

LOGİT MODEL İLE DİYABET HASTALIĞI BELİRLEYİCİLERİ ÜZERİNE BİR ÇALIŞMA

A STUDY ON DIABETES DETECTORS WITH THE LOGIT MODEL

Özet

Dünya sağlık örgütü verilerine göre tüm yaş grupları baz alındığında küresel ölçekte 422 milyon insanın Diyabet hastası olduğu bilinmektedir. Ayrıca 18 yaş üstü küresel diyabet prevalansı ise 1980'de %4.7 iken 2014 yılında %8.5 olarak gerçekleşmiştir. Diyabet hastalığının küresel ekonomiye yıllık maliyeti 825 milyar dolardır. Ayrıca Dünya Sağlık Örgütü diyabet raporuna göre toplam diyabet harcamalarının artması beklenmektedir. Bu nedenle Diyabet hastalarının ön tanılarının konulması hem hastalık prevalansının azalması hem de politika geliştiriciler için önem arz etmektedir. Bu amaçla çalışmada Glikoz, Yaş, Vücut Kitle Endeksi, Kan Basıncı, Triseps Deri Kıvrım Kalınlığı, Soy Ağacı ve İnsülin değişkenleri Logit regresyon analizi kullanılarak istatistiksel anlamda Diyabet hastalığını tetikleyen etkiler olup olmadığı araştırılmıştır. Tahmin edilen Logit regresyon model sonuçlarına göre Glikoz, Vücut Kitle Endeksi, Yaş ve Soy Ağacı değişkenlerinin istatistiksel olarak anlamlı olduğu sonucuna ulaşılmıştır. Analiz sonucunda İnsülin değişkeninin istatistiksel olarak anlamlı olmaması deneklerin özellikle Tip 2 Diyabet hastası oldukları yönünde ciddi bir sinyal vermektedir. İnsülin değişkeninin istatistiksel olarak anlamlı olmaması ile birlikte Kan Basıncı ve Triseps Deri Kıvrım Kalınlığı değişkenlerinin anlamlı olmaması ise beklentiler yönünde gerçekleşmiştir. Tahmin edilen Logit regresyon modelinin referans alınabilir bir model olup olmadığını test etmek adına model geçerlilik testleri uygulanmıştır. İlk olarak yapılan tanımlama hatası testi sonucunda modelin geçerli bir model olduğu ve Diyabet hastası olma durumunu tetikleyen prediktörlerin yeterli olduğu sonucuna ulaşılmıştır. Açıklayıcı değişkenler arasında çoklu doğrusal bağlılık problemini araştırmak için Varyans Büyütme Faktörü ve Art İndeks kriterleri dikkate alınarak açıklayıcı değişkenler arasında çoklu doğrusal bağlılık probleminin olmadığı sonucuna ulaşılmıştır. Logit regresyon modelinin tahmin başarısı için sınıflandırma tablosu dikkate alınmıştır. Bu doğrultuda modelin genel tahmin başarısı %77,66 olarak gerçekleşmiştir. Tahmin edilen modelin istatistiksel olarak anlamlı ve tüm testleri başarı ile geçebilmiş olması ile modelin Diyabet hastası olma olasılığı olan kişilerin ön tanılarının konulabilmesi ve bu doğrultuda karar vericiler için referans alınabilir bir model olduğu sonucuna ulaşılmıştır.

Anahtar Kelimeler: Gestasyonel Diyabet, Tip 1 Diyabet, Tip 2 Diyabet, Logistic Regresyon

Abstract

According to the data of the World Health Organization, it is known that 422 million people have diabetes on a global scale when all age groups are taken as basis. In addition, the global prevalence of diabetes over the age of 18 was 4.7% in 1980 and 8.5% in 2014. The annual cost of diabetes to the global economy is 825 billion dollars. In addition, according to the World Health Organization diabetes report, total diabetes expenditures are expected to increase. For this reason, pre-diagnosis of diabetes patients is important for both the decrease in the prevalence of the disease and the policy developers. For this purpose, in this study, the variables of Glucose, Age, Body Mass Index, Blood Pressure, Triceps Skin Fold Thickness, Pedigree and Insulin were statistically investigated whether the effects of diabetes are triggering by using Logistic regression analysis. According to the estimated Logistic regression model results, it was concluded that the variables of Glucose, Body Mass Index, Age and Pedigree were statistically significant. As a result of the analysis, the insulin variable is not statistically significant, which gives a serious signal that the subjects are especially Type 2 Diabetes patients. Although the insulin variable was not statistically significant, the variables of Blood Pressure and Triceps Skin Fold Thickness were not statistically significant. Model validity tests were applied to test whether the estimated Logistic regression model is a reference model. As a result of the first identification error test, it was concluded that the model is a valid model and that the predictors that trigger the diabetes patient status are sufficient. In order to investigate the multiple linear dependence problem among the explanatory variables, it was concluded that there was no multiple linear connectedness problem among the explanatory variables by taking the Variance Inflation Factor and Cond Index indicators into consideration. The classification table was taken into account for the prediction success of the logistic regression model. Accordingly,

the overall prediction success of the model was 78.93%. With the predicted model being statistically significant and successfully passing all tests, it was concluded that the model was able to make preliminary diagnoses of people who are likely to have diabetes, and that it is a reference model for decision-makers.

Key Words: Gestational Diabetes, Type 1 Diabetes, Type 2 Diabetes, Logit Model

1. GİRİŞ

Karın bölgemizin en arka kısmında bulunan yaklaşık olarak 15 cm. uzunluğunda bir çok önemli görevi bulunan organımız Pankreas salgı bezimiz bulunmaktadır. Bu salgı bezinin yeterli miktarda insülin hormonu üretememesi ya da ürettiği insülin hormonunun yeterli ve/veya bir şekilde kullanılamaması ve genel olarak kandaki şeker seviyesinin normalin üzerinde olması durumunda ayrıca bununla birlikte normalde şeker içermemesi gereken idrarda şekere rastlanması durumuna Diyabet adı verilmektedir (TSHGM,2011: 20). Şeker hastalığının bugün insanlığa verdiği zarara bakıldığında Uluslararası Diyabet Federasyonundan elde edilen istatistiksel bilgiler çerçevesinde Diyabet Atlası 2017 yılı tahminlerinin Küresel ölçekte her 11 yetişkinden 1'nin diyabet hastası olduğu ve 2 diyabetli yetişkinden 1'ine teşhis konulmadığı şeklindedir (Türk Diyabet Cemiyeti, n.d.: 1).

