

# Farklı Sayılarda Ağaç Türetildiğinde ve Çapraz Geçerlilikte “k” Parametresi Değiştirildiğinde Random Forest Performans Sonuçlarının İncelenmesi

## Examination of Random Forest Performance Results Generating Different Number of Trees and Changing “k” Parameter in Cross Validation

Songül ÇINAROĞLU<sup>a</sup>

<sup>a</sup>Sağlık İdaresi Bölümü,  
Hacettepe Üniversitesi  
İktisadi ve İdari Bilimler Fakültesi,  
Ankara

Geliş Tarihi/Received: 11.06.2015  
Kabul Tarihi/Accepted: 27.08.2015

Yazışma Adresi/Correspondence:  
Songül ÇINAROĞLU  
Hacettepe Üniversitesi  
İktisadi ve İdari Bilimler Fakültesi,  
Sağlık İdaresi Bölümü, Ankara,  
TÜRKİYE/TURKEY  
cinaroglus@hacettepe.edu.tr

**ÖZET Amaç:** Makine öğrenmesi yöntemleri içerisinde sınıflama ve regresyon amacıyla çok sayıda ağaç türetilerek oluşturulan yöntemlerden birisi de Random Forest (RF) yöntemidir. Bu yöntemde farklı sayılarda ağaç türetildiğinde bunun sınıflama performans sonuçlarına nasıl yansıdığı literatürde tartışılan konulardan birisidir. Bu nedenle bu çalışmada OECD’ye üye ülkeler sağlık harcamaları bakımından sınıflandırılarak farklı sayılarda ağaç türetildiğinde ve çapraz geçerlilikte “k” parametresi değiştirildiğinde RF performans sonuçlarının gözlenmesi amaçlanmıştır. **Gereç ve Yöntemler:** Oluşturulan veri setinde k-kat çapraz geçerlilik uygulanarak, “k” parametresinin yüksek ( $k \geq 13$ ) ya da düşük ( $k < 13$ ) tutulması durumunda ve RF yönteminde farklı sayılarda (50, 100, 150, 200, 250) ağaç türetildiğinde bunun ROC eğrisi esas alınarak incelenen RF performans sonuçlarında bir farklılık yaratıp yaratmadığı Mann Whitney U testi ile incelenmiştir. **Bulgular:** İncelenen veri seti için farklı sayıda ağaç türetilmesi ve çapraz geçerlilikte “k” parametresinin yüksek ya da düşük tutulmasının Random Forest (RF) performans sonuçlarında istatistiksel olarak anlamlı bir farklılık ortaya koymadığı ( $p > 0.05$ ) görülmüştür. Oluşturulan karar ağacı incelendiğinde sağlık harcamalarının tahmin edilmesinde en fazla bilgi kazancı sağlayan değişkenin ise algılanan sağlık statüsü olduğu görülmüştür. **Sonuç:** Araştırma sonuçları RF yönteminde farklı sayıda ağaç türetilmesinin performans sonuçlarında bir farklılık yaratmadığını ortaya koymaktadır. Bu konu ile ilgili olarak bundan sonra yapılacak araştırmalar için farklı tür ve büyüklüklerde veri setlerinin kullanılması durumunda bunun RF sınıflama performansı üzerinde nasıl bir etkiye bulunduğunun incelenmesi önerilmektedir.

**Anahtar Kelimeler:** Random forest; makina öğrenmesi; veri madenciliği; çapraz geçerlilik; ROC eğrisi

**ABSTRACT Objective:** Random Forest (RF) is one of machine learning techniques which is used for classification and regression with generating number of trees. There is a debate in the literature about how generating different number of trees reflects classification performance of this method. For this reason the aim of this study is to observe RF performance results by generating different number of trees and changing “k” parameter in cross validation while classifying OECD countries according to health expenditures. **Material and Methods:** In this dataset k-fold cross validation was implemented and Mann Whitney U test was used whether there is a difference in RF performance results using AUC when “k” parameter was high ( $k \geq 13$ ) or low ( $k < 13$ ) and while generating different number of trees (50, 100, 150, 200, 250). **Results:** Results of this study shows that generating different number of trees in RF not makes any significant changes ( $p > 0.05$ ) in performance results. It was seen that perceived health status was a variable which has more information gain for predicting health expenditures. **Conclusion:** It is advisable for future studies related with this subject to examine performance results of different datasets which are in different types and sizes.

**Key Words:** Random forest; machine learning; data mining; cross validation; ROC curve

Türkiye Klinikleri J Biostat 2015;7(2):108-18

doi: 10.5336/biostatic.2015-46743

Copyright © 2015 by Türkiye Klinikleri

**M**akine öğrenmesi yöntemleri içerisinde farklı algoritmalar kullanılarak oluşturulan karar ağaçlarından sıklıkla yararlanılmaktadır.<sup>1</sup> Bunlar sınıflama, karar teorisi ve tahminleme gibi amaçlarla

kullanılabilmektedir. Karar ağaçlarını oluşturacak değişkenler ise sürekli ya da kategorik türde olabilmektedir. Karar ağaçları verilerin sınıflandırılmasında, görsel sunum olanağı sunan oldukça etkili sınıflama yöntemlerindedir.<sup>2</sup> Farklı algoritmalar kullanılarak oluşturulan karar ağaçları içerisinde çok sayıda karar ağacının bir araya gelmesi ile meydana gelen bir karar ağacı türü ise Random Forest (RF) dir.<sup>3,4</sup> RF yöntemi ile yalnızca sınıflama yapmak değil, sınıflamanın yanı sıra regresyona dayalı tahminler yapmak da mümkün olabilmektedir.<sup>5</sup>

Karar ağaçlarının oluşturulmasında yaygınlıkla kullanılan yöntemlerden birisi de çapraz geçerlilik yöntemidir. Makina öğrenmesi yöntemleri içerisinde popüler bir yöntem olan çapraz geçerlilikten model geçerliliğinin yükseltilmesinde yararlanılmaktadır.<sup>6</sup> Model geçerliliğinin değerlendirilmesine yönelik olarak her zaman yeni veri elde edilmesi olanaklı olmayabilir. Bu nedenle başka yöntemlerden yararlanmak yoluna gidilebilmektedir. Çapraz geçerlilik bu farklı yöntemlerden birisidir. Çapraz geçerlilik verinin bölünmesine denilmektedir.<sup>7</sup> Bu yöntemde veri setinin rastgele kaç parçaya ayrılacağını belirleyen ise “k” parametresidir. k=2 olduğunda veri rastgele iki parçaya ayrıldıktan sonra parçalardan birisi için geliştirilen ayırma fonksiyonlarında diğer parçanın gözlem değerleri yerine konur ve bu grup için grup üyeliği kestiriminde bulunulur. Sonuçta, ayırma fonksiyonu geliştirilmeyen parçanın gerçek grup üyelikleri ile kestirilen grup üyeliklerinin tutarlılığı bir çapraz tablo aracılığı ile incelenir. Buna veriyi yarıya bölme yöntemi denir.<sup>7</sup> Çapraz geçerlilikte minimum 2 olmak üzere farklı sayıda “k” belirlenerek performans sonuçlarındaki değişim incelenebilmektedir. RF yönteminde de yararlanılan çapraz geçerlilikte, veri setinin kaç parçaya bölüneceğini belirten “k” parametresinin yükseltilmesi durumunda, sınıflama performans sonuçlarının bundan nasıl etkilendiği konusu ise literatürde tartışılmalı konulardan birisidir.<sup>8</sup>

