



Kullanılabilirlik Sezgiselleri ile Problemlerinin İlişkilendirilmesi: Makine Öğrenmesi Kullanımı

Aycan Pekpazar^{1*}, Çiğdem Altın Gümüşsoy²

¹ İstanbul Teknik Üniversitesi, İşletme Fakültesi, Endüstri Mühendisliği Bölümü, İstanbul, Türkiye (ORCID: 0000-0001-9329-6936)

² İstanbul Teknik Üniversitesi, İşletme Fakültesi, Endüstri Mühendisliği Bölümü, İstanbul, Türkiye (ORCID: 0000-0003-2925-0954)

(3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications June 11-13, 2021)

(DOI: 10.31590/ejosat.946150)

ATIF/REFERENCE: Pekpazar, A. & Gümüşsoy Altın, Ç. (2021). Kullanılabilirlik Sezgiselleri ile Problemlerinin İlişkilendirilmesi: Makine Öğrenmesi Kullanımı. *Avrupa Bilim ve Teknoloji Dergisi*, (26), 47-52.

Öz

Kullanılabilirlik, müşteri memnuniyeti ve marka sadakati üzerinde önemli etkisi olan ve kullanıcı arayüzlerinin ne kadar kullanıcı dostu olduğunu gösteren önemli bir faktördür. Bu sebeple arayüzlerin ve sistemlerin tasarım, geliştirme ve denetimleri aşamasında kullanılabilirlik problemlerinin tespit edilebilmesi için pek çok yöntem geliştirilmiştir. Kullanılabilirlik değerlendirme yöntemlerinden bir tanesi de sezgisel değerlendirme yöntemidir. Uzmanlar tarafından gerçekleştirilen sezgisel değerlendirme, genel kullanılabilirlik prensipleri olarak tanımlanan sezgisellere dayalı bir yöntemdir. Sezgisel değerlendirme süreçlerinde kullanılacak sistemlere özgü sezgisellerin geliştirilmesi ise uzman görüşlerine dayalı uzun ve zorlu bir süreçtir. Makine öğrenmesi ve yapay zekâ teknolojileri pek çok alanda olduğu gibi kullanılabilirlik alanında da yeni sezgisellerin geliştirilmesi ile ilgili süreçlerin otomasyonu konusunda kullanılabilir. Bu çalışmanın amacı kullanılabilirlik problemlerine dayalı olarak yeni sezgisel geliştirme süreçlerini etkinleştirmek için veri madenciliği ve makine öğrenmesi tekniklerinin kullanılmasıdır. Bu amaçla Türkiye'nin önde gelen dijital platformlarından birisi olan Digtürk'ten TV ve set üstü cihaz arayüzünün yazılımcılar tarafından değerlendirmesi sonucunda elde edilen 3695 kayıt temin edilmiştir ve kayıtlar incelenerek toplamda 2752 kullanılabilirlik problemi belirlenmiştir. Elde edilen kullanılabilirlik problemleri literatürde yaygın bir şekilde kullanılan Nielsen'in on sezgiseli ile eşleştirilmiştir. Çalışma kapsamında öncelikle kullanılabilirlik problemlerinin kullanılabilirlik sezgiselleri açısından belirli örüntülere sahip olup olmadığı ilişkilendirme kuralları tekniği ile araştırılmıştır. Ayrıca kullanılabilirlik problemlerinin sezgisellerle eşleştirilmeleri çeşitli makine öğrenmesi algoritmaları (naive bayes, lojistik regresyon, hızlı geniş marjın, derin öğrenme, rastgele orman, gradyan arttırma ağaçları, destek vektör makineleri teknikleri) yardımıyla tahmin edilmiştir. Sınıflandırıcıların validasyonu için tekrarlı holdout tekniği kullanılmıştır. Veri seti farklı eğitim/test oranlarına (50:50, 55:45, 60:40, 65:35, 70:30, 75:25, 80:20, 85:15, 90:10, 95:5) bölünmüş ve modellerin performansları doğruluk oranı ve F1-skor metrikleri kullanılarak karşılaştırılmıştır. Çalışma sonucunda sınıflandırma algoritmalarının doğruluk oranları %90'ın üzerinde, F1-skor değerleri de genel olarak %75 değerinin üzerinde gerçekleşmiştir. Sınıflandırma algoritmaları arasında gradyan arttırma ağaçlarının diğer algoritmalara göre daha iyi performans sergilediği gözlemlenmiştir.

Anahtar Kelimeler: TV, Set üstü cihaz, Kullanılabilirlik sezgiselleri, Makine öğrenmesi, Veri madenciliği.

