

ÜNİVERSİTE ÖĞRENCİLERİNİN DERS BAŞARISINI ETKİLEYEN FAKTÖRLERİN LOJİSTİK REGRESYON ANALİZİ İLE TAHMİNLENMESİ*

Şengül CAN**

Tuncer ÖZDİL***

Cengiz YILMAZ****

Özet

Büyük veri yığınları içerisinde gizli ve faydalı bilgiye ulaşmayı sağlayan veri madenciliği uygulamaları işletme, finans, tıp, eğitim gibi birçok alanda faydalı, önemli bilgilerin açığa çıkarılmasında kullanılmaktadır. Büyük veri yığınlarından elde edilecek gizli kalmış faydalı bilgiler eğitim alanında da eğitim öğretim kalitesinin artırılması ve çeşitli eğitim-öğretim sorunlarının çözümünde önemli katkılar sağlayabilmektedir. Öğretmenlerin ve derslerin öğrenciler tarafından değerlendirilmesi pek çok yükseköğretim kurumunda gerçekleştirilen bir uygulamadır. Bu çalışma kapsamında veri madenciliği tekniklerinden lojistik regresyon analiziyle dönem sonu ders değerlendirme anket verileri kullanılarak öğrencilerin ders başarılarını etkileyen faktörler araştırılmıştır. 5820 adet üniversite öğrencisine uygulanan anket bilgisi üzerinden iki ayrı ders başarı değişkeni tanımlanmıştır. Ders başarı değişkenlerini etkileyen likert ölçekli sorulara verilen cevaplarla ders başarıları arasındaki ilişki lojistik regresyon analizi ile tahmin edilmiştir. Elde edilen analiz sonuçlarına göre, öğrenci ders başarısını ders tekrar sayısının yanı sıra öğrencilerin mesleki gelişimlerini arttıran, öğrencilere yeni bakış açıları kazandıran uygulamaların daha çok etkilediği görülmüştür. Ayrıca dersi ilk kez alan öğrencilerin görüşlerinin dersi iki ve daha fazla sayıda tekrarlayan öğrencilere göre genel olarak daha olumlu olduğu fark edilmiştir. Bunun yanı sıra anketin uygulanma şeklinin değiştirilerek final sınavından önceki son derste uygulanmasının öğrencinin daha objektif değerlendirme yapmasını sağlayacağı kanaatine ulaşılmıştır.

Anahtar Kelimeler: Eğitimde veri madenciliği, Lojistik Regresyon Analizi, Öğrenci Başarısını Etkileyen Faktörler

Jel Kodları: I23, C25.

* Bu çalışma 2. Uluslararası Sosyal Bilimler Araştırmaları Kongresi'nde (19-21 Ekim 2017, Bakü, Azerbaycan) sözlü olarak sunulmuştur. Gelen görüş ve öneriler doğrultusunda geliştirilmiş ve yeni bölüm eklenmiştir.

**Manisa Celal Bayar Üniversitesi, BAUM, sengulcan54@gmail.com, orcid.org/ 0000-0003-4022-0393.

***Manisa Celal Bayar Üniversitesi, İİBF, tozdil12@hotmail.com, orcid.org/ 0000-0002-3199-6766.

****Manisa Celal Bayar Üniversitesi, İİBF, yilmazce@hotmail.com, orcid.org/ 0000-0002-7745-1710.

Date of submission: 08-11-2017

Date of acceptance: 23-03-2018

ESTIMATION OF THE FACTORS AFFECTING THE SUCCESS OF THE UNIVERSITY STUDENTS BY LOGISTIC REGRESSION ANALYSIS

Abstract

Data mining applications, which provide confidential and useful information from large data stacks, are used to expose important information useful in many fields such as business, finance, medicine, education. In the field of education, hidden useful information that can be obtained from large data piles is also; Increase the quality of education and training, and provide important contributions to the solution of various education and training problems. In this study, data mining techniques using logistic regression analysis and end-of-term course evaluation questionnaires were used to investigate the factors affecting student's course success. Two separate course success variables were defined through questionnaires applied to 5820 university students. The relationship between the answers to the likert scale questions affecting course success variables and the course success was estimated by logistic regression analysis. According to the results of the analysis, it is seen that the student lecture success is more influenced by the number of lesson repetition, as well as the applications which increase the professional development of the students and give new perspective to the students. It was also found that the views of the students who took the course for the first time were generally more positive than those of two or more repeat students. In addition to this, we have reached the opinion that by changing the way the questionnaire is applied, the student will be able to make a more objective assessment of before the final exam.

Keywords: Data Mining in Education, Logistic Regression Analysis, Factors Affecting Student Achievement,

JEL Codes: I23, C25.

I. GİRİŞ

Bilişim Teknolojilerindeki gelişmeler veri depolama kapasitesini arttırarak elektronik ortamda büyük veri yığınlarının (big data) oluşmasına yol açmıştır. Ancak içerisindeki faydalı bilgi çıkarılmadıkça büyük miktardaki veri anlamlı değildir. Büyük miktardaki veri yığından gizli ve anlamlı bilginin çıkartılarak geleceğe yönelik tahminler yapılması veri madenciliği teknikleri ile mümkündür. Veri madenciliği uygulamaları işletme, finans, tıp, eğitim gibi birçok alanda faydalı, önemli bilgilerin açığa çıkarılmasında kullanılmaktadır. Faydalı bilginin kullanımı yalnızca teknoloji alanında değil diğer sosyal alanlarda da önemli ciddi sonuçların elde edilmesini sağlamaktadır.

Büyük veri yığınlarından elde edilecek gizli kalmış faydalı bilgiler eğitim alanında da; eğitim öğretim kalitesinin arttırılması ve çeşitli eğitim öğretim sorunlarının çözümünde önemli katkılar sağlayabilmektedir. Bilginin yalnızca teknoloji değil diğer alanlarda da ön plana çıkması, her alanda kaliteli eğitimin önemini ortaya çıkarmıştır. Pek çok derste eğitim ve öğretim aşamalarının etkin olmaması, dersin öğrenci tarafından tekrar alınmasına neden

olmaktadır. Bu sorunlar; derslerin içeriğine uygun olan eğitim ve öğretim yöntemlerinin uygulanması, ders sürelerinin iyi değerlendirilmesi ve öğrencilerin derslere olan ilgilerinin artırılmasıyla çözülebilir (Özdil, Urdaletova & Yılmaz, 2010).

Üniversite öğrencilerinin yıl sonunda başarısız sonuçlarla karşılaşmalarının önlenmesi için başarıyı etkileyen faktörlerin belirlenmesi gerekmektedir (Kurt & Erdem, 2012). Başarıyı etkileyen faktörlerin önceden bilinmesiyle başarısızlık nedenlerinin de kontrol edilebileceği düşünülmektedir (Koç, Avşaroğlu & Sezer, 2004).

Bu çalışmanın temel amacı, yıl sonunda üniversite öğrencilerine uygulanan ders değerlendirme anket verileri kullanılarak; öğrencilerin ders başarılarını etkileyen faktörlerin belirlenmesidir. Buradan elde edilen bilgiler ışığında derslerin başarı oranlarının artırılmasına yönelik önerilerin de ortaya çıkarılması amaçlanmaktadır.

II. VERİ MADENCİLİĞİ

Dünyadaki globalleşme nedeniyle, veriye ulaşmanın yanı sıra veriden elde edilebilecek maksimum bilgiyi çıkarmada veri(data) oldukça önemli hale gelmiştir (Savaş, Topaloğlu & Yılmaz, 2012). Veri madenciliğinde hedeflenen temel amaç; büyük miktardaki veriden analiz edilerek çıkarım yapılamayacak anlamlı bilgiye ulaşmaktır (Coşkun & Baykal, 2011).

Veri madenciliğinin tanımında sözü edilen büyük veri ise; tek bir iş istasyonunda arşivlenemeyecek kadar büyük veri kümelerini temsil etmektedir. Veri madenciliği temel olarak; büyük veri setleri içerisindeki ilişki veya desenler üzerinden faydalı bilgiye erişim için yazılım tekniklerinin kullanılması olarak tanımlanmaktadır. Veri kümesi içerisindeki kural, ilişki veya niteliklerin belirlenmesi için bilgisayar yazılımları kullanılmaktadır. Başka bir ifadeyle veri madenciliği; veri seti içerisindeki ilişki, desen ve kuralların istatistiki açıdan önemli taraflarının bilgisayar yazılımları kullanılarak keşfedilmesidir (Baykasoğlu, 2005).

