

Veri Madenciliği ile Özel Öğrenme Güçlüğü'nün Tahmini*

Murat URFALIOĞLU¹

Sami ŞAHİN²

Gönderim Tarihi: 11.12.2021

Yayın Tarihi: 27.05.2022

Makale Türü: Araştırma Makalesi

Öz

Özel öğrenme güçlüğü (ÖÖG), bireyin zekâsı ve yaşı ile ilişkili olmayan gelişimsel bir problemdir. Öğrenme güçlüğü'nün hangi sebeplerle olduğu net ifade edilmemekle birlikte kalıtsal ve çevresel etkilerle ortaya çıktığı düşünülmektedir. Özel öğrenme güçlüğü olan bireylere tanı konulması tanı sonrası eğitim süreci kadar önemlidir.

Bu çalışmanın amacı özel öğrenme güçlüğü olan bireylere tanı konulmasında veri madenciliği tekniklerinden yararlanılarak uygun sınıflandırma yöntemini belirlemektir. Bu kapsamda Özel öğrenme Güçlüğü olan 392 ve özel öğrenme güçlüğü olmayan 107 öğrenci olmak üzere toplamda 499 öğrenciden eğitsel tanı formu, veli bilgi formu ve özel öğrenme güçlüğü performans belirleme formu ile veriler toplanmıştır. Bu veriler veri seti haline getirilerek gerekli hesaplamaları ve modelleri elde etmek için DecisionTable, RandomForest, NaiveBayes, JRip, DecisionStump, One R, RandomTree, REP Tree Part sınıflandırma algoritmaları; WEKA programında uygulanarak doğru sınıflama oranı, kappa istatistiği, ortalama mutlak hata değerleri, ROC eğrisi değerleri sonuçları karşılaştırılmıştır.

Uygulamanın sonucu olarak özel öğrenme güçlüğü tespitinde en iyi sonucu veren algoritma tespit edilmeye çalışılmıştır. Bu kapsamda 9 farklı sınıflandırma yönteminden doğru sınıflama oranı, kappa istatistiği ve ortalama mutlak hata değerleri bakımından en iyi sonuçlar Random Forest yöntemiyle elde edilmiştir. Kıyaslama sonucunda Random Forest algoritmasından %90.18 başarı elde edilmiştir.

Anahtar Kelimeler: Veri madenciliği, Özel Öğrenme Güçlüğü, WEKA

Prediction of Specific Learning Difficulties with Data Mining *

Abstract

Special learning disability (SLD) is a developmental problem that is not related to the intelligence and age of the individual. Although the reasons for the learning disability are not clearly stated, it is thought that it arises with hereditary and environmental effects. Diagnosing individuals with special learning disabilities is as important as the post-diagnosis education process.

The aim of this study is to determine the appropriate classification method by using data mining techniques in diagnosing individuals with special learning disabilities. In this context, data were collected from a total of 499 students, 392 students with special learning difficulties and 107 students without special learning difficulties, with educational diagnosis form, parent information form and special learning disability performance determination form. The results of correct classification rate, kappa statistics, mean absolute error values, ROC curve values were compared by

¹Sorumlu Yazar : Murat Urfalıoğlu, Doktora, Gazi Üniversitesi Bilişim Enstitüsü Yönetim Bilişim Sistemleri, Türkiye, muraturfal60@gmail.com, ORCID ID: 0000-0002-2893-8135

² Sami Şahin, Prof. Dr., Gazi Üniversitesi Eğitim Fakültesi Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü, Türkiye, sami@gazi.edu.tr, ORCID ID: 0000-0003-1873-5243

*Bu çalışma doktora tezinden üretilmiştir.

applying DecisionTable, RandomForest, NaiveBayes, JRip, DecisionStump, One R, RandomTree, REP Tree Part classification algorithms

in the WEKA program to obtain the necessary calculations and models by converting this data into a data set. As a result of the application, it was tried to determine the algorithm that gave the best result in determining the specific learning disability. In this context, the best results in terms of correct classification rate, kappa statistics and mean absolute error values from 9 different classification methods were obtained with Random Forest method. As a result of the comparison, 90.18% success was obtained from the Random Forest algorithm.

