Avrupa Bilim ve Teknoloji Dergisi
Özel Sayı, S. 64-73, Ekim 2019
© Telif hakkı EJOSAT'a aittir
Araştırma Makalesi

www.ejosat.com ISSN:2148-2683

European Journal of Science and Technology
Special Issue, pp. 64-73, October 2019
Copyright © 2019 EJOSAT
Research Article

Toplu Yemek Üretiminde Günlük Talep Tahmini için Alternatif Bir Yaklaşım: Öğrenci Regresyon

Derya Yergök[1*], Mehmet Acı[2]

[1] Mersin Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Mersin, Türkiye (ORCID: 0000-0002-4729-8540)
[2] Mersin Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Mersin, Türkiye (ORCID: 0000-0002-7245-8673)

(Bu yayın HORA 2019 kongresinde sözlü olarak sunulmuştur.)

(İlk Geliş Tarihi 1 Ağustos 2019 ve Kabul Tarihi 22 Ekim 2019)

(**DOI:** 10.31590/ejosat.636368)

ATIF/REFERENCE: Yergök, D. & Acı, M. (2019). Toplu Yemek Üretiminde Günlük Talep Tahmini için Alternatif Bir Yaklaşım: Öğrenci Regresyon. *European Journal of Science and Technology*, (Özel Sayı), 64-73.

Öz

Bu çalışmada, toplu yemek üretilen yerlerde günlük yemek talebini tahmin etmek ve dolayısıyla aşırı üretimden kaçınmak ve daha az gıda üretimi ile yemeklerin bitmesini önlemek amaçlanmıştır. Bunun için gauss süreç regresyon (GSR) ve regresyon ağaçları olmak üzere iki farklı regresyon modeli tasarlanmış ve modellerin başarısı karşılaştırılmıştır. Araştırmada, üniversite yemekhanesinden elde edilen veriler kullanılarak yemek yiyenlerin sayısı tahmin edilmiştir. Geliştirilen tahmin modellerinden GSR modelinde rasyonel kuadratik GSR, kare üstel GSR, matern 5/2 GSR, üstel GSR yöntemleri; regresyon ağaçları modelinde ise fine tree, medium tree, coarse tree yöntemleri kullanılmıştır. En iyi sonuçlar üstel GSR ile fine tree modellerinden elde edilmiştir.

Anahtar Kelimeler: Regresyon, Gauss süreç regresyon, Regresyon ağacı, Yemek talep tahmini.

An Alternative Approach for Daily Demand Prediction in Mass Meal Production: Regression Learner

Abstract

In this paper, it is aimed to estimate the daily demand for meal in the places where mass food is produced and thus to avoid overproduction and to prevent the end of meals with less food production. For this purpose, two different regression models were designed as gauss process regression (GPR) and regression trees and the success of the models were compared. In the study, the number of people eating was estimated using data obtained from the university dining hall. For the developed GPR prediction model rational quadratic GPR, square exponential GPR, matern 5/2 GPR and exponential GPR methods were used. In the regression trees prediction model fine tree, medium tree and coarse tree methods were used. The best results were obtained from the exponential GPR and fine tree models.

Keywords: Regression, Gauss process regression, Regression tree, Meal demand prediction.

[*] Sorumlu Yazar: Mersin Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Mersin, Türkiye, ORCID: 0000-0002-4729-8540, deryayergok@gmail.com

1. Giriş

Günümüzde gıda sektörü; gelişen sanayi, küreselleşen dünya, büyük ve küçük işletmelerin artış göstermesi, çeşitli eğitim kurumlarının ve yeni alışveriş mekanlarının açılması ile her geçen gün daha da büyümektedir. Bu büyüme ile gıda şirketleri de ivme kazanmaktadır. Sürekli işler ve organizasyonlarda catering şirketleri, alışveriş merkezlerinde restoranlar öne çıkarken, okullarda ve işyerlerinde ise yemekhaneler ön plana çıkmaktadır. Gıda pazarı gelişmesini sürdürdükçe, şirketleri etkileyen en önemli problem maliyet olmaktadır. Yemek üreten her şirket maliyetini en düşük düzeye indirmek istemektedir. Bu durumda, maliyetlerde en önemli rolü, yemek yiyen kişilerin tam sayısının belirlenmesi oynamaktadır. Yemek yiyecek kişi sayısındaki hata ile maliyet ve atıklarda artış olabileceği gibi yemek yetişmemesi gibi sorunlarla da karşılaşılabilir.

Bu çalışma ile hedeflenen, üniversite yemekhanesine ait olan verileri kullanarak, hazırlanacak yemeğe göre o gün yemek yiyecek kişi adedini tahmin etmektir. Buna bağlı olarak çok veya az yemek üretildiğinde ortaya çıkabilecek sorunların önlenmesi amaçlanmaktadır. Problemin çözümü için önerilen makine öğrenmesi yöntemlerine başvurulmuştur.

Makine öğrenmesi, matematiksel ve istatistiksel yöntemler kullanarak mevcut verilerden çıkarımlar yapan, bu çıkarımlarla bilinmeyene dair tahminlerde bulunan, modelleme ve algoritmalardan oluşan yapay zekanın bir alt dalıdır. Makine öğrenmesinin temel amacı, doğru kestirimler yapmaktır. Fakat kestirim fonksiyonlarını genellikle yorumlamak belirli bir olasılık modeli ile ilişkilendirmek zor olabilmektedir. Büyük veri analizi yapan araştırmacılar, istenilen bilgileri elde etmek ve kestirim yapmak amacıyla çeşitli makine öğrenmesi algoritmaları kullanmaktadırlar. Bu algoritmalar, kestirim yapmak için verilerin öğrenme yollarına göre denetimli (supervised) öğrenme ve denetimsiz (unsupervised) öğrenme olarak iki gruba ayrılmaktadırlar. Denetimsiz öğrenme, verilerdeki önceden bilinmeyen yapı ve ilişkileri bulmak için kullanılan yöntemleri ifade etmektedir. Denetimsiz öğrenmenin amacı, veri içindeki benzer örnek gruplarını bulmak ve veriler hakkında daha fazla bilgi edinmek için verilerdeki temel yapıyı veya dağılımı modellemektir. Denetimli öğrenme ise eğitim veri kümesine dayanan makine öğrenme modelini oluşturma sürecini ifade etmektedir. Denetimli öğrenme olarak adlandırılmasının nedeni, verilerin büyük bir kısmından oluşan eğitim verisinden bir algoritma öğrenme sürecinin bulunması ve öğrenme sürecinin test verisi ile denetlenmesidir. Denetimli makine öğrenmesi öncelikle tahmin ve kestirim problemlerine odaklanmaktadır. Denetimli makine öğrenmesi konuları regresyon ve sınıflandırma olarak iki alt grupta ele alınmaktadır. Sınıflandırma nitel değişkenlerin modellenmesi ve tahmin edilmesi için nitel kategorilere gözlemler atama ile ilgili denetimli öğrenme çeşididir. Regresyon ise, nicel değişkenlerin tahmin ve kestirimi için kullanılan denetimli öğrenmedir (Akay, 2018).

