

Öğrenci Ders Değerlendirme Verileriyle Ders Zorluğu ve Memnuniyeti Tahmini

Adem Halit Karagöz, Cem Civelek, Yavuz Selim Kaya

Bilgisayar Mühendisliği, Teknoloji Fakültesi, Gazi Üniversitesi, Ankara, Türkiye

ahalit.karagoz@gazi.edu.tr

Bilgisayar Mühendisliği, Teknoloji Fakültesi, Gazi Üniversitesi, Ankara, Türkiye

cem.civelek.@gazi.edu.tr

Bilgisayar Mühendisliği, Teknoloji Fakültesi, Gazi Üniversitesi, Ankara, Türkiye

yselem.kaya@gazi.edu.tr

ÖZET

Bu çalışmada, Türkiye’deki üniversite öğrencilerinin ders değerlendirme anketleri analiz edilerek, dersin zorluk düzeyi ve öğrenci memnuniyeti makine öğrenmesi algoritmalarıyla tahmin edilmiştir. UCI Machine Learning Repository’den alınan 5.820 gözlemden oluşan “Turkey Student Evaluation” veri seti kullanılmıştır. Anketlerde tüm sorulara aynı şekilde yanıt veren öğrenciler (“straight-liner”) %51 oranında tespit edilmiş ve analizlerde hem ayrı hem birlikte değerlendirilmiştir. Sayısal veriler Z-skora standardize edilmiş, kategorik değişkenler One-Hot kodlama yöntemiyle dönüştürülmüştür.

Modelleme sürecinde Rastgele Orman, Karar Ağacı, K-En Yakın Komşu, Lojistik Regresyon, SVM ve Gradient Boosting algoritmaları uygulanmıştır. Ders zorluğu tahmininde en yüksek doğruluk %58,9 ile Gradient Boosting ile elde edilirken, öğrenci memnuniyetinde en başarılı sonuç %93,7 doğrulukla Lojistik Regresyon modelinde gözlenmiştir. Özellik önem analizine göre, ders içeriği ve öğretim elemanının etkileşimi gibi faktörler, memnuniyet üzerinde belirleyici olmuştur.

Bu çalışma, öğrenci geri bildirimlerini veri temelli bir yaklaşımla analiz ederek, öğretim kalitesini iyileştirmeye yönelik stratejilere katkı sunmayı amaçlamaktadır. Ayrıca, tekdüze yanıt davranışının model performansı üzerindeki etkileri detaylı şekilde ele alınarak literatüre özgün bir katkı sağlanmıştır.

Anahtar Kelimeler: Makine öğrenmesi, öğrenci memnuniyeti, ders zorluğu, straight-liner, eğitim verisi, sınıflandırma algoritmaları, özellik önem analizi, Türkiye Öğrenci Değerlendirme veri seti

1. Giriş

Son yıllarda üniversitelerde, eğitimle ilgili kararların daha sağlıklı alınabilmesi için veri temelli planlama ve analiz yöntemlerine olan ilgi hızla artmıştır. Pek çok yükseköğretim kurumu, öğretim kalitesini değerlendirmek amacıyla öğrenci memnuniyet anketlerini temel ölçüt olarak benimsemiştir.

Dönem sonunda uygulanan, çok sayıda sorudan oluşan ve Likert ölçeğine dayanan bu anketler, nicel anlamda zengin veriler sunsa da, binlerce yanıtın geleneksel yollarla tek tek incelenmesi hâlâ oldukça zaman alıcıdır ve yorumların sübjektifliğine neden olabilir. Bu durum, öğretim elemanlarının derslerini hızlıca geliştirme ve anlık geri bildirimlere göre hareket etme yetisini kısıtlamaktadır.

Bu nedenle, elde edilen verilerin daha objektif, ölçeklenebilir ve gerçek zamanlı analizine olan ihtiyaç her zamankinden daha fazladır. Burada devreye yapay zekâ ve makine öğrenmesi tabanlı analiz araçları girmektedir.

Bu çalışmada, UCI Machine Learning Repository’de yayımlanan ve 5820 öğrenciden toplanan verilerle oluşturulmuş “Turkey Student Evaluation” veri seti incelenmektedir. Veri seti, 28 anket sorusunun yanı sıra ders kodu, derse devam durumu, dersi tekrar etme sayısı ve öğrencilerin algıladığı zorluk düzeyi gibi değişkenleri de içermektedir.