Association between Usability Heuristics and Problems: Use of Machine Learning

Abstract

Usability is an important factor showing that how user-friendly user interfaces are and it has a significant impact on customer satisfaction and brand loyalty. For this reason, many methods have been developed to identify usability problems during the design, development,

* Sorumlu Yazar: İstanbul Teknik Üniversitesi, İşletme Fakültesi, Endüstri Mühendisliği Bölümü, İstanbul, Türkiye, ORCID: 0000-0001-9329-6936, kayaayca@itu.edu.tr

and evaluation of user interfaces and systems. One of these usability evaluation methods is the heuristic evaluation method. Heuristic evaluation performed by experts is a method based on heuristics defined as general usability principles. The development of system-specific heuristics to be used in heuristic evaluation processes is a long and challenging process based on expert opinions. Machine learning and artificial intelligence technologies can be used in usability evaluation to automate the processes related to the development of new heuristics as used in many areas. This study aims to use data mining and machine learning techniques to make new heuristic development processes based on usability problems efficiently. Therefore, 3695 problems of a TV and set-top box interface determined by the software developers were obtained from Digiturk, which is one of Turkey's leading digital platforms. By examining the problems obtained, in total 2752 usability problems were determined. The usability problems were mapped with Nielsen's ten heuristics, which are widely used in the literature. Firstly, whether the usability problems have certain patterns in terms of usability heuristics was investigated with the association rules technique. Furthermore, the mappings of usability problems with heuristics were predicted using various machine learning algorithms (naive bayes, logistic regression, fast large margin, deep learning, random forest, gradient boosted trees, support vector machines techniques). The repeated holdout technique was used for the validation of classifiers. The data set was split into different training / test ratios (50:50, 55:45, 60:40, 65:35, 70:30, 75:25, 80:20, 85:15, 90:10, 95:5) and the performance of the models were compared using accuracy rate and F1-score metrics. As a result of the study, the accuracy rates of the classification algorithms were above 90%, and the F1-score values were generally above 75%. Among the classification algorithms, gradient boosted trees generally perform better than the other algorithms.

Keywords: TV, Set-top box, Usability heuristics, Machine learning, Data mining.

1. Giriş

Kullanılabilirlik kullanıcı arayüzlerinin ne kadar kolay kullanılabilirliğini gösteren bir kalite özelliğidir (Nielsen, 1993). ISO tarafından ise kullanılabilirlik; "Belirli kullanıcıların belirli ortamlarda belirli görevleri etkili, verimli ve memnuniyet içerisinde gerçekleştirme seviyesi" olarak tanımlanmıştır (ISO,1998). Nielsen ve Molich (1990) tarafından geliştirilen sezgisel değerlendirme, en yaygın kullanılan kullanılabilirlik inceleme yöntemlerinden birisidir (Quinones ve Rusu, 2017). Sezgisel değerlendirme yönteminde uzmanlar "genel kullanılabilirlik ilkeleri" olarak tanımlanan kullanılabilirlik sezgisellerini kullanarak kullanıcı arayüzlerini inceler ve kullanılabilirlik problemlerini tespit ederler (Quinones ve Rusu, 2017; Nielsen, 1994). Yeni kullanılabilirlik sezgisellerinin geliştirilmesi için literatürde önerilen farklı formal metodolojiler mevcuttur (Hub ve Capkova, 2010; Rusu vd., 2011; Hermawati ve Lawson, 2015; Jimenez vd., 2017; Quinones vd., 2018; Kaya vd., 2021). Yeni kullanılabilirlik sezgisellerinin geliştirilmesi için kullanılabilirlik problemlerinin incelenmesi ve mevcut sezgisellerle eşleştirilerek mevcut sezgisellerin yeni sistemlere özgü hale getirilmesi ve eşleştirilemeyen problemlerin de yeni sezgisel geliştirilmesi için kullanılması literatürde kullanılan tekniklerden birisidir (Hermawati ve Lawson, 2015; Kaya vd., 2021). Ayrıca kullanılabilirlik problemlerinin en çok hangi sezgiselle ilişkili olduğunu bulmak yazılımcılara ve proje yöneticilerine tasarımlarında en çok hangi hususlara dikkat etmeleri gerektiği konusunda yol gösterir. Ancak kullanılabilirlik problemleri ile sezgisellerin eşleştirilmesi kullanılabilirlik uzmanları tarafından yapılmaktadır ve belirlenen kullanılabilirlik problemi sayısı arttıkça sezgisellerle problemleri eşleştirmek uzun sürmektedir. Bu sebeple bu çalışma kapsamında sezgisel geliştirme sürecinin etkin hale getirilmesi için kullanılabilirlik problemleri ile kullanılabilirlik sezgiselleri arasındaki ilişkilerin tahmin edilmesi için veri madenciliği ve makine öğrenmesi teknikleri kullanılmıştır.

