

**T.C.
BAHÇEŞEHİR ÜNİVERSİTESİ**

**METİN MADENCİLİĞİ İLE DOKÜMANLAR
ARASINDAKİ BENZERLİKLERİN
BULUNMASI**

Yüksek Lisans Tezi

SELÇUK DÖVEN

İSTANBUL, 2013

T.C.
BAHÇEŞEHİR ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİ TEKNOLOJİLERİ YÜKSEK LİSANS PROGRAMI

**METİN MADENCİLİĞİ İLE DOKÜMANLAR
ARASINDAKİ BENZERLİKLERİN
BULUNMASI**

Yüksek Lisans Tezi

SELÇUK DÖVEN

Tez Danışmanı: Yrd. Doç. Dr. M. Alper TUNGA

İSTANBUL, 2013

T.C.
BAHÇEŞEHİR ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİ TEKNOLOJİLERİ YÜKSEK LİSANS PROGRAMI

Tezin Adı: Metin Madenciliği İle Dokümanlar Arasındaki Benzerliklerin Bulunması
Öğrencinin Adı Soyadı: Selçuk DÖVEN
Tez Savunma Tarihi: 07/06/2013

Bu tezin Yüksek Lisans tezi olarak gerekli şartları yerine getirmiş olduğu
_____ Enstitüsü tarafından onaylanmıştır.

Doç. Dr. Tunç BOZBURA
Enstitü Müdürü
_____ İmza _____

Bu tezin Yüksek Lisans tezi olarak gerekli şartları yerine getirmiş olduğunu onaylarım.

Yrd. Doç. Dr. M. Alper TUNGA
Program Koordinatörü
_____ İmza _____

Bu Tez tarafımızca okunmuş, nitelik ve içerik açısından bir Yüksek Lisans tezi olarak yeterli görülmüş ve kabul edilmiştir.

_____ Jüri Üyeleri _____

_____ İmzalar _____

Tez Danışmanı
Yrd. Doç. Dr. M. Alper TUNGA

Ek Danışman

Üye
Yrd. Doç. Dr. Tevfik AYTEKİN

Üye
Yrd. Doç. Dr. Y. Batu SALMAN

TEŐEKKÖR

Yüksek lisansım boyunca katkılarını, yönlendirici desteğini ve anlayışını hiçbir zaman esirgemeyen danışmanım Sayın Yrd. Doç. Dr. M. Alper TUNGA, tez için geliőtirdiğim uygulamada tecrübelerinden yararlandığım arkadaşım Bilgisayar Müh. Bahadır ŐENYAYLA, tez çalışmamda ve tüm okul hayatım boyunca maddi-manevi tüm destekleri için kız kardeşim Semra BİLGİN ve aileme teşekkür ederim.

İstanbul, 2013

Selçuk DÖVEN

ÖZET

METİN MADENCİLİĞİ İLE DOKÜMANLAR ARASINDAKİ BENZERLİKLERİN BULUNMASI

Selçuk Döven

Bilgi Teknolojileri Yüksek Lisans Programı

Tez Danışmanı: Y. Doç. Dr. M. Alper Tunga

Haziran 2013, 92 Sayfa

Metin madenciliği, yapısal olmayan kaynaklar (metin, resim vb.) içinde çeşitli yöntemler kullanarak veriler arasında sınıflandırma ve benzerlik gibi işlemler yapan bir uygulama alanıdır. Bu tezde bahsedilen metin madenciliği ile dokümanlar arasındaki benzerliklerin bulunmasını (örüntü tarama) bir uygulama üzerinde anlatmak için masaüstü uygulaması geliştirilmiştir. Uygulama, yalnızca algoritmaların çalışma mantığını gösterecek şekilde basit yapı olarak değil gerçek uygulamalar yapabilecek kapsamlı bir şekilde oluşturulmuştur. Uygulamayı diğer tez uygulamalarından ayıran özelliği, yalnızca iki doküman arasında değil kullanıcının ihtiyacına göre n sayıda doküman arasındaki benzerliklerin karşılaştırılmasını yapılabilmesidir. Uygulamaya yüklenen bir doküman içerisinde bulunan her cümlenin diğer dokümanlar içerisinde bulunan tüm cümleler ile benzerlik hesaplaması yapılmaktadır. Yine uygulama, sonuç kısmında karşılaştırılan cümlelerin sayısal olarak sonuçlarını göstermekte, her karşılaştırılan dosyayı ve karşılaştırılan bütün cümlelerin tamamını sayısal sonuçları ile birlikte sonuç tablosunda göstermektedir. Bu sayede kullanıcının cümlelere bakarak hangi cümlelerin hangi algoritma için nasıl bir sonuç ürettiğinin görülmesi sağlanmıştır. Uygulama için en çok kullanılan metin madenciliği benzerlik hesaplama algoritmaları olan kosinüs (cosine) ve jaccard algoritmaları kullanıp başarıları test edilmiştir.

Anahtar Kelimeler: Veri Madenciliği, Metin Madenciliği, Benzerlik, Kosinüs, Jaccard

ABSTRACT

FINDING SIMILARITIES BETWEEN DOCUMENTS USING TEXT MINING TECHNIQUES

Selçuk Döven

Graduate Program In Information Technology

Thesis Supervisor: Y. Doç. Dr. M. Alper Tunga

June 2013, 92 Pages

Text mining is a field of application that makes classification and comparison between data in unstructured sources (text, picture, etc.) by using several methods. A desktop application was developed to explain the details of this study with an example that shows the way of finding similarities between documents by using text mining. This application is not only a simple programme that shows working principle of algorithms, but also an extensive programme that can be applied on real applications. The discriminative ability of this application is that it can compare the similar parts of (n) number of documents according to the user's need. Similarities between every sentence in a document that is installed into the programme and all sentences in other documents can be calculated. Additionally, this application shows all of the folders and sentences that are compared, with the numerical outcomes in the result part. In this way, the user can understand from the outcomes that what kind of sentences and what kind of algorithms may produce results like that.

For the application, the most widely used algorithms; cosine and jaccard algorithms are used and their performance are tested.

Keywords: Data Mining, Text Mining, Similar, Cosine, Jaccard

İÇİNDEKİLER

TABLolar.....	viii
ŞEKİLLER.....	ix
KISALTMALAR.....	x
1.GİRİŞ.....	1
2. VERİ MADENCİLİĞİ YÖNTEM BİLİMİ.....	6
2.1 VERİ TABANLARI VE VERİ AMBARLARI.....	6
2.1.1 Veri Tabanları	6
2.1.2 Veri Ambarları.....	7
2.1.2.1 Veri ambarı özellikleri.....	8
2.1.3 Veri Tabanları Ve Veri Ambarlarının Karşılaştırılması.....	8
2.2 VERİ MADENCİLİĞİ.....	8
2.2.1 Veri Madenciliği Tanımları.....	9
2.2.2 Veri Madenciliği Kullanım Amacı ve Kullanım Alanları.....	11
2.2.3 Veri Madenciliği Süreci.....	15
2.2.3.1 Problemin tanımlanması.....	16
2.2.3.2 Verilerin hazırlanması.....	16
2.2.3.3 Modelin kurulması ve kullanılması.....	17
2.2.3.4 Modelin izlenmesi.....	17
2.2.4 Veri Madenciliği Yazılımları.....	17
2.2.5 Veri Madenciliği Modelleri.....	20
2.2.6 Veri Madenciliği Yöntemleri.....	20
2.2.6.1 Sınıflandırma.....	20
2.2.6.2 Kümeleme.....	22
2.2.6.3 Birliktelik kuralları.....	25
3.METİN MADENCİLİĞİ.....	27
3.1 METİN MADENCİLİĞİ ADIMLARI.....	28
3.1.1 Temizleme İşlemleri.....	28
3.1.2 Durdurma Kelimelerin Çıkarılması.....	28
3.1.3 Gövdeleme.....	29
3.1.4 Kelimelerin Yapısallaştırılması.....	30

3.2 WEB MADENCİLİĞİ.....	31
3.3 METİN SINIFLANDIRMA.....	32
3.4 METİN BENZERLİK HESAPLAMA YÖNTEMLERİ.....	33
3.4.1 Sözlük Kullanarak.....	33
3.4.2 Sözlük Kullanmayarak.....	35
3.5 METİN MADENCİLİĞİ UYGULAMA ALANLARI.....	37
3.6 METİNLERİN HATALI DEĞERLENDİRİLME DURUMLARI..	37
4. YÖNTEM VE UYGULAMA.....	39
4.1 GELİŞTİRİLEN UYGULAMA.....	39
4.3 VERİLERİN TEMİZLENMESİ.....	47
4.4 KELİMELEİN DÖNÜŞÜMÜ.....	48
4.5 CÜMLELERİN KARŞILAŞTIRILMASI VE UYGULAMANIN KARMAŞIKLIĞI.....	50
5. SONUÇ VE ÖNERİLER.....	52
KAYNAKÇA.....	57
EKLER	
Ek 1 Namık KEMAL' in hayatı 1.....	61
Ek 2 Namık KEMAL' in hayatı 2.....	63
Ek 3 Namık KEMAL' in hayatı 3.....	65
Ek 4 Namık KEMAL' in hayatı 4.....	71
Ek 5 Namık KEMAL' in hayatı 5.....	73
Ek 6 Namık KEMAL' in hayatı 6.....	75
Ek 7 Namık KEMAL' in hayatı 7.....	78
Ek 8 Namık KEMAL' in hayatı 8.....	85
Ek 9 Namık KEMAL' in hayatı 9.....	90
Ek 10 Namık KEMAL' in hayatı 10.....	91
ÖZGEÇMİŞ.....	92

TABLÖLAR

Tablo 3.1: Kelime kökleri.....	29
Tablo 3.2: Balkanet Sözlüğü.....	34
Tablo 5.1: “0” Değerler.....	53
Tablo 5.2: “1” Değerler.....	54
Tablo 5.3: “0 ile 1 arası” Değerler.....	54

ŞEKİLLER

Şekil 2.1: Veri ambarı mimarisi	7
Şekil 2.2: Veri madenciliği uygulama alanları.....	11
Şekil 2.3: Veri madenciliği süreci.....	15
Şekil 2.4: Karar ağaçları.....	21
Şekil 2.5: Kümeleme.....	24
Şekil 3.1: Vektör uzay modeli.....	30
Şekil 4.1: Geliştirilen uygulamanın 1. Versiyonu.....	39
Şekil 4.2: 1. Versiyon verilerin işlenmemiş görüntüsü.....	40
Şekil 4.3: 1. Versiyon verilerin işlenmiş görüntüsü.....	41
Şekil 4.4: Geliştirilen uygulamanın 2. Versiyonu.....	41
Şekil 4.5: Dosya ekleme ekranı.....	42
Şekil 4.6: Dosyaların tablodaki görüntüsü.....	43
Şekil 4.7: Sonuç sekmesi.....	43
Şekil 4.8: Sonuç gruplandırma (jaccard).....	44
Şekil 4.9: Excel olarak kayıt.....	45
Şekil 4.10: Excel tablosu.....	45
Şekil 4.11: Toplam satır ve toplam süre.....	46
Şekil 4.12: Filtrelenen kelimeler ve karakterler.....	47
Şekil 5.1: Uygulamaya yüklenen dosyalar.....	52
Şekil 5.2: Uygulamanın sonuç görüntüsü.....	53

KISALTMALAR

VTYS	:	Veritabanı Yönetim Sistemleri
OLAP	:	On-line Analytical Processing
CURE	:	Clustering Using Reipresentatives
SQL	:	Structured Query Language
XML	:	Extensible Markup Language
HTML	:	Hyper Text Markup Language
FBI	:	Federal Bureau of Investigation

1. GİRİŞ

İnsanođlu var olduđundan beri her dönem yeni ihtiyaçlara gereksinim duymuřtur. Bu gereksinimler ilk bařlarda barınma, yemek gibi en temel olan ihtiyaçlardı. Daha sonra bu ihtiyaçlar sanatın, keřiflerin, reformların, buluşların ve teknolojinin gelişimi ile birlikte deđişim gösterdi. İnsanlık tarihindeki önemli deđişimler (sanayi devrimi, dünya savařları vs.) insanların ihtiyaçlarını sürekli olarak bir üst seviyeye çıkardı.

Özellikle 1990'lı yıllarda bilgisayarların hızlı gelişimine ve internetin son kullanıcılara (akademik veya askeri kurumlar harici) açılmasına paralel olarak inanılmaz boyutlarda bir veri artışı olmuřtur. İşletmeler müşterileri, ürünleri ve satışları ile ilgili olan bilgileri barkod teknolojisiyle veritabanlarına kaydetmeye, son kullanıcılar ise internet aracılığıyla hem kendi bilgilerini paylaşmaya bařlamıř hem de diđer insanların ve kurumların bilgilerine ulaşmaya bařlamıřtır. Ama günümüzde artık yalın haliyle veriler önceki deđerini yitirmiř durumdadır. Çünkü artık kişisel bilgisayarların ve internetin inanılmaz gelişimiyle hemen hemen herkes istediđi veriye ulaşabilir hale gelmiřtir. Elde olan verilerin işlenerek kullanılabilir hale gelmesine yani bilgi haline gelmesine (anamlı hale gelmesine) ihtiyaç duyulmuřtur. Özellikle kar amaçlı olan işletmeler (özel işletmeler) rakiplerinden bir adım önde olabilmek için bu verilerden yararlanarak satışlarını artırmaya, daha çok müşteri kazanmaya çalışmak istemektedirler.

1990'lı yıllarda bu gelişimlere paralel olarak veri madenciliđi yöntemleri hızlı bir şekilde kullanılmaya başlanmıřtır. Veri madenciliđi bugün; spor, sanat, bilim, eğitim, pazarlama, meteoroloji gibi neredeyse hayatımızın her alanında yer almaktadır. Veri madenciliđi veriler arasındaki gizli kalmıř ilişkilere, önceden tahmin edilemeyecek olan bilgilere ulaşmaktadır. Veri madenciliđi hem kendisinin hem de kendisiyle ilişki içerisinde olan diđer disiplinlerin (yapay zeka, makine öğrenmesi, istatistik vb.) gelişimiyle daha farklı alanlarda faaliyet göstermeye ve daha derin bilgilere ulaşmaya bařlamıřtır. Veri madenciliđinin uygulama alanları içerisinde en önemlilerden birisi de metin madenciliđidir. Çünkü veri madenciliđi yapılandırılmıř (sayısal) veriler üzerinde çeřitli algoritmaları ve yöntemleri kullanarak işlem yapar. Metin madenciliđi ise yapılandırılmamıř (resim, metin gibi sayısal olmayan) veriler üzerinde işlem yapar. Bugün birçok kaynakta veri madenciliđinde kullanılabilecek verilerin yüzde 80 civarında yapılandırılmamıř verilerden olduđu belirtilmiřtir. Bu da metin madenciliđini çok önemli kılmaktadır. Metin madenciliđi ile örneđin, internetteki haberler

sınıflandırılıp internette benzer haberler bulunabilir veya metinlerin (iki akademik yazının) birbirine olan benzerlikleri hesaplanabilir. Ülkemizde Metin madenciliği ile ilgili yapılan çalışmalar veri madenciliğinin diğer uygulama yöntemlerine göre daha azdır (sınıflandırma, kümeleme ve birliktelik kuralları).

Metin madenciliği teknikleri, veri madenciliği yöntemlerinden olan kümeleme yöntemine benzemektedir. Kümeleme yönteminde, veri yığını içerisindeki tüm verilerin birbirine olan uzaklığını öklid ve manhattan gibi uzaklık ölçüleri ile ölçülerek birbirine en yakın veriler aynı kümede toplanmaktadır. Metin madenciliğinde de benzer bir yaklaşım bulunmaktadır. Yalnız metin madenciliğinde yapılan işlemler kümeleme yönteminde olduğu gibi sayısal değerlerde yapılmamaktadır. Metin madenciliğinde veriler (metinler) çeşitli temizleme ve dönüştürme işlemlerinden geçirildikten sonra veri madenciliği algoritmalarının kullanımına hazır hale getirilmektedir.

Metin madenciliğinde benzerlik arama, aslında veritabanında örüntü arama işlemlerine benzemektedir. Bunlardan bazıları; ses taraması, parmak izi taraması, dna benzerliği ve yüz tanıma yöntemleridir. Benzerlik yöntemi de tıpkı bu yöntemler gibi, elinde olan örneğe en benzer değeri veritabanında bulmaya çalışır.

Bu tez kapsamında dokümanların birbirine olan benzerlikleri hesaplayabilmek için, .Net Platformunda C# programlama dili (4.0 versiyonu) ile bir masaüstü uygulama geliştirilmiştir. .Net, Microsoft firmasının yazılım geliştirmek için oluşturduğu bir uygulama geliştirme platformudur. Bu platform içerisinde Visual Basic, C++ ve C# gibi birçok programlama dili bulunmaktadır. Platformun içerisinde bulunan programlama dilleri ile masaüstü, web ve mobil uygulamalar geliştirilebilmektedir. .Net Platformu tüm Windows işletim sistemlerinde çalıştırılabilmektedir. Diğer platformlarda çalıştırılabilmesi için çalışmalar devam etmektedir.

Bu tez kapsamında anlatılan metin madenciliği ile dokümanlar arasındaki benzerliklerin bulunması (örüntü tarama) uygulaması için, bir masaüstü (deskop) uygulama geliştirilmiştir. Uygulama yalnızca algoritmaların çalışma mantığını gösterecek şekilde basit uygulama olarak değil, gerçek uygulamalar yapabilecek, kapsamlı bir şekilde oluşturulmuştur. Uygulamayı diğer tez uygulamalarından ayıran özelliği; yalnızca iki doküman arasında değil kullanıcının ihtiyacına göre n sayıda doküman arasındaki benzerliklerin karşılaştırılmasını yapılabilmektedir. Uygulamaya yüklenen bir doküman içerisinde bulunan her cümlenin diğer dokümanlar içerisinde bulunan tüm cümleler ile

(kartezyen çarpımı) benzerlik hesaplaması yapılmaktadır. Yine uygulama, sonuç kısmında karşılaştırılan cümlelerin sayısal olarak sonuçlarını (kosinüs ve jaccard algoritmaları için) göstermekte, her karşılaştırılan dosyayı ve karşılaştırılan bütün cümlelerin tamamını sayısal sonuçları ile birlikte sonuç tablosunda göstermektedir. Bu sayede kullanıcının cümlelere bakarak hangi cümlelerin hangi algoritma için nasıl bir sonuç ürettiğinin görülmesi sağlanmıştır. Metin madenciliğinde kullanılan dokümanlar içerisinde bazı verilerin (Türkçe’ de edatlar/bağlaçlar ve harf olmayan karakterler) işleme alınmayıp, temizlenmesi gerekmektedir. Bu karakterlerin kullanılan uygulamaya göre genişletilip veya kısıtlanabilmesi için, geliştirilen uygulamada bu kısıtlarının dışarıdan kullanıcı tarafından yönetilmesi için yönetim paneli bulunmaktadır. Uygulamanın çalışmasının sonucunda karşılaştırma yapılan cümle sayısı ve uygulamanın çalışmayı tamamladığı toplam süre yine sonuç kısmında kullanıcıya gösterilmektedir. Uygulama için en çok kullanılan metin madenciliği benzerlik hesaplama algoritmaları olan kosinüs (cosine) ve jaccard algoritmaları kullanıp başarıları test edilmiştir.

Uygulamada kullanılan kosinüs algoritması, doküman içerisinde geçen kelimelerin ağırlığını alarak bir değer üretmektedir. Diğer kullanılan algoritma olan jaccard ise kelimenin doküman içerisinde geçip geçmeme durumuna bağlı olarak yalnızca 0 veya 1 değerini üretmektedir. Birbirinden farklı şekilde değer üreten iki ayrı algoritma seçilerek dokümanların benzerlikleri iki şekilde hesaplanmıştır.

Tez içerisinde bulunan yöntemlerin ve algoritmaların çalışmasını incelemek için bir uygulama gerçekleştirilmiştir. Uygulama için Türk yazarlarından Namık KEMAL’ in hayat hikâyesi internet üzerinden araştırılıp, farklı sitelerden alınan bilgiler pdf ve txt formatında saklanmıştır. Daha sonra dokümanlar uygulama tarafından incelenerek dokümanlar arasında benzerlikler aranmıştır ve iki ayrı algoritmanın aynı cümleler için verdiği farklı sonuçlar irdelenmiştir.

Metin madenciliği hakkında yapılan bazı çalışmalar şunlardır:

Rauka (2012) tezinde, Türkçe ve İngilizce metinleri analiz edip seçilen 5 farklı konuda kategorize etmiş, daha sonra aynı metinleri diğer dillere çevirerek tekrar sınıflandırma yapmıştır. Metinlerin çeviri yapıldıktan sonra sınıflandırmalardaki değişiklikleri ve bilgisayardaki çeviricilerin veri madenciliğindeki etkileri incelenmiştir.

Karaca (2012) tezinde, web madenciliğinde sınıflandırma teknikleri kullanarak bir uygulama gerçekleştirmiştir. Uygulama, haber sitelerinden alınan spor, ekonomi gibi köşe yazılarının otomatik olarak uygulanmasına ilişkin bir çalışmadır.

Seçkin (2011) tezinde, üç siyasi parti liderinin grup toplantı konuşmalarını metin madenciliği sınıflandırma yöntemini kullanarak hangi metnin hangi siyasi lidere ait olduğunu bulmaya çalışmıştır. Çeşitli sınıflandırma algoritmaları test edilerek başarıları karşılaştırılmıştır.

Çelikyay (2010) tezinde, Türkçede en çok kullanılan harf ve kelimelerin birbirini izleme sıklıkların ikili, üçlü şekilde metin madenciliği yöntemiyle belirlemeye çalışmıştır.

Dolgun ve diğ. (2009) makalelerinde, web madenciliği yöntemleriyle müşteri profilleri çıkarmışlardır. Yapısal olan veri ile yapısal olmayan veriler arasındaki farkları bulmuşlardır.

Işık ve Çamurcu (2008) makalelerinde, metin madenciliği benzerlik algoritmalarını kullanarak, algoritmaların başarılarını karşılaştırmışlardır.

Dursun (2008) tezinde, Türkçe metinlerin benzerliklerin karşılaştırılmasında sık yapılan hata durumlarını modelleyerek yeni bir hesaplama yöntemi geliştirmiş ve diğer yöntemlerle farklı durumlarda karşılaştırma yapmıştır. Bu yöntem özellikle yazım yanlışları üzerinde durarak metinleri daha iyi hesaplamaya çalışmaktadır.

Yücesoy (2007) tezinde, Türkçenin de içerisinde bulunduğu balkanet kütüphanesinden yararlanarak, kavramsal sözlük kullanarak metinlerin benzerliklerini hesaplamıştır. Ayrıca Wu-Palmer yöntemiyle bu yöntem karşılaştırılmıştır.

Haberal (2007) tezinde, web kullanım madenciliğine ilişkin bir çalışma yapmıştır. Çalışmasında bir internet sitesini web madenciliği teknikleriyle analizi yapılmıştır. En çok hangi sayfalara, hangi kullanıcı grupları, hangi yoğunlukta, hangi ülkelerden ziyaretler yapılıyor şeklinde bilgilere ulaşılmaya çalışılmıştır.

Tezin ikinci bölümünde; veri ambarlarının tanımı, oluşturulması ve mimarisi hakkında gerekli bilgiler verilmektedir. Veri madenciliğinin tanımı, kullanım alanları ve yöntemleri ile ilgili detaylı bilgiler verilmektedir.

Tezin üçüncü bölümünde tezin konusu olan metin madenciliği detaylı şekilde anlatılmaktadır. Metin madenciliğinin tanımı, kullanım alanları, metin madenciliği adımları ve yöntemleri anlatılmaktadır.

Tezin dördüncü bölümünde tez için geliştirilen uygulama ve metin madenciliği yöntemleri uygulama üzerinden anlatılmaktadır.

Tezin beşinci bölümünde uygulama sonunda elde edilen sonuçlar ve öneriler belirtilmiştir.

2. VERİ MADENCİLİĞİ YÖNTEM BİLİMİ

Veri madenciliği uygulamaları içerisinde birçok yapıyı ve disiplini barındırmaktadır. Veri madenciliği yöntemlerin uygulanabilmesi için veri tabanları ve veri ambarları gibi yapılar, verileri işlemek için matematik, istatistik gibi diğer alanlardan yöntemler kullanılmaktadır. Ayrıca yapay zeka ve makine öğrenmesi gibi diğer disiplinlerden algoritmalar kullanılmaktadır. Bu bölümde bu yapılara ve disiplinlere ait tanımlamalara yer verilmektedir.

2.1 VERİ TABANLARI VE VERİ AMBARLARI

Bu bölümde veritabanları ve veri ambarlarına dair tanımlar yapıлып, bu yapılarının özellikleri işlenecek olup, daha sonra iki yapı arasındaki farklara değinilmektedir.

2.1.1 Veri Tabanları

Bilgisayarların gelişimine paralel olarak verileri depolamak için kullanılan dosyalama sistemleri, daha yüksek miktarda ve daha düzenli şekilde veri depolamak için yerlerini veri tabanlarına bırakmıştır.

Veritabanı, verileri belirli kurallar ve ilişkiler düzeni içerisinde depolamaya verilen isimdir. Veri tabanında veriler tablolarda çeşitli tanımlayıcı anahtarlar aracılığı ile tutulmaktadır. Veri tabanları ile tekrarlı kayıtların önüne geçilmektedir. Tablo kavramının karşılığı dosya kavramı ile çok benzerdir. Tablolar da dosyalar gibi sıralı kayıtlardan oluşmaktadır. Fakat tablolar veri tabanında çeşitli ilişkiler kurularak birbiri ile ilişkilendirilir. Veri tabanlarının birçok çeşidi bulunmaktadır. Bugün bunlar içerisinde en çok kullanılanı ilişkisel veritabanlarıdır. İlişkisel veritabanları, veritabanı içerisindeki tabloları çeşitli anahtarlar ile birbiri ile ilişkilendirmektedir. İlişkisel veritabanları daha fazla veriyi, daha az alanda çok iyi yönetmemize olanak sağlamaktadır. Bugün veritabanlarını daha iyi yönetebilmek için İlişkisel Veritabanı Yönetim Sistemleri (VTYS) kullanılmaktadır. VTYS ile veriler daha rahat yönetilebilir ve izlenebilir hale gelmektedir. Kullanıcılar için roller ve yetkiler tanımlanabilir (Gözüdeli 2010, ss.39-40).

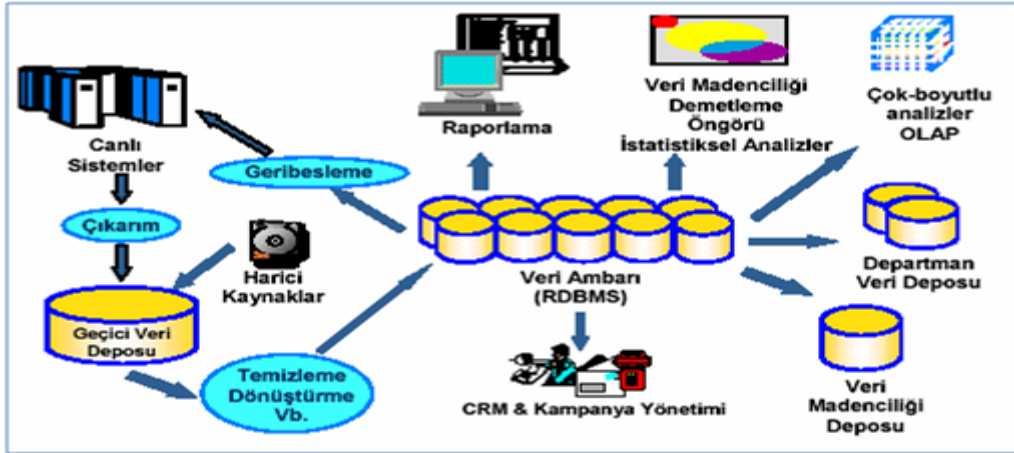
Ticari olarak kullanılan ilişkisel veritabanlarının haricinde tasarım ve mühendislik gibi alanlarda kullanılan başka veritabanları da bulunmaktadır. Bunlardan bazılarını Dolgun (2006) şu şekilde açıklamıştır;¹

- a) Nesne tabanlı veritabanları
- b) Uzaysal veritabanları
- c) Zaman serisi veritabanları
- d) Metin veritabanları

2.1.2 Veri Ambarları

Veri ambarları birçok kaynakta bulunan (veritabanı veya doküman) verinin birleşik bir şema altında depolanmasını sağlayan bir veri kaynağıdır. Veriler, veri ambarına kaydedilirken veri temizleme, veri dönüştürme, veri bütünleştirme gibi işlemlerden geçirilmektedir. Veri ambarları karar vermede kullanılacak olan bilgilerin belirli süre aralıklarında, konu odaklı olarak tek bir çatı altında toplanmasıdır. Veri ambarındaki veriler müşteri, tedarikçi, ürün gibi konular üzerinden yapılandırılarak depolanır. Veri ambarlarında veriler analiz ederken ve özetlerken OLAP (On-line Analytical Processing) yöntemleri kullanılır. OLAP verileri çok boyutlu olarak analiz eder.² Tipik bir veri ambarı mimarisi Şekil 2.1’ de gösterilmiştir.

Şekil 2.1: Veri ambarı mimarisi



Kaynak: Dolgun, M. Ö., (2006). Büyük alışveriş merkezleri için veri madenciliği uygulamaları. Yüksek Lisans Tezi. Ankara: Hacettepe Üniversitesi FBE.

¹ Dolgun, M. Ö., (2006). Büyük alışveriş merkezleri için veri madenciliği uygulamaları. Yüksek Lisans Tezi. Ankara: Hacettepe Üniversitesi FBE.

² Bilen, H., (2009). Bankacılık sektöründe personel seçimi ve performans değerlendirmesine ilişkin veri madenciliği uygulaması. Yüksek Lisans Tezi. Ankara: Gazi Üniversitesi FBE.

2.1.2.1 Veri ambarı özellikleri

Veri ambarları, birbirleriyle bütünleşik olmayan uygulamaların birleşmesini sağlar. Veri ambarların ilk özelliği konuya yönelik olmasıdır. Klasik veritabanlarında veriler işlemler üzerinden yürürken veri ambarlarında konular üzerinden yürümektedir. Örneğin bir okul için; öğrenci, dersler, notlar gibi... Veri ambarlarında veriler bütünleşik olmalıdır. Örneğin cinsiyet alanı mutlaka E, K veya 0,1 gibi tek bir türden gösterilmelidir. Farklılıklar giderilip tek bir türde birleştirilmelidir. Veri ambarlarında veriler belirli zaman aralıklarında muhafaza edilirler. Bunun sebebi geçmiş dönemdeki verilere bakarak gelecek dönemler için tahminlerde bulunmaktır. Veri ambarlarının diğer bir önemli özelliği de yalnız okunabilir olmasıdır. Veriler silinemez veya güncellenemez yalnızca okuma işlemi yapılarak sorgulamalar yapılabilir (Özkan 2008, ss. 22-26).

2.1.3 Veri Tabanları ve Veri Ambarlarının Karşılaştırılması

Veri tabanları ve veri ambarları hem yapı hem de amaç açısından birbirlerinden ayrılmaktadırlar. Veritabanları günlük işlemler üzerinden çalışmaktadır. Günlük faturalar, mal giriş-çıkışları ve yeni kayıtlar gibi. Veri ambarları ise kuruluşa yön gösterici niteliktedir. Kuruluşun geleceğine ilişkin kararlar almak için kullanılır. Veri ambarlarının modellenmesi boyutsal model olarak düşünülür. Veri tabanları ise genelde ilişki düzeyinde tablolarla gerçekleştirilir. Veritabanlarının üzerinde yapılan sorgulamalar genellikle standarttır. Önceden belirlenmiş sorgular çalıştırılır. Veri ambarlarında ise önceden planlanmamış ve ani gelişebilecek türdendir. Veritabanları günlük kayıtların bulunduğu yapılardır. Veri ambarları ise tarihsel dönemler düzeyinde yapılandırılmıştır. Bu yönüyle karar destek sistemlerinin içersindedir (Özkan 2008, ss.32-34).

2.2 VERİ MADENCİLİĞİ

Teknolojinin hızla gelişmesi ve insan ihtiyaçlarının giderek daha özel hizmetlere yönelmesiyle birlikte birçok disiplin kendini sürekli geliştirmektedir. Bunlardan biri de hiç kuşkusuz veri madenciliğidir. Veri madenciliği müşteri profili çıkarmaktan, insanda oluşabilecek hastalıkları tahmin etmeye kadar birçok alanda hizmet vermektir. Bu nedenle veri madenciliğin fayda sağladığı onlarca farklı alanı düşününce onu bir iki cümle ile tanımlamaya çalışmak onun sadece küçük bir kısmına odaklanmak olur.

