EDITOR
PROF. DR. NESLİHAN DOĞAN-SAĞLAMTİMUR

15TH INTERNATIONAL ISTANBUL SCIENTIFIC RESEARCH CONGRESS ON LIFE, ENGINEERING, AND APPLIED SCIENCES PROCEEDINGS BOOK

ISBN: 978-625-6879-40-9



PARTICIPANT COUNTRIES (14 COUNTRIES)
Türkiye, India, Iran, Algeria, Azerbaijan, Malaysia, Morocco, Bangladesh, Portugal, Nigeria, Saudi Arabia, Spain, Philippines.

ISBN: 978-625-6879-40-9

15th INTERNATIONAL İSTANBUL SCIENTIFIC RESEARCH CONGRESS ON LIFE, ENGINEERING, AND APPLIED SCIENCES

DECEMBER 09-11, 2023
ONLINE & IN-PERSON PARTICIPATION
ZOOM & İSTANBUL, TURKIYE

CONGRESS PROCEEDINGS BOOK

EDITORS

Prof. Dr. Neslihan DOGAN-SAGLAMTIMUR

BZT AKADEMİ YAYINEVİ®
TURKEY, GERMANY
TR: +90543 671 0123 GR: +491774586777
internationalistanbulcongress@gmail.com
https://www.internationalistanbulcongress.com/

All rights reserved
BZT AKADEMİ YAYINEVİ®
BZT ACADEMY PUBLISHING HOUSE
Publishing Date: 13.12.2023
ISBN: 978-625-6879-40-9

DAY (1) / GÜN (1) | HALL (4) / SALON (4)

Eğitim Hakları Salonu

MEETING ID: 851 1965 4118
PASSWORD: 970571

	09.12.2023	Time: 13^{30}-15^{15} (Turkey Local time GMT+3)	MODERATOR & HEAD OF SESSION: *Prof. Dr. Ahmet YÜCESAN* *CO-COORDINATOR* *Spc. Aydan RÜSTEMLİ*
	Authors	**Affiliation**	**Topic title**
13^{30}-13^{45}	Filiz KARDİYEN Elif BOZTÜRK KILINÇ	*Gazi University, Türkiye*	ANALYZING CUSTOMER COMMENTS OF A CHINESE RESTAURANT USING TEXT MINING TECHNIQUES AND MACHINE LEARNING METHODS
13^{45}-14^{00}	Buse ÖZDEMİR Necla GÜNDÜZ	*Gazi University, Türkiye*	FUZZY ROUGH SET THEORY FOR FEATURE SELECTION IN GENE DATA
14^{00}-14^{15}	Emine KARAKURT Aysel AYDIN KOCAEREN	*Çanakkale Onsekiz Mart University, Türkiye*	DEVELOPMENT OF POLYMERIC THIOPHENE BASED MEMBRANE FE 3+ SENSOR
14^{15}-14^{30}	Ahmet YÜCESAN Feryal DEMİR	*Süleyman Demirel University, Türkiye*	A STUDY ON CERTAIN SPACELIKE FRAMED HELIX TYPES
14^{30}-14^{45}	Ahmet YÜCESAN Feryal DEMİR	*Süleyman Demirel University, Türkiye*	ON THE SPACELIKE FRAMED RECTIFYING CURVES
14^{45}-15^{00}	Ahmet YÜCESAN Tunahan TURHAN Gözde ÖZKAN TÜKEL	*Süleyman Demirel University, Türkiye*	ON THE DARBOUX-TYPE ROTATION AXIS OF A CURVE IN A MYLLER CONFIGURATION
15^{00}-15^{15}	Ahmet YÜCESAN	*Süleyman Demirel University, Türkiye*	A CHARACTERIZATION OF SPACELIKE FRAMED CURVES

ISBN: 978-625-6879-40-9

CONTENTS

FARKLI GEOMETRİLERE SAHİP METAL ESASLI GIDA AMBALAJ MALZEMELERİNDE GERİLME VE DEFORMASYONUN VON MISES YÖNTEMİ VE ÜÇ BOYUTLU MODELLEME İLE ANALİZİ 1