Ayrıca şeker hastalığının vücut üzerindeki risklerine bakıldığında ise ciddi bir tablo ile karşılaşmaktadır. Ölüm riski, gerekli teşhis ve tedavi yöntemlerinin etkin bir şekilde sağlanamaması durumunda ise göz problemleri, böbrek problemleri ve sinir sistemi problemleri gibi insan vücudunun önemli organlarına kalıcı ve ciddi zararlar verebildiği görülmektedir (TEMD, 2013: 16). Uluslararası Sağlık Örgütlerinin açıklamaları incelendiğinde ise şeker hastalığının farklı sınıflara ayrıldığı görülmektedir. Özellikle Dünya Sağlık Örgütünün sınıflaması baz alındığında ise 3 farklı sınıfın olduğunu görmekteyiz. Bu sınıflar Tip 1. Diyabet, Tip 2. Diyabet ve Gestasyonel Diyabetir (The 6 Different Types of Diabetes, 2018: 1).

Tip 1. Diyabet incelendiğinde vücuttaki bağışıklık sisteminde oluşan bozukluktan dolayı vücut antikor üretimine başlamakta ve bu üretilen antikorlar insülin hücrelerini yok etmektedirler. Bu antikorların vermiş olduğu zarar ile birlikte vücutta insülin hücreleri üreten beta hücreleri tükenmekte ve böylece karbonhidrat metabolizması için insülin üretilmemektedir. Bu duruma ise Tip 1 Diyabet adı verilmektedir (Türkiye Diyabet Vakfı, 2020: 1). Tip 1 Diyabeti olan hastaların mutlaka dışarıdan insülin desteğinde bulunmaları gerekmektedir. Diyabet hastalarının yaklaşık %5-10'unun Tip 1 Diyabet hastası oldukları görülmektedir. Yapılan çalışmalar doğrultusunda Tip 1 Diyabet hastalarının % 7'nin 15 yaşından önce bu hastalığa sahip olduğu sonucuna ulaşılmıştır. (Türkiye Diyabet Vakfı, 2020: 1).

Tip 2 Diyabet pankreasın yeterli miktarda insülin salgılayamaması veya salgılanan insülinin yeterli derecede kullanılamaması nedeniyle kan şekerinin yükselmesi durumudur. Bu tür Diyabetli hastalar için erişkin diyabet hastaları denmektedir. Sebebi ise genellikle 40 yaşından sonra görülmesidir (Kalbini Dinlesen, 2020: 1).

Gestasyonel Diyabet ise hamilelik durumunda ortaya çıkan bir diyabet tipidir. Çeşitli derecelerde intoleransı olarak tanımlanmaktadır. Tip 2 Diyabet risk grubu içerisinde yer almaktadır. Doğum sonrası genellikle düzeldiği görülmüştür. Ancak hamile değilken dahi diyabet gelişimi için yüksek risk taşımaktadır (Oğuz, 2016: 1)

Dünya sağlık örgütü verilerine göre tüm yaş grupları baz alındığında toplam olarak 422 milyon diyabetli insanın olduğu bilinmektedir. Ayrıca 18 yaş üstü küresel diyabet prevalansı ise 1980'de %4.7 iken 2014 yılında %8.5 olarak gerçekleşmiştir (World Health Organization, 2020:

1). Yapılan uluslararası çalışmalar sonucunda diyabet hastalığının erken tanısının bizlere ne kadar elzem bir durum olduğunu göstermektedir. 21. yüzyılın vebası olarak adlandırılan Diyabet hastalığının erken dönemde tanısının konup önlenmesi, aynı zamanda ekonomik olarak bu alana yönlendirilecek harcamaların tasarruf edilmesi düşük ve orta gelirli ülkeler için çok daha önemli bir durumdur. Avustralya da bir çalışma doğrultusunda diyabet hastalığının erken teşhisinin başarısız olması durumunda Avustralya sağlık sistemine 700 milyon dolardan fazlaya mal olabileceği görülmüştür (Bilgin, 2021: 2). Bu yükün maliyeti küresel çapta 825 milyar dolardır (Öcal & Önsüz, 2018: 5). Diyabetli popülasyonun %75,4'ü düşük ve orta gelirli ülkelerde yaşamaktadır. Dolayısıyla düşük ve orta gelirli ülkelerin artan bu hastalık harcamalarını karşılayamaması riski ile karşı karşıyadırlar. Ayrıca Dünya Sağlık Örgütü 2016 diyabet raporuna göre toplam küresel diyabet harcamalarında artışın devam etmesi beklenmektedir. Yapılan laboratuvar ve literatür çalışmaları doğrultusunda belirlenen risk faktörleri ve sonradan ortaya çıkabilecek hastalıkların erken dönemlerde de var olabileceği göz önüne alındığında, olası ölümlerin ve organ kayıplarının önlenmesi, yaşam kalitesinin iyileştirilmesi ve sağlık sistemi üzerindeki ekonomik harcamaların azaltılması için hastalık önleme ve sağlığı geliştirme çalışmalarının yapılması bir gerekliliktir.

Bu çalışmanın amacı Diyabet hastalığını tetikleyen risk faktörlerinin Binary Logit model ile araştırılıp risk oranlarının tespit edilmesidir. Böylece Diyabet hastalığı erken tanısının konulabilmesi ve gerekli önlemlerin alınabilmesine yol gösterilmiş olacaktır.

