conferenceseries.com
1150th Conference

September 2017 I Volume 6 I Issue 5
ISSN: 2324-9307
Journal of
Computer Engineering & Information Technology

Proceedings of

4th International Conference on

BIG Data ANALYSIS AND Data MINING

September 07-08, 2017 Paris, France



Conference Series - America
One Commerce Center-1201, Orange St. #600, Wilmington, Zip 19899, Delaware, USA
P: +1-702-508-5200, F: +1-650-618-1417

Conference Series - UK
Kemp House, 152 City Road, London EC1V 2NX, UK
Toll Free: +1-800-014-8923

Scientific Program

Data Mining 2017

Day 1 Septmber 07, 2017

8:00-08:45 Registrations

Sunset -1

conferenceseries.com **08:45-09:00** **Opening Ceremony**

Keynote Forum

09:00-09:30
Title: Big data, data hubs, sensors, iot and precision agriculture – Their effect on data collection, analysis, forecasting and survey processes
Dr. Michael Valivullah, NASS, US Department of Agriculture, USA

09:30-10:00
Title: Big data analysis and approximations in information systems
En-Bing Lin, Central Michigan University, USA

10:00-10:30
Title: Maintenance of engineering systems by big data
Petra Perne, Institute of Computer Vision and Applied Computer Sciences, Germany

Group Photo 10:30-10:35

Networking and Refreshments Break 10:35-10:50 @ Foyer

10:50-11:20
Title: Data Mining in Big Data Analytics: Exploiting Resolution Scale, Addressing Bias, Having Analytical Focus
Fionn Murtagh, University of Huddersfield, UK

Plenary Session

11:20-11:50
Title: Big data – Big challenge for public administration – experiences of pilot project in ministry of public administration of republic slovenia
Karmen Kern Pipan, Ministry of Public Administration, Republic Slovenia

11:50-12:20
Title: A complementary square-error clustering criterion and initialization of k-means
Boris Mirkin, National Research University Higher School of Economics, Russia

12:20-12:50
Title: Process before-after order data verification/prediction systems based on paraconsistent annotated logic program Bf-EVALPSN
Kazumi Nakamatsu, University of Hyogo, Japan

Lunch Break 12:50-13:30 @ NCafé

13:30-14:00
Title: Smart data acquisition and analysis for deep learning based systems
Pavel kisilev, Huawei Research, Israel

14:00-14:30
Title: Predicting corporate misconduct
Bennett B Borden, Drinker Biddle & Reath, USA

14:30-15:00
Title: Exact data mining from inexact data
Nikolaos Freris, New York University Abu Dhabi, UAE

Sessions: Data Mining Applications in Science, Engineering, Healthcare and Medicine | Data Mining Methods and Algorithms | Artificial Intelligence | Big Data Applications | Big Data Algorithm | Data Privacy and Ethics | Data Mining Analysis | Business Analytics | Optimization and Big Data | New visualization techniques | Clustering

Session Chair: Michael Valivullah, CTO, NASS, USA
Session Co-chair: Petra Perne, Institute of Computer Vision and applied Computer Sciences, Germany

Session Introduction

15:00-15:20
Title: Scalable robust clustering method for large and sparse data
Tommi Kärkkäinen, University of Jyväskylä, Finland

15:20-15:40
Title: Data is not wrong, but your use of it might be!
Fred Jacquet, Evangelist, France

15:40-16:00	**Title: An automatic data prediction system with business applications** **Morgan C. Wang,** University of Central Florida, USA
	Networking and Refreshments Break 16:00-16:15 @ Foyer
16:15-16:35	**Title: A novel linear-time and memory saving approach for visual exploratory data analysis: data embedding and graph visualization** **Witold Dzwinel,** AGH University of Science and Technology, Poland
16:35-16:55	**Title: Multi-objective optimization via visualization** **Nicolas Wesner,** Mazars Actuariat, France
16:55-17:15	**Title: Multi-criteria Decision-making for purchasing cell phones using machine learning approach** **S N Mohanty,** KIIT University, India
17:15-17:35	**Title: Big data - applied in healthcare context, the case study of diabetes** **Jenny Lundberg,** Linnaeus University, Sweden
	Young Researchers Forum
17:35-17:55	**YRF: Cost-Sensitive deep neural networks to address class imbalanced problem** **Kruy Seng,** Lingnan University, Hong Kong
17:55-18:15	**YRF: Data driven intelligent analytics in education domain** **Prajakta Diwanji,** FHNW University of Applied Sciences and Arts, Switzerland
18:15-18:35	**YRF: The domus-lig experimental living-lab methodology and the emoz platform: overview of spontaneous and ecological corpora - from small smart data to big data** **SASA Yuko,** University Grenoble Alps, France
	Panel Discussion

Day 2 Septmber 08, 2017

Sunset -1

Keynote Forum

	Introduction
09:00-09:30	**Title:Extensions of dynamic programming: Applications for decision trees** **Mikhail Moshkov,** King Abdullah University of Science and Technology, Saudi Arabia
09:30-10:00	**Title: Disrupting retail analysis with patent applied wi-fi based advanced analytics solution - BonAir** **Ergi Sener,** Bonbon Tech, Turkey
10:00-10:30	**Title: Effective choice and ranking of alternatives in search and recommendation problems** **Fuad Aleskerov,** National Research University Higher School of Economics, Russia
	Networking and Refreshments Break 10:30-10:50 @ Foyer
10:50-11:20	**Title: Data enabled approaches to sensitivity analysis, calibration and risk visualization in general circulation models** **Omar M. Knio,** King Adullah University of Science and Technology, Saudi Arabia
	Plenary Session
11:20-11:50	**Title: Big data winners and losers: The ethics of algorithms** **Bennett B. Borden,** Drinker Biddle & Reath, USA
11:50-12:20	**Title: SQL on big data – Technology, architecture and innovations** **Sumit Pal,** Independent Consultant , USA
12:20-12:50	**Title: Regulatory learning: What framework can be proposed to ensure that the financial environment can be controlled?** **Bertrand Hassani,** Université Paris 1, France
	Lunch Break 12:50-13:30 @ NCafé

13:30-14:00	**Title: Two applications of variable neighborhood search in data mining** **Pierre Hansen,** GERAD and HEC Montréal, Canada
14:00-14:30	**Title: From knowledge to wealth: Intellectual property and technology licensing** **Parameshwar P Iyer,** Indian Institute of Science, India
14:30-15:00	**Title: Big data analytics and visualization for healthcare – A glance into the AEGLE platform features via its presentation layer** **Nada Philip,** Kingston University London, UK
15:00-15:30	**Title: Computerizing mathematical text, debugging proof assistants and object versus metal level** **Fairouz Kamareddine,** Heriot-Watt University, UK

Data Mining Applications in Science, Engineering, Healthcare and Medicine | Data Mining Methods and Algorithms | Data Mining Tools and Software | Big Data Applications | Data Mining Tasks and Processes | Data Privacy and Ethics | Big Data Technologies | Social network analysis | Business Analytics | Search and data mining | Clustering

Session Chair: Fionn Murtagh, University of Huddersfield, UK
Session Co-chair: Narasimha Murty Yedidi, Electronic Arts, USA

Session Introduction

15:30-15:50	**Title: An innovation framework for big data** **Knut Hinkelmann,** FHNW University of Applied Sciences and Arts, Switzerland
15:50-16:10	**Title: Prune the inputs, increase data volume, or select a different classification method – A strategy to improve accuracy of classification** **Alla Sapronova,** Uni Research Computing, Norway
16:10-16:30	**Title: Business analytics applications in budget modelling to improve network performance** **Eleni Rozaki,** Institute of Technology Tallaght, Ireland

Networking and Refreshments Break 16:30-16:45 @ Foyer

16:45-17:05	**Title: Powerful and novel multivariate statistical approaches in big data sets and in data mining with applications to bio-, medical-, and material-Informatics** **Derrick K. Rollins,** Iowa State University, USA
17:05-17:25	**Title: Contextual exploration (EC3): A strategy for the detection, extraction and visualization of target data** **Christophe Jouis,** LIP6 (CNRS -UPMC Sorbonne Université), France
17:25-17:45	**Title: Objective function-based clustering via near-Boolean optimization** **Giovanni Rossi,** University of Bologna, Italy
17:45-18:05	**Title: Data mining via entropy: Application on trade database** **Abdul Basit,** State Bank of Pakistan, Pakistan

Young Researchers Forum

18:05-18:25	**YRF: Implementation and performance comparison of wilson, holdout, and multiedit algorithm for edited k-nearest neighbor** **Febriana Misdianti,** University of Manchester School of Computer Science, UK
18:25-18:45	**YRF: Deep learning classification for large behavioral data** **Sofie De Cnudde,** Antwerp University, Belgium

Video Presentation

18:45-19:00	**Title: ADaGe – Automated dataset generator** **Joseph Bonello,** University of Malta, Malta
19:00-19:15	**Title: Data mining applied to robotics** **Jose Manuel Lopez-Guede,** University of the Basque Country, Spain

Panel Discussion

Awards & Closing Ceremony

List of Open Access Journals

Agri, Food & Aqua

Advances in Crop Science and Technology	2329-8863
Advances in Dairy Research	2329-888X
Agrotechnology	2168-9881
Aquaculture Research & Development	2155-9546
Arabidopsis C. Elegans and Zebrafish	-
Biofertilizers & Biopesticides	2155-6202
Crop Research	2454-1761
Experimental Food Chemistry	-
Fisheries & Livestock Production	2332-2608
Fisheries and Aquaculture Journal	2150-3508
Fisheriessciences	1307-234X
Food & Industrial Microbiology	-
Food & Nutritional Disorders	2324-9323
Food Processing & Technology	2157-7110
Food: Microbiology, Safety & Hygiene	-
Forest Research	2168-9776
Horticulture	2376-0354
International Biodiversity, Bioprospecting and Development	2376-0214
Marine Science: Research & Development	2155-9910
Medicinal & Aromatic Plants	2167-0412
Nutrition & Food Sciences	2155-9600
Plant Pathology & Microbiology	2157-7471
Poultry, Fisheries & Wildlife Sciences	2375-446X
Probiotics & Health	2329-8901
Research & Reviews: Journal of Agriculture and Allied Sciences	2347-226X
Research & Reviews: Journal of Food and Dairy Technology	2321-6204
Rice Research	2375-4338
Traditional Medicine and Clinical Naturopathy (Homeopathy & Ayurvedic Medicine-2167-1206)	-

Business & Management

Accounting & Marketing	2168-9601
Arabian Journal of Business and Management Review	2223-5833
Business & Financial Affairs	2167-0234
Business & Hotel Management	2324-9129
Business and Economics Journal	2151-6219
Defense Studies & Resource Management	2324-9314
Entrepreneurship & Organization Management	2169-026X
Global Economics	2375-4389
Hotel & Business Management	2169-0286
International Journal of Accounting Research	-
International Journal of Economics and Management Science	2162-6359
Internet Banking & Commerce	1204-5357
Review of Public Administration and Management	2315-7844
Stock & Forex Trading	2168-9458
Tourism & Hospitality	2167-0269

Chemical Engineering

Advanced Chemical Engineering	2090-4568
Bioprocessing & Biotechniques	2155-9821
Chemical Engineering & Process Technology	2157-7048
Thermodynamics & Catalysis	2157-7544

Chemistry

Analytical & Bioanalytical Techniques	2155-9872
Analytical & Electrochemical Insights	-
Bioenergetics: Open Access	2167-7662
Chemical Informatics	-
Chemical Sciences Journal	2150-3494
Chromatography & Separation Techniques	2157-7064
Clinical & Medical Biochemistry: Open Access	-
Clinical Chemistry: Open Access	-
Environmental & Analytical Toxicology	2161-0525
Environmental Analytical Chemistry	-
Glycobiology	2168-958X
Herbal Medicine: Open Access	-
Immuno Chemistry: Open Access	-
Industrial Chemistry: Open Access	-
International Journal of Applied Biology and Pharmaceutical Technology	0976-4550
International Journal of Drug Development & Research	0975-9344
Mass Spectrometry: Open Access	-
Medicinal Chemistry	2161-0444
Modern Chemistry & Applications	2329-6798
Natural Products Chemistry & Research Journal	2329-6836
Neuro Chemistry: Open Access	-
Organic & Inorganic Chemistry	-
Organic Chemistry: Current Research	2161-0401
Pharmaceutical Analytical Chemistry: Open Access	-
Physical Chemistry & Biophysics	2161-0398
RROIJ: Medicinal Chemistry	-
Structural Chemsitry & Crystallography Communication	-
Trends in Green Chemistry	-
Vitamins & Minerals	2376-1318

Clinical

Ageing Science	2329-8847
Ancient Diseases & Preventive Remedies	2329-8731
Anesthesia & Clinical Research	2155-6148
Annals of Clinical and Laboratory Research	2386-5180
Arrhythmia: Open Access	-
Atherosclerosis: Open Access	-
Cell Biology: Research & Therapy	2324-9293
Cellular & Molecular Pathology	-
Clinical & Experimental Cardiology	2155-9880
Clinical & Experimental Dermatology Research	2155-9554
Clinical & Experimental Nephrology	-
Clinical & Experimental Oncology	2324-9110
Clinical & Experimental Ophthalmology	2155-9570
Clinical & Experimental Orthopaedics	-
Clinical & Experimental Pathology	2161-0681
Clinical & Molecular Endocrinology	-
Clinical and Experimental Psychology	-
Clinical and Experimental Transplantation	-
Clinical Case Reports	2165-7920
Clinical Depression	-
Clinical Dermatology Research Journal	-
Clinical Diabetes & Practice	-
Clinical Nutrition & Dietetics	-
Clinical Oncology and Practice	-
Clinical Pediatrics	-
Clinical Pediatrics & Dermatology	-
Clinical Psychiatry	-
Clinical Research & Bioethics	2155-9627
Clinical Research On Foot & Ankle	2329-910X
Clinical Respiratory: Open Access	-
Clinical Toxicology	2161-0495
Clinical Trials	2167-0870
Clinics in Mother and Child Health	2090-7214
Cosmetology & Orofacial Surgery	-
Cosmetology & Trichology	-
Dermatitis	-
Diabetes Case Reports	-
Dialysis and Clinical Practice	-
Drug Intoxication & Detoxification : Novel Approaches	2327-4557
Dual Diagnosis: Open Access	-
Eye & Cataract Refractive Surgery	-
Forensic Toxicology & Pharmacology	2325-9841
Glaucoma: Open Access	-
HIV & Retro Virus	-
Immunooncology	-
Insights in Pediatric Cardiology	-

Intensive and Critical Care	-
International Journal of Anesthesiology & Pain Medicine	-
International Journal of Cardiovascular Research	2324-8602
International Journal of Digestive Diseases	-
International Journal of Ophthalmic Pathology	2324-8599
Interventional Cardiology: Open Access	-
JBR Journal of Clinical Diagnosis and Research	2376-0311
Optometry: Open Access	-
Phonetics & Audiology	-
Speech Pathology & Therapy	-
Stem Cell Research & Therapy	2157-7633
Toxicology: Open Access	-
Vasculitis	-

Earth & Environmental Sciences

Advances in Recycling	-
Astrobiology & Outreach	2332-2519
Biodiversity & Endangered Species	2332-2543
Biodiversity Management & Forestry	2327-4417
Bioremediation & Biodegradation	2155-6199
Biosafety	2167-0331
Climatology & Weather Forecasting	2332-2594
Coastal Zone Management	-
Earth Science & Climatic Change	2157-7617
Ecosystem & Ecography	2157-7625
Entomology, Ornithology & Herpetology	2161-0983
Expert Opinion On Environmental Biology	2325-9655
Fundamentals of Renewable Energy and Applications	2090-4541
Geography & Natural Disasters	2167-0587
Geoinformatics & Geostatistics: An Overview	2327-4581
Geology & Geosciences	2329-6577
Geophysics & Remote Sensing	2169-0049
Hydrogeology & Hydrologic Engineering	2325-9647
Hydrology: Current Research	2157-7587
Industrial Pollution Control	-
Innovative Energy Policies	2090-5009
International Journal of Evolution	2324-8548
International Journal of Waste Resources	2252-5211
Marine Biology & Oceanography	2324-8661
Oceanography: Open Access	2332-2632
Oil & Gas: Open Access	-
Petroleum & Environmental Engineering	2157-7463
Plant Physiology & Pathology	2329-955X
Pollution Effects & Control	2375-4397
Research & Reviews: Journal of Ecology and Environmental Sciences	-

EEE

Electrical & Electronic Systems	2332-0796
Electrical Engineering & Electronic Technology	2325-9833

Engineering

Advances in Automobile Engineering	2167-7670
Advances in Robotics & Automation	2168-9695
Aeronautics & Aerospace Engineering	2168-9792
Applied Bioinformatics & Computational Biology	2329-9533
Applied Mechanical Engineering	2168-9873
Architectural Engineering Technology	2168-9717
Automatic Control of Physiological State and Function	2090-5092
Biochips & Tissue Chips	2153-0777
Bioengineering & Biomedical Science	2155-9538
Biomusical Engineering	2090-2719
Biosensors & Bioelectronics	2155-6210
Biosensors Journal	2090-4967
Civil & Environmental Engineering	2165-784X
Computer Engineering & Information Technology	2324-9307
Computer Engineering and Information Technology	2324-9307
Defense Management	2167-0374
Fashion Technology & Textile Engineering	2329-9568
Global Journal of Technology and Optimization	2229-8711
Global Research in Computer Science	2229-371X
Industrial Engineering & Management	2169-0316
Information Technology & Software Engineering	2165-7866
International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering	2278-8875
International Journal of Advancements in Technology	0976-4860
International Journal of Biomedical Data Mining	2090-4924
International Journal of Innovative Research in Computer and Communication Engineering	2278-1021
International Journal of Innovative Research in Science, Engineering and Technology	2319-8753
International Journal of Sensor Networks and Data Communications	2090-4886
International Journal of Swarm Intelligence and Evolutionary Computation	2090-4908
Irrigation & Drainage Systems Engineering	2168-9768
Lasers, Optics & Photonics	-
Lovotics	2090-9888
Membrane Science & Technology	2155-9589
Molecular Imaging & Dynamics	2155-9937
Nuclear Energy Science & Power Generation Technology	2325-9809
Research & Reviews: Journal of Engineering and Technology	2319-9873
Steel Structures & Construction	-
Telecommunications System & Management	2167-0919
Textile Science & Engineering	2165-8064

General Science

Computer Science & Systems Biology Journal	0974-7230
Ergonomics	2165-7556
Research and Development	-
International Journal of Advance Innovations, Thoughts & Ideas	2277-1891
Metrology	-
Research & Reviews: Journal of Botanical Sciences	2320-0189
Research & Reviews: Journal of Chemistry	2319-9849
Tomography	-

Genetics & Molecular Biology

Advanced Techniques in Biology & Medicine	2379-1764
Advancements in Genetic Engineering	2169-0111
Advances in Molecular Diagnostics	-
Biochemistry & Analytical Biochemistry	2161-1009
Biochemistry & Molecular Biology Journal	-
Biochemistry & Physiology	2329-9029
Biological Systems	2329-6577
Biotechnology & Biomaterials	2155-952X
Bipolar Disorder: Open Access	-
Cell & Developmental Biology	2168-9296
Cell Science & Therapy	2157-7013
Cell Signaling	-
Cellular & Molecular Medicine: Open Access	-
Chemical Biology & Therapeutics	-
Clinical Epigenetics	-
Cloning & Transgenesis	2168-9849
Current Synthetic and Systems Biology	2332-0737
Cytology & Histology	2157-7099
Down Syndrome & Chromosome Abnormalities	-
Electronic Journal of Biology	-
Enzyme Engineering	2329-6674
Fertilization: in Vitro	2375-4508
Fungal Genomics & Biology	2165-8056
Gene Technology	2329-6682
Genetic Syndromes & Gene Therapy	2157-7412
Hereditary Genetics: Current Research	2161-1041
Human Genetics & Embryology	2161-0436
Insights in Cell Science	-
Insights in Stem Cells	-
International Journal of Genomic Medicine	2332-0672
Metabolomics: Open Access	2153-0769
Metabonomics & Metabolites	2325-9736
Microbial & Biochemical Technology	1948-5948
Microbial Methods & Assays Open Access	-
Molecular and Genetic Medicine	1747-0862
Molecular Biology	2168-9547
Molecular Biomarkers & Diagnosis	2155-9929
Molecular Cloning & Genetic Recombination	2325-9787
Nanomedicine & Biotherapeutic Discovery	2155-983X
Next Generation: Sequencing & Applications	-
Phylogenetics & Evolutionary Biology	2329-9002

Physiobiochemical Metabolism	2324-8793
Plant Biochemistry & Physiology	2329-9029
Proteomics & Enzymology	-
Single Cell Biology	2168-9431
Tissue Science & Engineering	2157-7552
Transcriptomics: Open Access	2329-8936
Translational Biomedicine	2172-0479

Health Care

Diversity and Equality and Health and Care	2049-5471
Health Care: Current Reviews	2375-4273
Health Science Journal	1791-809X
Pregnancy & Child Health	2376-127X
Primary Health Care	2167-1079
Quality in Primary Care	1479-1072
Tropical Diseases & Public Health	2329-891X
Women'S Health, Issues & Care	2325-9795

Immunology

Advances in Antibiotics & Antibodies	-
Allergy & Therapy	2155-6121
Autoimmune Diseases: Open Access	-
Clinical & Cellular Immunology	2155-9899
Cytokine Biology	-
Immunobiology	-
Immunogenetics: Open Access	-
Immunome Research	1745-7580
Immunotherapy: Open Access	-
Infectious Diseases & Immunological Techniques	2325-9752
Inflammatory Bowel Diseases & Disorders	-
Innate Immunity & Immunological Disorders	-
Interdisciplinary Journal of Microinflammation	-
Lupus: Open Access	-
Molecular Immunology	-
Osteoarthritis	-
Reproductive Immunology	-
Rheumatology: Current Research	2161-1149
Sarcoidosis	-
Vaccines & Vaccination	2157-7560

Informatics

Data Mining in Genomics & Proteomics	2153-0602
Glycomics and Lipidomics	2153-0637
Health & Medical Informatics	2157-7420
Proteomics & Bioinformatics	0974-276X
Theoretical and Computational Science	2376-130X

