



**TOBB ETÜ**  
Ekonomi ve Teknoloji Üniversitesi

**TOBB EKONOMİ VE TEKNOLOJİ ÜNİVERSİTESİ**

Bilgisayar / Yapay Zekâ Mühendisliği

**BİL 495 — YAP 495**

# Project Proposal

*Proje Önerisi*

Kapsamlı proje teklifi — onaylanmış konu üzerine kurulan dönem sözleşmesi

|                       |                                                                                     |
|-----------------------|-------------------------------------------------------------------------------------|
| <b>Proje Adı</b>      | Giyilebilir Cihaz Verileriyle Kişiselleştirilmiş Sağlık Asistanı                    |
| <b>Takım Üyeleri</b>  | Gülsüm Yıldırım<br>Nurefşan Olfaz<br>Remzi Mert Tehneldere<br>Bektaş Batuhan Kestek |
| <b>Akademik Dönem</b> | 2026 Yaz                                                                            |
| <b>Teslim Tarihi</b>  | 21.06.2026                                                                          |

## 1. Problem Tanımı ve Motivasyon

Giyilebilir teknoloji son on yılda kitlesel bir dönüşüm geçirmiştir. 2026 itibarıyla dünya genelinde 600 milyonu aşkın akıllı saat kullanıcısının bulunduğu tahmin edilmektedir (DemandSage, 2026); bu cihazlar kalp atış hızı, kalp atış hızı değişkenliği (HRV), uyku evreleri ve günlük aktivite gibi fizyolojik sinyalleri kesintisiz biçimde kaydetmektedir. Ne var ki bu zengin veri akışı büyük ölçüde değersiz kalmaktadır: mevcut uygulamaların büyük çoğunluğu bu verileri yalnızca geriye dönük grafikler ve istatistikler olarak görselleştirmekte, kullanıcının sağlığında anlamlı bir değişim yaşandığında bunu proaktif olarak tespit edip bildirememektedir.

Bu tablonun temelinde yapısal bir sorun yatmaktadır. Ticari giyilebilir cihazlar ve üçüncü taraf uygulamalar, fizyolojik değerlendirmeyi popülasyon ortalamalarından türetilmiş genel eşikler üzerine kurmaktadır. Oysa dinlenik nabız ve HRV gibi göstergeler bireyler arasında son derece geniş bir aralıkta değişmektedir: 48 atım/dakikalık bir dinlenik nabız antrenmanlı bir birey için olağan, başkası için endişe verici olabilir; sağlıklı HRV aralığı yaşa, cinsiyete ve kondisyon düzeyine göre büyük farklılıklar gösterir. Genel eşiklere dayanan sistemler bu nedenle ya gereğinden sık yanlış alarm üretmekte ya da gerçek fizyolojik sapmaları gözden kaçırmaktadır. Bir kişinin kendi geçmiş verisinden öğrenilen kişisel normdan sapma, popülasyon normundan sapma kadar anlamlı olsa da birincisini hesaba katan tüketici odaklı bir sistem bugün piyasada ya yoktur ya da kapalı, denetlenemez algoritmalar arkasına gizlenmiştir.

Sorunun bir başka boyutu kadın sağlığına ilişkindir. HRV ve dinlenik nabız, adet döngüsünün fazlarına göre sistematik dalgalanmalar göstermektedir; luteal fazda HRV'nin doğal olarak düştüğü bilinmektedir (Lin et al., 2026). Mevcut uygulamaların büyük çoğunluğu bu fizyolojik gerçeği görmezden gelmekte, regl döngüsü bilgisini anomali değerlendirmesine dahil etmemekte ve dolayısıyla kadın kullanıcılar için yapısal bir yanlış-alarm kaynağı oluşturmaktadır. Yalnızca fizyolojik ölçümler değil, bu ölçümlerin yorumlanma biçimi de toplumsal bir eşitsizlik içermektedir.

Son olarak, veriyi anlama sorunsalı mevcuttur. Giyilebilir cihazlar gitgide daha fazla veri sunmakta; ancak ortalama kullanıcı bu verileri yorumlayacak tıbbi ya da fizyolojik bilgiden yoksun kalmaktadır. Mevcut uygulamaların sunduğu sayısal göstergeler veya renk kodlu skorlar, kullanıcıya "ne yapması gerektiği" konusunda eyleme geçirilebilir bir yönlendirme sağlamamaktadır.

Yukarıda tanımlanan boşluktan en çok etkilenen kullanıcı profili; Apple Watch, Samsung Galaxy Watch veya eşdeğer bir giyilebilir cihaza sahip, sağlığını proaktif biçimde takip etmek isteyen 18-50 yaş aralığındaki bireylerdir. İkincil kullanıcı kitlesi ise kronik stres yaşayanlar, düzenli egzersiz yapan bireyler ve fizyolojik döngülerini daha iyi anlamak isteyen kadınlardır. Sistem, klinik ortam gerektirmeyen ve tıbbi teşhis aracı olmayan bir karar-destek ve farkındalık asistanı olarak konumlandırılmıştır.

Geliştirilen sistem bu boşlukları üç katmanda doldurmayı amaçlamaktadır: her kullanıcı için kişisel bir fizyolojik temel çizgi (baseline) oluşturarak sapmayı bu referansa göre tespit etmek; regl döngüsü fazını bağlamsal bir sinyal olarak değerlendirerek kadın kullanıcılar için daha isabetli bir değerlendirme sunmak; ve makine öğrenmesi modellerinin sayısal çıktılarını büyük dil modeli (LLM) aracılığıyla sade, anlaşılır ve eyleme geçirilebilir bir dile dönüştürmek. Böylece sistem, genel popülasyon eşikleri yerine kişinin kendi alışılmış değerlerini referans alan, şeffaf algoritmalarla çalışan ve gerçek kullanıcıların elindeki cihazlarla doğrudan entegre olan bir sağlık farkındalık asistanı işlevi görmektedir.

## 2. Önceki Çalışmalar / Literatür

Giyilebilir sensör verileriyle fizyolojik durum tespiti, son on yılda hem akademik hem ticari alanda yoğun ilgi gören bir araştırma konusuna dönüşmüştür. Bu bölümde önerilen sistemle doğrudan ilgili akademik çalışmalar ve ticari ürünler incelenmekte; ardından Tablo 1'de sunulan karşılaştırma çerçevesinde mevcut yaklaşımların ortak eksiklikleri ve bu projenin özgün katkısı tartışılmaktadır.

### *Akademik Çalışmalar*

Giyilebilir stres tespiti literatürünün temel referans noktası, Schmidt ve arkadaşları (2018) tarafından yayımlanan WESAD veri kümesidir. Bu çalışmada Empatica E4 ve RespiBAN cihazlarından toplanan çok modlu fizyolojik sinyaller üzerinde XGBoost ve Destek Vektör Makineleri başta olmak üzere çeşitli sınıflandırıcılar test edilmiş; stres/non-stres ikili sınıflandırmasında %85'i aşan doğruluk oranlarına ulaşılmıştır. Ne var ki bu çalışmada modeller tüm katılımcıların verisi üzerinde eğitilmektedir; kişiselleştirme boyutu ele alınmamaktadır. Li ve arkadaşları (2024), tüketici giyilebilir cihaz verisi üzerinde kişiselleştirilmiş ve genelleştirilmiş modelleri doğrudan karşılaştırmış; kişiselleştirilmiş yaklaşımın doğruluğu %95'in üzerine taşıdığını, genel modelin ise %67 civarında kaldığını göstermiştir. Bu bulgu, her kullanıcının kendi geçmiş verisinden öğrenilen kişisel bir referansa dayanan sistem tasarımını doğrudan desteklemekte ve önerilen projenin temel motivasyonunu akademik zemine oturtmaktadır.

