

PISA 2018 Fen Bilimleri Puanlarının Değerlendirilmesinde Eğitsel Veri Madenciliğinin Kullanımı

Esra Uğuz^{*1}, Seren Şahin², Ramazan Yılmaz³

Anahtar Sözcükler

PISA 2018
Karar Ağaçları
Eğitsel Veri
Madenciliği
Veri Analizi
Fen Bilimleri
Makale Hakkında
Gönderim Tarihi
26 Şubat 2021
Kabul Tarihi
03 Ağustos 2021
Yayın Tarihi
24 Aralık 2021
Makale Türü
Araştırma Makalesi

Öz

Bu çalışma, PISA 2018 fen bilimleri puanlarının; anne-baba eğitim durumu, fen öğrenmek için haftalık harcanan zaman, okulda bilgi iletişim teknolojileri (BİT) kullanımı ve öğrenci algılanan BİT yeterliği değişkenlerinden yararlanılarak veri madenciliği algoritmalarından olan karar ağaçları ile değerlendirilmesini amaçlamaktadır. Ayrıca veri madenciliği analiz programı Rapid Miner ile 6890 öğrenciden oluşan Türkiye örneklemini kullanılarak fen bilimleri puanı bağımlı değişkeninin bağımsız değişkenler ile arasındaki ilişkiye K-nn, naive bayes ve random forest algoritmaları kullanılarak bakılmıştır. Veri analizi öncesinde fen puanı sonuçlarının normal dağılımı bozulmayacak şekilde kayıp veri temizliği yapılmış, analiz 6001 veri üzerinden gerçekleştirilmiştir. Anne-baba eğitim durumunun fen puanı başarısında anlamlı bir farka sahip olmadığı, fen öğrenmek için haftalık harcanan zaman ve okulda BİT kullanımının fen puanı başarısında pozitif bir etkiye sahip olduğu, algılanan BİT yeterliğinin ise fen puanı başarısı ile arasında negatif bir ilişki tespit edilmiştir. Seçilen bağımsız değişkenlerin öğrenci fen başarı durumunu tahmin oranı K-nn algoritmasında %77, naive bayes algoritmasında %55.06, random forest algoritmasında ise %62.22 olarak saptanmıştır

The Use of Educational Data Mining in the Evaluation of PISA 2018 Scores of Science

Keywords

PISA 2018
Decision Trees
Educational Data
Mining
Data Analysis
Science Lesson

Article Info

Received
February 26, 2021
Accepted
August 03, 2021
Published
December 24, 2021
Article Type
Research Paper

Abstract

This study is based on PISA 2018 science scores; it aims to evaluate the educational status of parents, the time spent for science learning, the use of ICT at school, and the perceived ICT proficiency variables, with data mining algorithms. Data mining analysis programs Rapid Miner with 6890 students of relations between Turkey sample science with variable K-nn, evaluating their naive Bayes and research methods with random forest algorithms and decision methods were evaluated. When the data analysis science score was tested normal, normal test results were given, the analysis was performed with 6001. There was a significant difference in the success of the science score of the parents' education status, the time spent to learn science and the ICT use score at school had a negative relationship with science achievement. The rate of predicting student science achievement of selected independent variables was 77% in K-nn algorithm, 55.06% in naive bayes algorithm and 62.22% in random forest algorithm.

Atıf: Uğuz, E., Şahin, S., & Yılmaz, R. (2021). PISA 2018 fen bilimleri puanlarının değerlendirilmesinde eğitsel veri madenciliğinin kullanımı. *Bilgi ve İletişim Teknolojileri Dergisi*, 3(2), 212-227.

<https://doi.org/10.53694/bited.887425>

Cite: Uğuz, E., Şahin, S., & Yılmaz, R. (2021). The use of educational data mining in the evaluation of PISA 2018 scores of science. *Journal of Information and Communication Technologies*, 3(2), 212-227.

<https://doi.org/10.53694/bited.887425>

* Sorumlu Yazar/Corresponding Author: esra_uguz_1982@hotmail.com

¹ M.Sc. Student, Bartın University, Science Faculty, Bartın/Turkey, esra_uguz_1982@gmail.com, <https://orcid.org/0000-0001-5137-6444>

² M.Sc. Student, Bartın University, Science Faculty, Bartın/Turkey, serender.trabzon@gmail.com, <https://orcid.org/0000-0002-2740-6217>

³ Assoc. Prof. Dr., Bartın University, Science Faculty, Bartın/Turkey, ramazanyilmaz067@gmail.com, <https://orcid.org/0000-0002-2041-1750>

Extended Abstract

Introduction

PISA (Program for International Student Assessment); it is large-scale educational research that brings together academic skills such as reading skills, mathematics, science, financial literacy, as well as student's perception and attitude, family-school and environment. (Akyüz & Pala, 2010; Kamaliyah, Zulkardi, & Darmawijoyo, 2013). First organized in 2000 after the 2003 PISA exam Turkey has also participated regularly (Anchor, 2015). The last PISA exam was held in 2018. 15-year-old students living in 79 countries and economies and still attending a school participated in the study. PISA survey conducted in 2018 have been included in the 79 countries with 37 OECD countries, Turkey 454 math scores at 42, 39 and 466 points in 468 points reading science was ranked 40th (OECD, 2019a). Each assessment lasted two hours, and multiple choice and open-ended questions were asked. They also completed a 35-minute questionnaire asking for information about themselves, their homes, schools, and families (OECD, 2019c). Exam scores and survey results obtained in this study are effective in guiding the education policies of countries. For this reason, the analysis of exam and survey results is of great importance in determining the reasons for student success and failure, increasing student success, detecting defects in education and training environments, and creating more effective education and training environments.

Making sense of large amounts of raw data with analysis programs and using them for predicting the future is called data mining (Thuarisingham, 2003). In the field of education, with the use of mass media tools, large data heaps have been formed. Finding and processing data to be used for the benefit of all education stakeholders among these raw data will play a role in the arrangement of new generation educational environments (Özbay, 2015). Using data mining methods, PISA exam, which is one of the international large-scale exams, and making meaningful results from quite a lot of data about students and schools will also help to increase the education quality of countries.

