

## Veri Madenciliği Birliktelik Kuralları ile Binaların Risk Durumlarının Analizi: Kaynaşlı, Düzce Örneği

Levent SABAH<sup>\*,a</sup>, Hüseyin BAYRAKTAR<sup>b</sup>

<sup>a,\*</sup> Düzce Üniversitesi Rektörlük Bilgi İşlem Daire Başkanlığı, ORCID:0000-0002-6911-4749, DÜZCE 81620, TÜRKİYE

<sup>b</sup> Düzce Üniversitesi Kaynaşlı M.Y.O. Yapı Ressamlığı Bölümü, ORCID: 0000-0001-7277-0838, DÜZCE 81900, TÜRKİYE

### MAKALE BİLGİSİ

Alınma: 03.12.2019  
Kabul: 28.04.2020

#### **Anahtar Kelimeler:**

Birliktelik Kuralları,  
Veri Madenciliği,  
Yerel Deprem Puanı,  
Coğrafi Bilgi  
Sistemleri, Kaynaşlı

#### **\*Sorumlu Yazar:**

e-posta: leventsabah@  
duzce.edu.tr

### ÖZET

Çok çeşitli verinin analiziyle anlamlı sonuçlar elde edilmesine yönelik teknikler veri madenciliği kavramını ifade etmektedir. Birliktelik kuralları, veri madenciliğinin temel tekniklerinden birisidir. Genel olarak farklı olayların, birlikte gerçekleşebilme durumlarının tespit edilmesinde kullanılmaktadır. Analizi yapılan veri setinde bulunan verilerin birbirleri arasında bulunan birliktelik bağıntıları bulunabilmektedir.. Böylece bu bağıntılar arasında eğer bir yada daha fazla durum var ise, bu durumların sonucunda belli durumlar olabilir şeklinde çıkarımlar yapılabilmektedir. Bu çalışmada Kaynaşlı (Düzce) ilçesinde yer alan 2112 binaya ait deprem riski puanlama verileri kullanılmıştır. Bu verisetindeki sözel veriler; binalar ile ilgili kat yüksekliği, kullanım durumu, bina malzemesi, 1999 depremi öncesinde ya da sonrasında inşaa edilme durumu ve elde edilen risk puanlarından oluşmaktadır. Daha detaylı birliktelik analizi yapılması için coğrafi bilgi sistemleri kullanılarak kesişim analizi yapılmış ve veri setinin mekansal verileri elde edilmiştir. Bunlar binaların buldukları konumlara bağlı olarak oluşturulan eğim, jeoloji, taşıma gücü, zemin titreşim, zemin sınıfı, zemin büyütme ve yer altı suyu verileridir. Elde edilen sonuçlara göre bir binanın yerel deprem puanı 93.5'den büyük ve zemin titreşimi 0.10-0.20 aralığında ise düşük riskli, bina 2 katlı ve düşük riskli ise yerel deprem puanı 93.5'den büyük ve meskenler gibi çıkarımlar yapılabilmektedir.

<https://dx.doi.org/10.30855/gmbd.2020.01.07>

## Analysis of Risk Situations of Buildings with Data Mining Association Rules: Kaynaşlı, Düzce Example

### ARTICLE INFO

Received: 03.12.2019  
Accepted: 28.04.2020

#### **Keywords:**

Association Rules,  
Data Mining, Local  
Earthquake Score,  
Geographical  
Information Systems,  
Kaynaşlı

#### **\*Corresponding**

#### **Authors**

e-mail:  
leventsabah@  
duzce.edu.tr

### ABSTRACT

Techniques for obtaining meaningful results by analyzing a wide variety of data express the concept of data mining. Association rules are one of the basic techniques of data mining. It is generally used in determining the different events and their occurrence. The data in the analyzed dataset can have correlation relations between each other. Thus, if there are one or more situations between these relations, it can be inferred that there may be certain situations as a result of these situations.. In this study, earthquake risk scoring data of 2112 buildings in Kaynaşlı (Düzce) district were used. Verbal data of this dataset; Consists of floor height, usage status, building material, construction status before or after the 1999 earthquake and the obtained risk points. In order to perform more detailed association rules analysis, intersection analysis was performed using geographic information systems and spatial data of the data set were obtained. These are the slope, geology, bearing capacity, ground vibration, ground class, ground magnification and groundwater data created depending on the locations of the buildings. According to the results obtained, the local earthquake score of a building is greater than 93.5 and the ground vibration is in the range of 0.10-0.20, the result is low risk. If the building has 2 floors and low risk, it can be deduced that the local earthquake score is greater than 93.5 and residential.

<https://dx.doi.org/10.30855/gmbd.2020.01.07>

## 1. INTRODUCTION (GİRİŞ)

Günümüzde veriler her saniyede hızlı bir şekilde büyüdüğü için, verilerin işlenmesi büyük bir zorluk haline gelmiştir [1]. Veri madenciliği, büyük veri tabanlarından bilgi veya kalıp çıkarmak için kullanılan bilgi keşfi işlemlerinden biridir [2]. Veri madenciliği işleminin genel amacı bir veri setinden bilgi elde etmek ve daha sonra kullanmak üzere anlaşılabilir bir yapıya dönüştürmektir [3].