RF sınıflama yönteminin çalışma prensibi, her bir niteliğin bir “düğüm” tarafından temsil edildiği en son yapının “yaprak” en üst yapının “kök”, bunların arasında kalan yapıların ise “dal” olarak adlandırıldığı bir çalışma sistemi vardır. Bu yöntem

kullanılarak sınıflama yapılabilmekte, regresyon ağaçları oluşturulabilmektedir. Veri setinde kategorik bir sınıf değişkeni söz konusu olduğunda sınıflama yapılmakta, sürekli sayısal bir değişken olduğunda ise regresyon yöntemi uygulanmaktadır.<sup>9</sup>

RF yönteminde ağaçların oluşturulmasında kullanılan algoritma CART (Classification and Regression Tree) algoritmasıdır.<sup>4</sup> CART algoritması veri setinin hangi değişkenden başlayarak dallara ayrılacağına “Bilgi Kazancı”ni kullanarak karar veren bir algoritmadır. Bu algoritma Gini’ye dayalı ikili bölme işlemine dayanmaktadır ve dallara ayrılmak için seçilen değişkenin uygun test kriteri “Gini katsayısı” ile belirlenmektedir.<sup>10</sup> CART karar ağacı, ikili olarak özyinelemeli biçimde bölünen bir yapıya sahiptir.<sup>11</sup>

Gini yöntemi, m. değişkenden dallara ayrılma gerçekleşmeden önce ve alt dallara ayrıldıktan sonra bölünen veri için Gini değerlerinin hesaplandığı bir yöntemdir. Buna göre bölünme olmadan önceki verinin Gini değerleri ile bölünme olduktan sonraki verinin Gini değerleri arasındaki fark alınmaktadır. Ormanda m. değişken kullanılarak oluşturulan her bir ağaç için bölünme olmadan önceki Gini değerleri ile bölünme olduktan sonraki Gini değerleri arasındaki farklar hesaplanarak tüm ağaçlar arasındaki farklar toplanmaktadır. Bulunan değer m. değişkenin Gini önem derecesini vermektedir. Bu yöntemle elde edilen değişken önem derecesi, çoğu zaman standart yöntem ile bulunan değişken önem derecesine paralel bir değer olmaktadır.<sup>3</sup>

RF yönteminin karar ağacı algoritmaları içerisinde tercih edilen ve yaygınlıkla kullanılan bir yöntem olması ve üstün yönleri arasında; hızlı bir sınıflama yöntemi olması, veri setindeki değişikliklere kolay adapte olabilen ve sınıflama doğruluğunu garanti eden bir yöntem olması, bu yöntemle istenilen sayıda ağaç oluşturulabilmesi bulunmaktadır.<sup>4</sup> RF yönteminin dezavantajlı yönleri arasında ise, çok fazla sayıda ağaç barındırdığından sınıflamanın doğruluğuna ilişkin bir güven aralığı vermemesi bulunmaktadır.<sup>9</sup>

Sınıflamaya dayanan problemlerin çözümünde diğer veri madenciliği algoritmaları ile karşılaştırı-

rıldığına oldukça avantajlı yönleri bulunan RF yönteminde sınıflama performansının değerlendirilmesinde farklı ölçütlerden yararlanılması mümkün olmakla birlikte bu yöntemlerden birisi de ROC eğrisi altında kalan alandır.<sup>12</sup> ROC eğrisi bir testin ayırt etme yeteneğinin incelendiği bir performans ölçüdür. Bu eğri tıp alanında bir hastalığa tanı koymada kullanılan tanı testlerinin performanslarının değerlendirilmesi ve tanı koyabilmek için kesim noktasının belirlenmesi amacıyla kullanılmaktadır. ROC eğrisi incelendiğinde bu eğrinin dikey ekseninde doğru pozitif (duyarlılık), yatay ekseninde ise yanlış pozitif (seçicilik) değerlerinin yer aldığı görülmektedir.<sup>7</sup>

ROC eğrisinin temel performans ölçüsü ise eğri altında kalan alandır. ROC eğrisi altında kalan alanın alabileceği en büyük değer 1'dir. Eğri altında kalan alanın 1'e eşit olması, testin bireyleri kesin olarak ayırt ettiğini göstermektedir. Eğri altında kalan alan 1'e yaklaştıkça, yeni testin ayırt ediciliğinin arttığı, 0.50'ye yaklaştıkça testin ayırt edici olmadığı söylenebilmektedir.<sup>13</sup> Sonuç olarak bir tanı testinin ayırt etme yeteneği ROC eğrisi üzerinde çok iyi açıklanabilmektedir. Tanı testi ne kadar iyi ise eğri o kadar yukarıya (yüksek duyarlılık bölgesi) ve sola (düşük yalancı pozitif oranı bölgesi) doğru kayar. Bir tanı testi için ROC eğrisi altında kalan alan, etkinlik düzeyine bağlı olarak 0.50 ile 1.00 arasında değişen değerler alır. Bu değer büyüdükçe tanı testinin ayırt etme yeteneği artmaktadır.<sup>14</sup>

Veri madenciliğinde sınıflama problemlerinin çözümünde kullanılan, çok sayıda karar ağacının bir araya gelmesi ile oluşan, sınıflama performansının değerlendirilmesinde çoğunlukla ROC eğrisi altında kalan alanın kullanıldığı RF yönteminde, ağaç sayısının arttırılmasının performans sonuçlarına nasıl yansıdığı konusu literatürde tartışmalı bir konu olmaya devam etmektedir. Bu konuda yapılan araştırmaların bir kısmında ağaç sayısındaki artışın performans artışı sağlayacağı belirtilmektedir. Bununla ilgili olarak Oshiro ve ark. tarafından yapılan bir araştırmada sağlık alanında farklı veri setleri kullanılmıştır.<sup>15</sup> Araştırma sonucunda ağaç sayısını arttırmanın performans sonuçlarını her zaman yükseltmediği bulunmuştur. Bu araştırmada farklı "k" parametrelerine göre çapraz geçerlilik uygulan-

dığında ve RF yönteminde farklı sayılarda ağaç türetildiğinde bunun RF yönteminde yaygınlıkla kullanılan bir performans ölçüsü olan ROC eğrisi altında kalan performans sonuçlarında anlamlı bir farklılık yaratıp yaratmadığının incelenmesi amaçlanmıştır. Bu amaçla OECD'ye üye ülkelerde sağlık harcamalarını tahmin edici değişkenler belirlenecek ve sağlık harcamalarının tahmin edilmesinde en fazla bilgi kazancı sağlayan değişkenleri belirlemeye yönelik bir karar ağacı oluşturulacaktır.