Key Words: Data Mining, Specific Learning Difficulties, WEKA

Giriş

Özel öğrenme güçlüğü; yazılı ya da sözlü olarak anlamak ve kullanabilmek için gerekli olan bilgi alma süreçlerinin birinde veya birkaçında ortaya çıkan dinleme, konuşma, okuma, yazma, heceleme, dikkat yoğunlaştırma ya da matematiksel işlemleri yapma güçlüğü olarak açıklanmaktadır (MEB, 2006).

ÖÖG bireyin yaşı, zayıf kültür ve kısıtlı eğitim imkânları ile ilişkili olmayan gelişimsel bir problemdir. Öğrenme güçlüğü nelerin etkilediği net ifade edilmemekle birlikte kalıtsal ve çevresel etkilerle ortaya çıktığı düşünülmektedir (Tucker vd., 2003). Pekel (2010) ve Sakız (2011) yaptığı araştırmaya göre ÖÖG tanısı almış çocukların üst biliş bilgisi, bilme hissi ve öğrenme kararlarının bu tanıyı almamış yaşlılarına göre daha düşük olduğu görülmüştür.

Rehberlik ve Araştırma Merkezlerinde (RAM), Özel Eğitim Hizmetleri Yönetmeliği ve Özel Eğitim Hakkında Kanun Hükmünde Kararname (573 sayılı KHK, 1997) kapsamında işlemler yürütülmektedir (Millî Eğitim Bakanlığı [MEB], 2020). Rehberlik ve Araştırma merkezlerinden özel eğitim bölümünde ÖÖG çeken çocuklara tanı konulurken birçok kişisel bilgi alınmaktadır. Toplanan bilgiler arasında eğitsel tanı formu verileri, veliden alınan veli bilgi formu ve performans belirleme formları bulunmaktadır. Yapılan testler sonucunda çocukların akademik becerileri değerlendirilmektedir (MEB, 2020).

Veri yığınları arasında gizli kalmış, bir kısım analizler sonucu geçerli ve işe yarar bilgi elde edilmesine veri madenciliği denir (Aydemir, 2017). Veri madenciliği büyük miktardaki verilerden önceden bilinmeyen ancak potansiyel olarak kullanışlı bilginin çıkarılması olup; bu da kümeleme, veri özetleme, değişikliklerin analizi, sapmaların tespiti gibi belirli sayıda teknik yaklaşımları içerir (Işıklı, 2009). Veri madenciliği; istatistiksel ve matematiksel yöntemlerin kullanılarak, anlamlı yeni korelasyonların, örüntülerin ve trendlerin büyük miktarlardaki mevcut olan verilerden analiz edilerek sorunların çözülmesi ve yeni anlamlı bilgilerin keşfedilmesi sürecidir (Frank, 2009; Group, 2013). Veri madenciliğinin asıl amacı bir veri setinden bilgiyi üretmek ve onu daha ileri kullanımlar için anlaşılır net bir veriye dönüştürmektir (Baker vd, 2012).

Veri madenciliği teknikleri, büyük veriler arasından bilinmeyen bilgiyi keşfetmek için karmaşık algoritmalar kullanarak verileri analiz etmek ve ayıklamak için kullanılır. (Hong ve Wu, 2011). Bu teknoloji birçok alanda kullanılmaktadır. Hızla gelişen teknoloji, birçok alanda hayati önem taşıyan erken tanılamaya katkı sağlamaktadır. Teknolojinin sürekli gelişmesi birçok alanda olduğu gibi sağlık alanında da teşhis ve tedavilerin daha profesyonel şekilde uygulanmasına olanak sağlamaktadır (Tsumoto vd., 2017).

Materyal ve Yöntemler

Çalışmanın bu kısmında sınıflandırma yöntemleri, verilerin elde edilmesi ve modelin oluşturulma aşamalarından bahsedilmiştir.

Sınıflandırma Yöntemleri

Veri madenciliği birçok farklı uygulaması bulunmakla beraber farklı amaçlarda da kullanılabilir. Veri madenciliği; kümeleme, veri tanımlama, ilişkilendirme, tahmin etme ve sınıflandırma gibi bazı temel uygulamalar yaygın olarak kullanıldığı görülmektedir (Larose, 2005). Bu çalışmada kullanılan veri madenciliği bir sınıflandırma uygulamasıdır.