Veri madenciliği, önceden bilinmeyen ve karar verme bilgisi için potansiyel değeri ve çok miktarda veriden kuralları belirlemektir. Veri madenciliği pazarının sürekli genişlemesiyle birlikte, veri madenciliği ile ilgili çalışmalara sürekli olarak yer verilmektedir. Veri madenciliği, uygulamalı bir bilim ve en başından beri pratik bir teknolojidir. Bu nedenle, güçlü deneyselliği nedeniyle veri madenciliğinin bilgi analizinde uygun veri madenciliği araçlarını seçmek çok önemlidir [4].

Birliktelik kuralları, büyük hacimli işlemlerde ilginç ilişkilendirme veya korelasyon ilişkisi ortaya çıkarmaya yönelik veri madenciliği analiz yöntemlerinden biridir [5]. Agrawal [6], ilk olarak 1993 yılında müşteri işlemleri veri tabanında madencilik kümeleri madenciliğini ve 1994 yılında klasik Apriori algoritmasını önerdi [6], [7], [8]. Apriori tabanlı algoritmaların iki adımı vardır. İlk adım, işlemlerden ayarlanan sık öğeyi bulmaktır. İkinci adım, birliktelik kuralını oluşturmaktır [5].

Birliktelik kuralı, devasa veri setindeki eğilimleri ve örüntüleri tanımlamaya yardımcı olur [9], [10], [11]. Destek ve güven, birliktelik kuralı madenciliğinin iki ana parametresidir [2], [12]. Birliktelik kuralları ile veri setinde bulunan örnekler arasındaki temel ilişkiler ve bağlantılar hızlı ve basit bir şekilde keşfedilebilmektedir. Örnek sayısı ve öznitelik sayısı arttıkça veriler arasındaki ilginç ilişkileri ortaya çıkarma ihtimali artabilmektedir.

Bir ilişkilendirme kuralında, A ve B, ayrı öge kümeleridir.  $A \rightarrow B$ 'in destek ve  $A \rightarrow B$ 'in güven değerleri sırasıyla kullanıcı tarafından belirlenen minimum destek ve minimum güvenden daha az olmadığı  $A \rightarrow B$  olarak oluşturulmaktadır [11].

Çalışmamızda bina verileri arasında muhtemel bağlantıların bulunması amaçlanmıştır. Bu bağlamda veri madenciliği tekniklerinden birliktelik kurallarını kullanarak, Kaynaşlı (Düzce) ilçesinde yer alan 2112

binaya ait sözel ve mekansal öznitelikler analiz edilmiştir.

## 2. MATERYAL VE YÖNTEM (MATERIALS AND METHODS)

Çalışmada veri madenciliği araçlarından Weka [13] ve Orange [14] ile Bayraktar'ın doktora tez çalışmasından [15] Düzce ili Kaynaşlı ilçesinde 7 mahallede 2112 binanın sokak taraması yöntemiyle depreme dayanım açısından oluşturduğu envanter kullanılmıştır. Bina verileri kimlik bilgisi (ID), mahalle, sokak, kapı no, kat yüksekliği, kullanım türü, yapım tarihi, malzeme, güçlendirme durumu, deprem öncesi mi?, deprem sonrası mı?, açıklama, ek açıklama, uygunluk ve risk puanı özniteliklerinden oluşmaktadır. Risk puanı "Yerel Deprem Puanı" [15] şeklinde oluşturulan yöntemle hesaplanarak binaların düşük, orta veya yüksek seviyede risk durumları tanımlanmaktadır. Bu envanter ile birliktelik kurallarının uygulanması için veri ön işleme yapılmıştır. Sonrasında bina özniteliklerine ek olarak mekansal veriler binalarla eşleştirilmiştir. Yani Açık Kaynak Kodlu Coğrafi Bilgi Sistemi QGIS [16] kullanılarak binaların kapalı alan (poligon) şeklinde bulunan mekansal verileri ile Kaynaşlı ilçesine ait eğitim, jeoloji, taşıma gücü, yer altı suyu, zemin, zemin büyültme, zemin sınıfı, zemin titreşim mekansal katmanları üzerinde kesişim (intersection) işlemi yapılarak her binaya ait bu katmanlarda eşleştiği veriler elde edilerek birliktelik analizinde kullanılmak üzere veri ön işleme adımları uygulanmıştır.

### 2.1. Veri Madenciliği (Data mining)

Teknolojik gelişmeler sayesinde artan bir ivme şeklinde veri üretimi artmaya devam etmektedir. Bununla birlikte farklı birçok kaynaktan elde edilen verilerin anlamlı bilgiye dönüştürülmeleri için veri madenciliği teknikleri devamlı olarak gelişmekte ve yeni teknikler ortaya çıkmaktadır. Veri madenciliği sayesinde yığın haline gelmekte olan verilerden anlamlı sonuçlar, ilk bakışta görülemeyen, bilinmeyen ya da öngörülemeyen bağlantılar ve tahmin yöntemleri elde edilmektedir. Veriler üzerinden veri madenciliği tekniklerinin uygulanabilmesi için verilerin bazı veri ön işleme adımlarından geçmesi gerekebilmektedir [17], [18]. Genel olarak veri ön işleme adımları 4 adım şeklinde ifade edilebilir.