Materials Science

Bioceramics Developments and Applications	2090-5025
Material Sciences & Engineering	2169-0022
Nano Research & Applications	-
Nanomaterials & Molecular Nanotechnology	2324-8777
Nanomedicine & Nanotechnology	2157-7439
Plastic & Polymer Sciences	-
Powder Metallurgy & Mining	2168-9806
Research & Reviews: Journal of Material Sciences	2321-6212

Mathematics

Applied & Computational Mathematics	2168-9679
Biometrics & Biostatistics	2155-6180
Generalized Lie Theory and Applications	1736-4337
Physical Mathematics	2090-0902
Research & Reviews: Journal of Statistics and Mathematical Sciences	-

Medical

Abnormal and Behavioural Psychology	-
Acta Psychopathologica	-
Acta Rheumatologica	-
Addictive Behaviors , Therapy & Rehabilitation	2324-9005
Adenocarcinoma	-
Advances in Cancer Prevention	-
Advances in Genetic Engineering & Biotechnology	-
Advances in Weight Loss Management & Medical Devices	-
Aerobics & Fitness	-
Aesthetic & Reconstructive Surgery	-
Aids & Clinical Research	2155-6113
Air and Water Borne Diseases	2167-7719
Alternative & Integrative Medicine	2327-5162
Analgesia & Resuscitation : Current Research	2324-903X
Anaplastology	2161-1173
Anatomy & Physiology: Current Research	2161-0940
Andrology & Gynecology: Current Research	2327-4360
Andrology	2167-0250
Angiology: Open Access	2329-9495
Annals of Behavioural Science	-
Applied and Rehabilitation Psychology: Open Access	-
Archives in Cancer Research	2254-6081
Archives of Medicine	1989-5216
Archives of Surgical Oncology	-
Archivos De Medicina	1698-9465
Arthritis	2167-7921
Asthma and Bronchitis	-
Athletic Enhancement	2324-9080
Autacoids & Hormones	2161-0479
Biology and Medicine	0974-8369
Biomedical Engineering & Medical Devices	-
Biomedical Sciences	2254-609X
Bioterrorism & Biodefense	2157-2526
Blood	-
Blood & Lymph	2165-7831
Blood Disorders & Transfusion	2155-9864
Blood Pressure: Open Access	-
Bone Marrow Research	2329-8820
Bone Reports & Recommendations	-
Brain Tumors	-
Breast Cancer: Current Research	-
Cancer Biomarkers	-
Cancer Clinical Trials	-
Cancer Diagnosis	-
Cancer Medicine & Anticancer Drugs	-
Cancer Science & Therapy	1948-5956
Cancer Surgery	-
Carcinogenesis & Mutagenesis	2157-2518
Cardiovascular Diseases & Diagnosis	2329-9517
Cardiovascular Pathology: Open Access	-
Celiac Disease: Open Access	-
Cervical Cancer: Open Access	-
Chemotherapy	2167-7700
Chest Diseases	-
Childhood & Developmental Disorders	-
Childhood Obesity	-
Chronic Obstructive Pulmonary Disease: Open Access	-
Colorectal Cancer: Open Access	-
Communication Disorders, Deaf Studies & Hearing Aids	2375-4427
Community Medicine & Health Education	2161-0711
Complex Diseases and Treatment	-
Contraceptive Studies	-
Critical Care Obstetrics & Gynecology	-
Current Trends in Gynecologic Oncology	-
Dental Health: Current Research	-
Dental Implants and Dentures: Open Access	-
Dentistry	2161-1122
Depression and Anxiety	2167-1044
Dermatology Case Reports	-
Diabetes & Metabolism	2155-6156
Diabetes Medication and Care	-
Diabetic Complications and Medicine	-
Drug Abuse	-
Emergency Medicine	2165-7548
Endocrinology & Diabetes Research	-
Endocrinology & Metabolic Syndrome	2161-1017
Epidemiology: Open Access	2161-1165
Evidence based Medicine and Practice	-
Family Medicine & Medical Science Research	2327-4972
Forensic Biomechanics	2090-2697
Forensic Medicine	-

Forensic Nursing: Open Access	-
Forensic Odontology	-
Forensic Psychology	-
Forensic Research	2157-7145
Gastrointestinal & Digestive System	2161-069x
Gastrointestinal Cancer and Stromal Tumors	-
General Medicine	2327-5146
General Practice	2329-9126
Genetic Disorders & Genetic Reports	2327-5790
Genital System & Disorders	2325-9728
Geriatric Psychiatry	-
Gerontology & Geriatric Research	2167-7182
Gynecology & Obstetrics	2161-0932
Gynecology & Obstetrics- Case Report	-
Haematology & Thromboembolic Diseases	2329-8790
Hair: Therapy & Transplantation	2167-0951
Head and Neck Cancer Research	-
Health & Medical Economics	-
Health Care Communications	-
Health Economics & Outcome Research: Open Access	-
Health Education Research & Development (Biosafety & Health Education: Open Access-2332-0893)	-
Health Systems and Policy Research	2254-9137
Heart Transplant and Surgery	-
Heavy Metal & Chelation Therapy	-
Hepatology and Gastrointestinal Disorders	-
Hospital & Medical Management	-
Hypertension- Open Access	2167-1095
Hypo & Hyperglycemia	2327-4700
Imaging and Interventional Radiology	-
Medical Implants & Surgery	-
Informatics and Data Mining	-
Insights in Biomedicine	-
Insights in Medical Physics	-
Integrative Oncology	2329-6771
Internal Medicine	2165-8048
International Journal of Clinical & Medical Imaging	2376-0249
International Journal of Collaborative Research on Internal Medicine & Public Health	-
International Journal of Emergency Mental Health and Human Resilience	1522-4821
International Journal of Mental Health & Psychiatry	2327-4654
International Journal of Pediatric Neurosciences	-
International Journal of Physical Medicine & Rehabilitation	2329-9096
International Journal of Public Health and Safety	-
International Journal of School and Cognitive Psychology	-
Interventional Pediatrics	-
Invasive Cardiology Future Medicine	-
JBR Journal of Interdisciplinary Medicine and Dental Sciences	2376-032X
Kidney	-
Kidney Transplant	-
La Prensa Medica	0032-745X
Laser Surgery and Therapy	-
Leukemia	2329-6917
Liposuction	-
Liver	2167-0889
Liver: Disease & Transplantation	2325-9612
Lung Cancer Diagnosis & Treatment	-
Lung Diseases & Treatment	-
Malaria Control & Elimination	2090-2778
Maternal and Pediatric Nutrition	-
Medical & Surgical Pathology	-
Medical & Surgical Urology	2168-9857
Medical and Clinical Reviews	-
Medical Case Reports	-
Medical Diagnostic Methods	2168-9784
Medical Toxicology and Clinical Forensic Medicine	-
Melanoma and Skin Diseases	-
Mental Health in Family Medicine	2327-4972
Mental Illness and Treatment	-
Metabolic Syndrome	2167-0943
Molecular & Medical Histology	-
Molecular Medicine & Therapeutics	2324-8769
Neonatal Biology	2167-0897
Neonatal Studies	-
Neonatal Medicine	-
Neoplasm	-
Nephrology & Therapeutics	2161-0959
Neurobiotechnology	-
Neuroinfectious Diseases	2314-7326
Neurooncology: Open Access	-
Neurosurgery & Cardiac Surgery	-
Novel Physiotherapies	2165-7025
Nuclear Medicine & Radiation Therapy	2155-9619
Nutritional Disorders & Therapy	2161-0509
Obesity & Eating Disorders	-
Obesity & Weight Loss Therapy	2165-7904
Occupational Medicine Health Affairs	2329-6879
Omics Journal of Radiology	2167-7964
Oncology & Cancer Case Reports	-
Oncology Translational Research	-
Oral Health and Dental Management	2247-2452
Oral Health Case Reports	-
Oral Hygiene & Health	2332-0702
Orthodontics & Endodontics	-
Orthopedic & Muscular System: Current Research	2161-0533
Orthopedic Oncology	-
Osteoporosis & Physical Activity	2329-9509
Otolaryngology:Open Access	2161-119X
Otology & Rhinology	2324-8785
Pain & Relief	2167-0846
Pain Management & Medicine	-
Palliative Care & Medicine	2165-7386
Pancreatic Disorders & Therapy	2165-7092
Pediatric Care	-
Pediatric Dental Care	-
Pediatric Emergency Care and Medicine- Open Access	-
Pediatric Nephrology Practice	-
Pediatric Neurology and Medicine	-
Pediatric Nursing: Open Access	-
Pediatric Oncology: Open Access	-
Pediatric Physiotherapy	-
Pediatric Psychology and Psychiatry	-
Pediatrics & Therapeutics	2161-0665
Periodontics and Prosthodontics: Open Access	-
Pigmentary Disorders	2376-0427
Prevention Infection Control: Open Access	-
Preventive Medicine	-
Primary & Acquired Immunodeficiency Research	2324-853X
Prostate Cancer	-
Psoriasis & Rosacea Open Access	-
Psychiatry	2378-5756
Psychological Abnormalities in Children	2329-9525
Psychology & Psychotherapy	2161-0487
Pulmonary & Respiratory Medicine	2161-105x
Rare Disorders & Diseases	-
Regenerative Medicine	2325-9620
Reproductive Endocrinology & Infertility	-
Reproductive System & Sexual Disorders	2161-038x
Research & Reviews: Journal of Dental Sciences	2320-7949
Research & Reviews: Journal of Medical and Health Sciences	2319-9865
Research Journal of Biology	2322-0066
Sleep Disorders & Therapy	2167-0277
Sleep Disorders : Treatment & Care	2325-9639
Spine	2165-7939
Spine & Neurosurgery	2325-9701
Spine Research	-
Sports Medicine & Doping Studies	2161-0673
Sports Nutrition and Therapy	-
Steroids & Hormonal Science	2157-7536
Stroke Research & Therapy	-
Journal of Surgery [Jurnalul de Chirurgie]	1584-9341
Surgery: Current Research	2161-1076
The Headache Journal	-
The International Journal of Apitherapy	-
The Pancreas	1590-8577
Therapeutic Care and Physical Rehabilitation	-

Thrombosis and Circulation	-
Thyroid Disorders & Therapy	2167-7948
Translational Medicine	2161-1025
Transplant Reports : Open Access	-
Transplantation Technologies & Research	2161-0991
Trauma & Acute Care	-
Trauma & Treatment	2167-1222
Traumatic Stress Disorders & Treatment	2324-8947
Tropical Medicine & Surgery	2329-9088
Tumor Diagnostics and Reports	-
Universal Surgery	2254-6758
Vascular Medicine & Surgery	2329-6925
Vitiligo & Dermatomyositis	-
Voice Medicine & Surgery	-
Women's Health Care	2167-0420
Wound Medicine and Tissue Repair	-
Yoga & Physical Therapy	2157-7595

Microbiology

Advances in Influenza Research	-
Antimicrobial Agents	-
Antivirals & Antiretrovirals	1948-5964
Applied Microbiology: Open Access	-
Archives of Clinical Microbiology	1989-8436
Bacteriology and Parasitology	2155-9597
Clinical Infectious Diseases & Practice	-
Clinical Microbiology: Open Access	2327-5073
Colitis & Diverticulitis	-
Emerging Infectious Diseases	-
Fermentation Technology	2167-7972
Fibromyalgia: Open Access	-
Forensic Pathology	-
Hepatitis	-
Human Papillomavirus	-
Infectious Diseases and Diagnosis	-
Infectious Diseases and Therapy	2332-0877
Medical Microbiology & Diagnosis	2161-0703
Medical Mycology: Open Access	-
Meningitis	-
Mycobacterial Diseases	2161-1068
Pediatric Infectious Diseases: Open Access	-
Research & Reviews: Journal of Microbiology and Biotechnology	2320-3528
Research & Reviews: Journal of Inflammation	-
Research & Reviews: Journal of Pathology & Epidemiology	-
Virology & Mycology	2161-0517

Neuroscience

Addiction Research & Therapy	2155-6105
Alzheimers Disease & Parkinsonism	2161-0460
Autism-Open Access	2165-7890
Brain Disorders & Therapy	2168-975X
Child & Adolescent Behavior	2375-4494
Clinical & Experimental Neuroimmunology	-
Dementia & Mental Health	-
Epilepsy Journal	-
Insights in Clinical Neurology	-
International Journal of Neurorehabilitation	2376-0281
Multiple Sclerosis	2376-0389
Neurological Disorders	2329-6895
Neurology & Neurophysiology	2155-9562
Neurology and Neuroscience	2171-6625
Neuropsychiatry	-
Neuroscience & Clinical Research	-
Schizophrenia Journal	-

Health Care & Nursing

Advanced Practices in Nursing	-
Community & Public Health Nursing	-
Nursing & Care	2167-1168
Nursing & Clinical Research	-
Patient Care	-
Perioperative & Critical Intensive Care Nursing	-
Research & Reviews: Journal of Nursing and Health Sciences	-

Pharmaceutical Sciences

Advances in Pharmacoepidemiology & Drug Safety	2167-1052
Alcoholism & Drug Dependence	2329-6488
Bioanalysis & Biomedicine	1948-593X
Biochemistry & Pharmacology: Open Access Journal	2167-0501
Bioequivalence & Bioavailability	0975-0851
Biomarkers in Drug Development	2327-4441
Biomarkers Journal	-
Biomolecular Research & Therapeutics	2167-7956
Cardiovascular Pharmacology: Open Access	2329-6607
Clinical & Experimental Pharmacology	2161-1459
Clinical Pharmacology and Biopharmaceutics	2167-065X
Current Trends in Nutraceuticals	-
Developing Drugs	2329-6631
Diagnostic Techniques & Biomedical Analysis	-
Drug Designing: Open Access	2169-0138
Drug Metabolism & Toxicology	2157-7609
in Silico & in Vitro Pharmacology	-
Molecular Enzymology and Drug Targets	-
Molecular Pharmaceutics & Organic Process Research	2329-9053
Pharmaceutica Analytica Acta	2153-2435
Pharmaceutical Care & Health Systems	2376-0419
Pharmaceutical Microbiology	-
Pharmaceutical Regulatory Affairs: Open Access	2167-7689
Pharmaceutical Sciences & Emerging Drugs	-
Pharmaceutics & Drug Delivery Research	2325-9604
Pharmacoeconomics: Open Access	-
Pharmacogenomics and Pharmacoproteomics	2153-0645
Pharmacognosy & Natural Products	-
Pharmacokinetics & Experimental Therapeutics	-
Pharmacological Reports	-
Pharmacovigilance	2329-6887
Research & Reviews: Journal of Hospital and Clinical Pharmacy	-
Research & Reviews: Journal of Pharmaceutical Analysis	2320-0812
Research & Reviews: Journal of Pharmaceutical Quality Assurance	-
Research & Reviews: Journal of Pharmaceutics and Nanotechnology	2347-7857
Research & Reviews: Journal of Pharmacognosy and Phytochemistry	2321-6182
Research & Reviews: Journal of Pharmacy and Pharmaceutical Sciences	2320-1215
Virology & Antiviral Research	2324-8955

Physics

Astrophysics & Aerospace Technology	2329-6542
Research & Reviews: Journal of Pure and Applied Physics	2320-2459
Vortex Science and Technology	2090-8369

Social & Political Sciences

Anthropology	2332-0915
Arts and Social Sciences Journal	2151-6200
Civil & Legal Sciences	2169-0170
Forensic Anthropology	-
Global Media Journal	1550-7521
Intellectual Property Rights: Open Access	2375-4516
Mass Communication & Journalism	2165-7912
Political Science & Public Affairs	2332-0761
Research & Reviews: Journal of Educational Studies	-
Research & Reviews: Journal of Social Sciences	-
Socialomics	2167-0358
Sociology & Criminology	2375-4435

Veterinary Sciences

Animal Nutrition	-
Primatology	2167-6801
Research & Reviews: Journal of Veterinary Sciences	-
Research & Reviews: Journal of Zoological Sciences	2321-6190
Veterinary Science & Medical Diagnosis	2325-9590
Veterinary Science & Technology	2157-7579

Impact Factors* (IF)

Journal Name	Pubmed Short Name	Impact Factor
Biological Systems: Open Access	Biol Syst Open Access	0.76
Journal of Biotechnology & Biomaterials	J Biotechnol Biomater	1.94
Journal of Psychology & Psychotherapy	J Psychol Psychother	1.3
Advanced Techniques in Biology & Medicine	Adv Tech Biol Med	1.08
AIDS & Clinical Research	J AIDS Clin Res	2.7
Autism Open Access	Autism Open Access	3.52
Biochemistry & Physiology: Open Access	Biochem Physiol	1.03
Diversity Equality in Health & Care	Divers Equal Health Care	2.49
Drug Designing: Open Access	Drug Des	6
Fungal Genomics & Biology	Fungal Genom Biol	1.15
International Journal of Genomic Medicine	Int J Genomic Med	0.67
Journal of Addiction Research & Therapy	J Addict Res Ther	2.86
Journal of Alzheimers Disease & Parkinsonism	J Alzheimers Dis Parkinsonism	1.18
Journal of Fertilization: In Vitro	JFIV Reprod Med Genet	1
Journal of Genetic Syndromes & Gene therapy	J Genet Syndr Gene Ther	2.34
Journal of Microbial & Biochemical Technology	J Microb Biochem Technol	2.5
Journal of Nursing & Care	J Nurs Care	1.6
Journal of Osteoporosis and Physical Activity	J Osteopor Phys Act	0.66
Journal of Yoga & Physical Therapy	J Yoga Phys Ther	1.17
Molecular Biology	Mol Biol	1.85
Neurology & Neurophysiology	J Neurol Neurophysiol	0.77
Primary health care	Prim Health Care	1
Quality in Primary Care	Qual Prim Care	3.88
Tissue Science & Engineering	J Tissue Sci Eng	2.72
Biochemistry & Analytical Biochemistry	Biochem Anal Biochem	2.6
Molecular and Genetic Medicine	J Mol Genet Med	2.89
Advancements in Genetic Engineering	Adv Genet Eng	1
Enzyme Engineering	Enz Eng	2.3
Depression and Anxiety	J Depress Anxiety	1
Human Genetics & Embryology	Human Genet Embryol	1.2
Current Synthetic and Systems Biology	Curr Synthetic Sys Biol	0.8
Hereditary Genetics: Current Research	Hereditary Genet	1.2
International Journal of Emergency Mental Health and Human Resilience	Int J Emerg Ment Health	6.5
Spine	J Spine	1.9
Cloning & Transgenesis	Clon Transgen	1.5
Journal of Medical Microbiology & Diagnosis	J Med Microb Diagn	1.9
Biosensors Journal	Biosens J	0.33
Defense Management	J Def Manag	0.5
Review of Public Administration and Management	Review Pub Administration Manag	0.2
Single cell biology	Single Cell Biol	1
Gerontology & Geriatric Research	J Gerontol Geriatr Res	1
Neuroinfectious Diseases	J Neuroinfect Dis	2.4
Cell Science & Therapy	J Cell Sci Ther	1.37
Molecular Biomarkers & Diagnosis	J Mol Biomark Diagn	2.1
Brain Disorders & Therapy	Brain Disord Ther	1.6
Clinical Case Reports	J Clin Case Rep	1.2
Gene Technology	Gene Technol	0.83
Socialomics	J Socialomics	2.3
Journal of Trauma and Treatment	J Trauma Treat	0.6
Translational Biomedicine	Transl Biomed	1.06
Journal of Neurology and Neuroscience	J Neurol Neurosci	0.88
Research & Reviews: Journal of Botanical Sciences	J Bot Sci	0.33
Journal of Psychiatry	J Psychiatry	2.32
Anaplastology	Anaplastology	0.73
Tropical Medicine & Surgery	Trop Med Surg	0.4
Orthopedic & Muscular System: Current Research	Orthop Muscular Syst	0.32
Pediatrics & Therapeutics	Pediat Therapeut	1.32
Sports Medicine & Doping Studies	J Sports Med Doping Stud	1.45
Journal of Oral Hygiene & Health	J Oral Hyg Health	0.52
Emergency Medicine	Emerg Med (Los Angel)	0.875
Journal of Transplantation Technologies & Research	J Transplant Technol Res	1.39
Journal of Hypertension: Open Access	J Hypertens (Los Angel)	0.92
International Journal of Waste Resources	Int J Waste Resour	1.95
Surgery: Current research	Surgery Curr Re	0.587
Oral Health and Dental Management	Oral Health Dent Manag	1.23
International Journal of Advancement technology	Int J Adv Tech	5.08
Translational Medicine	Transl Med (Sunnyvale)	1.312
Air and Water Borne Diseases	Air Water Borne Diseases	0.6
Journal of Coastal Zone Management	J Coast Zone Manag	0.54
Biology and Medicine	Biol Med (Aligarh)	3.07
Journal of Bioterrorism and Biodefense	J Bioterror Biodef	0.38
Journal of Tropical Diseases & Public Health	J Trop Dis	0.83
Journal of Surgery	Journal of Surgery [Jurnalul de chirurgie]	0.08
Nephrology & Therapeutics	J Nephrol Ther	0.318
Journal of Fundamentals of Renewable Energy and Applications	J Fundam Renewable Energy Appl	1.41
Advances in Pharmacoepidemiology & Drug Safety	Adv Pharmacoepidemiol Drug Saf	1.37
Bioanalysis & Biomedicine	J Bioanal Biomed	1.67
Biochemistry & Pharmacology: Open Access	Biochem Pharmacol (Los Angel)	2.09
Bioequivalence & Bioavailability	J Bioequiv Availab	1.88
Biomolecular Research & Therapeutics	J Biomol Res Ther	1.67
Cardiovascular Pharmacology: Open Access	Cardiol Pharmacol	1.77
Clinical & Experimental Pharmacology	Clin Exp Pharmacol	1.83
Clinical Pharmacology & Biopharmaceutics	Clin Pharmacol Biopharm	1.69
Data Mining in Genomics & Proteomics	J Data Mining Genomics Proteomics	2
Drug Metabolism & Toxicology	J Drug Metab Toxicol	1.37
Ergonomics	J Ergonomics	1.38
Glycomics & Lipidomics	J Glycomics Lipidomics	1.82
Health & Medical Informatics	J Health Med Inform	1.98
Metabolomics: Open Access	Metabolomics (Los Angel)	3.03
Nanomedicine & Biotherapeutic Discovery	J Nanomedine Biotherapeutic Discov	2.69
OMICS Journal of Radiology	OMICS J Radiol	0.54
Pharmaceutica Analytica Acta	Pharm Anal Acta	1.83
Pharmaceutical Regulatory Affairs: Open Access	Pharm Regul Aff	1.88
Pharmacogenomics & Pharmacoproteomics	J Pharmacogenomics Pharmacoproteomics	1.69
Pharmacovigilance	J Pharmacovigil	2.65
Phylogenetics & Evolutionary Biology	J Phylogenetics Evol Biol	2.76
Proteomics & Bioinformatics	J Proteomics Bioinform	2.55
Advances in Automobile Engineering	Adv Automob Eng	1.750
Advances in Robotics & Automation	Adv Robot Autom	0.813
Arts and Social Sciences Journal	Arts Social Sci J	1.231
Bioceramics Developments and Applications	Bioceram Dev Appl	0.958
Business & Financial Affairs	J Bus & Fin Aff	2.000
Generalized Lie Theory and Applications	J Generalized Lie Theory Appl	1.750
Irrigation & Drainage Systems Engineering	Irrigat Drainage Sys Eng	4.286
Industrial Engineering & Management	Ind Eng Manage	0.474
Aeronautics & Aerospace Engineering	J Aeronaut Aerospace Eng	1.407
Applied & Computational Mathematics	J Appl Computat Math	0.581
Architectural Engineering Technology	J Archit Eng Tech	1.071
Accounting & Marketing	J Account Mark	0.500
Aquaculture Research & Development	J Aquac Res Development	1.272
Bioengineering & Biomedical Science	J Bioeng Biomed Sci	1.235
Biometrics & Biostatistics	J Biomet Biostat	1.272
Biosensors & Bioelectronics	J Biosens Bioelectron	2.137
Civil & Environmental Engineering	J Civil Environ Eng	1.294
Cytology & Histology	J Cytol Histol	0.569
Civil & Legal Sciences	J Civil Legal Sci	0.286
Ecosystem & Ecography	J Ecosyst Ecogr	1.806
Electrical & Electronic Systems	J Elec Electron Syst	0.533
Earth Science & Climatic Change	J Earth Sci Clim Change	2.082
Geography & Natural Disasters	J Geogr Nat Disast	0.800
Hotel & Business Management	J Hotel Bus Manage	1.600
Information Technology & Software Engineering	J Inform Tech Soft Engg	2.789
Molecular Imaging & Dynamics	J Mol Imaging Dynam	2.091

