



Bu çalışma "World Happiness Report" tarafından katılımcıların kendi yaşamlarına ilişkin derecelendirmelerine dayanan ulusal mutluluk makalelerini ve sıralamalarını içeren bir yayından alınan parametre ve değişkenler dahilinde katılımcıların mutluluklarının ve aynı zamanda yaşam kalitelerini çeşitli faktörlerle ilişkilendirilmesi amaçlanmıştır. Veriler "World Happiness Report" web sitesi üzerinden elde edilmiştir. Avrupa ve Orta Asya bölgesindeki verilerin elde ettiğimiz 31 ülke 496 gözlem arasındaki katılımcıların mutluluk derecelendirmelerinin gayri safi yurt içi hasıla, cömertlik, güven, yaşam tercihlerini yapma özgürlüğü, sosyal destek, beklenen sağlıklı yaşam beklentisi, mutluluk hali, gülme ve eğlenme gibi duygulanımlar, ülke içerisindeki görülme sıklığı, otokorelasyon testleri uygulanarak yaklaşımda bulunulmak istenmiştir. Yapılan çeşitli regresyon modellerine göre modelin anlamlı olduğunu görülmektedir. Bu veri seti için sabit etkiler modeli genel olarak anlamlıdır ve bağımsız değişkenler bağımlı değişken üzerinde istatistiksel olarak anlamlı etkilere sahiptir. Ayrıca, sabit etkili modelin ülkeler arasındaki farklılıkları açıklamak için oldukça başarılı olduğu söylenebilir.

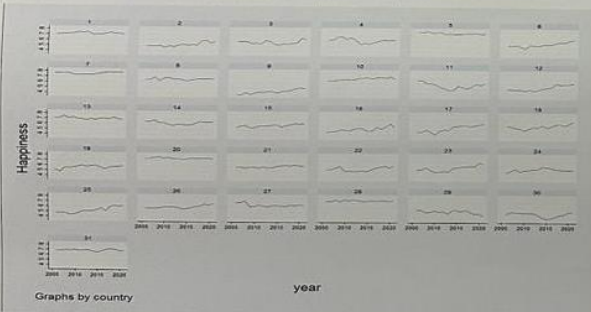
Mutluluk Analizi

Mutlu olmak, farklı mercilerdeki kişilerin görüşlerine göre gerek içimizde olup biten gerekse belirli koşullara bağlı olarak gelişen bir duygulanım olarak görülmektedir. Genel olarak, mutluluk normalde psikolojik açıdan bir ruh hali olarak düşünülür. Geçmişte yapılan çalışmaları bazılarında göre ise yaşamdaki kararları almadaki özgürlük ve sosyal destek değişkenlerinin ölçüm sonuçlarının yüksek çıkan ülkelerde yaşayan insanların mutluluk ölçümlerinde daha yüksek değerler gözlenmiştir (M. Mujiya Ulkhaq 'Arga Adayatama, 2020). World Happiness Report, mutluluğu sadece ruh halinden daha fazlası olarak görme eğilimindedir, çünkü araştırmacılar birçok sosyoekonomik faktörlerin, insanların belirli bir andaki duygulanımına ötesine geçtiğini savunmaktadır (Rodrigo Porto Carreiro Neves, 2021). Bu sonuçların bazılarını hakkında konuşulacak olursa, insanların yaşadıkları ülkelerde gözlemlenen durumların açıklanabilmesi açısından katılımcıların mutluluk ölçümleri, birtakım değişkenler ile tasvir edilebilir.



Şekil 1. Değişkenler Arası Korelasyon Isı Haritası

Bu korelasyonlar, değişkenler arasındaki ilişkileri gösterir, ancak nedensellik hakkında bir şey söylemez. Korelasyon, iki değişken arasındaki ilişkinin gücünü ve yönünü ölçer, ancak bu ilişkinin nedensel olup olmadığını belirlemez. Bu nedenle, korelasyon sonuçlarına dayanarak dikkatli bir şekilde çıkarımlarda bulunmak önerilir.



Şekil 2. Ülkelerin Mutluluk Grafikleri

Yukarıdaki grafiklerde ülkelerin 2006-2021 yılları arasındaki mutluluk puanlarının (10 üzerinden) dağılımı görülmektedir. Grafiğin üzerinden kısaca birkaç yorum yapacak olursak ülke numarası 5 (Danimarka) ve 13 (İrlanda) kelerinin mutluluk puan ortalamaları en yüksek iken ülke numarası 9 (Gürcistan) ve 30 (Ukrayna) ülkelerinin en düşük ortalamalarına sahiptir denilebilir. Ülkeler arasında mutluluk ölçümlerinin dalgalanmalarına bakıldığında 11 (unanistan) ve 27 (İspanya) numaralı ülkeler ilk yıllarda panel içerisinde bahsi geçen diğer ülkelerin ölçümlerine zaran daha yüksek oranda bir düşüş göstermiştir. Pozitif olarak sıralamalardan bahsedecek olursak 2015 ve sonrasında 30 (krayna) numaralı ülkenin mutluluk ölçüm farkı diğer ülkelere nazaran daha çoktur denilebilir.

Panel veri regresyon modeli sonuçlarına bakıldığında bağımlı değişken olan mutluluk değişkeninin diğer bağımsız değişkenlerle olan regresyon modeli aşağıda belirtildiği gibidir:

$$it) = -5.1118 + 0.6518 \times GDP(it) + 1.4578 \times SSP(it) + 0.5709 \times CNG(it) + 1.8196 \times PE(it) - 560 \times NE(it) - 0.0457 \times PC(it) + 0.1017 \times G(it) + 0.02844 \times HLEB(it) + 0.7293 \times FMLC(it) + u + e(it)$$

R² modelin açıklama gücünü gösterir. Bu açıdan, within R² 0.5849, between R² 0.8573, ve overall R² 0.8154 rak hesaplanmıştır. Böylece bu model hakkında konuşulduğu zaman bağımsız değişkenlerin bağımlı değişkeni gayet iyi şekilde açıkladığını söylenebilir.

KAYNAKÇA

[1] Rodrigo Porto Carreiro Neves (2021), *Representing Happiness: A comparative study of media coverage of the World Happiness Report* Chapter (pp.20)
[2] Jan-Emmanuel De Neve, & George Ward (2017). *Happiness at work to the World Happiness Report 2017* pp. (146-8)
[3] Ulkhaq, M. M., & Adayatama, A. (2020). Clusterine countries accordin to the World Happiness Report 2017

```

*** F test that all u_i = 0: F(30, 496) = 21.77      Prob > F = 0.0000
*** Wald chi2(10) = 301.94      Prob > chi2 = 0.0000
*** One-step results
***
R                Coef.    Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
R                -4.89961    0.541144      -9.06    0.000    -5.98018    -3.81904
GDP              -4833916    2286070      -21.16    0.000    -9414171    -1412751
SSP              1718237    4001889       0.43    0.668    -6302205    3911140
HLEB             912465    2021273       0.45    0.650    -3320293    1405162
FMLC            2807658    2371408       1.18    0.237    -7509254    1904828
G                6280959    1448114       4.34    0.000    3345978    9215980
PC              -110434    298395      -0.37    0.711    -6983774    6575006
PE              1407849    3766719       0.37    0.709    -6066881    3251082
NE              -1983379    3960004      -0.50    0.616    -8789522    4802144
CNG             3476174    1875262       1.85    0.066    709592    6246416
CNG              -4.13716    1.877451      -2.18    0.033    -7.85666    -0.417653

```

Tablo 1. Sabit Etkili Grup İç Panel Veri Regresyon Modeli Analizi

Çıktımızı yorumlayacak olursak mutluluk ve diğer bağımsız değişkenlerimiz arasında negatif yönlü orta dereceden biraz altında bir korelasyon gözlenmektedir denilebilir. (-%30). (Tablo 1.)

Modelin istatistikleri, R² within değeri 0.5883 olarak bulunmuştur. Bu durum bağımsız değişkenlerin bağımlı değişken üzerindeki varyansın yaklaşık %58.83'ünü karşıladığı anlamına gelmektedir. R² between değerimiz ise 0.8311 olarak bulunmuştur. Bu sonuçlar altında ülkeler arasındaki değişkenliğin büyük bir kısmını açıklanabildiğini söyleyebiliriz. R² overall sonucu ise 0.7927 olarak bulunmuştur. R² overall ifadesi bize bütün veri setindeki toplam değişkenliğin açıklanma oranını ifade eder. Bu değer toplam varyansın %79.27'sinin açıklanabildiğini bize göstermektedir. Varyansın %70.654395'i birimler arası farklılıklardan kaynaklanmaktadır.

```

*** F test that all u_i = 0: F(30, 496) = 21.77      Prob > F = 0.0000
*** Wald chi2(10) = 301.94      Prob > chi2 = 0.0000
*** One-step results
***
R                Coef.    Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
R                -4.89961    0.541144      -9.06    0.000    -5.98018    -3.81904
GDP              -4833916    2286070      -21.16    0.000    -9414171    -1412751
SSP              1718237    4001889       0.43    0.668    -6302205    3911140
HLEB             912465    2021273       0.45    0.650    -3320293    1405162
FMLC            2807658    2371408       1.18    0.237    -7509254    1904828
G                6280959    1448114       4.34    0.000    3345978    9215980
PC              -110434    298395      -0.37    0.711    -6983774    6575006
PE              1407849    3766719       0.37    0.709    -6066881    3251082
NE              -1983379    3960004      -0.50    0.616    -8789522    4802144
CNG             3476174    1875262       1.85    0.066    709592    6246416
CNG              -4.13716    1.877451      -2.18    0.033    -7.85666    -0.417653

```

Tablo 2. Arellano Bond Dinamik Panel Veri Tahmini

Dinamik panel veri analizi, panel veri setlerindeki zaman bağlamında değişken bağımsız değişkenlerin etkilerini modellemek üzere kullanılmaya başlanmıştır. Aynı zamanda en az bir bağımsız değişkenin zaman içinde değişen bir etkiye sahip olduğunu göstermektedir. Bu testin p-değeri çok küçük olduğu için, en az bir bağımsız değişkenin zaman içinde değişen bir etkiye sahip olduğunu sonuca varılabilir.

SONUÇLAR

Elde edilen bulgulara göre mutluluk ölçümlerinin en çok görüldüğü ülkelerin gelir seviyelerinin mutluluk ölçümleri kendilerine nazaran daha düşük olan ülkelere gelir seviyelerine oranla daha yüksek olduğu gözlemlenmiştir. Yaşam tercihlerini yapma özgürlüğü hususundaki ölçümler de mutluluk ile ilgili fikir edinebilmemiz konusunda bize doğru orantılı sonuçlar vermiştir. Yolsuzluğun görülme sıklığı ile ilgili ülkelerde yaşayan insanların mutluluk ölçümlerinin arasında bir ters orantı olduğu gözlemlenmiştir. Bağımsız değişkenlerin kendi aralarındaki ilişkilerine bakıldığında beklenen sağlıklı yaşam beklentisi ile gayri safi yurt içi hasıla değişkenleri arasında kuvvetli bir ilişki olduğu gözlemlenmektedir. Yolsuzluğun ülke içerisindeki görülme sıklığı değişkeni ile yaşam tercihlerini yapma özgürlüğü değişkeni arasında ise negatif yönlü kuvvetli bir ilişki söz konusudur. Aynı zamanda yaşam tercihlerini yapabileme özgürlüğü ile mutluluk, gülme ve eğlenme gibi duygulanımların ölçümleri değişkenleri arasında kuvvetli pozitif bir ilişki görülmektedir. Katılımcıların cevaplarına göre mutlulukları ile beklenen sağlıklı yaşam beklentisi, yaşamdaki tercihlerini yapabileme özgürlüğü, gayri safi yurt içi hasıla, mutluluk hali, gülme ve eğlenme gibi duygulanımların değişkeni, cömertlik ve güven değişkenleri ile orta seviye ve üzeri düzeyde pozitif bir ilişki gözlemlenmektedir. Ülke içerisinde görülme sıklığı ile endişe, öfke ve üzüntü gibi duygulanımların değişkeni ile mutluluk değişkeni arasında ise negatif yönlü orta düzeyde bir ilişki gözlemlenmektedir. Panel veri regresyon modeli analizlerine göre de bağımsız değişkenlerin bağımlı değişkeni iyi bir seviyede açıklayabildiği söylenebilir.

Dinamik panel veri tahminlerinin sonuçlarına göre, bu bir önceki dönemdeki bağımlı değişkenin, mevcut dönemdeki bağımlı değişkeni anlamlı bir şekilde etkilediğini gösterir. Aynı zamanda en az bir bağımsız değişkenin zaman içinde değişen bir etkiye sahip olduğunu göstermektedir. Bu testin p-değeri çok küçük olduğu için, en az bir bağımsız değişkenin zaman içinde değişen bir etkiye sahip olduğunu sonuca varılabilir.

Tüm bunlara ek olarak, mutluluk analizindeki bu çalışmanın geliştirilebilmesi, ilerleyen zamanlardaki yapılacak çalışmalara destek olabilmesi için modelimizin karmaşıklarıyla eğitim düzeyleri, refah seviyelerinin ölçümü, istihdam durumları gibi değişkenler araştırmanın içine dahil edilebilir. Yapılan geçmiş çalışmalarda aynı zamanda makroekonomik düzeyde, ulusal işsizlik seviyeleri ile negatif korelasyon göstermektedir (Jan-Emmanuel De Neve, George Ward, 2017).



FAKÜLTESİ

İSTATİSTİK BÖLÜMÜ

TEMEL BİLEŞEN ANALİZİ VE KÜMELEME ANALİZİ KULLANARAK NBA BASKETBOLUNUN DEĞİŞİMİNİN İNCELENMESİ

Bora Kağan MUMCU 19023079

Danışman: Dr. Öğr. Üyesi Erhan ÇENE

ÖZET

Bu çalışmada, NBA'de 1979-2017 dönemi arasında oyuncuların pozisyonlarına göre yapılan temel bileşen analizi ve kümeleme analizi, oyuncuların ve oyunun evrimine ışık tutmuştur. Temel bileşen analizi, veri setindeki ilişkileri açıklamak üzere kullanılmıştır. Elde edilen bileşenler, oyuncuların istatistiksel özelliklerini temsil etmiş ve bu özelliklerin zaman içinde nasıl değiştiğini göstermiştir. Kümeleme analizi ise, oyuncuların benzer özelliklere sahip gruplara ayırma için kullanılmış ve farklı dönemlerde hangi pozisyonlarda oynayan oyuncuların benzer özellikler sergilediğini ortaya koymuştur.

PCA (TEMEL BİLEŞEN ANALİZİ)

PCA (Principal Component Analysis - Temel Bileşen Analizi), çok boyutlu veri setlerini daha az boyutta bir yapıya dönüştürmek için kullanılan istatistiksel bir yöntemdir. Bu yöntem, veri setindeki değişkenler arasındaki ilişkileri anlamak ve veri setindeki değişkenliği azaltmak amacıyla kullanılır. Pearson'ın 1901 tarihli çalışması, PCA'nın değişkenler arasındaki ilişkiyi anlamak için kullanılan temel prensiplerini ortaya koymuştur.

PCA'nın temel amacı, veri setindeki değişkenler arasındaki korelasyonu azaltarak, veri setini daha az sayıda yeni değişken ya da bileşen ile ifade etmektir. Bu bileşenler, orijinal değişkenlerin bir kombinasyonu olarak elde edilir ve birbirleriyle lineer olarak ilişkilidirler.

PCA'nın çalışma prensibi, veri setindeki varyansı maksimize etmeye dayanır. Varyansın maksimum olduğu yönlere, en önemli bileşenler olarak kabul edilir. Bu bileşenler, veri setindeki değişkenliğin büyük bir kısmını temsil eder.

K-MEANS CLUSTERING (KÜMELEME ANALİZİ)

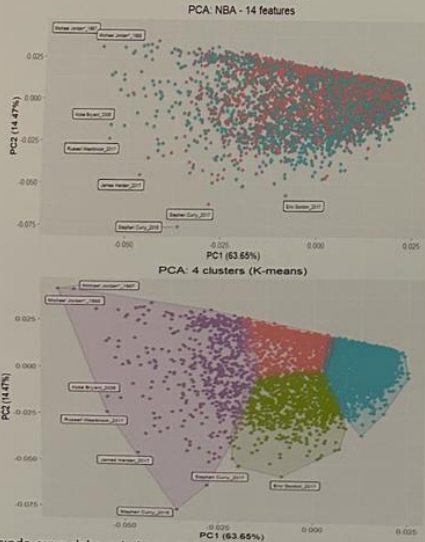
K-means kümeleme analizi, veri noktalarını belirli özelliklere göre gruplamak için kullanılan bir makine öğrenimi ve veri madenciliği tekniğidir. Bu yöntem, veri noktalarını verilen bir sayıda küme veya grup olarak böler. Bu gruplar, benzerlik ölçütlerine dayanarak oluşturulur.