### 2.1.1. Veri Ön İşleme Adımları (*Data Preprocessing Steps*)

Bu bölümde veri madenciliğinde yapılacak analizlerin kalitesini ve anlamlı sonuçlar elde etme olasılığını arttıran veri ön işleme adımları anlatılacaktır. Şekil 1 de veri ön işleme adımları gösterilmektedir. Ham veri üzerinde veri ön işleme yapıldıktan sonra yapılandırılmış veriler elde edilmektedir.



Şekil 1. Yapılandırılmış verinin oluşturulması [19]  
(Generation of structured data)

#### 2.1.1.1. Veri temizleme (*Data cleaning*)

Birçok satır ve sütundan oluşan veri seti üzerinde tutarlı bir modelleme yapmak için öncelikli olarak eksik veriler sonuca etkisi az ya da hiç olacağı düşünülmüyorsa silinebilir, sabit bir değer verilebilir ya da sütunda bulunan diğer verilerin ortalamaları hesaplanarak eksik veri bu şekilde tamamlanabilir.

#### 2.1.1.2. Veri bütünleştirme (*Data integration*)

Büyük veri kavramında farklı kaynaklardan farklı yapıda birçok veri bulunmaktadır. Bu farklı verilerin belirlenecek bir yapıda bütünleştirilmesi ile farklı kaynaklardan toplanan veriler aynı amaca uygun kullanılabilir hale getirilmelidir. Veri bütünleştirmeye örnek olarak bir kolonda tutulacak verilerin farklı kaynaklarda farklı kodlamalar şeklinde tutulması örnek olarak gösterilebilir (İl kolonunun bazı kaynaklarda plaka numarası ile ya da direkt ismiyle tutulması gibi). Çalışmamızda binalar 1999 depremi öncesi ya da sonrasında inşa edildiği bilgisi 2 farklı kolonda tutulmakta iken tek kolonda tanımlanmıştır.

#### 2.1.1.3. Veri dönüştürme (*Data conversion*)

Veri dönüştürme işlemi, verinin içeriğinin aynı kalmasıyla birlikte şeklinin dönüştürülmesi olarak tanımlanabilir. Veri dönüştürme işlemi kullanılacak modele uygun biçimde yapılmalıdır. Çalışmamızda bina envanteri ile ilgili “güçlendirme yapıldı mı?” kolonunda eğer binada güçlendirme yapılmış ise sadece evet şeklinde veri bulunmaktayken veri dönüştürme işlemi ile veri 0 ve 1 şeklinde kodlanmıştır. 1 güçlendirme için evet yazan verileri

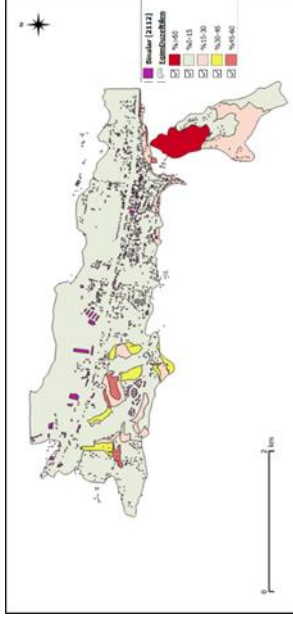
temsil ederken 0 veri kısmı boş olan satırlara eklenen kodlanmış veridir.

#### 2.1.1.4. Veri indirgeme (*Data reduction*)

Veri madenciliği araçları ile analiz edilecek ve modellenecek verilerden amaca en uygun verilerin seçimi anlamlı bir sonuç elde edilebilmesi için önem teşkil etmektedir. Çalışmamızdaki veri setinden her binaya ait başta adres bilgileri olmak üzere (mahalle, sokak, kapı numarası, açıklama, ek açıklama) sonuca etki etmeyecek olan sütunlardaki veriler tamamen silinmiştir.

## 2.2. Mekânsal Verinin Elde Edilmesi (*Obtaining Spatial Data*)

Çalışmamızda mekansal verileri elde etmek için Açık Kaynak Kod bir Coğrafi Bilgi Sistemi (CBS) olan QGIS kullanılmıştır. Bir CBS uygulamasıyla, kullanıcılar bilgisayarlarında sayısal haritalar açabilir, bir haritaya eklemek için yeni mekansal bilgiler oluşturabilir, ihtiyaçlarına göre özelleştirilmiş basılı haritalar oluşturabilir ve mekansal analizler yapabilir [20]. Veri madenciliği analizlerinde verinin kalitesi önemlidir. Bunun için büyük veri üzerinde veri ön işleme adımları uygulanarak elde edilen veri setinin kurgulanacak modele göre anlamlı sonuçlar ortaya çıkarması için yeterli sayıda örneklem ve öz nitelik olmalıdır. Bundan dolayı çalışmamızda kullandığımız bina envanteri veri setine ek öz nitelikler eklemek için her binanın mekansal kesişim analizi ile bulunduğu yerdeki eğim, jeoloji, taşıma gücü, yer altı suyu, zemin, zemin büyültme, zemin sınıfı, zemin titreşim katmanları kullanılmıştır. Böylece daha fazla öz nitelik ile veriler arasındaki ilginç bağlantıları keşfetme olasılığı arttırılmıştır. Şekil 2’ de binalar eğim katmanı ile birlikte gösterilmektedir. Benzer şekilde diğer katmanlar üzerinde de kesişim analizler yapılmıştır.