Veri madenciliği için tercih edilen teknikler mevcut veriler üzerinden çıkarım yapmayı sağlamanın yanında geleceğe yönelik tahminlerin yapılmasına da olanak tanımaktadır (Küçükşille, 2009). Bu yönleriyle veri madenciliği, bu teknikleri tercih eden kişi ve kurumların da etkili kararlar almalarına yardımcı olmaktadır. Ayrıca bu teknikler çözümü uzun zaman alan sorulara da çözüm sunmaktadır (Savaş, Topaloğlu & Yılmaz, 2012).

Veri madenciliği uygulamalarının etkin olarak gerçekleştirilebilmesi için; farklı tipte olan verilerle birlikte değerlendirilmesi, tercih edilen algoritmanın etkinlik ve ölçeklenebilirliği, elde edilen sonuçlarda yararlılık, kesinlik ve anlamlılık gibi niteliklerin sağlanması, keşfedilen kuralların farklı biçimlerle temsil edilebilmesi, dağılık kaynaklarda bulunan

31 Üniversite Öğrencilerinin Ders Başarısını Etkileyen Faktörlerin Tahminlenmesi

verilerle işlem yapabilmesi ayrıca gizlilik ve veri güvenliği özelliklerinin de sağlanması gerekmektedir (Dener, Dörterler & Orman, 2009).

Veri madenciliği bir istatistiksel yöntemler serisi gibi görülse de birkaç yönüyle geleneksel istatistik tekniklerinden ayrılmaktadır. Veri madenciliğindeki temel amaç; kolayca görsel sunum veya mantıksal kuralların oluşturulabileceği nitel modellerin çıkartılmasıdır. Bu yönüyle istatistik, yapay zeka, veri tabanları, makine öğrenmesi ve veri görselleştirme gibi alanlarla yakın olarak ilişkilidir.

Veri madenciliği tekniklerinin bazı uygulama alanları aşağıdaki gibi özetlenebilir (Baykasoğlu, 2005):

- Bilimsel ve Mühendislik Verileri: Laboratuvar ve bilgisayar ortamında üretilen büyük miktardaki bilimsel verinin anlamlandırılması zorunluluğu nedeniyle veri madenciliği önemli bir yere sahiptir.
- Sağlık Verileri: Özellikle tarama test sonuçlarını kullanarak çeşitli hastalıklarda ön tanı konulması, acil servis hastalarında öncelik ve risk tespiti gibi pek çok alanda yaygın olarak kullanılmaktadır.
- İşletme Verileri: İş süreci sırasında üretilen pek çok çeşit ve sayıdaki veriler yönetimdeki bütün karar verme aşamalarında kullanılmaktadır. Karar sürecinde kullanılan bu veriler çalışanların performans ölçümleri, işe alımlarda tanımlanabilecek kural tespitleri gibi pek çok faydalı bilgiye dönüştürülebilmektedir.
- Finans: Bankacılık sektöründe kredi ve kredi kartı dolandırıcılığının tahmin edilmesi, kâr analizlerinin yapılması gibi alanlarda; borsa ve finansal kuruluşlarda ise hisse senedi fiyat tahminleri ve portföy verimliliği gibi alanlarda kullanılabilir.
- Eğitim Sektörü: Öğrenci bilgilerinin analiz edilerek başarı ve başarısızlık nedenleri, başarıyı arttırmaya yönelik uygulamaların tespit edilmesi gibi amaçlarla kullanılmaktadır.
- İnternet ve Web Siteleri: Web sayfalarından kullanıcı erişim desenlerinin keşfedilerek analiz işlemlerinin gerçekleştirilmesi için kullanılmaktadır.
- Perakendecilik ve Marketçilik Verileri: Bu alanda en sık tercih edilen yöntemlerin başında sepet analizi gelmektedir. Sepet analizi temelde; müşterilerin satın aldıkları ürünler arasında bulunan ilişkileri kurmak ve bu ilişkiler aracılığıyla kârlılığı ve satış oranlarını arttırmaktır (Küçüksille, 2009).

III. VERİ MADENCİLİĞİ TEKNİKLERİNDEN LOJİSTİK REGRESYON ANALİZİ

Logit Regresyon özellikle ikili bağımlı değişken için tasarlanmış doğrusal olmayan bir regresyon modelidir. Uygun dönüşümler ile doğrusallaştırılabilen bir doğrusal olmayan modeldir. Literatürde, logit regresyon aynı zamanda “Lojistik Regresyon” olarak da adlandırılmaktadır (Stock & Watson, 2007). Modelde bağımlı değişken iki kategori ile ifade ediliyorsa, model “İkili Lojistik Regresyon Modeli”, ikiden fazla kategori ile ifade ediliyorsa, “Çoklu Lojistik Regresyon Modeli” olarak adlandırılır (Leech, Barrett & Morgan, 2004). İkili lojistik regresyon modelinde bağımlı değişkenin gözlenen değeri iki olası durumu ifade etmek üzere, olayın meydana gelmesi durumunda 1, meydana gelmemesi durumunda 0 değerini alır. Modeldeki bağımsız değişkenler sürekli ve/veya kategorik değişkenler olabilmektedir ve bağımsız değişkenlerin ikili veya üçlü etkileşimleri ortak değişken olarak modele dâhil edilebilmektedir (Işığışık, 2003).

Lojistik regresyon gözlemlerin gruplara ayrılmasında kullanılan çok değişkenli analiz yöntemlerinden biridir. Varsayım kısıtı olmaması nedeni ile kullanım rahatlığının yanısıra, çözümlenmeden elde edilen modelin matematiksel olarak çok esnek olması yönüne olan ilgiyi arttırmaktadır (Tatlıdil, 1996).

Bağımlı değişkenin kategorik olması durumunda, logit regresyon modelinin kullanılması standart yöntemlere göre çok daha uygundur (Leech, Barrett & Morgan, 2004). Lojistik regresyon analizini kullanmanın amacı istatistikte kullanılan herhangi bir model kurma tekniği ile aynıdır. Bağımlı değişken ile açıklayıcı değişkenler dizisi arasındaki ilişkinin tanımlanmasında en uygun ve makul modelin bulunması amaçlanır. Bağımlı değişkeni Y , bağımsız değişkeni X olan ikili basit lojistik regresyon modelini açıklayabilmek için lojistik dağılım fonksiyonundan yararlanılır (Gürüş & Çağlayan, 2000). Bu fonksiyon aşağıdaki biçimdedir.

$$P_i = E(Y = 1 \mid X_i) = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_i)}}$$

Burada, P_i , bağımsız değişken, X_i veri iken i . bireyin belirli bir seçim yapma olasılığını (i . birey için Y ' nin 1 ve 0 değerini alma olasılığını) göstermektedir. Bu fonksiyonda, P_i ' nin hem bağımsız değişken (X)' e göre hem de parametrelere (β 'lara) göre doğrusal dışı olduğu açıkça görülmektedir. Kümülatif lojistik olasılık dağılım fonksiyonu alt sınırı sıfır üst sınırı bir olan “S” şeklinde bir eğriyi göstermektedir. Logit model doğrusal olmayan bir yapıda görünmekle birlikte, uygun dönüşümlerle doğrusallaştırılabilir. Doğrusallaştırılan bu model L

33 Üniversite Öğrencilerinin Ders Başarısını Etkileyen Faktörlerin Tahminlenmesi

“Logit” olarak adlandırılır ve bu nedenle bunun gibi modellere “Logit Model” denilir (Gujarati, 2003). Literatürde bağımsız değişkenin birden fazla olduğu durumda ikili çoklu logit regresyon modeli söz konusu olmaktadır. Bu durumda uygun dönüşümler için

$$P_i = E(Y = 1 \mid X_i) = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_i + \beta_3 X_2 + \dots + \beta_k X_k)}}$$

ile ifade edilen çoklu lojistik dağılım fonksiyonundan yararlanılır.