Farklı fakültelerden birçok dersi kapsayan bu veri seti, önceki çalışmalarda kullanılan daha dar kapsamlı örneklem gruplarına kıyasla daha genellenebilir sonuçlar üretme potansiyeline sahiptir. Ancak veride dikkat çeken önemli bir durum, katılımcıların %51’inin anketin tüm sorularına aynı cevabı vermesidir. “Straight liner” olarak adlandırılan bu tür yanıtlar, öğrencilerin düşük çaba gösterdiğini gösterebilir ve analizlerde sapmaya yol açarak modelin doğruluğunu olumsuz etkileyebilir.

Bu çalışmada, dersin zorluk seviyesi (1–5 arasında) ve anket yanıtlarından türetilmiş üç seviyeli öğrenci memnuniyeti kategorileri tahmin edilmektedir. Bu amaçla, Rastgele Orman, Karar Ağacı, K-En Yakın Komşu, Gradient Boosting, Lojistik Regresyon ve çizgisel Destek Vektör Makineleri (SVM) gibi makine öğrenmesi yöntemleri karşılaştırılmıştır. Ayrıca, straight liner davranışının model başarısı üzerindeki etkisi de detaylı bir şekilde analiz edilmiştir.

Araştırma, Türkiye bağlamında hem öğrenci memnuniyetini hem de ders zorluk seviyesini birlikte ele alan ilk makine öğrenmesi temelli çalışma olması açısından öncü bir niteliğe sahiptir. Ders kodu ve derse tekrar sayısı gibi bilgilerle anket verilerini birleştiren çok kaynaklı özellik mühendisliği yaklaşımı, hem tahmin gücü hem de modelin açıklanabilirliği açısından önemli katkılar sunmaktadır.

Bu çalışmadan elde edilecek sonuçların, üniversite yönetimlerine ders bazlı erken uyarı sistemleri kurma ve dönem ortasında olağandışı zorluk artışlarına müdahale etme konusunda önemli fırsatlar sunması beklenmektedir.

Makalenin kalan bölümleri şu şekilde yapılandırılmıştır: İkinci bölüm veri seti, ön işleme adımları ve kullanılan algoritmalar; üçüncü bölüm bulgular ve tartışmalar; dördüncü bölüm ise sonuçlar ve geleceğe yönelik araştırma önerilerini içermektedir.

2. Materyal ve Yöntem

Bu bölümde, çalışmanın temelini oluşturan veri setinin özellikleri, sınıflandırma sürecinde kullanılan makine öğrenmesi algoritmalarının detayları ve konuyla ilişkili önceki çalışmaların değerlendirildiği literatür taraması sırasıyla ele alınmaktadır.

2.1. Veri Seti

Çalışmada UCI Machine Learning Repository’deki “*Turkey Student Evaluation*” veri seti kullanılmıştır. Veri kümesi 5 820 öğrenci gözleminden oluşur ve 28 adet Likert ölçekli anket maddesi (Q1–Q28, 1 = “kesinlikle katılmıyorum”, 5 = “kesinlikle katılıyorum”) ile dört ders-bağımlı nitelik (difficulty, attendance, nb.repeat, class) içerir . Kod akışında önce eksik değer analizi yapılmış, eksik veri olmadığı gözlenmiştir . Ardından “*straight-liner*” (SL) olarak adlandırılan, tüm sorulara aynı yanıtı veren öğrenciler %51 oranında tespit edilmiştir . SL örneklerinin neden olabileceği varyans daralmasını incelemek amacıyla veri iki alt küme halinde (SL ve SL dışı, yani DY) analiz edilmiştir.

Ön işleme adımlarında *instr* sütunu veri setinde mevcutsa çıkarılmış, *class* özniteliği kategorik tipe dönüştürülmüştür. Sayısal özniteliklere StandardScaler ile Z-skor dönüşümü, kategorik özniteliklere One-Hot Encoder ile eşlenik gösterim uygulanmış ve bu işlemler ColumnTransformer içinde birleştirilmiştir . Eğitim ve test bölünmesi, hedef dağılımını korumak için %80-%20 stratified split biçiminde gerçekleştirilmiştir .