When the literature is examined, it is seen that perceptions and attitudes are included in the analysis results along with academic success in large-scale exams such as PISA and TIMSS (Güzel, İş, & Berberoğlu, 2010). It was determined in the study by Pajares and Miller (1994) that there is a positive relationship between academic success and efficacy beliefs. While determining the independent variables of this study, the literature was used at this point. The effect of information communication technologies (ICT) efficacy perception on the science exam score was examined.

With this study, which of the independent variables of the mother's education status, father's education status, time spent to learn science, the use of ICT at school and the perceived ICT proficiency of the student, which are thought to have an effect on the science exam results, and the accuracy of the science exam result using these variables intended. It was investigated how students are classified in terms of their achievement status according to the variables of mother's education status, father's education level, time spent to learn science, use of ICT at school, and perceived ICT competence variables, which are accepted as independent variables. It is thought that the data mining methods used in this study will guide the design of new generation learning environments in the field of education, contribute to the determination of areas for needs, and can be turned into opportunities by determining the causes of failure. The success of predicting the dependent variable of the independent variables will contribute to the literature with the correct choice of variable selection.

Method

PISA 2018 science exam results and questionnaire data answered by students were downloaded from oecd.org/pisa and the variables were determined by examining. Then, the data to be analysed was extracted and analysed using data mining methods. Suggestions were made based on the findings and results obtained as a result of the analysis.

Findings

It seems that the majority of them cluster at 2.5 or below the average maternal education level. It is understood that the success rate of those whose mother's education level is between 2.5-3.5 is higher than those with 2.5 and below. However, when looking at the decision tree in general, the increase in mother's education level does not make a significant difference in science score success.

It is seeming that the majority of them cluster at 1.5 and above the average education level of fathers. It is understood that the success rate of those whose father's education level is between 3.5-4.5 is higher than those with 3.5 and below. However, when looking at the overall decision tree, the increase in father's education level does not make a significant difference in science score success.

It seems that the majority stacked 77.5 minutes above the time spent learning science per week. Although there is a positive relationship between the length of study period and science achievement, it is not significant enough to make a difference.

It is seen that the ICT scale used by the majority at school is stacked at 1.616 and above. Looking at the decision tree in general, it can be said that the increase in the ICT scale used in school positively affects the success status.

It is seen that the majority cluster above the perceived ICT competence scale -1.566. It can be said that all those with an ICT scale above 1,976 are successful, and there is a negative correlation in science score achievement, although it is not significant when looking at the overall decision tree.

When the test data with $k = 9$ and train were parsed according to a certain rate, he predicted the science score success with 77% accuracy using the data of mother's education status, father's education level, time spent to learn science per week, bit used in school, perceived bit efficacy data.

Naive Bayes algorithm; he predicted the science score success with 55.06% accuracy by using the data of mother's education status, father's education status, time spent to learn science, bit used at school, perceived bit efficacy.

Random Forest algorithm: He predicted science score success with 62.22% accuracy by using the data of mother's education status, father's education level, time spent to learn science, bit used at school, perceived bit efficacy.

Discussion and Conclusion

Since this study aims to evaluate PISA 2018 science score results with educational data mining methods, k-nn, naive bayes, random forest algorithms were used to determine the adequacy of the science score success of selected independent variables in the study. The K-nn algorithm correctly predicted the science score success by 77% using the variables of parents' education status, time spent learning science per week, use of ICT at school, and perceived ICT efficacy.

Turkey 2018 PISA sample of the mother's education level, the argument does not have a significant impact on the success of science scores, maternal education level seems to be more motivated student achievement under the value of 0.5. Father education level independent variable also does not have a significant effect on science score achievement. It should be considered that with the increase in the education level of the parents, the quality time spent with the parents has an effect on success.

A positive correlation was found between the time spent learning science per week and achievement in science scores. It can be said that this positive relationship is not meaningful enough to make a difference. Although it is understood from the fact that the success rate is 100% for employees over 485 minutes, the efficiency of working time is also important with working time.

It can be said that the increase in ICT used in school has a positive effect on science score achievement. It should not be forgotten that the exam in PISA 2018 is computer-aided. Benefiting from ICT resources during the lessons has become a necessity of the age. It is seen that there is a negative relationship between perceived ICT competence and science score achievement. Increasing ICT competence decreases the science score success. This situation shows that the use of ICT is used for purposes other than increasing the academic success. It is important to be able to guide our students in the use of ICT at school and to raise awareness of parents on this issue.

Giriş

PISA (**P**rogramme for **I**nternational **S**tudent **A**ssessment); okuma becerisi, matematik, fen, finansal okuryazarlık gibi akademik becerilerle beraber öğrencinin kendini yeterlik algısı ve tutumunu, aile-okul ve çevresi ile ilgili verileri bir araya getiren geniş çaplı bir eğitim araştırmasıdır. (Akyüz & Pala, 2010; Kamaliyah, Zulkardi, & Darmawijoyo, 2013). İlk defa 2000 yılında yapılan PISA sınavlarına 2003 yılından sonra Türkiye de düzenli olarak katılmıştır (Demir, 2015). Son PISA sınavı 2018 yılında yapılmış, sınava 37 OECD ülkesi ile birlikte 79 ülke dâhil olmuş, Türkiye 454 puanla matematikte 42., 468 puanla fende 39. ve 466 puanla okumada 40. sırada yer almıştır (OECD, 2019a). Öğrencilere bilgisayar destekli testler uygulanarak yapılan sınavda, her bir değerlendirme iki saat sürmüş, çoktan seçmeli ve açık uçlu sorular sorulmuştur. Ayrıca sınava giren öğrenciler; kendileri, evleri, okulları ve aileleri hakkında bilgiler istenen 35 dakikalık bir anket doldurmuşlardır (OECD, 2019c). Bu çalışmada alınan sınav puanları ve anket sonuçları, ülkelerin eğitim politikalarına yön vermede etkili olmaktadır. Bu sebeple sınav ve anket sonuçlarının analizi öğrenci başarı ve başarısızlık nedenlerinin tespit edilmesi, öğrenci başarılarının artırılması, eğitim-öğretim ortamlarındaki aksaklıkların tespit edilmesi, daha etkili eğitim-öğretim ortamlarının oluşturulmasında büyük önem taşımaktadır.