Earth Science & Climatic Change	J Earth Sci Clim Change	2.082
Geography & Natural Disasters	J Geogr Nat Disast	0.800
Hotel & Business Management	J Hotel Bus Manage	1.600
Information Technology & Software Engineering	J Inform Tech Soft Engg	2.789
Molecular Imaging & Dynamics	J Mol Imaging Dynam	2.091
Petroleum & Environmental Engineering	J Pet Environ Biotechnol	2.839
Stock & Forex Trading	J Stock Forex Trad	0.300
Textile Science & Engineering	J Textile Sci Eng	0.667
Tourism & Hospitality	J Tourism Hospit	1.190
Telecommunications System & Management	J Telecommun Syst Manage	0.800
Physical Mathematics	J Phys Math	4.500
Nanomedicine & Nanotechnology	J Nanomed Nanotechnol	4.68
Arabian Journal of Business and Management Review	Arab J Bus Manage Rev	1.42
Research and Reviews: Journal of Engineering and Technology	Engineering and Technology	0.14
Journal of Material Sciences & Engineering	J Material Sci Eng	1.31
Journal of Mass Communication & Journalism	J Mass Communicat Journalism	0.62
Journal of Powder Metallurgy & Mining	J Powder Metall Min	0.71
Journal of Applied Mechanical Engineering	J Appl Mech Eng	1.65
Archives of Clinical Microbiology		0.35
Dentistry	Dentistry	1.22
Journal of Diabetes & Metabolism	J Diabetes Metab	1.77
Otolaryngology: Current Research	Otolaryngol (Sunnyvale)	0.22
Journal of Metabolic Syndrome	J Metabolic Synd	1.27
Journal of Primatology	J Primatol	0.53
Journal of Thyroid Disorders & Therapy	Thyroid Disorders Ther	0.43
Jounal of Novel Physiotherapies	J Nov Physiother	1.24
Journal of Stem Cell Research & Therapy	J Stem Cell Res Ther	2.78
Anatomy & Physiology: Current Research	Anat Physiol	1
Pancreatic Disorders & Therapy	Pancreat Disord Ther	0.54
Journal of Cancer Science & Therapy	J Cancer Sci Ther	4.203
Journal of Biomedical Sciences		0.2
Journal of Nutritional Disorders & Therapy	J Nutr Disord Ther	1.46
Medical & Surgical Urology	Med Surg Urol	0.3
Journal of Biochips & Tissue Chips	J Biochip Tissue Chip	1.7
Journal of Liver	J Liver	0.08
Journal of Family Medicine and Medical Research	Fam Med Med Sci Res	0.78
Gynecology & Obstetrics	Gynecol Obstet (Sunnyvale)	0.52
Journal of Integrative Oncology	J Integr Oncol	1.67
Journal of Neonatal Biology	J Neonatal Biol	0.55
Journal of Glycobiology	J Glycobiology	0.8
Journal of Blood & Lymph	J Blood Lymph	0.12
Journal of Arthritis	J Arthritis	1.87
Journal of Membrane Science & Technology	J Membra Sci Technol	1.18
Medicinal Chemistry	Med Chem (Los Angeles)	2.64
Journal of Physical Chemistry & Biophysics	J Phys Chem Biophys	0.75
Organic Chemistry: Current Research	Organic Chem Curr Res	1.94
Journal of Bioprocessing & Biotechniques	J Bioprocess Biotech	1.74
Journal of Environmental & Analytical Toxicology	J Environ Anal Toxicol	2.58
Journal of Chemical Engineering & Process Technology	J Chem Eng Process Technol	1.21
Journal of Computer Science & Systems Biology	J Comput Sci Syst Biol	1.62
Journal of Analytical & Bioanalytical Techniques	J Anal Bioanal Tech	2.16
Journal of Plant Biochemistry & Physiology	J Plant Biochem Physiol	2.28
Journal of Chromatography & Separation Techniques	J Chromatogr Sep Tech	1.78
Journal of Thermodynamics & Catalysis		0.91
Community Medicine & Health Education	J Community Med Health Educ	1.27
Epidemiology: Open Access	Epidemiology (Sunnyvale)	1.35
Obesity & Weight Loss Therapy	J Obes Weight Loss Ther	0.94
Pain & Relief	J Pain Relief	1.14
Palliative Care & Medicine	J Palliat Care Med	0.88
Steroids & Hormonal Science	J Steroids Horm Sci	0.65
Gastrointestinal & Digestive System	J Gastrointest Dig Syst	0.43
Hair: Therapy & Transplantation		0.6
Andrology	Andrology (Los Angel)	1.16
Endocrinology & Metabolic Syndrome	Endocrinol Metab Syndr	1.12
Internal Medicine		2.48
Sleep Disorders & Therapy	J Sleep Disord Ther	0.5
Nuclear Medicine & Radiation Therapy	J Nucl Med Radiat Ther	0.88
Alternative & Integrative Medicine	Altern Integr Med	1.11
Pulmonary & Respiratory Medicine	J Pulm Respir Med	1.01
Occupational Medicine Health Affairs	Occup Med Health Aff	0.85
Reproductive System & Sexual Disorders	Reprod Syst Sex Disord	1.25
Medical Diagnostic Methods		0.29
Blood Disorders & Transfusion	J Blood Disord Transfus	0.5
General Medicine	Gen Med (Los Angel)	0.86
Bioenergetics: Open Access	Bioenergetics	3.1
Chemotherapy: Open Access	Chemotherapy (Los Angel)	1.8
Clinical & Experimental Pathology	J Clin Exp Pathol	1.54
Carcinogenesis & Mutagenesis	J Carcinog Mutagen	1.9
Clinical Research & Bioethics	J Clinic Res Bioeth	0.95
Vaccines & Vaccination	J Vaccines Vaccin	1.8
Immunome Research	Immunome Res	7.1
Clinical & Experimental Ophthalmology	J Clin Exp Ophthalmol	1.11
Clinical & Experimental Dermotology Research	J Clin Exp Dermatol Res	0.5
Clinical & Experimental Cardiology	J Clin Exp Cardiolog	1.33
Clinical Microbiology: Open Access	Clin Microbiol	0.7
Anesthesia & Clinical research	J Anesth Clin Res	0.7
Mycobacterial Diseases	Mycobact Dis	0.9
Clinical Toxicology	J Clin Toxicol	1.39
Clinical Trials & Research	J Clin Trials	1.33
Antivirals & Antiretrovirals	J Antivir Antiretrovir	1.27
Fermentation Technology	Ferment Technol	3.44
Clinical & Cellular immunology	J Clin Cell Immunol	2.019
Allergy & Therapy	J Allergy Ther	0.762
Bacteriology & Parasitology	J Bacteriol Parasitol	2.025
Rheumatology: Current Research	Rheumatology (Sunnyvale)	1.522
Virology & Mycology	Virol Mycol	0.69
Clinics in Mother and Child Health	Clinics Mother Child Health	0.432
Womens Health Care	J Womens Health Care	0.79
Marine Science: Research & Development	J Marine Sci Res Dev	0.45
Plant Pathology & Microbiology	J Plant Pathol Microbiol	1.75
Geology & Geophysics	J Geol Geophys	0.91
FisheriesSciences	J Fisheries Sci	0.51
Fisheries and Aquaculture Journal	Fish Aquac J	0.69
Bioremediation & Biodegradation	J Bioremediat Biodegrad	2.1
Advances in Crop Science and Technology	Adv Crop Sci Tech	0.39
Journal of Remote Sensing & GIS	J Geophys Remote Sens	0.77
Biofertilizers & Biopesticides	J Biofertil Biopestic.	1.19
Hydrology: Current Research	Hydrol Current Res	1.12
Probiotics & Health	J Prob Health	0.69
Veterinary Science & Technology	J Veterinar Sci Technolo	2.5
Medicinal & Aromatic Plants	Med Aromat Plants	2.02
Forest Research	Forest Res	1.69
International Journal of Sensor Networks and Data Communications	Sensor Netw Data Commun	1.66
Innovative Energy Policies	Innov Energ Policies	0.88
Biodiversity & Endangered Species	J Biodivers Endanger Species	0.25
Biosafety	Biosafety	0.49
Agrotechnology	Agrotechnol	0.69
Journal of Traditional Medicine and Clinical Naturopathy	J Tradition Med Clin Naturopth	0.49
Nutrition & Food Sciences	J Nutr Food Sci	1.14
Entomology, Ornithology & Herpetology	Entomol Ornithol Herpetol	1.26

Impact Factor Calculation:

Impact Factor was established by dividing the number of articles published in 2012 and 2013 with the number of times they are cited in 2014 based on Google search and the Scholar Citation Index database. If 'X' is the total number of articles published in 2012 and 2013, and 'Y' is the number of times these articles were cited in indexed journals during 2014 than, impact factor = Y/X







4th International Conference on

BIG DATA ANALYSIS AND DATA MINING

September 07-08, 2017 | Paris, France

Supporting Journals

Data Mining 2017

Supporting Journals

Journal of Computer Engineering & Information Technology
www.scitechnol.com/computer-engineering-information-technology.php

Journal of Computer Science & Systems Biology
www.omicsonline.org/computer-science-systems-biology.php

Journal of Data Mining in Genomics & Proteomics
www.omicsonline.org/data-mining-in-genomics-proteomics.php



Meet Inspiring Speakers and Experts at our 3000+ Global Events Every year on Medical, Pharma, Engineering, Science, Technology, Business and 40 Varient fields

AGRI, FOOD & AQUA

18th Global Summit on
Food & Beverages
Oct 02-04, 2017 Chicago, USA

8th International Conference on
Fisheries & Aquaculture
Oct 02-04, 2017 Toronto, Canada

9th Global Food Safety Conference
Dec 04-06, 2017 Atlanta, USA

20th Global Food Processing & Technology Summit Dec 11-13 , 2017 Philadelphia, USA

10th International Conference on
Agriculture & Horticulture
Oct 02-04, 2017 London, UK

7th International Conference on
Aquaculture & Fisheries
Oct 19-21, 2017 Madrid, Spain

19th International Conference on
Food Processing & Technology
Oct 23-25, 2017 Paris, France

2nd International Conference on
Food Microbiology
Nov 09-11, 2017 Madrid, Spain

World Aqua Congress
Oct 23-24, 2017 Dubai, UAE

5th International Food Safety, Quality & Policy Conference
Nov 27-28, 2017 Dubai, UAE

ALTERNATIVE HEALTHCARE

23rd International Conference on
Herbal and Alternative Remedies for Diabetes and Endocrine Disorders
Nov 02-04, 2017 Thailand, Bangkok

BIOCHEMISTRY

10th International Conference and Exhibition on
Metabolomics & Systems Biology
Oct 16-17, 2017 Baltimore, USA

3rd International Conference on
Genetic and Protein Engineering
Nov 08-09, 2017 Las Vegas, USA

9th International Conference and Expo on
Proteomics Oct 23-25, 2017 Paris, France

9th International Conference on
Bioinformatics
Oct 23-24, 2017 Paris, France

3rd International Conference on
Lipid Science & Technology
Dec 11-12, 2017 Madrid, Spain

3rd International Conference on
Transcriptomics
Oct 30- Nov 01, 2017 Thailand, Bangkok

CARDIOLOGY

21st International Conference on
Clinical & Experimental Cardiology
Nov 06-07, 2017 Las Vegas, USA

20th European Cardiology Conference
Oct 16-18, 2017 Budapest, Hungery

22nd World Cardiology Conference
Dec 11-12, 2017 Madrid, Spain

16th World Cardiology Congress
December 08-10, 2017 Dubai, UAE

3rd Global Summit on
Heart Diseases
Nov 02-04, 2017 Thailand, Bangkok

CHEMICAL ENGINEERING

7th International Congress and Expo on
Biofuels & Bioenergy
Oct 02-04, 2017 Toronto, Canada

3rd International Conference on
Chemical Engineering
Oct 02-04, 2017 Chicago, USA

7th International Conference and Exhibition on
Biopolymers and Bioplastics
Oct 19-21, 2017 San Francisco, USA

7th World Congress on
Petrochemistry and Chemical Engineering
Nov13-15, 2017 Atlanta, USA

2nd World Biodiesel Congress & Expo
Dec 04-05, 2017 Atlanta, USA

6th International Congress and Expo on
Biofuels, Bioenergy & Bioeconomy
Dec 04-06, 2017 Sao Paulo, Brazil

International Conference on
Renewable Energy and Resources
Oct 05-07, 2017 Kuala Lumpur, Malaysia

CHEMISTRY

4th International Conference on
Past and Present Research Systems of Green Chemistry Oct 16-18, 2017 Atlanta, USA

7th International Conference on
Medicinal Chemistry & Computer Aided Drug Designing
Nov 02-04, 2017 San Antonio, USA

2nd International Conference on
Nuclear Chemistry
Nov 06-07, 2017 Las Vegas, USA

2nd International Conference and Exhibition on
Polymer Chemistry
Nov 06-08, 2017 Las Vegas, USA

2nd International Conference on
Pharmaceutical Chemistry
Oct 02-04, 2017 Barcelona, Spain

5th International Conference and Expo on
Separation Techniques
Oct 23-25, 2017 Paris, France

10th Annual Chemistry Congress
Oct 18-19, 2017 Osaka, Japan

6th Global Congress on
Mass Spectrometry
Oct 18-19, 2017 Osaka, Japan

7th Global Mass Spectrometry Congress
Dec 14-16, 2017 Dubai, UAE

DENTISTRY

24th Euro Congress on **Dental & Oral Health**
Oct 19-20, 2017 Budapest, Hungery

International Conference on
Dentistry & Dental Marketing
Oct 05-06, 2017 Las Vegas, USA

29th Annual World Congress on
Dental Medicine & Dentistry
Oct 16-18, 2017 NewYork, USA

37th Global Summit on
Dental Surgeons & Dental Materials
Nov 02-04, 2017 San Antonio, USA

38th Annual Congress on
World Dentistry
Nov 06-08, 2017 San Antonio, USA

26th American Dental Congress
Dec 04-06, 2017 Atlanta, USA

39th World Dental Congress Summit
Dec 04-06, 2017 Sao Paulo, Brazil

37th Asia Pacific Dental and Oral Care Congress
Nov 20-22, 2017 Australia, Melbourne

9th Clinical Dermatology Congress
Oct 16-18, 2017 NewYork, USA

13th International Conference and Exhibition on
Cosmetic Dermatology and Hair care
Oct 26-27, 2017 Paris, France

23rd Asia Pacific Dermatology Conference
Oct 26-28, 2017 Osaka, Japan

DIABETES AND ENDOCRINOLOGY

International Conference on
Diabetes, Metabolism & Obesity
Nov15-17, 2017 Las Vegas, USA

International Conference on
Diabetes and Endocrinology
Dec 06-08, 2017 Atlanta, USA

21th International Conference on **Diabetes**
Oct 05-06, 2017 London, UK

2nd International Conference on
Herbal and Alternative Remedies for Diabetes and Endocrine Disorders
Nov 02-04, 2017 Thailand, Bangkok

25th Global Diabetes and Medicare Expo
Dec 11-12, 2017 Dubai, UAE

ENVIRONMENTAL SCIENCES

3rd Annual Congress on
Pollution and Global Warming
Oct 16-18, 2017 Atlanta, USA

4th International Conference on **GreenEnergy & Expo** Nov 06-08, 2017 Las Vegas, USA

5th International Conference on
Recycling: Reduce, Reuse and Recycle
Nov 06-08, 2017 Las Vegas, USA

2nd International Conference on
Pollution Control and Sustainable Environment
Oct 10-11, 2017 London, UK

4th World Conference on
Climate Change
Oct 19-21, 2017 Madrid, Spain

2nd World Congress on
Climate Change and Global Warming
Oct 16-17, 2017 Dubai, UAE

EEE & ENGINEERING

5th International Conference and Exhibition on
Mechanical & Aerospace Engineering
Oct 02-04, 2017 Las Vegas, USA

4th World Congress and Exhibition on
Construction and Steel Structure
Oct 16-18, 2017 Atlanta, USA

7th International Conference on
Nuclear Engineering
Oct 16-18, 2017 Atlanta, USA

International Conference on **Applied Energy**
Oct 23-24, 2017 Orlando USA

Meet Inspiring Speakers and Experts at our 3000+ Global Events Every year on Medical, Pharma, Engineering, Science, Technology, Business and 40 Varient fields

6th International Conference on **Biostatistics & Bioinformatics** Nov13-14, 2017 Atlanta, USA

2nd Global Summit on **Fluid Dynamics & Aerodynamics** Oct 19-20, 2017 Madrid, Spain

International Conference on **Mechatronics, Automation and Intelligent Materials** Oct 23-25, 2017 Paris, France

3rd World Congress on **Robotics and Artificial Intelligence** Oct 23-24, 2017 Osaka, Japan

GASTROENTEROLOGY

2nd International Conference on **Hepatology & Gastroenterology** Nov 13-14, 2017 Las Vegas, USA

2nd International Conference on **Digestive Diseases** Oct 16-17, 2017 London, UK

10th International Conference on **Gastroenterology** Oct 30- Nov 01, 2017 Bangkok, Thailand

GENETICS & MOLECULAR BIOLOGY

13th World Biotechnology Congress Oct 19-20, 2017 NewYork, USA

2nd World Biotechnology Congress Dec 04-05, 2017 Sao Paulo, Brazil

7th International Conference on **Tissue Engineering and Regenerative Medicine** Oct 02-04, 2017 Barcelona, Spain

10th International Convention on **Stem Cell and Biobanking** Oct 23-24, 2017 Osaka, Japan

GEOLOGY & EARTH SCIENCE

2nd International Convention on **Geophysics and Geo technics** Nov 08-09, 2017 Las Vegas, USA

2nd International Convention on **Geosciences and Remote Sensing** Nov 08-09, 2017 Las Vegas, USA

Annual Congress on **Soil Sciences** Dec 04-05, 2017 Madrid, Spain

5th International Conference on **Oceanography and Marine Biology** Oct 16-17, 2017 Seoul, South Korea

HEALTHCARE MANAGEMENT

6th International Conference on **Epidemiology & Public Health** Oct 23-25, 2017 Paris, France

2nd International Conference on **Health & Hospital Management** Nov 06-07, 2017 Vienna, Austria

International Conference on **Medical Education** Nov 06-08, 2017 Vienna, Austria

12th World Congress on **Healthcare and Medical Tourism** Oct 16-17, 2017 Dubai, UAE

IMMUNOLOGY

3rd Antibodies and Bio Therapeutics Congress Nov 08-09, 2017 Las Vegas, USA

5th International conference on **HIV/AIDS, STDS & STIS** Nov 13-14, 2017 Las Vegas, USA

9th World Congress and Expo on **Immunology** Nov 02-03, 2017 Atlanta, USA

3rd International Conference on **Immunity, Inflammation and Immunotherapies** Nov 02-03, 2017 Atlanta, USA

2nd International Conference on **Autoimmunity** Nov 09-10, 2017 Madrid, Spain

World Immunology Congress Dec 14-15, 2017 Dubai, UAE

INFECTIOUS DISEASES

13th World Congress on **Infection Prevention and Control** Oct 26-27, 2017 Milano, Italy

3rd Annual Congress on **Rare Diseases and Orphan Drugs** Oct 30-Nov 01, 2017 San Antonio, USA

3rd International conference on **Flu & Emerging Infectious Diseases** Nov 06-07, 2017 Las Vegas, USA

7th Asia Pacific STD and Infectious Diseases Congress Oct 23-25, 2017 Osaka, Japan

MATERIALS SCIENCE

3rd International Conference on **Polymer Science and Engineering** Oct 02-04, 2017 Chicago, USA

2nd International Conference on **Applied Crystallography** Oct 16-18, 2017 Chicago, USA

13th International Conference and Exhibition on **Materials Science and Engineering** Nov13-15, 2017 Las Vegas, USA

14th International Conference on **Functional Energy Materials** Dec 06-07, 2017 Atlanta, USA

International Conference on **Advanced Materials and Nanotechnology** Oct 26-28, 2017 Osaka, Japan

MICROBIOLOGY

2nd International conference on **Human Papillomavirus** Nov13-14, 2017 Las Vegas, USA

Global Veterinary Microbiology Summit & Expo Oct 02-04, 2017 Las Vegas, USA

11th World Summit on **Medical Microbiology** Oct 02-04, 2017 Las Vegas, USA

World Summit on **Nosocomial and Healthcare Associated Infections** Oct 02-04, 2017 Las Vegas, USA