2. LİTERATÜR ÖZETİ

Karataş ve Özkan (2020), Diabetes Mellitus (DM) hastalarındaki preferik polinöropatiler (DNP) yaşam kalitesini olumsuz etkileyen faktörler üzerine bir çalışma gerçekleştirmişlerdir. Çalışmada amaçlanan diyabetik preferik polinöropatiler açısından riskli hastaların önceden tanılarının konması ve bu tanıların atlanması önüne geçilmesidir. Çalışma 2018 Ağustos 2019 tarihleri arasında Karabük Üniversitesi Eğitim ve Araştırma Hastanesi FTR ve iç hastalıkları polikliniğine başvuran hastalar üzerine yapılmıştır. Toplam denek 399 Diabetes Mellitus hastasıdır. Bunların 173'ü erkek ve 226'sı kadındır. DNP tanıları LANSS ölçeği ile konulduktan sonra DNP pozitif ve negatif olanlar şeklinde ikiye ayrılmışlardır. DNP için yapılan logit regresyon analizi kapsamında bağımsız değişken olarak vücut kitle endeksi, HbA1c düzeyi, DM süreleri, oral antidiyabetik kullanımları ile DNP arasında istatistiksel anlamda bir ilişki bulunamamıştır. Ancak hastalık süresinin 61 aydan fazla olması, 60 yaş ve üstü ve obezitenin DNP üzerinde istatistiksel anlamda etkili olduğu sonucuna ulaşılmıştır.

Ertur ve diğ. (2020), Diyabetik preferik nöropati tanısı olan hastaların sıklık durumu ve bu hastalığa etki eden faktörler üzerine bir araştırma yapmışlardır. Çalışmada Tip 2 Diyabetli 336 denek kullanılmıştır. Anamnezleri, demografik ve antropometrik bilgiler ve deneklerin kullandıkları ilaçlar kayıt altına alınmıştır. Diyabetik nöropati tanısının konulabilmesi için MNSI skorlaması ve Semmes Weinstein monofilama testi uygulanmıştır. Çalışmada Hb1Ac, açlık kan şekeri ve vücut kitle endeksi sırasıyla %8.7, 182 ± 89 mg/dl ve 32 ± 5.8 kg/m² olarak bulunmuştur. DPN sıklığı %38.7 olarak saptanmıştır. Diyabetik nöropati olan grupla diyabetik nöropati olmayan grup arasında kadın cinsiyet, yaş, diyabet süresi, açlık kan şekeri, Hb1Ac, hipertansiyon öyküsü, aterosklerotik kardiyovasküler hastalık öyküsü, diyabetik retinopati ve insülin kullanımı istatistiksel anlamda farklı bulunmuştur. Çalışma sonucunda Tip 2 Diyabetli hastalarda yaklaşık

olarak %40 oranında diyabetik preferik nöropati saptanmıştır. Bu bulgular doğrultusunda rutin çalışma ve hasta tanı belirlemenin ne denli önemli olduğu sonucuna ulaşılmıştır.

Unal ve diğ. (2020), Tip 1 Diabetes Mellitus tanılı çocuk ve ergen hastalarda otoimmün tiroid hastalığı sıklığını ve prediktörlerle olan ilişkilerini incelemişlerdir. 722 Diabetes Mellitus hastası retrospektif olarak incelenmiştir. Hastaların yaş, cinsiyet, ortalama Hb1Ac, tiroid uyarıcı hormon, serbest tiroksin, antiroid peroksidaz ve anti-tiroglobülin düzeyi, vücut ağırlığı standart deviasyon skoru, boy SDS ve vücut kitle endeksi prediktör olarak kullanılmıştır. Analiz sonucunda Tip 1 Diyabet Mellitus hastalarda yaş ve diyabet süresinin arttıkça Hashimoto tiroid sıklığının arttığı sonucuna ulaşılmıştır. Özellikle bu durum kız cinsiyette ve pubertal dönemde daha sık görüldüğü gözlenmiştir. Tip 1 Diabetes Mellitus hastalarda çölyak varlığının Hashimoto tiroid riskini arttırdığı sonucuna ulaşılmıştır. Tiroid disfonksiyonunun metabolik kontrol, büyüme ve gelişme üzerindeki olumsuz etkilerinin önlenmesi açısından belli aralıklarla tiroid fonskiyon testleri ve otoantikörlerin izlenmesi gerektiği sonucuna ulaşılmıştır.

3. YÖNTEM

Nitel tercih modelleri ailesinden olan logit modeli özellikle iki sınıflı bağımlı değişkenler için kullanılan doğrusal olmayan bir modeldir. Logit model birikimli dağılım fonksiyonuna dayanmaktadır. Literatürde logit model yerine daha çok lojistik regresyon ifadesinin kullanıldığı görülmektedir. Logit modelin literatürde daha çok durum kontrolü, risk ölçüm ve tespit çalışmalarında kullanıldığı görülmektedir. Özellikle bağımlı değişkenin iki sınıflı olduğu durumlarda kullanılmaktadır. Örnek olarak elde edilecek tahmin sonucunun evet-hayır, var-yok, başarılı-başarısız, riskli-risksiz ve hasta-sağlıklı şeklindeki durumlarda logit modeli gibi ikili seçim modelleri kullanılır (Long, 1997: 116). Örneklerden de görüldüğü üzere logit modelin temel amacı iki sınıflı bağımlı değişkenlerin tahminin de kullanılan bir istatistiksel tekniktir. Ancak model içerisinde yer alacak açıklayıcı değişkenlerin sürekli veya süreksiz değişkenler olmasında herhangi bir kısıt bulunmamaktadır.

Logit modelin bizlere sunmuş olduğu bu avantajdan dolayı gözlemlerin gruplara ayrılmasında ve buradan yola çıkarak yeni değişkenlerin oluşturulabilmesinde bizlere olanak tanımaktadır. Logit modelin tahminin kolay olması ve esnek bir yapıya sahip olmasından dolayı literatürde en çok kullanılan nitel tercih modeli olduğu gözlenmektedir. Logit modelle amaçlanan bağımsız değişkenlerin bağımlı değişkene göre beklenen olasılık değerinin tahmin edilmesidir (Greene, 2016: 688).