2.2.1 Veri Madenciliği Tanımları

Veri madenciliği eldeki verilerden anlamlı bilgiler çıkararak kişi ve kuruluşları rakiplerinden bir adım önde tutmak için kullanılır. Veri madenciliği ile şirketlerin çok büyük stratejik kararları alınırken, doğru karar alması için önemli bilgiler elde etmesini sağlamaktadır.³

Veri madenciliği, işlenmemiş verinin kullanımda sunamadığı bilgiyi çıkaran veri analiz süreci olarak tanımlanabilir. Veri madenciliği istatistik, örüntü tanıma, makine öğrenme gibi disiplinlerin kullanıldığı, büyük veri tabanlarında önceden tahmin edilemeyen ilişkiler analizi olarak tanımlanabilir.⁴

Veri madenciliği veri kaynağında çeşitli sorgulamalar yaparak yararlı bilgi, desen ve eğilimleri bulmaya çalışır. Birçok kurumun amacı geçmiş deneyimlerini analiz ederek kurumun kar marjını yükseltmektir. Veri madenciliği bu analizleri işlemlerini yapmak için; istatistik, matematik, modelleme teknikleri ve çeşitli yazılımları kullanır. Veri madenciliği; birliktelik kuralları, sıralama ve sınıflandırma işlemlerini yaparken kurallar bulmaya çalışır.⁵

Veri madenciliği eldeki verilerden çok net olmayan, üstü kapalı olan fakat faydalı olabilecek bilgilere erişilmesidir. Veriler içerisindeki ilişkilerin, desenlerin ve kuralların istatistiksel olarak önemli eşikte bulunan yapıların keşfedilip kullanılmasıdır. Amaç daha önceden fark edilmemiş olan desenleri ortaya çıkarmaktır. İstatistik verilerin toplanıp analiz edilmesi, sınıflandırılıp özetlenmesi işlemlerinin yapılmasıdır. Bu işlemler sonunda ana kütle hakkında bilgiler elde edilmeye çalışılmaktadır. Veri madenciliğinin amacı da istatistik biliminin amacı ile aynıdır. Veri madenciliği istatistiğin biliminden en önemli farklarından birisi insan merkezli olmasıdır. Sorunu çözmek için karar verme sürecini destekleyen, çözüm için gerekli bilgileri sağlamaya çalışan bir araçtır.⁶

3 Karaca, M. F., (2012). Metin madenciliği yöntemleri ile haber sitelerindeki köşe yazılarının sınıflandırılması. Yüksek Lisans Tezi. Karabük: Karabük Üniversitesi FBE.

4 Akbulut, S., (2006). Veri madenciliği teknikleri ile bir kozmetik markanın ayrılan müşteri analizi ve müşteri segmentasyonu. *Yüksek Lisans Tezi*. Ankara: Gazi Üniversitesi FBE.

5 Kalıkov, A., (2006). Veri madenciliği ve bir e-ticaret uygulaması. *Yüksek Lisans Tezi*. Ankara: Gazi Üniversitesi FBE.

6 Tiryaki, S., (2006). Lojistik alanında bir veri madenciliği uygulaması. *Yüksek Lisans Tezi*. İstanbul: İstanbul Teknik Üniversitesi FBE.

Veri madenciliği, büyük miktardaki veriler içerisinde kuralların, örüntülerin ve çeşitli modellerin bulunmasıdır. Diğer bir ifade ile veri madenciliği, veri tabanı veya veri ambarı gibi çok büyük ölçüdeki veri yığının içerisindeki gizli birliktelikleri ve desenleri bulmak için, yapay zekâ ve istatistik gibi disiplinlerin yöntemlerin karmaşık arama yeteneği denilebilir. Keşif, bir veri kümesi içerisinde gizli örüntüleri önceden belirlenmiş bir fikir ve hipotez olmadan bulma sürecidir (Emel ve Taşkın 2005, s.223). Veriler içerisinde faydalı bilgilerin bulunması sürecine “bilgi keşfi” denilmektedir. Veri madenciliği bu keşif içerisinde önemli bir yere sahiptir. Veri madenciliği, özel algoritmaların veri yığınları içerisindeki örüntüleri belirlemek için kullanılan bir yöntemdir. Veri madenciliği, büyük miktardaki veri içerisinde gelecekle ilgili tahmin yapmamızı sağlayacak bilgilerin, bilgisayar uygulamaları kullanarak aranmasıdır (Timor ve Şimşek 2008, ss. 4-5).

Veri madenciliği, zeki yöntemler kullanılarak büyük miktardaki veriden anlamlı bilgilerin çıkarılması sürecidir. Daha sonra çıkarılan bu anlamlı bilgiler, içlerinden yararlı olanları belirlemek için değerlendirmeye alınmaktadır. Veri madenciliği yapay zeka, makine öğrenmesi, istatistik ve veri tabanları gibi alanların kesişimine dayalı bir yöntemdir (Gürbüz ve diğ. 2009, s. 74).

Genel olarak bilgisayar destekli bir bilgi çözümleme işlemi denilebilir.⁷

Veri madenciliği verileri analiz etmek için; sınıflandırma, kümeleme, birliktelik kuralları gibi birçok yöntem ve algoritma geliştirmiştir. Bu yöntemlerin birçoğu istatistik tabanlıdır. Bazı algoritmalar değişik şekillerde birden fazla konuda kullanılabilirler. Örneğin yapay sinir ağları hem sınıflandırmada hem de kümeleme modellerinde kullanılabilirler (Timor ve diğ. 2011, s. 129).

Veri madenciliğinde kullanılan yöntemler, işin niteliğine göre stratejik konularda karar almaya yardımcı olurlar. İşlenmemiş veriler içerisindeki gizli kalmış örüntüleri karar alma sürecinde tahmini bilgiye dönüştürürler.⁸

⁷ Sıramkaya, E., (2005). Veri madenciliğinde bulanık mantık uygulaması. *Yüksek Lisans Tezi*. Konya: Selçuk Üniversitesi FBE.

⁸ Tosun, T., (2006). Veri madenciliği teknikleriyle kredi kartlarında müşteri kaybetme analizi. *Yüksek Lisans Tezi*. İstanbul: İstanbul Teknik Üniversitesi FBE.

2.2.2 Veri Madenciliği Kullanım Amacı ve Kullanım Alanları

Veri madenciliği, birçok amaçta ve de birçok alan üzerinde araştırma yapıp, karar alma sürecine katkıda bulunmak için çalışmalar yapmaktadır. Veri madenciliğinin uygulama alanlarının yüzdesel dağılımı Şekil 2.2' de gösterilmiştir.

Şekil 2.2: Veri madenciliği uygulama alanları

Son 3 yıl içinde veri madenciliğinin uygulandığı alanlar	
Bankacılık (51)	12%
Biyoteknoloji / Genetik (11)	3%
Kredi skorlama (35)	8%
CRM (52)	12%
Doğrudan pazarlama (34)	8%
e-Ticaret (11)	3%
Eğlence/ Müzik (4)	1%
Sahtekarlık tespiti (31)	7%
Şans oyunu (2)	0,01 %
Kamu uygulamaları (12)	3%
Sigortacılık (24)	6%
Yatırım / Hisse senedi (5)	1%
Junk email / Anti-spam (5)	1%
Sağlık/ İK (15)	4%
İmalat (19)	5%
Tıp/ Farmakoloji (12)	3%
Perakende (25)	6%
Bilim (17)	4%
Güvenlik / Anti-terörizm(5)	1%
Telekomünikasyon (23)	5%
Seyahat (8)	2%
Web (9)	2%
Diğer (11)	3%

Kaynak: Akbulut, S., (2006). Veri madenciliği teknikleri ile bir kozmetik markanın ayrılan müşteri analizi ve müşteri segmentasyonu. Yüksek Lisans Tezi. Ankara: Gazi Üniversitesi FBE.

Bu kullanım amaçlarından bazılarını Tiryaki (2006) şu şekilde açıklamıştır;⁹

- Bir işletme kendi müşterisiyken rakibine giden müşterilerle ilgili analizler yaparak rakiplerini tercih eden müşterilerinin özelliklerini elde edebilir ve buradan hareketle gelecek dönemlerde kaybetme olasılığı olan müşterilerin

⁹ Tiryaki, S., (2006). Lojistik alanında bir veri madenciliği uygulaması. *Yüksek Lisans Tezi*. İstanbul: İstanbul Teknik Üniversitesi FBE.

kimler olabileceği yolunda tahminlerde bulunarak onları kaybetmemek, kaybettiklerini geri kazanmak için farklı stratejiler geliştirebilir.

- b) Mevcut müşterilerin işletme tarafından daha iyi tanınmasını sağlayabilir. Özellikle finans sektöründe mevcut müşterilerinin sınıflara ayrılarak çıkarılacak kredi risk davranış modellerinin yeni başvuruda bulunan müşterilere uygulanmasını sağlayarak riski minimize edebilir. Bir anlamda kredi risk puanlamasının altyapısının oluşturulmasında kullanılabilir.
- c) Mevcut müşterilerin ödeme performansları incelenerek kötü ödeme performansı gösteren müşterilerin ortak özellikleri belirlenerek, benzer özelliklere sahip tüm müşteriler için yeni risk yönetim politikaları oluşturulabilir.
- d) En karlı mevcut müşteriler belirlenerek, potansiyel müşteriler arasından en karlı olabilecekler belirlenebilir. Karlı müşteriler tespit edilerek onlara özel kampanyalar uygulanabilir. En masraflı müşteriler daha masrafsız müşteri haline dönüştürülebilir. Örneğin en çok bankacılık işlemi yapanlar ortaya çıkarılıp bunlar şube bankacılığı yerine daha masrafsız internet bankacılığına yönlendirilebilir.
- e) Mevcut müşteriye tanıyarak işletmelerin müşteri ilişkileri yönetimlerinde düzenleme ve geliştirmeler yapılabilir. Bu sayede firmanın müşterilerini daha iyi tanıyarak müşteri gibi düşünme kapasitelerinin artırılması sağlanabilir. Bunun da işletmelere pazarda avantaj sağlayacağı unutulmamalıdır.
- f) Geçmiş ve mevcut yapı analiz edilerek geleceğe yönelik tahminlerde bulunulabilir. Özellikle ciro, karlılık, pazar payı gibi analizlerde veri madenciliği çok rahat kullanılabilir.
- g) Mevcut müşteriler üzerinde firma ürünlerinin çapraz satış kapasitesinin artırılması sağlanabilir. Mesela, firmanın X ürününü alan müşterilerin çok büyük bir bölümünün Y ürününü de aldıkları biliniyorsa, buna yönelik pazarlama stratejileri geliştirilebilir.
- h) Piyasada oluşabilecek değişikliklere mevcut müşteri portföyünün vereceği tepkinin firma üzerinde oluşturabileceği etkinin tespitinde kullanılabilir.

- i) Operasyonel süreçte oluşabilecek olası kayıpların veya suistimallerin tespitinde kullanılabilir.
- j) Kurum teknik kaynaklarının en uygun şekilde kullanılmasını sağlamakta kullanılabilir.
- k) Firmanın finansal yapısının, makro ekonomik değişmeler karşısındaki duyarlılığı ve oluşabilecek risklerin tespitinde kullanılabilir.
- l) Günümüzde var olan yoğun rekabet ortamında firmaların hızlı ve kendisi için en doğru kararı almalarını sağlayabilir.

Veri madenciliğinin kullanım alanlarını Özçınar (2006) şu şekilde açıklamıştır:¹⁰

Pazarlama:

- a) Müşteri gruplandırmasında
- b) Müşterilerin demografik özellikleri arasındaki bağlantıların kurulmasında
- c) Çeşitli pazarlama kampanyalarında
- d) Mevcut müşterilerin elde tutulması için geliştirilecek pazarlama
- e) Stratejilerinin oluşturulmasında
- f) Pazar sepeti analizinde
- g) Çapraz satış analizleri
- h) Müşteri değerlendirme
- i) Müşteri ilişkileri yönetiminde
- j) Çeşitli müşteri analizlerinde
- k) Satış tahminlerinde

Bankacılık:

- a) Farklı finansal göstergeler arasındaki gizli korelasyonların bulunmasında
- b) Kredi kartı dolandırıcılıklarının tespitinde
- c) Müşteri sınıflandırılmasında
- d) Kredi taleplerinin değerlendirilmesinde
- e) Usulsüzlük tespiti
- f) Risk analizleri

¹⁰ Özçınar, H., (2006). Kpss sonuçlarının veri madenciliği yöntemleriyle tahmin edilmesi. *Yüksek Lisans Tezi*. Denizli: Pamukkale Üniversitesi FBE.

- g) Risk yönetimi
- h) Sigortacılık
- i) Yeni poliçe talep edecek Müşterilerin tahmin edilmesinde
- j) Sigorta dolandırıcılıklarının tespitinde
- k) Riskli müşteri tipinin belirlenmesinde

Perakendecilik:

- a) Satış noktası veri analizleri
- b) Alış-veriş sepeti analizleri
- c) Tedarik ve mağaza yerleşim optimizasyonu

Borsa:

- a) Hisse senedi fiyat tahmini
- b) Genel piyasa analizleri
- c) Hisse tespitlerinde
- d) Alım-satım stratejilerinin optimizasyonu
- e) Telekomünikasyon
- f) Kalite ve iyileştirme analizlerinde
- g) Hatların yoğunluk tahminlerinde
- h) Sağlık ve ilaç
- i) Test sonuçlarının tahmini
- j) Ürün geliştirme
- k) Tıbbi teşhis
- l) Tedavi sürecinin belirlenmesinde

Endüstri:

- a) Kalite kontrol analizlerinde
- b) Lojistik
- c) Üretim süreçlerinin optimizasyonunda
- d) Bilim ve mühendislik
- e) Deneysel veriler üzerinde modeller kurarak bilimsel ve teknik problemlerin çözümlenmesi

Eđitim:

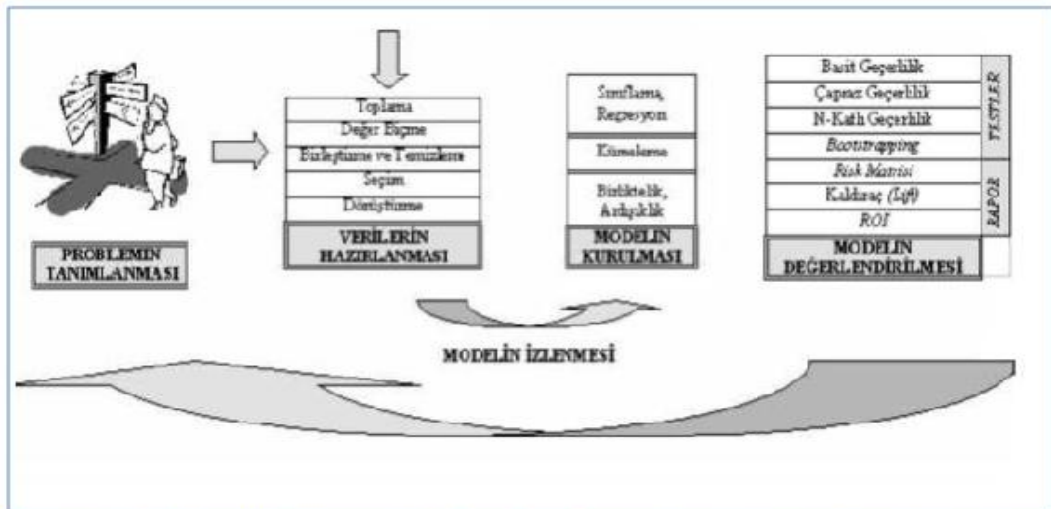
- a) Öğrenci davranışlarının öngörülmesi
- b) Öğrencilerin ders seçme eğilimlerinin belirlenmesi

Veri madenciliğinin örüntü arama işlemlerinden birisi de parmak izi taramasıdır. Parmak izi veya daktiloskopi kişilerin kimlik saptamasında kullanılan en eski ve en güvenilir biyometrik tanıma yöntemlerinden birisidir. Türkiye’ de ilk defa 1910 yılında kullanılan bu yöntem, FBI (Federal Bureau of Investigation) tarafından 1970’ lerde parmak izini suçlu tespitinde kullanmak için ilk otomatik sistemi kurmuştur. İnsan parmaklarının deri yüzeyinin kendine özgü kabartmaları, röfleyin ve çukurları bulunmaktadır. Bu kabartmalar ve çizgiler, karmaşıklığı ve tekrarlanmaz bir yapıyı oluşturmaktadır. Parmak izinde 60.000.000.000’ da 1 benzerliğin olması pratik olarak bu çizgilerin tekrarlanma olasılığının olmadığını göstermektedir (Nabiyev 2012, ss. 497-499).

2.2.3 Veri Madenciliği Süreci

Veri madenciliği; veriler arasındaki gizli kalmış, önceden kestirilemeyecek olan ilişkilerin veya benzerliklerin ortaya çıkarılma sürecidir. Bu süreç gerçekleştirilirken, problemin belirlenmesi, verilerin temin edilip dönüştürülmesi gibi çeşitli aşamalardan geçilir. Bu süreçlerin titiz ve doğru bir şekilde gerçekleştirilmesi daha sağlıklı sonuçlar elde edilmesini sağlar. Veri madenciliği süreci Şekil 2.3’ de gösterilmiştir.

Şekil 2.3: Veri madenciliği süreci



Kaynak: Muslu, D., (2009). Sigortacılık Sektöründe risk analizi: veri madenciliği uygulaması. Yüksek Lisans Tezi. İstanbul: İstanbul Teknik Üniversitesi FBE.

2.2.3.1 Problemin tanımlanması

Veri madenciliği çalışmalarının ilk adımı problemin doğru bir şekilde tanımlanmasıdır. Yapılan çalışmanın başarılı olabilmesi için hangi işletme için ne yapılacağına açık bir şekilde tanımlanması gerekmektedir. Sorunun doğru tespit edilmemesi durumunda sonuçların beklenen faydayı karşılamamasına sebep olur. Veri madenciliği algoritmalarının ürettiği tahminlerin doğru veya yanlış olması durumunda karşılaşılabilecek olan kayıpların veya kazanç tahminlerinin bu aşamada belirtilmesi gerekmektedir.¹¹

2.2.3.2 Verilerin hazırlanması

Veri madenciliği sürecinin en önemli aşamalarından birisi verilerin hazırlanması aşamasıdır. Çünkü uygulamayı yapacak olan analistin zamanının neredeyse yüzde 50 ile yüzde 75' ini alır. Bu aşamada, elde edilen sayısal verinin iyi analiz edilip veri madenciliği algoritmalarının doğru şekilde kullanılabilmesi şeklinde hazırlamak gerekmektedir.¹²

Verilerin hazırlanması; verilerin temizlenmesi, verilerin indirgenmesi ve verilerin yeniden yapılandırılması gibi süreçleri kapsamaktadır.

Verilerin temizlenmesi: Verilerin temizlenmesi, kayıp, yanlış veya gürültülü verilerin ortadan kaldırılmasıdır. Kayıp veriler, bazı bilgilerin hiç veri kaynağına girilmemiş olmasıdır. Yanlış veriler, bazı değerlerin gerçekle bağdaşmayacak şekilde uç veri olarak girilmiş olmasıdır. Örneğin, yaş alanına doğum tarihinin 1046 olarak girilmesidir. Bazen ürün kodu veya ismi gibi bilgiler yanlış girilmiş olabilir. Bu gibi sorunları giderebilmek için şu işlemler yapılabilir:

- a) Kayıp verilerin bulunduğu kayıtları veritabanından çıkarmak
- b) Kayıp verileri elle tekrar doldurmak
- c) Tüm kayıtlara aynı bilgileri girmek. Bu işlemi yapmak verilerin sonucu önemli ölçüde değiştirebilmektedir.

¹¹ Muslu, D., (2009). Sigortacılık sektöründe risk analizi: veri madenciliği uygulaması. *Yüksek Lisans Tezi*. İstanbul: İstanbul Teknik Üniversitesi FBE.

¹² Uslu, N. C., (2009). Veri madenciliği ile bilgisayar ağlarında yeni bir saldırı tespit algoritması. *Yüksek Lisans Tezi*. Ankara: Gazi Üniversitesi FBE.

- d) Kayıp olan verilere tüm verilerin ortalama deęerini girmek
- e) Regresyon yöntemi kullanarak, dięer deęişkenlerin yardımı ile kayıp olan verilerin tahmin edilmesi

(Silahtaroglu 2008, ss. 19-21).

Verilerin indirgenmesi: Bazı durumlarda işlenecek verilerin sayısı yapılacak olan uygulamanın çalışma süresini uzatacak olabilir. Eğer sonuçlara etki edeceği düşünülüyor ise verilerin sayısı azaltılabilir.

Verilerin yeniden yapılandırılması: Veri madenciliğinde kullanılan model ve algoritmalar belirli türdeki veriler ile işlem yapabilmektedir. Bazı algoritmalar sadece sayısal yapıdaki verilerle işlem yaparken bazıları yalnızca kategorik türdeki verileri kullanır. Karar ağaçları aralıklı deęerleri kullanır. Örneğin ücret deęişkeni, 550 ile 15.000 TL arasında deęer alıyorsa bu deęerleri 550-1.000, 1.000-5.000, 5.001-10.000, 10.001-15.000 gibi aralıklarda bölebiliriz (Silahtaroglu 2008, s. 24).

2.2.3.3 Modelin kurulması ve kullanılması

Problem için en doęru modelin bulunması gerekmektedir. Bunun için birçok model oluşturmak ve bunları deneyerek en iyi modeli kullanmak gerekmektedir. Kullanılan modeller direk kullanılacağı gibi başka uygulamaların alt parça olarak da kullanılabilir. Örneğin model kredi deęerlendirme, risk analizi için doğrudan kullanılabilirken, promosyon uygulamalarında alt parça olarak kullanılabilir.¹³

2.2.3.4 Modelin izlenmesi

Zaman içerisinde sistemin ürettięi veriler deęişebilir. Kullanılan modelin izlenmesi ve gerekiyorsa yeniden düzenlemesi gerekmektedir.¹⁴

2.2.4 Veri Madencilięi Yazılımları

Veri madencilięi birçok alanda ve birçok konuda bilgi keşfi sürecinin önemli bir parçasıdır. Sınıflandırma, kümeleme ve birliktelik analizi gibi yöntemler ile verileri

¹³ Onat, A., (2008). Veri madencilięinin web tabanlı uygulamalarda insan uyumluluklarının tespiti üzerine bir çalışma. *Yüksek Lisans Tezi*. Konya: Selçuk Üniversitesi FBE.

¹⁴ Dolgun, M. Ö., (2006). Büyük alışveriş merkezleri için veri madencilięi uygulamaları. *Yüksek Lisans Tezi*. Ankara: Hacettepe Üniversitesi FBE.

analiz etmektedir. Bu yöntemlerin hepsinin kendi içinde kullandığı birçok algoritma bulunmaktadır. Veri madenciliği ile bilgi keşfi için bu algoritmaların kullanıldığı birçok uygulama geliştirilmiştir. Bu uygulamaları Ataseven (2008) şu şekilde tanımlamıştır:¹⁵

a) Darwin

Sağlayıcı: Oracle Corporation

Kullandığı Teknikler: Karar ağaçları (CART), K-Ortalama, K-En Yakın Komşu, sinir ağları, regresyon

Çalıştığı Platformlar: Windows, Sun Solaris, HP-UX

“Oracle Data Mining Suite” ismiyle de anılan Darwin yazılımı, Windows tabanlı kullanışlı bir ara yüz sağlamaktadır. Farklı bir kaç Veri Madenciliği algoritması içermektedir. ODBC üzerinde birçok ilişkisel veritabanı sistemine bağlanabilir. Sunucu – İstemci yaklaşımını kullanmaktadır. Ayrıca yapılan çalışmaların C, C++ ve Java kodlarını export edilebilmektedir.

<http://www.oracle.com/technology/documentation/darwin.html>

b) DBMiner

Sağlayıcı: DBMiner Technologies Inc.

Kullandığı Teknikler: Karar ağaçları, K-ortalama

Çalıştığı Platformlar: Windows

DBMiner, Simon Fraser Üniversitesi tarafından bir bilimsel çalışma olarak geliştirilmiştir . Çevrimiçi Analitik İşleme (Online Analytical Mining) özelliğine sahiptir. Bu özellik sayesinde OLAP üzerinden dinamik olarak Veri Madenciliği teknikleri uygulanabilmektedir. Arayüz olarak kullanıcılara GUI veya DMSQL seçenekleri sunulmaktadır.

<http://www.dbminer.com>

c) Intelligent Miner

Sağlayıcı: IBM Corporation

Kullandığı Teknikler: Karar ağaçları (modifiye edilmiş CART), K-ortalama, sinir ağları, regresyon

¹⁵ Ataseven, S., (2008). Üniversitelerin adaylar tarafından tercih edilme desenlerini veri madenciliği yöntemleri ile belirleyen bir model önerisi. *Yüksek Lisans Tezi*. İstanbul: İstanbul Kültür Üniversitesi FBE.

Çalıştığı Platformlar: Windows, Solaris, AIX, OS/390, OS/400

IBM firması tarafından geliştirilen Intelligent Miner, DB2 veritabanları veya düz metin dosyaları üzerinde Veri Madenciliği teknikleri kullanır. Ayrıca ODBC üzerinden diğer ilişkisel DBMS'ler üzerindeki verilere erişebilir. Sunucu – istemci yaklaşımını kullanır ve kullanıcılara veri madenciliği işlemleri için basit bir GUI sunmaktadır. Metin analiz araçları, arama makinası gibi özellikler eklenmiştir.

<http://www-306.ibm.com/software/data/iminer/>

d) Enterprise Miner

Sağlayıcı: SAS Institute Inc.

Kullandığı Teknikler: Karar ağaçları (CART ve CHAID), K-en yakın komşu, regresyon, sinir ağları.

Çalıştığı Platformlar: İstemci (Windows), Sunucu (Unix, Windows)

Enterprise Miner, SAS firması tarafından geliştirilen Veri Madenciliği çözümleri sunan üründür. Bilgi keşfi sürecini, firmanın kendi belirlediği SEMMA isimli bir süreç ile sağlamaktadır. Simge tabanlı bir GUI ile süreç akışı kullanıcılar tarafından yönetilebilmektedir. Veriden örneklem alma, veriyi görselleştirme, filtreleme, dönüştürme amaçlı araçları içinde barındırmaktadır.

e) Clementine

Sağlayıcı: SPSS Inc.

Kullandığı Teknikler: Karar ağaçları, sinir ağları, regresyon, K-ortalama, K-en yakın komşu, Naive-Bayes, faktör ve ayırma analizleri, temel bileşenler analizi, kümeleme, kohonen ağları

Çalıştığı Platformlar: Windows, HP/UX, IBM AIX, Sun Solaris.

Clementine, SPSS firmasının Veri Madenciliği çözümleri sunan üründür. Kendisini diğer yazılımlardan ayıran en önemli özelliği, veri madenciliğine 1994 yılından itibaren getirmiş olduğu GUI yaklaşımıdır. Açıklayıcı simgelerin kullanımı ile veri akış süreci yaratılmakta ve işleme konmaktadır. Her simgenin, veri keşfi süreci anlamında, bir işlevi bulunmaktadır. Veriye erişim, verinin hazırlanması, görsellik ve modelleme aşamaları için birbirinden farklı simgelere GUI üzerinden erişilmektedir. Clementine geniş veri kümeleri üzerinde veri madenciliğini istemci sunucu yaklaşımı ile yapar. Gerektiğinde, veri erişim isteklerini SQL sorgularına dönüştürebilir. Geniş veri tipi desteğine sahiptir. Elde edilen sonuçların değişik biçimlerdeki export

seenekleri bulunmaktadır.

2.2.5 Veri Madenciliđi Modelleri

Bir model, bir veri kümesinin yüksek seviyedeki global bir tarifidir. Geniş bir örnek yani bir perspektif sunar. Tanımlayıcı veya sonuç olarak çıkarılmış olabilir ki bu da bize verinin elde edildiđi popülasyon hakkında ya da gelecek muhtemel veri deđerleri hakkında bazı çıkarımlar yapma olanađı tanır (Hand ve diđ. 2001, ss. 165-170).

Veri madenciliđi, denetimli öğrenme ve denetimsiz öğrenme olmak üzere iki ayrı modelleme bulundurmaktadır.

Denetimli öğrenim, örnekten öğrenme olarak da tanımlanabilir. Bir uzman tarafından ilgili sınıflar belirli kriterlere göre ayrılarak, her sınıf için ayrı örnekler verilir. Sistemin amacı, verilen örneklerden hareket ederek her bir sınıfa ait ilişkilerin bulunması ve bu özelliklerin kural cümleleri ile ifade edilmesidir. Denetimli öğrenmede seçilen algoritmaya uygun olarak mevcut verilerin bir kısmı modelin öğrenimi, diđer kısmı ise modelin geçerliliđinin testi için kullanılır. Modelin öğrenimi, öğrenim kümesi kullanıldıktan sonra test kümesini kullanarak modelin dođruluk derecesi belirlenir (Ođuzlar 2011, ss. 62-63).

Denetimsiz öğrenme modelinde çıktı veya bađlı deđişken yoktur. Denetimli öğrenmenin tersine bu öğrenimde önceden belirlenmiş sınıflar bulunmamaktadır. Denetimsiz öğrenme, kümeleme analizlerinde olduđu gibi örneklerin analiz edilip bu örnekler arasındaki benzerliklerin bulunmasını amaçlamaktadır. Denetimsiz öğrenim veya kümelemede benzer örüntüler gruplandırılır (Ođuzlar 2011, s. 69).

Sınıflandırma, denetimli öğrenmeyi; kümeleme ve birliktelik yöntemleri ise denetimsiz öğrenmeyi temsil etmektedir.

2.2.6 Veri Madenciliđi Yöntemleri

Veri madenciliđi, bilgi keşfinde kullanmak için genel olarak üç temel yöntemi kullanmaktadır. Bunlar; sınıflandırma, kümeleme ve birliktelik kurallarıdır. Her yöntem için geliştirilmiş birçok algoritma bulunmaktadır.

2.2.6.1 Sınıflandırma

Verileri içerdikleri özelliklere göre ayrıştırma işlemi sınıflandırma olarak tanımlanır. Sınıflandırma, bir öğrenme algoritması kullanır. Tüm veriler kullanılır ve bir öğrenme işlemi yapılır. Diđer bir ifade ile sınıflandırma, hangi sınıfa ait olduđu bilinmeyen bir kayıtın sınıfının belirlenmesidir. Örnek olarak basit bir sınıflandırma yaparak,

“ödemelerini üç gün içinde yapabilen müşteriler” ve “ödemelerini üç günden sonra yapanlar”. Daha sonra yeni bir durum (bir karar durumu) karşılanmış olduğumuzda, belirlenen bu sınıflara bakarak karar verilebilir (Özkan 2008, ss. 45-51).

Sınıflandırma bir veriyi daha önceden belirlenmiş bir sınıfa atama yöntemidir. Örneğin web alanında, site yöneticileri sınıflandırma yöntemini kullanarak müşterilerin hangi sınıfta veya kategoride bir profile sahip olduğu belirlenebilir (Takçı ve Soğukpınar 2002, s. [sayfa yok]).

Sınıflandırma yöntemi için birçok algoritma geliştirilmiştir:

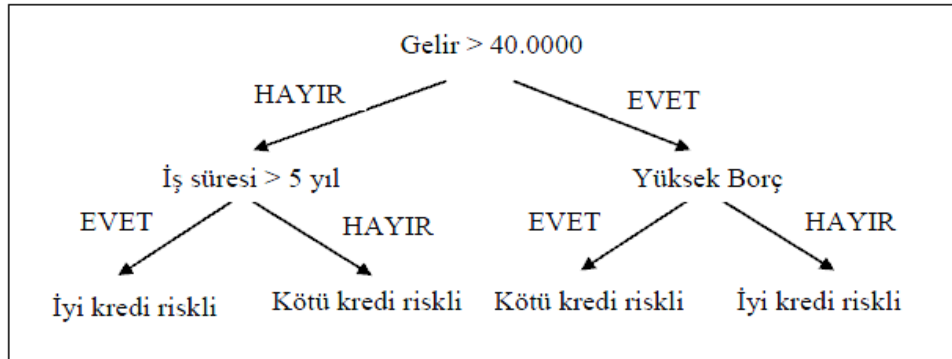
a) Karar Ağaçları

Silahtaroglu' un (2008, ss. 46-49) açıkladığı gibi:

Karar ağaçları sınıflandırma problemlerinde en çok kullanılan algoritmalarından biridir. Diğer yöntemlerle kıyaslandığında karar ağaçlarının yapılandırılması ve anlaşılması daha kolay denilebilir. Bu teknikte sınıflandırma için bir ağaç oluşturulur; daha sonra, veritabanındaki her bir kayıt bu ağaca uygulanır ve çıkan sonuca göre de bu kayıt sınıflandırılır. Temel olarak iki adımdan oluştuğu söylenebilir: Birincisi ağacın kurulması, ikincisi de verilerin teker teker ağaca uygulanarak sınıflandırmanın gerçekleşmesi şeklindedir.... Karar ağaçlarını oluştururken kullanılan algoritmanın ne olduğu önemlidir. Kullanılan algoritmaya göre ağacın şekli değişebilir. Değişik ağaç yapıları da farklı sınıflandırma sonuçları verecektir.... Örneğin veritabanında bulunan cevap evet/hayır gibiyse 2 eşit parçaya, evet/hayır/belki gibi üç değişkenliyse mümkün olduğunca 3 eşit parçaya bölünmesi istenmektedir. Buradaki amaç, en kısa yoldan istenilen yanıtı ya da sınıfa ulaştırmaktır.... Bu işlem yapılırken en önemli nokta hangi değişkenin ilk düğüm, yani “kök düğüm” olacağıdır. Burada öyle bir değişken seçilmeli ki yanıt ne olursa olsun, diğer değişkenlerle kıyaslandığında elimizdeki veritabanı kabaca iki eşit parçaya bölünebilmelidir.

Karar ağaçları için geliştirilmiş birçok algoritma bulunmaktadır. Bunlardan bazıları şunlardır: ID3, C 4.5, CART, SLIQ, SPRINT. Karar ağaçlarına ait bir örnek Şekil 2.4' de gösterilmiştir.