Logit modelde Odds ve Odds oranı önemli kavramlardır. Odds (veya olabilirlik, üstünlük) bir olasılık oranıdır. Odds (üstünlük), gerçekleşen olay sayısının gerçekleşmeyen olay sayısına oranı olarak tanımlanır. Başka bir ifadeyle bağımlı değişkenin, bağımsız değişkenin etkisi ile gözlemlenme olasılığının birinin diğerine oranla kaç kat daha fazla ya da az olarak ortaya çıkabileceğini göstermektedir (Girginer & Cankuş, 2008). Burada, ifade edilen üstünlük oranının logaritması alınmakta ve logaritması alınan bu ifade logit olarak adlandırılmaktadır. Olasılık, üstünlük ve logit kavramlarının tam olarak aynı şeyi ifade eden üç farklı yol olduğunun anlaşılması önemlidir. Üstünlük oranının açık bir yorumu vardır (Morgan & Teachman, 1988). 1’den büyük bir üstünlük oranı olayın gerçekleşmesinin olabilirliğinin arttığını, 1’den küçük bir üstünlük oranı ise olayın gerçekleşmesinin olabilirliğinin azaldığını gösterir. Lojistik regresyon denkleminde P incelenen olayın gözlenme olasılığını göstermektedir. İncelenen iki farklı olayın üstünlük değerlerinin birbirine oranına ise üstünlük oranı denir. Lojistik regresyon denkleminde üstünlük oranı, Exp (B) olarak ifade edilir. Olasılık oranı (Odds), bir olayın meydana gelme olasılığının meydana gelmeme olasılığına oranı olduğuna göre; exp (B) Y değişkeninin Xi değişkeninin etkisi ile kaç kat daha fazla ya da % kaç oranında fazla gözlenme olasılığına sahip olduğunu belirtir (Girginer & Cankuş, 2008).

Lojistik regresyon analizinde üstünlük oranları bağımlı ve bağımsız değişkenler arasındaki ilişkiyi açıklayan en önemli katsayılarıdır. Bu katsayıların doğru bir şekilde anlaşılması ve yorumlanması modelden elde edilecek sonuçların da doğruluk ve güvenilirliğini arttıracaktır. Bunu çalışmada kullandığımız araştırma modelini destekleyecek basit bir örnekle açıklayabiliriz. Örneğin 18 tane öğrencinin derse devam ile ders başarısı arasındaki ilişkiyi inceleyelim. Burada ders başarısı; dersten geçtiyse, başarılı olduysa 1, kaldıysa, başarısız olduysa 0 olacak şekilde iki sonuçlu dikotomik bağımlı değişkendir. Derse devam ise ders başarısını etkilediği düşünülen yönetmeliğe göre devam şartını sağlayarak hiç ders kaçırmayarak derslerin hepsine giren öğrencilerin 1 girmeyenlerin yani devamsızlık sınırı içinde kalsa bile az ya da çok ders kaçıranın ise 0 olarak kodlandığı iki sonuçlu dikotomik bağımsız değişkendir. 18 öğrencinin ders başarılarının, bağımsız değişken derse devam durumuna göre dağılımları Tablo 1’deki gibi olsun.

Tablo I: Üstünlük (Odds) Oranları Örneği

Bağımsız değişken ders devamı	Bağımlı değişken ders başarısı		Toplam
	Başarılı (1)	Başarısız (0)	
Derslerin hepsine girdi (1)	6	2	8
Derslerin hepsine girmedi (0)	3	7	10
Toplam	9	9	18

Burada derslerin hepsine girip başarılı olma olasılığının üstünlük değeri olasılıklar yardımıyla açıklanacak olursa;

- Bu değer derslere girip başarılı olma olasılığının (6/8), derslere girip başarılı olmama olasılığına (2/8) oranı olacaktır. Bu durumda üstünlük (odds) değeri; $[(6/8)/(2/8)]=(6/8)*(8/2)=6/2=3$ olacaktır. Yani bunun anlamı derse girip başarılı olma olasılığının üstünlüğü derse girip başarılı olmama olasılığından 3 kat daha fazladır şeklindedir.
- Benzer şekilde, derslerin hepsine girmeyip başarılı olma olasılığının üstünlük (odds) değeri; derse girmeyip başarılı olma olasılığının (3/10), derse girmeyip başarısız olma olasılığına (7/10) bölümüyle elde edilir. Bu durumda üstünlük değeri; $[(3/10)/(7/10)]=(3/10)*(10/7)=3/7=0,4285$ olacaktır. Yani burada da derse girmeyip başarılı olma olasılığının başarısız olma olasılığına oranı 0,4285'dir. Bir başka deyişle derse girmeden başarılı olma olasılığı olmama olasılığının 0,4285 katıdır. Ya da 7'de 3'dür.

Şimdi yukarıda verilen lojistik regresyon denklemine göre bu üstünlük değerlerinin logaritmaları alınırsa o zaman bağımlı ve bağımsız değişken arasındaki ilişkiyi doğrudan açıklayan üstünlük oranı (odds ratio) elde edilmiş olur.

$$P_i = (Başarı = 1/Devam) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 = -0,84746 + 1,098612x_1$$

Burada devam değişkeninin ders başarısı üzerindeki etkisini açıklayan üstünlük oranı 1.098612 çıkmıştır. Yani derse devamdaki 1 birimlik artış öğrencinin başarılılar arasına girme yani başarılı olma olasılığını üstünlük oranı 1,098612 kat arttıracaktır. Öğrenci derse devam etmediğinde yani $x_1=0$ olduğunda, derse devam etmeden başarılı olma olasılığına ilişkin üstünlük oranı -0,84746 olmaktadır. Bu durumda da, derse devam edilmezse yani devamsızlıktaki 1 birimlik artış öğrencinin başarılılar arasına girme yani başarılı olma olasılığını 0,84746 kat azaltacaktır. Katsayı negatif olduğundan arada ters yönlü etkileşim

35 Üniversite Öğrencilerinin Ders Başarısını Etkileyen Faktörlerin Tahminlenmesi

vardır. Devamsızlık azaldıkça başarılı olma olasılığı da artacaktır. Çalışmamızda oluşturulan regresyon modellerinde burada anlatıldığı şekliyle elde edilen üstünlük oranları yorumlanarak sonuçlar elde edilmiştir.

Lojistik regresyon analizi; bağımlı değişken üzerinde risk faktörü oluşturan bağımsız değişkenler ile bağımsız değişkenlerin bağımlı değişken değerini tahmin etmede ne derece etkide bulunduğunu tespit etmek için kullanılmaktadır. Lojistik regresyon analizi ile veriler olasılık kurallarına uygun olarak belirli sınıflara ayrılır ve bağımsız değişkenlerin etkileri incelenerek verilerin sınıflandırılması işlemleri gerçekleştirilir (Şekeroğlu, 2010).

Lojistik regresyon analizinin kullanılmasındaki amaç; minimum sayıda değişken kullanılarak uyum derecesi en iyi olan bağımlı ve bağımsız değişkenler arasındaki ilişkinin tanımlanması ve kabul edilen bir model kurulmasıdır (Bircan, 2004). Lojistik regresyon analizinde uygulanan adımlar aşağıdaki gibidir (Şekeroğlu, 2010):

- Modeldeki öncü grup üyeleri tespit edilir.
- Modele dâhil edilecek değişkenler belirlenir.
- Modelde bulunan parametreler tahmin edilir.
- Tahmin edilen parametrelerin anlamlılığı incelenir.
- Tahmin edilen parametreler aracılığıyla her bir gözlemin geldiği grup tahmin edilir.
- Modelde uyum iyiliği testleri yapılırken doğru sınıflandırma yüzdesi ve yapay R^2 kriterleri kullanılır.

IV. DERS BAŞARISINI ETKİLEYEN FAKTÖRLERİN LOJİSTİK REGRESYON ANALİZİYLE ARAŞTIRILMASI

Kalitenin devamlı olarak iyileşmesini hedefleyen yükseköğretim kurumları için beşeri sermaye sürekli bir analiz konusudur. Bu sebeple; öğrenci başarısının tahmin edilmesi yükseköğretim kurumları için oldukça önemlidir. Çünkü öğretim sürecindeki kalite; öğrencilerin ihtiyaçlarını karşılama yeteneğidir. Eğitim sürecinde katılımcıların yükümlülüklerini uygun etkinliklerle yerine getirmeleri sonucunda uyarlanması ve kullanılması gereken büyük miktarda veri oluşmaktadır. Bu veriler bilgiye dönüştürülerek; öğrenciler, eğitimciler ve yönetim kadrosundaki tüm katılımcıların memnuniyeti sağlanabilir. Eğitim sürecinde yer alan tüm paydaşlar yükseköğretim sisteminde bulunan verilerden veri madenciliği uygulamaları ile faydalanabilmektedir (Osmanbegovic & Suljić, 2012).

