



İŞYERİNDE DEVAMSIZLIK YAPMAYA ETKİ EDEN DEMOGRAFİK FAKTÖRLERİN VERİ MADENCİLİĞİ TEKNİKLERİ İLE ARAŞTIRILMASI INVESTIGATION OF DEMOGRAPHIC FACTORS THAT AFFECT ABSENTEEISM AT WORKPLACE BY DATA MINING TECHNIQUES

Sengül CAN*

Öz

Gelişen teknoloji büyük veriden anlamlı bilgi çıkarılmasını zorlaştırmaktadır. Bu nedenle veri madenciliği teknikleri kullanılmaktadır. Bu çalışma iş yerinde devamsızlık yapan personel verilerini, veri madenciliği teknikleri kullanılarak sınıflandırılmıştır. İş yerinde devamsızlık üzerinde demografik değişkenlerin etkileri araştırılmıştır. Devamsızlık davranışında en önemli etkenin eğitim olduğu görülmüştür. Devamsız personelin %77.7 oranında yüksekokul veya doktora eğitim derecesine sahip olduğu ancak lisans ve yüksek lisans seviyesindeki personelde devamsızlık davranışının daha az görüldüğü tespit edilmiştir. Devamsızlık alışkanlığı için diğer belirleyici faktörün ise yaş olduğu görülmüştür. 47 yaşından daha genç personelin daha fazla devamsızlık yaptığı ve 47 yaşından büyük personelde ise alkol kullanımının %59 oranında işe gelmeyi olumsuz etkilediği görülmüştür.

Anahtar Kelimeler: İşyerinde Devamsızlık, Veri Madenciliği, Karar Ağacı.

Abstract

Developing technology makes it difficult to produce meaningful information in large quantities. For this reason, data mining techniques are used. This study classified data on absenteeism in the workplace using data mining techniques. Later, the effects of demographic variables on staff absenteeism at the workplace were investigated. It was seen that education is the most important factor on absenteeism. It was found that 77.7% of the non-attendance staff had a college or doctoral degree, but less non-attendance behavior was observed in the undergraduate and graduate level staff. The other determining factor for absenteeism is age. It was seen that younger staff than 47 years old had more absenteeism, and in the staff who is older than 47 years, alcohol use had a negative effect on coming to work.

Keywords: Absenteeism, Data Mining, Decision Tree.

Giriş

Veri madenciliği teknikleri pek çok farklı alanda ve farklı amaçlarla kullanılmakta olup verinin tanımlanması, sınıflandırılması, kümelenmesi, ilişkilendirilmesi ve tahmin edilmesi gibi bazı temel uygulamalarda sıklıkla kullanılmaktadır. Veri madenciliği teknikleri için bilgi teknolojileri kullanılıyor olsa da elde edilen sonuçların yorumlanması da oldukça önemlidir. Dolayısıyla analiz süreci tümüyle bilgisayar ortamında ve otomatik olarak gerçekleşmemektedir (Seyrek ve Ata, 2010, 71).

Günümüzde iletişim teknolojileri, bilgisayar teknolojileri ve bilgi teknolojilerinde yaşanan değişimler işletmelerin ellerinde bulunan verilerde artışa neden olmuştur. Depolanan verilerdeki artış ise hem veri analizine yönelik araçların gelişmesine hem de firmaların veri analizine olan ilgisinin artmasına neden olmuştur. Ortaya çıkan bu gelişmeler sonucu hem akademik hem de iş dünyasında veri madenciliği uygulamaları yaygınlaşmıştır (Rygielski vd., 200, 483).

İşletmeler personelin işe devamsızlığı nedeniyle her yıl milyarlarca dolar zarara uğramaktadır. Bu nedenle günümüzde işletmelerin personel ile olan ilişkilerinde çözümleri gereken en önemli sorunların başında devamsızlık gelmektedir. İşyerinde devamsızlık hem işletme verimliliğini etkilemekte hem de işletmenin istihdam ve insan kaynakları politikalarını etkilemektedir (Bacak ve Yiğit, 2010, 30). Bu çalışma kapsamında işyerinde devamsızlık yapan personelin daha çok hangi nedenlerle devamsızlık yaptığı ve kişinin yaş cinsiyet, eğitim düzeyi vb. faktörlerin devamsızlık üzerine bir etkisi olup olmadığı araştırılacaktır. Çalışma kapsamında önce işyerinde devamsızlık kavramı ve ekonomik sonuçları daha sonra veri madenciliğinin ne olduğu ve çalışma prensipleri anlatılacaktır. Daha sonra veri madenciliği teknikleri ile devamsızlık nedenleri ve etki eden faktörler araştırılacaktır.

* Öğr. Gör., Manisa Celal Bayar Üniversitesi, Araştırma Girişimcilik ve Yenilikçilik Koordinatörlüğü.



2. İşyerinde Devamsızlık Kavramı ve Ekonomik Sonuçları

Çalışanın bir mazeret göstermeksizin ve haber vermeden işe gelmemesi işyerinde devamsızlık olarak tanımlanmaktadır (Tütüncü ve Demir, 2003, 152). Farklı bir tanımda ise işgörenin çalışma plan/programına uygun zamanlarda işyerine gelmemesi işyerinde devamsızlık olarak tanımlanmaktadır (Eren, 2007, 267). İşgörenlerin haklı bir mazeret veya sağlık sorunları sebebiyle işyerine gelmemesi durumunun ise devamsızlık olup olmayacağıyla çeşitli görüşler olmakla birlikte; işgörenin hastalanması veya hastalık dışındaki bazı psikolojik, sosyal ve ekonomik sebeplerle mesai saatleri içerisinde işyerinde bulunmamasıyla ilgili bir görüş birliği bulunmaktadır (Bacak ve Yiğit, 2010, 30).

