



TÜBİTAK–2209-A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI

Başvuru formunun Arial 9 yazı tipinde, her bir konu başlığı altında verilen açıklamalar göz önünde bulundurularak hazırlanması ve ekler hariç toplam 20 sayfayı geçmemesi beklenir (Alt sınır bulunmamaktadır). Değerlendirme araştırma önerisinin özgün değeri, yöntemi, yönetimi ve yaygın etkisi başlıkları üzerinden yapılacaktır.

ARAŞTIRMA ÖNERİSİ FORMU

2. Dönem Başvurusu

2209/A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

A. GENEL BİLGİLER

Başvuru Sahibinin Adı Soyadı: Mehmet Cihan Sakman
Araştırma Önerisinin Başlığı: TCGA Klinik ve Moleküler Verilerine Dayanarak Akciğer Adenokarsinom Ve Akciğer Skuamöz Hücre Karsinomu Hastalarının Risk Tahmini İçin Mobil Uygulama Geliştirilmesi
Danışmanın Adı Soyadı: Tuğba Süzek
Araştırmanın Yürütüleceği Kurum/Kuruluş: Muğla Sıtkı Koçman Üniversitesi

ÖZET

Türkçe özetin araştırma önerisinin (a) özgün değeri, (b) yöntemi, (c) yönetimi ve (d) yaygın etkisi hakkında bilgileri kapsamı beklenir. Her bir özet 450 kelime veya bir sayfa ile sınırlandırılmalıdır. Bu bölümün en son yazılması önerilir.

Özet

Akciğer kanseri dünya çapında en çok rastlanan ve ölüme en çok sebebiyet veren kanser türüdür. Dünya Sağlık Örgütü (DSÖ) akciğer kanserinin 2012 yılında toplam 1.8 milyon yeni olgu ile tüm kanserlerin %12.9'unu meydana getirdiğini bildirmiştir[1]. Ülkemizde Sağlık Bakanlığının 2012 yılı verilerine göre erkeklerde en sık; kadınlarda ise beşinci sıklıkta görülen kanser türüdür ve her yıl yaklaşık 30,000 yeni vakanın teşhis edildiği tahmin edilmektedir[2]. Akciğer kanseri sırasıyla Küçük Hücreli Akciğer Kanseri (SCLC) ve Küçük Hücreli Olmayan Akciğer Kanseri (NSCLC) olarak iki sınıfa ayrılabilir. Akciğer Adenokarsinomu (LUAD) ve Akciğer Skuamöz Hücreli Karsinomu (LUSC), akciğer kanseri hastalarının %85'inde ortaya çıkan Küçük Hücreli Olmayan Akciğer Kanserinin (NSCLC) üç alt tipinden ikisidir.

Akciğer kanseri vakalarının büyük çoğunluğunu oluşturan LUAD ve LUSC kanseri hastalarının sağlık durumları erken evrede tespit edilir ve kanserin gidişatına göre bir tedavi uygulanırsa bu hastaların hayatta kalma olasılıkları artırılabilir. Mobil cihazların hayatımıza yerleşmesinden sonra sağlık alanında birçok mobil uygulama üretilmiş ve hastanın sağlığı güvence altına alınmaya çalışılmıştır. Kanser vakalarının tespiti hakkında birçok Makine Öğrenmesi projesi yapılmış ancak ülkemizde veya dünyada bu iki akciğer kanserinin hem klinik verilerini (sağkalım ve ilaç) hem de moleküler verilerini (gen ifadesi ve mutasyon) entegre bir şekilde eğitimiyle uygulamaya geçirilmiş herhangi klinik destek mobil uygulaması mevcut değildir. Bu alanda üretilen ve piyasaya sürülen ONCOassist uygulaması klinisyenler için verileri bir araya getirip kanser vakaları hakkında bir özet bilgi sunmaktadır. Ancak bu uygulamada herhangi bir Makine Öğrenmesi söz konusu değildir.

Şu an klinikte hastanın risk grubu hakkında sadece kanser evreleme ("staging") ile klinisyen ve hasta bilgilendirilmektedir. Önerdiğimiz proje sonunda ortaya kullanıcı dostu bir mobil uygulama çıkarılması amaçlanmaktadır. Bu mobil uygulama sayesinde klinisyenlerin hastalarının, hastalıklarının risk miktarı hastaya özgü klinik parametrelere (tümör lokasyonu ,sigara kullanım miktarı, yaş, cinsiyet, kullanılan ilaçlar) göre hastalığın hangi risk grubunda olduğu hakkında bilgi sahibi olmaları amaçlanmaktadır. Kanser tedavisi gören hastaların da aynı uygulama üzerinden, kendisine özgü parametreler göz önüne alındığında hastalıklarının hangi risk grubunda olduğuna dair bilgi edinmeleri mümkün hale gelecek ve kullanmakta oldukları ilaçlar hakkında gerekli bilgileri edinebileceklerdir.