Büyük miktardaki ham verinin analiz programları ile anlamlandırılması ve geleceği tahmin amacıyla kullanılmasına veri madenciliği (VM) denilmektedir (Thuarisingham, 2003). Eğitim alanında da kitlesel iletişim araçlarının kullanımı ile büyük veri yığınları oluşmuştur. Bu ham veriler içinde tüm eğitim paydaşlarının yararına kullanılacak verilerin bulunması ve işlenmesi yeni nesil eğitim ortamlarının düzenlenmesinde rol oynayacaktır (Özbay, 2015). VM yöntemlerini kullanarak uluslararası geniş ölçekli sınavlardan biri olan PISA sınavı ile öğrenci ve okula ilişkin oldukça fazla verinin içerisinden anlamlı sonuçlar çıkarmak da ülkelerin eğitim kalitesini arttırmak adına yardımcı olacaktır.

Alan yazın incelendiğinde, Aksu ve Güzeller (2016), Güre, Kayri ve Erdoğan (2020), Sağlam, Pekiş ve Yılmaz (2020) yaptığı çalışmalarda PISA sonuçlarını veri madenciliği yöntemlerini kullanarak incelemişlerdir. PISA ve TIMSS gibi geniş çaplı sınavlarda akademik başarı ile beraber algı ve tutumların da analiz sonuçlarına dâhil edildiği görülmektedir (Güzel & Berberoğlu, 2010). Akademik başarı ile yeterlik inancı arasında pozitif bir ilişki olduğu Pajares ve Miller (1994) tarafından yapılan çalışmada belirlenmiştir. Bu çalışmanın bağımsız değişkenleri belirlenirken bu noktada alan yazından faydalanılmıştır. Bilgi iletişim teknolojileri (BİT) yeterlik algısının fen sınav puanı üzerindeki etkisi incelenmiştir. PISA gibi geniş çaplı sınavlarda fen başarısına etki eden değişkenleri faktör analizi kullanarak belirleyen Anıl (2009) çalışmasında baba eğitim durumu, tutum ve bilgisayar kullanımının en yordayıcı değişkenler olduğunu tespit etmiştir. PISA fen bilimleri sonuçları kullanılarak VM ile fen okuryazarlığının yordanmasında fen öğrenme süresi değişkeninin özellikle karar ağaçları algoritmasında en başarılı değişkenlerden olduğu belirlenmiştir (Aksu, 2018). Anne-baba eğitim durumu, fen öğrenmek için haftalık harcanan zaman, okulda BİT kullanımı ve öğrenci algılanan BİT yeterliği daha önce yapılan çalışmalarla paralellik gösterip göstermediğinin belirlenmesi için bu çalışmadaki değişkenlere dâhil edilmiştir.

Bu çalışma ile

- Anne eğitim durumunun öğrenci başarısına anlamlı etkisi var mıdır?
- Baba eğitim durumunun öğrenci başarısına anlamlı etkisi var mıdır?
- Fen öğrenmek için haftalık harcanan zamanın öğrenci başarısına anlamlı etkisi var mıdır?

- Okulda BİT kullanım miktarının öğrenci başarısına anlamlı etkisi var mıdır?
- Öğrenci algılanan BİT yeterlik miktarının öğrenci başarısına etkisi var mıdır?

sorularına cevap aranmıştır.

Ayrıca fen sınav sonuçları üzerinde etkisi olduğu düşünülen bu bağımsız değişkenlerden yararlanılarak fen sınav sonucunun doğruluk tahmini yapılması amaçlanmıştır.

Bu çalışmada kullanılan VM yöntemlerinin ihtiyaçlara yönelik alanların belirlenmesine katkı sağlayacağı, okul dışı öğrenme ortamlarının iyileştirilmesine katkıda bulunacağı, eğitim alanında yeni nesil öğrenme ortamlarının tasarımına rehberlik edeceği, başarısızlık nedenlerinin belirlenip fırsata dönüştürülebileceği düşünülmektedir.

Yöntem

PISA 2018 fen sınav sonuçları ve öğrencilerin yanıtladığı anket verileri oecd.org/pisa sayfasından indirilmiş ve incelenerek değişkenler belirlenmiştir. Sonrasında analiz edilecek veriler ayıklanmış ve veri madenciliği yöntemleri ile analiz edilmiştir. Analiz sonucunda elde edilen bulgular ve sonuçlara dayalı olarak önerilerde bulunulmuştur.

Araştırmanın Deseni

Bu araştırma, fen sınav puanlarının seçilen bağımsız değişkenler ile ilişkisini incelemeye dayandığından tarama modellerinden ilişkisel tarama modeli ile yapılmış bir çalışmadır. İlişkisel tarama modeli, birden çok değişken arasındaki etkileşimi neden-sonuç ilişkisi kurarak inceleyen bir araştırma yöntemidir (Karasar, 2006).

Evren ve Örneklem / Çalışma Grubu / Katılımcılar

PISA 2018 araştırmasına Türkiye’den 1.038.993 evren içinden 6.890 öğrenci dâhil olmuştur (OECD, 2019b). Örneklemi oluşturan öğrenciler tabakalı seçki yöntemi ile belirlenmiş 15 yaş grubu öğrencilerin kaydının bulunduğu her okuldan eşit olasılıkla seçilmiş 42 öğrenciden oluşmaktadır (OECD, 2019b). Bu çalışmada; Türkiye örneklemi oluşturan 6.890 öğrenci verisi incelenmiştir.

Veri Toplama Araçları

PISA 2018 çalışmasında fen sınav puanları üzerinde etkisi olduğu düşünülen bağımsız değişkenlerin belirlenmesi amacıyla öncelikle sınavda kullanılan öğrenci anketleri incelenmiştir. PISA öğrenci anketinde yer alan soruların incelenmesinin ardından fen sınav puanlarına etkisi olduğu düşünülen bağımsız değişkenler; anne eğitim durumu, baba eğitim durumu, fen öğrenmek için haftalık harcanan zaman, okulda BİT kullanımı ve öğrenci algılanan BİT yeterliği olarak belirlenmiştir. Anketten elde edilen bağımlı ve bağımsız değişkenler ile ilgili bilgiler Tablo 1’de gösterilmiştir.

