SETSCI Conference Proceedings,
4 (1): 445-460, 2019

ISAS 2019
3rd International Symposium on Innovative Approaches in Scientific Studies
Section: Engineering and Natural Sciences
April 19 – 21, 2019, Ankara, Turkey

Karar Ağaçları ile Otistik Spektrum Bozukluğu Tanısı Koyma

Büşra DUYGU, Merve KOCAOĞLU, Doç. Dr. Aysun COŞKUN

Bilgisayar Mühendisliği, Fen Bilimleri Enstitüsü, Gazi Üniversitesi, Ankara, Türkiye

busra.duygu13@gmail.com, mervegullukocaoglu@gmail.com, aysunc@gazi.edu.tr

Özet

Çalışmamız; sosyal etkileşim, tekrarlayan davranışlar, kısıtlı ilgi alanları ve iletişim sorunu olarak ortaya çıkan ve ömür boyu devam eden bir gelişim bozukluğu olan Otizmin erken tanısı üzerine gerçekleştirilmiştir. Karar ağacı algoritmaları ile Otistik Spektrum Bozukluğu tanısı koymada ilişkili özellikle on davranış tespiti sorusu ile %95,73 başarıyla doğru tanı koymada etkili olduğu gözlemlenmiştir. Ayrıca çalışmada, etnik yapının hastalığa etkisi incelenmiş ve sonuçlarda yüksek oranda bir fark tespit edilmemiştir.

Anahtar Kelimeler: karar ağaçları, otizm, veri madenciliği

Abstract

The study was carried out on the diagnosis of Autism, which is a lifelong developmental disorder, occurring as social interaction, repetitive behaviors, restricted interests and communication problems. It has been observed that, with ten problem determination problems, 95.73% have been found to be effective in diagnosing the autistic spectrum disorder with decision tree algorithms. In addition, the effect of ethnicity on the disease was examined and no significant difference was found in the results.

Keywords: decision tree, autism, data mining

1. Giriş

Otizm; kısıtlı ilgi alanları, sözlü iletişim ve davranışta bozukluk olarak ortaya çıkan ve ömür boyu gelişimi etkileyen bir hastalıktır. İlk olarak Leo Kanner tarafından 1943 yılında irdelenen konu; şizofreni rahatsızlığından farklı olduğu açıklanmış, çocuklarda tekrarlayan hareket, ekolali, ısrarcı olmaları, aynılığı sürdürme durumları ve saplantılar, otizm tanımında ölçüt olarak verilmiştir [1].

Yakın zamanda otizm, nörobiyolojik bir etiyolojiye sahip olduğu anne babanın tutumu ve sosyal çevrenin etkisi olmadığı görüşü ağır basmaktadır. Otizm tanısı için literatürde 3 ana koşul ortaya atılmıştır[2];

- Toplumsal alandan gelişim ve iletişim bozukluğu

- Sınırlı ilgi alanı ve yineleyici davranışlar
- Bu alanlardaki bozuklukların 30 ay öncesinden görülmesi

Otizmin tam olarak nedeni bilinmemesine rağmen, yapılan çalışmalarda, hastalığın oluşmasında ki en büyük faktörün genetik yapı olduğu gözlemlenmiştir [3]. Erken teşhisin hastalığın ilerleyişini olumlu yönde etkilemektedir. Hem otizm hem de diğer genetik alanlardaki rahatsızlıkların erken teşhisinde uygulanabilecek yöntem çalışmaları, teknolojinin de gelişmesiyle ivme kazanmıştır. Bu çalışma kapsamında otizmin erken teşhisinde etkili faktörlerin birbiriyle ilişkisini incelemektedir. Çalışmada ele alınan veri setinin boyutu ve miktarı göz önünde bulundurulduğunda verilerin incelenmesinde veri madenciliği yöntemlerinden faydalanılmaktadır. Görsellik yönünden zengin olması, maliyetinin düşük olması ve anlaşılabilirliğinin yüksek olması sebepleriyle karar ağaçları(KA) yöntemi kullanılmıştır.

Çalışma beş bölümden oluşmaktadır. Bu giriş bölümünden sonra, ikinci bölümde literatür özeti yer almaktadır. Üçüncü bölümde çalışmada kullanılan veri , özellikleri ve analiz teknikleri hakkında temel bilgiler sunulmaktadır. Dördüncü bölümde veri seti ve analiz sonuçları , beşinci bölümde ise sonuç ve tartışma yer almaktadır.

2.Literatür Taraması

Literatürde karar ağaçları sağlık alanında farklı problemlerin çözümü için başarılı bir şekilde kullanılmaktadır. Bu çalışmalardan bazılarını şu şekilde özetlemek mümkündür. Hepatit hastalığı üzerinde çalışan Uhm ve ark., 194 hasta ve 28 tekli nükleotit polimorfizmlerini (TNP) farklı değişik KA yöntemleriyle incelemiştir[4]. 2004 yılında yapılan başka bir çalışmada Huang ve ark., hipertansiyonlu hastalardan 4529"unun genetik varyasyonları üzerinde ID3, ADTree ve C4.5 gibi KA yöntemlerini uygulamış ve çalışmalarında bu yöntemlerin performans karşılaştırmasını yapmışlardır[5] .2010 yılında Anunciacao ve ark., göğüs kanseri alanında 258 hasta üzerinde gerçekleştirdikleri çalışmada kullanılan KA yöntemlerinden en yüksek kesinlik ölçütü C4.5 ağacı vermiştir[6]. Otizm spektrum bozukluğu üzerinde yapılan bir başka çalışmada 36 hasta bilgisi üzerinde genetik varyasyona bağlı teşhiste KA ve destek vektör makinesi yöntemleri karşılaştırılmış ve çıkan sonuçlarda iki yöntemin birbirine yakın sonuçlar verdiği görülmüştür[7]

Çalışmalarında[8] Otizm Spektrum Bozukluğu üzerinde erken teşhisin ve bilinçlenmenin üzerinde durup, hızlı erişebilir basit kullanımlı bir araç sağlamaktır. Ortak ASD tarama özelliklerinin belirlenmesini amaçlayarak ASDTests mobil uygulamasından elde edilen veriler üzerinde Variable Analysis Algoritması kullanılarak gerçekleştirilmiştir. Bu çalışmaya benzer olarak [9] 2012 de yaptıkları çalışmalarında ADTree algoritmasını kullanarak, otizm tanı yönteminde kullanılan öğe sayısını azaltarak, teşhisin hızlandırmak üzerine gerçekleştirilmiştir. Daha önce tespit için kullanılabilen 29 özellikten sadece 8 tanesinin kullanarak tanının konulmasında etkili olduğunu ileri sürmüşlerdir. Üzerinde durdukları 8 özellik ADTree algoritması sonucuna göre belirlenmiş olup diğer algoritmalarda başka özelliklerinde karar vermede etkili olduğu gözlemlenmiştir. Bu durum çalışmanın doğruluğunun kanıtlanmasını zorlaştırmıştır.

Literatürdeki otizm üzerine yapılan çalışmaların çoğunluğunda otizm erken dönemde teşhis edilmesinin öneminden bahsetmektedir. Bu çalışmalardan [10] Otizm hastalığının erken teşhisinin zor olduğu ve kesin tedavisinin günümüzde halen bulunamadığından, hastalığın gelişimi için belirtilerin iyi analiz edilmesi gerektiği vurgulanmaktadır. Çalışmada, halk sağlığı hemşirelerine yönelik çocuğun gelişimi ve davranışı izlenimi ile otizm tanı ve tedavisi üzerine yapılmıştır. Başka bir çalışmada [11] ise İsveç'te tüm çocuklara

30 aylık ücretsiz sağlık muayenesi planlaması ile 12 gözlemden istatiksel sonuç çıkarmıştır. Verileri; daha önce teşhisi Otizm Spektrum Bozukluğu olarak tespit edilmiş 37 çocuk ve down sendromlu 23 hasta ile test edilmiş ve otizm tanısı üzerinde 9 gözlem verisi üzerinde durulmuştur. Bu gözlemler; el sallamak, ismi söylendiğinde tepki verme, göz teması gibi çocuğun sosyal çevreye verdiği tepkilerin analizleridir. Çalışmamızda kullandığımız veri seti bu gözlemleri de içermektedir.