6th Annual Conference on **Microbiology** Oct 16-17, 2017 Baltimore, USA

11th World Congress on **Virology** Oct 16-17, 2017 Baltimore, USA

World Yeast Congress Dec 06-07, 2017 Sao Paulo, Brazil

6th Clinical Microbiology Conference Oct 26-27, 2017 Paris, France

4th World Congress and Expo on **Applied Microbiology** Nov 09-11, 2017 Madrid, Spain

2nd International Conference and Summit on **Industrial and Pharmaceutical Microbiology** Oct 23-25, 2017 Osaka, Japan

10th International Congress on **Clinical Virology** Dec 04-05, 2017 Dubai, UAE

Annual Congress on **Microbes and Infection** Dec 04-05, 2017 Dubai, UAE

NANOTECHNOLOGY

22th International Conference and Expo on **Nanoscience and Molecular Nanotechnology** Nov 13-14, 2017 Vienna, Austria

19th International Conference on **Nanotek and Expo** Nov13-15, 2017 Atlanta, USA

17th Nanotechnology products and Summit Nov13-15, 2017 Atlanta, USA

15th World Medical Nanotechnology Congress & Expo Oct 18-19, 2017 Osaka, Japan

3rd Biomedical Engineering and Expo Nov 07-08, 2017 Barcelona, Spain

NEPHROLOGY

16th European Nephrology Conference Oct 02-04, 2017 Barcelona, Spain

17th World Nephrology Conference Oct 18-19, 2017 Dubai, UAE

NEUROSCIENCE

3rd International Conference on **Spinal Surgery** Oct 16-17, 2017 Chicago, USA

8th International Conference and Exhibition on **Addiction Research & Therapy** Nov 13-15, 2017 Las Vegas, USA

16th International Conference on **Neuro Cognitive Disorders** Oct 10-11, 2017 London, UK

9th International Conference on **Alzheimer's Disease & Dementia** Oct 16-18, 2017 Madrid, Spain

6th International Conference on **Brain Disorders and Therapeutics** Nov 06-08, 2017 Madrid, Spain

18th Global Neurologists Annual Meeting on **Neurology and Neuro Surgery** Nov 16-17, 2017 Vienna, Austria

15th International Conference on **Neuroscience** Oct 16-17, 2017 Osaka, Japan

13th Global Neurologists Annual Meeting on **Neurology and Neuro Surgery** Nov 27-28, 2017 Dubai, UAE

NURSING

3rd International Conference on **Reproductive Health** Oct 05-06, 2017 Chicago, USA

11th Global Healthcare and Fitness Summit Oct 16-18, 2017 San Francisco, USA

46th Global Nursing and Healthcare Conference Dec 06-07, 2017 Sao Paulo, Brazil

Meet Inspiring Speakers and Experts at our 3000+ Global Events Every year on Medical, Pharma, Engineering, Science, Technology, Business and 40 Varient fields

40th International Conference on **Nursing & Healthcare**
Oct 16-18, 2017 NewYork, USA

4th International Conference on **Gynecology & Obstetrics** Oct 02-04, 2017 Barcelona, Spain

13th International Conference on **Midwifery and Women's Health**
Oct 02-04, 2017 London, UK

41st Euro Nursing & Medicare Summit
Oct 26-28, 2017 Paris, France

27th Surgical Nursing & Nurse Education Conference Oct 16-17, 2017 Dubai, UAE

Asia-Pacific Nursing and Medicare Summit
Oct 05-07, 2017 Kuala Lumpur, Malaysia

World Congress on **Nursing Pharmacology and Nursing Education**
Nov 20-21, 2017 Melbourne, Australia

NUTRITION & OBESITY

16th International conference and Exhibition on **Obesity & Weight Management**
Nov16-17, 2017 Atlanta, USA

17th World Fitness Expo
Nov 16-17, 2017 Atlanta, USA

6th International Conference and Exhibition on **Probiotics, Functional and Baby Foods**
Oct 02-03, 2017 London, UK

15th Global Obesity Meeting
Oct 23-24, 2017 Dubai, UAE

18th Global Dieticians & Nutritionists Annual Meeting
Oct 02-03, 2017 Kuala Lumpur, Malaysia

16th Obesity Medicine Conference
Oct 30- Nov 01, 2017 Thailand, Bangkok

ONCOLOGY & CANCER

25th World Congress on **Cancer Therapy**
Oct 18-20, 2017 Baltimore, USA

10th Annual World Congress on **Biomarkers & Clinical Research**
Oct 18-20, 2017 Chicago, USA

5th International Conference on **Medical Imaging and Radiology**
Oct 19-20, 2017 NewYork, USA

15th World Oncologists Annual Conference
Oct 19-20, 2017 NewYork, USA

World Medical and Clinical Oncology Congress Nov13-15 , 2017 Las Vegas, USA

5th World Congress on **Breast Cancer**
Oct 16-18, 2017 San Francisco, USA

9th International Conference on **Hematology and Hematological Oncology**
Nov 08-09, 2017, Las Vegas, USA

11th International Conference on **Hematologic Oncology**
Oct 05-06, 2017 London, UK

25th World Cancer Conference
Oct 19-21, 2017 Madrid, Spain

International Conference on **Epigenetics**
Nov 06-07, 2017 Madrid, Spain

9th International Conference on **Biomarkers**
Oct 16-17, 2017 Osaka, Japan

14th Asia Pacific Oncologists Annual Meeting
Oct 26-28, 2017 Osaka, Japan

World Cancer Convention
Nov 27-28, 2017 Dubai, UAE

International Conference on **Cancer Diagnostics**
Nov 27-28, 2017 Dubai, UAE

International Conference on **Epigenetic Research**
Oct 26-28, 2017 Osaka, Japan

OPHTHALMOLOGY

20th World Ophthalmology Summit
Dec 04-05, 2017 Sao Paulo, Brazil

18th European Ophthalmology Congress
Dec 07-09, 2017 Madrid, Spain

2nd World Ophthalmology Conference
Oct 23-25, 2017 Osaka, Japan

17th Global Ophthalmology and Glaucoma Conference
Nov 27-28, 2017 Dubai ,UAE

PALLIATIVE CARE

8th International Conference on **Geriatric Medicine & Gerontological Nursing**
Oct 30-Nov 01, 2017 San Antonio, USA

PATHOLOGY

7th International Conference on **Predictive, Preventive and Personalized Medicine & Molecular Diagnostics**
Oct 5-6, 2017 Chicago, USA

6th Experts Meeting on **Medical Case Reports**
Oct 16-18, 2017 San Francisco, USA

2nd International conference on **Digital Pathology**
Nov 02-03, 2017 San Antonio, USA

PEDIATRICS

2nd World Congress Pediatric Oncology & Pediatric Medicine
Oct 05-06, 2017 Las Vegas, USA

2nd Annual Congress on **Infancy, Child Nutrition & Development (ICND)**
Oct 19-21, 2017 Atlanta,USA

5th Annual Conference on **Translational Medicine** Nov15-16, 2017 Las Vegas, USA

20th International Conference on **Neonatology and Perinatology**
Dec 04-06, 2017 Madrid, Spain

PHARMACEUTICAL SCIENCES

10th International Conference and Exhibition on **Biologics and Biosimilars**
Oct 16-17, 2017 Baltimore, USA

11th World Drug Delivery Summit
Oct 16-18, 2017 NewYork, USA

5th International Conference on **Clinical Pharmacy**
Oct 23-24, 2017 Orlando, USA

International Conference on **Biotech Pharmaceuticals**
Oct 23-25, 2017 Paris, France

European Biopharma Congress
Nov 16-17, 2017 Vienna, Austria

3rd Global Summits on **Herbal & Traditional Medicine**
Oct 18-20, 2017 Osaka, Japan

9th Annual Congress on **Drug Formulation & Drug Design**
Oct 19-21, 2017 Seoul, South Korea

10th International Conference on **Neuropharmacology and Neuropharmaceuticals**
Oct 23-24, 2017 Dubai, UAE

3rd World Congress on **Medicinal Plants and Natural Products Research**
Oct 02-04, 2017 Kuala Lumpur, Malaysia

6th Global Congress on **Mass Spectrometry**
Oct 18-19, 2017 Osaka, Japan

17th International Conference on **Nanomedicine and Nanotechnology in Health Care**
Nov 23-24, 2017 Melbourne, Australia

10th International Conferences on **Immunopharmacology and Immunotoxicology**
Nov 20-21, 2017 Melbourne, Australia

PHYSICAL THERAPY & REHABILITATION

6th International Conference on **Physiotherapy**
Nov 27-28, 2017 Dubai, UAE

PHYSICS

3rd International Conference on **Theoretical and Condensed Matter Physics**
Oct 19-21, 2017 NewYork, USA

2nd International Conference on **Astrophysics and Particle Physics**
Oct 30-Nov1, 2017 San Antonio, USA

8th International Conference and Exhibition on **Lasers, Optics & Photonics**
Nov 02-04, 2017 San Antonio, USA

2nd International Conference on **Atomic and Nuclear Physics**
Nov 8-9, 2017 Las Vegas, USA

International Conference on **High Energy Physics**
Dec 11-12, 2017 Madrid, Spain

PSYCHIATRY

2nd Experts Meeting on **Forensic Psychology and Criminology**
Oct 02-03, 2017 London, UK

24th International Conference on **Psychiatry & Psychosomatic Medicine**
Oct 02-04, 2017 London, UK

25th World Summit on **Psychology, Psychiatry & Psychotherapy**
Oct 19-20, 2017 San Francisco, USA

5th International Conference on **Counseling Psychology** Oct 16-17, 2017 Osaka, Japan

International Conference on **Psychiatry and Mental Health**
Nov 20-21, 2017 Melbourne, Australia

10th World Psychiatrists Meet
Dec 07-08, 2017 Dubai, UAE

RESPIRATORY

5th International Conference and Exhibition on **Lung and Respiratory Care**
Oct 19-20, 2017 San Francisco, USA

SURGERY

2nd International Conference on **Ear, Nose and Throat Disorders**
Oct 16-18, 2017 Madrid, Spain

4th International Conference and Exhibition on **Rhinology and Otology**
Oct 18-20, 2017 Dubai, UAE

Meet Inspiring Speakers and Experts at our 3000+ Global Events Every year on Medical, Pharma, Engineering, Science, Technology, Business and 40 Varient fields

TOXICOLOGY
12th International Conference on **Environmental Toxicology and Ecological Risk Assessment** Oct 19-20, 2017 Atlanta,USA
11th International Congress on **Toxicology and Risk Management** Oct 10-12, 2017 London, UK
International Conference on **Occupational Toxicology and Industrial Health** Oct 16-17, 2017 Dubai, UAE
VACCINES
19th World Congress on **Vaccines, Therapeutics for Infectious and Emerging Diseases** Oct 02-03, 2017 Chicago, USA
27th Asia Pacific Vaccines & Vaccination Conference Oct 05-07, 2017 Kuala Lumpur, Malaysia
29th Global Vaccines & Vaccination Summit And Expo Nov 30-Dec 1, 2017 Dubai, UAE
VETERINARY
8th International Conference on **Animal Health and Veterinary Medicine** Oct 02-04, 2017 Toronto, Canada
9th Global Veterinary Summit Nov15-16, 2017 Las Vegas, USA







4th International Conference on

BIG DATA ANALYSIS AND DATA MINING

September 07-08, 2017 | Paris, France

Keynote Forum

Day 1

Data Mining 2017

Michael Valivullah, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-013



Michael Valivullah

National Agricultural Statistics Service, USA

Big data, data hubs, sensors, IoT and precision agriculture – their effect on data collection, analysis, forecasting and survey processes

The National Agricultural Statistics Agency (NASS) of the United States Department of Agriculture (USDA) is in a mission to provide timely, accurate and useful statistics on US agriculture. Precision agriculture (PA) is a farming management principle that measures and responds to variability in crop conditions and animal health using sensors, robots, satellites and global positioning systems (GPS). The advent of precision agriculture has provided agricultural producers an unprecedented amount of data for use in data mining and big data analytics in farming operations. As more and more farmers are using automatic and remote sensing tools to collect data to be more productive, efficient and profitable, it is in the best interest of NASS to collect data from these sources (sensors, agribots, farm data hubs, drones, etc.,) and not burden farm producers by asking for the same data in NASS surveys. Automatic data collection (machine to machine) will also eliminate manual data entry errors. NASS needs to develop new survey and data collection processes, algorithms to validate, and analyze and process this data. Data analysis will require data scientists familiar with sophisticated algorithms, artificial intelligence, decision models, predictive analytics, etc. NASS also needs data dissemination tools with sophisticated big data processing capabilities and data visualization abilities. This presentation will discuss opportunities and challenges in dealing with big data in precision agriculture.

Biography

Michael Valivullah is currently serving as the Chief Technology Officer (CTO) at US Department of Agriculture's (USDA) National Agricultural Statistics Service (NASS). He served as Director of Information Technology (IT) and Chief Information Officer (CIO) at NASS and the Directorate of Science and Technology in the US Department of Homeland Security (DHS). He is a member of the Senior Executive Service (SES) in the US Federal Government. He has worked for 14 years in the public sector and over 15 years in the private sector and non-profits.

michael.valivullah@nass.usda.gov

En-Bing Lin, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-013



En-Bing Lin
Central Michigan University, USA

Big data analysis and approximations in information systems

With the increasing use of advanced technology, the amount of data in our world has been exploding. Big data analytics can examine large data sets and uncover hidden patterns. On the other hand, poor quality of big data results in some inaccurate insights or compliance failures that give rise to partially complete information systems. In order to obtain complete information systems, we use Rough Set Theory (RST), which was introduced by Pawlak in 1982 as a way to deal with data analysis based on approximation methods in information systems. The theory has many applications in a number of different areas, such as engineering, environment, banking, medicine, bioinformatics, pattern recognition, data mining, machine learning and others. RST is intrinsically a study of equivalence relations on the universe (a set of object). In fact, rough sets can be used to represent ambiguity, vagueness and general uncertainty. Given some relations between objects in the set, we can construct lower and upper approximations of the objects. We intend to use some advanced computing methods to determine lower and upper approximations and find several properties of the characteristics of objects within RST, as well as to extend RST to generalized RST. This line of research has to do with some developments in big data analytics. Traditional algorithm cannot satisfy the needs of big data computing. In this presentation, we will show some advanced computing methods that can solve our problems effectively. We will also present several examples to illustrate the concepts introduced in this presentation.

Biography

En-Bing Lin is a Professor of Mathematics at Central Michigan University, USA. He has been associated with several institutions including Massachusetts Institute of Technology, University of Wisconsin-Milwaukee, University of California, Riverside, University of Toledo, UCLA, and University of Illinois at Chicago. He received his PhD from Johns Hopkins University. His research interests include Data Analysis, Applied and Computational Mathematics, and Mathematical Physics. He has supervised a number of graduate and undergraduate students. He serves on the editorial boards of several journals. He has organized many special sessions at regional IEEE conference.

enbing.lin@cmich.edu

Notes:

Petra Perne, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-013



Petra Perne

Institute of Computer Vision and Applied Computer Sciences, Germany

Maintenance of engineering systems by big data

The ubiquitous availability of high quality data European industry gathers, allows to optimize manufacturing processes even more and to stay competitive. However, while the data are rich enough to include those elements needed for optimization, the even increasing volume, veloctiy and variety of the data make mining effectively increasingly difficult. The paper addresses the special challenges in developing scalable algorithm and infrastructures for creating responsive analytical capabilities that produce timely prediction and monitoring alerts in industrial environments. We will describe a platform that can handle the special needs of the data and has a reach enough tool of data mining techniques. Case-based reasoning is used to combine streaming data of different types (sensor data, time series, maintenance logs etc.) as well. Special time series algorithm will be developed allowing the efficient analysis of the machine data. It will be deployed and validated in three industrial cases where data-driven maintenances are expected to have a significant impact: High-tech medical equipment, high-tech manufacturing of hard disks and structural health monitoring.

Biography

Petra Perne is the Director of the Institute of Computer Vision and Applied Computer Sciences IBaI. She received her Diploma degree in Electrical Engineering and PhD degree in Computer Science for the work on “Data reduction methods for industrial robots with direct teach-in-programing”. Her habilitation thesis entitled “A methodology for the development of knowledge-based image-interpretation systems". She has been the Principal Investigator of various national and international research projects. She received several research awards for her research work and has been awarded with three business awards for her work on bringing intelligent image interpretation methods and data mining methods into business. Her research interest includes “Image analysis and interpretation, machine learning, data mining, big data, machine learning, image mining and case-based reasoning”.

pperner@ibai-institut.de

Notes:

Fionn Murtagh, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-013



Fionn Murtagh

University of Huddersfield, UK

Data mining in big data analytics: Exploiting resolution scale, addressing bias, having analytical focus

The benefits and also the challenges of big data analytics can be addressed in innovative ways. It is known that analytical focus is important. Considering just as an analogy for our analytics, how a microscope or a telescope bring about observation and measurement at very fine scales and at very gross scales, we can take that analogy as being associated with the resolution scale of our analysis. Another challenge is the bias in big data. But we may calibrate our analytical process with a big data framework or infrastructure. A further challenge of an ethical nature is how respresentativity replaces the individual. So, we want to rehabilitate the individual. Important opportunities arise from contextualization that can be associated with the resolution scale of our analytics, and it can also be supported by full account taken of appropriate contexts. The innovation that stems from the different facets of our analytical procedures can be of great benefit. Here, we seek to discuss many such themes that are always in the context of interesting and important case studies. The main case studies for us here include the following: analytics of mental health and associated well-being; social media analytics based on Twitter; questionnaire and survey analytics with many respondents.

Biography

Fionn Murtagh is a Professor of Data Science and was Professor of Computer Science, including Head of Department at many universities. He was Editor-in-Chief of the *Computer Journal* (British Computer Society) for more than 10 years, and is an Editorial Board Member of many journals. With over 300 refereed articles and 30 books authored or edited, his fellowships and scholarly academies include: Fellow of British Computer Society (FBCS), Institute of Mathematics and Its Applications (FIMA), International Association for Pattern Recognition (FIAPR), Royal Statistical Society (FRSS), and Royal Society of Arts (FRSA). He is an elected Member of Royal Irish Academy (MRIA) and Academia Europaea (MAE).

f.murtagh@gold.ac.uk

Notes:







4th International Conference on

BIG DATA ANALYSIS AND DATA MINING

September 07-08, 2017 | Paris, France

Plenary Talk

Day 1

Data Mining 2017

Karmen Kern Pipan, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014



Karmen Kern Pipan

Ministry of Public Administration, Slovenia

Big data – Big challenge for public administration – Experiences of pilot project in Ministry of Public Administration of Republic Slovenia

The pilot project - Big data analysis for HR efficiency improvement has been established as part of the development oriented strategy supporting ICTs as an enabler of development of data driven public administration in Republic Slovenia. It has been run within Ministry of Public Administration of Republic Slovenia in collaboration with EMC Dell as an external partner. This pilot project has been launched aiming to learn what big data tool installed on Slovenian State Cloud Infrastructure could enable in terms of the research of HR data of our ministry to improve our efficiency. Therefore, anonymized internal data sources containing time management, HR database, finance database and public procurement had been combined with external resources using postal codes of employees and weather data to identify potentials for improvement and possible patterns of behavior. The results showed that there is considerable potential for improvement in the field of HR and lowering costs in the field of public procurement within our ministry.

Biography

Karmen Kern Pipan started her career in the private sector in the field of Informatics and Telecommunications and later continued in public administration for different positions at Metrology Institute of Republic Slovenia as Quality Manager and Head of Department for Quality and Business Excellence. Currently, she works as Secretary and Project Manager at Ministry of Public Administration of Republic Slovenia in IT Directorate. As an expert, she is involved in the field of Data Management, Business Intelligence and Big Data Analytics to improve data based decision making at Slovenian Public Administration. She collaborates a lot with successful Slovenian companies and international professional sphere to identify good practices in this field. She leads an inter-ministerial task force for the preparation of the Public Administration Development Strategy 2015-2020. She has published several papers in reputed national and international conferences and journals.

Karmen.Kern-Pipan@gov.si

Boris Mirkin, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014



Boris Mirkin[1,2]

[1]Higher School of Economics, Russia

[2]Birkbeck University of London, UK

A complementary square-error clustering criterion and initialization of K-means

Clustering is a set of major data analysis techniques. The square-error clustering criterion underlies most popular clustering methods including K-means partitioning and Ward agglomeration. For the K-means, the square-error criterion to be minimized is the sum of squared Euclidean distances from all the objects to their respective cluster centers/means, W(S,c), where S is the sought partition of the set of objects and c is the set of within-cluster means. The method's popularity stems from the simplicity of computation and interpretation. Yet there is a catch: the user is to specify both the number of clusters and the initial locations of cluster centers, which can be of an issue sometimes. To tackle the problem, the current author proposes using the complementary criterion. It is not difficult to prove that there is a complementary criterion, B(S,c), to be maximized, such that W(S,c)+B(S,c)=T where T is the data scatter. The complementary criterion B(S,c) is the sum of individual cluster contributions, equal each to the product of the cluster's cardinality and the squared Euclidean distance from the cluster's center to 0. Therefore, the complementary criterion leads to a set of anomalous clusters, which can be found either one-by-one or in parallel. Our experiments show that methods emerging in this perspective are competitive, and frequently superior, to other initialization methods.

Biography

Boris Mirkin holds a PhD in Computer Science and DSc in Systems Engineering from Institute for Systems Studies of the Academy of Sciences and State Committee of Science and Technology of the USSR. He published a dozen monographs and a hundred refereed papers. He develops methods for clustering and interpretation of complex data within the data recovery perspective. Currently, these approaches are being extended to automation of text analysis including use of hierarchical ontologies.

mirkin@dcs.bbk.ac.uk

Notes:

Kazumi Nakamatsu, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014



Kazumi Nakamatsu

University of Hyogo, Japan

Process before-after order data verification/prediction systems based on paraconsistent annotated logic program Bf-EVALPSN

Paraconsistent logic is a well-known formal logic that can deal with contradiction in the framework of a logical system consistently. One of paraconsistent logics called annotated logic was proposed by Prof. Newton da Costa et al. and its logic program was been developed by Prof. V S Subrahmanian et al. later as a tool of dealing with data in knowledge bases. Some years later a kind of paraconsistent annotated logic program has been developed for dealing with non-monotonic reasoning such as default reasoning by Kazumi Nakamatsu. Recently, a paraconsistent annotated logic program called Extended Vector Annotated Logic Program with Strong Negation (EVALPSN) that can deal with conflict resolving, defeasible deontic reasoning, plausible reasoning, etc. has been developed and already been applied to various intelligent controls and safety verification systems such as pipeline valve control, traffic signal control, railway interlocking safety verification, etc. Furthermore, most recently one specific version of EVALPSN called Before-after EVALPSN (Bf-EVALPSN) that can deal with before-after relations between processes (time intervals) has been developed. In this lecture, I introduce how EVALPSN and Bf-EVALPSN deal with contradictory data with a small example and can be applied to intelligent control or safety verification of sensed data.