Araştırmanın mobil uygulamaya dönüştürülmesine kadarki evrede izlenecek başlıca adımlar: Bilinen en geniş halka açık kanser hasta verilerini içeren Kanser Genom Atlas ağı (TCGA) 33 kanser tipinden 11000 hastanın verilerini içermektedir (<https://www.cancer.gov/tcga>). TCGA verilerinden elde edilen LUAD-LUSC akciğer kanseri veri kümesinin Python yazılım dili kullanılarak gerekli kütüphaneler aracılığıyla ön işleminden geçirilmesi, Python "scikit-learn" kütüphanesi kullanılarak Ensemble Learning(Kolektif Öğrenme) makine öğrenmesi algoritması ile hastanın, risk durumu ile ilgili en doğru tahminlerin yapılması, kullanıcılar tarafından kullanımı kolay olabilecek bir mobil uygulama arayüz tasarımının yapılması ve yazılımın mobil uygulamaya uyarlanması, Mobil uygulamanın Muğla Sıtkı Koçman Üniversitesi Tıp Fakültesi onkoloji uzmanı klinisyenler tarafından test edilmesi ve geri dönüş alınarak bu yönde iyileştirmeler yapılması, mobil uygulamanın Google Play Store'da basılacak bir apk haline getirilip ücretsiz kullanıma sunulmasıdır.

2209/A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

Anahtar Kelimeler: TCGA, klinik destek yazılımları, makine öğrenmesi

1. ÖZGÜN DEĞER

1.1. Konunun Önemi, Araştırma Önerisinin Özgün Değeri ve Araştırma Sorusu/Hipotezi

Araştırma önerisinde ele alınan konunun kapsamı ve sınırları ile önemi literatürün eleştirel bir değerlendirmesinin yanı sıra nitel veya nicel verilerle açıklanır.

Özgün değer yazılırken araştırma önerisinin bilimsel değeri, farklılığı ve yeniliği, hangi eksikliği nasıl gidereceği veya hangi soruna nasıl bir çözüm geliştireceği ve/veya ilgili bilim veya teknoloji alan(lar)ına kavramsal, kuramsal ve/veya metodolojik olarak ne gibi özgün katkılarda bulunacağı literatüre atıf yapılarak açıklanır.

Önerilen çalışmanın araştırma sorusu ve varsa hipotezi veya ele aldığı problem(ler)i açık bir şekilde ortaya konulur.

LUAD (adenokarsinoma) ya da LUSC (skuamoz hücreli) akciğer kanseri hastası klinikte tedavisi birden çok klinisyen tarafından yönetilen bir süreçtir. Genelde hastanın şikayetiyle ulaştığı Göğüs Hastalıkları uzmanı hastayı Radyoloji Uzmanına sevk ve sonrasında cerrahi müdahale için Genel Cerrahi uzmanına sevk ile teşhis süreci başlar. Genel Cerrahi uzmanının Patoloji uzmanı tarafından konulan teşhis ile beraber hastayı Tıbbi Onkoloji uzmanına yönlendirmesiyle de tedavi süreci başlar. Bu birden çok klinisyenden çoğu zaman farklı kliniklerden randevu alınması gereken süreçte hasta kendi risk düzeyi hakkında bir belirsizlik içinde kalmaktadır. Birden çok klinisyenin ve yardımcı personelin içinde olduğu süreçte insan hatalarının ve gecikmelerinin en aza indirilmesi için hem klinisyenlerin hem de hastanın hastaya özgü risk durumunu öğrenebileceği klinik destek yazılımları hem tedavi kalitesinin artırılması hem de randevu süreçlerinin hızlandırılmasında yardımcı olacaktır.

Yazılımımız bir klinisyenin LUAD ve LUSC akciğer kanseri hastaları için cinsiyet/stage/yaş/sigara kullanımı/ilac gibi klinik verileri basit bir arayüzle girdiğinde risk sınıflandırması yapan bir klinik destek mobil uygulaması olarak piyasadaki ilk örneği olacaktır. Arka planda eğiteceğimiz model için sağkalım ve ilaç gibi klinik verilere ek olarak mutasyon ve gen ifadesi verileri kullanılacaktır. Geliştireceğimiz bu yazılım sayesinde klinisyen (Göğüs Hastalıkları uzmanı ya da Tıbbi Onkoloji uzmanı) teşhis alan hastanın düşük risk ya da yüksek risk grubu hakkında bir tahmin sonucunu göz önüne alarak farklı tedavi yaklaşımı planlayabilecektir. Aynı zamanda kanserin herhangi bir seviyesinde bulunan hastaların doktorlar tarafından hangi ilaçla tedavi edildiğine ve hastaların bu süreçte nasıl tepki verdiklerine dair hastayı bilgilendirici bir arayüz tasarımı yapılacaktır.