Şekil 2.4: Karar ağaçları



Kaynak: Bilen, H., (2009). Bankacılık sektöründe personele seçimi ve performans değerlendirmesine ilişkin veri madenciliği uygulaması. Yüksek Lisans Tezi. Ankara: Gazi Üniversitesi FBE.

b) Genetik Algoritmalar

Genetik algoritmalar Darwin tarafından geliştirilen “evrim teorisi” ne dayalıdır. Algoritma ilk önce popülasyon adı verilen bir öğrenme kümesi ile başlatılır. Bir popülasyondan alınan sonuçlar daha iyi olacağı düşünülen yeni bir popülasyon oluşturmak için kullanılır. Evrim süreci tamamlandığında bağımlılık kuralları veya sınıf modelleri ortaya konmuş olur.¹⁶

c) Yapay Sinir Ağları

Yapay sinir ağları, insanlığın doğayı taklit etmesi çabalarından birisi olarak tanımlanabilir. Yapay sinir ağları, basit biyolojik sinir sistemlerinin çalışma şeklini canlandırmak için tasarlanmışlardır. Bir yapay sinir ağının öğrenme yeteneği, kullanılan ağırlık oranıyla doğrudan orantılıdır. Sistem için kullanılan girdiler dışarıdan elde edilen bilgilerdir. Toplam fonksiyonu, bir hücreye gelen net girdi miktarı olarak tanımlanabilir. Aktivasyon fonksiyonu, öğrenilme sonucu oluşan değerlerin ortaya çıkarılması için kullanılan bir fonksiyondur. Çıktı ise, aktivasyon fonksiyonundan elde edilen değer olarak tanımlanabilir.¹⁷

d) Bayesyen Sınıflandırma

Bayesyen sınıflandırma tekniği, hali hazırda sınıflandırılmış verileri kullanarak yeni bir verinin mevcut sınıflandırmalardan herhangi birine girmesini hesaplayabilen bir sınıflandırma yöntemidir (Silahtaroglu 2008, s. 60).

e) K- En Yakın Komşu

Sınıflandırma işlemi yapılırken veritabanı içerisinde bulunan her verinin diğer verilere olan uzaklığı hesaplanır. Her bir veri için k adet veri ile birbirine olan uzaklığı hesaplanır. K sayısı konunun uzmanı tarafından belirlenir.

2.2.6.2 Kümeleme

Kümeleme yöntemi tıpkı sınıflandırma yönteminde olduğu gibi verileri gruplara ayırma işlemidir. Sınıflandırma yönteminde sınıflar önceden belirlenmektedir. Kümeleme yönteminde ise önce tanımlı sınıflar bulunmamaktadır. Verilerin hangi gruplara

¹⁶ Akbulut, S., (2006). Veri madenciliği teknikleri ile bir kozmetik markanın ayrılan müşteri analizi ve müşteri segmentasyonu. *Yüksek Lisans Tezi*. Ankara: Gazi Üniversitesi FBE

¹⁷ Bilen, H., (2009). Bankacılık sektöründe personel seçimi ve performans değerlendirmesine ilişkin veri madenciliği uygulaması. *Yüksek Lisans Tezi*. Ankara: Gazi Üniversitesi FBE.

ayrılacağı verilerin birbirine olan benzerliklere göre değişmektedir. Belirlenen gruplara literatürde kümeleme denilmektedir. Kümeleme yöntemi biyoloji, pazarlama, ekonomi gibi birçok alanda kullanılmaktadır. Örneğin elimizde bir süpermarkete ait bir veritabanında müşterilere ait bilgilerin olduğunu düşünelim. Müşterileri yaşlarına göre kümelemek isteyebiliriz. Yaşları birbirine yakın olanlar aynı kümede yer alacaktır. Müşterilerin yaşlarının 20, 22, 24, 26, 27, 38, 42, 46, 50 olduğunu varsayarsak, yaşı 20 olan müşteri ile 27 olan aynı grupta yer alabilir. Fakat yaş aralıkları 19, 20, 21, 22, 23, 24, 25, 26, 27 olduğunu varsayarsak, 20 ile 27 yaşlarındaki iki kişi farklı kümelerde yer alır. Bu da bize verilerin sadece kendi değerlerinin değil, içinde bulunduğu veri yığını içerisindeki diğer değerlerle olan benzerliğine göre kümelendiğini gösterir (Silahtaroglu 2008, s. 99).

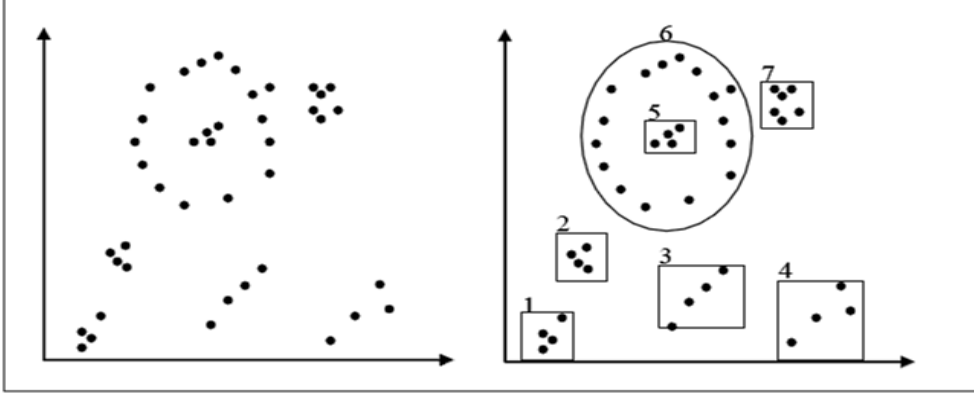
Kümeleme yöntemi için hangi algoritma kullanılırsa kullanılsın süreç hep aynı şekilde işlemektedir. Her kayıt, veri grubunda bulunan bütün kümelerle karşılaştırılır ve onlara olan uzaklığı hesaplanır. Kendisine en yakın olan kümeye atanır ve bu kümenin değerini değiştirir. Optimum duruma kadar kayıtlar yeniden atanır ve kümelerin merkezleri ayarlanır.¹⁸

Kümeleme yöntemlerinin amacı, veritabanında bulunan verilerin aralarındaki benzerliklere bakılarak ayrı kümelere atanmasıdır. Bunun sonucunda ortak özelliklere sahip küme üyeleri aynı kümelere atanacak, birbirine benzemeyen küme üyeleri diğer kümede kalacaktır. En yaygın kullanılan kümeleme algoritması k-ortalama algoritmasıdır (Timor ve diğ. 2011, s. 131).

Verilerin kümelere atanması yapılırken kullanılan en önemli kavramlar “benzerlik ve uzaklık” kavramlarıdır. Kümeleme ile kavramı Şekil 2.5’ de gösterilmiştir.

¹⁸ Akbulut, S., (2006). Veri madenciliği teknikleri ile bir kozmetik markanın ayrılan müşteri analizi ve müşteri segmentasyonu. *Yüksek Lisans Tezi*. Ankara: Gazi Üniversitesi FBE.

Şekil 2.5: Kümeleme



Kaynak: Bilen, H., (2009). Bankacılık sektöründe personel seçimi ve performans değerlendirmesine ilişkin veri madenciliği uygulaması. Yüksek Lisans Tezi. Ankara: Gazi Üniversitesi FBE

Silahtaroglu' nun (2008, s.100) açıkladığı gibi:

Veritabanındaki veriler kümelere ayrılırken, benzerlik ve uzaklık kavramlarından yararlanır. Bu, veritabanındaki her bir kaydın diğer bir kayıtlarla olan benzerliğini ya da her bir kaydın veritabanındaki diğer kayıtlardan olan uzaklığı olduğu gibi oluşturulan gerçek ve aday kümeler arasındaki mesafe ve benzerliği de içerir. Sözelimi, veriler birbirine olan uzaklığına göre başlangıçta 8 ayrı kümeye ayrılırlarsa, bu 8 ayrı kümenin gerçekten farklı özelliklere sahip birer küme olup olmadığının hesaplanması gerekecektir. Bu durumda, oluşturulmuş bu kümeler arasındaki mesafe/benzerlik de ölçülmelidir. Birbirlerinden pek farklı olmayan kümeler birleştirilerek tek bir küme haline dönüştürülebilir.... Bu nedenle kümeler arasındaki mesafenin ölçülmesiyle iki veya daha fazla kümenin birleşmesi söz konusu olduğu gibi aynı zamanda, bir kümeden birden fazla küme üretilmesi de söz konusudur. Bunun için de sürekli olarak kümelerin büyüklüğü ve çapı ölçülmelidir.

Kümeleme algoritmaları; hiyerarşik, bölümlenmeli, yoğunluğa dayalı, grid temelli gibi birçok farklı alt bölüme ayrılabilir.

a) K Ortalamalar Algoritması

Diğer kümeleme algoritmalarıyla karşılaştırıldığında k ortalamalar algoritması büyük veritabanlarının kümelenebilmesinde kullanılan etkin bir algoritmadır. Yeni bir vaka ortaya çıktığında, algoritma tüm verileri tarayarak buna en çok benzeyen durumların bir alt kümesini oluşturur ve onları çıktığı tahmin etmek için kullanır. Bu algoritmanın adımları şöyledir:

- I. Veri grubu rassal olarak k adet başlangıç kümesine ayrılır.
- II. Veri grubunda yer alan veriler, merkezi kendisine en yakın olan kümeye atanır.
- III. Her atamanın sonunda küme merkezi yeniden hesaplanır.
- IV. Veri grubundaki tüm verilerin ataması yapılmaya kadar 2. ve 3. adımlar tekrarlanır.

K deęerinin ve uzaklık ölçüsünün modelin kalitesi üzerinde büyük oranda etkisi bulunmaktadır. Bu nedenle bu deęerleri dikkatli seçmek gerekmektedir.¹⁹

b) Cure Algoritması

Kümeleme işlemleri yapılırken, oluşturulan kümelerin kalitesini etkileyen en önemli faktörlerden birisi de ana veri grubu içerisinde bulunan verilerden uzakta duran ve hiçbir kümeye ait olmaması gereken uç verilerdir. CURE(Clustering Using Reprisentatives) bu uç verilerin oluşturulan kümelerin kalitesini etkilememesi için 1998 yılında geliştirilmiş bir algoritmadır. CURE algoritması her veriyi ayrı bir küme olarak görür ve her adımda bu verileri birbirine olan yakınlıklarına göre birleştirir veya ayrı küme olarak tanımlar. Her bir küme (veya veri) için c adet temsilci seçilir. Seçilen bu noktalar kümelerin fiziksel şeklini geometrik olarak tanımlar. Sonra bu noktalar bir alfa katsayısı ile merkezine doğru kaydırılır. Bundan sonra iki küme arasındaki uzaklık, her bir kümeye ait olan en yakın temsilci çifti uzaklığıdır (Silahtaroglu 2008, s.109).

2.2.6.3 Birliktelik kuralları

Birliktelik kuralları veri madenciliğinin tanımlayıcı yöntemlerindedir. Birliktelik yöntemleriyle büyük veritabanları içerisindeki ilginç birliktelik örüntüleri ortaya çıkarılabilir. Ortaya çıkarılan bu örüntüler pazarlama ve karar verme gibi birçok konuda fayda sağlamaktadır. Birliktelik kurallarının en yaygın kullanıldığı alanlardan birisi de pazar sepeti analizidir. Bu yöntemle yapılan alış-verişler sırasında satın alınan farklı ürünler arasındaki birliktelikler ve ilişkiler keşfedilebilmektedir (Ay ve Çil 2008, s. 17).

Birliktelik kuralları, bir koşul kümesi ile kısmi bir sonuca ulaşma durumudur. Sürecin sonunda en iyi birlikteliklerin bir tablosu oluşmaktadır (Oğuzlar 2004, s. 315).

Birliktelik kuralları oluşturulurken ilk önce destek ve güven seviyeleri oluşturulmalıdır.

Destek seviyesi Y içindeki işlemlerin ne kadarının X' i kapsadığıdır. Bir veritabanında, birlikteliklerin ortaya çıkması için kullanıcının vereceği en

¹⁹ Akbulut, S., (2006). Veri madenciliği teknikleri ile bir kozmetik markanın ayrılan müşteri analizi ve müşteri segmentasyonu. *Yüksek Lisans Tezi*. Ankara: Gazi Üniversitesi FBE

küçük destek seviyesi ve en küçük güven seviyesinden daha büyük güven ve destek seviyelerine sahip kuralların tespit edilebilmesidir. En küçük destek seviyesinde sahip olan verilere geniş nesne kümesi, diğerlerine ise küçük nesne kümeleri denilmektedir. Bağlantı analizi yapıldıktan sonra ortaya çıkan sonuç aşağıdaki gibi ifade edilir.

Yaş (kişi, “20-30”) → satın alır (kişi, “LCD tv”)[Destek= %2, Güven=%13]

Yukarıdaki ilişki analizi, yaşı 20-30 arasındaki kişilerin LCD TV almaya yatkın olduğu, bugüne kadar alış veriş yapmış müşterilerin %2’ sinin bu yaş aralığında olduğunu ve bu yaş aralığındaki kişilerin %13’ ünün LCD TV aldığını ifade etmektedir (Silahtaroglu 2008, ss. 84-85).

Birliktelik kurallarının bulunması için; AIS, SETM, APRIORI gibi birçok algoritma geliştirilmiştir. Bunlardan en çok bilinen ve kullanılan algoritma apriori algoritmasıdır.

Apriori algoritması, arama uzayını indirgemede aşağıdaki teoremi

kullanır: Eğer bir öge küme sıkça tekrarlanıyorsa kendisinin altkümelerinin hepsi de sık sık tekrarlanmaktadır. Bunun anlamı şudur: i öge kümesini kullanma potansiyel i+1 öge kümeyi üretmeyi mümkün kılmaktadır. i+1 aday öge kümenin her bir alt kümesi tekrarlanmalıdır. Burada veritabanının tekrarlı taraması yapılarak tüm tekrarlı öge kümeleri bulmak mümkündür. i’inci veritabanı taraması süresince Apriori algoritması i öge kümenin meydana çıkma sayısını sayar ve i geçişin sonunda i+1 öge içeren adayları üretir.²⁰

²⁰ Sıramkaya, E., (2005). Veri madenciliğinde bulanık mantık uygulaması. *Yüksek Lisans Tezi*. Konya: Selçuk Üniversitesi FBE.

3.METİN MADENCİLİĞİ

İnternetin hayatımıza girmesiyle ve dijital depolama aygıtlarının kapasitelerinin hızla artmasıyla depolanan veri miktarı çok büyük boyutlara ulaşmıştır. Bilgisayarlar hayatımıza girmeden önce dokümanlarla ilgili her türlü işlem elle yapılmaktaydı. Bugün internette 2 milyardan fazla web sayfası olduğu düşünülürse bu bilgilere eski yöntemlerle ulaşmak neredeyse imkansızdır.²¹

Yapısal veriler, buldukları yapı içerisinde organize edilebilen ve tanımlanabilen veriler için kullanılan bir tanımlamadır. En yaygın kullanılan yapısal veri kaynakları SQL (Structured Query Language) ve Access gibi veri tabanlarıdır. Bunlar kolon ve satır düzeyinde tanımlanıp verilere kayıt bazlı ulaşılabilir. Yapısal olmayan verilerin ise tanımlanabilir bir yapısı yoktur. En bilinen türleri; resim dosyaları, word, text dokümanları, web logları ve e-postalardır. Birçok kurumun verileri yapısal olmayan formattadır. Merrill Lynch, potansiyel olarak kullanılabilir verilerin yüzde 80' inin yapısal olmayan formatta olduğunu belirtmiştir. Metin ve web madenciliği uygulamaları, veri madenciliği algoritmalarının verileri işlemesi için kullanılan bir ön aşama gibi tanımlanabilir. Metin madenciliği çok büyük kapasitedeki belgelerin analiz edilip gizli örüntülerin elde edilmesi şeklinde tanımlanabilir. Web madenciliği ise web sayfalarının içerik, yapı ve bağlantı (link) bakımından incelenip analiz edilmesi şeklinde tanımlanabilir (Dolgun ve diğ. 2009, s. 49).

Bu bölümde tezin konusu olan metin madenciliğini ve onunla aynı grupta yer alan web madenciliğinin kullanılması için gereken ön işlemleri ve çeşitli algoritmalar anlatılmıştır.

²¹ Seçkin, K., (2011). Metin madenciliğinde kullanılan yöntemlerin karşılaştırılması: Siyasi parti liderlerinin grup genel toplantı konuşmaları ile bir uygulama. *Yüksek Lisans Tezi*. Sakarya: Sakarya Üniversitesi SBE.

3.1 METİN MADENCİLİĞİ ADIMLARI

Veri madenciliğinde olduğu gibi, metin madenciliğinde de verilerden anlamlı bilgiler elde edebilmek için verilerin çeşitli aşamalardan geçmesi gerekmektedir. Verilerin doğru şekilde analiz edilmesi, algoritmaların doğru bir şekilde kullanılabilecek hale getirilmesi, doğru sonuçlar alınması açısından çok önemlidir. Metin benzerlik, metin sınıflandırma ve web madenciliği uygulamaları için metinlerin temizlenip yapılandırılması gerekmektedir.

3.1.1 Temizleme İşlemleri

Metin madenciliğinde, dokümanların işlenmesinde yapılacak ilk adım metinlerin öğrenme algoritmaları, sınıflandırma algoritmaları ve benzerlik algoritmaları için uygun formata getirilmesidir. Bu nedenle bir doküman grubunun üzerinde işlem yapmadan önce dokümanların temizlenmesi ve ayrıştırılması gerekmektedir. Bu işlem eğer web sayfaları üzerinde yapılacak ise sayfa, XML (Extensible Markup Language) ve HTML (Hyper Text Markup Language) gibi her türlü etiketten ayrıştırılmalıdır. Daha sonra sayfada bulunan tüm harfler küçük harfe çevrilir. Harf olmayan karakterlerin yerine boşluk karakteri yerleştirilir.²²

En son işlem olarak doküman, kullanılacak olan yöntemle göre kelimelere ve cümlelere ayrılır. Eğer metin benzerliği yöntemi uygulanacaksa doküman cümlelere, eğer metin sınıflandırma yöntemi uygulanacaksa kelimelere göre ayrıştırılır.

3.1.2 Durdurma Kelimelerin Çıkarılması

Uygulamada kullanılacak kelimeler belirlendikten sonra doküman içerisinde çok sık geçen fakat kendi başlarına bir anlamları olmayan kelimeler doküman içerisinden çıkarılır. Türkçede kullanılan edatlar / bağlaçlar bunlara örnek olabilir (ama, fakat, için, sonra, ile gibi). Bu kelimelerin bilgiye erişim sistemlerinde gerekli olmadığı bu sistemlerle yapılan ilk çalışmalardan itibaren bilinmektedir. Ayrıca bu kelimeler dokümanlarda çok fazla yer tutarak işlem

²² Rauka, L., (2012). Metinsel veri madenciliğinde bilgisayarlı çeviriciler. *Yüksek Lisans Tezi*. Trabzon: Karadeniz Teknik Üniversitesi FBE.

hızını düşürmektedir. Bu kelimeler genelde istatistiksel yöntemlerle bulunabilse de genelde hepsi bir listede tutulmaktadır.²³

3.1.3 Gövdeleme

Kelimelerin temizleme, ayrıştırma işlemleri bittikten sonra kelimelerin kökleri bulunur. Yani koşucular, koşucu, koşmak, koş gibi kelimeler aynı kök grubundan oldukları için bir arada toplanmış olurlar. Fakat kök bulma işleminde karşılaşılabilecek iki sorun vardır. Birincisi farklı kelimeler aynı anlam ve kök grubuna dâhil olabilir. İkinci önemli sorun ise kelimelerin köklerine ulaşmaya çalışırken çok az ekin çıkarılmasıdır. Bu durumda da sistem konuya uygun olan kelimeleri uygun değilmiş gibi algılayacaktır. Kelime kökleri bulmak için kullanılan örnek bir yapı Tablo 3.1’ de gösterilmiştir.

Tablo 3.1: Kelime kökleri

Yapar	Yapmak
Yapılır	Yapmak
Yapabilir	Yapmak
Yapılırsa	Yapmak
Yaparsa	Yapmak
Yapabilmek	Yapmak

Kaynak: Rauka, L., (2012). Metinsel veri madenciliğinde bilgisayarlı çeviriciler. Yüksek Lisans Tezi. Trabzon: Karadeniz Teknik Üniversitesi FBE.

Bu yöntemin dezavantajı çok fazla saklama alanı gerektirmesidir. Başka bir yöntem ise mevcut dokümanların oluşturulan bir sözlük içerisinde her kelimenin baş harfi alınarak ardışık farklılıklar karşılaştırılmaya başlar. Kökü bulunacak kelimenin sözlük içinde farklı bir kelime olarak bulunan n harfi kelimenin kökü olarak kabul edilir. Bu yöntemler her dil için geçerli yöntemlerdir. Veri kümesi İngilizce metinlerden oluşan çalışmalarda, Porter

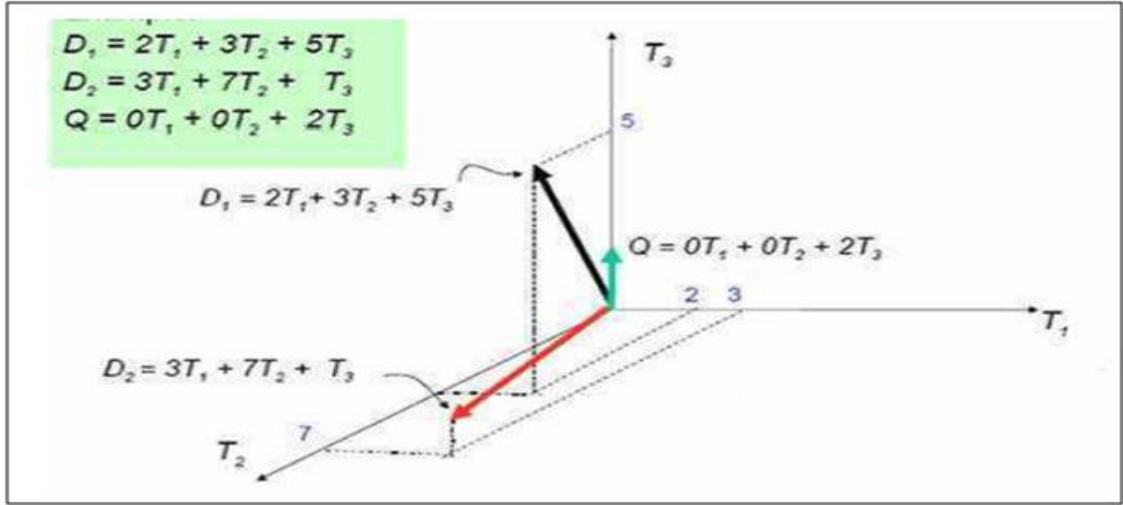
²³ Çelikyay, E. K., (2010). Metin madenciliği yöntemiyle türkçede en sık kullanılan ve birbirini takip eden harflerin analizi ve birliktelik kuralları. *Yüksek Lisans Tezi*. İstanbul: Beykent Üniversitesi FBE.

Stemmer algoritması basitlik ve performans bakımından diğerlerinden daha iyi olduğundan en sık başvurulan algoritmadır.²⁴

3.1.4 Kelimelerin Yapısallaştırılması

Dokümanların yapısallaştırılmasını Seçkin (2011, ss. 26-29) şöyle açıklamıştır: Metin halinde bulunan verilerin sınıflandırılması ya da üzerinde hesaplama işlemleri yapılabilmesi için bilgisayar ortamında incelenmesini sağlayacak farklı gösterim şekillerine gösterilmesi gerekmektedir. Metinlerin yapısal olmayan formdan sayısal hale getirilerek dönüştürülmesini sağlayan en yaygın yöntem vektör uzay modelidir. Her doküman, anahtar kelimelerden oluşan $M \times N$ boyutlu bir vektör olarak ifade edilir. M tüm dokümanlardaki farklı bir kelime ve N de elde bulunan tüm doküman sayısıdır. Bu vektördeki her eleman değeri, bir kelimenin o dokümanda ne sıklıkla geçtiğini sayısal olarak ifade eder. Vektör uzay modeli Şekil 3.6' da verilmiştir.

Şekil 3.1: Vektör uzay modeli



Kaynak: Seçkin, K., (2011). Metin madenciliğinde kullanılan yöntemlerin karşılaştırılması: Siyasi parti liderlerinin grup genel toplantı konuşmaları ile bir uygulama. Yüksek Lisans Tezi. Sakarya: Sakarya Üniversitesi SBE

Ağırlıklandırma, bir kelimenin bir doküman içerisindeki bulunma sıklığının gösterir.²⁵

²⁴ Rauka, L., (2012). Metinsel veri madenciliğinde bilgisayarlı çeviriciler. *Yüksek Lisans Tezi*. Trabzon: Karadeniz Teknik Üniversitesi FBE.

²⁵ Seçkin, K., (2011). Metin madenciliğinde kullanılan yöntemlerin karşılaştırılması: Siyasi parti liderlerinin grup genel toplantı konuşmaları ile bir uygulama. Yüksek Lisans Tezi. Sakarya: Sakarya Üniversitesi SBE.

Kelimelerin ağırlıklandırılmasını Miner ve diğ. (2012 s.50) şu şekilde açıklamışlardır:

Kelimelerin ağırlıklandırılması binary veya integer olarak yapılabilir. Binary türünde yapılan işlemlerde dokümanlar cümle içinde geçip-geçmeme durumlarına göre yalnızca 1 veya 0 değerlerini alırlar. Integer olarak yapılan işlemlerde ise kelimenin diğer cümle içerisinde kaç kere geçmiş olduğuna bakılır. Eğer kelime cümle içerisinde birden fazla geçiyorsa 2, 3, 4 şeklinde ağırlıklandırılır. Bu durumu şu örnekle açıklayabiliriz:

Document1: My dog ate my homework.

Document2: My cat ate the sandwich.

Document3: A dolphin ate my homework.

Bu vektör uzayında toplam 15 adet kelime bulunuyor. Bunlardan 9 adet kelime farklıdır. Bunları alfabetik sırayla şöyle gösterebiliriz:

A(1), ate (3), cat (1), dolphin (1), dog (1), homework (2), my (3), sandwich (1), the (2).

Bu kelimeleri geliştirilen uygulamalarda işlemek için aşağıdaki şekilde dizi haline getirilmektedir:

Binary vektör için;

Doc1: 0,1,0,0,1,1,1,0,0 (“my” kelimesi cümle içerisinde iki kere geçmiştir fakat binary formatında işlem gördüğü için “1” olara işleme alınmıştır.)

Doc2: 0,1,1,0,0,0,1,1,1

Doc3: 1,1,0,1,0,1,0,0,1

Integer vektör için;

Doc1: 0,1,0,0,1,1,2,0,0 (“my” kelimesi cümle içerisinde iki kere geçmiştir. Bu nedenle “2” olarak değerlendirilmiştir.)

Doc2: 0,1,1,0,0,0,1,1,1

Doc3: 1,1,0,1,0,1,0,0,1

(Miner ve diğ. 2012, s. 50).

3.2 WEB MADENCİLİĞİ

İnternet inanılmaz bir şekilde yaygınlaşarak kullanıcıların dünyasındaki en önemli unsurlardan birisi olmuştur. İnternet siteleri artık kullanıcıları tanımak istemektedir. Üye bilgilerini kullanarak, onları yönlendirmekte ve daha fazla bilgi sahibi olmaktadır

(kullanıcı bilgileri, çerezler vs...). Bu sayede daha fazla müşteriye ve kâra sahip olmak istemektedirler. Web madenciliği genel olarak web sayfalarında bulunan bilgilerin analiz edilip yorumlanmasıdır. Web madenciliği ilk kez Oren ETZONI tarafından dile getirilmiştir.²⁶

Web madenciliği üç alt başlıkta incelenebilir. Bu durumu Rouka (2012) şu şekilde tanımlamıştır:

a) Web İçerik Madenciliği

Web içerik madenciliği veri madenciliğiyle ilgili olduğu için veri madenciliğinin birçok tekniğini kullanmaktadır. Ayrıca veriler metin şeklinde olabildiği için metin madenciliğiyle de iç içedir. Web içerik madenciliği hem yapılandırılmış hem de yapılandırılmamış veriler ile çalışmaktadır. Web içerik madenciliği çalışmanın amacına göre web sayfasının hyperlinklere bağlı olarak sınıflandırılması, belirli bir alan adının yapısal hyperlink ağını veya web sitesinin yapısını gösteren raporlar üretebilir.

b) Web Yapı Madenciliği

Web yapı madenciliği Graph teorisini kullanarak web sayfaları arasındaki linkleri takip eder ve bilgi üretir.

c) Web Kullanım Madenciliği

Web kullanım madenciliği, kullanıcıların sitedeki geçmiş hareketlerinden veya log dosyalarından faydalanarak bilgi çıkarma işlemidir. Web kullanım madenciliği internet kullanıcılarının internette ne aradıklarını tespit etmeye çalışmaktadır. İstemcilerden gelen her kayıt, her istek log dosyalarında tutulmaktadır. Bu log dosyalarında kullanıcı desenleri bulunmaktadır. Bu bilgiler sayesinde kullanıcılar hakkında detaylı bilgiye ulaşılabilir.²⁷

3.3 METİN SINIFLANDIRMA

Metin sınıflandırma, önceden belirlenmiş sınıflara doküman atamaktadır. Sınıflandırma yapılmadan önce sınıfların belirlenmesi gerekmektedir. Dokümanların vektörel ifadeleri kullanılarak elde edilen veriler benzerlik ölçüm sonuçlarına ve uygulanacak

²⁶ Haberal, İ., (2007). Veri madenciliği algoritmalarını kullanarak web günlük erişimlerinin analizi. *Yüksek lisans Tezi*. Ankara: Başkent Üniversitesi FBE.

²⁷ Rauka, L., (2012). Metinsel veri madenciliğinde bilgisayarlı çeviriciler. *Yüksek Lisans Tezi*. Trabzon: Karadeniz Teknik Üniversitesi FBE.

algoritmalarla sınıflandırma gerçekleşir. Dokümanların uzunluğu metin madenciliğinde en önemli sorunlardan birisidir. Dokümanlar ön işlemlerden geçirilerek ve boyutları azaltılarak bu sorunlar giderilebilir. Metin sınıflandırma uygulamalarında eğitim dokümanları yaratılmaktadır. Dokümanların sınıflandırılması, bu eğitim dokümanları baz alınarak yapılmaktadır. Kullanılan eğitim dokümanının azlığı sözcüklerin tam doğru tespit edilememesine neden olur ve sınıflandırma başarısını düşürür.²⁸

3.4 METİN BENZERLİK HESAPLAMA YÖNTEMLERİ

Metin madenciliğinde metinlerin ağırlıklarının hesaplanması için iki yöntem bulunmaktadır. Birinci yöntem kelimelerin kavramsal olarak veya kelime kökü olarak bir sözlük ile karşılaştırılıp ağırlıklandırıldığı veya bir sınıfa atandığı yöntemdir. İkinci yöntem ise kelimelerin yalın haliyle doküman içerisinde geçip-geçmeme durumunun kontrol edildiği ve kelimelerin kullanılacak algoritmalar için yapılandırıldığı yöntemdir.

3.4.1 Sözlük Kullanarak

a) Balkanet ve Türkçe Kavramsal Sözlük

Balkanet projesi; Bulgaristan, Çek Cumhuriyeti, Yunanistan, Romanya, Türkiye ve Sırbistan'ın oluşturduğu Balkan ülkelerinin dillerine ait kelimelerin, bilgisayar ortamında işlenerek sözlükler arasında anlamsal ilişkiler barındıran bir sözlük çalışmasıdır. Sözlükteki anlam ilişkileri, kavramlar aracılığıyla tanımlanmaktadır. Bu sebeple hızlı ve sağlıklı bir yapı oluşturmaktadır. Balkanet bünyesindeki bütün sözcüklerin ilişkileri tablo xxx gösterilmiştir.²⁹ Balkanet sözlüğüne ait bir örnek tablo 3.2' de gösterilmiştir.

²⁸ Karaca, M. F., (2012). Metin madenciliği yöntemleri ile haber sitelerindeki köşe yazılarının sınıflandırılması. *Yüksek Lisans Tezi*. Karabük: Karabük Üniversitesi FBE

²⁹ Yücesoy, B., (2007). Türkçe dokümanlar için anlamsal benzerlik hesaplama yöntemi. *Yüksek Lisans Tezi*. İstanbul: İstanbul Teknik Üniversitesi FBE.