## GEREÇ VE YÖNTEMLER

### 2.1. VERİ SETİ

Araştırmada veri seti olarak OECD'ye üye olan 34 ülkenin 2011 yılı itibariyle toplam sağlık harcamaları ile bu ülkelerde sağlık harcamalarını belirlediği düşünülen değişkenler kullanılmıştır. Araştırma değişkenlerinin seçilmesinde OECD ülkelerinde sağlık ile ilgili göstergelerin kayıt altına alındığı istatistik kayıtlarından yararlanılmıştır.<sup>16</sup> Bu kayıtlarda, sağlık ile ilgili gösterge grupları sağlık statüsü, doğrudan sağlık ile ilgili olmayan göstergeler, sağlık ile ilgili kaynaklar, sağlık hizmetlerinin kullanımı, sağlıkta kalite göstergeleri, demografik göstergeler, ekonomik göstergeler ve sağlık harcamaları olmak üzere sekiz başlık altında toplanmıştır. Bu araştırmada ise OECD ülkelerinin sağlık harcamaları bakımından sınıflandırılmasında kullanılacak olan bağımlı değişken sağlık harcamaları olarak belirlenmiştir. OECD'ye üye ülkelerin sağlık harcamaları bakımından sınıflandırılmasında belirleyici değişkenlerin seçiminde 2011 yılı itibariyle tüm gösterge gruplarında tam veriye sahip olan değişkenler seçildikten sonra tahmin edici değişkenler arasında çoklu bağlantı sorununa yol açacak derecede yüksek korelasyonun varlığı incelenmiş ve yüksek korelasyona sahip olan değişkenler elenmiştir. Analizlerde kullanılmak üzere yapılan ön inceleme sonrasında bu araştırmada kullanılmak üzere yer verilen tamamı sayısal türde olan değişkenler ve bu değişkenlerin ölçüm birimi ile ilgili bilgiler Tablo 1'de görülebilmektedir.

### VERİLERİN ANALİZİ

Verilerin analizine tanımlayıcı bilgilerin elde edilmesi ile başlanmıştır. Daha sonra bu araştırma için

**TABLO 1:** Araştırma kapsamında kullanılan değişkenlere ait açıklamalar.

Değişken	Değişkenin Türü	Açıklama	Ölçüm Birimi	Değişken Türü	Kısaltma	Yıl
Doğuşta Beklenen Yaşam Süresi	Sağlık Statüsü	Ülke Genelinde Tüm Nüfus için Doğumda Beklenen Yaşam Süresi	Yıl	Sürekli Sayısal	Dog_Yas_Sur	2011
Algılanan Sağlık Statüsü	Sağlık Statüsü	15 Yaş Üzeri Nüfusta Sağlık Statüsünün Çok İyi ya da İyi Olarak Algılanmaların Yüzdesi	Oran (%)	Sürekli Sayısal	Alg_Sag_Stat	2011
Hekim Sayısı	Sağlıkla İlgili İnsan Kaynakları	Aktif Olarak Çalışan Toplam Hekim Sayısı	Sayı	Kesikli Sayısal	Hek_Say	2011
Hastane Sayısı	Sağlık ile İlgili Fiziki Kaynaklar	Ülke Genelinde Toplam Hastane Sayısı	Sayı	Kesikli Sayısal	Hast_Say	2011
Hastane Hizmetlerinin Kullanımı	Sağlık Hizmetlerinin Kullanımı	Taburcu Olan Hasta Sayısı	Sayı	Sürekli Sayısal	Hast_Kul	2011
Bağışıklama	Sağlık Hizmetlerinin Kullanımı	15 Yaş Üzerinde Difteri, Tetanoz ve Boğmaca Nedeniyle Aşılanan Çocuk Yüzdesi	Oran (%)	Sürekli Sayısal	Bağışıklama	2011
Alkol Tüketimi	Sağlıkla İlgili Diğer Göstergeler	15 Yaş Üzerinde Kişi Başı Tüketilen Alkol Miktarı	Litre	Sürekli Sayısal	Alk_Tuk	2011
Milli Gelir	Ekonomik Göstergeler	Satın Alma Paritesine Göre (\$) Kişi Başına Düşen Milli Gelir	Satın Alma Gücü Paritesi (\$)	Sürekli Sayısal	MG	2011
Sağlık Harcaması	Ekonomik Göstergeler	Satın Alma Gücü Paritesine Göre (\$) Kişi Başı Sağlık Harcaması	Satın Alma Gücü Paritesi (\$)	Sürekli Sayısal	SH	2011

OECD, Health at a Glance 2013 OECD Indicators, 2013.

sağlık harcamalarının tahmin edilmesinde kullanılan değişkenler arasındaki ilişkiler korelasyon analizi ile incelenmiştir. Bu çalışmada OECD'ye üye ülkelerin sağlık harcamaları bakımından sınıflandırılmasında bir makina öğrenmesi yöntemi olan RF yöntemi kullanılmıştır. RF bir sınıflama yöntemi olarak kullanıldığında bağımlı değişkenin kategorik türde olması gerekmektedir. Bu nedenle verilerin analize hazır hale getirilmesi aşamasında sürekli sayısal bir değişken olan ülke genelinde toplam sağlık harcaması değişkeni iki sınıflı bir kategorik değişken haline dönüştürülmüştür. Değişkenler arasındaki ilişkiler incelendikten ve RF sınıflama yöntemi için analize hazır hale getirildikten sonra bir istatistik ve veri madenciliği yazılım programı olan Orange programına aktarılmıştır. Bu araştırma kapsamında RF yönteminde ağaç sayısının artırılması durumunda bunun performans sonuçlarına nasıl yansıdığına incelenmesi için, sırasıyla 50, 100, 150, 200 ve 250 ağaç türetilmiştir. RF yönteminde CART (Classification and Regression Tree) algoritması kullanılarak ağaç oluşturulmaktadır.

Bu yöntemde her bir düğümde iki adet dal bulunmaktadır ve bu algoritma kullanılarak hedef değişkenin sürekli olması durumunda regresyon, kategorik olması durumunda ise sınıflama yapılabilmektedir. Bu algorithmda veri setinin dallara ayrılmasında “bilgi kazancı”, dallar ayırma kriteri olarak ise “Gini Katsayısı” kullanılmaktadır.<sup>16</sup>

Bu çalışmada farklı sayıda ağaç türetildiğinde bunun performans sonuçları üzerindeki etkisinin incelenmesinde ROC eğrisi altında kalan alan bir performans ölçüsü olarak dikkate alınmıştır. Farklı sayıda ağaç türetildiğinde ROC eğrisi altında kalan alan esas alınarak değerlendirilen performans sonuçlarının nasıl değiştiği, makina öğrenmesi yöntemleri içerisinde bir model geçerliliği ölçüsü olan k-katlı çapraz geçerlilik uygulanarak ve farklı sayılarda “k” parametresi belirlenerek gözlemlenmiştir. Bu çalışmada sınıflama performansının belirlenmesinde ROC eğrisi altında kalan alan kullanılmıştır. Performans ölçütü olarak ROC eğrisi altında kalan alanın tercih edilmesinin nedeni, genel sınıflama performansının değerlendirilmesi için en uygun sınıflama performans ölçütü olarak kabul edilmektedir.<sup>17,18</sup>

Bu çalışmada bir optimizasyon yöntemi olarak kullanılan k-parça çapraz geçerlilik algoritması incelendiğinde, bu yöntemde her defasında bir gözlem dışarıda bırakılarak geride kalan veriler ile modelin oluşturulduğu görülmektedir. Bu yöntem tahmin hatasının minimize edildiği bir optimizasyon yöntemidir ve çapraz geçerlilik algoritması şu şekilde ifade edilmektedir.<sup>19</sup>

$$\text{ÇG}(\lambda) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}^{-i})^2$$

$$\hat{y}^{-i} = x_i^T \hat{\beta}^{-i}(\lambda)$$

min ÇG( $\lambda$ )'dir.