Biyoinformatik Anabilim Dalından akademik danışmanlarımız Tuğba Süzek, PhD (Bilgisayar Mühendisi ve Biyoinformatik uzmanı) ve Talip Zengin, MSc (Moleküler Biyolog ve Biyoinformatik uzmanı) tarafından yayınlanmış ve basım aşamasında olan iki makale yayınında akciğer kanserinin iki çeşidi olan LUAD ve LUSC verilerinden risk gruplarını hesaplayan bir öğrenim modeli geliştirilmiştir[3,4]. Çalışmamızın iki ana amacı 1) bu risk modellerine dayanarak hasta kohortlarının klinik parametrelerini derin öğrenme temelli tahmin modelleri geliştirmek 2) klinisyen ve hastalara TCGA verilerini ve risk analizlerimizi kullanıma sunmak için kullanıcı dostu bir mobil uygulama geliştirmektir.

Bu alanda yapılan benzer çalışmaları sıralayacak olursak:

-Mikroskopik patoloji görüntü özellikleriyle LUAD ve LUSC -adenokarsinoma ya da skuamoz hücreli- akciğer kanseri prognozunu tahmin eden yazılım[5]

-Küçük Hücreli Olmayan Akciğer Kanserinin (NSCLC) patolojik evresinin teşhisi ve tahmini için kullanılabilen görüntüleme biyobelirteçlerini, CT görüntü özellik analizine dayalı çoklu makine öğrenme algoritmaları kullanılarak keşfeden yazılım[6]

-TCGA'dan alınan RNA-sıralama verilerini kullanarak Ensemble Learning yardımı ile kanser hastalarının hayatta kalma sürelerini tahmin etmeye çalışan yazılım[7]

-Projemizle en çok benzerlik gösteren bir uygulama olan ONCOassist, tüm kanser çeşitleri için Mobil ve Web arayüzlerinde klinisyenler tarafından hastanın istenilen bilgilerinin girilmesi durumunda, klinisyene hastanın gelecek 5 yıl içerisindeki yaşama şansına dair istatistiksel bilgi sunan bir yazılımdır.[8]

Sıraladığımız benzer yazılımlara rağmen yaptığımız araştırmalar sonucunda LUAD ve LUSC kanseri

2209/A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI ARAŞTIRMA ÖNERİSİ FORMU

hastalarının klinik verilerinin girildiğinde, bu hastaların yaşam süreleriyle ilgili bir makine öğrenmesi algoritması kullanan ve klinisyenleri bilgilendiren herhangi bir mobil uygulama başvuru tarihi itibarıyla bilginiz dahilinde mevcut değildir.

Akciğer kanserinin ülkemizde ve dünya çapında sebep olduğu ölümler göz önüne alındığında, temel klinik veriler yardımıyla hastalığın seyir riskini tahmin ederek insan hatalarını azaltma amaçlı bir uygulamanın önemi, bu yönde klinisyenler ve hastalar tarafından kullanıma hazır herhangi bir mobil uygulamanın olmaması bu projeyi özgün kılmaktadır. Böyle bir mobil uygulamanın hayata geçirilmesi kanser hastaları ve klinisyenlerin hasta açısından belirsizlikler içeren zor bir sürecin yönetiminde insan hatalarını ve gecikmeleri azaltarak tedavi sürecini kolaylaştıracak bir çözüm haline gelecektir.

1.2. Amaç ve Hedefler

Araştırma önerisinin amacı ve hedefleri açık, ölçülebilir, gerçekçi ve araştırma süresince ulaşılabilir nitelikte olacak şekilde yazılır.