EDM (educational data mining), eğitim sistemleri tarafından derlenen ham verileri, faydalı bilgiye dönüştürme süreci olarak tanımlanmaktadır. Eğitim sürecinde kullanılan veriler ile

öğrencilerin modellenmesi, eğitimin yapılandırılması ve öğrenci başarı notunun tahmin edilmesi gibi çeşitli işlemler yapılabilmektedir. Sınıflandırma, kümeleme ve birliktelik analizi gibi pek çok veri madenciliği tekniği; yeni bilgiler ortaya çıkarmak adına eğitsel veri madenciliği alanında kullanılmaktadır (Kumar & Vijayalakshmi, 2012).

Bu çalışmada da üniversite öğrencilerinin ders başarılarını etkileyen faktörlerin neler olduğu araştırılarak ders başarısını arttırabilmek için eğitim öğretim ortamında neler yapılabileceği konusunda taraflara önerilerde bulunmaktadır.

IV.I. Araştırmada Kullanılan Veri Setinin Özellikleri

Araştırma kapsamında kullanılan veri seti, Türkiye’de eğitim öğretim veren bir devlet üniversitesinin dönem sonunda öğrencilere uyguladığı ders değerlendirme anket verileridir. Çalışma kapsamında kullanılan veriler, üç farklı öğretim üyesi için öğrenciler tarafından yanıtlanan 5820 adet anket verisinden oluşmaktadır. Anket; öğrencilerin aldıkları dersleri içerik, eğitim öğretim ortamı, öğretim üyesi performansı gibi yönlerden değerlendiren likert ölçekli 1 ve 5 arasında verilen puanlarla tanımlanmış 28 soru ve 5 tane de demografik bilgileri içeren kategorik olmak üzere toplam 33 sorudan oluşmaktadır (Ek-1).

Araştırmada kullanılan veriler yukarıda belirtildiği gibi bir devlet üniversitesinin isteyen tüm kullanıcıların faydalanabileceği şekilde tasarlanan web sayfasından elektronik ortamda elde edilmiştir. Veri seti içerisinde derslerin ve ders sınavlarının zorluk derecesiyle ve öğrencilerin sınav başarı notları ve sınavların yapılma şekliyle ilgili herhangi bir bilgi yoktur. Araştırma tamamen kurum dışından yürütüldüğünden dolayı sadece anket verileri kullanılmıştır. Ders içerikleri, derslerin adları, öğretim üyelerinin nitelikleri, eğitim öğretim alanı bilinmemektedir. Tüm analizler eldeki anket verileriyle gerçekleştirilmiştir. Araştırmada kullanılan zorluk değişkeni anket soruları içerisinde yer alan, öğrencinin dersle ilgili zorluk algılamasını tanımlayan likert ölçekli soruya verilen cevaplardan elde edilmiştir. Ders başarı değişkenleri ise tamamen eldeki anket verileri kullanılacak şekilde izleyen bölümde de ayrıntılı olarak açıklandığı gibi tanımlanmıştır. Kuşkusuz burada belirtilen derslerin zorlukları, sınavların yapılma şekli, başarı notları gibi ek bilgiler araştırmaya dâhil edildiğinde çok daha ilginç, güvenilir sonuçlar elde edilecektir.

Genellikle sosyal bilimlerde anket verilerinin güvenilirliği ile ilgili araştırmalarda kullanılan Cronbach alfa katsayısı görüş ve tutum bildiren 28 soru için hesaplanmış 0.992 çıkmıştır (Can, 2017). Buradan istatistiksel olarak anket verilerinde içsel tutarlılığın sağlandığını verilerin oldukça güvenilir veriler olduğunu söyleyebiliriz.

37 Üniversite Öğrencilerinin Ders Başarısını Etkileyen Faktörlerin Tahminlenmesi

Ayrıca Gündüz vd.'nin veri seti üzerinde yaptıkları faktör analizi sonuçlarına göre; görüş ve tutum bildiren 28 soru başlıca iki temel faktöre ayrılmıştır. Bunlardan ilki faktör-1; dersi veren öğretim üyesi performansını tanımlayan, 13 ile 28 arasındaki sorulardan oluşmaktadır. Faktör-2 ise, kalan 1-12 arasındaki soruları kapsayan dersin işlenme şekli ve eğitim öğretim ortamıyla ilgili değerlendirmelerden oluşan faktördür (Gündüz & Fokoue, 2015).

Araştırmada öğrencilerin ders başarı durumunu tanımlayan iki farklı başarı değişkeni oluşturulmuştur. Bu değişkenlerden ilki, öğrencinin ders tekrar sayısı olarak belirlenmiştir, eğer öğrenci dersi ilk alışımda geçtiyse başarılı iki ve üzerinde tekrar aldıysa başarısız olarak değerlendirilmiştir. İkinci başarı değişkeni ise; öğrencinin dersi değerlendirmeye yönelik verdiği cevaplardan 9-12 numaralı sorular üzerinden tanımlanmıştır.

Söz konusu sorular; “9.Ders büyük ölçüde zevkliydi ve derse katılmak için istek duydum.”, “10.Ders hakkındaki dönem başındaki beklentilerim karşılandı.”, “11.Ders benim mesleki gelişimime katkıda bulundu.” ve “12.Ders benim hayata ve dünyaya bakışıma yeni bir bakış açısı kazandırdı.” ifadelerinden oluşmaktadır. Ders başarı değişkeninde sorulara 3 puan ve üzerinde verenlerin dersle ilgili beklentilerini karşıladıkları düşünülerek 4 soru üzerinden ($4*3=12$) 12 puan ve üzerinde verenler için dersin başarılı değerlendirildiği 3'ten az puan verenlerin dersten beklentilerinin karşılanmadığı düşüncesiyle toplam 12 puanın altında verenler için dersin başarısız olarak değerlendirildiği düşünülmüştür.

IV.II. Veri Setinin Analize Elverişli Hale Getirilmesi

Veri setinin analize elverişli hale getirilmesiyle ilgili yapılan incelemede; veri seti içerisinde herhangi bir eksik veya gürültülü veri olmadığı görülmüştür. Veri seti üzerinde ilk başarı kriterini tanımlamaya yönelik olarak Tablo 2’de yer alan güncelleme işlemi yapılmıştır.

Tablo II. Ders Tekrar Sayılarını Update Eden Cümle

```
UPDATE OGRENCI_ANKET
SET DERS_TKRAR_SAYI = '1.Tekrar'
WHERE DERS_TKRAR_SAYI = 1

UPDATE OGRENCI_ANKET
SET DERS_TKRAR_SAYI = '2veUzeri'
WHERE DERS_TKRAR_SAYI IN (2,3)
```

İlk başarı kriterinin tanımlanmasının ardından likert ölçekli anket verisinde dersin başarısını belirlemeye yönelik olarak 9-12 numaralı “Ders büyük ölçüde zevkliydi ve derse katılmak için istek duydum.”, “Ders hakkındaki dönem başındaki beklentilerim karşılandı.”,

“Ders benim mesleki gelişimime katkıda bulundu.” ve “Ders benim hayata ve dünyaya bakışıma yeni bir bakış açısı kazandırdı.” soruları üzerinden oluşturulan başarı kriteri için veri tabanından Tablo 3’de görülen sql cümlesi yardımıyla yeni veri seti oluşturulmuştur.

