlemeye imkân vermemektedir. Çünkü Bölüm 1.4’te anlatıldığı gibi C-indeks değerinin 0,50 olması rastgele bir tahmini ifade etmektedir. Ancak, katsayıların Tekdüze dağılımdan türetildiği ikinci durumda, yöntemlerin C-indeks değerleri 0,60’ın üzerinde olup oldukça yüksek bulunmuş ve senaryolardaki değişimle birlikte performansları karşılaştırılabilir hale gelmiştir.

Bu çalışmada elde edilen bulgulara göre, Cox regresyon yönteminden elde edilen C-indeks değerleri RSF ve DeepSurv yöntemlerine göre daha yüksek bulunmuştur. Ancak, senaryo bazlı incelendiğinde Cox regresyonun hemen hemen

tüm senaryolarda benzer sonuçlar verdiği görülmüştür. Bu nedenle, Cox regresyon yönteminde senaryolardaki farklılığın sonuçlara olan etkisi gözlemlenememiştir. Ayrıca, Cox regresyon yönteminden elde edilen sonuçların tüm senaryolarda yüksek olmasının nedeni olarak verilerinin türetilmesinde kullanılan fonksiyonun Cox regresyon tabanlı olmasından kaynaklandığı düşünülmektedir.

Tüm yöntemlerde örneklem büyüklüğü arttıkça yöntemlerin başarısının arttığı gözlemlenmiştir. Küçük örneklem büyüklüğü DeepSurv ve Cox regresyonun performanslarını olumsuz anlamda çok etkilemese bile, RSF yönteminde bu etki net olarak görülmüştür. Örneklem büyüklüğü 100 olduğu durumda RSF'nin performansı diğer yöntemlere göre daha düşük bulunmuştur. Bunun yanı sıra, küçük örneklem büyüklüğüne sahip veri setlerinde sansür oranı 0,90 iken her senaryo için gerçekleştirilen 100 tekrarın yarısından fazlasında ilgilenilen olayın gerçekleşmediği görülmüş ve bu senaryolara ilişkin sonuçlar sunulmamıştır.

Sağkalım süresi ve bağımsız değişken sayısına ilişkin yöntemlerin performanslarına ilişkin bir çıkarım yapılamamakla birlikte genellikle bu parametrelerdeki değişimin DeepSurv ve RSF'nin performanslarına benzer, Cox regresyonun performansına ise bu iki yönteme zıt şekilde etkisi olduğu gözlemlenmiştir. Bağımsız değişken sayısına göre sonuçları özetlemek gerekirse, bağımsız değişken sayısı 10 iken çoğu senaryoda DeepSurv ve RSF sonuçları, 20 iken de Cox regresyon sonuçları daha yüksek bulunmuştur. Sağkalım süresine göre ise, DeepSurv ve RSF yöntemlerinin çoğunlukla sağkalım süresi 120 olduğu durumda daha yüksek C-indeks değerine sahip olduğu görülmüştür. Cox regresyon yönteminde ise diğer iki yöntemin aksine genellikle sağkalım süresi 60 iken daha başarılı sonuçlar elde edilmiştir.

Sansür oranındaki artış/azalışla birlikte Cox regresyonun çoğu senaryoda benzer sonuçlar verdiği, ancak DeepSurv ve RSF'nin sansür oranındaki değişime duyarlı olduğu görülmüştür. Sansür oranındaki artış, DeepSurv ve RSF yöntemlerinin performanslarını azaltmıştır. Tüm sansür oranlarında DeepSurv sonuçları RSF yöntemine göre daha yüksek bulunmuştur.

Gerçek veri setinde yapılan analizler sonucunda DeepSurv, RSF ve Cox regresyon yöntemlerinin validasyon setlerinin her üçünde de 0,54 civarında C-indeks

değeri elde edilmiştir. Eğitim setinde ise DeepSurv yönteminin performansı diğer iki yönteme göre daha yüksek bulunmuştur. Bu durum, veri setinde daha çok kişide ilgilenilen olay gerçekleşmiş olsaydı, DeepSurv yönteminin performansının başarılı olacağını göstermiştir.

Son olarak, işlemci türleri performans açısından karşılaştırıldığında GPU'nun CPU'ya göre beş kat hızlı olduğu görülmüştür. Maliyetleri ise birbirine yakındır.

Sağkalım analizinde kullanılan Cox regresyon analizi gibi geleneksel yöntemler değişkenlerin sağkalım eğrisine etkisini ortaya koymak için tasarlanmıştır. Eğer amaç sağkalım süresini tahmin etmek ise klasik yöntemler yerine makine öğrenmesi yöntemleri tercih edilmelidir. Özetle, klasik istatistiksel yöntemler temel olarak bağımsız değişkenlerin bağımlı değişken ile ilişkisini ortaya koymaya odaklanırken, makine öğrenmesi yöntemleri sınıflama tahmini ile ilgilenir.