Kanser hastaları için tanı konulma aşamasından tedavi sürecinin sonuna kadar tüm süreçte alınan kararlar, insan hataları, gecikmeler tedavinin başarısında çok kritik önem taşımaktadır. Bu sürecin hasta ve doktoru tarafından iyi yönetilmesi hasta için hayati önem taşımaktadır. Her yıl binlerce insana akciğer kanseri tanısı konulmakta ve kanser vakalarının büyük çoğunluğunu oluşturan bu kanser binlerce insanın hayatını kaybetmesine neden olmaktadır. Tedavi süreci bir hayli zordur ve kanser hücrelerinin sürekli mutasyon geçirmeleri bu süreci daha da karmaşık hale getirmektedir. Yazılımımızın amacı hastanın cinsiyet/yaş/sigara kullanımı/kanser evresi/kullandığı ilaç/tümör lokasyonu gibi klinik verilerin yanı sıra mutasyon ve gen ifadesi verilerini de göz önüne alarak makine öğrenmesi yardımı ile hastanın düşük veya yüksek risk grubunda olup olmadığına dair bir tahminde bulunmak ve hastanın tedavi sürecinde klinisyene yardımcı bir mobil uygulama geliştirmektir. Aynı zamanda bir diğer amacımız ise hastanın kullandığı ilaçlar hakkında bilgilendirilmesi, bu ilaçları kullanan diğer TCGA kohortları hakkında bilgi edinmesi ve tedavi sürecinde daha güvenli bir şekilde doktoruyla işbirliğinin artırılmasıdır.

Bu amaçla kullanılacak bir mobil uygulama elbette hastanın durumu hakkında kesin ve doğru bir yargıda bulunmasa da hem klinisyene hem de hastaya geniş kohortlar üzerinden hesaplanan tahminler üzerinden yol gösterici bir rehber niteliği taşıyacaktır.

2. YÖNTEM

Araştırma önerisinde uygulanacak yöntem ve araştırma teknikleri (veri toplama araçları ve analiz yöntemleri dahil) ilgili literatüre atıf yapılarak açıklanır. Yöntem ve tekniklerin çalışmada öngörülen amaç ve hedeflere ulaşmaya elverişli olduğu ortaya konulur.

Yöntem bölümünün araştırmanın tasarımı, bağımlı ve bağımsız değişkenleri ve istatistiksel yöntemleri kapsamı gerekir. Araştırma önerisinde herhangi bir ön çalışma veya fizibilite yapıldıysa bunların sunulması beklenir. Araştırma önerisinde sunulan yöntemlerin iş paketleri ile ilişkilendirilmesi gerekir.

İP 1: Veri Ön İşleme:

Proje kapsamında makine öğrenme algoritması kurulmadan önce elimizdeki verilerden en iyi şekilde yararlanmamız gerekir. TCGA klinik verilerinde kirli veriler mevcut olup literatürde yapılmış önceki çalışmalara benzer şekilde veri setini düzenlemek, eksik verilerin bulunduğu sütunları silmek ve öğrenme algoritmasından en iyi şekilde yararlanılmasını sağlayacak verileri tespit etmek için veri seti ön işlemeden geçirilecektir. Veri ön işleme sırasında içerisinde herhangi bir veri bulunmayan sütunlar veri setinden atılacaktır. Birbiri ile örtüşen (aynı verileri içeren) sütunlar tespit edilip veri setinden çıkartılacaktır ve kullanılması hesaplanan makine öğrenme algoritmaları için gereken veri düzeni hazırlanacaktır.

Sağkalım verilerinin pozitif (uzun yaşayan) ve negatif (az yaşayan) kümelerinin dengeli olması ve aşırı örnekleme (oversampling) ya da az örnekleme (undersampling) önüne geçmek için gerekli örnekleme algoritmaları çalıştırılarak veri seti eğitim ve test veri setleri olarak ikiye ayrılacaktır.

İP 2: Öğrenme Algoritmalarının Uygulanması:

Bu iş paketinde, veri ön işleme sürecinden geçmiş verilere Python “sklearn” kütüphanesi kullanılarak Ensemble Learning(Kolektif Öğrenme) uygulanacak ve LUAD ve LUSC akciğer kanseri hastalarının: kullandıkları ilaç, sigara içme durumu, mutasyona uğrayan genleri, hastalığa yakalandıklarından beri geçen zaman ve cinsiyet gibi özellikleri göz önünde bulundurularak ne kadar süre daha hayatta kalabileceklerini tahmin edecek bir makine öğrenme algoritması kurulacaktır. Ensemble Learning için üç temel öğrenme algoritması kullanılacaktır ve bu algoritmalarından elde edilecek oylar sonucunda daha doğru bir tahmin elde edilmesi planlanmaktadır. Bu üç temel öğrenme algoritması: Yapay Sinir Ağları(Artificial Neural Network), Rastgele Orman(Random Forest) ve Karar Ağaçları (Decision Tree) olarak belirlenmiştir.

İP 3: Mobil Aplikasyon Arayüzü Tasarımı:

Makine öğrenme modeli oluşturulduktan sonra amacına hizmet etmesi adına Chaquopy kullanılarak öğrenilen model Android cihazlar için geliştirilecek bir mobil uygulama ile birleştirilecektir. Klinisyenlerin günlük klinikte kullanımının kolay olması amacıyla, sade ve anlaşılması kolay bir arayüz tasarlanacaktır.