Tablo III: Verilerin Ders Başarı Kriterine Göre Listelenmesi

```
select OGRETIM_UYESI,
DERS_KODU,
DERS_TKRAR_SAYI,
DERSE_KATILIM,
OGRENCI_ZORLUK_TANIMLAMASI,
ANSWER_Q1,
ANSWER_Q2,
ANSWER_Q3,
ANSWER_Q4,
ANSWER_Q5,
ANSWER_Q6,
ANSWER_Q7,
ANSWER_Q8,
ANSWER_Q9,
ANSWER_Q10,
ANSWER_Q11,
ANSWER_Q12,
(case when ANSWER_Q9+ANSWER_Q10+ANSWER_Q11+ANSWER_Q12 > 11 then 'Basarili' else 'Basarisiz' end )
BASARI_DURUM
from dbo. OGRENCI_ANKET
```

IV.III. Başarı Kriteri Olarak Ders Tekrar Sayısı ve Faktör-2 Soruları İle Gerçekleştirilen Lojistik Regresyon Analizi

Araştırmanın bu aşamasında 1. tekrarda dersi geçen öğrenciler başarılı 2 ve üzeri tekrar alan öğrenciler başarısız olarak değerlendirilerek iki sonuçlu dikotomik bir başarı değişkeni tanımlanmıştır. Gerçekleştirilen analizde bağımlı değişken ders_tkrar_sayı, bağımsız değişkenler ise 13-28 numaralı sorular dışında kalan tüm değişkenler yani 1-12 no’lu sorulara verilen cevaplardır. Bu değerlendirme kapsamında 5820 öğrenciden 4909 tanesinin başarılı, 911 tanesinin ise başarısız olarak ayrıldığı görülmüştür (Tablo 4). Veri seti cross-validation (çapraz doğrulama) tekniği ile 10 eşit parçaya ayrılmıştır. Bu aşamada yapılan bölümlenme işleminde veri setinin 9 parçası eğitim 1 parçası test için kullanılmaktadır.

Tablo IV: Başarı Kriteri Ders Tekrar Sayısı Olan 2.Faktör Soruların Lojistik Regresyon Sonuçları

=== Classifier model ===	
Variable	1.Tekrar
=====	
OGRETIM_UYESI=1.OgretimUyesi	0.2432
OGRETIM_UYESI=2.OgretimUyesi	0.7823
OGRETIM_UYESI=3.OgretimUyesi	-0.7376
DERS_KODU=Ders-1	-0.7929
DERS_KODU=Ders-2	0.4301
DERS_KODU=Ders-3	0.621
DERS_KODU=Ders-4	0.2675
DERS_KODU=Ders-5	1.4117
DERS_KODU=Ders-6	-0.86

39 Üniversite Öğrencilerinin Ders Başarısını Etkileyen Faktörlerin Tahminlenmesi

DERS_KODU=Ders-7	-0.6709
DERS_KODU=Ders-8	-0.2509
DERS_KODU=Ders-9	-0.5135
DERS_KODU=Ders-10	0.5467
DERS_KODU=Ders-11	-0.7199
DERS_KODU=Ders-12	0.2229
DERS_KODU=Ders-13	-0.2177
DERSE_KATILIM	0.2574
OGRENCI_ZORLUK_TANIMLAMASI	-0.3636
ANSWER_Q1	0.0036
ANSWER_Q2	-0.0756
ANSWER_Q3	-0.1473
ANSWER_Q4	-0.0506
ANSWER_Q5	0.1446
ANSWER_Q6	0.009
ANSWER_Q7	-0.1696
ANSWER_Q8	-0.0741
ANSWER_Q9	-0.0107
ANSWER_Q10	0.1067
ANSWER_Q11	-0.0931
ANSWER_Q12	-0.0470
Intercept	2.7265
Odds Ratios...	
Variable	1.Tekrar
=====	
OGRETIM_UYESI=1.OgretimUyesi	1.2754
OGRETIM_UYESI=2.OgretimUyesi	2.1864
OGRETIM_UYESI=3.OgretimUyesi	0.4782
DERS_KODU=Ders-1	0.4525
DERS_KODU=Ders-2	1.5373
DERS_KODU=Ders-3	1.8608
DERS_KODU=Ders-4	1.3066
DERS_KODU=Ders-5	4.103
DERS_KODU=Ders-6	0.4232
DERS_KODU=Ders-7	0.5112
DERS_KODU=Ders-8	0.7781
DERS_KODU=Ders-9	0.5984
DERS_KODU=Ders-10	1.7276
DERS_KODU=Ders-11	0.4868
DERS_KODU=Ders-12	1.2497
DERS_KODU=Ders-13	0.8044
DERSE_KATILIM	1.2936
OGRENCI_ZORLUK_TANIMLAMASI	0.6952
ANSWER_Q1	1.0157
ANSWER_Q2	1.0027
ANSWER_Q3	0.9796
ANSWER_Q4	1.0013
ANSWER_Q5	1.0465
ANSWER_Q6	1.0145
ANSWER_Q7	0.9833
ANSWER_Q8	1.0244
ANSWER_Q9	1.0500
ANSWER_Q10	1.1038

```

ANSWER_Q11                1.0243
ANSWER_Q12                1.0349

=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances   4895      84.1065 %
Incorrectly Classified Instances   925      15.8935 %
Kappa statistic                 0.0611
Mean absolute error             0.2408
Root mean squared error         0.3493
Relative absolute error         91.1632 %
Root relative squared error     96.14 %
Total Number of Instances       5820

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
      0.988   0.948   0.849   0.988   0.913   0.709   1.Tekrar
      0.052   0.012   0.435   0.052   0.092   0.709   2veUzeri
Weighted Avg.  0.841   0.802   0.784   0.841   0.784   0.709

=== Confusion Matrix ===

  a  b  <-- classified as
4848 61 |  a = 1.Tekrar
 864 47 |  b = 2veUzeri

```

Tablo 4'te görülen üstünlük (Odds) oranlarının incelenmesi önemlidir; çünkü Odds (olabilirlik veya üstünlük) gerçekleşen olay sayısının gerçekleşmeyen olay sayısına oranını temsil etmektedir (Özdil, Urdaletova & Yılmaz, 2010). Başka bir ifadeyle bağımlı değişkenin, bağımsız değişkenin etkisi ile gözlemlenme olasılığının birinin diğerine oranla kaç kat daha fazla ya da az olarak ortaya çıkabileceğini göstermektedir (Girginer & Cankuş, 2008). Yani 1'den büyük üstünlük oranları bağımlı değişkeni olumlu yönde etkileyerek başarılı olma olasılığını üstünlük oranı kadar arttırmakta, 1'den küçük katsayılar ise bağımlı değişkeni olumsuz etkileyerek başarılı olma olasılığını azaltmakta veya başarısız olma olasılığını üstünlük oranı kadar arttırmaktadır. Regresyon analizine ait üstünlük (Odds) oranları incelendiğinde ise (Tablo 4);

- 1 ve 2 numaralı öğretim üyelerinin üstünlük oranları 1'den büyük çıkmıştır. 3 no'lu öğretim üyesinin üstünlük oranı 1'den küçüktür. Bunun anlamı eğer öğrenciler dersi 3 no'lu öğretim üyesinden almak yerine 1 veya 2 no'lu öğretim üyesinden alırlarsa başarılı olma yani dersi ilk tekrarda geçme olasılıkları 2 no'lu öğretim üyesi için 2,1864 kat 1 no'lu öğretim üyesi için 1,2754 kat artacaktır. Ancak 3 no'lu öğretim üyesinden almaları durumunda dersi ilk tekrarda geçme olasılıkları 0,4782 kat azalacaktır. Yani görelî

41 Üniversite Öğrencilerinin Ders Başarısını Etkileyen Faktörlerin Tahminlenmesi

olarak burada 2 no'lu öğretim üyesi performansı en iyi olan öğretim üyesi olarak çıkmıştır.

