

18th Industrial Conference on Advances in Data Mining (ICDM 2018)

Poster Proceedings

New York, New York, USA
11-12 July 2018

Editors:

Petra Perner

ISBN: 978-1-5108-9513-3

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2018) by ibai-publishing
All rights reserved.

Printed with permission by Curran Associates, Inc. (2020)

For permission requests, please contact ibai-publishing
at the address below.

ibai-publishing
Arno-Nitzsche-Str. 45
04277 Leipzig, Germany

info@ibai-publishing.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Content

Effective Mechanism of Mapping Labels in Financial Tables using Machine Learning <i>Yan Chen, Agrima Srivastava, and Dakshina Murthy Malladi</i>	1
Evaluation of Community Structure Finding Algorithms on Social Network Data Sets <i>Abdul Qadar Kara and Harun Pirim</i>	7
Applying Chromathics to Data Processing and Computations <i>Tony Nolan, Warwick Graco, Emily Nolan, Stewart Turner, Garry Mitchell, and Charles Palmer</i>	12
Development of High-Speed Engineering Data Transfer Technique <i>Zixian Zhang, Ichiro Kataoka, and Yixiang Feng</i>	26
Improving Sales Forecasting with Customer Behavior Analysis <i>Yusuke Yamaura, Yiou Wang, and Takeshi Onishi</i>	30
Electricity Short Term Load Forecasting <i>Gassan Abujumra and Mohamed Bouguessa</i>	40
Echo State Network Optimized by Cross-Entropy for Short-Term Load Forecasting of a Large Power Plant <i>Gabriel Trierweiler Ribeiro, Flavia Bernardo Pinto, Viviana Cocco Mariani, and Leandro dos Santos Coelho</i>	45
Identification of Human Activity Change using Time Series Analysis <i>Yulei Pang and Xiaozhen Xue</i>	52
Flow Prediction Versus Flow Simulation Using Machine Learning Algorithms <i>Milan Cisty</i>	56
Improving Sales Forecasting with Customer Behavior Analysis <i>Yusuke Yamaura1, Yiou Wang2, and Takeshi Onishi</i>	76
Deep Learning in Large-Scaled Time Series Forecasting <i>Chuanyun Zang</i>	86
Research on evaluation indicators weigh computing method of scientific research institutions based on Linked Open Data <i>Shiyin Jiang</i>	89

Extracting Rate-changes in Transcriptional Regulation by Word Embedding with Sentence Structure and Domain Knowledge in Deep Neural Networks <i>Wenting Liu and Yilei Zhang</i>	94
Applications of Data Mining in Insurance Sector: Explorations into Techniques of Leveraging Big Data <i>Darshan Desai and Om Desai</i>	101
Clustering Professional Baseball Players with SOM and Deciding Team Reinforcement Strategy with AHP <i>Kazuhiro Kohara and Shota Enomoto</i>	108
A Novel Parallel Algorithm for Frequent Itemsets Mining in Large Transactional Databases <i>Huan Phan and Bac Le</i>	120
Using Clusters in network threat detection <i>Tsigkritis Theocharis, Groumas Georgios, and Schneider Moti</i>	137
Promoting Connectivity: Social Network Analysis to Support Entrepreneurs <i>Cristina Feniser, Ken Brown, Arik Sadeh, Javier Bilbao, and Gabriel Plesa</i>	152
Ensemble of Heterogeneous Regressors Applied to Forecasting in Cosmetics Industry <i>Leandro dos Santos Coelho, Agrima Srivastava, Frederico Gonzalez Colombo Arnoldi, and Donald Neumann</i>	165
Segmenting Biosignals using Hierarchical Clustering <i>David Yuan and Vangelis Metsis</i>	170
Research on evaluation indicators weigh computing method of scientific research institutions based on Linked Open Data <i>Shiyin Jiang</i>	180
A New Approach for Mining Representative Patterns <i>Abeda Sultana, Hosneara Ahmed, and Chowdhury Farhan Ahmed</i>	184
Extracting Rate-changes in Transcriptional Regulation by Word Embedding with Sentence Structure and Domain Knowledge in Deep Neural Networks <i>Wenting Liu and Yilei Zhang</i>	199
Rise, Fall, and Implications of the New York City Medallion Market <i>Sherraina Song</i>	206

Parsimonious Modeling for Binary Classification of Quality in a High Conformance Manufacturing Environment <i>Carlos A. Escobar Diaz and Ruben Morales-Menendez</i>	221
Dynamic Classifier and Sensor Using Small Memory Buffers <i>Gelbard R.and Khalemsky A.</i>	236
An Efficient Approach for Mining Weighted Sequential Patterns in Dynamic Databases <i>Sabrina Zaman Ishita, Faria Noor, and Chowdhury Farhan Ahmed</i>	247
Detection of IP Gangs: Strategically Organized Bots <i>Tianyue Zhao and Xiaofeng Qiu</i>	262
Identification of Human Activity Change using Time Series Analysis <i>Yulei Pang and Xiaozhen Xue</i>	274
A Dynamic Ensemble Learning Approach for Online Anomaly Detection in Alibaba Datacenters <i>Wanyi Zhu, Xia Ming Huafeng Wang, Junda Chen, Lu Liu, and Zhengong Cai</i>	278
A Partial Demand Fulfilling Capacity Constrained Clustering Algorithm to Static Bike Rebalancing Problem <i>Yi Tang and Bi-Ru Dai</i>	293
A Novel Parallel Algorithm for Frequent Itemsets Mining in Large Transactional Databases <i>Huan Phan and Bac Le</i>	308
Data mining for municipal financial distress prediction <i>David Alaminos, Sergio M. Fernández, Francisca García, and Manuel A. Fernández</i>	323
Understanding Customers and Their Grouping via WiFi Sensing for Business Revenue Forecasting <i>Vahid Golderzahl and Hsing-Kuo Pao</i>	335
Social Media Sentiment Analysis Based on Domain Ontology and Semantic Mining <i>Daoping Wang, Liangyue Xu, and Amjad Younas</i>	350
Fault Diagnosis of Transformers using Machine Learning Technique: An Application of Support Vector Machine <i>D. Lam, L. Cuevas, and B. Chattopadhyay</i>	362

Learning to Rank and Discover for E-commerce Search <i>Anjan Goswami, Chengxiang Zhai, and Prasant Mohapatra</i>	374
Prediction of Re-tweeting Activities in Social Networks Based on Event Popularity and User Connectivity <i>Sayan Unankard</i>	389
Ensemble of Heterogeneous Regressors Applied to Forecasting in Cosmetics Industry <i>Leandro dos Santos Coelho, Viviana Cocco Mariani, Frederico Gonzalez Colombo Arnoldi, and Donald Neumann</i>	401
Enhancing Outlier Detection by An Outlier Indicator <i>Xiaqiong Li, Xiaochun Wang, Xia Li Wang</i>	406
A Method of Biomedical Knowledge Discovery by Literature Mining Based on SPO Predications: A Case Study of Induced Pluripotent Stem Cells <i>Zheng-Yin Hu, Rong-Qiang Zeng, Xiao-Chu Qin, Ling Wei, and Zhiqiang Zhang</i>	419
Online Evaluation of Classifier Accuracy, False Acceptance Rate and False Rejection Rate <i>Sabit Hassan1, Shaden Shaar, Bhiksha Raj, and Saquib Razak</i>	431