Sınıflandırma ve kategori üyeliği makine öğrenimi ve istatistiklerinde bilinen gözlemleri (veya örnekleri) içeren bir veri seti temelinde yeni bir gözlemin kategorilerden hangisine ait olduğunu belirleme sorunudur (Rana, 2015). Veri setine uygulanan sınıflandırma yöntemleri kısaca açıklanmıştır.

Random Forest, Random Forest, ağaç tipi sınıflandırıcılar topluluğu olarak tanımlanabilir. İçerisinde oylama metodunu barındıran bir sınıflama yöntemidir. Torbala yöntemine ek olarak Rastgelelik özelliği eklenerek torbalama yöntemlerinin geliştirilmiş bir versiyonudur. Birçok karar ağacının bir araya gelmesiyle oluşmaktadır. Rastgele Orman algoritmasında kullanıcıdan ağaç sayısı ve ağaç yapısının oluşturulması için her düğümde kullanılmak üzere iki değişken istenir. Her bir karar ağacı tarafından oylanarak sınıf belirlenir. Karar ağaçları, birbirinden bağımsızdır (Breiman 2001).

Decision Table, Karar tablosu algoritmasında, bir ağaç yapısı oluşturulduktan sonra ağacın yaprakları seviyesinde sınıf etiketleri bulunmakta ve bu yapraklara giden ve kök düğümünden çıkan dallar ile özniteliklere ait işlemler gerçekleştirilmektedir (Da Silva, 2017; Chandrasekar, 2017). Her özneliğin birbirinden koşulsal bağımsız olduğu ve öğrenilmek istenen kavramın tüm bu özniteliklere koşulsal bağlı olduğu bir Bayes ağı olarak düşünülmektedir (Domingos & Pazzani, 1997; Hand & Yu, 2007).

Naive Bayes, Naive Bayes Sınıflandırıcı adını İngiliz matematikçi Thomas Bayes'ten alır. Bu yöntem niteliklerin hangi sınıfa hangi olasılıkla ait olduğunu değerlendirir ve tespit eder. Niteliklerin hepsi aynı derecede önemli ve birbirinden bağımsızdır (Lantz, 2013). Niteliklerin bağımsız olarak ele alınması daha iyi sonuç alınmasını sağlamaktadır. (Dougherty vd.,1995).

Jrip, JRip algoritması karar ağacı algoritmalarından olup çok yaygın ve etkili bir tekniktir. IREP'in optimize edilmiş bir versiyonu olarak Cohen tarafından geliştirilmiştir. JRip, eğitim verilerindeki örneklerin belirli yargılarını bir sınıf olarak ele almaktadır. Bu sınıfın tüm üyelerini kapsayan kurallar kümesini bularak aşağıdan yukarıya kuralları öğrenen bir yöntemdir (Sonawani vd., 2013). Kural tabanlı olup azaltılmış hata budama ile birliktelik kurallarına dayanır (Wikibooks, 2021).

One R, One R yöntemi, Holte (1993) tarafından geliştirilmiştir. Bu yöntem hata tabanlı kural çıkarım anlayışına dayanmaktadır. Sınıflandırmanın daha iyi sonuçlanması için iyileştirmeler yapmaktadır. Sınıflandırmada en iyi sonucu veren ve hata oranı en az olan en iyi kriteri seçmektedir. Ancak tek niteliğe odaklandığı için One R yüzeysel bir yöntem olarak görülmektedir (Uzun, 2005).

RandomTree, Random Tree, her düğümde K rasgele özneliğe sahip bir dizi olası ağaçtan rasgele oluşturulmuş bir ağaçtır. Bu bağlamda rastgele, ağaç kümesindeki her ağacın eşit örnekleme şansına sahip olduğu anlamına gelir. Rastgele ağaçlar, verimli bir şekilde oluşturulabilir ve büyük rastgele ağaç kümelerinin kombinasyonu genellikle doğru modellere yol açar. Son yıllarda makine öğrenmesi alanında Rastgele ağaçlar üzerinde kapsamlı bir araştırma yapılmıştır (Ali vd.,2012).