3.Kullanılan Veri ve Yöntemler

3.1.Karar Ağaçları

Son yıllarda en yaygın kullanılan sınıflandırma ve tanımlama yöntemlerinden biri karar ağaçlarıdır. Çok fazla tercih edilmesindeki en önemli sebep ise ağaç yapılarının oluşturulmasında izlenen kuralların anlaşılabilir ve sade olmasıdır[12]. Güvenilirliklerinin yüksek olması, veri tabanı sistemlerine entegrasyonunun kolay olması da tercih edilmesinde etkilidir [13].

Karar ağaçları kök düğüm, dallar ve yapraklardan oluşmaktadır. Karar ağaçlarındaki yapraklar sınıflandırmanın oluştuğu yerler olup dallar ise sonucu ifade etmektedir. Kök düğümden yaprak düğümlere doğru tüme varım yöntemiyle ağaç oluşturulur[14] .Karar ağacı iki temel işlem adımı içerir. Bu işlemler bölme ve budama işlemleridir. KA oluştururken en önemli adım hangi öznitelik değerlerini baz alıp dallanmanın oluşturulacağına karar vermektir. Literatürde bu sorunu çözmek için bir çok yaklaşım geliştirilmiştir. Bilgi kazancı[15] , Gini indeksi[16], Towing kuralı[16] ve Ki-Kare olasılık tablo istatistiği[17] yaklaşımları bunlardan en önemlileridir. Entropi kurallarını içeren bilgi kazancı ve bilgi kazancı oranı yaklaşımında eldeki tüm öznitelikler özyineli olarak test edilir ve bilgi kazancı en yüksek olan öznitelik dallanma için seçilir[18].

Formüller ile açılayacak olursak :

Bir veri setinin $A_1, A_2, \ldots, A_n$ şeklinde birkaç tane sınıftan oluştuğu ve 'nin sınıf değerini gösterdiği düşünülürse, bir sınıfa ait olasılık $P_i=(A_i/|T|)$ olur ve sınıflara ait Entropi ;

$$Entropi(T) = -\sum_{i=0}^{n} Pi \log_2(Pi) \qquad (1)$$

Şeklinde hesaplanır. Veri setindeki B özniteliğine göre T sınıf değerleri $T_1,T_2,\ldots T_n$ şeklinde alt kümelere ayrıldığı göz önünde bulundurulsun. B öznitelik değeri kullanılarak T sınıf değerlerinin bölünmesi sonucunda elde edilecek kazanç,

$$Kazanç(B,T) = Entropi(T) - \sum_{i=0}^{n} \frac{|Ti|}{|T|} Entropi(Ti) \qquad (2)$$

Şeklinde hesaplanır. T kümesi için B özniteliğinin değerini belirlemesinde bölümleme bilgisi kullanılır(eşitlik 3).

$$Bilgi(B) = -\sum_{i=0}^{k} \frac{|Ti|}{|T|} \log_2\left(\frac{|Ti|}{|T|}\right) \quad (3)$$

Kazanç oranı,

$$Kazanç\ oranı = \frac{Kazanç(B,T)}{Bilgi(B)} \quad (4)$$

şeklinde hesaplanır[19].

Diğer önemli adım ise budamadır. Budama, bir ağaçın istenmeyen alt düğümlerini veya alt ağaçlarını ayıklama işlemidir[20].
Karar ağacı algoritmasında ortaya çıkan sorulardan biri, işlemler sonucunda elde edilen final ağacının en uygun büyüklüğüdür. Çok büyük bir ağaç, eğitim verilerinin üzerinde çok fazla risk oluşturur ve yeni örneklere genellemenin zayıf olmasına neden olur. Tam tersi küçük bir ağaç örnek alanı hakkında önemli yapısal bilgileri yakalayamayabilir. Bununla birlikte, bir ağaç algoritmasının ne zaman durması gerektiğini söylemek zordur çünkü tek bir ekstra düğüm eklenmesinin hatayı önemli ölçüde azaltıp azalmayacağını söylemek mümkün değildir. Bu sorun ufuk etkisi(horizon effect) olarak bilinir. Ortak bir çözüm ise, her düğüm az sayıda örnek içerene kadar ağacı büyütmek ve ardından ek bilgi sağlamayan düğümleri kaldırmak için budama kullanmaktır [21] .Budama işlemi sonucunda daha az karmaşık ve danlaşılabilir bir ağaç elde edilmiş olur. Böylece sadeleşen ağaç yapısının işlem karmaşıklığı da azalmış olur. Budama yapılırken iki yöntem kullanılır : ilki ağacın oluşturulma aşamasında küçültülmesine karar verilen ön budama işlemi, diğeri ise ağaç yapısı oluşturulduktan sonrasında gerçekleştirilen son budama işlemidir[22].

3.2.Kullanılan Veri

Çalışmada Kaggle adlı veri seti paylaşım sitesinden "Yeni yürümeye başlayan çocuklar için Otistik Spektrum Bozukluğu Tarama Verileri" kullanılmıştır. Çalışma yeni yürümeye başlayan 319 kız ve 735 erkek çocuk ebeveyni ile gerçekleştirilmiştir.

Çalışmada, katılımcılara bir takım sorular sorulmuştur:

Soru Kod	Soru
Q1	Adını söylendiğinizde çocuğunuz size bakıyor mu?
Q2	Çocuğunuzla göz teması kurmanız sizin için ne kadar kolay?
Q3	Çocuğunuz bir şey istediğini belirtmek ister mi? (örneğin erişilemeyen bir oyuncak)
Q4	Çocuğunuz ilgilendiği şeyi sizinle paylaşıyor mu? (örneğin, ilginç bir manzaraya işaret etmek)
Q5	Çocuğunuz numara yapar mı? (ör. bebek bakımı, oyuncak telefonda konuşma)
Q6	Çocuğunuz aradığınız yeri takip ediyor mu?

Q7	Siz veya aile içindeki bir başkası gözle görülür bir şekilde üzülürse, çocuğunuz onları rahatlatmak istediğinin belirtilerini gösteriyor mu? (örneğin saçları okşayarak, kucaklayarak)
Q8	
Q9	Çocuğunuz jest ve mimiklerini kullanıyor mu? (örneğin el sallamak)
Q10	Çocuğunuz görünür bir amaç olmadan hiçbir şeye bakmıyor mu?

Tablo1: Çalışmada kullanılan anket soruları

Özellik	Minimum	Maximum	Mean	StdDev
Age	12	36	27.867	7.98
Q1	0	1	0.564	0.496
Q2	0	1	0.449	0.498
Q3	0	1	0.401	0.49
Q4	0	1	0.512	0.5
Q5	0	1	0.525	0.5
Q6	0	1	0.577	0.494
Q7	0	1	0.65	0.477
Q8	0	1	0.459	0.499
Q9	0	1	0.49	0.5
Q10	0	1	0.586	0.493

Tablo 2: Soruların minimum, maksimum, ortalama ve standart sapma bilgileri

Özellik	İçerik	Count
Cinsiyet	Kadın	319
	Erkek	735
Etnik	Middle east	188
	White Europen	334
	Hispanic	40
	Asian	299
	South Asian	60
	Native Indian	3
	Others	35
	Latino	26
	Mixed	8
	Pacifica	8
Jaundice	Yes	288
	No	766
Famil_mem_withASD	Yes	170
	No	884
Who completed	Family member	1018

	Health Care Professional	29
	Self	4
	Others	3
Class/ ASD Traits	Yes	728
	No	326

Tablo 3:Verilerin kategorileri ve sayısal değerleri

Çalışmada kullanılan veriler, hastalık şüphesi olan kişinin davranış bilgisini sorgulama yönelik 10 soru ve kişisel ve sosyal çevresine ait (yaşı, cinsiyeti, etnik kökeni, testi tamamlayan kişi, ailede otizmli bir başka üyenin olup olmaması gibi) sorulara cevap aranmıştır. Veriler üzerinde temizleme ve ayıklama işleminden sonra 728 hasta verisi ve 326 sağlıklı veri bilgisi kullanılmıştır. 735 erkek çocuk verisi ve 319 kız çocuk verisi ile tamamlanan veri setinde 319 veri ile en çok White Europen etnik köken katılımı sağlanmıştır. Verilerin dağılımları Tablo 2. ve 3. de gösterilmiştir.

