

12th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2020)

Volume 1: KDIR

Online

2 – 4 November 2020

Editors:

Ana Fred

Joaquim Filipe

ISBN: 978-1-7138-4065-7

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© (2020) by SCITEPRESS – Science and Technology Publications, Lda.
All rights reserved.

Printed with permission by Curran Associates, Inc. (2021)

For permission requests, please contact SCITEPRESS – Science and Technology Publications, Lda.
at the address below.

SCITEPRESS – Science and Technology Publications, Lda.
Avenida de S. Francisco Xavier, Lote 7 Cv. C,
2900-616 Setúbal, Portugal

Phone: +351 265 520 185

Fax: +351 265520 186

info@scitepress.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

CONTENTS

INVITED SPEAKERS

KEYNOTE SPEAKERS

- Context-aware Knowledge Management for Socio-Cyber-Physical Systems: New Trends towards Human-machine Collective Intelligence 5
Alexander Smirnov, Nikolay Shilov and Andrew Ponomarev
- From Activity- to Data-centric Approaches to Business Process Management: Challenges, Technologies, Applications 19
Manfred Reichert
- Hybrid Intelligence: AI Systems That Collaborate with People Instead of Replacing Them 21
Frank van Harmelen
- Persistent Identification for Interoperable Data Management and Preservation 23
Stefan Decker

PAPERS

FULL PAPERS

- Conversation Management in Task-oriented Turkish Dialogue Agents with Dialogue Act Classification 29
O. Fatih Kilic, Enes B. Dundar, Yusufcan Manav, Tolga Cekic and Onur Deniz
- Saudi Stock Market Sentiment Analysis using Twitter Data 36
Amal Alazba, Nora Alturayef, Nouf Alturayef and Zainab Alhathloul
- Improving Word Association Measures in Repetitive Corpora with Context Similarity Weighting 48
Aleksi Sahala and Krister Lindén
- The Max-Cut Decision Tree: Improving on the Accuracy and Running Time of Decision Trees 59
Jonathan Bodine and Dorit S. Hochbaum
- Large-scale Retrieval of Bayesian Machine Learning Models for Time Series Data via Gaussian Processes 71
Fabian Berns and Christian Beecks
- Bottom-up Discovery of Context-aware Quality Constraints for Heterogeneous Knowledge Graphs 81
Xander Wilcke, Maurice de Kleijn, Victor de Boer, Henk Scholten and Frank van Harmelen
- A Recurrent Neural Network and Differential Equation based Spatiotemporal Infectious Disease Model with Application to COVID-19 93
Zhijian Li, Yunling Zheng, Jack Xin and Guofa Zhou
- CoExDBSCAN: Density-based Clustering with Constrained Expansion 104
Benjamin Ertl, Jörg Meyer, Matthias Schneider and Achim Streit
- Geographic Feature Engineering with Points-of-Interest from OpenStreetMap 116
Adelson de Araujo, João Marcos do Valle and Nélío Cacho
- Amharic Document Representation for Adhoc Retrieval 124
Tilahun Yeshambel, Josiane Mothe and Yaregal Assabie