Burada  $\hat{\beta}^{-i}(\lambda)$ ,  $(x_i, y_i)$  dışında diğer gözlemlere dayanan model tahminidir,  $\lambda$  ise model seçiminde ayar parametresi olarak kullanılmaktadır.

Yapılan gözlemlerde istatistiksel testlerin uygulanmasında genel bir kural olan gözlem sayısının değişken sayısının en az 5 katı kadar olması kuralının bir gereği olarak 5 farklı sayıda ağaç için en az 2, en fazla 26 kat çapraz geçerlilik uygulama sonuçları kaydedilmiştir.

Verilerin analizinde ayrıca bu araştırma kapsamında OECD ülkelerinin toplam sağlık harcamaları bakımından sınıflandırılmasında ön plana çıkan değişkenlerin belirlenmesi amacıyla modelde en fazla bilgi kazancını sağlayan değişkenlerin görsel sunumuna imkan veren karar ağacından yararlanılmıştır. "k" parametresinin yüksek ya da düşük tutulması durumunda RF yönteminde farklı sayıda ağaç türetilerek oluşturulan ROC eğrisi altında kalan alana ait performans sonuçlarının anlamlı bir

farklılık gösterip göstermediğinin incelenmesinde ise iki ortalama arasındaki farkın önemlilik testinin non-parametrik karşılığı olan Mann-Whitney U testi kullanılmıştır.

## BULGULAR

### TANIMLAYICI BULGULAR

Araştırmada OECD ülkelerinin RF yöntemi kullanılarak sağlık harcamaları bakımından sınıflandırılmasında kullanılan değişkenlere ait bilgiler Tablo 2'de sunulmuştur. Buna göre sağlık harcamalarının tahmin edilmesinde kullanılan 34 ülkeye ait tahmin edici değişkenler içerisinde doğuştan beklenen yaşam süresinin ortalaması 80,15(±2,46), algılanan sağlık statüsünün ortalama 66,94 (±11,51), hekim sayısının ortalama 113.310,32 (±143.735,72), hastane sayısı ortalama 1280,65 (±1862,56), hastane hizmetlerinin kullanımı 4664635,68 (±4536116,94), bağışıklama ortalama 96,06 (±3,20), alkol tüketimi ortalama 9,35 (±2,11), Milli Gelir ortalama 35436,06 (±13802,03) ve sağlık harcaması ortalama 3394,44 (±1597,56)'dir (Tablo 2).

### BAĞIMSIZ DEĞİŞKENLER ARASINDA ÇOKLU BAĞLANTININ İNCELENMESİ

Sınıflandırma problemlerinin çözümünde hedef değişkene ait değerlerin tahmin edilmesi amacıyla oluşturulan modellerde bağımsız değişkenler arasında çoklu bağlantı sorununun varlığına işaret edecek derecede yüksek bir korelasyonun olmaması beklenmektedir. Ayrıca değişkenler arasındaki ilişkilerin incelenmesinde verilerin bir ya da daha fazlasının normal dağılım göstermemesi du-

**TABLO 2:** Araştırma değişkenlerine ilişkin tanımlayıcı bilgiler.

	Değişken	N	Min.	Mak.	Ort.	SS.	
Tahmin Edici Değişkenler	Doğuştan Beklenen Yaşam Süresi	34	74	83	80,15	2,46	
	Algılanan Sağlık Statüsü	34	34	88	66,94	11,51	
	Hekim Sayısı	34	1121	809492	113310,32	143735,72	
	Hastane Sayısı	34	8	8605	1280,65	1862,56	
	Hastane Hizmetlerinin Kullanımı	34	78704	19868738	4664635,68	4536116,94	
	Bağışıklama	34	83	99	96,06	3,20	
	Alkol Tüketimi	34	2	12	9,35	2,11	
	Milli Gelir	34	16984	88781	35436,06	13802,03	
	Bağımlı Değişken	Sağlık Harcamaları	34	937	8483	3394,44	1597,56

**TABLO 3:** Tahmin edici değişkenler arasındaki ilişkileri gösteren korelasyon matrisi.

Değişkenler	$r_s$	Doğuşta Beklenen Yaşam Süresi	Hekim Sayısı	Hastane Sayısı	Hastane Hizmetlerinin Kullanımı	Alkol Tüketimi	Milli Gelir	Algılanan	
								Sağlık Statüsü	Bağışıklama
Doğuşta Beklenen Yaşam Süresi	$r_s$	1							
Hekim Sayısı	$r_s$	-0,142	1						
Hastane Sayısı	$r_s$	0,128	0,507**	1					
Hastane Hizmetlerinin Kullanımı	$r_s$	0,156	0,453**	0,497**	1				
Alkol Tüketimi	$r_s$	-0,174	0,226	-0,270	-0,189	1			
Milli Gelir	$r_s$	0,528**	-0,096	-0,017	-0,156	0,027	1		
Algılanan Sağlık Statüsü	$r_s$	0,586**	-0,329	-0,074	-0,083	-0,198	0,614**	1	
Bağışıklama	$r_s$	-0,119	-0,083	0,115	0,120	-0,051	-0,303	-0,408*	1

$r_s$ : Spearman Korelasyon Katsayısı; \*  $p < 0,05$ ; \*\*  $p < 0,01$ .

rumunda Spearman korelasyon katsayısının kullanımı önerilmektedir. Bu çalışmada ise doğuşta beklenen yaşam süresi değişkeni ile milli gelir değişkenlerinin dağılımları normal dağılıma uygun olmadığından bu çalışmada bağımsız değişkenler arasındaki ilişkiler Spearman korelasyon katsayısı kullanılarak incelenmiş ve sonuçlar Tablo 3'de sunulmuştur (Tablo 3).

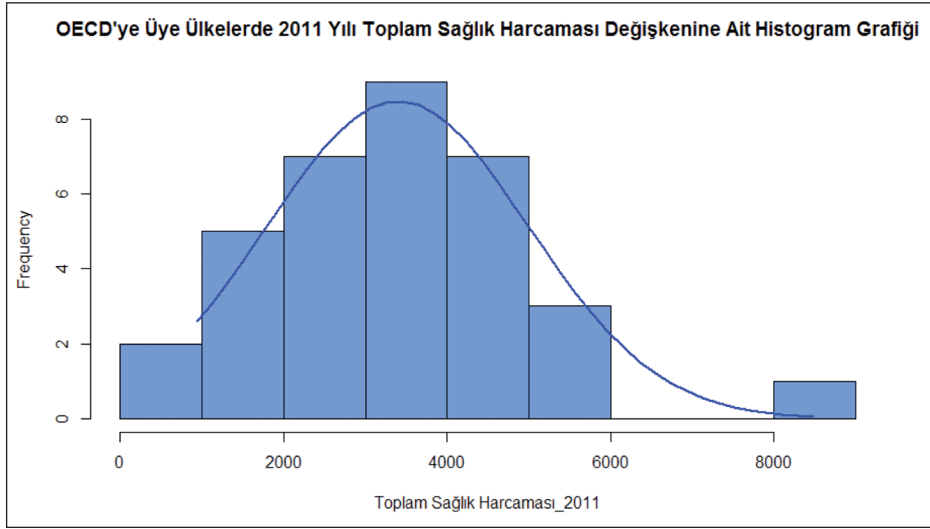
Tablo 3'de gösterilen ve Spearman korelasyon katsayılarının kullanıldığı korelasyon tablosu incelendiğinde bağımsız değişkenler arasında çoklu bağlantı sorununun varlığına işaret edecek derecede çok güçlü korelasyonların ( $r_s \geq 0.70$ ) mevcut olmadığı görülmektedir. Bu durum bu araştırmada bu bağımsız değişkenlere yer verilebileceğini göstermektedir. Bağımsız değişkenler arasındaki ilişkide çoklu bağlantının varlığı araştırıldıktan sonraki aşamada, bu araştırmada hedef değişken olarak belirlenen ve bir sürekli sayısal değişken olan sağlık harcaması değişkeninin kategorik forma dönüştürülmesi gerçekleştirilmiştir.