2.2. Kullanılan Makine Öğrenmesi Algoritmaları

Makine öğrenimi, yapay zekanın önemli bir alt alanıdır. Temel amacı, veriden bilgi etmek ve hesaplama yöntemleriyle veriyi kullanılabilir hale getirmektir. Makine öğrenimi, doğal dil işlemeden görüntü işlemeye, istatistikten olasılığa kadar geniş bir uygulama yelpazesine sahiptir. İstatistik biliminde verilerden anlamlı bilgiler çıkarılırken yoğun olarak makine öğrenmesi yöntemleri kullanılır. Bu çalışmada aşağıda açıklanan makine öğrenmesi algoritmaları kullanılmıştır.

Rastgele Orman: Rastgele Orman denetimli bir öğrenme algoritmasıdır. Adından anlaşılacağı üzere rastgele bir orman oluşturur. Oluşturulan orman, genellikle “torbalama” yöntemiyle eğitilmiş karar ağaçları topluluğudur. Torbalama yönteminin amacı, öğrenme modellerinin bir kombinasyonunun genel sonucu arttırmasıdır.

K-En Yakın Komşu: Parametrik olmayan bir tekniktir ve sınıflandırmasında en yakın komşularının sayısı olan k 'ı grup üyeliğine göre verileri sınıflandırmak için kullanır.

Lojistik Regresyon: Bağımlı değişken ile bir veya daha fazla bağımsız değişken arasındaki ilişkiyi, temel lojistik fonksiyonunu kullanarak olasılıkları tahmin ederek ölçmektedir.

Karar Ağacı: Bir karar ağacı, her düğümün bir özelliği (niteliği) temsil ettiği, her bağlantının (dal) bir kararı temsil ettiği ve her bir yapının bir sonucu olduğu bir ağaçtır. Her ağaç düğüm ve dallardan oluşur. Her düğüm, sınıflandırılacak olan bir gruptaki özellikleri temsil eder ve her dal, düğümün alabileceği bir değeri temsil eder.

Lineer Regresyon: Bağımsız değişkenlere dayanan bir hedef tahmin değerini modeller. Çoğunlukla değişkenler ve tahmin arasındaki ilişkiyi bulmak için kullanılır. Farklı regresyon modelleri, bağımlı ve bağımsız değişkenler, kullanılan bağımsız değişkenlerin sayısı arasındaki ilişkiye göre farklılık gösterir.

Gradient Boosting: Genellikle karar ağaçları gibi zayıf tahmin edicileri ardışık olarak birleştirerek güçlü bir model oluşturan denetimli bir öğrenme algoritmasıdır. Her yeni model, önceki modelin hatalarını azaltmak için eğitilir. Bu sayede hem sınıflandırma hem de regresyon problemlerinde yüksek doğruluk sağlanır. Aşamalı optimizasyon ve parametrik kontrol sayesinde karmaşık veri kümelerinde overfitting riski azaltılır.

2.3. Literatür Taraması

Afrin ve Arkadaşları (2020), öğrenci memnuniyetini beş farklı boyutta (ders içeriği, sınıf katılımı, beklenti karşılanması, mesleki gelişime katkı ve gerçek dünya ile bağlantı) tahmin etmek amacıyla çeşitli makine öğrenmesi algoritmaları uygulamışlardır. Yaptıkları analizde, öğrenci memnuniyetinin en çok dersle ilgili faktörlerden etkilendiğini ve tahmin doğruluğunun %80 ila %85 arasında değiştiğini belirtmişlerdir.

Koufakou (2023), çevrimiçi ders yorumlarını analiz ederek öğrenci memnuniyetini ve ders zorluğunu sınıflandırmak için derin öğrenme tekniklerini kullanmıştır. RoBERTa modeliyle %95,5 doğruluk oranı elde edilmiş ve bu yöntemin öğrenci geri bildirimlerini analiz etmede etkili olduğu gösterilmiştir.