Biography

Kazumi Nakamatsu received Doctor of Science degree at Kyushu University, Japan. He is a Professor at University of Hyogo, Japan. He contributed over 150 journal/conference papers and book chapters, and edited/authored 12 books published by prominent publishers. He has chaired various international conferences/workshops, and he has been a program committee member/chair of academic conferences. He has contributed numerous invited lectures at conferences/academic organizations. He received some conference/best paper awards at some international conferences.

nakamatu@shse.u-hyogo.ac.jp

Notes:

Pavel Kisilev, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014



Pavel Kisilev

Toga Networks Ltd., Israel

Smart data acquisition and analysis for deep learning based systems

Deep Learning (DL) becomes the method of choice in many fields of computer science, including computer vision, Natural Data Processing (NLP), autonomous systems, and many others. While DL methods are known for their superior performance, in order to achieve it, large amount of training examples are required. Furthermore, the quality of training examples affects largely the performance of DL training and learnt model quality. In particular, if the training examples represent well the real world phenomena, a good model is expected to learn. If, on the other hand, the examples are highly correlated and represent only sparse knowledge about phenomena, the learnt model quality will be low. In this talk, we present several general principles and methods to diagnose the data quality, and also the suitability of DL architecture to model the data at hand. We also propose several methods to pre-process raw data to better suit the requirements of DL systems. We show several examples of applications of our framework to various datasets, including known large image datasets with millions of images, binary sequence sets, gene datasets, and others. We show the efficacy of the proposed methods to analyze and predict the performance of DL methods on a given data.

Biography

Pavel Kisilev serves as the CTO of Artificial Intelligence at Huawei Research Center in Israel. He completed his Graduation at Technion, Israel Institute of Technology in 2002 with a PhD in Electrical Engineering. Before joining Huawei, he was a Lead Scientist at IBM Research from 2011-2016, a Senior Research Scientist at HP Labs from 2003-2011, and a Research Associate at Technion. His research interests include computer vision, deep learning, general statistical methods, and inverse problems. He is an author of over 50 filed patents, three book chapters, and of nearly 50 papers in top journals and conferences in computer science.

Pavel.Kisilev@toganetworks.com

Notes:

Bennett B Borden, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014



Bennett B Borden

Drinker Biddle & Reath, USA

Predicting corporate misconduct

Corporate misconduct costs $3 Trillion worldwide every year to prevent, detect and deal with its consequences. If we can predict when someone will purchase a product, click on an ad, or how they will vote for a candidate, why can't we predict when he or she will engage in some form of fraud or other misconduct? Well, perhaps we can. In this session, Chief Data Scientist Bennett Borden, from the law firm Drinker Biddle & Reath will present his work on developing algorithms to predict corporate misconduct and how this technology is being used today and how it will likely be used in the future.

Biography

Bennett B Borden is a Chief Data Scientist at Drinker Biddle & Reath. He is a globally recognized authority on the legal, technology and policy implications of information. His ground-breaking research into the use of machine-based learning and unstructured data for organizational insight is now being put to work in data-driven early warning systems for clients to detect and prevent corporate fraud and other misconduct. He received his Master of Science in Data Analytics at New York University and his JD at Georgetown University.

bennett.borden@dbr.com

Notes:

Nikolaos Freris, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014



Nikolaos Freris

New York University Abu Dhabi, UAE

Exact data mining from inexact data

Big data pertain to multiple facets of modern science and technology enlisting biology, physics, social networks, financial analysis, smart cities and many more. Despite the overwhelming amount of accessible data alongside the abundance of mining schemes, the prelude of data mining faces a key challenge in that the data are hardly ever available in their original form. Common operations such as compression, anonymization and right protection may significantly affect the accuracy of the mining outcome. We will discuss the fundamental balance between data transformation and data utility under prevalent mining operations such as search, K-nearest neighbors and clustering. In specific, we will illustrate classes of data transformation – information extraction methods where it is actually feasible to acquire the exact mining outcome even when operating on the transformed domain. This talk will feature three specific problems: Optimal distance estimation of compressed data series; nearest neighbor preserving watermarking and; cluster preserving compression. We provide provable guarantees of mining preservation, and further highlight the efficacy and efficiency of our proposed methods in a multitude of datasets: weblogs, VLSI images, stock prices, videos, and images from anthropology, natural sciences, and handwritings.

Biography

Nick Freris is an Assistant Professor of Electrical and Computer Engineering (ECE), and Director of Cyber-physical Systems Laboratory (CPSLab) at New York University Abu Dhabi. He received Diploma in ECE at National Technical University of Athens in 2005; MS degree in ECE in 2007; MS in Mathematics in 2008 and; PhD in ECE at University of Illinois at Urbana-Champaign. His work was recognized with 2014 IBM High Value Patent award. He is a senior member of IEEE, and member of SIAM and ACM.

nf47@nyu.edu

Notes:







4th International Conference on

BIG DATA ANALYSIS AND DATA MINING

September 07-08, 2017 | Paris, France

Scientific Tracks & Abstracts

Day 1

Data Mining 2017

Major Sessions:

Data Mining Applications in Science, Engineering, Healthcare and Medicine | Data Mining Methods and Algorithms | Artificial Intelligence | Big Data Applications | Big Data Algorithm Data Privacy and Ethics | Data Mining Analysis | Business Analytics | Optimization and Big Data | New visualization techniques | Clustering

Session Chair
Michael Valivullah
CTO, NASS, USA

Session Co-Chair
Petra Perne
Institute of Computer Vision and applied Computer Sciences, Germany

Session Introduction

Title: **Scalable robust clustering method for large and sparse data**
Tommi Kärkkäinen, University of Jyväskylä, Finland

Title: **Data is not wrong, but your use of it might be!**
Fred Jacquet, Evangelist, France

Title: **An automatic data prediction system with business applications**
Morgan C. Wang, University of Central Florida, USA

Title: **A novel linear-time and memory saving approach for visual exploratory data analysis: data embedding and graph visualization**
Witold Dzwinel, AGH University of Science and Technology, Poland

Title: **Multi-objective optimization via visualization**
Nicolas Wesner, Mazars Actuariat, France

Title: **Multi-criteria decision-making for purchasing cell phones using machine learning approach**
S.N.Mohanty, KIIT University, India

Title: **Big data - applied in healthcare context, the case study of diabetes**
Jenny Lundberg, Linnaeus University, Sweden

YRF Session

YRF: **Cost-Sensitive deep neural networks to address class imbalanced problem**
Kruy Seng, Lingnan University, Hong Kong

YRF: **Data driven intelligent analytics in education domain**
Prajakta Diwanji, FHNW University of Applied Sciences and Arts, Switzerland

YRF: **The domus-lig experimental living-lab methodology and the emoz platform: overview of spontaneous and ecological corpora - from small smart data to big data**
SASA Yuko, University Grenoble Alps, France

Tommi Kärkkäinen, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Scalable robust clustering method for large and sparse data

Tommi Kärkkäinen
University of Jyväskylä, Finland

Clustering is the most common unsupervised, descriptive analysis technique to reveal hidden patterns and profiles from a dataset. There exists large number of different clustering algorithms, but approaches that specifically address clustering of sparse datasets are still scarce, even if real world datasets are many times characterized by missing values with unknown sparsity pattern. Typical approaches in the knowledge discovery process is to either completely omit the observations with missing values or use some imputation method to fill in the holes of data. However, the throw data away approach does not utilize all possible data and the imputation necessarily introduces assumptions about the unknown density of the data. Moreover, by the well-known curse-of-dimensionality results, such assumptions are no more valid in the high dimensional spaces. The purpose of this presentation is to describe and summarize a line of research that addresses the sparse clustering problems with the available data strategy and robust prototypes. The strategy allows one to utilize all available data without any additional assumptions. The actual prototype-based clustering algorithm, the k-spatial medians, relies on the computation of a robust prototype as cluster centroid, again argumenting on non-Gaussian within-cluster error in comparison to the classical k-means method. As with any prototype-based algorithm, the initialization step of the locally improving relocation algorithm has an important role and should be designed to handle the sparse data. Such an approach is proposed and the scalability of a distributed implementation of the whole algorithm is tested with openly available large and sparse datasets.

Biography

Tommi Kärkkäinen has completed his PhD at University of Jyväskylä in 1995 and worked as a Full Professor in the Faculty of Information Technology since 2002. He has been and is serving in many positions of administration and responsibility at the faculty and the university level. He has published over 150 research papers, led dozens of R&D projects, and supervised over 20 PhD theses.

tommi.karkkainen@jyu.fi

Notes:

Fred Jacquet, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Data is not wrong, but your use of it might be!

Fred Jacquet
Solutions Architect & Evangelist | Cloud & Big Data Analytics - Artificial Intelligence - Digital Transformation, Talend

Some companies are born in the data intelligence era, some need to redefine their IT infrastructure to stay competitive in the world of improved analytics, scale, speed, production and economics. The old data warehousing and business insights tools helped to find the answers to what happened? but are not able to answer the new questions asked by more and more forward thinking companies. Traditional warehousing tools such as ETL, RBMS and OLAP databases cannot provide answers to questions such as what's happening right now or what will happen? The new era of analytics demands speed, scale and reduced costs from every IT team. This presentation will take you through the considerations and steps of modernizing your data warehouse to become ready for big data analytics and evaluate if the data lake is right for you and your business needs. Don't be one of the many companies that failed to grasp this opportunity to leapfrog their competition. After all, over 50% of companies that were originally on the Fortune 500 list have vanished since 2000, because these are failed to innovate and keep up with the changes in the market.

Biography

Fred Jacquet has over 20 years of experience in the IT industry and has CTO and Architect roles within a variety of leading data-driven companies. His main area of expertise is in Business Intelligence, Data Integration and Big Data. He is committed to help organizations on their mission to become successfully data driven through evangelization, education and enablement.

fred.w.jacquet@gmail.com

Notes:

Morgan C Wang, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

An automatic intelligent model building system with business applications

Morgan C Wang
University of Central Florida, USA

An automatic intelligent model building system was developed. This system has five components: data exploration component, data preparation component, model building/validation/selection component, result automatic generation/data scoring component, and model understanding component. This system support four type of target variables – binary, nominal, count, and numerical. The data preparation component can be used to fix data problems such as missing values, skewness, and high cardinality. The modeling component can automatically identify parameter that leads to build a better model. Currently it supports neural network, decision trees, gradient boosting, rand forest and many regression algorithms. After the optimal model selected, the user can further test the model performance or use the selected model to score new data. This system also attempts to open the black box to allow the user to see some insight of the modeling results such as interaction among predictors, important predictors, how to alter predictors to change the predicted values. This system is intended to be used by company personal without extensive model building training since there is a gap between the demand and supply of well-trained analysts. Two case studies will be discussed. The first case is to use this system on establish optimal pricing of an auto insurance firm in China. The second case is to develop a risk score system for an internet lending firm. The results from both cases are very positive and encouraging.

Biography

Morgan C Wang received his PhD at Iowa State University in 1991. He is the funding Director of Data Mining Program and Professor of Statistics at University of Central Florida. He has published one book and over 80 papers in referee journals and conference proceedings on topics including interval analysis, meta-analysis, computer security, business analytics, health care analytics and data mining. He is the elected member of International Statistical Association and member of American Statistical Association and International Chinese Statistical Association.

Chung-Ching.Wang@ucf.edu

Notes:

Witold Dzwinel, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

A novel linear-time and memory saving approach for visual exploratory data analysis: Data embedding and graph visualization

Witold Dzwinel
AGH University of Science and Technology, Poland

Data embedding (DE) and graph visualization (GV) methods are very congruent tools used in exploratory data analysis for visualization of complex data such as high-dimensional data and complex networks, respectively. However, high computational complexity and memory loads of existing DE and GV algorithms (based on t-SNE concept from one hand, and force-directed methods from the other), considerably hinders visualization of truly large and big data consisting of as many as M~106+ data objects and N~103+ dimensions. In this paper, we demonstrate the high computational efficiency and robustness of our approach to data embedding and interactive data visualization. We show that by employing only a small fraction of distances between data objects, one can obtain very satisfactory reconstruction of the topology of N-D data in 2D in a linear-time O (M). The IVHD (Interactive Visualization of High-Dimensional Data) method quickly and properly reconstructs the N-D data topology in a fraction of computational time required for the state-of-art DE methods such as bh-SNE and all its clones. Our method can be used for both metric and non-metric (e.g. large graphs) data visualization. Moreover, we demonstrate that even poor approximations of the nearest neighbor (NN) graph, representing high-dimensional data, can yield acceptable data embedding. Furthermore, some incorrectness in the nearest neighbor list can often be useful to improve the quality of data visualization. This robustness of IVHD, together with its high memory and time efficiencies, meets perfectly the requirements of big and distributed data visualization, when finding the accurate nearest neighbor list represents a great computational challenge.

Biography

Witold Dzwinel holds Full Professor position at AGH University of Science and Technology, Department of Computer Science in Krakow. His research activities focus on "Computer modeling and simulation methods employing discrete particles". Simultaneously, he is doing research in interactive visualization of big data and machine learning algorithms. He is the author and co-author of about 190 papers in computational science, computational intelligence and physics.

dzwinel@agh.edu.pl

Notes:

Nicolas Wesner, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Multi-objective optimization via visualization

Nicolas Wesner
Mazars Actuariat, France

This paper presents a very simple and intuitive multi-objective optimization method that makes use of interactive visualization techniques. This approach stands mid-way between the brush and link technique, a visual method used in operational research for exploratory analysis of multidimensional data sets and interactive multi-criteria decision methods that use the concept of reference point. Multiple views of the potential solutions on scatterplots allow the user to directly search acceptable solutions in bi-objective spaces whereas a Venn diagram displays information about the relative scarcity of potential acceptable solutions under distinct criteria. Those very intuitive data visualization techniques allow for comprehensive interpretation and permit to communicate the results efficiently. More generally the combination of information visualization with data mining allows the user to specify what he is looking for, yields easily reportable results and respects human responsibility. An application to the visual steering of genetic algorithms in a multi-criteria strategic asset allocation optimization problem is presented.

Biography

Nicola Wesner has completed his PhD in Economics at Paris X Nanterre University in 2001 and he is an Associate Actuary since 2011. He is the Head of the Pension department at Mazars Actuariat, an audit and advisory firm. He has published many papers in reputed journals and has specialized periodicals on various subject such as econometrics, quantitative finance, insurance and pension, and data mining.

nicolas.wesner@mazars.fr

Notes:

S N Mohanty, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Multi-criteria decision-making for purchasing cell phones using machine learning approach

S N Mohanty
KIIT University, India

The process of cell-phone selection for purchasing is a multi-criteria decision-making (MCDM) problem with conflicting and diverse objective. In this study, various techniques using machine learning approach is discussed. To begin, participants responded to a questionnaire having different latest features available in a cell-phone. Seven independent input variables cost, talk-time, rear camera, weight, size, memory and operating system were then derived from participant respondents. Linguistic terms such as low, medium and high were used to represent each of the input variables. Using Mamdani approach both traditional fuzzy reasoning tool (FLC) and neuro-fuzzy system (ANFIS) were designed for three input and one output process. The neuro-fuzzy system was trained using a back-propagation algorithm. Compared to two traditional fuzzy reasoning tools, artificial neural network approach (ANN), the neuro-fuzzy system could provide better accuracy for selecting a cell-phone for personal use.

Biography

S N Mohanty completed his PhD from IIT Kharagpur, India in the year 2014, with MHRD Scholarship from Govt. of India. He has recently joined as Assistant Professor in the School of Computer Science & Engineering at KIIT University. His research areas include Data mining, Big Data Analysis, Cognitive Science, Fuzzy Decision Making, Brain-Computer Interface, Cognition, and Computational Intelligence. He has received two best paper awards during his PhD at IIT Kharagpur from International Conference at Benjing, China, and the other at International Conference on Soft Computing Applications organized by IIT Roorkee in the year 2013. He has published five papers in international journals and has been elected as member of Institute of Engineers and IEEE Computer Society. He is the Reviewer of IJAP and *IJDM*.

snmohantyfcs@kiit.ac.in

Notes:

Jenny Lundberg, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Big data applied in healthcare context: The case study of diabetes

Jenny Lundberg[1, 2]
[1]Linnaeus University, Sweden
[2]Lund University, Sweden

As the industry 4.0 era gives us extensive IoT opportunities to provide evidence and context based data, opening for new approaches and methods to meet societal challenges. The handling of chronic diseases is global and poses challenges to the current health systems. The incidence of the chronic disease diabetes is of epidemic character. In 2015, 415 million in the world have diabetes and it is estimated that by 2040, 642 million in the world will have diabetes. More specifically: Diabetes is a heterogeneous group of conditions that all result in, and defined by, rise in plasma glucose level if not well treated. If it is untreated, death is sooner or later, if later with a lot of unpleasant complications over time. There are two main types of diabetes, type 1 (10% of all); and type 2 (85-90% of all). Diabetes places extremely high demands on the individual in terms of self-care, and a lot of complications can occur. It is a well-known fact that this creates serious health condition and high social costs. Potentially, this can be prevented with new methods for better support for self-care. Given developments, recent advances in mobile computing, sensor technology, big data can be used to better understand diabetes, measurements and data. To overcome some of the problems in this area, open data, such as social media, special designed apps, sensors and wearables can be used to find proactive ways and methods of diabetes treatment.

Biography

Jenny Lundberg has completed her PhD in Computer Science at 2011 at BTH, the most profiled IT University in Sweden. She is a Senior Lecturer at Linnaeus University and a Researcher at Lund University. She has extensive international research and education collaboration experiences. Her research interest is in health applications and work in close cooperation with clinical researchers in healthcare with Big data, e- and m-health approaches and techniques.

jenny.lundberg@lnu.se

Notes:







4th International Conference on

BIG DATA ANALYSIS AND DATA MINING

September 07-08, 2017 | Paris, France

Young Researcher Forum

Day 1

Kruy Seng et al., J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Cost-sensitive deep neural networks to address class imbalanced problem

Kruy Seng and **Man-Leung Wong**
Lingnan University, Hong Kong

Class imbalance is one of the common issues in data mining. Most data in real-life applications are imbalanced, that is, the number of examples belonged to a class is significantly greater than those of the others. Abnormal behaviors or rare events are the main reasons that cause the distribution of classes to be imbalanced. The scarcity of minority class examples leads the standard classifiers to be more focused on the majority class examples while ignores the other. As a result, classification becomes more challenging. A number of classification algorithms were proposed in the past decades; however, they are accuracy-oriented and unable to address class imbalanced problem. Much research has been conducted to address this problem. However, it is still an intensive research topic. In this work, we propose an approach called class-based cost-sensitive deep neural networks to perform classification on imbalanced data. The costs of misclassification of each type of errors are treated differently and are incorporated into the training process of deep neural networks. We also generalize the method by reducing the effort of hyper parameter selection by adapting evolutionary algorithm search for optimal cost vector setting and network structure. Finally, experiments will be conducted to analyze the performance and compare with other existing methods.

Biography

Kruy Seng is currently an MPhil student in Department of Computing and Decision Sciences at Lingnan University. His research interests include Data Mining in Business Application and Machine Learning.

kruyseng@ln.hk

Notes:

Prajakta Diwanji et al., J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Data driven intelligent analytics in education domain

Prajakta Diwanji, Knut Hinkelmann and **Hans Friedrich Witschel**
FHNW University of Applied Sciences and Arts Northwestern, Switzerland

In recent times, there is a steady growth in student's personal as well as academic data in the education field. Many universities and institutes have adopted information systems like virtual learning environments, learning management systems and social networks that collect student's digital footprints. This data is both large in volume and diversity. Learning analytics offers tools to facilitate the understanding of different parameters related to student's engagement/motivation, learning behavior, performance, teaching content and learning environment. Such information could help teachers better prepare for the classroom sessions and to deliver personalized or adaptive learning experiences. This, in turn, could enhance student performance. The current literature research states that there is a shift of focus from classroom based learning to a more anytime, anywhere learning; as well as from a teacher as a sole knowledge contributor to agent or learner as a contributor towards learning. The use of intelligent digital tutors/chatbots has taken the learning process to a new level of student engagement, interaction, and learning. Such intelligent data analysis tools/systems make use of data analysis techniques like machine learning, natural language processing etc. along with artificial/cognitive intelligence techniques. The research work identifies the current challenges faced by universities in learning/teaching processes in a real world context and tries to address these problems using data driven intelligent analysis methods. The main goal would be to focus on preparing students as well as lecturers effectively for the classroom lectures; to understand learning needs of students beforehand; to address those needs proactively in a timely manner.

Biography

Prajakta Diwanji is working as a Researcher in Information Systems at University of Applied Sciences and Arts, Northwestern Switzerland (FHNW). She is a first-year Doctoral student at University of Camerino, Italy. Her research interest is in the area of intelligent data analytics in education domain. She has completed her Master's degree in Business Information Systems from FHNW, Switzerland, and Masters in Computer Science at University of Pune, India. She has more than seven years of work experience in IT industry where she has taken up several challenging roles. During this tenure, she has worked with international companies like Roche Pharma, Switzerland and IBM, India etc.

prajakta.diwanji@fhnw.ch

Notes:

SASA Yuko, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

The Domus-LIG Experimental Living-Lab methodology and the EmOz platform: overview of spontaneous and ecological corpora - from Small Smart Data to Big Data

SASA Yuko
University Grenoble Alps, France

The availability of real-life data through the IoT and ICT technologies, completed by the augmenting computers' calculability have developed the paradigm of Big Data, to process an enormous amount of information to automate various system as speech technologies and robotics. The motivation resided on the "intelligence" of data to handle the naturalness and variability of the human behaviors. The actual machine learning techniques, as DNN, let the computer mathematics to approximate and to generalize the cognitive processes more than modelizing them. The quantity of data aim is then to cover the knowledge, which is poorly explicit. They maybe rely on too many implicit mechanisms locked up into black boxes. The Domus-LIG Experimental Living-Lab methodology and the EmOz platform, a wizard of oz tools interfaced with a robot, both developed to induce, observe and collect spontaneous and ecological interactional data of human-robot communication, particularly with socio-affective values. The robot is thus a measuring instrument of the human multimodal speech features' effects, strongly hypothesized. This methodology works on agile loops of processing to control the data contents and format them in short and rapid processes. This evolving corpus is the basis for an iterative machine learning, dedicated to automatic recognition systems. Different studies leading to the EEE (EmOz Elderly Expressions) or the GEE (Gesture EmOz Expressions) corpora this approach. This bootstrap is an invitation to discuss the possible mechanisms to move from Small Smart Data to relevant Big Data.

