



# Düzce Üniversitesi Bilim ve Teknoloji Dergisi

*Araştırma Makalesi*

## Shiny ile R Uygulaması Geliştirme ve Müşteri Kayıp Analizine İlişkin Web Destekli Bir Örnek Çalışma

Fatih KAYAALP\*, Muhammet Sinan BAŞARSLAN

*Bilgisayar Mühendisliği Bölümü, Mühendislik Fakültesi, Düzce Üniversitesi, Düzce, TÜRKİYE*  
*\*fatihkayaalp@duzce.edu.tr*

### ÖZET

Veri madenciliğindeki problemlerin çözümünde çeşitli platformlar kullanılır. R programlama dili de bu platformlar içinde yaygın olarak kullanılanlardan biridir. R programlama dili açık kaynak kodlu olması ve sunduğu topluluk desteği sayesinde araştırmacıların ilgi odağındadır. Shiny, R kodlarının bir web arayüzü aracılığıyla paylaşılmasını sağlayan bir R paketidir. Müşteri Kayıp Analizi, hizmet almakta olan müşterilerden hangilerinin hizmet alımını durdurmayı düşündüğünü tahmin etmek amaçlı olarak kullanılan bir analizdir. Özellikle abonelik temelli sektörlerden olan haberleşme, ödemeli TV gibi sektörlerde müşteri analizlerinde yaygın olarak kullanılmaktadır. Bu çalışmada, telekomünikasyon sektöründeki müşteri veri seti üzerinde C4.5 karar ağacı algoritması kullanılarak R paketlerinden biri olan Shiny vasıtasıyla web destekli olarak gerçekleştirilmiş olan müşteri kayıp analizi sunulmaktadır. Sunulan çalışma, uygulamanın gerçek hayattan alınmış bir veri seti üzerinde yapılmış olması, bu analizin yerel bilgisayarda yapılması yerine erişim izni olan kullanıcılar tarafından web üzerinden erişilerek uzaktan yapılabilmesine imkan sağlaması ve telekomünikasyon alanında müşteri kayıp analizi hakkında daha önceden R ile web tabanlı olarak yapılmamış olması özelliklerinden dolayı özgünlük taşımaktadır.

**Anahtar Kelimeler:** *Veri madenciliği, R dili, Müşteri Kayıp analizi, Shiny, Telekomünikasyon*

## A Web-based Customer Churn Analysis Study With Shiny On R

### ABSTRACT

Various platforms are used to solve problems in data mining. The R programming language is also one of the widely used within these platforms. The R programming language is open source, and thanks to community support it is the on focus of researchers. Shiny is an R package that allows R code to be shared via web interface. Customer Churn Analysis is used to predict the customers who are going to stop receiving services. Customer churn analysis is widely used in customer analysis on communications sector which is one of the subscription-based sectors. A customer churn analysis has been performed and presented in the paper by applying C4.5 decision tree algorithm on telecommunications customer data which is web-supported by Shiny, one of the R packages. The study is different than others as it has been applied on a real dataset, allowing the analysis to be made possible remotely on the web by remote users who have permissions to access instead of doing on local computer and a similar paper has not been seen in literature on telecommunicatons churn analysis before with R.

**Keywords:** *Data Mining, R Language, Customer Churn Analysis, Shiny, Telecommunications*

## I. GİRİŞ

Veri madenciliği, büyük ve karmaşık veriden anlamlı ve kurallı bilgiler elde edilmesi yani bilgi keşfidir. Veri madenciliğinde kullanılan ücretli ve ücretsiz programlar vardır. Bu programlar içinde açık kaynak kodlu olması ve sunduğu topluluk desteği sayesinde R programlama dili araştırmacılar arasında etkin olarak kullanılmaktadır [1]. Müşteri Kayıp Analizi; cep telefonu operatörleri, sigorta şirketleri gibi abone bazlı sektörlerde çalışan şirketlerden hizmet almakta olan müşterilerden hangilerinin kısa vadede hizmet alım sözleşmesini bitireceğini tahmin etmek üzere yapılan veri madenciliği esaslı bir veri analiz işlemidir. Bu analiz sonucunda hizmet sözleşmesini bitirmesi beklenen müşteriler belirlenir ve bu müşterilerin hizmet akdini uzatabilmek için genellikle çeşitli kampanyalar düzenlenir. Çünkü eldeki müşteriyi tutmak, yeni müşteri kazanmaktan her zaman daha ucuzdur.