Tablo 3.2: Balkanet Sözlüğü

İlişki	Açıklama	Örnek
Synonymy	Eş anlam	(canlılık, hayatiyet)
Near_antonym	Yakın zıt anlam	(almak) - (satmak)
Category_domain	Aynı kategoride olma	(hücre) - (biyoloji)
Also_see	Ayrıca ilişkisi	(giyinmek) - (şık giyinmek)
Be_in_state	Durumunda olma	(uzun:1) - (boy pos:2, yükseklik:3)
Subevent	Alt olay	(bulmak) - (mantık yürütmek)
Causes	Sebeplik olma	(zayıflatmak) - (zayıflamak)
Hypernym (superordinat)	Üst kavram	(reçel) - (gül reçeli)
Hyponym (subordinat)	Alt kavram	(gül reçeli) - (reçel)
Verb_group	Fiil grubu	(act:5 play:8 act-as:2) - (dissemble:3 pretend:2 act:9)
Similar_to	Yakın anlam	(müteakip) - (sıradaki) - (bir sonraki)
Usage_domain	Aynı kullanım alanı	
Holo_member	Üye - bütünü	(Enterobacteriales) - (bakteri)
Mero_member	Bütün - üyesi	(bakteri) - (Enterobacteriales)
Holo_part	Kısmi Parça - bütünü	(zeytin) - (zeytin ağacı)
Mero_part	Bütün - kısmi parçası	(zeytin ağacı) - (zeytin)
Derived_From	Türetilmiş	(hızlıca) - (hızlı)
Region_domain		
Holo_madeof	Hammadde - bütünü	(kâğıt) - (kitap)
Mero_madeof	Bütün - hammaddesi	(masa) - (tahta)

Kaynak: Yücesoy, B., (2007). Türkçe dökümanlar için anlamsal benzerlik hesaplama yöntemi. Yüksek Lisans Tezi İstanbul: İstanbul Teknik Üniversitesi FBE.

b) Zemberek

Zemberek, Tspell projesinin geliştirilmiş halidir. Türk diline ait bilgi işlem problemlerinin çözülmesi için geliştirilmiş platform bağımsız bir doğal dil işleme kütüphanesi oluşturulması amacıyla başlamıştır. Türkçe kelimeleri denetleme, kelime

önerme, oluşturma gibi işlemler için oluşturulmuştur. İçerisinde yaklaşık olarak 20.000 kök ve 6.000 adet özel isim içeren bir sözlüktür.³⁰

3.4.2 Sözlük Kullanmayarak

Sözlük kullanılmadan kelimelerin birbirine olan benzerliğini hesaplayan birçok uzaklık ölçüsü bulunmaktadır.

Işık ve Çamurcu' nun pearson ve kosinüs algoritmalarını (2008, s.39) açıkladığı gibi:

a) Pearson

Pearson ilişkisinde, kosinüs benzerliğindeki gibi iki vektörün benzerliğinde vektörlerin aralarındaki açıya göre karşılaştırma yapılır. Kosinüs benzerliğinden farkı iki vektörün iç çarpımı yapılmadan önce her birinin ayrı ayrı ortalama değerleri hesaplanır ve her ortalama değer ait olduğu vektörün tüm elemanlarından çıkarılır. d ve d^* birbirinden farklı iki belgeyi temsil eden çok boyutlu vektör ise aralarındaki pearson benzerliği Denklem 3.1 ile hesaplanır.

$$S^{Pearson}(d, d^*) = \frac{1}{2} \left(\frac{\sum_{i=1}^n (d_i - \bar{d}_i)(d_i^* - \bar{d}_i^*)}{\sqrt{\sum_{i=1}^n (d_i - \bar{d}_i)^2 \sum_{i=1}^n (d_i^* - \bar{d}_i^*)^2}} + 1 \right) \quad (3.1)$$

Çıkan sonucu [0,1] aralığında tutmak için normalizasyon işlemi yapılmaktadır. Bunun için hesaplanan değere 1 eklenip daha sonra da 2'ye bölmek yeterli olacaktır. Pearson ilişkisi değeri 1'e yaklaştıkça iki vektörün birbirlerine olan benzerlikleri de artar.

b) Kosinüs

Belge kümelemede çok kullanılan vektör tabanlı bir ölçüt olan kosinüs (cosine) Benzerliği ile iki vektör arasındaki açının kosinüs değeri hesaplanarak vektörlerin benzerliği bulunur. Vektör boyutundan etkilenmemesi, kosinüs benzerliğinin güçlü bir özelliğidir. Farklı çok sayıda kelimeler içeren benzer içerikteki belgeleri kolaylıkla tespit eder. Denklem 2 de görüldüğü gibi, vektörlerin skaler çarpımlarının, genliklerine bölünmesiyle iki vektör arasındaki açı elde edilir. İki vektör arasındaki açı ne kadar 0'a yaklaşırsa, açının kosinüs değeri 1'e yaklaşır ve iki vektörün birbirlerine olan benzerlikleri de artar.

Denklem 2 de iki vektör arasındaki açının kosinüs değerinin hesabında, d ve d^* birbirinden farklı iki belgeyi temsil eden çok boyutlu vektörleri ve “•” vektörlerin iç

³⁰ Dursun, B., (2008). Türkçe metinlerin benzerliğinin hesaplanması için yeni bir yöntem. *Yüksek Lisans Tezi*. İstanbul: Yıldız Teknik Üniversitesi FBE

çarpımını, $|d|$ ise vektörün uzunluğunu temsil etmektedir. Kosinüs benzerliği Denklem 3.2' de gösterilmiştir.

$$\cos(\theta) = \frac{d \cdot d^*}{|d| |d^*|} = \frac{\sum_{i=1}^n d_i d_i^*}{\sqrt{\sum_{i=1}^n (d_i)^2} \sqrt{\sum_{i=1}^n (d_i^*)^2}} \quad (3.2)$$

c) Jaccard

Bu benzerlik ölçme yöntemi de kosinüs yöntemi gibi doküman benzerliğinde sıkça kullanılmaktadır. Doküman içerisinde aranacak kelimelerin yoğunluğu bu algoritma için önemli değildir. Aranılan kelimenin ilgili dokümanda bulunup bulunmama durumuna göre yalnızca 0 veya 1 değerini üretmemektedir. Jaccard benzerliği Denklem 3.3' de gösterilmiştir.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (3.3)$$

d) Levenshtein Edit Distance

Bu kavram, iki metnin birbirine benzemesi için kaç birim hamle sonrasında eşit olacaklarının ölçüsüdür. Bu birim hamleler ekleme/silme/değiştirme işlemlerinden herhangi birisi olabilir. Örnek olarak kitap ve kitap kelimelerinin birbirine eşit hale getirmek için gerekli işlem sayısı 2' dir. Bu iki metnin Edit Distance' ı 2' dir. Çünkü kitap != kitap (iki metin eşit değil)

kitip != kitap (ilk metinde bulunan a harfi i harfi ile değiştiriliyor, işlem sayısı 1)

kitap = kitap (ilk metinde bulunan ikinci i harfi a harfi ile değiştiriliyor, işlem sayısı 2)

Yani toplamda 2 adet *değiştirme* işlemi sonucunda iki metin birbirinin aynı oldu.³¹

³¹ Dursun, B., (2008). Türkçe metinlerin benzerliğinin hesaplanması için yeni bir yöntem. *Yüksek Lisans Tezi*. İstanbul: Yıldız Teknik Üniversitesi FBE.

3.5 METİN MADENCİLİĞİ UYGULAMA ALANLARI

Metin madenciliği, metinlerin bulunduğu her platformda kullanılabilir.

Bu konuda Karaca (2012) şu alanları işaret etmiştir:³²

- I. Müşteri ilişkileri yönetimi
- II. Sahtekârlık tespiti
- III. Sağlık alanı
- IV. Pazar araştırmaları
- V. Metinlerden bilgi çıkarımı
- VI. Doküman özetleme
- VII. Doküman sınıflandırma
- VIII. Benzer içerikleri belirleme
- IX. Web içerikleri sınıflandırma
- X. Yazar tanıma sistemleri
- XI. Soru-cevap sistemleri

3.6 METİNLERİN HATALI DEĞERLENDİRİLME DURUMLARI

Metin madenciliğinde hangi yöntem kullanılırsa kullanılsın, metinlerin yapılandırılmasında yanlış dönüşüm işlemleri olabilir. İki kelime aynı iken farklı gibi görünüp yanlış sınıflara atanabilir veya diğer kelime ile benzerliği kontrol edilirken farklı kelime gibi görünebilir. Bu durum sonuçları önemli ölçüde etkileyebilmektedir. Metin madenciliğinde, dokümanların temizlenme işlemlerinden sonra dokümanlar mutlaka doğru şekilde yapılandırılmalıdır. Bu tür sorunları aşmak için çeşitli kavramsal/anlamsal veya kelimeleri köklerine göre ayrıştıran kütüphanelerden ve de doğal dil işleme yöntemlerinden faydalanılabilir.

³² Karaca, M. F., (2012). Metin madenciliği yöntemleri ile haber sitelerindeki köşe yazılarının sınıflandırılması. *Yüksek Lisans Tezi*. Karabük: Karabük Üniversitesi FBE.

a) Büyük Harf veya Küçük Harf Karşılaştırılması

Metin madenciliğinde, dokümanlar belirli sınıflara atama yapılırken veya benzerlikleri hesaplanırken iki kelimenin birbirine olan benzerliği, geliştirilen uygulamalarda (yazılım yapılan dile göre değişiklik gösterebilir) “=” operatörü ile karşılaştırılır ve kullanılacak algoritmaya göre sayısal bir değer alır. Bu karşılaştırma yapılırken operatör iki kelimenin de aynı olup olmadığına bakmaktadır. Örneğin; Veritabanı==Veritabanı eşitliği “true” sonuç döndürecektir. Fakat Veritabanı==veritabanı “false” sonuç döndürecektir. Çünkü hem bilgisayarlar hem de yazılım yapılan uygulamalar küçük/büyük harfe duyarlıdır. Bu sorunu aşmak için işlem yapılacak dokümanların içindeki verilerin tamamı küçük harfe çevrilir.

b) Eşsesli veya Eşanlamlı Kelimeler

Dokümanların yapılaşdırılmasında karşılaşılan bir diğer önemli sorun ise eşsesli ve eşanlamlı verilerdir. Aynı anlama gelen siyah ve kara veya farklı anlamlara gelen yaz kelimeleri dokümanların sınıflandırılması veya benzerliklerinin hesaplanmasında yanlış sonuçlara ulaşılmasına neden olabilir. Bu tür sorunları aşmak için doğal dil işleme yöntemlerinden, çeşitli dil için oluşturulmuş kütüphanelerden faydalanılarak kelimelerin aynı veya farklı olduğu tespit edilebilmektedir.

Metinlerin işlenmesi konusunda, Dursun (2008) şu durumları tespit etmiştir.³³

c) Kısaltma Kullanılmasından Kaynaklı Sorunlar

Metin eşleştirme çalışmalarında karşılaşılan en büyük sorunlardan birisi de birkaç kelimedenden oluşan bir kelime bloğunun baş harflerinden veya içerisindeki kelimelerden birkaçının kısaltılmış olmasıdır. Örneğin bir mahalle ismi olan Kazım Karabekir Mahallesi, K.karabekir Mahallesi veya K.karabekir Mah. gibi çok farklı şekillerde gösterilebilir. Bu gibi durumlar yanlış eşleştirme sonuçları üretebilmektedir.

d) Karakterlerin Klavyedeki Dizilişinden Kaynaklı Sorunlar

Bu durum bilgisayarda veri girişi yaparken hızlı yazı yazmak adına klavyede bir tuşa basacakken yanlışlıkla yanındaki diğer tuşa basılması şeklinde olmaktadır. Örneğin klavyede “Erzurum” kelimesi yazmak istenirken “z” harfi yerine yanlışlıkla klavyede “z” harfine yakın olan “x” harfine basılabilir. Bu da dokümanların işlenmesi sırasında yanlış sonuçlara yol açabilmektedir.

³³ Dursun, B., (2008). Türkçe metinlerin benzerliğinin hesaplanması için yeni bir yöntem. *Yüksek Lisans Tezi*. İstanbul: Yıldız Teknik Üniversitesi FBE.

4. YÖNTEM VE UYGULAMA

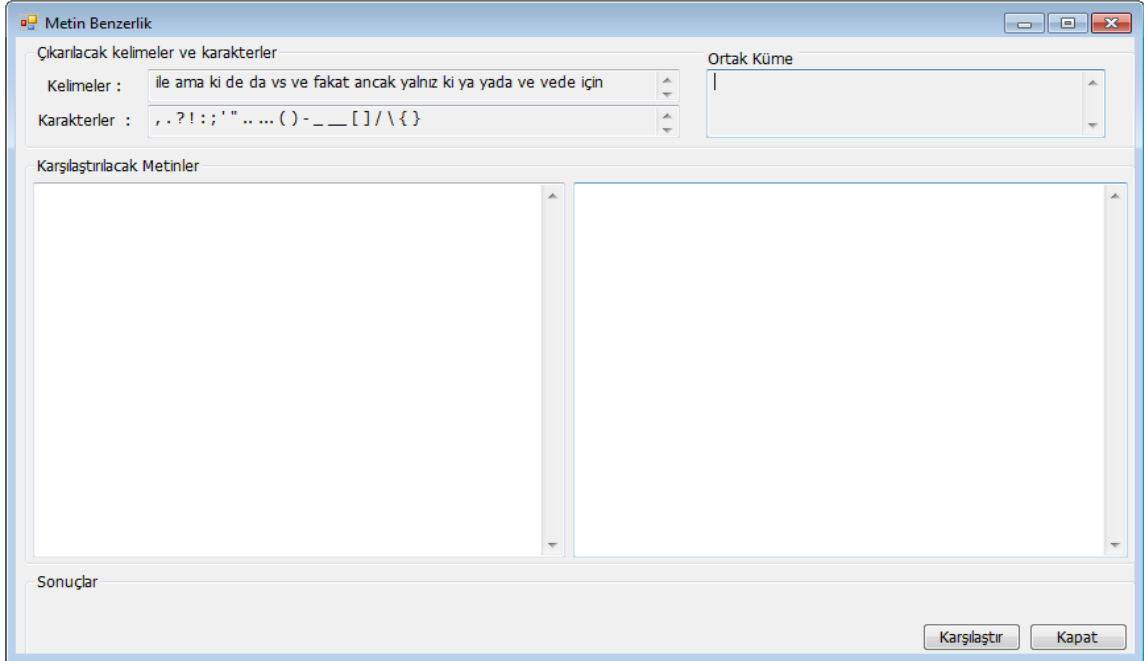
Bu bölümde ilk önce tez için geliştirilen uygulama tanıtılacak olup, daha sonra uygulamada kullanılan yöntemler ve adımlar anlatılmaktadır.

4.1 GELİŞTİRİLEN UYGULAMA

Bu tez kapsamında dokümanların birbirine olan benzerlikleri hesaplayabilmek için, .Net Platformunda C# programlama dili (4.0 versiyonu) ile bir masaüstü uygulama geliştirilmiştir. .Net, Microsoft firmasının yazılım geliştirmek için oluşturduğu bir uygulama geliştirme platformudur. Bu platform içerisinde Visual Basic, C++ ve C# gibi birçok programlama dili bulunmaktadır. Platformun içerisinde bulunan programlama dilleri ile masaüstü, web ve mobil uygulamalar geliştirilebilmektedir. .Net Platformu tüm Windows işletim sistemlerinde çalıştırılabilmektedir. Diğer platformlarda çalıştırılabilmesi için çalışmalar devam etmektedir.

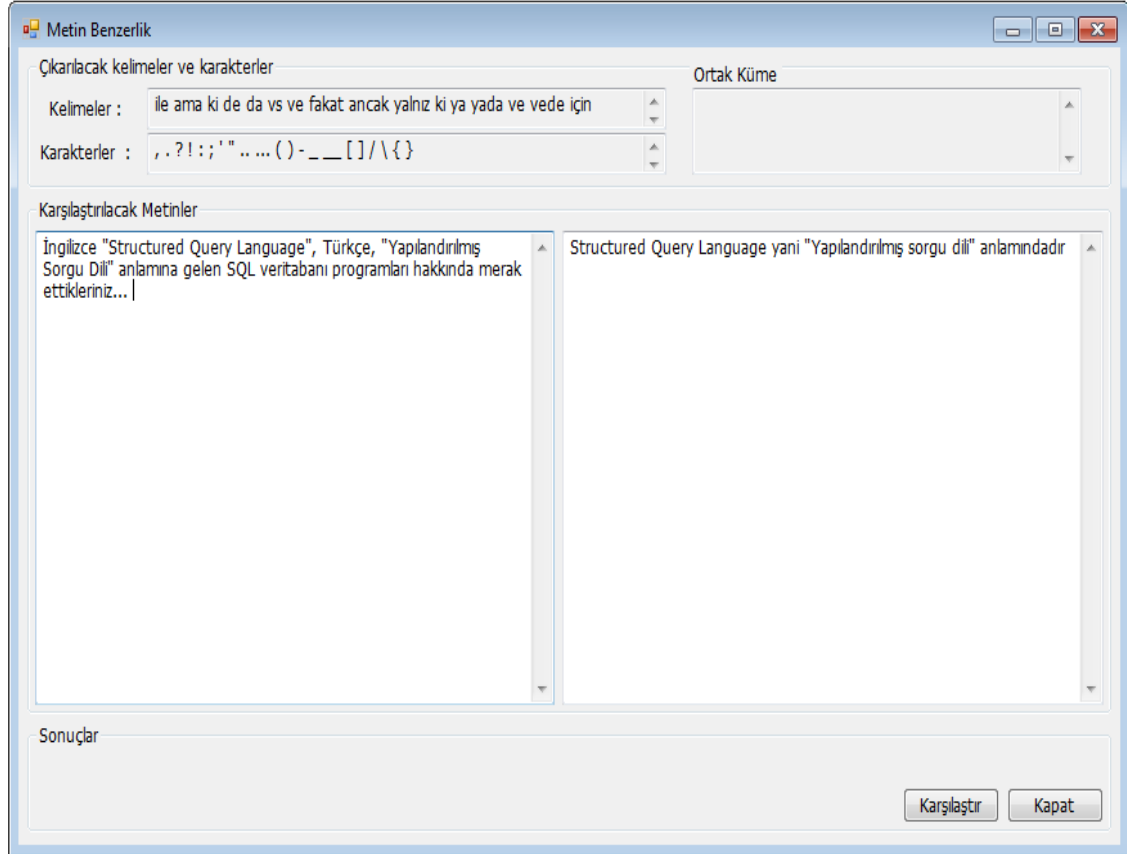
Uygulama iki aşamada geliştirilmiştir. İlk geliştirilen uygulamanın amacı verilerin benzerliklerin hesaplanmasını basit bir örnek üzerinden analiz etmektir. Birinci aşamada yalnızca iki cümlelerin birbirine olan benzerliklerini kosinüs (cosine) ve jaccard algoritmaları için ayrı ayrı hesaplayan bir uygulama gerçekleştirilmiştir. İlk geliştirilen uygulama Şekil 4.1' de gösterilmiştir.

Şekil 4.1: Geliştirilen uygulamanın 1. versiyonu



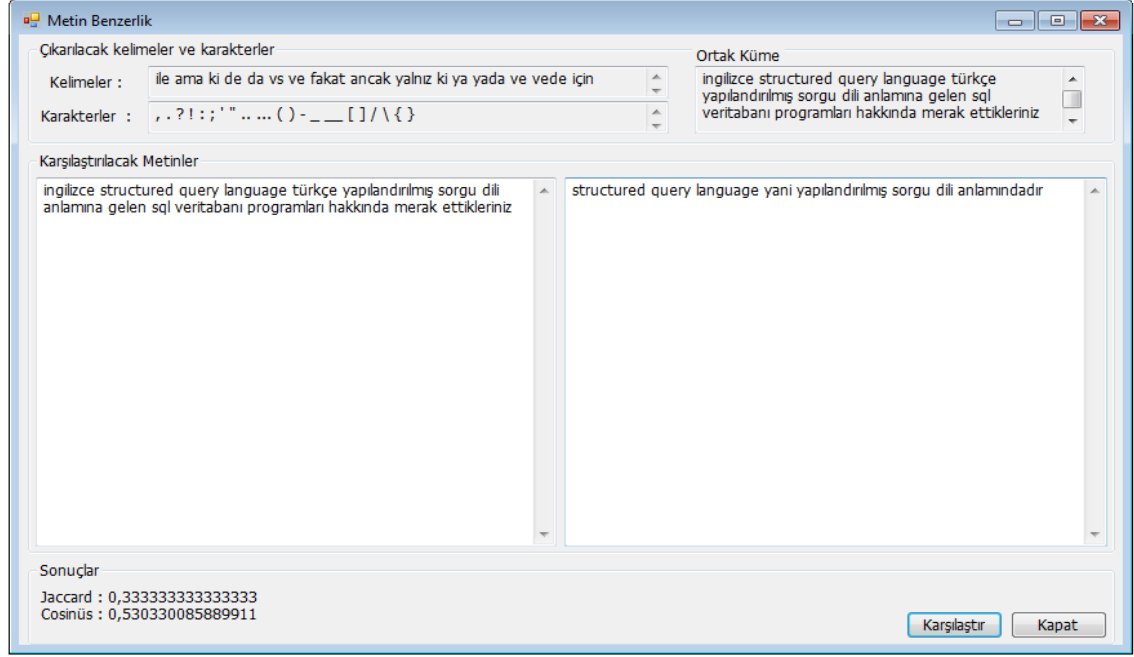
Uygulamada, karşılaştırılan cümlelerin ilk durumu ile gerekli temizleme işlemlerinden sonraki durumlarını görmek için cümlelerin yükleneceği iki adet metin kutusu bulunmaktadır. Dokümanlarda bulunan ve cümlelerin anlamlarını değiştirmeyen kelimelerin (edatlar ve bağlaçlar) ve karakterlerin bulunduğu iki adet textbox kutusu bulunmaktadır. Uygulama burada bulunan metinleri dokümanların içerisinden çıkarmakta ve hesaplamaların dışında tutmaktadır. Cümlelerin karşılaştırılması sırasında hesaplamalarda kullanılan bir küme oluşturulmaktadır. Yine bu küme, dokümanların işlenişinin görülmesi açısından ortak küme adında bir metin kutusunda gösterilmiştir. Son olarak dokümanların benzerlik sonuçlarının gösterildiği kosinüs ve jaccard isimli değişkenlerde gösterilmiştir. Uygulamanın verileri işlememiş örnek bir görüntüsü Şekil 4.2' de gösterilmiştir.

Şekil 4.2: 1. Versiyon verilerin işlenmemiş görüntüsü



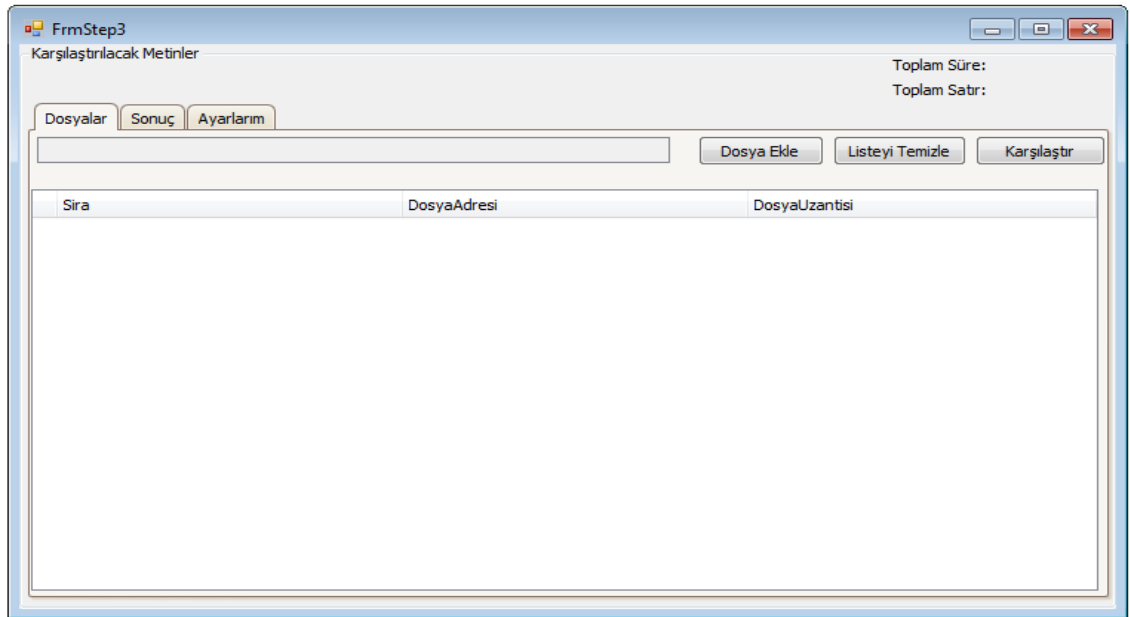
Uygulamanın verileri işlendikten sonraki görüntüsü Şekil 4.3' de gösterilmiştir.

Şekil 4.3: 1. Versiyon verilerin işlenmiş görüntüsü



Uygulamanın bir bütün olarak çalıştığı 2.versiyon ise Şekil 4.4' de gösterilmiştir. Bu aşamada geliştirilen uygulama, yalnızca iki cümlenin doğruluğunu hesaplamak için basit bir uygulama düzeyinde geliştirilmemiştir. Uygulama, metin madenciliği benzerlik uygulamalarında kullanılabilir düzeyde kodlanmıştır. Geliştirilen uygulama yalnızca iki cümle veya iki sayfanın karşılaştırılması için değil dinamik sayıdaki sayfanın birbirine olan benzerliğini hesaplayacak şekilde geliştirilmiştir. Uygulamanın ilk açıldığı andaki görüntüsü Şekil 4.4' de gösterilmiştir.

Şekil 4.4: Geliştirilen uygulamanın 2. Versiyonu



Uygulama üzerinde 3 tane sekme bulunan bir tablo bulunmaktadır. Bunlar; dosyalar, sonuç ve ayarların sekmeleridir. Dosyalar sekmesinde uygulamaya yüklenen dosyaların sıra numarasını, dosyaların yüklendiği adresleri ve dosyaların uzantılarının gösterildiği bir tablo bulunmaktadır. Uygulama “.txt” ve “.pdf” olan dosya türlerini destekleyecek şekilde geliştirilmiştir. Bu sekmede bulunan üç butonun görevleri şunlardır.

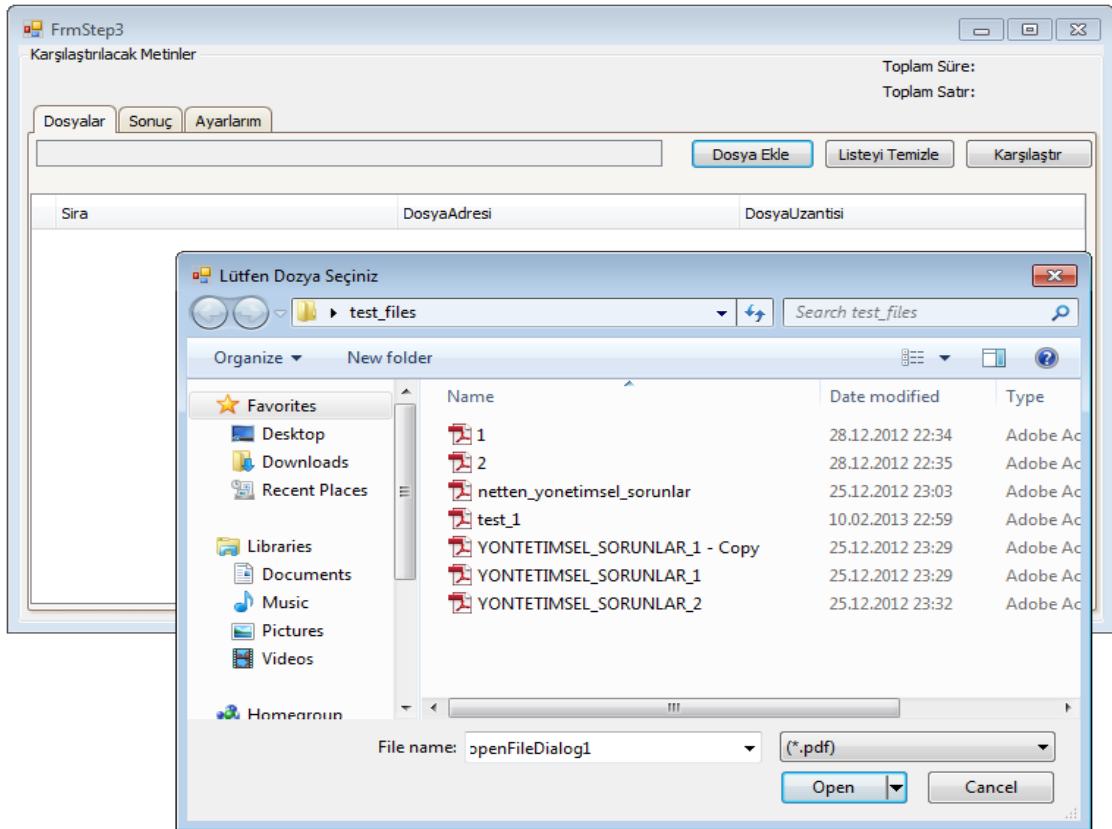
Dosya Ekle: Uygulamada kullanılacak olan dosyaların karşılaştırma yapılması için dosya tablosuna yüklenmesini sağlamaktadır.

Listeyi Temizle: Tabloda bulunan dosyaların silinmesini sağlamaktadır.

Karşılaştır: Gerekli temizleme ve hesaplama işlemlerinin yapılıp, sonuçların bulunduğu sonuç sekmesine gidilmesini sağlamaktadır.

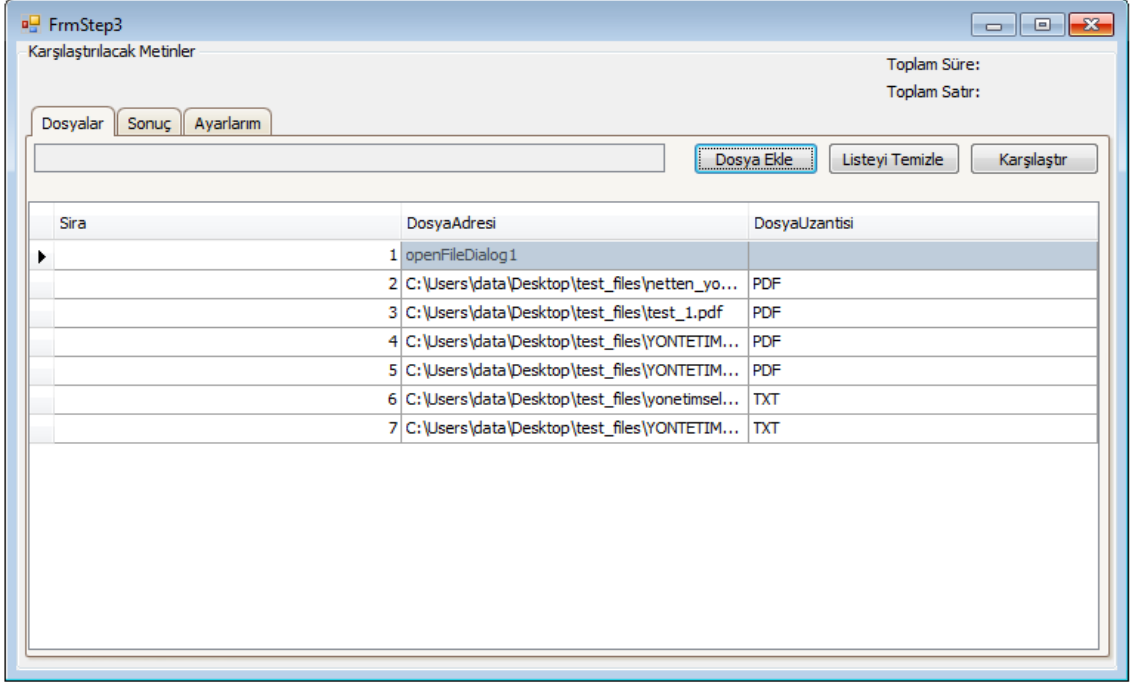
Dosyalar sekmesinde bulunan dosya ekle butonuna basıldığında istenen dosya (aynı anda birden fazla dosyada seçilebilmektedir) uygulamaya yüklenebilmektedir. Bu durum Şekil 4.5’ de gösterilmektedir.

Şekil 4.5: Dosya ekleme ekranı



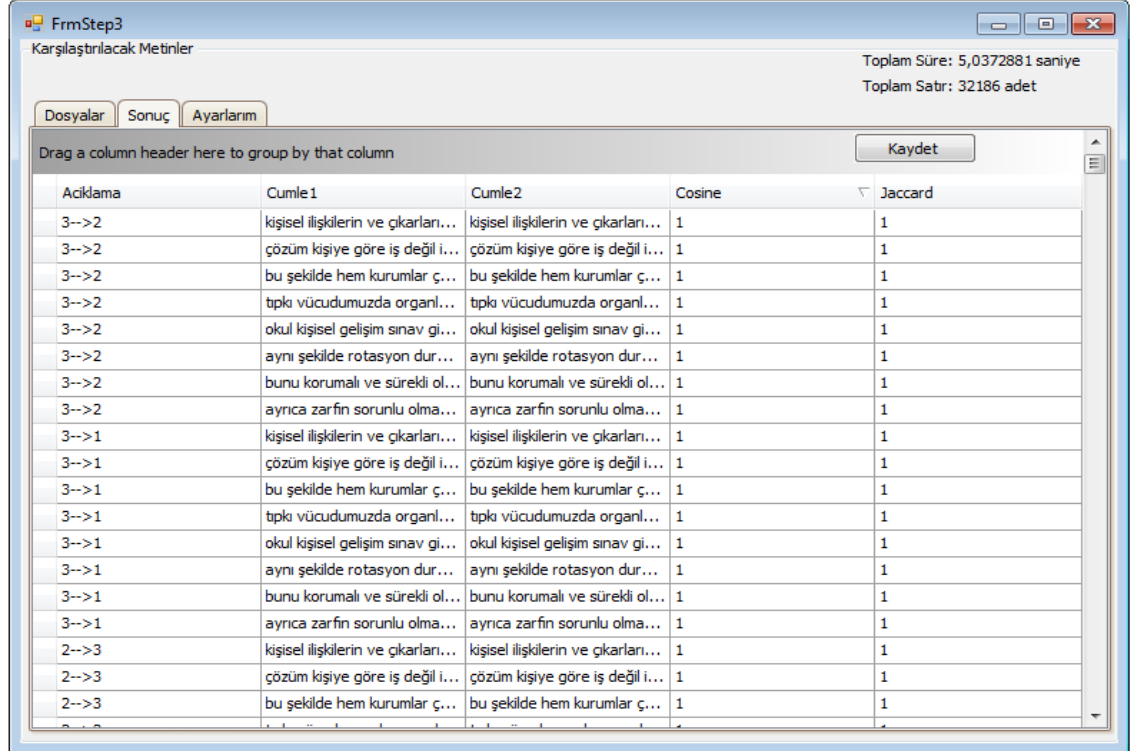
Dosyaların yüklendikten sonraki dosyalar sekmesindeki görüntüsü Şekil 4.6’ da gösterilmektedir.