Algoritma genellikle şu adımları takip eder:

1. Küme Merkezi Belirleme
2. Noktaları Kümelere Atama
3. Yeni Küme Merkezleri Hesaplama
4. Yeniden Atama ve Merkez Güncelleme

Bu süreç, kümeleme işleminin verilen bir kriteri optimize ederek (genellikle küme içi varyansı minimize ederek) tamamlanır.

UYGULAMA



1980'lerin ortasında oyun daha çok iki sayılı isabetler üzerinden dönerken bu durum, 2010'ların sonuna doğru üç sayılı isabetlere kaymıştır.

Mor renkte belirtilen küme süperstarlar (4), yeşil küme all-starlar (2), kırmızı küme rol oyuncular (1), mavi küme ise bench oyuncuları (3) olarak belirlenmiştir.

| Küme | Süre | Sayı | Ribaunt | Asist | Blok | Top Çalma | YOGD* | Verimlilik |
|----------------|------|------|---------|-------|-------|-----------|--------|------------|
| Süperstar | 36.5 | 21.6 | 4.38 | 5.82 | 0.376 | 1.57 | 3.24 | 19.7 |
| All-Star | 30.1 | 12.9 | 3.12 | 3.75 | 0.225 | 1.06 | 1.33 | 14.8 |
| Rol Oyuncusu | 29.2 | 12.2 | 3.06 | 4.89 | 0.225 | 1.31 | 0.961 | 15.0 |
| Bench Oyuncusu | 17.6 | 6.0 | 1.78 | 2.21 | 0.127 | 0.666 | -0.103 | 11.4 |

(*) YOGD: Yedek Oyuncuya Göre Değer, Bir oyuncunun o pozisyonda olan bench oyuncusuna kıyasla takımına ne kadar katkıda bulunduğunu gösteren bir istatistiktir.



Oyunun forvet pozisyonu tarafında iki sayılı ve üç sayılı oyun planı yıllara göre dengeli bir konumda. Mor renkte belirtilen küme süperstarlar (4), yeşil küme üç sayılı uzmanlar (2), kırmızı küme rol oyuncuları (1), mavi küme ise bench oyuncuları (3) olarak belirlenmiştir.



Oyunun pivot pozisyonu tarafında iki sayılı oyun planı daha ağır basmaktadır. Mor renkte belirtilen küme ofansif starlar (4), mavi küme defansif starlar (3), yeşil küme rol oyuncuları (1) olarak belirlenmiştir.

KAYNAKÇA

- Jolliffe, I. T. (2002). Principal Component Analysis. Springer Series in Statistics.
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. Pattern Recognition Letters, 31(8), 651-666.
- da Silva, M. D. M., de Souza, J. A. D., ve Teixeira, A. C. P. (2015). Clustering basketball players using principal component analysis and k-means. International Journal of Data Science and Analytics, 2(1), 12.



ÖZET

Kişilerin günlük hayatta birden fazla parametreye bağlı olabilecek psikolojik iyi oluşlarının, aktif olarak egzersiz yapıp yapmaması açısından ele alarak bu bağlamda anlamlı bir fark olup olmadığı araştırılmaktadır. Bu çalışmada, psikolojik iyi oluş ölçeği kullanılması yanı sıra cevapların doğruluğu ne kadar yansıttığı tamamen kişilerin cevaplarının gerçeği yansıttığı varsayılarak kabul edilmiştir. Bunların yanı sıra, psikolojik iyi oluş subjektif bir olgu olduğundan ve örneklemimiz herhangi bir kısıtlama ile seçilmediğinden, sonuçlarımızda olabildiğince çeşitli ve farklı çıkabilmektedir. Belirtildiği gibi, psikolojik iyi oluş subjektif bir olgu olsa dahi hayatımızdaki birçok subjektivite de olduğu gibi ortak bir paydada buluşabilmemiz mümkündür. Bu açıdan yapılan çalışmada kişilerin psikolojik iyi oluşunun düzenli egzersiz yapma açısından ortak bir payda oluşturup oluşturmadığını araştırılmaktadır.

Psikoloji ve Psikolojik İyi Oluş Kavramı

Psikoloji tanım gereği insanların içgüdüsel davranışlarını ve zihni düşüncelerini araştırın bu konular hakkında veri toplayan bilim dalıdır. Psikolojik iyi oluş kişilerin kendi hayatlarında var olan durumları yorumlamaları, bu yönde verdiği tepkiler ve genel olarak hissiyatlarını içeren bir ölçek düzeyindedir. Kişilerin kendileri hakkındaki düşünceleri, başkalarının kişinin kendisi hakkındaki düşünceleri, kişilerin kendi hayatına ve hayatlarındaki durumlara bakış açısını gösteren bir olgudur.

Psikolojik İyi Oluş Ölçeği

Bu çalışmada kullanılmış olduğumuz ölçek Ed Diener ve arkadaşlarının gerçekleştirmiş olduğu psikolojik iyi oluş ölçeğidir. Diener'a göre, psikolojik iyi oluş, kişinin yaşamının belirli bir döneminde duygusal ve bilişsel olarak iyi hissetmesi durumunu ifade eder. Diener ve meslektaşları, psikolojik iyi oluşun ölçümü için çeşitli anketler ve ölçekler geliştirmişlerdir. Bu ölçekler, insanların kendi yaşamlarının kalitesini değerlendirebilmelerine ve araştırmacıların bu konuda veri toplamalarına yardımcı olmuştur. Sekiz sorudan oluşan testimizde cevap şıkları en düşük 1 "Kesinlikle katılmıyorum" ve en yüksek 7 "Kesinlikle katılıyorum" şeklindedir. Teste tabii olan kişi 7 puana ne kadar yakın ise psikolojisinin o kadar iyi olduğu varsayılmaktadır. Bu sayede insanların psikolojik iyi oluşları ölçilebilir hale gelir.

Egzersiz Nedir ?

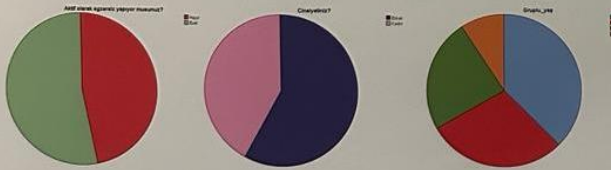
Egzersiz, planlı ve düzenli olarak yapılan bedensel aktivitelerdir. Bu faaliyetler, genellikle vücut sağlığını korumak, fiziksel formu artırmak, kasları güçlendirmek, esnekliği artırmak veya genel olarak sağlıklı bir yaşam tarzını desteklemek amacıyla gerçekleştirilir. Yürüyüş, koşu, yüzme, bisiklete binme, ağırlık antrenmanları, egzersiz salonlarında yapılan çalışmalar, yoga, pilates gibi çeşitli aktiviteler, egzersiz kapsamına girer. Bu aktivitelerin düzenli olarak yapılması, vücut kondisyonunu artırabilir, kalp sağlığını destekleyebilir, stresi azaltabilir, kilo kontrolüne yardımcı olabilir ve genel yaşam kalitesini yükseltebilir.

Egzersiz ve Psikolojik İyi Oluş Kavramı

Fiziksel aktivitenin psikolojik ve fizyolojik sağlığa birçok olumlu etkisi vardır. Araştırmalar, spor ve egzersizin kişilerin psikolojik sağlığını geliştirdiğini, stresi azalttığını ve olumlu duygularını artırdığını göstermektedir. Egzersizin kalp damar sağlığı, azami oksijen tüketimi ve benzeri fizyolojik yararlarının yanında kontrol algısında artış, öz yeterlilik ve benzeri psikolojik konularda da olumlu etkisi vardır.

Araştırma Verileri

Bu araştırmada kullanılan veriler anket yöntemi ile toplanılmıştır. Totalde 222 katılımcımız bulunmaktadır. Bu katılımcıların demografik bilgileri ve dağılımları aşağıdaki görseller ile açıklanmaktadır. Araştırmanın verileri toplanırken herhangi bir sınırlama konulmadan kişilerin verdiği cevaplar doğru söyledikleri varsayılarak gerekli test ve işlemler yapılmıştır. Bunların sonuçlarında doğan çıktılar üzerinden bir sonuca varılarak araştırma tamama erdirilmiştir.



Yukarıdaki pasta grafiklerde görüleceği üzere katılımcıların %46,8'i egzersiz yapmamakta ve %53,2'si egzersiz yapmaktadır. Bunun yanı sıra katılımcıların %41,9'u kadın ve %58,1' erkektir. Bunların yanı sıra katılımcılarımızın yaş dağılımı %37,4 ile 18-23, %29,5 ile 29-39, %23,5 ile 40-50 ve %9,6 ile 51+ şeklindedir.



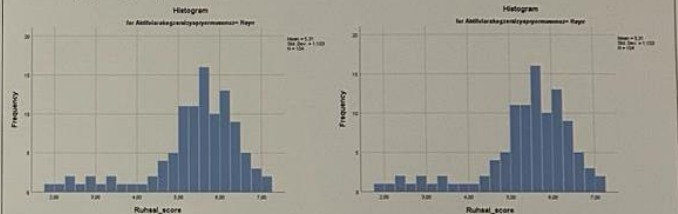
Katılımcılarımızın eğitim durumları ve kiloları hakkındaki düşünceleri ise yukarıdaki pasta grafiklerinde gösterilmektedir. Katılımcıların %10,8'i kendini zayıf, %56,8'i kendini normal ve %32,4 ü kendini kilolu hissetmektedir. Bunun yanı sıra katılımcılarımızın %1,8'i ilkökul, %6,8'i ortaokul, %23,0'i lise, %14,4'ü ön

Araştırma Sonuçları

Öncelikle araştırmamızda toplanan verilere normale uygun olup olmadığını sınamak için ki-kare testi uygulanmış ve elde ettiğimiz veriler aşağıdaki gibi bulunmuştur

| Tests of Normality | | Kategoriye Göre | | Shapiro-Wilk | | | |
|--------------------|---------------------------------|-----------------|------|--------------|------|------|------|
| | Statistik | df | Sig. | Statistik | df | Sig. | |
| Ruhsal_scora | Shapiro-Wilk | .977 | 154 | .200 | .882 | 154 | .200 |
| | Kolmogorov-Smirnov ^a | .183 | 118 | .000 | .234 | 118 | .000 |

Tabloda görüldüğü üzere Sig. değerimiz 0,05 değerinden küçük olduğu için psikolojik iyi oluş ölçeğinin spor yapma açısından yorumunun normal dağılıma uygun olmadığı için parametrik olmayan testler üzerinden yapılacakları saptıyoruz.



Dağılım şekilleri ise yukarıdaki histogram grafiklerinde görülmektedir

| Tablo 3.c | | Tablo 3.d | | Tablo 3.e | |
|------------------------|------------------|------------------------|------------------|------------------------|------------------|
| | Psikolojik_scora | | Psikolojik_scora | | Psikolojik_scora |
| Mann-Whitney U | 9346,500 | Kruskal-Wallis H | 9,380 | Mann-Whitney U | 5305,500 |
| Z | -1,144 | df | 2 | Wilcoxon W | 13882,500 |
| Asymp. Sig. (2-tailed) | .248 | Asymp. Sig. (2-tailed) | .026 | Asymp. Sig. (2-tailed) | .043 |
| | | Asymp. Sig. | .200 | | |

Yukarıda tablo 3.c'de egzersiz yapmanın psikolojik iyi oluş üzerindeki etkisini Mann-Whitney U testi ile ölçtüğümüzde elde ettiğimiz değerleri görmekteyiz. Burada Sig. değerimiz 0,05'den büyük bir değer aldığı için araştırmamızın sonucu olarak kişilerin egzersiz yapmalarının psikolojileri üzerinde anlamlı bir etkisi yoktur diyebilmekteyiz. Tablo 3.d'de ise kişilerin kiloları hakkında düşündükleri ile psikolojik iyi oluş arasında anlamlı bir ilişki var diyebiliriz. Bu çıkarımı yaptıktan sonra hangi gruptan dolayı bu farkın oluştuğunu anlamak için bir de anova testi yapmaktayız. Tablo 3.d'nin yanındaki görselde görüldüğü gibi kendini normal hissedeni insanların psikolojik skorları kendini zayıf veya kilolu hissedenlere oranla daha olumlu sonuçlar vermektedir. Burada vücut kitle indeksine değil kişinin kendini nasıl hissettiğine bakıldığından endeksi yüksek bir insan da düşük bir insan da kendini normal olarak nitelendirilebilir. Bunun yanı sıra egzersiz yapmak bir insanı çeşitli sebeplerden iyi hissettireceği gibi başka bir kişiyi kötü hissettirebilir. Test istatistiklerimizde bu yönde çıkmıştır. Örneğin bir insan kendini yorgun hissettiği için egzersiz yapmaktan kaçınıp psikolojisini başka hobi veya aktiviteleri ile yüksek tutabilir.

Egzersiz yanı sıra kişilerden toplanılan çeşitli demografik bilgiler ile de psikolojik iyi oluş karşılaştırmaları yapılmıştır.

| Tablo 3.f | | Tablo 3.g | |
|------------------|------------------|------------------------|------------------|
| | Psikolojik_scora | | Psikolojik_scora |
| Kruskal-Wallis H | 9,380 | Mann-Whitney U | 5305,500 |
| df | 3 | Wilcoxon W | 13882,500 |
| Asymp. Sig. | .026 | Asymp. Sig. (2-tailed) | .043 |

Tablo 3.e de kişilerin eğitim düzeylerinin psikolojik iyi oluşları üzerindeki etkisi Kruskal-Wallis testi ile ölçülmüş ve Sig. değeri 0,075 ile 0,05 den büyük olması nedeniyle, kişilerin eğitim düzeylerinin psikolojik iyi oluş skorları arasında anlamlı bir fark olmadığını göstermektedir. Aynı şekilde tablo 3.f de ise kişilerin cinsiyet değişkeninin psikolojik skoru üzerindeki etkisi olup olmadığına bakılmış, test istatistiklerimiz 0,141 çıktığından dolayı cinsiyetinde psikolojik skora bir etkisi olmadığı saptanmıştır.

Literatürdeki birçok çalışma aksi yönde bulgular elde etmiş olsa da, günümüz insanları giderek harekete daha az ihtiyaç duymakta. İnsanlar işleri başta olmak üzere eğlence ve benzeri alanlarda da artık daha az hareket etmeye meyillidir. Bunu ile birlikte günümüz teknolojileri ile birlikte bu istekleri hem karşılanıyor hem de daha fazlası insanlara imkan olarak sunuluyor. Eskiden fiziksel olan oyun, eğlence, iş ve benzeri alanlardaki aktiviteler gün geçtikçe azalmakta. Bu yüzden egzersizin de psikolojideki rolü etkisini yitirmiş olabilir.

KAYNAKÇA

[1] Durusoy, E., & Mutuş, R. (2021). Yeşil egzersizin kronik ağrıya, fiziksel ve mental sağlığa etkileri. İstanbul Gelişim Üniversitesi Sağlık Bilimleri Dergisi, (14), 351-362.
 [2] Frontera, W. R., Slovik, D. M., & Dawson, D. M. (Eds.). (2006). Exercise in rehabilitation medicine. Human Kinetics.
 [3] ARDİC, F. (2014). Egzersizin psikolojik iyi oluş üzerindeki etkisi üzerine yapılan araştırmaların değerlendirilmesi.



ÖZET

Bu çalışma "World Happiness Report" tarafından katılımcıların kendi yaşamlarına ilişkin derecelendirmelerine dayanan ulusal mutluluk makalelerini ve sıralamalarını içeren bir yayından alınan parametre ve değişkenler dahilinde katılımcıların mutluluk derecelendirmelerini aynı zamanda yaşam kalitelerini çeşitli faktörlerle ilişkilendirilmesi amaçlanmıştır. Veriler "World Happiness Report" web sitesi üzerinden elde edilmiştir. Avrupa ve Orta Asya bölgesindeki verilerin elde ettiğimiz 31 ülke 496 gözlem arasında katılımcıların mutluluk derecelendirmelerinin gayri safi yurt içi hasıla, cömertlik, güven, yaşam tercihlerini yapma özgürlüğü, sosyal destek, beklenen sağlıklı yaşam beklentisi, mutluluk hali, gülme ve eğlenme gibi duygulanımlar, ülke içerisindeki görülme sıklığı ve son olarak endişe, öfke ve üzüntü gibi duygulanımların ölçümlerine ne kadar bağlı olduklarının, çeşitli regresyon modelleri, durağanlık testleri, eş varyans, değişim varyans ve birimler arası korelasyon tahmin testleri, dinamik panel veri tahmini, otokorelasyon testleri uygulanarak yaklaşımda bulunulmak istenmiştir. Yapılan çeşitli regresyon modellerine göre modelin anlamlı olduğu görülmüştür. Bu veri seti için sabit etkiler modeli genel olarak anlamlıdır ve bağımsız değişkenler bağımlı değişken üzerinde istatistiksel olarak anlamlı etkilere sahiptir. Ayrıca, sabit etkili modelin ülkeler arasındaki farklılıkları açıklamak için oldukça başarılı olduğu söylenebilir.