Tablo 1. Değişken Bilgileri

DEĞİŞKEN TÜRÜ	DEĞİŞKEN ADI	DEĞİŞKEN ARALIKLARI
Bağımlı Değişken	Fen sınav puanları	PV1SCIE,PV2SCIE,PV3SCIE,PV4SCIE,PV5SCIE, PV6SCIE,PV7SCIE,PV8SCIE,PV9SCIE,PV10SCIE
Bağımsız Değişkenler	(MISCED) Anne eğitim durumu	SEVİYE 1-6

(FISCED)	Baba eğitim durumu	SEVİYE 1-6
(SMINS)	Haftada fen öğrenmek için harcanan zaman	HAFTADA 0<ZAMAN<1710 DAKİKA
(USESCH)	Okulda BİT kullanımı	-1.716<PUAN<3.304
(COMPICT)	Algılanan BİT yeterliği	-2.603<PUAN<2.065

Tablo1’de görüldüğü gibi bağımlı değişken fen sınav puanları, 10 sorunun cevabına verilen puanlardan oluşmaktadır. Bağımsız değişkenler anne ve baba eğitimi 6 seviyeden oluşmakta ve seviye 1 ilkokul, seviye 2 ortaokul, seviye 3 lise, seviye 4 ön lisans, seviye 5 lisans ve seviye 6 lisansüstü eğitime karşılık gelmektedir. Haftada fen öğrenmek için harcanan zaman, okulda BİT kullanımı ve algılanan BİT yeterliği aralıkları verilen ölçek puanları yansıtmaktadır.

Veri Toplama Süreci

Bu çalışmada kullanılan veriler, OECD’nin resmi internet sayfası olan <http://www.oecd.org/pisa/data/2018database/> adresinden elde edilmiştir. SPSS veri dosyası formatında yer alan öğrenci anketinden ülke kodu TR (792) olan 6890 öğrenciye ilişkin veriler analiz kapsamında veri kaynağı olarak kullanılmıştır.

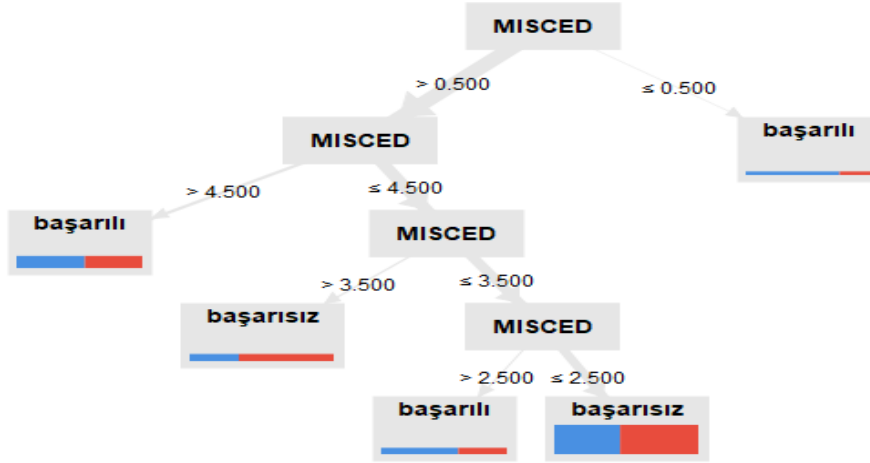
Veri Analizi

Araştırma, PISA 2018 Türkiye örnekleme üzerinden yürütülmüştür. 6890 veriden oluşan örnekleme analiz öncesi veri ön işleme yapılmış ve kayıp veriler fen puanlarının normal dağılımı bozulmayacak şekilde temizlenerek analiz 6001 veri üzerinde tamamlanmıştır. Rapid Miner analiz programına alınan veride her öğrenci için fen puanlarının ortalaması hesaplanmış 468 puan ve üzeri başarılı, 468 puan altı başarısız olarak sınıflanmıştır. Anne eğitim durumu, baba eğitim durumu, haftada fen öğrenmek için harcanan zaman, okulda kullanılan BİT, algılanan BİT yeterliği bağımsız değişkenleri kullanılarak veriye uygulanan K-nn, Naive Bayes, Random Forest algoritmalarının bağımlı değişken fen puanlarını başarılı veya başarısız şeklinde sınıflama performansı tespit edilmiştir. Bir değişkenin başka bir değişkenle arasındaki ilişkiyi göstermenin anlaşılır yollarından biri olan karar ağaçları da kullanılmıştır (Karaibrahimoğlu, 2014).

Bulgular

1. Anne Eğitim Durumu- Fen Puanı Başarı Durumu İlişkin Karar Ağacı

Anne eğitim durumu ve fen puanı başarı durumuna ilişkin karar ağacı yapısı Şekil 1’de gösterilmiştir.



Şekil 1. Anne Eğitim Durumu- Fen puanı Başarısına İlişkin Karar Ağacı

MISCED > 0.500

| MISCED > 4.500: başarılı {başarılı=684, başarısız=573}

| MISCED ≤ 4.500

| | MISCED > 3.500: başarısız {başarılı=230, başarısız=443}

| | MISCED ≤ 3.500

| | | MISCED > 2.500: başarılı {başarılı=353, başarısız=222}

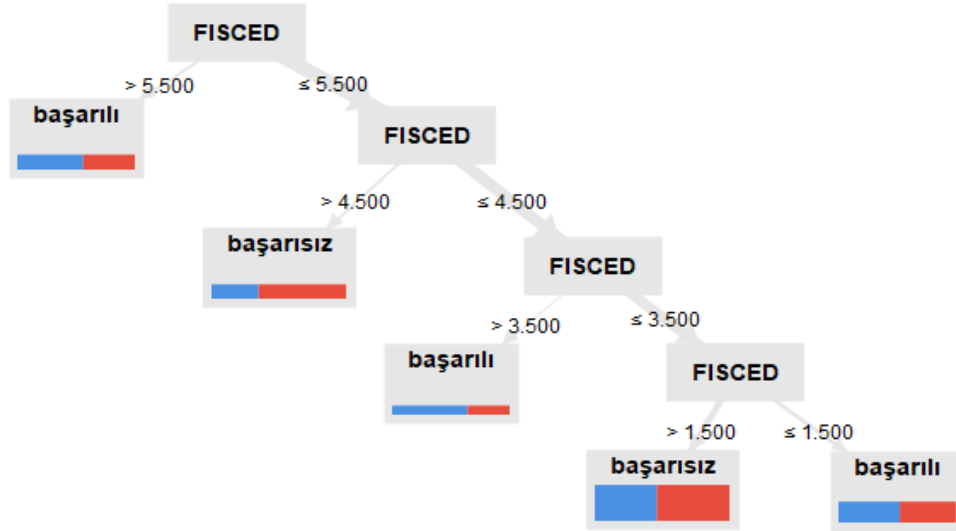
| | | MISCED ≤ 2.500: başarısız {başarılı=1482, başarısız=1758}