İP 4: Öğrenme Mekanizmasının Test Edilmesi ve Geri Bildirim Alınması:

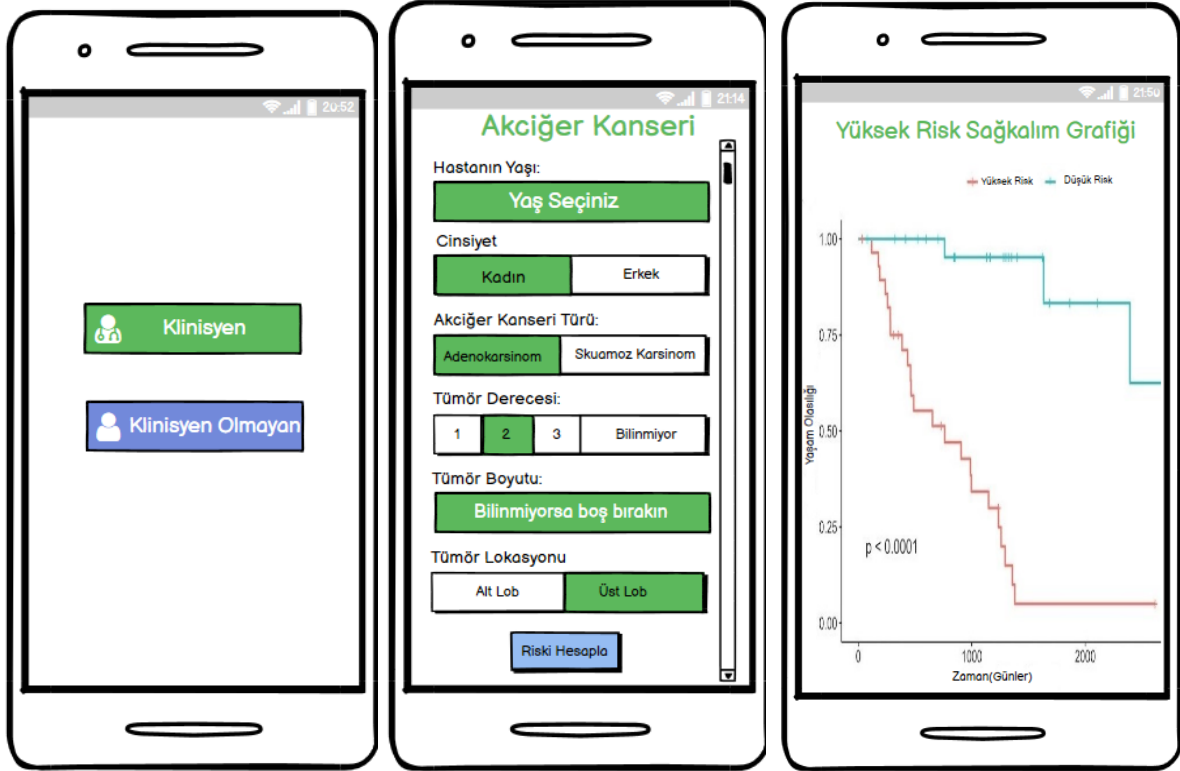
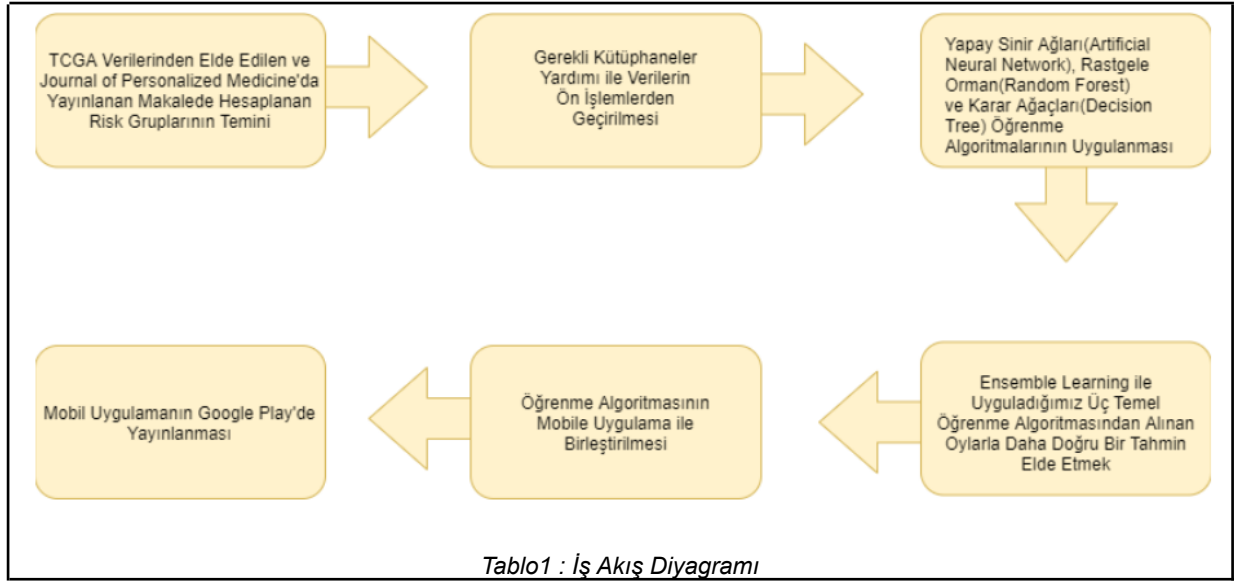
Bu aşamada kurulan algoritma ve tasarlanan mobil arayüz hazır hale getirilmiş olacaktır. Mobil uygulamamız Muğla Sıtkı Koçman Tıp fakültesindeki Göğüs Hastalıkları uzmanı , Tıbbi Onkoloji uzmanı ve Balıkesir Sevgi hastanesinden Genel Cerrah Dr.Erol Önal ile test ederek kullanıcı arayüzünün kullanıcı dostu hale getirilebilmesi için geri bildirim alınacaktır. Son kullanıcımız olan klinisyenler tarafından uygulamamız başarılı bulunduğu takdirde proje sonrasındaki bir sonraki versiyon için 20 gerçek hasta verisi ile sistemin test edilmesi için Muğla Sıtkı Koçman Üniversitesi insan araştırmaları etik kuruluna etik kurul başvurusu yapılacaktır. Göğüs Hastalıkları uzmanı, Tıbbi Onkoloji uzmanı ve Genel Cerrahi Uzmanı Op.Dr. Erol Önal’dan alınan geri bildirimler yardımı ile yazılıma son hali verilecek ve kullanıma hazır hale getirilecektir.