REP Tree, REP Tree, REP Tree'de, bölme kriteri olarak bilgi kazancı ile karar ağacı oluşturulur ve budamak için azaltılmış hata budama kullanılır. Hız için optimize edilmiştir. Yalnızca sayısal öznelikler için değerleri bir kez sıralar. Karar ağacı algoritmalarından olan C45'in yaptığı gibi eksik değerleri işlemek için kullanılır (Ali vd.,2012).

Part, Part, PART algoritması kural tabanlı sınırlayıcı olarak geçmektedir. Doğru kurallar oluşturmak için global optimizasyonu yürütmeyen, ancak ayrı ayrı uygulanan ve fethet stratejisi olan nispeten basit bir algoritmadır. Örneğin bir kural oluşturur, kapsadığı örnekleri kaldırır ve artık örnekler kalmayana kadar dinlenme örnekleri için özylenelemeli bir kural oluşturmaya devam eder (Mohamed, 2017).

Veri Seti

Çalışmada ÖÖG'li öğrencilerden RAM'larda toplanan eğitsel değerlendirme istek formu (EDİF), veli görüşme formu (VGF) ve ÖÖG performans belirleme formu (ÖÖGPBF) verileri kullanılmıştır. ÖÖG olmayan öğrencilerden de yine aynı formlar ile veri elde edilmiştir. Toplamda 260 öznelik değerlendirmeye alınmıştır. ÖÖG olan 392 ve olmayan 107 öğrenciden EDİF, VGF ve ÖÖGBF verileri alınmıştır. Bu kapsamda her bir öğrenciden 260 öznelik olmak üzere toplam 499 öğrencinin verisi Microsoft Exel ortamında toplanmıştır. Toplanan veriler WEKA 3.8.5 programına yüklenmek üzere “.Arff” dosyasına dönüştürülmüştür.

Tablo 1 Veri seti cinsiyet dağılımı

Cinsiyet	ÖÖG	Normal	Toplam
Erkek	258	59	317
Kadın	134	48	182
Toplam	392	107	499

Veri seti içerisinde özel öğrenme güçlüğü olan ve olmayan öğrencilerin cinsiyet dağılımı tablo 1'de verilmiştir.

Tablo 2 Veri seti sınıfa göre dağılımı

Sınıf	ÖÖG	Normal	Toplam
1. Sınıf	69	22	91
2. Sınıf	82	24	106
3. Sınıf	84	23	107
4. Sınıf	34	38	72
5. Sınıf	23	-	23
6. Sınıf	55	-	55
7. Sınıf	26	-	26
8. Sınıf	9	-	9
9. Sınıf	5	-	5
10. Sınıf	5	-	5
Toplam	392	107	499

Veri seti içerisinde özel öğrenme güçlüğü ve normal öğrencilerin sınıflara göre dağılımı tablo 2’de verilmiştir.

Bulgular

ÖÖG verileri öğrencinin özel öğrenme güçlüğü tahmin etmek amacıyla bir veri madenciliği programı olan WEKA 9.8.5 ile birden fazla sınıflandırma algoritması kullanılmıştır. Bu amaçla doğru sınıflama oranı, ortalama mutlak hata, kappa istatistiği ve ROC eğrisi değerleri karşılaştırılarak tablo 3’de verilmiştir.

Tablo 3 Sınıflama Algoritmaları ve Sonuçları

Algoritma	Doğru Sınıflama Oranı	Kappa İstatistiği	Ortalama Mutlak Hata Değerleri	ROC Eğrisi Altındaki Alan
RandomForest	90.18	0.74	0.18	0.84
DecisionStump	88.57	0.69	0.19	0.80
OneR	88.57	0.69	0.11	0.81
REPTree	88.37	0.69	0.19	0.80
DecisionTable	87.37	0.66	0.21	0.82
NaiveBayes	86.17	0.64	0.13	0.81
Part	81.36	0.55	0.19	0.78
RandomTree	80.36	0.53	0.19	0.77
JRip	78.55	0.00	0.33	0.48

Tablo 3 incelendiğinde doğruluk oranı en yüksek algoritmanın Random Forest olduğu görülecektir. İkinci yüksek orana ise Decision Stump algoritması olduğu görülmektedir. Ardından One R ve REP Tree algoritmaları gelmektedir. Bu bilgilerden yola çıkarak özel öğrenme güçlüğünü en iyi tahminleyen algoritmanın Random Forest algoritmasıdır.