Sağkalım analizlerinde, hazardların orantısallığı gibi varsayımlar klasik istatistiksel yöntemlerde bir kısıt olarak ortaya çıkarken, makine öğrenmesi yöntemlerinde böyle bir kısıt yoktur. Temel olarak, makine öğrenmesi yöntemlerinde bu varsayımların kontrol edilmesine gerek yoktur.

5. SONUÇ VE ÖNERİLER

Bu tez çalışmasında sağkalım analizine DÖ tabanlı bir yaklaşım olan DeepSurv uygulanmış ve sonuçlar makine öğrenmesi yöntemlerinden RSF ve klasik istatistiksel yöntemlerden Cox regresyon ile karşılaştırılmıştır.

Yapılan benzetim çalışması sonucunda, bağımsız değişken katsayılarının tekdüze dağılımdan türetildiği senaryolarda tüm yöntemlerin tahminleme başarısı yüksek bulunmuştur. Örneklem büyüklüğü arttıkça yöntemlerin performansı artmıştır. Küçük örneklem büyüklüğünde ise RSF'nin başarısının düşük olduğu görülmüştür. Senaryolardaki farklılaşmanın etkisi Cox regresyonda gözlemlenemezken, RSF ve DeepSurv yöntemlerinde parametrelerdeki değişimin etkisi açık bir şekilde ortaya konmuştur. Sansür oranının artması DeepSurv ve RSF yöntemlerinin tahminleme başarısını azaltmıştır. Tüm senaryolarda, DeepSurv yönteminden elde edilen sonuçlar RSF'ye göre daha yüksek bulunmuştur.

Bağımsız değişken katsayılarının paketteki mevcut haliyle kullanılıp normal dağılımdan türetildiği durumda ise üç yöntem de 0,50-0,60 civarında bir C-indeks değerine sahiptir ve tüm senaryolarda benzer sonuç verdikleri gözlemlenmiştir.

Gerçek veri setinde yapılan analizler sonucunda DeepSurv, RSF ve Cox regresyon yöntemlerinin her üçünün de uygulanabileceği sonucuna varılmıştır. DeepSurv yönteminde eğitim setinde performansın yüksek olması, ilgilenilen olayın gerçekleştiği kişi sayısının daha yüksek olduğu bir veri setinde çalışılıyorsa, bu yöntemle başarılı sonuçlar elde edilebileceğini göstermiştir.

Sonuç olarak, DÖ oldukça hızlı gelişen bir alandır. DeepSurv algoritmasında çok sayıda hiperparametre kullanılmasına rağmen, mimarisi anlaşılır ve basittir. Hiperparametre sayısının çok olması, model başarısını arttırmak için avantajlıdır. İlerideki çalışmalarda bu tez çalışmasındakine benzer senaryolarda, hiperparametrelerin de etkisi araştırılabilir.

ÖZET

Sağkalım Analizi için Derin Öğrenme Yaklaşımı

Tıp alanında sıklıkla kullanılan sağkalım analizi, ilgilenilen sonuç değişkeninin; olayın gerçekleştiği ana kadar geçen süre olduğu veriler için kullanılan istatistiksel yöntemlerin bütünüdür. Sağlık alanında, veri kümelerinin boyutunun artması ve derin öğrenmenin yaygınlaşması ile birlikte, sağkalım analizinde makine öğreniminin kullanımına olan ilgi artmaktadır. Son zamanlarda, derin öğrenme yöntemleri, çeşitli uygulama alanlarında sağkalım analizi problemlerini çözmeye konusunda büyük ilgi görmüştür.

Bu çalışmada sağkalım analizine derin öğrenme tabanlı bir yaklaşım olan DeepSurv uygulanmış ve sonuçlar makine öğrenmesi yöntemlerinden RSF ve klasik istatistiksel yöntemlerden Cox regresyon ile karşılaştırılmıştır. Bu amaçla, türetilmiş veri setleri kullanılarak bir benzetim çalışması gerçekleştirilmiştir. Bu kapsamda örneklem büyüklüğü, sağkalım süresi, bağımsız değişken sayısı, sansür oranının çeşitli düzeylerinin ele alındığı ve bağımsız değişkenlere ilişkin katsayıların tekdüze dağılımdan türetildiği toplam 60 senaryo oluşturulmuştur. Oluşturulan her senaryo ise 100 kez tekrarlanmıştır. Belirlenen senaryolara ek olarak, bağımsız değişken katsayılarına ilişkin ön belirlemenin yapılmadığı, R programlama dilindeki “coxed” paketindeki mevcut haliyle bırakıldığı 60 ayrı senaryo daha oluşturulmuştur. Ayrıca, bu çalışmada evre I akciğer adenokarsinomu veri seti kullanılarak da DeepSurv, RSF ve Cox regresyon yönteminin karşılaştırılması yapılmıştır. Yöntemlerin karşılaştırılmasında Harrell’in C-indeksi değeri kullanılmıştır. Verilerin türetilmesi, RSF ve Cox regresyon yöntemlerinin uygulanması R programlama dili (ver. 3.6.3) ile gerçekleştirilmiştir. Derin öğrenmeye ilişkin analizler ise Python programlama dilinde gerçekleştirilmiştir.