R programlama dili sürekli geliştirilen binlerce pakete sahiptir. Sunulan çalışmada, bu paketlerden biri olan Shiny kullanımı yapılmıştır[2]. Shiny, yazılan R kodlarının bir web arayüzü aracılığıyla paylaşılmasını sağlayan bir R paketidir[3]. Bu sayede Shiny ile R programlama dilinde yazılan kodlar, HTML, CSS ya da JavaScript bilgisine gerek duymadan web arayüzlerinden kullanıcıların erişimine sunulabilmekte ve web destekli uygulamalar geliştirilerek uygulama dinamik bir yapıda web üzerinden kullanılabilir.

Bu çalışmada veri madenciliği tekniklerinden C4.5 Karar ağacı algoritması kullanılarak Shiny ile Müşteri kayıp analizi anlatılmıştır. İkinci bölümde Shiny uygulaması gerçekleştirilmiş olan Telekomünikasyon veri seti ve C4.5 veri madenciliği algoritması anlatılmıştır. Üçüncü bölümde Shiny hakkında temel bilgiler ve gerçekleştirilmiş olan Shiny uygulaması anlatılmıştır. Dördüncü ve son bölümde de elde edilen sonuçlar verilmiştir.

## II. TELEKOMÜNİKASYON VERİ SETİ ve C4.5 KARAR AĞACI ALGORİTMASI

Bir telekomünikasyon şirketinden alınan müşteri verilerinden yararlanılarak müşterilerin ayrılma ihtimallerini tahmin etmek üzere Shiny'den yararlanarak dinamik bir Müşteri Kayıp Analizi çalışması yapılmıştır. Bu çalışmada kullanılan veri setinde, 8000 müşteri kaydının kişisel detaylar haricindeki temel bilgileri ile beraber 6 aylık haberleşmeleriyle alakalı verileri mevcuttur. Bu veri setinin, gerekli ön işleme ve değişken seçme çalışmaları sonrasında ortaya çıkan hali üzerinde C4.5 karar ağacı algoritması kullanılarak müşterilerin 7. ayda ayrılma oranlarına dair karar ağaçları ile bir sınıflandırılma gerçekleştirilmiştir.

### *A. VERİYİ ANLAMA VE HAZIRLAMA*

Telekomünikasyon şirketinden alınmış olan veri setinde 8000 müşteri kaydı ve 20 değişken bulunmaktadır. Bu 20 değişkenin isimleri, açıklamaları ve veri tipleri tablo 1'de görülmektedir.

**Tablo 1:** Telekomünikasyon veri seti ilişkin tüm değişkenler, gösterim biçimleri ve türleri

Nitelik	Açıklaması	Veri tipi
age	Yaş	Nümerik
age_of_line	Müşterilik süresi	Nümerik
tariff_type	Tarif tipi (postpaid, prepaid)	Nominal
device_type	Cihaz tipi, smartphone, laptop vs.	Nominal
last_reload_year	Yükleme yapılan son tarih (prepaid aboneler için)	Nominal
mmo_count_07	Aylık kendi aboneleriyle konuşma sayısı (arama)	Nümerik
mmo_duration_07	Aylık kendi aboneleriyle konuşma süresi (arama)	Nümerik
mmt_count_07	Aylık aboneleriyle konuşma sayısı (aranma)	Nümerik
mmt_duration_07	Aylık kendi aboneleriyle konuşma süresi (aranma)	Nümerik
mmo_non_count	Diğer aboneleri ile aylık konuşma sayısı (arama)	Nümerik
mmo_non_duration	Diğer aboneleri ile aylık konuşma süresi (arama)	Nümerik
mmt_non_count	Diğer aboneler ile aylık konuşma sayısı (aranma)	Nümerik
mmt_non_duration	Diğer aboneleri ile aylık konuşma süresi (aranma)	Nümerik
mmo_total_count	Aylık toplam konuşma sayısı (arama)	Nümerik
mmo_total_duration	Aylık toplam konuşma süresi (arama)	Nümerik
mmt_total_count	Aylık toplam konuşma sayısı (aranma)	Nümerik
mmt_total_duration	Aylık toplam konuşma süresi (aranma)	Nümerik
msmo_count_07	Aylık sms atma sayısı	Nümerik
callcenter_count_07	Aylık çağrı merkezi şikayet arama sayısı	Nümerik
Churn_2013_07	Abone churn durumu	İkili