Anomali tespiti cephesinde Mishra ve arkadaşları (2020), Fitbit ve Apple Watch verilerini kullanan iki istatistiksel yöntem geliştirmiştir: dinlenik nabızın kişisel ortalamadan sapmasını ölçen RHR-Diff ve kalp atış hızını adım sayısına oranlayan HROS-AD. Bu yöntemler, COVID-19 enfeksiyonunu semptom başlangıcından ortalama dört gün önce tespit etmeyi başarmıştır. Önerilen sistem, bu iki yöntemi doğrudan benimseyerek hastalık-olabilirlik tespiti bileşeninin çekirdeğini oluşturmaktadır. Denetimli modellerin, anomali tespiti motorlarının ve büyük dil modellerinin tek bir ardışık mimaride birleştirilmesi ise en yakın akademik örnekle, Gabrielli ve arkadaşlarının AI on the Pulse sistemiyle (2025) örtüşmektedir. Bu çalışma, UniTS zaman serisi modeliyle tespit edilen anomalileri bir LLM katmanının doğal dile çevirdiği uçtan uca bir boru hattı önermektedir; ancak sistem çevresel sensörler gerektirmekte ve tüketici giyilebilir cihazlarla doğrudan çalışmamaktadır. LLM entegrasyonu açısından Cosentino ve arkadaşlarının geliştirdiği Google PH-LLM (2024), Gemini Ultra'yı uyku ve fitness verisi üzerinde ince ayara tabi tutarak giyilebilir veriyi yorumlayan ilk büyük ölçekli dil modeli çalışmaları arasındadır; ancak periyodik raporlama odaklıdır ve otonom anomali tespiti motorunu barındırmamaktadır. Son olarak Lin ve arkadaşları (2026), mcPHASES veri kümesini Fitbit Sense'ten elde edilen HRV ve RHR verilerini hormonla doğrulanmış döngü faz etiketleriyle eşleştirerek oluşturmuştur; bu, önerilen sistemin regl döngüsü modülünün akademik dayanağını sağlayan açık erişimli tek veri kümesidir.

### *Ticari Ürünler*

Ticari ekosistemde en yakın rakip Bevel Health'tir (Bevel, 2025). Uygulama, HealthKit verilerini kısa ve uzun dönemli hareketli ortalamalarla karşılaştıran bir baseline mekanizması ve "Bevel Intelligence" adlı LLM koçunu birleştirmektedir; ancak yalnızca iOS platformunu desteklemekte, algoritması kapalı kutu olarak kalmakta ve anomali tespiti için denetimsiz öğrenme yöntemi kullanmamaktadır. Welltory, OpenAI altyapısıyla desteklenen yapay zeka koçu ve kişisel baseline karşılaştırmasıyla dikkat çekmektedir (Welltory, 2025); ne var ki kullanıcı raporları LLM'in zaman zaman gecikmeli veriler üzerinden tutarsız

çıktılar ürettiğini göstermektedir. Garmin'in Firstbeat analitiği, HRV ve solunum hızını birleştiren kişiselleştirilmiş bir stres puanı üretmekte (Firstbeat, 2016); ancak algoritmik çıktıyı açıklayan bir dil modeli bulunmamakta ve öneriler statik şablon metinlerinden oluşmaktadır. WHOOP ise 30 günlük hareketli ortalamaya dayalı güçlü bir kişisel baseline sunan ve OpenAI tabanlı WHOOP Coach'u barındıran en olgun ticari platform olmakla birlikte (WHOOP, 2025), yalnızca kullanıcı sorusu geldiğinde yanıt veren reaktif bir asistan olarak çalışmakta ve proaktif anomali açıklaması yapmamaktadır.

### Karşılaştırma

| Kaynak / Ürün            | Kişisel Baseline | Denetimsiz Anomali | LLM Açıklama | iOS + Android | Adet Döngüsü | Şeffaf Algoritma |
|--------------------------|------------------|--------------------|--------------|---------------|--------------|------------------|
| Schmidt ve ark. (2018)   | ×                | ×                  | ×            | —             | ×            | ✓                |
| Li ve ark. (2024)        | ✓                | ×                  | ×            | —             | ×            | ✓                |
| Mishra ve ark. (2020)    | ✓                | ✓                  | ×            | —             | ×            | ✓                |
| Gabrielli ve ark. (2025) | ✓                | ✓                  | ✓            | —             | ×            | ✓                |
| Cosentino ve ark. (2024) | ✓                | ×                  | ✓            | —             | ×            | ×                |
| Lin ve ark. (2026)       | ✓                | ×                  | ×            | —             | ✓            | ✓                |
| Bevel Health             | ✓                | ×                  | ✓            | ×             | ×            | ×                |
| Welltory                 | ✓                | ×                  | ✓            | ✓             | ×            | ×                |
| Garmin                   | ✓                | ×                  | ×            | ✓             | ×            | ×                |
| WHOOP                    | ✓                | ×                  | ✓ (reaktif)  | ✓             | ×            | ×                |
| <b>Bu proje</b>          | ✓                | ✓                  | ✓ (proaktif) | ✓             | ✓            | ✓                |

Tablo 1. Mevcut akademik çalışmalar ve ticari ürünlerle karşılaştırma.

Tablo 1'de görüldüğü üzere, mevcut akademik çalışmalar ya yüksek doğruluklu denetimli sınıflandırma modelleri sunarken kişisel baseline ve LLM entegrasyonunu ihmal etmekte, ya da LLM kullanırken denetimsiz anomali tespitini atlamaktadır. Ticari ürünler ise kişisel baseline mekanizmalarını kapalı algoritmalar ardına saklamakta, Android kullanıcılarını büyük ölçüde dışarıda bırakmakta ve adet döngüsü gibi bağlamsal sinyalleri otomatik olarak baseline hesabına dahil etmemektedir. Önerilen sistem, bu boşlukları denetimli stres sınıflandırması ile denetimsiz anomali tespitini birleştiren hibrit mimarisi, HealthKit ve Health Connect üzerinden iOS ve Android platformlarını kapsayan cihaz-bağımsız tasarımı ve regl döngüsü fazını baseline hesabına otomatik dahil eden şeffaf bütünlük katmanıyla kapatmayı hedeflemektedir.

### 3. Çözüm Yaklaşımı ve Yöntem

#### 3.1 Temel Yaklaşım: Kişisel Temel Çizgi (Baseline)

Sistemin merkezindeki fikir, sağlık değerlendirmesini popülasyon ortalamalarına dayanan genel eşikler yerine her kullanıcıya özel bir temel çizgi üzerine kurmaktır. Bu tercih fizyolojik bir gerçeğe dayanır: dinlenik nabız ve kalp atış hızı değişkenliği (HRV) gibi göstergeler bireyler arasında geniş bir aralıkta değişir. Bir kişi için tamamen normal olan bir dinlenik nabız değeri başka biri için olağandışı olabilir; benzer şekilde sağlıklı HRV aralığı yaşa, cinsiyete ve genel kondisyona göre büyük farklılık gösterir. Genel eşiklere dayanan sistemler bu nedenle ya çok sık yanlış alarm verir ya da gerçek sapmaları kaçırır. Kişisel baseline yaklaşımı, her kullanıcının kendi geçmiş verisinden öğrenilen normal aralığı referans alarak bu sorunu çözer (Nkurikiyeyezu et al., 2019). Sistem, kullanıcının ilk birkaç haftalık verisinden kişisel normal aralığını öğrenir ve bu çizgiden anlamlı, sürekli sapmaları tespit eder.

#### 3.2 Sistemin Yapay Zekâ Bileşenleri

Sistem dört analitik bileşenden oluşur: iki denetimli model, bir denetimsiz istatistiksel tespit katmanı, bir kural tabanlı hesaplama ve bunları tamamlayan bir büyük dil modeli (LLM) yorumlama katmanı. Ayrıca kadınlar için regl döngüsü fazı opsiyonel bir bağlamsal sinyal olarak sisteme entegre edilir. Her bileşenin yöntem seçimi farklı bir gerekçeye dayanır ve aşağıda ayrı ayrı açıklanmıştır.