- Dersler için yapılan incelemede ise 2,3,4,5,10 ve 12 no'lu dersler üstünlük oranları 1'den büyük olan derslerdir. Yani bu öğrenciler bu dersleri seçerlerse bu derslerden ilk tekrarda geçme olasılıkları oldukça yüksektir. 5 no'lu ders 4.1030 üstünlük oranıyla ilk tekrarda başarılı olunma olasılığı en yüksek olan derstir. Üstünlük oranı 1'den küçük olan derslerde ise en küçük oran (0,8044) 13 numaralı derste elde edilmiştir. Bu ders ise seçilmesi halinde ilk tekrarda başarılı olma olasılığı en az olan derstir.
- Derse devam değişkenine ait üstünlük oranı ise 1,2936'dır. Bu da olumlu olarak diğer şartlar sabitken derse devam arttıkça öğrencilerin aldıkları dersti ilk tekrarda geçme olasılıklarının 1,2936 kat arttığı anlamına gelir.
- Zorluk tanımlaması dikkate alındığında da, burada üstünlük oranı 0,6952 olarak elde edilmiştir. Yani zorluk derecesi arttıkça dersin ilk tekrarda geçilme olasılığı da azalmakta veya dersin 2 ve daha çok sayıda tekrarda geçilme olasılığı artmaktadır. Bu da beklentilere uygun şekilde öğrencilerin zor olarak algıladıkları derslerden daha zor başarılı oldukları sonucunu bize vermektedir.
- Faktör-2 yani dersin içerik, yapıma şekli ve dersten beklentilerle ilgili sorulara verilen yanıtlar için yapılan incelemede ise 3 ve 7 no'lu 2 soru dışında tüm sorularda üstünlük oranları 1'den büyük çıkmıştır. Üstünlük oranı 1'den küçük olan sorularda da oranlar sırasıyla 0,9796 ve 0,9833 olarak aslında 1'e çok yakın değerlerdir. Bu iki soru dışındaki 10-no'lu "Ders hakkındaki dönem başındaki beklentilerim karşılandı." ifadesi üstünlük oranı (1,1038) en yüksek olan ifade olup dersti ilk tekrarda geçme olasılığını en çok etkileyen ifadedir. Yani öğrencilerin dersten beklentileri arttıkça dersti ilk tekrarda geçme olasılıkları da artmaktadır. Başarıyı az da olsa olumsuz etkileyen iki ifade 3 no'lu "Ders için harcanan enerji dönem başında verilen kredi miktarımı karşılıyordu" ve 7 no'lu "Ders saha çalışmasına, laboratuvar çalışması ve uygulama yapmaya uygundu." ifadeleridir.

Lojistik regresyon analizi sonucunda yapılan tahminlerde doğru sınıflandırma oranı %84,1 gibi oldukça yüksek bir değer çıkmıştır. Yani 5820 adet verinin 4895 tanesi doğru sınıflandırılırken yalnızca 925 tanesi yanlış sınıflandırılmıştır. Confusion Matrix incelendiğinde 61 adet verinin a yani 1.Tekrar olarak sınıflandırılması gerekirken b yani 2 ve üzeri olarak sınıflandırıldığı, aynı zamanda 864 adet verinin b yani 2 ve üzeri olması gerekirken 1.Tekrar olarak sınıflandırıldığı görülmüştür.

IV.IV. Dersin Başarı Kriteri İçin Lojistik Regresyon Analizi

Bu aşamada 9,10,11 ve 12 numaralı sorular dikkate alınarak dersin başarılı olarak verilip verilmediğiyle ilgili bir başarı değişkeni tanımlanmıştır. Söz konusu sorular; “9.Ders büyük ölçüde zevkliydi ve derse katılmak için istek duydum.”, “10.Ders hakkındaki dönem başındaki beklentilerim karşılandı.”, 11.Ders benim mesleki gelişimime katkıda bulundu.” ve “12.Ders benim hayata ve dünyaya bakışıma yeni bir bakış açısı kazandırdı.” şeklinde öğrencilerin dersle ilgili memnuniyetlerini tanımlayan sorulardır. Bu sorulara verilen cevaplar 1: katılmıyorum, 5.katılıyorum olacak şekilde 1 ve 5 arasındaki puanlarla alınmıştır. Dersin beklentileri karşılayacak şekilde verilip verilmediğiyle ilgili olarak tanımlanan ders başarı değişkeninde sorulara 3 puan ve üzerinde verenlerin dersle ilgili beklentilerini karşıladıkları düşünülerek 4 soru üzerinden ($4*3=12$) 12 puan ve üzerinde verenler için dersin başarılı değerlendirildiği 3’ten az puan verenlerin dersten beklentilerinin karşılanmadığı düşüncesiyle toplam 12 puanın altında verenler için dersin başarısız olarak değerlendirildiği düşünülmüştür. Buna göre puan toplamının 12’den az ve çok olmasına göre iki sonuçlu dikotomik ders başarı değişkeni tanımlanmıştır. Lojistik regresyon analizi için bağımlı değişken *basari_durum*, bağımsız değişkenler ise 13-28 numaralı sorular dışındaki tüm değişkenlerdir. 5820 öğrencinin 3945’inin dersleri başarılı bulduğu 1875’inin ise dersleri başarısız olarak nitelendirdiği görülmektedir.

Tablo V: Dersin Başarı Kriteri İçin Lojistik Regresyon Sonuçları

=== Classifier model===	
Variable	Basarili
=====	
OGRETIM_UYESI=1.OgretimUyesi	-0.1605
OGRETIM_UYESI=2.OgretimUyesi	0.1719
OGRETIM_UYESI=3.OgretimUyesi	-0.0574
DERS_KODU=Ders-1	0.1854
DERS_KODU=Ders-2	-0.1419
DERS_KODU=Ders-3	-0.2527
DERS_KODU=Ders-4	0.6296
DERS_KODU=Ders-5	-0.798
DERS_KODU=Ders-6	0.1741
DERS_KODU=Ders-7	0.0526
DERS_KODU=Ders-8	0.33
DERS_KODU=Ders-9	0.3984
DERS_KODU=Ders-10	-0.2369
DERS_KODU=Ders-11	0.5432
DERS_KODU=Ders-12	-0.5265
DERS_KODU=Ders-13	-0.0908
DERS_TKRAR_SAYI	-0.0045
DERSE_KATILIM	0.1443
OGRENCI_ZORLUK_TANIMLAMASI	-0.1236
ANSWER_Q1	0.0470
ANSWER_Q2	0.4359

43 Üniversite Öğrencilerinin Ders Başarısını Etkileyen Faktörlerin Tahminlenmesi

ANSWER_Q3	0.2791					
ANSWER_Q4	-0.0638					
ANSWER_Q5	-0.0788					
ANSWER_Q6	-0.4792					
ANSWER_Q7	0.3601					
ANSWER_Q8	0.1397					
Intercept	1.4802					
Odds Ratios...						
Variable	Basarili					
=====						
OGRETIM_UYESI=1.OgretimUyesi	0.8517					
OGRETIM_UYESI=2.OgretimUyesi	1.1876					
OGRETIM_UYESI=3.OgretimUyesi	0.9442					
DERS_KODU=Ders-1	1.2036					
DERS_KODU=Ders-2	0.8677					
DERS_KODU=Ders-3	0.7767					
DERS_KODU=Ders-4	1.8768					
DERS_KODU=Ders-5	0.4502					
DERS_KODU=Ders-6	1.1902					
DERS_KODU=Ders-7	1.054					
DERS_KODU=Ders-8	1.391					
DERS_KODU=Ders-9	1.4895					
DERS_KODU=Ders-10	0.7891					
DERS_KODU=Ders-11	1.7215					
DERS_KODU=Ders-12	0.5907					
DERS_KODU=Ders-13	0.9132					
DERS_TKRAR_SAYI	1.0045					
DERSE_KATILIM	1.1553					
OGRENCI_ZORLUK_TANIMLAMASI	0.8837					
ANSWER_Q1	1.0721					
ANSWER_Q2	1.0396					
ANSWER_Q3	1.0783					
ANSWER_Q4	1.0350					
ANSWER_Q5	0.9989					
ANSWER_Q6	0.9964					
ANSWER_Q7	1.1245					
ANSWER_Q8	2.2094					
=== Stratified cross-validation ===						
=== Summary ===						
Correctly Classified Instances	5332 91.6151 %					
Incorrectly Classified Instances	488 8.3849 %					
Kappa statistic	0.8055					
Mean absolute error	0.1186					
Root mean squared error	0.2489					
Relative absolute error	27.1621 %					
Root relative squared error	53.2603 %					
Total Number of Instances	5820					
=== Detailed Accuracy By Class ===						
TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class

	0.95	0.155	0.928	0.95	0.939	0.964	Basarili
	0.845	0.05	0.889	0.845	0.867	0.964	Basarisiz
Weighted Avg.	0.916	0.121	0.916	0.916	0.916	0.964	
=== Confusion Matrix ===							
a	b	<-- classified as					
3747	198	a = Basarili					
290	85	b = Basarisiz					