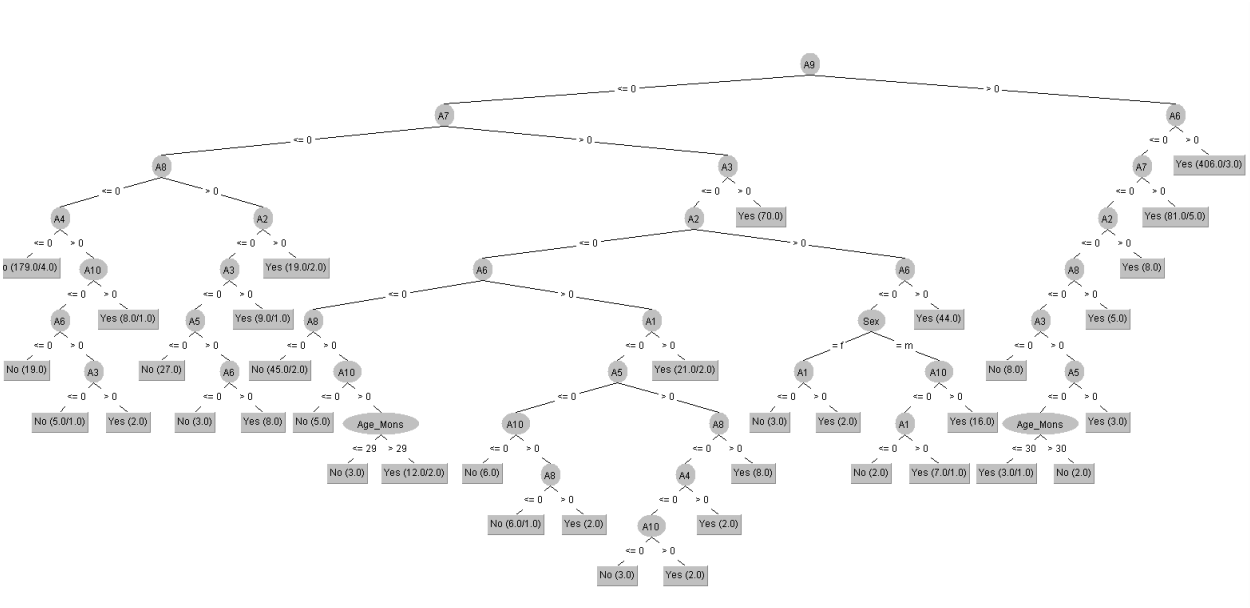
Sağlık; toplumsal yaşantıyı etkilediği gibi toplumsal ve ekonomik olaylarda etkilenen bir olgu olarak kabul edilmektedir. Literatürde sağlığı etkileyen en önemli toplumsal faktörler olarak aile, sosyal sınıf, din, coğrafi bölge, cinsiyet ve etnik köken olarak verilmiştir [23].

Toplumsal faktörlerden etnik yapı hakkında [23] çalışmada, Siyahi kökenlilerin beyazlara oranla zihinsel hastalık teşhisi sonucuna varılma olasılıklarının yüksek olduğu, akıl hastanelerinde beş kişiden birinin siyahi olduğu ve diğerlerini azınlık etnik grubun oluşturulduğu belirtilmektedir. Yine aynı çalışmada kalp hastalıklarında, hipertansiyon, bebek ölümleri, sickle cell gibi rahatsızlıklarda etnik kökenin etkisi olduğu da belirtilmiştir. Çalışmamızda toplumsal etki olan etnik yapı konusunda da ayrıntılı inceleme yapılmıştır.

4.Veri Analizi

Waikato Üniversitesinde geliştirilmiş olan WEKA Programının (Waikato Environment for Knowledge Analysis) 3.8.3 sürümü kullanılmıştır. WEKA Programı, birçok sınıflandırma, kümeleme ve birliktelik kurallarına ait algoritmayı desteklemektedir ve tamamen açık kaynak kodlu bir yazılımdır. WEKA, arff dosya tipinin yanı sıra data, csv, c45, libsvm, dat, bsi, xrff, xrff.gz dosya tiplerini de desteklemektedir. Mevcut veriseti ile yapılan uygulama sonucunda izleyen tablodaki sonuçlar algoritmaların performans derecelerine göre sıralanarak sunulmuştur:

İlk olarak çalışmada karar ağacı algoritmalarının doğruluğunu ve sonucunu kötü etkilememesi sebebiyle veri üzerinde temizleme işlemleri yapılmıştır. Bunlardan “Who completed” bilgisinin hastalık tespiti üzerinde bir etkisi olmadığı gerekçesiyle veri setinden kaldırılmıştır. Veri setinde Q1,Q2,…Q10 sorularının toplamını ifade eden Qchat-10 score alanı da kaldırılmıştır. Bu şekilde her bir sorunun karar ağacındaki etkisi detaylı incelenmiştir.