Yapılan benzetim çalışması sonucunda, bağımsız değişken katsayılarının tekdüze dağılımdan türetildiği senaryolarda tüm yöntemlerin tahminleme başarısı yüksek bulunmuştur. Örneklem büyüklüğü arttıkça yöntemlerin performansı artmıştır. Sansür oranının artması DeepSurv ve RSF yöntemlerinin tahminleme başarısını azaltmıştır. Bağımsız değişken katsayılarının paketteki mevcut haliyle kullanılıp normal dağılımdan türetildiği durumda ise üç yöntem sonucunda da 0,50-0,60 civarında bir C-indeksi değeri elde edilmiştir. Gerçek veri setinde yapılan analizler sonucunda ise DeepSurv, RSF ve Cox regresyon yöntemleri benzer sonuçlar vermiştir.

Anahtar Sözcükler: Cox regresyon, DeepSurv, Harrell’in C-indeksi, Random Survival Forest, Sağkalım Analizi

SUMMARY

A Deep Learning Approach for Survival Analysis

Time to event analysis is used often within medical research. Generally, survival analysis is a collection of statistical procedures for data analysis for which the outcome variable of interest is time until an event occurs. In the healthcare field, with the increasing size of datasets and the popularity of deep learning, interest in the use of machine learning in survival analysis has increased. Recently, deep learning methods have received great interest in solving survival analysis problems in various application areas.

In this study, DeepSurv, a deep learning-based approach, was applied to survival analysis and the results were compared with RSF from machine learning methods and Cox regression from classical statistical methods. For this purpose, a simulation study has been carried out using simulated data sets. In this context, a total of 60 scenarios were designed, in which various levels of sample size, survival time, number of independent variables and the probability of censored data were discussed and the coefficients of the independent variables were simulated from the uniform distribution. Each scenario was repeated 100 times. In addition to these scenarios, 60 different scenarios were created in which the coefficients of independent variables were used with the default in “coxed” package in R programming language. In addition, a comparison of DeepSurv, RSF and Cox regression methods was made using the stage I lung adenocarcinoma dataset in this study. Harrell's C-index value was used to compare the methods. The data simulation, RSF and Cox regression methods were performed with the R programming language (ver. 3.6.3). Analysis of deep learning was carried out in the Python programming language.

As a result of the simulation study, the performance of all methods was found to be high in scenarios where the coefficients of independent variables are simulated from the uniform distribution. As the sample size increased, the performance of the methods increased. The increase in the probability of censored data decreased the prediction success of DeepSurv and RSF methods. In the case where the independent variable coefficients are used as null, the C-index value of around 0.50-0.60 was obtained as a result of all three methods. According to the analyzes made on the real data set, the results of DeepSurv, RSF and Cox regression methods are similar.

Keywords: Cox regression, DeepSurv, Harrell's C-index, Random Survival Forest, Survival Analysis

KAYNAKLAR

- ABADI M, AGARWAL A, BARHAM P, BREVDO E, CHEN Z, CITRO C, CORRADO GS, DAVIS A, DEAN J, DEVIN M (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*.
- AGARWAL A, AKCHURIN E, BASOGLU C, CHEN G, CYPHERS S, DROPPA J, EVERSOLE A, GUENTER B, HILLEBRAND M, HOENS R (2014). An introduction to computational networks and the computational network toolkit. *Tech. Rep. MSR-TR-2014-112*.
- BICE N, KIRBY N, BAHR T, RASMUSSEN K, SAENZ D, WAGNER T, PAPANIKOLAOU N, FAKHREDDINE M (2020). Deep learning-based survival analysis for brain metastasis patients with the national cancer database. *Journal of Applied Clinical Medical Physics*, **21(9)**: 187-192.
- BINDER H, SCHUMACHER M (2008). Allowing for mandatory covariates in boosting estimation of sparse high-dimensional survival models. *BMC Bioinformatics*, **9(1)**: 14.
- BRADBURN MJ, CLARK TG, LOVE SB, ALTMAN DG (2003). Survival analysis part II: multivariate data analysis—an introduction to concepts and methods. *British Journal of Cancer*, **89(3)**: 431-436.
- BREIMAN L (2001). Random forests. *Machine Learning*, **45(1)**: 5-32.
- BYUN SS, HEO TS, CHOI JM, JEONG YS, KIM YS, LEE WK, KIM C (2021). Deep learning based prediction of prognosis in nonmetastatic clear cell renal cell carcinoma. *Scientific Reports*, **11(1)**: 1-8.
- CHING T, ZHU X, GARMIRE LX (2018). Cox-nnet: an artificial neural network method for prognosis prediction of high-throughput omics data. *PLoS Computational Biology*, **14(4)**, e1006076.
- CORTES C, VAPNIK V (1995). Support vector machine. *Machine Learning*, **20(3)**: 273-297.
- COX DR (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, **34(2)**: 187-202.
- CUTLER SJ, EDERER F (1958). Maximum utilization of the life table method in analyzing survival. *Journal of Chronic Diseases*, **8(6)**: 699-712.
- ELMAN JL (1990). Finding structure in time. *Cognitive Science*, **14(2)**: 179-211.