### HEDEF DEĞİŞKENİN KATEGORİK FORMA DÖNÜŞTÜRÜLMESİ

Araştırmada bağımlı değişken olarak kullanılacak olan sağlık harcaması sürekli sayısal bir değişken olduğundan bu değişkenin kategorik hale getirilmesi amacıyla bir kesim noktasının belirlenmesine ihtiyaç duyulmuştur. Sınıflama problemlerinde hedef değişkenin az sayıda kategoriden oluşması durumunda performans sonuçlarının daha yüksek olacağı hakkında genel bir görüş mevcut olmakla birlikte, bu araştırmada 34 OECD ülkesi içerisinde

toplam sağlık harcaması ortalamasının üzerinde olan ve altında bulunan ülkelerden iki grup oluşturulması düşünülmüştür. Bu amaçla sürekli sayısal bir değişken olan toplam sağlık harcaması değişkeninin dağılımı incelenmiştir. Literatürde sağlık harcamalarının dağılımını inceleyen sağlık ekonomisi alanında uzman araştırmacılar bu harcamaların genellikle aşırı derecede sağa çarpık özellik gösterdiklerini belirtmektedirler.<sup>20</sup> Sağlık harcamalarının bu aşırı sağa çarpıklık özelliği ile başa çıkabilmek amacıyla logaritmik dönüşüm uygulandığında genellikle dağılımın normale yakın bir görünüm sergilediği görülmektedir.<sup>21</sup> Bu bilgilerden yola çıkarak bu araştırmada bağımsız değişken olan sağlık harcaması değişkeninin dağılımı incelenmiştir (Şekil 1).

Şekil 1'de görüldüğü gibi sağlık harcaması değişkeninin dağılımının normal dağılıma yakın bir dağılım göstermesi nedeniyle bu değişkene ait değerlerin kategorik forma dönüştürülmesinde ortalama değerlerin bir kesim noktası olarak kullanılabilmesine karar verilmiş ve dönüşüm uygulamaya gerek duyulmamıştır. Bu bilgi doğrultusunda 2011 yılı için ortalama toplam sağlık harcaması 3394,44 \$'dan daha az olan ülkeler 1, bu harcama tutarına eşit ve bundan daha fazla miktarda sağlık harcamasına sahip olan ülkeler 2 olarak kodlanmıştır. Hedef değişkenin kategorik forma dönüştürülmesinden sonraki aşamada ise OECD'ye üye ülkelerin sağlık harcamaları bakımından, RF yöntemine göre farklı sayıda ağaç türetilerek sınıflandırılmasına geçilmiştir.



ŞEKİL 1: Toplam sağlık harcaması değişkeni histogram grafiği.

### RF YÖNTEMİNDE FARKLI SAYILARDA AĞAÇ TÜRETİLDİĞİNDE VE k-KAT ÇAPRAZ GEÇERLİLİKTE “k” PARAMETRESİ YÜKSELTİLDİĞİNDE ROC EĞRİSİ PERFORMANS SONUÇLARI

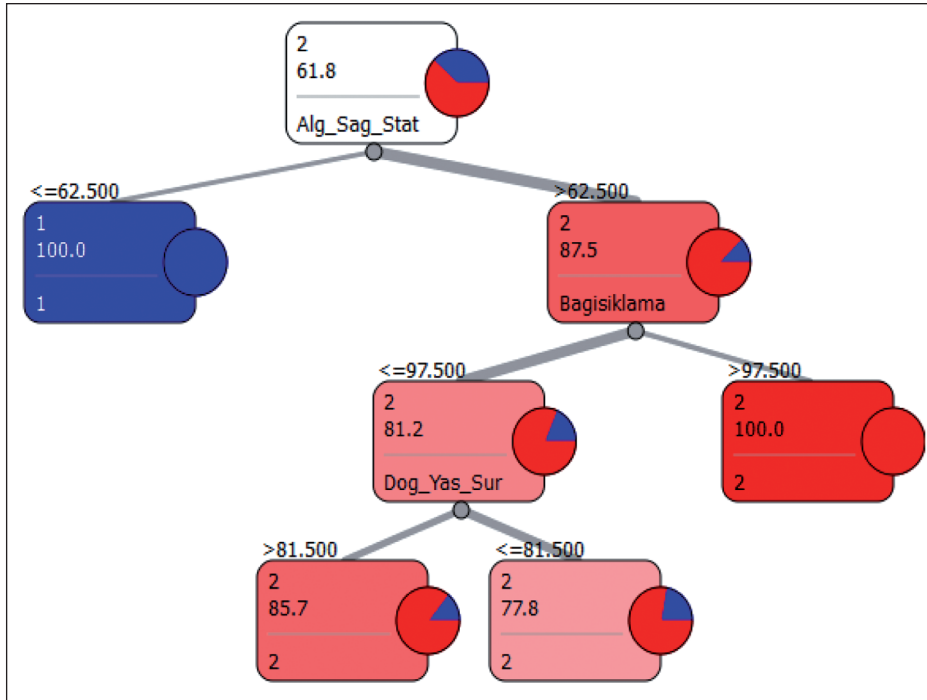
Tablo 4’de 50, 100, 150, 200 ve 250 ağaç türetilerek RF yönteminde oluşturulan karar ağaçlarının 2 ile 26 arasında değişen “k” parametrelerine göre çapraz geçerlilik uygulandığında ROC eğrisi performans sonuçları görülmektedir. Bu araştırmada RF yöntemine göre farklı sayılarda ağaç türetildiğinde ve k-kat çapraz geçerlilikte “k” parametresi yükseltildiğinde bunun ROC performans sonuçları üzerinde istatistiksel olarak anlamlı bir farklılık yaratıp yaratmadığının test edilebilmesi amacıyla, istatistiksel testlerin uygulanmasında genel bir kural olarak gözlem sayısının değişken sayısının en az 5 katı kadar olması gerektiği yaklaşımı<sup>7</sup> esas alınmıştır. Bu nedenle 5 farklı sayıda ağaç türetildiğinde elde edilen sonuçlar k=2’den k=26’ya kadar 25 farklı “k” parametresine göre incelenmiştir. Tablo 4’de sunulan performans sonuçları incelendiğinde ve ROC eğrisi altında kalan alanın 1’e yaklaşması durumunda sınıflama performansının oldukça yüksek olduğu göz önünde bulundurulduğunda, bu araştırmada elde edilen RF performans sonuçlarının (ROC Eğrisi Altında Kalan Alan > 0,85) son derece yüksek olduğu görülmektedir. Bu durum OECD’ye üye olan 34 ülkenin sağlık harcamaları bakımından sınıflandırılmasında dikkate alınan değişkenlerin sınıflama

performansının son derece yüksek olduğunu ortaya koymaktadır (Tablo 4).