Regresyon kavramı ilk olarak Francis Galton ismindeki bir İngiliz istatistikçisi tarafından 19. yüzyılda, bir biyolojik inceleme yapılırken ortaya atılmıştır. Francis Galton bu araştırmaları sırasında regresyonu kavram ve yöntem olarak geliştirmiştir. Karl Pearson, R.A. Fisher ve Udny Yüle bu yöntemi daha genel istatistiksel alanlara uyarlayarak geliştirmişlerdir. Adrien Marie Legendre tarafından 1805'te bulunan En Küçük Kareler Yöntemi regresyon yönteminin ilk şeklidir. Daha sonra 1809'da C.F. Gauss aynı yöntemi açıklamıştır. Bu iki bilim adamı, Güneşin etrafındaki uyduların yörüngelerini tespit etmek için astronomi gözlemleri yaparken En Küçük Kareler yöntemini ortaya çıkartmışlardır. Regresyon Analizi; bağımlı (dependent) ve bağımsız (independent) değişkenler arasındaki sayısal ilişkiyi incelemek amacıyla kullanılan bir yöntemdir. Bağımlı değişken (açıklanan değişken, yanıt değişkeni, sonuç değişkeni olarak da adlandırılır) bağımsız değişkenden etkilendiği kabul edilen değişkendir. Bağımsız değişken (açıklayıcı değişken, etken değişken olarak ta adlandırılır) ise bağımlı değişkeni etkilediği düşünülen değişkendir. Regresyon Analizinde; bir tane bağımlı ve bir tane de bağımsız değişken var ise Basit Regresyon Analizi, bir tane bağımlı değişken, birden fazla bağımsız değişken var ise Çoklu Regresyon Analizi, birden fazla bağımlı değişken var ise Çok Değişkenli Regresyon analizi yöntemleri uygulanır. Değişkenler arasındaki ilişki doğrusal ise Doğrusal Regresyon Analizi, değilse Eğrisel Regresyon Analizi olarak adlandırılır (Deniz ve Koç, 2019).

Gauss süreç regresyon (GSR) modelleri çok değişkenli dağılımlı rasgele değişkenlerin sınırlı bir koleksiyonuna sahip parametrik olmayan çekirdek tabanlı olasılık modelleridir. Her doğrusal kombinasyon eşit dağılmıştır. Gaussian süreçleri, Carl Friedrich Gauss'un ismini almıştır çünkü Gauss dağılımının çok değişkenli normal dağılımların sonsuz boyutlu bir genellemesi olduğu kavramına dayanmaktadır. Gauss süreçleri istatistiksel modelleme, çoklu hedef değerlere regresyon ve haritalamanın daha yüksek boyutlarda analiz edilmesinde kullanılır (Zhang ve ark., 2018). GSR çok değişkenli normal dağılımların sonsuz boyutlu genellemesi ile parametrik olmayan çekirdek bazlı olasılık teknikleridir. Gauss süreçleri istatistiksel modellemede, çoklu hedef değerlere regresyonda ve haritalamanın daha yüksek boyutlarda analizinde kullanılır. Farklı çekirdekli dört değişik modeli vardır (Zhang ve Leatham, 2018). Gauss süreçleri, veri modellemesi için esnek bir parametrik olmayan araç sağlayan stokastik süreçlerin bir ailesidir. En temel ayarda, bir Gauss süreci, sınırlı bir gözlem kümesine dayanan gizli bir işlevi modellemektedir. Gauss süreci, çok değişkenli bir Gauss dağılımının sonsuz sayıda boyuta uzatılması olarak görülebilir, burada herhangi bir sonlu boyut kombinasyonu, ortalama ve kovaryans fonksiyonlarını tamamen belirten çok değişkenli bir Gauss dağılımına neden olur. Ortalama ve kovaryans fonksiyonunun seçimi (aynı zamanda çekirdek olarak da bilinir), ilgilenilen gizli fonksiyona düzgünlük varsayımları uygular ve ilgili gözlem verisi noktaları arasındaki, X öklid mesafenin bir fonksiyonu olarak Y çıktı gözlemleri arasındaki korelasyonu belirler (Fairbrother ve ark., 2018).

Araştırmada kullanılan veri kümesi, test ve eğitim veri kümesi olarak iki gruba ayrılmıştır ve öngörülen model eğitim verisi ile eğitildikten sonra test verileri ile tahminler elde edilmiştir. Bu işlemler MATLAB yazılım ortamında yapılmıştır. MATLAB'in istatistik ve makine öğrenmesi araç kutusunda bulunan Regression Learner App ile denetimli makine öğrenmesini kullanarak verileri tahmin etmek için önerilen regresyon modelleri eğitilmiştir.

Çalışmada dört farklı GSR modeli ve üç farklı regresyon ağacı modeli tasarlanmıştır. GSR modelinde rasyonel kuadratik, kare üstel, matern 5/2 ve üstel yöntemleri ve kernel fonksiyonları kullanılmıştır. Yöntemler arasında en iyi sonuç üstel GSR ile elde edilmiştir. Regresyon ağacı modelindeki yöntemler; fine tree, medium tree, coarse tree yöntemleridir. Regresyon ağacı modelinde kullanılan minimum yaprak sayısı fine tree için 1, medium tree için 3 ve coarse tree için 9 şeklindedir. Yöntemler arasında en iyi sonuç fine tree ile elde edilmiştir.

Önerilen modellerin doğruluğu; Ortalama Kare Hata (OKH), Kök Ortalama Kare Hata (KOKH), Ortalama Mutlak Hata (OMH) ve Çoklu Korelasyon Katsayısı (ÇKK) ile belirlenmiştir.

2. Materyal ve Metot

2.1. Veri Kümesi

Veri kümesi 241 adet satır ve 9 adet sütun değerlerinden oluşmaktadır. Bu değerler kullanılmadan önce, içindeki metin sayısal değerlere dönüştürülmüştür. Ek olarak, benzer veriler için bir gruplandırma işlemi gerçekleştirilmiştir. Bu şekilde, ağın öğrenme yeteneği arttırılmıştır. Veri kümesinde 9 özellik bulunmaktadır. Bunlardan 8 tanesi (hafta günleri, maaş günü, çorbalar, ana yemek, yan yemek, ramazan, tatil ve sınav haftası) girdi verisi ve 1(kişi sayısı) çıktı verisidir. Tablo 1, nitelik içeriğinin bir örneğini göstermektedir. Sayısal değerlere dönüştürülen bir veri kümesi örneği Tablo 2' de verilmiştir.