- FADNAVIS RA (2019). Application of Machine Learning For Survival Analysis- A Review. *IOSR Journal of Engineering*, **9(5)**: 56-60.
- FARAGGI D, SIMON R (1995). A neural network model for survival data. *Statistics in Medicine*, **14(1)**: 73-82.
- FOTSO S (2018). Deep neural networks for survival analysis based on a multi-task framework. *arXiv preprint arXiv:1801.05512*.
- FRANÇOIS C (2017). Deep learning with Python. In: Manning Publications Company.
- GENSHEIMER MF, NARASIMHAN B (2018). A Scalable Discrete-Time Survival Model for Neural Networks. *arXiv preprint arXiv:1805.00917*.
- GIUNCHIGLIA E, NEMCHENKO A, VAN DER SCHAAAR M (2018). Rnn-surv: A deep recurrent model for survival analysis. International Conference on Artificial Neural Networks. Springer. 23-32.
- GORDON L, OLSHEN RA (1985). Tree-structured survival analysis. *Cancer Treatment Reports*, **69(10)**: 1065-1069.
- HARRELL FE, CALIFF RM, PRYOR DB, LEE KL, ROSATI RA (1982). Evaluating the yield of medical tests. *JAMA*, **247(18)**: 2543-2546.
- HOCHREITER S, SCHMIDHUBER J (1997). Long short-term memory. *Neural Computation*, **9(8)**: 1735-1780.
- HOSMER DW, LEMESHOW S, MAY S (2002). Applied survival analysis: regression modeling of time to event data. Wiley New York, NY.
- HUANG Z, JOHNSON TS, HAN Z, HELM B, CAO S, ZHANG C, SALAMA P, RIZKALLA M, YU CY, CHENG J (2020). Deep learning-based cancer survival prognosis from RNA-seq data: approaches and evaluations. *BMC Medical Genomics*, **13**: 1-12.
- HUNG AJ, CHEN J, GHODOUSSIPOUR S, OH PJ, LIU Z, NGUYEN J, PURUSHOTHAM S, GILL IS, LIU Y (2019). A deep-learning model using automated performance metrics and clinical features to predict urinary continence recovery after robot-assisted radical prostatectomy. *BJU International*, **124(3)**: 487-495.
- ISHWARAN H, KOGALUR UB, BLACKSTONE EH, LAUER MS (2008). Random survival forests. *The Annals of Applied Statistics*, **2(3)**: 841-860.
- ISHWARAN H, KOGALUR UB, KOGALUR MUB (2021). Package ‘randomForestSRC’.

- JIA Y, SHELHAMER E, DONAHUE J, KARAYEV S, LONG J, GIRSHICK R, GUADARRAMA S, DARRELL T (2014). Caffe: Convolutional architecture for fast feature embedding. *Proceedings of the 22nd ACM International Conference on Multimedia*.
- KATZMAN JL, SHAHAM U, CLONINGER A, BATES J, JIANG T, KLUGER Y (2018). DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Medical Research Methodology*, **18(1)**: 24.
- KAYAALP K, SÜZEN A (2018). Derin Öğrenme ve Türkiye’deki Uygulamaları. *IKSAD International Publishing House*.
- KIM DW, LEE S, KWON S, NAM W, CHA IH, KIM HJ (2019). Deep learning-based survival prediction of oral cancer patients. *Scientific Reports*, **9(1)**: 1-10.
- KLEINBAUM DG, KLEIN M (2010). *Survival analysis*. Springer.
- KROPKO J, HARDEN JJ (2019). coxed: An R Package for Computing Duration-Based Quantities from the Cox Proportional Hazards Model. *The R Journal*, **11(2)**: 38-45.
- LEBLANC M, CROWLEY J (1993). Survival trees by goodness of split. *Journal of the American Statistical Association*, **88(422)**: 457-467.
- LECUN Y, BOTTOU L, BENGIO Y, HAFNER P (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, **86(11)**: 2278-2324.
- MA X, TAO Z, WANG Y, YU H, WANG Y (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies*, **54**: 187-197.
- Microsoft Azure. Sanal Makine (VM) Nedir?. Erişim Adresi: [<https://azure.microsoft.com/tr-tr/overview/what-is-a-virtual-machine/>]. Erişim Tarihi: 23/04/2021.
- MIOTTO R, WANG F, WANG S, JIANG X, DUDLEY JT (2018). Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*, **19(6)**: 1236-1246.
- MORELLI D, DOLEZALOVA N, PONZO S, COLOMBO M, PLANS D (2021). Development of digitally obtainable 10-year risk scores for depression and anxiety in the general population. *arXiv preprint arXiv:2104.10087*.
- MORGAN M (2019). BiocManager: access the bioconductor project package repository. R package version 1.30.10.
- NVIDIA. Deep Learning Frameworks. Erişim Adresi: [<https://developer.nvidia.com/deep-learning-frameworks>]. Erişim Tarihi: 10/02/2021.