MISCED ≤ 0.500: başarılı {başarılı=191, başarısız=65}

Anne eğitim durumu 0.5'in altında iken başarı oranı %74, 0.5-2.5 aralığında %45, 2.5-3.5 aralığında %61, 3.5-4.5 aralığında %34, 4.5 üzeri %54'tür. Anne eğitim durumunun fen başarısını pozitif yönde etkilemesi için anne eğitim durumu arttıkça başarı oranının da artması beklenir. Fakat genele bakıldığında anne eğitim durumu ile fen başarısı arasında anlamlı bir ilişki bulunamamıştır.

2. Baba Eğitim Durumu- Fen Puanı Başarı Durumu İlişkin Karar Ağacı

Baba eğitim durumu ve fen puanı başarı durumuna ilişkin karar ağacı yapısı Şekil 2'de gösterilmiştir.



Şekil 2. Baba Eğitim Durumu-Fen Puanı Başarısına İlişkin Karar Ağacı

FISCED > 5.500: başarılı {başarılı=521, başarısız=407}

FISCED ≤ 5.500

| FISCED > 4.500: başarısız {başarılı=307, başarısız=568}

| FISCED ≤ 4.500

| | FISCED > 3.500: başarılı {başarılı=331, başarısız=184}

| | FISCED ≤ 3.500

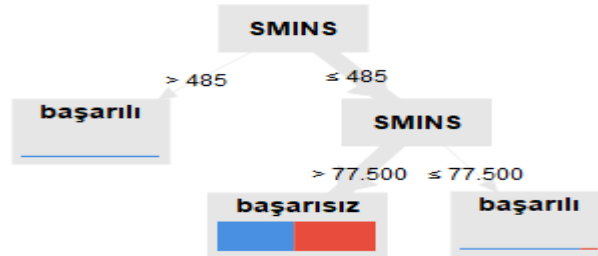
| | | FISCED > 1.500: başarısız {başarılı=1089, başarısız=1270}

| | | FISCED ≤ 1.500: başarılı {başarılı=692, başarısız=632}

Baba eğitim durumu 1.5'in altında başarı oranı %52, 1.5-3.5 aralığında %46, 3.5-4.5 aralığında %64, 4.5-5.5 aralığında %35, 5.5 üzeri %56'dır. Baba eğitim durumunun artışı genel olarak fen başarısı üzerinde anlamlı bir etki oluşturmamaktadır.

3. Fen Öğrenmek İçin Haftalık Harcanan Zaman- Fen Puanı Başarı Durumu İlişkin Karar Ağacı

Fen öğrenmek için haftalık harcanan zaman ve fen puanı başarı durumuna ilişkin karar ağacı yapısı Şekil 3'te gösterilmiştir.



Şekil 3. Fen Öğrenmek İçin haftalık Harcanan Zaman – Fen Puanı Başarısına İlişkin Karar Ağacı

SMINS > 485: başarılı {başarılı=30, başarısız=0}

SMINS ≤ 485

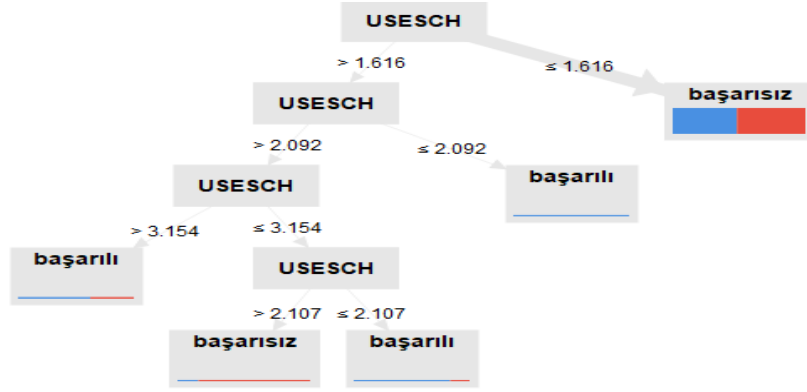
| SMINS > 77.500: başarısız {başarılı=2869, başarısız=3055}

| $SMINS \leq 77.500$: başarılı {başarılı=41, başarısız=6}

Fen öğrenmek için haftalık harcanan zaman 77.50 dakikanın altında başarı oranı %87, 77.50-485 dakika aralığında %48, 485 dakika üzeri %100'dür. Çalışma süresinin uzunluğu ile fen başarı durumu arasında pozitif bir ilişki bulunsa da fark oluşturacak kadar anlamlı değildir.

4. Okulda Kullanılan BİT Miktarı – Fen Puanı Başarı Durumuna İlişkin Karar Ağacı

Okulda kullanılan BİT ve fen puanı başarı durumuna ilişkin karar ağacı yapısı Şekil 4'te gösterilmiştir.



Şekil 4. Okulda Kullanılan BİT Miktarı – Fen Puanı Başarısına İlişkin Karar Ağacı