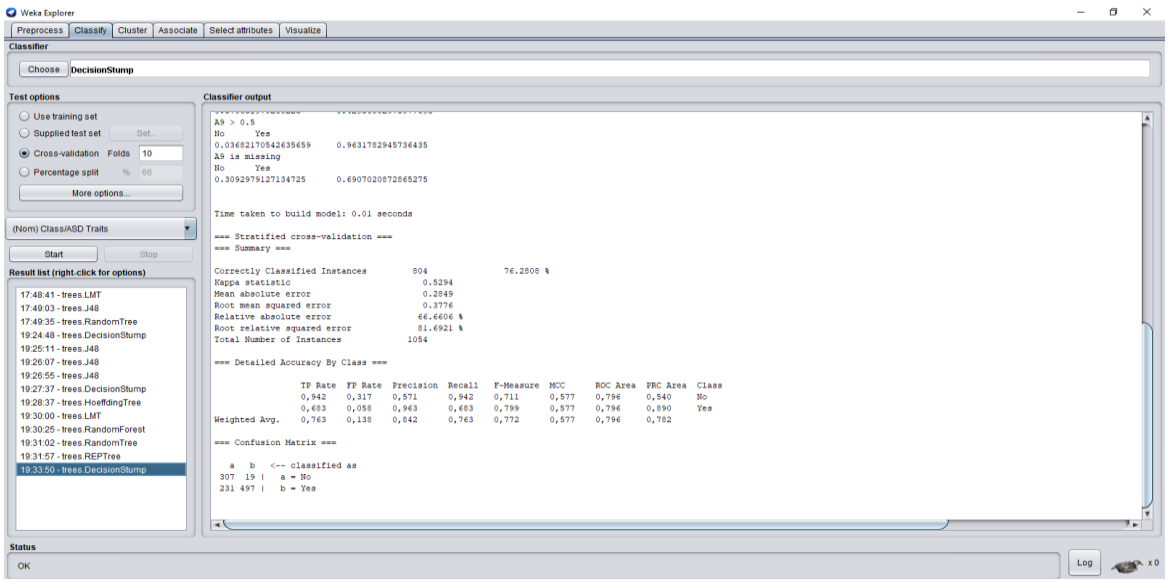


Şekil 1

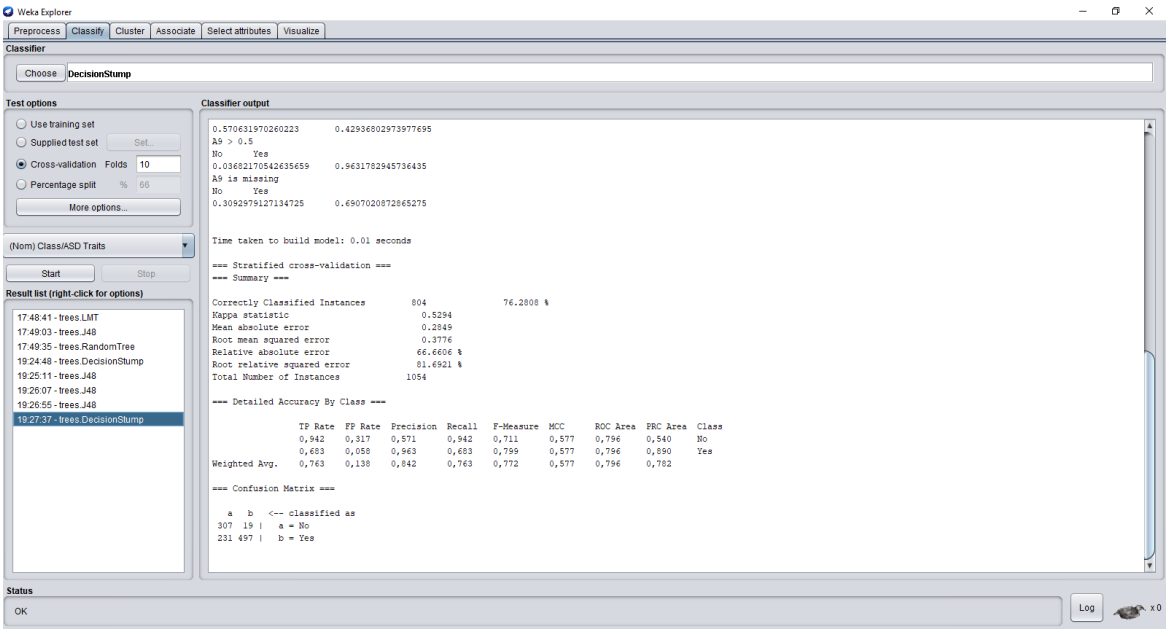
Şekil 1 kullanılan algoritmalardan j48 'in karar ağacı dallanma görüntüsüdür. Kök düğüm olarak seçilen "Q9: Çocuğunuz jest ve mimiklerini kullanıyor mu?" sorusunun bilgi kazancının en yüksek olduğu görülmektedir. Şekil içerisinde alt dallanmalar ve nihai sonuçları gösteren yapraklarda mevcuttur.

Çalışmada hastalık üzerinde etnik yapının etkisi araştırılmış ve çeşitli karar ağaçları ile test edilmiştir. Bazı algoritmalarda etnik yapının hastalık tespiti üzerinde etkili olduğu görülürken, DecisionStump, J48 algoritmalarında etnik yapının hastalık tespiti üzerinde etkisi olmadığı sonucuna ulaşılmıştır.

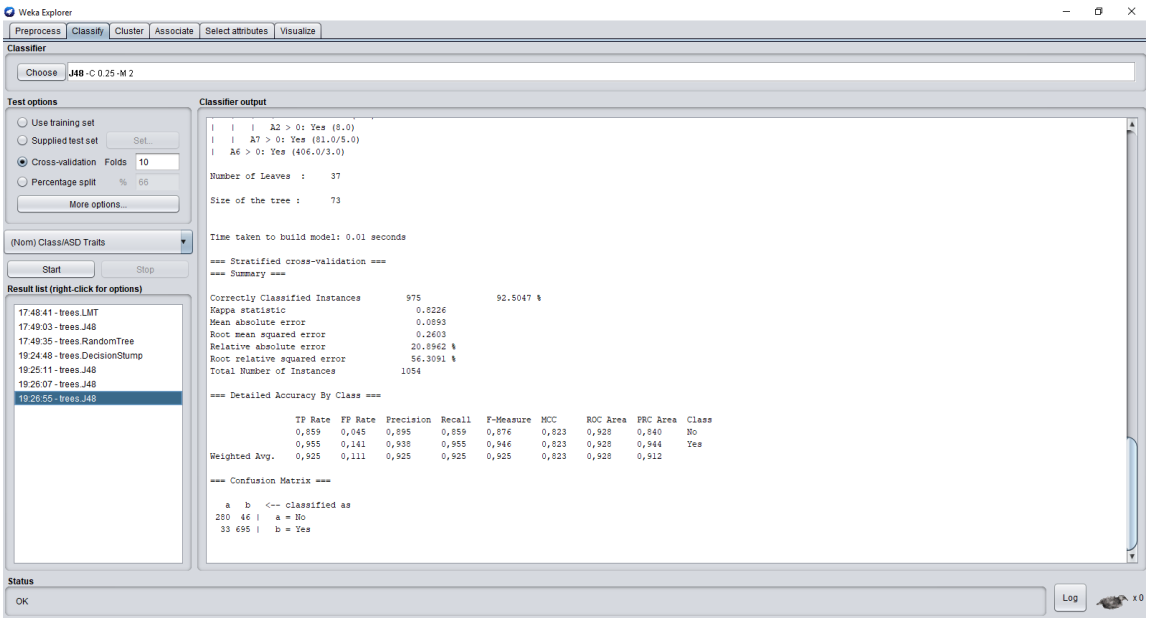
Çalışmanın Weka uygulaması üzerindeki sonuçların detaylı gösterimi paylaşılmıştır.