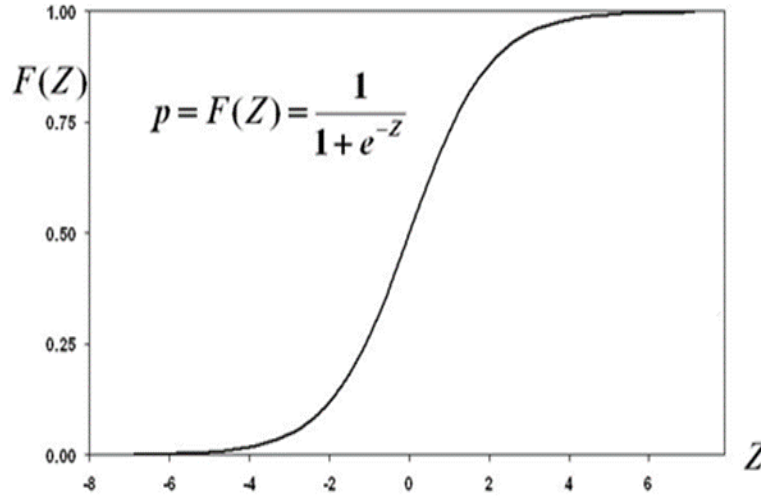
Logit model;

$$P_i = F(Z_i) = \beta\chi + U_i = \frac{1}{1+e^{-Z_i}} = \frac{1}{1+e^{-(\beta\chi)}} \quad (1)$$

şeklinde logistic dağılım fonksiyonundan türetilmektedir. Buradaki P_i değeri, bağımsız değişken X_i veri iken i . bireyin belirli bir seçim yapma olasılığını göstermektedir. Z_i değişkeni, $-\infty$ 'a yaklaştıkça olasılık değeri $P(Y_i=1)$ 'in 0'a yaklaştığını, buna karşılık Z_i değişkeni $+\infty$ 'a yaklaştıkça $P(Y_i=1)$ olasılık değeri 1'e yaklaşmaktadır. Bu da aslında P_i ile Z_i arasındaki ilişkinin doğrusal

olmadığını göstermektedir. Doğrusal olmama durumunun var olması sebebiyle buradaki parametreler en küçük kareler yöntemiyle tahmin edilememektedir. Buna alternatif olarak maksimum olabilirlik yöntemi tahmin tekniği olarak sıkça kullanılmaktadır (Wooldridge, 2019: 578). Maksimum olabilirlik yöntemiyle elde edilen tahminler küçük örneklerde sapmasız, etkin olmayan ve normal dağılmayan parametrelerdir. Büyük örneklerde ise tutarlı, asimptotik etkin ve normal tahminlerdir. Logit model birikimli olasılık dağılımından türetilmiş lojistik dağılım fonksiyonudur. Birikimli dağılım fonksiyonu Şekil 1'deki gibidir.

Şekil 1. Birikimli Dağılım Fonksiyonu



Modeldeki katsayıların doğrusallaştırılması için aşağıdaki yöntem uygulanmaktadır.

$$\frac{P_i}{1-P_i} = \frac{1+e^{z_i}}{1+e^{-z_i}} = e^{z_i} \quad (2)$$

ifadesi elde edilir. Buradan elde edilen oranın doğal logaritması alınarak lojistik fonksiyon doğrusal regresyon analizi kapsamında kullanılabilir.

$$L_i = \ln\left(\frac{P_i}{1-P_i}\right) = Z_i = \beta\chi + u_i \quad (3)$$

elde edilir. L_i doğrusaldır ve Logit ismini almaktadır. (Güriş & Çağlayan Akay, 2018: 671)

Logit modelin bazı özellikleri;

- P_i olasılığı 0 ile 1 arasında değer alırken Logit değeri $-\infty$ ile $+\infty$ arasında değer almaktadır.
- L_i doğrusaldır. Ancak olasılıklar doğrusal değildir.
- Logit modelin hata terimlerinin binom dağılımlı olmasından dolayı değişen varyansa sahiptir.

- Logitin artı olması durumunda fark oranı artmaktadır. Eksi olduğu durumda ise fark oranı düşmektedir.(Gujarati, 2016: 236)

4. VERİ SETİ

Çalışmada kullanılan veri seti ABD Ulusal Diyabet Sindirim ve Böbrek Hastalıkları Enstitüsünden elde edilmiştir. Çalışma kapsamında kullanılan veriler ABD Arizona Eyaletinde yaşayan kadınlar üzerinde diyabet hastalığı araştırması sonucu elde edilmiştir. Veri seti 2001 yılında derlenmiştir. Diyabet hastalığının ön tanısının konulabilmesi için yapılacak olan bu çalışmada veri seti oluşturulurken 768 gözleme ulaşılmıştır. Ancak İnsülin değişkeninin bir hayli gözlem eksikliği problemi barındırması sebebiyle çalışmada kullanılan gözlem sayısı 394 olarak gerçekleşmiştir. Çalışmada kullanılan değişkenlerin tanımları Tablo 1’de verildiği gibidir.

Tablo 1: Değişken Tanımları

Değişken	Tanımı
Glucose	Oral Glikoz Tolerans Testi
Blood Pressure	Diastolik Kan Basıncı
SkinThickness	Trisepts Deri Kıvrım Kalınlığı
BMI	Vücut Kitle Endeksi(VKİ)
DiabetesPedigreeFunction	Soyağacı
Age	Yaş
Insulin	İnsülin
Outcome/Diabetes Mellitus	Diyabet Hastası

4.1.Değişkenlerle İlgili Genel Bilgiler

Glikoz tolerans testi zaman içinde, genellikle iki saat olmak üzere, birden fazla kan örneğinin alınmasını içeren bir kan testidir. Diyabet hastalığı teşhisi için kullanılmaktadır. Kan testi sonucu ise üç farklı sınıfa ayrılmaktadır. Diyabet hastalığı için normal kabul edilen durum 2 saatlik glikoz seviyesinin 140 mg/dL’den az olması durumudur. Bozulmuş sonuçlar için iki saatlik glikoz seviyesinin 140 ila 200 mg/dL arasında olmasıdır. Anormal tanısall sonuçlar ise iki saatlik glikoz seviyesinin 200 mg/dL seviyesinin üzerinde olması durumudur.

Diastolik kan basıncı Tansiyon arteriyel sistolik ve diastolik olmak üzere iki bileşenden oluşur. Sistolik değer, kalp kasıldığında kalbden damarlara doğru atılan kanın damar duvarında yaptığı basınçtır. Diastolik değer ise kalp gevşediğinde hâlâ damar duvarında mevcut olan basınçtır. Normal bir kan basıncı 120 mmHg değerinin altında sistolik ve 80 mmHg değerinin altında diastoliktir. Yüksek kan basıncı ise 120-129 mmHg değerinin üzerinde sistolik ve 80 mmHg değerinin altında diastoliktir. Vücut kitle endeksi, bir diğer adıyla kilo boy indeksi vücudumuzdaki tahmini yağ oranını gösteren bir değerdir. İnsan kilosunun boyuna göre normal olup olmadığının anlaşılması için kullanılmaktadır.