Telekomünikasyon veri setine ilişkin tablo 1’de görülen veri setinin veri türleri ve genel dağılımı Şekil 1’de görülmektedir.

```
> str(telekomkat)
'data.frame': 8000 obs. of 20 variables:
 $ age                : int  50 37 29 47 22 55 31 37 53 49 ...
 $ age_of_line       : num  2935 49 2935 2935 49 ...
 $ tariff_type       : Factor w/ 2 levels "Konturlu","Faturalı": 1 2 1 1 1 1 2 1 1 1 ...
 $ device_type       : Factor w/ 8 levels "Bilinmeyen","Mobil Tel",...: 5 5 1 5 5 2 5 2 2 ...
 $ last_reload_year  : int  2014 1900 1900 2014 2014 2014 2014 2013 1900 ...
 $ mmo_count_07      : num  7 32 1 1 8 0 148 0 0 19 ...
 $ mmo_duration_07   : Factor w/ 9 levels "0-2 saat","2-4 saat",...: 1 1 1 1 1 1 2 1 1 1 ...
 $ mmo_non_count_07  : num  3 15 29 98 61 117 136 23 0 1 ...
 $ mmo_non_duration_07 : Factor w/ 11 levels "0-500 saat","500-1000 saat",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ mmt_count_07      : num  54 27 9 93 139 12 29 19 0 48 ...
 $ mmt_duration_07  : Factor w/ 27 levels "0-500 saat","500-1000 saat",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ mmt_non_count_07  : num  3 15 29 98 61 117 136 23 0 1 ...
 $ mmt_non_duration_07 : Factor w/ 7 levels "0-500 saat","500-1000 saat",...: 1 1 1 1 1 1 1 1 1 ...
 $ mmo_total_count_07 : num  10 128 325 253 0 171 284 23 0 114 ...
 $ mmo_total_duration_07 : Factor w/ 22 levels "0-500 saat","500-1000 saat",...: 1 1 1 1 1 2 4 1 1 2 ...
 $ mmt_total_count_07 : num  136 56 0 54 0 70 117 29 0 62 ...
 $ mmt_total_duration_07 : num  10796 0 0 3330 10574 ...
 $ msmo_count_07     : int  1644 35 29 2 1644 17 1644 5 0 16 ...
 $ callcenter_count_07 : int  0 0 1 0 1 0 2 1 0 0 ...
 $ churn_2013_07     : Factor w/ 2 levels "H","E": 1 1 1 1 1 1 1 1 1 1 ...
```