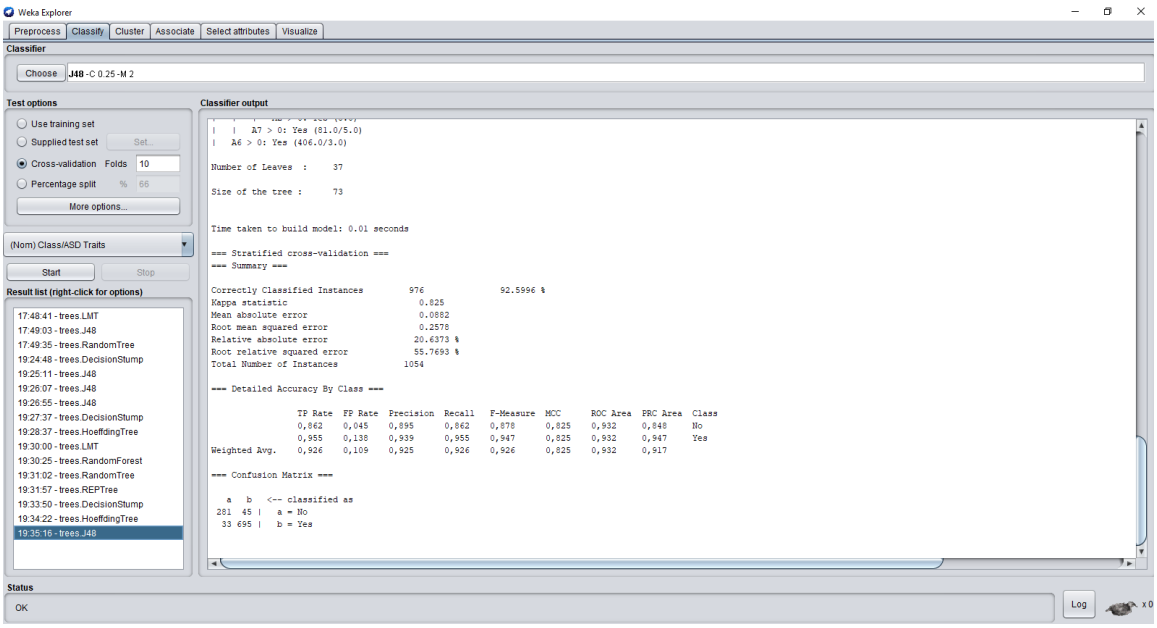
Şekil 1. DecisionStump etnik yapı dahil



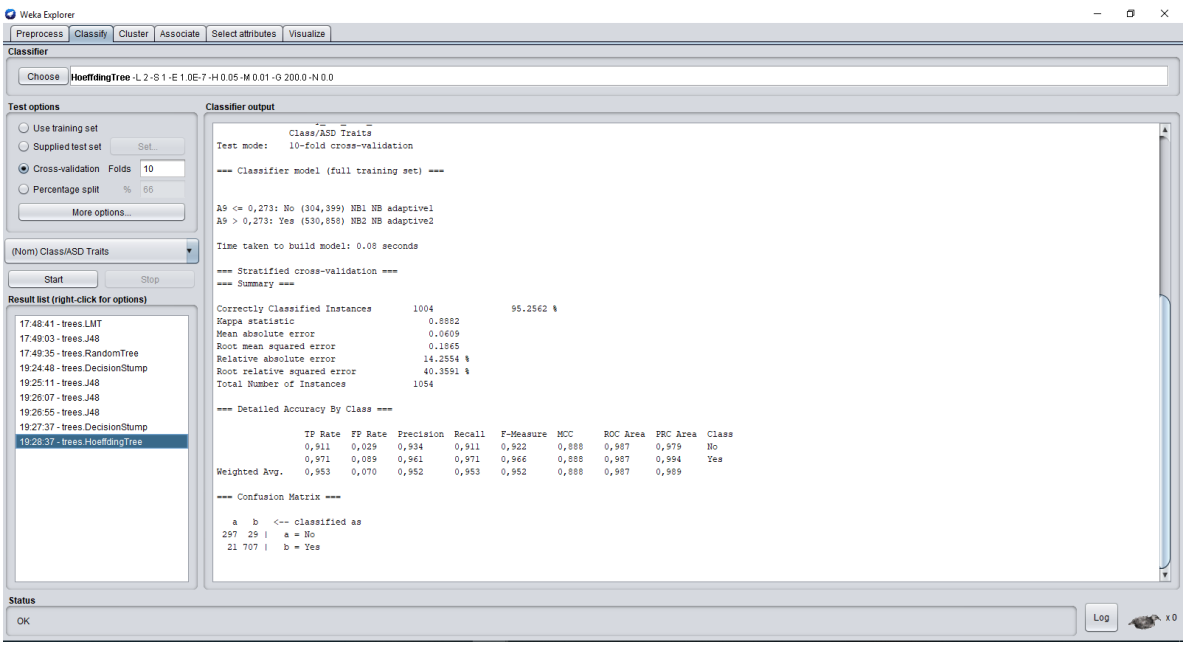
Şekil 2 DecisionStump etnik yapı dahil değil