- PARVAT A, CHAVAN J, KADAM S, DEV S, PATHAK V (2017). A survey of deep-learning frameworks. International Conference on Inventive Systems and Control (ICISC).
- PHI M (2018). Illustrated guide to lstm's and gru's: A step by step explanation. Erişim Adresi: [<https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21>]. Erişim Tarihi: 03/03/2021.
- RAIVO K (2019). Pheatmap: pretty heatmaps. R package version 1.0.12.
- ROSENBLATT F (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, **65(6)**: 386.
- SCHOBER P, VETTER TR (2018). Survival analysis and interpretation of time-to-event data: the tortoise and the hare. *Anesthesia and Analgesia*, **127(3)**: 792.
- SCHRÖDER MS, CULHANE AC, QUACKENBUSH J, HAIBE-KAINS B (2011). survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. *Bioinformatics*, **27(22)**: 3206-3208.
- SHATNAWI A, AL-BDOUR G, AL-QURRAN R, AL-AYYOUB M (2018). A comparative study of open source deep learning frameworks. 9th International Conference on Information and Communication Systems (ICICS).
- ŞEKER A, DIRI B, BALIK HH (2017). Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme. *Gazi Mühendislik Bilimleri Dergisi (GMBD)*, **3(3)**: 47-64.
- TEAM TTD, AL-RFOU R, ALAIN G, ALMAHAIRI A, ANGERMUELLER C, BAHDANAU D, BALLAS N, BASTIEN F, BAYER J, BELIKOV A (2016). Theano: A Python framework for fast computation of mathematical expressions. *arXiv preprint arXiv:1605.02688*.
- THERNEAU TM, GRAMBSCH PM (2000). The cox model. In Modeling survival data: extending the Cox model. Springer. 39-77.
- TRINH L, MINGO S, VANSTRUM EB, SANFORD DI, MA R, NGUYEN JH, LIU Y, HUNG AJ (2021). Survival Analysis Using Surgeon Skill Metrics and Patient Factors to Predict Urinary Continence Recovery After Robot-assisted Radical Prostatectomy. *European Urology Focus*.
- VACHHANI B, BHAT C, DAS B, KOPPARAPU SK (2017). Deep Autoencoder Based Speech Features for Improved Dysarthric Speech Recognition. Interspeech. 1854-1858.

- VOULODIMOS A, DOULAMIS N, DOULAMIS A, PROTOPAPADAKIS, E (2018). Deep learning for computer vision: A brief review. *Computational Intelligence and Neuroscience*.
- WANG P, LI Y, REDDY CK (2019). Machine learning for survival analysis: A survey. *ACM Computing Surveys (CSUR)*, **51(6)**: 1-36.
- ZHANG HH, LU W (2007). Adaptive Lasso for Cox's proportional hazards model. *Biometrika*, **94(3)**: 691-703.
- ZHAO L, FENG D (2019). DNNSurv: Deep Neural Networks for Survival Analysis Using Pseudo Values. *arXiv preprint arXiv:1908.02337*.
- ZHU X, YAO J, HUANG J (2016). Deep convolutional neural network for survival analysis with pathological images. IEEE International Conference on Bioinformatics and Biomedicine (BIBM).
- ZOU J, HUSS M, ABID A, MOHAMMADI P, TORKAMANI A, TELENTI A (2019). A primer on deep learning in genomics. *Nature Genetics*, **51(1)**: 12-18.