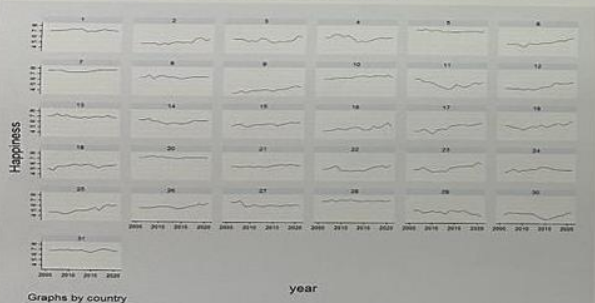
Mutluluk Analizi

Mutlu olmak, farklı mercilerdeki kişilerin görüşlerine göre gerek içimizde olup biten gerekse belirli koşullara bağlı olarak gelişen bir duygulanım olarak görülmektedir. Genel olarak, mutluluk normalde psikolojik açıdan bir ruh hali olarak düşünülür. Geçmişte yapılan çalışmaların bazılarına göre ise yaşamdaki kararlar almada özgürlük ve sosyal destek değişkenlerinin ölçüm sonuçlarının yüksek çıkan ülkelerde yaşayan insanların mutluluk ölçümlerinde daha yüksek değerler gözlenmiştir (M. Mujiya Ulkhaq :Arga Adayatama, 2020). World Happiness Report, mutluluğu sadece ruh halinden daha fazlası olarak görme eğilimindedir, çünkü araştırmacılar birçok sosyoekonomik faktörlerin, insanların belirli bir andaki duygularının ötesine geçtiğini savunmaktadır (Rodrigo Porto Carreiro Neves,2021). Bu koşulların bazıları hakkında konuşulacak olursa, insanların yaşadıkları ülkelerde gözlenen durumları açıklanabilmeleri açısından katılımcıların mutluluk ölçümleri, birtakım değişkenler ile tasvir edilebilir.



Şekil 1. Değişkenler Arası Korelasyon Isı Haritası

Bu korelasyonlar, değişkenler arasındaki ilişkileri gösterir, ancak nedensellik hakkında bir şey söylemez. Korelasyon, iki değişken arasındaki ilişkinin gücünü ve yönünü ölçer, ancak bu ilişkinin nedensel olup olmadığını belirlemez. Bu nedenle, korelasyon sonuçlarına dayanarak dikkatli bir şekilde çıkarımlarda bulunmak önemlidir.



Şekil 2. Ülkelerin Mutluluk Grafikleri

Yukarıdaki grafiklerde ülkelerin 2006-2021 yılları arasındaki mutluluk puanlarının (10 üzerinden) dağılımı görülmektedir. Grafiğin üzerinden kısaca birkaç yorum yapacak olursak ülke numarası 5 (Danimarka) ve 13 (İrlanda) ülkelerinin mutluluk puan ortalamaları en yüksek iken ülke numarası 9 (Gürcistan) ve 30 (Ukrayna) ülkelerinin en düşük puan ortalamalarına sahiptir denilebilir. Ülkeler arasında mutluluk ölçümlerinin dalgalanmalarına bakıldığında 11 (Yunanistan) ve 27 (İspanya) numaralı ülkeler ilk yıllarda panel içerisindeki bahsi geçen diğer ülkelerin ölçümlerine nazaran daha yüksek oranda bir düşüş göstermiştir. Pozitif olarak sıralamalardan bahsedecek olursak 2015 ve sonrasında 30 (Ukrayna) numaralı ülkenin mutluluk ölçüm farkı diğer ülkelere nazaran daha çoktur denilebilir.

Panel veri regresyon modeli sonuçlarına bakıldığında bağımlı değişken olan mutluluk değişkeninin diğer bağımsız değişkenler olan regresyon modeli aşağıda belirtildiği gibidir:

$$H(it) = -5.1118 + 0.6518 \times GDP(it) + 1.4578 \times SSP(it) + 0.5709 \times CNG(it) + 1.8196 \times PE(it) - 2.5560 \times NE(it) - 0.0457 \times PC(it) + 0.1017 \times G(it) + 0.02844 \times HLEB(it) + 0.7293 \times FMLC(it) + u_i + \epsilon_t(it)$$

R² modelin açıklama gücünü gösterir. Bu çıktıda, within R² 0.5849, between R² 0.8573, ve overall R² 0.8154 olarak hesaplanmıştır. Böylece bu model hakkında konuşulduğu zaman bağımsız değişkenlerin bağımlı değişkeni gayet iyi bir şekilde açıkladığı söylenebilir.

KAYNAKÇA

[1] Rodrigo Porto Carreiro Neves (2021), *Representing Happiness: A comparative study of media coverage of the World Happiness Report* Chapter (pp.20)
 [2] Jan-Emmanuel De Neve, & George Ward (2017). *Happiness at work to the World Happiness Report 2017* pp. (146-178)
 [3] Ulkhaq, M. M., & Adayatama, A. (2020). Clustering countries according to the World Happiness Report 2019. pp. 117

```

xtreg H GDP HLEB FMLC SSP PC PE NE G CNG, fe
Fixed-effects (within) regression
Group variable: country
Number of obs = 496
Number of groups = 31
Obs per group: min = 14
max = 16
avg = 16.0
Wald chi2(10) = 502.86
Prob > chi2 = 0.0000
F(10,485) = 76.28
R-squared = 0.8154
Adj R-squared = 0.8000
Total # var: 11
Instruments for differenced equation
GDP: type: beta(1)
Standardized GDP: D_GDP: D_SSP: D_HLEB: D_FMLC: D_G: D_PC: D_PE: D_NE: D_CNG
Instruments for level equation
Standardized_come

```

Tablo 1. Sabit Etkili Grup İçli Panel Veri Regresyon Modeli Analizi

Çıktımızı yorumlayacak olursak mutluluk ve diğer bağımsız değişkenlerimiz arasında negatif yönlü orta dereceden biraz altında korelasyon gözlenmektedir denilebilir. (-%30), (Tablo 1.)

Modelin istatistikleri, R² within değeri 0.5883 olarak bulunmuştur. Bu durum bağımsız değişkenlerin bağımlı değişken üzerinde varyansın yaklaşık %58.83'ünü karşıladığı anlamına gelmektedir. R² between değeri ise 0.831 olarak bulunmuştur. Bu sonuç altında ülkeler arasındaki değişkenliğin büyük bir kısmını açıklanabildiğini söyleyebiliriz. R² overall sonucu ise 0.7927 olarak bulunmuştur. R² overall ifadesi bize bütün veri setindeki toplam değişkenliğin açıklanma oranını ifade eder. Bu değer topl varyansın %79.27'sinin açıklanabildiğini bize göstermektedir. Varyansın %70.65-4395'i birimler arası farklılıklardan kaynaklanmaktadır.

```

xtabond H GDP SSP HLEB FMLC G PC PE NE CNG, lags(1) artests(2)
Arellano-Bond dynamic panel-data estimation
Number of obs = 496
Group variable: country
Time variable: year
Number of instruments = 115
Wald chi2(10) = 502.86
One-step results Prob > chi2 = 0.0000

```

Tablo 2. Arellano Bond Dinamik Panel Veri Tahmini

Dinamik panel veri analizi, panel veri setlerindeki zaman bağlamında değişken bağımsız değişkenlerin etkilerini modellemek üzere kullanıldığı ölçümler de mutluluk ile ilgili fikir edinebilmemiz konusunda bize doğru orantılı sonuçlar vermiştir. Yolsuzluğun görülme yaygınlığı ile ilgili ülkelerdeki yaşayan insanların mutluluk ölçümlerinin arasında bir ters orantı olduğu gözlemlenmiştir. Bağımsız değişkenlerin kendi aralarındaki ilişkilerine bakıldığında beklenen sağlıklı yaşam beklentisi ile gayri safi yurt içi hasıla değişkenleri arasında kuvvetli bir ilişki olduğu gözlemlenmektedir. Yolsuzluğun ülke içerisindeki görülme yaygınlığı değişkeni ile yaşam tercihlerini yapma özgürlüğü değişkeni arasında ise negatif yönlü kuvvetli bir ilişki söz konusudur. Aynı zamanda yaşam tercihlerini yapabileme özgürlüğü ile mutluluk, gülme ve eğlenme gibi duygulanımların ölçümleri değişkenleri arasında kuvvetli pozitif bir ilişki görülmektedir. Katılımcıların cevaplarına göre mutlulukları ile beklenen sağlıklı yaşam beklentisi, yaşamdaki tercihlerini yapabileme özgürlüğü, gayri safi yurt içi hasıla, mutluluk hali, gülme ve eğlenme gibi duygulanımların değişkeni, cömertlik ve güven değişkenleri ile orta seviye ve üzeri düzeyde pozitif bir ilişki gözlenmektedir. Ülkede işçilerin görülme sıklığı ve son olarak endişe, öfke ve üzüntü gibi duygulanımların değişkeni ile mutluluk değişkeni arasında ise negatif yönlü orta düzeyde bir ilişki gözlenmektedir. Panel veri regresyon modeli analizlerine göre de bağımsız değişkenlerin bağımlı değişkeni iyi bir seviyede açıklayabildiği söylenebilir.

SONUÇLAR

Elde edilen bulgulara göre mutluluk ölçümlerinin en çok görüldüğü ülkelerin gelir seviyelerinin mutluluk ölçümleri kendilerine nazaran daha düşük olan ülkelerin gelir seviyelerine oranla daha yüksek olduğu gözlemlenmiştir. Yaşam tercihlerini yapma özgürlüğü hususundaki ölçümler de mutluluk ile ilgili fikir edinebilmemiz konusunda bize doğru orantılı sonuçlar vermiştir. Yolsuzluğun görülme yaygınlığı ile ilgili ülkelerdeki yaşayan insanların mutluluk ölçümlerinin arasında bir ters orantı olduğu gözlemlenmiştir. Bağımsız değişkenlerin kendi aralarındaki ilişkilerine bakıldığında beklenen sağlıklı yaşam beklentisi ile gayri safi yurt içi hasıla değişkenleri arasında kuvvetli bir ilişki olduğu gözlemlenmektedir. Yolsuzluğun ülke içerisindeki görülme yaygınlığı değişkeni ile yaşam tercihlerini yapma özgürlüğü değişkeni arasında ise negatif yönlü kuvvetli bir ilişki söz konusudur. Aynı zamanda yaşam tercihlerini yapabileme özgürlüğü ile mutluluk, gülme ve eğlenme gibi duygulanımların ölçümleri değişkenleri arasında kuvvetli pozitif bir ilişki görülmektedir. Katılımcıların cevaplarına göre mutlulukları ile beklenen sağlıklı yaşam beklentisi, yaşamdaki tercihlerini yapabileme özgürlüğü, gayri safi yurt içi hasıla, mutluluk hali, gülme ve eğlenme gibi duygulanımların değişkeni, cömertlik ve güven değişkenleri ile orta seviye ve üzeri düzeyde pozitif bir ilişki gözlenmektedir. Ülkede işçilerin görülme sıklığı ve son olarak endişe, öfke ve üzüntü gibi duygulanımların değişkeni ile mutluluk değişkeni arasında ise negatif yönlü orta düzeyde bir ilişki gözlenmektedir. Panel veri regresyon modeli analizlerine göre de bağımsız değişkenlerin bağımlı değişkeni iyi bir seviyede açıklayabildiği söylenebilir.

Dinamik panel veri tahminlerinin sonuçlarına göre, bu bir önceki dönemdeki bağımlı değişkenin, mevcut dönemdeki bağımlı değişkeni anlamlı bir şekilde etkilediğini gösterir. Aynı zamanda en az bir bağımsız değişkenin zaman içinde değişen bir etkiye sahip olduğunu göstermektedir. Bu testin p-değeri çok küçük olduğu için, en az bir bağımsız değişkenin zaman içinde değişen bir etkiye sahip olduğu sonucuna varılabilir.

Tüm bunlara ek olarak, mutluluk analizindeki bu çalışmanın geliştirilebilmesi, ilerleyen zamanlardaki yapılacak çalışmalara destek olabilmeleri için modelimizin karmaşıklarıyla eğitim düzeyleri, refah seviyelerinin ölçümü, istihdam durumları gibi değişkenler araştırmanın içine dahil edilebilir. Yapılan geçmiş çalışmalarda aynı zamanda makroekonomik düzeyde, ulusal işsizlik seviyeleri ile negatif korelasyon göstermektedir (Jan-Emmanuel De Neve & George Ward 2017).



Bu çalışma "World Happiness Report" tarafından katılımcıların kendi yaşamlarına ilişkin derecelendirmelerine dayanan ulusal mutluluk makalelerini ve sıralamalarını içeren bir yayından alınan parametre ve değişkenler dahilinde katılımcıların mutluluklarının ve aynı zamanda yaşam kalitelerinin çeşitli faktörlerle ilişkilendirilmesi amaçlanmıştır. Veriler "World Happiness Report" web sitesi üzerinden elde edilmiştir. Avrupa ve Orta Asya bölgesindeki verilerin elde ettiğimiz 31 ülke 496 gözlem arasında katılımcıların mutluluk derecelendirmelerinin gayri safi yurt içi hasıla, cömertlik, güven, yaşam tercihlerini yapma özgürlüğü, sosyal destek, beklenen sağlıklı yaşam beklentisi, mutluluk hali, gülmeye ve eğlenmeye gibi duygulanımlar, ülke içerisindeki görülme sıklığı, otokorelasyon testleri uygulanarak yaklaşımda bulunulmuştur. Yapılan çeşitli regresyon modellerine göre modelin anlamlı olduğu görülmüştür. Bu veri seti için sabit etkiler modeli genel olarak anlamlıdır ve bağımsız değişkenler bağımlı değişken üzerinde istatistiksel olarak anlamlı etkilere sahiptir. Ayrıca, sabit etkiler modelin ülkeler arasındaki farklılıkları açıklamak için oldukça başarılı olduğu söylenebilir.

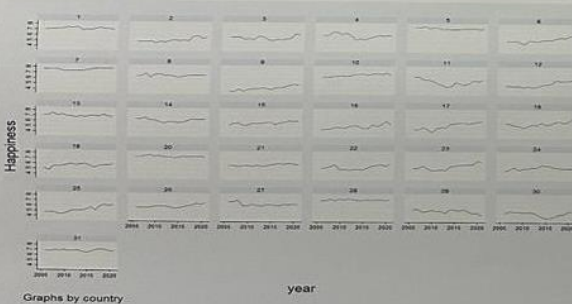
Mutluluk Analizi

Mutlu olmak, farklı mercilerdeki kişilerin görüşlerine göre gerek içimizde olup biten gerekse belirli koşullara bağlı olarak gelişen bir duygulanım olarak görülmektedir. Genel olarak, mutluluk normalde psikolojik açıdan bir ruh hali olarak düşünülür. Geçmişte yapılan çalışmaların bazılarında göre ise yaşamdaki kararları almada özgürlük ve sosyal destek değişkenlerinin ölçüm sonuçlarının yüksek çıkan ülkelerde yaşayan insanların mutluluk ölçümlerinde daha yüksek değerler gözlemlenmiştir (M. Mujiya Ulkhaq; Arga Adayatama, 2020). World Happiness Report, mutluluğu sadece ruh halinden daha fazlası olarak görme eğilimindedir, çünkü araştırmacılar birçok sosyoekonomik faktörlerin, insanların belirli bir andaki duygularının ötesine geçtiğini savunmaktadır (Rodrigo Porto Carreiro Neves, 2021). Bu koşulları bazılar hakkında konuşulacak olursa, insanların yaşadıkları ülkelerde gözlemlenen durumların açıklanabilmesi açısından katılımcıların mutluluk ölçümleri, birtakım değişkenler ile tasvir edilebilir.