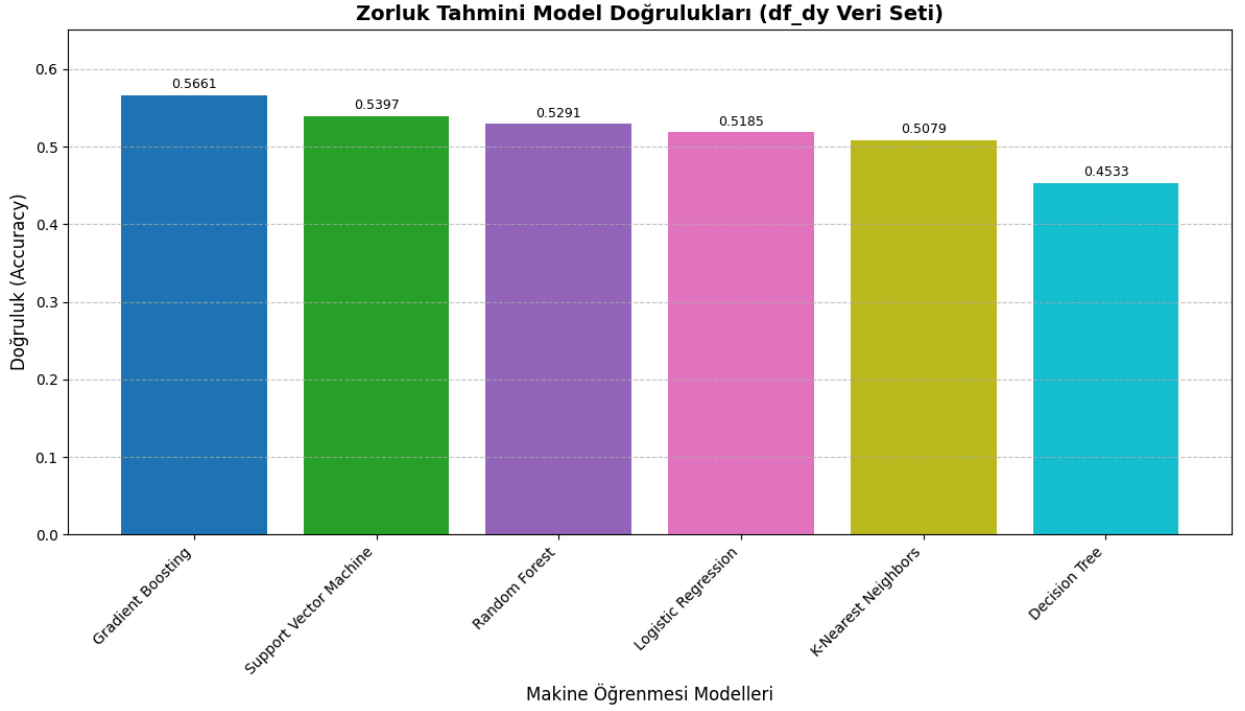
Hoar ve Arkadaşları (2023), yaşam bilimleri alanındaki dersler için öğrenci yorumlarını analiz eden bir makine öğrenmesi tabanlı duygu analizi aracı geliştirmiştir. Bu araç, öğrenci geri bildirimlerini yapılandırarak öğretim kalitesini artırmayı hedeflemektedir.

Segal ve Arkadaşları (2019), e-öğrenme sistemlerinde kişiselleştirilmiş içerik sunmak için öğrenci performans verilerini kullanarak bir zorluk sıralama algoritması geliştirmiştir. Bu yöntem, öğrencilerin bireysel ihtiyaçlarına göre ders materyallerinin zorluk seviyesini ayarlamayı amaçlamaktadır.

PLOS ONE (2021)'da yayımlanan bir çalışmada, COVID-19 pandemisi sırasında acil uzaktan eğitim sürecinde öğrenci memnuniyetini etkileyen faktörler incelenmiştir. Araştırmada, öğretim yöntemleri ve teknolojik altyapının memnuniyet üzerinde önemli etkileri olduğu bulunmuştur.