Kullanılabilirlik değerlendirmesi uzun zaman alan ve çok fazla insan gücü kullanımı gerektiren pahalı bir süreç olmasına rağmen (Ivory ve Hearst, 2001; Boza vd., 2014; El-Halees, 2014), bu sürecin otomasyonu ile ilgili literatürde sınırlı sayıda çalışmalar yapılmıştır., González vd. (2006), González vd. (2008), Sagar ve Saha (2016), Etemadi vd. (2018) ilişkilendirme

kuralları ve karar ağaçları gibi teknikleri kullanarak kullanılabilirlik problemlerinin sahip olduğu örüntüleri belirlemeye çalışmışlardır. El-Halees (2014) kullanıcıların yazılımlar hakkında etkinlik, etkililik ve memnuniyet faktörleri doğrultusunda yazmış oldukları görüşlerini maksimum entropi ve kNN (en yakın k komşu) sınıflandırıcılarını kullanarak negatif ve pozitif olmak üzere sınıflandırılmıştır. Wu vd. (2014) ürünlerle ilgili çevrimiçi müşteri yorumlarının hangilerinin kullanılabilirlikle ilişkili olduğunu belirlemek üzere apriori algoritmasını kullanarak ilişkilendirme kurallarını elde edilmiş ve müşteri yorumlarını negatif ve pozitif olarak sınıflandırmıştır.

Oztekin vd. (2013) e-öğrenme sistemlerinin kullanılabilirliklerinin değerlendirilmesi için makine öğrenmesi temelli bir değerlendirme sistemi geliştirmişlerdir. Destek vektör makineleri, yapay sinir ağları ve karar ağaçları teknikleri ile birlikte çoklu lineer regresyon yöntemlerini kullanarak genel sistem kullanılabilirliğini etkileyen faktörleri belirlemişlerdir. Boza vd. (2014) web sitelerinin beş kullanılabilirlik faktörü (öğrenilebilirlik, etkililik, hatırlanabilirlik, hata önleme, memnuniyet) açısından değerlendirilmeleri ile ilgili kuralları ilişkilendirme kuralları ve karar ağaçları tekniklerini birlikte kullanarak belirlemişlerdir. Chamba-Eras vd. (2017) üniversite web sitelerini kNN algoritması ve Öklidyen uzaklığı değerlerini kullanarak kullanılabilirlik düzeyleri açısından kümelere ayırmıştır.

Bu çalışma kapsamında ise kullanılabilirlik sezgiselleri geliştirme sürecinin iyileştirilmesi amacıyla veri madenciliği ve makine öğrenmesi teknikleri kullanılmıştır. Bu amaçla Türkiye'nin önde gelen dijital platformlarından birisi olan Digiturk Firması'ndan elde edilen TV ve set üstü cihaz arayüzü ile ilgili kullanılabilirlik problemleri literatürde yaygın olarak kullanılan ve Nielsen tarafından geliştirilen 10 sezgiselle eşleştirilmiştir (Nielsen, 1995). İlk aşamada kullanılabilirlik problemlerinin sezgisellerle eşleştirilmesi açısından sahip olduğu örüntüyü belirlemek üzere ilişkilendirme kuralları tekniği kullanılmıştır. İkinci aşamada ise naive bayes, logistik regresyon, hızlı geniş marjın, derin öğrenme, rastgele orman, gradyan artırma ağaçları, destek vektör makineleri teknikleri kullanarak kullanılabilirlik problemleri sezgisellerle eşleştirilmelerine göre sınıflandırılmıştır. Literatürdeki çalışmalardan farklı olarak bu çalışma kullanılabilirlik problemlerinin kullanılabilirlik sezgiselleri ile ilişkilerini veri madenciliği ve makine öğrenmesi teknikleri ile inceleyen ilk çalışmadır.

Çalışmanın ikinci bölümünde çalışmanın metodolojisi, üçüncü bölümde ise yapılan uygulama çalışmasından bahsedilmiştir. Son bölümde ise çalışmanın sonuçları, kısıtları ve gelecek çalışma önerileri tartışılmıştır.

2. Materyal ve Metot

Bu çalışmada üç aşamalı bir metodoloji uygulanmıştır: Birinci adımda TV ve set üstü cihaz arayüzleri ile ilgili toplanan kullanılabilirlik problemleri Nielsen (1995) tarafından geliştirilen kullanılabilirlik sezgiselleri ile eşleştirilmiştir. İkinci adımda ise kullanılabilirlik problemlerinin sezgisellerle aralarında bulunan ilişkileri ve örüntüleri belirlemek üzere öncelikle ilişkilendirme kuralları kullanılmıştır. Üçüncü adımda ise kullanılabilirlik problemlerinin sezgisellerle olan ilişkileri sınıflandırma algoritmaları kullanılarak tahmin edilmiştir. Analizler için RapidMiner Studio 9.9 programı kullanılmıştır.