Şekil 1. Değişkenler Arası Korelasyon Isı Haritası

Bu korelasyonlar, değişkenler arasındaki ilişkileri gösterir, ancak nedensellik hakkında bir şey söylemez. Korelasyon, iki değişken arasındaki ilişkinin gücünü ve yönünü ölçer, ancak bu ilişkinin nedensel olup olmadığını belirlemez. Bu nedenle, korelasyon sonuçlarına dayanarak dikkatli bir şekilde çıkarımlarda bulunmak önemlidir.



Şekil 2. Ülkelerin Mutluluk Grafikleri

Yukarıdaki grafiklerde ülkelerin 2006-2021 yılları arasındaki mutluluk puanlarının (10 üzerinden) dağılımı görülmektedir. Grafikten üzerinde kısaca birkaç yorum yapacak olursak ülke numarası 5 (Danimarka) ve 13 (İrlanda) ülkelerinin mutluluk puan ortalamaları en yüksek iken ülke numarası 9 (Gürcistan) ve 30 (Ukrayna) ülkelerinin en düşük puan ortalamalarına sahiptir denilebilir. Ülkeler arasında mutluluk ölçümlerinin dalgalanmalarına bakıldığında 11 Yunanistan) ve 27 (İspanya) numaralı ülkeler ilk yıllarda panel içerisindeki bahsi geçen diğer ülkelerin ölçümlerine nazaran daha yüksek oranda bir düşüş göstermiştir. Pozitif olarak sıçramalardan bahsedince olursak 2015 ve sonrasında 30 (Ukrayna) numaralı ülkenin mutluluk ölçüm farkı diğer ülkelere nazaran daha çoktur denilebilir.

Panel veri regresyon modeli sonuçlarına bakıldığında bağımlı değişken olan mutluluk değişkeninin diğer bağımsız değişkenlerle olan regresyon modeli aşağıda belirtildiği gibidir:

$$\hat{H}(it) = -5.1118 + 0.6518 \times GDP(it) + 1.4578 \times SSP(it) + 0.5709 \times CNG(it) + 1.8196 \times PE(it) - 0.5560 \times NE(it) - 0.0457 \times PC(it) + 0.1017 \times G(it) + 0.02844 \times HLEB(it) + 0.7293 \times FMLC(it) + ui + e(it)$$

R² modelin açıklama gücünü gösterir. Bu kıtımda, within R² 0.5849, between R² 0.8573, ve overall R² 0.8154 olarak hesaplanmıştır. Böylece bu model hakkında konuşulduğu zaman bağımsız değişkenlerin bağımlı değişkeni gayet iyi bir şekilde açıkladığını söylenebilir.

KAYNAKÇA

- [1] Rodrigo Porto Carreiro Neves (2021), *Representing Happiness: A comparative study of media coverage of the World Happiness Report* Chapter (pp.20)
- [2] Jan-Emmanuel De Neve, & George Ward (2017), *Happiness at work to the World Happiness Report 2017* pp. (146-178)
- [3] Ulkhaq, M. M., & Adayatama, A. (2020). Clustering countries according to the World Happiness Report 2019. pp.(137-140).

| Model | Coef. | Std. Err. | t | P> t | [95% Conf. Interval] |
|-------|---------|-----------|--------|-------|----------------------|
| H | | | | | |
| GDP | 0.6518 | 0.0688 | 9.46 | 0.000 | 0.5142 0.7894 |
| SSP | 1.4578 | 0.0588 | 24.79 | 0.000 | 1.3392 1.5964 |
| CNG | 1.8196 | 0.0588 | 30.94 | 0.000 | 1.7008 1.9384 |
| PE | 1.8196 | 0.0588 | 30.94 | 0.000 | 1.7008 1.9384 |
| NE | -0.5560 | 0.0588 | -9.46 | 0.000 | -0.6748 -0.4372 |
| PC | -0.0457 | 0.0588 | -0.78 | 0.438 | -0.1644 0.0730 |
| G | 0.1017 | 0.0588 | 1.73 | 0.089 | 0.0000 0.2034 |
| HLEB | 0.7293 | 0.0588 | 12.40 | 0.000 | 0.6105 0.8481 |
| FMLC | 0.7293 | 0.0588 | 12.40 | 0.000 | 0.6105 0.8481 |
| _cons | -5.1118 | 0.0588 | -86.77 | 0.000 | -5.2294 -5.0042 |

Tablo 1. Sabit Etkili Grup İçin Panel Veri Regresyon Modeli Analizi

Çıktımızı yorumlayacak olursak mutluluk ve diğer bağımsız değişkenlerimiz arasında negatif yönlü orta dereceden bir azaltıcı korelasyon gözlemlenmektedir denilebilir. (-%30), (Tablo 1.)

Modelin istatistikleri, R² within değeri 0.5883 olarak bulunmuştur. Bu durum bağımsız değişkenlerin bağımlı değişken üzerindeki varyansın yaklaşık %58.83'ünü karşıladığı anlamına gelmektedir. R² between değerimiz ise 0.831 olarak bulunmuştur. Bu sonuçlar altında ülkeler arasındaki değişkenliğin büyük bir kısmını açıklanabildiğini söyleyebiliriz. R² overall sonucu ise 0.7927 olarak bulunmuştur. R² overall ifadesi bize bütün veri setindeki toplam değişkenliğin açıklanma oranını ifade eder. Bu değer toplam varyansın %79.27'sinin açıklanabildiğini bize göstermektedir. Varyansın %70.654395'i birimler arası farklılıklardan kaynaklanmaktadır.

| Model | Coef. | Std. Err. | t | P> t | [95% Conf. Interval] |
|-------|---------|-----------|--------|-------|----------------------|
| H | | | | | |
| GDP | 0.6518 | 0.0688 | 9.46 | 0.000 | 0.5142 0.7894 |
| SSP | 1.4578 | 0.0588 | 24.79 | 0.000 | 1.3392 1.5964 |
| CNG | 1.8196 | 0.0588 | 30.94 | 0.000 | 1.7008 1.9384 |
| PE | 1.8196 | 0.0588 | 30.94 | 0.000 | 1.7008 1.9384 |
| NE | -0.5560 | 0.0588 | -9.46 | 0.000 | -0.6748 -0.4372 |
| PC | -0.0457 | 0.0588 | -0.78 | 0.438 | -0.1644 0.0730 |
| G | 0.1017 | 0.0588 | 1.73 | 0.089 | 0.0000 0.2034 |
| HLEB | 0.7293 | 0.0588 | 12.40 | 0.000 | 0.6105 0.8481 |
| FMLC | 0.7293 | 0.0588 | 12.40 | 0.000 | 0.6105 0.8481 |
| _cons | -5.1118 | 0.0588 | -86.77 | 0.000 | -5.2294 -5.0042 |

Tablo 2. Arellano Bond Dinamik Panel Veri Tahmini

Dinamik panel veri analizi, panel veri setlerindeki zaman bağlamında değişken bağımsız değişkenler etkilerini modellemek üzere kullanılmış olup beklenen sağlıklı yaşam beklentisi, yaşam tercihlerini yapma özgürlüğü, cömertlik, yolculuğun yaygınlığı, güven gibi bağımsız değişkenlerin bağımlı değişken üzerinde anlamlı bir zaman etkisi olmadığı sonucuna varılır. Aynı zamanda endişe, üzüntü, öfke gibi duygulanım değişkeni, mutlu olma hali, gülmeye ve eğlenmeye gibi duygulanımların, sosyal destek, gayri safi yurt içi hasıla (borderline olarak) ve gecikmeli bağımlı değişken (L1H) arasında mevcut dönemdeki bağımlı değişkeni üzerinde zaman bağlamında anlamlı bir etkisi olduğu anlamına gelmektedir. (Tablo 2.)

SONUÇLAR

Elde edilen bulgulara göre mutluluk ölçümlerinin en çok görüldüğü ülkelerin gelir seviyelerinin mutluluk ölçümleri kendilerine nazaran daha düşük olan ülkelerin gelir seviyelerine oranla daha yüksek olduğu gözlemlenmiştir. Yaşam tercihlerini yapma özgürlüğü hususundaki ölçümler de mutluluk ile ilgili fikir edinebilmeye konusunda bize doğru oranlı sonuçlar vermiştir. Yolculuğun görülme yaygınlığı ile ilgili ülkelerdeki yaşayan insanların mutluluk ölçümlerinin arasında bir ters orantı olduğu gözlemlenmiştir. Bağımsız değişkenlerin kendi aralarındaki ilişkilerinde bakıldığında beklenen sağlıklı yaşam beklentisi ile gayri safi yurt içi hasıla değişkenleri arasında kuvvetli bir ilişki olduğu gözlemlenmektedir. Yolculuğun ülke içerisindeki görülme yaygınlığı değişkeni ile yaşam tercihlerini yapma özgürlüğü değişkeni arasında ise negatif yönlü kuvvetli bir ilişki söz konusudur. Aynı zamanda yaşam tercihlerini yapabileme özgürlüğü ile mutluluk, gülmeye ve eğlenmeye gibi duygulanımların ölçümleri değişkenleri arasında kuvvetli pozitif bir ilişki gözlemlenmektedir. Katılımcıların cevaplarına göre mutlulukları ile beklenen sağlıklı yaşam beklentisi, sosyal destek ve güven değişkenleri ile orta seviye ve üzeri düzeyde pozitif bir ilişki gözlemlenmektedir. Ülkeler içerisindeki yolculuğun yaygınlığı ile endişe, öfke ve üzüntü gibi duygulanımların analizi ile mutluluk değişkeni arasında ise negatif yönlü orta düzeyde bir ilişki gözlemlenmektedir. Panel veri regresyon modeli analizlerine göre de bağımsız değişkenlerin bağımlı değişkeni iyi bir şekilde açıklayabildiği söylenebilir.

Dinamik panel veri tahminlerinin sonuçlarına göre, bu bir önceki dönemdeki bağımlı değişkenin, mevcut dönemdeki bağımlı değişkeni anlamlı bir şekilde etkilediğini gösterir. Aynı zamanda en az bir bağımsız değişkenin zaman içinde değişen bir etkiye sahip olduğunu görülmektedir. Bu testin p-değeri çok küçük olduğu için, en az bir bağımsız değişkenin zaman içinde değişen bir etkiye sahip olduğu sonucuna varılabilir.

Tüm bunlara ek olarak, mutluluk analizindeki bu çalışmanın geliştirilebilmesi, ilerleyen zamanlardaki yapılacak çalışmalara destek olabilmeye için modelimizin karmaşıklıkların eğitim düzeyleri, refah seviyelerinin ölçümü, istihdam durumları gibi değişkenler araştırmanın içine dahil edilebilir. Yapılan geçmiş çalışmalarda aynı zamanda makroekonomik düzeyde, ulusal işsizlik seviyeleri ile negatif korelasyon göstermektedir (Jan-Emmanuel De Neve; George Ward, 2017).



YFU FEN EDEBİYAT FAKULTESİ

İSTATİSTİK BÖLÜMÜ

DÖVİZ KURU VE HİSSE SENEDİ FİYATLARI ARASINDAKİ İSTATİSTİKSEL İLİŞKİ

Sinan Ulusan 18023001

Danışman: Doç. Dr. Fatma Sevinç KURNAZ

ÖZET

Bu çalışmada, Türkiye' deki 2018 yılı Aralık ayı ile 2023 yılı Aralık ayı arasındaki enflasyon oranı ile Pegasus, Şişecam ve Coca Cola şirketlerinin borsadaki hisse senedi fiyatları karşılaştırılmıştır. Bütün bu şirketlerin sırasıyla USD döviz oranı ile hisse senedi fiyatları ile döviz kuru değeri oranı aylık olarak kıyaslanmıştır. Aylık olarak kıyaslanan değerlerin korelasyonuna bakılmıştır, ek olarak nedensellik analizi ve son olarak ise eş bütünleşme analizi uygulanmıştır. Sonuç olarak üç büyük hisse senedi ile USDTRY arasındaki ilişkiyi bakılmıştır.

BORSA

Borsa, devlet denetimindeki alım satım işlemlerinin gerçekleştirildiği bir piyasadır. Dünya genelinde ilk borsa, 1409'da Brugge'da kurulmuş olup bu dönemde 'borsa' terimi ortaya çıkmıştır. Türkiye'de, Osmanlı İmparatorluğu döneminde 1873-1874'te kurulan Dersaadet Tahvilat Borsası ile resmîyet kazanmıştır. Türkiye'deki hisse senedi alım satımı, Borsa İstanbul bünyesinde gerçekleştirilmekteydi. Borsa İstanbul, 1986'da İstanbul Menkul Kıymetler Borsası olarak kurulmuş ve 2013'te Borsa İstanbul adını almıştır.

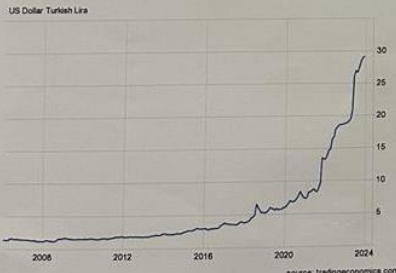
Hisse senedi piyasaları, ülkelerin ekonomik büyümesine önemli katkılarda bulunur. Şirketler, hisse senetleri piyasası aracılığıyla sermaye artışı elde ederken, ülkeler de yabancı sermaye piyasalarından döviz kazanır. Şirketler, hisse senedi ihracı ederek yeni ortaklar bulup öz kaynak sağlamaktadır.

USD TRY DÖVİZ KURU

Döviz kuru bir ülkenin para biriminin diğer ülkenin para birimi karşısındaki piyasa değerini ifade eder. Parite ise, iki ülkenin para birimleri arasındaki uzun vadeli denge değerini ifade eder. USDTRY ABD dolarının Türk lirası karşısındaki değerini gösteren bir paritedir. Türkiye'de en çok işlem yapan parite konumundadır. Bu sebeple, USDTRY iki ülkenin merkez bankası politikaları, siyasi ve ekonomik gelişmelerden önemli ölçüde etkilenir.

Amerikan dolarının değeri Türkiye'deki birçok vatandaş için ekonomik durumu günlük olarak izlenmesi gereken temel göstergelerden biridir. Türk lirasının tarih boyunca geçirdiği değer evrimi ve Türkiye'nin şimdiye kadar benimsemiş olduğu döviz politikalarını sıkça gündeme getirmektedir.

Türkiye'deki 2002 - 2023 Arasındaki USDTRY Kur Değişimi



ZAMAN SERİLERİ

Zaman serileri analizi, istatistiksel verilerin zaman içindeki değişimlerini inceleyen bir analiz yöntemidir. Zaman serileri aynı zamanda istatistiksel analiz ve ekonomi gibi farklı alanlarda kullanılan bir analiz yöntemidir.

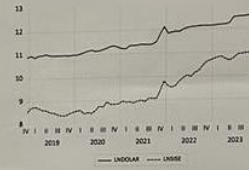
Korelasyon analizi, bağımlı ve bağımsız değişkenler arasındaki ilişkiyi, bu ilişkinin pozitif ve negatif yönlü olduğunu ve şiddeti ile ilgili bilgiler elde edebilmek için kullanılan bir istatistiksel analiz yöntemidir.

Durağanlık, zaman serilerinde oldukça önemli bir kavramdır. Çoğu istatistiksel test yorumlanırken, veri setinin durağan olduğu varsayılır. Eğer veri seti durağan değilse veri seti durağan hale getirilir.

Nedensellik Analizi, iki değişken arasında zamana bağlı bir ilişkinin nedenselliğinin istatistiksel olarak belirlenmesidir.

Zaman serisi verilerinde, yapılan ekonomik analizlerde değişkenler durağan olmadığında eş bütünleşme analizine başvurulmaktadır. Durağan olmayan zaman serilerinde uygulanan regresyon analizlerinin anlamlı bir ilişkileri olabilmeleri için incelenen değişkenlerin aralarında eş bütünleşme ilişkisi bulunmalıdır.

UYGULAMA



Yandaki tabloya göre SISE hisse senedi fiyatları ve enflasyon oranı verileri 2018 yılı Aralık ayı ile 2023 yılı Aralık ayı arasında yükseliş trendi göstermektedir.

| Covariance | LNDOLAR | LNSISE |
|-------------|----------|----------|
| Probability | 0.304061 | |
| LNSISE | 0.481945 | 0.809050 |
| | 0.0000 | |

Değişkenlerin logaritması alınarak aralarındaki korelasyon kat sayısına bakılmıştır. Tabloda da görüldüğü üzere değişkenlerin aralarındaki korelasyon kat sayısı %48'dir. Değişkenlerin kuvvetli bir ilişkisi mevcut görülmemektedir.