Şekil 4.6: Dosyaların tablodaki görüntüsü



Karşılaştırılmak istenen dosyalar seçilip uygulamaya yüklendikten sonra karşılaştır butona basılır ve metinlerin birbirine olan benzerliklerinin hesaplandığı yöntem çalışır. Bütün temizleme ve dönüşüm işlemleri bittikten sonra tüm sonuçlar, sonuç sekmesinde gösterilmektedir. Bu durum Şekil 4.7' de gösterilmiştir.

Şekil 4.7: Sonuç sekmesi



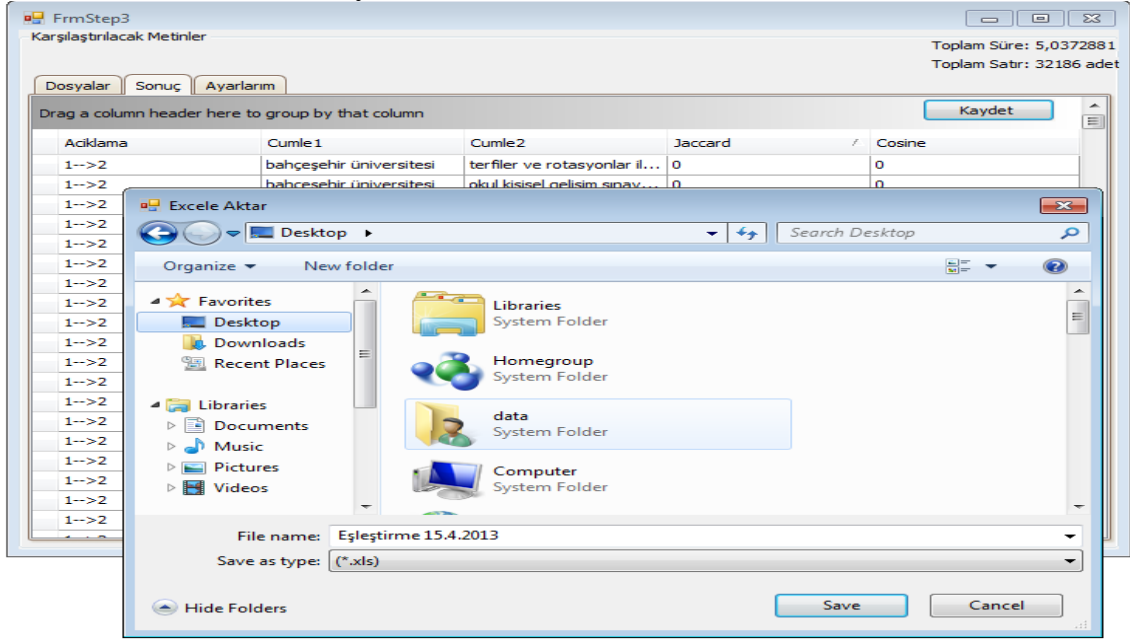
Sonuç sekmesinde bulunan kolonlarda; karşılaştırılan metinlerin tabloya ekleniş numaralarını, dokümanlar içerisinde karşılaştırma yapılan cümleleri, karşılaştırmalar sonucunda iki ayrı algoritma (kosinüs ve jaccard) için karşılaştırma sonuçları görülebilmektedir. Sonuç sekmesinde bulunan devexpress firmasına ait tablo (grid) sayesinde ulaşılan sonuçları cümle1, cümle2, kosinüs veya jaccard algoritmasına göre sıralayabiliriz. Bu tablo ile istendiği takdirde sonuçlar istenilen alana göre gruplandırılabilme imkanı sunmaktadır. Örneğin sonuçların jaccard algoritması için gruplandırılmış görüntüsü Şekil 4.8’ de gösterilmiştir.

Şekil 4.8: Sonuç gruplandırma (jaccard)

Acıklama	Cumle1	Cumle2	Cosine
Jaccard: 0			
Jaccard: 0,0188679245283019			
Jaccard: 0,02			
Jaccard: 0,0204081632653061			
Jaccard: 0,0212765957446809			
Jaccard: 0,0217391304347826			
1-->2	bir personel bir üst seviyeye yüks...	örnek çalışanların kariyer planlam...	0,101476513723761
1-->2	örnek çalışanların kariyer planlam...	bir personel bir üst seviyeye yüks...	0,101476513723761
1-->3	bir personel bir üst seviyeye yüks...	örnek çalışanların kariyer planlam...	0,101476513723761
1-->3	örnek çalışanların kariyer planlam...	bir personel bir üst seviyeye yüks...	0,101476513723761
2-->1	bir personel bir üst seviyeye yüks...	örnek çalışanların kariyer planlam...	0,101476513723761
2-->1	örnek çalışanların kariyer planlam...	bir personel bir üst seviyeye yüks...	0,101476513723761
2-->3	bir personel bir üst seviyeye yüks...	örnek çalışanların kariyer planlam...	0,101476513723761
2-->3	örnek çalışanların kariyer planlam...	bir personel bir üst seviyeye yüks...	0,101476513723761
3-->1	bir personel bir üst seviyeye yüks...	örnek çalışanların kariyer planlam...	0,101476513723761
3-->1	örnek çalışanların kariyer planlam...	bir personel bir üst seviyeye yüks...	0,101476513723761
3-->2	bir personel bir üst seviyeye yüks...	örnek çalışanların kariyer planlam...	0,101476513723761
3-->2	örnek çalışanların kariyer planlam...	bir personel bir üst seviyeye yüks...	0,101476513723761
Jaccard: 0,0222222222222222			

Sonuçlar kısmındaki verilerin kaydedilip, daha sonra çeşitli analizler yapılması için kaydet isimli bir buton bulunmaktadır. Kaydet isimli buton sonuçlar kısmındaki tabloyu Excel şeklinde kaydetme imkanı sunmaktadır. Excel’e kayıt görüntüsü Şekil 4.9’ da gösterilmiştir.

Şekil 4.9: Excel olarak kayıt



Sonuç tablosunun Excel olarak kaydedildiği görüntü Şekil 4.10' da gösterilmiştir.

Şekil 4.10: Excel tablosu

Eşleştirme 15.4.2013 [Uyu]					
Giriş Ekle Sayfa Düzeni Formüller Veri Gözden Geçir Görünüm Load Test Team					
Access'ten Web'den Metinden Diğer Varolan Dış Veri Al Kaynaklardan Bağlantılar					
Tümünü Yenile Bağlantıları Düzenle Bağlantılar Sırala Filtre Sırala ve Filtre Uyu					
A1 Acıklama					
	A	B	C	D	E
25845	3-->1	örneğin bir işletmede	tıpkı vücudumuzda organlar gibi	0,037037037	0,080064077
25846	3-->1	iki çalışanda yaptığı işten istediği mutluluğu	çözüm bölümlerin hepsinin birbirine ilişki	0,037037037	0,066226618
25847	3-->1	iki çalışanda yaptığı işten istediği mutluluğu	çözüm çalışanların eğitimine yeteneklerine	0,037037037	0,132453236
25848	3-->1	çözüm çalışanların eğitimine yeteneklerine	örnek çalıştığım kurumda insan kaynakları	0,037037037	0,066226618
25849	3-->1	sorunun türü organizasyon	insan kaynakları bölümü yalnızca yeni personel	0,037037037	0,142133811
25850	3-->1	sorunun türü organizasyon	çoğu zaman kurum için yararlı olabilecek işler ve	0,037037037	0,071066905
25851	3-->1	sorunun türü organizasyon	3sorun çalışanların yeteneklerine ve	0,037037037	0,071066905
25852	3-->1	sorunun türü organizasyon	örnek çalışanların niteliklerine ve	0,037037037	0,071066905
25853	3-->1	yöneticiler altındaki personellere yetkileri	işletmelerdeki işletme fonksiyonları ve yönetim	0,037037037	0,066666667
25854	3-->1	yöneticiler altındaki personellere yetkileri	bunu korumalı ve sürekli olarak geliştirmelidir	0,037037037	0,081649658
25855	3-->1	yöneticiler altındaki personellere yetkileri	terfiler ve rotasyonlar ilişkiler üzerinden	0,037037037	0,081649658
25856	3-->1	yöneticiler altındaki personellere yetkileri	sorunun türü kordinasyon ve planlama	0,037037037	0,081649658
25857	3-->1	çözüm yöneticilerin yetki dağılımının net bir şekilde	85 liralık zarf üretim bölümünde kullanılırken	0,037037037	0,071428571
25858	3-->1	çözüm yöneticilerin yetki dağılımının net bir şekilde	bu durumda şirketin çıkarları ikinci plana	0,037037037	0,071428571
25859	3-->1	çözüm yöneticilerin yetki dağılımının net bir şekilde	kalp diğer solumun bu organların doğru	0,037037037	0,071428571
25860	3-->2	satın alma bölümünün üretim bölümü için kefi	en az pazarlama bölümü kadar önemli bir vere	0,037037037	0,064549722

Uygulamada kullanıcının amacına ve ihtiyacına göre ikiden fazla sayıda dosya işleme alınabilir. Uygulama bir dosyada bulunan bir cümleyi diğer bütün dosya içerisindeki cümleler ile karşılaştırmaktadır. Buda çok büyük bir işlem adedi ve işlem süresi yaratmaktadır. Bu nedenle uygulama için harcanan süreyi ve uygulamada benzerlik hesaplaması yapılan cümleleri gösterebilmek için uygulamada toplam süre ve toplam satır isimli iki adet label kullanılmıştır. Bu durum Şekil 4.11’ de gösterilmiştir.

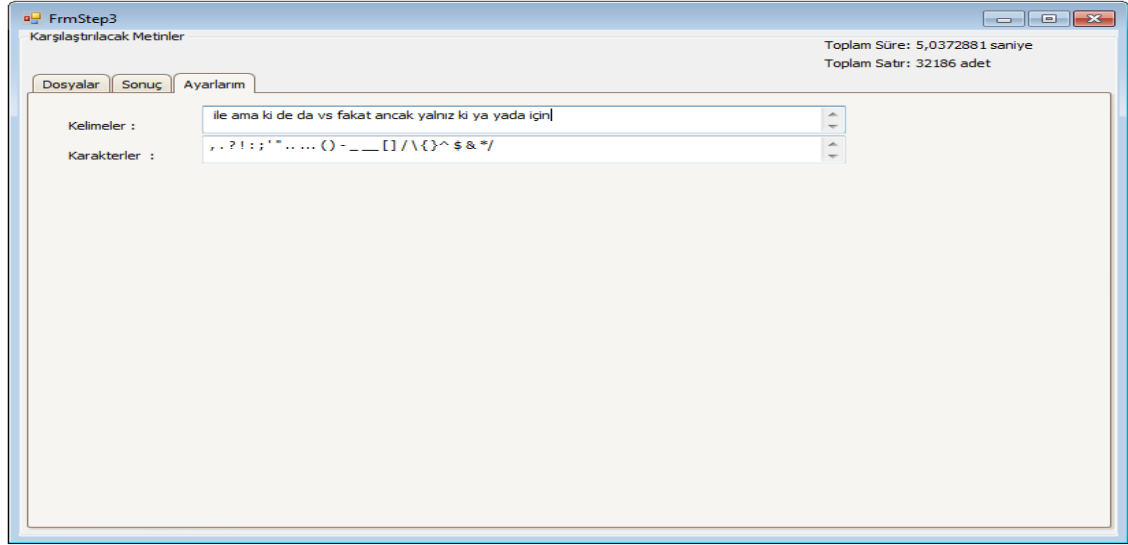
Şekil 4.11: Toplam satır ve toplam süre

Acıklama	Cumle1	Cumle2	Jaccard	Cosine
3-->2	2 sorunun olduğu bölüm arge	aynı şekilde rotasyon duru...	0,037037037037037	0,0710669054518702
3-->2	2 sorunun olduğu bölüm arge	örnek işletmelerde en büyü...	0,037037037037037	0,0710669054518702
3-->2	şuan bu durumun çok çok u...	yöneticiler altındaki person...	0,037037037037037	0,0608580619450184
3-->2	arge bölümü olarak yenilik	satın alma bölümü örneğin ...	0,037037037037037	0,0605227532668802
3-->2	en az pazarlama bölümü ka...	satın alma bölümünün üreti...	0,037037037037037	0,0645497224367903
3-->2	en az pazarlama bölümü ka...	örnek çalıştığım kurumda ins...	0,037037037037037	0,0721687836487032
3-->2	en az pazarlama bölümü ka...	örneğin bir işletmede makin...	0,037037037037037	0,0721687836487032
3-->2	en az pazarlama bölümü ka...	çözüm yöneticilerin geri gid...	0,037037037037037	0,0721687836487032
3-->2	personel atamalarında ve g...	çalışanların daha üst düzey ...	0,037037037037037	0,0710669054518702
3-->2	örneğin 2005 yılında çalışm...	satın alma bölümü örneğin ...	0,037037037037037	0,0605227532668802
3-->2	personellerin	bu şekilde hem kurumlar çalı...	0,037037037037037	0,0690065559342354
3-->2	bu hem	yöneticiler altındaki person...	0,037037037037037	0,0533760512683624
3-->2	yönetim fonksiyonları sor...	çözüm bölümlerin hepsinin b...	0,037037037037037	0,0662266178532522
3-->2	yönetim fonksiyonları sor...	çözüm çalışanların eğitime...	0,037037037037037	0,132453235706504
3-->2	bu durumda şirketin çıkarlar...	kalp ciğer solunum bu orga...	0,037037037037037	0,0714285714285714
3-->2	çoğu zaman kurum için yara...	üst yönetimde bulunan bir ...	0,037037037037037	0,12104550653376
3-->2	sorunun türü kontrol ve per...	terfi almalarında okul ingilizc...	0,037037037037037	0,0645497224367903
3-->2	örnek sistemdeki her parça...	örnek kurumumuzda satın a...	0,037037037037037	0,0714285714285714
3-->2	örnek sistemdeki her parça...	örnek büyük şirketlerde öze...	0,037037037037037	0,0648203723552164

Son sekme olan ayarlarım kısmında ise uygulamada kullanılan filtrelemeler yer almaktadır. Metin madenciliğinde verilerin işlenmesi için ilk önce gereksiz kelime gruplarından ve karakterlerden temizlenmesi gerekmektedir. Bu nedenle, sonuçlar kısmında bulunan iki textbox kutusu ile bu temizleme işlemleri gerçekleştirilmektedir. Çıkarılacak kelime grupları uygulamada işlenecek dile göre değişebilmektedir. Örneğin Türkçede dokümanlar için “fakat” kelimesi dokümanlardan çıkarılması gerekirken İngilizce kelimelerden “the” kelimesi dokümanlardan çıkarabilmektedir. Bu nedenle bu filtreleme kısmı kullanıcının tercihine bırakılmıştır.

Bu durum Şekil 4.12’ de gösterilmiştir.

Şekil 4.12: Filtrelenen kelimeler ve karakterler



4.3 VERİLERİN TEMİZLENMESİ

Metin madenciliğinde dokümanların benzerliklerinin hesaplanmasında önemli konulardan birisi metinlerin metin madenciliği için kullanılan algoritmalara uygun hale getirilmesidir. Bunun için ilk önce Türkçede bulunan edat/bağlaçların dokümanlardan çıkarılması gerekmektedir. Bu kelimeler, cümle içerisinde kullanıldığında cümlenin anlamını değiştirmeyen fakat diğer cümlede olmaması durumunda cümleler karşılaştırıldığında sonuçları değiştirebilecek kelimelerdir. Aynı zamanda kelimelerin benzerliği ile ilgili hiçbir ilgisi bulunmayan “. , () %” gibi işaretlerde dokümanların içerisinden çıkarılmalıdır. Bunun için geliştirilen uygulamada, dokümanlarda filtreleme yapacak iki adet textbox kutusu bulunmaktadır. Bu kutular içerisinde bulunan karakterler ve kelimeler uygulama hesaplama işlemine başlamadan dokümanlardan çıkarılmaktadır. Bu durum şu kodlarla yapılmaktadır.

```
public string MetniTemizle(string v1In,string v1KarakterListesi,string v1CikacakMetinListesi)
{
    string[] strKarakterList = v1KarakterListesi.Split(' ');

    v1In = v1In.ToLower().Trim().Replace("\r\n", " # ");
    foreach (string stri in strKarakterList)
    {
        v1In = v1In.Replace(stri, "");
    }

    //v1In.Replace("")
    List<string> strMetin1 = v1In.ToLower().Split(' ').ToList();

    string[] strList = v1CikacakMetinListesi.Split(' ');

    foreach (string strA in strList)
```

```

{
    for (int iM = 0; iM < strMetin1.Count; iM++)
    {
        string vlgg=strMetin1[iM].Trim();

        if (vlgg == "")
        {
            strMetin1.RemoveAt(iM);
            iM--;
            continue;
        }

        if ((iM < strMetin1.Count) && (vlgg == strA.Trim()))
        {
            strMetin1.RemoveAt(iM);
            iM--;
        }
    }
}

StringBuilder str = new StringBuilder();
foreach (string ii in strMetin1)
{
    str.Append(ii + " ");
}

return str.ToString();
}

```

MetinTemizle isimli metod ile metoda verilen metinler (cümleler) küçük harfe çevrilip, temizleniyor ve bölünerek dizi haline getiriliyor. Daha sonra eğer filtrelene kelime veya karakterler bu cümleler içerisinde bulunuyor ise bu kelimeler siliniyor veya boşluklarla yer değiştiriliyor.

Örneğin uygulamaya işlenmesi için yüklenen şu cümle;

1-Veritabanı Yönetim Sistemleri, çok kullanışlıdır fakat doğru şekilde yönetilmeleri çok önemlidir.

Uygulama cümleyi işledikten sonra şu şekilde dönüşmektedir;

veritabanı yönetim sistemleri çok kullanışlıdır doğru şekilde yönetilmeleri çok önemlidir

Cümle küçük harfe çevrilmiş, cümle içerisinde bulunan gereksiz kelimeler ve karakterler silinmiştir.

4.4 KELİMELERİN DÖNÜŞÜMÜ

Doküman içerisinde bulunan verilerin temizleme işlemleri bittikten sonra cümleler doküman içerisinde tespit edilmelidir.

```

public string SatirlariIsaretle(string strMetin)
{
    string[] karakter = { "..", "...", ".", "?", "!" };

    foreach (string oldVal in karakter)
    {
        strMetin = strMetin.Replace(oldVal, "#");
    }

    return strMetin;
}

```

Cümleler “. , ? !” gibi işaretler aratarak işaretlenir ve doküman içerisinde parçalara ayrılır. Daha sonra dokümanlar içerisinde yer alan tüm veriler metin madenciliği algoritmalarının kullanımına hazır hale gelmesi için sayısal formata çevrilmiştir. Dokümanlar içerisinde karşılaştırma yapılacak tüm cümleler, ilk önce birer dizi (küme) haline getirilir. Daha sonra iki dizi tekrarsız üyeler olarak (eşsiz üyeler) birleştirilerek ortak küme dizisi oluşturulur.

```

public List<string> OrtakKumeyiOlustur(string v1OrtakKume, string
v1Metin1, string v1Metin2)
{
    List<string> strOut = new List<string>();

    string[] str1 = v1Metin1.Split(' ');
    foreach (string v11 in str1)
    {
        if (strOut.Find(p => p == v11) == null)
        {
            strOut.Add(v11);
        }
    }

    string[] str2 = v1Metin2.Split(' ');
    foreach (string v12 in str2)
    {
        if (strOut.Find(p => p == v12) == null)
        {
            strOut.Add(v12);
        }
    }

    v1OrtakKume = "";
    foreach (string cc in strOut)
    {
        v1OrtakKume += cc + " ";
    }

    return strOut;
}

```

Oluşturulan ortak kümede bulunan tüm dizi elemanları benzerliği hesaplanacak dizi elemanları ile karşılaştırılarak binary (yalnızca 0 ve 1) veya integer (0,1 veya cümle

içerisinde bulunma adedi) değerlere çevrilirler. Örneğin jaccard benzerlik algoritması için şu şekilde hesaplanma yapılmıştır:

```
public string JaccardIleHesapla(List<string> OrtakKume, string[] Kume_1,
string[] Kume_2)
{
    string[] m1 = JaccardKarsilastir(OrtakKume, Kume_1).Split(',');
    string[] m2 = JaccardKarsilastir(OrtakKume, Kume_2).Split(',');

    int eslesenSayisi = 0;
    for (int i = 0; i < m1.Length; i++)
    {
        if (m1[i] == m2[i]) { eslesenSayisi += 1; }
    }

    double kume_1_EslesenElemanSayisi = m1.Where(m => m == "1").Count();
    double kume_2_EslesenElemanSayisi = m2.Where(m => m == "1").Count();

    double yanlizKume1 = kume_1_EslesenElemanSayisi - eslesenSayisi;
    double yanlizKume2 = kume_2_EslesenElemanSayisi - eslesenSayisi;

    return (eslesenSayisi / (eslesenSayisi + yanlizKume1 +
yanlizKume2)).ToString();
}
```

4.5 CÜMLELERİN KARŞILAŞTIRILMASI VE UYGULAMANIN KARMAŞIKLIĞI

Doküman içerisinde bulunan tüm verilerin temizlenme, cümlelere bölünme ve hesaplanma işlemleri bittikten sonra tüm veriler uygulama içerisinde bir tabloda gösterilmektedir. Uygulamaya çok sayıda dosya yüklendiği ve her dosyada birden çok cümle olduğu için uygulamanın bunları işlemesi ve dönüşüm işlemlerinden geçirmesi oldukça yüksek bir zaman almaktadır. Uygulama içerisinde bir dosyada bulunan her cümle diğer dosyalar içerisinde ki tüm cümleler ile karşılaştırılmakta ve cümlelerin birbirine olan benzerliği hesaplanmaktadır. Bu durum oldukça karmaşık bir durum yaratmaktadır. Bu işlemleri gerçekleştirmek için uygulama içerisinde iç içe dört adet döngü kurulmuştur.³⁴

³⁴ İç içe döngülerin formatı word belgesine tam sığmadığı için sola yaslı olarak gösterilmiştir.

```

foreach (dsnDosya v1AnaMetin in DosyaList)
{
foreach (dsnDosya v1AltMetin in DosyaList.Where(p => p.Sira != v1AnaMetin.Sira))
{
string AnaMetinFull = v1AnaMetin.DosyaObj;
string AltMetinFull = v1AltMetin.DosyaObj;

foreach (string AnaMetinSatir in AnaMetinFull.Split('#'))
{
foreach (string AltMetinSatir in AltMetinFull.Split('#'))
{
string iStr1 = fnUtil.Metin_BoslukAl(AnaMetinSatir);
string iStr2 = fnUtil.Metin_BoslukAl(AltMetinSatir);

if (iStr2 != "" && iStr1 != "" && iStr1.Count()>1 && iStr2.Count()>1)
{
List<string> strOrtakKume = fnUtil.OrtakKumeyiOlustur(ortakKume, iStr1, iStr2);

string sonucCos = fnUtil.CosineIleHesapla(strOrtakKume, iStr1.Split(' '), iStr2.Split('
'));
string sonucJacc = fnUtil.JaccardIleHesapla(strOrtakKume, iStr1.Split(' '), iStr2.Split('
'));

OranResult rc = new OranResult();
rc.Aciklama = v1AnaMetin.Sira.ToString() + "-->" + v1AltMetin.Sira.ToString();
rc.Cumle1 = iStr1;
rc.Cumle2 = iStr2;
rc.Cosine = sonucCos;
rc.Jaccard = sonucJacc;
recList.Add(rc);
sayac++;
}
}
}
}
}

```

5. SONUÇ VE ÖNERİLER

Tez konusu olan metin madenciliği ile dokümanlar arasında benzerliklerin bulunması konusunda bir uygulama gerçekleştirilmiştir. Uygulama tez için geliştirilen Metin Benzerlik 5.6 uygulaması ile yapılmıştır. Yazar Namık KEMAL' in hayat hikayesini anlatan 10 adet doküman (10 adet dokümanın tamamı ekler kısmında verilmiştir), .txt ve .pdf formatlarında kaydedilmiştir. Daha sonra bu hayat hikayeleri arasındaki benzerlikler (kosinüs ve jaccard algoritmaları ile) bulunmaya çalışılmış ve her bir algoritma için varılan sonuçlar irdelenmiştir. Uygulamadan elde edilen sonuçlar şunlardır:

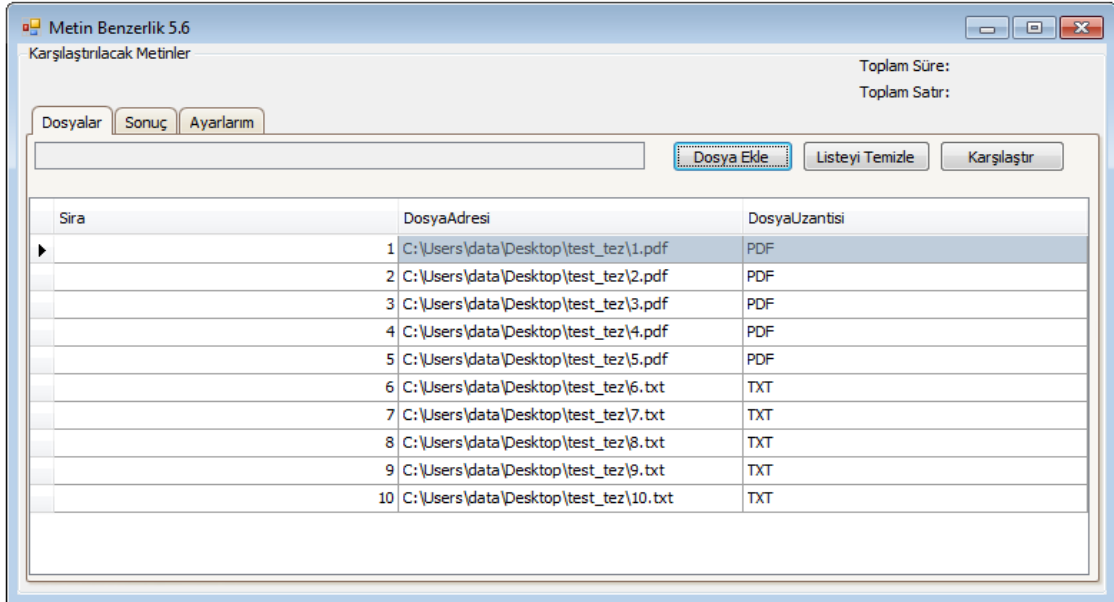
Dokümanların uygulamaya hazırlanması sırasında bazı kelimeler (edatlar ve bağlaçlar) ve karakterler uygulama tarafından dokümanlardan temizlenmiştir (bu kısıtlar kullanıcının isteği ve ihtiyacı doğrultusunda uygulama anında değiştirilebilmektedir).

Kelimeler: ile ama ki de da vs fakat ancak yalnız ki ya yada ve vede çünkü ne

Karakterler: , . ? ! : ; ' " () - _ __ [] / \ { } % & * -

Uygulamaya eklenen dosyalar Şekil 5.1' de gösterilmiştir.

Şekil 5.1: Uygulamaya yüklenen dosyalar



The screenshot shows the 'Metin Benzerlik 5.6' application window. The title bar reads 'Metin Benzerlik 5.6'. The main window has a tabbed interface with 'Dosyalar' (Files) selected. Below the tabs, there are buttons for 'Dosya Ekle' (Add File), 'Listeyi Temizle' (Clean List), and 'Karşılaştır' (Compare). The main area contains a table with the following data:

Sıra	DosyaAdresi	DosyaUzantisi
1	C:\Users\data\Desktop\test_tez\1.pdf	PDF
2	C:\Users\data\Desktop\test_tez\2.pdf	PDF
3	C:\Users\data\Desktop\test_tez\3.pdf	PDF
4	C:\Users\data\Desktop\test_tez\4.pdf	PDF
5	C:\Users\data\Desktop\test_tez\5.pdf	PDF
6	C:\Users\data\Desktop\test_tez\6.txt	TXT
7	C:\Users\data\Desktop\test_tez\7.txt	TXT
8	C:\Users\data\Desktop\test_tez\8.txt	TXT
9	C:\Users\data\Desktop\test_tez\9.txt	TXT
10	C:\Users\data\Desktop\test_tez\10.txt	TXT

Uygulama sonucunda 10 adet doküman için uygulama 53,1560404 saniye çalışmış olup, 445.198 adet cümleyi diğer cümleler ile karşılaştırmıştır. Sonuçların çıktısı görüntüsü Şekil 5.2' de gösterilmiştir.

Şekil 5.2: Uygulamanın sonuç görüntüsü

Acıklama	Cumle1	Cumle2	Cosine	Jaccard
10-->6	yaklaşık üç yıl avrupada k...	burada arapça farsça öğr...	0	0
10-->6	yaklaşık üç yıl avrupada k...	bir süre sonra kaymakamli...	0,0924500327042049	0,0476190476190476
10-->6	yaklaşık üç yıl avrupada k...	sofya'da geçirdiği iki yıl na...	0,0808452083454443	0,04
10-->6	yaklaşık üç yıl avrupada k...	1857'de istanbul'a gelen n...	0,0563436169819011	0,027027027027027
10-->6	yaklaşık üç yıl avrupada k...	öte yandan dönemin düşü...	0	0
10-->6	yaklaşık üç yıl avrupada k...	şinasi yakınlık kurması ona...	0	0
10-->6	yaklaşık üç yıl avrupada k...	adı dönemin reformcu olar...	0	0
10-->6	yaklaşık üç yıl avrupada k...	'yeni osmanlılar' adlı gizli ö...	0	0
10-->6	yaklaşık üç yıl avrupada k...	tasfiri efkar ali suavi'nin g...	0	0
10-->6	yaklaşık üç yıl avrupada k...	çok geçmeden gazete kap...	0	0
10-->6	yaklaşık üç yıl avrupada k...	namık kemal erzurum vali ...	0	0
10-->6	yaklaşık üç yıl avrupada k...	bu evrede padişah abdüla...	0	0
10-->6	yaklaşık üç yıl avrupada k...	bu çağrını üzerine namık ke...	0	0

Uygulama sonucunda dokümanlar arasındaki benzerliklerin incelenmesi için benzerlikler işlenip matris şeklindeki tablolarda gösterilmiştir. Tablo 5.1, Tablo 5.2 ve Tablo 5.3' de Kosinüs ve jaccard algoritmalarının dokümanlar içerisinde bulmuş olduğu 1, 0 ve 1 ile 0 arasındaki değerlerin toplam sayıları verilmiştir.

Tablo 5.1: "0" Değerler

KOSİNÜS VE JACCARD İÇİN BULUNAN "0" DEĞERLER											TOPLAM
	1	2	3	4	5	6	7	8	9	10	
1	0	2799	5859	2234	1084	1964	5762	6331	2587	1065	29685
2	2799	0	7181	2802	1293	2432	7087	7815	3155	1310	35874
3	5859	7181	0	5863	2816	4949	14470	16362	6737	2749	66986
4	2234	2802	5863	0	1086	1964	5771	6336	2588	1065	29709
5	1084	1293	2816	1086	0	910	2631	3019	1260	510	14609
6	1964	2432	4949	1964	910	0	4695	5449	2343	909	25615
7	5762	7087	14470	5771	2631	4695	0	15668	6651	2674	65409
8	6331	7815	16362	6336	3019	5449	15668	0	7185	3008	71173
9	2587	3155	6737	2588	1260	2343	6651	7185	0	1226	33732
10	1065	1310	2749	1065	510	909	2674	3008	1226	0	14516

Tablo 5.2: "1" Değerler

KOSİNÜS VE JACCARD İÇİN BULUNAN "1" DEĞERLER											TOPLAM
	1	2	3	4	5	6	7	8	9	10	
1	1		1	48	1		1	8			59
2		1									0
3	1		1	1				1			3
4	48		1	1	1		1	8			59
5	1			1	1			1			3
6						1					0
7	1			1			1	6			8
8	8		1	8	1		6	1	1		25
9								1	1		1
10										1	0

Tablo 5.3: "0 ile 1 arası" Değerler

KOSİNÜS VE JACCARD İÇİN BULUNAN "0 - 1 ARASINDA" DEĞERLER											TOPLAM
	1	2	3	4	5	6	7	8	9	10	
1	-	301	740	218	165	336	987	661	113	135	3656
2	301	-	1003	298	257	420	1283	865	193	178	4798
3	740	1003	-	736	484	1123	3350	2117	391	419	10363
4	218	298	736	-	163	336	978	656	112	135	3632
5	165	257	484	163	-	240	744	480	90	90	2713
6	336	420	1123	336	240	-	1515	991	141	195	5297
7	987	1283	3350	978	744	1515	-	3226	639	566	13288
8	661	865	2117	656	480	991	3226	-	374	352	9722
9	113	193	391	112	90	141	639	374	-	70	2123
10	135	178	419	135	90	195	566	352	70	-	2140

Uygulama sonucundaki benzerlik sonuçları şunlardır:

Diğer sayfalar ile benzerliğin (1.00 benzerlik) en fazla olduğu sayfalar 59 adet benzerlik ile 1 ve 4 numaralı sayfalar, ikinci olarak 25 adet benzerlik ile 8 numaralı sayfadır. 2,6 ve 10 numaralı sayfaların bire bir benzerlik taşıdığı hiçbir sayfa bulunmamaktadır (Her sayfanın kendisi ile olan bire-bir benzerliği tabloda gösterilmiştir fakat bu toplam sonuca yansıtılmamıştır). Birbirlerine en çok benzeyen sayfalar, 48 adet benzerlikle 1 ve 4 numaralı sayfalardır.