2.1. Kullanılabilirlik Sezgiselleri

Sezgisel değerlendirme, genel kullanılabilirlik ilkeleri olarak tanımlanan kullanılabilirlik sezgiselleri kullanılarak uzmanlar tarafından yürütülen bir kullanılabilirlik değerlendirme sistemidir. Nielsen (1995) tarafından geliştirilen on kullanılabilirlik sezgiseli

literatürde sezgisel değerlendirme ve yeni sezgisel geliştirilmesinde en sık kullanılan sezgisellerdir (Tablo 1). Bu çalışma kapsamında kullanılabilirlik problemleri Tablo 1’de açıklamaları verilen sezgiseller ile eşleştirilmiştir.

2.2.İlişkilendirme Kuralları

İlişkilendirme kuralları veri etiketlemeye gerek duyulmadan veriler arasındaki ilişkileri ve veri kümesindeki örüntüleri eğer/sonra şeklinde kurallar olarak belirlemeye yarayan gözetimsiz bir veri madenciliği tekniğidir (Sagar ve Saha, 2016; Dökeroglu vd.,2018). İlişkilendirme kurallarının belirlenmesinde destek ve güven olmak üzere iki temel parametre kullanılmaktadır. ($X \Rightarrow Y$) kuralının destek değeri hem X hem Y’yi (XUY) içeren işlemler yüzdesi olarak tanımlanmaktadır. Diğer bir deyişle destek değeri, X ve Y nesnelerinin birlikte bulunma olasılıklarını göstermektedir. Bir kuralın güven değeri ise (XUY)’yi içeren işlemlerin sayısının X’i içeren işlemler sayısına oranıdır (Srikant ve Agrawal, 1995). Güven değeri tespit edilen ilişkinin kesinlik derecesini göstermektedir ve koşullu olasılık olarak tanımlanabilir (Han vd., 2011). Bu çalışma kapsamında ilişkilendirme kurallarının üretilmesinde FP-growth algoritması kullanılmıştır. Ayrıca çalışma kapsamında minimum destek değeri 0,05; minimum güven değeri ise 0,75 olarak alınmıştır.

Tablo 1. Kullanılabilirlik sezgiselleri ve açıklamaları (Nielsen, 1995)

Kullanılabilirlik sezgiselleri	Açıklamalar
H1: Sistem durumunun görünürlüğü	Tasarım, makul bir süre içinde uygun geri bildirimler yoluyla kullanıcıları her zaman neler olup bittiğinden haberdar etmelidir.
H2: Sistem ile gerçek dünya arasındaki uyum	Tasarım, kullanıcının dilini konuşmalıdır. Teknik jargon yerine kullanıcının aşına olduğu kelimeleri, kelime öbeklerini ve kavramları kullanmalıdır. Tasarım, gerçek dünya kurallarını takip etmeli, bilgileri doğal ve mantıksal bir sırayla sunmalıdır.
H3: Kullanıcı kontrolü ve özgürlük	Kullanıcılar çok sık yanlışlıkla eylemleri gerçekleştirdikleri için uzun bir süreçten geçmek zorunda kalmadan istenmeyen eylemi terk etmek için açıkça işaretlenmiş bir "acil çıkış" a ihtiyaçları vardır.
H4: Tutarlılık ve standartlar	Kullanıcılar, farklı kelime, durum veya eylemlerin aynı anlama gelip gelmediğini merak etmemelidir. Tasarım, şirket ve sektör standartlarını takip etmelidir.
H5: Hataları önleme	İyi hata mesajları önemlidir, ancak en iyi tasarımlar sorunların ilk etapta oluşmasını dikkatlice önlemelidir. Tasarımda hataya açık durumlar ortadan kaldırmalı veya kontrol edilmelidir. Eylemi gerçekleştirmeden önce kullanıcılara bir onay seçeneği sunulmalıdır.
H6: Hatırlamak yerine tanıma	Öğeler, eylemler ve seçenekler görünür hale getirilerek kullanıcının bellek yükü en aza indirilmelidir. Kullanıcı, arayüzün bir bölümünden diğerine geçerken bilgileri hatırlamak zorunda kalmamalıdır. Tasarımı kullanmak için gerekli bilgiler (örn. Alan etiketleri veya menü öğeleri) görünür olmalı veya gerektiğinde kolayca erişilebilir olmalıdır.
H7: Esneklik ve kullanım verimliliği	Acemi kullanıcılardan gizlenen kısayollar, uzman kullanıcıların etkileşimlerini hızlandırabilir, böylece tasarımın hem deneyimsiz hem de deneyimli kullanıcılara uygun hale getirilmesi sağlanabilir. Tasarım, kullanıcıların sık eylemleri kişileştirmesine izin vermemelidir.
H8: Estetik ve minimalist tasarım	Arayüzler ilgisiz veya nadiren ihtiyaç duyulan bilgileri içermemelidir. Bir arayüzdeki her ekstra bilgi, ilgili bilgi birimleri rekabet eder ve görelî görünürlüğünü azaltır.
H9: Kullanıcıların hataları tanınmasına ve düzeltilmesine yardımcı olma	Hata mesajları sade bir dille (hata kodları olmadan) ifade edilmeli, sorunu tam olarak belirtmeli ve yapıcı bir çözüm önermelidir.
H10: Yardım ve dokümantasyon	Sistemin herhangi bir ek açıklamaya ihtiyaç duymaması en iyisidir. Ancak, kullanıcıların görevlerini nasıl tamamlayacaklarını anlamalarına yardımcı olacak belgeler sağlamak gerekli olabilir.