Kişisel verileri koruma kanunu (KVKK) kapsamında test verileri tamamen istemci tarafında klinisyenin kendi mobil telefonunda işlenecek ve hiç bir şekilde test verisi SuzekLab sunucularına aktarılmayacaktır.

İP 5: Uygulamanın Yaygınlaştırılması:

Geliştirdiğimiz prototip uygulamayı Google Play’de basılabilecek şekilde bir apk haline getireceğiz. Uygulamamızın tanıtımı için European Medical and Biological Engineering 2021 gibi saygın uluslararası bir konferansa tam metin bildiri hazırlayacağız.

2209/A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU



Resim1: Klinisyen Girişi Mobil Arayüzü



Resim2: Klinisyen Olmayan Girişi Mobil Arayüzü

2209/A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

3 PROJE YÖNETİMİ

3.1 İş- Zaman Çizelgesi

Araştırma önerisinde yer alacak başlıca iş paketleri ve hedefleri, her bir iş paketinin hangi sürede gerçekleştirileceği, başarı ölçütü ve araştırmının başarısına katkısı "İş-Zaman Çizelgesi" doldurularak verilir. Literatür taraması, gelişme ve sonuç raporu hazırlama aşamaları, araştırma sonuçlarının paylaşımı, makale yazımı ve malzeme alımı ayrı birer iş paketi olarak gösterilmemelidir.

Başarı ölçütü olarak her bir iş paketinin hangi kriterleri sağladığında başarılı sayılacağı açıklanır. Başarı ölçütü, ölçülebilir ve izlenebilir nitelikte olacak şekilde nicel veya nitel ölçütlerle (ifade, sayı, yüzde, vb.) belirtilir.

İŞ-ZAMAN ÇİZELGESİ (*)

İP No	İş Paketlerinin Adı ve Hedefleri	Kim(ler) Tarafından Gerçekleştirileceği	Zaman Aralığı (.-.. Ay)	Başarı Ölçütü ve Projenin Başarısına Katkısı
1	Veri Ön İşleme	Yürütücü (Mehmet Cihan Sakman)	0-3 ay	Başarı ölçütü:Hastaların en az %20sinde (100 hasta) öğrenmesi modeli için hazırlanan
2	Öğrenme Algoritmalarının Uygulanması	Yürütücü (Mehmet Cihan Sakman)	3-8 ay	Başarı ölçütü:En az 3 makine öğrenmesi algoritması %30
3	Mobil Aplikasyon Arayüzü Tasarımı	Yürütücü (Mehmet Cihan Sakman)	8-9 ay	Başarı ölçütü:En az bir platformda (Android ya da Apple) .Katkısı: %5
4	Öğrenme Mekanizmasının Test Edilmesi ve Geri Bildirim Alınması	Yürütücü (Mehmet Cihan Sakman)	9-11 ay	Başarı ölçütü: Eğittiğimiz modelin test verisinde en Katkısı: %20
5	Uygulamanın Yaygınlaştırılması	Yürütücü (Mehmet Cihan Sakman)	11-12 ay	Başarı ölçütü:Sonuçların bir ingilizce bildiri hazırlanması sözlü bildiri olarak sunulması

(*) Çizelgedeki satırlar ve sütunlar gerektiği kadar genişletilebilir ve çoğaltılabilir.