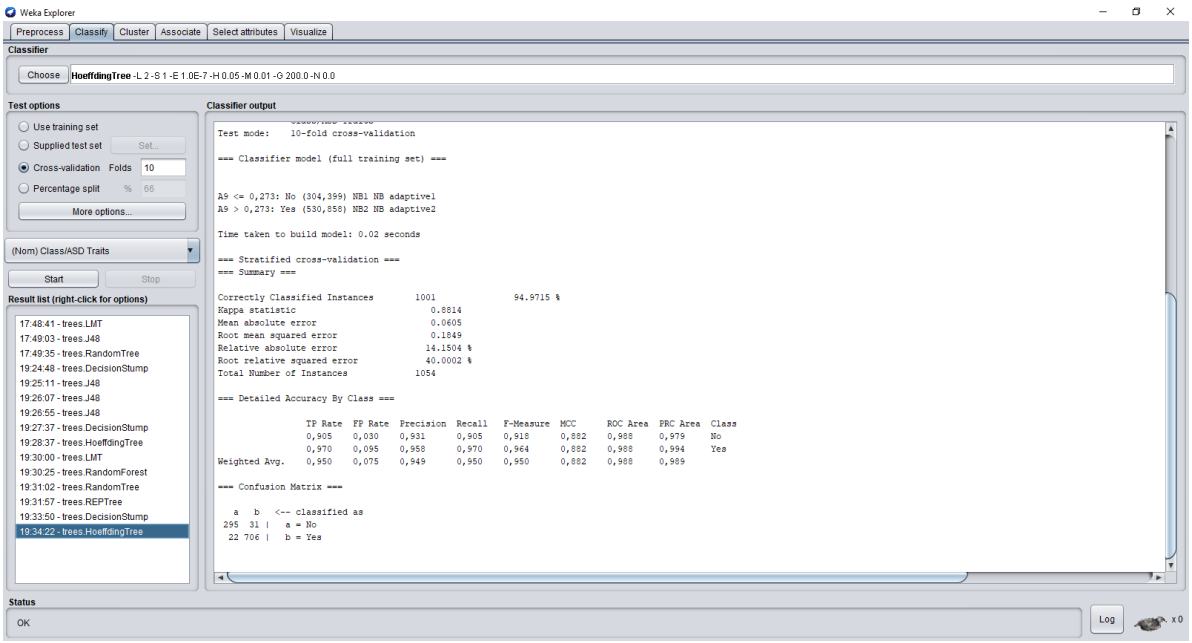
Şekil 3. J48 etnik yapı dahil



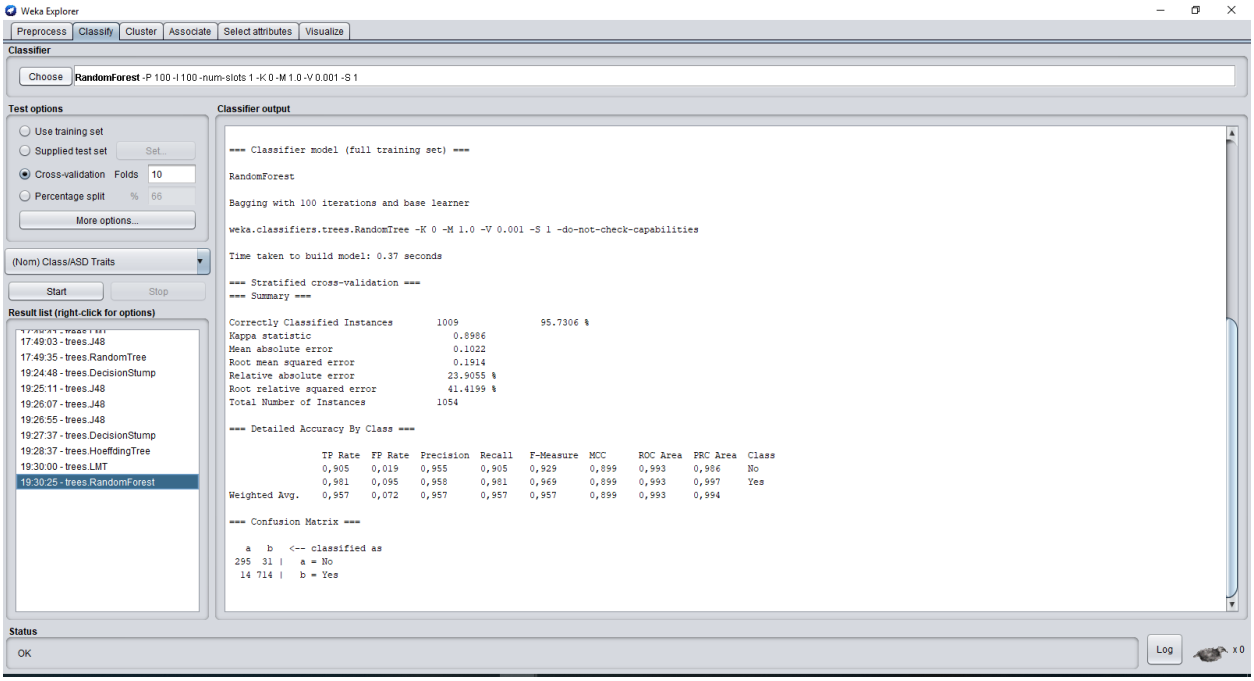
Şekil 4. J48 etnik yapı dahil değil



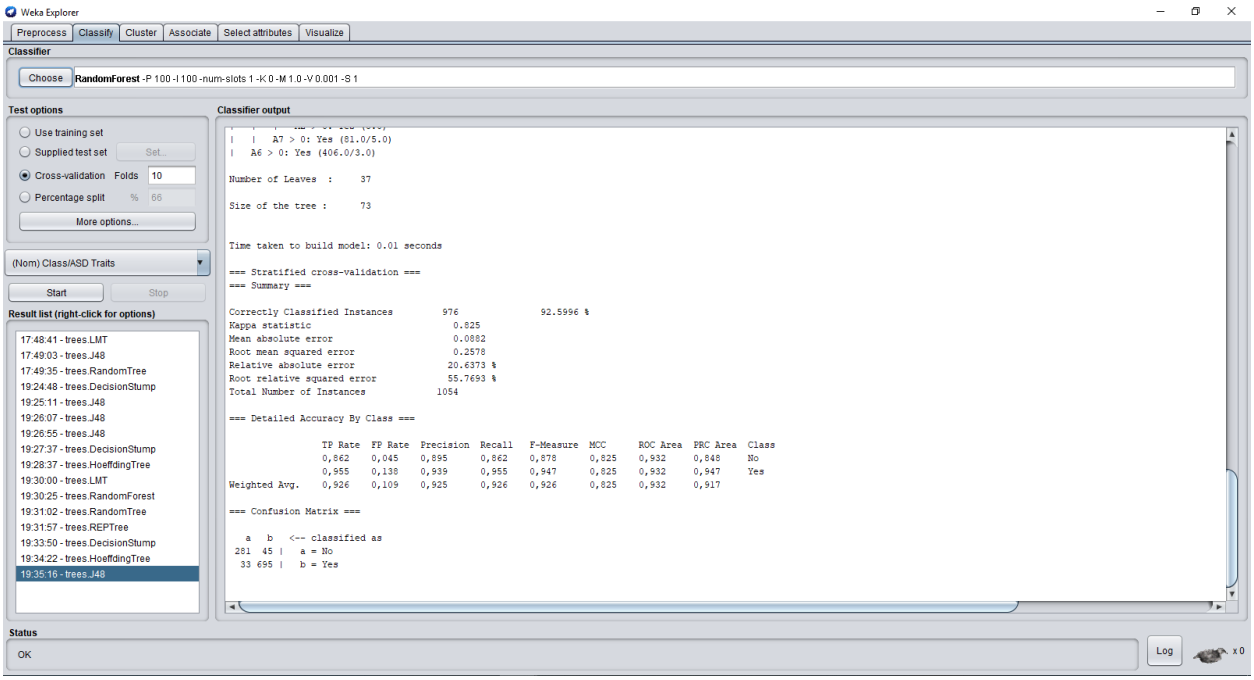
Şekil 5. HoeffdingTree etnik yapı dahil



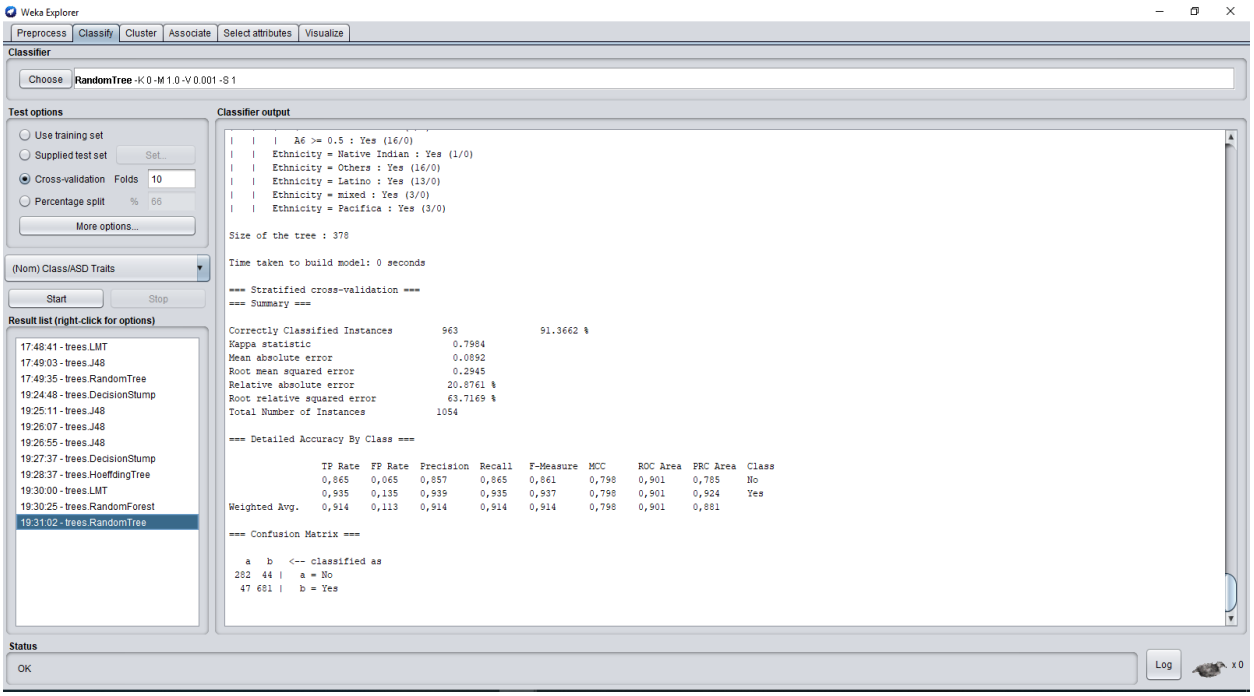
Şekil 6.HoeffdingTree etnik yapı dahil değil