#### 3.3 Model 1: Denetimli Stres Sınıflandırma

**Özet:** Bu bileşen, giyilebilir cihazlardan elde edilen fizyolojik ölçümler ve denetimli makine öğrenmesiyle stres tespitinin yapılabilirliğini araştırır. Akıllı saatler, kalp atış hızı ve HRV gibi stres göstergelerini gerçek zamanlı ve müdahalesiz biçimde ölçerek bu tespiti mümkün kılar. Stresin erken tespiti, fiziksel ve zihinsel sağlık üzerindeki uzun vadeli olumsuz sonuçların önlenmesine katkı sağlayabilir. Etiketli WESAD veri seti üzerinde öznitelik tabanlı bir XGBoost sınıflandırıcısı eğitilir ve modelin laboratuvar dışı genellenebilirliği LifeSnaps veri setiyle sınanır.

**Yöntem seçimi gerekçesi:** XGBoost üç nedenle seçilmiştir. Birincisi, ağaç tabanlı bir gradyan artırma (gradient boosting) yöntemi olarak öznitelik mühendisliğiyle hazırlanmış tablo verisinde güçlü performans gösterir (Chen & Guestrin, 2016) ve fizyolojik sinyallerden çıkarılan öznitelikler bu yapıya uygundur. İkincisi, küçük ve orta ölçekli veri setlerinde (WESAD 15 katılımcılıdır) aşırı öğrenmeye karşı derin öğrenme yöntemlerinden daha dayanıklıdır ve dahili düzenleme mekanizmaları içerir. Üçüncüsü, eğitim ve çıkarım maliyeti düşüktür; bu, sınırlı süre ve hesaplama kaynağıyla çalışan bir ekip için uygundur.

**Veri ve öznitelikler:** Model, ham fizyolojik sinyalden kısa zaman pencerelerinde (yaklaşık 60 saniye) çıkarılan özniteliklerle beslenir: ortalama kalp atış hızı, RMSSD ve SDNN gibi zaman-alanı HRV metrikleri. Eğitim, WESAD veri setinin yalnızca bilek (Empatica E4) verisiyle yapılır; göğüs cihazı verisi ve EDA (elektrodermal aktivite) bilinçli olarak hariç tutulur, çünkü bu sinyaller gerçek kullanıcının cihazından (HealthKit/Health Connect) elde edilemez. Bu, eğitim verisi ile çıkarım verisi arasındaki tutarlılığı korur ve gerçek dünyada işe yaramayacak bir özelliğe bağımlılığı önler.

**Değerlendirme:** Modelin gerçek-yaşam koşullarına genellenip genellenmediği, laboratuvar dışında toplanmış LifeSnaps veri setiyle doğrulanır. WESAD kontrollü bir laboratuvar protokolünden gelirken,

LifeSnaps 71 katılımcıdan dört ay boyunca doğal yaşam koşullarında toplanmıştır (Yfantidou et al., 2022); bu nedenle modelin laboratuvar dışındaki davranışını sınamak için uygun bir doğrulama kaynağıdır. Önemli bir nüans: LifeSnaps'teki stres etiketi kullanıcı tarafından doğrulanmış bir "stresli/değil" etiketi değil, Fitbit'in kapalı algoritmasıyla üretilmiş üretilmiş bir skordur; bu nedenle LifeSnaps doğrulaması kesin bir doğruluk ölçümü olarak değil, modelin gerçek-yaşam verisindeki davranışını değerlendiren bir kıyaslama olarak yorumlanır. Tüm değerlendirmelerde, aynı kişinin verisinin hem eğitimde hem testte bulunmasını önlemek için kişi-bazlı bölme (subject-wise split) yöntemi kullanılır.

### **3.4 Model 2: Denetimsiz Anomali ve Hastalık-Olabilirlik Tespiti**

Bu bileşen, kullanıcının kişisel baseline'ından çok-günlük sapmaları tespit eder ve olağandışı bir fizyolojik durumu (aşırı yorgunluk, hastalık öncesi dönem gibi) erken uyarı olarak işaretler. Bir makine öğrenmesi modeli eğitmek yerine istatistiksel yöntemlere dayanması bilinçli bir tasarım kararıdır; çünkü "normal" kavramı her kullanıcının kendi geçmişinden öğrenilir, dolayısıyla dışarıdan etiketli hastalık verisine ihtiyaç duyulmaz. Bu yaklaşım hem etiketli veri eksikliği sorununu ortadan kaldırır hem de gizliliği güçlendirir.

Anomali tespitine kişisel z-skor tabanlı basit bir yaklaşımla başlanacak; bu temel doğrulandıktan sonra RHR-Diff ve HROS-AD yöntemleri ileride bonus olarak devreye alınması planlanmaktadır. Kullanılan yöntemler hakemli literatürde belirti-öncesi enfeksiyon tespitinde doğrulanmış iki tekniktir. Birincisi RHR-Diff: dinlenik nabızın, 28 günlük kayan bir pencereyle oluşturulan kişisel ortalamaya göre anlamlı ve sürekli yükselip yükselmediğini standartlaştırılmış kalıntılarla ölçer. İkincisi HROS-AD: kalp atış hızını adım sayısına oranlayarak (Heart Rate Over Steps) yeni bir öznitelik üretir; bu, aktivite etkisini doğal olarak normalize eder ve "yüksek aktiviteyle birlikte yüksek nabız" (egzersiz, normal) ile "düşük aktiviteyle birlikte yüksek nabız" (anlamlı sinyal) durumlarını ayırt eder. Bu iki yönteme HRV düşüş kontrolü eklenir; çünkü literatür HRV'deki düşüşün enfeksiyon başlangıcının güvenilir bir öncü göstergesi olduğunu göstermektedir (Mishra et al., 2020).

Sistemin yalnızca tek bir anormal okuma üzerine değil, birkaç gün süren kalıcı sapma üzerine uyarı vermesi temel bir tasarım ilkesidir; bu çok-günlük kuralı, egzersiz, akut stres anı gibi kısa süreli olayları doğal olarak filtreler. Yöntemler, 71 katılımcılı, dört aylık gerçek-yaşam verisi içeren LifeSnaps veri setiyle yanlış-alarm açısından doğrulanır. Sistem hiçbir zaman belirli bir hastalığı isimlendirmez; yalnızca "vücudunuz olağandışı bir yük altında olabilir" anlamına gelen bir olabilirlik sinyali üretir.

### **3.5 Model 3: Denetimli Aktivite Tanıma**

Bu model, kullanıcının gün içindeki aktivite türünü ve yoğunluğunu (yatma, oturma, yürüme, koşma ve farklı yoğunluk seviyeleri) sınıflandırır. Yöntem olarak Random Forest seçilmiştir (Breiman, 2001); gerekçesi, ilgili literatürde dakika bazlı giyilebilir cihaz verisiyle (Apple Watch ve Fitbit) yapılan aktivite sınıflandırmasında en yüksek doğruluğu veren yöntemlerden biri olarak raporlanması ve yorumlanabilir öznitelik önemi sağlamasıdır. Model, dakika bazında kalp atış hızı, adım, mesafe ve aktif kalori öznitelikleriyle beslenir; eğitim için Apple Watch ve Fitbit cihazlarıyla toplanmış, etiketli, halka açık bir aktivite veri seti kullanılır (Fuller et al., 2020).

Bu bileşen iki işlev görür: kullanıcıya günlük aktivite ve enerji harcaması özeti sunar; ve daha önemlisi, Model 2'nin ürettiği uyarılara bağlam katarak yüksek bir nabızın egzersizden mi yoksa anlamlı bir baseline sapmasından mı kaynaklandığının ayırt edilmesine yardımcı olur.