$USESCH > 1.616$

| $USESCH > 2.092$

| | $USESCH > 3.154$: başarılı {başarılı=73, başarısız=44}

| | $USESCH \leq 3.154$

| | | $USESCH > 2.107$: başarısız {başarılı=3, başarısız=16}

| | | $USESCH \leq 2.107$: başarılı {başarılı=5, başarısız=1}

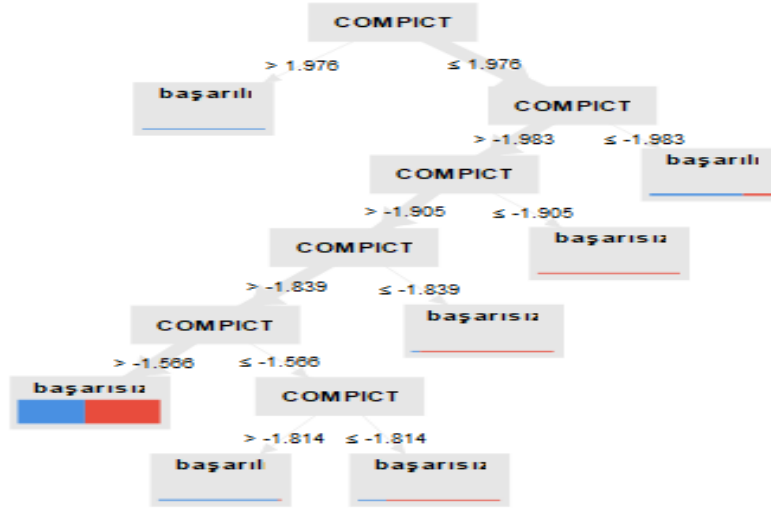
| $USESCH \leq 2.092$: başarılı {başarılı=40, başarısız=0}

$USESCH \leq 1.616$: başarısız {başarılı=2819, başarısız=3000}

Okulda kullanılan BİT miktarı 1.616'nın altında iken başarılı öğrenci oranı %48, 1.616-2.092 aralığında %100, 2.092-2.107 aralığında %83, 2.107-3.154 aralığında %15, 3.154 üzeri %61'dir. Karar ağacının genelinde anlamlı bir başarı artışı görünmese de okulda kullanılan BİT miktarı artışının başarı durumunu pozitif etkilediği söylenebilir.

5. Öğrenci Algılanan BİT Yeterliği – Fen Puanı Başarı Durumuna İlişkin Karar Ağacı

Öğrenci algılanan BİT yeterliği ve fen puanı başarısına ilişkin karar ağacı yapısı Şekil 5'te gösterilmiştir.



Şekil 5. Öğrenci Algılanan BİT Yeterliği – Fen Puanı Başarısına ilişkin Karar ağacı

COMPACT > 1.976: başarılı {başarılı=6, başarısız=0}

COMPACT ≤ 1.976

| COMPACT > -1.983

| | COMPACT > -1.905

| | | COMPACT > -1.839

| | | | COMPACT > -1.566: başarısız {başarılı=2644, başarısız=2957}

| | | | COMPACT ≤ -1.566

| | | | | COMPACT > -1.814: başarılı {başarılı=80, başarısız=2}

| | | | | COMPACT ≤ -1.814: başarısız {başarılı=1, başarısız=4}

| | | COMPACT ≤ -1.839: başarısız {başarılı=1, başarısız=15}

| | COMPACT ≤ -1.905: başarısız {başarılı=0, başarısız=17}

| COMPACT ≤ -1.983: başarılı {başarılı=208, başarısız=66}

Öğrenci algılanan BİT yeterliği miktarı -1.983'ün altında iken başarılı öğrenci oranı %74, (-1.983)-(-1.905) aralığında %0,(-1.905)-(-1.839) aralığında %6, (-1.839)-(-1.814) aralığında %20, (-1.814)-(-1.566) aralığında %97, (-1.516)-(-1.976) aralığında %47, 1.976 üzeri %100'dür. Öğrenci algılanan BİT yeterliği arttıkça fen başarısı üzerinde anlamlı bir artış oluşmamaktadır. Karar ağacının geneline bakıldığında yeterliğin fen başarısını negatif etkilediği söylenebilir.

Knn Algoritması Sınıflama Sonuçları

k=9 ve train ile test verisi belirli orana göre ayrıştırıldığında, anne eğitim durumu, baba eğitim durumu, fen öğrenmek için haftalık harcanan zaman, okulda kullanılan bit miktarı, öğrenci algılanan bit yeterliği verileri kullanılarak fen puanı başarısını %77 doğrulukla tahmin etmiştir.

accuracy: 77.00%

	true başarılı	true başarısız	class precision
pred. başarılı	704	236	74.89%
pred. başarısız	178	682	79.30%
class recall	79.82%	74.29%	

Şekil 6. Rapid Miner Programı K-nn Algoritması Sınıflama Sonucu

Şekil 6’da Rapid Miner programında K-nn algoritmasının sınıflama sonucu matris şeklinde gösterilmiştir. Bu matrise göre başarılı olup algoritma tarafından başarılı olarak sınıflandırılanlar 704 kişi, başarısız olup başarılı olarak sınıflandırılanlar ise 236 kişidir ve başarılıları doğru tahmin oranı %74.89’dur. Veride başarılı olup başarısız olarak sınıflandırılanlar 178 kişi, başarısız olup başarısız olarak sınıflandırılanlar 682 kişidir ve başarısızları doğru tahmin oranı %79.30’dur. Başarılı ve başarısızları doğru tahmin ortalaması alınınca doğruluk tahmin oranı %77 bulunmuştur.

Naive Bayes Algoritması Sınıflama Sonuçları

Naive Bayes algoritması; anne eğitim durumu, baba eğitim durumu, haftada fen öğrenmek için harcanan zaman, okulda kullanılan BİT miktarı, öğrenci algılanan BİT yeterliği verilerini kullanarak fen puanı başarısını %55.06 doğrulukla tahmin etmiştir.

accuracy: 55.06%

	true başarılı	true başarısız	class precision
pred. başarılı	359	286	55.66%
pred. başarısız	523	632	54.72%
class recall	40.70%	68.85%	

Şekil 7. Naive Bayes Algoritması Doğruluk Sonucu

Şekil 7’de Rapid Miner programında Naive Bayes algoritmasının sınıflama sonucu matris şeklinde gösterilmiştir. Bu matrise göre başarılı olup algoritma tarafından başarılı olarak sınıflandırılanlar 359 kişi, başarısız olup başarılı olarak sınıflandırılanlar ise 286 kişidir ve başarılıları doğru tahmin oranı %55.66’dır. Veride başarılı olup başarısız olarak sınıflandırılanlar 523 kişi, başarısız olup başarısız olarak sınıflandırılanlar 632 kişidir ve başarısızları doğru tahmin oranı %54.72’dir. Başarılı ve başarısızları doğru tahmin ortalaması alınınca doğruluk tahmin oranı %55.06 bulunmuştur.