0 ile 1 Değerleri arasındaki en çok benzerlik 13.288 adet benzerlik ile 7 numaralı sayfadır. En az benzerlik taşıyan sayfa ise 2123 adet benzerlik ile 9 numaralı sayfadır.

Birbirleri arasında en fazla benzerliğin olduğu sayfalar ise 3350 adet benzerlik ile 3 ve 7 numaralı sayfalardır.

Uygulama sonucunda kosinüs ve jaccard algoritmalarının her ikisi de aynı miktarda “0” değeri ve aynı miktarda “1” değeri bulmuştur. Bunun nedeni hiçbir ortak (kesişen kelimenin) benzerlik olmadığı durumda iki algoritmada “0” değerini üretmektedir. Jaccard algoritması karşılaştırma yaptığı her iki cümlede de aynı kelimelerin olması durumunda “1” değerini üretmektedir. Kosinüs algoritması ise kelimenin geçme sıklığına uygun bir değer üretmektedir. Yani bir kelimenin cümlede 1 kez geçmesi diğer cümlede ise daha fazla geçmesi sonucu etkilemektedir. Kosinüs algoritmasının “1” sonucu ürettiği her değer için jaccard algoritması da “1” değerini üretebilir. “0” ve “1” arasındaki değerler için ise farklı sonuçlar ortaya çıkmaktadır.

Yapılan uygulama sonucunda cümlelerin birbiri ile karşılaştırılmasında jaccard algoritmasının daha etkili olabileceği söylenebilir. Bunun nedeni jaccard algoritması, kelimenin var olup olmama durumuna göre değer vermektedir. Cümlelerin yapı gereği genelde çok uzun olmaması sebebi ile kelimelerin birden fazla geçme ihtimali daha düşüktür. Yapılan uygulama sayfaların birbiri arasındaki genel benzerliğini bulmak olduğunda kosinüs algoritmasının daha etkili olabileceği düşünülebilir. Çünkü iki ayrı sayfadaki, anlatılan konularda geçen kelimelerin yoğunluğunu ölçmek çok önemlidir. Kosinüs algoritması bu konuda daha etkili sonuçlar verebilecektir.

Bu tez uygulamasında amaç kosinüs ve jaccard algoritmalarının başarılarını uygulama üzerinde gözlemlemektir. Geliştirilen uygulama genel olarak dokümanların birbirleri arasındaki benzerlikleri bulmakla beraber, uygulama esnasında geliştirilen yöntemle ilgili eksiklikler ve kullanılan dilin yapısı ile ilgili çeşitli sorunlarla karşılaşmıştır. Bu sorunlar giderildiğinde güvenilirliği daha yüksek sonuçlar alınabilecektir. Uygulama ile ilgili şu düzenlemeler yapılabilir:

Örneğin; **ilk romanı “intibah” 1876’da yayınladı** cümlesindeki ‘ ve “ ” gibi işaretlerin normalde uygulama tarafından silinmesi gerekmektedir. Fakat uygulama bu karakterleri algılayamamıştır. Bu tür kısıtlar uygulamanın hem çok sağlıklı çalışmamasına hem de gereksiz eşleştirme işlemleri yapmasına neden olabilmektedir. Bu tür sorunlar, kullanılan dosya türlerinin farklı yapıda karakter, link, resim vb... tutmasından kaynaklanmaktadır. Uygulama geliştirildikten sonra kodların her dosya türü için ayrı ayrı sonuçlar verebileceği gözlemlenmiştir. Bu sorunu aşmak için uygulamada

kullanılan kodların her dosya türü için ayrı ayrı uygulanıp test edilmesi gerekmektedir. Başka bir sorun ise kısaltmalardan kaynaklıdır. Örneğin “Dr. Ali Bey” isimli cümle işleme alındığında sanki iki ayrı cümleymiş gibi işleme alınmaktadır. Çünkü geliştirilen uygulama cümleleri, yapısı gereği cümlelerin bitme işaretleri olan “. , ? !” gibi işaretlerden bölmektedir. Uygulama burada bulunan “.” İşaretini cümle sonu gibi algılamaktadır. Aynı şekilde cümle içersinde geçebilecek olan “mahalle” kelimesi “mah.” olarak geçtiğinde farklı kelimelermiş gibi algılanabilecektir. Bu durum doğal dil işleme yöntemlerinin uygulama alanına girmektedir.

Geliştirilen uygulama genel olarak tez, makale ve ödev tarzı uygulamaların karşılaştırılması için tasarlanmıştır. Uygulama internet sitelerinin benzerliğini hesaplayacak şekilde geliştirilebilir. Sayfaların skaler çarpımını yapacak şekilde algoritmalar kullanılabilir. Bunun için uygulamanın temizleme işlemleri tekrar gözden geçirilmelidir. Çünkü internet sitelerini içerisinde yüksek miktarda resim, link, formül, flash uygulamalar vb... işleme alınmayacak yapılar bulunmaktadır. Ayrıca internet sitelerinin kaydedildikten sonra içerisinden her türlü html, xml ve css türü işaretleyicilerin temizlenmesi gerekmektedir. Geliştirilen uygulama verileri uygulama içerisinde dizilerde tutup daha sonra Excele aktarmaktadır. Kullanılan dokümanların sayısı ve dokümanlar içerisindeki cümlelerin artmasıyla kimi zaman 500.000 satırı aşan sonuçlar ortaya çıkmaktadır. Veriler sql veritabanına (oracle, mssql, mysql vb...) kaydedilerek hem daha hızlı işlem yapılabilir hem de sql sorgularıyla daha sağlıklı ve karmaşık sorgulamalar yapılabilir.

KAYNAKÇA

Kitaplar

- Gözüdeli, Y., 2010. *Yazılımcılar için sql server 2008 r2 ve veritabanı programlama*. 5. Baskı. Ankara: Seçkin Yayıncılık.
- Hand, D., Mannila, H., Smyth, P., 2001. *Principles of data mining*. USA: The Mit Press
- Miner, G., Delen, D., Elder, J., Fast, A., Hill, T., ve Nisbet, R. A., 2012. *Practical text mining and statistical analysis for non-structured text data applications*. USA: Academic Press.
- Nabiyev, V. V., 2012. *Yapay zeka*. 4. Baskı. Ankara: Seçkin Yayıncılık.
- Oğuzlar, A., 2011. *Temel metin madenciliği*. Bursa: Dora Yayıncılık.
- Özkan, Y., 2008. *Veri madenciliği yöntemleri*. İstanbul: Papatya Yayıncılık.
- Silahtaroglu, G., 2008. *Kavram ve algoritmalarıyla temel veri madenciliği*. İstanbul: Papatya Yayıncılık.

Sürelî Yayınlar

- Ay, D. ve Çil, İ., 2008. Migros Türk a.ş.de birliktelik kurallarının yerleşim düzeni planlamada kullanılması. *Sakarya Üniversitesi Endüstri Mühendisliği Dergisi*. **21** (2), pp. 14-29.
- Dolgun, M. Ö., Özdemir G. T., Oğuz D., 2009. Veri madenciliğinde yapısal olmayan verinin analizi: metin ve web madenciliği. *İstatistikçiler Dergisi*. (2), pp.48-58.
- Emel, G. K., ve Taşkın, Ç., 2005. Veri madenciliğinde karar ağaçları ve bir satış analizi uygulaması. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi*. **6** (2), pp. 222-239.
- Gürbüz, F., Özbakır L., Yapıcı H., 2008. Türkiyede bir havayolu işletmesine ait parça söküm raporlarına ilişkin veri madenciliği uygulaması. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*. **24** (1), pp. 73-78.
- Işık, M. ve Çamurcu, A. Y., 2008. Web belgeleri kümelemede benzerlik ve uzaklık ölçüleri başarılarının karşılaştırılması. *Marmara Üniversitesi Fen Bilimleri Enstitüsü Dergisi*. (20), pp. 35-49.
- Oguzlar, A., 2004. Veri madenciliğinde birliktelik kuralları. *Uludağ Üniversitesi , İİBF, Ekonometri Bölümü*. **6** (22), pp. 315-321.
- Takçı, H.ve Soğukpınar, İ., 2002. Kütüphane kullanıcılarının erişim örüntülerinin keşfi. *Bilgi Dünyası*. **3** (1).
- Timor, M. ve Şimşek, U. T., 2008. Veri madenciliğinde sepet analizi ile tüketici davranışı modellemesi. *İstanbul Üniversitesi İşletme Fakültesi Yönetim*. (59), pp. 3-10.
- Timor, M., Ezerçe A., Gürsoy U. T., 2011. Müşteri profili ve alışveriş davranışlarını belirlemede kümeleme ve birliktelik kuralları analizi: perakende sektöründe bir uygulama. *İstanbul Üniversitesi İşletme Fakültesi Sayısal Yöntemler Anabilim Dalı*. (22), pp. 128-147.

Diğer Yayınlar

- Akbulut, S., (2006). Veri madenciliği teknikleri ile bir kozmetik markanın ayrılan müşteri analizi ve müşteri segmentasyonu. *Yüksek Lisans Tezi*. Ankara: Gazi Üniversitesi FBE.
- Ataseven, S., (2008). Üniversitelerin adaylar tarafından tercih edilme desenlerini veri madenciliği yöntemleri ile belirleyen bir model önerisi. *Yüksek Lisans Tezi*. İstanbul: İstanbul Kültür Üniversitesi FBE.
- Bilen, H., (2009). Bankacılık sektöründe persole seçimi ve performans değerlendirmesine ilişkin veri madenciliği uygulaması. *Yüksek Lisans Tezi*. Ankara: Gazi Üniversitesi FBE.
- Çelikyay, E. K., (2010). Metin madenciliği yöntemiyle türkçede en sık kullanılan ve birbirini takip eden harflerin analizi ve birliktelik kuralları. *Yüksek Lisans Tezi*. İstanbul: Beykent Üniversitesi FBE.
- Dolgun, M. Ö., (2006). Büyük alışveriş merkezleri için veri madenciliği uygulamaları. *Yüksek Lisans Tezi*. Ankara: Hacettepe Üniversitesi FBE.
- Dursun, B., (2008). Türkçe metinlerin benzerliğinin hesaplanması için yeni bir yöntem. *Yüksek Lisans Tezi*. İstanbul: Yıldız Teknik Üniversitesi FBE.
- Haberal, İ., (2007). Veri madenciliği algoritmalarını kullanarak web günlük erişimlerinin analizi. *Yüksek lisans Tezi*. Ankara: Başkent Üniversitesi FBE.
- Kalikov, A., (2006). Veri madenciliği ve bir e-ticaret uygulaması. *Yüksek Lisans Tezi*. Ankara: Gazi Üniversitesi FBE.
- Karaca, M. F., (2012). Metin madenciliği yöntemleri ile haber sitelerindeki köşe yazılarının sınıflandırılması. *Yüksek Lisans Tezi*. Karabük: Karabük Üniversitesi FBE.
- Muslu, D., (2009). Sigortacılık Sektöründe risk analizi: veri madenciliği uygulaması. *Yüksek Lisans Tezi*. İstanbul: İstanbul Teknik Üniversitesi FBE.
- Onat, A., (2008). Veri madenciliğinin web tabanlı uygulamalarda insan uyumluluklarının tespiti üzerine bir çalışma. *Yüksek Lisans Tezi*. Konya: Selçuk Üniversitesi FBE.

Diğer Yayınlar

- Özçınar, H., (2006). Kpss sonuçlarının veri madenciliği yöntemleriyle tahmin edilmesi. *Yüksek Lisans Tezi*. Denizli: Pamukkale Üniversitesi FBE.
- Rauka, L., (2012). Metinsel veri madenciliğinde bilgisayarlı çeviriciler. *Yüksek Lisans Tezi*. Trabzon: Karadeniz Teknik Üniversitesi FBE.
- Seçkin, K., (2011). Metin madenciliğinde kullanılan yöntemlerin karşılaştırılması: Siyasi parti liderlerinin grup genel toplantı konuşmaları ile bir uygulama. *Yüksek Lisans Tezi*. Sakarya: Sakarya Üniversitesi SBE.
- Sıramkaya, E., (2005). Veri madenciliğinde bulanık mantık uygulaması. *Yüksek Lisans Tezi*. Konya: Selçuk Üniversitesi FBE.
- Tiryaki, S., (2006). Lojistik alanında bir veri madenciliği uygulaması. *Yüksek Lisans Tezi*. İstanbul: İstanbul Teknik Üniversitesi FBE.
- Tosun, T., (2006). Veri madenciliği teknikleriyle kredi kartlarında müşteri kaybetme analizi. *Yüksek Lisans Tezi*. İstanbul: İstanbul Teknik Üniversitesi FBE.
- Uslu, N. C., (2009). Veri madenciliği ile bilgisayar ağlarında yeni bir saldırı tespit algoritması. *Yüksek Lisans Tezi*. Ankara: Gazi Üniversitesi FBE.
- Yücesoy, B., (2007). Türkçe dökümanlar için anlamsal benzerlik hesaplama yöntemi. *Yüksek Lisans Tezi*. İstanbul: İstanbul Teknik Üniversitesi FBE.

EKLER

EK 1: Namık KEMAL' in hayatı 1

Kaynak: <http://mirayson95.blogcu.com/namik-kemal-tevfik-fikret-in-hayati-ve-fikirleri/10728277> (Erişim Tarihi: 20/04/2013)

21 Aralık 1840'ta Tekirdağ'da doğdu, 2 Aralık 1888'de Sakız Adası'nda öldü. Asıl adı Mehmed Kemal. Namık adını ona şair Eşref Paşa verdi. Babası, II. Abdülhamid döneminde müneccimbaşılık yapmış olan Mustafa Asım Bey. Annesini küçük yaşında yitirince çocukluğunu dedesi Abdüllâtif Paşa'nın yanında, Rumeli ve Anadolu'nun çeşitli kentlerinde geçirdi. Bu yüzden özel öğrenim gördü. Arapça ve Farsça öğrendi. 18 yaşında İstanbul'a babasının yanına döndü. 1863'te Babıali Tercüme Odası'na kâtip olarak girdi. Dört yıl çalıştığı bu görev sırasında dönemin önemli düşünür ve sanatçılarıyla tanışma olanağı buldu. 1865'te kurulan ve daha sonra yeni Osmanlılar Cemiyeti adıyla ortaya çıkan İttifak-ı Hamiyet adlı gizli derneğe katıldı. Bir yandan da Tasvir-i Efkar gazetesinde hükümeti eleştiren yazılar yazıyordu. Gazete, Yeni Osmanlılar Cemiyeti'nin görüşleri doğrultusunda yaptığı yayım nedeniyle 1867'de kapatıldı.

Sürgünler dönemi

Namık Kemal, İstanbul'dan uzak olması için Erzurum'a vali muavini olarak atandı. Bu göreve gitmeyi erteledi ve Mustafa Fazıl Paşa'nın çağrısı üzerine Ziya Paşa'yla birlikte Paris'e kaçtı. Bir süre sonra Londra'ya geçerek Mustafa Fazıl Paşa'nın parasal desteğiyle Ali Suavi'nin Yeni Osmanlılar adına çıkardığı "Muhbir" gazetesinde yazmaya başladı. Ama Ali Suavi'yle anlaşamadı, Muhbir'den ayrıldı. 1868'de gene Fazıl Paşa'nın desteğiyle "Hürriyet" gazetesini çıkardı. Çeşitli anlaşmazlıklar yüzünden, Avrupa'da desteksiz kalınca, 1870'te zaptiye nazırı Hüsnü Paşa'nın çağrısıyla İstanbul'a döndü. Nuri, Reşat ve Ebüzziya Tevfik beylerle birlikte 1872'de "İbret" gazetesini kiraladı. Aynı yıl burada çıkan bir yazısı üzerine gazete 4 ay kapatıldı. İstanbul'dan uzaklaştırılmak için Gelibolu mutasarrıflığına atandı. Orada yazmaya başladığı "Vatan Yahut Silistre" oyunu, 1873'te Gedikpaşa Tiyatrosu'nda sahnelendi. Oyunu izleyenler galeyana gelip olay çıkardı. Namık Kemal birçok arkadaşıyla birlikte tutuklandı. Bu kez kalebentlikle Magosa'ya sürgüne gönderildi.

Türk Edebiyatı'nda İlkleri

1876'da I. Meşrutiyet'in ilanından sonra İstanbul'a döndü. Şura-yı Devlet (Danıştay) üyesi oldu. Kanun-î Esasi'yi (Anayasa) hazırlayan kurulda görev aldı. 1877 Osmanlı-Rus Savaşı çıkınca Meclis-i Mebusan kapatıldı, Namık Kemal tutuklandı. Midilli Adası'na sürüldü. 1879'da Midilli mutasarrıfı oldu. Aynı görevle 1884'te Rodos, 1887'de Sakız Adası'na gönderildi. Ertesi yıl burada öldü ve Gelibolu'da Bolayır'da gömüldü. Şiirlerini küçük yaşlardan itibaren yazdı. Şinasi'yle tanışmaya değin, şiirlerinde tasavvuf etkileri görülür. Bu dönemde özellikle Yenişehirli Avni, Leskofçalı Galib gibi şairlerden etkilendi. En önemli özelliklerinden biri, Türk şiirini Divan şiirinin etkisinden kurtarmaya çalışması. "Vatan Şairi" diye de isimlendirildi. Tiyatroya özel bir önem verdi, altı oyun yazdı. Bir yurtseverlik ve kahramanlık oyunu olan Vatan Yahut Silistre, Avrupa'da da ilgi uyandırdı ve beş dile çevrildi. İlk romanı "İntibah" 1876'da yayınladı. Ruhsal çözümlerinin, bir olayı toplumsal ve bireysel yönleriyle görmeye çalışmasının yanı sıra, dış dünya betimlemeleriyle de İntibah Türk romanında bir başlangıç sayılır. Romanı ve tiyatroyu toplumsal yaşama soktuğu gibi, edebiyat eleştirisini de Türkiye'ye ilk getiren kişilerden biri oldu. En önemli eleştiri eserleri Tahrib-i Harâbât ile Takip. Gazeteci olarak da Türk kültürü içinde önemli bir yeri var. Döneminin hemen hemen bütün yenilik yanlısı ve ilerici gazetelerinde yazıları yayımlandı. Siyasal ve toplumsal sorunlardan edebiyat, sanat, dil ve kültür konularına dek çok çeşitli alanlarda yazdığı makalelerin sayısı 500 kadar.

EK 2: Namık KEMAL' in hayatı 2

Kaynak: <http://www.merakname.com/namik-kemal-kimdir/> (Erişim Tarihi: 20/04/2013)

Namık Kemal kimdir? Namık Kemal ünlü Türk yazarların başında gelir. Aynı zamanda devlet adamı, şair, roman ve tiyatro yazarı ve gazetecidir. Türk milliyetçiliği öncülerindedir. Hayatı 1840 yılında Tekirdağ'da doğmuştur. Asıl adı Mehmet Kemal'dir. Küçük yaşta annesini kaybetmesi nedeniyle dedesinin yanında yaşamaya devam etti. Dedesinin çeşitli illerdeki görevi nedeniyle önce Bayezid Rüştüyesi sonra Valide Mektebine gitti. Ancak dedesinin görevi nedeni ile yarım kaldı eğitimleri ve özel eğitimler almak zorunda kaldı. Bu sayede Arapça ve Farsça dillerini öğrendi. Aynı zamanda Vaizzade Seyid Mehmet Hamdi Efendi'den divan edebiyatı dersleri aldı. Bunların yanı sıra ilgi alanları olan avcılık, atıcılık, cirit dersleri aldı. 1855 yılında Bulgaristan/Sofya'da Binbaşı Eşref Bey, şiirlerini okumuştur. Bunun üzerine Mehmet Kemal'e "yazıcı, şair" anlamlarına gelen "Namık" ismini vermiştir. Yine Sofya'da evlenmiş, üç çocuk sahibi olmuştur. 1858 yılında İstanbul'a dönmüştür. Bab-ı Ali Tercüme Odası'nda önce stajyer olarak başladı. 1863 yılında da kâtip oldu. Osmanlı Cemiyeti'ne mensup gizli bir dernek olan İttifak-ı Hamiyet'e katıldı. Tasvir-i Efkâr gazetesinde eleştiri yazıları yazdı. Bu yazılarda Yeni Osmanlı Cemiyeti'nin görüşlerini bildiren yazılar yayınlaması üzerine 1867 yılında gazete kapatıldı ve Erzurum'a atandı. Erzurum'a gitmeyi reddeden Namık Kemal, Mustafa Fazıl Paşa'nın daveti ile Paris'e kaçtı. Mustafa Fazıl Paşa ile birlikte Londra'da "Muhbir" gazetesini çıkardılar. Bir süre sonra çıkan anlaşmazlık nedeni ile Namık Kemal Paris'e geri döndü. Burada Fransız hükümetini yaptığı "Genç Osmanlıların ülkeyi terk etmesi" yönündeki çağrı nedeniyle Londra'ya geri döndü. Burada Hürriyet Gazetesi'ni çıkardı. Gazetenin anlaşmazlıklar üzerine kapanması nedeni ile 1870 yılında Sadrazam Ali Paşa ile barıştı ve İstanbul'a döndü. "Diyojen" isimli mizahi bir dergide imzasız fıkralar yayımladı. 1872 yılında İbret Gazetesi'ni çıkardı. Gazetede kendini eleştiren yazıları nedeni ile dönemin sadrazamı Mahmut Nedim Paşa gazeteyi kapatarak Namık Kemal'in İstanbul'dan uzaklaştırılması kararını almıştır. Bunun üzerine Namık Kemal, Gelibolu'ya atanmıştır. Burada kaldığı süre içinde "Vatan yahut Silistre" ve "Evrak-ı Perişan" eserlerini yazdı. Aynı zamanda başmuharrir anlamına gelen "BM" imzası ile İbret gazetesinde, isminin baş harflerinden oluşan "N.K." imzası ile Hadika gazetesinde yazılarını yazmaya devam etti. 1872 yılı sonlarına doğru Osmanlı Hükümeti tarafından açığa alındı. Namık Kemal,

İstanbul'a geri dönerek İbret gazetesinin başına geçti. Gazetede yazılan yazılar sebebi ile gazete tekrar kapatılarak Namık Kemal hakkında soruşturma başlatıldı. 1873 yılında Namık Kemal tiyatroya yönelerek Gelibolu'da yazdığı "Vatan yahut Silistre" oyununu sergiledi. Oyunun halkı etkileyerek olaylar çıkmasına neden olması sonucu İbret gazetesinde yazılar yayımlandı. Gazete bir daha açılmamak üzere kapatılarak Namık Kemal Magosa'ya sürüldü. Burada 38 ay kaldığı süreç içinde sık sık sıtma ve birçok hastalığa yakalandı. İstanbul'a döndüğünde II. Abdülhamit'in kurduğu komisyona üye oldu. Ancak mecliste yaptığı açıklama nedeni ile 6 ay hapis cezasına çarptırıldı. Davadan beraat etmesine rağmen adada ikamet etmek zorunda bırakıldı. Namık Kemal Midilli Adasını tercih etti. Burada 5 yıl mutasarrıf olarak görev yaptı. Görevi sırasında birçok Türk Okulu açtı, kaçakçılık ile mücadele etti. 1882 yılında çalışmalarındaki başarıları nedeniyle Nişan-i Osmanlı Madalyasına layık görüldü. Şikâyet üzerine 1884 yılında Midilli'den Rodos adasına atandı. Yine buradaki çalışmaları nedeniyle İmtiyaz madalyasına layık görüldü. 1887 yılında İngiliz ve Yunanlıların şikâyetleri nedeniyle Sakız Adasına atandı. Buradaki kuru hava nedeniyle hastalanarak 1888 yılında vefat etti. Naaşı adaya defnedilse de, daha sonra vasiyeti üzerine Gelibolu'ya nakledildi. Eserleri 1873 yılında Gelibolu'da "Vatan yahut Silistre" isimli tiyatro oyununu yazdı. Aynı yıl Zavallı Çocuk, 1874'te Akif Bey, 1875 yılında Gülnihal isimli tiyatro oyunlarını kaleme aldı. 1873-1876 yılları arasında Magosa Sürgününde İntibah romanını yazdı. Bir diğer romanı 1880 yılında ilk basımı yapılan Cezmi'dir. 1885 yılında diğer bir tiyatro eseri Celalettin Harzam Şah'ı ardından 1910 yılında Karabela'yı yazdı. Yine 1885 yılında eleştiri alanında Tahrib-i Harabat ve Takip, 1887 yılında İrfan Paşa'ya Mektup, 1888 yılında Mukaddeme-i Celal, son olarak Renan Müdafaaamesi isimli eserleri kaleme aldı. Birçok roman, tiyatro ve eleştiri eserlerinden sonra kendini tarihi eserlere vererek 1871 yılında Devr-i İstila, 1872 yılında Barika-i Zafer ve Evrak-ı Perişan eserlerini, 1874 yılında Kanije ve Silistre Muhasarası eserlerini yazmıştır. Osmanlı Tarihi ve Büyük İslam Tarihi adlı eserleri yazarın ölümünden sonra yayımlanmıştır.

EK 3: Namık KEMAL' in hayatı 3

Kaynak: http://tr.wikipedia.org/wiki/Namık_Kemal (Erişim tarihi 16/05/2013)

Namık Kemal (21 Aralık 1840, Tekirdağ - 2 Aralık 1888, Sakız Adası), Türk milliyetçiliğinin öncülerinden, Genç Osmanlı hareketi mensubu yazar, gazeteci, devlet adamı ve şair.

Yurtseverlik, hürriyet, millet kavramlarına bağlı bir Tanzimat Devri aydınıdır. Bu kavramları Türk fikir hayatına ve edebiyatına sokan kişi kabul edilir[1]. Heyecanlı, kavgacı kişiliği, akıcı, parlak üslubu nedeniyle devrinin diğer yazarlarından daha fazla tanındı[2]. “Vatan Şairi” ve “Hürriyet Şairi” olarak anılan Namık Kemal, şiirin yanı sıra tenkit, biyografi, tiyatro, roman, târih ve makale türlerinde eserler verdi. Özellikle Türk Edebiyatının ilk edebi romanı olan "İntibah" ve batılı anlamda Türk Edebiyatının ilk tiyatro eseri olan "Vatan Yahut Silistre" eserleriyle ünlüdür. Türkiye Cumhuriyeti'nin kurucusu Mustafa Kemal Atatürk'ü eserleri ve fikirleriyle etkiledi.

Yaşam öyküsü [değiştir]

21 Aralık 1840 tarihinde Tekirdağ'da dünyaya geldi. Babası Yenişehirli Mustafa Asım Bey, annesi Fatma Zehra Hanım'dır.

Tekirdağ'daki evlerinin civarında bulunan tekkenin şeyhi Tokatlı Hafız Ali Rıza Efendi kendisine “Mehmet Kemal” adını verdi[1]. Çocukluğu annesinin babası Abdülatif Paşa'nın yanında geçti. Abdülatif Paşa, Tekirdağ (Tekfurdağ) sancağında vali yardımcısı idi ; Afyonkarahisar sancağına tayin edildiğinde ailece Afyon'a taşındılar. 1848 yılında annesi Fatma Zehra Hanım'ı Afyon'da kaybetti[1]. Mehmet Kemal, yaşamını büyükbabasının yanında sürdürdü.

Abdülatif Paşa'nın değişik kentlerde görev yapması nedeniyle düzenli bir eğitime devam edemedi. Özel dersler aldı ve kendi kendini yetiştirmeye çalıştı. Arapça ve Farsça öğrendi. Dedesi Afyon'daki vali yardımcılığı görevinin ardından ailesiyle İstanbul'a gelmişti. Orada, 3 ay Bayezid Rüştüyesine ve ardından 9 ay Valide Mektebi'ne devam etme fırsatı buldu[3]. Dedesinin Kars'a mutasarrıf olarak atanması sebebiyle 1,5 yıl Kars'ta yaşadı. Karlı şair ve müderris Vaizzade Seyid Mehmet Hamid Efendi'den divan edebiyatını öğrendi. Avcılık, atıcılık, cirit dersleri aldı[1]. Kars'ta görevi sona eren dedesi ile 1854'te İstanbul'a döndü.

1855'te babasının Bulgaristan Filibe mal müdürü, dedesinin Sofya kaymakamı oluşu ile Sofya'ya gitti. Sofya'da evlerine ziyarete gelen dedesinin arkadaşı şair Binbaşı Eşref

Bey, şiirlerini okuduktan sonra Mehmet Kemal'e yazıcı, kâtip anlamlarındaki "Namık" adını verdi. O günden sonra Namık Kemal olarak anılmaya başladı. 18 yaşına kadar kaldığı Sofya'da komşuları Niş Kadısı Mustafa Ragıp Efendi'nin kızı Nesime Hanım ile evlendi[2]. Bu evlilikten Feride ve Ulviye adında iki kızı ve Ali Ekrem adında bir oğlu dünyaya geldi.

İstanbul yılları [değiştir]

1857'de İstanbul'a döndü ve Bab-ı Ali Tercüme Odası'nda stajyer olarak memurluğa başladı. 1858'de büyükannesi Mahmude Hanım'ı, 1859'da büyükbabası Abdülatif Paşa'yı kaybetti. Babasının ikinci evliliğini yaptığı Dürrüye Hanım'ın Kocamustafapaşa'daki evinde yaşadı[1]. Babasının bu evliliğinden Naşit adında bir kardeşi oldu[2]. 1859'da Gümrük Kalemi'nde çalışmaya başladı.

İlk şiirlerini Sofya'da yazan Namık Kemal, İstanbul'a geldiğinde kısa sürede şairler arasında tanınmıştı. Henüz Batı edebiyatı ile bir teması yoktu. İstanbul'da divan edebiyatı geleneğini takip ettiren şairlerle tanıştı. Arap ve Fars edebiyatlarını öğrenmeye çalıştı. Leskofçalı Galip Bey adlı şair ile yakın dostluk kurdu. Bu şairin başkanlığında kurulan Encümen-i Şuara adlı şairler topluluğuna katıldı.

1863'ten itibaren dört yıl yeniden Tercüme Odası'nda görev aldı. Bu yeni görevi sırasında Batı'yı tanıyan kimselerle tanışma imkânı buldu ve gözlerini batı kültürüne çevirdi. Edebiyatta batılılaşmanın ilk adımlarını atan İbrahim Şinasi ile tanışması hayatını değiştirdi. Sanat ve hayat görüşü değişti. Batı edebiyatını öğrenmeye başladı, ilgisi nesire yöneldi. Tarih ve hukuk alanında kendini geliştirmeye çalıştı. Tercüme odasının bir kâtibinden Fransızca dersleri aldı. Tasvir-i Efkâr'da fıkra ve tercüme yazılar kaleme aldı. İlk defa Şinasi'de gördüğü "hak, millet, vatan, hürriyet, millet meclisi" gibi kelimeleri yaygınlaştırdı[3].

Genç Osmanlılar [değiştir]

1865'te Şinasi, Tasvir-i Efkâr Gazetesi'ni kendisine bırakarak Fransa'ya gidince Namık Kemal, tek başına gazeteyi çıkardı. Aynı dönemde İttifak-i Hakimiyyet adlı (daha sonra Yeni Osmanlılar Cemiyeti adını alacak) gizli derneğin kurucuları arasına girdi (Sağırismetbeyzâde Mehmet Bey'in öncülüğündeki derneğin diğer kurucuları Menâpirzâde Nuri Bey, Kayazade Reşat Bey, Mir'at Mecmuası sahibi Mustafa Refik Bey, Suphipaşazade Ayetullah Bey ve Ziya Bey'dir). Derneğin amacı bir anayasa hazırlanmasını ve parlamenter bir yönetim sistemi kurulmasını sağlamaktı. Namık

Kemal gazetesinde, bu görüşler doğrultusunda ve hükümet aleyhine şiddetli makaleler yayınladı. "Şark Meselesi" üzerine yazdığı bir makale, gazetenin 1867'de kapatılmasına ve kendisinin Erzurum vali muavini olarak atanmasına yol açtı.