### 3.6 Uyku Skoru: Kural Tabanlı Hesaplama

Uyku kalitesi için ayrı bir denetimli model eğitmek yerine, yerleşik uyku tıbbi metriklerine dayanan açık-formüllü bir uyku skoru hesaplanır. Girdi olarak uyku verimliliği (yatakta geçen süreye oranla gerçek uyku), toplam uyku süresi, gece uyanma sayısı ve (cihaz destekliyorsa) derin/REM/hafif uyku evre oranları kullanılır; bu metrikler ağırlıklandırılarak 0-100 aralığında bir skora dönüştürülür. Bu tercihin iki gerekçesi vardır: birincisi, görev için güvenilir ve doğrudan uygulanabilir bir eğitim altyapısı sınırlıdır, kural-tabanlı hesaplama daha basit ve sağlamdır; ikincisi, açık formül, ticari ürünlerin (Fitbit, Oura) kapalı algoritmalarına karşı projenin temel farklılaşma noktalarından biri olan şeffaflığı güçlendirir.

### 3.7 Bağlamsal Sinyal: Regl Döngüsü Fazı (Opsiyonel)

Kullanıcı, HealthKit/Health Connect'in üreme sağlığı bölümünde regl döngüsü bilgisini girmişse, sistem bu veriye erişerek döngünün hangi fazında (regl, foliküler, ovulasyon, luteal) olduğunu belirleyebilir. Bu bilgi iki amaçla kullanılır: Model 2'nin yanlış alarmlarını azaltmak (örneğin luteal fazda beklenen HRV düşüşünü anomali olarak işaretlememek) ve LLM yorumuna bağlam sağlamak (örneğin regl döneminde gözlenen yüksek nabız normalleştirici bir dille açıklamak). Kullanıcı açıkça izin vermeden bu veriye hiç erişilmez. Üreme sağlığı verisinin son derece hassas bir kategori olması nedeniyle bu güvence sistemin gizlilik tasarımının ayrılmaz bir parçasıdır.

### 3.8 LLM Yorumlama Katmanı

Yukarıdaki bileşenlerin sayısal çıktıları (stres durumu, anomali skoru, aktivite özeti, uyku skoru) ile varsa regl fazı bilgisi, hazır bir büyük dil modeli API'sine yapılandırılmış bir istem (prompt) olarak verilir. LLM bu çıktıları tek bir sade, anlaşılır paragrafta birleştirir. Bu katman eğitim gerektirmez; yalnızca dikkatli istem tasarımıyla yönlendirilir. Tasarım ilkesi, LLM'in teşhis koymaması, normalleştirici ve farkındalık artırıcı bir dil kullanması ve gerektiğinde kullanıcıyı bir uzmana yönlendirmesidir. Örneğin sistem "gribiniz var" demez; bunun yerine "dinlenme nabızınız son birkaç gündür kişisel ortalamanızın üzerinde, bu vücudunuzun bir şeyle uğraştığının işareti olabilir, kendinize dikkat edin ve belirtiler sürerse bir uzmana danışın" gibi temkinli bir dil kullanır.

### 3.9 Kapsayıcı Tasarım Kararları

İki temel karar tüm sistemi kapsar. Birincisi, eğitim ve çıkarım ayrımı: tüm modeller yalnızca bir kez, halka açık veri setleri üzerinde çevrimdışı olarak eğitilir; kullanıcının kişisel verisi modeli yeniden eğitmek için asla kullanılmaz, yalnızca çıkarım anında kişiselleştirme amacıyla işlenir. Bu, sistemin temel gizlilik garantisidir. İkincisi, donanım bağımsızlığı: proje kendi giyilebilir cihazını üretmez; bunun yerine kullanıcının zaten sahip olduğu cihazdan (Apple Watch, Galaxy Watch ve benzeri) HealthKit/Health Connect aracılığıyla veri okur. Bu, sistemi marka-bağımsız kılar, donanım geliştirme yükünü ortadan kaldırır ve gerçek bir kullanım senaryosuna uygun, ölçeklenebilir bir çözüm sağlar. Üçüncüsü, gizlilik odaklı LLM entegrasyonu: kullanıcının ham fizyolojik verisi hiçbir zaman üçüncü taraf LLM API'sine gönderilmez. API'ye yalnızca modellerin ürettiği anonim sayısal çıktılar (stres durumu, anomali skoru, uyku skoru, aktivite özeti) iletilir; bu çıktılar tek başına bir bireyi tanımlamaya yetmez. Regl fazı gibi hassas bağlamsal bilgiler ise yalnızca kullanıcının açık onayıyla ve yalnızca sayısal faz kodu olarak (0–3)

aktarılır. Bu tasarım, KVKK kapsamındaki özel nitelikli kişisel veri (sağlık ve biyometrik veri) yükümlülüklerini minimize etmeyi amaçlamaktadır. İleri aşamada, cihaz üzerinde çalışan küçük dil modeli (SLM) kullanımı değerlendirilerek API bağımlılığı tamamen ortadan kaldırılabilir ve KVKK yükümlülükleri daha da güçlendirilebilir.

## 4. Mimari Taslağı

### 4.1 Genel Bakış

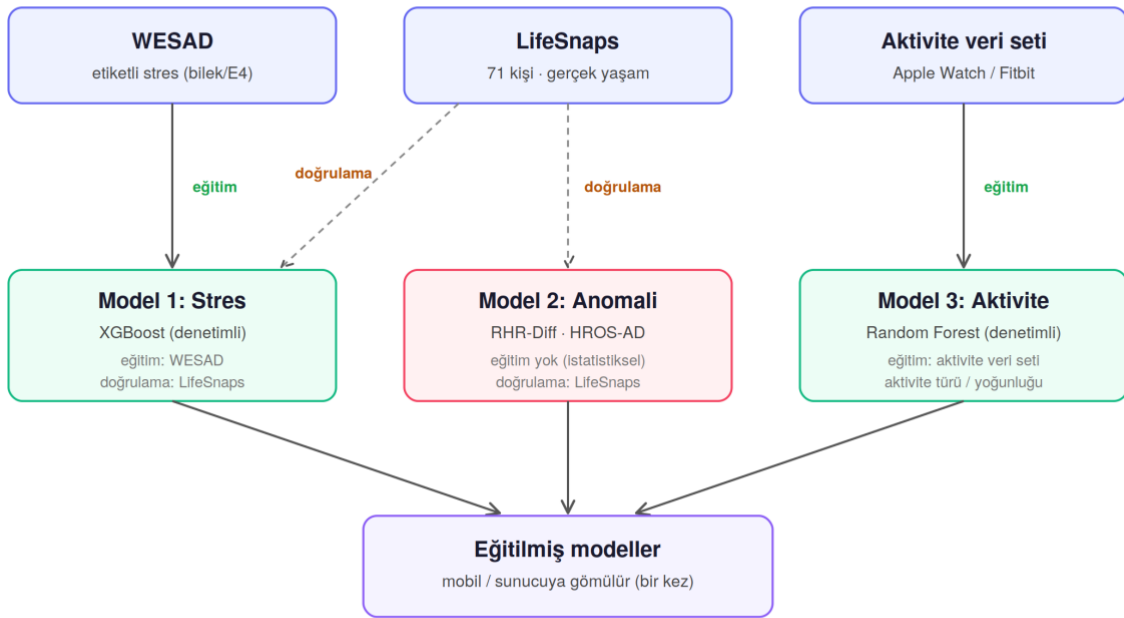
Sistem, birbirinden kesin biçimde ayrılmış iki çalışma hattından oluşur. Çevrimdışı eğitim hattı yalnızca halka açık veri setlerini kullanır ve bir kez çalıştırılır; ortaya çıkan modeller mobil uygulamaya veya sunucuya gömülür. Çalışma zamanı hattı ise bu eğitilmiş modelleri kullanarak gerçek kullanıcının cihazından gelen veriyi her gün işler. Bu ayrım, sistemin temel gizlilik garantisidir: kullanıcının kişisel verisi model eğitiminde asla kullanılmaz, yalnızca çıkarım anında kişiselleştirme amacıyla işlenir.

