

Kalp Hastalığı Tahmini

Halil Çevik*

Özet

Makine öğrenmesi, kalp hastalığı tahmininde önemli bir rol oynayan birçok çalışmanın yapıldığı bir alan haline gelmiştir. Makine öğrenmesi, kalp hastalığı riskini tahmin etmek için büyük miktarda veriyi analiz etme yeteneğine sahip karmaşık algoritmalar kullanır. Makine öğrenimiyle kalp hastalığı tahmini, erken teşhis ve tedavi için önemli bir araç olabilir. Bu tür modeller, doktorlara ve sağlık profesyonellerine hastaların risklerini belirlemeleri ve uygun önlemleri alabilmeleri konusunda yardımcı olabilir. Ancak, bu modellerin geliştirilmesi ve doğruluklarının artırılması için daha fazla araştırma ve doğrulama gerekmektedir.

1 Giriş

Koroner arter hastalığı (KAH), kalp krizi veya ani, ölüm riskini arttıran bir durumdur. KAH, kalp kasını besleyen kalp kaslarına oksijen ve besin sağlayan koroner damarların kolesterol plakları tarafından daralması veya tam tıkanması ile meydana gelir. KAH, dünyada erişkin yaşlarda en sık ölüm nedenidir ve ölümlerin yaklaşık yarısından sorumludur. Bu nedenle KAH'ın tanınması ve önlem alınması çok önemlidir.

KAH hastalığının önceden tanınması için öncelikle KAH için risk faktörlerinin iyi bilinmesi gerekir. En yaygın risk faktörleri yaşlılık, sigara, hipertansiyon (yüksek tansiyon), hiperlipidemi (yüksek kolesterol), obezite (şişmanlık), diyabet (şeker hastalığı) ve hareketsiz yaşam tarzıdır (sedanter yaşam). Ayrıca, ailede KAH öyküsü olanlar, erkekler ve kadınlara göre daha yüksek bir risk altındadır. Yaş, KAH geliştirme riskinde önemli bir faktördür. Yaşın ilerlemesi, arterlerin sertleşmesine ve daralmasına yol açarak KAH riskini artırır. Sigara içiciliği, KAH riskini artıran en önemli risk faktörlerinden biridir. Sigara içmek, kan damarlarının daralmasına ve plak oluşumuna yol açarak, koroner arterlerin tıkanmasına neden olabilir. Hipertansiyon, koroner arterlerin daralmasına ve tıkanmasına neden olarak KAH riskini artırır. Bu nedenle, hipertansiyonu olan kişilerin tansiyonlarını kontrol altında tutmaları önemlidir. Hiperlipidemi, KAH riskini artıran bir diğer önemli faktördür. LDL olarak da bilinen kötü kolesterol, koroner arterlerin tıkanmasına neden olabilir ve kolesterol plaklarının içeriğinde ana etkindir. Obezite, KAH riskini artıran bir diğer önemli faktördür. Vücut kitle indeksinin

*15080069, [Github Repo](#)

(VKİ) yüksek olması, yüksek kan basıncı, yüksek kolesterol ve diyabet riskini artırır. Diyabet, KAH riskini artıran bir diğer önemli faktördür. Yüksek kan şekeri seviyeleri, damarları ve kalp kasını etkileyerek, KAH riskini artırır. Sedanter yaşam, KAH riskini artıran bir diğer önemli faktördür. Egzersiz yapmamak, obezite, yüksek tansiyon ve yüksek kolesterol riskini artırarak KAH geliştirme riskini artırır.

1.1. Çalışmanın Amacı:

Framingham çalışmasında elde edilen ve github'dan elde edilen verisetini kullanarak 10 yıl içinde KAH gelişmesinde açıklayıcı olan risk faktörlerini tanımlamaktır. Verisetinde toplamda 16 değişkende 4,241 gözlem değeri olduğu için rassal örneklem ile alınacak 200 gözlemde çalışma yapılacaktır.

1.2. Literatür:

KAH'nın insan hayatında önemi oldukça yüksektir ve birçok risk faktörünün birlikteliği ilgili en bilinen çalışmalardan biri Framingham Çalışmasıdır (dawber1951epidemiological). ABD'de Farmingham şehrinde 1948'den beri devam eden ve üçüncü nesil katılımcılar ile devam eden çalışmada, ilk kez olarak bu risk faktörleri belirlenmiş ve tıbbi literatüre elde edilen kanıtlar sunulmuştur.

Framingham'da elde edilen bu verilerle geliştirilen Framingham Risk Puanlaması ile hastaların sahip olduğu risk faktörlerini kullanarak 10 yıl içinde kardiyovasküler hastalık gelişme olasılığını tespit edebilmek mümkün olmuştur (wilson1998prediction).