Şekil 2. Kaynaşlı'daki Bina ve Eğim Verilerinin Birlikte Gösterimi (*Demonstration of Building and Slope Data in Kaynaşlı*)

### 2.3. Birliktelik analizi (*Association analysis*)

Birliktelik analizi, genel anlamda belli bir destek ve güven eşik değeri yardımıyla birlikte sık görülen işlemlerin, kayıtların veya verilerin tespit edilip analiz edilmesidir. Böylece destek ve güven eşik değerleri kullanılarak veri seti içindeki en önemli ilişkiler analiz edilerek tanımlanabilmektedir.

Birliktelik kurallarının genel bir örneği olarak market sepeti incelenecek olursa birlikte en çok satılan ürünler veya ürün grupları satış işlemleri içerisinde tespit edilebilmektedir. Böylece bu ürün grupları market içindeki raflarda birbirlerine daha yakın bir şekilde bir araya getirilmektedir [3], [17], [21].

Çalışmamızda bina verileri ile birliktelik analizi yapılmadan önce veri ön işleme adımları uygulanmıştır. Daha sonrasında bina envanter verilerine ek öznitelikler Kaynaşlı ilçesine ait sayısal haritalar QGIS coğrafi bilgi sistemi ile kesişim analizi yapılarak veri setinin son hali elde edilmiştir. Birliktelik analizi için Weka ve Orange veri madenciliği programları kullanılmıştır. Birliktelik analizinde yaygın bir algoritma olan Apriori algoritması bu iki programda uygulanmıştır.

Apriori, boolean ilişki kuralları için geçerli bir veri madenciliği algoritmasıdır. Algoritmanın ismi, bilgileri bir önceki adımdan aldığı için “prior”

anlamında “Apriori”dir. Bu algoritma özünde iteratif (tekrarlayan) bir niteliğe sahiptir ve hareket bilgileri içeren veri setinde sık geçen öge kümelerinin keşfedilmesinde kullanılmaktadır. Bu algoritmada temel yaklaşımı eğer k-öge kümesi minimum destek ölçütünü sağlıyorsa, bu kümenin alt kümeleri de minimum destek ölçütünü sağlamaktadır. Bir öğeler kümesinin destek değeri, alt kümesinin destek değerinden büyük olmamaktadır [22], [23].

#### 2.3.1. Tanımlar (*Definitions*)

Elde edilen birliktelik kuralları ile birlikte bu kurallar ile ilgili çeşitli ölçütler elde edilmektedir. Bunlar Destek (Support), Güven (Confidence), Kapsam (Coverage) ve Kaldırma (Lift) dir. Burada destek ve güven değerlerinin değeri büyüdükçe elde edilen birliktelik kuralının o kadar güçlü olduğu çıkarımı yapılabilir. Analiz işlemlerinde uygulanacak algoritma için destek ve güven parametrelerinin uygun bir aralıkta seçimi önemlidir. Çok büyük değerlerde bazı kurallar tespit edilemeyebilir çok küçük seçildiğinde ise benzer birçok kural ortaya çıkabilir ve amaca uygun olarak bir kural tespit edilemeyebilir. Geçerli birliktelik kurallarının oluşturulabilmesi için minimum destek ve güven (eşik) değerlerinin sağlanması gereklidir.

##### 2.3.1.1. Destek (*Support*)

Bir kuralın tüm veri setinde ne kadar tekrarlandığını belirtir.

$$\text{Destek}(A \rightarrow B) = \frac{\text{Sayı}(A,B)}{N} \quad (1)$$

##### 2.3.1.2. Güven (*Confidence*)

Örnek bir A değişkeninin B değişkeni ile birlikte bulunma olasılığıdır.

$$\text{Güven}(A \rightarrow B) = \frac{\text{Sayı}(A,B)}{\text{Sayı}(A)} \quad (2)$$

Birliktelik kurallarının oluşturulmasında destek ve güven dışında çok sayıda farklı istatistikî ölçütler de kullanılır. Bu ölçütlerin birçoğu destek ve güven ölçütlerine bağlı hesaplanır [24].