ÖZGEÇMİŞ

I- Bireysel Bilgiler

Adı : İrem
Soyadı : KAR
Doğum yeri ve tarihi :
Uyruğu : T.C.
Medeni durumu : Bekar
E-posta :
Adres :

II- Eğitimi

Doktora : Ankara Üniversitesi, Tıp Fakültesi, Biyoistatistik AD
(2017 -)
Yüksek Lisans : Ankara Üniversitesi, Tıp Fakültesi, Biyoistatistik AD
(2015 - 2017)
Lisans : Çukurova Üniversitesi, İktisadi İdari Bilimler Fakültesi,
Ekonometri Bölümü (2010 - 2013)
Lisans : Çukurova Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik
Bölümü (2008 - 2013)
Lise : Seyhan Rotary Anadolu Lisesi (2004 - 2008)
Yabancı Dil : İngilizce (YDS:70, YÖKDİL:83,75)

III- Ünvanları

Araştırma Görevlisi

IV- Mesleki Deneyimi

Araştırma Görevlisi, Ankara Üniversitesi, Tıp Fakültesi, Biyoistatistik AD
(05.02.2015 -)

V- Üye Olduğu Bilimsel Kuruluşlar

Biyoistatistik Derneği

International Biometrics Society - Eastern Mediterranean Region (IBS-EMR)

VI- Bilimsel İlgi Alanları

Uluslararası Hakemli Dergilerde Yayımlanan Makaleler

1. Yalçındağ Fatime Nilüfer, Temel Emine, Şekkeli Mehmet Zahid, Kar İrem (2021). Macular structural changes and factors affecting final visual acuity in patients with Behçet uveitis. Graefe's Archive for Clinical and Experimental Ophthalmology, 259, Doi: 10.1007/s00417 020-04958-4.
2. Göktaş Ayşe, Kar İrem, Varlı Murat (2020). The Impact of Activity-role Participation Oncognitive Status in Geriatric Individuals. Saudi Journal for Health Sciences.
3. Duman Berker, Çolak Burçin, Özdemir Sueda Nur, Özkasap Necati Serhat, Sabahi İrem Sanem, Şahinoğlu Mehmet Bahadır, Taşçı Didar, Tatar Mustafa Mert, Ünalp Canan, Kar İrem, Kumbasar Hakan (2019). Tıp Eğitiminin Ruhsal Hastalıklara Yönelik Tutumlara Etkisi. Türkiye Aile Hekimliği Dergisi, 23(4), 141-149.
4. Yiğit Öznur, Tokgoz Yilmaz Suna, Tahir Emel, Bajin Münir Demir, Kar İrem, Sennaroğlu Levent (2019). Wideband Tympanometry Results of Bone Cement Ossiculoplasty. American Journal of Otolaryngology, 40(4), 512-519., Doi: 10.1016/j.amjoto.2019.04.009.
5. Özgür Onur, Bahadır Burak, Doğan İhsan, Eroğlu Ümit, Aktan Serhat, Kar İrem, Çakıcı Gizem, Bayatlı Eyüp, Kahilogullari Gökmen, Ünlü Mustafa Ağahan (2019). The Effect of Phenyramidol on Neural Development in Early Chicken Embryo Model. Turkish Neurosurgery, 29(6), 851-855., Doi: 10.5137/1019-5149.JTN.26158-19.1.
6. Çakır Akay Gül Aslıhan, Yıldırım Dilsun, Erdoğan Fatma Gülrü, Kar İrem (2019). Evaluation of Long Term Prognosis in Patients Treated with Narrowband Ultraviolet B: A Retrospective Study for 8 Years. Türkiye Klinikleri Journal of Medical Sciences, 39(4), 353-361., Doi: 10.5336/medsci.2019-66821.

7. Çalışkan Hasan, Karabulut Gözde, Ergün Ahmet, Cihan Koray Hamza, Demirtaş Deniz, Tan Atakan Şafak, Saeed Ashraf Mohammad, Şahin Samet, Alsobai Osama Jamal Saleh, Kalay İncilay, Kar İrem (2019). A Correlational Study on Itching and Quality of Life in Patients with Dermatitis. *Acta Physiologica*, 227(S722), 124-124.
8. Kimyon Cömert Günsu, Türkmen Osman, Kar İrem, Yücel Özge, Kılıç Çiğdem, Boran Nurettin, Başaran Derman, Karalok Alper, Turan Taner (2019). Hormone Therapy Following Surgery in Low-grade Endometrial Stromal Sarcoma: Is It Related to a Decrease in Recurrence Rate?. *Journal of the Chinese Medical Association*, 82(5), 385-389., Doi: 10.1097/JCMA.0000000000000039.
9. Kimyon Cömert Günsu, Türkmen Osman, Kar İrem, Sınacı Selcan, Yılmaz Ergani Seval, Karalok Alper, Başaran Derman, Turan Ahmet Taner (2018). Independent Predictors of Survival in Endometrium Cancer: Platelet-to-lymphocyte Ratio and Platelet/neutrophil/monocyte-to-lymphocyte Ratio. *Journal of the Turkish-German Gynecological Association*, 19(2), 78-86., Doi: 10.4274/jtgga.2017.0112