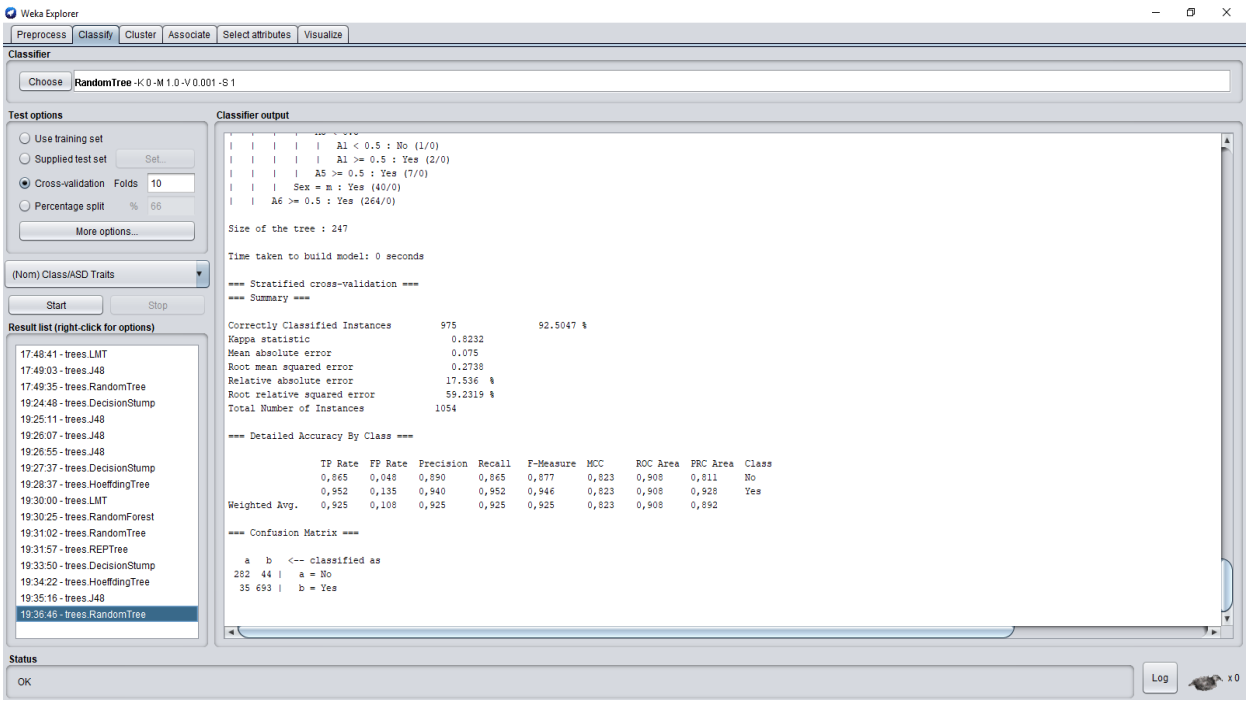
Şekil 7.RandomForest etnik yapı dahil



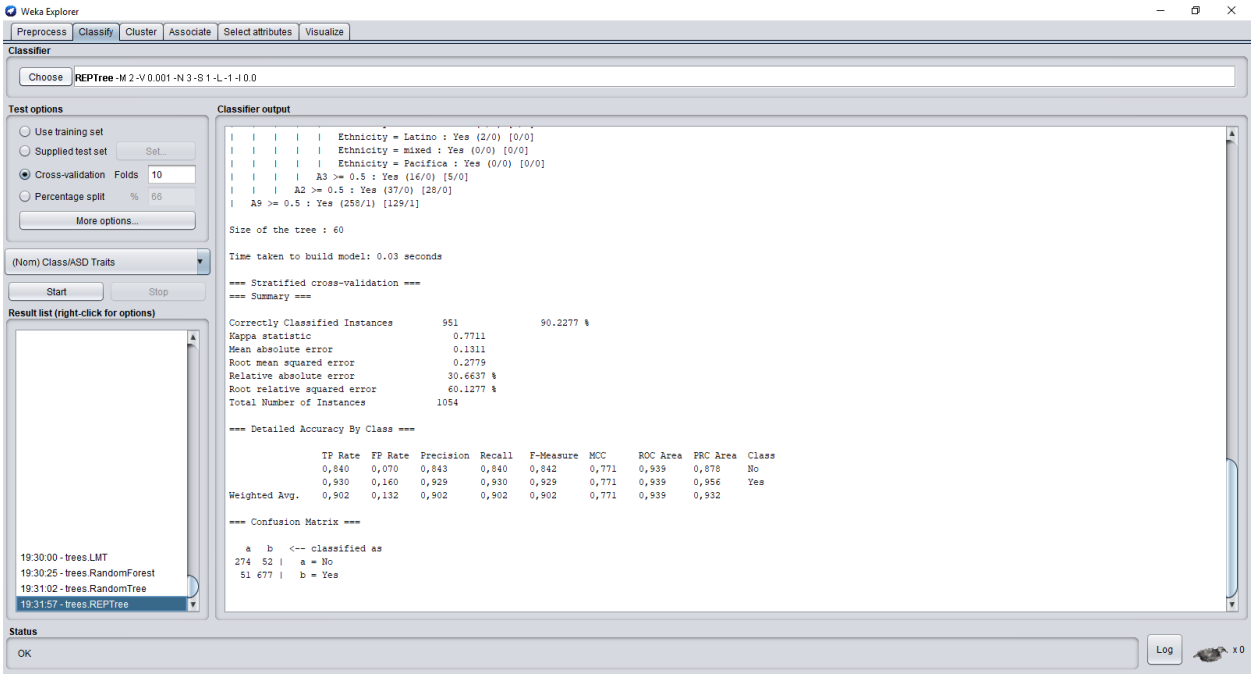
Şekil 8.RandomForest etnik yapı dahil değil



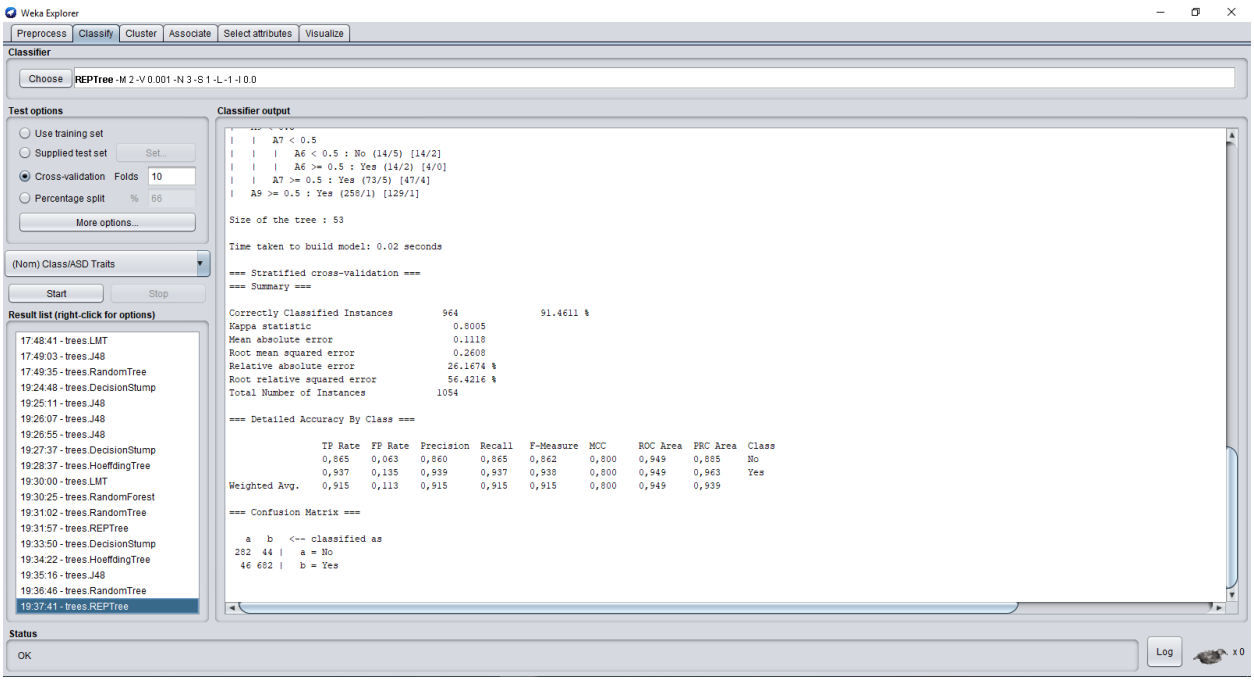
Şekil 9.RandomTree etnik yapı dahil



Şekil 10.RandomTree etnik yapı dahil değil



Şekil 11.REPTree etnik yapı dahil



Şekil 12.REPTree etnik yapı dahil değil

5.Sonuç ve Öneriler

Çalışmada karar ağaçlarına ait RandomForest, J48, RandomTree, Hoeffding Tree, DecisionStump,REPTree gibi pek çok algoritma denenmiş, başarım dereceleri tabloda karşılaştırılmıştır.

Algoritma	Etnik Bilgisi	Doğru sınıflandırılan Örnek	Kappa İstatistiği	Ortalama Mutlak Hata	Ortalama Hata Karekök	Göreli Mutlak Hata %	Göreli Hata Karekök %	TP Oranı	FP Oranı	F-Ölçütü
Decision Stump	dahil	804	0.529	0.284	0.377	66.6	81.69	0.942	0.317	0.711
	dahil değil	804	0.529	0.284	0.377	66.6	81.69	0.942	0.317	0.711
J48	dahil	975	0.822	0.089	0.260	20.89	56.30	0.859	0.045	0.876
	dahil değil	976	0.825	0.088	0.257	20.63	55.76	0.862	0.045	0.878
REPTree	dahil	951	0.771	0.131	0.277	30.66	60.12	0.840	0.070	0.842

	dahil değil	964	0.800	0.111	0.260	26.16	56.42	0.86	0.063	0.862
RandomForest	dahil	1009	0.898	0.102	0.191	23.90	41.41	0.905	0.019	0.929
	dahil değil	976	0.825	0.088	0.257	20.63	55.76	0.862	0.045	0.878
RandomTree	dahil	963	0.798	0.089	0.294	20.87	63.71	0.865	0.065	0.861
	dahil değil	975	0.823	0.075	0.273	17.53	59.23	0.865	0.048	0.877
Hoeffding Tree	dahil	1004	0.888	0.060	0.186	14.25	40.35	0.911	0.029	0.922
	dahil değil	1001	0.881	0.060	0.184	14.15	40.00	0.905	0.030	0.918

Otizm tanısı için literatürde çok fazla çalışma yer almaktadır. Bu çalışmalar incelendiğinde otizm tanısının konulmasının yorucu ve uzun bir süreci beraberinde getirdiği görülmektedir. Çalışmamızda toplanan on davranış tespiti sorusu ve diğer bilgiler ile Otistik Spektrum Bozukluğunun tespitinin mümkün olduğu %95,73 başarıyla gerçekleştirilebileceği gösterilmiştir. Hem başarının yüksek olması hem de birkaç dakikada sonuca ulaşılması otizmin erken teşhisinde etkili olacağı öngörülmektedir.