2.3.Sınıflandırma

Sınıflandırma, veri sınıflarını tanımlayan modelleri çıkarmak için kullanılan gözetimli bir veri analizi tekniğidir (Han vd., 2011). Sınıflandırma analizi, öğrenme ve sınıflandırma aşamalarından oluşur. Sınıflandırma modelleri, öğrenme aşamasında eğitim verisinde bulunan etiketleri belli olan verileri analiz eder ve verilerdeki özelliklere dayanarak sınıflara ilişkin modeller üretir ve sınıflandırma aşamasında bu modellere göre test verisindeki verilerin sınıf etiketlerini tahmin eder (Han vd., 2011). Bu çalışma kapsamında naive

bayes, genelleştirilmiş lineer model, derin öğrenme, lojistik regresyon, hızlı geniş marjın, rastgele orman, gradyan artırma ağaçları ve destek vektör makinesi yöntemleri kullanılarak sınıflandırma modelleri oluşturulmuş ve sınıf etiketleri tahmin edilmiştir. Modellerin validasyonu için tekrarlı holdout yöntemi kullanılmıştır ve veri seti farklı oranlarda eğitim/test verisi olarak ayrılmıştır. Veri seti 50:50, 55:45, 60:40, 65:35, 70:30, 75:25, 80:20, 85:15, 90:10 ve 95:5 olmak üzere on eğitim/test parçasına ayrılarak değerlendirilmiştir. Tekrarlı holdout yönteminde örneğin veri seti 80:20 oranında ayrıldıysa, model

verinin %80'i ile eğitildikten sonra kalan %20'lik test verisi genel veri setinin özelliklerini taşıyan 7 farklı alt kümeye ayrılmıştır. Model her bir alt küme üzerinde tahminleri yaptıktan sonra 7 farklı alt kümeye ait performansların ortalamaları alınarak modelin genel performansı ve standart sapması hesaplanmıştır. Sınıflandırma algoritmalarının performanslarının değerlendirilmesinde doğruluk oranı ve F1-Skor parametreleri kullanılmıştır. Doğruluk oranı, algoritma tarafından doğru sınıflandırılan test örneklerinin yüzdesi olarak tanımlanmaktadır (Kılınç vd., 2016). F1-Skor ise duyarlılık (precision) ve hassasiyet (recall) değerlerinin harmonik ortalamasıdır ve bu iki parametrenin etkisini birleştirmek için kullanılır.

3. Araştırma Sonuçları ve Tartışma

3.1. Veri

Türkiye'nin önde gelen dijital TV platformu olan Digiturk kendi TV ve set üstü cihaz arayüzlerini tasarlamaktadır. Çalışma kapsamında Digiturk'te geliştirilen arayüzlerle ilgili yazılımcıların belirlediği 3695 problem Digiturk'ten temin edilmiştir. Problemlerden kullanılabilirlikle ilgili olmayanlar ve benzer anlama gelecek şekilde yazılan tekrar eden problemler

veri setinden çıkarıldıktan sonra toplamda 2752 problem kalmıştır.

3.2. İlişkilendirilme Kuralları

Öncelikle metin verisi sözcüklere ayırma (tokenization), karakter dönüştürme, etkisiz sözcükleri filtreleme, uzunluğa göre sözcükleri filtreleme ve kök bulma ön işleme süreçlerinden geçirilmiştir. Daha sonra her sözcüğün bir kayıta bulunup bulunmaması durumunu gösteren 0-1 değerlerinden oluşan vektör dokümanı elde edilmiştir. Sonraki adımda FP-Growth algoritması kullanılarak her bir sözcüğün destek değerleri hesaplanmış ve ilişkilendirme kuralları elde edilmiştir.

İlişkilendirme modeli her bir sezgisel için ayrı ayrı çalıştırılmıştır ve problemler için ilişkilendirilme kuralları Tablo 2'deki gibi elde edilmiştir. Tablo 2'de gördüğümüz kuralları yorumlayacak olursak örneğin H1 sezgiseli için (enduser ⇒ display) ilişki kuralının destek değeri %2,7, güven değeri ise %100 olarak bulunmuştur. Diğer bir deyişle, veri setinin %2,7'sinde "enduser" ve "display" kelimeleri birlikte kullanılmıştır ve "enduser" kökünden türetilen bir kelimenin bulunduğu kullanılabilirlik problemlerinin tümünde "display" kökünden türetilmiş en az bir kelime bulunmaktadır.