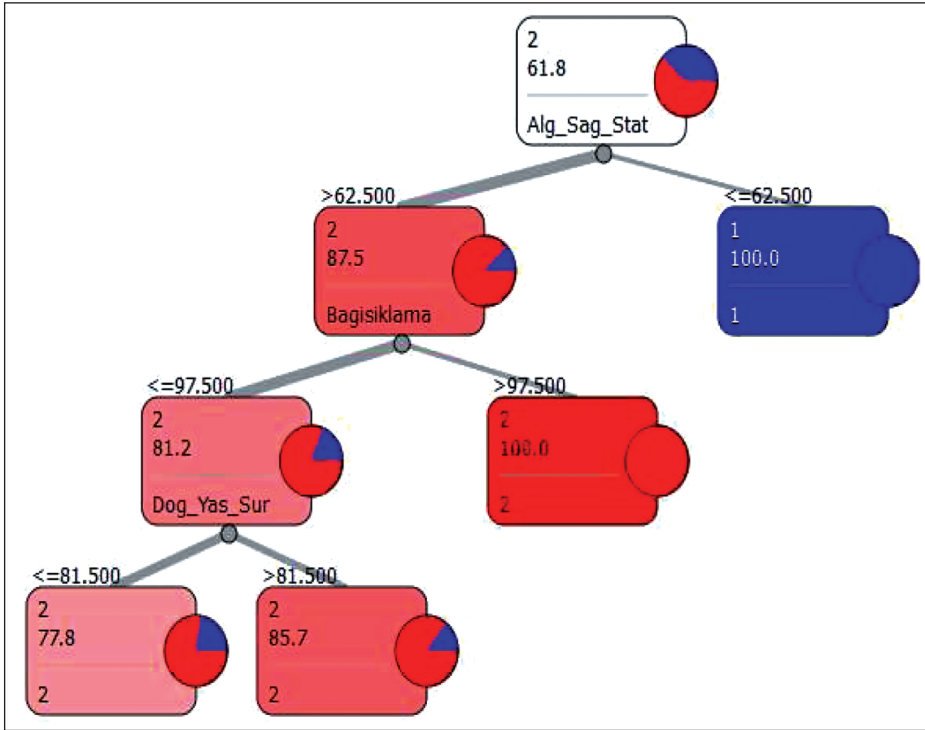
Tablo 4 incelendiğinde, RF yönteminde farklı sayılarda ağaç türetildiğinde ve k-kat çapraz geçerlilikte “k” parametresi yükseltildiğinde ROC eğrisi altında kalan alan performans sonuçları incelendiğinde k=13 kat ve k=15 kat çapraz geçerlilik uygulandığında en yüksek sınıflama performanslarının (ROC Eğrisi altında kalan alan=1) elde edildiği görülmektedir. Bu doğrultuda OECD’ye üye ülkelerin sağlık harcamaları bakımından sınıflandırılmasında oldukça yüksek performans gösteren bu sınıflamada sağlık harcamalarının öngörülmesinde hangi değişkenlerin ön plana çıktığının belirlenmesine ve bu nedenle karar ağaçlarının incelenmesine gerek duyulmuştur. Farklı sayılarda ağaç türetilerek oluşturulan RF yöntemi sonuçları içerisinde 250 ağaç türetilerek oluşturulan RF sonuçları değerlendirilmiştir. Bunun nedeni literatürde genel bir yaklaşım olarak kabul edilen ağaç sayısını artırılması durumunda performans sonuçlarının yükseleceği bilgisidir. Buna göre, 250 ağaç türetilerek ve ROC eğrisi performansı en yüksek olan sırasıyla k=13 ile k=15 parça çapraz geçerlilik uygulanarak elde edilen karar ağaçlarının incelenmesine karar verilmiştir. Şekil 2 ve Şekil 3’de sırasıyla k=13 ve k=15 katlı çapraz geçerlilik uygulandığında ve RF yöntemine göre 250 ağaç türetildiğinde oluşan karar ağaçları görülmektedir (Şekil 2, Şekil 3).

**TABLO 4:** RF yönteminde farklı sayılarda ağaç türetildiğinde ve çapraz geçerlilikte "k" parametresi arttırıldığında ROC eğrisi altında kalan alana ilişkin performans sonuçları.

Çapraz Geçerlilikte Kullanılan "k" Parametresi	RF Yöntemine Göre Türetilen Ağaç Sayısı				
	RF_50_Ağaç	RF_100_Ağaç	RF_150_Ağaç	RF_200_Ağaç	RF_250_Ağaç
k=2	0,9167	0,9444	0,9444	0,9375	0,9444
k=3	0,9111	0,9111	0,9204	0,9111	0,9111
k=4	0,9187	0,9187	0,9187	0,9187	0,9187
k=5	0,9833	0,9833	0,9833	0,9833	0,9833
k=6	0,9537	0,9537	0,9537	0,9352	0,9537
k=7	0,9524	0,9286	0,9524	0,9524	0,9524
k=8	0,9792	0,9792	0,9792	0,9792	0,9583
k=9	0,9167	0,9167	0,9167	0,9444	0,9444
k=10	0,975	0,975	0,975	1	0,975
k=11	0,9545	0,9545	0,9545	0,9545	0,9545
k=12	0,875	0,875	0,875	0,8755	0,875
k=13	1	1	1	1	1
k=14	0,9643	0,9643	0,9643	0,9643	0,9643
k=15	1	1	1	1	1
k=16	0,9688	0,9688	0,9688	0,9688	0,9688
k=17	0,9412	0,9412	0,9412	0,9412	0,9412
k=18	0,9654	0,9481	0,955	0,9585	0,955
k=19	0,9273	0,9377	0,9446	0,9481	0,9343
k=20	0,9481	0,9377	0,9412	0,9377	0,9343
k=21	0,9516	0,9377	0,9446	0,9446	0,9308
k=22	0,9308	0,9308	0,9308	0,9273	0,9273
k=23	0,955	0,9412	0,9481	0,9412	0,9412
k=24	0,9343	0,9308	0,9273	0,9343	0,9343
k=25	0,9343	0,9377	0,9273	0,9273	0,9239
k=26	0,9343	0,9343	0,9308	0,9273	0,9273

**ŞEKİL 2:** RF yöntemine göre 250 ağaç türetildiğinde ve k=13 parça çapraz geçerlilik uygulandığında sağlık harcamalarının tahmin edici değişkenler.





ŞEKİL 3: RF yöntemine göre 250 ağaç türetildiğinde ve k=15 parça çapraz geçerlilik uygulandığında sağlık harcamalarının tahmin edici değişkenler.