Tablo 1. Veri Kümesi İçeriğinin Bir Örneği

Nitelik	**Sayı Aralığı**	**Açıklama**
Haftanın Günleri (A)	1-5	Pazartesi, Salı, Çarşamba, Perşembe, Cuma
Maaş Günü (B)	0-1	Maaş günü ya da değil
Çorbalar (C)	0-9	Tavuk çorbası, Mercimek çorbası vs.
Ana Yemek (D)	0-14	Et yemeği, Tavuk yemeği vb.
Yan Yemek (E)	0-5	Pirinç, Soslu makarna vb.
Ramazan (F)	0-1	Ramazan ya da değil
Tatil (G)	0-1	Tatil ya da değil
Sınav Haftası (H)	0-1	Sınav haftası ya da değil
Kişi Sayısı (I)	0-9	Yemekhanede yemek yiyen kişi sayısı

Tablo 2. Veri Kümesi Örneği

A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
3	0	3	3	1	0	1	0	1
4	0	3	12	0	0	1	0	2
5	0	5	2	4	0	1	0	2
1	0	0	5	4	1	1	1	0
2	0	3	10	1	1	1	1	0
3	0	3	5	1	1	1	1	0
4	0	5	14	5	1	1	1	0

2.2. Rasyonel Kuadratik GSR Model

Rasyonel Kuadratik GSR çekirdeği, farklı ölçeklerde değişen verileri modellememize izin verir. Rasyonel Kuadratik GSR algoritması uzaysal istatistiklerde, geostatistik, makine öğrenmesi, görüntü analizi ve metrik uzaylar üzerinde çok değişkenli istatistiksel analizlerin yapıldığı diğer alanlarda kullanılır (Zhang ve ark., 2018).

2.3. Kare Üstel GSR Model

Kare Üstel GSR, sonsuz sayıda temel fonksiyona sahip radyal temelli fonksiyon regresyon modelinin bir fonksiyon uzay ifadesidir. Kare Üstel GSR, Öklid mesafesinin kare olması dışında Üstel GSR ile aynıdır. Kare Üstel GSR'yi kullanan etkileyici bir özellik, temel işlevlerin iç ürünlerini çekirdeklerle değiştirmesidir. Ayrıca, süreksizlikleri iyi idare eder (Zhang ve ark., 2018).

2.4. Matern 5/2 GSR Model

Matern 5/2 çekirdeği, sabit çekirdeğin spektral yoğunluklarını alır ve radyal tabanlı fonksiyon (RBF) çekirdeğinin Fourier dönüşümlerini oluşturur. Matern 5/2 çekirdeği, yüksek boyutlu uzaylar için ölçüm problemlerinin konsantrasyonuna sahip değildir (Zhang ve ark., 2018).

2.5. Üstel GSR Model

Üstel GSR, Öklid mesafesinin kare olmaması dışında, Kare Üstel GSR ile aynıdır. Üstel GSR, temel işlevlerin iç ürünlerini, Kare Üstel GSR'den daha yavaş olan çekirdeklerle değiştirir. Üstel GSR düzgün işlevlerini en az hatayla iyi bir şekilde işler, ancak süreksizliklerle iyi şekilde başa çıkamaz (Zhang ve ark., 2018).

2.6. Regresyon Ağacı Modeli

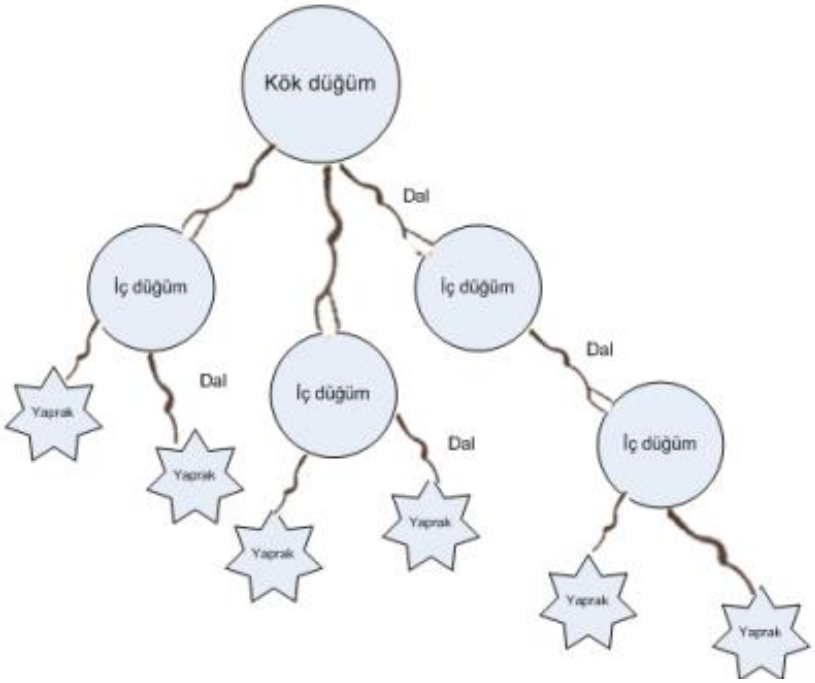
Ağaç yapılı sınıflandırma ve regresyon, geçen on yılda popülerliği büyük ölçüde artmış olan, parametrik olmayan, hesaplama açısından yoğun yöntemlerdir. Hem çok sayıda vaka hem de çok sayıda veri içeren veri kümelerine uygulanabilir. Değişkenler ve aykırı değerlere karşı son derece dirençlidirler (Sutton, 2005).

Sınıflandırma ve regresyon ağacı yöntemi, analiz edilen verilerin özelliğine göre farklı isimlerle anılmaktadır. Bağımlı değişkenin kategorik olması durumunda Sınıflandırma Ağacı, sürekli değişken olması durumunda da Regresyon Ağacı kullanılmaktadır (Kayri ve Boysan, 2008).

Karar ağaçları (KA), bir problemi oluşturan veri setlerinin yapısına göre bir ağaç yapısı şeklinde sınıflandırma ve regresyon modelleri oluşturmaktadır. Söz konusu ağaç yapılarının oluşturulmasında kullanılan karar kurallarının anlaşılabilir olması yöntemin kullanımını yaygın hale getirmiştir. KA sınıflandırma ve regresyon probleminin çözümünde çok aşamalı ve ardışık bir yaklaşım ile karmaşık yapıdaki verileri aşamalı bir hale dönüştürerek basit bir karar verme işlemini gerçekleştirmektedir. Karar ağacı yöntemi ile sınıflandırma veya regresyon işlemi neticesinde tahmin edilecek hedef öznitelikler ayrık verilerden veya belirli kategorilerden oluşuyorsa kullanılan model sınıflandırma ağacı, öznitelik verileri sürekli değişkenlerden oluşuyorsa model regresyon ağacı olarak adlandırılmaktadır (Kavzoğlu ve ark., 2012).

Regresyon ağaçları, sürekli bir hedef değişkeni ile ilgilenen özel bir karar ağacı türüdür. Bu yöntemler, verimli bir özyinelemeli bölümleme algoritması aracılığıyla indüksiyon gerçekleştirir. Ağacın her bir düğümdeki test seçimi genellikle en küçük kareler hata kriteri ile yönlendirilmektedir. İkili ağaçlar, her bir ağaç düğümünde iki yönlü bölünme olduğunu düşünür (Torgo, 1997).