Tablo 5'te yer alan lojistik regresyon analiz sonuçlarında bulunan üstünlük (Odds) oranları incelendiğinde;

- Öğretim üyeleri bakımından yapılan değerlendirmede, 2 numaralı öğretim üyesinin üstünlük oranı 1'den büyük çıkmıştır. Yani dersin bu öğretim üyesi tarafından verilmesi dersin başarılı olma/verilme olasılığını 1,876 kat arttıracaktır. Aynı değişken (2. Öğretim üyesi) diğer modelde de öğrencilerin ders başarısını olumlu yönde etkilemiştir. Diğer öğretim üyelerinin üstünlük oranları 1'den küçük çıkmıştır. Bunun anlamı derslerin bu iki öğretim üyesi tarafından verilmesi halinde derslerin başarılı olma/verilme olasılığı sırasıyla 0,8517 ve 0,9442 kat azalacaktır.
- Dersler için yapılan değerlendirmede ise 1,4,6,7,8,9 ve 11 no'lu derslerin üstünlük oranları 1'den büyüktür. Yani bu derslerin seçilmesi genel olarak eğitim kurumunda derslerin başarılı verilme olasılığını arttıracaktır. Örneğin 4, 9 ve 11 numaralı derslerde en büyük katsayılar elde edilmiştir. Bu üstünlük oranları sırasıyla 1,8768 ve 1,7215 ve 1,4895'dir. Yani bu derslerin seçilmesi ders başarılı olma olasılığını katsayılar kadar arttıracaktır. Üstünlük oranı 1'den küçük olan 2, 3, 5, 10, 12 ve 13 no'lu derslerin seçilmesi ise dersin başarılı verilme/olma olasılığını azaltacaktır. 5 no'lu ders en küçük üstünlük oranıyla (0,4502) dersin başarılı olmama olasılığını en çok arttıran derstir.
- Ders tekrar sayısı değişkeninde katsayı 1'e yakın olmakla birlikte 1'den büyüktür (1,0045). Bu da ders başarısını olumlu yönde etkileyen bir değişkendir. Yani tekrar sayısı arttıkça dersin başarılı algılanma olasılığı da artmaktadır.
- Derse katılım değişkenine ait üstünlük oranı 1,1553'tür. Burada da derse devam arttıkça öğrencilerin derse başarılı algılama olasılığı artmaktadır. Burada da ders tekrar sayısı ile tutarlı bir sonuç ortaya çıkmıştır. Yani öğrencinin derse olan ilgisi arttıkça dersin başarılı verildiği algısı da artmakta güçlenmektedir.
- Faktör-2'ye ait dersin içerik, verilme şekli, dersle ilgili beklentileri tanımlayan sorulara verilen cevapların ders başarısına etkisi incelendiğinde; 5 ve 6. no'lu sorular dışındaki tüm sorularda üstünlük oranları 1'den büyük çıkmıştır. Özellikle 8, 7 ve 3 numaralı

45 Üniversite Öğrencilerinin Ders Başarısını Etkileyen Faktörlerin Tahminlenmesi

sorularda en yüksek oranlar elde edilmiştir. Bu oranlar sırasıyla 2,2094, 1,1245 ve 1,0783'tür. Bu ifadeler; “8.Sınav, proje ve quizler öğrenmeye yardımcı oldu.”, “7.Ders saha çalışmasına, laboratuvar çalışması ve uygulama yapmaya uygundu.” ve “3.Ders için harcanan enerji dönem başında verilen kredi miktarını karşılıyordu.” şeklindedir. Bu unsurlardaki gelişmeler dersin başarılı verilme olasılığını “üstünlük oranları” kadar kat arttıracaktır. Üstünlük oranlarının 1'den küçük olduğu 5 ve 6 no'lu ifadeler ise, “5.Sınav, proje ve quizler öğrenmeye yardımcı oldu.” ve “6.Ders kitabı ve ders için verilen diğer kaynaklar yeterince günceldi.” ifadeleridir. Yani sınav, quiz ve ödevler ve güncellik ders başarısını etkilemeyip beklenenin aksine olasılık olarak azaltmaktadır.

Elde edilen lojistik regresyon sonuçları incelendiğinde; 9 parça veri seti üzerinde gerçekleştirilen öğrenme işlemi ve 1 parça üzerinde yapılan test sonrasında yapılan tahminlerde %91,6 oranında başarı sağlanmıştır. 5820 adet veriden 5332 tanesi doğru sınıflandırılırken yalnızca 488 tanesi yanlış sınıflandırılmıştır. Confusion Matrix incelendiğinde de, 198 adet verinin a yani Başarılı olarak sınıflandırılması gerekirken b yani Başarısız olarak sınıflandırıldığı, aynı zamanda 290 adet verinin b yani Başarısız olması gerekirken Başarılı olarak sınıflandırıldığı görülmüştür.

V. SONUÇ

Öğrencilerin birinci tekrarda dersi başarılı olarak tamamlamalarını etkileyen faktörlerin belirlenmesinde; 1 ve 2 numaralı öğretim üyelerinden alınan derslerin öğrencilerin başarılı olma olasılığını arttırdığı, 5, 3 ve 10 numaralı derslerin başarılı olma olasılığını en çok arttıran dersler olduğu, 13 numaralı dersin ise başarısız olma olasılığını en çok arttıran yani başarılı olma olasılığını en çok azaltan ders olduğu anlaşılmıştır. Öğrencinin derse katılımındaki artışın derste başarılı olma ihtimalini de arttırdığı görülmüştür. “10.Ders hakkındaki dönem başındaki beklentilerim karşılandı.”, “5. Sınıfta yapılan tartışmalar, ödevler, uygulamalar ve çalışmalar tatmin edici idi.” ve “1.Sınıfta yapılan tartışmalar, ödevler, uygulamalar ve çalışmalar tatmin edici idi.” ifadeleri başarılı olma olasılığını en çok arttıran ifadeler olmuştur.

Dersin beklentileri karşılayacak şekilde başarılı verilip verilmediğiyle ilgili yapılan analizde ise burada da 2 numaralı öğretim üyesinden alınan derslerin, dersin başarılı olarak değerlendirilme olasılığını arttırdığı anlaşılmıştır. 4, 11 ve 9 numaralı derslerin seçilmesi halinde dersin başarılı algılanma olasılığının artacağı, ancak 13 numaralı dersin başarılı olma olasılığını azaltan ders olduğu ortaya çıkmıştır. Hem ders tekrar sayısındaki artışın hem de derse devamın, öğrencilerin dersi başarılı olarak değerlendirme olasılığını arttırdığı anlaşılmıştır.

Elde edilen lojistik regresyon sonuçlarına göre; öğrencilerin derse katılım için istek duyması, öğrencilerin mesleki gelişimlerini arttırabilmeleri ve farklı bakış açıları kazanmalarının öğrencilerin başarısını ölçmede daha iyi sonuçlar verdiği görülmektedir. Öğrencilerin bu kazanımları elde ettiklerinde dersi de başarılı olarak değerlendirdikleri tespit edilmiştir. Bu nedenle derslerin başarısını arttırmaya yönelik olarak; öğrencilere yeni bakış açılarının kazandırılabilceği bilgilendirmeler yapılabileceği ve farklı uygulamalar gerçekleştirilebileceğinin gerekli olduğu düşünölmektedir. Ayrıca öğrencilerin derse katılımını arttıracak faaliyetlerin gerçekleştirilmesinin derse devamı arttıracacağı ve buna bağılı olarak dersin başarısının da bundan olumlu yönde etkileneceği anlaşılmıştır. Bunun yanı sıra, öğrencilerin ders tekrar sayılarına göre dersi başarılı ya da başarısız olarak değerlendirdiği sonucuna ulaşılmıştır. Yani öğrenciler eğer ilk tekrarda dersi geçtilerse dersi başarılı olarak değerlendirmiş ancak iki ve üzeri tekrar aldılarsa dersi başarısız olarak değerlendirmişlerdir. Bunun yanı sıra anketin uygulanma şekli değiştirilerek final sınavından önceki son derste uygulanmasının öğrencinin daha objektif değerlendirme yapmasını sağlayacağı kanaatine ulaşılmıştır.