Tablo 2. İlişkilendirme kuralları

Sezgisel No	Öncül	Sonuç	Destek	Güven
H1: Sistem görünürlüğü	enduser	display	0,027	1
H2: Sistem ile gerçek dünya arasındaki uyum	display, language	Turkish	0,046	1
	display, menu, Turkish	language	0,037	1
	display, menu, language	Turkish	0,037	1
	week	this	0,033	1
	week	Digiturk	0,033	1
H3: Kullanıcı kontrolü ve özgürlük	key, display	press	0,046	1
	key, live	press	0,046	1
	key, stb, return	press	0,036	1
H4: Tutarlılık ve standartlar	enduser	display	0,056	1
	now on Digiturk	display	0,042	1
	available	display	0,038	1
	rec	display	0,035	1
	time	display	0,035	1
H5: Hataları önleme	return	stb	0,111	1
	update	message	0,063	1
	press, return	stb	0,055	1
	text, update	message	0,052	1
	stb, info	return	0,052	1
H6: Hatırlamak yerine tanıma	do	exist	0,067	1
	enduser	display	0,057	1
	option, do	exist	0,043	1
	now on Digiturk	display	0,039	1
	event, icon	display	0,035	1
H7: Esneklik ve kullanım verimliliği	work, ok	key	0,103	1
	focus	key	0,093	1
	work, menu	key	0,082	1
	exist	play	0,072	1
	option	play	0,072	1
H8: Estetik ve minimalist tasarım	text	overlap	0,078	1
	icon	overlap	0,059	1
	overlap, problem	seen	0,059	1
	overlap, seen	problem	0,059	1
H9: Kullanıcıların hataları tanınmasına ve düzeltmesine yardımcı olma	message	error	0,416	1
	code	error	0,299	1
	display, message	error	0,260	1
	display, code	error	0,221	1
	text	error	0,156	1
H10: Yardım ve dokümantasyon	record	info	0,185	0,97
	key	info	0,115	0,97
	live	info	0,108	0,97
	stb	info	0,108	0,97
	press, key	info	0,083	0,96

3.3.Sınıflandırma

Bir kullanılabilirlik sezgiseli birden fazla problemle eşleşebildiği gibi bir problem de birden fazla sezgiselle ilişkili olabilir. Bu sebeple sınıflandırma modelleri her bir sezgisel için ayrı ayrı çalıştırılmıştır. Ayrıca sınıflandırıcı performanslarını kıyaslamak için veri seti on farklı eğitim/test parçasına ayrılarak çalıştırılmış ve doğruluk ve F1-skor değerleri hesaplanmıştır. Tablo 3'te her bir sezgisel için her bir sınıflandırıcının ulaştığı en iyi doğruluk oranları ve F1-skorları sunulmuştur. Her bir sezgisel için sonuçlar

incelendiğinde en yüksek performansa sahip sınıflandırıcıların tümünün %90'ın üzerinde doğruluk oranlarına sahip olduğu görülmüştür. Ayrıca H10 sezgiseli hariç her bir sezgisel için sonuçlar incelendiğinde en yüksek performansa sahip sınıflandırıcıların tümünün %77,67 ile %100 arasında F1-skoru değerine ulaştığı gözlemlenmiştir. Genel olarak doğruluk oranları ve onlara karşılık gelen F1-skor değerlerine göre gradyan artırma ağaçları diğer algoritmalara göre daha iyi bir performans sergilemiştir.

Tablo 3. Kullanılabilirlik sezgiselleri için en iyi performans gösteren sınıflandırma modeli sonuçları