Namık Kemal, hükümet tarafından gönderildiği Erzurum'a gitmek yerine Ziya Paşa ile birlikte Paris'e kaçtı. O ve arkadaşlarını Paris'te yaşayan Mısırlı prens Mustafa Fazıl Paşa davet etmiş ve maddi himayesine almıştı. Mısır valisi Kavalalı Mehmet Ali Paşa'nın torunu olan ancak Sultan Abdülaziz'in bir fermanıyla Mısır yönetimindeki haklarından mahrum edilen Mustafa Fazıl Paşa, kendisini Yeni Osmanlılar Cemiyeti'nin reisi ilan etmiş ve Avrupa'ya davet ettiği örgüt üyelerinin finansörlüğünü üstlenmiş birisiydi[4]. M. Fazıl Paşa'nın desteğiyle Londra'da "Muhbir" adlı gazeteyi çıkardılar ancak Namık Kemal, Ali Suavi ile yaşadığı anlaşmazlık üzerine Muhbir'den ayrıldı. Aynı yıl Sultan Abdülaziz Uluslararası Paris Sergisi'ni görmek üzere şehre gelince Fransız hükümeti Genç Osmanlılar'ı ülkeyi terk etmeye davet etti. Namık Kemal, bazı arkadaşlarıyla birlikte Londra'ya gitti ve orada "Hürriyet Gazetesi"ni çıkardılar. Bu arada Mustafa Fazıl Paşa, Paris'e gelen Abdülaziz'le ilişkilerini düzeltmiş ve onunla İstanbul'a dönmüştü. Giderken gazeteyi çıkarmaya devam etmelerini, desteğinin süreceğini söylediye de İstanbul'a döndükten sonra fikrini değiştirdi ve geçici olarak Hürriyet'i kapatmalarını istedi[5]. Bunun üzerine Namık Kemal ile Ziya Paşa gazeteyi kendi imkânları ile çıkarmayı denediler. Bir süre sonra arkadaşları ile arası bozulan Namık Kemal vazgeçti ve 1870'te Sadrazam Âli Paşa ile barışıp yurda döndü.

Sürgün yaşamı [değiştir]

Siyasetten uzak durmak, yazı yazmamak koşuluyla affedilmiş olan[6] Namık Kemal, İstanbul'a döndükten sonra "Diyojen" adlı mizah dergisinde imzasız fıkralar yazdı [2]; Sadrazam Ali Paşa'nın ölümünden sonra 1872'de "İbret Gazetesi"ni çıkararak yeniden muhalefete başladı. Gazete sık sık kapatıldı ve sonunda sadrazam Mahmut Nedim Paşa'yı eleştiren yazılar yüzünden Namık Kemal, İstanbul'dan uzaklaşması için mutasarrıf olarak Gelibolu'ya atandı.

Birkaç ay kaldığı Gelibolu'da "Vatan yahut Silistre" adlı oyunu ile "Evrâk-ı Perişan" adlı eserini tamamladı. Gelibolu'nun bazı sorunları ile ilgilendi ve su davasını halletti[1]. Rumeli fatihi Gazi Süleyman Paşa'nın Bolayır'daki kabrini ziyaret etti. Ebüzziya Tevfik Bey'e burada gömülmeyi vasiyet etti.

Namık Kemal, bir yandan da "İbret Gazetesi"'ne "BM" (Baş muharrir) ve Ebuzziya'nın çıkardığı "Hadika" Gazetesine "N.K" imzası ile yazı göndermeye de devam ediyordu. Gelibolu'da salgın haline gelen kuduz hastalığını önlemek için köpekleri sürgün etmesi bahane edilerek Gelibolu mutasarrıflığı görevinden alındı.

Vatan Yahut Silistre [değiştir]

Osmanlı hükümeti tarafından açığa alınan Namık Kemal 1872'nin son günlerinde Gelibolu'dan İstanbul'a döndü, İbret'in başına geçti. Çok geçmeden bir makalesi nedeniyle hakkında soruşturma açılıp gazetesi tekrar kapatılınca tiyatro ile ilgilenmeye başladı. Vatan yahut Silistre oyunu, 1 Nisan 1873 gecesini İstanbul'da Güllü Agop'un Gedikpaşa'daki tiyatrosunda sahnelendi. Oyunun sahnelenmesi halkı coşturup olaylar çıkmasına neden olmuştu. Bu konuda İbret'te yayımlanan yazılardan sonra gazete bir daha çıkmamak üzere kapatıldı; Namık Kemal ve dört arkadaşı yargılanmadan sürgüne gönderildiler[6]. Namık Kemal Mağusa'ya, Ahmet Mithat ile Ebuzziya Tevfik Bey Rodos'a, Menapirzade Nuri ve Bereketzade Hakkı Beyler de Akka'ya sürüldü[7].

Magosa (Kıbrıs) Sürgünlüğü [değiştir]

Namık Kemal'in Mağusa (Kıbrıs) sürgünlüğü 38 ay sürdü. Mağusa'da son derece olumsuz koşullar altında yaşamak zorunda kaldı, pek çok kez sıtmaya ve başka hastalıklara yakalandı[6]. Edebiyatçı Namık Kemal, birkaçı dışında eserlerinin tamamını bu dönemde Kıbrıs'ta vermişti.

Midilli sürgünlüğü [değiştir]

Sürgün dönüşü İstanbul'da bir kahraman gibi karşılandı[2]. Tahta çıkışından 93 gün sonra akıl bozukluğu gerekçesiyle indirilen V. Murat'ın yerine Osmanlı tahtına oturan II. Abdülhamit, ilk Osmanlı Anayasası'nı oluşturmak için bir komisyon kurdu. Namık Kemal, bu komisyonun bir üyesi oldu. Ancak şair, padişahın aleyhine bir tehdit beyiti yazıp bunu mecliste okuyunca mahkemede yargılandı. Söylediği Arapça beyit, "Bir şey, ikilendi mi, muhakkak üçlenir de" anlamındaydı ve tıpkı Abdülaziz ve V. Murat gibi Abdülhamit'in de tahttan indirilebileceğini ima ediyordu. Namık Kemal, asayışı bozduğu gerekçesiyle suçlu bulunup 6 ay hapis cezasına çarptırıldıysa da sonradan beraat etti. Girit Adası'nda ikamete mecbur edildi. Kendi isteği üzerine ikameti Midilli Adası'na çevrildi. 2,5 yıl sonra Midilli mutasarrıfı olarak görevlendirildi. Midilli'de tanıdığı genç yaşta Hüseyin Hilmi Paşa'yı ömrü boyunca koruyup destekledi. Hüseyin Hilmi Paşa, yıllar sonra 1909'da sadrazamlığa kadar yükselmiştir.

1879'dan itibaren 5 yıl süren Midilli'deki görevi sırasında kaçakçılıkları önledi; hazine gelirini arttırdı. 20 Türk ilkokulu açtı. Türk'lerin hayat seviyesini yükseltti. Adalarda yaşayan Türk ahalisinin sorunlarını dile getiren bir rapor hazırlayıp Bâb-ı Âli'ye sundu.[1] 1882'de Nişan-i Osmanlı madalyası ile ödüllendirildi. "Vaveyla", "Murabba", "Vatan Mersiyesi" gibi şiirlerini burada yazdı. Magosa'da yazmaya başladığı Celaleddin Harzemşah adlı eserini tamamladı. Bu eser, okunmak için yazılmış 15 perdelik tarihi bir oyundur. Harzemşahlar Devleti'nin son hükümdarı Celaleddin Harzemşah etrafında gelişen oyunda İslam birliği düşüncesini işledi. Abdülhamit, bu eserinden ötürü onu bâlâ rütbesi ile ödüllendirdi.

Namık Kemal'in Midilli'de kaçakçılıkla mücadelesinden çıkarları zarar görenlerin şikâyetinden sonra 1884'te Rodos mutasarrıfı oldu. Rodos adasındaki çalışmaları da padişahın imtiyaz madalyası ile ödüllendirildi. Rodos'ta, Osmanlı tarihi hakkında eser yazmaya başladı. İngiliz ve Yunanlılar'ın şikayeti üzerine 1887'de Rodos'taki görevi sona erdi. Sakız Adası mutasarrıfı oldu.

Ölümü [değiştir]

Sakız Adası'nın kuru havası nedeniyle rahatsızlanan Namık Kemal, 2 Aralık 1888 günü 48 yaşında hayatını kaybetti. Adada bir caminin haziresine defnedildi. Arkadaşı Ebüziyya Tevfik, şairin Bolayır'da gömülme arzusunu padişah II. Abdülhamit'e iletince naaşı Gelibolu'ya nakledildi. Bolayır'da Orhan Gazi'nin oğlu Şehzade Gazi Süleyman Paşa'nın türbesinin yanına gömüldü. Birkaç yıl sonra Sultan Abdülhamit bir türbe yaptırdı. Türbenin planını Tevfik Fikret çizdi. 1912 Mürefte-Şarköy depreminde sütunlar zedelendiği için halen mermer kaplı bir kabirde bulunmaktadır..

Namık Kemal'in ölümünden sonra II. Abdülhamit, şairin oğlu Ali Ekrem'i sarayda görevlendirdi, babası Mustafa Asım'ı ise saraya müneccimbaşı tayin etti.[8]

Sanatçı özellikleri [değiştir]

Tanzimat döneminin en önemli düşünce, sanat ve siyaset adamlarından birisidir. "Toplum için sanat" anlayışını benimsemiştir. Sanatı, toplumun Batılılaşması için bir araç olarak kullanmıştır. Eserlerini halkın anlayabileceği sade bir dille yazmayı amaçlamıştır. Divan edebiyatının süslü-sanatlı düz yazısı yerine, belli bir düşünceyi iletmeyi amaçlayan yeni bir düzyazıyı kullanmıştır. Eserlerinde noktalama işaretlerini kullanmıştır. Gençliğinde Divan Edebiyatı tarzında şiirler yazmış, Avrupa'ya gittikten sonra yeni edebiyatı benimsemiş ve o yolda yapıtlar vermiştir. Namık Kemal, Fransız

edebiyatını örnek almış, romantizmin etkisinde kalmıştır. Şiirleri biçim bakımından eski, konu bakımından yenidir. Yurt, ulus, özgürlük gibi konuları işlemiştir. Ayrıca şiirlerinde mücadeleci tipte bir insan yaratmıştır. Tiyatroyu “eğlencelerin en faydalısı” olarak nitelemiş, halkın eğitilmesinde okul gibi görmüş, sahne dili ve tekniği yönünden başarılı yapıtlar vermiştir.

EK 4: Namık KEMAL' in hayatı 4

Kaynak: http://www.edebiyatogretmeni.net/namik_kemal.htm (Erişim Tarihi: 16/05/2013)

21 Aralık 1840'ta Tekirdağ'da doğdu, 2 Aralık 1888'de Sakız Adası'nda öldü. Asıl adı Mehmed Kemal. Namık adını ona şair Eşref Paşa verdi. Babası, II. Abdülhamid döneminde müneccimbaşılık yapmış olan Mustafa Asım Bey. Annesini küçük yaşında yitirince çocukluğunu dedesi Abdüllâtif Paşa'nın yanında, Rumeli ve Anadolu'nun çeşitli kentlerinde geçirdi. Bu yüzden özel öğrenim gördü. Arapça ve Farsça öğrendi. 18 yaşında İstanbul'a babasının yanına döndü. 1863'te Babialı Tercüme Odası'na kâtip olarak girdi. Dört yıl çalıştığı bu görev sırasında dönemin önemli düşünür ve sanatçılarıyla tanışma olanağı buldu. 1865'te kurulan ve daha sonra yeni Osmanlılar Cemiyeti adıyla ortaya çıkan İttifak-ı Hamiyet adlı gizli derneğe katıldı. Bir yandan da Tasvir-i Efkâr gazetesinde hükümeti eleştiren yazılar yazıyordu. Gazete, Yeni Osmanlılar Cemiyeti'nin görüşleri doğrultusunda yaptığı yayın nedeniyle 1867'de kapatıldı.

Sürgünler dönemi

Namık Kemal, İstanbul'dan uzak olması için Erzurum'a vali muavini olarak atandı. Bu göreve gitmeyi erteledi ve Mustafa Fazıl Paşa'nın çağrısı üzerine Ziya Paşa'yla birlikte Paris'e kaçtı. Bir süre sonra Londra'ya geçerek Mustafa Fazıl Paşa'nın parasal desteğiyle Ali Suavi'nin Yeni Osmanlılar adına çıkardığı "Muhbir" gazetesinde yazmaya başladı. Ama Ali Suavi'yle anlaşamadı, Muhbir'den ayrıldı. 1868'de gene Fazıl Paşa'nın desteğiyle "Hürriyet" gazetesini çıkardı. Çeşitli anlaşmazlıklar yüzünden, Avrupa'da desteksiz kalınca, 1870'te zaptiye nazırı Hüsnü Paşa'nın çağrısıyla İstanbul'a döndü. Nuri, Reşat ve Ebüzziya Tevfik beylerle birlikte 1872'de "İbret" gazetesini kiraladı. Aynı yıl burada çıkan bir yazısı üzerine gazete 4 ay kapatıldı. İstanbul'dan uzaklaştırılmak için Gelibolu mutasarrıflığına atandı. Orada yazmaya başladığı "Vatan Yahut Silistre" oyunu, 1873'te Gedikpaşa Tiyatrosu'nda sahnelendi. Oyunu izleyenler galeyana gelip olay çıkardı. Namık Kemal birçok arkadaşıyla birlikte tutuklandı. Bu kez kalebentlikle Magosa'ya sürgüne gönderildi.

1876'da I. Meşrutiyet'in ilanından sonra İstanbul'a döndü. Şura-yı Devlet (Danıştay) üyesi oldu. Kanun-î Esasi'yi (Anayasa) hazırlayan kurulda görev aldı. 1877 Osmanlı-Rus Savaşı çıkınca Meclis-i Mebusan kapatıldı, Namık Kemal tutuklandı. Midilli Adası'na sürüldü. 1879'da Midilli mutasarrıfı oldu. Aynı görevle 1884'te Rodos, 1887'de Sakız Adası'na gönderildi. Ertesi yıl burada öldü ve Gelibolu'da Bolayır'da gömüldü. Şiirlerini küçük yaşlardan itibaren yazdı. Şinasi'yle tanışmaya değin, şiirlerinde tasavvuf etkileri görülür. Bu dönemde özellikle Yenişehirli Avni, Leskofçalı Galib gibi şairlerden etkilendi. En önemli özelliklerinden biri, Türk şiirini Divan şiirinin etkisinden kurtarmaya çalışması. "Vatan Şairi" diye de isimlendirildi. Tiyatroya özel bir önem verdi, altı oyun yazdı. Bir yurtseverlik ve kahramanlık oyunu olan Vatan Yahut Silistre, Avrupa'da da ilgi uyandırdı ve beş dile çevrildi. İlk romanı "İntibah" 1876'da yayınladı. Ruhsal çözümlerinin, bir olayı toplumsal ve bireysel yönleriyle görmeye çalışmasının yanı sıra, dış dünya betimlemeleriyle de İntibah Türk romanında bir başlangıç sayılır. Romanı ve tiyatroyu toplumsal yaşama soktuğu gibi, edebiyat eleştirisini de Türkiye'ye ilk getiren kişilerden biri oldu. En önemli eleştiri eserleri Tahrib-i Harâbât ile Takip. Gazeteci olarak da Türk kültürü içinde önemli bir yeri var. Döneminin hemen hemen bütün yenilik yanlısı ve ilerici gazetelerinde yazıları yayımlandı. Siyasal ve toplumsal sorunlardan edebiyat, sanat, dil ve kültür konularına dek çok çeşitli alanlarda yazdığı makalelerin sayısı 500 kadar.

EK 5: Namık KEMAL' in hayatı 5

Kaynak: <http://www.soruegitim.com/namik-kemal-hayati-hakkinda/>

(Erişim Tarihi: 16/05/2013)

21 Aralık 1840'ta Tekirdağ'da doğdu, Asıl adı Mehmed Kemal'dir, Namık adını ona şair Eşref Paşa vermiştir. Babası, II.Abdülhamid döneminde müneccimbaşılık yapmış olan Mustafa Asım Bey'dir. Annesini küçük yaşta yitirince çocukluğunu dedesi Abdüllatif Paşa'nın yanında, Rumeli ve Anadolu'nun çeşitli kentlerinde geçirdi, Arapça ve Farsça öğrendi.

18 yaşlarında İstanbul'a babasının yanına dönen Namık Kemal, 1863 yılında Babıali Tercüme Odası'na katip olarak girdi. Bu görevi sırasında dönemin önemli düşünür ve sanatçılarıyla tanışma olanağı buldu. Namık Kemal ilk şiirlerini çocuk denecek yaşlarda yazmaya başlamıştır. Eski ve yeni kuşaktan şairlerin bir araya gelerek kurdukları Encümen Şuara'ya ve kimi Divan şairlerine nazireler yazmıştır. Diri bir sesle konuşması ve yapıtlarına kattığı yeni kavramlarla Türk Şiirini Divan Şiirinin edilgen edasından kurtarmıştır. Bütün bu nitelikler onun **Vatan Şairi** olarak anılmasına yol açmıştır.

Tiyatro türüne önem veren Namık Kemal, altı oyun yazmıştır. Bir yurtseverlik ve kahramanlık oyunu olan "**Vatan Yahut Silistire**"yalnız ülke için değil, Avrupada'da ilgi uyandırmış ve beş dile çevrilmiştir. İlk romanı olan "İntibah" 1876 yılında yayımlanmıştır. Romanı ve tiyatroyu toplumsal yaşama soktuğu gibi, edebiyat elestrisini de Türkiye'ye ilk getiren kişilerden biri olmuştur.

Gazeteci olarak da Türk kültürü içinde önemli bir yer alır. Dönemin hemen hemen bütün yenilik yanlısı ve ilerici gazetelerinde yazmıştır. Siyasal ve toplumsal sorunlardan edebiyat, sanat, dil ve kültür konularına dek çok çeşitli alanlarda yazdığı makalelerin sayısı 500 kadardır.

"Vatan Yahut Silistire" oyunu 1873'te Gedik Paşa Tiyatrosu'nda sahnelendiğinde halkı coşturup olaylara neden olması ve bu haberi gazetede yazması üzerine tutuklanan Namık Kemal, Magosa'ya sürgüne gönderildi. 1876 yılında I.Meşrutiyet'in ilanından sonra İstanbul'a döndü, Şura-yı Devlet(Danıştay) üyesi oldu. II. Abdülhamid'in Meclis-i Mebusan'ı kapatması üzerine tutuklandı. Beş ay tutuklu kaldıktan sonra Midilli

Adası'na sürüldü. 1879 yılında Midilli mutasarrıfı oldu, aynı görevle 1884'te Rodos, 1887'de Sakız Adası'na gönderildi.

Namık Kemal, 2 Aralık 1888'de Sakız Adası'nda hayata veda etti ve Gelibolu Bolayır'da defnedildi.

,

EK 6: Namık KEMAL' in hayatı 6

Kaynak:http://mebk12.meb.gov.tr/meb_iys_dosyalar/46/02/734364/icerikler/namik-kemal8217in-hayati_239933.html (Erişim Tarihi: 16/05/2013)

Namık Kemal, 21 ARALIK 1840'ta Tekirdağ'da doğdu. Müneccimbaşı Mustafa Asım ile Fatma Zehra Hanım'ın oğludur. Sekiz yaşındayken annesinin ölümü üzerine dedesi Abdüllatif Paşa'nın yanına alındı. Bir yıl İstanbul'da Beyazıt ve Valide rüştiyelerinde okudu (1849). Rumeli ve Anadolu'nun çeşitli merkezlerinde görevler alan dedesiyle birlikte dolaşır. Kars'ta buldukları yıllar (1851-1854) müderris ve şair Seyyit Mehmet Hamit Efendi tarafından yetiştirildi. Bir yıl sonra İstanbul'a döndü. Burada Arapça ve Farsça öğrenimi gördü. Bir süre sonra kaymakamlık görevi verilen dedesiyle birlikte Sofya'ya gitti (15 Mayıs 1855).

Sofya'da geçirdiği iki yıl Namık Kemal'e yeni öğrenim olanakları kazandırmış, Fransızca'ya ve ilk şiir denemelerine bu şehirde başlamıştır. 1857'de İstanbul'a gelen Namık Kemal bir süre sonra Tercüme Odası'na girdi ve hem edebiyat, hem düşün adamı kişiliğinin oluşması yolunda kendisine çok şey kazandıran bir ortam içinde yetişme olanakları buldu, Fransızcasını ilerletti.

Öte yandan dönemin düşün ve sanat adamlarının toplantılarına katılmaya başlamış, aralarında Leskofça'lı Galip, Hersekli Arif Hikmet, Şeyh Osman Şems gibi ünlü şairlerinde bulunduğu Encümen-i Şûra çevresindeki kişilerle tanışmıştı. Şinasi ile yakınlık kurması ona 1862'den itibaren Tasvir-i Efkâr'da yazma olanağı sağladı. Adı dönemin reformcu olarak bilinen aydınları arasında duyulmaya başladı.

'Yeni Osmanlılar' adlı gizli örgütün İstanbul'daki ilk toplantısına katıldığı zaman (Haziran 1865) Namık Kemal 25 yaşında bir gençti. Tasfir-i Efkâr ve Ali Suavi'nin çıkardığı Muhbir (1866) gazetelerinde, 'Yeni Osmanlılar Cemiyeti'nin gizli toplantılarında alınan kararlar doğrultusunda yapılan yayın, özellikle Ali Suavi'nin 'Millet Meclisi Usulü'nden söz açan yazısı Bab-I Ali'yi baskıya yöneltti. Çok geçmeden gazete kapatıldı. Namık Kemal Erzurum vali muavinliğine atanarak İstanbul'dan uzaklaştırıldı. Bu evrede Padişah Abdülaziz'e karşı olanlardan Mısırlı Hidiv ailesinden Prens Mustafa Fazıl Paşa, Namık Kemal ve Ziya Paşayı Paris'e çağırdı. Bu çağrı üzerine Namık Kemal ve Ziya Paşa Paris'e kaçtılar (17 Mayıs 1867). Bir süre sonra orada toplanan dokuz ihtilalci, Paşanın başkanlığında 'Yeni Osmanlılar Cemiyeti'nin ilk yönetim kurulunu meydana getirdiler.

Paris'te uzun süre sürekli olarak Muhbir (31 Ağustos) gazetesine yazıyor, bir yandan da felsefe, edebiyat, hukuk, toplum bilim alanlarında kendisini eğitmeye, dönemin bilim adamlarıyla tanışarak onlardan yararlanmaya çalışıyordu. Bir süre sonra Ali Suavi ile görüş ayrılığına düşünce Ziya Paşa ile birlikte Londra'ya geçen Namık Kemal, orada çıkardıkları Hürriyet gazetesinde (29 Haziran 1868) yazmaya başladı. Bu gazetede yazıları 64.sayıya kadar sürdü.

Namık Kemal'in İstanbul'a döndükten sonra uzun süre Sadrazam Ali Paşanın ölümüne kadar (7 Eylül 1871) sustuğu söylenebilir. Sadrazamın ölümünden sonra, önce birkaç arkadaşıyla birlikte 'İstikbal' adlı bir gazete çıkarmak istediysede hükümet izin vermeyince 'Hadika' gazetesinde ve Teodar Kasap'ın yayımladığı Türkçe ilk güldürü dergisi olarak bilinen Diyojen'de yazmaya başladı. Bu derginin de kapatılması üzerine arkadaşlarıyla Aleksan Sarrafyan adlı Ermeni'nin sahibi olduğu 'İbret' gazetesini kiradılar. Ne var ki, 13 Haziran 1872'de başlayan bu yayın, 9 Temmuz 1872 günü hükümetçe yasaklandı. Namık Kemal de Gelibolu mutasarrıflığına atanarak İstanbul'dan uzaklaştırıldı. Kısa süre sonra görevinden azledilerek dönüşü serbest bırakıldı (25 Aralık 1872). Bu tarihten sonra bazı yazılarından seçmeleri 'Evrak-ı Perişan' adlı kitapta toplayıp yayınlama isteğine ise izin verilmedi.

V.Murat'ın tahta getirmesinden hemen sonra çıkarılan af sonucu öteki sürgünlerle birlikte İstanbul'a dönen (7 Haziran 1876) Namık Kemal'e Şura-yı Devlette ve Kanun-i Esasiyi hazırlayacak kurulda görev verildi.

Bu arada V.Murat'ın sinir hastalığı artmış, Mithat Paşa ve arkadaşları Şehzade II.Abdülhamit'i (Meşrutiyetin ilanına engel olmayacağı üzerine söz alarak) tahta çıkardılar (31 Ağustos 1876). Çok geçmeden, II.Abdülhamit Mebuslar Meclisi açıldığı gün Sadrazam Mithat Paşa ile Devlet Şurası üyesi Namık Kemal'i tutuklamaktan çekinmedi. Namık Kemal maaşı ödenmek koşuluyla Midilli adasına sürgün edildi (1877). İki yıl sonra (Avrupa'ya kaçması ihtimali göz önünde tutularak) bâlâ rütbesiyle mutasarrıflık verildi.

Beş yıl süren (1879-1884) Midilli'deki yaşamında Namık Kemal'in kendisini tamamen Osmanlı Tarihi'ni yazmaya verdiği söylenebilir. Ama ne var ki, beş yıllık emeğine konan yasak onu çok sarstı. Ölümünden az önce Midilli'den Rodos'a (1884), oradan da Sakız'a (1888) atanmıştı. Manevi güçsüzlüğü arttıkça, yakalandığı zatürreye karşı

koyma gücünü bulamadı. Kısa süren hastalık sonucu öldü. Ölümünden sonra naşı Bolayır'a Tefvik Fikret'in çizdiği mezara taşındı.

EK 7: Namık KEMAL' in hayatı 7

Kaynak:<http://www.dilededebiyat.net/turk-edebiyati-tarihi/bati-etkisindeki-turk-edebiyati/tanzimat-edebiyati/tanzimat-donemi-sanatcileri/namik-kemal/namik-kemal-2>
(Erişim Tarihi: 16/05/2013)

Tanzimat Edebiyatı Dönemi şairlerimizden olan Namık Kemal, hürriyet, vatan, millet konularındaki eserleriyle Tanzimat Devrinin en gür sesli şairi olur, toplum için sanat ilkesine bağlanır, Türklüğün ve Osmanlı Devleti'nin düşmanlarına karşı maneviyatın yükseltilmesi yolunda didinir, böylece vatan şairlerimizin en büyüğü sayılır, diktaya karşı siyasi hürriyetleri elde etmek uğrunda mücadeleleriyle hürriyet kahramanı olur, her çağda vatanseverliğin ve hürriyetçiliğin sembolü ve bayrağı haline gelir.

HAYATI

21 Aralık 1840'ta Tekirdağ'da dünyaya gelen Namık Kemal, Müneccimbaşı Mustafa Asım Efendi'nin oğludur. İki yaşındayken annesini kaybeden Namık Kemal, çocukluğunu annesinin babası Abdülatif Paşa'nın yanında geçirir. Bu arada, dedesi Abdülatif Paşayla birlikte Kars ve Sofya'ya gider, aynı zamanda özel öğrenim görerek hem genel bilgisini attırır, hem de Fransızca'yı öğrenir. 1856'da evlendikten sonra, 1857'de İstanbul'a gelen Namık Kemal, daha o zaman küçük bir divanı dolduracak kadar şiir yazmış bulunur.

1860 yılında Babıâli Tercüme Odası'na memur olarak giren Namık Kemal, bu sırada Şinasi ile tanışır, onun teşvikiyle Tasvir-i Efkâr gazetesinde yazmaya başlar; Şinasi'nin iki yıl sonra Paris'e gitmesi üzerine gazeteyi tek başına yönetmeye başlar. Ama gazetede yazılan hürriyet fikirlerini canlandırıcı nitelikte ve "istipdat"a karşı halkta tepki uyandırıcı mahiyette görülür ve Tasvir-i Efkâr gazetesi kapatılır, Gazetenin yazarlarından yakalananlar sürgün edilir, diğerleri de yabancı memleketlere kaçmak zorunda kalır. Bu arada Namık Kemal de, Ziya Paşayla birlikte Paris'e kaçar.

1867'de Paris'e geldikten bir yıl sonra Londra'da Hürriyet gazetesini çıkararak Meşrutiyet kurmak için çalışan Yeni Osmanlılar Cemiyeti'nin amaçlarını yansıtmaya başlar. İstanbul'dan siyasi hükümlülere af çıktığı haberleri gelince, 1870 yılında yurda dönen Namık Kemal, aynı yıl, arkadaşlarıyla birlikte İbret gazetesini yayımlar. Ama on dokuz sayı sonra İbret gazetesi de kapatılır ve Namık Kemal Gelibolu Mutasarrıflığına atanarak İstanbul'dan uzaklaştırılır. Bir süre sonra Gelibolu Mutasarrıflığından da azledilen Namık Kemal, tekrar İstanbul'a döner. 1873'de, Vatan Yahut Silistre adındaki

piyesinin Gedikpaşa Tiyatrosu'nda oynayışı halkı coşturunca, yakalanır ve Kıbrıs'taki Magosa Zindanı'na sürülür. Zindana girerken, bileklerine takılmış zincirlerin çıkardığı sesler için "Bu ses, milletimin hürriyet sesidir!" der.

Namık Kemal, Magosa sürgününde otuz sekiz ay kalır. Hürriyet aşkı ve vatan sevgisi daha da kuvvetlenir, ama ömrü boyunca sürgünlerde geçen vücudu zayıflar. Ağızdan ağza dolaşan şiirleri, memlekette hürriyet ve Meşrutiyet yolundaki çalışmalara hız verir. 1876'da Abdülaziz'in tahttan indirilmesi ve Anayasayı ilan edeceğine, Meşrutiyet kuracağına söz veren 2. Abdülhamit'in Padişah olması üzerine, Namık Kemal, Magosa Zindanı'ndan, kendisini hürriyet kahramanı olarak karşılayan İstanbul'a mutlulukla döner. Ülküsü gerçekleşir, millet Anayasa'lı idareye kavuşacaktır, bu hürriyet düzeni sayesinde Türk gücü tekrar canlanacak ve imparatorluğu yükseltecektir... Bu umutlarla, Anayasa'nın hızlanması işinde Ziya Paşayla birlikte çalışır. Kısa zamanda Anayasa ilan edilir, 1. Meşrutiyet İdaresi kurulur... Ama 2. Abdülhamit, 1877'de Osmanlı-Rus Savaşı çıkınca Meclis-i Mebusan'ı kapatır ve uzun yılların çalışmalarıyla, mücadeleleriyle ulaşılan devlet yönetimi son bulur. Namık Kemal, aynı yıl, yakalanarak Midilli adası'na sürülür... Bir süre sonra 1879'da Midilli Mutasarrıflığına atanan Namık Kemal, 1884'de Rodos, 1887'de Sakız Adası Mutasarrıflığına getirilir. Sakız Adası Mutasarrıfı iken, zatürreeye tutulur. Sürgünlerde sarsılan vücudu hastalığı yenemez, hürriyetçiliğin, vatanseverliğin bu dev temsilcisi 2 Aralık 1888'de henüz kırk sekiz yaşındayken, hayata gözlerini yumar. Cenazesi Sakız'dan Balayır'a nakledilir ve vasiyeti gereğince, Rumeli'ye ilk geçen Türklerin Serdarı Süleyman Paşanın Türbesinin yanına gömülür.

EDEBİ KİŞİLİĞİ

Namık Kemal, Tanzimat Döneminde Avrupai Türk Edebiyatının nazımda ve nesirde en büyük isimlerinden biridir. Tanzimat Edebiyatının vatan ve hürriyet şairidir. Vatanseverliğe güç kazandıran şiirleri ve piyesleri, Osmanlı İmparatorluğu'na canlılık, hamlecilik, savaçılık yolunda yeni bir ruh kazandırır. Namık Kemal, vatan aşkıyla hayatını o kadar doldurmuştu ki bir gün kendisinden bir şiir rica eden kızı Feride'ye bile şu satırları yazdırmıştır:

Nine koynundan uyurken çıktım,
Geldim amma bu cihana, bıktım.
Bir kızım var sana kurban olsun,

Ođlumun merkodi al kan olsun,
Bu da Osmanlılara Őan olsun!
Sen gidersen yine gitmez Őerefın,
A vatan, Kâbe midir her tarafın?

Namık Kemal'in hürriyetçiliđi işleyen, hürriyet sesini getiren Őiirleri, dikta idaresine karşı, kiři hürriyetlerini, toplum hürriyetlerini savunur; ama bütün kiřilerin ayrı ayrı hür olmalarını ister. Türk milletinin, hür Türkler sayesinde dođrulacađına, yükseleceđine ve zafer kazanacađına inanır. Uzun yıllar sonra Atatürk, Namık Kemal'in bu ilkesi yolunda, Kurtuluř Savařı'nı, Türk milletinin egemenliđini temsil eden T.B.M.M sayesinde ve bu düzeni yaratan bir lider olarak kazanır.

ŐAİR NAMIK KEMAL

Namık Kemal ilk Őiirlerini çocuk denecek yařlarda yazmaya bařlar. İstanbul'a geldikten sonra eski ve yeni kuřaktan Őairlerin bir araya gelerek kurdukları Ercümen-i Őuara'ya ve kimi divan Őairlerine nazireler yazar. Őinasi'yle tanışıncaya deđin, Őiirlerinde tasavvuf etkileri görülür. Bu dönemde özellikle Yeniřehirli Avni, Leskofçalı Galib gibi Őairlerden etkilenmiřtir. Őinasi'yle tanışmasından sonra Őiirlerindeki içerik de deđiřir. Günlük konuřma dilinden alıntılarını yanı sıra, o zamana deđin geleneksel Türk Őiirinde görülmemiř olan "hürriyet kavgası", "esaret zinciri", "vatan", kalb-i millet" gibi yepyeni kavramlarla birlikte, dođrudan dođruya düřüncenin aktarılmasını amaçlayan bir tür "manzum nesir" oluřturur. Bosna Hersek Savařları, 93 Harbi gibi olayların yarattıđı sonuçlar, onun yazdıđı vatan Őiirlerini etkiler. Bu Őiirlerin en tanınmıřları arasında "Vaveyla", "Vatan Mersiyesi", "Vatan Őarkısı" ve "Hürriyet Kasidesi" yer alır.