KA, verilen bir problemin yapısına bağlı olarak bir ağaç yapısı şeklinde sınıflandırma ve regresyon modeli oluşturmaktadır. Ağaç yapıları kurallardan meydana gelmektedir. Kuralların anlaşılabilir olması yöntemin kullanımını kolay ve uygulanabilir kılmıştır. KA, bir problemin çözümünde, basit bir karar verme işlemini, çok aşamalı ve ardışık işlemlerden sonra gerçekleştirmektedir. Sınıflandırma ağacında tahmin edilecek öznitelikler ayrık verilerden oluşurken, regresyon ağacında öznitelik verileri sürekli değişkenlerden oluşmaktadır. Basit bir regresyon ağacı yapısı Şekil 1'de gösterilmiştir. Bu yapıda her bir öznitelik bir düğüm tarafından temsil edilmiştir. Ağaç yapısının en üst kısmı kök, en alt kısmı yapraklardan oluşurken; kök, yaprak ve düğümler arasındaki ilişkiler dal olarak tanımlanmaktadır. KA yapısı oluşturulurken verilere bir takım sorular sorularak karar kuralları oluşturulur. Sorular kök düğümünde sorulmaya başlanır, yapraklara kadar ağacın dallanması devam eder (Şengür ve Tekin, 2013).



Şekil 1. Basit bir regresyon ağacı yapısı (Şengür ve Tekin, 2013)

MATLAB'in araç kutusunda bulunan Regression Learner uygulamasında, üç tip regresyon ağacı vardır: fine tree, medium tree, coarse tree. Fine tree ve medium tree benzer KOKH'lere sahipken, coarse tree daha az doğrudur.

3. Bulgular ve Tartışma

Tüm tahmin modelleri ölçülen beş performans açısından değerlendirildi, (1) ÇKK değeri, hedef ve öngörülen değerler arasındaki korelasyonu ölçmek için, (2) OKH ve (3) KOKH değerleri, hataların karelerinin ortalamasını ölçmek için, (4) OMH değeri ise tahminlerin hedef değerlere yakınlığını ölçmek için kullanıldı. Bu performans ölçütlerinin denklemleri, sırasıyla (1), (2), (3) ve (4) denklemlerinde verilmektedir (Witten ve Frank, 2005).

$$ÇKK = \sqrt{1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - O_m)^2}} \quad (1)$$

$$OKH = \frac{1}{n}[\sum_{i=1}^{n}(O_i - P_i)^2] \quad (2)$$

$$KOKH = \sqrt{\frac{1}{n}[\sum_{i=1}^{n}(O_i - P_i)^2]} \quad (3)$$

$$OMH = \frac{1}{n}\sum_{i=1}^{n}|O_i - P_i| \quad (4)$$

Burada *n,* test için kullanılan veri noktası sayısı ise, P_i öngörülen değer, O_i gözlenen değer ve O_m gözlemlenen değerlerin ortalamasıdır.

Tablo 3 ve 4'te verilen sonuçlara göre, aşağıdaki sonuçlar elde edilmiştir:

- GSR modeli için en iyi sonuçlar üstel GSR yönteminde KOKH 1.1231, OKH 1.2613, OMH 0.8772 ve ÇKK 0.90 bulunmuştur.
- Regresyon ağacı modeli için en iyi sonuç fine tree yönteminde KOKH 1.0327, OKH 1.0664, OMH 0.73486 ve ÇKK 0.9165 bulunmuştur.

Tablo 3. GSR Modeli ile Elde Edilen Sonuçlar

Modeller	**KOKH**	**OKH**	**OMH**	**ÇKK**
Rasyonel Kuadratik	1.3659	1.8657	1.0723	0.8485
Kare Üstel	1.3659	1.8657	1.0723	0.8485
Matern 5/2	1.3588	1.8464	1.0680	0.8485
Üstel	**1.1231**	**1.2613**	**0.8772**	**0.90**

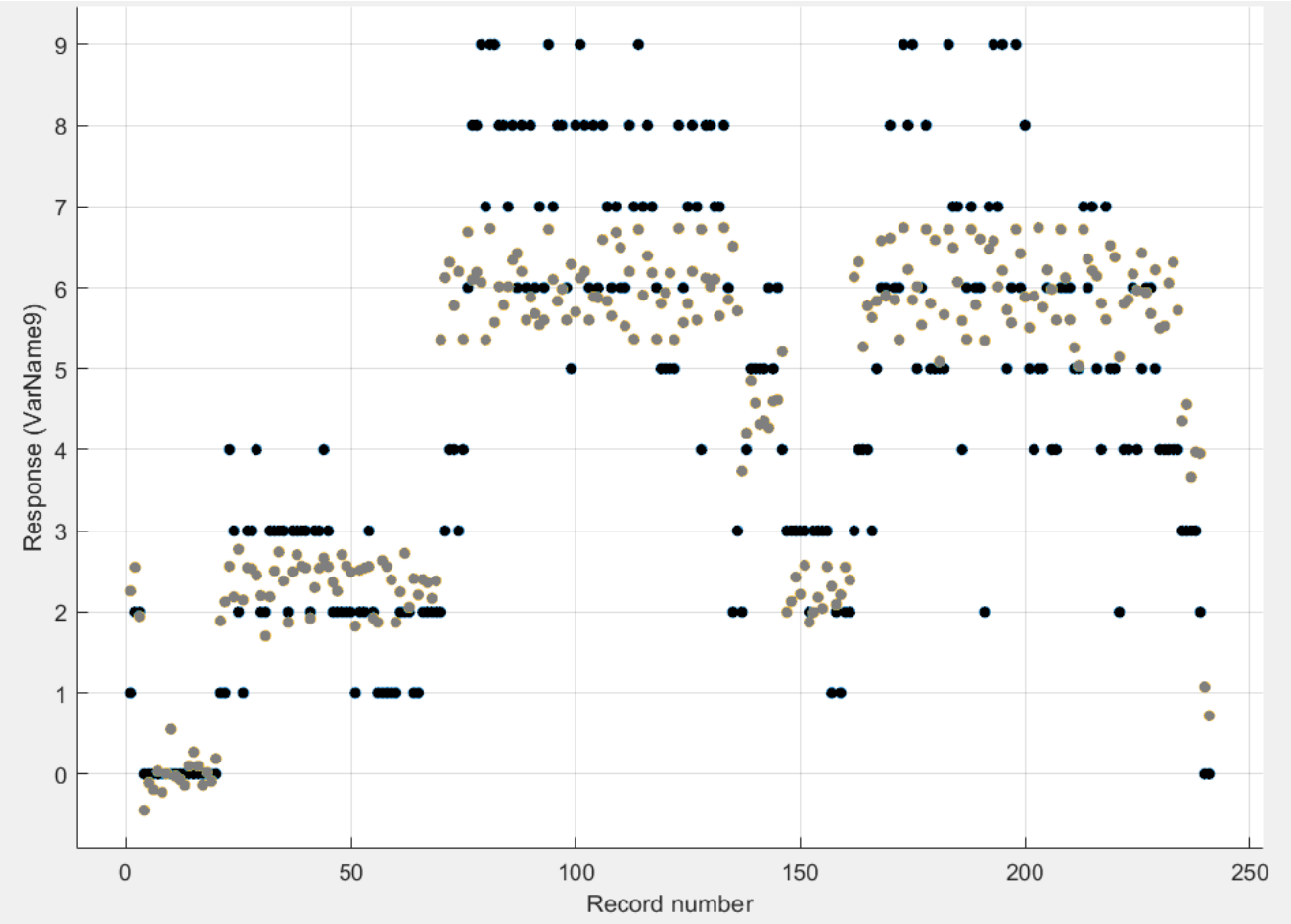
Tablo 4. Regresyon Ağacı Modeli ile Elde Edilen Sonuçlar