**Şekil 1:** Veri türlerinin RStudio'daki tür ve genel dağılımları

Şekil 2’de telekomünikasyon veri setinin özet bilgisi görülmektedir. Şekil 2’de yer alan niteliklerin minimum değeri (Min), maksimum değeri (Max), birinci çeyreklik değeri (1st Qu.), medyanı (Median), ortalaması (Mean) ve üçüncü çeyreklik değeri (3st Qu.) görülmektedir.

```
> summary(telekomkat)
  age      age_of_line      tariff_type      device_type      last_reload_year      mmo_count_07
Min.   :16.00   Min.    : 49   Konturlu:5461   Mobil Tel   :3701   Min.    :1900   Min.    : 0.00
1st Qu.:28.00   1st Qu.: 266   Faturali:2539   Akilli Telefon:3254   1st Qu.:1900   1st Qu.: 0.00
Median :36.00   Median : 705                                     Bilinmeyen  : 786   Median :2013   Median : 15.00
Mean   :38.13   Mean   :1117                                     Usb Modem   : 229   Mean   :1981   Mean   : 40.63
3rd Qu.:46.00   3rd Qu.:1476                                     Tablet PC   : 17   3rd Qu.:2014   3rd Qu.: 54.00
Max.   :73.00   Max.   :2935                                     Modul       : 11   Max.   :2014   Max.   :1049.00
                                         (Other)    : 2
mimo_duration_07  mimo_non_count_07      mimo_non_duration_07  mmt_count_07      mimo_duration_07
0-2 saat :7245   Min.    : 0.00   0-500 saat :7213   Min.    : 0.00   0-500 saat :7346
2-4 saat : 591   1st Qu.: 2.00   500-1000 saat : 589   1st Qu.: 4.00   500-1000 saat : 415
4-6 saat : 107   Median : 24.00   1000-1500 saat : 128   Median : 23.00   1000-1500 saat : 108
6-8 saat : 46   Mean   : 63.87   1500-2000 saat : 34   Mean   : 42.62   1500-2000 saat : 64
8-10 saat : 6   3rd Qu.: 82.00   2000-2500 saat : 16   3rd Qu.: 59.00   2000-2500 saat : 28
10-12 saat : 2   Max.   :1155.00   2500-3000 saat : 7   Max.   :629.00   2500-3000 saat : 16
(Other) : 3                                     (Other)    : 13                                     (Other)    : 23
mmt_non_count_07      mmt_non_duration_07  mmo_total_count_07      mimo_total_duration_07
Min.    : 0.00   0-500 saat :7506   Min.    : 0.0   0-500 saat :6104
1st Qu.: 2.00   500-1000 saat : 412   1st Qu.: 6.0   500-1000 saat :1176
Median : 24.00   1000-1500 saat : 58   Median : 58.0   1000-1500 saat : 378
Mean   : 63.87   1500-2000 saat : 16   Mean   : 107.6   1500-2000 saat : 163
3rd Qu.: 82.00   2000-2500 saat : 5   3rd Qu.: 153.0   2000-2500 saat : 64
Max.   :1155.00   2500-3000 saat : 1   Max.   :1533.0   2500-3000 saat : 51
                                         (Other)    : 64
mmt_total_count_07  mmt_total_duration_07  mmo_count_07      callcenter_count_07  churn_2013_07
Min.    : 0   Min.    : 0   Min.    : 0   Min.    : 0.0000   H:7857
1st Qu.: 19   1st Qu.: 1981   1st Qu.: 1   1st Qu.: 0.0000   E: 143
Median : 69   Median : 10447   Median : 15   Median : 0.0000
Mean   : 97   Mean   : 19009   Mean   : 403   Mean   : 0.9441
3rd Qu.:138   3rd Qu.: 24524   3rd Qu.: 172   3rd Qu.: 1.0000
Max.   :883   Max.   :808596   Max.   :1644   Max.   :26.0000
```

Şekil 2: Veri setinin özet bilgisi

### III. SHINY PAKETİ İLE KARAR AĞAÇLARI ESASLI MÜŞTERİ KAYIP ANALİZİ UYGULAMASI

Veri madenciliği yöntemlerinden sınıflandırma algoritmaları ile yapılan çalışmaların daha dinamik hale gelmesi Shiny paketi ile yapılabilmektedir. Shiny ile web tabanlı dinamik veri madenciliği uygulamaları gerçekleştirilebilmektedir. Bu bölümde ilk olarak veri madenciliği algoritmaları ile web tabanlı dinamik bir çalışma gerçekleştirmek için yapılması gereken işlemler anlatılmıştır. Daha sonra Telekomünikasyon müşterilerine yönelik kayıp analizi için gerçekleştirilen işlemler anlatılmıştır. Bu çalışma boyunca RStudio [4], R Shiny paketi [2], performans değerlendirme yöntemlerinden hold-out için caret [5] paketi kullanılmıştır.