### 4.2 Çevrimdışı Eğitim Hattı

Bu hatta üç açık veri seti işlenir. WESAD ile Model 1 eğitilir ve gerçek yaşam verisi olan LifeSnaps ile doğrulanır. Halka açık Apple Watch/Fitbit aktivite veri seti Model 3'ü eğitir. Model 2 eğitim gerektirmez; istatistiksel yöntemleri (RHR-Diff, HROS-AD) yine LifeSnaps gerçek yaşam verisiyle doğrulanır. Eğitilmiş modeller mobil uygulamaya/sunucuya gömülür. Uyku skoru ve regl fazı bu hatta yer almaz, çünkü eğitim gerektirmeyip çalışma zamanında hesaplanırlar.

**Şekil 1. Çevrimdışı Eğitim Hattı**

Yalnızca halka açık veri setleri — bir kez çalıştırılır



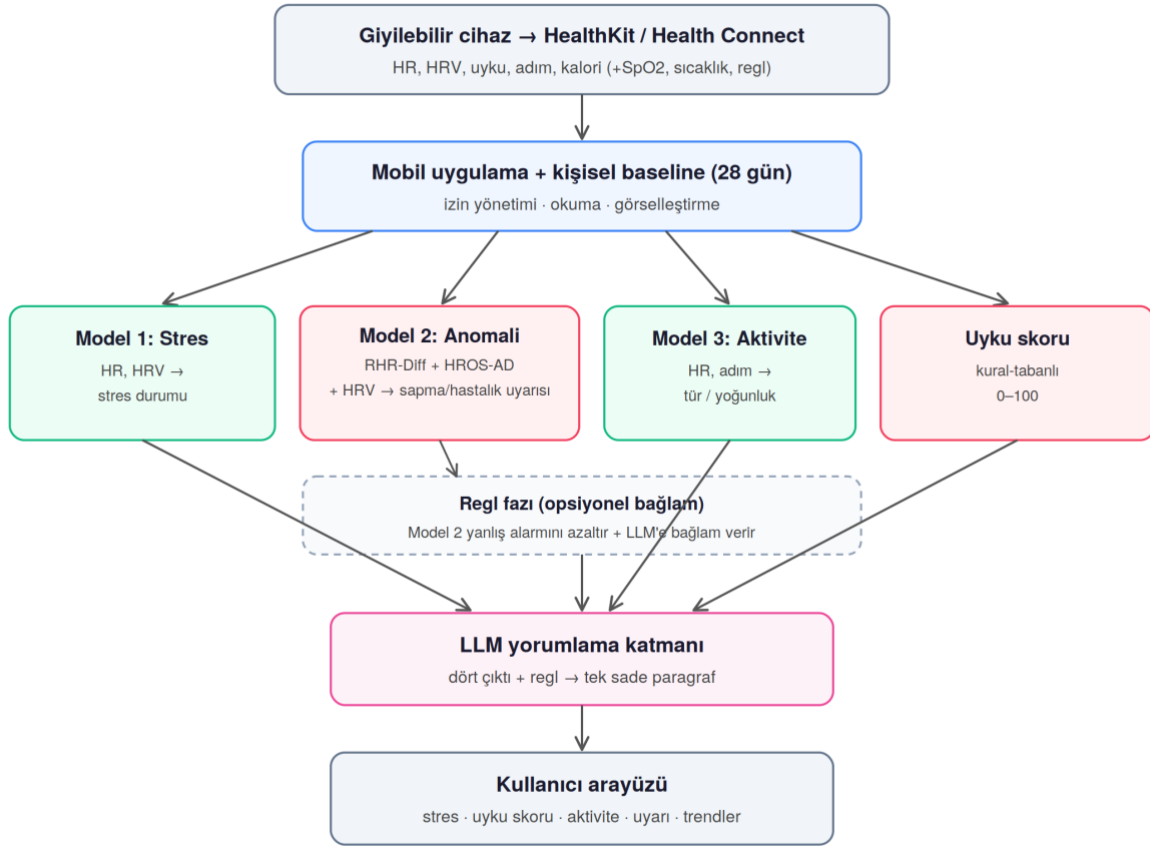
Uyku skoru ve regl fazı eğitim gerektirmez — çalışma zamanında hesaplanır.

#### 4.3 Çalışma Zamanı Hattı

Kullanıcının izniyle HealthKit (iOS) veya Health Connect (Android) üzerinden çekilen sinyaller mobil uygulamada toplanır ve kişisel baseline (28 günlük kayan pencere) ile karşılaştırılır. Dört bileşen paralel çalışır: Model 1 anlık stresi, Model 2 baseline'dan sapmayı, Model 3 aktivite türünü hesaplar; uyku skoru kural tabanlı olarak üretilir. Varsa regl fazı, hem Model 2'nin yanlış alarımını azaltır hem de LLM'e bağlam sağlar. LLM katmanı tüm çıktıları tek bir sade paragrafta birleştirir ve kullanıcıya içgörü, uyarı ve gerektiğinde uzmana yönlendirme olarak sunar.

### Şekil 2. Çalışma Zamanı (Runtime) Mimarisi

Gerçek kullanıcı verisi — yalnızca çıkarımda işlenir



Tেধিস কোমজ; ফার্কিন্দালিক সাগ্গার বে গেরেক্টিগ্গিন্দে উজমানে য়োন্ডেরির্.

ক্টিশেল বেরি য়ালিন্জকা চ্কারিম্দা িশলেনির্, মডেল ংগ্টিমিন্দে কুলানিলমাজ.

#### 4.4 Veri Katmanı ve Sinyaller

Çalışma zamanında okunan sinyaller üç kademeye ayrılır. Çekirdek sinyaller her zaman kullanılır: kalp atış hızı (heartRate, restingHeartRate), HRV (heartRateVariabilitySDNN), uyku (sleepAnalysis), aktivite (stepCount, activeEnergyBurned, distanceWalkingRunning) (Apple Inc., t.y.; Google, t.y.). Bağlamsal sinyaller cihaz sağladığında kullanılır: SpO2 (oxygenSaturation), solunum hızı (respiratoryRate), cilt sıcaklığı, regl döngüsü fazı. EDA, ham ECG/EMG ve anket/biyokimyasal veriler tamamen kapsam

dışındır. Mimari, HealthKit ve Health Connect'te her veri türü için kullanıcıdan ayrı izin alındığından ve kullanıcı her an geri çekebileceği için eksik veriye dayanıklı tasarlanmıştır.

#### 4.5 Eğitim ve Doğrulama Veri Setleri

| Veri seti                             | Özellikleri                                                | Kullanım  | İlgili Bileşen     |
|---------------------------------------|------------------------------------------------------------|-----------|--------------------|
| WESAD                                 | 15 kişi, laboratuvar, etiketli stres; bilek/E4, BVP→HR/HRV | Eğitim    | Model 1            |
| LifeSnaps                             | 71 kişi, 4 ay, gerçek yaşam, Fitbit Sense                  | Doğrulama | Model 1 ve Model 2 |
| Apple Watch/Fitbit Aktivite Veri Seti | Dakika bazında HR/adım/mesafe/kalori, etiketli aktivite    | Eğitim    | Model 3            |

#### 4.6 Modüllerin Sorumlulukları

| Bileşen                      | Sorumluluk / İşlev                                                                                 |
|------------------------------|----------------------------------------------------------------------------------------------------|
| <b>Mobil Uygulama</b>        | Veri erişimi, kişisel baseline'ın yerel tutulması ve görselleştirilmesi.                           |
| <b>ML Çıkarım Motoru</b>     | Dört bileşenin hesaplanması.                                                                       |
| <b>LLM Katmanı</b>           | Yalnızca yorumlama ve sade dil üretimi (Teşhis koyma yetkisi yoktur).                              |
| <b>Katmanlı Yapı (Genel)</b> | Her bileşenin bağımsız test edilebilmesini ve birinin değişmesinin diğerlerini bozmamasını sağlar. |