İşyerinde devamsızlık gerek nesnel gerekse kuramsal bakış açısından önemli çalışma alanları içerisinde görülmektedir. Devamsızlık kavramının kuramsal boyutunda; devamsızlık yapan işgörenlerin ortak yönlerinin neler olduğuna değinilirken, nesnel boyutunda; işyerlerine olan mali yükü dikkate alınmaktadır. İşgören devamsızlık yaptığında, yapılması gereken iş yarım kalabilir veya deneyimi yetersiz bir çalışana gördürülebilir (Jex, 2002, 126). İşyerinde yapılan devamsızlıklar çoğunlukla hastalık veya iş göremezlik durumuna bağlı olsa da çalışanlar başka nedenlerle de işe gelmiyor olabilir. İşgörenin yaptığı devamsızlıkların cezalandırılması yerine sağlık politikaları ve sürekliliğin teşvik edilmesi yaygınlaşan bir uygulama olmuştur. Ekonomik Kalkınma ve İşbirliği Örgütü'nün yaptığı çalışmalarda hastalık yardımlarının işsizlik maliyetlerinden 2,5 kat fazla olduğu görülmektedir (Avrupa Yaşam ve Çalışma Koşullarını İyileştirme Vakfı, 2010, 1).

Gerçekleştirilen çalışmalarda, işverenlerin üstlendikleri toplam sağlık ve verimlilikle ilgili maliyetlerin yarısından daha azını tıbbi harcamaların oluşturduğu görülmektedir. Sağlıkla ilgili diğer maddi giderleri ise; devamsızlık, iş göremezlik giderleri, çalışanların tazminatları, personel devir hızı ve verimlilikle ilgili diğer kayıplar oluşturmaktadır (Goetzl vd., 2003, 5). Devamsızlık işyerine maliyeti olan bir işgören davranışıdır. İngiltere'de yapılan bir çalışma işgören devamsızlığının her yıl 8,4 milyar £ maliyeti olduğunu göstermektedir. Devamsızlık; iş akışında ve işgören gücünün planlanmasında aksama ve verimlilik azalışına neden olduğundan kalitenin de düşük olmasını beraberinde getirmektedir ve bu durum işletmeler için ciddi bir mali yük olmaktadır. Bazı ülkeler devamsızlık maliyetinin hesaplanmasında farklı yöntemler geliştirmiş olsalar da bazı ülkeler henüz herhangi bir yöntem tespit edememişlerdir. İşletmelerin verimlilik faktörlerinin değerlendirilmesi için dikkate alınacak bazı araçlar ise Tablo 1 ile gösterilmektedir (Kandemir, 2014, 31).

Tablo 1: Seçilmiş Risk Faktörleri ve Maliyet Değerlendirme Araçları

Risk Faktörü	Sponsor Kurum	Web sayfası
Alkol bağımlılığı	The George Washington University Medical Center	http://www.alcoholcostcalculator.org/roi/
Astım	Agency for Healthcare Research and Quality	http://statesnapshots.ahrq.gov/asthma/
Depresyon Diyabet	American Diabetes Association	www.depressioncalculator.com www.diabetesarchive.net/advocacyand-legalresources/cost-of-diabetes.jsp
Diyabet	Agency for Healthcare Research and Quality	www.ahrq.gov/populations/diabcostcalc
Migren	The Pharmaceutical Research and Manufacturers of America (PhRMA)	www.migrainecalculator.com/welcome.asp
Hareketsizlik	East Carolina University	www.ecu.edu/picostcalc
Tütün Kullanımı	American Health Insurance Plans	www.businesscaseroi.org/roi/default.aspx
Obezite	Business Group on Health	www.businessgrouphealth.org/healthto pics/obesitycostcalculator.cfm

Kaynak: Kandemir, 2014, 68.

İşyeri devamsızlıkları sıfıra indirilemese de sağlık problemlerine yönelik olarak alınacak önlemler hastalık sebebiyle devamsızlık oranını azaltacaktır. Bazı incelemelerde ise insanlarda devamsızlık eğilimi bulunabileceği bu kişilerin yaptıkları işten hoşlanmayan, kişisel ilişkilerinde başarısız, grup içerisinde kaynaşamayan ve yalnız insanlar oldukları görülmüştür. Bu kişilerin işyerinde yaşadıkları tatminsizlik nedeniyle işe gelmekten kaçındıkları ve daha iyi imkânlar bulduklarında işten ayrıldıkları görülmektedir. Bu türdeki devamsızlıkları önleme ise imkânsız olarak görülmektedir (Eren, 2007, 273).

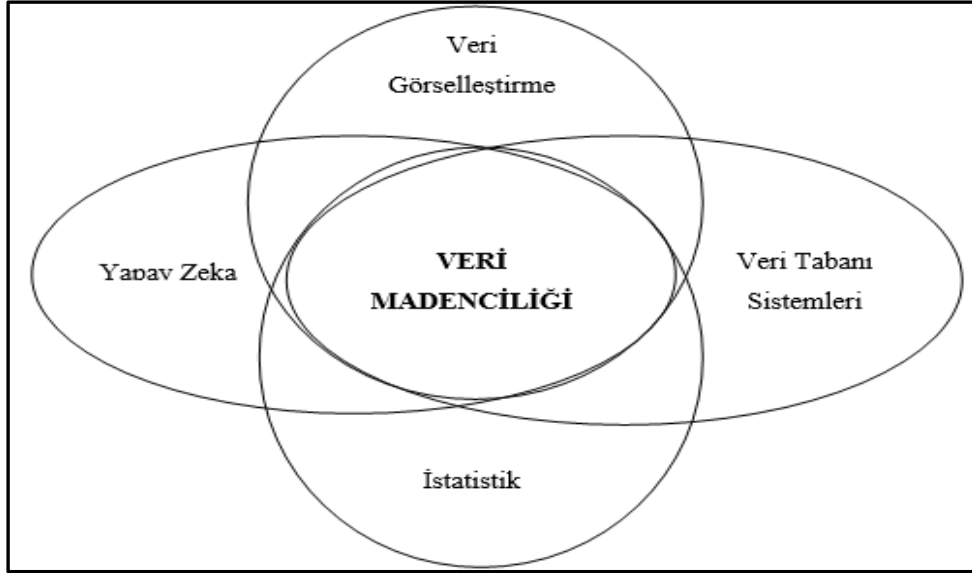
3. Veri Madenciliği ve Çalışma Prensipleri

İçerisinde bulunduğumuz bilgi ve teknoloji çağında, verilerin değeri her geçen gün daha da önem kazanmaktadır. Verilerin günümüzde çok önemli olmasının en temel nedenlerinden biri, hızlı bir şekilde

işlenebilmeleri ve gelişmiş veri analizlerinin yapılabilmesidir. Veri analizine bağlı olarak, daha anlamlı sonuçlar üretilmekte ve veriler daha etkili bir şekilde kullanılmaktadır (Sönmez vd., 2018, 1109).