Biography

Yuko SASA is a young researcher in the field of social robotics, finishing a PhD the LIG - computer science lab and financed by the Labex Persyval-lab. She completed a computational linguistics Master in 2012, and a Gerontechnology Master in 2013, at Grenoble Alps University. Her supervisors are V. Aubergé (LIG), G. Feng (Gipsa-Lab) and Y. Sagisaka (Waseda University). She is enrolled in several academic committees and was selected for international research programs as the French-American Doctoral Exchange Seminar on Cyber-Physical Systems (FadEx - French Embassy), or the ROW (Research Opportunities Week - Technical University of Munich).

yuko.sasa@univ-grenoble-alpes.fr

Notes:







4th International Conference on

BIG DATA ANALYSIS AND DATA MINING

September 07-08, 2017 | Paris, France

Keynote Forum

Day 2

Data Mining 2017

Mikhail Moshkov, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-013



Mikhail Moshkov

King Abdullah University of Science and Technology, Saudi Arabia

Extensions of dynamic programming: Applications for decision trees

In the presentation, we consider extensions of dynamic programming approach to the investigation of decision trees as algorithms for problem solving, as a way for knowledge extraction and representation, and as classifiers which, for a new object given by values of conditional attributes, define a value of the decision attribute. These extensions allow us to describe the set of optimal decision trees, to count the number of these trees, to make sequential optimization of decision trees relative to different criteria, to find the set of Pareto optimal points for two criteria, and to describe relationships between two criteria. The applications include the minimization of average depth for decision trees sorting eight elements (this question was open since 1968), improvement of upper bounds on the depth of decision trees for diagnosis of 0-1-faults in read-once combinatorial circuits over monotone basis, existence of totally optimal (with minimum depth and minimum number of nodes) decision trees for Boolean functions, study of time-memory tradeoff for decision trees for corner point detection, study of relationships between number and maximum length of decision rules derived from decision trees, study of accuracy-size tradeoff for decision trees which allows us to construct enough small and accurate decision trees for knowledge representation, and decision trees as classifiers, outperform often decision trees constructed by CART.

Biography

Mikhail Moshkov is a Professor in CEMSE division at King Abdullah University of Science and Technology (KAUST), Saudi Arabia since October 1, 2008. He completed his Master's degree at Nizhny Novgorod State University; Doctorate degree at Saratov State University and; Habilitation at Moscow State University. From 1977 to 2004, he was at Nizhny Novgorod State University. Since 2003, he worked at Institute of Computer Science, University of Silesia, Poland and; at Katowice Institute of Information Technologies since 2006. His main areas of research are complexity of algorithms, combinatorial optimization, and machine learning. He is an author and co-author of five research monographs published by Springer.

mikhail.moshkov@kaust.edu.sa

Ergi Sener, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-013



Ergi Sener
Bonbon Tech, Turkey

Disrupting retail analysis with patent applied Wi-Fi based advanced analytics solution - BonAir

In recent years, the increasing importance of big data has also led to big expectations. Par-ticularly with the introduction of the concept of Internet of Things (IoT), each object is linked to the internet and with the continuous increase in mobile and digital applications and ser-vices; data has been gathered at a surprising rate from various sources. When used and evaluated correctly, data has become a crucial competitive weapon, so in the technology world, data is frequently expressed as new gold or new oil. However, data does not represent a value by itself; value is formed as a result of processing data to solve a unique problem or fulfill a need. BonAir makes sense of big data by analyzing the data collected from customer visits, customer behaviors and customer profiles and uncovers the potential of big data and lead to provide competitive advantages for clients. BonAir is a disruptive technology innovation that understands and analyses in-store customer behavior (wait time, service time, visit frequencies etc.) without being connected to any Wi-Fi, or open Bluetooth or without having a smart phone application. With its unique technology, BonAir aims to perform real-time behavior-based analysis. Based on their needs, customers can be directed at the right time to the right location with optional app integration as well. BonAir solution is currently the widest Wi-Fi based analytics network in Turkey with more than 5.000 sensors deployed in the field. Some of the clients include Akbank, Benetton Turkey, BMW Turkey, Skechers Turkey, Lacoste Turkey, Burberry Turkey, etc.

Biography

Ergi Sener began his career as the Co-founder and Business Development Director of New Tone Technology Solutions in 2007 with the partnership of Sabancı Universi-ty's Venture Program. Between 2009 and 2013, he worked as a CRM specialist at Garanti Payment Sys-tems and as a Senior Product Manager in the New Technology Business Division of Turkcell. In 2013, he joined MasterCard as a business development and innovation manager for emerg-ing markets and managed the SEE cluster. He was also one of the Co-founders and the Man-aging Director of Metamorfoz ICT. He is currently acting as the CEO of Bonbon Tech - the biggest IoT focused new generation analytics company of Turkey.

ergi@bonbon.tech

Notes:

Fuad Aleskerov, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-013



Fuad Aleskerov

Higher School of Economics, Russia

Effective choice and ranking of alternatives in search and recommendation problems

The problem of the high computational complexity of most accurate algorithms in search, rank, and recommendation applications is critical when we deal with large datasets. The quadratic complexity may be unadmissible. Thus, the task is to develop efficient algorithms by consistent reduction of information and by the use of linear algorithms on the first steps. The problem of whether functions of several variables can be expressed as superposition of functions of fewer variables was firstly formulated by Hilbert in 1900 as the Hilbert's thirteenth problem. The answer to this general question for the class of continuous functions was given in 1957 by Arnold and Kolmogorov. For the class of choice functions, this matter was studied only by our team. A new effective method for search, ranking, and recommendation problems in large datasets is proposed based on superposition of choice functions. The developed algorithms have low computational complexity so they can be applied on big data. One of the main features of the method is the ability to identify the set of efficient options when one deals with large number of options or criteria. Another feature of the method is the ability to adjust its computational complexity. The application of the developed algorithms to the Microsoft LETOR dataset showed 35% higher efficiency comparing to the standard techniques (for instance, SVM). The proposed methods can be applied, for instance, for the selection of effective options in search and recommendation systems, decision support systems, Internet networks, traffic classification systems and other relevant fields.

Biography

Fuad Aleskerov is a leading Scientist in Mathematics and multicriterial choice and decision making theory. He is the Head of the International Laboratory of Decision Choice and Analysis and the Head of the Department of Mathematics for Economics of National Research University Higher School of Economics (Moscow, Russia). He has published 10 books, many articles in leading academic journals. He is a member of several scientific societies, Editorial Board Member of several journals, founder and head of many conferences and workshops. He has been an invited speaker on numerous international conferences, workshops and seminars.

fa201204@gmail.com

Notes:

Omar M Knio, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-013



Omar M Knio

King Abdullah University of Science and Technology, Saudi Arabia

Data enabled approaches to sensitivity analysis, calibration and risk visualization in general circulation models

This talk discusses the exploitation of large databases of model realizations for assessing model sensitivities to uncertain inputs and for calibrating physical parameters. Attention is focused on databases of individual realizations of ocean general circulation model, built through efficient sampling approaches. Attention is then focused on the use of sampling schemes to build suitable representations of the dependence of the model response on uncertain input data. Non-intrusive spectral projections and regularized regressions are used for this purpose. Bayesian inference formalism is then applied to update the uncertain inputs based on available measurements or observations. We illustrate the implementation of these techniques through extreme-scale applications, including inference physical parametrizations and quantitative assessment and visualization of forecast uncertainties.

Biography

Omar M Knio completed his PhD from MIT in 1990. He held a Post-doctoral position at MIT, before joining the Mechanical Engineering Faculty at Johns Hopkins University in 1991. In 2011, he joined Mechanical Engineering and Materials Science Department at Duke University. In 2013, he joined AMCS Program at KAUST, where he served as Deputy Director of the SRI Center for Uncertainty Quantification in Computational Science and Engineering. He has co-authored over 100 journal papers and two books.

omar.knio@kaust.edu.sa

Notes:







4th International Conference on

BIG DATA ANALYSIS AND DATA MINING

September 07-08, 2017 | Paris, France

Plenary Talk

Day 2

Bennett B Borden, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014



Bennett B Borden

Drinker Biddle & Reath, USA

Big data winners and losers: The ethics of algorithms

Analytic models are playing an increasing role in the development, delivery and availability of goods and services. Who gets access to what goods or services and at what price are increasingly influenced by algorithms. This may not matter when we're talking about a $0.25 coupon for a candy bar, but what about public goods and services like education, healthcare, and energy distribution? What about predicting who will get a job or how we will police our society? In this session, we will explore the socioeconomic impact of algorithms, the ethics of big data, and how to work ethics into our analytics projects.

Biography

Bennett B Borden is a Chief Data Scientist at Drinker Biddle & Reath. He is a globally recognized authority on the legal, technology and policy implications of information. His ground-breaking research into the use of machine-based learning and unstructured data for organizational insight is now being put to work in data-driven early warning systems for clients to detect and prevent corporate fraud and other misconduct. He received his Master of Science in Data Analytics at New York University and his JD at Georgetown University.

bennett.borden@dbr.com

Sumit Pal, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014



Sumit Pal

Independent Consultant , Boston, USA

SQL on big data – Technology, architecture and innovations

This presentation gives an exhaustive overview of the architectural and technological underpinnings of SQL on Hadoop solutions of today. It covers architectures of low latency SQL engines on Hadoop for structured, unstructured and streaming analytics as well as SQL for operational systems and operational analytics on Hadoop. The talk also covers the innovations happening in the space with probabilistic engines like BlinkDB to GPU based engines like MapD. With the rapid adoption of Hadoop in the enterprise it has become all important to build SQL engines on Hadoop for all kinds of workloads for almost all kind of end users and use cases. From low latency analytics based SQL to ACID based semantics on Hadoop for operational systems, to SQL for handling unstructured and streaming data, SQL is fast becoming the ligua-franca in the big data world too. The talk focuses on the exciting tools, technologies and innovations and their underlying architectures and the exciting road ahead in this space. This is a fiercely competitive landscape with vendors and innovators trying to capture mindshare and piece of the pie – with a whole suite of innovations like – index based SQL solutions in Hadoop to OLAP with Apache Kylin and Tajo to BlinkDB and MapD. Topics to be discussed in this presentation includes: Why SQL on Hadoop; challenges of SQL on Hadoop; SQL on Hadoop architectures for low latency analytics (Drill, Impala, Presto, SparkSQL, JethroData); SQL on Hadoop architecture for semi-structured data; SQL on Hadoop architecture for streaming data and operational analytics and; innovations (OLAP on Hadoop, probabilistic SQL engines, GPU based SQL solutions).

Biography

Sumit Pal has more than 22 years of experience in the software industry. He is a big data, visualisation and data science Consultant; a software architect and big data enthusiast and builds end-to-end data-driven analytic systems. He has worked for Microsoft (SQL server development team), Oracle (OLAP development team) and Verizon (Big Data analytics team). Currently, he works for multiple clients advising them on their data architectures and big data solutions and does hands on coding with Spark, Scala, Java and Python. He has spoken at big data conferences in Boston, Chicago, Las Vegas and Vancouver. He has extensive experience in building scalable systems across the stack from middle-tier, data tier to visualization for analytics applications, using BigData, NoSQL DB. He has expertise in DataBase Internals, Data Warehouses, Dimensional Modeling, Data Science with Scala, Java and Python and SQL. He started his career being part of SQLServer Development Team at Microsoft in 1996-97 and then as a Core Server Engineer for Oracle Corporation at their OLAP Development team in Boston, MA, USA. He has also worked at Verizon as an Associate Director for Big Data Architecture, where he strategized, managed, architected and developed platforms and solutions for analytics and machine learning applications. He has also served as Chief Architect at ModelN/LeapfrogRX (2006-2013) where he architected the middle tier core analytics platform with open source olap engine (Mondrian) on J2EE and solved some complex dimensional ETL, modelling and performance optimization problems. He has MS and BS in Computer Science.

palsumitpal@gmail.com

Notes:

Bertrand Hassani, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014



Bertrand Hassani[1, 2]

[1]University College London, UK

[2]Paris 1 University, France

Regulatory learning; what framework can be proposed to ensure that the financial environment can be controlled?

The arrival of big data strategies is threatening the latest trends in financial regulation related to the simplification of models and the enhancement of comparability of approaches chosen by the various entities. Indeed, the intrinsic dynamic philosophy of big data strategies is almost incompatible with the current legal and regulatory framework in particular the one related to model risk governance. Besides, as presented in our application to credit scoring, the model selection may also evolve dynamically forcing both practitioners and regulators to develop libraries of models, strategies allowing to switch from one to the other and supervising approaches allowing financial institutions to innovate in a risk mitigated environment. The purpose of this paper is therefore to analyse the issues related to the big data environment and in particular to machine learning models and to propose solutions to regulators, analyzing the features of each and every algorithm implemented, for instance a logistic regression, a support vector machine, a neural network, a random forest and gradient boosting.

Biography

Bertrand Hassani is an Associate Researcher at Paris 1 University and University College London. He wrote several articles dealing with Risk Measures, Data Combination, Scenario Analysis and Data Science. He was the Global Head of Research and Innovations of Santander Group for the risk division and is now the Chief Data Scientist of Capgemini Consulting. In this role, he aims at developing novel approaches to measure risk (financial and non-financial), to know the customers, to improve predictive analytics for supply chain (among others) relying on methodologies coming from the field of data science (data mining, machine learning, A.I., etc.).

bertrand.hassani@gmail.com

Notes:

Pierre Hansen, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014



Pierre Hansen

GERAD and HEC Montréal, Canada

Two applications of variable neighborhood search in data mining

Many problems can be expressed as global or combinatorial optimization problems, however due the vast increase in the availability of data bases, realistically sized instances cannot be solved in reasonable time. Therefore, one must be often content with approximate solutions obtained by heuristics. These heuristics can be studied systematically by some general frameworks or metaheuristics (genetic search, tabu search, simulated annealing, neuron networks, ant colonies and others). Variable Neighborhood Search (VNS) proceeds by systematic change of neighborhood in the descent phase towards a local minimum and in a perturbation phase to get out of the corresponding valley. VNS heuristics have been developed for many classical problems such as TSP, quadratic assignment, p-median, and others. Instances of the latter problem with 89600 entities in the Euclidean plane have been solved with an ex-post error not larger than 3%. In the last two decades, several discovery systems for graph theory have been proposed (Cvetkovic's Graph; Fajtlowicz's Graffiti; Caporossi and Hansen's AutoGraphiX (AGX)). AGX uses VNS to find systematically extremal graphs for many conjectures judged to be of interest. Aouchiche systematically studied relations between 20 graph invariants taken by pairs, considering the four basic operations (-, +, /, x). Conjectures were found in about 700 out of 1520 cases. A majority of these 1520 cases were easy and solved automatically by AGX. Examination of the extremal graphs found suggests open conjectures to be solved by graph theoretical means. This led to several tens of papers by various authors mainly from Serbia and from China.

Biography

Pierre Hansen is a Professor of Operations Research in the Department of Decision Sciences of HEC Montréal. His research is focused on “Combinatorial optimization, metaheuristics and graph theory”. With Nenad Mladenovic, he has developed the Variable Neighborhood Search metaheuristic, a general framework for building heuristics for a variety of combinatorial optimization and graph theory problems. He has received EURO Gold Medal in 1986 as well as other prizes. He is a member of Royal Society of Canada and author or co-author of 400 scientific papers. His first paper on VNS was cited almost 3000 times.

pierre.hansen@gerad.ca

Notes:

Parameshwar P Iyer, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014



Parameshwar P Iyer

Indian Institute of Science, India

From knowledge to wealth: Intellectual property and technology licensing

The Office of Intellectual Property and Technology Licensing (IPTeL) at the Indian Institute of Science is geared towards identifying and sourcing the IP at IISc, facilitating the steps towards protection of the IP, continuously safeguarding and prosecuting the IP, and finally enabling the commercial and social utilization of the IP, through appropriate means of licensing and technology transfer.

The specific activities and roles of IPTeL are as follows:

- Identifying the sources of Intellectual Property at the Indian Institute of Science;
- Inviting Invention Disclosures from the Inventor(s);
- Evaluation and Assessment of the above Invention Disclosure(s);
- Assisting the Inventor(s) with prior art search;
- Assigning appropriate Patent Attorney(s) to assist the Inventor(s) in drafting claims and specifications for the Invention(s);
- Informing the Inventor(s) of the various stages of progress during the filing, prosecution, and maintenance of the Intellectual Property;
- Assisting in the efforts of the Inventor(s) towards Technology Licensing, and other forms of Technology Transfer to bring the Invention to "commercial" and/or "social" practice.

Some of the interesting "inventions" from the Institute, and the accompanying "technology transfer" practices are highlighted in this paper. These include:

- Portable Washing Machine for Rural Areas
- Flood-resistant Septic Tank
- Optical nano sensor
- Electrical Gradient Augmented Fluid Filtration Apparatus
- Patient Transfer Device
- Electrochemical Test Cell

Biography

Parameshwar P Iyer holds a Bachelore degree from IIT Kharagpur, Master degree from University of Illinois, and a Doctorate from University of California. He has over 43 years of teaching, research, and consultancy experience in Engineering and Technology Management. He has held several important academic and administrative positions, including being the Founder Professor of an IIM in India. He has over 50 publications and four books to his credit. He has received several professional awards, including a prestigious Fellowship from Unversity of Illinois. Currently, he is a Principal Research Scientist at Indian Institute of Science.

piyer@mgmt.iisc.ernet.in

Nada Philip, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014



Nada Philip

Kingston University London, UK

Big data analytics and visualization for healthcare – A glance into the AEGLE platform features via its presentation layer

Recently there has been great interest in health data and on how to produce value out of it so that it can support optimized integrated clinical decision making about the health of individuals and reduce cost. Many initiatives exist at the European level and worldwide, pinpointing the importance and usefulness of healthcare big data. One of these initiatives is the AEGLE EU project which aims to generate value from the healthcare data value chain data with the vision to improve translational medicine and facilitate personalized and integrated care services and thus improving overall healthcare at all levels, to promote data-driven research across Europe and to serve as an enabler technology platform, enabling business growth in the field of big data analytics for healthcare. In this talk, we will be looking into the main technical features of the AEGLE platform through its presentation layer that act as the user interface to the Big Data analytics and visualization infrastructure and algorithms of the AEGLE platform. In particular, we will be looking into the type 2 diabetes (T2D) use case and present some of the initial results of scenarios that aims at using analytics and visualization tools to discover knowledge and the causing factors of diabetes complications and their correlation to medication modality as well as biological markers while providing the potential to act as a predictive tool of such complications.

Biography

Nada Philip is an Associate Professor in the field of Mobile Health at Kingston University London. She is the Founder of Digital Media for Health Research Group. She has been involved in many national and international mHealth projects in the areas of "Chronic disease management (e.g. diabetes, COPD and Cancer), big data analytics for health, decision support systems, telehealth and telecare in developing countries, social robotics, medical video streaming, namely; OTELO, Delphi, LTE-Health, WELCOME and AEGLE". She is the author and co-author of more than 50 journals, peer reviewed conferences and book chapters. She is a member of the review panels for many grants bodies, transactions, and conferences.

N.Philip@kingston.ac.uk

Notes:

Fairouz Kamareddine, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014



Fairouz Kamareddine

Heriot-Watt University, UK

Computerizing mathematical text, debugging proof assistants and object versus metal level

Mathematical texts can be computerized in many ways that capture differing amounts of the mathematical meaning. At one end, there is document imaging, which captures the arrangement of black marks on paper, while at the other end there are proof assistants (e.g., Mizar, Isabelle, Coq, etc.), which capture the full mathematical meaning and have proofs expressed in a formal foundation of mathematics. In between, there are computer type setting systems (e.g., LATEX and presentation MathML) and semantically oriented systems (e.g., Content MathML, OpenMath, OMDoc, etc.). In this talk, we advocate a style of computerization of mathematical texts which is flexible enough to connect the different approaches to computerization, which allows various degrees of formalization, and which is compatible with different logical frameworks (e.g., set theory, category theory, type theory, etc.) and proof systems. The basic idea is to allow a man-machine collaboration which weaves human input with machine computation at every step in the way. We propose that the huge step from informal mathematics to fully formalized mathematics be divided into smaller steps, each of which is a fully developed method in which human input is minimal. We also propose a method based on constraint generation and solving to debug proof assistants written in SML and reflect on the need and limitation of mixing the object and meta-level.

Biography

Fairouz Kamareddine has been involved in a number of worldwide consultancy assignments, especially on education and research issues working for United Nations and the EU. Her research interest includes Interface of Mathematics, Logic and Computer Science. She has held numerous invited positions at universities worldwide. She has well over 100 published articles and a number of books. She has played leading roles in the development, administration and implementation of interdisciplinary, international, and academic collaborations and networks.