## 5. Hedef Kullanıcı, Kapsam ve Başarı Kriterleri

### 5.1 Kapsam İçi (In-Scope)

Sistem, kullanıcının zaten sahip olduğu giyilebilir cihazdan HealthKit (iOS) veya Health Connect (Android) üzerinden veri çeken, tek platformda çalışan bir mobil prototip olarak geliştirilecektir. Kapsam içindeki bileşenler: çekirdek sinyallerin (HR, HRV, uyku, aktivite/adım/kalori) okunması, görselleştirilmesi ve her kullanıcı için kişisel temel çizgi oluşturulması; dört analitik bileşen (XGBoost tabanlı stres sınıflandırma, istatistiksel anomali/hastalık-olabilirlik tespiti, Random Forest tabanlı aktivite tanıma, kural-tabanlı uyku skoru); tüm modellerin halka açık veri setleri üzerinde bir kez çevrimdışı eğitilmesi ve kişi bazlı bölme yöntemiyle değerlendirilmesi; modellerin sayısal çıktılarını sade dile çeviren bir LLM yorumlama katmanı. SpO2, solunum hızı ve cilt sıcaklığı, cihaz desteklediğinde ve kullanıcı izin verdiğinde devreye giren bağlamsal sinyaller olarak kapsam içindedir; bu sinyaller mevcut olmadığında

sistem yalnızca çekirdek sinyallerle çalışmaya devam eder. Regl döngüsü fazı ise yalnızca kullanıcının açık onayıyla erişilen, tamamen isteğe bağlı bir bağlam katmanıdır.

### 5.2 Kapsam Dışı (Out-of-Scope)

Sistem bir teşhis aracı değildir. Kapsam dışı: belirli bir hastalığın (grip, COVID-19 vb.) isimlendirilmesi veya teşhis edilmesi; EDA tabanlı stres ölçümü (Apple Watch'ta donanım yok, HealthKit/Health Connect'te standart tip değil); ham ECG/EMG analizi, aritmi/kalp hastalığı tespiti, kan basıncı ve kan şekeri gibi tıbbi ölçümler; klinik düzeyde uyku bozukluğu (uykusuzluk, uyku apnesi) teşhisi; kendi giyilebilir donanımımızın üretimi; çoklu platform desteği (MVP tek platform); kullanıcı verisiyle modelin yeniden eğitilmesi. Ham sinyal üzerinde LSTM ve autoencoder gibi gelişmiş derin öğrenme yöntemleri, zaman elverirse değerlendirilecek ileri aşama hedefleridir, MVP kapsamında değildir.

### 5.3 Ölçülebilir Başarı Kriterleri

|   | Kriter                                                                                                                                | Hedef                                                   |
|---|---------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------|
| 1 | Model 1 stres sınıflandırma başarısı (WESAD, ikili, kişi bazlı bölme)                                                                 | $\geq \%80$ doğruluk ve makro-F1                        |
| 2 | Model 1 gerçek yaşam genellemesi (LifeSnaps doğrulamasında başarı)                                                                    | makul performans / referans skorla uyum                 |
| 3 | Model 2 anomali değerlendirmesi:<br>(a) sağlıklı günlerde yanlış pozitif oranı<br>(b) hastalık bildirilen günlerde sinyal yükselmesi  | (a) $\leq \%10$ yanlış alarm<br>(b) $\geq \%60$ örtüşme |
| 4 | Model 3 aktivite sınıflandırma doğruluğu (kişi bazlı bölme)                                                                           | $\geq \%80$ doğruluk                                    |
| 5 | Uçtan uca gecikme (veri çekimi $\rightarrow$ ekranda gösterim)                                                                        | $\leq 5$ saniye                                         |
| 6 | LLM çıktısı güvenliği: teşhis içermeyen, normalleştirici dil kullanan yanıtların oranı (örnek çıktılar üzerinde ekip değerlendirmesi) | $\geq \%90$ uygun çıktı                                 |

Bu hedef değerler, ilgili literatürdeki benzer çalışmaların raporladığı performans aralıkları esas alınarak belirlenmiştir ve ilk deney sonuçları doğrultusunda revize edilebilir.

## 6. Çalışma Planı ve İş Bölümü

Tablo 2. Çalışma Planı ve İş Bölümü

| Rapor                 | Tarih  | Batuhan                                                  | Gülsüm                          | Nurefşan                                                  | Mert                       |
|-----------------------|--------|----------------------------------------------------------|---------------------------------|-----------------------------------------------------------|----------------------------|
| <b>Proposal</b>       | 21 Haz | Problem, literatür, kapsam, başarı kriterleri, kaynaklar | Çözüm yaklaşımı, mimari (öncü)  | Mimari (ortak), çalışma planı & iş bölümü, kalite kontrol | Riskler, beklenen çıktılar |
| <b>Specifications</b> | 28 Haz | Introduction, etik sorunlar, references                  | Description (fonksiyonlar, out- | Requirements (FR+NFR)                                     | Constraints                |

| Rapor             | Tarih  | Batuhan                                                                  | Gülsüm                                                        | Nurefşan                             | Mert                                                            |
|-------------------|--------|--------------------------------------------------------------------------|---------------------------------------------------------------|--------------------------------------|-----------------------------------------------------------------|
|                   |        |                                                                          | of-scope)                                                     |                                      |                                                                 |
| <b>Analysis</b>   | 12 Tem | Introduction, current system, overview, senaryolar, glossary, references | Use case model, dynamic models, UI mockup (öncü)              | Class model, UI mockup (ortak), NFR  | Functional requirements, pseudo requirements                    |
| <b>PKE-Plan</b>   | 31 Tem | Beklenen etkiler                                                         | Standartlara katkı                                            | Kalite kontrol                       | Kullanılan standartlar, proje kısıtları (öncü)                  |
| <b>HLD</b>        | 6 Ağu  | Purpose, design goals, definitions/acronyms, glossary, references        | Subsystem decomposition, HW/SW mapping, global control (öncü) | Subsystem services, overview (ortak) | Persistent data, access control & security, boundary conditions |
| <b>Web Sitesi</b> | 6 Ağu  | Proje açıklaması, içerik metinleri, landing özet                         | Site kurulumu, teknik altyapı                                 | Rapor linkleri, kalite kontrol       | Takım sayfası, GitHub linki, danışman bilgisi, iletişim         |

Tablo 3. 496 Modül Sorumlulukları

| Modül                       | Sorumlu          | Açıklama                                                |
|-----------------------------|------------------|---------------------------------------------------------|
| Denetimli stres modeli      | Gülsüm, Nurefşan | XGBoost eğitimi, değerlendirme, WESAD pipeline          |
| Denetimsiz anomali modeli   | Mert, Batuhan    | RHR-Diff, HROS-AD, z-skor baseline, LifeSnaps doğrulama |
| Mobil platform entegrasyonu | Gülsüm, Batuhan  | HealthKit / Health Connect bağlantısı, izin yönetimi    |
| Arayüz ve LLM katmanı       | Mert, Nurefşan   | Kullanıcı arayüzü, prompt tasarımı, LLM entegrasyonu    |
| Sistem entegrasyonu ve test | Tüm ekip         | Modüllerin birleştirilmesi, uçtan uca test              |

## 7. Riskler ve Önlemler

### 7.1 Olasılık × Etki Değerlendirmesi

|    | Risk                                                | Olasılık | Etki | Öncelik |
|----|-----------------------------------------------------|----------|------|---------|
| R1 | EDA olmadan stres modelinin doğruluğu literatürdeki | Yüksek   | Orta | Yüksek  |

|     |                                                                                              |        |        |        |
|-----|----------------------------------------------------------------------------------------------|--------|--------|--------|
|     | tam-modalite sonuçların altında kalması                                                      |        |        |        |
| R2  | HealthKit'te HRV'nin seyrek/düzensiz gelmesi nedeniyle baseline kurulamaması                 | Orta   | Yüksek | Yüksek |
| R3  | Üç modelin paralel geliştirilmesinin dönem süresini aşması                                   | Orta   | Yüksek | Yüksek |
| R4  | Mobil platform (HealthKit/Health Connect) entegrasyonunun beklenenden zor olması             | Orta   | Orta   | Orta   |
| R5  | Anomali tespitinde yanlış alarmlar (seyahat, irtifa, regl, vb.)                              | Orta   | Orta   | Orta   |
| R6  | Veri setlerinin dar demografisi nedeniyle modelin geniş kitleye genellenememesi              | Yüksek | Düşük  | Orta   |
| R7  | LifeSnaps stres etiketinin Fitbit'in kapalı algoritmasından gelmesi (doğrulama belirsizliği) | Orta   | Düşük  | Düşük  |
| R8  | Gerçek giyilebilir cihaza erişim olmaması nedeniyle canlı test yapılamaması                  | Orta   | Düşük  | Düşük  |
| R9  | LLM çıktısının yanlış/teşhis-niteliğinde ifadeler üretmesi                                   | Yüksek | Yüksek | Orta   |
| R10 | Yeterli Geçmiş Veri Oluşamaması                                                              | Yüksek | Orta   | Yüksek |
| R11 | Cihaz arası veri değişkenliği                                                                | Yüksek | Orta   | Yüksek |