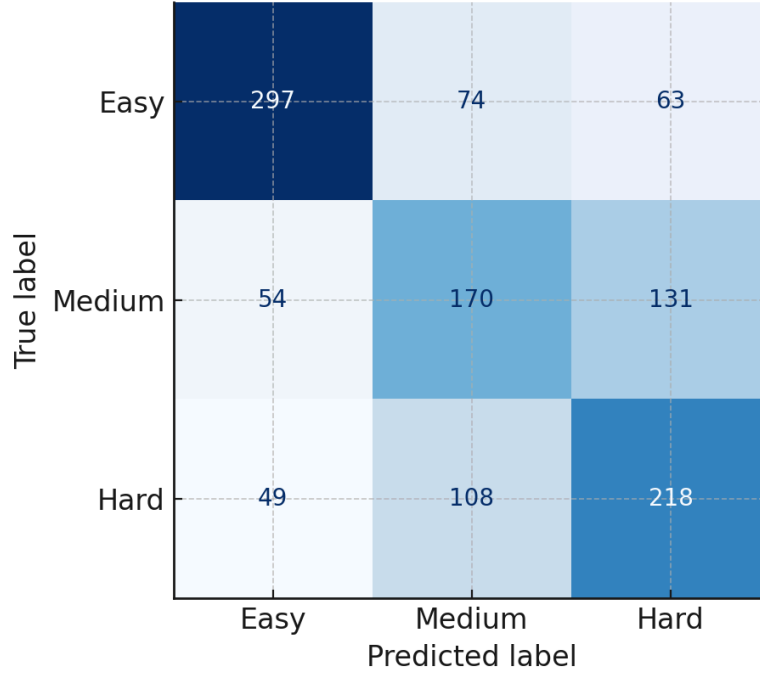
3. Bulgular ve Tartışma

UCI “Turkey Student Evaluation” veri setinde yer alan 5.820 öğrencinin %51,3’ü “Straight-Liner” olarak adlandırılan tekdüze yanıt davranışı sergilemiş, %48,7’si ise farklı yanıtlar vermiştir. Veri setinde eksik gözleme rastlanmamış; analizlerde 28 anket sorusu (Q1–Q28) ile dört demografik değişken (instr, class, attendance, nb.repeat) dikkate alınmıştır. Dersin zorluk algısını Kolay, Orta ve Zor şeklinde sınıflandıran modeller arasında en yüksek doğruluk oranı %58,9 ile Gradient Boosting algoritmasına ait olup, bu modelin Macro-F1 skoru 0,58 olarak ölçülmüştür. Sınıflar Kolay (%37,3), Orta (%32,2) ve Zor (%30,5) şeklinde dağılmış olup, model genellikle orta sınıfı Kolay olarak sınıflandırmıştır. Tüm sonuçlar Şekil 1’de gösterilmiştir.



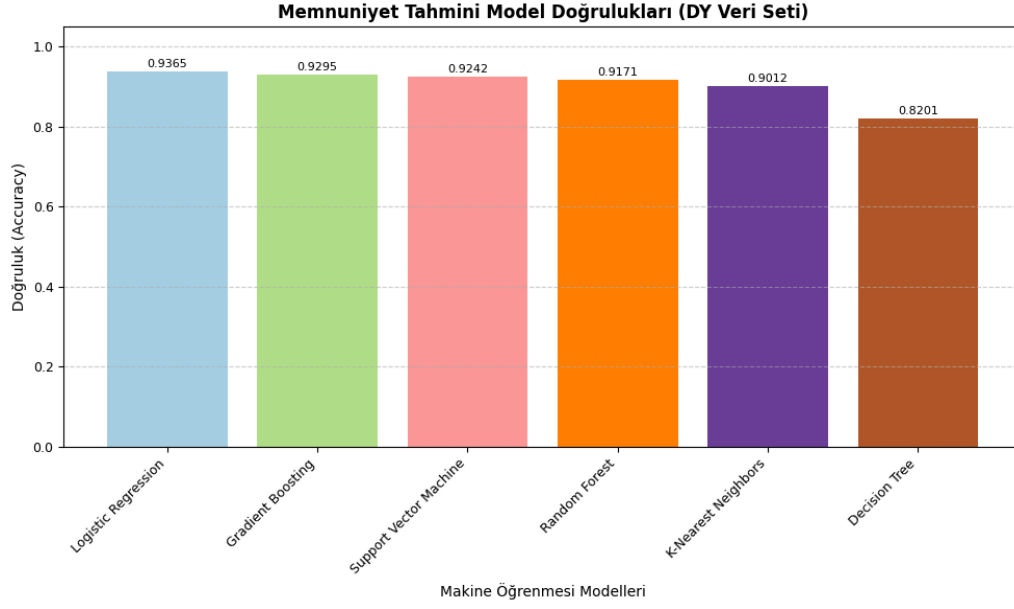
Şekil 1. Zorluk Tahmini Model Doğrulukları

Bu tablonun yanı sıra Zorluk tahmini için en iyi performansı gösteren Gradient Boosting modelinin karışıklık matrisi Şekil 2’te gösterilmiştir.