Modeller	**KOKH**	**OKH**	**OMH**	**ÇKK**
Fine Tree	**1.0327**	**1.0664**	**0.73486**	**0.9165**
Medium Tree	1.1043	1.2196	0.81627	0.90
Coarse Tree	1.2943	1.6752	0.93261	0.8660

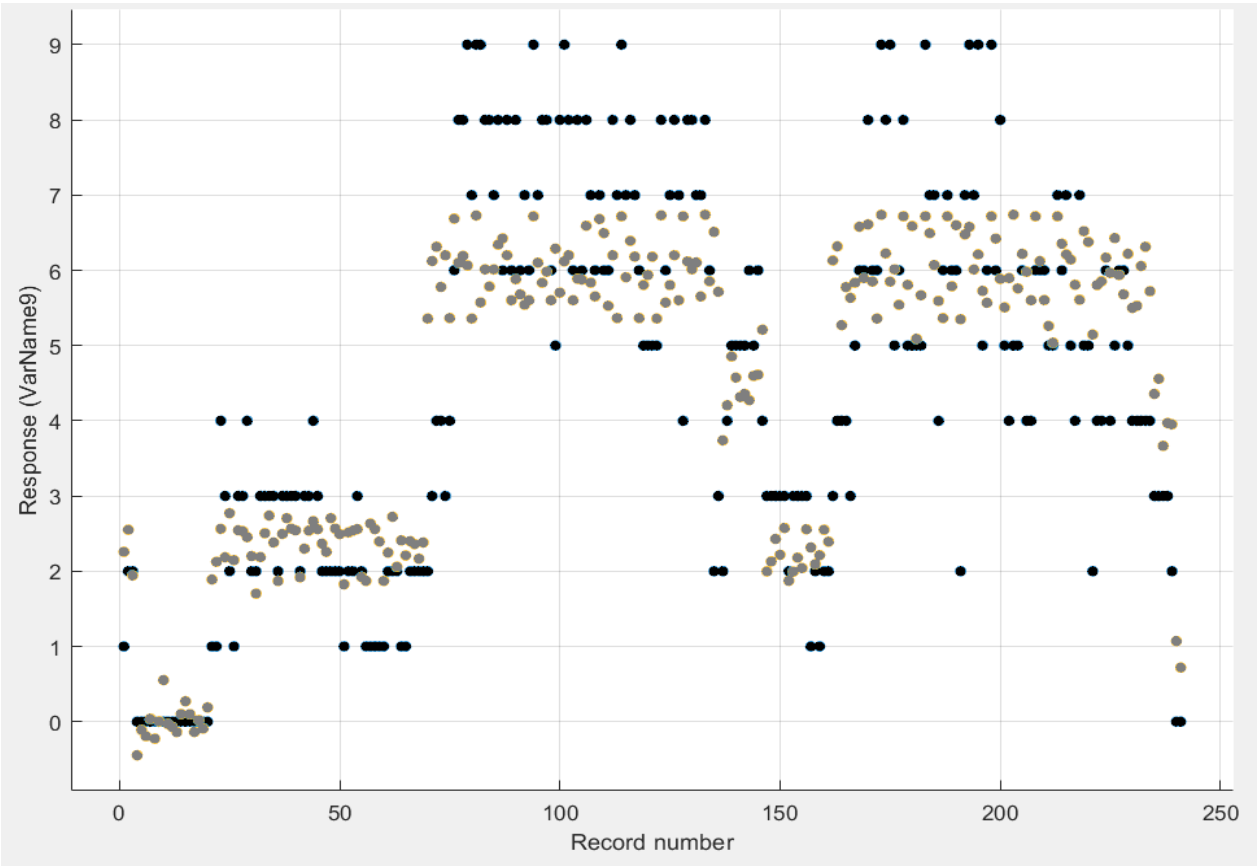
GSR modelindeki yöntemlerin, veri kümesini eğitme süreleri rasyonel kuadratik için 6.4832, kare üstel için 5.6001, matern 5/2 için 6.517 ve üstel için 6.3179 saniyedir. Regresyon ağacı modelindeki yöntemlerin, veri kümesini eğitme süreleri ise fine tree için 1.1545, medium tree için 0.85503 ve coarse tree için 0.83487 saniyedir. Bu süreler her iki model için de veri kümesinin ilk eğitildiği zamanki değerlerdir. Yöntemler tekrar tekrar çalıştırıldığında eğitim süreleri daha kısa olmaktadır.

GSR modelindeki yöntemler için MATLAB'in mevcut parametre değerleriyle oynanmadan yüksek sonuçlara ulaşılmıştır. Modellerin parametreleri veri kümesine ve probleme uygun olarak değiştirildiğinde daha iyi sonuçlar elde edildiği görülmüştür. Regresyon ağacı modelindeki yöntemlerde fine tree için minimum yaprak sayısı MATLAB'de varsayılan olarak 4, medium tree için 12 ve coarse tree için 3'dır. Çalışmada bu parametreler fine tree için 1, medium tree için 3 ve coarse tree için 9 yapılmıştır.