Kappa istatistiği -1 ile 1 arasında değer alabilen beklenen ve gözlenen değerlerin karşılaştırıldığı sayısal bir değerdir. 0'dan küçük olması durumunda uyum olmadığı 1'e yaklaştıkça ise tam bir uyumun olduğunu gösterir. 0 ile 0.20 arası düşük; 0.20 ile 0.40 arası ortalamanın altı; 0.40 ile 0.60 arası ortalama; 0.60 ile 0.80 arası güçlü; 0.80 ve üzeri mükemmel olarak tanımlanmaktadır (Viera ve Garrett, 2005). Bu kapsamda Tablo 3'te sınıflandırma algoritmalarının kappa istatistiği incelendiğinde tam uyuma en yakın değer Random Forest algoritmasının aldığı görülmektedir. Buna göre RandomForest algoritmasının kappa istatistik değeri .74 değeri olarak güçlü uyum olduğunu belirtebiliriz. Güçlü olarak sınıflandırabileceğimiz diğer algoritmalar ise sırasıyla Decision Stump, One R, REP Tree, Decision Table, Naive Bayes, Part, Random Tree, JRip algoritmalarıdır.

Ortalama mutlak hata değeri, tahmin edilen veri sayısı ile gerçekte o değere sahip olan veri sayısı arasındaki farkın mutlak değerlerinin ortalamasıdır. Bu değer sıfıra yakın olması tahmin edilen ve gerçekte olması gereken verinin o kadar birbirine yakın olması anlamına gelmektedir. Bundan dolayı bu hata değerinin küçük olması daha iyi bir anlama gelmektedir. Bu kapsamda Tablo 3'e bakıldığında ortalama mutlak hata değerinin One R algoritmasında en küçük olduğu görülmektedir. One R algoritması hata tabanlı kural çıkarım anlayışına dayanmakta olup hata oranı en az olan en iyi kriteri seçmesinden dolayı bu değer en küçük çıkması beklenebilir.

ROC eğrileri sınıflandırıcının performansını hatanın türüne ve sınıf dağılımına bakmaksızın betimlemeye çalışmaktadır. ROC eğrisi altındaki alan bir parametrenin iki grubun ne kadar iyi ayırt edilebildiğinin bir ölçüsüdür (Sinha ve May, 2005). Grafiğin altındaki alan ne kadar büyük ve 1'e yaklaşıyor ise o kadar başarılı bir sınıflandırma olduğu söylenir. Tablo 3'de açıkça görüldüğü üzere ROC eğrisi altındaki alana göre sınıflandırma başarımı en yüksek olan algoritma Random Forest algoritmasıdır.

Tartışma ve Sonuçlar

Çalışma kapsamında ÖÖG tanılmasında kullanılabilecek uygun bir sınıflandırma algoritması geliştirilmesi amaçlanmıştır. Böylece bu algoritma kullanılarak geliştirilen yazılım ile öğrencileri özellikle ilkokula başladığında özel öğrenme güçlüğü olup olmadığı tahmin edilerek RAM'lara yönlendirilmesi daha sistematik olacaktır. Böylece erken tanı konularak eğitim sürecinde gerekli tedbirlerin alınması sağlanacaktır.

Veri madenciliği uygulamalarında yaygın olarak kullanılan WEKA 9.8.5 programı aracılığıyla veriler 9 farklı yöntem ile analiz edilmiş ve tahmin başarı oranları karşılaştırılmıştır. Karşılaştırma sonucu en yüksek doğruluk oranı Random Forest ile elde edilmiştir. Yine kappa istatistik sonuçlarına da bakıldığında random forest algoritmasının 0.60 ile 0.80 arasında değeri ile bu çalışmada en güçlü algoritma olmuştur.