## 7.2 Riskler ve Önlemler

*R1 — EDA eksikliği ve stres doğruluğu:* WESAD'ın orijinal yüksek doğruluğu büyük ölçüde EDA sinyalinin gelir; biz EDA'yı (gerçek cihazdan alınamadığı için) bilinçli olarak dışarıda bırakıyoruz, dolayısıyla doğruluğumuz tam-modalite sonuçların altında kalabilir.

*Önlem:* Modeli iki varyantla (tam-modalite, yalnızca araştırma kıyası için; ve deploy-edilebilir bilek-only) eğitip farkı raporlamak; eksikliği kişiselleştirme ve çoklu-sinyal füzyonuyla telafi etmek.

*R2 — Seyrek HRV ve baseline:* HealthKit'te HRV pasif/sürekli değil, çoğunlukla "Breathe/Mindfulness" egzersizleri sırasında ölçülür; bu, baseline için yeterli örnek toplamayı zorlaştırabilir.

*Önlem:* Sistemi HRV'ye tek başına bağımlı kılmamak; dinlenik nabız, uyku ve aktiviteyi de baseline'a katmak (sistem zaten çok-sinyalli). Yeterli HRV verisi olmayan kullanıcıda HRV'siz çalışacak bir "azaltılmış mod" tanımlamak. Mümkünse uyku-tabanlı HRV veren cihazları (Whoop/Oura ekosistemi) gelecek çalışma olarak not etmek.

*R3 — Üç modelin zaman riski:* İki denetimli model (Model 1, Model 3) + istatistiksel katman + uyku skoru + LLM, 4 kişilik bir ekip için iddialı bir yük olabilir.

*Önlem:* Net bir MVP/stretch ayrımı. MVP çekirdeği: Model 1 + Model 2 + uyku skoru + LLM. Model 3 (aktivite tanıma) ve regl bağlamı, zaman elverirse eklenecek; gecikme olursa bunlar stretch'e kaydırılır, MVP teslimi riske atılmaz. Faz bazlı takvim ve haftalık ara hedeflerle erken sapma tespiti.

*R4 — Mobil entegrasyon zorluğu:* Gerçek bir HealthKit/Health Connect entegrasyonu izinler, gerçek cihaz ve platform-spesifik kısıtlar nedeniyle beklenenden zaman alabilir.

*Önlem:* Tek platforma odaklanmak; gerçek cihaz erişimi sınırlıysa, WESAD/LifeSnaps gerçek verisini HealthKit/Health Connect'e zaman-damgalı örnekler olarak yazıp tüm pipeline'ı sentetik-enjeksiyonla demo etmek. Entegrasyonu projenin erken fazına almak.

*R5 — Yanlış alarmlar:* Seyahat/jet-lag, yüksek irtifa, regl döngüsü gibi çok-günlük faktörler anomali modelini yanıltabilir.

*Önlem:* Çok-günlük kalıcı sapma eşiği (kısa olayları zaten filtreler); regl fazını bağlamsal sinyal olarak modele katmak; seyahat/irtifa gibi faktörleri LLM yorum katmanında "olası açıklama" olarak sunup kesin alarm vermemek. Eşik değerlerini LifeSnaps doğrulamasıyla kalibre etmek.

*R6 — Dar demografi:* WESAD (genç yetişkinler) ve LifeSnaps (çoğunlukla genç-orta yaş) demografisi dar olduğundan model yaşlı/kronik hastalıklı bireylere genellenemeyebilir. WESAD yalnızca 15 katılımcı içermekte olup bu, kişi-bazlı çapraz doğrulama fold sayısını sınırlandırmaktadır; ancak bu veri kümesi bilek tabanlı giyilebilir stres tespiti literatüründe en yaygın kullanılan akademik referans veri setlerinden biridir.

*Önlem:* Hedef kullanıcı 'genç-orta yaş, genel olarak sağlıklı yetişkinler' olarak açıkça tanımlanacaktır. WESAD yalnızca 15 katılımcı içerdiğinden elde edilen sonuçlar geniş popülasyona doğrudan genellenemeyebilir; bu riski azaltmak amacıyla kişi-bazlı veri bölme kullanılacak, LifeSnaps üzerinde ek doğrulamalar yapılacak ve sonuçlar yalnızca hedef kullanıcı kitlesi için yorumlanacaktır. Geniş demografi gelecek çalışma olarak belirtilecektir.

*R7 — LifeSnaps etiket belirsizliği:* Model 1'in LifeSnaps doğrulamasında kullanılan stres etiketi Fitbit'in kapalı algoritmasından gelir, gerçek insan etiketi değildir.

*Önlem:* LifeSnaps doğrulaması 'kesin doğruluk' değil 'referans/kıyas' olarak konumlandırılacak; kişi-bazlı bölme uygulanarak ek doğrulama yapılacaktır. Ana performans değerlendirmesi WESAD'ın etiketli verisiyle gerçekleştirilecektir. LifeSnaps, gerçek yaşam koşullarındaki davranış örüntülerini gözlemlemek ve anomali tespiti bileşenini destekleyici şekilde değerlendirmek amacıyla kullanılacaktır.

*R8 — Gerçek cihaz erişimi:* Ekipte gerçek bir akıllı saat/bileklik bulunmayabilir, bu da canlı testi sınırlar.

*Önlem:* Gerçek cihaz erişimi sınırlı kalırsa, WESAD ve LifeSnaps gerçek verisinden HealthKit/Health Connect'e zaman-damgalı örnek enjeksiyonu ile tüm pipeline demo edilebilir. Bu yöntem, uçtan uca sistemin işleyişini kanıtlamak için yeterlidir ve cihaz yokluğu projenin teslimini engellemez. Ekip üyelerinde kişisel cihaz varsa bonus olarak gerçek veri testi yapılacaktır.

*R9 — LLM hatalı/teşhis niteliğinde çıktı:* LLM, istenmeyen biçimde teşhis koyan veya yanlış bir sağlık ifadesi üretebilir.

*Önlem:* İstem (prompt) tasarımında katı kurallar (teşhis yok, normalleştirici dil, uzmana yönlendirme); çıktıyı şablona sınırlamak; sayısal sonuçları LLM'in "uydurmasına" izin vermeyip yapılandırılmış girdi olarak vermek; örnek çıktılar üzerinde uzman/kullanıcı değerlendirmesi (başarı kriteri #6 bununla bağlantılı).

*R10 — Yeterli Geçmiş Veri Oluşamaması:* Kişiselleştirilmiş baseline oluşturulabilmesi için kullanıcının belirli bir süre veri üretmesi gerekir; yeni veya düzensiz kullanıcılar için anomali tespiti başlangıç döneminde düşük performans gösterebilir.

*Önlem:* İlk kullanım döneminde yalnızca temel analizler sunulacak, baseline güven seviyesi kullanıcıya bildirilecek ve veri biriktikçe kişiselleştirme kademeli olarak artırılacaktır.