Random Forest Algoritması Sınıflama Sonuçları

Random Forest algoritması; anne eğitim durumu, baba eğitim durumu, haftada fen öğrenmek için harcanan zaman, okulda kullanılan BİT miktarı, öğrenci algılanan BİT yeterliği verilerini kullanarak fen puanı başarısını %62.22 doğrulukla tahmin etmiştir.

accuracy: 62.22%

	true başarılı	true başarısız	class precision
pred. başarılı	296	94	75.90%
pred. başarısız	586	824	58.44%
class recall	33.56%	89.76%	

Şekil 8. Random Forest Algoritması Doğruluk Sonucu

Şekil 8’de Rapid Miner programında Random Forest algoritmasının sınıflama sonucu matris şeklinde gösterilmiştir. Bu matrise göre başarılı olup algoritma tarafından başarılı olarak sınıflandırılanlar 296 kişi, başarısız olup başarılı olarak sınıflandırılanlar ise 94 kişidir ve başarılıları doğru tahmin oranı %75.90’dır. Veride başarılı olup başarısız olarak sınıflandırılanlar 586 kişi, başarısız olup başarısız olarak sınıflandırılanlar 824 kişidir ve başarısızları doğru tahmin oranı %58.44’tür. Başarılı ve başarısızları doğru tahmin ortalaması alınca doğruluk tahmin oranı %62.22 bulunmuştur.

Tartışma ve Sonuç

Bu çalışma, PISA 2018 fen puanı sonuçlarını belirlenmiş değişkenler çerçevesinde eğitsel veri madenciliği yöntemleri ile değerlendirmeyi amaçladığı için çalışmada seçilmiş bağımsız değişkenlerin fen puanı başarı durumunu tahmin yeterliliğini tespit amacıyla k-nn, naive bayes, random forest algoritmaları kullanılmıştır. K-nn algoritması anne-baba eğitim durumu, fen öğrenmek için haftalık harcanan zaman, okulda BİT kullanımı ve öğrenci algılanan BİT yeterliği değişkenlerini kullanarak fen puanı başarı durumunu %77, naive bayes %55, random forest %62 oranında doğru tahmin etmiştir.

Gürsakar (2012) yaptığı çalışmada varyans analizi sonucunda anne ve baba eğitim durumunun artışı ile fen başarısının paralellik gösterdiğini söylemektedir. PISA 2018 Türkiye örnekleme üzerinde anne eğitim durumu bağımsız değişkeninin fen puanı başarısında anlamlı bir etkiye sahip olmadığı, anne eğitim durumu 0.5 değerinin altında bulunan öğrencilerin başarı motivasyonunun daha fazla olduğu görülmektedir. Baba eğitim durumu bağımsız değişkeni de fen puanı başarısında anlamlı bir etkiye sahip değildir. Bu değişkenler açısından literatürden farklı sonuçların ortaya çıkması eğitim politikaları sonucu zaman içinde kişilerin kendini geliştirmeye yönelik çalışmalara daha fazla önem vermesi ile bir fark oluşturma durumundan çıkması olabilir. Anne-baba eğitim durumunun artışı ile beraber ebeveynlerle geçirilen nitelikli zamanın başarıda etkisinin olduğu göz önünde bulundurulmalıdır.

Kaya ve Kaya (2018) yaptığı çalışmada belli bir zaman diliminde yapılan fen ödevlerinin ders başarısını arttırdığını, azının ve fazlasının ise başarıyı düşürdüğünü tespit etmiştir. Fen öğrenmek için haftalık harcanan zaman, ödev için ayrılan süre gibi düşünüldüğünde fen puanı başarısını belli bir aralıkta arttırdığı, fazlasının ise düşüşe sebep olduğu tespit edilmiştir. Başarının şans olmadığı 485 dakika üzeri çalışanlarda başarı oranının %100 olmasından anlaşılabilir da çalışma süresiyle birlikte çalışma süresinin verimliliği de önem arz etmektedir.

Derslerde yeterli miktarda teknoloji kullanılarak motivasyonu arttırmak ve öğrenmeyi kalıcı hale getirmek mümkün olabilir (Durak & Yılmaz, 2019). Bu çalışmada da okulda kullanılan BİT artışının fen puanı başarısında pozitif bir etkiye sahip olduğu söylenebilir. PISA 2018’de sınavın bilgisayar destekli gerçekleştirildiği unutulmamalıdır. Ders sırasında BİT kaynaklarından yararlanmak artık çağın gereği haline dönüşmüştür. Bu

çalışmanın yazıldığı dönem covid salgın dönemine denk gelmekte, bu dönemde öğrencilerin okulda BİT kullanımının ve öğrenci BİT yeterliklerinin daha fazla arttığı düşünülürse PISA 2022’de BİT ile ilgili bu değişkenlerin sonuçları daha net bir şekilde görülecektir. Şenel (2020) çalışmasında evde BİT kullanımının PISA 2018 matematik sınav sonuçları üzerinde anlamlı bir yordayıcı olduğunu açıklamıştır. Bu araştırma literatürle uyularak, öğrenci algılanan BİT yeterliği ile fen puanı başarısı arasında anlamlı fark oluşturacak kadar negatif bir ilişki olduğunu göstermektedir. BİT yeterliğinin artışı fen puanı başarısını düşürmektedir. Bu durum BİT kullanımının ders başarısını artırma amacı dışında kullanıldığını düşündürmektedir. Kaya ve Kaya (2018), internette ders başarısını arttırmaya yönelik geçirilen zamanın fen başarısına olumlu katkısı olduğunu tespit etmiştir. Günümüzde öğrenci BİT kullanım yeterliğinin artışı, bilgisayar oyunu bağımlılığı ile paralellik göstermektedir (Sağlam, Pekiye, & Yılmaz, 2020). BİT kullanımı konusunda öğrencilerimize okulda rehberlik edebilmek, ebeveynlerin bu konuda bilinçlendirilmesi önem arz etmektedir.

Yayın Etiği Bildirimi / Research Ethics

Bu çalışmada etik dışı bir durum yoktur, araştırma yapılırken yayın etiği konusuna dikkat edilmiştir. / There is no unethical situation in this study, attention was paid to the issue of publication ethics while doing the research.