Augmented Dickey-Fuller Unit Root Test on D(LNSISE)

| Null Hypothesis: D(LNSISE) has a unit root | t-Statistic | Prob.* |
|--|-------------|--------|
| Exogenous Constant | -6.806435 | 0.0000 |
| Lag Length: 0 (Automatic - based on SIC, maxlags=10) | -3.549590 | |
| Augmented Dickey-Fuller test statistic | -6.806435 | 0.0000 |
| Test critical values: | | |
| 5% level | -2.911730 | |
| 10% level | -2.593551 | |

SISE değişkeninin durağanlığı Düzeltilmiş Dickey - Fuller testi ile kontrol edilmiştir ve durağan olmadığı anlaşılmıştır. Ardından değişken 1. seviyede durağanlaşmıştır.

VAR Residual Serial Correlation LM Tests

| Lag | LRE* stat | df | Prob. | Rao F-stat | df | Prob. |
|-----|-----------|----|--------|------------|-----------|--------|
| 1 | 6.207945 | 4 | 0.1841 | 1.582334 | (4, 88.0) | 0.1842 |
| 2 | 5.841638 | 4 | 0.2113 | 1.462444 | (4, 88.0) | 0.2113 |
| 3 | 7.170664 | 4 | 0.1271 | 1.845886 | (4, 88.0) | 0.1272 |

Modelin gecikme uzunluğunu inceledikten sonra kurulan modelin otokorelasyonu incelenmiştir. Gecikme uzunluğundaki olasılık değerine bakıldığında otokorelasyonun olmadığı tespit edilmiştir. Bu sayede Nedensellik Analizi incelenebilir.

Dependent variable: LNDOLAR

| Excluded | Chi-sq | df | Prob. |
|----------|----------|----|--------|
| LNSISE | 8.022008 | 10 | 0.6267 |
| All | 8.022008 | 10 | 0.6267 |

USDTRY oranı bağımlı değişken olarak tanımlandığında yapılan Granger Nedensellik Analizinde ortaya çıkan olasılık değeri %95 güven düzeyinde değişkenler arasında nedensellik rastlanmamıştır.

Date: 12/03/23 Time: 20:23

| Hypothesized No. of CE(s) | Eigenvalue | Trace Statistic | Trace Critical Value | Prob.** |
|---------------------------|------------|-----------------|----------------------|---------|
| None* | 0.185615 | 20.43838 | 20.26164 | 0.0473 |
| At most 1 | 0.136781 | 8.528700 | 9.164548 | 0.0689 |

Değişkenler arasında uzun vadeli ilişkiyi tespit etmek amacıyla Johansen Eş Bütünleşme Analizi yapılmıştır. Ortaya çıkan sonuçta %95 güven seviyesinde değişkenler arasında Eş Bütünleşme olduğu anlaşılmıştır.

Yapılan tez çalışmasında, hisse senedi fiyatları ve USDTRY döviz kuru değerleri kıyaslanmıştır. İlk olarak E-views programı kullanılarak yapılan regresyon analizi anlamsız çıkmıştır. Bu sebeple veriler incelenerek düzensiz artışlar ve azalışlar olan ilgili dönemlere gölge değişken atanmıştır. Ardından nedensellik analizi yapılabilmeleri için değişkenlerin otokorelasyon varlığı ve değişken varyans varlığı incelenmiştir. Ardından nedensellik analizi uygulanmıştır. Sonrasında Eş Bütünleşme Analizi yapılmıştır. Yapılan analizlerde nedensellik arasında Pegasus hariç Eş Bütünleşme olduğu anlaşılmıştır. Total sonuçta göre ise 2019 yılı Ocak değişkenleri arasında Pegasus hariç Eş Bütünleşme olduğu anlaşılmıştır. Total sonuçta göre ise 2019 yılı Ocak Üretim faaliyetlerinin düşmesi ve döviz kurunun düzensiz artışı sebebiyle gerçekleşen fiyatlandırmalardan dolayı düşüşü göstermiştir, buna rağmen deprem felaketi zamanında ise PGSUS ve SISE hisseleri yükseliş göstermeye devam etmiştir. CCOLA hisse senetlerinin düzensiz artış ve düşük yaşadığı bu periyootta oldukça düzenli artış ve azalış içeren bir tablo göstermiştir.



ÖZET

Bu tez çalışmasında, 303 hastanın öz niteliklerinin olduğu, Kaggle'dan alınan kalp hastalığı veri setinin Makine Öğrenmesi (ML) algoritmaları uygulanmıştır. Sırasıyla Lojistik Regresyon, Naive Bayes, kNN ve Rastgele Orman makine öğrenmesi algoritmaları kullanılarak doğruluk (accuracy), duyarlılık (recall), kesinlik (precision) ve f1-score performans ölçüm metrikleri bulunmuştur. En başarılı olan algoritma Lojistik Regresyon, en başarısız olan algoritma kNN olmuştur.

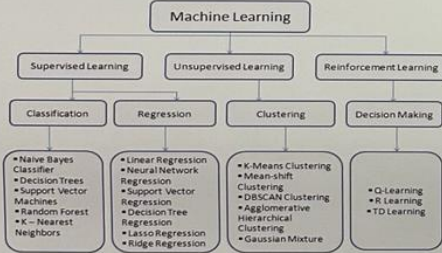
KALP HASTALIĞI

Günümüzde kalp hastalıkları dünyada öne gelen ölüm nedenleri olarak karşımıza çıkmaktadır. Her yıl ortalama 17 milyon kişi kalp hastalıkları sebebiyle hayatını kaybetmektedir. Türkiye'de ölümlerin %40'ı son üç senede kalp hastalığına bağlı olarak ortalama 500 bin insan hayatını yitirmiştir. Kalp hastalıkları kalpte meydana gelen ve kalbi etkileyen herhangi bir bozukluğu kapsayan bir terim olarak karşımıza çıkmaktadır.

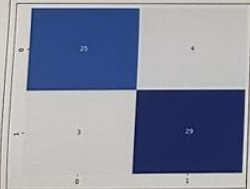


MAKİNE ÖĞRENMESİ

Makine öğrenmesinin (ML) gelişmesiyle birlikte bir çok alanda kullanılmakla, sağlık sektöründe de oldukça fazla görülmektedir. Bu teze Lojistik Regresyon, Naive Bayes, kNN, Rastgele Orman algoritmaları kullanılmıştır.



LOJİSTİK REGRESYON



Accuracy of Logistic Regression: 88.52459016393442

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.89 | 0.86 | 0.88 | 29 |
| 1 | 0.88 | 0.91 | 0.89 | 32 |
| accuracy | | | 0.89 | 61 |
| macro avg | 0.89 | 0.88 | 0.88 | 61 |
| weighted avg | 0.89 | 0.89 | 0.89 | 61 |

Lojistik Regresyon modelinden Accuracy %88.52 oranı ile hastalığı teşhis etmekte başarılı olmuştur.

Naive Bayes

17. yüzyılda yaşayan matematikçi Thomas Bayes tarafından alan Naive Bayes, bir sınıflandırma algoritmasıdır. Yalın tasarımına ve görünüşte basitleştirilmiş varsayımlara rağmen gerçek dünya durumlarında beklenenden daha iyi sonuçlar vermiştir.

Accuracy of Naive Bayes model: 86.88524590163934

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.84 | 0.90 | 0.87 | 29 |
| 1 | 0.90 | 0.84 | 0.87 | 32 |
| accuracy | | | 0.87 | 61 |
| macro avg | 0.87 | 0.87 | 0.87 | 61 |
| weighted avg | 0.87 | 0.87 | 0.87 | 61 |

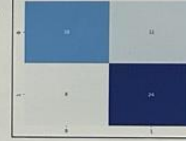
$$\text{Formülü şu şekildedir; } P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Naive Bayes algoritmasında %86.89 doğruluk oranı elde edilmiştir.

kNN (K-Nearest Neighbour)

kNN yani K en Yakın Komşu algoritmaları, 1967 yılında Thomas M. Cover ve Peter E. Hart tarafından önerilmiştir. Algoritma, sınıfları belli olan bir örnek kümesindeki verilerden yararlanılarak kullanılmaktadır.

kNN algoritmasının doğruluk oranı %68.85'tir. Diğer algoritmalarla kıyasla en düşük doğruluk oranı elde edilmiştir.

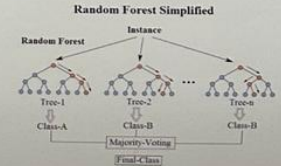


Accuracy of K-NeighborsClassifier: 68.85245901639344

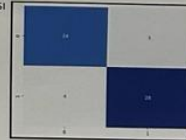
| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.69 | 0.62 | 0.65 | 29 |
| 1 | 0.69 | 0.75 | 0.72 | 32 |
| accuracy | | | 0.69 | 61 |
| macro avg | 0.69 | 0.69 | 0.69 | 61 |
| weighted avg | 0.69 | 0.69 | 0.69 | 61 |

RASTGELE ORMAN (RANDOM FOREST)

Rastgele Orman yöntemi, Leo Breiman tarafından 2001 yılında geliştirilmiş bir yapay öğrenme tekniğidir. İsminde geçen Random ifadesinin arkasında, gözlemlere ve değişkenlere göre rastgele seçim yapılmasından kaynaklı olarak makine öğrenmesinin temel kaygılarının olan genellebilirlik ve aşırı öğrenme gibi konularda farkındalık yaratması vardır.



Rastgele orman algoritması doğruluk oranının %85.25 olduğu görülmektedir.



Accuracy of Random Forest: 85.24590163934425

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.86 | 0.83 | 0.84 | 29 |
| 1 | 0.85 | 0.88 | 0.86 | 32 |
| accuracy | | | 0.85 | 61 |
| macro avg | 0.85 | 0.85 | 0.85 | 61 |
| weighted avg | 0.85 | 0.85 | 0.85 | 61 |

| ALGORİTMA | DOĞRULUK (ACCURACY) |
|--------------------|---------------------|
| Lojistik Regresyon | %88.52 |
| Naive Bayes | %86.89 |
| kNN | %68.85 |
| Rastgele Orman | %85.25 |

Bu çalışmanın sonucunda Makine Öğrenmesi (ML) algoritmalarından en yüksek değeri alan, Lojistik Regresyon olmuştur. En düşük değeri de Naive Bayes almıştır.

KAYNAKÇA

- <https://www.kaggle.com/code/alkidiarete/heartattack-accuracy-92/notebook>
- <https://www.ahmetarnaz.com/kalp-damar-hastaliklari-nedir/>
- <https://aws.amazon.com/tr/what-is/logistic-regression/>



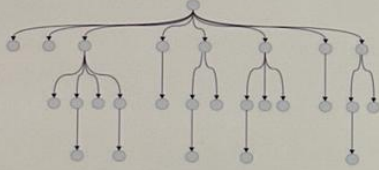
ÖZET

Bu çalışma bankaların çeşitli ürünlerinde rekabet avantajı sağlamak amacıyla makine öğrenimi ve yapay zekâ teknolojilerinin kullanımını vurgulamaktadır. Bu teknolojilerin, müşterileri daha iyi tanıma, müşterilere doğru ürün pazarlayarak müşteri sadakatini artırma ve gelirleri büyüme konularında bankalara yardımcı olduğu belirtilmektedir. Bu çalışma, sınıflandırma yöntemlerini kullanarak banka müşterilerinin vadeli mevduata katılma olasılığını tahmin etmeyi amaçlamaktadır. Bankacılık sektöründe müşteri segmentasyonu ve vadeli mevduat tercihlerini anlama, faiz oranları ve kampanya kararlarını etkileme konularında önemli rol oynamaktadır. Çalışma, karar ağaçları, Rastgele Orman, LightGBM, XGBoost ve Adaboost gibi makine öğrenmesi algoritmalarını kullanarak müşterilerinin vadeli mevduata katılma olasılığının tahminlerini gerçekleştirir.

KULLANILAN ALGORİTMALAR

Karar Ağacı

Karar ağacı algoritması, sınıflandırma algoritmalarından biridir. Bir karar ağacı tanımlanmış hedef değişkenine sahiptir. Model yukarıdan aşağıya böl ve yönet metodu ile çalışır. Veri kümesi, karar verme kuralları uygulanarak daha küçük kümeler bölünür. Karar ağaçları kök, düğüm ve yapraklardan oluşur. Bölünme en iyi bölünmeyi gösteren kök hücre ile başlar. Bu işlem hedefe ulaşıncaya kadar tekrarlı bir şekilde yinelenir.



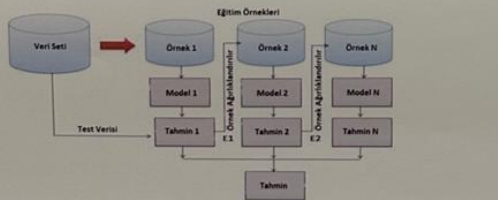
Rastgele Orman

Rastgele Orman algoritması, bir ensemble (topluluk) öğrenme yöntemi olarak karşımıza çıkar. Bu algoritma, birden fazla karar ağacının bir araya gelerek güçlü bir model oluşturmasına dayanır. Rastgele Orman, her bir ağacın farklı alt kümesiyle eğitilmesini ve tahminlerin topluca alınmasını içerir.



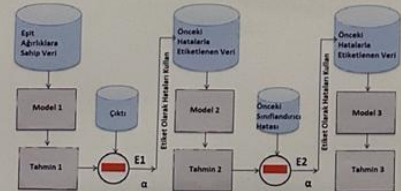
Boosting

Boosting, makine öğrenimi ve istatistikte kullanılan bir ensemble öğrenme tekniğidir. Ensemble öğrenme, bir dizi zayıf öğreniciyi (genellikle düşük doğruluklu modeller) bir araya getirerek güçlü bir öğrenici elde etmeyi amaçlar. Boosting zeminde, zayıf öğreniciler sırasıyla eğitilir ve her bir öğrenici, önceki öğrenicilerin hatalarını düzeltmek için. Bu çalışmada bir tane Boosting algoritması kullanılmıştır. Kullanılan algoritma AdaBoost'dur.



Gradient Boosting

Gradient Boosting, bir tahmin modelinde, genellikle karar ağaçları gibi zayıf tahmini modellerin topluluk formunu üreten, regresyon ve sınıflandırma problemleri için bir makine öğrenme tekniği olarak tanımlanmaktadır. Gradyan artıncı karar ağaçlarında, birçok zayıf öğreniciyi birleştirerek güçlü bir öğrenici bulmaya çalışır. Buradaki zayıf öğreniciler, bireysel karar ağaçlarıdır. Tüm ağaçlar seri olarak bağlanır ve her bir ağaç, bir önceki ağacın hatasını en aza indirmeye çalışır. Bu çalışmada iki tane Gradient Boosting algoritması kullanılmıştır. Bunlar XGBoost ve LightGBM'dir.



VERİ SETİ

Bu veri seti; Portekizli bir bankacılık kurumunun doğrudan pazarlama kampanyaları (telefon aramaları) ile ilgilidir. Sınıflandırma hedefi, müşterinin vadeli mevduata abone olup olmayacağını tahmin etmektir. Veri seti; mevduat pazarlama kampanyasına dahil olan 41189 birey ve bu bireylere ait 15 değişkenlerden, sosyoekonomik durumu gösteren 5 değişken ve hedef değişkenden oluşmaktadır. Üzere toplamda 21 değişken vardır. Kayıp veri yoktur.

Banka müşteri verileri:

- Yaş
- Meslek
- Medeni durum
- Eğitim seviyesi
- Kredi ödemesinde gecikme durumu
- Konut kredisi durumu
- Kişisel kredi durumu

Diğer özellikler:

- Kampanya: bu kampanya sırasında gerçekleştirilen temas sayısı
- Pdays: müşterinin son ziyaretinden sonra geçen gün sayısı.
- Önceki kampanya iletişimi: bu kampanyadan önce ve bu kampanya için gerçekleştirilen temas sayısı.
- Sonuç: önceki pazarlama kampanyasının sonucu.