2209/A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

3.2 Risk Yönetimi

Araştırmanın başarısını olumsuz yönde etkileyebilecek riskler ve bu risklerle karşılaşıldığında araştırmanın başarıyla yürütülmesini sağlamak için alınacak tedbirler (B Planı) ilgili iş paketleri belirtilerek ana hatlarıyla aşağıdaki Risk Yönetimi Tablosu'nda ifade edilir. B planlarının uygulanması araştırmanın temel hedeflerinden sapmaya yol açmamalıdır.

RİSK YÖNETİMİ TABLOSU*

İP No	En Önemli Riskler	Risk Yönetimi (B Planı)
1	SuzekLab'da mevcut yüksek performanslı sunucular modelimizi eğitmeye yetersiz olması	Model eğitimi için kullanılacak veriler halka açık olan veriler olacağı için modeli eğitmek için Google Colab da kullanılabilir. Test için gerçek hasta verisi kullanıma ihtimalinde yurtdışı bir bulut ortamı olan Colab kullanılamaz. SuzekLab sunucuları test için yetersiz olursa Tübitak Truba sisteminden hesap açılacaktır. Tübitak Tübitak projesi olan öğrencilere Truba hesabı açmaya izin vermektedir. Truba hesabındaki sunucular kullanılacaktır.
2	Test seti için gerekli etik kurulun alınamaması	Etik kurul vaktinde alınamazsa test veri seti olarak akademik danışmanın bağlantısının olduğu NIH'ten 2020 yazında temin iznini aldığı dbGap verisi (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000728.v1.p1) ya da sağkalım verisi içeren bir GEO veriseti kullanılacaktır. DbGap verisi hem klinik hem de RNAseq verisi içermekte ama klinik veriler içinde sağkalım verisi içermeme ve klinisyen memnuniyetini test etme imkanı olmaması sebebiyle test için öncelikli tercih edilmemiştir.

(*) Tablodaki satırlar gerektiği kadar genişletilebilir ve çoğaltılabilir.

3.3. Araştırma Olanakları

Bu bölümde projenin yürütüleceği kurum ve kuruluşlarda var olan ve projede kullanılacak olan altyapı/ekipman (laboratuvar, araç, makine-teçhizat, vb.) olanakları belirtilir.

ARAŞTIRMA OLANAKLARI TABLOSU (*)

Kuruluşta Bulunan Altyapı/Ekipman Türü, Modeli (Laboratuvar, Araç, Makine-Teçhizat, vb.)	Projede Kullanım Amacı
Süzek Lab	Bu projede Muğla Üniversitesi Araştırma Laboratuvarları Merkezinde bulunan Süzek laboratuvarı (SuzekLab) olanaklarından faydalanılacaktır. Laboratuvar da proje yöneticisi lisans öğrencileri için ayrı bir ofis alanı ve 8 yüksek performanslı (Dell Poweredge R530 128GB RAM, 8TB sabit disk, Ubuntu 20.04) sunucular mevcuttur. Yazılım bu yüksek performanslı Linux sunucularında yazılacaktır.
Kütüphane	Basılı ve internet üzerinden sunulan literatüre ulaşmak için üniversitedeki kütüphane imkanları da kullanılacaktır.

(*) Tablodaki satırlar gerektiği kadar genişletilebilir ve çoğaltılabilir.

2209/A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

4. YAYGIN ETKİ

Önerilen çalışma başarıyla gerçekleştirildiği takdirde araştırmadan elde edilmesi öngörülen ve beklenen yaygın etkilerin neler olabileceği, diğer bir ifadeyle yapılan araştırmadan ne gibi çıktı, sonuç ve etkilerin elde edileceği aşağıdaki tabloda verilir.