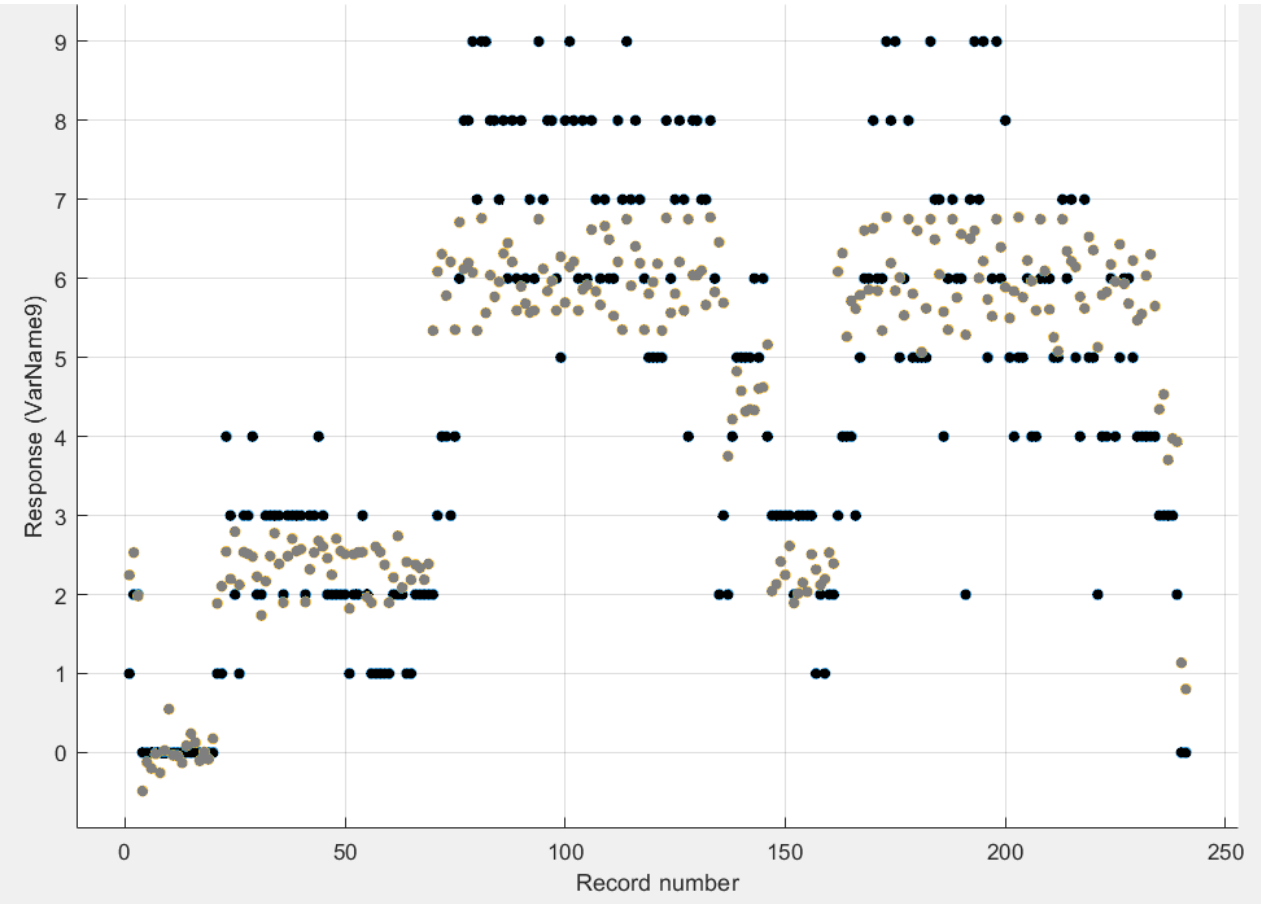
Veri kümesi 241 satır ve 9 sütundan oluşmaktadır. 9. Sütun sonuç sütunudur, yemek yiyecek kişi sayısının tahmini değerlerini bulundurur. Şekil 2, 3, 4, 5, 6, 7 ve 8 GSR ve regresyon ağacı modellerinin performans grafiklerini göstermektedir. Grafiğin X ekseni 241 satırdan oluşan tüm veriyi, Y ekseni sonuç verisini göstermektedir. Siyah ile gösterilen alanlar gerçek değerleri, gri ile gösterilen alanlar tahmin edilen değerleri ifade etmektedir. Üstel GSR ve fine tree modellerinde en iyi performans sonuçlarına ulaşılmıştır.



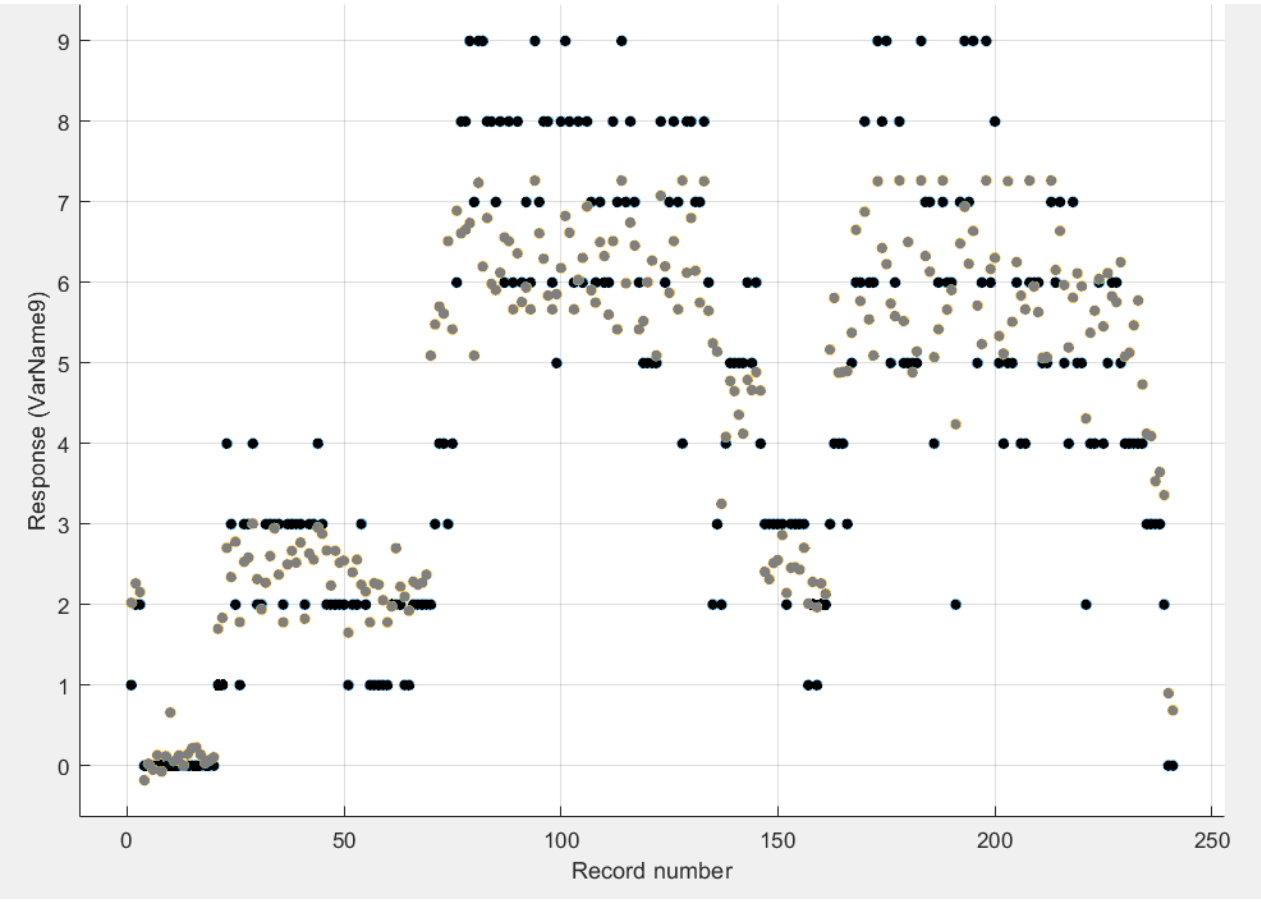
Şekil 2. Rasyonel kuadratik GSR modelinin cevap grafiği