Ulusal Hakemli Dergilerde Yayımlanan Makaleler

1. Göktaş Ayşe, Kar İrem, Uyanık Mine, Varlı Murat (2020). Geriatrik Bireylerin Aktivite-Rol Katılımlarının Yaşam Kalitesi Üzerine Etkisi. *Hacettepe Üniversitesi Sağlık Bilimleri Fakültesi Dergisi*.
2. Bakırarar Batuhan, Kar İrem, Gökmen Derya, Elhan Atilla Halil, Genç Volkan (2019). The Prediction of Breast Biopsy Outcomes Using Two Data Mining Algorithms Based on Parameter Variations. *Türkiye Klinikleri Journal of Biostatistics*, 11(2), 133-142., Doi: 10.5336/biostatic.2019-64754.
3. Üstün Sertaç, Kurt Selin, Erdinli Dilek, Ekebil Yağmur, Eriş Erdeniz, Kesik Sena, Arslan Abdullah, Kar İrem (2019). The Relationship Between Short-term Memory And Fasting Blood Glucose. *Journal of Ankara University Faculty of Medicine*, 72(1), 29-33., Doi: 10.4274/atfm.galenos.2019.29591.

Ulusal ve Uluslararası Kongrelerde Sunulan Bildiriler ve Posterler

1. Özgür Onur, Bahadır B, Doğan İhsan, Eroğlu Ümit, Aktan Serhat, Kar İrem, Çakıcı Gizem, Bayatlı Eyüp, Kahiloğulları Gökmen, Ünlü Mustafa Ağahan

- (2019). Erken Dönem Tavuk Embriyo Modelinde Feniramidol'un Nöral Gelişim Üzerine Etkisi. Türk Nöroşirurji Derneği 33. Bilimsel Kongresi, 29(2), 299 (Özet Bildiri/Poster).
2. Kar İrem, Bakırarar Batuhan, Elhan Atilla Halil, Köse Serdal Kenan (2018). Çok Yüzeyle Rasch Modeli ve Bir Uygulaması. 20. Ulusal ve 3. Uluslararası Biyoistatistik Kongresi (Tam Metin Bildiri/Sözlü Sunum).
 3. Bakırarar Batuhan, Kar İrem, Elhan Atilla Halil (2018). Rserve ile Kümeleme Algoritmalarına ait Web Tabanlı Yazılım Geliştirilmesi. 20. Ulusal ve 3. Uluslararası Biyoistatistik Kongresi (Tam Metin Bildiri/Sözlü Sunum).
 4. Bakırarar Batuhan, Kar İrem, Doğanay Erdoğan Beyza, Elhan Atilla Halil (2018). Eksik Veri Yerine Değer Atama Yöntemlerinin Performanslarının Veri Madenciliği Sınıflama Algoritması Kullanılarak Değerlendirilmesi. 20. Ulusal 3. Uluslararası Biyoistatistik Kongresi (Tam Metin Bildiri/Sözlü Sunum).
 5. Özcan Gökçen, Gündüz Ahmet Kaan, Kar İrem, Biçer Özlem (2018). Koroid Malign Melanomunda Cyberknife Radyocerrahi Sonuçları ve Radyasyon Retinopatisi için Risk Faktörleri. Türk Oftalmoloji Derneği 52. Ulusal Kongresi (Tam Metin Bildiri/Sözlü Sunum).
 6. Kimyon Cömert Günsu, Turkmen Osman, Kar İrem, Yücel Çelik Özge, Kılıç Çiğdem, Boran Nurettin, Basaran Derman, Karalok Alper, Turan Ahmet Taner (2018). Hormone Therapy Following Surgery in Low Grade Endometrial Stromal Sarcoma: Is it Related with Decrease in Recurrence Rate?. XII. Turkish-German Gynecologic Congress, 19(Suppl 1), 161-163. (Özet Bildiri/Poster).
 7. Göktaş Ayşe, Kar İrem, Varlı Murat (2018). Yaşlı Bireylerde Yaşam Kalitesini Etkileyen Etmenler. 10.uluslararası Protez Ortez Kongresi (Özet Bildiri/Poster).
 8. Göktaş Ayşe, Kar İrem, Varlı Murat (2018). The Impact of Activity-Role Participation on Cognitive Status in Geriatric Individuals. 14. Avrupa Birliği Geriatri Tıp Birliği Kongresi (Özet Bildiri/Poster).
 9. Göktaş Ayşe, Kar İrem, Varlı Murat (2018). Geriatrik Bireylerde Aktivite-Rol Katılımlarının Yaşam Kalitesi Ve Yaşam Memnuniyeti Üzerine Etkisi. 11.Akademik Geriatri Kongresi (Özet Bildiri/Poster).
 10. Bakırarar Batuhan, Kar İrem, Gökmen Derya, Doğanay Erdoğan Beyza, Elhan Atilla Halil (2017). Demonstration Of A Computerized Adaptive Testing