Kullanılabilirlik Sezgiseli	Doğruluk oranı (%) ± Standart sapma (%)	F1- skoru (%) ± Standart sapma (%)
H1: Sistem görünürlüğü	92,54 ± 0,93 Gradyan artırma ağaçları (85/15)	83,33 ± 1,88 Gradyan artırma ağaçları (85/15)
H2: Sistem ile gerçek dünya arasındaki uyum	97,63 ± 0,93 Gradyan artırma ağaçları (85/15)	81,85 ± 13,93 Gradyan artırma ağaçları (85/15)
H3: Kullanıcı kontrolü ve özgürlük	98,95 ± 2,35 Gradyan artırma ağaçları (95/5)	93,33 ± 14,91 Gradyan artırma ağaçları (95/5)
H4: Tutarlılık ve standartlar	95,95 ± 2,27 Gradyan artırma ağaçları (90/10)	77,67 ± 13,72 Gradyan artırma ağaçları (90/10)
H5: Hataları önleme	96,44 ± 2,27 Gradyan artırma ağaçları (90/10)	80,29 ± 18,47 Gradyan artırma ağaçları (90/10)
H6: Hatırlamak yerine tanıma	97,00 ± 4,47 Hızlı geniş marjin (95/5)	83,33 ± 23,57 Hızlı geniş marjin (95/5)
H7: Esneklik ve kullanım verimliliği	99,00 ± 2,24 Genelleştirilmiş lineer model, hızlı geniş marjin, karar ağaçları, gradyan artırma ağaçları (95/5)	80 ± 0,00 Genelleştirilmiş lineer model, hızlı geniş marjin, karar ağaçları, gradyan artırma ağaçları (95/5)
H8: Estetik ve minimalist tasarım	100,00 ± 0,00 Genelleştirilmiş lineer model, hızlı geniş marjin, karar ağaçları, gradyan artırma ağaçları, destek vektör makineleri (95/5)	100 ± 0,00 Genelleştirilmiş lineer model, hızlı geniş marjin, karar ağaçları, gradyan artırma ağaçları, destek vektör makineleri (95/5)
H9: Kullanıcıların hataları tanınmasına ve düzeltilmesine yardımcı olma	100,00 ± 0,00 Gradyan artırma ağaçları (90/10, 95/5)	100 ± 0,00 Gradyan artırma ağaçları (90/10, 95/5)
H10: Yardım ve dokümantasyon	92,95 ± 4,44 Hızlı geniş marjin (95/5)	58,82 ± 0,00 Hızlı geniş marjin (95/5)

4. Sonuç

Bu çalışma kapsamında kullanılabilirlik sezgiselleri geliştirme sürecini etkin hale getirebilmek için veri madenciliği ve makine öğrenmesi tekniklerinin nasıl kullanılabilceği gösterilmiştir. Öncelikle TV ve set üstü cihaz arayüzü ile ilgili kullanılabilirlik problemleri ilişkilendirme kuralları tekniği ile incelenmiş ve kullanılabilirlik sezgiselleri ile ilişkileri açısından sahip oldukları örüntüler belirlenmiştir. Ayrıca kullanılabilirlik problemlerinin kullanılabilirlik problemleri ile ilişkileri naive bayes, logistik regresyon, hızlı geniş marjin, derin öğrenme, rastgele orman, gradyan artırma ağaçları, destek vektör makineleri teknikleri kullanılarak tahmin edilmiştir. Kullanılabilirlik problemlerinin sezgisellerle eşleştirilmelerinin tahmin edilmesi konusunda sınıflandırıcıların doğruluk oranları %90 değerinin üzerinde, F1-skorları ise %75 değerinin üzerinde gerçekleşmiştir. Genel olarak gradyan artırma ağaçları tekniği diğer sınıflandırıcılara kıyasla daha iyi sonuçlar vermiştir. Çalışma sonuçlarına göre sezgisel geliştirme süreçlerinde makine öğrenmesi teknikleri süreçleri hızlandırarak kullanılabilirlik uzmanlarına destek olabilir.

Çalışmanın bazı kısıtları da bulunmaktadır. Bu çalışmada sadece TV ve set üstü cihaz arayüzleriyle ilgili kullanılabilirlik problemleri toplanmıştır. Farklı sistemlerle ve arayüzlerle ilgili kullanılabilirlik problemleri ile sezgiseller arasındaki ilişkilerin incelenmesi makine öğrenmesi tekniklerinin performanslarını objektif olarak görmek açısından faydalı olacaktır. Ayrıca sonuçlar veri setinin küçük ve dengesiz olması sebebiyle sınıflandırıcıların kullanılabilirlik problemlerinin bazı sezgisellerle ilişkilerini tahmin etmek konusunda düşük performans sergilediğini göstermiştir. Bu sebeple daha çok veri toplanarak benzer çalışmalar yapılması sayesinde daha güvenilir örüntüler elde edilebileceği gibi sınıflandırma algoritmaları da daha iyi performans sergileyebilecektir.

5. Teşekkür

Bu çalışma Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) tarafından TÜBİTAK 3001 programı ile desteklenmiştir (Proje numarası: 217M143, 2018). TÜBİTAK'a katkılarından dolayı teşekkür ederiz.