Uygulama gerçekleştirildikten sonra, web ortamında paylaşılması shinyapps.io ortamından olabilmektedir. Shinyapps.io, RStudio tarafından yönetilen ve Shiny uygulamalarının barınmasına imkan sağlayan bir platformdur [6]. Bu platforma ücretli ve ücretsiz kayıt olma işlemini gerçekleştirdikten sonra shinyapps.io ‘daki başlama klavuzunda yer alan direktiflere göre yüklenebilir.[7]

#### A. VERİYİ ANLAMA VE HAZIRLAMA

Shiny paketi, R programla dili için kullanıcılara kodlama anlamında avantajlar sunan RStudio yazılım geliştirme ortamına ait bir projedir[8]. Shiny uygulamalarında kullanılan dosyalar da R uzantılı dosyalar olsa da bu dosyaların çalıştırılması ve bir araya getirilmesi normal R çalıştırma prensibinden farklıdır. Bu noktadaki fark, hem “kullanıcı” hemde “sunucu” taraflı olmasıdır. Kullanıcı arayüzü kullanıcıya gösterilen metin, grafik ya da işlem başlatacak buton gibi bileşenleri içermesinin yanı sıra kullanıcıdan alınan girdilere göre çıktıları da gösterir. Kullanıcı arayüzündeki formlar sunucu tarafı ile

ilişkilendirilir. Yani kullanıcı arayüzünde kullanıcı forma veri girdiğinde bu veriler sunucu tarafına gönderilir. Sunucu tarafında R kodları ile yapılmak istenen amaca yönelik işlemler yapılarak kullanıcı arayüzünde gösterilmek üzere oluşturulan çıktılar kullanıcı arayüzüne gönderilir.

İlk olarak Rstudio’da uygulama klasörü açılır ve bu klasörün içine üç farklı dosya oluşturulur. Bunlar “ui.R”, ”server.R” ve ”global.R” dosyalarıdır. “Global.R” dosyası, “server.R” ve “ui.R” dosyaları tarafından erişilen bir dosyadır.

“Global.R” dosyasında kullanılacak kütüphaneler, veri setinin okunacağı dosya dizini veya web url’nin bildirilmesi ve veri seti üzerinde yapılması gereken çeşitli işlemler bulunur. “ui.R” dosyasında kullanıcıların kişisel bilgi, aylık telekomünikasyon bilgisi gibi bilgileri girmeye ve bu bilgiler ile yapılan sınıflandırma sonucunu görmeye imkan sağlamak için tasarlanmış bir arayüz bulunur. ”server.R” dosyasında da kullanıcı arayüzünden gelen bilgiler kullanılarak tanımlanan analizler yapılır.

### B. Shiny’de Karar Ağaçları ile Müşteri Kayıp Analizi

Uygulamaya geçmeden önce yukarıda isimleri geçen “Global.R”, “ui.R”, “server.R” dosyaları oluşturulur. “Global.R” dosyasında yapılan işlemler; telekomünikasyon müşteri veri setini dosyadan çağırma, kullanılacak R paketleri ve veri seti sütun isimleri ile ilgili kodlar olarak sıralanabilir. Şekil 3’de oluşturulan “Global.R” dosyası görülmektedir.

```
global.R
1 library(shiny)
2 library(caret)
3 library(e1071)
4 library(RWeka)
5
6 musteriler<-read.table("C:/Users/sinan/Desktop/sonn/data/as.csv",header = FALSE, sep="," ,dec=".",encoding = "UTF-8")
7
8
9 colnames(musteriler)<-c("age","age_of_line","tariff_type","device_type","last_reload_year","mmo_count_07",
10 "mmo_duration_07","mmo_non_count","mmo_non_duration",
11 "mmt_count_07","mmt_duration_07","mmt_non_count","mmt_non_duration",
12 "mmo_total_count_07","mmo_total_duration_07","mmt_total_count_07","mmt_total_duration_07",
13 "mmo_count_07","callcenter_count_07","churn_2013_07")
```