Veriler yapısal bir formatta ise ham veriden bilgi çıkarma süreci veri madenciliği teknikleri ile gerçekleşmektedir. Büyük veri kümeleri veya veri tabanlarından gizli kalmış ve faydalı olacak bilginin çıkarılmasına olanak tanıyan veri madenciliği teknikleri; istatistik, makine öğrenmesi, veri yönetimi ve veri tabanı, yapay zekâ alanlarının kesişim noktasıdır. Bu alanların hepsi de veri ile ilgilidir ve pek çok ortak noktaları bulunmaktadır (Aytekin vd., 2018, 783).

Şekil 1: Veri Madenciliğinin İlişkide Olduğu Disiplinler



Kaynak: Baykasoğlu, 2005:1.

Genellikle büyük miktardaki veri setinden ilişkilerin aranması ve anlaşılır biçimde yorumlanmasını esasına dayanan veri madenciliğindeki ilişkiler desen veya model olarak tanımlanmaktadır. Bu türdeki veriler genel olarak veri madenciliği dışında farklı amaçlar için toplanır. Bir bankada tutulan müşteri kayıtları veri madenciliği için kayıt altına alınmaz ancak bu veriler üzerinde veri madenciliği teknikleri ile müşterilerin harcama alışkanlıklarının belirlenmesi örnek olarak gösterilebilir (Aytekin vd., 2018, 783).

Bir süreç olarak işleyen veri madenciliğinde sürecin temel unsurunu gerçekleştiren uygulamadır. Hedeflenen sonuca ulaşabilmek için sürecin tüm adımları doğru bir şekilde gerçekleştirilmelidir. Bilgi keşfi süreci olan veri madenciliği aşağıdaki adımlardan meydana gelmektedir (Küçük Yılmaz vd., 2016, 25):

- Verinin temizlenmesi
- Veri bütünlüğünün sağlanması
- Analiz için uygun verilerin belirlenmesi
- Verinin analize uygun hale getirilmesi
- Veri madenciliği algoritmalarının kullanılması
- Örüntülerin tanımlanması
- Bilginin son kullanıcıya sunulması.

Veri madenciliğinde kullanılan modeller tahminleyici ve tanımlayıcı olmak üzere ikiye ayrılırlar. Tahminleyici modeller önceden sonuçları bilinen veriler üzerinden bir model geliştirilmesi ve geliştirilen modelin sonucu bilinmeyen bir veri seti üzerinde uygulanarak sonuçların tahmin edilmesidir. Tanımlayıcı modeller ise; karar vermeye yönelik olarak veri içerisindeki örüntülerin tanımlanmasıdır. Veri madenciliğinde kullanılan temel modeller (Can, 2017, 36);

- Sınıflandırma (classification)
- Kümeleme (clustering)
- Birliktelik kuralları (association rules)

Bu temel modellere ek olarak;

- Tahmin (estimation)
- Öngörme (prediction)
- Zaman serisi analizi (time series analyses)

- Sıralılık keşfi (sequence discovery)

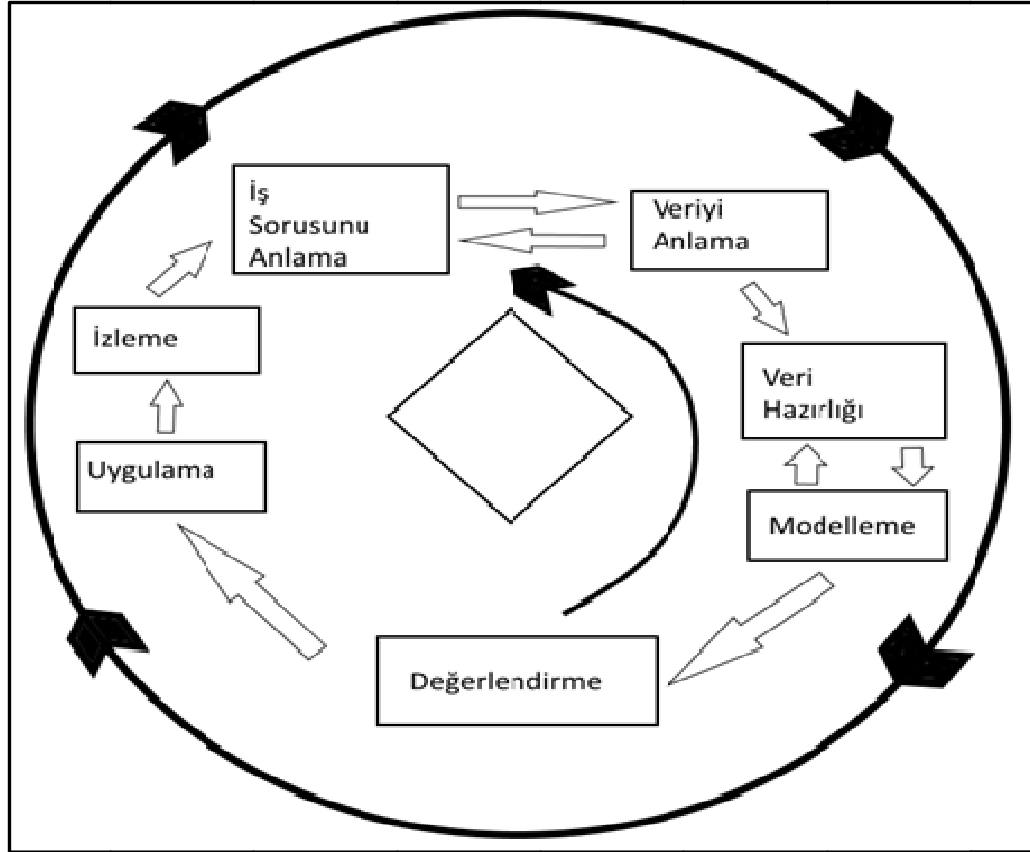
Eğer anlaşılır olabiliyorsa; tahmin edici modeller tanımlayıcı, tanımlayıcı modeller de tahmin edici olarak kullanılabilir (Can, 2017, 36).