2. Veri:

Çalışmada kullanılan veriseti Framingham çalışma kapsamında elde edilen gözlemlerden sağlanmıştır. Çalışmada kullanılan verisetinde toplam 4240 gözlem ve 16 değişken bulunmaktadır. İlk olarak eksik verilere bakılınca toplam toplam 582 gözlemde eksik veri olduğu görülmüştür ve bu eksik gözlemler verisetinden çıkarılmıştır. Böylece çalışma için 3658 gözlem olan veriseti kullanılacaktır.

10 sene içinde kalp hastalığı geçirip geçirmeme durumuna göre kesikli veri türünde 1 bağımlı değişken (10 yılda kalp hastası olma (TenYearCHD) vardır. 15 bağımsız değişkenden 7'si kesikli (erkek cinsiyet(male), eğitim durumu(education), sigara(currentSmoker), tansiyon ilacı kullanımı (BPMeds), önceden inme öyküsü (prevalentStroke), hipertansiyon öyküsü (prevalentHyp) ve şeker hastalığı (diabetes)), 8'i ise sürekli (yaş (age), günlük içilen sigara adedi(cigsPerDay), kolesterol seviyesi (totChol), sistolik kan basıncı (sysBP), diastolik kan basıncı (diaBP), vücut kütle indeksi (BMI), nabız (heartRate) ve kan şekeri seviyesi (glucose)) veri türündendir.

Sürekli veriler ise incelendiğinde ortalama yaş 49.5 (SS 8.6), ortalama günlük içilen sigara adedi 9.0 (SS 11.9), ortalama kolesterol seviyesi 236.8 (SS 44.1), ortalama sistolik kan basıncı 132.4 (SS 22.1), ortalama diastolik kan basıncı 82.9 (SS 12.0), ortalama vücut kütle indeksi

25.8 (SS 4.1), ortalama nabız 75.7 (SS 12.0) ve ortalama kan şekeri seviyesi 81.9 (SS 23.9) olarak bulundu.

Tablo 1: Sürekli Verilerin Özet İstatistikleri

	Ortalama	Std Der	Min	Ortanca	Maks
age	49.551941	8.562029	32.00	49.00	70.0
BMI	25.782802	4.065601	15.54	25.38	56.8
cigsPerDay	9.025424	11.921591	0.00	0.00	70.0
diaBP	82.917031	11.974258	48.00	82.00	142.5
glucose	81.852925	23.904164	40.00	78.00	394.0
heartRate	75.730727	11.981525	44.00	75.00	143.0
sysBP	132.370558	22.086866	83.50	128.00	295.0
totChol	236.847731	44.097681	113.00	234.00	600.0

Kesikli türde olan 1 bağımlı değişken ve 8 bağımsız değişken incelenmiştir. Bağımsız değişken olan 10 yılda kalp hastası olan gözlem sayısı 557 (%15.0) olarak saptanmıştır. 1623 (%44,0) hastanın erkek, 1789 (%49,0) hastanın sigara içmeye devam ettiği, 111 (%3,0) hastanın hipertansiyon için ilaç kullandığı, 21 (%0,6) hastanın önceden inme geçirdiği, 1140 (%31,0) hastanın hipertansiyon hastası olduğu ve 99 (%2,7) hastanın ise şeker hastası olduğu görülmüştür.

Tablo 2: Kesikli Verilerin Özet İstatistikleri

Değişken	Sayı (Yüzde)
male	1,623 (44%)
education	NA
1	1,526 (42%)
2	1,101 (30%)
3	608 (17%)
4	423 (12%)
currentSmoker	1,789 (49%)
BPMeds	111 (3.0%)
prevalentStroke	21 (0.6%)
prevalentHyp	1,140 (31%)
diabetes	99 (2.7%)
TenYearCHD	557 (15%)