SHORT PAPERS

A Feature Space Transformation to Intrusion Detection Systems <i>Roberto Saia, Salvatore Carta, Diego Reforgiato Recupero and Gianni Fenu</i>	137
Estimating Personalization using Topical User Profile <i>Sara Abri, Rayan Abri and Salih Cetin</i>	145
Ontology-based Methods for Classifying Scientific Datasets into Research Domains: Much Harder than Expected <i>Xu Wang, Frank Van Harmelen and Zhisheng Huang</i>	153
Using Affective Features from Media Content Metadata for Better Movie Recommendations <i>John Kalung Leung, Igor Griva and William G. Kennedy</i>	161
Diverse Group Formation based on Multiple Demographic Features <i>Mohammed Alqahtani, Susan Gauch, Omar Salman, Mohammed Ibrahim and Reem Al-Saffar</i>	169
Session Similarity based Approach for Alleviating Cold-start Session Problem in e-Commerce for Top-N Recommendations <i>Ramazan Esmeli, Mohamed Bader-El-Den and Hassana Abdullahi</i>	179
Large Scale Intent Detection in Turkish Short Sentences with Contextual Word Embeddings <i>Enes Burak Dünder, Osman Fatih Kılıç, Tolga Çekiç, Yusufcan Manav and Onur Deniz</i>	187
Identifying the k Best Targets for an Advertisement Campaign via Online Social Networks <i>Mariella Bonomo, Armando La Placa and Simona E. Rombo</i>	193
FLE: A Fuzzy Logic Algorithm for Classification of Emotions in Literary Corpora <i>Luis-Gil Moreno-Jiménez, Juan-Manuel Torres-Moreno, Hanifa Boucheneb and Roseli S. Wedemann</i>	202
Towards Strength-sensitive Social Profiling in Ego Networks <i>Asma Chader, Hamid Haddadou, Leila Hamdad and Walid-Khaled Hidouci</i>	210
Weak Ties Are Surprising Everywhere <i>Yaakov Exman, Asaf Yosef and Omer Ganon</i>	218
Moving towards a General Metadata Extraction Solution for Research Data with State-of-the-Art Methods <i>Benedikt Heinrichs and Marius Politze</i>	227
Extracting Body Text from Academic PDF Documents for Text Mining <i>Changfeng Yu, Cheng Zhang and Jie Wang</i>	235
TAGWAR: An Annotated Corpus for Sequence Tagging of War Incidents <i>Nancy Sawaya, Shady Elbassuoni, Fatima K. Abu Salem and Roaa Al Feel</i>	243
Target Evaluation for Neural Language Model using Japanese Case Frame <i>Kazuhito Tamura, Ikumi Suzuki and Kazuo Hara</i>	251
Filtering a Reference Corpus to Generalize Stylometric Representations <i>Julien Hay, Bich-Liên Doan, Fabrice Popineau and Ouassim Ait Elhara</i>	259
Enhanced Active Learning of Convolutional Neural Networks: A Case Study for Defect Classification in the Semiconductor Industry <i>Georgios Koutroulis, Tiago Santos, Michael Wiedemann, Christian Faistauer, Roman Kern and Stefan Thalmann</i>	269

Social Media as an Auxiliary News Source <i>Stephen Bradshaw, Colm O’Riordan and Riad Cheikh</i>	277
A Survey of Sensor Modalities for Human Activity Recognition <i>Bruce X. B. Yu, Yan Liu and Keith C. C. Chan</i>	282
Personalised Recommendation Systems and the Impact of COVID-19: Perspectives, Opportunities and Challenges <i>Rabaa Abdulrahman and Herna L. Viktor</i>	295
R-peak Detector Benchmarking using FieldWiz Device and Physionet Databases <i>Tiago Rodrigues, Hugo Silva and Ana Fred</i>	302
Generating Adequate Distractors for Multiple-Choice Questions <i>Cheng Zhang, Yicheng Sun, Hejia Chen and Jie Wang</i>	310
Unsupervised Descriptive Text Mining for Knowledge Graph Learning <i>Giacomo Frisoni, Gianluca Moro and Antonella Carbonaro</i>	316
Stock Trend Prediction using Financial Market News and BERT <i>Feng Wei and Uyen Trang Nguyen</i>	325
Analysing the Effect of Platform and Operating System Features on Predicting Consumers’ Purchase Intent using Machine Learning Algorithms <i>Ramazan Esmeli, Alaa Mohasseb and Mohamed Bader-El-Den</i>	333
Historical Document Processing: A Survey of Techniques, Tools, and Trends <i>James Philips and Nasseh Tabrizi</i>	341
AUTHOR INDEX	351