F.D.Kamareddine@hw.ac.uk

Notes:







4th International Conference on

BIG DATA ANALYSIS AND DATA MINING

September 07-08, 2017 | Paris, France

Scientific Tracks & Abstracts

Day 2

Data Mining 2017

Major Sessions:

Data Mining Applications in Science, Engineering, Healthcare and Medicine | Data Mining Methods and Algorithms | Data Mining Tools and Software | Big Data Applications | Data Mining Tasks and Processes | Data Privacy and Ethics | Big Data Technologies | Social network analysis | Business Analytics | Search and data mining | Clustering

Session Chair
Fionn Murtagh
University of Huddersfield, UK

Session Co-Chair
Narasimha Murty Yedidi
Electronic Arts, USA

Session Introduction

Title: **An innovation framework for big data**
Knut Hinkelmann, FHNW University of Applied Sciences and Arts, Switzerland

Title: **Prune the inputs, increase data volume, or select a different classification method – a strategy to improve accuracy of classification**
Alla Sapronova, Uni Research Computing, Norway

Title: **Business analytics applications in budget modelling to improve network performance**
Eleni Rozaki, Institute of Technology Tallaght, Ireland

Title: **Powerful and novel multivariate statistical approaches in big data sets and in data mining with applications to bio-, medical-, and material-Informatics**
Derrick K. Rollins, Iowa State University, USA

Title: **Contextual exploration (ec3): A strategy for the detection, extraction and visualization of target data**
Christophe Jouis, LIP6 (CNRS -UPMC Sorbonne Université), France

Title: **Objective function-based clustering via near-Boolean optimization**
Giovanni Rossi, University of Bologna, Italy

Title: **Data mining via entropy: Application on trade database**
Abdul Basit, State Bank of Pakistan, Pakistan

YRF Session

YRF: **Implementation and performance comparison of wilson, holdout, and multiedit algorithm for edited k-nearest neighbor**
Febriana Misdianti, University of Manchester School of Computer Science, UK

YRF: **Deep learning classification for large behavioral data**
Sofie De Cnudde, Antwerp University, Belgium

Knut Hinkelmann et al., J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

An innovation framework for big data

Knut Hinkelmann, Prajakta Diwanji and **Hans Friedrich Witschel**
FHNW University of Applied Sciences and Arts Northwestern, Switzerland

Big data has challenged existing business models and provided new business opportunities. Companies capture large volume and variety of transactional data, which contains information about their customers, suppliers and operations. Advanced data analysis techniques can help to gain knowledge about patterns, trends and user behaviors from such large datasets. This knowledge empowers businesses to innovate new products, services and business models. Scientific literature discusses use cases of value creation, data analysis techniques and technologies supporting big data. According to recent studies, the main challenge faced by companies is the proper utilization of the knowledge extracted via data analysis to create meaningful innovations. The current innovation frameworks like Google's Design Sprint guide organizations to create innovative IT applications in an agile manner. Design thinking oriented innovation frameworks like the one from Beckman and Barry (2007) place a strong emphasis on observations, e.g. to understand customer behavior and to identify their (implicit) needs. In today's digitalized world, however, the observation of such behavior requires analyzing digital traces of on-line transactions and combining it with data from different sources. We therefore propose to develop an innovation framework for big data that helps companies to exploit the knowledge generated from such data present within or outside the organization. This framework will provide the best practices, data analysis tools and technologies that can guide the companies to innovate from big data. In order to give meaning to the identified patterns, the data analysis is combined with background knowledge represented in ontology.

Biography

Knut Hinkelmann is Professor for Information Systems and Head of the Master of Science in Business Information Systems at the FHNW University of Applied Sciences and Arts Northwestern Switzerland. He is a Research Associate at the University of Pretoria (South Africa) and Adjunct Professor at the University of Camerino (Italy). Prior to that he has worked at the German Research Center for Artificial Intelligence (DFKI) and as Product Manager Innovation for Insiders Information Management

knut.hinkelmann@fhnw.ch

Notes:

Alla Sapronova et al., J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Prune the inputs, increase data volume, or select a different classification method – A strategy to improve accuracy of classification

Alla Sapronova[1] and **Mark Sytse Ybema**[2]
[1]Uni Research Computing, Norway
[2]Sustainovate AS, Norway

Classification, the process of assigning data into labeled groups, is one of the most common operations in data mining. Classification can be used in predictive modeling to learn the relation between desired feature-vector and labeled classes. When the data set contains arbitrary big number of missed data and/or the amount of data samples is not adequate to the data complexity, it is important to define a strategy that allows reaching highest possible classification accuracy. In this work, authors present results on classification-based predictive model's accuracy for three different strategies: input pruning, semi-auto selection of various classification methods, and data volume increase. Authors suggest that a satisfactory level of model's accuracy can be reached when preliminary input pruning is used. The presented model is connecting fishing data with environmental variables. Even with limited number of samples, the model is able to resolve the type of the fish with up to 92% of accuracy. The results of using various classification methods are shown and suggestions are made towards defining the optimal strategy to build an accurate predictive model, opposed to common trial-and-error method. Different strategies for input pruning that assure information's preservation are described.

Biography

Alla Sapronova has completed her PhD at Moscow State University, Russia and Post-doctoral studies at UniFob, University of Bergen, Norway. She is the Head of Data Science at Center for Big Data Analysis, Uni Research, Norway. She has published more than 15 papers in reputed journals.

alla.sapronova@uni.no

Notes:

Eleni Rozaki, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Business analytics applications in budget modeling to improve network performance

Eleni Rozaki
Institute of Technology Tallaght, Ireland

Efficient and effective network performance monitoring is a vital part of telecommunications networks. Finding and correcting faults, however, can be very costly. In order to address network issues associated with financial losses a model in business analytics for budget planning is presented. The model that is presented is based on previous work of network fault detection, along with cost considerations and customer segmentation. This work focuses on data mining techniques to show the cost probability of the distribution of network alarms based on budget planning classification rules using predictive analytics to determine the minimum bandwidth costs that are possible with network optimization. The results of the tests performed show that reductions in optimization costs are possible; some test cases are clustered, of which the results were used to create a performance based budget model. The results also find out the clients' demographics, customers' churn and simultaneously the financial cost of network optimization in order to review an efficient budget process and improve expenditure prioritisation.

Biography

Eleni Rozaki completed an Honours degree in Economics and MSc degree in Quantitative methods and Informatics at University of Bari, Italy. Her PhD degree is in the area of "Data mining and business analytics in telecommunication networks at Cardiff University, United Kingdom". She has experience as a Data Analyst in IT and in the telecommunication industry in Ireland. She is currently working as an Associate Lecturer at Institute of Technology Tallaght, National College of Ireland and Dublin Institute of Technology. Her current research interests include "Business analytics, predictive modeling, decision support systems and data mining techniques".

Eleni.Rozaki@it-tallaght.ie

Notes:

Derrick K Rollins, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Powerful and novel multivariate statistical approaches in big data sets and in data mining with applications to bio-, medical-, and material-informatics

Derrick K Rollins
Iowa State University, USA

Advanced statistical methodologies have key roles to contribute in data mining and informatics for large data sets. Over the years, our research has developed a number of statistical techniques exploiting multivariate analysis and methodologies in many applications including bioinformatics, specifically, microarray data sets in a number of applications, medical informatics, including disease diagnosis and discovery, and in material informatics, including the development and evaluation of material properties and testing techniques. In this talk, we present the tools and methodologies that we have developed over the years and discuss their attributes and strengths. The two primary multivariate statistical methodologies that we have exploited have been principal component analysis (PCA) and cluster analysis (CA). This talk will break this technique down for the non-expert and then demonstrate their strengths in handling large data sets to extract critical information that can be exploited in analysis, inference, diagnosis and discovery.

Biography

Derrick K Rollins is a Professor of Chemical and Biological Engineering and Statistics. He completed his Graduation in Statistics and Chemical Engineering from Ohio State University. He worked at DuPont chemical company for seven years and four months as a Faculty Intern. He received various awards that include: The 2012 Tau Beta Pi McDonald Mentoring Award, the 1996 American Association for the Advancement of Science (AAAS) Mentor Award and National Science Foundation Presidential Faculty Fellows Award. His research areas include “Blood glucose monitoring, modeling and control, and (medical-, bio-, and material-) informatics”.

drollins@iastate.edu

Notes:

Christophe Jouis, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Contextual Exploration (EC3): A strategy for the detection, extraction and visualization of target data

Christophe Jouis
LIP6 (CNRS -UPMC Sorbonne Université), France

EC3 is intended to extract relevant information from large heterogeneous and multilingual text data, in particular in WEB 3.0. The project is based on an original method: contextual exploration. EC3 does not need syntactic analysis, statistical analysis or a "general" ontology. EC3 uses only small ontologies called "linguistic ontologies" that express the linguistic knowledge of a user who must concentrate on the relevant information from one point of view. This is why EC3 works very quickly on large corpus, whose components can be both whole books as well as "short texts": SMS to books. At the output, EC3 offers a visual representation of information using an original approach: the "Memory Islands". EC3 is implanted in the ACASA / LIP6 team. EC3 is tested on the large digitized corpus provided by the Labex OBVIL «Observatoire de la Vie Littéraire», in partnership with the Bibliothèque Nationale de France (http://obvil.paris-sorbonne.fr/). OBVIL intends to develop all the resources offered by digitization and computer applications to examine French literature from the sixteenth to the twentieth century, English and American literature, Italian literature, Spanish literature, in its most traditional formats and media Or the most innovative.

Biography

Christophe Jouis is assistant professor at the University Paris Sorbonne Nouvelle, France. He received a Ph.D. in Applied Mathematics at the “Ecole des Hautes Etudes en Sciences Sociales” (EHESS); and CAMS (“Centre d’Analyse et de Mathématiques Sociales”), OPTION: Science, Logic, Linguistics. From 2000 to 2004 he was associate professor in the Department of Computer Science at the University of Quebec at Trois-Rivieres (Canada), under the direction of Professor Ismail Biskri. In 2005, he joined the LIP6 ("Laboratoire d'Informatique de Paris 6), affiliated with the University Pierre et Marie Curie (UMPC) and the CNRS (France). Within the LIP6, he is currently a member of the research team ACASA (“Cognitive Agents and Automated Symbolic Learning”), led by Professor Jean-Gabriel Ganascia. His research interests are in natural language processing (NLP), cognitive sciences, ontology, typicality, data mining and information retrieval.

christophe.jouis@lip6.fr

Notes:

Giovanni Rossi, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Objective function-based clustering via near-Boolean optimization

Giovanni Rossi
University of Bologna, Italy

Objective function-based clustering is here looked as a maximum-weight set partitioning combinatorial optimization problem, with the instance given by a pseudo-Boolean (set) function assigning real-valued cluster scores (or costs, in case of minimization) to data subsets, while on every partition of data the global objective function takes the value given by the sum over clusters (or blocks) of their individual score. The instance may thus maximally consist of 2n reals, where n is the number of data, although in most cases the scores of singletons and pairs also fully determine the scores of larger clusters, in which the pseudo- Boolean function is quadratic. This work proposes to quantify the cluster score of fuzzy data subsets by means of the polynomial MLE (multi-linear extension) of pseudo-Boolean functions, thereby translating the original discrete optimization problem into a continuous framework. After analyzing in these terms, the modularity maximization problem, two further examples of quadratic cluster score functions for graph clustering are proposed, while also analyzing greedy (local and global) search strategies.

Biography

Giovanni Rossi completed his PhD in Economics and joined Computer Science department at University of Bologna. His research interests include "Discrete applied mathematics, combinatorial geometries, lattice and pseudo-Boolean functions, network community structure and graph clustering".

giovanni.rossi6@unibo.it

Notes:

Abdul Basit et al., J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Data mining via entropy: Application on trade database

Abdul Basit[1], **Muhammad Saeed**[1], **Asif Ali**[1] and **Zafar Iqbal**[2]
[1]State Bank of Pakistan, Pakistan
[2]National College of Business Administration & Economics, Pakistan

Entropy is a mathematical tool to gather maximum information about understanding distribution, systems, surveys and databases. We are introducing entropy as a tool of data mining which provides the maximum information about the trade behavior in different regions. In this study, we also derived a new entropy measure for data mining. This study will lead us to explore the new avenues of business and investment in Pakistan. China is the biggest player of global trade from Asian region. To expand the scope of competitiveness, China is continuously investing in different projects around the world. The China–Pakistan Economic Corridor (CPEC) is one of the major projects. The corridor is considered to be an extension of China's economic ambition One Belt-One Road initiative (OBOR). In future China wants to expand her trade with the world using the CPEC to enhance the scope of competitiveness. Pakistan also believes in open trade and continuously trying to enhance trade with the world. To attain maximum advantage of CPEC, Pakistan needs to explore the opportunities for the investors and business communities. In this study, we will develop linkages between the trends of our industries, commodities and their future demand in different regions.

Biography

Abdul Basit has completed his MS in Social Sciences at SZABIST, Pakistan. Currently, he is the PhD research scholar in the discipline of Statistics at National College of Business Administration & Economics Lahore, Pakistan. He is an Assistant Director of Statistics & DWH Department of State Bank of Pakistan. He has published four research papers in the journals and many articles were presented in national and international conferences.

basit_ravian917@hotmail.com

Notes:







4th International Conference on

BIG DATA ANALYSIS AND DATA MINING

September 07-08, 2017 | Paris, France

Video Presentation
Day 2

Data Mining 2017

Joseph Bonello et al., J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Automated dataset generator

Joseph Bonello, Sheryl Camilleri and **Conrad Attard**
University of Malta, Malta

Many IT organizations today spend more than 40% of the project's budget on software testing (Capgemini, 2015; Tata, 2015). One way of making the testing process easier is to automatically generate synthetic data that would be similar to the data inputted in the real environment. The benefits offered by an automated process are invaluable since real datasets can be hard to obtain due to data sensitivity and confidentiality issues. Automated Dataset Generator (ADaGe) enables the end-user to define and tweak a dataset definition through an interface in a way that satisfies the system's requirements. Our tool enables the automatic generation of small to large amount of synthesized data that can be outputted to a file or streamed to a real-time processing system.

Biography

Joseph Bonello is an Assistant Lecturer at University of Malta. He has worked on various EU related projects. He worked on various IT projects for the Department of Customs. He has managed the development of the IRRIIS project for a local member of the consortium. He has experience in developing commercial applications including work for local banking firms and government. He completed his BSc (Hons) in IT in 2002 at University of Malta and Master's degree at University of Malta on Automated Detection and Optimization of NP-Hard Scheduling Problems in Manufacturing.

joseph.bonello@um.edu.mt

Notes:

Jose Manuel Lopez-Guede, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Data Mining applied to Robotics

Jose Manuel Lopez-Guede
University of the Basque Country (UPV/EHU), Spain

One of the key activities of Data Mining is to discover and make clear hidden relations and working rules in complex systems. Robotics is a complex scope in which first prinicples approaches have been used to solve straight problems, but that approach is not enought to deal with complex problems, where more intelligence-based approaches are needed. Data Mining can be used for autonomous learning of control algorithms for Linked Multicomponent Robotic Systems (L-MCRS), because it is an open research field. Single Robot Hose Transport (SRHT) is a limit case of this kind of systems, when one robot moves the tip of a hose to a desired position, while the other hose extreme is attached to a source position. Reinforcement Learning (RL) algorithms have been applied to learn autonomously the robot control in SRHT, specifically Q-Learning and TRQ-Learning have been applied with success. However storing the state-action value functional information in tabular form produces large and intractable data structures, and using the Data Mining approach the problem can be addressed by discivering and learning the state-action values of Q-table by Extreme Learning Machines (ELM), obtaining a data reduction because the number of ELM parameters is much less than the Q-table's size. Moreover, ELM implements a continuous map which can produce compact representations of the Q-table, and generalizations to increased space resolution and unknown situations.

Biography

Jose Manuel Lopez-Guede received the Ph.D. degree in Computer Sciences from Basque Country University. He got 3 investigation grants and worked in a company 4 years. Since 2004 he worked as full time Lecturer and since 2012 as assoc. prof. He has been involved in 24 competitive projects and published more than 100 papers, 25 on Educational Innovation and the remaining in specific research areas. He has 20 ISI JCR publications, more than 15 other journals and more than 40 conferences. He has belonged to more than 10 organizing committees of international conferences and to more than 15 scientific committees.

jm.lopez@ehu.es

Notes:







4th International Conference on

BIG DATA ANALYSIS AND DATA MINING

September 07-08, 2017 | Paris, France

Young Researcher Forum

Day 2

Febriana Misdianti, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Implementation and performance comparison of Wilson, holdout, and Multiedit algorithm for edited k-nearest neighbor

Febriana Misdianti
University of Manchester School of Computer Science, UK

K-nearest neighbor (k-NN) classifier is widely used for classifying data in various segments. However, k-NN classifier has high computational cost because it uses linear search through all training data. In naïve implementation, k-NN will go through all training set i to compute its distance d from input data (O(nd)). Then, it loops again for all training set n to find k smallest results (O(nk)). So, the overall time complexity for k-NN is O(nd+nk). Thus, it is not suitable to classify multidimensional data with a huge number of training set. Meanwhile, k-NN needs a large number of samples in order to work well. There have been several ideas proposed to increase the time performance of k-NN in predicting a test data. One of the popular ideas is by reducing the number of TS in a model so that it would cut the testing time as the number of data that need to be explored becomes smaller. The aim of this experiment is to implement k-NN editing algorithms that cut the number of training data so that it become faster in predicting an input. This experiment implements three editing algorithms, namely, Wilson's editing, holdout editing, and Multiedit algorithm, and also compare the performance of them.

Biography

Febriana Misdianti is a Post-graduate student at University of Manchester. She has completed her Bachelor degree of Computer Science at Universitas Indonesia and has two years of working experience in startup companies in Jakarta and Singapore. She has been winning a lot of competitions related to computer science and has published a paper about data security in a reputable journal.

febriana.misdianti@gmail.com

Notes:

Sofie De Cnudde, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-014

Deep learning classification for large behavioral data

Sofie De Cnudde
Antwerp University, Belgium

Deep learning has demonstrated significant performance improvements over traditional state-of-the-art classification techniques in knowledge discovery fields such as computer vision and natural language processing. Moreover, the representation-learning nature of the deep learning architecture enables intuitive interpretation of these complex models. The predictive analysis of very high-dimensional human behavioral data (originating from contexts such as e-commerce) could highly benefit from a complex classification model on the one hand and from an intuitive insight in the many fine-grained features on the other hand. The research question we investigate is whether and how deep learning classification can extend its superior results to this omnipresent type of big data. We present the following three contributions. First and foremost, the results of applying deep learning on large, sparse behavioral data sets demonstrate as good as or better results compared to shallow classifiers. We find significant performance improvements over linear support vector machines, logistic regression with stochastic gradient descent and a relational classifier. Moreover, we shed light on hyper-parameter values to facilitate the adoption of deep learning techniques in practice. The results demonstrate that an unsupervised pre-training step does not improve classification performance and that a tanh non-linearity achieves the best predictive performance. Lastly, we entangle the meaning of the neurons in a manner that is intuitive for researchers and practitioners, and show that the separate neurons identify more nuances in the many fine-grained features compared to the shallow classifiers.

Biography

Sofie De Cnudde is currently finishing her PhD at the University of Antwerp in Belgium. Her PhD analyses and compares the performance of wide and deep classification techniques on large and sparse human behavioral data. She has given tutorials on the use of deep learning for large-scale classification and has currently finished a research stay at an e-commerce company in London. Her work has been published in Expert Systems with Applications, Decision Support Systems and has been presented at conferences such as IFORS.

Sofie.DeCnudde@uantwerpen.be

Notes:







4th International Conference on

BIG DATA ANALYSIS AND DATA MINING

September 07-08, 2017 | Paris, France

e-Posters

Data Mining 2017

Anne Marie Smith, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-015

Injecting big data into enterprise data management programs in accordance with the DAMA-DMBOK©

Anne Marie Smith
EW Solutions, USA

Although many organizations have established the basic foundation of an enterprise data management framework, they do not know how to include the emerging concepts of big data management into the established components. Therefore, these organizations struggle with implementing big data management into their organization's data management activities. This presentation will address the injection of big data concepts and technologies into the normal operation of an organization's enterprise data management framework. Using the DAMA-DMBOK © wheel, the presentation will outline a methodology for including big data into an enterprise's approach to data management. Learning objectives include Big data foundations – part of an enterprise data management program; big data management implementation phases and steps – a practical approach to a methodology; challenges to a successful enterprise data management program; risks through addition of big data; cultural and organizational alignment; adapting for cultural change with big data management.

Biography

Anne Marie Smith completed her PhD at North Central University and Post-doctoral studies at La Salle University. She is the Vice President and Chief Methodologist of EW Solutions, a premier enterprise data management and data science organization. She has published more than 25 papers in reputed journals and has been serving as an Editorial Board Member for a variety of publications.

amsmith@ewsolutions.com

Notes:

Claudia Kleinschrodt et al., J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-015

Optimization of the data transfer of tool models by manipulation of the transfer data files

Claudia Kleinschrodt, Reinhard Hackenschmidt and **Frank Rieg**
University of Bayreuth, Germany

Data exchange is the foundation of a network world. Through the global operating range of companies, complex supplier structures, media or system breaks within processes, the topic also gains relevance. This applies to all sectors in which data is generated. In consideration of the manufacturing industry, the exchange of tool models plays a crucial role. For instance, for the management of server-based catalogues or the virtual commissioning of machines, it is important to transfer the developers' 3D models to the operators. Although many institutions and committees are developing standards and guidelines to ensure an accurate transfer of information, an automated error-free data exchange, necessary for the vision of a smart factory, is still not possible. Investigations show the variety of problems during the data exchange of 3D CAD models via the neutral data format STEP. Detailed analyses of the course of action can be used to identify and classify faults. Based on these findings, it is possible to develop remedies, which are adapted to the problem causes. In case of interface-related errors the manipulation of the transfer files is a useful method to increase compatibility and information content.

Biography

Claudia Kleinschrodt studied Environmental and Bio-engineering (Materials and Process technology) at University of Bayreuth. She completed her Diploma in 2014 and then worked as a Research Assistant at University of Bayreuth. Her research focuses on 3D data exchange of precision tools.

claudia.kleinschrodt@uni-bayreuth.de

Notes:

Si Fan, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-015

Learning Analytics: A Growing Interest for Big Data Analytics in Australian Higher Education

Si Fan
University of Tasmania, Australia

Big Data analytics is commonly used to inform decisions, particularly in the areas of business, health care, science, and agriculture (Daniel, 2015). In recent years, there has also been an increase in the use of Big Data analytics for educational purposes, driven and facilitated by various reasons. First, educational institutions are facing more pressure towards performance management, metrics and quantification. Secondly, learning tends to generate large data sets, especially in Learning Management Systems and Massively Open Online Courses. One most common form of Big Data analytics in higher education is Learning Analytics. Through the analysis of educational data, learning analytics applications enable educational institutions to build predictive models on student outcomes, especially to identify at-risk students (students will potentially drop out or fail) and to provide early warnings so that interventions can be made (Rienties et al., 2016; Timms, 2015). The project extracted learning data from a learning management system, which is based on the Brightspace platform (Desire2Learning, 2017), used at one Australian university. The data were analysed through Excel, R and Python, and models were developed to detect factors that can be used as predictors and measurements for student enagagement and performance, for instance, the length of time spent by students in online discussion boards, and for accessing online lectures and reading materials, as well as lecturer and tutor inputs. As an early exploration in the research field, this project demonstrates one example of how Learning Analytics can be used to support decision making in Australian higher education.