Vücut kitle endeksine ulaşabilmek için vücut ağırlığı boy uzunluğunun karesine bölünmektedir. Bu şekilde vücut kitle endeksine ulaşılır. Aynı zamanda buradan yola çıkarak ulaşmak istenen vücut kitle endeksi ile boy uzunluğunun karesi çarpılarak ideal kilomuzu da

hesaplayabilme fırsatımız bulunmaktadır. Vücut kitle endeksimizin 18,5 değerinin altında olması zayıf olduğumuzu, 18,5 ile 24,9 arasında olması ise normal ve sağlıklı bir kiloda olduğumuzu göstermektedir. 25,0-29,9 aralığı fazla kilolu olma durumunu ve 30,0 üzeri ise obezite seviyesine ulaşıldığını göstermektedir. Deri kıvrım kalınlığı ölçümlerinin prensibi deri altı yağ deposunun büyüklüğünü belirlemektir.

Kadınlar için ideal trisept deri kıvrım kalınlığı 18,0 mm'dir. Soyağacı değişkeni ise üzerinde test yapılan kişinin soyunda kaç kişinin diyabet tanısı içerdiğini göstermektedir. Metabolik sendrom olarak bilinen insülin direnci, insan vücudunun kaslarında, yağında ve karaciğerde bulunan hücrelerin insüline doğru şekilde tepki vermediği ve bu nedenle enerji için kandaki glikozu kullanamadığı durumdur.

Pankreas, kanda artan glikoz seviyelerini telafi etmek ve hücrelerde oluşan direnci kırmak için sürekli daha fazla insülin üretir. Bu da zamanla kan şekeri seviyesinin yükselmesine neden olur. Ortalamada her üç insandan birisini etkilediği görülen insülin sendromu obezite, yüksek tansiyon, yüksek kolesterol ve Tip 2 Diyabet gibi tıbbi sorunlara yol açabilmektedir. Yaş değişkeni analize katılan bireylerin yaşlarını ve Diabetes Mellitus değişkeninin 1 değerini adlığını durumlarda kişinin diyabet hastası olduğunu ve 0 olduğu durumlarda ise kişinin sağlıklı olduğunu göstermektedir.

5. AMPİRİK SONUÇLAR

5.1. Diyabet Hastalığı için Bağımsız Değişken Dağılımları

Diyabet hastası olma durumunu tetikleyen ve yapılan literatür çalışmaları sonucunda diyabet hastalığına etki edebilecek insülin, yaş, VKİ, glikoz, kan basıncı, soy ağacı, hamilelik sayısı ve deri kıvrım kalınlığı değişkenlerinin bağımlı değişkenin diyabet hastası olma ve olmama durumlarına göre ortalama farklılık değerleri araştırılmıştır. Tablo 2 bu istatistikleri yansıtmaktadır.

Tablo 2: Çıktıya Göre Ortalama Farklılıkları

Glikoz		İnsülin		Soy Ağacı	
Çıktı		Çıktı		Çıktı	
0	111,669	0	68,792	0	0,493
1	144,962	1	100,335	1	0,608

Tablo 2 incelendiğinde tabloda bulunan tüm değişkenlerin diyabet hastası olma ve olmama çıktılarına göre ortalamalarının farklılık gösterdiği sonucuna ulaşılmıştır. Özellikle insülin değişkeninin ve hamilelik değişkeninin ortalamalarının neredeyse çıktıya göre 2 kat derecede bir farklılık gösterdiği görülmektedir. Tüm değişkenler dikkate alındığında ise glikoz, hamilelik sayısı, insülin ve soy ağacı değişkenlerinin değerlerinin arttığı durumlarda diyabet hastası olma riskinin yüksek derecede arttığı görülmektedir.

Tablo 3: Çıktıya Göre Ortalama Farklılıkları

Yaş		Deri Kıvrım Kalınlığı		VKİ		Kan Basıncı	
Çıktı		Çıktı		Çıktı		Çıktı	

0	28,685	0	27,265	0	19,854	0	69,414
1	35,931	1	33,000	1	8,431	1	73,973

Tablo 3 incelendiğinde ise Tablo 2'ye nispeten burada bulunan değişkenlerin ortalama farklılık değerlerinin daha düşük olduğunu görmekteyiz. Hatta vücut kitle endeksi değişkeninin diyabet hastası olma durumunda değerinin neredeyse yarı değerine ulaştığı görülmektedir. Bunun yanında yaş, deri kıvrım kalınlığı ve kan basıncı değişkenlerinin ise bireyin diyabet hastası olması durumlarında arttığı görülmektedir. Ancak bu artış değerlerinin istatistiki anlamda bir etkisinin olup olmadığını görebilmek adına istatistiksel analizlerin yapılması gerekmektedir. Diyabet hastası olma durumunu etkileyen faktörler ayrıntılı bir şekilde incelenip yorumlandıktan sonra bir sonraki aşamada ise logit modeli ile araştırılmıştır.

5.2. Diyabet Hastası Olma Olasılığının Logit Model İle Analizi

Diyabet hastası olma durumunu etkileyen belirleyicileri tespit etmek ve incelemek amacıyla, yapılacak bu analizde verilerin yatay kesit verileri olması ve genellikle yatay kesit verilerinde değişen varyans probleminin olması sebebiyle dirençli hatalar dikkate alınarak Logit modeli tahmin edilmiş ve marjinal etkiler hesaplanmıştır.