##### 2.3.1.3. Kapsam (*Coverage*)

$A \rightarrow B$  şeklinde bir kuralın ne kadar sıklıkta olduğudur. Öncül destek (antecedent support) şeklinde de tanımlanmaktadır.

$$Kapsam(A \rightarrow B) = Destek(A) \quad (3)$$

### 2.3.1.4. Kaldırma (Lift)

A ve B değerlerinin bağımsız olup olmadıklarının gösteren oransal destek değeridir. Değişkenler birbirine bağımlı olmalıdırlar, aksi takdirde bu değerlerle ilişkili kurallar çıkarılamaz. Kaldıraç değeri 1 den büyük ise değişkenler birbirine bağımlıdır. Bulunan kuralın ne kadar ilgi çekici olduğunu belirlemek için kullanılır.

$$Kaldırma(A \rightarrow B) = \frac{Destek(A,B)}{Destek(A)*Destek(B)} \quad (4)$$

## 3. BULGULAR VE TARTIŞMA (FINDINGS AND DISCUSSION)

Veri madenciliği programlarından Orange ve Weka ile yapılan birliktelik analizlerinin sonuçları sırasıyla Şekil 3' de Orange programında yapılan analiz sonucu elde edilen bağlantılar gösterilmektedir. Şekil 4' de ise Weka programı ile yapılan birliktelik analiz sonucu elde edilen kurallar gösterilmektedir.

Orange programı ile bina envanteri veri seti üzerinde yapılan birliktelik analiz sonuçları Tablo 1' de gösterilmektedir. Sonuçlar incelendiğinde risk puanı 93.5'den büyük ve titreşim aralığı 0.10 - 0.20 ise binanın mesken ve düşük riskli olduğu söylenebilir. Bunun dışında 2 katlı ve düşük riskli binaların risk puanlarının 93.5'den büyük ve mesken oldukları gibi genelleştirmeler yapılabilir.

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent	Consequent
0.314	0.950	0.331	2.073	1.385	0.087	risk_puanı ≥ 93.5, ZeminTitresimWrtBina_d=0.10-0.20,	kullanım=mesken, riskderecesi=düşük
0.314	0.954	0.330	2.115	1.369	0.085	ZeminTitresimWrtBina_d=0.10-0.20,	kullanım=mesken, risk_puanı ≥ 93.5
0.301	0.933	0.323	2.158	1.338	0.076	kat_yuksekl=2, riskderecesi=düşük	kullanım=mesken, risk_puanı ≥ 93.5
0.301	0.912	0.330	2.079	1.330	0.075	kat_yuksekl=2, risk_puanı ≥ 93.5	kullanım=mesken, riskderecesi=düşük
0.314	0.901	0.349	1.966	1.313	0.075	ZeminTitresimWrtBina_d=0.10-0.20,	kullanım=mesken, riskderecesi=düşük
0.314	0.901	0.349	1.966	1.313	0.075	ZeminTitresimWrtBina_d=0.10-0.20,	kullanım=mesken, risk_puanı ≥ 93.5, riskderecesi=düşük
0.396	0.900	0.440	1.558	1.312	0.094	malzemesi=betonarme, risk_puanı ≥ 93.5	kullanım=mesken, riskderecesi=düşük
0.396	0.900	0.437	1.597	1.303	0.092	malzemesi=betonarme, riskderecesi=düşük	kullanım=mesken, risk_puanı ≥ 93.5
0.330	0.997	0.330	2.322	1.299	0.076	kat_yuksekl=1, risk_puanı ≥ 93.5	riskderecesi=düşük
0.316	0.905	0.349	1.997	1.299	0.073	ZeminTitresimWrtBina_d=0.10-0.20,	kullanım=mesken, risk_puanı ≥ 93.5
0.330	0.996	0.331	2.319	1.297	0.076	risk_puanı ≥ 93.5, ZeminTitresimWrtBina_d=0.10-0.20,	kullanım=mesken, risk_puanı ≥ 93.5
0.420	0.996	0.422	1.817	1.297	0.096	deprem_once_0_sonra_1=1.0, risk_puanı ≥ 93.5, EjinWrtBina_c= %0-15	riskderecesi=düşük
0.314	0.996	0.316	2.430	1.297	0.072	kullanım=mesken, deprem_once_0_sonra_1=1.0, risk_puanı ≥ 93.5,	riskderecesi=düşük
0.497	0.995	0.500	1.536	1.297	0.114	deprem_once_0_sonra_1=1.0, risk_puanı ≥ 93.5,	riskderecesi=düşük
0.366	0.995	0.367	2.089	1.296	0.084	kullanım=mesken, deprem_once_0_sonra_1=1.0, risk_puanı ≥ 93.5,	riskderecesi=düşük
0.440	0.995	0.442	1.736	1.296	0.100	kullanım=mesken, deprem_once_0_sonra_1=1.0, risk_puanı ≥ 93.5,	riskderecesi=düşük
0.339	0.902	0.376	1.852	1.294	0.077	malzemesi=betonarme, EjinWrtBina_c= %0-15, riskderecesi=düşük	kullanım=mesken, risk_puanı ≥ 93.5
0.437	0.991	0.440	1.743	1.292	0.099	malzemesi=betonarme, EjinWrtBina_c= %0-15, riskderecesi=düşük	kullanım=mesken, risk_puanı ≥ 93.5
0.376	0.991	0.380	2.021	1.292	0.085	malzemesi=betonarme, risk_puanı ≥ 93.5, EjinWrtBina_c= %0-15	riskderecesi=düşük
0.396	0.991	0.400	1.918	1.291	0.089	malzemesi=betonarme, risk_puanı ≥ 93.5, EjinWrtBina_c= %0-15	riskderecesi=düşük
0.339	0.990	0.343	2.239	1.290	0.076	kullanım=mesken, malzemesi=betonarme, risk_puanı ≥ 93.5,	riskderecesi=düşük
0.768	0.985	0.779	0.985	1.283	0.169	risk_puanı ≥ 93.5	riskderecesi=düşük