Alan yazın incelendiğinde PİSA başarısını tahmin etmede, hastalıkların tahmin edilmesinde ve Tam Kan Sayımı Sonuçlarından Covid-19 Test Sonuçlarının Tahmin edilmesinde Random Forest algoritmasının doğruluk oranı ile en yüksek başarı elde edilmiştir (Aksu, 2018; Sıtkı, 2020; Bozkurt, 2021). Aydemir'de (2017) öğrencilerin not ortalamalarına göre yaptığı tahmin algoritmalarında,

SMO; mezuniyet yılına göre yaptığı tahminlerde ise J4.8 ve NaiveBayes algoritmalarının en yüksek doğruluk oranı verdiği sonucuna varmıştır. Başer, Hökelekli ve Âdem (2020) ise ortaöğretimde öğrenci tahmin etmede en başarılı yöntem olarak One R'yi bulmuştur. Tüm bu bilgiler ışığından ileride geliştirilecek olan ÖÖG tahmin sisteminde kullanılacak sınıflandırma algoritmasının Random Forest olması en iyi sonuç almayı sağlayacaktır.

Kaynakça

- Aksu, G. (2018). PISA başarısını tahmin etmede kullanılan veri madenciliği yöntemlerinin incelenmesi (Yayımlanmamış doktora tezi). Hacettepe Üniversitesi, Ankara.
- Ali J, Khan R, Ahmad N, Maqsood I. (2012) "Random forests and decson trees", IJCSI Internatonal Journal of Computer Science Issues, vol.9:5, pp.1-7
- Baker, R. S., Yacef, K. (2009). The State of Educational Data Mining in 2009: *A Review and Future Visions. Journal of Educational Data Mining*, 1(1), 3-17. <https://jedm.educationaldatamining.org/index.php/JEDM/article/view/8>
- Baker, R. S., Gowda, S. M., Corbett, A. T., and Ocumpaugh, J. (2012). Towards automatically detecting whether student learning is shallow. *In International Conference on Intelligent Tutoring Systems*, 444-453
- Başer, S. H., Hökelekli, O., & Kemal, A. D. E. M.(2020) Ortaöğretimde Öğrenim Gören Öğrenci Performanslarının Veri Madenciliği Yöntemleri İle Tahmin Edilmesi. *Bilgisayar Bilimleri ve Teknolojileri Dergisi*, 1(1), 22-27.
- Bozkurt, A. (2021). *Veri Madenciliği Yöntemleri ile Tam Kan Sayımı Sonuçlarından Covid-19 Test Sonuçlarının Tahmini* (Master's thesis, Necmettin Erbakan Üniversitesi Fen Bilimleri Enstitüsü).ISO 690
- Breiman L., (2001), Random forests,machine learning, 2001 Kluwer Academic Publishers, 45(1), 5-32
- Breiman,L.(2001).Random Forest[Elektronik Sürüm]. Machine Learning, 45, 5-32
- Chandrasekar, P., Qian, K., Shahriar, H., & Bhattacharya, P., 2017. Improving the prediction accuracy of decision tree mining with data preprocessing. In 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC) (Vol. 2, pp. 481-484). IEEE.
- Da Silva, I. N., Spatti, D. H., Flauzino, R. A., Liboni, L. H. B., & dos Reis Alves, S. F. 2017. Artificial neural networks. Cham: Springer International Publishing.
- Data Mining Algorithms In R/Classification/JRip, https://en.wikibooks.org/wiki/Data_Mining_Algorithms_In_R/Classification/JRip, Son Erişim: 3 Eylül 2021
- Ecemiş, O., & Irmak, S. (2018). Paslanmaz Çelik Sektörü Satış Tahmininde Veri Madenciliği Yöntemlerinin Karşılaştırılması. *Kilis 7 Aralık Üniversitesi Sosyal Bilimler Dergisi*, 8(15), 148-169.
- Frank, E. (2009), *Data Mining Know It All*. USA: Morgan Kaufmann Publishers.
- Hong, T. P., & Wu, C. W. (2011). Mining Rules From an Incomplete Dataset with a High Missing Rate. *Expert Systems with Applications*, 38(4), 3931-3936.
- Işıklı, B., <http://www.burakişikli.wordpress.com> (Erişim Tarihi: 19.08.2013)
- J. Dougherty, R. Kohavi, and M. Sahami. Supervised and unsupervised discretization of continuous features. In A. Prieditis and S. Russel, editors, *Machine Learning: Proceedings of the 12th International Conference*, San Francisco, CA, 1995
- Lantz, B. (2019). *Machine learning with R: expert techniques for predictive modeling*. Packt publishing ltd.
- Larose, D. T.. (2005). *Discovering Knowledge in Data. An Introduction to Data Mining*. Hoboken, New Jersey: John Wiley & Sons, Inc.
- MEB, (2016). Millî Eğitim Bakanlığı Bilim ve Sanat Merkezleri Yönergesi. *Tebliğler Dergisi*, 79:2710.