Biography

Si Fan is a lecturer in the School of Education at the University of Tasmania, Australia. She completed her PhD at the same university in 2011. She has been involved in a number of research projects, and has a broad research interest in higher education, learning analytics, early childhood education, educational technologies, Big Data analytics for educational purposes, and online language education.

si.fan@utas.edu.au

Notes:

A Veerendhar, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-015

Model development for P-glycoprotein inhibitors using Random forest method

A Veerendhar
Indian Statistical Institute, India

Drug discovery is very expensive as it involves expenditure of large amount of money, manpower and time. It involves biological evalution of large set of diverse molecules. However, now a days drug discovery has been facilitated by numerous computational methodologies to shorten the lengthy process. QSAR (Quantitative structure-activity relationship) tools are well know for the prediction of activity of closely related analogs. However these methods significantly depends on the nature of training sets and have high predictive power when applied to compounds structurally related to the compounds included in the training set. And have poor predictive power when applied to structually diverse molecules. In the current study of P-glycoprotein inhibitors. Model development carried out on more than 150 compounds. The compounds along with activities has been devided into training sets (for model development) and test compounds (for validation). Descriptors generaed using DRAGONH. Random forest (Number of tress used is 1000) used for model development with Q2 more than 0.75.

Biography

Veerendhar has completed his PhD in Organic Chemistry from University of Hyderabad and Postdoctoral studies from The Hebrew University of Jerusalem, Israel. He has two US patents to his credit. He is working as a Principal Scientist in a premier Pharmaceutical industry and also persuing Business Analytics and Master Black Belt at Indian Statistical Institute. He has published his research work in eight reputed journals.

veerendharainelly@gmail.com

Notes:

Antonio J Tallon-Ballesteros, J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-015

Addressing the suitability of the activation function kind on feed-forward neural networks

Antonio J Tallon-Ballesteros
University of Seville, Spain

This paper analyses the suitability to include hyperbolic tangent units within neural networks based on Multilayer Perceptron (MLP) trained by means of a back propagation algorithm. The experimentation is conducted out in classical problems from the scope of health and also with data sets from high dimensional domains. As a main conclusion could be drawn that hyperbolic tangent are very handy. On the other hand, the training of this kind of neural networks is a drawback. The setting of the parameters is an issue. The results, compared to those of sigmoid activation function, are very promising and very often better. This is the starting point to design other kind of activation functions.

Biography

Antonio J. Tallón-Ballesteros completed his PhD in 2013 from the University of Seville. He has been a Lecturer since 2005. He got the accreditation to Senior Lecturer in 2013. He has recently been a Programme Co-chair of some conferences such IDEAL 2016, ICDIM 2016 and FSDM 2016.He has been an active member of the programme committe of a great deal of conferences in the last 10 years, such as HIS, ICADIWT, FUTURE COMPUTING, ISDA, IWINAC, INTECH and NaBIC. He has published more than 40 publications. Among them, the selection could consists of 6 papers which are listed in the Journal Citation Reports by Thomson Reuters.

atallon@us.es

Notes:







4th International Conference on

BIG DATA ANALYSIS AND DATA MINING

September 07-08, 2017 | Paris, France

Accepted Abstracts

Data Mining 2017

J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-015

Linking academia, industry, and government in developing disruptive big data analytics technologies

Borko Furht
Florida Atlantic University, USA

In this talk, we first introduce the NSF-sponsored Industry/University Cooperative Center for Advanced Knowledge Enablement at Florida Atlantic University (FAU), which presently has 34 industry members with about $5 million in funding. The center is successfully building a bridge linking academia, industry, and government in a coordinated research initiative. We present non-traditional, radical university arrangements that we implement in order to create an entrepreneurial university. We describe several applied research projects conducted within the center to develop disruptive technologies in the areas of big data analytics, machine learning, and multimedia data mining. The specific projects include modeling Ebola spread using big data analytics, developing innovative machine learning technologies for HPCC platform, medical image analysis using deep learning techniques, and deep learning algorithms for use cases in auto industry. These projects are initiated by industry partners who are members of the center and who are interested to apply the obtained research results and create successful commercial products.

bfurht@fau.edu

Vocational computing: A data mining application for the workplace

Fillia Makedon
University of Texas Arlington, USA

According to the US Department of Labor, thousands of workers die on the job each year because of accidents, or lack of training in using new technologies. Computational methods can be used to provide evidence-based quantitative assessments of worker ability and identify needs for training. Data mining methods can be applied to the analysis of performance multi-sensing interaction data collected while a person performs a certain work task. We describe the iWork smart service work-assessment system with recommendations of personalized intervention, following multimodal data mining of activity data. The service assesses mental, cognitive and physical skills of a worker for improved placement and informed decision-making. The proposed service takes advantage of recent advancements in robotics, sensing technologies, and intelligent communication platforms to enhance human ability to learn by interactive experiences. The service trains assistive workplace robots to provide personalized help to complete difficult cognitive and/or physical tasks in the workplace. A new machine learning methodology is described and demonstrated.

makedon@uta.edu

J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-015

Using data analytics to drive operational efficiency in an African call center

Megan Yates
Ixio Analytics, South Africa

The operations division of a large multinational company offering subscription-based services across Africa used Ixio Analytics to create a predictive model for call volumes in their inbound support call centre. The model previously used was overestimating call volumes, costing the operations division heavily in agent staffing and resulting in inefficient call scheduling. This situation had persisted for several years. Ixio Analytics used call volume data at half hour intervals from January 2012 to create a predictive time series model. We used an ensemble modeling approach, combining a time series forecast model and data splitting at key time intervals. This was the first known application of this modeling method to call volume forecasting. The model took into account seasonal, random and trend components in the data. Total call volumes for every 30 minute period, as well as call volumes for various types of calls (such as billing) were predicted. The Ixio model has achieved 94% accuracy since implementation. This is a significant improvement from the previous model, which achieved only 70% accuracy. The Ixio model has been implemented and used in the company's workforce planning. The accuracy of predictions has enabled efficient workforce planning and an increase in call scheduling. This is currently saving the company approximately USD 5 million annually. A subsequent iteration of the model now includes event data that has further improved performance.

megan@ixioanalytics.com

Comparative evaluation of the mutual information class of ICA with other classes: A case study on denoising EEG

Janett Williams
Brescia University, USA

Statement of the Problem: Independent component analysis (ICA) is a popular blind source separation (BSS) technique that has proven to be promising for the analysis of electroencephalography (EEG) data. A number of ICA approaches have been used for EEG data analysis, and even more ICA algorithms exist, however, the impact of using different algorithms on the results is largely unexplored. The purpose of this study is to examine the performance of different classes of algorithms when they are applied to EEG data.

Methodology & Theoretical Orientation: Four major classes of algorithms for spatial ICA were examined. They were information maximization, maximization of non-Gaussianity, joint diagonalization of cross-cumulant matrices, and second-order correlation based methods. Each algorithm performance was evaluated using several analysis techniques.

Findings: The results showed that the ICA algorithms using mutual information estimation proved to have better performance at separating signals.

Conclusion & Significance: This research supports the claims made by researchers that algorithms using mutual information estimators produced more accurate ICA algorithms. With this information, it is recommended that for creation of any new algorithms which has great separation performance it should use a mutual information based estimator.

janett.williams@brescia.edu

J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-015

Opinion analysis of the university library readers'

Oukrich Jaouad
Sultane Moulay Slimane University, Morocco

The last years were known to have a big particular interest about the opinion analysis. However, the evaluation of this task was not satisfactory because of the rarity or of the non-availability of the corpuses annotated by various languages and data processing methods have changed from analytics (ETL, SQL, and Reporting) into Big Data (Hadoop, NoSQL). This article describes our contribution on the opinion fine detection of the university library readers' in social media, mainly the treatment of Big Data, which is extracted from the Facebook. The approach adopted for the readers' opinion analysis (combination of the statistical and symbolic methods) will be carried out by using a Java Hadoop Framework. Firstly, we describe the construction of the vocabulary that allows characterizing a positive, neutral or a negative opinion applied to the study of readers' satisfaction. Secondly, we introduce the application results by using machine-learning techniques for the opinion classification task.

oukrichjaouad@gmail.com

Analyzing large scale recurrent event data using parametric frailty model: A divide-and-conquer approach

Jerry Q Cheng
Rutgers University, USA

Currently when analyzing large scale recurrent event data, we are facing challenges such as memory limitations, un-scalable computing time, etc. In this research, we propose a divide-and-conquer approach using parametric frailty models. Specifically, we randomly divide the data into many subsets and obtain the maximum likelihood estimator from each individual data set. Then, we propose a weighted method to combine these individual estimators as the final estimator. We show that this divide-and-conquer estimator is asymptotically equivalent to the estimator based on the full data. We conduct simulation studies to demonstrate the performance of this proposed method and apply it to a large real data set of repeated heart failure hospitalizations.

jcheng1@rwjms.rutgers.edu

J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-015

Big data challenges and opportunities in the developing world

Kassim S Mwitondi
Sheffield Hallam University, UK

Population growth and globalization bring numerous challenges - food security, environmental degradation, peace and security. Human-related activities and natural phenomena immensely impinge on our livelihood and sustainability of life on earth. The 17 UN Sustainable Development Goals (SDGs) exhibit the multivariate nature of sustainable development. In a typical developing country, these data attributes will be high-dimensional, multi-faceted and fragmented. This talk will explore and attempt to balance the technical and logistical challenges and opportunities of extracting knowledge from these attributes. It touches on the intricacy of harnessing and gaining meaningful insights for supporting informed decision-making processes. Illustrations are based on enhanced unsupervised and supervised modeling tools, methods and techniques and the results underline the need for interdisciplinary collaborative research. The capacity for these tools and methods to learn rules from data are typically constrained by the random nature of training and test data; diversity and disparity of models and related parameters and limitations in data sharing. Hence, the mantra of collaborative interdisciplinary research highlights typical challenges and opportunities that derive from shared knowledge and openness as to who is doing or did what, where, when and how. A prototype project idea Agro-Data Dynamics (ADD) - an open data-driven initiative for improving farming practices, productivity & human livelihood in East Africa is used to illustrate existing challenges and opportunities. A framework for agro-forestry data repositories on soils, vegetation, income, and other socio-economic and cultural attributes is proposed for formalizing what we know, what we don't, what we can and can't do.

k.mwitondi@shu.ac.uk

Extraction of association rules based on the genetic algorithm and data mining techniques

Lylia Betit
University of Science and Technology Houari Boumediene, Algeria

We are interested in our project to design and implement a tool dedicated to the automatic generation of association rules from a database. In order to generate the most relevant rules, we have considered in this work three different approaches: extraction of rules from three data mining algorithms (Apriori, Close, Chram); enrichment of approach one by the application of the genetic algorithm and; use of the genetic algorithm before applying any of the data mining algorithms. In addition, we proposed two measures of relevance: weighted dominance and mean support. A modeling with the UML language was done before developing the extraction tool that supports our three approaches. Several experiments were carried out with a basis related to the purchases of products. The results show that the third approach has generated the most relevant rules.

betit.lilya@gmail.com

J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-015

Data oriented analysis on atomic collisions

Maged Mostafa
MSA University, Egypt

Particle collision is a growing field that aims to discover the science of the future. Very few particle colliders are built, but the most powerful one is the Large Hadron Collider (LHC). This project is a data oriented analysis on the experiments made in the LHC which implies two main tasks. Firstly, studying the science behind the particle collisions, and secondly, analyzing the data resulting from collisions. The proposed analysis in this project finds tracks of Higgs boson particle. On 04 July 2012, ATLAS and CMS, the two biggest LHC detectors, announced the discovery of a new boson with a mass around 125 GeV, which by the preliminary data analysis appears to be the Higgs boson, which completes the Standard Model (SM), and fills its breaches, regarding how particles acquire mass. Agile methodology is used to manage the whole project. The data that author has worked on is released by CERN (European Organization for Nuclear Research).

maged.m.anwar@gmail.com

Memory management in modern data analytics platforms

Mayuresh Kunjir
Duke University, USA

The field of data processing is seeing an explosion of computation platforms such as Hadoop, Spark, HBase, Cassandra, Flink, Storm, and others. Some key observations of these platforms are: They are JVM-based and therefore developer friendly, stable, and secure; they run in multi-tenant environments where resources are isolated using containers, e.g. Yarn, Mesos, and Docker and; they use memory extensively both for data storage and computations. The central premise of this work is that data analytics will increasingly be done in memory on shared-nothing clusters running JVM inside containers. Given this, it is high time for a better understanding of memory management in these modern data analytics platforms. This work uses Thoth, a data-driven platform for multi-system cluster management developed at Duke University, to build a deep understanding of different interplays of memory-management decisions made at multiple levels such as the resource manager (e.g. Yarn), the application framework (e.g. Spark), and the JVM. An empirical study is carried out over a suite of industry-standard analytical workloads to show the impact of various memory management options on response time, efficiency, and reliability. A set of guidelines follows to ensure a safe and efficient execution of the workloads.

mayuresh.kunjir@gmail.com

J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-015

Relevance of the accuracy of financial data in research work

Narcisa Roxana Mosteanu
American University in the Emirates, UAE

This study aims to show how important data is in economic research, especially that which refers to public macroeconomic indicators. Data lead us to take financial decision. However, if the data is not accurate, the financial decision taken based on this can drive the business in an appropriate direction, not the one we considered to be. The research will present how, sometimes the data presented as official one is not reflecting the reality, and sometimes the way that the data are presented have a different mining as in the reality. Nowadays, researchers relied in their analyses more and more on the information provided by official websites. What the author faced in other research paper was that for the same indicators the data can be different, and, in other times, even if the data are the same, in the reality figures are dissimilar.

narcisa.mosteanu@aue.ae

Development of a big data environment-based interactive fashion design platform for optimizing textile supply chains

Xianyi Zeng
ENSAIT Textile Institute, France

For a modern textile company, under the pressure of international competitions, making quick and reactive decisions on organization of production and supply network have been increasingly important and critical for obtaining competitive advantages. The development of information technologies, like cloud computing, big data and virtual reality, provides the possibility of transforming the classical labor intensive international supply chain into digital supply chains, composed of networked units for design and small series local production. These new supply chains should be organized around consumers' personalized requirements and facilitate connections and interactions between all concerned parties. In this context, this paper gives a summary of the recent work realized in our research group for building the digital interactive garment design platform in big data environment. The main force of this platform is to provide a series of data-based services, including a designer-oriented and a consumer-oriented recommendation system, an interactive virtual garment fitting process, a human body modeling and classification procedure, a normalized sensory virtual garment evaluation procedure, and a feedback mechanism for adjusting design parameters according to consumer's and designer's perception and learning from past successful design cases. The treatment of human perception and emotions related to fashion products and their ambiances has been integrated into these data services. The intelligent techniques, including fuzzy logic, design trees, rough sets, and cognitive maps have been successfully used for formalizing concepts and modeling complex relations at different stages.

xianyi.zeng@ensait.fr

J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-015

The particularities of big data ethics

Renee Carpentier
University of Quebec in Montreal, Canada

The future technologies for knowledge discoveries in data are full of promises and benefits for research advancements in various fields of expertise such as medicine, environment and social sciences but also involve risks and ethics issues requiring a better understanding to avoid significant errors in our decision making. This presentation will concentrate on the ethics of big data in order to highlight their particularities mostly based on the technical capacities to connect different data sources making knowledge discoveries more effective. Ethics are part of the research world since a long time especially when human subjects are involved in order to ensure privacy and avoid abusive treatments during experimentations. When it comes to the ethics of big data, some particularities need to be understood in order to take the right decisions and elaborate good policies. We will see that these particularities of big data ethics include the unprecise ownership and control of data, the sharing and re-use of data across diverse contexts and the diversity of actors and interests involved bringing new power distributions between big data collectors, big data users, and big data generators. This context specific to big data requires new ways of thinking ethics such as the ethics of care recently elaborated in order to develop models of social property of data focusing on sharing and collaboration and considering big data as a social resource rather than a private resource.

carpentier.renee@courrier.uqam.ca

Big data analytics and knowledge engineering - The power of analytics in value creation

Soraya Sedkaoui
Khemis Miliana University, Algeria

Since the advent of IT and the internet, the amount of data stored in digital form has been rapidly growing. An increasing number of data silos are created across the world, which means that this growth will never stop. The purpose of this article is to show that the challenges in the age of data revolution focus on the integration and utilization of data. The objective of this work is to discuss the analytics dimension of big data and its contribution for decision-making. It shows that big data analytics offers strong support in the knowledge discovery and the decision-making process.

soraya.sedkaoui@gmail.com

J Comput Eng Inf Technol, 6:5
DOI: 10.4172/2324-9307-C1-015

Synergies between data mining and operations research: A new paradigm of business analytics

Pang Yan James
NUS Business School, Singapore

Both data mining and operations research approaches have long-established history in the data-driven decision making area. With the recent development in big data and business analytics, the data mining people and operations research people talk to each other more and more. The synergies between data mining and operations research are important for many real-life business analytics problems. In this talk, we will discuss the fundamental linkage between these two domains. This talk will also cover the typical industry use cases which require the synergies between data mining and operations research techniques for building the solutions. The open research topics in this area will be also discussed in this talk.

jamespang@nus.edu.sg

Medical information retrieval: Semantic approach for answering PICO clinical questions

Eya Znaidi
Paris 1 Panthéon Sorbonne University, France

Answering clinical questions give rise to numerous challenges among which term ambiguity and relevane estimation based on the distribution of the query facets in the documents.We address the issue of answering PICO (Patient/Problem, Intervention, Comparison, Outcome) clinical queries. Our contributions of this area include: A new document ranking model based on a prioritized aggregation operator that computes the global relevance score based on the relevance estimation of the semantic facet sub-queries and; leverages the importance of the facets according to the document and query under evaluation. Using this semantic representation, we proposed the use of a prioritized document relevance scoring operator that jointly uses the graded importance of the facet in an EBM task and the topical relevance of the document. The effectiveness of the proposed search approach is empirically evaluated using a clinical retrieval collection including 423 queries and more than 1.2 million of medical abstracts from PubMed. The experimental results show that our approach for PICO query answering significantly overpasses state-of-the-art document ranking models.

znaidi.eya@gmail.com

INDEX

A Veerendhar	83	Fionn Murtagh	23	Nikolaos Freris	31
Abdul Basit	69	Fred Jacquet	36	Omar M Knio	51
Alla Sapronova	64	Fuad Aleskerov	50	Parameshwar P Iyer	58
Anne Marie Smith	80	Giovanni Rossi	68	Pavel Kisilev	29
Bennett B Borden	30	Jenny Lundberg	41	Petra Perne	22
Bennett B Borden	54	Jose Manuel Lopez-Guede	73	Pierre Hansen	57
Bertrand Hassani	56	Joseph Bonello	72	Prajakta Diwanji	45
Boris Mirkin	27	Karmen Kern Pipan	26	S N Mohanty	40
Christophe Jouis	67	Kazumi Nakamatsu	28	SASA Yuko	46
Claudia Kleinschrodt	81	Knut Hinkelmann	63	Si Fan	82
Derrick K Rollins	66	Kruy Seng	44	Sofie De Cnudde	77
Eleni Rozaki	65	Michael Valivullah	20	Sumit Pal	55
En-Bing Lin	21	Mikhail Moshkov	48	Tommi Kärkkäinen	35
Ergi Sener	49	Morgan C Wang	37	Witold Dzwinel	38
Fairouz Kamareddine	60	Nada Philip	59		
Febriana Misdianti	76	Nicolas Wesner	39		

Meet Inspiring Speakers and Experts at our 3000+ Global Events with 1000+ Conferences, 1000+ Symposiums and 1000+ Workshops on Medical, Pharma, Engineering, Science, Technology, Business and 700+ Open Access Journals

Conferences | Publishing | Global Networking | Research | Business

Exhibitions | Collaborations | Careers | International Memberships | Advertisements & Promotions

Conferences By Continent

USA & Americas

USA | Brazil | Canada | Mexico

Asia-Pacific

Australia | China | India | Japan | Malaysia | New Zealand | Philippines | Singapore | South Korea

Europe & Middle East

Austria | Denmark | Finland | France | Germany | Italy | Netherlands | Norway | Poland | South Africa | Spain | Switzerland | Turkey | UAE | UK | Ukraine

Medical & Clinical Journals and Conferences

- Alternative Healthcare
- Cardiology
- Dentistry
- Dermatology
- Diabetes & Endocrinology
- Gastroenterology
- Healthcare Management
- Hematology
- Infectious Diseases
- Medical Ethics
- Health Policies
- Neuroscience
- Nephrology
- Nursing
- Obesity
- Oncology & Cancer
- Ophthalmology
- Palliativecare
- Pathology
- Pediatrics
- Physical Therapy Rehabilitation
- Psychiatry
- Pulmonology
- Radiology
- Surgery
- Vaccines
- Reproductive Medicine &
- Women Healthcare

Journals and Conferences by Subject

- Agri, Food & Aqua
- Biochemistry
- Business & Management
- Chemical Engineering
- Chemistry
- EEE & Engineering
- Environmental Science
- Genetics & Molecular Biology
- Geology & Earth Science
- Immunology
- Materials Science
- Microbiology
- Nanotechnology
- Nutrition
- Oncology & Cancer
- Petroleum
- Pharma Marketing & Industry
- Pharmaceutical Sciences
- Physics
- Veterinary

Contact us

Conference Series - America
One Commerce Center-1201, Orange St. #600
Wilmington, Zip 19899, Delaware, USA
P: +1-702-508-5200, F: +1-650-618-1417
E: america@conferenceseries.com

Conference Series - UK
Kemp House, 152 City Road
London EC1V 2NX, UK
Toll Free: +1-800-014-8923
E: europe@conferenceseries.com

Conference Series - Asia Pacific
6th Floor, North Block, Divyasree Building, Raidurg, Hyderabad
Telangana, INDIA-500032 P: +91-40-33432300
E: asia@conferenceseries.com

Toll Free

Australia - 1-800-651-097 | **France** - 0805-080048 | **Hyderabad** - 91-40-71279012

Hyderabad - 91-40-71279013 | **UK** - 0-800-014-8923

Los Angeles Downtown - 1-213-233-9462 | **USA/Canada** - 1-888-843-8169



Bookmark your dates

5th International conference on

Big Data Analysis and Data Mining

June 20-21, 2018 | Rome, Italy

E-mail: bigdata@enggconferences.com; dataanalysis@enggconferences.com

Website: datamining.conferenceseries.com