Tuğba Güngör Ertuğral, Çağla Çam

REALIZATION OF CASTING PART DEMAND FORECAST AND ITS APPLICATION IN A CASTING FACTORY 7

Selinay KAYALI, Merve CENGİZ TOKLU

KARMA TAŞIYICI SİSTEMLERE SAHİP BİNALARIN YAPISAL OLARAK İNCELENMESİ 9

Zehra MERGEN

Doç. Dr. Abdullah SÖNMEZ

SPEED CONTROL OF PERMANENT MAGNET SYNCHRONOUS MOTORS USING FRACTIONAL ORDER SLIDING MODE CONTROLLER 34

Burak Çetin, Ali Saygın

CONCEPTUAL DESIGN AND MATHEMATICAL MODELLING OF A HYBRID VTOL AIRCRAFT 35

Muhammed Emin Ozturk, Arif Ankarali

TÜRKİYE'DE FAALİYET GÖSTEREN BİR OTELİN MÜŞTERİLERİ TARAFINDAN YAPILAN YORUMLARIN METİN MADENCİLİĞİ ve MAKİNE ÖĞRENMESİ KULLANILARAK İNCELENMESİ 37

Doç.Dr. Filiz KARDİYEN, Gizem ÖZMERİÇ

TRANSFORMATIONS IN URBAN SPACE BY NEOLIBERAL POLICIES: ANKARA TRAIN STATION .. 40

Helin Ceran, Zeynep Uludağ, Gülşah Güleç

ÇİN RESTORANINA YAPILAN MÜŞTERİ YORUMLARININ METİN MADENCİLİĞİ TEKNİKLERİ VE MAKİNE ÖĞRENMESİ YÖNTEMLERİYLE ANALİZ EDİLMESİ 42

Doç. Dr. Filiz KARDİYEN, Elif BOZTÜRK KILINÇ

FUZZY ROUGH SET THEORY FOR FEATURE SELECTION IN GENE DATA 45

Buse ÖZDEMİR, Necla GÜNDÜZ

DETERMINATION OF GENOTOXIC AND ANTIGENOTOXIC EFFECTS OF EXOPOLYSACCHARIDE (EPSJD2) OBTAINED FROM *PICHIA KUDRIAVZEVII* JD2 BY CHROMOSOMAL ABNORMALITY TEST IN HUMAN LYMPHOCYTES *IN VITRO* AT 48 HOURS OF TREATMENT 47

Sultan Yıldırım, Deniz Yüzbaşıoğlu, Zehranur Yüksekdağ, Fatma Ünal

INCORPORATION OF YR15 AND YR5 GENES IN WHEAT FOR RESISTANCE AGAINST YELLOW RUST 49

Zemran MUSTAFA, Muhammad AASIM

ISBN: 978-625-6879-40-9

FUZZY ROUGH SET THEORY FOR FEATURE SELECTION IN GENE DATA

Buse ÖZDEMİR[1], Necla GÜNDÜZ[2]

[1]Gazi Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik, Ankara, Türkiye.

[2]Gazi Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik, Ankara, Türkiye.