Koyu renklerin gruplarda daha fazla sayıda ülkenin bulunduğu duruma işaret ettiği Şekil 2 ve 3'de RF yöntemine göre 250 ağaç türetildiğinde ve k=13 parça çapraz geçerlilik uygulandığında sağlık harcamalarının tahmin edilmesinde en fazla bilgi kazancı sağlayan değişkeninin algılanan sağlık statüsü olduğu görülmektedir. Bu sonuca göre Şekil 2'de OECD'ye üye olan 34 ülke 4 gruba ayrılmaktadır ve bu dört grubu oluşturan ülkelerin OECD'ye üye olan diğer ülkeler içerisinde toplam sağlık harcamaları ortalamasının daha yüksek olduğu, 2 ile kodlanmış olan ülkeleri temsil ettikleri, görülmektedir. Bu gruplardan birincisinde; algılanan sağlık statüsü  $\leq 62,5$  olan ülkeler yer alırken, ikinci grupta algılanan sağlık statüsü  $> 62,5$  ve bağışıklama  $>97,5$  olan ülkeler bulunmaktadır. Üçüncü grupta algılanan sağlık statüsü  $> 62,5$ , bağışıklama  $\leq 97,5$  ve doğumda beklenen yaşam süresi  $>81,5$  olan ülkeler yer almaktadır. Dördüncü ve son grup ise algılanan sağlık statüsü  $> 62,5$ , bağışıklama  $\leq 97,5$  ve doğumda beklenen yaşam süresi  $\leq 81,5$  olan ülkelerin oldukları görülmektedir.

Şekil 3'de ise OECD'ye üye olan ülkeler içerisinde toplam sağlık harcaması ortalaması diğer ül-

kelere göre daha düşük olan 1 ile kodlanmış olan ülkeler görülmektedir. Bu grupların birincisinde algılanan sağlık statüsü  $\leq 62,5$  olan ülkeler yer alırken, ikinci grupta algılanan sağlık statüsü  $> 62,5$ , bağışıklama  $\leq 97,5$  ve doğumda beklenen yaşam süresi  $\leq 81,5$  olan ülkeler bulunmaktadır. Üçüncü grupta algılanan sağlık statüsü  $> 62,5$ , bağışıklama  $\leq 97,5$  ve doğumda beklenen yaşam süresi  $> 81,5$  olan ülkeler yer alırken, dördüncü ve son grupta ise algılanan sağlık statüsü  $> 62,5$  ve bağışıklama  $> 97,5$  olan ülkelerin buldukları görülmektedir. RF yöntemine göre 250 ağaç içerisinde en yüksek performans gösteren ağaca ait grafikleri temsil eden Şekil 2 ile 3 incelendiğinde OECD'ye üye ülkelerin toplam sağlık harcamaları bakımından sınıflandırılmasında ön plana çıkan değişkenlerin başta algılanan sağlık statüsü olmak üzere, bağışıklama ile doğumda beklenen yaşam süresi olduğu görülmektedir.

Tablo 5'de RF yöntemine göre farklı sayılarda ağaç türetildiğinde ve k parça çapraz geçerlilikte "k" parametresi yüksek ( $k \geq 13$ ) ya da düşük ( $k < 13$ ) olduğunda ROC eğrisi altında kalan alana ait performans sonuçlarının istatistiksel olarak anlamlı bir farklılık gösterip göstermediği Mann Whitney U

**TABLO 5:** Farklı k-kat çapraz geçerliliklere göre farklı sayılarda ağaç türetildiğinde RF performans sonuçlarının karşılaştırılması.

RF Ağaç Sayısı	k <13		k ≥ 13		U	p
	N	Sıra No	N	Sıra No		
RF_50_Ağaç	11	11,45	14	14,21	60	0,35
RF_100_Ağaç	11	11,91	14	13,86	65	0,51
RF_150_Ağaç	11	12,27	14	13,57	69	0,66
RF_200_Ağaç	11	11,20	14	13,43	57	0,44
RF_250_Ağaç	11	13,18	14	12,86	75	0,91

testi ile incelenmiştir. Bu sonuçlara göre k-kat çapraz geçerlilikte "k"nın yüksek ya da düşük tutulmasına göre RF yönteminde sırasıyla 50 (U=60, p>0.05), 100 (U=65, p>0.05), 150 (U=69, p>0.05), 200 (U=57, p>0.05) ya da 250 (U=75, p>0.05) sayıda ağaç türetilerek elde edilen ROC eğrisi altında kalan alana ait performans sonuçlarının istatistiksel olarak anlamlı bir farklılık göstermediği (p>0.05) bulunmuştur (Tablo 5).

## TARTIŞMA

Karar ağaçlarından sınıflamaya dayanan problemlerin çözümünde sıklıkla yararlanılmaktadır. Bu amaçla kullanılan yöntemlerden birisi de birden çok sayıda karar ağacının bir araya gelmesi ile oluşturulan RF'dir. RF sınıflama performans sonuçlarının incelenmesinde en fazla kullanılan performans ölçütlerinden birisi ise ROC eğrisi altındadır. ROC eğrisi altında kalan alanın 1'e yaklaşması yöntemin sınıflama performansının oldukça yüksek olduğunu göstermektedir.

RF yönteminde performans sonuçlarının yükseltilmesi için ağaç sayısının artırılması gerektiğini savunan yaklaşımlar bulunsada bu konuda literatürde herkes tarafından kabul edilen tek bir görüş olmadığı bilinmektedir.<sup>22</sup> Bu konuyu inceleyen araştırmacılar arasında bulunan Lattinne ve ark.na göre ise bu görüşlerden farklı olarak, RF yönteminde ağaç sayısının minimum tutulması durumunda en doğru kestirimin elde edilebileceği belirtilmektedir.<sup>23</sup> Bu konu ile ilgili olarak Akman ve ark. sağlık ile ilgili farklı göstergeleri kullanarak RF sınıflama performans sonuçlarını incelenildiği bir araştırmada, ağaç sayısı artışının sınıflamada hata oranının azaltıldığı

ortaya konulmuştur.<sup>3</sup> Bu araştırma sonuçlarına göre ise önceki araştırmalarla uyumlu bir şekilde RF yönteminde ağaç sayısının artırılmasının ROC eğrisi performans sonuçlarına olumlu etkide bulunduğu görülmektedir. Ancak bu çalışmada diğer çalışmalardan farklı olarak bir optimizasyon yöntemi olarak kullanılan çapraz geçerlilikte farklı "k" parametrelerinin belirlenmesi durumunda bunun performans sonuçları üzerinde nasıl bir etkide bulunduğu incelendiğinde "k" parametresinin yüksek ya da düşük tutulması durumunda farklı sayıda ağaç türetilerek elde edilen ROC eğrisi performans sonuçlarının karşılaştırılması sonucunda istatistiksel olarak anlamlı bir farklılık ortaya çıkmadığı görülmektedir. Kohavi tarafından belirtildiği üzere 10 parça çapraz geçerliliğe kadar performans ölçütlerinde farklı sonuçlar elde edilmesi mümkün olmakta iken, 20 parça çapraz geçerlilikten sonra benzer performans sonuçları elde edileceği belirtilmektedir.<sup>24</sup> Bu çalışmada ise en fazla 26 kat çapraz geçerlilik uygulandıktan sonra elde edilen performans sonuçlarının "k" parametresinin yüksek ya da düşük tutulması durumunda farklılık göstermediği görülmüştür. Bu sonuçlar, bu çalışmada makine öğrenmesi yöntemlerinin uygulanmasında farklı bir veri seti olarak ele alınan OECD ülkeleri gibi gelişmiş ülke ekonomilerinde sağlık harcamalarını inceleyen araştırmacılar için bu ülkelerin sınıflandırılmasına yönelik olarak RF yönteminin kullanılması durumunda bir optimizasyon yöntemi olarak k-parça çapraz geçerlilik kullanımının performans sonuçları üzerinde anlamlı bir farklılık yaratmadığını ortaya koymaktadır.