Şekil 3. Orange programı ile elde edilen birliktelik kuralları (Association rules obtained with the Orange)



Tablo 1. Orange Programı ile elde edilen bazı birliktelik kuralları (Some association rules obtained with the Orange)

Destek	Güven	Kapsam	Kaldırma	Öncülli	Sonucu
0,301	0,933	0,323	1,338	2 Katlı, Düşük Riskli	Risk Puanı $\geq$ 93,5, Mesken
0,314	0,950	0,331	1,385	Risk Puanı $\geq$ 93,5, Zemin Titreşim=0.10-0.20	Mesken, Düşük Riskli

0,420	0,996	0,422	1,297	Depremden sonra inşa edilmiş, Risk Puanı $\geq$ 93,5, Eğim=0-15	Düşük Riskli
-------	-------	-------	-------	---	--------------

Şekil 4. Weka programı ile elde edilen birliktelik kuralları (Association rules obtained with the Weka)

Tablo 2' de Weka programı ile elde edilen bazı birliktelik kurallarına bakıldığında ise taşıma gücü 1.00 - 2.00 ise eğim %0 -15 aralığındadır ya da düşük riskli binaların zemin büyüklüğü 1.50-2.00 ve eğimi %0-15 aralığındadır şeklinde genelleme yapılabilir.

Tablo 2. Weka Programı ile elde edilen bazı birliktelik kuralları (Some association rules obtained with the Weka)

Destek	Güven	Kaldırma	Öncülli	Sonucu
0,3	0,94	1,08	Taşıma gücü = 1.00-2.00	Eğim %0-15
0,3	0,91	1,19	Zemin büyüklüğü 1.50-2.00	Düşük riskli
0,3	0,9	1,17	Eğim %0-15, zemin büyüklüğü 1.50-2.00	Düşük riskli

Association rules:

```

1. taşınmaz_ticari_ticari_10-2,00 boyut= 104 ==> Erişilebilirlik_0-15 972    conf=(0,84) > lift=(1,03) 1491 (0,03) 1461 conf=(2,07)
2. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1027    conf=(0,83) > lift=(1,14) 1491 (0,04) 1461 conf=(2,20)
3. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
4. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
5. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
6. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
7. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
8. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
9. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
10. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
11. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
12. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
13. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
14. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
15. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
16. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
17. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
18. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
19. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)
20. taşınmaz_ticari_ticari_10-2,00 boyut= 110 ==> Erişilebilirlik_0-15 1151 (1,03) 1491 (0,04) 1461 conf=(1,45)

```

Birliktelik analizi sonucunda Orange ve Weka ile çeşitli bağlantılar elde edilmiştir. Elde edilen birliktelik kuralları incelendiğinde her iki programda da veriler arasında bir takım bağlantılar bulunmuştur. Aynı veri seti ile farklı bağlantıların ortaya çıkması nedenlerinden biri olarak veri setinin boyutunun küçük olması, rasgele oluşturulacak başlangıç koşullarında farklı sonuçlar çıkarabilmektedir. Ayrıca çevrimiçi kaynaklar incelendiğinde ise [25] verilen cevapta test ve eğitim verileri arasındaki farkların nedenleri üzerinde durulmuştur. Burada farklı programlarda aynı modellerin farklı doğruluk değerleri verebilmelerinin nedeni olarak düşünülebilir. Ayrıca kodlama farklılıklarında etkili olabileceği belirtilmiştir. Aynı sınıflandırma modeli için farklı programların verileri mesela ondalıklı kısımlarının farklı uzunlukta alınması, yuvarlanması, eğitim verilerinin farklı seçilmesi gibi nedenlerle de aynı veri setinde farklı sonuçlar alınabilmektedir. Bununla birlikte elde edilen farklı sonuçlar arasından uzmanlar tarafından uygun görülen çıkarımlar yapılabilmektedir.

#### 4. SONUÇLAR VE ÖNERİLER (RESULTS AND SUGGESTIONS)

Bu çalışmanın amacı bina envanter veri tabanında birbiriyle ilişkili olabilecek verilerin veri madenciliği yöntemlerinden biri olan birliktelik kuralı analizinin uygulanmasıdır. Bunun sonucunda elde edilen bağlantıların genelleştirilmesi ile daha kapsamlı veri setlerinde genel çıkarımlar yapılabileceği, binalar ile ilgili daha fazla öznel bilgileri ve daha fazla bina verisi ile analiz işlemlerinden daha fazla ve farklı çıkarımlar elde edilebileceği düşünülmektedir.