Şekil 3: Global.R dosya görüntüsü

Kullanıcı ara yüzü sayfa tasarımı için fluidPage ve column fonksiyonları kullanılmıştır. FluidPage, yapısında sayfa ve sütunlar bulunduran bir sayfa yerleşim biçimidir. fluidRow bir satır; column ise bir satırdaki sütunları oluşturmak için kullanılmıştır. Şekil 4’de görüldüğü gibi uygulamanın başlığı title ifadesiyle verilirken; “window title” ise web tarayıcının başlık alanında görüntülenmek üzere verilmiştir. Ayrıca style ile stil düzenlemesi yapılmış ve h6, div, hr gibi çeşitli HTML komutları kullanılmıştır.

```
fluidPage(
  fluidRow(column(12,titlePanel(title="KARAR AĞAÇLARI ile MUSTERİ KAYIP ANALİZİ",
    windowTitle = "Shiny - Veri Madenciliği"),
    div("Basarısız Sinan M.,2017. KARAR AĞAÇLARI ile MUSTERİ KAYIP ANALİZİ",
      style="color:black;font-size:small"),
    h6("Müşteri Kayıp Analizi Uygulaması Bir Telekomünikasyon Veri Üzerinde Gerçekleştirilmiştir. "),
    hr()
  )),
)
```

Şekil 4: Ui.R’da satır ve sütun düzenlemeleri kodu

Şekil 5’de görüldüğü gibi telekomünikasyon veri setindeki tamsayı olan nitelikler numericInput ile tanımlanmış ve ilk değer olarak da 0 atanmıştır.

```

fluidRow(
  column(4,div(tags$h4(" TelekomunikasyonMusteri Kisisel Bilgisi ",style="font-size:small;background-color:#E0F2F7",
hr(),
selectInput(inputId="age",label="Yaş :",choices=levels(musteri$age)),
numericInput(inputId="age_of_line",label="Musteriilik Suresi (Gun):",value=0),
selectInput(inputId="tariff_type",label="Tarife Tipi :",choices=levels(musteri$tariff_type)),
selectInput(inputId="device_type",label="Cihaz Tipi :",choices=levels(musteri$device_type)),
numericInput(inputId="memo_count",label="Aylık Mesajlama Sayısı :",value=0),
numericInput(inputId="callcenter_count",label="Aylık Sikayet Sayısı :",value=0),
selectInput(inputId="last_reload_year",label="Son Yukeleme Miktarı :",choices=levels(musteri$last_reload_year))
)),

```

Şekil 5: Ui.R’da veri girişlerinin tanımlanma kodları

Kullanıcı kişisel ve aylık telekomünikasyon verilerine yönelik değerleri girdikten sonra “hesapla” butonu ile tahmin sonucunu, karar ağacını ve performans değerlendirmek için confusion matrisini görüntülemektedir. Bunun için actionbutton kullanılmıştır. Çıktı değerleri de verbatimTextOutput(“tahminSonucu) ile verilmiştir. Şekil 6’da kullanıcı arayüz tasarımı görülmektedir.

#### KARAR AGACLARI ile MUSTERİ KAYIP ANALIZI

Basarslan Sinan M. 2017. KARAR AGACLARI ile MUSTERİ KAYIP ANALIZI  
Musteri Kayıp Analizi Uygulanması Bir Telekomunikasyon Veri Üzerinde Gerçekleştirilmiştir

Şekil 6: Ui.R sayfa görüntüsü

Kullanıcı arayüzünden girilen değerlerin analiz edilerek çıktı döndürmesi, sonucu tarafındaki “server.R” dosyasındaki komutlarla olur. Kullanıcıdan form olarak istenen bilgiler Şekil 6’da görülmektedir. Kullanıcı bu değerleri doldurduktan sonra Şekil 7’de yer alan Analiz değişkenleri bölümünden eğitim veri seti olarak rasgele ayırma oranını belirler ve hesapla butonuna basar. Hesapla butonuna basması ile beraber kullanıcının girdiği değerler “server.R” dosyasına iletilir. “Server.R” dosyasında oluşturulan eğitimIndisleri() fonksiyonu içinde indislerin rasgele seçimi için createDataPartition() fonksiyonu kullanılmaktadır. Eğitim veri setinin yüzdesel ayrımının arka planda yapılmasını sağlayan kod bloğu şekil 7’de görülmektedir.

```

egtmIndisleri<-reactive({
  #egitim veriseti icin indisler belirlenir.
  set.seed(1)
  egitimIndisleri<-createDataPartition(y=musteri$churn_2013_07,p=as.numeric(input$holdout),list=FALSE)
  egitimIndisleri
})