Sosyal ve ekonomik bağlam özellikleri:

- İstihdam oranı
- Tüketici fiyat endeksi
- Tüketici güven endeksi
- Eurobond
- İletişim süresi, saniye cinsinden
- Çalışan sayısı

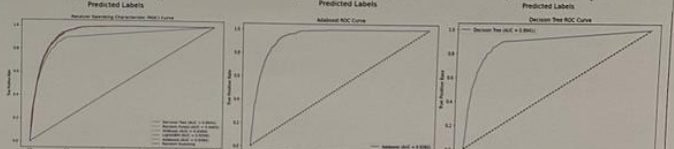
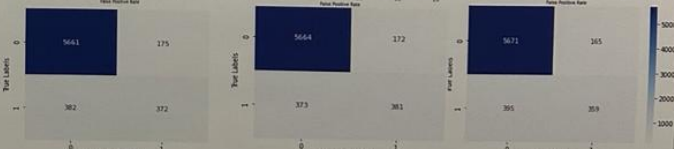
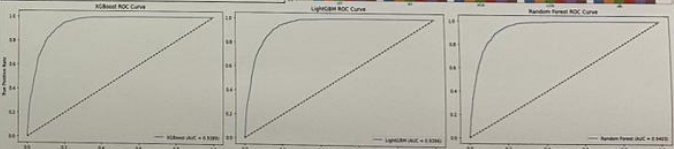
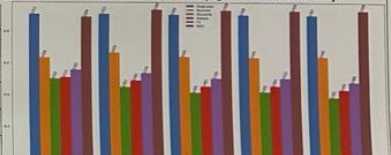


Bu çalışmada veri setinin %80'i eğitim verisi, %20'si test verisi olarak ele alınmış ve uygulamada Python açık kaynak yazılımı kullanılmıştır.

SONUÇ

Kullanılan sınıflandırma algoritmalarından elde edilen sonuçlar performans ölçütleri bakımından incelenmiştir. Performans ölçütleri olarak doğruluk, kesinlik, duyarlılık, Kappa, F1 ve ROC eğrisi değerleri kullanılmıştır. Karar ağaçları, Rastgele Orman, XGBoost ve LightGBM, Adaboost algoritmalarının performans ölçütleri aşağıdaki tabloda verilmiştir.

| Algoritma | Doğruluk | Keskinlik | Duyarlılık | Kappa | F1 | ROC |
|-----------|----------|-----------|------------|-------|-------|-------|
| DT | 0.911 | 0.636 | 0.503 | 0.513 | 0.561 | 0.894 |
| RF | 0.913 | 0.670 | 0.453 | 0.495 | 0.541 | 0.940 |
| XGB | 0.9087 | 0.645 | 0.420 | 0.461 | 0.509 | 0.938 |
| LGB | 0.9085 | 0.641 | 0.426 | 0.464 | 0.512 | 0.936 |
| AD | 0.9073 | 0.646 | 0.392 | 0.440 | 0.488 | 0.938 |



Veri seti: <https://www.kaggle.com/datasets/henriqueyamahata/bank-marketing>

KAYNAKÇA

- [1] Gülçin Y. 2019 XGBOOST VE KARAR AĞACI TABANLI ALGORİTMALARIN DİYABET VERİ SETLERİ ÜZERİNE UYGULAMASI
- [2] BURCU D. 2023 ADABOOST-BAGGING SINIFLANDIRICISINI İYİLEŞTİRMEYE YÖNELİK HİBRİT BİR ÖRNEKLEME ALGORİTMASI: SAĞLIK ALANINDA BİR UYGULAMA



ÖZET

Kişilerin günlük hayatta birden fazla parametreye bağlı olabilecek psikolojik iyi oluşlarının, aktif olarak egzersiz yapıp yapmamaları açısından ele alarak bu bağlamda anlamlı bir fark olup olmadığı araştırılmaktadır. Bu çalışmada, psikolojik iyi oluş ölçeği kullanılması yanı sıra cevapların doğruluğu ne kadar yansıtıyor tamamen kişilerin cevaplarının gerçeği yansıttığı varsayılarak kabul edilmiştir. Bunların yanı sıra, psikolojik iyi oluş subjektif bir olgu olduğundan ve örneklemimiz herhangi bir kısımla ile seçilmediğinden, sonuçlarımızda olabildiğince çeşitli ve farklı çıkabilmektedir. Belirtildiği gibi, psikolojik iyi oluş subjektif bir olgu olsa dahi hayatımızdaki birçok subjektivite de olduğu gibi ortak bir paydada buluşabilirimiz mümkündür. Bu açıdan yapılan çalışmada kişilerin psikolojik iyi oluşunun düzenli egzersiz yapma açısından ortak bir payda oluşturup oluşturmadığını araştırılmaktadır.

Psikoloji ve Psikolojik İyi Oluş Kavramı

Psikoloji tanım gereği insanların içgüdüsel davranışlarını ve zihni düşüncelerini araştıran bu konular hakkında veri toplayan bilim dalıdır. Psikolojik iyi oluş kişilerin kendi hayatlarında var olan durumları yorumlamaları, bu yönde verdiği tepkiler ve genel olarak hissiyatlarını içeren bir ölçek düzeyindedir. Kişilerin kendileri hakkındaki düşünceleri, başkalarının kişinin kendisi hakkındaki düşünceleri, kişilerin kendi hayatına ve hayatlarındaki durumlara bakış açısını gösteren bir olgudur.

Psikolojik İyi Oluş Ölçeği

Bu çalışmada kullanılmış olduğumuz ölçek Ed Diener ve arkadaşlarının gerçekleştirmiş olduğu psikolojik iyi oluş ölçeğidir. Diener'a göre, psikolojik iyi oluş, kişinin yaşamının belirli bir döneminde duygusal ve bilişsel olarak iyi hissetmesi durumunu ifade eder. Diener ve meslektaşları, psikolojik iyi oluşun ölçümü için çeşitli anketler ve ölçekler geliştirmişlerdir. Bu ölçekler, insanların kendi yaşamlarının kalitesini değerlendirebilmelerine ve araştırmacıların bu konuda veri toplamalarına yardımcı olmuştur. Sekiz sorudan oluşan testimizde cevap şıkları en düşük 1 'Kesinlikle katılmıyorum' ve en yüksek 7 'Kesinlikle katılıyorum' şeklindedir. Teste tabii olan kişi 7 puana ne kadar yakın ise psikolojisinin o kadar iyi olduğu varsayılmaktadır. Bu sayede insanların psikolojik iyi oluşları ölçümlenebilir.

Egzersiz Nedir ?

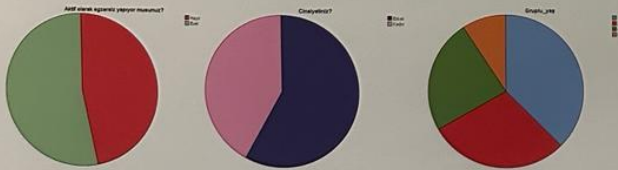
Egzersiz, planlı ve düzenli olarak yapılan bedensel aktivitelerdir. Bu faaliyetler, genellikle vücut sağlığını korumak, fiziksel formu artırmak, kasları güçlendirmek, esnekliği artırmak veya genel olarak sağlıklı bir yaşam tarzını desteklemek amacıyla gerçekleştirilir. Yürüyüş, koşu, yüzme, bisiklete binme, ağırlık antrenmanları, egzersiz salonlarında yapılan çalışmalar, yoga, pilates gibi çeşitli aktiviteler, egzersiz kapsamına girer. Bu aktivitelerin düzenli olarak yapılması, vücut kondisyonunu artırabilir, kalp sağlığını destekleyebilir, stresi azaltabilir, kilo kontrolüne yardımcı olabilir ve genel yaşam kalitesini yükseltebilir.

Egzersiz ve Psikolojik İyi Oluş Kavramı

Fiziksel aktivitenin psikolojik ve fizyolojik sağlığa birçok olumlu etkisi vardır. Araştırmalar, spor ve egzersizin kişilerin psikolojik sağlamlığını geliştirdiğini, stresi azalttığını ve olumlu duygularını artırdığını göstermektedir. Egzersizin kalp damar sağlığı, azami oksijen tüketimi ve benzeri fizyolojik yararlarının yanında kontrol algısında artış, öz yeterlilik ve benzeri psikolojik konularda da olumlu etkisi vardır.

Araştırma Verileri

Bu araştırmada kullanılan veriler anket yöntemi ile toplanılmıştır. Totalde 222 katılımcımız bulunmaktadır. Bu katılımcıların demografik bilgileri ve dağılımları aşağıdaki görseller ile açıklanmaktadır. Araştırmanın verileri toplanırken herhangi bir sınırlama konulmadan kişilerin verdiği cevapları doğru söyledikleri varsayılarak gerekli test ve işlemler yapılmıştır. Bunların sonuçlarında doğan çiktılar üzerinden bir sonuca varılarak araştırma tamama erdirilmiştir.



Yukarıdaki pasta grafiklerde görüleceği üzere katılımcıların %46,8'i egzersiz yapmamakta ve %53,2'si egzersiz yapmaktadır. Bunun yanı sıra katılımcıların %41,9'u kadın ve %58,1' erkektir. Bunların yanı sıra katılımcılarımızın yaş dağılımı %37,4 ile 18-28, %29,5 ile 29-39, %23,5 ile 40-50 ve %9,6 ile 51+ şeklindedir.



Katılımcılarımızın eğitim durumları ve kiloları hakkındaki düşünceleri ise yukarıdaki pasta grafiklerinde gösterilmektedir. Katılımcıların %10,8'i kendini zayıf, %56,8'i kendini normal ve %32,4 ü kendini kilolu hissetmektedir. Bunun yanı sıra katılımcılarımızın %1,8'i ilkökul, %6,8'i ortaokul, %23,0'i lise, %14,4'ü ün

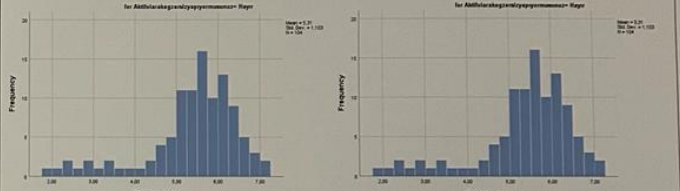
Araştırma Sonuçları

Öncelikle araştırmamızda toplanan verilere normale uygun olup olmadığını sınamak için ki-kare testi uygulanmış ve elde ettiğimiz veriler aşağıdaki gibi bulunmuştur

| Aktif olarak egzersiz yapıyor musunuz? | χ ² değeri | | Sıra | | Sıra | |
|--|-----------------------|-----|------|-----------|------|------|
| | Statistik | df | Sig. | Statistik | df | Sig. |
| Hesap | 157 | 154 | ,000 | 382 | 154 | ,000 |
| Evvel | 183 | 118 | ,000 | 836 | 118 | ,000 |

a. Likertler Significance Correction

Tabloda görüldüğü üzere Sig. değerimiz 0,05 değerinden küçük olduğu için psikolojik iyi oluş ölçeğinin spor yapma açısından yorumunun normal dağılıma uygun olmadığını parametrik olmayan testler üzerinden yapılabileceğini saptıyoruz.



Dağılım şekilleri ile yukarıdaki histogram grafiklerinde görülmektedir

| Tablo 3.c | Tablo 3.d | Tablo 3.e | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|-----------------|-----------------|----------------|----------|------------|----------|---|-------|------------------------|------|---|--|-----------------|------------------|-------|----|---|-------------|------|--|--|-----------------|------------------|--------|----|---|-------------|------|
| <table border="1"> <thead> <tr> <th></th> <th>Psikolojik_olcu</th> </tr> </thead> <tbody> <tr> <td>Mann-Whitney U</td> <td>1386,500</td> </tr> <tr> <td>Wilcoxon W</td> <td>1386,500</td> </tr> <tr> <td>Z</td> <td>-,134</td> </tr> <tr> <td>Asymp. Sig. (2-tailed)</td> <td>,448</td> </tr> </tbody> </table> | | Psikolojik_olcu | Mann-Whitney U | 1386,500 | Wilcoxon W | 1386,500 | Z | -,134 | Asymp. Sig. (2-tailed) | ,448 | <table border="1"> <thead> <tr> <th></th> <th>Psikolojik_olcu</th> </tr> </thead> <tbody> <tr> <td>Kruskal-Wallis H</td> <td>8,386</td> </tr> <tr> <td>df</td> <td>2</td> </tr> <tr> <td>Asymp. Sig.</td> <td>,009</td> </tr> </tbody> </table> | | Psikolojik_olcu | Kruskal-Wallis H | 8,386 | df | 2 | Asymp. Sig. | ,009 | <table border="1"> <thead> <tr> <th></th> <th>Psikolojik_olcu</th> </tr> </thead> <tbody> <tr> <td>Kruskal-Wallis H</td> <td>12,483</td> </tr> <tr> <td>df</td> <td>2</td> </tr> <tr> <td>Asymp. Sig.</td> <td>,000</td> </tr> </tbody> </table> | | Psikolojik_olcu | Kruskal-Wallis H | 12,483 | df | 2 | Asymp. Sig. | ,000 |
| | Psikolojik_olcu | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Mann-Whitney U | 1386,500 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Wilcoxon W | 1386,500 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Z | -,134 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Asymp. Sig. (2-tailed) | ,448 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Psikolojik_olcu | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Kruskal-Wallis H | 8,386 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| df | 2 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Asymp. Sig. | ,009 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Psikolojik_olcu | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Kruskal-Wallis H | 12,483 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| df | 2 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Asymp. Sig. | ,000 | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Yukarıda tablo 3.c'de egzersiz yapmanın psikolojik iyi oluş üzerindeki etkisini Mann-Whitney U testi ile ölçtüğümüzde elde ettiğimiz değerleri görmekteyiz. Burada Sig. değerimiz 0,05'den büyük bir değer olduğu için araştırmamızın sonucu olarak kişilerin egzersiz yapmalarının psikolojileri üzerinde anlamlı bir etkisi yoktur diyebilecekteyiz. Tablo 3.d'de ise kişilerin kiloları hakkında düşüncükleri ile psikolojik iyi oluşları arasındaki Kruskal-Wallis testi istatistikimizin Sig. değerinin 0,05 den aşağıda bir değer olduğu görülmektedir. Bu nedenle 3 farklı grup arasında psikolojik iyi oluş açısından anlamlı bir farklılık var diyebiliriz. Bu çıkarımı yaptıktan sonra hangi gruptan dolayı bu farkın oluştuğunu anlamak için bir de anova testi yapmaktayız. Tablo 3.d'nin yanındaki görselle görüldüğü gibi kendini normal hisseden insanların psikolojik skorları kendini zayıf veya kilolu hissedenlere oranla daha olumlu sonuçlar vermektedir. Buradan yola çıkarak kilo ile egzersizi bağdaştırmamız gerekmektedir. Burada vücut kitle indeksine değil kişinin kendini nasıl hissettiğine bakıldığında endeksi yüksek bir insan da düşük bir insan olarak kendini normal olarak nitelendirilebilir. Bunun yanı sıra egzersiz yapmak bir insanı çeşitli sebeplerden iyi hissettireceği gibi başka bir kişiye kötü hissettirebilir. Test istatistiklerimizde bu yönde çıkmıştır. Örneğin bir insan kendini yorgun hissettiği için egzersiz yaptıktan kaçınıp psikolojisinin başka hobi veya aktiviteleri ile yüksek tutabilir.

Egzersiz yanı sıra kişilerden toplanılan çeşitli demografik bilgiler ile de psikolojik iyi oluş karşılaştırmaları yapılmıştır.

| Tablo 3.f | Tablo 3.g | | | | | | | | | | | | | | | | | | |
|---|-----------------|-----------------|----------------|----------|------------|----------|---|-------|------------------------|------|---|--|-----------------|------------------|-------|----|---|-------------|------|
| <table border="1"> <thead> <tr> <th></th> <th>Psikolojik_olcu</th> </tr> </thead> <tbody> <tr> <td>Mann-Whitney U</td> <td>1305,500</td> </tr> <tr> <td>Wilcoxon W</td> <td>1305,500</td> </tr> <tr> <td>Z</td> <td>-,472</td> </tr> <tr> <td>Asymp. Sig. (2-tailed)</td> <td>,641</td> </tr> </tbody> </table> | | Psikolojik_olcu | Mann-Whitney U | 1305,500 | Wilcoxon W | 1305,500 | Z | -,472 | Asymp. Sig. (2-tailed) | ,641 | <table border="1"> <thead> <tr> <th></th> <th>Psikolojik_olcu</th> </tr> </thead> <tbody> <tr> <td>Kruskal-Wallis H</td> <td>8,990</td> </tr> <tr> <td>df</td> <td>5</td> </tr> <tr> <td>Asymp. Sig.</td> <td>,076</td> </tr> </tbody> </table> | | Psikolojik_olcu | Kruskal-Wallis H | 8,990 | df | 5 | Asymp. Sig. | ,076 |
| | Psikolojik_olcu | | | | | | | | | | | | | | | | | | |
| Mann-Whitney U | 1305,500 | | | | | | | | | | | | | | | | | | |
| Wilcoxon W | 1305,500 | | | | | | | | | | | | | | | | | | |
| Z | -,472 | | | | | | | | | | | | | | | | | | |
| Asymp. Sig. (2-tailed) | ,641 | | | | | | | | | | | | | | | | | | |
| | Psikolojik_olcu | | | | | | | | | | | | | | | | | | |
| Kruskal-Wallis H | 8,990 | | | | | | | | | | | | | | | | | | |
| df | 5 | | | | | | | | | | | | | | | | | | |
| Asymp. Sig. | ,076 | | | | | | | | | | | | | | | | | | |

Tablo 3.e de kişilerin eğitim düzeylerinin psikolojik iyi oluşları üzerindeki etkisini Kruskal-Wallis testi ile ölçtükümüz ve Sig. değeri 0,075 ile 0,05 den büyük olması nedeniyle, kişilerin eğitim düzeylerinin psikolojik iyi oluş skorları arasında anlamlı bir fark olmadığını göstermektedir. Aynı şekilde tablo 3.f de ise kişilerin cinsiyet değişikliğinin psikolojik skorları üzerinde etkisi olup olmadığını bakılmış, test istatistikimiz 0,141 çıktığından dolayı cinsiyetinde psikolojik skora bir etkisi olmadığını saptanmıştır.