Kaynakça

- Boza, B. C., Schiaffino, S., Teyseyre, A., & Godoy, D. (2014). An approach for knowledge discovery in a web usability context. In Proceedings of the 13th Brazilian Symposium on Human Factors in Computing Systems, 393-396.
- Chamba-Eras, L., Jacome-Galarza, L., Guaman-Quinche, R., Coronel-Romero, E., & Labanda-Jaramillo, M. (2017, April). Analysis of usability of universities Web portals using the Prometheus tool-SIRIUS. In 2017 Fourth International Conference on eDemocracy & eGovernment (ICEDEG), IEEE, 195-199.
- Dökeroğlu, T., Malık, Z. M. M., & Shadi, A. S. (2018). Gözetimsiz Makine Öğrenme Teknikleri ile Miktarla Dayalı Negatif Birliktelik Kural Madenciliği. Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 6(4), 1119-1138.
- El-Halees, A. M. (2014). Software Usability Evaluation Using Opinion Mining. JSW, 9(2), 343-349.
- Etemadi, V., Bushehrian, O., & Akbari, R. (2017). Association rule mining for finding usability problem patterns: A case study on StackOverflow. In 2017 International Symposium on Computer Science and Software Engineering Conference (CSSE), IEEE, 24-29.
- González, M. P., Granollers, T., & Lorés, J. (2006). A hybrid approach for modelling early prototype evaluation under user-centred design through association rules. In International Workshop on Design, Specification, and Verification of Interactive Systems, Springer, Berlin, Heidelberg, 213-219.
- González, M. P., Lorés, J., & Granollers, A. (2008). Enhancing usability testing through datamining techniques: A novel approach to detecting usability problem patterns for a context of use. Information and software technology, 50(6), 547-568.
- Han, J., Kamber, M., & Pei, J. (2011). Data mining concepts and techniques third edition. The Morgan Kaufmann Series in Data Management Systems, 83-124.
- Hermawati, S., & Lawson, G. (2015). A User-Centric Methodology to Establish Usability Heuristics for Specific Domains. In Proceedings of the International Conference on Ergonomics & Human Factors, (pp. 80-85). Northamptonshire, UK: April 13-16.
- Hub, M., & Capkova, V. (2010). Heuristic Evaluation of Usability of Public Administration Portal. In Proceedings of the International Conference on Applied Computer Science, (pp. 234-239). The Netherlands: University of Amsterdam, May 31-June 2.
- ISO (1998). Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs). Part 11: Guidance on Usability (ISO 9241-11:1998). Retrieved from <https://www.iso.org/standard/16883.html>
- Ivory, M. Y., & Hearst, M. A. (2001). The state of the art in automating usability evaluation of user interfaces. ACM Computing Surveys (CSUR), 33(4), 470-516.
- Jimenez, C., Allende Cid, H., & Figueroa, I. (2017). PROMETHEUS: PROcedural and METHodology for developing HEuristics of Usability. IEEE Latin America Transactions, 15(3), 541-549. <https://doi.org/10.1109/TLA.2017.7867606>
- Kaya, A., Gumussoy, C. A., Ekmen, B., & Bayraktaroglu, A. E. (2021). Usability heuristics for the set-top box and TV interfaces. Human Factors and Ergonomics in Manufacturing & Service Industries, 31(3), 270-290.
- Kılınc, D., Borandağ, E., Yücalar, F., Tunalı, V., Şimşek, M., & Özçift, A. (2016). KNN algoritması ve r dili ile metin madenciliği kullanılarak bilimsel makale tasnifi. Marmara Fen Bilimleri Dergisi, 28(3), 89-94.
- Nielsen, J., & Molich, R. (1990). Heuristic Evaluation of User Interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, (pp. 249-256). Washington, USA. <https://doi.org/10.1145/97243.97281>
- Nielsen, J. (1993). Usability Engineering, Academic Press.
- Nielsen, J. (1994). How to Conduct a Heuristic Evaluation. Retrieved August 8, 2020, from <https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/>
- Nielsen, J. (1995). 10 Usability Heuristics for User Interface Design. Retrieved August 10, 2020, from <http://www.nngroup.com/articles/ten-usability-heuristics/>
- Oztekin, A., Delen, D., Turkyilmaz, A., & Zaim, S. (2013). A machine learning-based usability evaluation method for eLearning systems. Decision Support Systems, 56, 63-73.
- Quinones, D., & Rusu, C. (2017). How to Develop Usability Heuristics: A Systematic Literature Review. Computer Standards & Interfaces, 53, 89-122. <https://doi.org/10.1016/j.csi.2017.03.009>
- Quinones, D., Rusu, C., & Rusu, V. (2018). A Methodology to Develop Usability/User Experience Heuristics. Computer Standards & Interfaces, 59, 109-129. <https://doi.org/10.1016/j.csi.2018.03.002>
- Rusu, C., Roncagliolo, S., Rusu, V., & Collazos, C. (2011). A Methodology to Establish Usability Heuristics. In Proceedings of the Fourth International Conferences on Advances in Computer-Human Interactions (ACHI 2011), (pp. 59-62). Gosier, Guadeloupe, France: February 23-28.
- Sagar, K., & Saha, A. (2016). Enhancing usability inspection through data-mining techniques: an automated approach for detecting usability problem patterns of academic websites. In International Conference on Intelligent Human Computer Interaction, Springer, Cham, 229-247.
- Srikant, R., & Agrawal, R. (1995). Mining generalized association rules. In 21st VLDB Conference Zurich, Switzerland, 407-419.
- Wu, M., Wang, L., Li, M., & Long, H. (2014). An approach of product usability evaluation based on Web mining in feature fatigue analysis. Computers & Industrial Engineering, 75, 230-238.