```

Şekil 7:Eğitim indislerinin seçiminde server.R tarafında çalışan kod

Eđitim veri seti oluřturulduktan sonra müşteri veri seti üzerinde oluřturulan eđitim indisleri çağrılarak C4.5 modelini oluřturmak için server.R'da bulunan ve Őekil 8'da kodu gürülen model() fonksiyonu çağrılır.

```

model<-reactive(
{
  #Eđitim veriseti oluřturulur.
  egitim<-musteri[egtmIndisleri(),]

  #C4.5 algoritması uygulanır, kullanıcının secimine göre oluřturulan modeller deđisir.

  C45_modeli<-J48(churn_2013_07 ~.,data=egitim)

  C45_modeli
}
)

```

Őekil 8: Model() fonksiyonu

Eđitim indisleri kullanılarak model oluřturulduktan sonra müşterilerin kaybına yönelik tahmini model sonucunun, kullanıcının hesapla butonuna bastıktan sonra yazdırılabilmesi için eventReactive() fonksiyonu kullanılır. Kullanıcı girdisi olarak istenen deđerler girildikten sonra bu deđerler ve müşteri ayrılma bilgisini tutan churn niteliđi sırayla yenimusteri deđiřkenlerine atılır. Ancak churn deđeri boş geçilir. C4.5 modelinin hesaplandıđı model() fonksiyonu ve yenimusteri deđiřkenleri predict fonksiyonuna verilerek tahminiSinif deđiřkenine atanır. Elde edilen sonuçlar tahminSonucu() nesnesi ile metinsel ifade olarak döndürülür. Bu iřlemler Őekil 9'da gürölmektedir.

```

tahminSonucu<-eventReactive(
  input$hesapla,{
    #Kullanıcının gireceđi bilgilerin tutulacađı bir deđisken tanımlanır
    #ve girilen deđerler sırasıyla ilgili alanlara atanır.

    yenimusteri<-musteri[0,]
    yenimusteri<-musteri[1,1]<-input$age
    yenimusteri<-musteri[1,2]<-as.numeric(input$age_of_line)
    yenimusteri<-musteri[1,3]<-input$tariff_type
    yenimusteri<-musteri[1,4]<-input$device_type
    yenimusteri<-musteri[1,5]<-input$last_reload_year
    yenimusteri<-musteri[1,6]<-as.numeric(input$mno_count_07)
    yenimusteri<-musteri[1,7]<-as.numeric(input$mno_duration_07)
    yenimusteri<-musteri[1,8]<-as.numeric(input$mmt_count_07)
    yenimusteri<-musteri[1,9]<-as.numeric(input$mmt_duration_07)
    yenimusteri<-musteri[1,10]<-as.numeric(input$mno_non_count_07)
    yenimusteri<-musteri[1,11]<-as.numeric(input$mno_non_duration_07)
    yenimusteri<-musteri[1,12]<-as.numeric(input$mmt_non_count_07)
    yenimusteri<-musteri[1,13]<-as.numeric(input$mmt_non_duration_07)

    yenimusteri<-musteri[1,14]<-as.numeric(input$mno_total_count_07)
    yenimusteri<-musteri[1,15]<-as.numeric(input$mno_total_duration_07)
    yenimusteri<-musteri[1,16]<-as.numeric(input$mmt_total_count_07)
    yenimusteri<-musteri[1,17]<-as.numeric(input$mmt_total_duration_07)

    yenimusteri<-musteri[1,18]<-as.numeric(input$mno_count_07)
    yenimusteri<-musteri[1,19]<-as.numeric(input$callcenter_count_07)

    yenimusteri<-musteri[1,20]<-NA

    #Girilen bilgilere göre yeni ornek için tahmin gercekleştirilir.
    tahminiSinif<-predict(model(),newdata=yeniOrnek[1,-20])

    #Elde edilen sonuc metinsel bir ifade ile doldurulur.
    if(tahminiSinif==E)
      "Musteri ayrılma riski yoktur."
    else
      "Musteri ayrılma riski vardır."
  })