Yapılan birliktelik analizleriyle elde edilen sonuçlara göre bir binanın yerel deprem puanı 93.5'den büyük ve zemin titreşimi 0.10-0.20 aralığında ise düşük riskli, bina 2 katlı ve düşük riskli ise yerel deprem puanı 93.5'den büyük ve meskendir gibi çıkarımlar yapılabilmektedir.

Birliktelik kurallarının kullanımı ile mevcut verilerin birbiri ile ilişki durumları ortaya çıkarılabilmektedir.

Daha büyük hacimli bir bina envanteri ile bina özelliklerinin hangi mekânsal özellikler ile ilişkili olabileceği üzerine daha detaylı bağlantılar çıkarılabileceği düşünülmektedir.

Seçilecek güven ve destek değerlerine göre kaldırma değeri en yüksek kurallar sıralanarak veri seti içinden ilginç bağlantılar elde edilebilir.

Birliktelik analizi yapılacak farklı veri madenciliği araçları ve algoritmaları ile daha farklı sonuçlar elde edilebilmektedir.

Analiz yapılacak verisetinde ilave öznel bilgiler kullanılarak veriler arasındaki daha farklı bağlantılar elde edilebilir. Ayrıca veri kümesinin eleman sayısının artırılması ile daha tutarlı çıkarımlar yapılabilir.

#### ÇIKAR ÇATIŞMASI BEYANI (CONFLICT OF INTEREST STATEMENT)

Yazarlar tarafından herhangi bir çıkar çatışması bildirilmemiştir.

#### REFERENCES (KAYNAKLAR)

[1] S. Daw, "Machine Learning Applications Using Waikato Environment for Knowledge Analysis," in Proc. of the 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 11-13 March 2020, Erode, India [Online].

Available: IEEE Xplore, <https://ieeexplore.ieee.org/abstract/document/9076531>

[2] R. Amornchewin and W. Kreesuradej, "Probability-based incremental association rule discovery algorithm," in Proc. of the International Symposium on Computer Science and Its Applications (CSA), 13-15 Oct. 2008, Hobart, ACT, Australia [Online]. Available: IEEE Xplore, <https://ieeexplore.ieee.org/document/4654088>. [Accessed: July, 15, 2019].

[3] J. K. Chahal, "Finding Association Rules in Medical Datasets," *International Journal of Scientific Research in Science and Technology*, vol. 2, no. 3, pp. 167-170, April 2019. Doi: <http://doi.org/10.32628/IJSRST>

[4] Y. Shen, J. Liu, and J. Shen, "The further development of Weka base on positive and negative association rules," in Proc. of the 2010 International Conference on Intelligent Computation Technology and Automation (ICICTA), 11-12 May 2010, Changsha, China [Online]. Available: IEEE Xplore, <https://ieeexplore.ieee.org/document/5523114>. [Accessed: July, 15, 2019].

[5] K. Rameshkumar, M. Sambath, and S. Ravi, "Relevant association rule mining from medical dataset using new irrelevant rule elimination technique," in Proc. of the 2013 International Conference on Information Communication and Embedded Systems (ICICES), 21-22 Feb. 2013, Chennai, India [Online]. Available: IEEE Xplore, <https://ieeexplore.ieee.org/document/6508351>. [Accessed: July, 15, 2019].

[6] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," in *ACM SIGMOD Record*, vol. 22, no. 2, pp. 207-216, Jun. 1993. Doi: <https://doi.org/10.1145/170036.170072>

[7] Y. Zhong and Y. Liao, "Research of mining effective and weighted association rules based on dual confidence," in Proc. of the 4th International Conference on Computational and Information Sciences (ICIS), 17-19 Aug. 2012, Chongqing, China [Online]. Available: IEEE Xplore, <https://ieeexplore.ieee.org/document/6301339>. [Accessed: July, 15, 2019].

[8] J. Feng, Q. Zeng, and Z. Zhang, "A method of mining the meta-association rules for dynamic association rule based on the model of AR-markov," in Proc. of the 2nd International Conference on Networks Security, Wireless Communications and Trusted Computing (NSWCTC), 24-25 Apr. 2010, Wuhan, Hubei, China [Online]. Available: IEEE Xplore,

<https://ieeexplore.ieee.org/document/5480826>.  
[Accessed: July, 15, 2019].

[9] S. Naredi and R. A. Deshmukh, "Improved extraction of quantitative rules using Best M Positive Negative Association Rules Algorithm," in Proc. of the 2015 IEEE International Conference on Electronics, Computing and Communication Technologies, (CONECCT), 10-11 July 2016, Bangalore, India [Online]. Available: IEEE Xplore,

<https://ieeexplore.ieee.org/document/7383857>.  
[Accessed: July, 15, 2019].

[10] X. Piao, Z. Wang, and G. Liu, "Research on mining positive and negative association rules based on dual confidence," in Proc. of the 5th International Conference on Internet Computing for Science and Engineering, (ICICSE), 1-2 Nov. 2010, Heilongjiang, China [Online]. Available: IEEE Xplore,

<https://ieeexplore.ieee.org/document/6076550>.