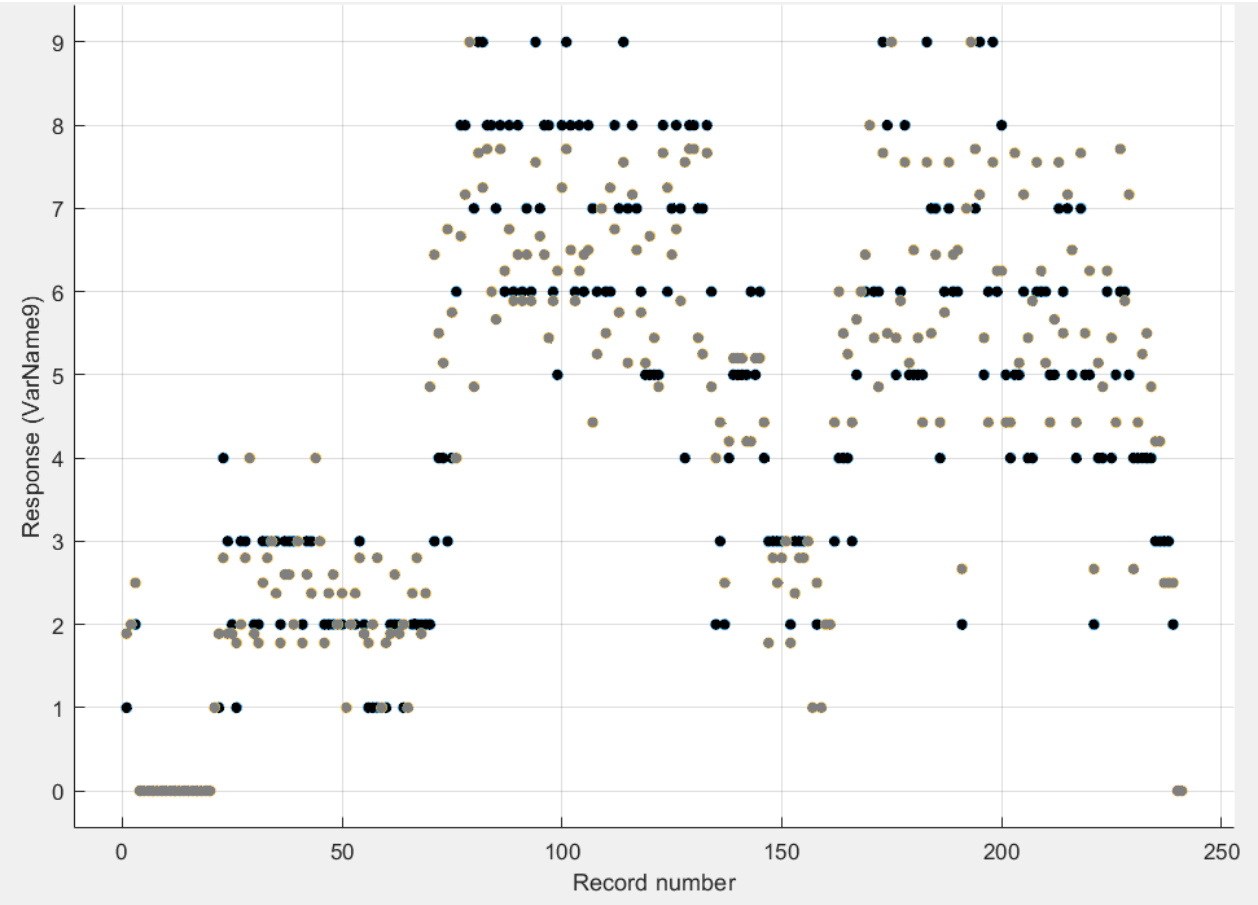
Şekil 3. Kare üstel GSR modelinin cevap grafiği



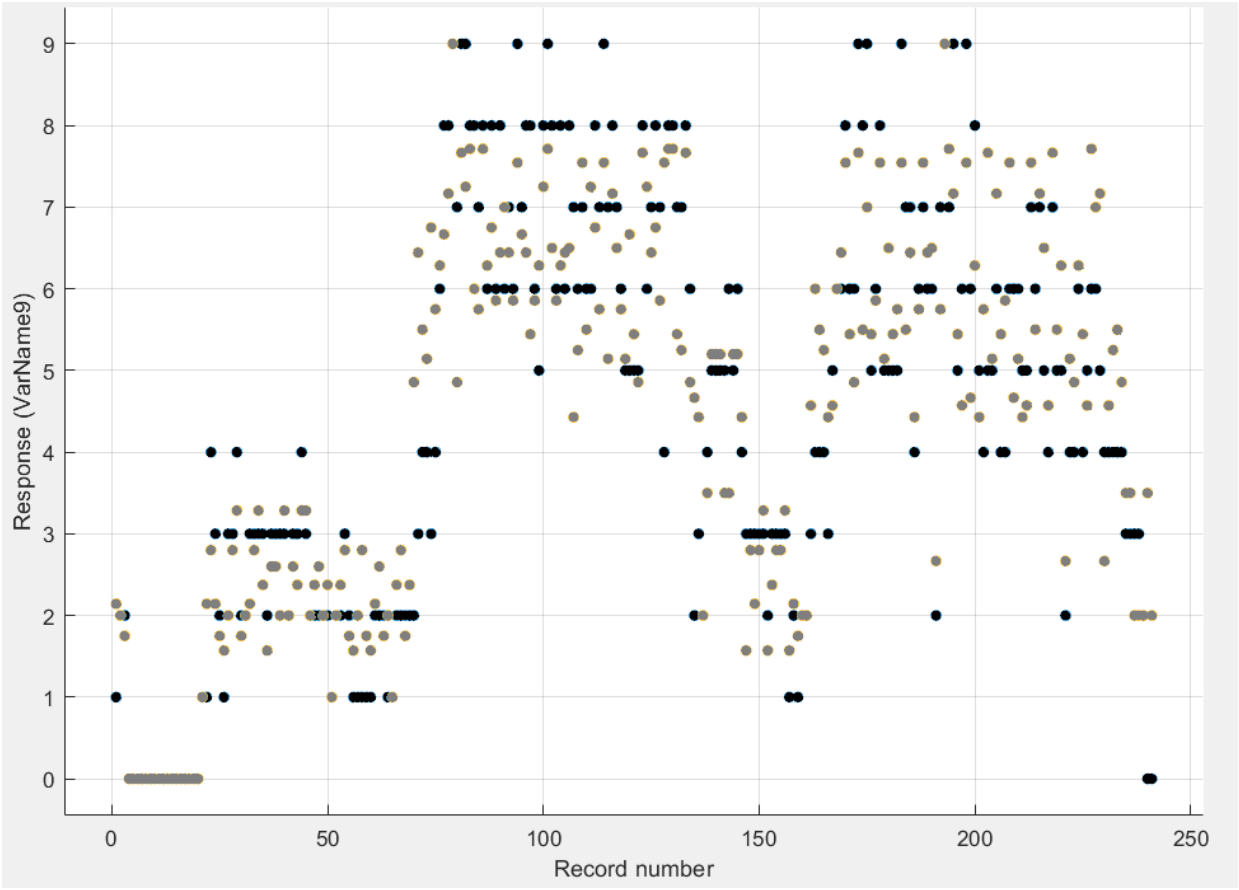
Şekil 4. Matern 5/2 GSR modelinin cevap grafiği