```

Őekil 9: Tahmin sonucunun hesaplanması

Őekil 9'daki gibi tahmin sonucunun eventReactive() fonksiyonu ile döndürüldüđü gibi karar ađacı ve performans hesabını gürsteren kontenjans tablosunun yer aldıđı confusion matrisi oluřturulması kodu Őekil 10'da gürölmektedir. Kontenjans tablosu Őekil 14'de gürölmektedir.



```

kararAgaci<-eventReactive (
  input$hesapla, (#Egitim sonucunda olusturulan karar agaci dondurulur
  model ()
)
)
performansHesabi<-eventReactive (
  input$hesapla, {
    #Test veri seti kullanilarak performans degerlendirilmesi yapilir
    tahminiSiniflar<-predict (model (), newdata=musteri [-egtmIndisleri (), -20])
    cm<-caret::confusionMatrix (data=tahminiSiniflar, reference=musteri [-egtmIndisleri (), 20])
    cm
  }
)
)

```

**Şekil 10:** Karar ağacı, performans hesabı

Bu işlemlerin kullanıcı arayüzünde görülmesini sağlayan kod bloğu da şekil 11’de görülmektedir.

```

#Tahmini sonucu model kararı ekrana yazdırılır
output$karar<-renderText (
  {
    tahminSonucu ()
  }
)
#Karar ağacı ekrana yazdırılır.
output$kararAgaci<-renderPrint (
  {
    kararAgaci ()
  }
)
#Performans degerlendirme sonucu bulgular ekrana yazdırılır.
output$performansDegerlendirme<-renderPrint (
  {
    performansHesabi ()
  }
)
)

```

**Şekil 11:** Sonuçların yazdırılması

Çıktıların ekrana yazdırılması ile uygulama tamamlanmış olur. Kullanıcı arayüzünde veri girme ve ekrana sonuçların yazdırılması ile ilgili görseller sırasıyla şekil 12 ve şekil’13 de görülmektedir.

**Şekil 12:** Kullanıcı ara yüzünde veri girilen kısım





## V. KAYNAKLAR

- [1] Kartal E, Balaban, E., “*M. E. Sınıflandırmaya Dayalı Makine Öğrenmesi Teknikleri Ve Mes Kardiyolojik Risk Değerlendirmesine İlişkin Bir Uygulama*”, Çağlayan Kitapevi, 1. BASKI, 2016
- [2] RStudio, 2015b, *Shiny*, <http://shiny.rstudio.com/>
- [3] Chang, W., Cheng, J., Allaire, J. J., Xie, Y. ve McPherson, J., 2015, *Shiny: Web Application Framework for R*, <http://CRAN.R-project.org/package=shiny>, (*Erişim tarihi: 10 şubat, 2016*).
- [4] RStudio, 2015a, *Home - RStudio*, <http://www.rstudio.com/>
- [5] Wing, M. K. C. From J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T.,.... Scrucca, L.(2015). *Caret: Classification and regression Training*. <http://CRAN.R-project.org.org/package=caret>
- [6] Shinyapps user guide. (2016). *Shinyapps.io user guide*. <http://docs.rstudio.com/getting-started.html> (*Erişim tarihi: 21 şubat, 2016*).
- [7] Shinyapps user guide installation. (2016). *Shinyapps.io user guide*. <http://docs.rstudio.com/shinyapps.io/getting-started.html#installation>, (*Erişim tarihi: 21 şubat, 2016*).
- [8] RStudio, 2015b, *Shiny*, <http://shiny.rstudio.com/>, (*Erişim tarihi: 21 şubat, 2016*).