ARAŞTIRMA ÖNERİSİNDEN BEKLENEN YAYGIN ETKİ TABLOSU

Yaygın Etki Türleri	Önerilen Araştırmadan Beklenen Çıktı, Sonuç ve Etkiler
Bilimsel/Akademik (Makale, Bildiri, Kitap Bölümü, Kitap)	Önerilen proje kapsamında yürütülmesi planlanan lisans bitirme projesi çalışmasının sonuçlarının European Medical and Biological Engineering Conference (EMBEC) 2021 gibi saygın bir uluslararası medikal enformatik konferansında basılması planlanmaktadır.
Ekonomik/Ticari/Sosyal (Ürün, Prototip, Patent, Faydalı Model, Üretim İzni, Çeşit Tescili, Spin-off/Start- up Şirket, Görsel/İşitsel Arşiv, Envanter/Veri Tabanı/Belgeleme Üretimi, Telif Konu Olan Eser, Medyada Yer Alma, Fuar, Proje Pazarı, Çalıştay, Eğitim vb. Bilimsel Etkinlik, Proje Sonuçlarını Kullanacak Kurum/Kuruluş, vb. diğer yaygın etkiler)	Mobil uygulama ile erken dönem hastalıklarda uzman gözünden kaçabilecek hatalara , hasta veri güvenliği riske atılmadan erken müdahale edilebilecektir. Hastalığın geç teşhisinin ülkemiz sağlık sistemine yükü azaltılarak hem hastanın kendi hem de devlet bütçesine katkıda bulunma potansiyelini içermektedir
Araştırmacı Yetiştirilmesi ve Yeni Proje(ler) Oluşturma (Yüksek Lisans/Doktora Tezi, Ulusal/Uluslararası Yeni Proje)	Proje kapsamında bir lisans öğrencisinin tezi için gerekli araştırmaların tamamlanması ve desteklenmesi planlanmıştır. Proje boyunca öğrenciler biyoinformatik, görüntü işleme, internet sitesi oluşturma alanlarında bilgiler edinecek, dolayısı ile Medikal Enformatik üzerine çok disiplinli bir çalışma yapacaktır. Öğrencimiz mezun olduğunda akademik çalışmalarına devam edebilmek veya sanayideki Araştırma-Geliştirme çalışmalarına katkıda bulunmak için hazır olacaktır.

5. BELİRTMEK İSTEDİĞİNİZ DİĞER KONULAR

Sadece araştırma önerisinin değerlendirilmesine katkı sağlayabilecek bilgi/veri (grafik, tablo, vb.) eklenebilir.

Güz 2020 döneminde Data Mining dersi kapsamında bu projeye giriş olarak LUAD-LUSC klinik verilerini kapsayan bir proje geliştirdim. Bu proje kapsamında veriyi ön işlemeden geçirerek temizledim ve öğrenme algoritmalarını kullanabilecek hale getirdim. Bunun yanı sıra veri seti ile ilgili bazı önemli detayları Python'da bulunan *matplotlib* kütüphanesi yardımı ile görselleştirdim. Aynı zamanda Güz 2020 döneminde almakta bulunduğum Mobile Application Development dersi kapsamında da Android için bir sosyal medya uygulaması projesi geliştirmekteyim. Bu tecrübe sayesinde de Makine Öğrenmesi yazılımımızı kolaylıkla, kullanıcı dostu arayüze sahip bir mobil uygulamaya çevireceğime inanıyorum.

2021 Bahar döneminde Muğla Üniversitesi Bilgisayar Mühendisliği bölümünde lisans seviyesinde verilmekte olan Applied Machine Learning dersini de alarak proje için gerekli alt yapı hakkında bilgi sahibi olacağım.

6. EKLER

EK-1: KAYNAKLAR

1-Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015 Mar 1;136(5):E359-86.

2209/A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

2-Akciğer Kanseri Yol Haritası – Türk Akciğer Kanseri Derneği.

3-Zengin T, Önal-Süzek T. Analysis of genomic and transcriptomic variations as prognostic signature for lung adenocarcinoma. *BMC Bioinformatics*. 2020 Sep 30;21(Suppl 14):368. doi: 10.1186/s12859-020-03691-3. PMID: 32998690; PMCID: PMC7526001.

4-Zengin T, Önal-Süzek T. Comprehensive profiling of genomic and transcriptomic differences between risk groups of lung adenocarcinoma and lung squamous cell carcinoma, under review, Jan 2021, *Journal of Personalized Medicine*.

5-Yu, KH., Zhang, C., Berry, G. *et al.* Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nat Commun* 7, 12474 (2016).

6-Yu, L., Tao, G., Zhu, L. *et al.* Prediction of pathologic stage in non-small cell lung cancer using machine learning algorithm based on CT image feature analysis. *BMC Cancer* 19, 464 (2019).

7-Qiu, Y.L., Zheng, H., Devos, A. *et al.* A meta-learning approach for genomic survival analysis. *Nat Commun* 11, 6350 (2020).

8-Leontina Postelnicu, Study seeks to identify key factors in adoption of mobile health solution for oncology.