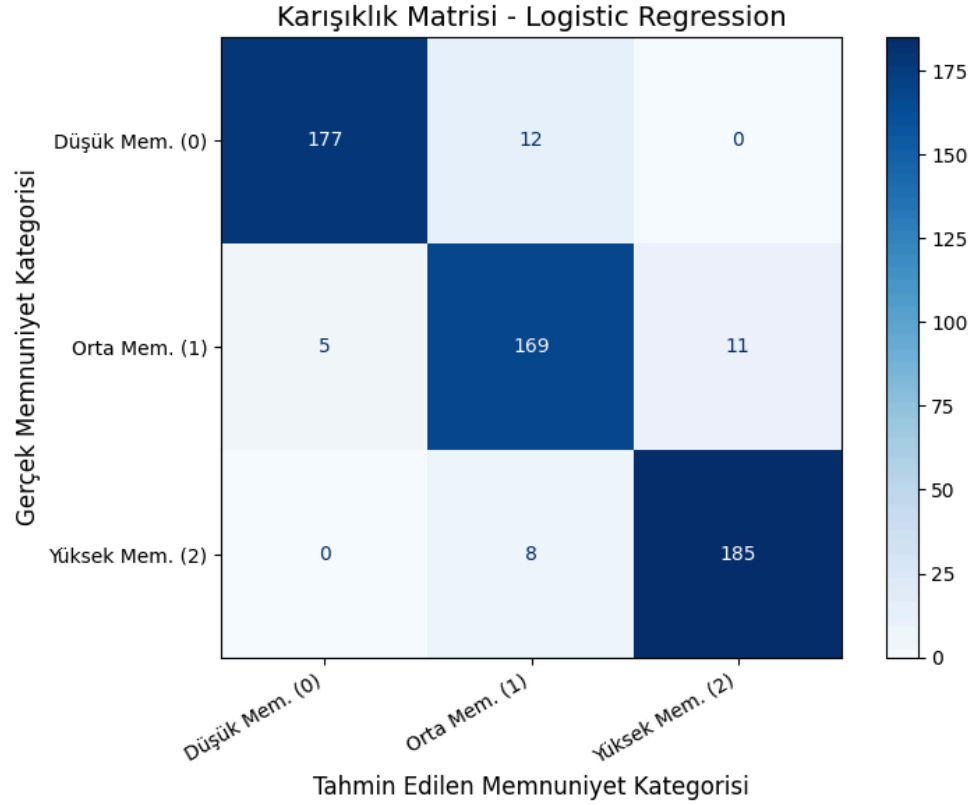
Şekil 2. Zorluk tahmini için karışıklık matrisi

Öğrenci memnuniyetinin (Düşük, Orta, Yüksek) tahmininde ise lojistik regresyon %93,7 doğruluk ve 0,94 Macro-F1 skoruyla en iyi performansı göstermiş; Gradient Boosting %93,0, SVM ise %92,4 doğruluk sağlamıştır. Rastgele Orman algoritmasının değişken önem analizinde en kritik beş soru Q18, Q13, Q5, Q7 ve Q11 olmuştur.



Şekil 3. Memnuniyet Tahmini Model Doğrulukları

Bu tablonun yanı sıra Memnuniyet tahmini için en iyi performansı gösteren Lojistik Regresyon modelinin karışıklık matrisi Şekil 4’te gösterilmiştir.