Literatürdeki birçok çalışma aksi yönde bulgular elde etmiş olsa da, günümüz insanları giderek hareket daha az ihtiyaç duymakta. İnsanlar işleri başta olmak üzere eğlence ve benzeri alanlarda da artık daha az hareket etmeye meyillidir. Bunu ile birlikte günümüz teknolojileri ile birlikte bu istekleri hem karşılanıyor hem de daha fazlası insanlara imkan olarak sunuluyor. Eskiden fiziki olan oyun, eğlence, iş ve benzeri alanlardaki aktiviteler gün geçtikçe azalmakta. Bu yüzden egzersizin de psikolojideki rolü etkisini yitirmiş olabilir.

KAYNAKÇA

[1] Durusoy, E., & Mutuş, R. (2021). Yeşil egzersizin kronik ağrıya, fiziksel ve mental sağlığı etkileri. İstanbul Gelişim Üniversitesi Sağlık Bilimleri Dergisi, (14), 351-362.
 [2] Frontera, W. R., Slovick, D. M., & Dawson, D. M. (Eds.). (2006). Exercise in rehabilitation medicine. Human Kinetics.



ÖZET

Havayolu ulaşımı günümüzde en yaygın ve popüler olan ulaşım türlerinden birisidir. Gerek zamandan tasarruf gerekse güvenlik bakımından ulaşımı kolaylaştırın yapısı sebebi ile insanlar özellikle uzak mesafeler için bu ulaşım türünü oldukça sık tercih etmektedir. Havayolu ulaşımı sadece yolcu taşımacılığı için değil aynı zamanda kargo taşımacılığı için de tercih kullanılmaktadır. Havayolu taşımacılığının bu denli yaygın olması sebebi ile günümüzde sektörde çok sayıda havayolu taşımacılığı yapan şirket bulunmaktadır. Bunlardan biri de Türk Hava Yolları'dır (Turkish Airlines). Türk Hava Yolları (THY) kargo ve yolcu taşımacılığı yapan bir hava yolu şirkettir. Havayolu şirketlerinin bu denli fazla olması sektörde rekabete yol açmıştır. Bu rekabet ile ortaya farklı hizmetler çıkmıştır. Bu çalışmada Havayolu yolcu taşımacılığında verilen hizmetler ve yolculara ait bilgileri birlikte yolcuların memnuniyet durumlarının yer aldığı, bir veri seti üzerinde memnuniyet değerleri analiz edilerek, makine öğrenmesi algoritmaları ile tahmin modeli oluşturmak amaçlanmıştır.

HAVA YOLU TAŞIMACILIĞI

Hava yolu taşımacılığı ulaşım araçlarından biri olan uçak ile yapılan taşımacılık işlemine verilen addir. Hava yolu taşımacılığı özellikle uzun mesafe için getirdiği ulaşım kolaylığı, zamandan tasarruf, konfor gibi artıları ile insanların en çok tercih ettiği taşımacılık türlerinden biridir.

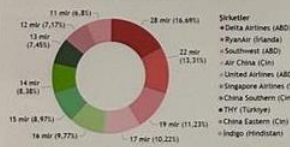
TÜRK HAVA YOLLARI

Türk Hava Yolları (THY), Türkiye'nin bayrak taşıyıcı ulusal hava şirkettir. Türk Hava Yolları'nın merkezi İstanbul'dadır. Uçuş ağı Avrupa, Orta Doğu, Uzak Doğu, Kuzey Afrika, Orta Afrika, Güney Afrika Kuzey ve Güney Amerika'ya kadar uzanmaktadır.

Amerika havacılık derecelendirme kurumu Aviate tarafından 2018 yılında verilen 5 yıldızlı küresel hava yolu ödülünü günümüze kadar taşımaktadır. SkyTrax'ın 2016 ödülleri göre arka arkaya 6. Kez Avrupa'nın en iyi hava yolu seçilmiştir. Geniş uçuş ağı sayesinde dünyanın en çok ülkesine sefer düzenleyen havayolu ünvanına sahiptir. Logoları, 1959 yılında Mesut Maniöglü tarafından tasarlandı. Logo 9.000 metreye ulaşan irtifada kıtalararası uçuş yapabilesiyle bilinen yaban kazı figüründen oluşmaktadır.



Borsa Değerlerine Göre Hava Yolları

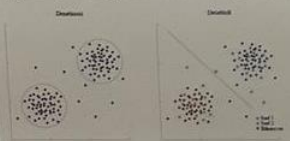


Türk Hava Yolları borsa değerlerine bakıldığında dünya çapında ilk ona girmektedir. Fiyosuna göre dünya havayolu şirketleri sıralamasında ise 12. sırada yer almaktadır. Türk Hava Yolları, 5 uçak ve 30'dan fazla çalışanla 1933 tarihinde Devlet Hava Yolları İşletmesi adıyla başlamıştır. Ülkesinin ilk yurtdışı uçuşunu İstanbul'dan Atina'ya 1947 yılında gerçekleştirdi. 1951'de 33 uçak filosu ile Lefkoşa, Beyrut ve Kahire gibi yeni rotalara uçuşmaya başladı. Hava yolu, bugünkü adı olan Türk Hava Yolları'nı 1955'te aldı ve Uluslararası Hava Taşımacılığı Birliği (IATA)'nın üyeleri arasında yerini aldı.

MAKİNE ÖĞRENMESİ

Makine öğrenimi, temel olarak yapay zekadan kaynaklanır. Yani makine öğreniminin yapay zekanın alt dalı olduğu söylenebilir. Makine öğrenimi, bilgisayar deneyimini öğrenmeye, örneklerle öğrenmeye ve bu öğrenme yeteneklerinin zaman içinde zeka sisteminin performansını geliştirmesine olanak sağlayan uyarlamalı mekanizmayı içerir.

Makine öğrenmesinin teminli algoritmalar oluşturmaktadır. Makine öğrenmesinin farklı çeşitlerinin olmasının sebebi ise farklı yapıdaki algoritmaların kullanılmasıdır. Makine öğrenmesinin teminli oluşturan ve ortaya makine öğrenmesi çeşitlerini çıkaran algoritmalarından bazıları şu şekildedir;

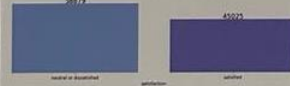


- Denetimli Makine Öğrenmesi (Supervised Algorithms)
- Denetimsiz Makine Öğrenmesi (Unsupervised Algorithms)
- Yarı Denetimli Algoritmalar (Semi - Supervised Algorithms)
- Takviye Algoritmaları (Reinforcement Algorithms)

UYGULAMA

Bu çalışmada havayolu yolcu memnuniyeti birtakım etkenler bazında incelenmiştir. Analiz kısmında sınıflandırma işlemi makine öğrenmesi yöntemi ile yapılmıştır. Veriyi analiz ederken Python programlama dilinin yanı sıra, grafiklerin oluşturulmasında PowerBI kullanılmıştır.

Verimizde 58,879 memnun değil veya kararsız gözlemleri varken 45,025 adet memnun gözlemleri yer almaktadır.

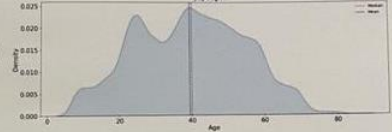


YOLCU PROFİLİ



Yolculara ait değerlere bakıldığında çoğunluğun kadın olduğu görülebilir. Aynı zamanda çoğunluk sadık müşteri kategorisindedir. Sınıf tercihinin bakıldığında ise en çok tercih edilen sınıf ekonomi plus sınıfı olmuştur. Sonrasında ekonomi en az tercih edilen ise bussines sınıfıdır.

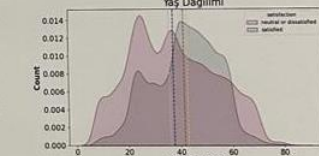
Yolcuların yaş dağılımına bakıldığında yolcuların çoğunun 20-40 yaşları arasında olduğu görülebilir. Yaş aralığı olarak oldukça geniş bir aralıkta yolcuya sahip olduğu da söylenebilir.



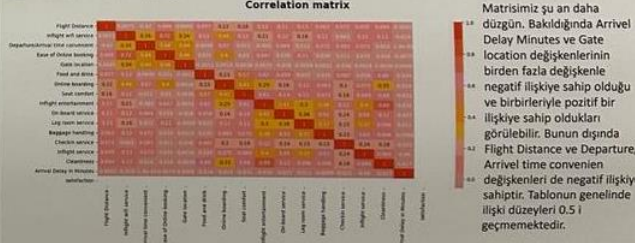
YOLCU MEMNUNİYETİ



Memnuniyet değerlerinin cinsiyet, müşteri tipi, sınıf ve yaş ile olan ilişkisi grafik haline getirilmiştir. Cinsiyet grafiğine bakıldığında kadın ve erkek olarak fark etmezsiniz memnun değil veya kararsız değerinin fazla olduğu görülebilir. Aynı durum sadık olan ve olmayan müşteriler için de geçerlidir. Sınıf bazında bakıldığında diğerlerinden farklı olarak bussines sınıfında memnuniyetin daha yüksek olduğunu görebiliriz. Memnuniyetin 40-60 yaşları arasında yüksek iken 20-40 yaşları arasında memnun değil veya kararsız durumu daha fazladır.



KORELASYON ANALİZİ



Matrisimiz şu an daha düzgün. Bakıldığında Arrival Delay Minutes ve Gate location değişkenlerinin birinden fazla değişimle negatif ilişkiye sahip olduğu ve birbirleriyle pozitif bir ilişkiye sahip oldukları görülebilir. Bunun dışında Flight Distance ve Departure/Arrival time convenient değişkenleri de negatif ilişkiye sahiptir. Tablonun genelinde ilişki düzeyleri 0.5'i geçmemektedir.

MODEL YAPIMI

Model yapımında makine öğrenimi algoritmaları olan xgboost ve catboost kullanılmıştır. Bu algoritmaların her ikisi de Gradient Boosting Machine (GBM) türüdür.

Xgboost modeli kurulup doğruluk skoruna bakıldığında modelin %96 oranında doğru sonuç verdiği görülmüştür. Ayrıca oluşturulmuş olan karışık matrisi değerleri incelendiğinde 17,130 değer doğru iken doğru tahmin edilmiş 408 değer doğru iken yanlış denilmiştir. 801 değer yanlışken doğru denilmiş 12,798 değer ise yanlışken yanlış denilmiştir. Catboost modeline bakıldığında ise modelin %97 doğruluk oranı ile iyi bir tahmin yaptığı görülmüştür.

```

[ ] : accuracy_score(y_test, preds)
0.9611715964929184
[ ] : confusion_matrix(y_test, preds)
array([[17130, 408],
       [ 801, 12798]])
[ ] : cb_model.score(catboost_train_data)
0.9766971353131347
    
```



KAYNAKÇA

- [1] <https://www.glojistik.com/haber-detay/havayolu-tasimaciliginin-tarihsel-gelistimi/172>
- [2] <https://www.turkishairlines.com/tr-int/basin-odasi/hakkimizda/hakkevimiz/>
- [3] https://www.melway.com.tr/tr/yapay-zeka-ai-takviye-ogrenimi-nedir_a.htm/



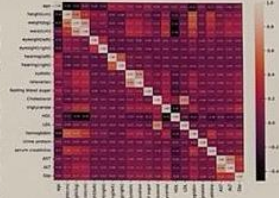
ÖZET

Sigara içmenin dünya genelinde ve ülkemizde önemli bir halk sağlığı sorunu oluşturduğu ve çeşitli sağlık sorunlarına, özellikle kardiyovasküler hastalıklar, kanser ve solunum yolu hastalıkları gibi ciddi sorunlara neden olduğu belirtilmiştir. Makine öğrenimi projesi, bireyleri temel sağlık biyolojik sinyal verilerine göre sigara içen ve içmeyen olarak sınıflandırmayı hedeflemektedir, özellikle sigarının kan basıncına olan etkisini ortaya çıkarmak amaçlanmıştır. Rastgele Orman modeli, %81,77 duyarlılık ve %80,18 kesinlik ile yüksek bir doğruluk oranı elde ederek, sigara içenleri doğru bir şekilde sınıflandırmada etkili olduğunu göstermektedir; bu bulgular, halk sağlığı müdahaleleri ve politikalarının geliştirilmesinde önemli bir katkı sağlamaktadır.

| Değişken Adı | Açıklama | Türü |
|-------------------------------------|---|-----------|
| Cinsiyet | Bireyin cinsiyeti, genellikle erkek veya kadın olarak belirtilir. | Kategorik |
| Yaş | Bireyin yaşı temsil eden sayısal bir değer. | Sürekli |
| Boy(cm) | Bireyin boy uzunluğu santimetre cinsinden. | Sürekli |
| Kilo(kg) | Bireyin ağırlığı kilogram cinsinden. | Sürekli |
| Bel(cm) | Bireyin bel çevresi santimetre cinsinden. | Sürekli |
| Görme Yeteneği(Sol) | Bireyin sol gözündeki görme yeteneği. | Sürekli |
| Görme Yeteneği(Sağ) | Bireyin sağ gözündeki görme yeteneği. | Sürekli |
| Duyuma Yeteneği(Sol) | Bireyin sol kulağındaki duyma yeteneği. | Sürekli |
| Duyuma Yeteneği(Sağ) | Bireyin sağ kulağındaki duyma yeteneği. | Sürekli |
| Sistolik | Kan basıncının sistolik ölçümü. | Sürekli |
| Diastolik | Kan basıncının diastolik ölçümü. | Sürekli |
| Açlık Kan Şekeri | Açlık durumundaki kan şekeri ölçümü. | Sürekli |
| Kolesterol | Kan kolesterol seviyesi. | Sürekli |
| Trigliserit | Kan trigliserit seviyesi. | Sürekli |
| HDL (Yüksek yoğunluklu lipoprotein) | Yüksek yoğunluklu lipoprotein, "iyi" kolesterol. | Sürekli |
| LDL (Düşük yoğunluklu lipoprotein) | Düşük yoğunluklu lipoprotein, "kötü" kolesterol. | Sürekli |
| Hemoglobin | Kan hemoglobin seviyesi. | Sürekli |
| İdrar Protein | İdrardaki protein miktarı. | Sürekli |
| Serum Kreatinin | Kandaki kreatinin seviyesi. | Sürekli |
| AST (Aspartat Aminotransferaz) | Karaciğer ve diğer dokulardaki hasarı değerlendiren bir enzim. | Sürekli |
| ALT (Alanin Aminotransferaz) | Karaciğer hasarını değerlendiren bir enzim. | Sürekli |
| GTP (Gamma-Glutamyl Transferaz) | Karaciğer, safra yolları ve pankreas sağlığını değerlendiren bir enzim. | Sürekli |
| Oral | Ağız sağlığı ile ilgili faktörleri ifade eder. | Kategorik |
| Diş Çürüğü | Diş çürüğü varlığını ifade eder. | Kategorik |
| Tartar | Diş taşı birikimi. | Kategorik |
| Sigara İçme Durumu | Bireyin sigara içip içmediğini ifade eder. | Kategorik |

Korelasyon Matrisi

Veri setindeki korelasyon grafiği, yaş, boy, kilo ve hemoglobin seviyeleri arasında güçlü bir ilişki olduğunu gösterirken, ağırlık ile bel, ağırlık ile HDL, bel ile HDL, kolesterol ile LDL, trigliserit ile HDL ve AST ile ALT arasında belirgin korelasyonlar tespit edilmiştir. Korelasyonlar üzerine yapılan değerlendirmeler, ayrıklı değerlerin olabileceğini ve bu durumun model performansını olumsuz etkileyebileceğini işaret etmektedir, bu nedenle veri setinin dikkatlice incelenmesi ve gerektiğinde ayrıklı değerlerin ele alınması önerilir. Ayrıca, tıp alanında uzman bir kişiye danışarak elde edilen göstergeler, korelasyonların klinik olarak anlamlı olup olmadığını değerlendirmekte önemli bir rol oynayabilir, bu da daha güvenilir ve anlamlı sonuçların elde edilmesine katkı sağlar.