Karar ağaçları algoritmaları ile veri setindeki otizm üzerindeki etkili faktörlerin birbirleriyle olan ilişkisi incelenmiş ve etnik yapının bu hastalık üzerinde etkisinin olup olmadığı araştırılmıştır. DecisionStump ve J48 algoritmalarının sonuçlarına bakıldığında başarım oranlarında gözle görülür bir değişim olmamıştır.

Literatürdeki sonuçlarla elde etmiş olduğumuz sonuçlar karşılaştırılmış ve sonuçların benzer özellikler gösterdiği ve bu benzerlikten yola çıkılarak otizmin tespitinde bu çalışmadan yararlanılabileceği sonucuna varılmıştır.

Çalışmamızda otizm veri seti az olmasına rağmen yüksek başarı elde edildiği görülmüştür. Bu konuya çözümler sunması açısından önemli bir çalışma olduğu değerlendirilmektedir.

Çalışmada incelenen bir alt başlık olarak etnik yapının otizm üzerindeki etkisi, dünyada bulunan tüm etnik yapılara ait otizmli hasta bilgilerinin yer alacağı bir veri seti ile daha doğru sonuçlar elde edilebileceği öngörülmektedir.

6. Kaynakça

[1] Özeren,Sultan,Gül.,2013.Otizm Spektrum Bozukluğu (OSB) ve Hastalığa Kanıt Penceresinden Bakış.Acıbadem Üniversitesi Sağlık Bilimleri Dergisi .4(2).pp.57-63

[2] Bodur,Ş., Soysal,Ş., Otizmin Erken Tanısı ve Önemi, 2004. cilt 13 ,sayı 10, 394-398

[3] Şener,E.F.,Özkul ,Y., 2013.OTİZMİN GENETİK TEMELLLERİ ,The Genetic Basis of Autism, Sağlık Bilimleri Dergisi (Journal of Health Sciences) 22(1) 86-92, 2013
[4] Uhmn S., Kim D.H., Ko Y.W., Cho S., Cheong J., Kim J., 2009. A Study on Application of Single Nucleotide Polymorphism and Machine Learning Techniques to Diagnosis of Chronic Hepatitis. Expert Systems. 26(1)
[5] Huang J., Lin A., Narasimhan B., Quertermous T., Hsiung C.A., Ho L.T., Grove J.S., Oliver M., Ranade K., Risch N.J., Olshen R.A., 2004. Tree-structured supervised learning and the genetics of hypertension. Proceedings of the National Academy of Sciences of the United States of America. July 12. 101(29):10529–10534
[6] Anunciacao O., Gomes B.C., Vinga S., Gaspar J., Oliveira A.L., Rueff J., 2010. A Data Mining Approach for the Detection of High-Risk Breast Cancer Groups. In: Rocha M.P., Riverola F.F., Shatkay H., Corchado J.M. Eds. Advances in Bioinformatics. Advances in Intelligent and Soft Computing, Springer, Berlin, Heidelberg. 74: 43-51
[7] JiaoY., Chen R., Ke X.,Cheng L., ChuK., Lun Z., Herskovits E.H., 2011. Predictive Models for Subtypes of Autism Spectrum Disorder Based on Single-Nucleotide Polymorphisms and Magnetic Resonance Imaging. Advances in Medical Sciences. 56: 334-342
[8] Thabtah,F., Kamalov,F.,Rajab,K.,2018.A new computational intelligence approach to detect autistic features for autism screening,Volume117, September 2018, Pages 112-124
[9] Wall DP, Kosmicki JA, DeLuca T, Harstad EB, Fusaro VA.,2012. Use of Machine Learning to Shorten Observation-based Screening and Diagnosis of Autism. Translational Psychiatry. 2012a;2(4):e100.
[10] Dur,Ş., Mutlu, B.,2017. Otizm Spektrum Bozukluğu ve Hemşirelik Yaklaşımı ,Autism Spectrum Disorder and Nursing Approach. Hemşirelikte Eğitim ve Araştırma Dergisi 2018;15 (1): 45-50
[11] Haglund,N., Dahlgren,SO.,Källén, K., Gustafsson, P., Råstam, M., 2015.The Observation Scale for Autism (OSA): A New Screeening Method to Detect Autism Spectrum Disorder before Age Three Years,2015, 3, 230-237
[12] Safavian S.R., Landgrebe D., 1991, "A survey of decision tree classifier methodology", IEEE Transactions on Systems Man and Cybernetics, 21, 660-674
[13] Göral,MA., "Kredi Kartı Başvuru Aşamasında Sahtecilik Tespiti İçin Bir Veri Madenciliği Modeli – Yük. Lisans Tezi İ.T.Ü. Fen Bilimleri Ens.. Ocak 2007-
AKPINAR, 2000, s. 17.
[14] Geetha, A. ve Nasira, G.M. (2014) "Data Mining for Meteorological Applications: Decision Trees for Modeling Rainfall Prediction" IEEE International Conference on Computational Intelligence and Computing Research.
[15] Quinlan J.R., 1993, "C4.5: Programs for Machine Learning", Morgan Kaufmann, San Mateo, CA,302 s
[16] Breiman L., Friedman J.H., Olshen R.A. and Stone C.J., 1984, "Classification and Regression Trees"Monterey, CA: Wadsworth, 358 s
[17] Mingers J., 1989, "An empirical comparison of pruning methods for decision tree induction",Machine Learning, 4, 227–243
[18] Yücebaş,S.,2018.Karmaşık Hastalıkların Teşhisinde Veri Madenciliği Yöntemlerinin Başarım Karşılaştırması, Çanakkale Onsekiz Mart University, Journal of Graduate School of Natural and

Applied Sciences, 2018:4,1, 14-27//// Quinlan J.R., 1986. Induction of Decision Trees. Machine Learning. 1(1):81-106

[19] Kavzoğlu, T., Çölkesen, İ., "Karar Ağaçları İle Uydu Görüntülerinin Sınıflandırılması: Kocaeli Örneği" Harita Teknolojileri Elektronik Dergisi 2010, 2(1) 36-45

[20] Bounsaythip, C. ve Esa, R. R. (2001). "Overview of Data Mining For Customer Behavior Modeling", VTT Information Technology Research Report, Version:1, s.1-53

[21] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. The Elements of Statistical Learning. Springer: 2001, pp. 269-272

[22] Breiman L., Friedman J.H., Olshen R.A. and Stone C.J., 1984, "Classification and Regression Trees" Monterey, CA: Wadsworth, 358 s

[23] Aytaç,Ö.,Kurtdaş,MC.,2015.Sağlık-Hastalığın Toplumsal Kökenleri ve Sağlık Sosyolojisi, Social Origins of Illness-Health and Health Sociology Fırat Üniversitesi Sosyal Bilimler Dergisi Fırat University Journal of Social Science Cilt: 25, Sayı: 1, Sayfa: 231-250, ELAZIĞ-2015