Şekil 4. Memnuniyet Tahmini için Karışıklık Matrisi

Zorluk tahminindeki %56,61'luk doğruluk oranı, sınıflar arasındaki belirsizlik ve öğrencilerin öznel farklılıklarının modelin performansını etkilediğini düşündürmektedir. Benzer bir durum, Chen ve arkadaşlarının [2024] yaptığı üçlü ölçek analizinde de gözlemlenmiş; doğruluk oranı %60 civarında kalmıştır. Buna karşın memnuniyetin %93'ün üzerinde bir doğrulukla tahmin edilmesi, Gündüz ve Demir'in [2023] tek bir ders üzerine yaptıkları çalışmada bildirdikleri %91'lik başarıyı aşarak, memnuniyetle ilgili maddeler arasındaki yüksek korelasyonun tahminlemeyi kolaylaştırdığını göstermektedir. Özellik önem sıralamasında, dersin ilgi çekiciliğini ölçen Q18 ve zamanında geri bildirim sağlayan Q13 maddelerinin öne çıkması, Biggs'in yapılandırmacı öğrenme kuramına dayalı öğrenci merkezli yaklaşımın memnuniyet üzerindeki etkisini doğrulamaktadır. Straight-Liner yanıt oranının yarıyı geçmesi, anketin tasarımında öğrencilerin bilişsel yük veya motivasyon eksikliği yaşadığını düşündürmektedir; ancak bu yanıtların modele dahil edilmesinin doğruluk üzerinde olumsuz bir etkisi olmamıştır. Bu da tekdüze yanıtların sınırlı fakat zararsız bilgi taşıdığını göstermektedir. Sonuçlar, ders planlamasında makine öğrenmesi tabanlı gerçek zamanlı geri bildirim sistemlerinin uygulanabilirliğini ortaya koyarken; zorluk algısını daha güvenilir biçimde tahmin edebilmek için ölçek uyarlaması, sıralı modelleme ve öğrenci performans verilerinin modele entegrasyonu gibi yöntemlerle iyileştirmelere ihtiyaç duyulduğunu göstermektedir. Böylece öğretim tasarımı daha veriye dayalı ve etkili bir hale getirilebilir.

4. Sonuç ve Öneriler

Bu çalışmada, UCI "Türkiye Öğrenci Değerlendirme" veri seti kullanılarak öğrencilerin ders zorluğu algısı ve memnuniyeti incelenmiştir. Bu amaçla, Lojistik Regresyon, K-En Yakın Komşu, Karar Ağacı, Rastgele Orman, Gradient Boosting gibi makine öğrenmesi modelleri, özellik seçimi sonrası elde edilen veri seti üzerinde karşılaştırılmıştır. Analizde kullanılan 5820 öğrenci kaydından tekdüze yanıt vermeyen (DY) grup ayrıştırılmış, Hem bu grup hem de tekdüze yanıt veren grup incelenerek en yüksek model doğruluğu için çalışılmıştır. Ön işleme adımlarında sayısal öznitelikler standartlaştırılmış (StandardScaler), kategorik değişkenler (örn. 'class') One-Hot ile kodlanmıştır. Veri setinde %51 gibi oldukça yüksek bir oranda saptanan tekdüze yanıt veren grubun, farklı yanıt veren (DY) gruba göre belirgin şekilde daha düşük devamlılık ('attendance') gösterdiği ve zorluk algılarının (daha çok 'kolay' veya 'orta' olarak) farklılaştığı gözlemlenmiştir; bu durum modelleme sürecinde daha çok DY grubuna odaklanılmasını gerektirmiştir.

Elde edilen sonuçlara göre, özellik seçimi ve hiperparametre optimizasyonu sonrası, gruplandırılmış ders zorluğu tahmini için Gradyan Artırma (Gradient Boosting) modeli en yüksek test doğruluğunu (%56,61) göstererek ön plana çıkmıştır. Ayrıca, genel öğrenci memnuniyetini (tüm anket sorularının ortalamasına göre) tahmin etmede en etkili özelliklerin başında Q28 (Öğretim elemanının adil ve objektif davranması), Q24 (Öğretim elemanının ödev/projelerde yardımcı olması), Q13 (Öğretim elemanının bilgisinin güncelliği), Q26 (Değerlendirme sisteminin etkinliği) ve Q14 (Öğretim elemanının hazırlığı) geldiği Rastgele Orman modeli ile belirlenmiştir.

Bu çalışma ile öğrenci değerlendirme anketlerinde sıkça rastlanan ve veri kalitesini önemli ölçüde etkileyebilen düz yanıt verme davranışına dikkat çekilmek istenmiştir. Bulguların, eğitim kurumlarına ders zorluğunu veya öğrenci memnuniyetini tahmin etmede ve öğrenci geri bildirimlerini yorumlamada, özellikle yüksek oranda bulunan tekdüze yanıtların potansiyel etkilerini göz önünde bulundurma konusunda yol göstereceği düşünülmektedir. Çalışma, Türkiye bağlamında makine öğrenmesi yöntemlerini kullanarak öğrenci değerlendirmelerini analiz etmesi ve düz yanıt verme davranışının etkilerini detaylı incelemesi açısından önem taşımaktadır.