Tablo 4: Diyabet Hastalığı Belirleyicilerinin İncelenmesi: Logit Regresyon Sonuçları

Değişkenler	Katsayı	Marjinal Olasılıklar
İnsülin	-0,00081 (-0,53)	-0,000117 (-0,53)
Soy Ağacı	-0,49485* (1,65)	0,070939* (1,67)
Yaş	-0,04779*** (3,20)	0,006851*** (3,32)
VKİ	-0,00002*** (-2,63)	0,000002*** (-2,67)
Deri Kıvrım Kalınlığı	-0,02035 (1,41)	0,002918 (1,42)
Kan Basıncı	-0,00349 (0,30)	0,000501 (0,30)
Glikoz	-0,03854*** (6,48)	0,005526*** (8,32)
Gözlem Sayısı:	394	
Olabilirlik Oranı:	- 175,4193	
Wald Chi2(7):	92,62***	

Pseudo R²: 0,2979

Not: Robust standart hatalar parantez içinde verilmiştir. *p<0,01; **p<0,05; *p<0,1**

$$DHO_i = -0,00081Insulin_i -0,49485SoyAğacı_i -0,04779Yaşı_i -0,00002VKİ_i -0,02035DKKkalınlığı_i -0,00349KanBasıncı_i -0,03854Glikoz_i \quad (4)$$

Tablo 4’de Diyabet hastası olma olasılığını etkileyen faktörlerin Logit model sonuçları verilmiştir. Logistic model sonuçlarına göre, Soy Ağacı, Yaş, Vücut Kitle Endeksi ve Glikoz değişkenleri istatistiksel olarak anlamlı bulunmuştur. Literatür çalışmalarından da görüldüğü üzere anlamlı değişkenler baz alındığında daha çok Tip 2 Diyabet hastası olma olasılığını etkileyen faktörlerin istatistiksel olarak anlamlı olduğu görülmektedir. İnsülin değişkeninin ise daha çok Tip 1 Diyabet hastası olma durumunu etkileyen bir faktör olduğu literatür sonucu bilinmektedir. Diyabet hastaları sınıfları incelendiğinde ise Diyabet hastalarının %90-95 oranında Tip 2 Diyabet hastası oldukları bilinmektedir. Diyabet hastalarından Tip 1 Diyabet hastası olanların ise sadece %5-10 aralığında olduğu bilinmektedir (Türkiye Diyabet Vakfı, 2020: 1).

Elde edilen logit model sonuçları incelendiğinde de bu durumu doğrulayıcı sonuçları elde ettiğimiz görülmektedir. İnsülin değişkeni ile birlikte hastaların daha çok Tip 2 Diyabet hastası oldukları göz önüne alınırsa Kan Basıncı ve Deri Kıvrım kalınlığı değişkenlerinin istatistiksel olarak anlamsız olmaları durumu beklentiler yönünde gerçekleşmiştir (Öcal, 2019:141). Bu çerçevede elde bulunan istatistiksel olarak anlamlı prediktörlerle yeni bir model tahmin edilip gerekli analiz ve değerlendirmeler yapılacaktır.

Tablo 5: Diyabet Hastalığı Belirleyicilerinin İncelenmesi: Logit Model Sonuçları

Değişkenler	Katsayı	Marjinal Etkiler
Soy Ağacı	0,52053992* (1,71)	0,0751843* (1,74)
Yaş	0,0506079*** (3,57)	0,0073096*** (3,74)
VKİ	-0,0000255*** (-3,74)	0,00000369*** (-3,86)
Glikoz	0,0372625*** (7,42)	0,005382*** (10,61)
Gözlem Sayısı:	394	
Olabilirlik Oranı:	- 176,63253	
Wald Chi2(7):	91,14***	
Pseudo R ² :	0,2931	

Not: Robust standart hatalar parantez içinde verilmiştir. *p<0,01; **p<0,05; *p<0,1**

$$DHO_i = 0,52053992SoyAğacı_i +0,0506079Yaşı_i -0,0000255VKİ_i +0,0372625Glikoz_i \quad (5)$$

İstatistiksel olarak anlamlı olan değişkenlerin Diyabet hastası olma durumunu etkileme oranları ise şu şekildedir. %10 anlamlılık seviyesinde, diğer değişkenler sabit iken Soy Ağacındaki

%1'lik artış Diyabet hastası olma oranını %0,0751843 kadar arttırmaktadır. Diğer değişkenler sabit iken %1 anlamlılık seviyesinde Yaş değişkeni %1 arttığında Diyabet hastası olma oranı %0,0073096 artmaktadır. Vücut kitle endeksi değişkeni içinde aynı şekilde %1 anlamlılık seviyesinde %1'lik bir artış gerçekleştiğinde Diyabet hastası olma oranı %0,00000369 değerinde artmaktadır. Glikoz değişkeni içinde aynı şartlar geçerli olmak üzere Glikoz değişkeninde %1'lik artış Diyabet hastası olma oranını %0,005382 oranında arttırmaktadır.

Tablo 4 incelendiğinde Diyabet hastalığı için tahmin edilen Logit regresyon modeli, dahil edilen açıklayıcı değişkenler doğrultusunda modelin iyi sonuçlar verdiği görülmektedir. Açıklayıcı değişkenlerden Soy Ağacı, Yaş, VKİ ve Glikoz değişkenlerinin istatistiksel anlamda önemli risk faktörleri olduğu sonucuna ulaşılmıştır.

5.3. Diyabet Hastalığı Modeli İçin Tanı Testleri

Diyabet hastalığı risk faktörleri için kurulmuş olan bu modelin geçerli bir model olup olmadığı sonucuna ulaşabilmek adına tanısal testlerinin yapılması gerekmektedir. Bu çerçevede modelin geçerli bir model ve modelde kullanılan değişkenler açısından gerekli değişkene ihtiyaç olup olmadığını analiz edebilmek için tanımlama testi uygulanacaktır. Ayrıca değişkenler arasında çoklu doğrusal bağıllık problemi olup olmadığı test edilecektir. Son olarak da modelimizin tahmin başarı metrikleri değerlendirilerek yüzdesel anlamda ne kadar bir başarı ile tahmin gerçekleştirdiği analiz edilecektir.

Tablo 6: Tanımlama Hatası Testi

Outcome	Katsayı	Std. Err.	Z	P
_hat	0,9302696	0,1072928	8,67	0,000
_hatsq	-0,1034637	0,0613625	-1,69	0,092
_cons	0,1547598	0,1708129	0,91	0,365

Tanımlama hatası testi sonucunda tahmin değerimiz(_hat) istatistiksel olarak anlamlı iken karesel tahmin değeri (_hatsq) istatistiksel olarak anlamsızdır. Tahmin değişkeninin istatistiksel olarak anlamlı olması durumu modelin geçerli bir model olduğunu ifade ederken, karesel tahmin değerinin istatistiksel olarak anlamsız olması ise modele herhangi bir değişken ekleme veya etkileşim değişkenine ihtiyaç olmadığını göstermektedir. Sonuç olarak modelimizin açıklayıcı değişkenler bakımından geçerli bir model olduğu sonucuna ulaşılmıştır.