*R11 — Cihaz Arası Veri Değişkenliği:* HealthKit/Health Connect'ten gelen veriler cihaza göre önemli farklılıklar gösterir; Apple Watch ile Samsung Galaxy Watch aynı metriği farklı sensör ve algoritmalarla üretebilir. Bu durum model genellenebilirliğini etkileyebilir.

*Önlem:* Model eğitiminde cihaz türü bir değişken olarak ele alınacak; mümkünse birden fazla cihaz türüyle test edilecektir.

## 8. Beklenen Çıktılar

### 8.1 Çalışan Sistem / Prototip

Tek platformda (iOS veya Android) çalışan, kullanıcının izniyle HealthKit/Health Connect üzerinden veri çeken bir mobil uygulama prototipi. Prototip; kişisel temel çizgi oluşturma, dört analitik bileşenin (stres sınıflandırma, anomali/hastalık-olabilirlik tespiti, aktivite tanıma, uyku skoru) çalıştırılması, sonuçların görselleştirilmesi ve LLM aracılığıyla sade dilde açıklanması işlevlerini içerir. Gerçek cihaz erişimi sınırlıysa, akış halka açık veri setlerinden HealthKit/Health Connect'e zaman-damgalı örnek enjeksiyonuyla demo edilebilir.

### 8.2 Eğitilmiş Modeller ve Değerlendirme Sonuçları

Halka açık veri setleri üzerinde eğitilmiş iki denetimli model (XGBoost stres sınıflandırıcı, Random Forest aktivite tanıma) ve istatistiksel anomali tespiti bileşeni (RHR-Diff, HROS-AD) ile kural-tabanlı uyku skoru. Her bileşen için kişi-bazlı bölme yöntemiyle elde edilmiş doğruluk (Accuracy), F1-Score, Precision, Recall ve yanlış alarm oranı gibi performans metrikleri ile LifeSnaps üzerindeki gerçek-yaşam doğrulama bulguları raporlanacaktır.

### 8.3 Açık Kaynak Depo

Veri ön-işleme, öznitelik çıkarımı, model eğitimi/değerlendirmesi ve mobil uygulama kodunu içeren, belgelenmiş bir Git deposu (GitHub). Depo; kurulum talimatları, kullanılan veri setlerine erişim bağlantıları (WESAD, LifeSnaps, aktivite veri seti), yöntem açıklamaları, örnek veri akışları ve sonuçların yeniden üretilebilmesi için gerekli adımları içeren bir README ile birlikte sunulur.

### 8.4 Raporlar

Dönem boyunca üretilecek yazılı çıktılar beş rapordan oluşmaktadır: bu Proje Önerisi, Spesifikasyon Raporu, Analiz Raporu, PKE-Plan ve Yüksek Seviye Tasarım Raporu. Her rapor, ilgili aşamanın rubric gereksinimlerini karşılayacak biçimde hazırlanacaktır. Yüksek Seviye Tasarım Raporu; sistem mimarisi, yöntem, değerlendirme sonuçları ve sınırlamaları içeren dönemin kapsamlı teknik çıktısı niteliğindedir. Raporlarda sistemin güçlü yönleri, sınırlılıkları ve gelecekte yapılabilecek geliştirmeler ayrıca tartışılacaktır.

### 8.5 Sunum ve Demo

Projeyi tanıtan bir final sunumu ve uygulamanın gerçek/simüle veriyle uçtan uca çalıştığını gösteren canlı bir demo. Demo; veri çekiminden başlayıp model çıkarımı ve LLM açıklamasının kullanıcı arayüzünde gösterilmesine kadar olan tam akışı sergiler.

### 8.6 Mimari ve Tasarım Dokümantasyonu

Sistemin iki ana akışını (çevrimdışı eğitim hattı ve çalışma zamanı mimarisi) gösteren diyagramlar, veri kaynağı–model eşlemesi, sinyal kademeleri (çekirdek/bağlamsal/kapsam dışı) ve modüllerin sorumluluk dağılımını içeren tasarım dokümantasyonu.

## 9. Kaynaklar

Apple Inc. (t.y.). *HealthKit*. Apple Developer Documentation.

<https://developer.apple.com/documentation/healthkit>

Bevel. (2025). *Bevel Health — The connected health coach*. <https://www.bevel.health>

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.

<https://doi.org/10.1023/A:1010933404324>

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.

<https://doi.org/10.1145/2939672.2939785>

Cosentino, G., Liu, S., Zeng, A., Tu, Z., Kleinerman, A., Small, S., DeCamp, L., Dodd, D., Hao, J., & Xu, Y. (2024). *Towards a personal health large language model*. arXiv. <https://arxiv.org/abs/2406.06474>

DemandSage. (2026). *Smartwatch statistics: Global users & market share*.

<https://www.demandsage.com/smartwatch-statistics/>

Firstbeat Technologies. (2016). *Heart rate variability and stress*. <https://www.firstbeat.com/wp-content/uploads/2016/07/FSN-ARTICLE.pdf>

Fuller, D., Rahimipour Anaraki, J., Simango, B., Dorani, F., Bozorgi, A., Luan, H., & Basset, F. (2020). *Using machine learning methods to predict physical activity types with Apple Watch and Fitbit data using indirect calorimetry as the criterion* [Veri seti]. Harvard Dataverse. <https://doi.org/10.7910/DVN/ZS2Z2J>

Gabrielli, L., Brutti, A., Farella, E., & Milanesio, M. (2025). AI on the pulse: Real-time health anomaly detection with wearable and ambient intelligence. *arXiv*. <https://arxiv.org/abs/2508.03436>

Google. (t.y.). *Health Connect*. Android Developers. <https://developer.android.com/health-and-fitness/guides/health-connect>

- Li, X., Dunn, J., Salins, D., Zhou, G., Zhou, W., Schüssler-Fiorenza Rose, S. M., & Sonecha, R. (2024). A comparison of personalized and generalized approaches to emotion recognition using consumer wearable devices. *JMIR AI*, 3, e52171. <https://doi.org/10.2196/52171>
- Lin, J., Li, T., Kalani, K., Malak, P., Tran, M., & Dolatabadi, E. (2026). mcPHASES: A dataset of physiological, hormonal, and self-reported events and symptoms for menstrual health tracking with wearables. *Scientific Data*, 13, 305. <https://doi.org/10.1038/s41597-026-06805-3>
- Mishra, T., Wang, M., Metwally, A. A., Bogu, G. K., Brooks, A. W., Bahmani, A., Alavi, A., Celli, A., Higgs, E., Dagan-Rosenfeld, O., Fay, B., Kirkpatrick, S., Kellogg, R., Gibson, M., Wang, T., Hunting, E. M., Mamic, P., Ganz, A. B., Tran, B., ... Snyder, M. P. (2020). Pre-symptomatic detection of COVID-19 from smartwatch data. *Nature Biomedical Engineering*, 4, 1208–1220. <https://doi.org/10.1038/s41551-020-00640-6>
- Nkurikiyeyezu, K., Yokokubo, A., & Lopez, G. (2019). *The effect of person-specific biometrics in improving generic stress predictive models*. arXiv. <https://arxiv.org/abs/1910.01770>
- Schmidt, P., Reiss, A., Dürichen, R., Marberger, C., & Van Laerhoven, K. (2018). Introducing WESAD, a multimodal dataset for wearable stress and affect detection. *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, 400–408. <https://doi.org/10.1145/3242969.3242985>
- Welltory. (2025). *Welltory — Manage your stress levels*. <https://welltory.com>
- WHOOP. (2025). *WHOOP recovery: How it works, key metrics, and tips*. <https://www.whoop.com/us/en/thelocker/how-does-whoop-recovery-work-101>
- Yfantidou, S., Karagianni, C., Efstathiou, S., Vakali, A., Palotti, J., Giakatos, D. P., Marchioro, T., Kazlouski, A., Ferrari, E., & Girdzijauskas, Š. (2022). LifeSnaps, a 4-month multi-modal dataset capturing unobtrusive snapshots of our lives in the wild. *Scientific Data*, 9, 663. <https://doi.org/10.1038/s41597-022-01764-x>