[11] W. Ouyang and Q. Huang, "Mining direct and indirect weighted fuzzy association rules in large transaction databases," in Proc. of the 6th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD, 14-16 Aug. 2009, Tianjin, China [Online]. Available: IEEE Xplore,

<https://ieeexplore.ieee.org/document/5358916>.  
[Accessed: July, 15, 2019].

[12] X. S. Xiong, L. Fan, and Z. Lei, "Ontology-Based Association Rule Quality Evaluation Using Information Theory," in Proc. of the 2010 International Conference on Computational and Information Sciences, 17-19 Dec. 2010, Chengdu, China [Online]. Available: IEEE Xplore,

<https://ieeexplore.ieee.org/document/5708913>.

[13] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software," in ACM SIGKDD Explorations Newsletter, vol. 11, no. 1, p. 10, Nov. 2009. Doi: <https://doi.org/10.1145/1656274.1656278>

[14] J. Demšar et al., "Orange: Data Mining Toolbox in Python," *Journal of Machine Learning Research*, vol. 14, no. 1, pp. 2349–2353, Jan. 2013. Doi: <https://dl.acm.org/doi/abs/10.5555/2567709.2567736>

[15] H. Bayraktar, "CBS ve Sokak Taraması Yöntemleri Kullanılarak Düzce- Kaynaşlı İlçesinin Afet Riski Yönünden Yerleşim Durumunun Belirlenmesi ve Yerel Afet Risk Yönetimi," Ph.D. dissertation, Gazi Üniversitesi, Ankara, Türkiye, 2014.

[16] "QGIS." [Online]. Available: <https://www.qgis.org/en/site/about>. [Accessed: July, 15, 2019].

[17] İ. B. Aydılek, "Veri Kümelerindeki Eksik Değerlerin Yeni Yaklaşımlar Kullanılarak Hesaplanması," Ph.D. dissertation, Selçuk Üniversitesi, Konya, Türkiye, 2013.

[18] D. Sağaltıcı, F. D. Alay, C. Efil, and N. İlhan, "Veri Madenciliği Yöntemleri İle Meteorolojik Verilerden Kayıp Güneş Işınım Değerlerinin Tahmini," *Harran Üniversitesi Mühendislik Dergisi*, vol. 03, no. 2, pp. 49–53, Oct. 2018.

[19] Imran, S. Ahmad, and D. H. Kim, "Quantum GIS Based Descriptive and Predictive Data Analysis for Effective Planning of Waste Management," *IEEE Access*, vol. 8, pp. 46193–46205, March 2020. Doi: 10.1109/ACCESS.2020.2979015

[20] Q. Cai, "Cause Analysis of Traffic Accidents on Urban Roads Based on an Improved Association Rule Mining Algorithm," *IEEE Access*, vol. PP, p. 1, April 2020. Doi: 10.1109/ACCESS.2020.2988288

[21] F. C. Özçakir and A. Y. Çamurcu, "Birliktelik Kuralı Yöntemi İçin Bir Veri Madenciliği Yazılımı Tasarımı Ve Uygulaması," *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, vol. 6, no. 12, pp. 21–37, Güz 2007/2.

[22] A. C. Gülce, "Veri Madenciliğinde Apriori Algoritması ve Apriori Algoritmasının Farklı Veri Kümelerinde Uygulanması," M.Sc. dissertation Trakya Üniversitesi, Edirne, Türkiye, 2010.

[23] H. Sever and B. Oğuz, "Veritabanlarında Bilgi Keşfine Formal bir Yaklaşım, Kısım 1: Eşleştirme Sorguları ve Algoritmalar," *Bilgi Dünyası*, vol. 3, no. 2, pp. 173-204, Jan. 2002.

[24] M. Hahsler, "A Probabilistic Comparison of Commonly Used Interest Measures for Association Rules." [Online]. Available: [https://michael.hahsler.net/research/association\\_rules/measures.html](https://michael.hahsler.net/research/association_rules/measures.html). [Accessed: July, 15, 2019].

[25] "Validation and Testing accuracy widely different." [Online]. Available: <https://stackoverflow.com/questions/48718663/validation-and-testing-accuracy-widely-different>.

[Accessed: July, 15, 2019].

## Levent SABAH

1986 yılında Eskişehir’ de doğmuştur. Düzce Üniversitesi Rektörlük Bilgi İşlem Daire Başkanlığı’ nda çalışmaktadır. Açık Kaynak Kodlu Coğrafi Bilgi Sistemleri ile web ve mobil uygulamalar geliştirmektedir. Güncel olarak salgın modellemeleri üzerine doktora çalışmaları yapmaktadır.



### **Hüseyin BAYRAKTAR**

1977 yılında Manisa’ da doğmuştur. Düzce Üniversitesi Kaynaşlı M.Y.O.’ da çalışmaktadır. Kentsel Planlama, Deprem, Planlamada CBS ve Bilgi Teknolojileri, Afet Risk Yönetimi ve Sakınım Planlaması üzerine çalışmaktadır.