- Meenakshi, M. and Geetika, G. (2014). Survey on classification methods using WEKA. *International Journal of Computer Applications*, 86 (18), 16–19.
- Milli Eğitim Bakanlığı [Ministry of National Education]. (2017). Özel eğitim ve rehberlik hizmetleri [Special education and guidance services]. <http://adana.meb.gov.tr/www/ozel-egitim-ve-rehberlik-hizmetleri/icerik/1415>
- Milli Eğitim Bakanlığı [Ministry of National Education]. (2020). Rehberlik ve Araştırma Merkezi Yönergesi [Guidance and Research Center Directive]. http://orgm.meb.gov.tr/meb_iys_dosyalar/2020_08/31230315_RAM_YONERGESY.pdf
- Mohamed W. N. H. W., Salleh M. N. M., and Omar A. H. A comparative study of reduced error pruning method and decision tree algorithms. In: *Control System, Computing and Engineering (ICCSCE), 2012 IEEE International Conference on. IEEE; 2012 Nov 23-25; Penang (Malaysia): IEEE; 2012; p.392-397.*
- Pekel D. Özel öğrenme güçlüğü olan ve olmayan çocukların üst bilişsel özelliklerinin karşılaştırılması, 2010; Yüksek Lisans Tezi. İstanbul Üniversitesi, İstanbul.
- Rana, D. (2015). One class SVM vs SVM classification, *International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064.*
- Sakız H.(2011). Öğrenme güçlüğü olan çocuklarda algılanan yaşam kalitesi, 2011; Yüksek Lisans Tezi. Boğaziçi Üniversitesi, İstanbul.
- Sıtkı, Y. H. (2020). Tıp Bilişiminde Veri Madenciliği Yöntemleri Kullanılarak Hastalıkların Tahmin Edilmesi. Yüksek Lisans Tezi. Hacettepe Üniversitesi, Ankara.
- Sinha, A. P., & May, J. H. (2005). Evaluating and tuning predictive data mining models using receiver operating characteristic curves. *Journal of Management Information Systems*, 21(3), 249-280.
- Sonawani, S., Mukhopadhyay, D. (2013) "A Decision Tree Approach to Classify Web Services using Quality Parameters"
- Tanha, J., van Someren, M., & Afsarmanesh, H. 2017. Semi-supervised self-training for decision tree classifiers. *International Journal of Machine Learning and Cybernetics*, 8(1), 355-370
- Tsumoto, S., Kimura, T., Iwata, H., & Hirano, S. (2017). Mining Text For Disease Diagnosis. *Procedia Computer Science*, 122, 1133-1140.
- Tucker J, Stevens LJ, Ysseldyke JE. Learning disabilities: The experts speak out. *J Learn Disabil* 1983; 16: 6-14.
2. Dyson LL. Children with learning disabilities within the family context: A Comparison with siblings in global self-concept, academic self-perception, and social competence. *Learn Disabil Res Pract* 2003; 18(1): 1- 9.
- Uzun, Y. (2005). Tıbbi veriler üzerinde makine öğrenme algoritmaları ve bulanık mantık ile kurallar öğrenme. (Yayımlanmamış yüksek lisans tezi).
- Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Fam med*, 37(5), 360-363.
- Yang Y., Pedersen J.O., A comparative study on feature selection in text categorization, *Proc. of the Fourteenth International Conference on Machine Learning, San Francisco CA - USA, 412-420, 1997.*