ORCID Code: https://orcid.org/0000-0003-2901-4206

ABSTRACT

Data mining methods, also known as knowledge discovery, knowledge mining and machine learning, are applied to large data sets. Big data is data that contains a lot of diversity and whose volume is rapidly increasing. Big data also refers to 3V (Volume, Velocity, Variety), that is, volume, velocity and variety. Examples of large data sets include determining customer behavior and product strategies in marketing; assisting with transportation, GPS navigation, traffic and weather alerts; in public administration, analysis of tax, defense and public health data; in business, streamlining management operations and optimizing costs; Regarding cyber security, detection and solutions of system vulnerabilities and cyber threats; In healthcare, examples include rapid access to medical records, accelerating treatment development, and identifying diseases and treatment methods with gene data. Gene data provides important information about cancers, tumors and genetic diseases (Giordani, 2009; Kasim et al., 2016; Vicente, 2009). However, in gene data analysis, which has a very large data structure and a wide variety of features, feature selection must be made first in order to reach more reliable and summary information, as in other large data groups. Genetic algorithms in data mining, information gain, symmetric uncertainty, Chi-square, An efficient feature selection algorithm (ERGS), improved feature selection based on effective range (IFSER) and Relief- F methods are some of the methods used in feature selection. In this study, it is aimed to compare the Relief-F and Chi-square algorithms, which are filtering methods for feature selection on gene data, and the Quick-Reduct algorithm, which is a fuzzy rough set method. Open source statistical program R-project v4.3.2 was used to analyze the data. FSelector was used for the Relief-F and Chi-square algorithms, and Rcpp and RoughSets packages were used for the Quick-Reduct algorithm.

Keywords: Machine learning, big data, gene data, feature selection, fuzzy rough set

Gen verilerinde özellik seçimi için bulanık kaba küme teorisi

ÖZET

Bilgi keşfi, bilgi madenciliği ve makine öğrenmesi gibi adlarla da anılan veri madenciliği yöntemleri büyük veri setlerine uygulanmaktadır. Büyük veri, fazla çeşitlilik içeren ve hacmi hızla artan verilerdir. Büyük veri aynı zamanda 3V (Volume, Velocity, Variety) yani hacim, hız ve çeşitliliği ifade eder. Büyük

ISBN: 978-625-6879-40-9

veri setlerine örnek olarak pazarlama konusunda, müşteri davranışı ve ürün stratejilerinin belirlenmesi; ulaşım konusunda, GPS navigasyonuna, trafik ve hava durumu uyarılarına yardımcı olunması; kamu yönetimi konusunda, vergi, savunma ve halk sağlığı verilerinin analizi; iş konusunda, yönetim operasyonlarının kolaylaştırılması ve maliyetlerin optimize edilmesi; siber güvenlik konusunda, sistem açıkları ve siber tehditlerin tespit edilmesi ve çözümleri; sağlık hizmetleri konusunda, tıbbi kayıtlara hızlıca erişim, tedavi gelişiminin hızlandırılması ve gen verileri ile hastalık ve tedavi yöntemlerinin belirlenmesi gibi örnekler verilebilir. Gen verileri, kanserler, tümörler ve genetik hastalıklar hakkında önemli bilgiler sunmaktadır (Giordani, 2009; Kasim vd., 2016; Vicente, 2009). Ancak çok büyük bir veri yapısına ve çok çeşitli özelliklere sahip olan gen veri analizlerinde, diğer büyük veri gruplarında olduğu gibi, daha güvenilir daha özet bilgilere ulaşabilmek amacıyla, öncelikle özellik seçimi yapılmalıdır. Veri madenciliği konusunda genetik algoritmalar, bilgi kazancı, simetrik belirsizlik, Ki-kare, etkili özellik seçim algoritması(An efficient feature selection algorithm, ERGS), etkili aralığa dayalı geliştirilmiş özellik seçimi (improved feature selection based on effective range, IFSER) ve Relief-F yöntemleri özellik seçiminde kullanılan yöntemlerden bazılarıdır. Bu çalışmada, gen verileri üzerinde özellik seçimi için filtreleme yöntemlerinden Relief-F ve Ki-kare algoritmaları ile bulanık kaba küme yönteminden Quick-Reduct algoritmasının karşılaştırılması amaçlanmıştır. Verilerin analiz edilmesinde açık kaynak kodlu istatistik programı R-project v4.3.2 kullanılmıştır. Relief-F ve Ki-kare algoritmaları için FSelector, Quick-Reduct algoritması için Rcpp ve RoughSets paketlerinden yararlanılmıştır.

Anahtar Kelimeler: Makine öğrenmesi, büyük veri, gen verileri, özellik seçimi, bulanık kaba küme