Araştırmacıların Katkı Oranı / Contribution Rate of Researchers

Yazarların katkı oranı eşittir. Makalenin planlanması, verilerinin elde edilmesi, veri analizinin gerçekleştirilmesi, makalenin girişi, bulgular ile tartışma ve sonuç bölümlerinin yazılmasında üç yazarın katkısı eşittir. / The authors' contribution rate is equal. The contributions of three authors are equal in planning the article, obtaining its data, performing data analysis writing the introduction of the article, findings, discussion and conclusion.

Çıkar Çatışması / Conflict of Interest

Bu çalışmada bir çıkar çatışması bulunmamaktadır. / There is no conflict of interest in this study.

Fon Bilgileri / Funding

Bu araştırma süresince veya öncesinde herhangi bir fondan destek alınmamıştır. / No funding was received during or before this research.

Etik Kurul Onayı / The Ethical Committee Approval

Bu çalışmada, tüm araştırmacılara açık, uluslararası veri tabanında yer alan veriler kullanıldığından etik kurul kararı gerektirmemektedir. / Since, an international database, which is open to all researchers is used, an approved ethical committee decision for research is not required to be submitted for this study.

Kaynakça / References

- Aksu, G., & Güzeller, C. O. (2016). PISA 2012 matematik okuryazarlığı puanlarının karar ağacı yöntemiyle sınıflandırılması: Türkiye Örnekleme. *Eğitim ve Bilim Dergisi*, 41(185), 101-122.
- Aksu, G. (2018). *PISA başarısını tahmin etmede kullanılan veri madenciliği yöntemlerinin incelenmesi. [Investigation of data mining methods used for estimating pisa success].* (Yayımlanmamış doktora tezi). Hacettepe Üniversitesi, Ankara.
- Akyüz, G., & Pala, M. N. (2010). PISA 2003 sonuçlarına göre öğrenci ve sınıf özelliklerinin matematik okuryazarlığına ve problem çözme becerilerine etkisi. *İlköğretim Online Dergisi*, 9(2), 668-678. <http://ilkogretim-online.org.tr>
- Anıl, D. (2009). Uluslararası öğrenci başarılarını değerlendirme programı (PISA)'nda Türkiye'deki öğrencilerin fen bilimleri başarılarını etkileyen faktörler. [Factors effecting science achievement of science students in programme for international students' achievement (PISA) in Turkey]. *Eğitim ve Bilim Dergisi*, 34(152), 88-100.
- Demir, E. (2015). Türkiye'de on beş yaş grubu öğrencilerin matematik okuryazarlık becerileri ile ilişkili duyuşsal özellikleri [Affective characteristics predicting 15-year-old students' mathematics literacy skills in Turkey]. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 48(2), 165-184.
- Durak, A., & Yılmaz, F. G. K. (2019). Artırılmış gerçekliğin eğitsel uygulamaları üzerine ortaokul öğrencilerinin görüşleri [Opinions of secondary school students on educational practices of augmented reality]. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 19(2), 468-481.
- Güre B. Ö., Kayri M., & Erdoğan F. (2020). PISA 2015 matematik okuryazarlığını etkileyen faktörlerin eğitsel veri madenciliği ile çözümlenmesi. *Eğitim ve Bilim Dergisi*, 45(202), 393-415.
- Gürsakal, S. (2012). PISA 2009 öğrenci başarı düzeylerini etkileyen faktörlerin değerlendirilmesi. [An evaluation of PISA 2009 student achievement levels' affecting factors]. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 17(1), 441-452.
- Güzel, Ç. İ., & Berberoğlu, G. (2010) Students' affective characteristics and their relation to mathematical literacy measures in the programme for international student assessment (PISA) 2003. *Eurasian Journal of Educational Research (EJER)*, (40), 93-113.
- Kamaliyah, K., Zulkardi, Z., & Darmawijoyo, D. (2013). Developing the sixth level of PISA-like mathematics problems for secondary school students. *Journal On Mathematics Education*, 4(1), 9-28. doi:10.22342/jme.4.1.559.9-28. <http://ejournal2.unsri.ac.id/index.php/jme/issue/view/87>
- Karaibrahimoğlu, A. (2014). *Veri madenciliğinden birliktelik kuralı ile onkoloji verilerinin analiz edilmesi: Meram Tıp Fakültesi Onkoloji örneği [Analyzing breast cancer data using association rule mining: meram faculty of medicine oncology department].* (Yayımlanmamış doktora tezi). Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.
- Karasar, N. (2006). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayın Dağıtım.

- Kaya, V. H., & Kaya, E. (2018). Fen başarısını artırmak için ödevler ve kurslar gerekli midir? *Anadolu Öğretmen Dergisi*, 2(1), 48-62.
- OECD. (2019a). PISA 2018 results (volume I), what students know and can do. PISA, *OECD Publishing*, Paris. <https://doi.org/10.1787/5f07c754-en>.
- OECD. (2019b). PISA 2018 results (volume III), what school life means for students' lives. PISA, *OECD Publishing*, Paris. <https://doi.org/10.1787/acd78851-en>.
- OECD. (2019c). What is PISA? in PISA 2018 results (volume II): Where all students can succeed: *OECD Publishing*, Paris. <https://doi.org/10.1787/65d76825-en>.
- Özbay, Ö. (2015). Veri madenciliği kavramı ve eğitimde veri madenciliği uygulamaları. *Uluslararası Eğitim Bilimleri Dergisi*, (5), 262-272.
- Pajares, F., & Miller, M. D. (1994). The role of self-efficacy and self-concept beliefs in mathematical problem-solving: A path analysis. *Journal of Educational Psychology*, 86, 193-203.
- Sağlam, Z., Pekiş, F. M. & Yılmaz, R. (2020). PISA 2018 araştırmasına etki eden duygusal faktörlerin veri madenciliği yöntemleri ile incelenmesi. [Investigation of emotional factors affecting PISA 2018 research with data mining methods]. *Bilgi ve İletişim Teknolojileri Dergisi/Journal of Information and Communication Technologies*, 2(2), 113-148.
- Şenel, C. H. (2020). Öğrencilerin bilgi ve iletişim teknolojilerine erişebilirliğinin matematik başarısına etkisi PISA 2018 Türkiye Örneği. *13.Uluslararası Eğitim Camiası Sempozyumu*, Rize.
- Thuraisingham, B. (2003). *Web data mining and applications in business intelligence and counter terrorism*. USA: CRC Press LLC, Boca Raton, FL.