4. Veri Madenciliği Teknikleri İle İşyerinde Devamsızlık Nedenlerinin Araştırılması

İşyerinde devamsızlık bilgilerini içeren veri seti Uci Machine Learning web sayfasından elde edilmiştir. Alınan veri setinde 21 farklı nitelik içeren 740 adet veri bulunmaktadır. Veri seti veri madenciliği analizi gerçekleştirmek üzere incelenmiş ve literatürde yaygın olarak kabul gören CRISP-DM sürecine uygun şekilde analiz işlemi gerçekleştirilmiştir (Göral, 2007, 17).

Veri madenciliği uygulamalarının birçok disiplini bir arada barındırması ve farklı uygulama alanlarındaki prosedür çeşitliliği nedeniyle standart bir metodoloji oluşturma aşamasında sorunlarla karşılaşmaktadır. Standardize edilen bir uygulama metodolojisi; teknolojik uygulamaları daha az maliyetli, güvenilirlik düzeyi daha yüksek, daha kullanışlı ve daha hızlı bir hale getirebilir. Bunlar dışındaki bir metodoloji, veri madenciliği uygulamalarını daha kolay uyum sağlayabilen ve anlaşılabilir bir biçime dönüştürecektir (Şekeroğlu, 2010, 23). CRISP-DM (Cross Industry Standard Process for Data Mining) modeli veri madenciliğinde en yaygın kullanılan süreçtir. CRISP-DM modeli; lider veri madenciliği kullanıcı ve tedarikçilerinin oluşturduğu bir birliktelik tarafından geliştirilmiştir (Göral, 2007, 17).

Şekil 2: Çapraz Endüstri Veri Madenciliği Standart Süreci



Kaynak: Ekinci, 2009:14 kullanılarak tekrar düzenlenmiştir.

CRISP-DM süreci temel olarak altı farklı aşamadan oluşmaktadır:

- İş anlama
- Veriyi anlama
- Veriyi hazırlama
- Modelleme
- Değerlendirme
- Gerçekleştirme

Sistem kurulurken en çok dikkat edilecek adımlar ilk dört adımdır. Algoritmanın uygulanması aşamasında karşılaşılabilecek sorunlar, tekrar bu aşamaya dönülerek verilerde yeniden düzenleme



yapılmasına neden olacaktır. Bu durumda veri hazırlama ve model kurma aşamalarında analistin zamanının %50-%85'ini bu işe harcaması gerekmektedir (Ayık vd., 2007, 442).

4.1.İşi Anlama

Çalışmanın bu adımı sürecin ilk adımıdır. Gerçekleştirilecek olan çalışma amacı ve gereksinimlerinin anlaşılması ve daha sonra da bu bilgilerin hedeflenen amacı net olarak ifade etmesi gerekmektedir. Bu aşamada verinin nasıl kullanılacağı konusunda karar verilmesi gerekmektedir. Bu nedenle çalışma sisteminin bu aşamasında veri setiyle ilgili açıklamalar irdelenmiş ve nasıl bir analiz yapılacağı üzerinde durulmuştur. Kümeleme ve sınıflandırma veri madenciliği modellerine uygun olan veri setinde devamsızlığa en çok etkisi olan nedenler ve yaş, cinsiyet, eğitim düzeyi, evcil hayvan beslenip beslenmediği gibi faktörlerin bu sürece etkisi olup olmadığının araştırılmasına karar verilmiştir.

4.2.Veri Anlama

Bu aşamada verinin anlaşılmasına yönelik olarak gerekli incelemeler yapılmıştır. Veri setinde yer alan 21 adet nitelik incelenmiştir. Niteliklerin temsil ettiği sayısal değerlere karşılık gelen ifadeler kontrol edildiğinde veri daha anlaşılır hale gelmiştir. Veriyle ilgili orijinal açıklamalar Şekil 3 ile gösterilmiştir.

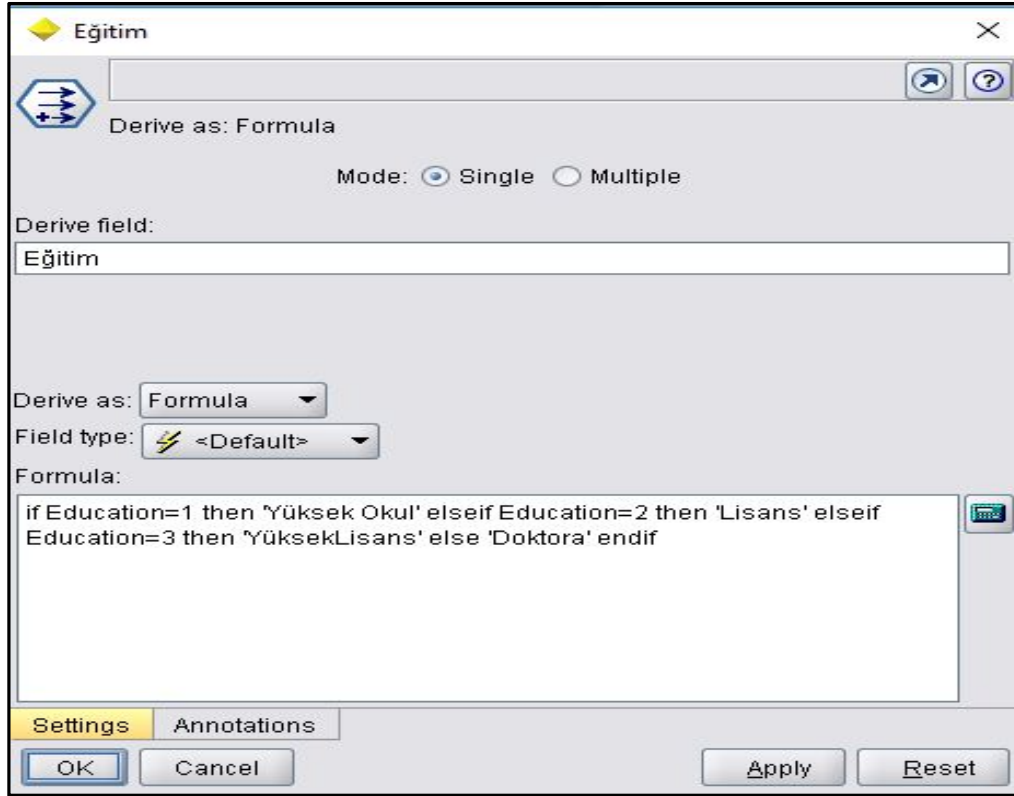
Şekil 3: Veri Setinde Yer Alan Niteliklerin Anlamları