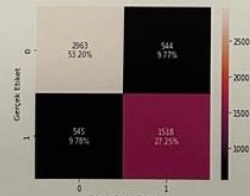
Lojistik Regresyon Modeli

Lojistik Regresyon modelinin test sonuçlarına göre elde edilen hata matrisi, sigara içmeyenleri %49,12 doğru ve %13,84 yanlış tahminle tanıma başarısının iyi olduğunu, ancak yanlış tahmin oranının dikkat çekici olduğunu göstermektedir. Sigara içenleri tahmin etme başarısında ise %26,03 doğru ve %11,01 yanlış tahminle modelin genel olarak kabul edilebilir bir başarı elde ettiği ancak iyileştirilebilecek alanlar olduğu belirtilmektedir. Hata matrisi, Lojistik Regresyon modelinin güçlü ve zayıf yönlerini anlamamıza yardımcı olurken, yanlış sınıflandırmaların nedenlerini anlamak ve modelin performansını artırmak için değerlendirmelerin önemli bir temel oluşturabileceğini vurgulamaktadır.

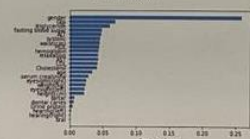


Karar Ağaçları Modeli

Karar Ağaçları modelinin test sonuçlarına göre elde edilen hata matrisi, sigara içmeyenleri tanıma başarısının %53,20 doğru ve %69,77 yanlış tahminle özellikle lojistik regresyon modeline göre daha yüksek olduğunu gösterirken, sigara içenleri tahmin etme başarısında %27,25 doğru ve %9,78 yanlış tahminle karar ağaçları modelinin genel performansını benzer bir seviyede tuttuğunu, ancak sigara içmeyenleri tanıma başarısındaki artışın modelin performansını iyileştirdiğini ortaya koymaktadır. Hata matrisi, karar ağaçları modelinin güçlü ve zayıf yönlerini anlama, yanlış sınıflandırmaların nedenlerini anlama ve modelin performansını geliştirmek için değerli bir değerlendirme aracı sunmaktadır.



Sütun grafiğine göre, karar ağaçları modeli için cinsiyet değişkeninin belirgin bir ağırlığı sahip olduğu ve diğer özelliklere kıyasla daha etkili olduğu gözlemlenirken, sağ kulak duymaması ve oral değişkenlerin sütun grafiğinde düşük öneme sahip olduğu görülmektedir. Bu durum, cinsiyetin modelin sınıflandırma sürecinde önemli bir ayırıcı ölçütü olduğunu gösterirken, sağ kulak duymaması ve oral değişkenlerin modelin sonucunu belirleme sürecinde daha az etkili olduğunu veya hiç etkisi olmadığını düşündürülebilir, bu nedenle modelin optimize edilmesi için bu özelliklerin değerlendirilmesi gerekebilir.

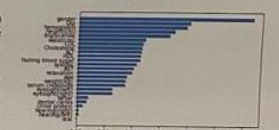


Rastgele Orman Modeli

Rastgele Orman modelinin test sonuçlarına göre elde edilen hata matrisi, sigara içmeyenleri %55,48 doğru ve %7,49 yanlış tahminle belirgin bir doğruluk oranına sahip olduğunu gösterirken, sigara içenleri tahmin etme başarısında diğer modellere göre önemli bir üstünlük sergilediğini ortaya koymaktadır. Hata matrisi, modelin güçlü ve zayıf yönlerini anlama, yanlış sınıflandırmaların nedenlerini anlama ve performansını artırma konusunda değerli bir araç oluşturmaktadır.



Rastgele Orman modelinin değerlendirdiği sütun grafiğine göre, cinsiyet, GTP, hemoglobin, boy ve trigliserit özellikleri modelin en önemli belirleyicileridir. Öte yandan, oral değişkeni belirgin bir öneme sahip değildir, bu da modelin sınıflandırma sürecinde daha az etkili olabileceğini işaret etmektedir. Bu analiz, modelin performansını optimize etmek için odaklanılması gereken özellikleri belirlemede önemli bir rehberlik sağlamaktadır.



SONUÇ

| Model | Doğruluk | F1 Skoru | Kesinlik Ölçütü | Duyarlılık |
|--------------------|----------|----------|-----------------|------------|
| Lojistik Regresyon | 0.751526 | 0.676937 | 0.652859 | 0.702860 |
| Karar Ağaçları | 0.804488 | 0.736000 | 0.736178 | 0.735822 |
| Rastgele Orman | 0.857630 | 0.809695 | 0.801806 | 0.817741 |

Rastgele Orman modeli, %85,76 ile en yüksek doğruluk oranını sağlayarak, sigara içen ve içmeyen bireyleri sınıflandırmada başarılı bir performans sergilemiştir, ayrıca %81,77 duyarlılık ve %80,18 hassasiyet elde etmiştir. Diğer modeller olan Lojistik Regresyon ve Karar Ağaçları ise sırasıyla %75,15 ve %80,44 doğruluk oranlarına sahiptir, ancak Rastgele Orman modelinin öne çıktığı görülmektedir. Elde edilen sonuçlar, makine öğrenimi modellerinin sağlık verilerini etkili bir şekilde sınıflandırabileceğini ve özellikle sigara içme durumunu belirlemede başarılı olduğunu göstererek, sağlık profesyonellerine daha etkili müdahalelerde bulunma ve sigara içenleri belirleme konusunda yardımcı olabileceğini vurgulamaktadır.

KAYNAKÇA

- [1] KARADAKOVAN, A. (1989). SİĞARA İÇEN VE İÇMEYEN BİREYLERDE KAN BASINCI DEĞERLERİNİN KARŞILAŞTIRILMASI. Ege Üniversitesi Hemşirelik Fakültesi Dergisi, 5(2), 15-23.
- [2] Okutan, O., Taş, D., Kaya, H., & Kartaloğlu, Z. (2007). Sigara içen sağlık personelinde nikotin bağımlılık düzeyini etkileyen faktörler. Tüberküloz ve Toraks Dergisi, 55(4), 356-363.
- [3] SABRİ ÖZYURT, M., DAYIOĞLU, H., BİNGÖL, N., & KILIÇ, F. (2004). SİĞARA KULLANIMININ KANDAKİ TOTAL KOLESTEROL, TRİGLİSERİD VE PROTEİN SEVİYELERİNE ETKİLERİ. Journal of Science and Technology of Dumlupınar University, (007), 45-60.



İSTATİSTİK BÖLÜMÜ

**E-TİCARET ÜZERİNDEN PERAKENDE ALIŞVERİŞ YAPAN
MÜŞTERİLERİN KÜMELEME ANALİZİ İLE SINIFLANDIRILMASI**

Mehmet Şihli Gökçe

18023020

Danışman: Arş. Gör. Dr. Coşkun PARİM

ÖZET

Bu tez çalışmasında, Pakistan'da bulunan bir e-ticaret şirketinin verileri kullanılmıştır. Veri kümesinin betimsel istatistikleri özetlenmiş ve analiz öncesinde gerekli ön işleme işlemleri uygulanmıştır. Ardından, temel bileşenler analizi (TBA) ve k-ortalamlar gibi denetimsiz makine öğrenmesi algoritmaları kullanılarak müşteri segmentasyonu gerçekleştirilmiştir. Analiz sonucunda kullanıcılar altı farklı kümeye ayrılmış ve her bir kümenin alışveriş alışkanlıkları detaylı bir şekilde incelenmiştir. Bu segmentasyonun işletmeye uzun vadeli müşteri sadakati ve karlılık sağlamada nasıl yardımcı olabileceği tartışılmıştır.

Veri Kümesinin Tanımlanması

| Müşteri No | Sipariş No | Kategori | Fiyat | Sepet Tutarı | Tarih |
|------------|------------|--------------------|-------|--------------|------------------------|
| 60018 | 574301 | Güzellik & Bakım | 4900 | 4900 | 2017-08-28 00:00:00 |
| 60020 | 574311 | Sağlık & Spor | 700 | 3850 | 2017-08-28 00:00:00 |
| 4112 | 574308 | Eğlence | 18960 | 18960 | 2017-08-28 00:00:00 |
| 60019 | 574309 | Telefon & Tablet | 18900 | 18900 | 2017-08-28 00:00:00 |
| 21928 | 574310 | Elektronik Aletler | 4440 | 4440 | 2017-08-28 00:00:00 |

Müşterilerin bir yıl içerisindeki alışveriş bilgilerini içeren veri kümesi 282337 satırdan ve Tablo 1'de yer alan 6 değişkenden oluşmaktadır. Her bir satır ilgili müşterinin bir siparişinde bulunan bir ürüne denk gelmektedir. Tablo 1'de görülebileceği gibi, veri kümesinde rastgele seçilen 64.189 benzersiz kullanıcının 2017-08-28 ile 2018-08-28 tarihleri arasında vermiş olduğu sipariş bilgileri yer almaktadır.

Betimsel İstatistikler

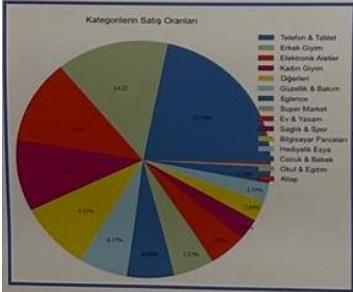
Veri kümesi, 282.348 sipariş, 64.189 kullanıcı ve 14 ana kategoriden oluşmaktadır. Kategori değişkenine göre ürünlerin sayısı şu şekildedir: Telefon & Tablet (61.352), Erkek Giyim (40.155), Elektronik Aletler (32.597), Kadın Giyim (27.842), Güzellik & Kişisel Bakım (17.423), Eğlence (17.194), Süper Market (14.747), Ev & Yaşam (13.682), Sağlık & Spor (8.314), Bilgisayar (8.024), Hediyeleşim Eşya (7.134), Çocuk & Bebek (6.290), Okul & Eğitim (1.050) ve Kitap Kategorisi (764).

| | Fiyat | Sepet Tutarı |
|------------|------------|--------------|
| n | 282348.00 | 282348.00 |
| Ortalama | 8568.13 | 11259.84 |
| Std. Sapma | 17453.57 | 26339.55 |
| Minimum | 0.00 | 1594.00 |
| 1.Kartil | 499.00 | 1234.16 |
| 2.Kartil | 1199.00 | 3125.00 |
| 3.Kartil | 9599.00 | 11553.45 |
| Maksimum | 1012626.00 | 1028751.00 |

Veri kümesindeki sayısal değişkenlerin betimsel istatistikleri yukarıdaki tabloda görülebileceği üzere, sepet tutarı 112.600 Amerikan Doları iken, ortalama ürün tutarı ise 8.568,13 Amerikan doları olarak görülmektedir.

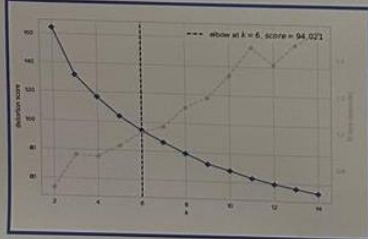
Veri Görselleştirme

Veri görselleştirme aşamasında daha anlamlı grafikler elde edebilmek adına, veri setinin bütünlüğünü bozan aykırı değerlerin bulunmadığı yeni bir veri seti oluşturulmuştur. Veri setinde boş değer kontrolü yapılmış ve boş bir değer bulunmamıştır. Ancak, sepet tutarı ve ürün tutarı 0 TL olan veriler tespit edilmiştir. Aykırı değerler, görselleştirme amaçlı arakçı veriden çıkarılmıştır. Bu işlem, aykırı değerlerin manipülasyonunun makine öğrenmesi modelinin tahminlerini kilyebileceği gereğinden kaynaklanmaktadır.

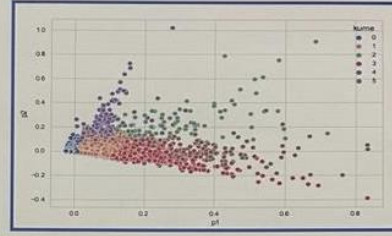


Kategorilerin Satış Oranları Grafiği incelendiğinde ise, en çok satış yapılan kategorinin %21,73 ile Telefon ve Tablet kategorisinin olduğu ve onu %14,22 ile Erkek Giyim kategorisinin izlediği görülebilmektedir. Ardından Kadın Giyim, Elektronik Aletler, Süper Market ve diğer kategorilerinin geldiği görülmektedir.

Kümeleme Analizi



Şekildeki kElbow Görselleştirme grafiği, makine öğrenmesi modelinin eğitim süresini de içererek en optimal küme sayısını düz çizgi ile göstermektedir.



Verilerin kümeler arası dağılımına bakıldığında, ilk kümede 51834 kullanıcı, ikinci kümede 10214 kullanıcı, üçüncü kümede 81 kullanıcı, dördüncü kümede 1364 kullanıcı, beşinci kümede 381 kullanıcı, altıncı kümede 315 kullanıcı bulunduğu görülmüştür.

| Küme | 0 | 1 | 2 | 3 | 4 | 5 |
|--------------------|---------|-----------|------------|-----------|-----------|-----------|
| Elektronik Aletler | 1834.07 | 11789.41 | 60645.22 | 14142.09 | 352751.46 | 22212.86 |
| Güzellik & Bakım | 243.90 | 95.30 | 14706.45 | 50.59 | 1704.21 | 63.59 |
| Kitap | 4.77 | 3.55 | 24.07 | 0.00 | 9.18 | 0.00 |
| Bilgisayar | 388.94 | 5286.48 | 164819.30 | 10933.69 | 15441.25 | 6096.86 |
| Eğlence | 932.87 | 20322.46 | 136458.67 | 8320.73 | 163803.73 | 7802.50 |
| Sağlık & Spor | 91.37 | 66.22 | 371.47 | 21.92 | 2492.66 | 0.00 |
| Ev & Yaşam | 161.24 | 110.38 | 532.79 | 138.98 | 2856.78 | 32.05 |
| Çocuk & Bebek | 75.38 | 53.80 | 332.41 | 109.27 | 670.97 | 0.00 |
| Erkek Giyim | 585.67 | 251.42 | 3867.47 | 106.28 | 5439.42 | 62.03 |
| Telefon & Tablet | 3113.92 | 64931.77 | 2604690.40 | 112456.17 | 189368.49 | 411036.93 |
| Okul & Eğitim | 8.85 | 0.88 | 0.00 | 0.00 | 27.97 | 0.00 |
| Hediyeleşim | 42.95 | 14.63 | 177.47 | 0.29 | 753.50 | 3.14 |
| Süpermarket | 161.42 | 216.48 | 8328.65 | 193.41 | 1477.99 | 1.02 |
| Kadın Giyim | 754.31 | 313.30 | 1719.79 | 321.09 | 4063.99 | 0.00 |
| Minimum | 2702.31 | 18774.33 | 22891.46 | 80400.89 | 7175.37 | 75321.78 |
| Maksimum | 4736.72 | 33430.97 | 351966.23 | 94164.76 | 68344.76 | 90365.75 |
| Ortalama | 3959.13 | 41741.47 | 168847.83 | 98027.22 | 55065.19 | 304638.12 |
| Toplam | 9271.02 | 104818.84 | 2996827.54 | 147666.20 | 741814.02 | 448263.35 |



Son olarak oluşturulan kümelere ait gözlem değerlerinin ortalama alışveriş alışkanlıklarından yola çıkarak kümeleme isimlendirilmiştir. Aynı zamanda bu kümelere pazarlama bakışıyla, karakteristik yapılarına özel ürün ve hizmet önerileri sunulması tavsiye edilmiştir.

KAYNAKÇA

[1] Akın, Y. K. (2008). Veri madenciliğinde kümeleme algoritmaları ve kümeleme analizi (Doctoral dissertation, Marmara Üniversitesi (Turkey)).
[2] Mahesh, B. (2018). Machine Learning Algorithms - A Review. International Journal of Science and Research (IJSR), 9(1), 381-386.
[3] Lestari, M. E. I. (2015). Penerapan algoritma klasifikasi Nearest Neighbor (K-NN) untuk mendeteksi penyakit jantung. Faktor Exacta, 7(4), 366-371.