Bu araştırmada elde edilen önemli bulgulardan bir diğeri ise RF yöntemine göre oluşturulan karar ağacında sağlık harcamalarının tahmin edilmesinde en fazla bilgi kazancı sağlayan değişkenin algılanan sağlık statüsü olması durumudur. Algılanan sağlık statüsü dışında sağlık harcamalarının tahmin edilmesinde ön plana çıkan diğer değişkenler arasında ise; bağışıklama ile doğumda beklenen yaşam süresi bulunmaktadır. Sağlık harcamalarını belirleyen değişkenler ile ilgili önceki yıllarda yapılmış çalışmalarda, OECD ülkeleri için Milli Gelirin önemli bir gösterge olduğu belirtilmiş olmasına rağmen bu araştırma sonucunda Milli Gelirin en

fazla bilgi kazancı sağlayan değişkenler arasında yer almadığı görülmektedir.<sup>25</sup> Diğer araştırmalardan farklı olarak bu sonuç, sağlık sistemleri için subjektif performans ölçümünde sıklıkla yararlanılan bir gösterge olan algılanan sağlık statüsü değişkeninin OECD ülkeleri gibi büyük bir çoğunluğu gelişmiş ve gelişmekte olan ekonomilere sahip ülkelerde önem verilmesi gereken bir gösterge olduğunu ortaya koymaktadır. Bu araştırmada OECD ülkelerinin sağlık harcamaları bakımından sınıflandırılmasında ön plana çıkan diğer değişkenler arasında ise doğumda beklenen yaşam süresinin olduğu görülmektedir. Doğumda beklenen yaşam süresi nüfusun yaş kompozisyonu hakkında fikir veren bir gösterge olmakla birlikte OECD ülkelerinde doğumda beklenen yaşam süresinin giderek yükseldiği bilinmektedir. Bu konu ile ilgili olarak OECD ülkelerinde nüfusun yaş kompozisyonundaki farklılığın sağlık harcamalarındaki farklılığın temel bir belirleyicisi olduğu belirtilmektedir.<sup>26</sup>

## SONUÇ

Bu araştırmada çapraz geçerlilik uygulanarak farklı sayılarda "k" parametresinin kullanılması durumunda elde edilen bu sonuçlar, çapraz geçerlilikte "k"nın artırılması durumunda, ne kadar ağaç türetilirse türetilmiş olursa olsun RF performans sonuçlarının farklılık göstermediğini ortaya koymaktadır. Sınıflama problemlerinin çözümünde sıklıkla kullanılan bir yöntem olan RF yönteminde farklı sayıda ağaç türetildiğinde performans sonuçlarının bundan nasıl etkilendiğinin OECD'ye üye ülkelerde sağlık harcamalarının tahmin edilmesine yönelik bir veri seti kullanılarak incelendiği bu araştırmadan farklı olarak bundan sonra yapılacak farklı araştırmalarda farklı büyüklüklerde ve farklı türlerde veri setleri kullanılabilir, farklı model geçerliliği yöntemleri uygulanabilir ve bunun performans sonuçları üzerinde nasıl bir etkide bulunduğunu incelenebilecektir.

## KAYNAKLAR

1. Witten IH, Frank E. What's it all about? Data Mining Practical Machine Learning Tools and Techniques. 2<sup>nd</sup> ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc; 2005. p.3-39.
2. Sathyadevan S, Nair RR Comparative analysis of decision tree algorithms: ID3, C4.5 and Random Forest. In: Jain LC, Mandal HSB, Mohapatra DP, eds. Computational Intelligence in Data Mining, Smart Innovation, Systems and Technologies. Proceedings of the International Conference on CIDM. 20-21 December 2014. India: Springer; 2015. p.549-562.
3. Akman M, Genç Y, Ankaralı H. [Random forests methods and an application in health science]. *Türkiye Klinikleri J Biostat* 2011;3(1):36-48.
4. Li Y. Random forest algorithm in big data environment. *CMNT* 2014;18(12A):147-51.
5. Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP. Random forest: a classification and regression tool for compound classification and QSAR modeling. *JCIM* 2003;43(6):1947-58.
6. Gerdts TA, Cai T, Schumacher M. The performance of risk prediction model. *BJ* 2008;50(4):457-79.
7. Alpar R. Uygulamalı Çok Değişkenli İstatistiksel Yöntemler. 3. Baskı. Ankara: Detay Yayıncılık; 2011. p.603.
8. Forman G, Scholz M. Apples to apples in cross validation studies: pitfalls in classifier performance measurement. *ACM Digital Library* 2010;12(1):49-57.
9. Díaz Uriarte R, de Andrés SA. Gene selection and classification of microarray data using random forest. *BMC Bioinformatics* 2006;7:3.
10. Emel GG, Taşkın Ç. [Decision trees in data mining and a sales analysis application]. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi* 2005;6(2):221-39.
11. Coşgun E, Karağaoğlu E. [Microarray gene expression data analysis with data mining methods]. *Hacettepe Tıp Dergisi* 2011;42:180-9.
12. Chattamvelli R. *Data Mining Methods*. 1sted. U.K. Oxford: Alpha Science International Ltd; 2009. p.411.
13. Alpar R. Uygulamalı İstatistik ve Geçerlilik-Güvenlilik: Spor, Sağlık ve Eğitim Bilimlerinden Örneklerle. 1. Baskı. Ankara: Detay Yayıncılık; 2010. p.339.
14. Akçay A, Demirel A. [Determination of optimal cut-off points of some blood parameters in the bitches with pyometra by likelihood ratios and ROC (Receiver Operating Characteristic) curve metod]. *J Fac Vet Med Univ Erciyes* 2011;8(3):153-63.
15. Oshiro TM, Perez PS, Baranauskas JA. How many trees in a random forest? *MLDM* 2012;(7376):154-168.
16. Rutkowski L, Jaworski M, Pietruczuk L, Duda P. The CART decision tree for mining data streams. *Information Sciences* 2014;266:1-15.
17. Hossin M, Sulaiman MN. A review on evaluation metrics for data classification evaluations. *IJDJP* 2015;5(2):1-11.
18. Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. *IPM* 2009;45(4):427-37.
19. Stone M. [Cross validity choice and assessment of statistical predictions]. *JRSA* 1974;36(2):111-47.
20. Bilger M, Manning WG. Measuring overfitting in nonlinear models: a new method and an application to health expenditures. *Health Econ* 2015;24(1):75-85.
21. Manning WG, Mullahy J. Estimating log models: to transform or not to transform? *J Health Econ* 2001;20(4):461-94.
22. Liaw A, Wiener M. Classification and regression by random forest. *R News* 2002;18-22.
23. Latinne P, Debeir O, Decaestecker C. Limiting the number of trees in random forests. *Multiple Classifier Systems LNCS* 2001;2096:178-87.
24. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Appears in the International Joint Conference on Artificial Intelligence (IJCAI)* 1995:1-7.
25. Gerdtham UG, Sogaard J, Andersson F, Jönsson B. An econometric analysis of health care expenditures: a cross-section study of the OECD countries. *JHE* 1992;11(1):63-84.