<p>1. Individual identification (ID) 2. Reason for absence (ICD). Absences attested by the International Code of Diseases (ICD) stratified into 21 categories (I to XXI) as follows:</p> <p>I Certain infectious and parasitic diseases II Neoplasms III Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism IV Endocrine, nutritional and metabolic diseases V Mental and behavioural disorders VI Diseases of the nervous system VII Diseases of the eye and adnexa VIII Diseases of the ear and mastoid process IX Diseases of the circulatory system X Diseases of the respiratory system XI Diseases of the digestive system XII Diseases of the skin and subcutaneous tissue XIII Diseases of the musculoskeletal system and connective tissue XIV Diseases of the genitourinary system XV Pregnancy, childbirth and the puerperium XVI Certain conditions originating in the perinatal period XVII Congenital malformations, deformations and chromosomal abnormalities XVIII Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified XIX Injury, poisoning and certain other consequences of external causes XX External causes of morbidity and mortality XXI Factors influencing health status and contact with health services.</p> <p>And 7 categories without (CID) patient follow-up (22), medical consultation (23), blood donation (24), laboratory examination (25), unjustified absence (26), physiotherapy (27), dental consultation (28).</p> <p>3. Month of absence 4. Day of the week (Monday (2), Tuesday (3), Wednesday (4), Thursday (5), Friday (6)) 5. Seasons (summer (1), autumn (2), winter (3), spring (4)) 6. Transportation expense 7. Distance from Residence to Work (kilometers) 8. Service time 9. Age 10. Work load Average/day 11. Hit target 12. Disciplinary failure (yes=1; no=0) 13. Education (high school (1), graduate (2), postgraduate (3), master and doctor (4)) 14. Son (number of children) 15. Social drinker (yes=1; no=0) 16. Social smoker (yes=1; no=0) 17. Pet (number of pet) 18. Weight 19. Height 20. Body mass index 21. Absenteeism time in hours</p>
--

4.3.Veri Hazırlama

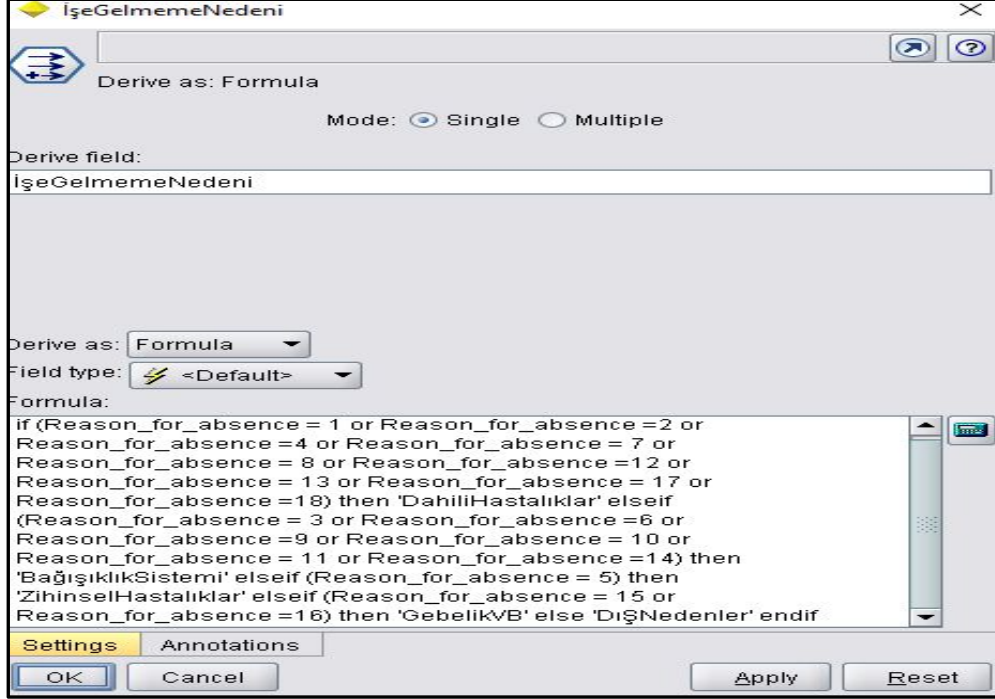
Bu aşamada amaç verinin analize uygun hale getirilmesini sağlamaktır. Bu aşamada eğer veri farklı kaynaklardan elde edilmişse gerekli dönüşümlerin yapılması, veri içerisinde eksiklikler varsa bu eksikliklerin giderilmesi için işlemlerin gerçekleştirilmesi gerekmektedir. Veri hazırlanması adımı verinin hazırlanması, dönüştürülmesi ve özellik seçimi gibi işlemler yapılmaktadır. Çalışma kapsamında kullanılan veri seti incelenmiş ancak herhangi bir eksik veri tespit edilmemiştir. Ancak verilerin tamamının sayısal olarak tutuluyor olmasının analiz sonrası yapılacak yorumlamayı zorlaştıracığı düşünüldüğü için veriler yorumlamaya uygun olarak dönüştürülmüştür. Verilerin dönüştürülmesi için SPSS Clementine programında yapılan dönüşümlerden eğitim bilgisinin dönüşümü Şekil 4 ile örneklendirilmiştir.

Şekil 4: Eğitim Bilgisinin Sayısal-Alfasayısal Dönüştürülmesi



Veri içerisinde yer alan Reason_for_absence bilgisi içerisinde toplamda 28 farklı nitelik yer almaktadır. Bu durumun veri üzerinde analiz yapılmasını zorlaştıracığı ve verinin yorumlanmasını karmaşıklaştıracığı düşünüldüğü için aşağıdaki şekilde gruplandırmaya tabi tutulmuştur.

Şekil 5: Devamsızlık Bilgisinin Sayısal-Alfasayısal Dönüştürülmesi



Burada devamsızlık nedenlerinin gruplandırılması; dâhili hastalıklar (1,2, 4, 7, 8, 12, 13, 17,18), bağıışıklık/ sindirim/ solunum sistemi rahatsızlıkları (3, 6, 9, 10, 11, 14), zihinsel hastalıklar (5), hamilelik (15,16), dış nedenler (19,20,21, 22, 23, 24, 25, 26, 27, 28) şeklinde yapılmıştır.

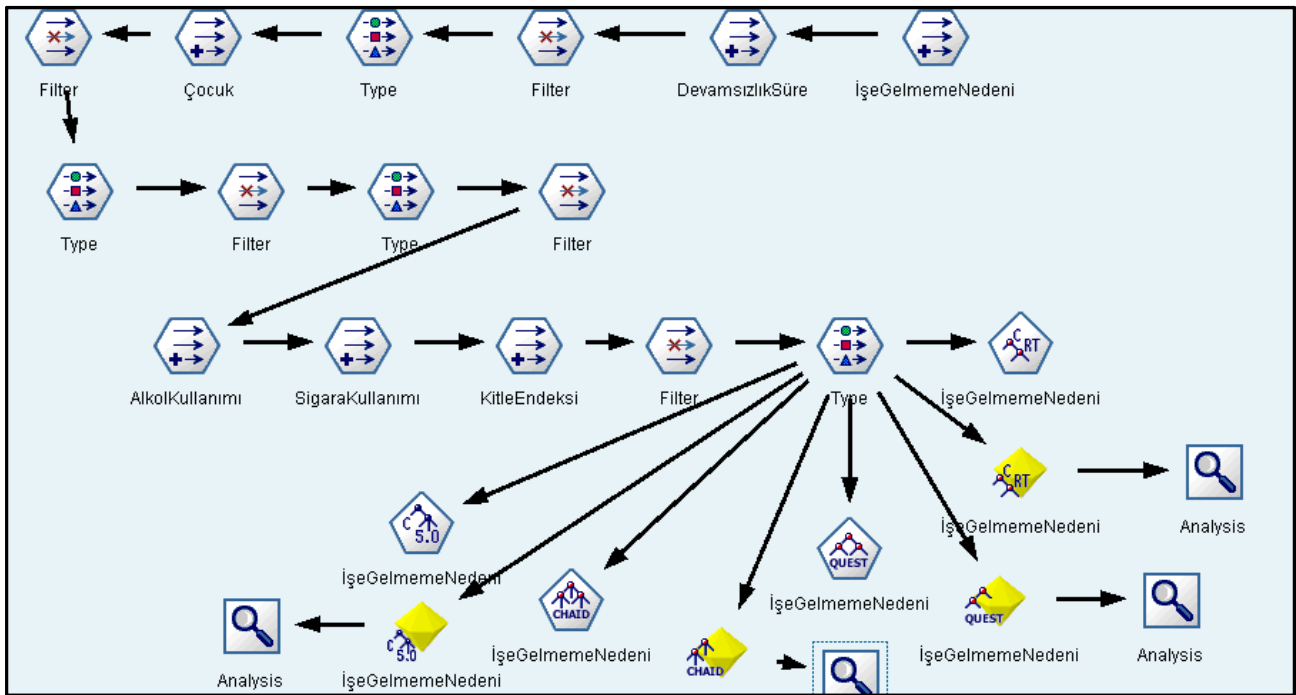
4.4. Modelleme

Elde bulunan veriler veri madenciliği tekniklerinin kullanımına uygun hale getirildikten sonra modelleme aşamasına geçilmiştir. Modelleme aşamasında karar ağacı tekniği tercih edilmiştir. Karar ağacı oluşturulurken veri seti içerisinde yer alan niteliklerin önem seviyeleri kontrol edilmiş ve önem seviyesi en yüksek olan veriler sınıflandırma için kullanılmıştır. Model oluşturulurken veriler rastgele olarak 80-20 oranında ayrılmış verilerin %80'i modelin eğitimi %20'si ise eğitimi yapılan modelin test edilmesi için kullanılmıştır. Modelin eğitim aşamasında kullanılan algoritmaların doğruluk oranlarına bakılarak en yüksek doğruluk oranına sahip algoritma olan C&R Tree algoritması sınıflandırma işlemi için kullanılmıştır.

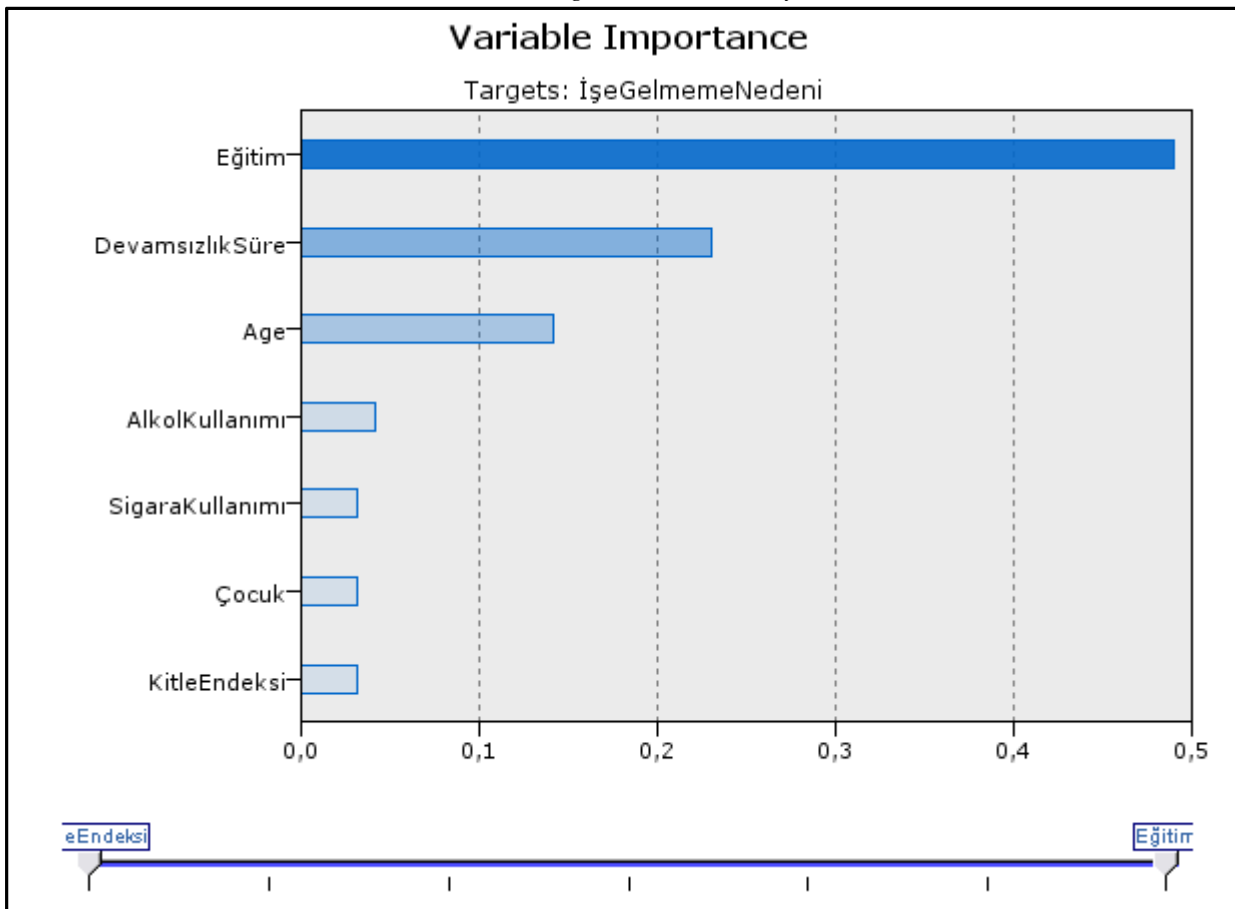
Şekil 6: C&R Tree Algoritmasına Ait Modelin Doğruluk Oranı

Results for output field İşeGelmemeNedeni		
Comparing \$R-İşeGelmemeNedeni with İşeGelmemeNedeni		
Correct	526	71,08%
Wrong	214	28,92%
Total	740	

Şekil 7: Devamsızlık Nedenleri Sınıflandırma Modeli



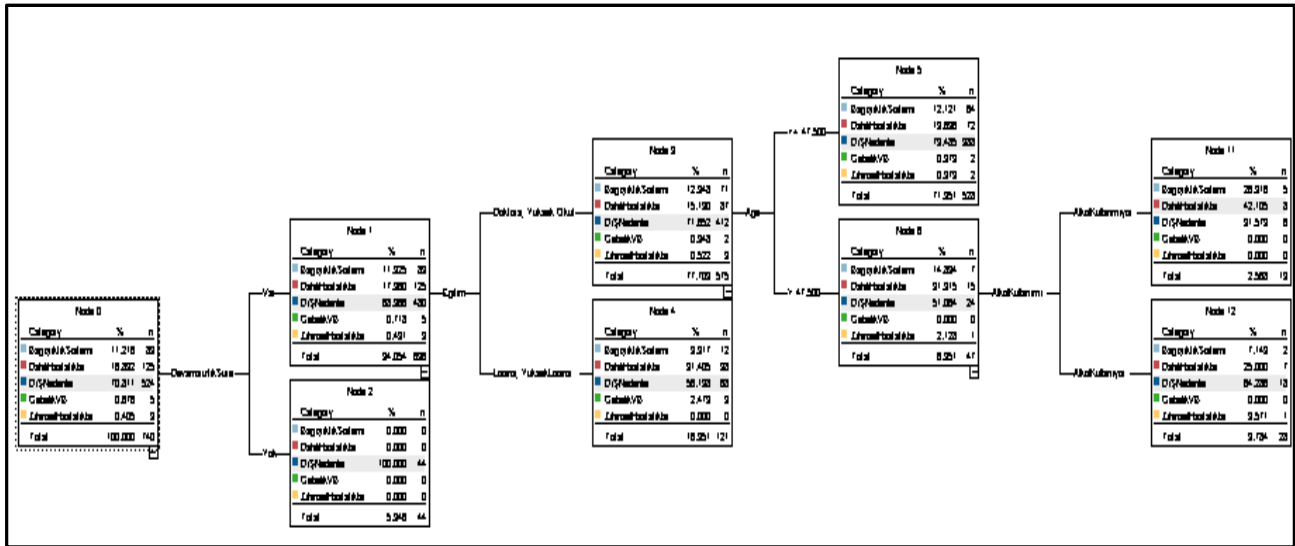
Şekil 8: Değişkenlerin Önem Seviyesi



Şekil 8' de görüldüğü gibi işe devamsızlık nedenlerinin sınıflandırıldığı ve C&R tree algoritmasının kullanıldığı model için en önemli değişkenin Eğitim, Devamsızlık Süre ve Yaş olduğu görülmektedir. Alkol

ve sigara kullanımı, çocuk sahibi olmak ve vücut kitle indeksinin ise daha düşük önem seviyesinde olduğu görülmektedir.

Şekil 9: C&R Tree Algoritması Kullanılarak Oluşturulan Karar Ağacı Genel Yapısı



Şekil 9'da C&R Tree algoritması ile oluşturulan karar ağacının genel yapısı bulunmaktadır.

Sonuç

Küreselleşen dünyada işletmeler büyük bir rekabet içerisinde var olmaya çalışmaktadırlar. Yoğun rekabet ortamında işletmelerin varlıklarını sürdürebilmeleri finansal yapı, pazarlama, satış sonrası destek, kalite yanı sıra personel yetkinliği de oldukça önemlidir. İşletmeler personel devamsızlıkları nedeniyle her yıl maddi zarara uğramaktadırlar. Bu nedenle işletmelerin çözümleri gereken en önemli sorunların başında personel devamsızlıkları gelmektedir. İşyerinde yapılan devamsızlık hem işletme verimliliğini hem de işletmelerin istihdam ve insan kaynakları politikalarını etkilemektedir.