Tablo 6'da ise çoklu doğrusal bağıllık göstergeleri bulunmaktadır. Çoklu doğrusal bağıllığı araştırmak adına varyans büyütme faktörü ve art indeks kriterleri kullanılmıştır. Varyans büyütme faktörü doğrultusunda VIF değerinin 5 ile 10 aralığında olması durumunda orta derece ve 10'dan büyük bir değer aldığında ise yüksek derecede çoklu doğrusal bağıllık probleminin olduğunu bizlere göstermektedir. Logit model analizimizde kullanılan açıklayıcı değişkenler için hesaplanan VIF değerleri dikkate alındığında ortalama VIF değeri 1,09 olarak gerçekleşmiştir. Gerçekleşen bu VIF değerinin açıklayıcı değişkenler açısından bir çoklu doğrusal bağıllık probleminin olmadığını göstermektedir.

Cond Index göstergesi ise 10 ile 30 arasında yer alan değerleri orta dereceli ve 30 değerinden büyük değerleri ise yüksek dereceli çoklu doğrusal bağıllık olarak tanımlamaktadır.

Art indeks değerleri doğrultusunda da model de kullanılan açıklayıcı değişkenlerimizin çoklu doğrusal bağıllık sorunu içermediği sonucuna ulaşılmıştır.

Tablo 7: Çoklu Doğrusal Bağıllık Göstergeleri

Değişken	VIF	Cond Index	
Soy Ağacı	1,01	1	1,0000
Yaş	1,14	2	2,5725
VKI	1,03	3	3,5384
Glikoz	1,16	4	8,3195
Ortalama VIF	1,09	5	12,0061

Tablo 7’de ise tahmin başarısı sınıflandırma tablosu bulunmaktadır. Bu tabloya göre model diyabet hastası olmayanları %87,50 düzeyinde bir başarı ile tahmin ederken diyabet hastası olma durumunu ise %57,69 düzeyinde bir başarı ile tahminleyebilmektedir. Modelin genel tahmin başarısı ise %77,66 olarak gerçekleşmiştir. Modelin genel başarı performansı açısından bakıldığında modelin açıklayıcı değişkenler iyi bir uyum sergilediği görülmektedir. Ayrıca yapılan diğer testlerle de birlikte modelimizin geçerli bir model olduğu sonucuna ulaşılmıştır.

Tablo 8: Sınıflandırma Tablosu

		Tahmin Edilen		Doğru Sınıflandırma Yüzdesi
		Diyabet Hastası		
		0	1	
Gözlenen Hasta	0	231	75	87,50
Gözlenen Hasta	1	264	130	57,69
Genel Yüzde				77,66

6. SONUÇ ve DEĞERLENDİRME

Çalışmada ABD Ulusal Diyabet Sindirim ve Böbrek Hastalıkları Enstitüsünden alınmış 394 denek verisi kullanılarak Diyabet hastası olma durumunu etkileyen belirleyiciler Logistic regresyon modeli kullanılarak araştırılmıştır. Çalışma kapsamında Glikoz, Yaş, Kan Basıncı, Triseps Deri Kıvrım Kalınlığı, İnsülin ve Soy Ağacı değişkenleri prediktör olarak kullanılmıştır.

Çalışmada prediktör olarak kullanılan değişkenlerin Diyabet Hastası olma durumunu ne şekilde etkilediklerini görebilmek adına Diyabet hastası olan ve Diyabet hastası olmayan deneklerin bilgileri ayrıştırılmıştır. Bu kapsamda Diyabet hastalığı belirleyicilerinin ortalama değerlerinin hasta olanlar ile olmayanlar arasında anlamlı farklılıkların olduğu sonucuna ulaşılmıştır. Glikoz, İnsülin, Yaş ve Deri Kıvrım Kalınlığı değişkenlerinin diyabet hastası olma oranı ile pozitif yönde ve etkili bir ortalama farklılığı değeri olduğu sonucuna ulaşılmıştır. Soy

Ağacı ve Kan Basıncı değişkenlerinin Diyabet hastası olma olasılığı ile pozitif yönlü fakat düşük bir ortalama farklılık değeri olduğu görülmüştür. Son olarak VKİ değişkeninin ise Diyabet hastası olma olasılığı ile negatif yönlü ve ciddi bir ortalama farklılık değeri olduğu sonucuna ulaşılmıştır.

Çalışmanın ikinci aşamasında Logit model tahmini gerçekleştirilmiştir. Sonuçlara göre, Glikoz, Yaş, VKİ ve Soy Ağacı değişkenlerinin istatistiki anlamda Diyabet hastası olma olasılığı üzerinde etkili oldukları görülmüştür. İnsülin değişkeninin istatistiksel olarak anlamsız olması ise hastaların Tip 2 Diyabet hastası olabilecekleri yönünde ciddi bir sinyal vermektedir. Bu doğrultuda Kan Basıncı ve Deri Kıvrım Kalınlığı değişkenlerinin istatistiksel anlamda etkisiz olmaları beklentiler yönünde gerçekleşmiştir. İstatistiki olarak önemeli bir prediktör olduğu sonucuna ulaşılan değişkenlerin Diyabet hastası olma durumunu etkileme oranları ise şu şekildedir. Diğer değişkenler sabit iken Soy Ağacındaki %1'lik artış Diyabet hastası olma oranını %0,0751843 kadar arttırmaktadır. Diğer değişkenler sabit iken Yaş değişkeni %1 arttığında Diyabet hastası olma oranı %0,0073096 artmaktadır. Vücut kitle endeksi değişkeni içinde aynı şekilde %1 anlamlılık seviyesinde %1'lik bir artış gerçekleştiğinde Diyabet hastası olma oranı %0,00000369 değerinde artmaktadır. Glikoz değişkeni içinde aynı şartlar geçerli olmak üzere Glikoz değişkeninde %1'lik artış Diyabet hastası olma oranını %0,005382 oranında arttırmaktadır. Diyabet hastalığını etkileyen faktörler için tespit ettiğimiz bu etkilerin geçerli bir etki olup olmadıklarını araştırmak adına Logit model için tanısal testler uygulanmıştır. Sonuçlar doğrultusunda modelin geçerli bir model olduğu ve prediktörlerin Diyabet hastalığını açıklamada yeterli değişkenler olduğu sonucuna ulaşılmıştır. Böylelikle tahmin edilmiş bu model Diyabet hastası olanların ön tanıların konulabilmesi için referans alınabilir bir modeldir.