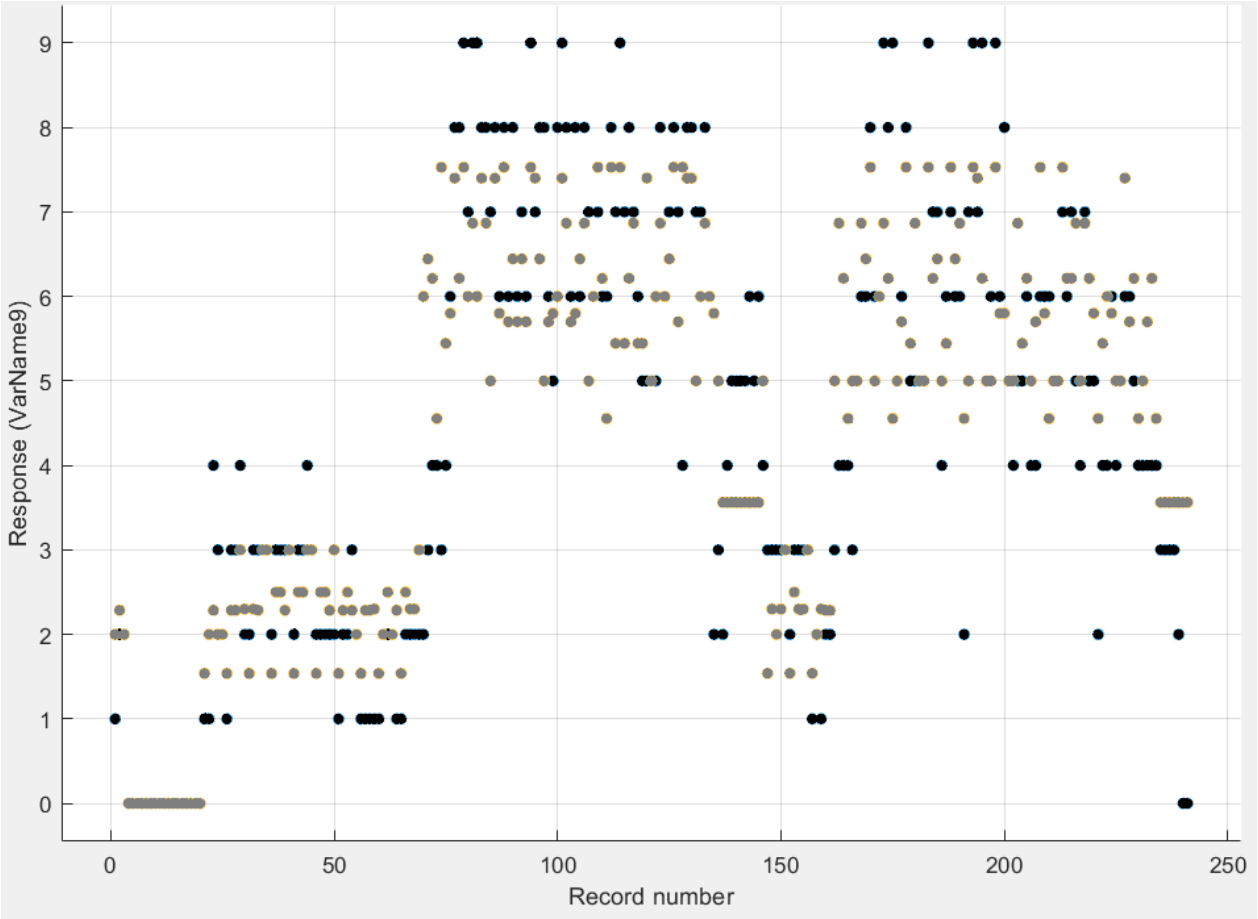
Şekil 5. Üstel GSR modelinin cevap grafiği



Şekil 6. Fine tree regression modelinin cevap grafiği



Şekil 7. Medium tree regression modelinin cevap grafiği



Şekil 8. Coarse tree regression modelinin cevap grafiği

4. Sonuçlar ve Öneriler

Bu çalışmada, üniversite yemekhanesinden elde edilen veriler kullanılarak hazırlanacak yemeğe göre yemek yiyecek kişi sayısını öngörmek için iki farklı regresyon modeli tasarlanmış ve tatmin edici sonuçlar elde edilmiştir. İki model ve toplam 7 yöntem arasından üstel GSR ve fine tree yöntemlerinin diğer yöntemlerden daha başarılı olduğu sonucuna ulaşılmıştır.

Başarılı sonuçlar, mevcut yöntemlerin iyileştirilebileceğini ve farklı yöntemler/modeller kullanılarak daha iyi sonuçlar alınabileceğini göstermektedir.

5. Teşekkür

Bu çalışma Mersin Üniversitesi Bilimsel Araştırma Projeleri Birimince 2019-1-TP2-3215 Proje Numarası ile desteklenmiştir.

Kaynakça

Akay, Ç. E. (2018). Ekonometride Yeni Bir Ufuk: Büyük Veri ve Makine Öğrenmesi. *Sosyal Bilimler Araştırma Dergisi*, *7*(2), 41-53.

Deniz, G, & Koç, S. (2019). Türkiye'de Ekonomik Büyüme ile Bazı Makro Değişkenler Arasındaki İlişki: Çoklu Doğrusal Regresyon Modeli Analizi.

Zhang, N, Xiong, J, Zhong, J, & Leatham, K. (2018, June). Gaussian Process Regression Method for Classification for High-Dimensional Data with Limited Samples. In *2018 Eighth International Conference on Information Science and Technology (ICIST)* (pp. 358-363). IEEE

Zhang, N, & Leatham, K. (2018, December). Neurodynamics-Based Nonnegative Matrix Factorization for Classification. In *International Conference on Neural Information Processing* (pp. 519-529). Springer, Cham.

Fairbrother, J, Nemeth, C, Rischard, M, & Brea, J. (2018). GaussianProcesses. jl: A Nonparametric Bayes Package for the Julia Language. *arXiv preprint arXiv:1812.09064*.

Sutton, C. D. (2005). Classification and Regression Trees, Bagging, and Boosting. *Handbook of statistics*, *24*, 303-329

Kayri, M, & Boysan, M. (2008). Bilişsel Yatkınlık ile Depresyon Düzeyleri İlişkisinin Sınıflandırma ve Regresyon Ağacı Analizi ile İncelenmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *34*(34), 168-177.

Kavzoğlu, T, Şahin, E. K, & Çölkesen, İ. (2012). Heyelan Duyarlılığının İncelenmesinde Regresyon Ağaçlarının Kullanımı: Trabzon Örneği. *Harita Dergisi*, *147*(3), 21-33.

Torgo, L. (1997, July). Functional Models for Regression Tree Leaves. In *ICML* (Vol. 97, pp. 385-393).

Şengür, D, & Tekin, A. (2013). Öğrencilerin Mezuniyet Notlarının Veri Madenciliği Metotları ile Tahmini. *Bilişim Teknolojileri Dergisi*, *6*(3), 7-16.

Witten, I. H, Frank, E, Hall, M. A, & Pal, C. J. (2016). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann.