

T.C.
DOKUZ EYLÜL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI
ÜRETİM YÖNETİMİ VE ENDÜSTRİ İŞLETMELERİ
YÜKSEK LİSANS TEZİ

**VERİ MADENCİLİĞİ TEKNİKLERİNİ KULLANARAK
ÜRETİM SÜRESİ TAHMİNİ VE BİR UYGULAMA**

Gökhan BİLEKDEMİR

Danışman
Doç. Dr. Özlem İPEKGİL DOĞAN

2010

YEMİN METNİ

Yüksek Lisans Tezi olarak sunduğum “ Veri Madenciliği Tekniklerini Kullanarak Üretim Süresi Tahmini ve Bir Uygulama ” adlı çalışmanın, tarafımdan, bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın yazıldığını ve yararlandığım eserlerin kaynakçada gösterilenlerden oluştuğunu, bunlara atıf yapılarak yararlanılmış olduğunu belirtir ve bunu onurumla doğrularım.

Tarih

.../.../.....

Gökhan BİLEKDEMİR

İmza

YÜKSEK LİSANS TEZ SINAV TUTANAĞI

Öğrencinin

Adı ve Soyadı : Gökhan BİLEKDEMİR
Anabilim Dalı : İşletme
Programı : Üretim Yönetimi Ve Endüstri İşletmeleri
Tez Konusu : Veri Madenciliği Tekniklerini Kullanarak Üretim Süresi
Tahmini Ve Bir Uygulama
Sınav Tarihi ve Saati :

Yukarıdaki kimlik bilgileri belirtilen öğrenci Sosyal Bilimler Enstitüsü'nün tarih ve sayılı toplantısında oluşturulan jürimiz tarafından Lisansüstü Yönetmeliği'nin 18. maddesi gereğince yüksek lisans tez sınavına alınmıştır.

Adayın kişisel çalışmaya dayanan tezini dakikalık süre içinde savunmasından sonra jüri üyelerince gerek tez konusu gerekse tezin dayanağı olan Anabilim dallarından sorulan sorulara verdiği cevaplar değerlendirilerek tezin,

BAŞARILI OLDUĞUNA	<input type="radio"/>	OY BİRLİĞİ	<input type="radio"/>
DÜZELTİLMESİNE	<input type="radio"/>	OY ÇOKUĞU	<input type="radio"/>
REDDİNE	<input type="radio"/>		

ile karar verilmiştir.

Jüri teşkil edilmediği için sınav yapılamamıştır. ***
Öğrenci sınava gelmemiştir. **

*Bu halde adaya 3 ay süre verilir.
**Bu halde adayın kaydı silinir.
***Bu halde sınav için yeni bir tarih belirlenir.

Tez burs, ödül veya teşvik programlarına (Tüba, Fulbright vb) aday olabilir.	Evet
Tez mevcut hali ile basılabilir.	<input type="radio"/>
Tez gözden geçirildikten sonra basılabilir.	<input type="radio"/>
Tezin basımı gerekliliği yoktur.	<input type="radio"/>

JÜRİ ÜYELERİ

İMZA

..... <input type="checkbox"/> Başarılı <input type="checkbox"/> Düzeltme <input type="checkbox"/> Red
..... <input type="checkbox"/> Başarılı <input type="checkbox"/> Düzeltme <input type="checkbox"/> Red
..... <input type="checkbox"/> Başarılı <input type="checkbox"/> Düzeltme <input type="checkbox"/> Red

ÖZET

Yüksek Lisans Tezi

Veri Madenciliği Tekniklerini Kullanarak Üretim Süresi Tahmini Ve Bir
Uygulama

Gökhan BİLEKDEMİR

Dokuz Eylül Üniversitesi

Sosyal Bilimler Enstitüsü

İşletme Anabilim Dalı

Üretim Yönetimi Ve Endüstri İşletmeleri

Küresel rekabet ortamı nedeni ile üretim işletmeleri kalite, maliyet, teslim zamanı ve esneklik gibi konularda rekabet etmek zorundadır. Üretim sürecinde siparişe göre üretim, stoklara tercih edilir. Bu işletmeler için düşük maliyet ve esneklik sağlar ancak bir siparişin üretim süresini doğru tahmin etmeyi de beraberinde getirir. Teslim süresini kısa vermek müşterileri kazanmayı sağlar ancak genellikle gecikmeyi beraberinde getirir. Teslim süresi uzun tutulduğunda ise teslimatın tamamlanması kolay olur ancak bu seferde müşteri daha kısa teslim süresi veren firmaya kayabilir.

Gelişen teknolojiye uygun olarak bilgisayar sistemlerinin ucuzlaması, kapasitelerinin artması, üretim ortamında sağlanan otomasyon ile üretim verilerin alınmasını kolaylaştırmıştır. Bu durum elde edilen verilerin arasından anlamlı bilgilerin çıkarılması ihtiyacını da beraberinde getirmektedir.

Veri madenciliği çok sayıda veri içerisinde anlamlı bilgiler çıkarılması işlemi olarak tanımlanabilir. Bu tezin amacı üretim süresi tahminini veri madenciliği tekniklerini kullanarak yapmaktır. Bu amaçla ilk bölümde veri madenciliği teknikleri, ikinci bölümde üretim süresi ve tahmin yöntemleri incelenmiştir. Üçüncü bölümde su sayacı üretim süreci incelenmiştir. Elde edilebilen veriler ışığında veri madenciliği teknikleri kullanılarak makinelerin üretim süresi tahmin edilmeye çalışılmıştır.

Anahtar Kelimeler: Veri Madenciliği, Üretim Süresi, Karar Ağaçları

ABSTRACT

Master Thesis

Manufacturing Lead Time Estimation using Data Mining Techniques

Gökhan Bilekdemir

Dokuz Eylül University

Institute of Social Sciences

Department of Business Administration

Program of Production Management and Industrial Administration

Due to the global competition environment, manufacturing firms bound to compete for issues such as quality, cost, delivery, and flexibility. In manufacturing process make-to-order production is preferred instead of making stocks. This yields low cost and flexibility however brings with accurate estimation of lead time. Giving short lead times gains customers however brings with the time delays. When long lead times is given, completion of the delivery will be easy but on the other hand customer tends to go the firm which gives short lead times.

Accordance with the developing technology, due to the cheaper computer systems, increasing capacity, automation in the manufacturing environment facilitates to get manufacturing data. This situation brings with the requirement of extracting meaningful knowledge from that data.

Data mining can be defined as extracting meaningful knowledge from large databases. The purpose of this thesis is estimating lead time using data mining techniques. For this purpose in the first section the data mining techniques, in the second section lead time and lead time estimation methods are investigated. In the third section manufacturing process of water meters is examined and in the light of the given data the lead time of machines is estimated by using data mining techniques.

Keywords: Data Mining, Lead Time, Decision Trees

İÇİNDEKİLER

YEMİN METNİ.....	ii
ÖZET.....	iv
ABSTRACT.....	v
İÇİNDEKİLER.....	vi
KISALTMALAR.....	ix
TABLO DİZİNİ.....	x
ŞEKİL DİZİNİ.....	xi
GİRİŞ.....	1

BİRİNCİ BÖLÜM VERİ MADENCİLİĞİ

1.1 Veri Madenciliği Kavramı.....	2
1.2 Veri Ambarı.....	5
1.2.1 Veri Pazarı (DataMart).....	6
1.2.2 Metadata.....	6
1.2.3 Veri Ambarı Bileşenleri ve Fonksiyonları.....	7
1.3 Bilgi Keşfi Süreci.....	7
1.4 Veri Ön İşleme.....	10
1.4.1 Veri Temizleme.....	10
1.4.2 Eksik Veri.....	10
1.4.3 Verilerin Düzeltilmesi.....	11
1.4.4 Veri Birleştirme.....	12
1.4.5 Veri Dönüştürme.....	12
1.4.6 Veri İndirgeme.....	13
1.5 Veri Madenciliği Modelleri.....	15
1.5.1 Tanımlayıcı Modeller.....	15
1.5.1.1 Kümeleme.....	15
1.5.1.2 Birliktelik Kuralları.....	20
1.5.2 Tahmin Edici Modeller.....	22
1.5.2.1 Sınıflandırma ve Tahmin.....	22

1.5.2.2 Karar Ağaçları.....	24
1.5.2.3 Yapay Sinir Ağları	29
1.5.2.4 Genetik Algoritmalar	32
1.5.2.5 En Yakın Komşu Metodu	33
1.5.2.6 Regresyon Analizi.....	34
1.5.2.7 Naive Bayes Metodu	34
1.6 Veri Madenciliği Uygulama Alanları	35
1.7 Üretimde Veri Madenciliği Uygulamaları.....	36
1.8 Veri Madenciliğinin Avantaj ve Dezavantajları	41

İKİNCİ BÖLÜM

ÜRETİM YÖNETİMİNDE ÜRETİM SÜRESİ

2.1 Üretim/İşlemler Planlaması ve Kontrolü.....	43
2.1 Ön Planlama	44
2.2 Planlama	45
2.3 Kontrol	49
2.4 Üretim Süresi	49
2.4.1 Üretim Yönetiminde Üretim Süresi Tahmininin Önemi.....	54
2.4.2 Üretim Süresi Tahmininde Kullanılan Yöntemler.....	55

ÜÇÜNCÜ BÖLÜM

KARAR AĞACI C4.5 ALGORİTMASINI KULLANARAK SU SAYACI ÜRETİM SÜRESİ TAHMİNİ

3.1 Araştırmanın Amacı	60
3.2 Araştırmanın Yöntemi	60
3.3 Su Sayacı Üretim Süreci	62
3.3.1 Maçahane Bölümünde Üretim Süreci	62
3.3.2 Dökümhane Bölümünde Üretim Süreci.....	63
3.3.3 Kesimhane Bölümünde Üretim Süreci	63
3.3.4 Boyahane Bölümünde Üretim Süreci.....	63
3.3.5 Talaşlı İmalat Bölümünde Üretim Süreci.....	63

3.3.6 Plastikhane Bölümünde Üretim Süreci	63
3.3.7 Mekanizma Montaj Bölümünde Üretim Süreci.....	64
3.3.8 Sayaç Montaj Bölümünde Üretim Süreci	64
3.3.9 Üst gövdenin sıkılması.....	65
3.3.10 Kalibrasyon Bölümünde Üretim Süreci	65
3.3.11 Paketleme Bölümünde Üretim Süreci.....	66
3.4CNC, Maça ve Transfer Makinelerinin Üretim Süresi Tahmini	70
3.4.1 Yazılım.....	71
3.4.2 CNC Makinesi ile ilgili oluşturulan karar ağacı ve sınıflama kuralları.....	73
3.4.3 Maça Makinesi ile ilgili oluşturulan karar ağacı ve sınıflama kuralları.....	76
3.4.4 Transfer Makinesi ile ilgili oluşturulan karar ağacı ve sınıflama kuralları	79
SONUÇ VE ÖNERİLER.....	83
KAYNAKLAR.....	85

KISALTMALAR

CART : Sınıflama ve Regresyon Ağaçları (Classification and Regression Trees)

CNC : Bilgisayarlı Nümerik Kontrol (Computer Numerical Control)

CR : Kritik Oran (critical ratio)

EDD : En Erken Teslim Süresi (earliest due date)

ERP : Kurumsal Kaynak Planlaması

FCFS : İlk Gelen İlk Yapılır (first come first served)

GB : GigaByte

KDD : Bilgi Keşfi Süreci

SPT : En Kısa İşlem Zamanı (shortest processing time)

VM : Veri Madenciliği

WEKA : Waikato Environment for Knowledge Analysis

TABLO DİZİNİ

Tablo 1. Örnek Müşteri Verileri.....	16
Tablo 2 Örnek Veri Seti.....	27
Tablo 3. Bölündükten Sonraki Veri Seti	29
Tablo 4. CNC Makineleri İşlem Süreleri.....	73
Tablo 5. Maça Makinesi İşlem Süreleri	76
Tablo 6 Transfer Tezgahları İşlem Süreleri	79

ŞEKİL DİZİNİ

Şekil 1. Veri Madenciliği	3
Şekil 2. Bilgi Keşfi Süreci.....	9
Şekil 3. Karar Ağacı	28
Şekil 4. Yapay Sinir Ağının Yapısı.....	32
Şekil 5. Geleneksel ve Modern Üretim Sistemlerinde Bilgi Akışı	36
Şekil 6. Üretimde Veri Madenciliği Uygulamalarının Tarihçesi	37
Şekil 7. Su Sayacı Üretim Süreci-1	67
Şekil 8. Su Sayacı Üretim Süreci-2	68
Şekil 9. Su Sayacı Üretim Süreci-3	69
Şekil 10. Su Sayacı Üretim Süreci-4	70
Şekil 11. Verilerin yazılım tarafından alınması.....	72
Şekil 12. Verilerin WEKA'ya aktarılması.....	72
Şekil 13. CNC Makinesi Karar Ağacı	74
Şekil 14. CNC Makinesi Karar Ağacı Kuralı.....	75
Şekil 15. CNC Makinesi Kural Test	76
Şekil 16. Maça Makinesi Karar Ağacı.....	77
Şekil 17. Maça Makinesi Karar Ağacı Kuralı.....	78
Şekil 18. Maça Makinesi Kural Test.....	78
Şekil 19. Transfer Tezgâhı Karar Ağacı	80
Şekil 20. Transfer Makinesi Karar Ağacı Kuralı.....	81
Şekil 21. Transfer Tezgahı Kural Test	82

GİRİŞ

Üretim süresi veya temin süresi, bir sürecin gerektirdiği hammaddelerin siparişinden son montajın tamamlanmasına kadar tüm sıralı aşamaların temin sürelerinin toplamı olarak da ifade edilebilir.

Üretim firmaları tedarik zincirlerinin verimliliğini arttırmaya ve stokları azaltmaya çalıştıkça zamanında teslim etmenin önemi çalışmalarını etkilemektedir. Zorlayıcı kalite standartlarının yanı sıra üreticiler, tedarikçilerinden taahhüt edilmiş teslim süresini ya da ciddi anlamda yaptırımları karşılamasını beklemektedirler.

Teslim süresini kısa vermek müşterileri kazanmayı sağlar ancak genellikle gecikmeyi de beraberinde getirir. Teslim süresi uzun tutulduğunda ise teslimatın tamamlanması kolay olur ancak bu seferde müşteri daha kısa teslim süresi verene kayabilir. Bu açmazı aşmak için satış departmanı fabrikanın durumu hakkında kesin bilgi sahibi olmalıdır.

Bir üretim sisteminin başarı göstergesi siparişin belirtilen zamanda karşılanıp karşılanmadığı ile belli olur. Kestirilen talebi tam zamanında karşılamak için, ardışık birçok karar içeren üretim planlaması yapmak gerekir. Bu kararları verebilmek için bilgi teknolojilerinden faydalanılır.

Veri Madenciliği ile büyük veri setlerinden anlamlı bilgi çıkarılması işlemi olarak tanımlanabilir. Teslim süresini gerçekçi verebilmek için verilen siparişe benzer geçmiş üretim verilerinden tahmin etmek amacı ile VM tekniklerinin kullanımı amaçlanmıştır.

BİRİNCİ BÖLÜM

VERİ MADENCİLİĞİ

1.1 Veri Madenciliği Kavramı

Çok sayıda veri içerisinde anlamlı bilginin çıkarılması işlemi olarak tanımlanan veri madenciliği, bazı yazarlar tarafından şu şekilde tanımlanmıştır:

Anand ve Büchner (2007) veri madenciliğini büyük veri setlerinden belli bir önemi olan, önceden bilinmeyen ve yararlı ve anlaşılır örüntülerin keşfi olarak tanımlamıştır.¹

Fayyad et al (1996) Veri Madenciliğini, kabul edilebilir sayısal verimlilik çerçevesinde, veri içerisinde belirli örüntüleri ortaya koyan veri analizi ve keşif algoritmalarının uygulanması ile oluşan Bilgi Keşfi Sürecinin bir adımı olarak tanımlar.²

Westpal ve Blaxton (1998) veri madenciliği fonksiyonlarını sınıflama, bölümlenme, tahmin ve tanımlama olarak kategorize eder.³

Yevich'ten aktaran Jothishankar (2004), veri madenciliğini işlemciye nasıl sorulacağı bilinmeyen soruların cevaplarını istemenin bir yolu olarak tanımlar.⁴

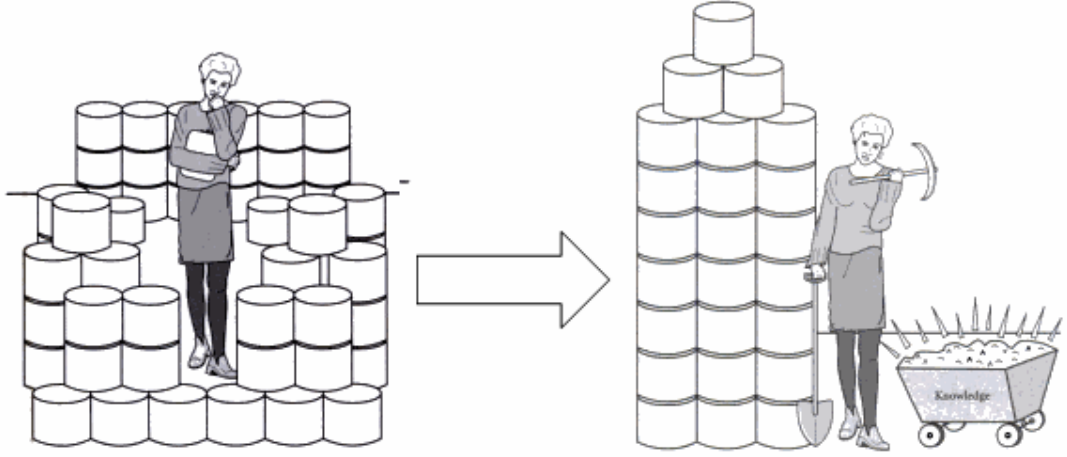
Geleneksel insan analizi büyük hacimli verilerin analizinde sınırlı olduğundan; veri madenciliği, veritabanı büyüklüğü ve karmaşıklığı insanın tek

¹ BÜCHNER, Alex G., Sarabjot S. Anand ve John G. Hughes. "Data Mining in Manufacturing Environments: Goals, Techniques and Applications", **Studies in Informatics and Control**, 1997, s. 1.

² FAYYAD, Usama, Gregory Piatetsky ve Shapiro, Padhraic Smyth. "Knowledge Discovery and Data Mining: Towards a Unifying Framework", 1996, <http://courses.cs.cornell.edu/cs478/2000sp/handoutDBLP>, (20.03.2009), s. 3.

³ WESTPAL, Christopher ve Teresa Blaxton. **Data Mining Solutions Methods and Tools for Solving Real-World Problems**, Wiley Computer Publishing, John Wiley & Sons, Inc., 1998, s. 14.

⁴ JOTHISHANKAR, M. C., Tong (Teresa) Wu, Johnie Roberts, Jiun-Yan Shiau. "Case Study: Applying Data Mining to Defect Diagnosis", **Journal of Advanced Manufacturing Systems**, Vol.3, No.1, 2004, s. 71.



Şekil 1. Veri Madenciliği⁵

başına gözlemlemek için çok büyük olduğu durumlarda veriyi analiz etmek ve yararlı bilgiyi çıkartmak için yöntemler ve araçlar geliştirmeyi amaçlar.⁶

Klasik istatistiksel uygulamalar ve veri madenciliği arasındaki en temel farklılık, veri kümesinin büyüklüğüdür. Bir istatistikçi için 'büyük' veri kümesi birkaç yüz veya bin veri içerir. Veri madenciliği ile uğraşan birileri için ise milyon veya milyarlık veri beklenmeyen bir sayı değildir. Bu tip büyük veri tabanları gerçek hayatta sıkça ortaya çıkmaktadır.⁷

Veri Madenciliği veri yığınlarının içinden işletme yöneticileri için en gerekli olanlarının seçilmesi, düzenlenmesi ve modellenmesi süreçlerini içerir. Bu noktada veri madenciliğini, karar verme mekanizmaları için yeni bilgiler üreten teknikler ve kavramlar bütünü olarak tanımlamak mümkündür.⁸

⁵ HAN, Jiawei ve Micheline Kamber. **Data Mining Concepts and Techniques** , Kitap, ikinci baskı, Morgan Kaufmann, San Francisco, 2006, ss. 4, 5.

⁶ JOTHISHANKAR, s. 70.

⁷ OĞUZLAR, Ayşe. "Veri Ön İşleme", **Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi**, Sayı: 21, 2003, s. 69.

⁸ BİÇEN, Pelin, "Veri Madenciliği: Sınıflandırma ve Tahmin Yöntemlerini Kullanarak Bir Uygulama", (Yayınlanmamış Yüksek Lisans Tezi), Yıldız Teknik Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul, 2002, s. 2.

Veri Madenciliği büyük veri kümelerinden bilginin çıkarılması anlamına gelmektedir. Kavram aslında yanlış tanımlanmaktadır. Nasıl ki büyük kaya ve kumların arasından altın çıkarılması kaya madenciliği ya da kum madenciliği yerine “altın madenciliği” olarak adlandırılıyorsa veri madenciliğinin de “verilerden bilgi madenciliği (knowledge mining from data)” olarak tanımlanması gerekli idi. Ancak bu ifadenin uzun olması ve “bilgi madenciliğinin (knowledge mining)” büyük verilerden bilgi çıkarılmasını tam olarak ifade edememesi nedeni ile içerisinde hem “veri (data)” hem de madencilik (mining) kavramlarını barındıran “veri madenciliği (data mining)” kavramını popüler hale getirmiştir.⁹

Veri Madenciliği, kullanıcının verdiği birçok kararları kapsayan çok sayıda adımı içinde bulunduran interaktif ve tekrarlayıcı (iteratif) bir işlemdir. Problemin tanımlanması ve anlaşılması ile başlar, sonuçların analizi ve bu sonuçları kullanarak bir avantaj kazanmak için stratejik planlama ile sona erer.¹⁰

Veri Madenciliği kavramı akademik çalışmalarda ve ticari pazarda farklı anlamlarda kullanılmaktadır. Akademik çevrelerde Veri Madenciliği kavramı Bilgi Keşfi Sürecinin (Knowledge Discovery from Data- KDD) bir parçası olarak tanımlanır. Bununla birlikte piyasada Veri Madenciliği bu işlemlerin tümünü kapsayacak şekilde tanımlanır.¹¹

Veri Madenciliği yeni, değerli, belli bir önemi olan bilginin büyük veri içerisinden araştırılmasıdır. Veri Madenciliği insan ve bilgisayarların ortak çabasıdır. En iyi sonuç problemin tanımlanmasında insan deneyiminin ve bilgisayarların araştırma yeteneklerinin dengelenmesi ile ortaya çıkar.¹²

Veri Madenciliği işleminin başarısı tasarımcının işe koyduğu, bilgi, yaratıcılık ve enerjiye bağlıdır. Aslında Veri Madenciliği bulmaca çözmek gibidir. Bulmacanın her bir parçası kendi içinde basit yapıdadır. Bununla birlikte bir araya geldiklerinde büyük bir sistemi oluştururlar.¹³

⁹ HAN, s. 5.

¹⁰ HA, Sung Ho ve Sang Chan Park. “Application of data mining tools to hotel data mart on the Intranet for database marketing”, **Expert Systems With Applications** 15, 1998, s. 4.

¹¹ HA, s. 4.

¹² KANTARDZIC, Mehmed. **Data Mining Concepts, Models, Methods, and Algorithms**, Kitap, Wiley – Interscience, USA, 2003, s. 2.

¹³ KANTARDZIC, s. 3.

1.2 Veri Ambarı

Veri ambarı ilişkili verilerin sorgulanabildiği ve analizlerinin yapılabildiği bir depodur. Bir veri ambarı, analizler ve sorgular için kullanılabilir, bütünleşmiş bilgi deposudur. Veri ve bilgiler, üretildiklerinde heterojen kaynaklardan elde edilirler. Veri ambarı, başlangıçta farklı kaynaklardan gelen verinin üzerinde daha etkili ve daha kolay sorguların yapılmasını sağlamaktadır.

Veri ambarları ilk olarak William Inmon tarafından 1980'lerde kullanılmıştır. William Inmon veri ambarlarını karar destek sistemlerini destekleyen, nesne tabanlı, tümleşik, zamana bağımlı, kalıcı veri kümesi olarak tanımlamıştır.¹⁴

Veri ambarları müşteriler, tedarikçiler, ürünler, satışlar gibi birçok kaynaktan gelen verilerle şekillenir. Veri ambarları işletmenin günlük işlemleri yerine karar vericiler için verilerin modellenmesi ve analizi üzerine yoğunlaşır.¹⁵

Veri ambarının asıl amacı, karar destek organlarından gelen verileri bir araya toplamaktır. İyi tasarlanmış bir veri ambarı, veri madenciliği sürecini çok kolaylaştırabilir. Veri ambarı ve veri madenciliğinde esas olan veridir. Verinin kalitesi ve miktarı bu iki sürecin performansını da yakından etkilemektedir. Performansı arttırmak için süreç içinde işlenen verinin kendi içinde bir hiyerarşisi vardır.

Bunlar;

- Operasyonel Veri
- Özet Veri
- Veri tabanı şeması
- Meta Veri
- İş Kuralları olarak sıralanır.¹⁶

Operasyonel Veri, verinin en ham olduğu kısımdır. Bu kısımdaki veriler temel işlemler için kullanılır.

¹⁴ DUNHAM, Margaret H. **Data Mining Introductory and Advanced Topics**, Kitap, Prentice Hall, New Jersey, 2003, s. 35.

¹⁵ HAN, s. 106.

¹⁶ BIÇEN, s. 21.

Özet Veri, soyutlama sürecinin ilk basamağıdır. Operasyonel veriye bağlı bir soyutlama yapıldığından sürekli güncellenmelidir. Kullanıcıya veriye yönelik statik bir görüş açısı kazandırır.

Veritabanı şeması, veriye yönelik bir altyapı şemasının hazırlandığı süreçtir. Veri tipleri, şemalar, tablolar ve indekslerin bulunduğu kısımdır.

Metaveri, fiziksel altyapıdan farklı olarak işletme terimleriyle açıklanmaya çalışılır.

İş kuralları ise veri madenciliği sürecine geçilmeye başlandığı adımdır. Bu adımda ilişkilerin nedenlerinin ve sonuçlarının çıkarımı yapılır ve anlamlı kurallar üretilir.¹⁷

Veri Ambarları, sağlık sektöründen coğrafi bilişim sistemlerine, işletmelerin pazarlama bölümünden üretime, geleceğe dönük tahminler yapmada, sonuçlar çıkarmada ve işletmelerin yönetim stratejilerini belirlemede kullanılmakta olan bir sistemdir. Pahalı bir yatırım maliyeti olsa bile sonuç olarak getirisi (yararı) bu maliyeti kat kat aşmaktadır.

1.2.1 Veri Pazarı (DataMart)

Datamartlar küçük boyutlu (1-10 GB) bölümsel ambarlardır. Datamartlarda verilerine ihtiyaç duyulan bölümlere ait veri ambarıdır. Datamartlar Veri ambarının alt kümesidir.

Organizasyonun (işletmenin) belirli kullanıcıları için ayrılmış ve onlara ait verileri içerir.

1.2.2 Metadata

Veri Ambarının en önemli bileşenlerinden biri metadadır. Veri Ambarında verilerin tanımlandığı kısımdır. Metadata “veri hakkında veri” anlamındadır. Metadata her veri elementinin anlamını, hangi elementlerin hangileriyle nasıl ilişkili olduğunu ve kaynak verisi ile erişilecek veri gibi bilgileri içermektedir.

¹⁷ BİÇEN, s. 22.

1.2.3 Veri Ambarı Bileşenleri ve Fonksiyonları

- Değişik platformlar üzerindeki işletimsel uygulamalara ait verilere erişim ve gerekli verilerin bu platformlardan alınması.
- Alınan verilerin temizlenmesi, tutarlı duruma getirilmesi, özetlenmesi, birleştirme ve birbirleriyle entegrasyonunun sağlanması.
- Dönüştürülen verilerin Veri Ambarı veya datamart ortamına dağıtımı
- Gönderilen verilerin bir veri tabanında toplanması
- Depolanan bilgi ile metadadata bulunan ilgili bilgilerin veri kataloğunda saklanması ve son kullanıcılara sunulması.
- Veri Ambarı veya Datamartta bulunan bilgileri uç kullanıcıların karar destek amaçlı kullanımının sağlanması.

1.3 Bilgi Keşfi Süreci

Veri Madenciliği sistemi binlerce hatta milyonlarca örüntü veya kural oluşturabilir. Bu örüntülerin küçük bir kısmı kullanıcı için ilginç gelecektir. Bir örüntünün ilginç olması için kullanıcı tarafından kolaylıkla anlaşılabilmesi, yeni veriler veya test verisi ile belli bir oranda uygun olabilmesi, işe yarar ve yeni olması gereklidir.

Bir örüntü aynı zamanda kullanıcının onaylayacağı bir hipotezi destekliyorsa da ilginçtir. İlginç örüntü bilgiyi temsil eder.¹⁸

Bilgi, ilginç ve faydalı olması muhtemel olan veriler arasındaki ilişkidir.¹⁹

Bilgi Keşfi Süreci, problemin tanımlanması anlaşılmasından sonra aşağıdaki adımları kapsar:

- Veri Temizleme (Gürültülü ve Tutarsız verilerin kaldırılması)
- Veri Birleştirme (Değişik veri kaynaklarının birleştirilmesi)
- Veri Seçimi (Analiz için gerekli olan verilerin veritabanından alınması)

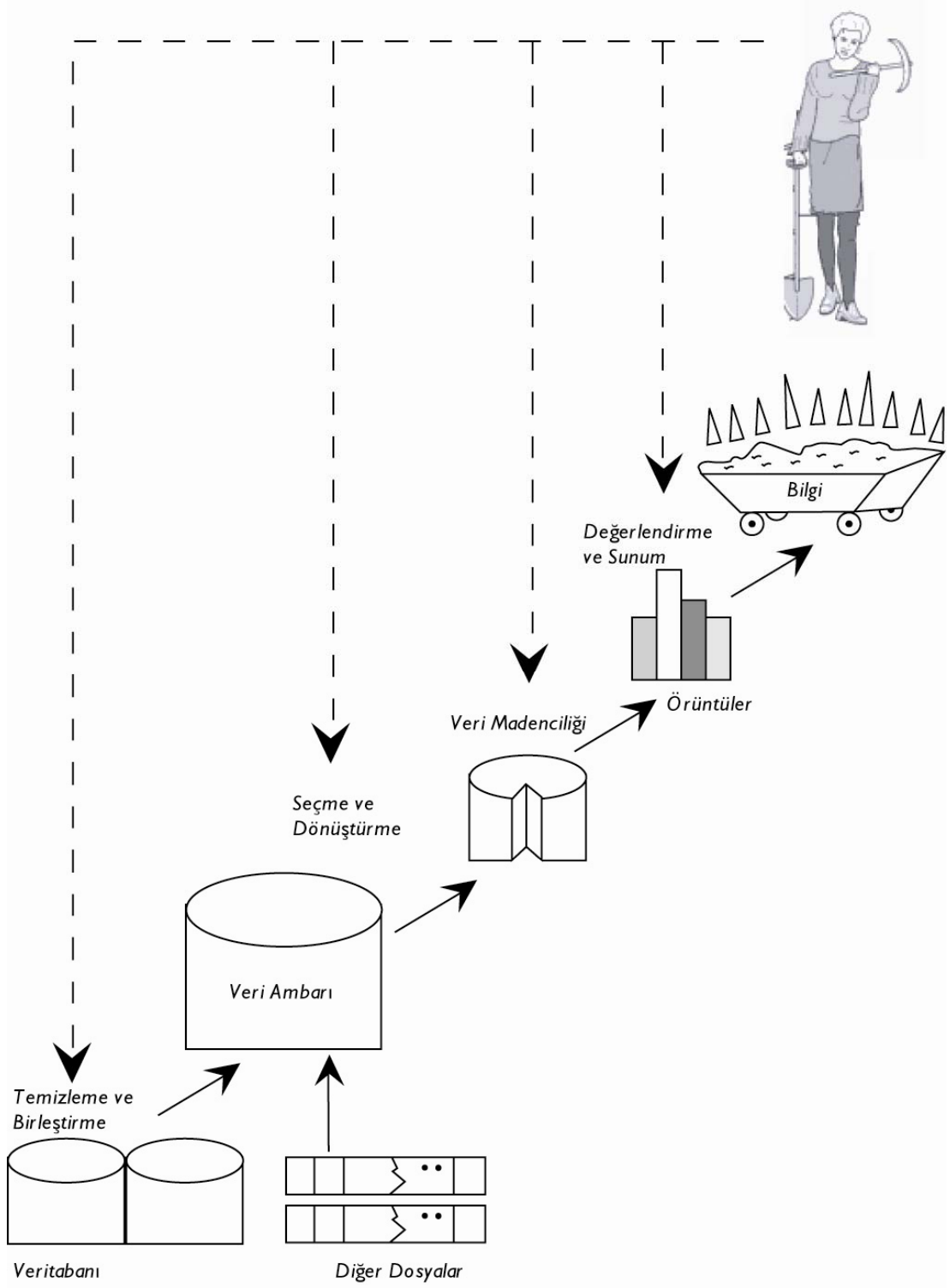
¹⁸ HAN, ss. 27, 28.

¹⁹ YURTSEVER, Ulaş. "Veri Madenciliği ve Uygulaması", (Yayınlanmamış Yüksek Lisans Tezi), Sakarya Üniversitesi Sosyal Bilimler Enstitüsü, Sakarya, 2002, s. 8.

- Veri Dönüştürme (Verilerin VM için uygun hale getirilmesi)
- Veri Madenciliği (veri çiftlerinin çıkarılması için zeki yöntemlerin uygulanma işlemi)
- Örüntü Değerlendirme (Bulunan örüntüler arasından işe yarar, anlamlı olanların tanımlanması)
- Bilgi Sunumu (elde edilen bilginin görselleme ve sunum tekniklerinin kullanılması ile kullanıcıya sunulması)

1'den 4'e kadar olan adımlar veri ön işlemenin değişik formlarıdır.²⁰

²⁰ HAN, s. 7.



Bilgi Keşfi Süreci

Şekil 2. Bilgi Keşfi Süreci²¹

²¹ Han, s. 6.

1.4 Veri Ön İşleme

Günümüzde veritabanları büyük boyutlarından ve birçok farklı kaynaktan gelmelerinden dolayı gürültülü, eksik, tutarsız veriler ile doludur. Verilerin kalitesiz olması veri madenciliğinden elde edilen sonuçların da kalitesiz olmasına yol açabilir.

Veri Madenciliğinde veri kümesinin büyüklüğünden kaynaklanan en fazla zaman alıcı aşama, verilerin ön işlemden geçirilmesi aşamasıdır. Veri Madenciliği uygulamalarında kaynakların %80'i verilerin ön işlemden geçirilmesi ve temizlenmesi süreçleri için harcanmaktadır.²²

Veri ön işleme veri kümelerinin seçilmesi, temizlenmesi, birleştirilmesi ve işlenmesi gibi işlemleri içerir.

Veri ön işleme teknikleri aşağıdaki gibi sıralanabilir:

1.4.1 Veri Temizleme

Veri Temizleme ile gürültülü ve uygunsuz verilerin temizlenme işlemi yapılır. Veri temizleme işlemindeki ilk adım tutarsızlıkları bulmaktır (Discrepancy detection). Tutarsızlıklar çeşitli nedenlerle ortaya çıkabilir. Örneğin sipariş ile ilgili bir veritabanında siparişi veren ile ilgili her bilgi yer almayabilir. Ya da sipariş veritabanı ile müşteri veritabanında tutarsızlıklar olabilir. Veri girişinde insan hataları olmuş olabilir.

Veri temizleme için verinin özelliklerini bilmek gerekir. Buna üstveri (metadata) denir. Bir başka ifade ile verinin içeriği hakkındaki veriye üstveri denir. Her özelliğin alabilecek değerleri ve uzunlukları bilinmelidir.

1.4.2 Eksik Veri

Satış ve müşteri verilerinin analizi yapıldığını düşünelim. Veri setinde müşteri geliri kaydının tutulmadığı durumda bu özellik için eksik veri nasıl tamamlanır?

Herhangi bir değişkene ilişkin eksik değerlerin doldurulması için farklı yollar vardır. Bunlardan bazıları aşağıda kısaca açıklanmaktadır²³:

²² OĞUZLAR, ss. 67, 76.

²³ OĞUZLAR, s. 71.

- Eksik değer içeren kayıt veya kayıtlar atılabilir. Bu metot genellikle sınıf etiketi eksik olduğu durumda yapılır. Bu metot satır birden fazla özellik eksik veri içermediği sürece verimli değildir.
- Eksik veri manüel olarak tamamlanabilir. Bu metot zaman alıcı bir yöntemdir ve büyük veri setlerinde uygulanabilir değildir.
- Eksik veri genel bir sabit ile doldurulur. Bütün eksik veriler “Bilinmiyor”, “∞” gibi aynı sabitle doldurulur. Bu yöntemde Veri Madenciliği yazılımı verilerin hepsinin ortak “Bilinmiyor” verisini içerdiği sonucunu çıkarabilir.
- Aynı sınıfa ait tüm örneklem için değişkenin ortalaması kullanılabilir. Örneğin aynı kredi risk kategorisine giren müşteriler için ortalama gelir değeri eksik değerler yerine kullanılabilir.
- Eksik değer ortalama değer ile doldurulur. Örneğin müşteri geliri eksik ise tüm müşterileri gelirlerinin ortalaması eksik değere yazılır.
- Var olan verilere dayalı olarak en uygun değer kullanılabilir. Burada sözü edilen en uygun değer belirlenmesi için regresyon veya karar ağacı gibi teknikler kullanılabilir. Örneğin yaşı x, eğitim düzeyi y olan bir kişi için ücret durumu, mevcut verilerden yukarıdaki tekniklerden birinin kullanılmasıyla tahmin edilebilir.

Genellikle var olan değerlere dayalı olarak en uygun değer kullanılması tercih edilir.²⁴

Yine bazı durumlarda eksik veri hata olmayabilir. Örneğin kredi ehliyet türünün tutulduğu bir veri tabanında ehliyeti olmayanlar ilgili kısmı boş bırakacaklardır.

1.4.3 Verilerin Düzeltilmesi

Veri temizleme işleminde gürültülü verilerin düzeltilmesi de gerekmektedir. Gürültü ölçülen bir değişkendeki rassal hata veya tutarsızlık olarak tanımlanabilir.²⁵ Veri düzeltme işlemi için çeşitli teknikler bulunmaktadır. Bunlardan bazıları²⁶:

²⁴ HAN, s. 62.

²⁵ HAN, s 62.

- Binning: Binning yöntemleri, küçükten büyüğe veya büyükten küçüğe sıralanmış verileri düzeltmek için kullanılır. Binning yönteminde öncelikle sıralanmış veriler eşit büyüklükteki bin'lere ayrılır. Daha sonra bin'ler, bin ortalamaları, bin medyanları veya bin sınırları yardımıyla düzeltilir.
- Kümeleme (Clustering): Aykırı değerler kümeler ile belirlenebilir
Benzer değerler aynı grup veya küme içinde yer alırken, aykırı değerler kümelerin dışında yer alacaktır.
- Regresyon (Regression): Veriler regresyon ile verilere bir fonksiyon uydurularak düzeltilebilir. Uydurulan fonksiyona uymayan noktalar aykırı değerlerdir.

1.4.4 Veri Birleştirme

Veri Birleştirme çeşitli veri kaynaklarındaki verilerin birleştirilmesi işlemidir. Farklı veritabanlarındaki veriler veri ambarlarında birleştirilirler. Veri birleştirme yapılırken birçok durumun dikkate alınması gereklidir. Örneğin analizci iki farklı veritabanında Abone_No ile Abone_İndeks alanlarının aynı olduğundan emin olması için her alanın adını, anlamını, veri tipini ve aralığını bilmesi gerekir. Buna veriye ilişkin veri anlamında üstveri (metadata) denir.

1.4.5 Veri Dönüştürme

Veri Dönüştürme verinin Veri Madenciliği için uygun hale dönüştürülmesi işlemidir. Veri dönüştürme aşağıdaki adımlardan bir veya birkaçını kapsar:

Verilerin Düzeltilmesi: Verilerdeki gürültünün kaldırılması için kullanılır. Binning, kümeleme ve regresyon tekniklerini kapsar.

Verilerin Kümelenmesi(Aggregation): Verilere kümeleme ya da özetleme işlemleri uygulanır. Örneğin günlük satış verileri aylık ve yıllık miktarları hesaplamak için kümelenir.

²⁶ Han, ss. 63, 64.

Verilerin Genelleştirilmesi (Generalization): Düşük seviyeli veriler yüksek seviyeli kavram ile yer değiştirilir. Örneğin sokak gibi kategorik özellikler şehir ya da ülke gibi yüksek seviyeli kavramlara genelleştirilir.

Özniteliklerin Oluşturulması (Attribute construction): Yeni öznitelikler oluşturulup veri madenciliği işlemini kolaylaştırmak için verilen öznitelik kümesine eklenir.

Verilerin Normalleştirilmesi (Normalization): Veriler belirlenmiş küçük aralıklara ölçeklenir. Değişik veri normalleştirme teknikleri bulunmaktadır. Bunlardan bazıları²⁷:

- Min-Max normalleştirilmesi ile orijinal veriler yeni veri aralığına doğrusal dönüşüm ile dönüştürülürler.
- Z Skor normalleştirmede değişkenin herhangi bir y değeri, değişkenin ortalaması ve standart sapmasına bağlı olarak bilinen Z dönüşümü ile normalleştirilir.
- Ondalık Ölçekleme ile normalleştirmede ele alınan değişkenin değerlerinin ondalık kısmı hareket ettirilerek normalleştirme gerçekleştirilir. Hareket edecek ondalık nokta sayısı, değişkenin maksimum mutlak değerine bağlıdır.

1.4.6 Veri İndirgeme

Veri indirgeme teknikleri, daha küçük hacimli olarak ve veri kümesinin indirgenmiş bir örneğinin elde edilmesi amacıyla uygulanır. Bu sayede elde edilen indirgenmiş veri kümesine veri madenciliği teknikleri uygulanarak daha etkin sonuçlar elde edilebilir.

Veri indirgeme yöntemleri aşağıdaki biçimde özetlenebilir²⁸:

I. Veri Birleştirme veya Veri Küpü (Data Aggregation or Data Cube): Veri birleştirme veya veri küpü yapılacak 2000-2003 yılları için çeyrek dönemlik satış tutarlarından oluşan bir veri kümesinin bulunduğunu varsayalım. Bu yıllar için yıllık satış tutarları tek bir tabloda toplandığında veri birleştirilmesi yapılmış olur. Sonuç

²⁷ OĞUZLAR, s. 73.

²⁸ OĞUZLAR, s. 73.

olarak elde edilen veri kümesinin hacmi daha küçüktür fakat yapılacak analiz için bir bilgi kaybı söz konusu değildir. Veri küpleri ise çok değişkenli birleştirilmiş bilginin saklandığı küplerdir. Örneğin bir firmanın satış tutarları yıllar, satışı yapılan ürünler ve firmanın farklı satış yerleri için aynı küp üzerinde gösterilebilir. Veri küpleri özet bilgiye herhangi bir hesaplama yapmadan hızlı bir biçimde erişilmesini sağlarlar.

II. Boyut indirgeme (Dimension Reduction): Veri madenciliği yapılacak veri kümesi bazen gereksiz olarak yüzlerce değişken içerebilir. Örneğin bir ürünün satışına ilişkin olarak düzenlenen bir veri kümesinde, tüketicilerin telefon numaraları gereksiz bir değişken olarak yer alabilir. Bu tür gereksiz değişkenler elde edilecek örüntüleri kalitesizleştirebileceği gibi veri madenciliği sürecinin yavaşlamasına da yol açacaktır. Gereksiz değişkenlerin elenmesi amacıyla ileri veya geri yönlü olarak sezgisel seçimler yapılabilir. İleri yönlü sezgisel seçimde orijinal değişkenleri en iyi temsil edecek değişkenler belirlenir. Ardından her bir değişken veya değişkenler grubunun, bu kümeye dahil edilip edilmeyeceği sezgisel olarak belirlenir. Geri yönlü sezgisel seçimde ise öncelikle değişkenlerin tüm kümesi ele alınır. Daha sonra gereksiz bulunan değişkenler kümeden dışlanarak, en iyi değişken kümesi elde edilmeye çalışılır. Boyut indirgeme amacıyla kullanılacak bir diğer yöntem ise karar ağaçlarıdır. Karar ağaçları ele alınacak çıktı değişkenini en iyi temsil edecek değişken kümesini verecektir.

III. Veri Sıkıştırma (Data Compression): Veri sıkıştırmada ise orijinal verileri temsil edebilecek indirgenmiş veya sıkıştırılmış veriler, veri şifreleme veya dönüşümü ile elde edilirler. Bu şekilde indirgenmiş veri kümesi, orijinal veri kümesini bir bilgi kaybı olacak biçimde temsil edebilecektir. Bununla beraber bilgi kaybı olmaksızın indirgenmiş veri kümesi elde edilmesine yarayacak bir takım algoritmalar da mevcuttur. Bu algoritmalar bir takım sınırlamalara sahip olduklarından sıkça kullanılamamaktadır. Bununla beraber temel bileşenler analizi gibi yöntemler, bir bilgi kaybına göz yumularak sıkıştırılmış veri kümesi elde edilmesinde kullanışlıdır.

1.4.7.4 Kesikli hale getirme (Discretization)

Kesikleştirme, bazı veri madenciliği algoritmaları yalnızca kategorik değerleri ele aldığından, sürekli verilerin kesikli değerlere dönüştürülmesini içerir. Bu şekilde sürekli verilerin kesikli değer aralıklarına dönüştürülmesiyle elde edilen kategorik değerler, orijinal veri değerlerinin yerine kullanılırlar. Bir kavram hiyerarşisi (concept hierarchy), verilen sürekli değişken için, değişkenin ayrıştırılması olarak

tanımlanabilir. Kavram hiyerarşileri, düşük düzeyli kavramların yüksek düzeyli kavramlarla değiştirilmesiyle verilerin indirgenmesinde kullanılır. Örneğin yaş değişkeni 1-15, 16-40, 40+ olacak biçimde daha yüksek kavram düzeyinde ifade edilebilir. Bu şekilde veri indirgemede detay bilgiler kayboluyorsa da, genelleştirilmiş veriler daha anlamlı olacak, daha kolay yorumlanabilecek ve orijinal verilerden daha düşük hacim kaplayacaktır.²⁹

1.5 Veri Madenciliği Modelleri

Genelde Veri Madenciliğinde kullanılan modeller, tanımlayıcı (descriptive) ve tahmin edici (predictive) olmak üzere iki ana başlık altında incelenmektedir. Tahmin edici modellerde, sonuçları bilinen verilerden hareket edilerek bir model geliştirilmesi ve kurulan bu modelden yararlanılarak sonuçları bilinmeyen veri kümeleri için sonuç değerlerin tahmin edilmesi amaçlanmaktadır. Tanımlayıcı modellerde ise karar vermeye rehberlik etmede kullanılacak mevcut verilerdeki örüntülerin tanımlanması sağlanmaktadır.³⁰

1.5.1 Tanımlayıcı Modeller

Karar vermeye rehberlik etmede kullanılacak mevcut verilerdeki örüntülerin tanımlanması sağlanmaktadır.

1.5.1.1 Kümeleme

Kümeleme analizi ilk olarak Tryon tarafından kullanılmıştır. Denetimsiz öğrenmenin kullanıldığı veri madenciliği ve çok değişkenli sınıflandırma tekniklerindedir. Önceden sınıflandırılmış ve eğitilmiş bir veri kümesi yoktur, buna ek olarak bağımsız ve bağımlı değişken gibi bir ayırım da söz konusu değildir. Kümeleme analizinde yapılan birbirine çok benzeyen nesnelere veya bireyleri aynı gruba yerleştirmektir.³¹

Kümeleme veriyi sınıflara veya kümelere bölme işlemidir. Sınıflamanın aksine kümelemede objenin sınıf etiketi bilinmemektedir.³²

²⁹ OĞUZLAR, ss. 73, 74.

³⁰ ÖZEKEŞ, Serhat. "Veri Madenciliği Uygulaması", (Yayınlanmamış Yüksek Lisans Tezi), Marmara Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, 2002, ss. 6, 7.

³¹ BIÇEN, s. 34.

³² HAN, s. 383.

Aynı sınıf içerisindeki objeler birbirleriyle karşılaştırıldığında benzer özellikler taşıırken diğer sınıflardaki objelerden farklı özellik taşırlar. Bu benzerlikler objenin nitelik değerlerine dayanarak çıkarılır ve uzaklık ölçütleri kullanılır.³³

İki nokta arasında uzaklığın ölçülmesinde en çok kullanılan yöntem Öklid uzaklığıdır.

$$\sqrt{(p_x - q_x)^2} = |p_x - q_x|$$

Kümelemenin belli başlı özelliklerini aşağıdaki şekilde özetleyebiliriz(sınıflandırmanın tersine)³⁴:

- Küme sayısı bilinmemektedir.
- Kümelere ilişkin olası bir bilgi olmayabilir.
- Kümeleme sonuçları dinamiktir.

Örnek: Bir firma müşterilerini benzer özelliklerine göre gruplandırmak ve çıkan sonuca göre pazarlama ve reklam kampanyalarını yönlendirmek istemektedir. Firma yönetiminin nasıl bu gruplar hakkında önceden tanımlanmış bir fikri bulunmuyor olsun. Firmanın müşterileri hakkında elindeki veriler Tablo1'deki gibidir.

Reklam kampanyasının türüne göre müşteri verilerinin hepsinin önemi yoktur. Örneğin çocuk giyim üzerine bir kampanya yürütülecekse çocuklu müşterilere yoğunlaşmak gerekecektir.

Tablo 1. Örnek Müşteri Verileri

Gelir (TL)	Yaş	Çocuk Sayısı	Medeni Hal	Eğitim
2000	35	3	Bekar	Lisans
1250	25	1	Evli	Lisans
1650	40	0	Bekar	Lisans
2500	20	0	Dul	Lisans
1250	25	3	Dul	Lise
5000	60	0	Evli	Lise
7500	30	0	Evli	Yüksek Lisans
15000	45	5	Evli	Yüksek Lisans
7500	50	2	Dul	Lise

³³ YURTSEVER, s. 39.

³⁴ DUNHAM, s. 127.

Tablodaki kümelemede ilk grup küçük çocukları olan eğitimi seviyesi yüksek olan gruptur. İkinci grup ilk gruba göre benzer bir grup olup çocukları yoktur. Üçüncü grup çocuğu olan lise mezunu olan bir gruptur. Son iki grup ise yüksek geliri olan ve en az lise mezunu olan bir gruptur. Yaş ve medeni duruma göre farklı gruplamalar da yapılabilir.

Kümeleme Değişkenleri

Kümeleme analizinde nesnelere sayısal değerler olması gerekir. Değişken tipleri; aralık ölçekli değişkenler, binary değişkenler, kategorik değişkenler, oran ölçekli değişkenler, sıralı değişkenler olabilir ya da bunların kombinasyonu olabilir.

İkili (Binary) değişkenlerde değişken sadece iki durum alabilir. 1 veya 0 olabilir.

Kategorik değişkenler ikiden fazla durum alabilir. Örneğin renk değişkeni: yeşil, sarı, beyaz vs. durumlarını alabilir.

Aralık ölçekli değişkenler lineer ölçeğin sürekli ölçümüdür. Yükseklik, ağırlık, uzunluk gibi örnekler verilebilir.

Sıralama değişkenleri büyüklüğün ne kadar olduğu belli olmadığı halde sıralama yapılabilen kategorik değişkenlerdir. Örnek olarak askeri rütbeler verilebilir (binbaşı, yarbay, albay)

Oran ölçekli değişkenler doğrusal olmayan (nonlinear) ölçeğin pozitif ölçümüdür. Örneğin 1000 TL geliri olan biri, 500 TL geliri olan birine göre iki kat maaş alıyor denilebilir.

Pek çok kümeleme algoritması bulunmaktadır. Kullanılacak olan kümeleme algoritmasının seçimi, veri tipine ve amaca bağlıdır. Genel olarak başlıca kümeleme yöntemleri şu şekilde sınıflandırılabilir³⁵:

- 1 - Bölme yöntemleri (Partitioning methods)
- 2- Hiyerarşik yöntemler (Hierarchical methods)
- 3- Yoğunluk tabanlı yöntemler (Density-based methods)

³⁵ ÖZEKEŞ, Serhat, "Veri Madenciliği Modelleri ve Uygulama Alanları", **İstanbul Ticaret Üniversitesi Dergisi**, Sayı:3, 2003 (dergi), s. 7.

4- Izgara tabanlı yöntemler (Grid-based methods)

5- Model tabanlı yöntemler (Model-based methods)

Bölme yöntemlerinde, n veri tabanındaki nesne sayısı ve k oluşturulacak küme sayısı olmak üzere bölme algoritması n adet nesneyi, k adet kümeye böler ($k \leq n$). Bölme algoritması aynı zamanda aşağıdaki koşulları da sağlamalıdır³⁶:

- Her küme en az bir nesne içermelidir.
- Her nesne sadece bir kümeye ait olmalıdır.

Kümeler tarafsız bölme kriteri olarak nitelendirilen bir kriteri uygun oluşturulduğu için aynı kümedeki nesnelere birbirlerine benzerken, farklı kümedeki nesnelere farklıdır.

En iyi bilinen ve en çok kullanılan bölme yöntemleri k -means yöntemi, k -medoids yöntemi ve bunların varyasyonlarıdır.

k -means yöntemi, ilk önce n adet nesneden rastgele k adet nesne seçer ve bu nesnelere her biri, bir kümenin merkezini veya orta noktasını temsil eder. Geriye kalan nesnelere her biri kendisine en yakın olan küme merkezine göre kümelere dağılırlar. Yani bir nesne hangi kümenin merkezine daha yakın ise o kümeye yerleşir.

Ardından her küme için ortalama hesaplanır ve hesaplanan bu değer o kümenin yeni merkezi olur. Bu işlem tüm nesnelere kümelere yerleşinceye kadar devam eder.

k -means yöntemi, sadece kümenin ortalaması tanımlanabildiği durumlarda kullanılabilir, kullanıcıların k değerini, yani oluşacak küme sayısını belirtme gerekliliği bir dezavantaj olarak görülebilir. Esas önemli olan dezavantaj ise dışarıda kalanlar (outliers) olarak adlandırılan nesnelere karşı olan duyarlılıktır. Değeri çok büyük olan bir nesne, dahil olacağı kümenin ortalamasını ve merkez noktasını büyük bir derecede değiştirebilir. Bu değişiklik kümenin hassasiyetini bozabilir. Bu sorunu gidermek için kümedeki nesnelere ortalamasını almak yerine, kümede ortaya en yakın noktada konumlanmış olan nesne anlamındaki medoid kullanılabilir.³⁷

³⁶ HAN, s. 398.

³⁷ ÖZEKEŞ, ss. 10, 11.

Bu işlem k-medoids yöntemi ile gerçekleştirilir. k-medoids kümeleme yönteminin temel stratejisi ilk olarak n adet nesnede, merkezi temsili bir medoid olan k adet küme bulmaktır. Geriye kalan nesnelere, kendilerine en yakın olan medoide göre k adet kümeye yerleşirler. Bu bölünmelerin ardından kümenin ortasına en yakın olan nesneyi bulmak için medoid, medoid olmayan her nesne ile yer değiştirir. Bu işlem en verimli medoid bulunana kadar devam eder.³⁸

Hiyerarşik yöntemler, verilen veri nesnelere hiyerarşik şekilde ayrıştırır. Hiyerarşik kümeleme yöntemleri, hiyerarşik ayrışmanın aşağıdan yukarıya veya yukarıdan aşağıya doğru olmasına göre agglomerative ve divisive hiyerarşik kümeleme olarak sınıflandırılabilir.³⁹

Agglomerative hiyerarşik kümelemede, hiyerarşik ayrışma aşağıdan yukarıya doğru olur. İlk olarak her nesne kendi kümesini oluşturur ve ardından bu atomik kümeler birleşerek, tüm nesnelere bir kümede toplanıncaya dek daha büyük kümeler oluştururlar.

Divise hiyerarşik kümelemede, hiyerarşik ayrışma yukarıdan aşağıya doğru olur. İlk olarak tüm nesnelere bir kümededir ve her nesne tek başına bir küme oluşturana dek, kümeler daha küçük parçalara bölünürler.

Hiyerarşik yöntemlerin dezavantajı bir adım (birleşme ya da ayrışma) gerçekleştirildikten sonra asla geri dönülemez. Bu yöntem farklı tercihlerin kombinasyonlarının söz konusu olmadığı durumlarda hesaplama maliyetlerini düşürür. Bununla birlikte bu teknik hatalı kararları düzeltemez. Hiyerarşik yöntemlerin kalitesini arttırmak için iki yöntem kullanılmaktadır.

- Chameleon'daki gibi her hiyerarşik bölmede nesne bağlantılarının dikkatli analizi yapılır.
- Agglomerative hiyerarşik kümeleme ile diğer yaklaşımlar birleştirilir. İlk olarak agglomerative algoritması ile nesnelere mikrom kümeler şeklinde kümelendirilir. Daha sonra diğer kümeleme algoritmaları kullanılarak makrokümeler oluşturulur.⁴⁰

³⁸ ÖZEKEŞ, s. 11.

³⁹ HAN, s. 408.

⁴⁰ HAN, s. 399.

Kümelerin gelişigüzel şekillerini ortaya çıkarmak için yoğunluk tabanlı yöntemler geliştirilmiştir. Bunlar kümeleri veri uzayında nesnelerin düşük yoğunluklu bölgelerden ayrılmış yoğunlaşmış bölgeler olarak kabul eder.⁴¹ Yoğunluk tabanlı kümelemede, alanlar veri yoğunluğunun fazla ve az olmasına göre belirlenir. Yoğunluk tabanlı kümeleme algoritmaları arasında DBSCAN, OPTICS, DENCLUE sayılabilir.

Izgara tabanlı kümeleme yaklaşımı çok çözünürlüklü (multiresolution) izgara veri yapısını kullanır. Nesne uzayını bütün kümeleme işlemlerinin gerçekleştirildiği sonlu sayıda hücrelerden oluşmuş izgara yapısı olarak nicelendirir. Bu yaklaşımın en önemli avantajı nesne sayısından bağımsız, her boyuttaki hücre sayısına bağlı hızlı işlem zamanıdır. Izgara tabanlı kümeleme algoritmaları arasında STING, WaveCluster sayılabilir.⁴²

Model tabanlı kümeleme analizinde ise verilen veri ile belli bazı matematiksel modellerin optimize edilmesini amaçlar. Çoğu yöntem verilerin olasılık dağılımlarının karışımı olarak üretildiği varsayımına dayanır.⁴³

Bu algoritmaların yanında genetik algoritmalar, yapay sinir ağları da kümeleme algoritmaları olarak kullanılabilir.

Kümeleme, marketlerde farklı müşteri gruplarının keşfedilmesi ve bu grupların alışveriş örüntülerinin ortaya konması, biyolojide bitki ve hayvan sınıflandırmaları ve işlevlerine göre benzer genlerin sınıflandırılması, şehir planlanmasında evlerin tiplerine, değerlerine ve coğrafik konumlarına göre gruplara ayrılması gibi uygulamalarda kullanılabilir. Kümeleme aynı zamanda Web üzerinde bilgi keşfi için dokümanların sınıflandırılması amacıyla da kullanılabilir.⁴⁴

1.5.1.2 Birliktelik Kuralları

Birliktelik kuralları büyük veri kümeleri arasında birliktelik ilişkileri bulurlar. Örneğin bir ürünün alınırken başka bir ürünün alınması birliktelik kuralıdır. Birliktelik kuralları sıklıkla perakende mağazalarında pazarlama ve tanıtım amacıyla kullanılır.

⁴¹ HAN, s. 418.

⁴² HAN, s. 424.

⁴³ HAN, s. 429.

⁴⁴ ÖZEKEŞ, dergi, s. 7

Birliktelik kuralları, büyük veri kümeleri arasında birliktelik ilişkileri bulurlar. Toplanan ve depolanan verinin her geçen gün gittikçe büyümesi yüzünden, şirketler veritabanlarındaki birliktelik kurallarını ortaya çıkarmak istemektedirler. Büyük miktardaki mesleki işlem kayıtlarından ilginç birliktelik ilişkilerini keşfetmek, şirketlerin karar alma işlemlerini daha verimli hale getirmektedir.

Birliktelik kurallarının kullanıldığı en tipik örnek market sepeti uygulamasıdır. Birliktelik kuralları “sepet analizi” olarak da tanımlanmaktadır. Bu işlem, müşterilerin yaptıkları alışverişlerdeki ürünler arasındaki birliktelikleri bularak müşterilerin satın alma alışkanlıklarını analiz eder. Bu tip birlikteliklerin keşfedilmesi, müşterilerin hangi ürünleri bir arada aldıkları bilgisini ortaya çıkarır ve market yöneticileri de bu bilgi ışığında daha etki satış stratejileri geliştirebilirler.

Bir yerel bir market senaryosu üzerinden örnek verelim. Örneğin insanların belirli günlerde belirli ürünleri almaya yönelik eğilimleri olsun. Mesela cumartesi akşamları bira, pazar günleri süt ve meyve suyu gibi belirli günlerde belirli ürünleri almaya eğilimli olsun. Örneğin bir müşteri süt satın alıyorsa, aynı alışverişte sütün yanında ekmek alma olasılığı nedir? Bu tip bir bilgi ışığında rafları düzenleyen market yöneticileri ürünlerindeki satış oranını arttırabilirler. Örneğin bir marketin müşterilerinin süt ile birlikte ekmek satın alan oranı yüksekse, market yöneticileri süt ile ekmek raflarını yan yana koyarak ekmek satışlarını arttırabilirler.

Birliktelik kuralının matematiksel modeli Agrawal, Imielinski ve Swami tarafından 1993 yılında sunulmuştur. Bu modelde, $I = \{i_1, i_2, \dots, i_m\}$ kümesine “ürünler” adı verilmektedir. D, veri bütünlüğündeki tüm hareketleri, T ise ürünlerin her bir hareketini simgeler. TID ise, her harekete ait olan tek belirteçtir.

Birliktelik kuralı şu şekilde tanımlanabilir;

$$A_1, A_2, \dots, A_m \rightarrow B_1, B_2, \dots, B_n$$

Bu ifadede yer alan, A_i ve B_j , yapılan iş veya nesnelere dir. Bu kural, genellikle

“ A_1, A_2, \dots, A_m ” iş veya nesnelere meydana geldiğinde, sık olarak “ B_1, B_2, \dots, B_n ” iş veya nesnelere aynı olay veya hareket içinde yer aldığını belirtir.

Birliktelik kuralı, kullanıcı tarafından minimum değeri belirlenmiş destek ve güvenilirlik eşik değerlerini sağlayacak biçimde üretilir. Bir ürün kümesindeki destek, D ile ifade edilen tüm hareketler içinde ilgili ürün kümesini içeren hareketlerin yüzdesidir. A ve B ürün kümelerinin, birliktelik kuralı “ $A \rightarrow B$ ” olarak gösterilirse, destek aşağıdaki gibi tanımlanır.

$$\text{Destek } (A \rightarrow B) = (\text{A ve B'nin bulunduğu satır sayısı}) / (\text{toplam satır sayısı})$$

$A \square B$ birliktelik kuralının güven değeri ise, A'yı içeren hareketlerin B'yi de içermeye yüzdesidir. Örneğin, bir kural % 85 güvenilirliğe sahip ise, A'yı içeren ürün kümelerinin % 85'i B'yi de içermektedir. İşe bağlı veri satırları verilmiş ise,

$(A \rightarrow B)$ güveni aşağıdaki gibi tanımlanır.

$$\text{Güven } (A \rightarrow B) = (\text{A ve B'nin bulunduğu satır sayısı}) / (\text{A'nın bulunduğu satır sayısı})$$

Güven değerinin % 100 olması durumunda, kural bütün veri analizlerinde doğrudur ve bu kurallara “kesin” denir.

Birliktelik kuralına ilişkin olarak geliştirilen bazı algoritmalarından en çok bilineni Apriori algoritmasıdır.

Birliktelik kurallarını bulmak için genellikle problem iki parçaya bölünür. Önce sık tekrarlanan öğeler bulunur. Bu öğelerin her biri en az, önceden belirlenen minimum destek sayısı kadar sık tekrarlanırlar. Sonra sık tekrarlanan öğelerden güçlü birliktelik kuralları oluşturulur. Bu kurallar minimum destek ve minimum güven değerlerini karşılamalıdır.

1.5.2 Tahmin Edici Modeller

Tahmin edici modellerde sonuçları bilinen test verilerinden hareketle geliştirilen model, yeni veri setlerine uygulanarak sonuç tahmin edilmeye çalışılır.

1.5.2.1 Sınıflandırma ve Tahmin

Sınıflandırma ve tahmin, önemli veri sınıflarını tanımlamada ya da gelecek trendleri tahmin etmede kullanılacak veri analizinin iki formudur. Sınıflandırma

kategorik (discrete, sıralanmamış) tanımları tahmin eder, tahmin modelleri ise sürekli fonksiyonları tahmin eder.⁴⁵

Sınıflama kuralı 3 adımdan oluşur.

Adım1 : Bulunan model eğitilen veriye dayanır, her veri önceden belli bir sınıf etiketine atanır. Bu verileri örnek veya obje denir. Eğitim kümesini oluşturan verilere eğitim örneği denir. Rastgele seçilirler. Sınıf etiketleri bilindiği için bu adıma denetli öğrenme “supervised learning” denir. Sınıf etiketlerinin, sınıf sayılarının bilinmediği durumlara denetimsiz öğrenme –unsupervised learning- denir.

Adım2 : Model sınıflama için kullanılır. Modelin tahmin gücü araştırılır. Test örneği rastgele seçilir ve öğrenme kümesinden bağımsızdır. Verilen deney kümesi üzerinde bir modelin doğruluğu doğru olarak sınıflanan deney örneklerinin yüzdesi ile ölçülür. Her deney örneği için bilinen sınıf etiketi modelin tahmin ettiği etiketle karşılaştırılır. Eğer modelin doğruluğu sadece öğrenme kümesine bakarak belirlenirse model verileri ezberler.

Adım3 : Eğer modelin doğruluğu yeterli görülürse model etiketi bilinmeyen verileri sınıflamakta kullanılır.⁴⁶

Sınıflandırma ve Tahmin modellerinde kullanılan başlıca teknikler şunlardır:

1. Karar Ağaçları
2. Yapay Sinir Ağları
3. Genetik Algoritmalar
4. K-En Yakın Komşu Metodu
5. Regresyon Analizi
6. Naive- Bayes Metodu

⁴⁵ HAN, s. 285.

⁴⁶ YURTSEVER, ss. 37, 38.

1.5.2.2 Karar Ağaçları

1970'lerin sonu 1980'lerin başında J.Ross Quinlan ID3 algoritması olarak bilinen karar ağacı algoritmasını geliştirmiştir. Quinlan daha sonra yeni danışmanlı öğrenme algoritmaları için ölçü olacak olan C4.5 algoritmasını sunmuştur.⁴⁷

Karar ağacı, örnekleri ağacın kökünden yapraklarına doğru sıralayarak sınıflandırır. Ağaçtaki her düğüm bazı değişkenleri test eder ve her dal bu değişkenin alabileceği değerlerden birine uygun düğümü gösterecek şekilde aşağıya doğru iner. Karar ağacı algoritması, ağacın kökünde hangi değişken ile test edilmesi gerektiği sorusu ile başlayarak yukarıdan aşağıya doğru ağacı oluşturur.

Karar kümelerini betimleyen, ağaç şeklinde yapılardır. Bu kararlar veri kümelerinin sınıflandırılmasıyla ilgili kuralları oluşturur. Bazı özel karar ağaçları, CART ve CHAID metotlarıdır.⁴⁸

Karar ağaçları, veri madenciliğinde kuruluşlarının ucuz olması, yorumlanmalarının kolay olması, veritabanı sistemleri ile kolayca entegre edilebilmeleri ve güvenilirliklerinin iyi olması nedenleri ile sınıflama modelleri içerisinde en yaygın kullanıma sahip tekniktir.

Karar ağacı, adından da anlaşılacağı gibi bir ağaç görünümünde, tahmin edici bir tekniktir. Ağaç yapısı ile kolay anlaşılabilen kurallar yaratabilen, bilgi teknolojileri işlemleri ile kolay uyum sağlayabilen en popüler sınıflama tekniğidir.

Karar ağacı karar düğümleri, dallar ve yapraklardan oluşur. Karar düğümü, gerçekleştirilecek testi belirtir. Bu testin sonucu ağacın veri kaybetmeden dallara ayrılmasına neden olur. Her düğümde test ve dallara ayrılma işlemleri ardışık olarak gerçekleşir ve bu ayrılma işlemi üst seviyedeki ayrımlara bağımlıdır. Ağacın her bir dalı sınıflama işlemi tamamlamaya adaydır. Eğer bir dalın ucunda sınıflama işlemi gerçekleşemiyorsa, o daim sonucunda bir karar düğümü oluşur. Ancak daim sonunda belirli bir sınıf oluşuyorsa, o dalın sonunda yaprak vardır. Bu yaprak, veri üzerinde belirlenmek istenen sınıflardan biridir. Karar ağacı işlemi kök düğümünden başlar ve yukarıdan aşağıya doğru yaprağa ulaşana dek ardışık düğümleri takip ederek gerçekleşir.

⁴⁷ HAN, s. 292.

⁴⁸ YURTSEVER, s. 42.

Karar ağacı tekniğini kullanarak verinin sınıflanması iki basamaklı bir işlemdir.

İlk basamak öğrenme basamağıdır. Öğrenme basamağında önceden bilinen bir eğitim verisi, model oluşturmak amacıyla sınıflama algoritması tarafından analiz edilir. Öğrenilen model, sınıflama kuralları veya karar ağacı olarak gösterilir. İkinci

Basamak ise sınıflama basamağıdır. Sınıflama basamağında test verisi, sınıflama kurallarının veya karar ağacının doğruluğunu belirlemek amacıyla kullanılır. Eğer doğruluk kabul edilebilir oranda ise, kurallar yeni verilerin sınıflanması amacıyla kullanılır.

Test verisine uygulanan bir modelin doğruluğu, yaptığı doğru sınıflamanın test verisindeki tüm sınıflara oranıdır. Her test örneğinde bilinen sınıf, model tarafından tahmin edilen sınıf ile karşılaştırılır. Eğer modelin doğruluğu kabul edilebilir bir değer ise model, sınıfı bilinmeyen yeni verileri sınıflama amacıyla kullanılabilir.

Karar Ağacı Oluşturma Algoritması

Karar ağacı algoritmasında giriş değerleri:

- Eğitim Verisi,
- Özellikler Listesi
- Özellik Seçim Metodudur.

Bir eğitim verisi kullanılarak karar ağacı yaratılırken şu basamaklar izlenir⁴⁹:

- Ağaç, eğitim verisini temsil eden tek bir düğüm olarak başlar
- Eğer veriler aynı sınıftansa düğüm yaprak haline gelir ve o sınıf ile etiketlenir.
- Aksi takdirde verilere sınıflara en iyi şekilde ayırarak alanı seçmek için bilgi kazancı olarak adlandırılan entropi tabanlı bir ölçüm kullanılır.
- Bu test alanındaki her bir öznelik(attribute) için dal yaratılır ve buna göre tüm veriler bölümlere ayrılır.

⁴⁹ ÖZEKEŞ, s. 24.

- Algoritma her bölümdeki veriler için karar ağacı oluşturmak amacı ile bu işlemi tekrarlar.
- Bu tekrarlı dallanmalar şu durumlardan biri doğru olunca sonlanır.
 - a) Verilerin düğümdeki tüm veriler aynı sınıfa ait olduğunda
 - b) Bölünecek başka bir öznitelik kalmadığında

D sınıf etiketli eğitim seti olmak üzere, m veri setindeki sınıf etiketinin sahip olduğu birbirinden farklı öznitelik, $C_i(i=1,2,3,\dots,m)$ olarak tanımlanabilir. $C_{i,D}$ veri setinin içinde C_i sınıfında tanımlı veriler, $|D|$ D veri setinin eleman sayısı, $|C_{i,D}|$ veri setinin eleman sayısıdır⁵⁰

Her öznitelik, eğitim örneklerinin sınıflandırmasına karar vermek için istatistiksel test kullanılarak değerlendirilir. En iyi öznitelik seçilir ve ağacın kök düğümünde test için kullanılır. Kök düğümünün dal sayısı, seçilmiş olan özniteliğin alabileceği değer sayısına göre değişir. Karar ağacı algoritmasında ana seçim ağaçtaki her düğümde hangi özniteliğin seçileceğidir. Öznitelik değerinin ölçüsü bilgi kazancı adı verilen istatistik bir değer ile belirlenir. Bilgi Kazancı için entropi adı verilen bir tanımlama kullanılır.

$$\text{Info}(D) = - \sum_{i=1}^m p_i \log_2(p_i)$$

Bir özniteliğin etkinlik ölçüsü bilgi kazancı terimi ile kullanılmaktadır. A değişkeninin beklenen bilgi kazancı $\text{Info}_A(D)$ aşağıdaki formül ile tanımlanır.

$$\text{Info}_A(D) = \sum_{j=1}^v \frac{|D_j|}{|D|} \times \text{Info}(D_j)$$

Bilgi kazancı Gain (A), sınıfların dağılımına bağlı olan orijinal bilgi gereksinimi ile A özniteliğinin beklenen bilgi kazancı $\text{Info}_A(D)$ arasındaki farktır.

⁵⁰ HAN, s. 297.

$$\text{Gain}(A) = \text{Info}(D) - \text{Info}_A(D)$$

Örnek⁵¹: Tablo.2 bir bilgisayar firmasının müşteri veritabanından alınan D eğitim verilerini göstermektedir. Bu örnekte her özneliğin değeri kesiklidir. Sınıf etiketinin (“bilgisayar alır”) “Evet” ve “Hayır” olmak üzere aldığı iki değer vardır. Dolayısıyla m=2 olmaktadır. C₁’in “Evet”e, C₂’nin “Hayır”a karşılık geldiği kabul edilirse eğitim verilerinde “Evet”e karşılık gelen 9 adet, “Hayır”a karşılık gelen 5 adet veri vardır. Bölme kriterini bulmak için her özelliğin bilgi kazancını hesaplamak gereklidir. D verilerini sınıflandırmak için beklenen bilgi hesaplamak için Entropi değeri:

$$\text{Info}(D) = -\frac{9}{14} \log_2\left(\frac{9}{14}\right) - \frac{5}{14} \log_2\left(\frac{5}{14}\right) = 0.940 \text{ bits. Olarak hesaplanır.}$$

“Yaş” özelliğinin bilgi kazancını bulmak için “Yaş” kategorisi için “Evet” ve “Hayır” dağılımına bakmak gereklidir. Örnekte “genç” özellik değeri için 2 “Evet” ve 3 “Hayır” sınıfı vardır. “Orta yaş” özellik değeri için 4 “Evet” ve 0 “Hayır”, “Yaşlı” özellik değeri için 3 “Evet” ve 2 “Hayır” kaydı bulunmaktadır. Tablo.3 bölümlenmeden sonraki veri setini göstermektedir.

Tablo 2 Örnek Veri Seti

sıra	yaş	gelir	Öğrenci	kredi notu	Sınıf: bilgisayar alır
1	genç	yüksek	hayır	vasat	Hayır
2	genç	yüksek	hayır	mükemmel	Hayır
3	orta yaş	yüksek	hayır	vasat	Evet
4	yaşlı	orta	hayır	vasat	Evet
5	yaşlı	düşük	evet	vasat	Evet
6	yaşlı	düşük	evet	mükemmel	Hayır
7	orta yaş	düşük	evet	mükemmel	Evet
8	genç	orta	hayır	vasat	Hayır
9	genç	düşük	evet	vasat	Evet
10	yaşlı	orta	evet	vasat	Evet
11	genç	orta	evet	mükemmel	Evet
12	orta yaş	orta	hayır	mükemmel	Evet
13	orta yaş	yüksek	evet	vasat	Evet
14	yaşlı	orta	hayır	mükemmel	Hayır

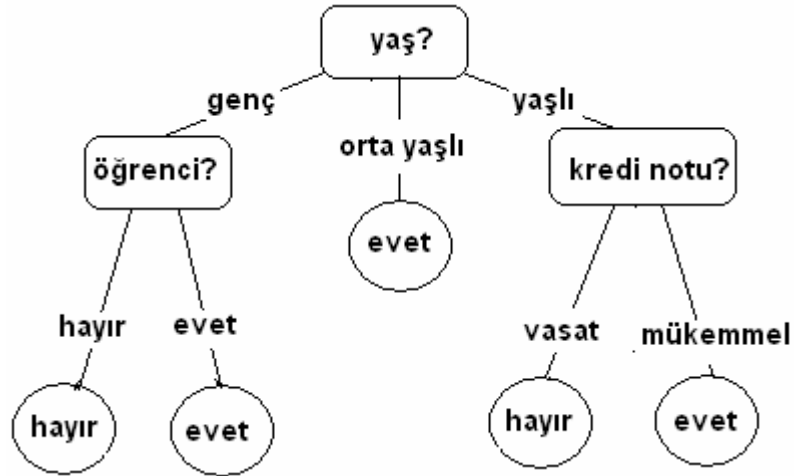
⁵¹ HAN, s. 298.

“Yaş” özellik değeri için beklenen bilgi kazancı:

$$\begin{aligned} \text{Info}_{age}(D) &= \frac{5}{14} \times \left(-\frac{2}{5} \log_2 \frac{2}{5} - \frac{3}{5} \log_2 \frac{3}{5} \right) \\ &+ \frac{4}{14} \times \left(-\frac{4}{4} \log_2 \frac{4}{4} - \frac{0}{4} \log_2 \frac{0}{4} \right) \\ &+ \frac{5}{14} \times \left(-\frac{3}{5} \log_2 \frac{3}{5} - \frac{2}{5} \log_2 \frac{2}{5} \right) \\ &= 0.694 \text{ bits. olarak hesaplanır.} \end{aligned}$$

Bu durumda “Yaş” özelliği için kazanç değeri:

$$\text{Gain}(age) = \text{Info}(D) - \text{Info}_{age}(D) = 0.940 - 0.694 = 0.246 \text{ bits. olarak bulunur.}$$



Şekil 3. Karar Ağacı

Tablo 3. Bölündükten Sonraki Veri Seti

YAŞ = Genç			
gelir	Öğren	kredi notu	Sınıf
yüksek	Hayır	Vasat	Hayır
yüksek	Hayır	mükemmel	Hayır
orta	Hayır	Vasat	Hayır
düşük	Evet	Vasat	Evet
orta	Evet	mükemmel	Evet

YAŞ =Yaşlı			
Gelir	öğrenc	kredi notu	Sınıf
Orta	hayır	vasat	Evet
Düşük	evet	vasat	Evet
Düşük	evet	mükemmel	Hayır
Orta	evet	vasat	Evet
Orta	hayır	mükemmel	hayır

YAŞ =Orta Yaşlı			
gelir	Öğrenc	kredi notu	Sınıf
yüksek	Hayır	vasat	Evet
düşük	Evet	mükemmel	Evet
Orta	Hayır	mükemmel	Evet
Yüksek	Evet	vasat	Evet

Aynı şekilde diğer özellikler için de kazançlar;

- Kazanç(Gelir)=0.029
- Kazanç(Öğrenci)=0.151.
- Kazanç(Kredi Notu)=0.048 bit

olarak hesaplandığında en yüksek bilgi kazancı “Yaş” olduğu için bölme işlemi için “Yaş” özelliği seçilir. N kökü “Yaş” olarak etiketlenir.

Aynı işlemler diğer özellikleri için de bölünen her tabloda tekrarlanır. Karar ağacının son hali Şekil.3'teki gibidir.

1.5.2.3 Yapay Sinir Ağları

Yapay sinir ağları insan beyninden esinlenerek geliştirilmiş, ağırlıklı bağlantılar aracılığıyla birbirine bağlanan ve her biri kendi belleğine sahip işlem elemanlarından oluşan paralel ve dağıtılmış bilgi işleme yapılarıdır. Yapay sinir

ağları bir anlamda paralel bilgi işleme sistemi olarak düşünülebilir. Yapay sinir ağlarına bu bilgiler ilgili olaya ait örnekler üzerinde eğitilerek verilir.⁵²

Yapay sinir ağları iki alan üzerinde yoğunlaşmıştır. Bunlardan birincisi ileri beslemeli çok katmanlı ağlar, ikincisi ise Hopfield ağlarıdır. Yapay sinir ağları bilgi sınıflama ve bilgi yorumlamanın da içinde bulunduğu çok değişik problemlerin çözümünde kullanılmaktadır. Yapay sinir ağlarının kullanıldığı alanlar şu şekilde sıralanabilir;

- Denetim
- Sistem Modelleme
- Ses tanıma
- El yazısı tanıma
- Parmak izi tanıma
- Elektrik işareti tanıma
- Meteorolojik yorumlama
- Otomatik araç denetimi
- Fizyolojik işaretleri (kalp fonksiyonları gibi) izleme, tanıma, ve yorumlama

Yapay sinir ağlarının başlıca uygulama alanları sınıflandırma, tahmin ve modelleme olarak ele alınabilir.⁵³

Yapay sinir ağları birbirine bağlı çok sayıda işlem elemanlarından oluşan genellikle paralel işleyen yapılar olarak adlandırılabilir. Yapay sinir ağlarındaki işlem elemanları (düğümler) basit sinirler olarak adlandırılır. Bir yapay sinir ağı ,birbirleriyle bağlantılı çok sayıda düğümlerden oluşur.⁵⁴

⁵² ELMAS, Çetin. **Yapay Sinir Ağları (Kuram, Mimari, Eğitim, Uygulama)**, Seçkin Yayıncılık, Ankara, 2003, ss. 22, 23.

⁵³ ELMAS, ss. 25, 26.

⁵⁴ ELMAS, s. 31.

İşlem elemanları (düğümler) biyolojik sınırlara göre basit olmasına karşın, biyolojik sınırların 4 temel işlevini taklit ederler. Girişler x_i ile gösterilir. Bu girişlerin her biri ağırlık w ile çarpılır. Bu ürünler eşik değeri θ_j ile toplanır ve sonucu oluşturmak için etkinlik işlevi ile işlem yapılır ve yi çıkışı alınır.⁵⁵

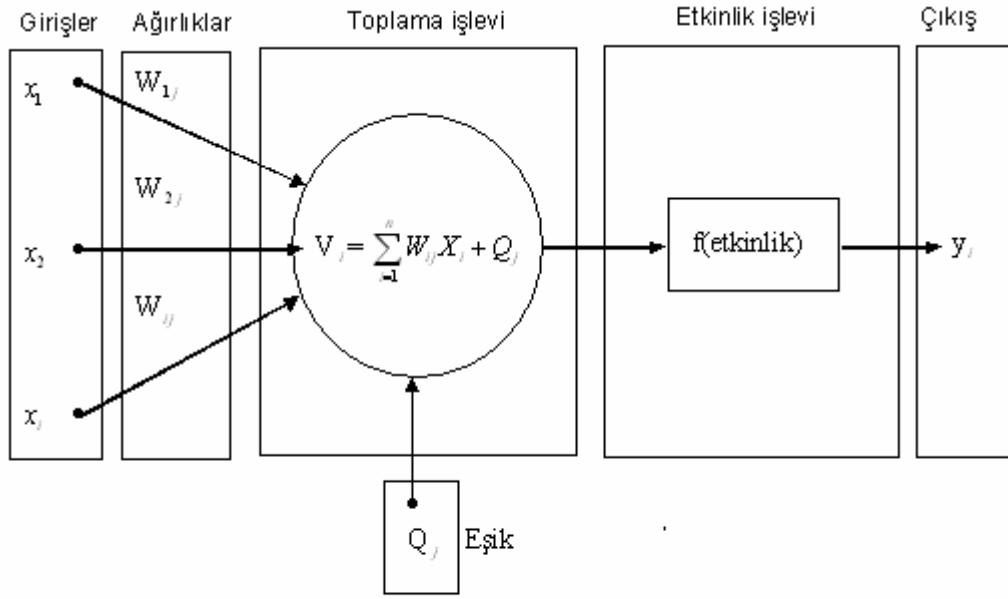
Girişler($x_1, x_2, x_3, \dots, x_n$) çevreden aldığı bilgiyi sinire getirir. Ağırlıklar (w_1, w_2, \dots, w_n) yapay sinir tarafından alınan girişlerin sinir üzerindeki etkisini belirleyen uygun katsayılarıdır. Her bir giriş kendine ait bir ağırlığa sahiptir.

Bir ağırlığın değerinin büyük olması, o girişin yapay sinire güçlü bağlanması ya da önemli olması, küçük olması zayıf bağlanması ya da önemli olmaması anlamına gelmektedir.

Toplama işlevi v_j , sinirde her bir ağırlığın ait olduğu girişlerle çarpımının toplamlarını eşik Q_j değeri ile toplayarak etkinlik işlevine gönderir. Bazı durumlarda toplama işlevi bu kadar basit bir işlem yerine en az, en çok, çoğunluk veya birkaç normalleştirme algoritması gibi çok daha karmaşık olabilir.

Toplama işlevinin sonucu etkinlik işlevinden $f(\text{etkinlik})$ geçirilip çıkışa iletilir. Bir etkinlik işlevinin kullanım amacı, zaman söz konusu olduğunda toplama işlevinin çıkışının değişmesine izin vermektir.

⁵⁵ ELMAS, s. 32.



Yapay bir sinir (düğüm)

Şekil 4. Yapay Sinir Ağının Yapısı⁵⁶

Çıkış işlevi etkinlik işlevi sonucunun dış dünyaya veya diğer sinirlere gönderildiği yerdir. Bir sinirin tek bir çıkışı vardır. Sinirin bu çıkışı, kendinden sonra gelen herhangi bir sayıdaki diğer sinirlere giriş olabilir.

1.5.2.4 Genetik Algoritmalar

Genetik algoritmalar, yönlendirilmiş rastgele araştırma algoritmalarının bir türüdür. Tabii seçme ve canlılarda bulunan genetik gelişimin benzetişimini gerçekleştirmektedir. Algoritma diğer evrimsel algoritmalar gibi araştırma uzayında bulunan çözümlerin bazılarının oluşturduğu bir başlangıç popülasyonunu kullanır. Başlangıç popülasyonu her jenerasyonda tabii seçme ve tekrar üreme işlemleri

⁵⁶ ELMAS, s. 32.

vasıtası ile art arda geliştirilir. En son kuşağın en uygun yani en kaliteli bireyi, problem için optimal çözüm olmaktadır.⁵⁷

Genetik algoritmaların temel adımları aşağıdaki şekilde özetlenebilir.

1. Çözümler için bir başlangıç fonksiyonu oluşturulur.
2. Popülasyondaki her çözümün uygunluk değeri hesaplanır.
3. Durdurma kriteri sağlanıyorsa araştırma durdurulur.
4. Durdurma kriteri sağlanmıyorsa tabii seleksiyon işlemleri uygulanır. Uygunluk değeri daha yüksek olan çözümler yeni popülasyonda daha fazla temsil edilirler. Çaprazlama işlemi uygulanır. Mevcut iki çözümden yeni iki yapı üretilir. Mutasyon işlemi uygulanır. Çözümlerde rastgele değişim uygulanır.
5. Tekrar 2. Adıma gidilir.

Genetik algoritmaların temsilinde ilk ve en çok kullanılan mekanizma (0,1) alfabesini kullanan ikili bit dizisidir.⁵⁸

Genetik algoritma uygulamaları üç ana gruba ayrılabilir. Bunların ilki deneysel uygulamalardır; mevcut diğer optimizasyon algoritmalarına karşı genetik algoritmalarının üstünlüğünü ispat etmek amacı ile belirli problemlerin çözümlerini bulmak için genetik algoritmaların kullanıldığı çalışmalardır. İkinci grup pratik uygulamalardır; genetik algoritmalar endüstri ve diğer gerçek problemlerin çözümlerini bulmak için kullanılır. Üçüncü grup ise veri madenciliğinde bilgi çıkarılma amacı ile kullanılan çalışmalardır.⁵⁹

1.5.2.5 En Yakın Komşu Metodu

Bu algoritmada sınıflandırma sırasında çıkarılan özelliklerden (feature extraction), sınıflandırılmak istenen yeni bireyin daha önceki bireylerden k tanesine yakınlığına bakılır.

⁵⁷ KARABOĞA, Derviş. **Yapay Zeka Optimizasyon Algoritmaları**, Kitap, Atlas Yayın Dağıtım, İstanbul, 2004. s. 78.

⁵⁸ KARABOĞA, s. 79.

⁵⁹ KARABOĞA, s. 90.

Örneğin $k = 3$ için yeni bir eleman sınıflandırılmak istensin. Bu durumda eski sınıflandırılmış elemanlardan en yakın 3 tanesi alınır. Bu elemanlar hangi sınıfa dahilse, yeni eleman da o sınıfa dahil edilir. Mesafe hesabında genelde öklid mesafesi (euclid distance) kullanılabilir.

KNN, eğitilmiş öğrenme algoritmasıdır ve amacı, yeni bir örnek geldiğinde varolan öğrenme verisi üzerinde sınıflandırma yapmaktır. Algoritma, yeni bir örnek geldiğinde, onun en yakın K komşusuna bakarak örneğin sınıfına karar verir.

1.5.2.6 Regresyon Analizi

Regresyon analizi bağımlı değişken ile bir veya daha çok bağımsız değişken arasındaki ilişkiyi incelemek amacıyla kullanılan bir analiz yöntemidir. Bir tek bağımsız değişkenin kullanıldığı regresyon tek değişkenli regresyon analizi, birden fazla bağımsız değişkenin kullanıldığı regresyon analizi de çok değişkenli regresyon analizi olarak adlandırılır.

Tek değişkenli regresyon analizi bir bağımlı değişken ve bir bağımsız değişken arasındaki ilişkiyi inceler. Tek değişkenli regresyon analizi ile bağımlı ve bağımsız değişkenler arasındaki doğrusal ilişkiyi temsil eden bir doğrunun denklemi formüle edilir.

Çok Değişkenli Regresyon Analizi içinde bir adet bağımlı değişken ve birden fazla bağımsız değişkenin bulunduğu regresyon modelleri çok değişkenli regresyon analizi olarak bilinir.

1.5.2.7 Naive Bayes Metodu

Naive Bayes algoritması her kriterin sonuca olan etkilerinin olasılık olarak hesaplanması temeline dayanan bir VM modelidir. Naive Bayes algoritmasında bir sonucun çıkma olasılığı o sonucu etkileyen tüm faktörlerin o sonucu sağlama olasılıklarının çarpımıdır. Aşağıdaki şekilde ifade edilebilir.

$$p(C | F_1, \dots, F_n) = \frac{p(C)p(F_1, \dots, F_n | C)}{p(F_1, \dots, F_n)}$$

1.6 Veri Madenciliği Uygulama Alanları

Veri madenciliği bankacılık, finans, pazarlama, iletişim, astronomi, biyoloji, sigorta, tıp gibi birçok alanda kullanılmaktadır. Örneğin pazarlamada mevcut müşterilerin satın alma davranışları, hangi müşteriye hangi promosyonların yapılacağı, sepet analizi gibi uygulamalarda kullanılmaktadır.

Bankacılık sektöründe veri madenciliği teknikleri genel olarak altı amaç dâhilinde kullanılmaktadır. Bunlar; müşteri segmentasyonu ve profillemeye, müşteri aşınmalarının tahmin edilmesi, kredi skorlama, kredi dolandırıcılıklarının (Fraud Detection), şirket analizi ve risk yönetimi ve kapasite planlama tahminidir.⁶⁰

Kredi Dolandırıcıları kredi kartı firmaları için önemli bir problemdir. Örneğin Visa ve MasterCard dolandırıcılıktan 1995 yılında 700 milyon dolar kaybetmiştir.⁶¹

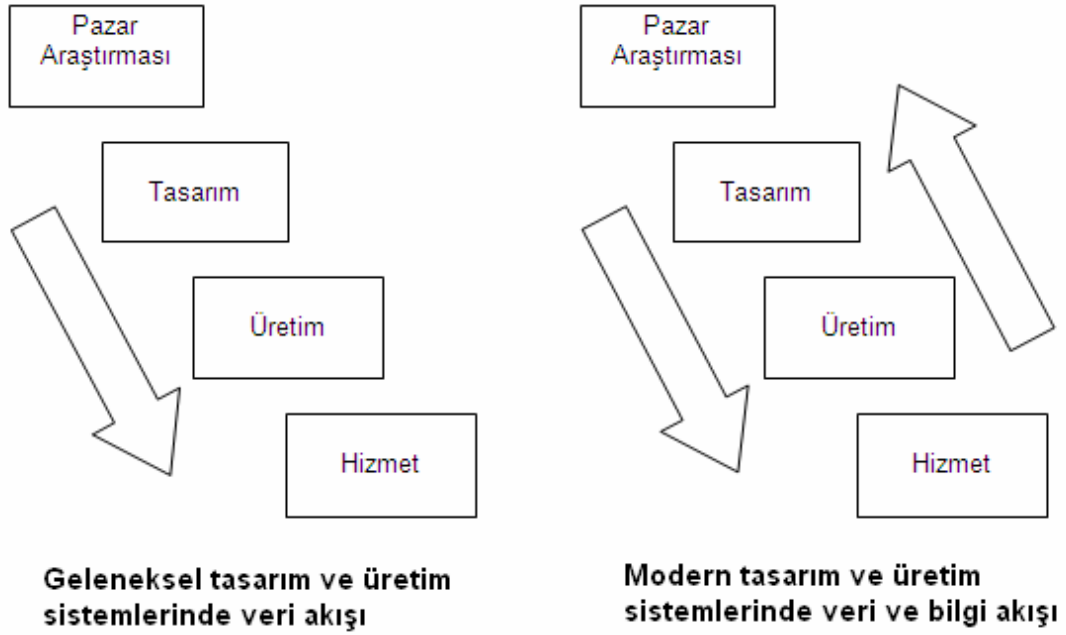
Perakende sektöründe de satış, müşterinin alışveriş geçmişi, nakliye, tüketim ve hizmet verilerinin toplandığı büyük veri setlerinin olmasından ötürü veri madenciliğinin yaygın kullanım alanı vardır. Özellikle internetin popülerleşmesi ve e-ticaretin yaygınlaşması ile toplanan veri miktarı hızla artmaktadır.⁶²

İletişim sektöründe veri transferi ve trafiğinin gittikçe artması ile sistem yükü, kaynak kullanımı, kullanıcı davranışları gibi konularda veri madenciliği teknikleri iletişim hizmetlerinin geliştirilmesine yardımcı olmaktadır.

⁶⁰ BIÇEN, s. 59.

⁶¹ KANTARDZIC, s. 328.

⁶² HAN, s. 651.



Şekil 5. Geleneksel ve Modern Üretim Sistemlerinde Bilgi Akışı⁶³

Veri Madenciliği, tıpta, biyolojide ve diğer bilimsel uygulamalarda, bilgisayar ve ağların güvenliği konularında da kullanılmasının yanı sıra suçluların bulunmasında, olası terör saldırılarının tahmini gibi sosyal konularda da kullanılmaktadır.⁶⁴

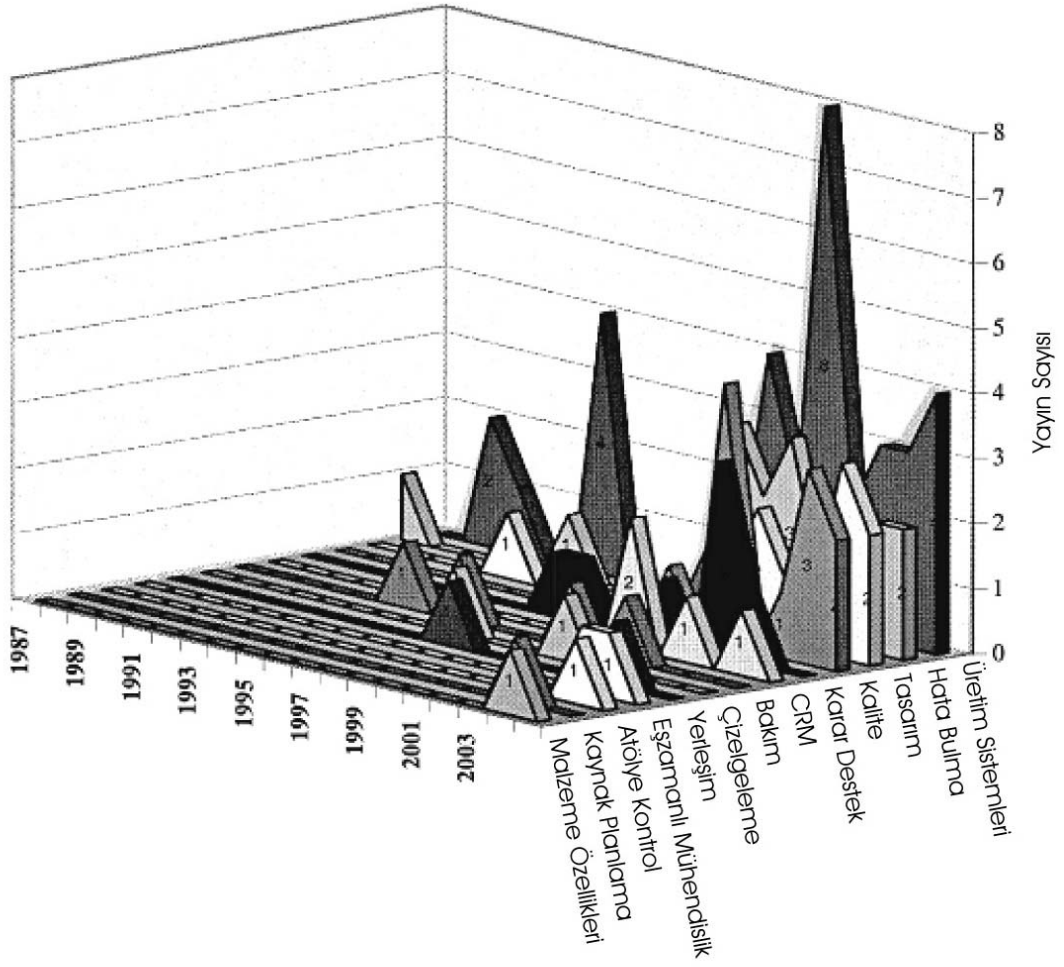
Kısacası büyük veri setlerinin bulunduğu ve analizinin gerektiği her sektörde veri madenciliği kullanım alanı bulmaktadır.

1.7 Üretimde Veri Madenciliği Uygulamaları

Geleneksel tasarım ve üretim sistemlerinde veri akışı tek yönlüdür. Modern sistemlerde ise veri akışı iki yönlüdür. Geleneksel ve modern üretim sistemlerinde bilgi akışı Şeki 5'te gösterilmiştir. Market analizi, tasarım, üretim ve hizmet döngüsündeki bu döngü bir veri yığını oluşturur. Bu verilerin önemi gün geçtikçe

⁶³ KUSIAK, Andrew. "Data Mining in Design of Products and Production Systems", **Proceedings of INCOM'2006:12th IFAC/IFIP/IFORS/IEEE Symposium on Control Problems in Manufacturing**, Vol. 1, 2006, Saint-Etienne, France, s. 50.

⁶⁴ DUNHAM s.17.



Üretimde Veri Madenciliği Uygulamalarının Tarihçesi

Şekil 6. Üretimde Veri Madenciliği Uygulamalarının Tarihçesi⁶⁵

artmaktadır. Veri madenciliği algoritmaları bu verilerden elde edilen bilgileri çıkarır ve yeni nesil ürün ve hizmetin gelişimine rehberlik eder.⁶⁶

Bir üretim ağında veri analizi, üretim ortamından toplanan verilerin planlama ve karar vermede kullanılma gerekliliğinin bulunması sonucunda çok önemli bir

⁶⁵ HARDING, s. 970.

⁶⁶ KUSIAK, s. 50.

görev haline gelmiştir. Geleneksel SQL tabanlı veri analizi uygulamaları ve teknolojisi gerekli olmakla birlikte yeterli değildir.⁶⁷

Üretimde veri madenciliği, teknikleri kullanımı 1990 larda başlamıştır. Veri madenciliği artık önleyici bakım, hata yakalama, tasarım, üretim, kalite kontrol, çizelgeleme ve karar destek sistemleri gibi üretim mühendisliğinin birçok farklı alanında kullanılmaktadır.⁶⁸

Harding et al (2006), üretimde veri madenciliği uygulamalarının tarihçesini ve yıllara göre yayınlanan makale sayısını Şekil6'daki gibi belirtmektedir.

Veri Madenciliği uygulamaları telekomünikasyon, tıp gibi çeşitli alanlarda araştırılmasına rağmen üretim alanında benzer ilgi olmamıştır. Bunun sebepleri aşağıdaki gibi açıklanabilir⁶⁹:

1. Üretim alanında çalışan araştırmacıların çoğunun veri madenciliği algoritmaları ve araçlarına yabancı olmaları
2. Veri Madenciliği araştırmacılarının çoğunun karmaşık üretim alanlarına yabancı olmaları
3. Veri Madenciliği ve üretim alanında çalışma yapan az miktardaki araştırmacının da üretim verilerine erişme imkânlarının olmayışı
4. Veri Madenciliğinin üretimde uygularken geçerliliğinin ve faydasının değerlendirilmesinin zor olması

Dan Baraha et al (2002) da veri madenciliğinin birçok alanda kullanılmasına rağmen tasarım ve üretimde işletmelerde uyum sağlamadığını belirtmektedir. Üretim firmaları arasında veri madenciliğine ilgi, veri madenciliğinin iş performansını değiştirmedeki potansiyelinin farkında olan zeki firmalarda ortaya çıkmaktadır. Örneğin Texas Instruments(hata önleme), Caterpillar (atıkların kalite kontrolü ve haklı şikayetlerin analizi), Ford (gerilme, gürültü ve vibrasyon analizi), Motorola

⁶⁷ LAU, H.C.W., Bing Jiang, W.B. Lee ve K.H. Lau. "Development of intelligent data mining system for a dispersed manufacturing network", **Expert Systems**, 2001 Vol.18, s. 175.

⁶⁸ HARDING, James A., M.Shahbaz, Srinivas ve Andrew Kusiak. "Data Mining in Manufacturing : A review", **Journal of Manufacturing Science and Engineering**, 2006, Vol. 128,

⁶⁹ WANG, Kesheng. "Applying data mining to manufacturing: the nature and implications", **Springer Science+Business Media, LLC 2007**, 2007, s. 489.

(CDMA Baz İstasyonu Yerleşimi), Boeing (Uçuş öncesi tanılama), ve Kodak (veri görselleme) gibi firmalar veri madenciliğini kullanmaktadırlar.⁷⁰

Büchner et al (1997) üretimde veri madenciliği uygulamaları üzerine örnekler vermiştir. Veri madenciliğinin üretimde en çok kullanıldığı alan arıza teşhisidir(fault diagnosis). Süreç ve kalite kontrol bir ürünün yaşam döngüsü ile de ilgilenir. Her ürünün yaşam döngüsü yaşam döngüsündeki her aşama dikkate alınır ve standartla karşılaştırılmak üzere ölçümler kaydedilir. Bu standartlardan sapmalar veri madenciliği için verimli bir kaynak oluşturur.⁷¹

C. Da Cunha (2006) ürün kalitesinin artırılmasında veri madenciliği yaklaşımı kullanılarak üretim verilerinden montaj sırasının belirlenmesini ve bu sayede üretimdeki hata riskinin en aza indirilmesini amaçlamıştır.⁷²

Andrew Kusiak (2006), alüminyum ve metal işleme, yarı iletken üretimi, elektronik dizgi, DNA üretimi, biyoteknoloji ve kimya sektörü, tıp, ilaç sektörü gibi üretim ve hizmet uygulamalarında veri madenciliği uygulamalarından örnekler vermiştir.⁷³

Backus et al (2006), da üretim süresi tahmininde veri madenciliği yaklaşımını incelemiştir. Sınıflama, K-En yakın komşu ve regresyon ağaçları yöntemlerinin birbirlerine göre avantaj ve dezavantajlarını araştırmıştır.⁷⁴

Menon et al (2005) , hızlı ürün geliştirme işlemi için metin veritabanlarının analizinde veri madenciliğini kullanmıştır. Ürün geliştirme sürecinin her aşamasında veriler çeşitli veritabanlarında tutulmaktadır. Örneğin planlama ve tasarım aşamasında veriler rapor formatında tutulmaktadır. Resim ve kelime şeklinde tutulan bu veritabanları multimedya veritabanları olarak da adlandırılmaktadır. Üretim ve hizmet aşamasında hem sayısal hem de metin veritabanları görülebilir. Buradan da

⁷⁰ BRAHA(ED), Dan. **Data Mining for Design and Manufacturing**, Preface, Springer, 2002, Hardcover,ISBN:1-4020-0034-0, s. xii.

⁷¹ BÜCHNER et al, ss. 6, 7.

⁷² CUNHA, C.Da, Bruno Agard ve Andrew Kusiak. "Data Mining for improvement of product quality", **International Journal of Production Research**, Vol. 44, No. 18-19, 2006, s. 4028.

⁷³ KUSIAK, Andrew. "Data Mining: manufacturing and service applications", **International Journal of Production Research**, Vol. 44, No. 18-19, 2006, s. 4175.

⁷⁴ BACKUS, Philip, Mani Janakiram, Shanin Mowzoon, George C. Runger ve Amit Bhargava, "Factory Cycle Time Prediction with a Data Mining Approach", **IEEE Transactions On Semiconductor Manufacturing**, Vol. 19, No. 2, 2006, s. 255.

anlaşılacağı gibi ürün geliştirme sürecinin her aşamasında metin veritabanları görülebilmektedir. Bu veritabanlarına veri madenciliği uygulanmaması firmalar için büyük kayıp olacaktır.⁷⁵

Pamela N. Ajoku ve Bart Nnaji veri madenciliğini üretkenliğin artırılmasında kullanmıştır. Örneğin bir makinenin verimliliği ki üretkenliği etkilemektedir, ister istemez kullanım sıklığına ve servis/bakım bilgisine bağlıdır. Belirtilen makine durumlarının verilerinden modeller geliştirilebilir ve çizelgeleme optimize edilip gereksiz arıza süreleri minimize edilebilir.⁷⁶

H.C.W. Lau et al (2001), dağınık üretim ağı için veri, madenciliği sistemi için kuralcı OLAP yaklaşımını kullanmıştır.⁷⁷

Gürdal Erdek et al (2004), üretim planlama için satış verilerinden görsel ve analitik veri madenciliğini kullanmıştır.⁷⁸

Sonuç olarak veri madenciliği uygulamaları gelişmiş üretim sistemlerinin aşağıdaki kısımlarında uygulanır⁷⁹:

1. Üretim sistemleri modellemesi
2. Üretim süreç kontrolü
3. Kalite Kontrol
4. İzleme ve teşhis
5. Güvenlik değerlendirme
6. Süreç planlama ve çizelgeleme

⁷⁵ MENON, Rakesh, Loh Han Tong ve S. Sathiyakeerthi. "Analyzing textual databases using data mining to enable fast product development processes", **Reliability Engineering and System Safety** **88**, 2005, s. 6.

⁷⁶ AJOKU, Pamela N. ve Bart Nnaji, "Improving Productivity in Manufacturing Environments Using Data Mining", <http://iweb.uta.edu/vchen/AIDM/AIDM-Ajoku.pdf>, 10.02.2009. s. 2.

⁷⁷ Lau, s. 175.

⁷⁸ ERTEK, Gürdal, Can Kuruca, Cenk Aydın, Besim Ferit Erel, Harun Doğan, Mustafa Duman, Mete Ocal ve Zeynep Damla Ok. "Visual and analytical mining of sales transaction data for production planning and marketing", **4th International Symposium on Intelligent Manufacturing Systems**, 2004, s. 848.

⁷⁹ WANG, s. 490.

7. Üretim karının optimizasyonu
8. Montaj seçimi
9. Robotların öğrenmesi
10. Malzeme İhtiyaç Planlanması
11. Önleyici makine bakımı

1.8 Veri Madenciliğinin Avantaj ve Dezavantajları

Veri Madenciliğinin avantajları aşağıdaki gibi sıralanabilir:

- Rekabette üstünlük kazanmak için daha iyi bilgi sağlar. Veri Madenciliği işletmenin kendisini, müşterisini ve pazarı anlaması için gerekli bilgiyi sağlamada güçlü bir araçtır.
- Veri Ambarlarına değer katar. İşletmeler veri ambarı yatırımlarının yüzde 10 ile 70 arasını veri madenciliği uygulayarak geri kazanabilirler.
- Bilimsel araştırmaları darboğazdan kurtarır. Sosyal bilimlerde gerçek veriler ile çalışmanın imkansız olduğu durumlarda veri madenciliği büyük veri sağlar.
- İşlem hızını artırır. Veri madenciliğinin büyük verileri hızlı bir şekilde organize ve analiz etme kabiliyeti işyeri verimliliğini artırır.
- Verileri kullanmada esneklik sağlar.
- İşletme maliyetlerini azaltır.
- Verileri analiz için çok fazla istatistiksel deneyim gerektirmez.

Veri Madenciliğinin avantajlarının yanı sıra dezavantajları da vardır. Bunlar:

- Kesin bir cevabı yoktur. Veri Madenciliği kullanışlı öngörüler bulur ancak kesin bir cevabı yoktur.
- Maliyeti yüksektir. Veri Madenciliği uygulaması geliştirmek maliyetlidir.
- Karmaşık ve uzun bir projedir. Veri madenciliği uygulamalarından verim almak uzun zaman alır.

- Veri Madenciliğinde kullanılan veriler mahremiyet içerebilir. World Wide Web dahil edildiğinde problem daha da kötüleşebilir.
- Kullanıcı ara yüzlerinin basit ve veri analizinin otomatik olmasına rağmen, istatistik, üretim yönetimi ve yönetim bilimleri altyapısı olan kullanıcılara daha uygundur.

İKİNCİ BÖLÜM

ÜRETİM YÖNETİMİNDE ÜRETİM SÜRESİ

2.1 Üretim/İşlemler Planlaması ve Kontrolü

Kestirilen talebi tam zamanında karşılamak için, ardışık birçok karar içeren üretim planlaması yapmak gerekir. Söz konusu ardışık kararlar, başlıca şu soruların yanıtlanmasıyla ilgilidir: Ne kadar hammadde, yardımcı madde, işletme malzemesi ve parça, ne zaman sipariş edilmelidir? Hangi tür ve nitelikte kaç kişi, geçici veya kalıcı olarak işe alınmalıdır? Hangi kişi, ne zaman hangi işe, makineye veya sürece verilmelidir? Zamanın boşa geçmesini önlemek için, nerede ne kadar envanter bulundurmalıdır? Stokta mal tutmadan tüketici talebini tam zamanında karşılamak için, her bir aşamada ne kadar stok bulundurulmalıdır? Üretim planlaması ve zamanlamasında, güncel bilgi teknolojilerinden yararlanır. Özellikle Kurumsal Kaynak Planlaması (Enterprise ResourcePlanning) ve Stoksuz Üretim Sistemi (Just-In-Time), bu alanda en çok kullanılan yazılım ve yaklaşımlar arasında sayılır. Üretim planlaması ve zamanlaması, üretim yönetiminin en büyük uğraş alanlarından birisidir.

Üretim yönetimi, işletmenin elinde bulunan malzeme, makine ve insan gücü kaynaklarının belirli miktarlardaki mamulün istenilen niteliklerde (kalitede), istenilen zamanda ve en düşük maliyetle üretimini sağlayacak biçimde bir araya getirilmesidir.⁸⁰

Üretim/İşlemler kontrolü istenilen çıktıyı istenilen zamanda teslim etmeye çaba harcar.

Üretim/İşlemler Planlaması istenilen zamanda, nicelikte ve kalitede maddelerin ya da hizmetlerinin üretimini yapılmasının sağlanması ve işlemlerin

⁸⁰ KOBU, Prof. Dr. Bülent. **Üretim Yönetimi**, Beta Yayınları, İstanbul, Kasım 2008, Genişletilmiş 14. Baskı., s. 5.

uygulamaya konulması için konunun kurumsal yanını yazılı, biçimsel ve matematiksel biçimde hazırlaması olarak tanımlanabilir.⁸¹

Kusar Üretim ve planlama ve kontrolü, en basit amacı belirlenmiş bir başlama zamanı ya da bitiş zamanı olan bir sipariş için işin ne zaman başlayacağı sorusuna yanıt bulma olarak tanımlamış ve 5 aşamaya bölmüştür.

İlk aşamada siparişin üretim planı oluşturulur. İkinci aşamada sipariş için malzeme ihtiyaç planlaması yapılır. Üçüncü aşamada sipariş üretim ve tedarik siparişi olarak ikiye bölünür. Dördüncü aşamada üç aşamalı çizelgeleme yapılır. İlk aşamada siparişin yapısına göre üretim süresi çizelgelemesi yapılır. Bu çizelgelemenin amacı kapasiteleri dikkate almadan işlemlerin başlama ve tamamlanma sürelerini çıkarmaktır. İkinci aşamada kabaca kapasite verilerine göre kapasite planı yapılır. Bu kaba planlamanın amacı kapasite yük profilini ve gerekli ayarlamaları görmektir. Üçüncü aşamada kapasiteye göre hassas kapasite çizelgelemesi yapılır. Beşinci aşamada üretim planlama ve kontrol üretim ve tedarik siparişlerinin zamanında yapılmasını kontrol için kullanılır.⁸²

Üretim/işlemler planlaması ve kontrolü işlevini Ön Planlama, Planlama ve Kontrol olarak üç ana grupta toplanabilir.

2.1 Ön Planlama

Üretim geliştirme ve tasarımı, satış öngörülmemesi, işyeri düzeni, araç politikası ve üretimin ön planlaması çalışmalarını kapsar.

Üretim yönetiminin amacı, dört üretim faktörü (miktar, kalite, zaman ve maliyet) için en uygun (optimum) değerlerin bulunmasına yönelmiştir. Ön planlama aşamasında dört üretim faktörü ile ilgili olanaklar araştırılıp gerekli önlemler alındıktan sonra üretim yönetiminin planlama aşamasında yer alan; rotalama, zaman tahminleri ve programlanma fonksiyonlarının yürütülmesine geçilir.

⁸¹ DEMİR, Prof. Dr. M. Hulusi ve Prof. Dr. Şevkinaz Gümüšoğlu. **Üretim Yönetimi (İşlemler Yönetimi)**, Beta Yayınları, İstanbul, Kasım 1998, Genişletilmiş 5. Baskı., s. 369.

⁸² KUSAR, Janez, Ales Brezovar, Janez Grum, Marko Starbek. "Realistic lead time scheduling of operations of orders", **International Journal of Machine Tools & Manufacture**, Vol 44, 2004, s. 1038.

2.2 Planlama

Üretim planlaması gelecekteki imalat ve hizmet faaliyetlerinin düzeylerini ve limitlerini belirleyen fonksiyon olarak tanımlanabilir. Üretim planının hazırlanması için yapılması gereken işler aşağıdaki gibi sıralanabilir.⁸³

- 1- Üretim planının kapsayacağı zaman aralığı tespit edilir
- 2- Ekonomik stok düzeyleri hesaplanır
- 3- Talep tahminleri yapılır
- 4- Plan dönemi başındaki ve sonundaki stok düzeyleri belirlenir
- 5- Başlangıç ve bitiş stokları arasındaki fark bulunur
- 6- Planlama dönemi içinde üretilmesi gereken miktar bulunur
- 7- Üretilmesi gereken miktar dönem dilimlerine dağıtılır

Bu gruba planlama kaynakları ile işin hazırlanması ve iş verme işlevleri girer. Bunlar:

Materyaller:

Üretimi yapılacak malın tüm teknik özellikleri çıkartılıp üretim esnasında gerekecek materyaller ve bunların nitelik ve nicelikleri belirlenir.

Yöntemler:

Üretim süreçlerin incelenmesi ve çözümlenmesi ile verilen koşullar altında optimum yöntem belirlenir.

İşgücü, makine ve araçlar:

Yöntem saptanırken işgücünün ve makinelerin niteliği belirlenir. İşçilere gerekli eğitim verilir, makinelerin bakım onarımı yapılır. Üretimin aksamaması için önemli yedek parçalar stoklanır.

Bir sonraki aşamada yapılacak işlerin sırası saptanır.

Rotalama:

⁸³ KOBU, ss. 452, 453.

Rotalama, üretim faaliyetlerinin hangi sıraya göre ve nasıl bir akış içerisinde yürütüleceğini belirler. Rotalamada makinelerin fabrika içerisindeki yerleşim düzeni dikkate alınarak mamulün izleyeceği yol belirlenir.

Yapılacak işlerin sırasının saptanması, öteki tüm planlamaların dayandığı bir temel üretim işlevidir.⁸⁴

Öncelik Kuralları ile İş Sıralama:

Üretim yönetiminde iş sıralama problemi, belli bir performans ölçütünü eniyileyecek iş akış sırasını (tezgahlara gelen işlerin yapılma sırasını) belirlemek olarak bilinir. İş sıralama problemi işlerin geliş şekline göre Akış Tipi ve Sipariş Tipi olmak üzere ikiye ayrılır. Akış Tipi iş sıralama problemlerinde belirli bir dönem için iş listesi bilinmektedir ve tüm işlerin makinelerde işlenme sırası aynıdır. Sipariş Tipi problemlerde ise iş listesi sürekli ve rastgele değişmekte, işler düzensiz aralıklarla atölyeye gelmektedir. Burada önemli olan nokta ise işlerin makinelerde işlenme sıralarının farklı olabilmesidir.⁸⁵

Öncelik kuralları (priority rules) herhangi bir teori ve matematik model kullanmadan sadece tecrübe, sağduyu veya işletme politikasına dayanarak oluşturulur. Bu kurallar yardımı ile yapılacak sıralamalar optimal değildir. Bir kuralı diğerinden daha iyi veya daha kötü olduğunu söyleyerek genelleme yapma olanağı da yoktur. Belirli bir iş grubu için en uygun sıralama yöntemi her kural için hesaplanan karakteristik ölçüleri kıyaslama yolu ile seçilir.

Uygulamada en çok rastlanan öncelik kurallarından dördü şunlardır:

İlk Gelen İlk Yapılır - FCFS (first come first served):

İşler siparişlerin alınış sırasına göre birinci sipariştten başlayarak sıralanır.

⁸⁴ Demir, s. 375.

⁸⁵ BASKAK, Murat ve Vural Erol, "Sipariş Tipi Atölyelerde İş Sıralama Problemi için bir genetik algoritma uygulaması", **YA/EM'2004-Yöneylem/Endüstri Mühendisliği-XXIV Ulusal Kongresi**, 15-18 Haziran 2004, Gaziantep – Adana.

En Kısa İşlem Zamanı - SPT (shortest processing time):

İşler işlem sırası en kısa olan işten başlayarak sıralanır.

En Erken Teslim Süresi - EDD (earliest due date):

İşler süresi teslim süresi en kısa işten başlayarak sıralanır.

Kritik Oran - CR(critical ratio):

İşler kritik oranı en küçük olan işten başlayarak sıralanır.

Öncelik kuralları ile sıralamada kullanılan ölçülerin tanımları aşağıda verilmiştir:

T_i : İşlem Süresi (operation time); bir birimin üretilme süresi. Bu süre bir parti malın üretilme süresidir. Bu süre ön işlem süresini ve ayarlama süresini içermez.

T_f : Akış Süresi (flowtime); işlem, gecikme ve bekleme sürelerinin toplamıdır

\bar{T}_i : Ortalama akış süresi; n iş sayısı olmak üzere ($\sum t_i/n$)'dir.

T_l : Gecikme (lateness); akış süresi ile teslim süresi arasındaki farktır. Negatif gecikmeler sıfır kabul edilir.

\bar{T}_l : Ortalama gecikme süresidir.

M_s : Yapım süresi (makespan); son işin bitiş zamanı ile ilk işin başlangıç zamanı arasındaki farktır. İlk işin başlama süresi hesaplamaları kolaylaştırmak için sıfır kabul edilir.

\bar{WIP} : Ara depolarda bekleyen (work in progress) ortalama iş sayısı

CR: Kritik oran (critical ratio); (teslim süresi / t_i)'dir.⁸⁶

⁸⁶ KOBU, s. 472.

İşlem zamanlarının öngörümlemesi:

Üretim yöntemi ve sırası saptandıktan sonra işlem zamanları hesaplanabilir. Bunun için iş ve zaman etüdüleri yapılır.

Yükleme ve Çizelgeleme:

Kesikli üretimde bulunan işletmelerde makinelerin ve çeşitli üretim dairelerinin boş kalmaması ve aynı zamanda birden çok işin kendilerine verilmemesi gerekir. Alınan siparişlerin işletmedeki bütün işçilerden ve makinelerden en büyük verim alınabilecek biçimde sıraya konulmalıdır. İşletmede bu sorun "Yükleme ve Çizelgeleme" ile çözülür.

Çizelgeleme problemlerinde üç ana amaç mevcuttur. Bu amaçlardan ilki teslim tarihi ile ilgilidir; müşterilerin siparişleri geciktirilmemeye çalışılır. İkinci amaç akış süreleri ile ilgilidir; işlem süreleri minimize edilmeye çalışılır. Üçüncü amaç ise iş merkezlerinin kullanımı ile ilgilidir; makine, teçhizat ve personel açısından iş merkezini en etkin kullanımı amaçlanır. Çizelgeleme tipleri statik ve dinamik olarak ikiye ayrılabilir.

1- İşlerin Geliş Şekli: Eğer çizelgelenecek işler zamanla değişiklik göstermiyorsa, sistem statik, zamanla yeni işler ortaya çıkıyorsa sistem dinamik olarak isimlendirilir. Statik modeller dinamik modellere göre daha kolay kontrol edilebilir bir yapıdadır ve daha geniş bir çalışma alanına uygulanmıştır. Ayrıca statik modellerin analizi, daha genel durumların çözümünde faydalı olabilecek faydalı girdiler ve sezgisel yaklaşımların bulunmasında etkili olmuştur.

2- Tezgah Sayısı: Atölyede yer alan tezgah sayısına tek ve çok tezgahlı olarak ikiye ayrılır.

3- İş Akışı: Eğer gelen bütün işler belli bir sırayı takip ediyorsa sıralama problemi akış tipi; farklı sıraları takip ediyorsa karışık iş akışından söz edilir.

4- Performans Ölçütü: Atölye performansını değerlendirmek için kullanılan ölçüt çizelgeleme problemlerinde önemli bir ol oynamaktadır. Bu ölçütlerden bazıları akış zamanı, üretim içi stok seviyesi, işlerin tezgahta bekleme süresi, işlerin ortalama gecikme süresi, geciken iş yüzdesi, tezgah ve iş gücü kullanım oranlarıdır.

Yükleme, işlerin iş merkezlerine atanması işlemidir. Zaman ve metot etütleri ile saptanan makine ve işçilik süreleri ve eldeki iş yükü göz önüne alınarak, ayrıntılı üretim programları hazırlanır. Bir üretim programında hangi işlerin hangi makinelerde kim tarafından ve ne zaman yapılacağı kesin olarak belirlenir. İşin yapılacağı makinelerin saptanması, daha önce başlanmış işlerin süregelmeleri ve işlem zamanlarının farklı olması nedeni ile çok karmaşık sorun yaratır. Literatürde **iş yükleme** adını alan bu problemin çözümü, orta büyüklükte üretim üniteleri için dahi, ileri düzeyde matematik yöntemlerinin uygulanmasını gerektirir. Yüklemede yapılacak hatalar, mamulün istenilen zamanda bitmemesi veya makineler aşırı yüklü iken bazılarının boş kalması gibi sonuçlar doğuracağından işletmenin doğrudan maddi kayıplara uğramasına yol açar.⁸⁷

İş Akışı, üretim/işlemler kontrolünde üretim emrinden sonraki aşama, materyalin depodan çıkarılması ve fabrikada ilgi dairelere, atölyelere dağıtımının yapılarak fiilen üretime başlaması ile olur.

2.3 Kontrol

İşlerin planlandığı ve zamanlandığı biçimde yürütülmesi için üretim sırasında oluşan aksaklıkların giderilmesi işlemidir.

2.4 Üretim Süresi

Üretim süresi veya temin süresi, bir sürecin gerektirdiği hammaddelerin siparişinden son montajın tamamlanmasına kadar tüm sıralı aşamaların temin sürelerinin toplamı olarak da ifade edilebilir.

Bir iş emrinin iş istasyonuna gelişinden ile o istasyondaki işin tamamlanmasına kadar geçen süre üretim süresi olarak bilinir. İşin siparişinden tamamlanmasına kadar üretim sisteminde geçirdiği süre ise toplam üretim süresi olarak tanımlanır.⁸⁸

⁸⁷ KOBU, s. 13.

⁸⁸ CHINCHOLKAR, Mandar M., Timoty Burroughs ve Jeffrey W. Hermann. "Estimating Manufacturing Cycle Time and Throughput in Flow Shops with Process Drift and Inspection", 2004, <http://www.isr.umd.edu/Labs/CIM/projects/dfp/flowshop.pdf>, (10.02.2009), s. 4.

Üretim süresi bir malzemenin işleme başlamadan önceki bekleme süresi ve işlem süresinin toplamıdır.⁸⁹

Üretim süresi tahmini teslim süresi tahmini ile işlem süresi arasındaki farktır. Gerçek üretim süresi ise gerçek teslim süresi ile işlem süresi arasındaki farktır. Teslim zamanı siparişin verilmesinden teslim edilmesine kadar geçen süredir.⁹⁰

Üretim yönetiminde zaman tahmini yapabilmek için aşağıdaki kavramların tanımlanması gereklidir.

İşlem Süresi: Bir kişi ya da bir makine tarafından bir işlemi yapmak ya da belli bir miktardaki ürünü elde etmek için harcanan sürenin oluşumu aşağıdaki şekildeki gibi ele alınabilir.

İş Kapsamı: Üretim süreci boyunca tanımlanmış faaliyetler, üretken birimlerin yönetsel ya da fiziksel nitelikli görevlerini(işlerini) yerine getirmeleriyle gerçekleştirilir. Buna, işin fonksiyonel kapsamı denir. Diğer taraftan, genellikle adam-saat veya makine-saat birimiyle ölçülen ve işin tamamlanma süresi olarak ifade edilen bir iş kapsamından daha bahsedilir. İş etüdü, her iki iş kapsamının analizinde de önemli bir yere sahiptir.

Diğer taraftan, tamamlanma süresi itibariyle iş kapsamı ise bir takım alt düzeylerden ve elemanlardan oluşur. Burada ilk düzey mevcut koşullar altında işin toplam tamamlanma süresidir. Toplam iş süresi olarak da ifade edilen bu düzey, işin tanımında yer alan başlangıç ve bitiş noktaları arasında geçen süreyi göstermektedir.

Her ürünün kendine özgü parametreleri vardır. Kurulum zamanı, işlem süresi ve bekleme oranı gibi. Kurulum zamanı işlem süresinden genellikle daha uzundur. Benzer özellikteki ürünlerin topluca üretilmesi ile kurulum zamanı kısaltılabilir. Bu

⁸⁹ LEE, Ho Woo, No Ik Park ve Se Won Lee. "Analysis of the manufacturing lead time in a production system with non-renewal batch input, threshold, policy and post operation", **Applied Mathematical Modeling 31**, 2007, s. 2161.

⁹⁰ SLOTNICK, Susan A. ve Matthew J. Sobel. "Manufacturing lead-time rules: Customer retention versus tardiness costs", **European Journal of Operational Research 163**, 2005, s. 826.

sayede üretim süresi kısalmış ve kapasite artmış. İş büyüklüğünü arttırmak için yapılmış süresini arttırır ve üretim süresi artar.⁹¹

Bu süre içerisinde iş etüdü açısından önemli olan üç özellik vardır. Bunlar:

- Doğrudan üretimde bulunan etken süre.
- İşçilere yasalar gereği verilen dinlenme payları ile işçi ve makinelere iş ölçümü gereği verilen dinlenme payları
- İş süresinin içindeki dinlenme paylarının dışındaki etken olmayan süre.

Her ürün ya da işlemin temel iş kapsamı vardır. İş kapsamı, adam-saat ya da makine-saat ile ölçülen belli bir ürün ve sürecin kapsadığı iş miktarıdır. Temel iş kapsamı daha fazla azaltılması olanaksız en düşük çalışma süresidir. Uygulamada temel iş kapsamı gerçekleşmez. Gerçek işlem süreleri temel iş kapsamından daha uzundur Buna aşırı iş kapsamı denir.⁹²

Üretim Süresi (Toplam İş Süresi) :Toplam İş Kapsamı ve Toplam Etken Olmayan Sürenin toplamından oluşur.

Toplam İş Kapsamı: Ürünün veya işlemin temel iş kapsamı, ürün belirlemelerindeki eksiklerden doğan ek iş kapsamı, etken olmayan üretim ya da işlem yöntemlerinden dolayı artan iş kapsamının toplamından oluşur.

Toplam Etken Olmayan İş Kapsamı: Yönetimin yetersizliğinden doğan etken olmayan süre, işçinin denetimi altındaki etken olmayan sürenin toplamından oluşur.

Kötü hazırlanmış modeller en ekonomik süreçlerin kullanılmasını önlerler. Ürün ve parçanın modeli en ekonomik süreçlerin ve yapım yöntemlerinin kullanılmasına olanak vermeyebilir. Bu, genellikle metal işleri endüstrisinde ve özellikle, büyük çapta üretim yapılan yerlerde görülür. Parçanın modeli yüksek verimli makinelerden yarar sağlayabilecek şekilde düzenlenmeyebilir.(Örneğin: Bir

⁹¹ KAMATH, Narasimha, Subir Bhattacharya, "Lead time minimization of a multi-product, single-processor system: A comparison of cyclic policies", **Int. J. Production Economics**, Vol. 106, 2007, s. 29.

⁹² INTERNATIONAL LABOUR OFFICE (Uluslararası çalışma Örgütü), çev. Zühal Akal, **İş Etüdü**, Kitap, Milli Prodüktivite Merkezi Yayınları: 29,Dördüncü Basım, Ankara, 1991, s. 14.

metal parça, tek bir parça halinde prese verilecek yerde, kesilecek, perçinlenecek ya da kaynak yapılacak şekilde hazırlanabilir.)

Çok çeşitli ürün üretmek ya da parçalarda standardizasyona gitmemek, iş gruplarının küçük olmasını gerektirir ve bu iş özel amaçlı yüksek verimli üretim makinelerinde yapılamaz, buna karşılık daha küçük ve çeşitli işlere uydurabilecek makinelerde yapılması bir zorunluluk haline gelir.

Çok yüksek ya da çok düşük olarak seçilmiş kalite standartları iş kapsamını arttırabilir. Mühendislikte, toleransların gerektiğinde daha dar tutulması daha fazla makine işi gerektirir. Bu halde, daha fazla geri çevrilen mal ve buna koşut olarak da malzeme kaybı olacaktır. Diğer taraftan, çok düşük kalite malzeme, istenen işin çıkarılmasında güçlükler doğurur ya da ürünü hazırlarken, kullanılabilmesini sağlamak için temizleme işi gibi ek işler ortaya çıkarır. Malzemenin kalitesi otomasyonda daha büyük önem kazanır.

Bir ürünün modeli öyle bir şekilde yapılır ki, bu ürüne son şeklini vermek için malzemenin fazlaca yontulması gerekir. Bu durum iş kapsamını artırır ve malzeme kaybına neden olur.

Sürecin etken olmayan biçimde işletilmesi üretimde ya da işlemde etken olmayan yöntemlerin kullanılması iş kapsamını aşağıda sıralanan biçimde etkilemektedir:⁹³

- Küçük bir torna işini büyük bir torna tezgâhına yapmak ya da dar bir kumaşı çok geniş bir tezgâhta dokumak gibi olması gerekenden daha büyük çıktısı olan yanlış tipte ya da büyüklükte bir makine kullanılmışsa;
- Süreç iyi işlemiyorsa, yani malzeme zamanında gelmiyorsa, malzemeni geliş hızı, akış oranı, ısı ve yoğunluğu ya da bu işlemi yürüten herhangi bir koşul uygun değilse, makine ve makine tesisatları iyi durumda değilse,
- Yanlış el araçları kullanılıyorsa,
- İşletmenin yerleşim düzeni, hareket, zaman, çaba kaybına yol açıyorsa,

⁹³ INTERNATIONAL LABOUR OFFICE, s. 18.

- Üretken işçinin çalışma yöntemleri hareket, zaman ve çaba kaybına neden oluyorsa.

Süreyle tanımlanan iş kapsamını, işlemin düzenli ortalama çalışma hızında yapıldığı varsayımına dayanmaktadır. İş hızındaki bir düşüş nedeni ile oluşan ek süre etken olmayan süre olarak ölçülür.

Gereğinden çok mal çeşidinin üretilmesini gerektiren pazarlama politikası ile her çeşidin kısa süreli üretimine neden olur ve değişik ürünlerin üretimine geçerken makineler boş kalır. İşçilerin herhangi bir işlemde hız ve ustalık kazanmalarına olanak yoktur.

Ürünler arasında ya da belli bir ürünün içindeki parçaları mümkün olduğu kadar standartlaştırmama, modellerin gereği gibi geliştirilmesini ya da müşterilerin isteklerinin baştan beri karşılanmasını sağlamada başarısızlık ile bu işin duraklamasına, malzeme kaybına olduğu kadar, makine ve insan saatlerinin kaybına neden olabilecek model değişikliklerine yol açar.

İş ve sırasının akışı planlamada başarısızlık ile bunun sonucu bir sıra hemen diğerini izleyemez, makineler ve işçiler sürekli olarak çalıştırılmaz. İş için gerekli olan hammaddenin, araçların ve diğer donatımın sağlanamaması dolayısıyla makinelerin ve işçilerin iş bekleme yüzünden boş kalmasına yol açar. Tesisat ve makinelerin en iyi biçimde bakımlarının yapılmaması ile bu makinelerin bozulmalarından dolayı gereksiz duruşlara yol açar.

Makinelerin ve tesisatın kötü koşullar altında çalışmasına göz yummak, böylelikle yapılan işin kötü olmasına ya da düzeltilmesi için geri verilmesine ve yeniden yapılmasına neden olmakla-ışın yeniden yapılması etken olmayan sürenin doğmasına yol açar.

İşçinin sürekli ve aksaksız çalışabileceği koşulları hazırlamakta başarısızlığa uğramakla, işçilerin güvenliği için gerekli önlemleri almamakla bu iş kazaları yoluyla zaman kaybına sebep olur.

İşçilerin olumlu ya da olumsuz davranışları etken olmayan süreyi aşağıdaki gibi etkilemektedirler:⁹⁴

⁹⁴ INTERNATIONAL LABOUR OFFICE, s. 21.

- İşçiler tarafından anlamlı bir neden olmadan çalışmaya ara vermek,
- Geç kalmak,
- İşe başlama saatinde işe hemen başlamamak,
- İş sırasında boş oturmak ya da isteyerek yavaş çalışmakla
- İsrafa ya da işi yeniden yapmaya yol açacak dikkatsiz çalışmaya,
- İşin tekrar yapılması zaman kaybına,
- Güvenlik kurallarına dikkat etmemek ve dikkatsizlik yüzünden kazalara yol açmakla

İş Ölçümü

İş ölçümünde amaç işin yapılış süresini tespit ederek üretim planlama ve diğer yönetim kademelerinin daha etkin çalışmalarını sağlamaktır. Bir işi oluşturan elemanlar malzeme, makine ve insan gücüdür. Malzeme ve makinenin değişkenliği az olduğu için ölçülmesi kolaydır. İnsan gücü ise sosyal çevre, kişisel yetenekler, eğitim tecrübe, teşvik, çalışma güvenliği gibi çok çeşitli değişkenlerin etkisi altındadır.⁹⁵

2.4.1 Üretim Yönetiminde Üretim Süresi Tahmininin Önemi

Bir üretim sisteminin başarı göstergesi siparişin belirtilen zamanda karşılanıp karşılanmadığı ile belli olur.

Zamanın satıcılar ve alıcılar için de aynı anlamı ifade etmesi dolayısıyla, rekabetçi avantaj elde etmede “zaman”, “maliyet” gibi stratejik bir güç olarak kabul görmüştür.⁹⁶

Sipariş ve teslimi arasındaki en kısa zaman dilimi müşterilerin değer verdiği bir olgudur. Bu süre, zamanında teslim şeklinde ölçülmektedir. Müşteriye zamanında teslim süresi yanında, üretim sürecindeki sürenin azaltılması da önemli bir amaçtır.

⁹⁵ KOBU, s. 411.

⁹⁶ FİLİZ, Atilla, **Üretim Yönetiminde VERİMLİLİK SIRLARI**, Sistem Yayıncılık: 588, Birinci Basım, İstanbul, 2008, s. 52.

En geniş tanıma göre tüm süreç, bir müşteriden siparişin alınması ile başlamakta ve siparişin müşteriye teslimi ile sona ermektedir.⁹⁷

Üretim sürecinde yer alan ve üretimin gerçekleştirilmesindeki faaliyet ve işlemlerin süresi, toplam üretim süresini oluşturur.⁹⁸ Üretim firmaları tedarik zincirlerinin verimliliğini arttırmaya ve stokları azaltmaya çalıştıkça zamanında teslim etmenin önemi çalışmalarını etkilemektedir. Zorlayıcı kalite standartlarının yanı sıra üreticiler, tedarikçilerinden taahhüt edilmiş teslim süresini ya da ciddi anlamda yaptırımları karşılamaını beklemektedirler.⁹⁹

Hatalı üretim süresi teklifinin ana sebebi eksik üretim kapasitesi bilgisidir. Üreticiler bu nedenle kesin teslim süresi vermeyi kolaylaştıran, üretim süresini tahmin etmek için gerçek zamanlı üretim verilerini kullanan ERP ya da benzer yazılım araçlarını araştırmaktadırlar. Bu nedenle kesin üretim süresi hesabı başarılı teslim süresi garantisi için önkoşuldur.¹⁰⁰

Teslim süresini kısa vermek müşterileri kazanmayı sağlar ancak genellikle gecikmeyi beraberinde getirir. Teslim süresi uzun tutulduğunda ise teslimatın tamamlanması kolay olur ancak bu seferde müşteri daha kısa teslim süresi verene kayabilir. Bu açmazı aşmak için satış departmanı fabrikanın durumu hakkında kesin bilgi sahibi olmalıdır.¹⁰¹

2.4.2 Üretim Süresi Tahmininde Kullanılan Yöntemler

Üretim süresinin hesaplanması için siparişin ürün ağacı, üretim akışı ve parti büyüklüğü bilinmelidir.¹⁰² Çoğu üretim sistem modellerinde hammaddelerin matematiksel kolaylık olması açısından poisson dağılımında geldiği kabul edilir.

İş emirlerinin yerine getirilmesinde; makine arızaları, devamsızlık, verimsiz çalışma, ani ve önemli siparişlerin araya sıkıştırılması gibi önceden kestirilmesi güç

⁹⁷ FİLİZ, Atilla, s. 53.

⁹⁸ FİLİZ, Atilla, s. 55.

⁹⁹ SLOTNICK, s. 825.

¹⁰⁰ SLOTNICK, s. 826.

¹⁰¹ SLOTNICK, s. 826.

¹⁰² Oracle Applications, "Manufacturing Lead Time Computations", http://download-west.oracle.com/docs/cd/A60725_05/html/compls/us/bom/ldtms03.htm, (28.11.2009)

nedenlerle aksamalar olur.¹⁰³. Bu nedenle üretim süresi tahmininde bu değişkenler de dikkate alınmalıdır.

Çoğu çalışmada üretim süresi statik bitiş tarihi kuralları ile belirlenir. Genelde kullanılan iki kural İşlem Süresi ve Bekleme Süresinin Toplamı (Processing Plus Waiting Time) kuralı ve Toplam İş Hacmi (Total Work Content) kuralıdır.¹⁰⁴

Bir siparişin süresi F_i siparişin gelişi ile tamamlanması arasındaki zamandır. Siparişin tahmini için birçok teknik vardır. Birçok kaynaktan beslenen üretim sistemlerinde tahmin yapmak istendiğinde birçok problemle karşılaşılır. Her siparişin süresi önemli derecede farklılık gösterse ve sabit bir değer zayıf bir tahmin olsa da genellikle pratikte sabit üretim süresi (fixed lead time) kullanılır.¹⁰⁵

Sabit üretim süresi Conway tarafından 1965 yılında rapor edilmiştir. CON metodu olarak da bilinir.

Üretimde bir işin ne kadar sürede bitirileceğinin tahmini önemli bir problemdir. Bir malın üretim süresi tahmini belli adımlarda daha önce üretilmiş benzer malların üretim süreleri kıyaslanarak yapılabilir. Üretim süresini tahmin metodları Chung ve Huang tarafından simulasyon, regresyon modellerinin ya da veri madenciliğinin kullanıldığı istatistiksel analiz metodları, kuyruk kuramının temelini oluşturduğu analitik metodlar ve bunların bir veya birden fazlasını bünyesinde barındıran hibrid metodlar olarak 4 bölüm olarak karakterize edilmiştir. Bunlardan en çok kullanılanı simulyondur.¹⁰⁶

Backus et al, üretim süresini malın üretime girmeden önceki bekleme süresi ve malın sistemde geçirdiği sürenin toplamı olarak tanımlar. Burada bekleme süresi malın sisteme girme süresi ile bir önceki malın sistemden çıkma süresi arasındaki farktır. Little Kanununa göre normal koşullar altında bir malın ortalama üretim süresi (Cycle Time), sisteme giren ortalama mal sayısının (Work In Progress) birim zamanda üretilen mal sayısına (Throughput) oranıdır. Bunun için Nirmal Govind ve Theresa M. Roeder (2006) üretimde kalan işlerin koşullu güven aralığı tahmini için

¹⁰³ KOBU, s. 13.

¹⁰⁴ ENNS, Silvanus T., "Lead time selection and the behaviour of work flow in job shops", **European Journal of Operational Research** **109**, 1998, s. 128.

¹⁰⁵ MOSES et al, s. 3490.

¹⁰⁶ BACKUS et al, s. 252.

Response Surface istatistiksel metodunu kullanılmasını önermektedir. Govind ve Roeder'in amacı sistemde işlemleri bitmeden rotası tam olarak bilinmeyen her bir iş için beklenen kalan zamanı tahmin etmektir. Tahmin için simülasyon yöntemi kullanılabilir. Bununla birlikte işin rotasının belirsiz olduğu durumlarda simülasyon kullanarak yapılan tahmin kesin olamamaktadır.¹⁰⁷

Goodwin et al (2004) de üretim süresi tahmininde Little Kanunu kullanan statik lineer modelleme ve simülasyon tekniklerinin kullanılabileceğini bununla birlikte tahmin işlemi için statik modellerin stokastik değişkenleri kapsamadığını, simülasyonun ise çok miktarda hesaplama gerektirdiğini söylemektedir.¹⁰⁸

Üretim süresini belirleme yaklaşımları analitik ve deneysel olarak sınıflandırılabilir. Analitik metodların sonuçları kesin olmakla birlikte tanımlanması çok zordur.¹⁰⁹

Da Yin Liao ve Chia-Nan Wang (2004) öncelikli 300-mm otomatik malzeme yükleme işlemi için teslim süresi tahmininde yapay sinir ağları tabanlı bir yöntem önermektedir.¹¹⁰

Mansour Abdoli (2005), İş Akış Zamanları için Deneysel Bayes Tahmin Yöntemlerini araştırmıştır.¹¹¹

Abdulrahman Alenezi et al (2007), sipariş tipi bir üretim sisteminde gerçek zamanlı üretim süresi tahmini için destek vektör makinesi sınıflandırma metodunu kullanmıştır. Klasik zaman serileri modelleri sadece tamamlanmış siparişlerin

¹⁰⁷ GOVIND, Nirmal ve Theresa M. Roeder. "Estimating Expected Completion Times with Probabilistic Job Routing", **Proceedings of the 2006 Winter Simulation Conference**, s. 1810.

¹⁰⁸ GOODWIN, Randall, Russell Miller, Eugene Tuv, Alexander Borisov, Mani Janakiram ve Sigal Louchheim. "Advancements and Applications of Statistical Learning/Data Mining in Semiconductor Manufacturing", **Intel Technology Journal**, Vol. 8, No. 4, 2004, s. 326.

¹⁰⁹ GOVIND, s.1805.

¹¹⁰ LIAO, Da-Yin ve Chia-Nan Wang. "Neural-Network Based Delivery Time Estimates for Prioritized 300-mm Automatic Material Handling Operations", **IEEE Transactions on Semiconductor Manufacturing**, Vol. 17, No.3, 2004, s. 325.

¹¹¹ ABDOLİ, Mansour. ve Choobineh, F. Fred., "Empirical Bayes forecasting methods for job flow times", **IIE Transactions**, **Temmuz**, **2005**, <http://www.entrepreneur.com/tradejournals/article/print/133516519.html>, (05.02.2009).

geçmiş verileri kullanılarak tahmin yapılır. Bununla birlikte yeni siparişin üretim süresi üretim sisteminin o anki durumundan da etkilenir.¹¹²

Moses et al üretim süresi tahmininde 6 farklı metodu incelemiştir.¹¹³

1- CON Metodu:

$$F_i = K$$

2- TWK Metodu (Toplam İş Kapsamı):

Bu metod Conway tarafından 1965 yılında rapor edilmiştir. Uzun işlem süresi gerektiren siparişlerin tamamlanma sürelerinin de uzun olacağı görüşüne dayanır. Literatürdeki yaygınlığından dolayı deneysel kontrol amacı ile kullanılır.

$$\hat{F}_i = K \sum_{j \in R_i} P_{ij}$$

Burada j, i siparişinin R_i rotasındaki işlemini ve p_{ij} siparişin işlem süresini ifade eder.

3- JIS (Jobs in System) Metodu (Sistemdeki İşler)

Güncel Sistem trafiğinin dikkate alınması üretim süresin tahmininin kesinliğini artırır. Little Kanunu da siparişler sabit oranda geldiğinde ortalama üretim süresinin uzun vadede proses içi stoklarla orantılı olduğunu doğrulamaktadır.

$$\hat{F}_i = KN + \sum_{j \in R_i} P_{ij}$$

Burada N sistemdeki sipariş (iş) sayısıdır.

¹¹² ALENEZI, Abdulrahman, Scott A. Moses ve Theodore B. Trafalis, "Real-Time prediction of order flowtimes using support vector regression", **Computers and Operations Research**, Vol. 35, No. 11, 2008, ss. 3489, 3495.

¹¹³ MOSES, Scott A., Hank Grant, Le Gruenwald ve Simin Pulat. "Real-time due-date promising by build-to-order environments", **International Journal of Production Research.**, Vol. 42, No.20, s. 4353.

4- WIS (Work in System) Metodu (Sistemdeki İşler)

Bu parametrik metod da JIS metoduna benzer. Ancak güncel sistem trafiği için daha detaylı bir tahmin kullanmaktadır. Sistem trafiği, siparişlerin sayısı ile değil de siparişlerin toplam kalan iş süreleri ile hesaplanır.

$$\hat{F}_i = K \sum_k p'_k + \sum_{j \in R_i} p_{ij}$$

Burada p_k sistemdeki her k siparişi için kalan işlem süresini ifade etmektedir.

5- ESF Metodu (Exponentially Smoothing Flow Time) :Üssel Düzeltme Üretim Süresi

Bu parametrik olmayan metod bir üretim sistemindeki üretim sürelerinin birbirleri ile ilişkili olduğunu, bu nedenle yeni gözlemlenmiş üretim sürelerinin gelecekteki üretim süreleri için iyi bir tahminleyici olduğunu baz alır. Bununla birlikte siparişin üretim süresinin varyansının büyüktür bu nedenle üssel düzeltme tahmin metodunu geliştirir.

$$\overline{F}_k = \alpha F_k + (1 - \alpha) \overline{F}_{k-1}$$
$$\hat{F}_i = \overline{F}_k$$

Burada \overline{F}_k en son tamamlanmış k siparişinin üretim süresi, F_k k siparişinin tamamlandıktan sonra düzeltilmiş üretim süresi, α düzeltme sabitidir.

6- IFS (Incremental Forward Simulation) İleri Artan Simulasyon Metodu

IFS Metodu her bir siparişin tüm karakteristiklerini dikkate alan karmaşık bir methodtur. Bununla birlikte her sipariş geldiğinde yeni bir simulasyon başlatır ve bu nedenle sistem büyüklüğü arttıkça yavaştır ve pratikte uygulanması zordur.

ÜÇÜNCÜ BÖLÜM

KARAR AĞACI C4.5 ALGORİTMASINI KULLANARAK SU SAYACI ÜRETİM SÜRESİ TAHMİNİ

3.1 Araştırmanın Amacı

Ürünlerin istenen zamanda, istenen kalitede, istenen miktarda ve uygun fiyatla pazara çıkartılabilmesinin önemli olduğu günümüz üretim ortamında işletmeler kendilerine rekabet üstünlüğü sağlayacak her türlü aracı kullanmak istemektedirler.

Bu doğrultuda bir ürünün teslim zamanının doğru tahmin edilmesi, üretim ortamında beklenmeyen duruşların tahmini ve buna göre oluşabilecek aksiliklerin önceden belirlenmesi büyük önem kazanmaktadır.

Gelişen teknoloji ile üretimde otomasyonun yaygınlaşması üretim ortamından da verilerin alınabilir hale getirmiştir. Bu verileri kullanarak kendilerine avantaj sağlayabilen işletmeler rakiplerine göre bir adım önde olacaklardır.

Veri madenciliği teknikleri işletmelerde genellikle pazarlama ve reklam faaliyetlerinde kullanılmaktadır. Bu araştırmanın amacı veri madenciliğini üretimde de kullanılabilirliğini göstermektir.

3.2 Araştırmanın Yöntemi

Üretim süresini tahmin etmek amacı ile;

I = İşin tipi

j = istasyon numarası

m_{ij} =i işinin j istasyonundan bir sonraki istasyona gönderilme süresi

w_{ij} =i işinin j istasyonunda bekleme süresi

p_{ij} =i işinin j istasyonundaki işlem süresi

$S\{i\}$ =i işinin rotasındaki istasyon kümesi olmak üzere aşağıdaki denklem kullanılabilir.¹¹⁴:

$$F_i = \sum_{j \in S\{i\}} (s_{ij} + m_{ij} + w_{ij} + p_{ij})$$

Araştırmada ilk olarak su sayacı üretim süreci incelenmiştir ve veri olarak işletmenin CNC, Transfer ve Maça makinelerinin 2007, 2008, 2009 yıllarına ait verileri elde edilmiştir. Bu veriler ışığında ürün olarak o makinelere gelen yarı mamuller kabul edilmiş ve bu ürünün bu makinelerde harcayacağı zaman tahmin edilmeye çalışılmıştır.

Araştırmada işin bir sonraki istasyona gönderilme süresinin (m_{ij}) bilindiği varsayılmıştır. İşin ilgili istasyonda birim miktardaki işin işlem süresi (w_{ij}) aşağıdaki şekilde hesaplanmıştır.

(Toplam_Çalışma_Zamanı-Toplam_Duruşlar) / Ürün_Adedi

Araştırmanın asıl amacı tezgah dönüşleri, arızalar ve diğer duruşlar gibi ürünün beklemesini gerektiren sürenin tahmin edilmesidir.

Araştırmada kurulması ve anlaşılması kolay olması açısından karar ağacı tekniği kullanılmıştır. Karar ağacının oluşturulmasında 2007 yılının verileri kullanılmış, kurulan karar ağacı 2008 ve 2009 yıllarının verileri ile doğruluğu test edilmiştir.

Karar ağacı oluşturulurken aşağıdaki özellikler dikkate alınmıştır.

- Ürünün Adı,
- Makine tipi
- Vardiyada (gece, gündüz) üretildiği,
- Tezgahtaki bir önceki ürün,

Sınıf etiketi olarak makinelerdeki toplam bekleme zamanı 0, 45 dakikadan büyük ve 45 dakikadan küçük olarak tanımlanmıştır.

¹¹⁴ BAYKASOĞLU, Adil, Mustafa Göçken ve Zeynep D. Unutmaz. "New Approaches to due date assignment in job shops", **European Journal of Operational Research** 187, 2008, s. 34.

Araştırmanın son aşamasında C++ Builder ile üretimden gelen verileri MySQL veritabanına aktaran ve daha sonra açık kaynak kodlu WEKA yazılımına aktarılacak hale getiren yazılım geliştirilmiştir.

3.3 Su Sayacı Üretim Süreci

Su sayacı; kullanıcı (abone) tarafından tüketilen suyun tam miktarının doğru şekilde ölçülerek kaydedilmesini sağlayan en ideal ölçüm aletidir. Su sayacı, metal gövdesi ve plastik aksamı olarak iki kısımda incelenebilir.

Su sayacının metal gövdesinin üretim süreci;

- Maçahane,
- Dökümhane,
- Kesimhane,
- Boyahane,
- Talaşlı İmalat olarak 5 bölümden oluşur.

Plastik aksamın üretim süreci;

- Plastikhane,
- Mekanizma Montaj olarak 2 bölümden oluşur.

Su sayacının üretiminde son aşama su sayacı gövdeleri ile plastik aksamın montajının yapılması, sayaç kalibrasyonlarının yapılması ve paketlenmesi işlemidir. Su sayacının üretim süreci Şekil 7, Şekil 8, Şekil 9, Şekil 10'da gösterilmektedir.

3.3.1 Maçahane Bölümünde Üretim Süreci

Maça kumu değirmene atılır, karıştırılmaya başlanır. Kuma, Serter ve Reçine azar azar ilave edilir ve 3 dakika karıştırılır. Değirmen boşaltılır ve temizlenir. Oluşan karışım makinanın deposuna yüklenir. Makinanın ayarı ilgili parçaya göre yapılır. Makine çalıştırılır. İlk parça basılır. Operatör ilk parçayı kontrol eder ve gerekiyorsa ince ayarları bir daha kontrol eder. Sorun yoksa imalata devam eder. Maçalar Alkol içerisinde karışım olarak bulunan Grafit çözeltilisine daldırarak yüzeylerinin grafit tabakasıyla kaplanması sağlanır. Grafitle kaplanmış maçalar kasalarda istiflenir. Üretime ve stok sahasına sevk edilir.

3.3.2 Dökümhane Bölümünde Üretim Süreci

Kalıba ilgili maçanın alt kısmı koyularak malın dışarı akması engellenir ve kalıp kapatılır. Kalıba ergimiş metal yavaş yavaş dökülür. 20 sn beklenir, makina çalıştırılır, kalıp açılır, grafitte gönderilir. Dökümü yapılacak parçanın maçası kalıba yerleştirilir. Maça kalıba tam oturmalı, boşluğu olmamalı ve kumları dökülmemelidir. Kalıba yerleştirilen maça üzerine 30 cm uzaklıktan dökülen kumları temizlemek için hava tutulur. Döküm işlemi yapılır. Döküm işlemi bitince yani kalıp dik hale gelince malın kalıp içinde soğuması beklenir. Bu süre yaklaşık olarak 3-4 sn' dir. Maça kalıntıları kalıp içinden temizlenip, pedala basılarak kalıp grafit banyosuna gönderilir. Kalıbın kontrolü yapılır.

3.3.3 Kesimhane Bölümünde Üretim Süreci

Kesim ayarları yapıldıktan sonra kesim motoru çalıştırılır. Kesilmiş olarak önüne gelen iş parçasını alınır. Tekrar yükleme yapılarak kesim yapılır.

3.3.4 Boyahane Bölümünde Üretim Süreci

Boya tavaları fırına yerleştirilir. Fırın istenilen sıcaklığa getirilip, boyanan mamullerin pişme zamanı tamamlanınca, önce brülör kapatılır. Fan şalteri yaklaşık 10 dakika boşta çalıştırılır, böylece fırın içindeki sıcak havanın sirküle edilmesi sağlanır, dolayısıyla cehennemlik soğutulmuş olur, mallar soğumaya alınır. Fırın her çalıştırıldığında bu işlemler tekrarlanır.

3.3.5 Talaşlı İmalat Bölümünde Üretim Süreci

Transfer Tezgahları

Tezgahın üzerindeki işleme kafalarındaki tüm uçlar ve diğer kesici takımlar kontrol edilir. Hangi istasyonların çalıştırılacağı seçilir. Çalışacak istasyonların motorları çalıştırılır.

3.3.6 Plastikhane Bölümünde Üretim Süreci

Basılacak parçanın malzeme cinsi tablodan seçilir. Kullanılacak hammadde makinanın mal haznesine boşaltılır. Termostatların sıcaklık değerleri, hammaddenin çalışma sıcaklığına ayarlanır. Enjeksiyon basıncı ve miktarı ayarlanır. Ocak çalışma sıcaklığına geldiğinde makina boşta çalıştırılıp mal alıp verdiği kontrol edilir. Manuel olarak mal aldırılıp bir kaç baskı yapılır. Baskıdan çıkan ürünler kontrol edilip,

baskının istenildiği gibi olup olmadığı gözlenir. Gözlemlere göre kapama basıncı, enjeksiyon basıncı, mal alma ve sıcaklık ince ayarları yapılır. Otomatik veya yarı otomatik baskıya başlanır. Kullanılan malzemeye göre makine ayarları farklıdır.

3.3.7 Mekanizma Montaj Bölümünde Üretim Süreci

Plastik Enjeksiyondan gelen mamul ve yarı mamuller montaja alınır. Üst plakalara ve tamburlara mekanizma montaj elemanları tarafından serigrafi yapılır. Serigrafiden sonra üst plakalara burç takılır. Montaj elmanı yıldız dişlileri ve tamburları üst plakaya yerleştirir. Mekanizmanın üst plakasına dişlileri sıralanışına göre dizilir ve alt plakası takılır. Pervanelere mil ve safir taşı çakılır. Pervane çanaklarına milleri çakılır. Platformlara safir taşı çakılır. Platforma pervane ve pervane çanağının montajı yapılır. Son kontrol masasında ibre ve hassas göstergeleri takılır. Mekanizmanın son kontrolünde de ibreler sıfırlanır. Mekanizmanın montajı ve son kontrolleri yapılır. Sayaç montaja gönderilmeye hazır bir şekilde rafta bekletilir.

3.3.8 Sayaç Montaj Bölümünde Üretim Süreci

Çok Hüzmeli Sayaç Montajı:

Çok hüzmeli alt gövdeler montaj masasına dizilir. Giriş taraflarına tüp filtre takılır. Ayar vidası yerine takılır ve by-pass deliğinin 3/4'ünü kapatacak şekilde içeri doğru vidalanır. Ayar tapası ve derlin conta, vidanın ardından takılır ve tapa sıkılarak conta sıkıştırılır, sızdırmazlık sağlanır. Check-valve çıkış tarafına takılır, içeri doğru yerine oturtulur. Mekanizma montajdan gelen mekanizmanın altına pervane çanak contası takılır. Komple mekanizma sağ elde ve sayaç alt gövdesi ters çevrilmiş olarak sol elde olmak üzere, komple mekanizma kaması alt gövde de bulunan kama yuvasına gelecek şekilde sayacın içine monte edilir. Mekanizma sıfırlanır. Mekanizma tutucu rondela ve çemberin dışına conta yerleştirilir. Polikarbonat camın yivleri contaya basacak şekilde (bombeli kısım yukarıya bakacak) konulur. Camın üzerine kaydırıcı pul yerleştirilir. Üst gövdeye kapak yerleştirilir. Alt ve üst gövdeleri birbirlerine montajlanır. Sayacın montajı tamamlandıktan sonra sayaçlar istenilen tork değerinde sıkılmak üzere Üst gövde sıkma makinasına gönderilir.

Tek Hüzmeli Sayaç Montajı:

Talaşlı imalattan işlenmiş alt gövdeler montaj masalarına dizilir. Tek hüzmeli alt gövdeye ayar vidası montaj yapılır. Tek hüzmeli alt gövdenin su girişine filtre yerleştirilir. Montaj elmanı alt gövdenin içine mekanizmayı yerleştirir. Mekanizma üzerine contasını yerleştirir. Sızdırmazlık contasının üzerine camını yerleştirir. Üst gövdeye kapak yerleştirir. Alt ve üst gövdeleri birbirlerine montajlar. Sayacın montajı tamamlandıktan sonra sayaçlar istenilen tork değerinde sıkılmak üzere Üst gövde sıkma makinasına gönderilir.

3.3.9 Üst gövdenin sıkılması

Ana şalter açılır ve hidrolik motor çalıştırılır. Üst gövde aparatı ilgili sayaca göre değiştirilir. İlgili sayaçta istenen sıklığa göre sıkma momenti için hidrolik motor basıncı ayarlanır.

Sayaç ters olarak aparata yerleştirilir. Rakor başları ile sayaç sıkma kolları birbiri ile çakışmamalıdır. Motor sıkma işlemine başlar. Sıkma işlemi devam ederken manometre yükselmeye başlar ve bir noktada işlem biter.

3.3.10 Kalibrasyon Bölümünde Üretim Süreci

Ayar masasına test edilecek anma çapına uygun bağlama aparatları yerleştirilir. Bağlantı bölgelerindeki o-ringlerin sağlam olduğu kontrol edilir. Sayaçlar panoya su akış yönüne dikkat edilerek, seri olarak yerleştirilir. Pnömatik vana açılır ve sıkıştırma butonu ile sayaçlar panoda sıkıştırılarak sabitlenir. Panoda bir su kaçağı olup olmadığı kontrol edilir. Test için gerekli miktardaki suyun akışı teraziden takip edilir, istenilen miktarda su akışı gerçekleştiğinde debimetre vanası kapatılarak test sonlandırılır.

Endeks okuma için, sayaçların son endeksleri ve terazide görülen gerçek su miktarı okunur. Teraziden okunan değer (kg) sıcaklığa göre suyun yoğunluğuna bölünerek hacimsel değeri bulunur.

Sayacın gösterdiği hacim ve gerçek hacim arasındaki farka göre sayacın hatası hesaplanır. Çok hassas ölçümler için suyun sıcaklığına göre yoğunluğu hesaba katılır.

$$[(\text{Son Endeks} - \text{İlk Endeks}) - \text{Gerçek Hacim}] * 100 / \text{Gerçek Hacim}$$

Kazandaki su tahliye edilir. Sayaçlar hata değerlerine göre, ayar vidalarından kalibre edilirler ve daha sonra kalibrasyonlarını kontrol etmek amacıyla test tekrarlanır. Sayaçların kalibrasyonu tamamlandıktan sonra, Qmax veya Qn, Qt ve Qmin debilerinde test edilerek hata değerleri tespit edilir. Basınç tahliye vanası açılıp kapatılarak sistemdeki basınçlı su salınır. Sistemde su basıncı kalmadığından emin olunduktan sonra sıkıştırma pistonu açılarak sayaçlar gevşetilir ve sayaçlar panodan sökülür.

3.3.11 Paketleme Bölümünde Üretim Süreci

Kurşun Takma:

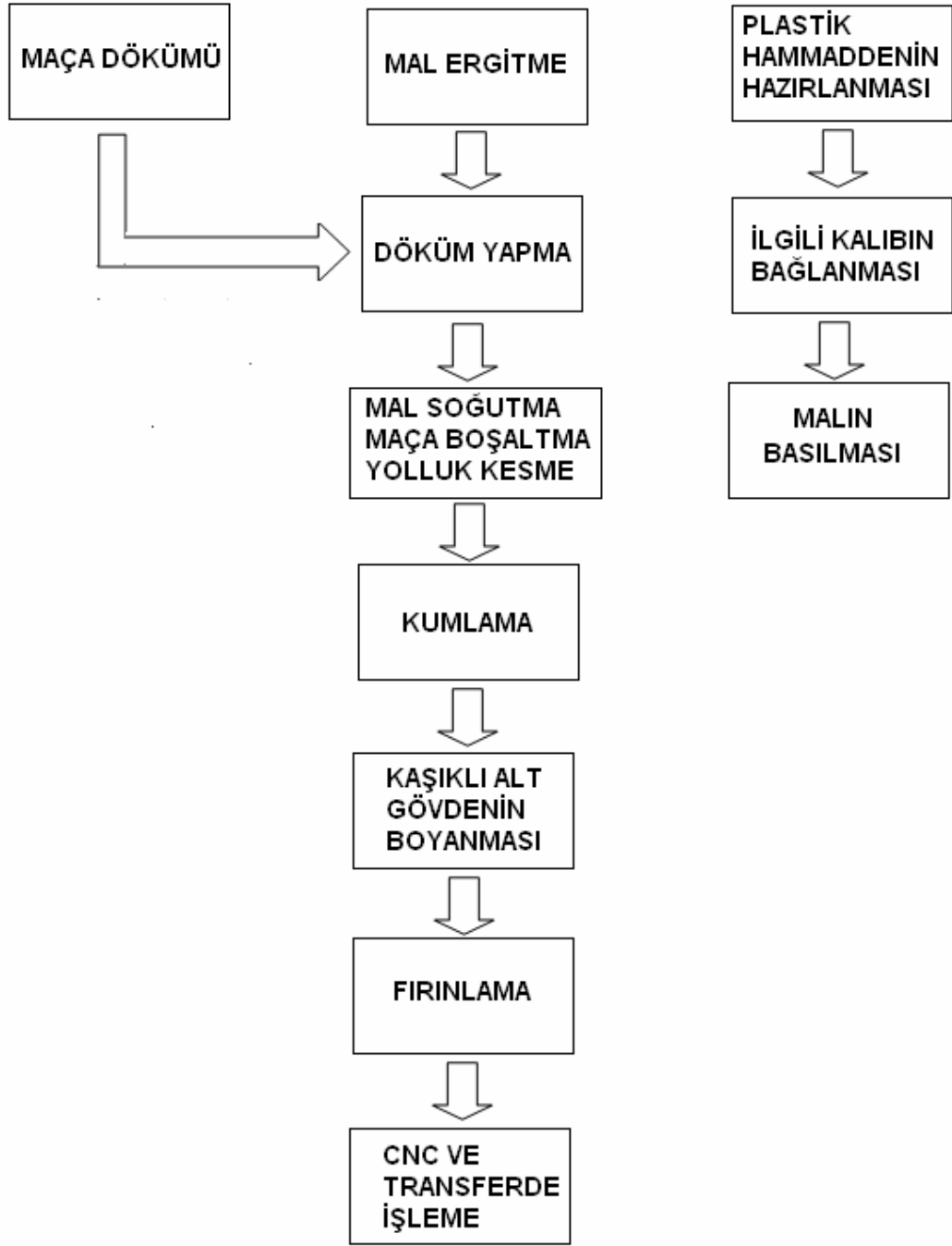
Mühür teli alt gövdenin ayar tapası deliğinden geçirilip ortalanır. Bir uç üst gövdedeki en yakın mühür deliğinden geçirilir. Diğer uç ile birleştirilip, 3-4 defa bükülür. Uçlar tekrar ayrılır. Kurşun iki delikli olan tarafından ayrı ayrı geçirilir. Tek delikli kısımdan birlikte çıkan iki uç 3 defa daha bükülür. Fazla tel kesilir.

Rafa Dizme:

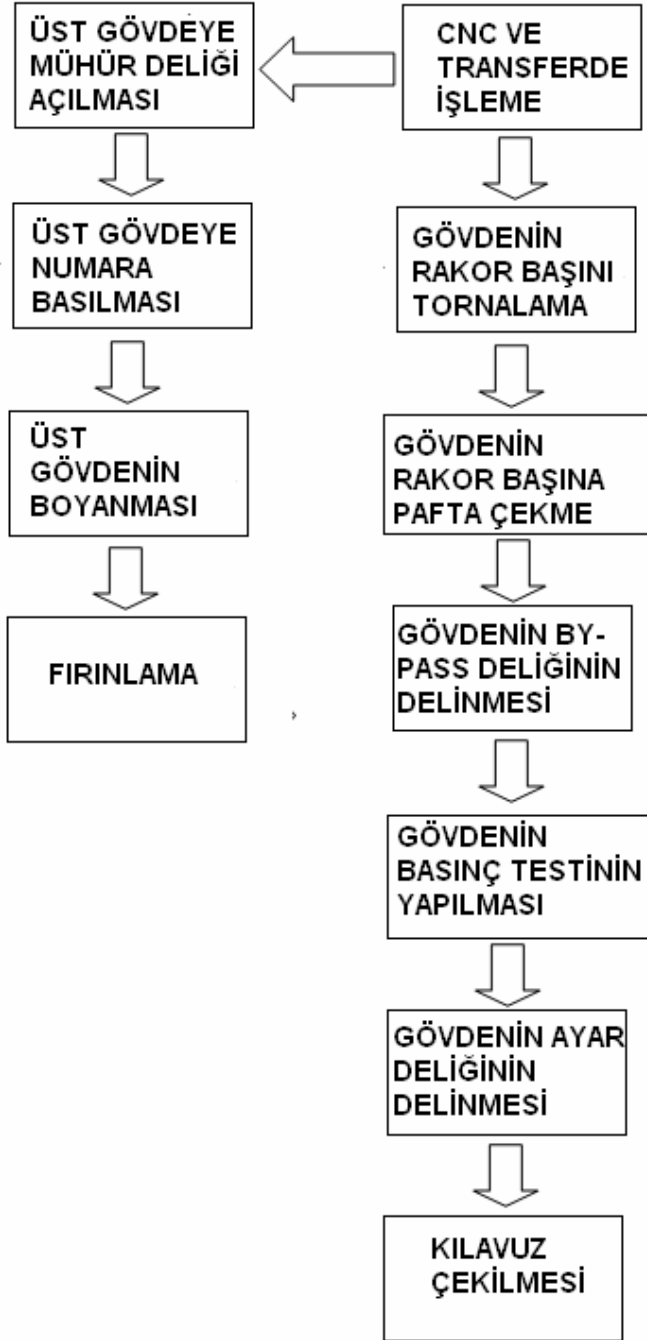
0'dan 10'a kadar olmak üzere 10'arlı guruplar halinde, mühür kurşunu solda olacak şekilde sayaçlar dizilir.

Paketleme:

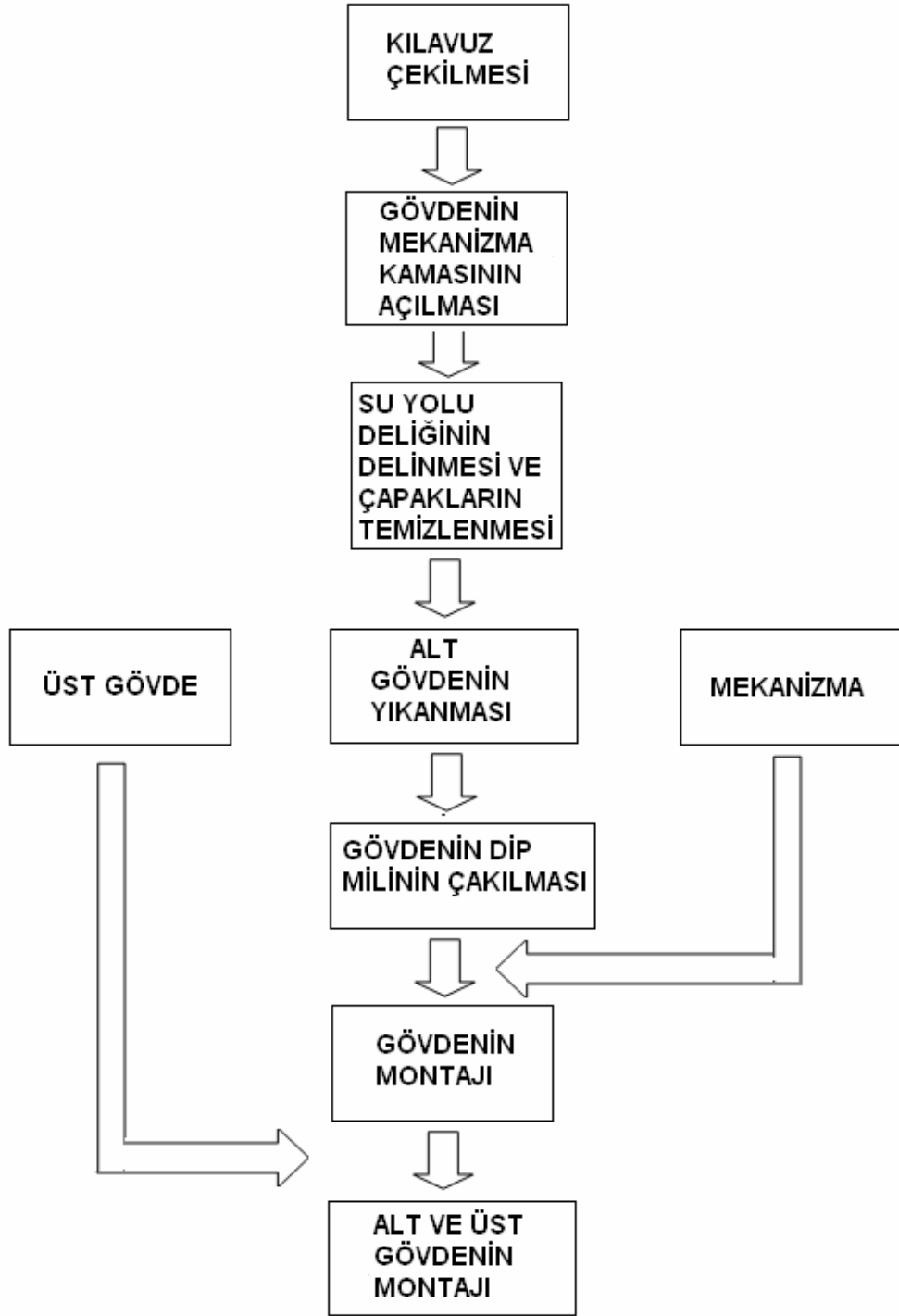
Mühürlenmiş sayaç kendine ait kutuya konur. Kutuya paketlenmiş durumda 2 takım rakor ve 2 plastik conta konur.(1 takım=1 somun +1 tüp). Garanti belgesi kutuya yerleştirilir. Kutunun kapağı kapatılır. Büyük kutunun tabanı bantlanarak hazırlanır. Kutulanmış 10 sayaç rafa dizilme sırasına göre (seri no. ya göre) büyük kutuya yerleştirilir. Büyük kutunun kapağı kapatılıp bantlanır. 10 arlı sayaç serisinin serisi dış kutunun üzerine yazılır. Kutunun enine iki adet olmak üzere makine yardımıyla şerit çekilir.



Şekil 7. Su Sayacı Üretim Süreci-1



Şekil 8. Su Sayacı Üretim Süreci-2



Şekil 9. Su Sayacı Üretim Süreci-3



Şekil 10. Su Sayacı Üretim Süreci-4

3.4 CNC, Maça ve Transfer Makinelerinin Üretim Süresi Tahmini

CNC, maça ve transfer makinelerinin üretim süreleri tahmin edilmek için aşağıdaki adımlar izlenmiştir.

- Verilerin firmadan Excel ortamında alınması,
- Verilerin MySQL veritabanına aktarılması,
- Geliştirilen yazılım ile eğitim verilerinin WEKA formatına aktarılması,
- WEKA yazılımı ile karar ağacının oluşturulması,
- Geliştirilen yazılım ile karar ağacına göre kuralların oluşturulması,
- Kuralların daha önce kullanılmamış yeni veri seti üzerinde denenmesi ve doğruluğunun hesaplanması,

Geliştirilen yazılım ile MySQL veritabanında tutulan veriler daha sonra karar ağacı oluşturmak üzere WEKA yazılımının istediği formata dönüştürülmüştür. WEKA

ile karar ağacı oluşturulduktan sonra ilgili karar ağacına istinaden kurallar oluşturulmuştur.

3.4.1 Yazılım

Üretim verilerini MySQL veritabanına aktaran yazılım C++ Builder ile geliştirilmiştir. Yazılımda ilk olarak üretim ortamından gelen verilerin MySQL veritabanına aktarımı sağlanmıştır. Geliştirilen yazılımın aktarma modülü Şekil 11'de gösterilmektedir.

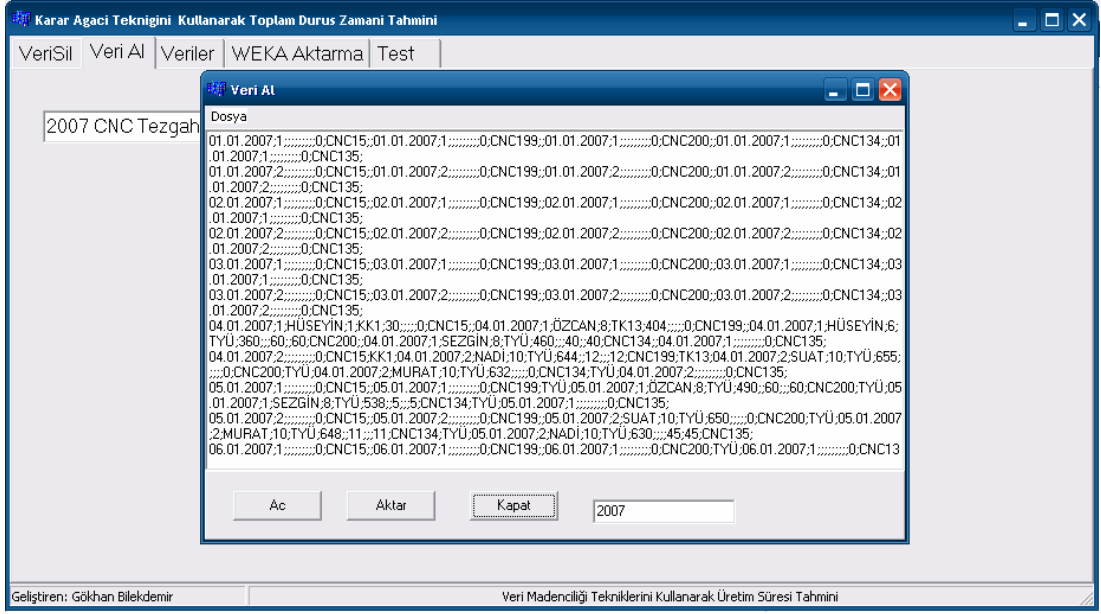
MySQL veritabanına gelen veriler üzerinde veri ön işleme teknikleri kullanılmış, eksik veya hatalı veriler düzeltilmiştir. Ürün çeşitliliğini sadeleştirmek için ürünler gruplandırılmıştır.

MySQL veritabanına aktarılan verilerden her makine için T ve D ürünleri için işlem süreleri belirlenmiştir.

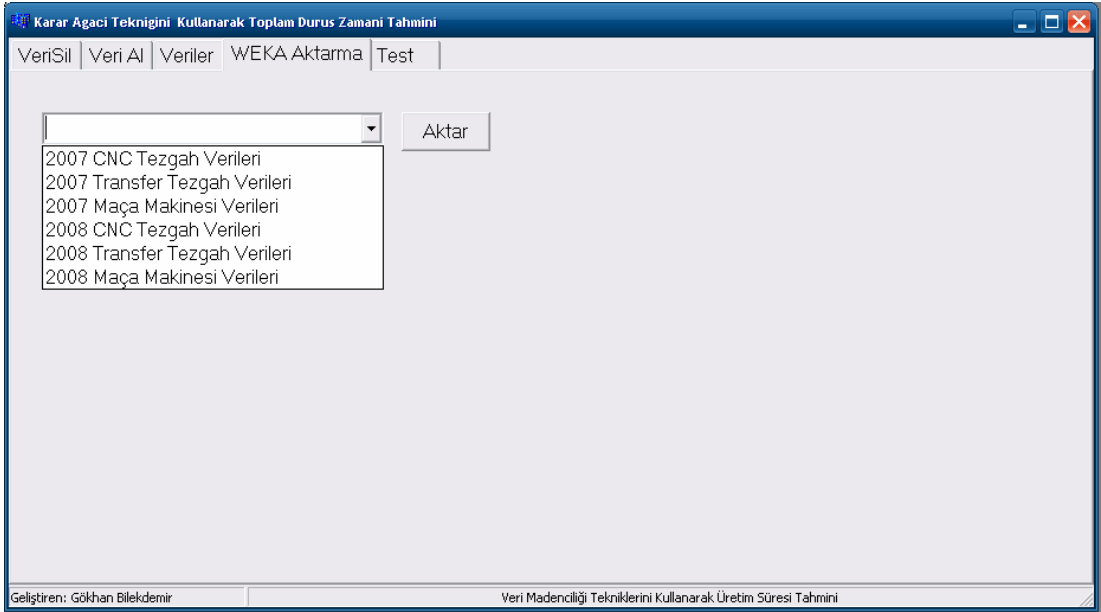
Daha sonra veriler yazılım tarafından WEKA yazılımının istediği formata getirilmiştir. Geliştirilen yazılımın WEKA aktarma menüsü Şekil 12'deki gibidir.

WEKA bir proje olarak başlayıp bugün dünya üzerinde birçok insan tarafından kullanılmaya başlanan bir Veri Madenciliği uygulaması geliştirme programıdır. WEKA java platformu üzerinde geliştirilmiş açık kodlu bir programdır. Arff, Csv, C4.5 formatında bulunan dosyalar WEKA'da import edilebilir. Herhangi bir text dosyasındaki verileri WEKA ile işlemek olanaksızdır. Ayrıca Jdbc kullanılarak veritabanına bağlanıp burada da işlemler yapılabilir. WEKA'nın içerisinde Veri İşleme, Veri Sınıflandırma, Veri Kümeleme, Veri İlişkilendirme özellikleri mevcuttur.¹¹⁵

¹¹⁵ DENER, Murat, Murat Dörterler ve Abdullah Orman, "Açık Kaynak Kodlu Veri Madenciliği Programları: WEKA'da Örnek Uygulama", Akademik Bilişim, 2009, Harran Üniversitesi, Şanlıurfa, <http://ab.org.tr/ab09/bildiri/42.pdf>, (10.01.2010), ss. 789, 790.



Şekil 11. Verilerin yazılım tarafından alınması



Şekil 12. Verilerin WEKA'ya aktarılması

WEKA yazılımına aktarılan verilere C4.5 algoritması ile karar ağacını oluşturan J48 algoritması seçilmiştir.

Eğitim verileri olarak 2007 yılının verileri seçilerek karar ağacı oluşturulmuştur. Oluşturulan karar ağacı ile ilgili veriler ektedir.

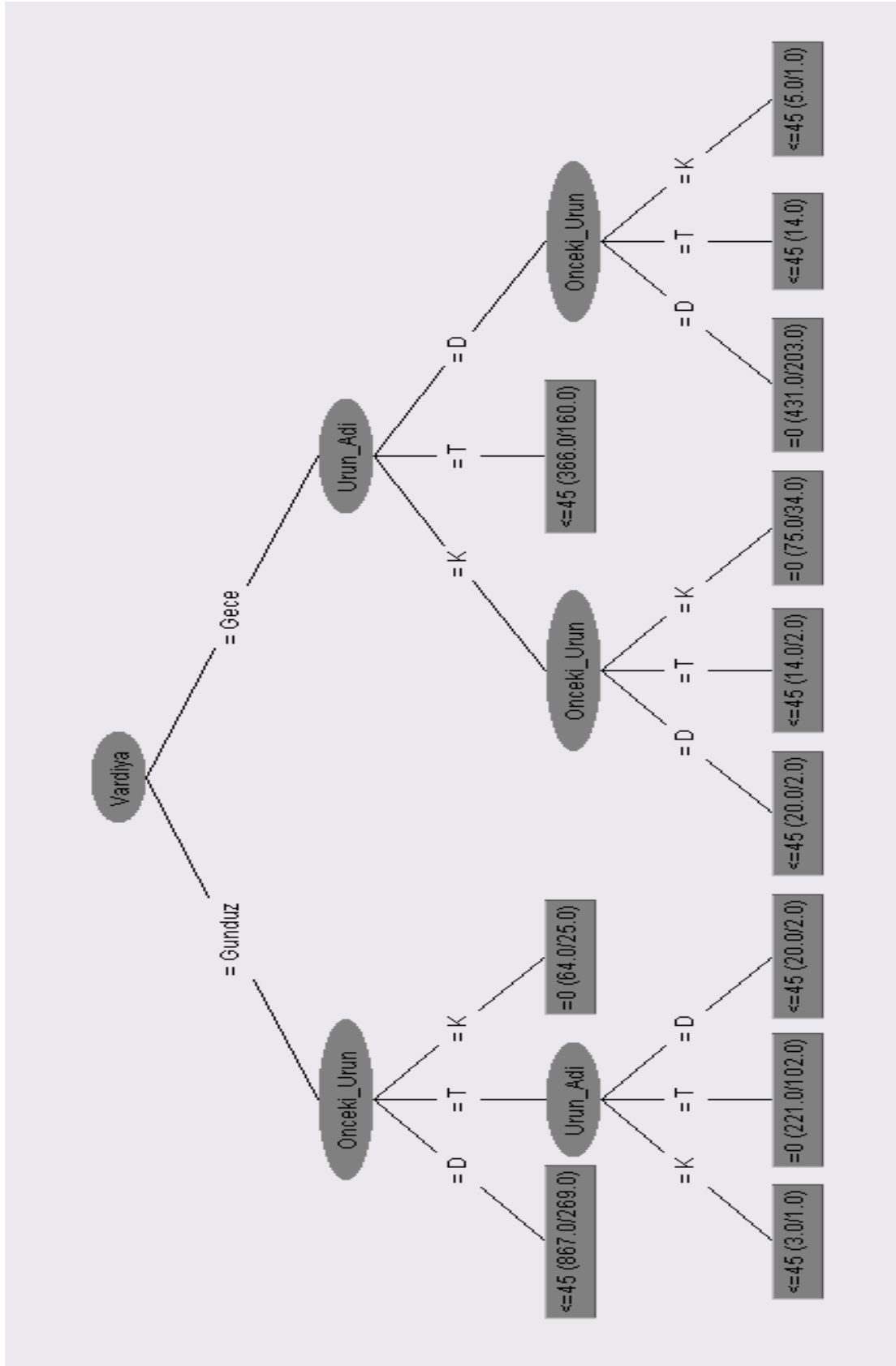
Eğitim verileri ile karar ağacı oluşturulduktan sonra kurallar oluşturularak elde edilen kuralların test verileri ile doğruluğunun test edilmesi gerekmektedir. Oluşturulan kurallar 2008 verileri ile test edilmiştir.

3.4.2 CNC Makinesi ile ilgili oluşturulan karar ağacı ve sınıflama kuralları

CNC makinelerinin ürünlere göre işlem süreleri Tablo 4'teki gibidir. CNC makineleri ilgili eğitim verileri WEKA yazılımına aktarılmıştır. WEKA tarafından oluşturulan karar ağacı Şekil 13'te gösterilmektedir. Bu karar ağacına göre “**IF THEN**” yapısına göre kurallar oluşturulmuştur. Oluşturulan karar ağacının kuralları Şekil 14'deki gibidir.

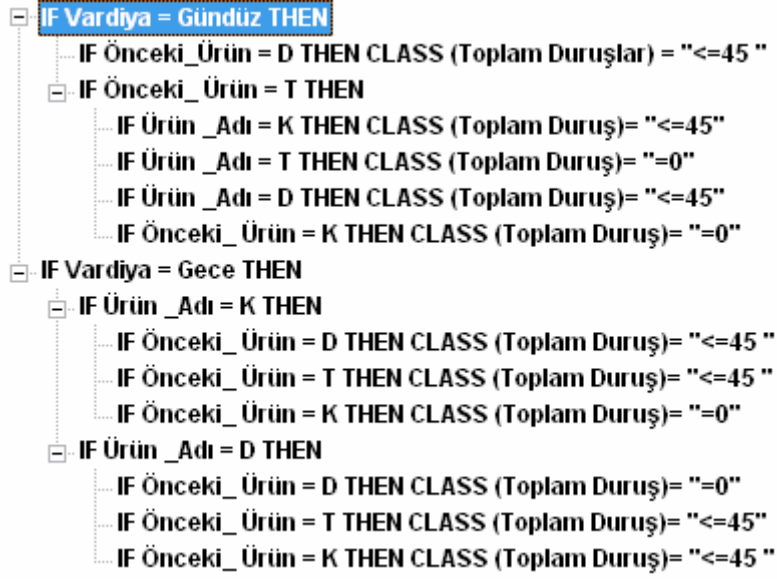
Tablo 4. CNC Makineleri İşlem Süreleri

CNC				
	2007		2008	
	T	K	T	K
CNC 135	1,11	0,92	0,85	0,78
CNC 15	1,04	1,56	0,89	1,15
CNC134	0,97	0,95	0,89	0,85
CNC199	1,06	0,89	0,90	0,78
CNC200	1,28	1,22	0,92	1,09



Şekil 13. CNC Makinesi Karar Ağacı

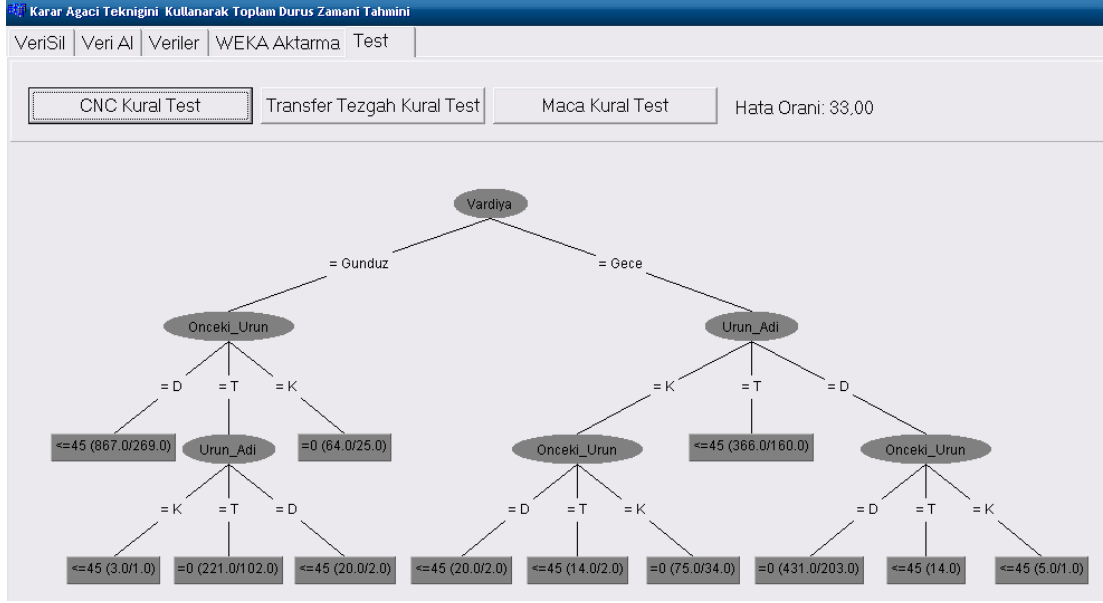
Karar ağacı ile oluşturulan kurallar aşağıdaki gibidir:



Şekil 14. CNC Makinesi Karar Ağacı Kuralı

CNC makinesi ile ilgili oluşturulan kurallar 2008 yılı verileri test edilmiş ve kurulan modelin 2008 verilerinde %67 oranında doğru tahmin ettiği gözlemlenmiştir. Geliştirilen yazılımın hesaplama menüsü Şekil 15'teki gibidir.

Bu karar ağacına göre oluşturulan kurala göre örneğin "T" ürününün üretilmesi planlandığında ürünün gece üretilmesi durumunda ve tezgahdaki ürün ne olursa olsun tezgah bekleme süresi 45 dakikadan küçük olacaktır. Aynı ürün gündüz üretilmesi durumunda tezgahdaki ürün "T" ise tezgah bekleme süresi 0, tezgahdaki ürün "D" ise bekleme süresi 45 dakikadan küçük, tezgahdaki ürün "K" olması durumunda ise 0 olacaktır.



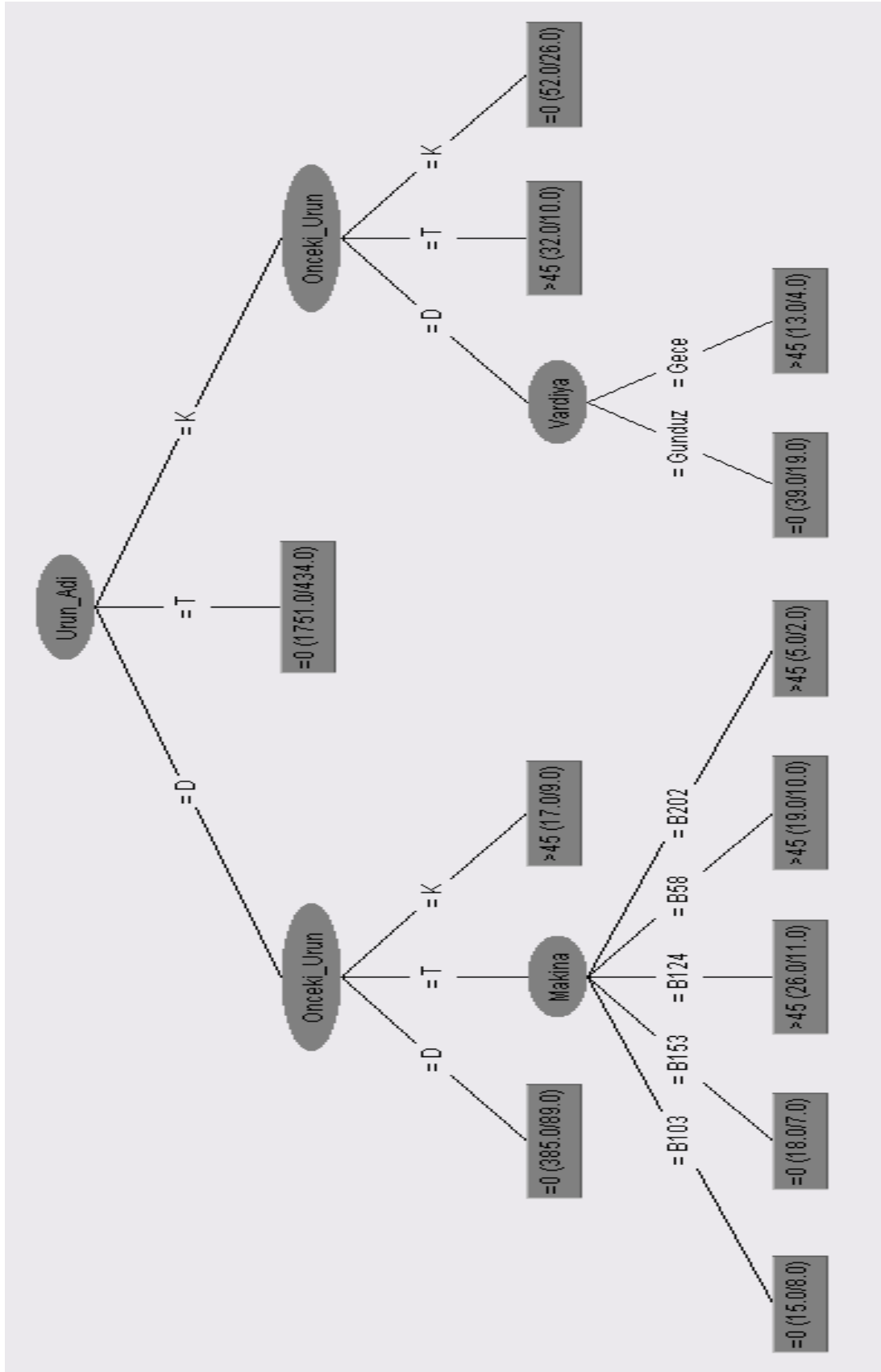
Şekil 15. CNC Makinesi Kural Test

3.4.3 Maça Makinesi ile ilgili oluşturulan karar ağacı ve sınıflama kuralları

Maça makinelerinin ürünlere göre işlem süreleri Tablo 5'teki gibidir. Maça makineleri ilgili eğitim verileri WEKA yazılımına aktarılmıştır. WEKA tarafından oluşturulan karar ağacı Şekil 16'da gösterilmektedir. Bu karar ağacına göre "IF THEN" yapısına göre kurallar oluşturulmuştur. Oluşturulan karar ağacının kuralları Şekil 17'deki gibidir.