Bu çalışma kapsamında Brezilya Nove de Julho Üniversitesi (Martiniano vd, 2012) tarafından bağışlanan işyeri devamsızlık verileri kullanılarak, personelin işyeri devamsızlığı üzerinde demografik faktörlerin etkileri araştırılmıştır. 740 adet veri içeren veri setinde personelin işe gelmediği ay, gün ve devamsızlık süreleri yanı sıra yaş, eğitim durumu, alkol ve sigara kullanım alışkanlıkları, çocuk sahibi olup olmadığı, boy-kilo ve vücut kitle indeksi gibi pek çok demografik bilgi yer almaktadır. Veri seti veri madenciliği tekniklerine uygun olarak incelenerek veri madenciliği algoritmalarında kullanılacak şekilde tekrar düzenlenmiştir. SPSS clementine programında farklı karar ağacı algoritmaları kullanılarak en yüksek doğruluk oranına sahip karar ağacı algoritması seçilmiş ve veri seti sınıflandırılmıştır.

Gerçekleştirilen analiz sonucunda veri seti içerisinde yer alan demografik verilerden devamsızlık davranışını etkileyen en belirleyici faktörün eğitim olduğu görülmüştür. Bunun yanında yaş ve alkol kullanımı da önemli değişkenler içerisinde yer almaktadır. Devamsızlık nedenleri içerisinde %70 oranında dış faktörler (yaralanma, sağlık kontrolü, kan bağıışı, diş muayenesi gibi) yer almaktadır. 740 personel içerisinde 696 kişi iş yerinde devamsızlık yaparken yalnızca 44 tanesi devamsızlık yapmamıştır. Devamsızlık alışkanlığı gözlenen personelin %77.7 oranında doktora veya yüksek okul düzeyinde eğitim aldığı görülmektedir. Ayrıca genç personelin 47 yaşından daha büyük personele göre çok daha fazla işyerinde devamsızlık yaptığı görülmüştür. 47 yaş üzerinde personelde ise devamsızlık alışkanlığının alkol kullanımıyla birlikte fazlaştığı görülmektedir. İlerleyen çalışmalarda demografik nitelikler faktörlere göre ayrılarak lojistik regresyon analiziyle veri tahminlemesi yapılması planlanmaktadır.

KAYNAKÇA

- Avrupa Yaşam ve Çalışma Koşullarını İyileştirme Vakfı. (2010). İşe Devamsızlık. Dublin, Ireland: The European Foundation for the Improvement of Living and Working Conditions (Eurofound).
- Ayık, Y.Z., Özdemir, A. ve Yavuz, U. (2007). Lise Türü Ve Lise Mezuniyet Başarısının, Kazanılan Fakülte İle İlişkinin Veri Madenciliği Tekniği İle Analizi. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*. 10(2), 441-454.
- Aytekin, Ç., Sütçü, C.S. ve Özfidan, U. (2018). Karar Ağacı Algoritması ile Metin Sınıflandırma: Müşteri Yorumları Örneği. *Uluslararası Sosyal Araştırmalar Dergisi*, 11(55), 782-792.
- Bacak, B. ve Yiğit, Y. (2010). İşe Devamsızlığın Nedenleri, Ekonomik Sonuçları ve Azaltılması İçin Alınması Gereken Önlemler. *Girişimcilik ve Kalkınma Dergisi*, 5(1), 29-44.



- Baykasoğlu, A. (2005). Veri Madenciliği ve Çimento Sektöründe Bir Uygulama. 7. Akademik Bilişim 2005, Gaziantep. 2-4 Şubat 2005.
- Can, Ş. (2017). Veri Madenciliği ve Eğitim Sektöründe Bir Uygulama. Manisa Celal Bayar Üniversitesi Sosyal Bilimler Enstitüsü, Manisa.
- Ekinci, T. (2009). Veri Madenciliği ve Telekomünikasyon Sektöründe Uygulaması. Yayınlanmamış Yüksek Lisans Tezi, Celal Bayar Üniversitesi Sosyal Bilimler Enstitüsü, Manisa.
- Eren, E. (2007). Örgütsel Davranış ve Yönetim Psikolojisi. İstanbul: Beta Yayınları.
- Göral, M. A. (2007). Kredi Kartı Başvuru Aşamasında Sahtecilik Tespiti İçin Bir Veri Madenciliği Modeli. Yayınlanmamış Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
- Jex, S. M. (2002). Organizational Psychology: A Scientist-Practitioner Approach. New York: John Wiley & Sons.
- Kandemir, A. (2014). Hastane Çalışanları Arasında Sağlık Sorunları Nedeniyle Yaşanan İşte Var Olamama Ve İşe Devamsızlık Davranışı İle Maliyetini Etkileyen Faktörlerin İncelenmesi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Küçük Yılmaz, A., Durak, M.Ş., Özdemir, E., Kiracı, K. Ve Tanrıverdi, G. (2016). Eskişehir İlinde Havayolu Yolcu Potansiyelinin Belirlenmesi Üzerine Bir Karar Ağacı Uygulaması. Anadolu Üniversitesi Bilim ve Teknoloji Dergisi B – Teorik Bilimler, 5(1), 22-33.
- Martiniano, A., Ferreira, R. P., Sassi, R. J., ve Affonso, C. (2012). Application of a neuro fuzzy network in prediction of absenteeism at work. In Information Systems and Technologies (CISTI), 7th Iberian Conference on (pp. 1-4). IEEE.
- Sönmez, F., Zontul, M., Kaynar, O. ve Tutar, H. (2018). Anomaly Detection Using Data Mining Methods in IT Systems: A Decision Support Application. Sakarya University Journal of Science, 22(4), 1109-1123.
- Şekeroğlu, S. (2010). Hizmet Sektöründe Bir Veri Madenciliği Uygulaması. Yayınlanmamış Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
- Rygielski, C., Wang, J.-C. ve Yen, D. C.. (2002). Data Mining Techniques for Customer Relationship Management. Technology in Society, 24 (4), 483-502.
- Seyrek, İ.H. ve Ata, H.A. (2010). Veri Zarflama Analizi ve Veri Madenciliği ile Mevduat Bankalarında Etkinlik Ölçümü. BDDK Bankacılık ve Finansal Piyasalar, 4(2), 67-84.
- Tütüncü, Ö. ve Demir, M. (2003). Konaklama İşletmelerinde İnsan Kaynakları Kapsamında İşgücü Devir Hızının Analizi ve Muğla Bölgesi Analizi. Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 5(2), 29-44.