Namık Kemal Őiirleriyle Őiir tekniđine büyük bir katkıda bulunmuř sayılmazsa da o günler için alışılmamıř diri bir sesle konuřmuř olması ve yapıtlarına kattıđı yeni kavramlarla Türk Őiirini Divan Őiirinin edilgen edasından kurtarır. Bütün bu nitelikler onun Vatan Őairi olarak anılmasına yol açar.

TİYATRO VE NAMIK KEMAL

Tiyatro türüne özellikle önem veren Namık Kemal, Hadika gazetesindeki bir yazısında tiyatronun her Őeyden evvel faydalı bir eđlence olduđunu söyler. Namık Kemal tiyatrolarında yaymak istediđi fikirleri ele alır. Oyunlarında Fransız romantiklerinin özellikle Victor Hugo'nun tesiri görülür. Tiyatroları, romantik tiyatronun bizdeki ilk örnekleridir. Namık Kemal'in altı tiyatro eseri vardır.

Vatan Yahut Silistre: Bir yurtseverlik ve kahramanlık oyunu olan bu eserini otuz üç yaşında kaleme alan Namık Kemal bu eserini 1873 yılında Güllü Agop Efendinin Gedikpaşa'daki tiyatrosunda temsil ettirir. 4 perdedir. Namık Kemal'in Türk sahnesi için yazdığı ilk tiyatro eseridir. Tiyatro tekniği bakımından zayıftır. Konu 2. Mahmut Devrindeki Şumnu Kuşatmasındaki bir olaydan alınır, 1854 Kırım Savaşı'ndaki Silistre Muharebesi'ne uygulanır. Eserde bir Osmanlı kahramanı olan İslam Beyin, kendisiyle beraber kıyafet savaşa katılan Zeynep isimli kızla, savaş sırasındaki vatanseverlikleri canlandırılır. Namık Kemal'in hayattayken oynanan tek piyesi budur. Bu eserin Namık Kemal'in hayatında büyük bir rolü vardır. Onun İstanbul'da küçük bir zümre arasında bile olsa uyandırdığı heyecan, şaire otuz sekiz aylık bir mahpusluk hayatına son vermesine sebep olur. Abdülaziz devrinin yenilik karşısındaki tereddütüne de son vermesine sebep olur. Bu eser yalnız ülke için değil, Avrupa'da da ilgi uyandırır ve beş dile çevrilir.

Zavallı Çocuk: Vatan Yahut Silistre ile aynı yıl yayınlanan bu eserde anneleri ve babaları tarafından sevemeyecekleri kimselerle evlenmek zorunda bırakılan gençlerin mutsuzlukları canlandırılır. Namık Kemal'in Magosa'da yazdığı bu eser üç perdeliktir.

Akif Bey: 1874'te yayımlana bu eserde vatan fikri hakimdir, daha olgun bir tiyatro ve insan anlayışıyla karşılaşılmaktadır. Namık Kemal'in Magosa'ya giderken tasarladığı, zindana konulduğu gece yazmaya başladığı beş perdelik bir eserdir. Eserde yurtsever bir deniz subayının göreve koştugu sırada karısının kendisine bağlılık göstermeyişini anlatırken, ahlaksal bir yorum da getirir.

Gülnihal: 1875'te yayımlanan bu eserde, zalim bir Sancak Beyine karşı halkın isyanı yansıtılır. Namık Kemal'in "Celaleddin Harzemşah" la beraber siyasi ve ahlaki kanaatlerini en çok anlattığı, en çok kendisi olduğu eserlerden bir sayılır. Piyenin asıl adı Raz-ı Dil'dir; fakat Namık Kemal sonradan hem birçok yerlerini kaldırmış, hem de adını değiştirmiştir. Namık Kemal'in tiyatroları içinde en başarılı olanı kabul edilir. Sahneler canlı, bağlantılar ustalıklıdır ve olaylar birbirine girmemiştir. Eserde Victor Hugo ve Şekspir'in tesiri vardır. Beş perde olan bu oyunun sahnelenmesinde pek çok bölüm sansür tarafından çıkarılır.

Celaleddin Harzemşah: Bu eser 1865'te yayınlanır. Eserin konusu, Harzemşahlar Devri Türk tarihinden alınır. On beş perdelik olan bu eser, daha çok, bir tarihi sahne romanı şeklinde yazılır. "Celal" Namık Kemal'in en şahsi eserlerindedir. İslam birliği düşüncesini kapsamlı bir biçimde sergiler. Namık Kemal bu oyunu oynanması için değil okunması için yazar.

Kara Bela: Konu Babür Sarayından alınmış bu eser Namık Kemal'in belki üslup ve kurtuluş itibariyle en zayıf eseridir.

ROMANCI NAMIK KEMAL

Namık Kemal'in ilk romanı olan İntibah 1876'da yayımlanır. Ruhsal çözümlerinin, bir olayı toplumsal ve bireysel yönleriyle görmeye çalışmasının yanı sıra, dış dünya betimlemesiyle de İntibah Türk romanında bir başlangıç sayılabilir. Eleştirmenler Namık Kemal'in bu romanda yüksek bir edebi düzey tutturamadığı görüşünde birleşirler. Dört yıl sonra yayımladığı Cezmi, tarihsel bir romandır. Kırım Şehzadesi Adil Günay'ın yaşadığı aşk ve Cezmi'nin onu kurtarmak isterken geçirdiği serüvenlerle gelişen romanda Namık Kemal'in tam anlamıyla Avrupa Romantizmi'nin etkisinde olduğu izlenir.

ELEŞTİRİLERİYLE NAMIK KEMAL

Namık Kemal romanı ve tiyatroyu toplumsal yaşama soktuğu gibi, edebiyat eleştirisini de Türkiye'ye ilk getiren kişilerden biri olur. En önemli eleştiri yapıtları Tahrir-i Harabat ile Takip'tir. Eleştirilerinde canlı, dolaysız bir üslup kullanılır. Tahrir-i Harabat Ziya Paşa'nın Harabat adlı güldestesinde karşı yazılmış sert bir eleştiri niteliğindedir. Takip de yine aynı güldestenin ikinci cildini eleştirir.

Mukaddeme-i Celal eleştirisinde Namık Kemal, Batı edebiyatıyla ile Doğu edebiyatını karşılaştırır, tiyatro roman türleri arasına durur.

TARİHÇİ NAMIK KEMAL

Namık Kemal'in ilgi duyduğu alanlardan birisi de tarihtir. Osmanlı İmparatorluğu'nun kuruluş ve yükseliş dönemlerini anlattığı Devr-i İstila yayımlandığında büyük ilgi görür. 1872'de çıkan Evrak-ı Perişan'da, Selahaddin Eyyubi, Fatih gibi tarihi kişilikleri, Barika-i Zafer'de İstanbul'un alınışını anlatır.

Ahmed Nafiz takma adıyla yayımladığı Silistre Muhasarası ve Kanije, yine Osmanlı tarihine ilişkin kahramanlık olaylarını ele alan kitaplardır. Namık Kemal'in tarih konusunda en kapsamlı çalışması olan Osmanlı Tarihi'nde, Hammer'in etkisinde

kaldığı, yapıtın bilimsel olmaktan çok, eğitici değer taşıdığı konusunda görüşler ileri sürülür. Yarım kalan bu yapıtın ilk basımı 2. Abdülhamid tarafından yasaklanır.

1975'te yayımlanan Büyük İslam Tarihi adlı yapıtındaysa Namık Kemal, İbn Haldun, İbn Rüşd gibi yazarlardan yararlanmış olduğunu belirtir.

GAZETECİ NAMIK KEMAL

Namık Kemal gazeteci olarak da Türk kültürü içinde önemli bir yer alır. Namık Kemal'in gazeteciliği bir zanaat değil, bir sanattır ve ilmi, hakiki, edebi manada gazetecilik... Döneminin hemen hemen bütün yenilik yanlısı ve ilerici gazetelerinde yazar. Siyasal ve toplumsal sorunlardan edebiyat, sanat, dil ve kültür konularına dek çok çeşitli alanlarda yazdığı makalelerin sayısı beş yüz kadardır. Bunlardan düzyazıdaki üstün yeteneğini ortaya koyduğu ve çok etkili bir üslup yarattığı kabul edilir.

Namık Kemal ilgili görüşler:

MUSTAFA KEMAL ATATÜRK: “Vatanın kurtuluşu ve istiklali için ölmeyi bugünkü nesle Namık Kemal öğretti.”

NİHAD SAMİ BANARLI: “Namık Kemal, vatan kelimesini, Türkçede ilk defa, bugünkü manasında kullanan şairdir. Milliyet ve hürriyet kelimelerini de Türkçeye o kazandırmıştır... Başlangıçta Divan kültürüyle yetişen Namık Kemal'in şiir lisanı... His, hayal, fikir, ahenk ve kültür bakımından zengin ve doyurucu bir lisandır... Namık Kemal, şiirde olduğu kadar, nesirde de kuvvetli bir üslupçudur... Bu üslup, fikrin heyecanla birleşmesinden doğan, haykırıcı bir üsluptur ve yazmaktan çok, insan topluluklarına işittirmek için ayarlanmış bir hitabet lisanı gibidir... Namık Kemal'de milliyet fikri, ilk bakışta bir İslam ve Osmanlı milliyetçiliği gibi görünür... Bununla beraber, eserleri dikkatle incelendiği zaman, kalbinde Türkler için çarpan tarafın ağır bastığı görülür. Ve anlaşılır ki, Namık Kemal'in İslam veya Osmanlı milliyetçiliği, imparatorluğun bütünlüğünü sarsmak için gözetilmiş bir tedbirdir ve aslı yine Türk milliyetçiliğidir... Namık Kemal'in hürriyet fikri... Fransız İhtilali'nin meydana koyduğu hürriyet anlayışının aynıdır... Namık Kemal, edebiyatımızın, Türk milletinin söylenen dili olması için çalışmış, halk Türkçesini anlayışla müdafaa etmiş, mühim bir kısım eserlerini, halkın konuştuğu dille yazmıştır... Bu güzel dilin, daha sade, daha tabii ve daha halk dili olabilmesi için mühim olarak şu çareleri düşünmüştür:

a) Türkçenin kaideleri mükemmel suretle tesbit ve tanzim edilmelidir,

b) Kelimelerimiz, halkın kullandığı kelimeler olmalıdır, c)Lisanı meydana getiren unsurlar, imla ve mana bakımından kuvvetle birleştirilmelidir, d) Dil, her bakımdan tabiileştirilmeli ve bu tabiiliğe engel olan külfetli sanatlardan uzaklaştırılmalıdır.”

TÜRKER ACAROĞLU: “Namık Kemal... Edebiyatın her türünden eser yazdı. Biçim ve ifade eski, ruh ve özce yenidir.”

ESERLERİ

ŞİİRLERİ: Şiirleri ilk defa “Namık Kemal, Hayatı ve Şiirleri” adını taşıyan bir kitapta toplanarak Saadettin Nüzhet Ergun tarafından yayınlamıştır.

OYUN: Vatan Yahut Silistre (1873, yeni harflerle 1940), Zavallı Çocuk (1873, yeni harflerle 1940), Akif Bey (1874, yeni harflerle 1958), Gülnihal (1875), Celaleddin Harzemşah (1885, yeni harflerle 1977), Kara Bela (1908)

ROMAN: İntibah (1876, yeni harflerle 1994), Cezmi (1880, yeni harflerle 1963)

ELEŞTİRİ: Tahrib-i Harabat (1865), Takip (1885), Renan Müdafaaanamesi (1908, yeni harflerle 1862’de Fuad Köprülü tarafından yayınlandı), İrfan Paşaya Mektup (1887), Mukaddeme-i Celal (1888)

TARİHİ KİTAPLAR: Devr-i İstila (1871), Barika-i Zafer (1872), Evrak-ı Perişan (1872, yeni harflerle 1973), Kanije (1874), Silistre Muhasarası (1874, yeni harflerle 1946), Osmanlı Tarihi (1889, ölümünden sonra, yeni harflerle 3 cilt, 1971-1974), Büyük İslam Tarihi (1975, ölümünden sonra).

EK 8:Namık KEMAL' in hayatı 8

Kaynak: <http://www.bilgicik.com/yazi/namik-kemal-biyografi-hayati-kim-kimdir/>

(Erişim Tarihi: 16/05/2013)

21 Aralık 1840'ta Tekirdağ'da doğdu, 2 Aralık 1888'de SakızAdası'nda öldü. Asıl adı Mehmed Kemal'dir, Namık adını ona şair Eşref Paşa vermiştir. Babası, II. Abdülhamid döneminde müneccimbaşılık yapmış olan Mustafa Asım Bey'dir. Annesini küçük yaşında yitirince çocukluğunu dedesi Abdüllâtif Paşa'nın yanında, Rumeli ve Anadolu'nun çeşitli kentlerinde geçirdi. Bu yüzden özel öğrenim gördü. Arapça ve Farsça öğrendi. 18 yaşlarında İstanbul'a babasının yanına döndü.

1863'te Babıali Tercüme Odası'na kâtip olarak girdi. Dört yıl çalıştığı bu görev sırasında dönemin önemli düşünürve sanatçılarıyla tanışma olanağı buldu. 1865'te kurulan ve daha sonra yeni Osmanlılar Cemiyeti adıyla ortaya çıkan İttifak-ı Hamiyet adlı gizli derneğe katıldı. Bir yandan da Tasvir-i Efkâr gazetesinde hükümeti eleştiren yazılar yazıyordu. Gazete, Yeni Osmanlılar Cemiyeti'nin görüşleri doğrultusunda yaptığı yayın sonucu 1867'de kapatıldı. Namık Kemal de İstanbul'dan uzaklaştırılmak için Erzurum'a vali muavini olarak atandı. Bu göreve gitmeyi çeşitli engeller çıkarıp erteledi ve Mustafa Fazıl Paşa'nın çağrısı üzerine Ziya Paşa'yla birlikte Paris'e kaçtı. Bir süre sonra Londra'ya geçerek M. Fazıl Paşa'nın parasal desteğiyle Ali Suavi'nin Yeni Osmanlılar adına çıkardığı Muhbir gazetesinde yazmaya başladı. Ama Ali Suavi'yle anlaşamaması üzerine Muhbir'den ayrıldı. 1868'de gene M. Fazıl Paşa'nın desteğiyle Hürriyet adı altında başka bir gazete çıkardı.

Çeşitli anlaşmazlıklarsonucu, Avrupa'da desteksiz kalınca, 1870'te zaptiye nazırı Hüsnü Paşa'nın çağrısı üzerine İstanbul'a döndü. Nuri, Reşat ve Ebüzziya Tevfik beylerle birlikte 1872'de İbret gazetesini kiraladı. Aynı yıl burada çıkan bir yazısı üzerine gazete hükümetçe dört ay süreyle kapatıldı. Namık Kemal gene İstanbul'dan uzaklaştırılmak için Gelibolu mutasarrıflığına atandı. Orada yazmaya başladığı Vatan Yahut Silistire oyunu, 1873'te Gedikpaşa Tiyatrosu'nda sahnelendiğinde halkı coşturup olaylara neden oldu. Bu haberi İbret gazetesinin yazması üzerine o sırada İstanbul'a dönmüş olan Namık Kemal birçok arkadaşıyla birlikte tutuklandı. Bu kez kalebentlikle Magosa'ya sürgüne gönderildi. 1876'da I. Meşrutiyet'in ilanından sonra İstanbul'a döndü. Şura-yı Devlet (Danıştay) üyesi oldu. Kanun-î Esasi'yi (Anayasa) hazırlayan

kurulda görev aldı. 1877 Osmanlı-Rus Savaşı çıkınca II. Abdülhamid Han'ın Meclis-i Mebusan'ı kapatması üzerine tutuklandı. Beş ay kadar tutuklu kaldıktan sonra Midilli Adası'na sürüldü. 1879'da Midilli mutasarrıfı oldu. Aynı görevle 1884'te Rodos, 1887'de Sakız Adası'na gönderildi. Ertesi yıl burada öldü ve Gelibolu'da Bolayır'da gömüldü.

Namık Kemal ilk şiirlerini çocuk denecek yaşlarda yazmaya başlamıştır. İstanbul'a geldikten sonra eski ve yeni kuşaktan şairlerin bir araya gelerek kurdukları Encümen-i Şuârâ'ya ve kimi Divan şairlerine nazireler yazmıştır. Şinasi'yle tanışmaya değin, şiirlerinde tasavvuf etkileri görülür. Bu dönemde özellikle Yenişehirli Avni, Leskofçalı Galib gibi şairlerden etkilenmiştir. Şinasi'yle tanışmasından sonra şiirlerindeki içerik de değişmiştir. Günlük konuşma dilinden alıntılarının yanı sıra, o zamana değin geleneksel Türk şiirinde görülmemiş olan "hürriyet kavgası", "esaret zinciri", "vatan", "kalb-i millet" gibi yepyeni kavramlarla birlikte, doğrudan doğruya düşüncenin aktarılmasını amaçlayan bir tür "manzum nesir" oluşturmuştur. Bosna-Hersek Savaşları, 93 Savaşı gibi olayların yarattığı sonuçlar, onun yazdığı vatan şiirlerini etkilemiştir. Bu şiirlerin en tanınmışları arasında "Vâveyla", "Vatan Mersiyesi", "Vatan Şarkısı" ve "Hürriyet Kasidesi" yer alır. Namık Kemal şiirleriyle şiir tekniğine büyük bir katkıdabulunmuş sayılmazsa da o günler için alışılmamış diri bir sesle konuşmuş olması ve yapıtlarına kattığı yeni kavramlarla Türk şiirini Divan şiirinin edilgen edasından kurtarmıştır. Bütün bu nitelikler onun Vatan Şairi olarak anılmasına yol açmıştır.

Tiyatro türüne özellikle önem veren Namık Kemal, altı oyun yazmıştır. Bir yurtseverlik ve kahramanlık oyunu olan Vatan Yahut Silistire yalnız ülke için değil, Avrupa'da da ilgi uyandırmış ve beş dile çevrilmiştir. Magosa'dayken yazdığı Gülnihal'de baskıya ve zulme karşı duyduğu tepkiyi dramatik bir biçimde dile getirmiştir. Oyunun sahnelenmesinde pek çok bölüm sansür tarafından çıkarılmıştır. Namık Kemal yine Magosa'da yazdığı Akif Bey'de, yurtsever bir deniz subayının göreve koştugu sırada karısının kendisine bağlılık göstermeyişini anlatırken, ahlaksal bir yorum da getirir. Zavallı Çocuk'ta görücü yoluyla evlenmeye karşı çıkar. On beş perdelik Celaleddin Harzemşah, Namık Kemal'in en beğendiği yapıtı olarak bilinir. Oyun, Moğollar'a karşı İslam dünyasını koruyan Celaleddin Harzemşah'ın kişiliği çevresinde gelişir. Bu eserde Namık Kemal, İslam birliği düşüncesini kapsamlı bir biçimde sergilemiştir. Namık

Kemal'in ilk romanı olan İntibah 1876'da yayımlanmıştır. Ruhsal çözümlerinin, bir olayı toplumsal ve bireysel yönleriyle görmeye çalışmasının yanı sıra, dış dünya betimlemeleriyle de İntibah Türk romanında birbaştlangıç sayılabilir. Eleştirmenler Namık Kemal'in bu romanda yüksek bir edebi düzey tutturamadığı görüşünde birleşirler. Dört yıl sonra yayımladığı Cezmi, tarihsel bir romandır. Kırım Şehzadesi Adil Giray'ın yaşadığı aşk ve Cezmi'nin onu kurtarmak isterken geçirdiği serüvenlerle gelişen romanda, Namık Kemal'in tam anlamıyla Avrupa Romantizmi'nin etkisinde olduğu izlenir.

Namık Kemal'in yaşamı boyunca ilgi duyduğu alanlardan birisi de tarihtir. Osmanlı İmparatorluğu'nun kuruluş ve yükseliş dönemlerini anlattığı Devr-i İstila yayımlandığında büyük ilgi görmüştür. 1872'de çıkan Evrak-ı Parişan'da, Selahaddin Eyyubi, Fatih gibi tarihi kişilikleri, Barika-i Zafer'de İstanbul'un alınışını anlatır. Ahmed Nâfiz takma adıyla yayımladığı Silistire Muhasarası ve Kanije, yine Osmanlı tarihine ilişkin kahramanlık olaylarını ele alan kitaplardır. Namık Kemal'in, tarih konusunda en kapsamlı çalışması olan Osmanlı Tarihi'nde, Hammer'in etkisinde kaldığı, yapıtın bilimsel olmaktan çok, eğitici değer taşıdığı konusunda görüşler ileri sürülmüştür. Yarım kalan bu yapıtın ilk basımı II. Abdülhamid tarafından yasaklanmıştır. 1975'te yayımlanan Büyük İslam Tarihi adlı yapıtındaysa Namık Kemal, İbn Haldun, İbn Rüşd gibi yazarlardan yararlanmış olduğunu belirtmiştir. Namık Kemal romanı ve tiyatroyu toplumsal yaşama soktuğu gibi, edebiyat eleştirisini de Türkiye'ye ilk getiren kişilerden biri olmuştur. En önemli eleştiri eserleri Tahrîb-i Harâbât ile Takîp'dir. Eleştirilerinde canlı, dolaysız bir üslup kullanmıştır. Tahrîb-i Harâbât, Ziya Paşa'nın Harâbât adlı güldestesine karşı yazılmış sert bir eleştiri niteliğindedir. Takîp de yine aynı güldestenin ikinci cildini eleştirir. Mukaddeme-i Celal eleştirisinde Namık Kemal, Batı edebiyatı ile Doğu edebiyatını karşılaştırmış, tiyatro, roman türleri üstünde durmuştur. Namık Kemal gazeteci olarak da Türk kültürü içinde önemli bir yer alır. Döneminin hemen hemen bütün yenilik yanlısı ve ilerici gazetelerinde yazmıştır. Siyasal ve toplumsal sorunlardan edebiyat, sanat, dil ve kültür konularına dek çok çeşitli alanlarda yazdığı makalelerin sayısı 500 kadardır. Bunlarda düzyazıdaki üstün yeteneğini ortaya koyduğu ve çok etkili bir üslup yarattığı kabul edilir.

ESERLERİ:

Oyun: Vatan Yahut Silistire, 1873 (yeni harflerle, 1940); Zavallı Çocuk, 1873 (yeni harflerle, 1940); Akif Bey, 1874 (yeni harflerle, 1958); Celaleddin Harzemşah, 1885 (yeni harflerle, 1977); Kara Belâ, 1908. Roman: İntibah, 1876 (yeni harflerle, 1944); Cezmi, 1880 (yeni harflerle, 1963). Eleştiri: Tahrib-i Harâbât, 1885; Takip, 1885; Renan Müdafaaanamesi, 1908 (yeni harflerle, 1962); İrfan Paşa'ya Mektup, 1887; Mukaddeme-i Celal, 1888. Tarihsel Yapıt: Devr-i İstila, 1871; Barika-i Zafer, 1872; Evrak-ı Perişan, 1872 (yeni harflerle, 1973); Kanije, 1874; Silistire Muhasarası, 1874 (yeni harflerle, 1946); Osmanlı Tarihi, (ö.s.), 1889 (yeni harflerle, 3 cilt, 1971-1974); Büyük İslam Tarihi, (ö.s.), 1975. Çeşitli: Rüya, 1893; Namık Kemal'in Mektupları, Ö.F. Akün (yay.), 1972.

HAKKINDA YAZILANLAR

1. Peçeye İsyan

Namık Kemal'in Torununun Anıları

(Unveiled)

Selma Ekrem

Anahtar Kitapları Yayınevi

Namık Kemal'i hepimiz tanırız. Tarihle ilgilenenler oğlu Ali Ekrem Bey'i de bilebilir. Oysa torunu Selma Ekrem'i yakın aile çevresi dışında bilen pek az çıkar. Üstelik bu insan, 1923 yılında, 21 yaşında bir genç kızken ABD'ye gitmek zorunda kalmış, 1986'da ölümüne kadar orada yaşamışsa. Bu kitabın yayınlanmasıyla birlikte onu herkes tanıyacak. Çünkü 1930 yılında

ABD'de ardarda dört baskı yapan ve büyük övgüler alan anıları tarihsel bir değer taşıyor. Zaman dilimi olarak 1902-1923 tarihleri arasını kapsayan anılar, çocuk ve genç kız gözüyle, çökmekte olan Osmanlı İmparatorluğu'na ilişkin gözlemlerini, tanıklıklarını yalın ve akıcı bir dille anlatıyor...

2. Boğaz'daki Aşiret

Mahmut Çetin

Edille Yayınları

“Boğaz'daki Aşiret” başlığı ister istemez “Boğaz Neresi” ve “Aşiret Kim” sorularını akla getiriyor. Evet Boğaz, bildiğimiz Boğaziçi. Genelde kırsal kesimle alakalı bir kavram olan aşiret kelimesi ise Boğaziçi'nde bir kast oluşturan büyükçe bir ailenin

tarihini anlatırken hassaten seçildi. Bir sülale tarihi diyebileceğimiz Boğaz'daki Aşiret yer yer Türk Solu tarihi, yer yer de

Batılılaşma Tarihi'nin belirli dönemlerini resmediyor. Aileler arasında evliliklerle kurulan bağların, sanata, ticarete, eğitime, bürokrasiye ve giderek bir yabancılaşma zihniyeti şeklinde hayata nasıl yansdığı eserdeki ipuçları yardımıyla daha iyi görülecektir zannediyoruz.

Boğaz'daki Aşiret, dört büyük ailenin birbirleriyle irtibatından oluşur. Eser bu sebeple dört bölüm olmuştur. Aile büyüklerinin asıl isimleri seçilerek de Konstantin'in Çocukları, Detrois'in Çocukları, Sotori'nin Çocukları, Topal Osman Paşa – Namık Kemal kanadı bölümleri ortaya çıktı. Boğaz'daki Aşiret! şenlikli bir kitap. Ali Fuat Cebesoy'dan Nazım Hikmet'e,

Oktay Rifat'tan Refik Erduran'a, Rasih Nuri İleri'den Ali Ekrem Bolayır'a, Zeki Baştımar'dan Sabahattin Ali'ye, Numan Menemencioğlu'ndan Abidin Dino'ya uzanan ilginç akrabalık zinciri.

Polonez, Hırvat, Alman, Macar ve Rum kökenli meşhurların, yerlilerle evliliklerinden oluşan "Boğaz'daki Aşiret" in, batılılaşma tarihinde oynadığı roller...

Kimlerin kimlikleri, Çıldır tan çizelgelerle soyağaçları. Ve dipnotlar! Onlar hiç bu kadar sevimli olmamışlardır.

EK 9:Namık KEMAL' in hayatı 9

Kaynak: <http://www.edebiyatfakultesi.com/namik-kemal.htm> (Erişim Tarihi: 16/05/2013)

Türk Edebiyatımızın önemli şair ve yazarlarından. Tekirdağ'da doğdu. Çocukluğu dedesinin ya-nda, Kars ve Sofya'da özel dersler alarak geçti. İstanbul'a dönünce Tercüme Odası'na memur oldu (1863). Şinasi ile Tasvir-i Efkâr'da yazılar yazdı. 1865'te aynı gazeteyi çıkardı. Yeni Osmanlılar Cemiyeti üyelerinin sü-rülme sebebiyle Ziya Paşa ile Paris'e kaçtı (1867). Londra'-da Ziya Paşa ile Hürriyet gazetesini çıkardı (1868). 1870'de İstanbul'a döndü, mutasarrıf olarak Gelibolu'ya gönderildi. Azledilirler tekrar İstanbul'a geldi. Namık Kemal 1877'de Midilli'ye sürüldü ve 1879'da oraya mutasarrıf oldu. Rodos (1884) ve Sakız (1887) adalarında da aynı görevi yaptı. Sakız'da öldü. Mezarı Bolayır'dadır. Namık Kemal, Tanzimat Edebiyatı'nın her türde eser veren gür sesli dâva adamı ve şâiridir. Sanatını toplumun hizmetine vermiş, şe-kil ve ifâde bakımından eski, öz ve ruh yönünden de yeni eserler vermiştir. Eserlerinde hürriyet, vatan, millete hiz-met, haksızlıkla savaş, adalet vb. fikirleri dile getirildi. Namık Kemal, dilde ve edebiyatta sadeleşmeyi savundu. Yeni nazım biçimleri denedi. Edebi-yatımızın batılılaşmasında rolü büyüktür. Şiirleri ilk defa Sadettin Nüzhet Ergun tarafından toplanıp "Namık Kemal- Hayâtı ve Şiirleri" adı ile yayımlandı (1933).

Piyesle-ri: 1. Vatan Yahut Silistre (1873), 2. Zavallı Çocuk (1873), 3. Akif Bey (1874), 4. Gülnihâl (1875), 5. Celâleddin Harzemşah (1885), 6. Karabelâ (1910).

Romanları: 1. İntibah (1876), 2. Cezmi (1880).

Tenkit eserleri: 1. Tahrîb-i Harabat (1885), 2. Tâkib (1885), 3. Renan Müdafaa-nâmesi (1962).

Tarih ve bi-yografi:1. Kaniye (1874), 2. Silistre Muhasarası (1874,1946), 3. Osmanlı Tarihi (Yeni baskısı: 1971,1974, üç cilt), 4. Büyük İslam Tarihi (Yeni baskı: 1975), 5. Evrak-ı Perişan (1871), 6. Yavuz Sultan Selim (1968 yeni baskı). Mektuplarını F. A. Tansel "Hususî Mektuplarına Göre Namık Kemal ve A. Ha-nı id'1 (1949.4 cilt) ile O. F. Akün "Namık Kemal'in Mektupla-rı" (1972) neşretti. Mehmet Kaplan'ın doktora tezinin konu-şu da "Namık Kemal Hayatı ve Eserleri" (1948)'dir. M. N Özön, "Namık Kemal ve İbret Gazetesi" (1938) adlı kitapta makalelerini topladı.

EK 10:Namık KEMAL' in hayatı 10

Kaynak: <http://1isinasli.blogspot.com/2013/02/namik-kemalin-hayati-kisaca-ozet.html>

(Erişim Tarihi: 16/05/2013)

"Vatan ve Hürriyet Şairi" olarak bilinen Namık Kemal, 1840 yılında Tekirdağ'da doğdu. Babası Mustafa Asım Bey, annesi ise Fatma Zehra Hanım'dır. Dedesi Abdülatif Paşa, eğitimiyle yakından ilgilendi. Çok küçük yaşlarda Arapça, Farsça ve Fransızca öğrendi.

Namık Kemal

1857'de Babıali Tercüme Odası'na girdi. Bu arada edebiyat çevrelerinde ismi duyulmaya başladı. Henüz 21 yaşında Encümen-i Şuara, yani Şairler Akademisi üyesi oldu.

Onun hayatındaki en önemli değişim, Şinasi ile tanışmasıydı. Şinasi'nin vasıtasıyla Tasvir-i Efkar'da yazılar yazdı. Bu sırada Batı edebiyatını daha yakından tanımaya başladı. Şinasi'nin 1865 yılında Fransa'ya gitmesi üzerine Tasvir-i Efkar'ı tek başına çıkarmaya başladı.

Kaleme aldığı bir yazı yüzünden gazete kapandı ve kendisi de Erzurum Valiliği'ne atandı. Fakat Namık Kemal, Erzurum'a gitmeyerek Ziya Paşa ile birlikte Paris'e kaçtı.

Yaklaşık üç yıl Avrupa'da kaldıktan sonra tekrar İstanbul'a döndü. Bu üç yıl içerisinde Batı edebiyatını daha yakından gözlemleme fırsatını buldu. İstanbul'a dönünce "İbret" adlı bir gazete çıkarmaya başladı. Yazdığı "Vatan Yahut Silistre" oyunundan dolayı Kıbrıs'taki Magosa'ya sürüldü. Birçok eserini, 38 ay süren bu sürgün yıllarında kaleme aldı.

1876'da af çıkınca İstanbul'a geldi. Yeni anayasa çalışmalarında "Şura-yı Devlet" üyesi oldu. 1879'da Midilli mutasarrıflığına tayin edildi.

Hayatının önemli bir bölümü sürgünlerle geçen Namık Kemal, 2 Aralık 1888 günü hayata gözlerini yumdu.

ÖZGEÇMİŞ

Adı Soyadı : Selçuk DÖVEN

Sürekli Adresi : Kazım Karabekir mahallesi Ozan caddesi No: 71 Ümraniye/İstanbul

Doğum Yeri ve Yılı : İstanbul/12.09.1983

Yabancı Dili : İngilizce

İlk Öğretim : Ahmet Yavuz İlköğretim Okulu-1997

Orta Öğretim : Ümraniye Teknik ve Endüstri Meslek Lisesi (Elektrik Bölümü)-2000

Lisans : Anadolu Üniversitesi (İşletme Fakültesi)-2011

Yüksek Lisans : Bahçeşehir Üniversitesi - Devam

Enstitü Adı : Fen Bilimleri Enstitüsü

Program Adı : Bilgi Teknolojileri Yüksek Lisans Programı

Yayımları :

Çalışma Hayatı : Yapı ve Kredi Bankası A.Ş (2005 - Devam)