REFERANSLAR

- Baykasoğlu, A. 2005. *Veri madenciliği ve çimento sektöründe bir uygulama*. 7. Akademik Bilişim, Gaziantep.
- Bircan, H. 2004. Lojistik regresyon analizi: tıp verileri üzerine bir uygulama. *Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 2: 185-208.
- Can, Ş. (2017). *Veri Madenciliği Ve Eğitim Sektöründe Bir Uygulama*. Yüksel Lisans Tezi, Manisa Celal Bayar Üniversitesi Sosyal Bilimler Enstitüsü, Manisa.
- Coşkun, C. & Baykal, A. 2011. *Veri Madenciliğinde Sınıflandırma Algoritmalarının Bir Örnek Üzerinde Karşılaştırılması*. Akademik Bilişim Konferansı Bildirileri, Malatya.
- Dener, M., Dörterler, M. & Orman, A. 2009. *Açık Kaynak Kodlu Veri Madenciliği Programları: Weka'da Örnek Uygulama*. Akademik Bilişim'09 - XI. Akademik Bilişim Konferansı Bildirileri, Harran Üniversitesi. 787-796.
- Girginer, N. & Cankuş, B. 2008. Tramvay Yolcu Memnuniyetinin Lojistik Regresyon Analiziyle Ölçülmesi: Etram Örneği. *Celal Bayar Üniversitesi, Yönetim ve Ekonomi Dergisi*. 15(1): 181-193.
- Gujarati, D.N. 2003. Basic Econometrics, Mc-Graw Hill. New York.
- Güriş, S. & Çağlayan, E. 2000 Ekonometri Temel Kavramlar, Der Yayınları, İstanbul.
- Gündüz, N. & Fokoue, E. 2015. Pattern Discovery in Students' Evaluations of Professors A Statistical Data Mining Approach. *To appear in the Journal of Applied Statistics*. 1501: 1-20.
- Işığışık, E. 2003. *Bebeklerin Doğum Ağırlıklarını ve Boylarını Etkileyen Faktörlerin Lojistik Regresyon Analizi ile Araştırılması*. VI. Ulusal Ekonometri ve İstatistik Sempozyumu, Gazi Üniversitesi, Ankara.
- Koç, M. Avşaroğlu, S. & Sezer, A. 2004. Üniversite Öğrencilerinin Akademik Başarıları İle Problem Alanları Arasındaki İlişki. *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*. 11: 483-498.
- Kumar, S.A. & Vijayalakshmi, M.M. 2012. *Mining Of Student Academic Evaluation Records in Higher Education*. International Conference on Recent Advances in Computing and Software Systems (pp. 67-70), SSN College of Engineering, Old Mahabalipuram Rd, Kalavakkam Chennai, April, India.
- Kurt, Ç. & Erdem O.A. (2012). Öğrenci Başarısını Etkileyen Faktörlerin Veri Madenciliği Yöntemleriyle İncelenmesi. *Politeknik Dergisi*. 15(2): 111-116.

- Küçükşille, E. 2009. *Veri Madenciliđi Süreci Kullanılarak Portföy Performansının Deđerlendirilmesi Ve İmkb Hisse Senetleri Piyasasında Bir Uygulama*. Doktora Tezi. Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü, Isparta.
- Leech, N. L., Barrett, K. C. & Morgan, G. A. 2004. *Spss For Intermediate Statistics: Use and Interpretation*, Lawrance Erlbaum Associates Publishers, Manwah New Jersey.
- Morgan, P. S. & Teachman, J. D. 1988. Logistic regression: Descriptions, examples, and comparisons. *Journal of Marriage and the Family* 50: 929-936.
- Osmanbegović, E. & Suljić, M. 2012. Data mining approach for predicting student performance. *Economic Review*. 10 (1): 3-12.
- Özdil, T., Urdaletova, A. & Yılmaz, C. 2010. *İktisadi ve İdari Bilimler Fakültesi Öğrencilerinin Ders Başarılarını Etkileyen Faktörlerin Lojistik Regresyon Analiziyle Araştırılması*. 2. Uluslararası Balkanlarda Sosyal Bilimler Kongresi. Düzenleyen: Sakarya Üniversitesi, Priştine Üniversitesi, Bozok Üniversitesi. Kosova.
- Savaş, S., Topalođlu, N. & Yılmaz, M. 2012. Veri Madenciliđi ve Türkiye'deki Uygulama Örnekleri. *İstanbul Üniversitesi Fen Bilimleri Dergisi*. 11(21): 1-23.
- Stock, J.H. & Watson, M.W., 2007. *Introduction to Econometrics*. Pearson Addison Wesley, Boston.
- Tatlıdil, H. 1996. *Uygulamalı Çok Deđerşkenli İstatistiksel Analiz*, Akademi Matbaası, Ankara.
- Şekerođlu, S. 2010. *Hizmet Sektöründe Bir Veri Madenciliđi Uygulaması*. Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.

EKLER

Ek-1: Araştırmada Kullanılan Anket Formu

Değişken	Açıklama	Değer İçeriği
Instr	Öğretim Üyesi	1,2,3
Class	Ders Kodu	1-13
Repeat	Ders Tekrar Sayısı	1,...
Attendance	Derse Katılım Sayısı	0,1,2,3,4
Difficulty	Zorluk Tanımı	1,2,3,4,5
Q1	Dönem ders içeriği, öğrenme tekniği ve dönem başındaki değerlendirme sistemi yeterli	1-5
Q2	Dönem başında dersin amaç ve hedefleri açıkça belirlendi mi?	1-5
Q3	Ders için harcanan enerji dönem başında verilen kredi miktarını karşılıyordu.	1-5
Q4	Konular sınıfta ilk gün açıklanan müfredata göre öğretildi.	1-5
Q5	Sınıfta yapılan tartışmalar, ödevler, uygulamalar ve çalışmalar tatmin edici idi.	1-5
Q6	Ders kitabı ve ders için verilen diğer kaynaklar yeterince günceldi.	1-5
Q7	Ders saha çalışmasına, laboratuvar çalışması ve uygulama yapmaya uygundu.	1-5
Q8	Sınav, proje ve quizler öğrenmeye yardımcı oldu.	1-5
Q9	Ders büyük ölçüde zevkliydi ve derse katılmak için istek duydum.	1-5
Q10	Ders hakkındaki dönem başındaki beklentilerim karşılandı.	1-5
Q11	Ders benim mesleki gelişimime katkıda bulundu.	1-5
Q12	Ders benim hayata ve dünyaya bakışıma yeni bir bakış açısı kazandırdı.	1-5
Q13	Öğretim üyesinin bilgisi yeterli ve günceldi.	1-5
Q14	Öğretim üyesi derse hazırlıklı geldi.	1-5
Q15	Öğretim üyesi ders planına uygun olarak işledi.	1-5
Q16	Öğretim üyesi dersi yeterince anlaşılır biçimde anlattı.	1-5
Q17	Öğretim üyesi derse zamanında geldi.	1-5
Q18	Öğretim üyesinin konuşmalarını takip etmek kolay ve anlaşılırdı.	1-5
Q19	Öğretim üyesi ders saatlerini etkin bir biçimde kullandı.	1-5
Q20	Öğretim üyesi dersi anlattı ve öğrencilere yardımcı olmak için istekliydi.	1-5
Q21	Öğretim üyesi öğrencilere karşı olumlu bir tavır içerisindeydi.	1-5
Q22	Öğretim üyesi öğrencilerinin dersle ilgili görüşlerine karşı açık ve saygılıydı.	1-5
Q23	Öğretim üyesi derse katılımı teşvik etti.	1-5
Q24	Öğretim üyesi dersle ilgili ev ödevi, proje verdi ve öğrencilere yardım etti.	1-5
Q25	Öğretim üyesi dersle ilgili sorulara derste ve ders dışında yanıt verdi.	1-5
Q26	Öğretim üyesinin değerlendirmesi (vize, final, quiz vb.) ders hedeflerini etkin bir şekilde ölçmüştür.	1-5
Q27	Öğretim üyesi sınav sorularının çözümlerini verip bunları öğrencilerle tartıştı.	1-5
Q28	Öğretim üyesi tüm öğrencilere objektif bir şekilde yaklaştı.	1-5