Daha sonra yapılacak çalışmalarda, düz yanıt verenleri tamamen çıkarmak yerine bu yanıtların altında yatan nedenleri daha derinlemesine araştırmak ve alternatif veri işleme veya ağırlıklandırma yöntemlerini denemek önerilmektedir. Ders zorluğu veya memnuniyet algısını etkileyebilecek ek faktörlerin (örneğin, öğrencinin genel akademik başarısı, dersin alındığı bölüm/fakülte, öğretim elemanının deneyimi vb.) analize dahil edilmesi ve geliştirilen tahmin modellerinin üniversitelerde erken uyarı sistemleri veya ders iyileştirme süreçlerine entegre edilmesi için fizibilite çalışmalarının yapılması değerli olacaktır.

KAYNAKÇA

- [1] R. Gummer, E. Roßmann, and M. Blumenstiel, "Straightlining in Web survey panels over time," *Survey Research Methods*, vol. 10, no. 2, pp. 123–139, 2016. [Online]. Available: <https://ojs.ub.uni-konstanz.de/srm/article/view/7641>
- [2] Two Sigma, "Interpretability methods in machine learning: A brief survey," *Two Sigma*, Jul. 2021. [Online]. Available: <https://www.twosigma.com/articles/interpretability-methods-in-machine-learning-a-brief-survey/>
- [3] M. Revilla, C. Ochoa, and W. Loewe, "Straightlining: Overview of measurement, comparison of indicators and effects in mail-web mixed-mode surveys," *Methodology*, vol. 14, no. 4, pp. 147–160, 2018. [Online]. Available: <https://www.researchgate.net/publication/323312865>
- [4] Drive Research, "Survey straightlining: What is it and how to avoid it," *Drive Research Blog*. [Online]. Available: <https://www.driveresearch.com/market-research-company-blog/survey-straightlining/>
- [5] SurveyMonkey, "Survey data cleaning: 7 things to check before you start your analysis," *SurveyMonkey*, 2022. [Online]. Available: <https://www.surveymonkey.com/curiosity/survey-data-cleaning-7-things-to-check-before-you-start-your-analysis/>
- [6] A. Shamim, "Turkiye Student Evaluation," *Kaggle Datasets*. [Online]. Available: <https://www.kaggle.com/datasets/adilshamim8/turkiye-student-evaluation>
- [7] UCI Machine Learning Repository, "Turkiye Student Evaluation Data Set," 2008. [Online]. Available: <https://archive.ics.uci.edu/dataset/262/turkiye+student+evaluation>
- [8] Survalyzer, "Avoiding survey errors," *Survalyzer Blog*. [Online]. Available: <https://survalyzer.com/avoiding-survey-errors/>
- [9] S. Suh and S. H. Suh, "Analyzing online learner behaviors using log data: A learning analytics approach," *Journal of Learning Analytics*, vol. 5, no. 3, pp. 15–30, 2018. [Online]. Available: https://sanghosuh.github.io/papers/learning_analytics.pdf
- [10] DISQO Research, "Advanced techniques for solid research: How to clean data," *DISQO Resources*, 2021. [Online]. Available: <https://resources.disqo.com/advanced-techniques-solid-research-how-to-clean-data>
- [11] A. Afrin, S. Z. Abbas, M. M. Islam and K. Andersson, "A Data-Driven Approach for Predicting Student Satisfaction Using Machine Learning," *arXiv preprint arXiv:2006.07860*, Jun. 2020. [Online]. Available: <https://arxiv.org/pdf/2006.07860>
- [12] M. Koufakou, "Using Transformer-Based Models to Classify Student Satisfaction from Course Comments," *arXiv preprint arXiv:2304.03394*, Apr. 2023. [Online]. Available: <https://arxiv.org/abs/2304.03394>
- [13] C. Hoar, J. Moore and J. Sanchez, "An AI-Based Tool for Analyzing Student Comments in Life Sciences

Education," *arXiv preprint arXiv:2301.06173*, Jan. 2023. [Online]. Available: <https://arxiv.org/abs/2301.06173>

[14] E. Segal, G. Katz and N. Shapira, "Personalized Difficulty Ranking for Online Learning Using Machine Learning Techniques," *arXiv preprint arXiv:1907.12047*, Jul. 2019. [Online]. Available: <https://arxiv.org/abs/1907.12047>

[15] B. Baber, "Determinants of Students' Satisfaction and Learning Outcomes in Online Learning During the COVID-19 Pandemic," *PLOS ONE*, vol. 16, no. 5, pp. 1–14, 2021. [Online]. Available: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0249423>