Tablo 5. Maça Makinesi İşlem Süreleri

MAÇA				
	2007		2008	
	T	K	T	K
B103	0,76	0,40	0,67	0,45
B153	0,60	0,43	0,55	0,34
B124	0,68	0,36	0,57	0,35
B58	0,81	0,43	0,68	0,39
B202	0,58	0,36	0,61	0,49



Şekil 16. Maça Makinesi Karar Ağacı

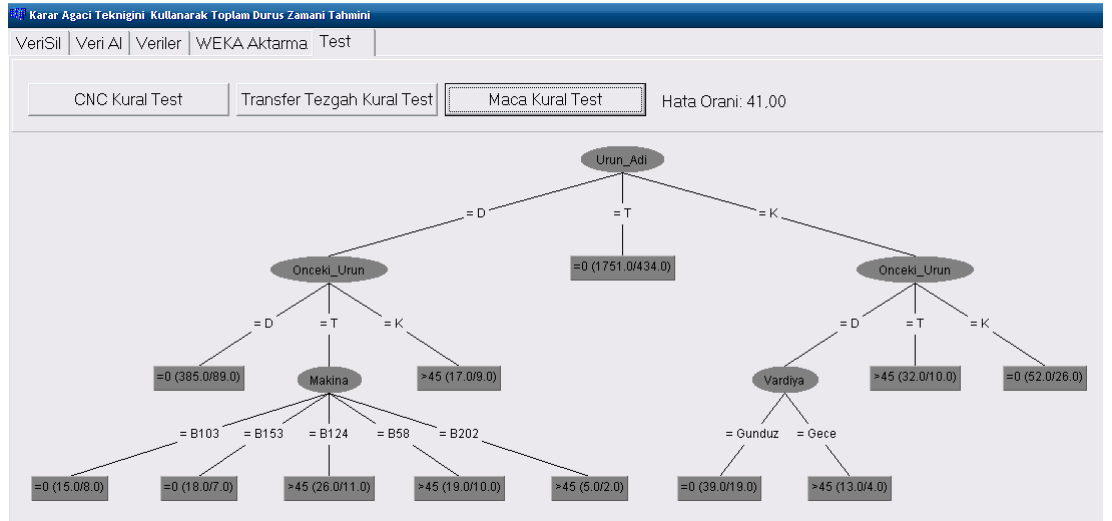

```

IF Ürün_Adi = D THEN
  IF Önceki_Ürün = D THEN CLASS (Toplam Duruş)="0"
  IF Önceki_Ürün = T THEN
    IF Makine = B103 THEN CLASS (Toplam Duruş)="0"
    IF Makine = B153 THEN CLASS (Toplam Duruş)="0"
    IF Makine = B124 THEN CLASS (Toplam Duruş)=">45"
    IF Makine = B58 THEN CLASS (Toplam Duruş)=">45"
    IF Makine = B202 THEN CLASS (Toplam Duruş)=">45"
  IF Önceki_Ürün = K THEN CLASS (Toplam Duruş)=">45"
  IF Ürün_Adi = T THEN CLASS (Toplam Duruş)="0"
  IF Ürün_Adi = K THEN
    IF Önceki_Ürün = D THEN
      IF Vardiya = Gündüz THEN CLASS (Toplam Duruş)="0"
      IF Vardiya = Gece THEN CLASS (Toplam Duruş)=">45"
    IF Önceki_Ürün = T THEN CLASS (Toplam Duruş)=">45"
    IF Önceki_Ürün = K THEN CLASS (Toplam Duruş)="0"

```

Şekil 17. Maça Makinesi Karar Ağacı Kuralı

Maça makinesi ile ilgili oluşturulan kurallar 2008 yılı verileri test edilmiş ve kurulan modelin 2008 verilerinde %59 oranında doğru tahmin ettiği gözlemlenmiştir. Geliştirilen yazılımın hesaplama menüsü Şekil 18'deki gibidir.



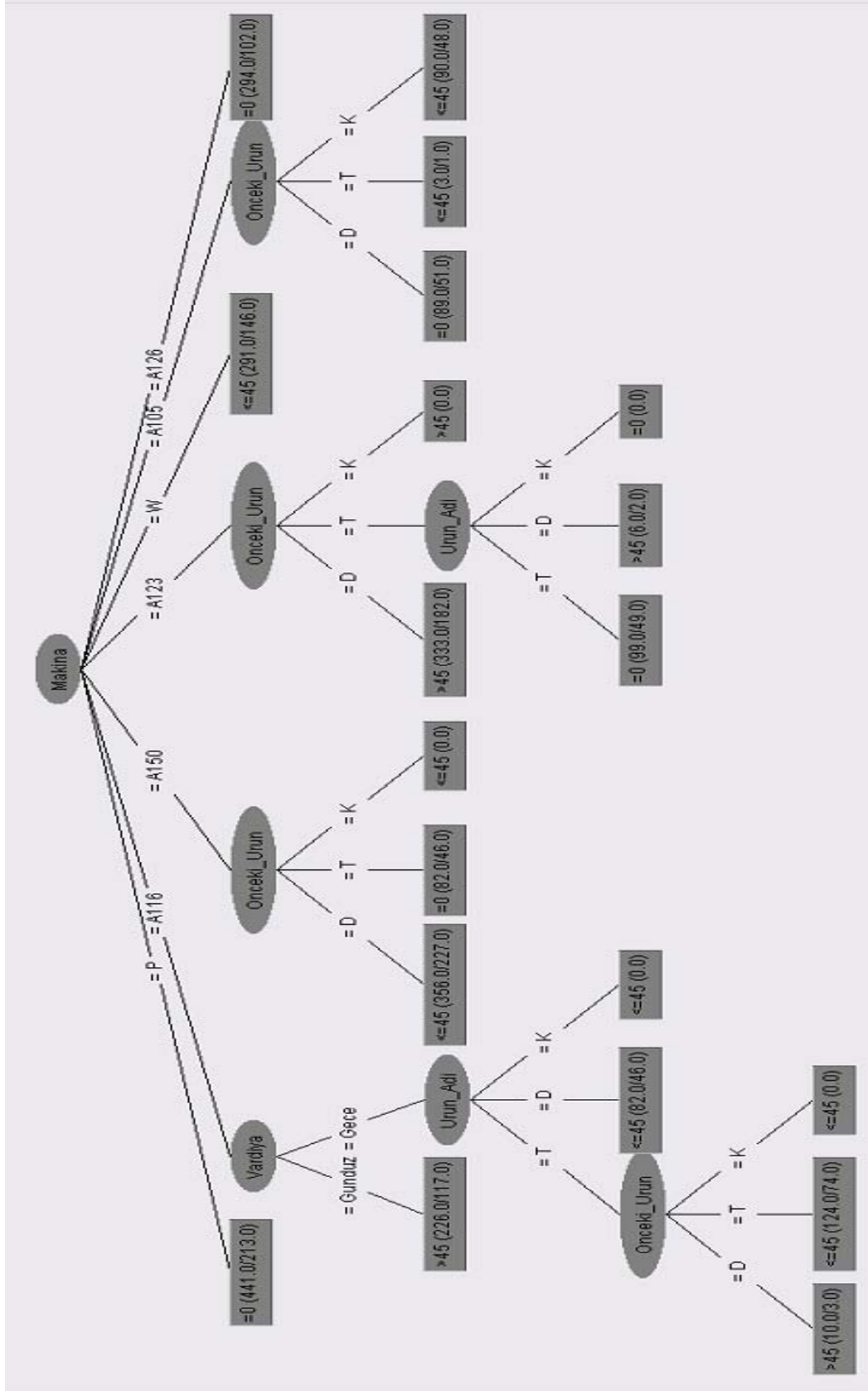
Şekil 18. Maça Makinesi Kural Test

3.4.4 Transfer Makinesi ile ilgili oluşturulan karar ağacı ve sınıflama kuralları

Transfer makinelerinin ürünlere göre işlem süreleri Tablo 6'daki gibidir. Transfer makineleri ilgili eğitim verileri WEKA yazılımına aktarılmıştır. WEKA tarafından oluşturulan karar ağacı Şekil 19'da gösterilmektedir. Bu karar ağacına göre "IF THEN" yapısına göre kurallar oluşturulmuştur. Oluşturulan karar ağacının kuralları Şekil 20'deki gibidir.

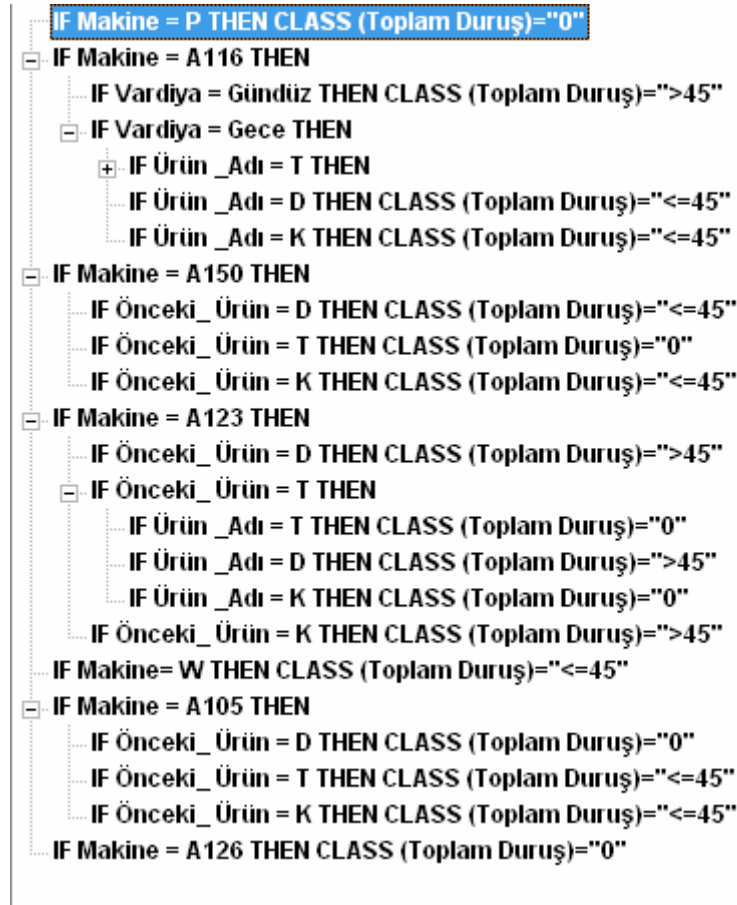
Tablo 6 Transfer Tezgahları İşlem Süreleri

TRANSFER				
	2007		2008	
	T	K	T	K
W	0,80	-	0,81	0,84
P	0,80	-	0,86	-
A116	0,82	-	0,92	-
A123	0,80	-	1	-
A150	0,92	-	0,75	0,53
A105	0,74	0,78	0,32	0,86
A126	0,40	-	0,38	-



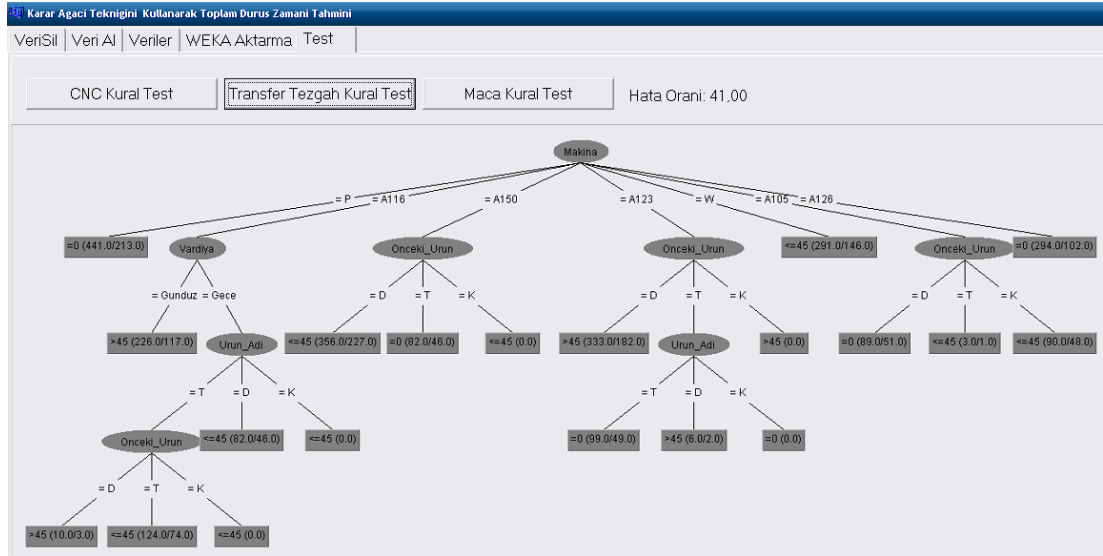
Şekil 19. Transfer Tezgâhı Karar Ağacı

Karar ağacı ile oluşturulan kurallar aşağıdaki gibidir:



Şekil 20. Transfer Makinesi Karar Ağacı Kuralı

Transfer makinesi ile ilgili oluşturulan kurallar 2008 yılı verileri test edilmiş ve kurulan modelin 2008 verilerinde %59 oranında doğru tahmin ettiği gözlemlenmiştir. Geliştirilen yazılımın hesaplama menüsü Şekil 21'deki gibidir.



Şekil 21. Tranfer Tezgahı Kural Test

SONUÇ VE ÖNERİLER

Veri Madenciliği çok sayıda veri içerisinde anlamlı bilgilerin çıkarılması işlemidir. Bilgi keşfi sürecinin bir parçası olan veri madenciliği, birçok disiplin ile etkileşim içerisinde. Bu nedenle çok çeşitli veri madenciliği modelleri bulunmaktadır. Bu modeller tanımlayıcı ve tahmin edici modeller olarak iki kısımda incelenebilir.

Gün geçtikçe zorlaşan rekabet ortamında işletmeler, müşterilerin beklentilerini tam olarak karşılamak için kaliteli, düşük maliyetli ve zamanında teslim edilen ürünler üretmek istemektedirler. Bu doğrultuda gelen siparişin teslim süresini doğru tahmin etmek işletmeler için önem kazanmaktadır.

Bu uygulamanın amacı veri madenciliği tekniklerinden karar ağacı tekniğini kullanarak üretim süresini tahmin etmektir. Bu amaçla üretim ortamından gelen veriler doğrultusunda karar ağacı tekniği ile sınıflandırma kuralları oluşturulmaya çalışılmıştır. Daha sonra bu sınıflama kuralları kullanılarak bir sipariş durumunda ürünün yapısı, ne zaman üretileceği gibi özellikler dikkate alınarak ilgili siparişin belli makinelerdeki bekleme süreleri tahmin edilmeye çalışılmıştır.

CNC makinelerinden elde edilen veriler incelendiğinde ürünün hangi makinede üretildiğinin bekleme süresine etkisinin olmadığı görülmüştür. Ürünün bekleme süresini etkileyen etmenler olarak, ürünün cinsi, gece veya gündüz vardiyasında yapılması, o an tezgahta işlenen ürünün cinsi sayılabilir.

Karar ağacında bir ürünün bekleme süresinin tezgahdaki ürüne bağlı olduğu görülmektedir. Örneğin gündüz vardiyasında T ürünü, gece vardiyasında da D ve K ürünlerinin bekleme süreleri incelendiğinde sipariş edilen ürün ile tezgahdaki ürün aynı ise bekleme süresi 0 olarak tanımlanmıştır. Tezgahta aynı ürün işleneceği için makine ayar süresi ya da tezgah dönüşleri olmayacağı için bekleme süresi de olmayacaktır.

Bu uygulamada karar ağacının okunabilir hale getirmek için ürünler üç grupta toplanmıştır. Gerçek üretim ortamında su sayacının ürün çeşitliliği çok fazladır. Bu durum karar ağacının kesinliğini azaltmaktadır.

Veri ön işleme aşamasında 5 adet CNC makinesinden toplam 2007 yılı verileri olarak 3100 veri alınmıştır. Bu verilerden toplam çalışma zamanı 0'dan büyük olan 1518 adedi dikkate alınmıştır.

2007 yılı CNC verilerinde 65 farklı ürün tipi, 2007 yılı maça verilerinde 101 farklı ürün tipi, 2007 transfer makinesi verilerinde 69 farklı ürün tipi gözlemlenmiştir. Bu ürünler T tipi K tipi ve D tipi olmak üzere üç ana gruba toplanmıştır.

2007 yılı CNC verilerinde 22, 2007 yılı Maça verilerinde 37, 2007 yılı Transfer makinesi verilerinde 31 farklı operatör ismi geçmektedir. Uygulamada operatör alanı dikkate alınmamıştır.

Geçmiş verilerde yapılan işin hangi operatör tarafından yapıldığı verisi bulunmasına rağmen uygulamada bu özellik dikkate alınmamıştır. İşletmeden operatörlerle ilgili ne kadar zamandır o işletmede çalıştığı ya da eğitim seviyesi gibi veriler elde edilebilmiş olsa idi bu durumun dikkate alınması çok daha kararlı bir sınıflama kuralının oluşturulmasını sağlamış olacaktır.

Vardiya saatleri gece ve gündüz olarak iki kısım olarak tanımlıdır.

Makine ayar süreleri, tezgah dönüşleri ve diğer duruşlar toplam duruşlar adı altında toplanmıştır.

Bu çalışmada veri madenciliği tekniklerini kullanarak üretim süresinin tahmin edilebileceği gösterilmiştir. İşletmeler geçmiş veriler ışığında makinelerini ve üretim sistemlerinin modellerini oluşturabilirler. Üretim süresini etkileyen faktörlerin hangi nedenlerden oluştuğunu analiz edip, gerekli önlemleri alarak teslim sürelerinin daha doğru verebilirler. Ancak üretim süresinin tahmininde performansın artırılabilmesi için makine verilerinin yanında işçiler ile ilgili verilerin de toplanması ve hesaba katılması gereklidir. Bu nedenle işletmeler analizi yapmadan önce gerekli zaman ve metot etütlerini yaparak gerekli verileri toplamaları gereklidir.

Bu çalışmada karar ağaçları ile sınıflandırma tekniği kullanılarak tahmin yapılmaya çalışılmıştır. Karar ağaçları ile sürekli tipteki değişkenlerin tahmini yapılamamasından ötürü tahmin için çalışma yapay sinir ağları, çizelgeleme problemlerinin çözümü için de genetik algoritmalar gibi başka veri madenciliği teknikleri kullanılarak geliştirilebilir.

KAYNAKLAR

ABDOLİ, Mansour. ve Choobineh, F. Fred., "Empirical Bayes forecasting methods for job flow times", **IIE Transactions**, Temmuz, 2005, <http://www.entrepreneur.com/tradejournals/article/print/133516519.html>, (05.02.2009).

AJOKU, Pamela N. ve Bart Nnaji, "Improving Productivity in Manufacturing Environments Using Data Mining", <http://ieweb.uta.edu/vchen/AIDM/AIDM-Ajoku.pdf>, 10.02.2009.

ALENEZI, Abdulrahman, Scott A. Moses ve Theodore B. Trafalis, "Real-Time prediction of order flowtimes using support vector regression", **Computers and Operations Research**, Vol. 35 , No. 11, 2008, ss. 3489-3503.

BACKUS, Philip, Mani Janakiram, Shanin Mowzoon, George C. Runger ve Amit Bhargava, "Factory Cycle Time Prediction with a Data Mining Approach", **IEEE Transactions On Semiconductor Manufacturing**, Vol. 19, No. 2, 2006, ss. 252-258.

BASKAK, Murat ve Vural Erol, "Sipariş Tipi Atölyelerde İş Sıralama Problemi için bir genetik algoritma uygulaması", **YA/EM'2004-Yöneylem/Endüstri Mühendisliği-XXIV Ulusal Kongresi**, 15-18 Haziran 2004, Gaziantep – Adana.

BAYKASOĞLU, Adil, Mustafa Göçken ve Zeynep D. Unutmaz. "New Approaches to due date assignment in job shops", **European Journal of Operational Research** 187, 2008, ss. 31-45.

BİÇEN, Pelin, "Veri Madenciliği: Sınıflandırma ve Tahmin Yöntemlerini Kullanarak Bir Uygulama", (Yayınlanmamış Yüksek Lisans Tezi), Yıldız Teknik Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul, 2002.

BRAHA(ED), Dan. **Data Mining for Design and Manufacturing**, Preface, Springer, 2002, Hardcover, ISBN:1-4020-0034-0.

BÜCHNER, Alex G., Sarabjot S. Anand ve John G. Hughes. "Data Mining in Manufacturing Environments: Goals, Techniques and Applications", **Studies in Informatics and Control**, 1997, ss. 319–328.

CHINCHOLKAR, Mandar M., Timoty Burroughs ve Jeffrey W. Hermann. "Estimating Manufacturing Cycle Time and Throughput in Flow Shops with Process Drift and Inspection", 2004, <http://www.isr.umd.edu/Labs/CIM/projects/dfp/flowshop.pdf>, (10.02.2009).

CUNHA, C.Da, Bruno Agard ve Andrew Kusiak. "Data Mining for improvement of product quality", **International Journal of Production Research**, Vol. 44, No. 18-19, 2006, ss. 4027-4041.

INTERNATIONAL LABOUR OFFİCE (Uluslararası Çalışma Örgütü). **İş Etüdü Kitabı**, çev. Zühal Akal, Milli Prodüktivite Merkezi Yayınları: 29, Dördüncü Basım, Ankara, 1991.

DEMİR, Prof. Dr. M. Hulusi ve Prof. Dr. Şevkinaz Gümüšođlu. **Üretim Yönetimi (İşlemler Yönetimi)**, Beta Yayınları, İstanbul, Kasım 1998, Genişletilmiş 5. Baskı.

DENER, Murat, Murat Dörterler ve Abdullah Orman, "Açık Kaynak Kodlu Veri Madenciliđi Programları: WEKA'da Örnek Uygulama", Akademik Bilişim, 2009, Harran Üniversitesi, Şanlıurfa, <http://ab.org.tr/ab09/bildiri/42.pdf>, (10.01.2010), ss. 787-796.

DUNHAM, Margaret H. **Data Mining Introductory and Advanced Topics**, Kitap, Prentice Hall, New Jersey, 2003.

ELMAS, Çetin. **Yapay Sinir Ağları (Kuram, Mimari, Eğitim, Uygulama)**, Seçkin Yayıncılık, Ankara, 2003.

ENNS, Silvanus T. "Lead time selection and the behaviour of work flow in job shops", **European Journal of Operational Research** **109**, 1998, ss. 122-136.

ERTEK, Gürdal, Can Kuruca, Cenk Aydın, Besim Ferit Erel, Harun Dođan, Mustafa Duman, Mete Ocal ve Zeynep Damla Ok. "Visual and analytical mining of sales transaction data for production planning and marketing", **4th International Symposium on Intelligent Manufacturing Systems**, 2004, ss. 848-859.

FAYYAD, Usama, Gregory Piatetsky ve Shapiro, Padhraic Smyth. "Knowledge Discovery and Data Mining: Towards a Unifying Framework", 1996, <http://courses.cs.cornell.edu/cs478/2000sp/handoutDBLP>, (20.03.2009).

FİLİZ, Atilla. **Üretim Yönetiminde VERİMLİLİK SIRLARI**, Sistem Yayıncılık, İstanbul, Nisan 2008,

GOODWIN, Randall, Russell Miller, Eugene Tuv, Alexander Borisov, Mani Janakiram ve Sigal Louchheim. "Advancements and Applications of Statistical Learning/Data Mining in Semiconductor Manufacturing", **Intel Technology Journal**, Vol. 8, No. 4, 2004, ss. 325-336.

GOVIND, Nirmal ve Theresa M. Roeder. "Estimating Expected Completion Times with Probabilistic Job Routing", **Proceedings of the 2006 Winter Simulation Conference**, ss. 1804-1810.

HA, Sung Ho ve Sang Chan Park. "Application of data mining tools to hotel data mart on the Intranet for database marketing", **Expert Systems With Applications** **15**, 1998, ss. 1-31.

HAN, Jiawei ve Micheline Kamber. **Data Mining Concepts and Techniques** , Kitap, ikinci baskı, Morgan Kaufmann, San Francisco, 2006.

HARDING, James A., M. Shahbaz, Srinivas ve Andrew Kusiak. "Data Mining in Manufacturing : A review", **Journal of Manufacturing Science and Engineering**, 2006, Vol. 128, ss. 969 -976.

JOTHISHANKAR, M. C., Tong (Teresa) Wu, Johnie Roberts, Jiun-Yan Shiau. "Case Study: Applying Data Mining to Defect Diagnosis", **Journal of Advanced Manufacturing Systems**, Vol.3, No.1, 2004, ss. 69-83.

KAMATH, Narasimha, Subir Bhattacharya, "Lead time minimization of a multi-product, single-processor system: A comparison of cyclic policies", **Int. J. Production Economics**, Vol. 106, 2007, ss. 28-40.

KANTARDZIC, Mehmed. **Data Mining Concepts, Models, Methods, and Algorithms**, Kitap, Wiley – Interscience, USA, 2003.

KARABOĞA, Derviş. **Yapay Zeka Optimizasyon Algoritmaları**, Kitap, Atlas Yayın Dağıtım, İstanbul, 2004.

KOBU, Prof. Dr. Bülent. **Üretim Yönetimi**, Beta Yayınları, İstanbul, Kasım 2008, Genişletilmiş 14. Baskı.

KUSAR, Janez, Ales Brezovar, Janez Grum, Marko Starbek. "Realistic lead time scheduling of operations of orders", **International Journal of Machine Tools & Manufacture**, Vol 44, 2004, ss.1037-1046.

KUSIAK, Andrew. "Data Mining: manufacturing and service applications", **International Journal of Production Research**, Vol. 44, No. 18-19, 2006, ss. 4175-4191.

KUSIAK, Andrew. "Data Mining in Design of Products and Production Systems", **Proceedings of INCOM'2006:12th IFAC/IFIP/IFORS/IEEE Symposium on Control Problems in Manufacturing**, Vol. 1, 2006, Saint-Etienne, France, ss. 49-53.

LAU, H.C.W., Bing Jiang, W.B. Lee ve K.H. Lau. "Development of intelligent data mining system for a dispersed manufacturing network", **Expert Systems**, 2001 Vol.18, ss. 175-185.

LEE, Ho Woo, No Ik Park ve Se Won Lee. "Analysis of the manufacturing lead time in a production system with non-renewal batch input, threshold, policy and post operation", **Applied Mathematical Modeling** **31**, 2007, ss. 2160-2171.

LIAO, Da-Yin ve Chia-Nan Wang. "Neural-Network Based Delivery Time Estimates for Prioritized 300-mm Automatic Material Handling Operations", **IEEE Transactions on Semiconductor Manufacturing**, Vol. 17, No.3, 2004, ss. 324-332.

MENON, Rakesh, Loh Han Tong ve S. Sathiyakeerthi. "Analyzing textual databases using data mining to enable fast product development processes", **Reliability Engineering and System Safety** **88**, 2005, ss. 171-180.

MOSES, Scott A., Hank Grant, Le Gruenwald ve Simin Pulat. "Real-time due-date promising by build-to-order environments", **International Journal of Production Research**., Vol. 42, No.20, ss. 4353-4375.

OĞUZLAR, Ayşe. "Veri Ön İşleme", **Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi**, Sayı: 21, 2003, ss. 67-76

Oracle Applications, "Manufacturing Lead Time Computations", http://download-west.oracle.com/docs/cd/A60725_05/html/comnl/us/bom/ldtms03.htm, (28.11.2009)

ÖZEKEŞ, Serhat, "Veri Madenciliği Modelleri ve Uygulama Alanları", **İstanbul Ticaret Üniversitesi Dergisi**, Sayı:3, 2003, ss. 65-82.

ÖZEKEŞ, Serhat. "Veri Madenciliği Uygulaması", (Yayınlanmamış Yüksek Lisans Tezi), Marmara Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, 2002,

SLOTNICK, Susan A. ve Matthew J. Sobel. "Manufacturing lead-time rules: Customer retention versus tardiness costs", **European Journal of Operational Research** **163**, 2005, ss. 825-856.

WANG, Kesheng. "Applying data mining to manufacturing: the nature and implications", **Springer Science+Business Media, LLC 2007**, 2007, ss. 487-495.

WESTPAL, Christopher ve Teresa Blaxton. **Data Mining Solutions Methods and Tools for Solving Real-World Problems**, Wiley Computer Publishing, John Wiley & Sons, Inc., 1998.

YURTSEVER, Ulaş. "Veri Madenciliği ve Uygulaması", (Yayınlanmamış Yüksek Lisans Tezi), Sakarya Üniversitesi, Sosyal Bilimler Enstitüsü, Sakarya, 2002.