Yirmi birinci yüzyılın vebası olarak adlandırılan Diyabet hastalığı insanlık ve gelecek adına büyük bir tehlike arz etmektedir. Vücut yapısı üzerinde bıraktığı kalıcı hasarlar ve ölüm oranları hızla artmaktadır. Bununla birlikte özellikle düşük ve orta gelirli ülkeler için ekonomik anlamda ciddi bir tehdit olarak yer almaktadır. Bu çerçevede Diyabet hastalığına etki eden faktörler ve bunlarla birlikte hastalarının ön tanıların konulmaları bir zorunluluk arz etmektedir. Böylelikle hasta olma olasılığı olan bireyler için ön tanı konulup kişinin hayatını tehdit eden bu riskin etkilerini minimize etme durumu söz konusudur. Ayrıca küresel anlamda ciddi bir ekonomik yük oluşturan bu olgunun önüne geçilmiş olacaktır.

KAYNAKÇA

- Bilgin, G., (2021), Makine Öğrenmesi Algoritmaları Kullanarak Erken Dönemde Diyabet Hastalığı Riskinin Araştırılması, *Zeki Sistemler Teori ve Uygulamaları Dergisi*, 4 (1), 55-64.
- ERTUR, E., VURAL KESKİNLER, M., ÇAKIR, İ. B., ERBAKAN, A., OĞUZ, A., (2020), Tip 2 Diyabetli Hastalarda Diyabetik Periferik Nöropati Sıklığı, İlişkili Faktörler ve Farkındalık Durumunun Değerlendirilmesi, *Kocaeli Üniversitesi Sağlık Bilimleri Dergisi*, 6(3), 180–185.
- GÜRİŞ, S., ÇAĞLAYAN AKAY, E., (2018), *Ekonometri Temel Kavramlar*, İstanbul, Der Yayınları, Beşinci Basım, ISBN: 978-975-353-529-8.
- Gujarati, D., (2016), *Örneklerle Ekonometri*, Ankara: BB101 Yayınları, Birinci Basım, ISBN: 978-605-9802-08-06.
- Greene, H. W., (2016), *Ekonometrik Çözümleme*, Ankara, Palme Yayıncılık, Yedinci Basım, ISBN: 978-605-355-542-1
- Kalbini Dinlesen (2020, Haz 29), Kimler şeker hastalığı riski altındadır? [online],

<https://www.kalbinidinlesen.com/kalp-hastaliklari-ile-yasamak/seker-hastaligi/kimler-seker-hastaligi-riski-altindadir>, [Eriřim Tarihi: 15 Mart 2021].

- Karatař, H. G., Ertınmaz zkan, A., (2020), Diyabetik Polinropatiyi Predikte Eden Bağımsız Klinik ve Demografik Faktrler. *Turkish Journal of Diabetes and Obesity*, 1(370), 11–15.
- Long J. Scott., (1997), *Regression Models For Categorical And Limited Dependent Variables*, California, Sage Publications, Seventh Edition, ISBN: 0-8039-7374-8.
- Oğuz, A., (2016), Gestasyonel Diyabet, *Kahramanmarař St İmam niversitesi Tıp Fakltesi Dergisi*, 11(1), 26–29.
- cal, E., & nsz, M., (2018), Diyabet Hastalığının Ekonomik Yk, *Trk Dnyası Uygulamave Arařtırma Merkezi Halk Saėlıėı Dergisi*, 3(1), 24–31.
- Satman, G., Imamoėlu, G., Yılmaz, C., Akalın, S., Salman, S., (2018), Diabetes Mellitus ve Komplikasyonlarının Tanı, Tedavi ve Gzlem Kılavuzu, Ankara, TEMD Yayınları, Onuncu Basım, ISBN: 978-605-4011-32-2.
- TSHGM, (2011), *Trkiye Diyabet nleme ve Kontrol Programı Eylem Planı (2011- 2014)*. Ankara: Saėlık Bakanlığı Yayın No:816.
- cal, I., (2019), Obezite ve Tip-2 Diyabetin Tedavisinde Manyetik Alan ve Alternatif Yaklařımlar, *Arřiv Kaynak Tarama Dergisi*, 28(2), 140–155.
- The 6 Different Types of Diabetes, (2018), The Diabetic Journey[online], <https://thediabeticjourney.com/the-6-different-types-of-diabetes>, [Eriřim Tarihi: 16 Mart 2021].
- Trk Diyabet Cemiyeti (n.d.). Diyabet İstatistikleri[online], 7. Diyabet Atlası, <http://www.diabetcemiyeti.org/c/diyabet-istatistikleri-2>, [Eriřim Tarihi : 19 Mart].
- Trkiye Diyabet Vakfı (2020), Diyabet Hakkında Her řey[online], Diyabet Nedir ?, <https://www.turkdiab.org/diyabet-hakkinda-hersey.asp?lang=TR&id=46>, [Eriřim Tarihi: 17 Mart].
- Unal, E., Demiral, M., cal, M., Baysal, B., & Ozbek, M. N., (2020), Tip 1 Diyabetli Hastalarda Otoimmn Tiroid Hastalığı Sıklığı, *Gncel Pediatri*, 18(2), 251–262.
- Wooldridge, J., (2019), *Ekonometriye Giriř Modern yaklařım*, Ankara, Nobel Akademik Yayıncılık, Drdnc Basım, ISBN:978- 605-133-375-5.
- World Health Organization (2020, Haz 8). Diyabet Hakkında nemli Gerekler[online], řeker Hastalığı, <https://www.who.int/news-room/fact-sheets/detail/diabetes>, [Eriřim Tarihi: 16 Mart 2021].