3. Yöntem ve Veri Analizi:

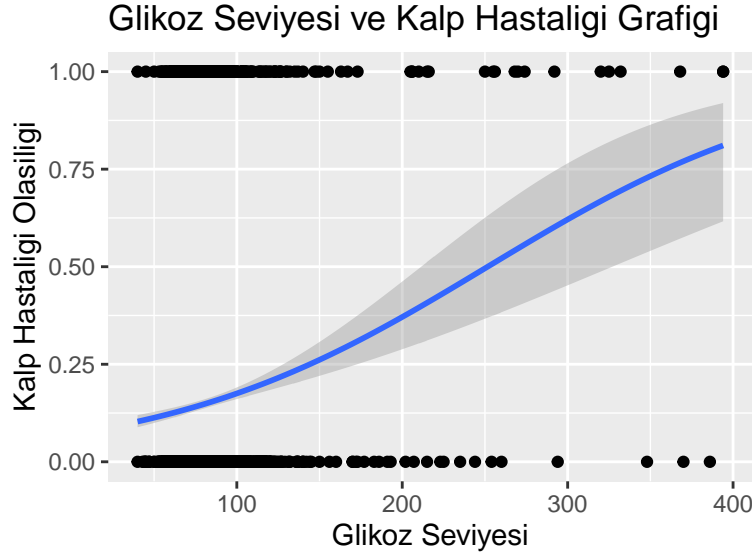
Framingham çalışmasından elde edilen verisetini kullanarak gözlemlerde 10 sene içinde kalp hastalığı gelişme olasılığını tahmin eden model oluşturmak için lojistik regresyon metodu kullanıldı. Lojistik regresyon modelinin kullanılmasının nedeni, bağımlı değişken kalp hastası olacak ve olmayacak diye ikiye ayrıldığı içindir.

Model incelendiğinde 10 yıl içinde kalp hastalığı gelişmesini açıklayan değişkenlerin erkek cinsiyet, yaş, günlük içilen sigara sayısı, kolesterol seviyesi, sistolik kan basıncı ve kan şekeri seviyesi olduğu görülmüştür. Logistik regresyon modeli şu şekilde bulunmuştur:

$$Pr(Kalpast = 1|X) = \frac{\exp(\beta_0 + \beta_1 erkek + \beta_2 ya + \beta_3 sigara + \beta_4 kolest + \beta_5 sistkan + \beta_6 glikoz)}{1 + \exp(\beta_0 + \beta_1 erkek + \beta_2 ya + \beta_3 sigara + \beta_4 kolest + \beta_5 sistkan + \beta_6 glikoz)}$$

Anlamli çıkan deęişkenler yorumlandığında ise erkek cinsiyette olmanın, kadın cinsiyete göre 1.707 kat daha fazla kalp hastalığı oluşmasına neden olduğu bulunmuştur. Yaşta 1 birim artış 1.0642 kat, günlük içilen sigara sayısında 1 birim artış 1.0181 kat, kolesterol seviyesinde 1 birim artış 1.002 kat, sistolik kan basıncında 1 birim artış 1.0156 kat ve kan şekeri seviyesinde 1 birim artış 1.0072 kat daha fazla kalp hastalığı olma olasılığını arttırdığı görülmüştür.

```
## `geom_smooth()` using formula = 'y ~ x'
```



4. Tartışma:

10 sene içinde kalp hastalığı geçirmeyi öngören risk faktörleri incelendiğinde hastanın erkek olması, yaşı, günde içtiği sigara sayısı, kan kolesterol seviyesi, sistolik kan basıncı ve kan şekeri seviyesinin anlamlı olduğu bu çalışmada görüldü. Bu risk faktörlerinin bilinmesi ve kişilerde iyi bilinmesi, kişinin kalp hastalığı riskleri açısından bilgilendirilmesi gerekmektedir. Ayrıca risk faktörlerinin kontrol altına alınması, sigarayı bırakması, kolesterol seviyesini azaltan diet ve tedavi uygulaması, kan basıncının kontrole alınması ve kan şekeri seviyesinin diet veya ilaç ile düşürülmesiyle 10 seneli kalp krizi geçirme olasılığı azaltılabilir.

5. Kaynakça:

1. Dawber, T. R., Meadors, G. F., & Moore Jr, F. E. (1951). Epidemiological approaches to heart disease: the Framingham Study. *American Journal of Public Health and the Nations Health*, 41(3), 279-286.
2. Wilson, P. W., D'Agostino, R. B., Levy, D., Belanger, A. M., Silbershatz, H., & Kannel, W. B. (1998). Prediction of coronary heart disease using risk factor categories. *Circulation*, 97(18), 1837-1847.
3. Akbulut, T., Oğuz, M., Şaylık, F., Şipal, A., & Ural, D. (2022). Comparison of Framingham, SCORE, PROCAM and TEKHARF risk scores for prediction of 10 year cardiovascular disease risk in patients with essential hypertension. *Kocaeli Medical Journal*, 11(2), 138–148. [doi:10.5505/ktd.2022.92603](https://doi.org/10.5505/ktd.2022.92603)
4. Topuz, A. N., & Bozdemir, N. (2022). Sağlıklı bir popülasyonda kardiyovasküler hastalık risk faktörleri bilgi düzeyi, Framingham skoru ve kardiyak belirteçlerin değerlendirilmesi. *Cukurova Medical Journal*, 47(3), 1086–1094. [doi:10.17826/cumj.1103219](https://doi.org/10.17826/cumj.1103219)