- Application Over A Simulated Data. 10th International Statistics Congress (Özet Bildiri/Sözlü Sunum).
11. Kar İrem, Bakırarar Batuhan, Doğanay Erdoğan Beyza, Gökmen Derya, Köse Serdal Kenan, Elhan Atilla Halil (2017). A Comparison Of Maximum Likelihood And Expected A Posteriori Estimation In Computerized Adaptive Testing. 10th International Statistics Congress (Özet Bildiri/Sözlü Sunum).
 12. Köse Serdal Kenan, Kar İrem, Bakırarar Batuhan (2017). Bulanık Çıkarsama Sistemleri ile Lojistik Regresyon Sınıflandırıcılarının Performanslarının Apandisit Verisinde Karşılaştırılması. 19. Ulusal Biyoistatistik Kongresi ve 2. Uluslararası Biyoistatistik Kongresi (Özet Bildiri/Poster)
 13. Bakırarar Batuhan, Kar İrem, Elhan Atilla Halil (2017). Rserve ile Veri Madenciliği Algoritmalarının Web Uygulaması. 19. Ulusal Biyoistatistik Kongresi ve 2. Uluslararası Biyoistatistik Kongresi (Özet Bildiri/Sözlü Sunum).
 14. Bakırarar Batuhan, Kar İrem, Elhan Atilla Halil (2017). LMS Yönteminin Uyum İyiliğinin Ölçülmesinde Kullanılan Yöntemlerin Karşılaştırılması. 19. Ulusal Biyoistatistik Kongresi ve 2. Uluslararası Biyoistatistik Kongresi (Özet Bildiri/Poster).
 15. Kar İrem, Bakırarar Batuhan, Köse Serdal Kenan (2017). Bulanık Çıkarsama Sistemleri ile Veri Madenciliği Yöntemlerinin Sınıflama Performansının Benzetim Çalışması ile Karşılaştırılması. 19. Ulusal Biyoistatistik Kongresi ve 2. Uluslararası Biyoistatistik Kongresi (Özet Bildiri/Sözlü Sunum).
 16. Bakırarar Batuhan, Kar İrem, Elhan Atilla Halil, Köse Serdal Kenan (2017). Big Data and Usage in Data Mining Applications on Medicine. 9th EMR-IBS and Italian Region Conference, 1-172. (Özet Bildiri/Poster).
 17. Kar İrem, Bakırarar Batuhan, Köse Serdal Kenan (2016). Veri Madenciliği Sınıflandırıcılarının Performanslarının Karşılaştırılması. 18.Ulusal Biyoistatistik Kongresi ve 1. Uluslararası Biyoistatistik Kongresi (Özet Bildiri/Poster).
 18. Bakırarar Batuhan, Kar İrem, Yavuz Yasemin (2016). Büyük Veriye Genel Bakış Sağlık Alanında Örnek Uygulama. 18.Ulusal Biyoistatistik Kongresi ve 1. Uluslararası Biyoistatistik Kongresi (Özet Bildiri/Sözlü Sunum).
 19. Kar İrem, Bakırarar Batuhan, Köse Serdal Kenan (2015). İki Sürekli Değişken Arasındaki Uyum Araştırmalarında En Uygun Kesim Noktasının Tespiti ROC

- Analizi mi Tek Değişkenli Lojistik Regresyon Analizi mi. 17. Ulusal Biyoistatistik Kongresi (Özet Bildiri/Poster).
20. Bakırarar Batuhan, Kar İrem, Yavuz Yasemin, Köse Serdal Kenan (2015). Veri Madenciliği ne Genel Bakış Ve Çok Katmanlı Algılayıcı Yönteminin WEKA Programında İncelenmesi Sağlık Alanında Bir Uygulama. 17. Ulusal Biyoistatistik Kongresi (Özet Bildiri/Sözlü Sunum).
21. Efe Ceren, Kar İrem, Ünal İlker, Sertdemir Yaşar, Burgut Hüseyin Refik (2015). Three Way ROC Analysis Theory and Application. 8th Conference of Eastern Mediterranean Region International Biometric Society (Özet Bildiri/Poster)

Yazılan Ulusal/Uluslararası Kitaplardaki Bölümler

1. Bilimsel Araştırma Yöntemleri, Bölüm adı:(Nitel Araştırmalar) (2018)., Kar İrem, Gökmen Derya, Ankara Üniversitesi Basımevi, Editör: Gökmen Derya, Basım sayısı:1, Sayfa Sayısı 130, ISBN:978-605-136-384-4, Türkçe(Bilimsel Kitap).

Seminerler

1. “Bulanık Mantık ve Bir Uygulaması”, Ankara Üniversitesi, Haziran 2017
2. “Covariate-Adjusted ROC”, Ankara Üniversitesi, Mayıs 2019

Ödüller

1. Poster Birincilik Ödülü, 17.Ulusal Biyoistatistik Kongresi, 2015
2. En İyi İkinci Sözlü Sunum, 20. Ulusal ve 3. Uluslararası Biyoistatistik Kongresi, 2018