2017 International Conference on Machine Learning and Data Science (MLDS 2017)

Noida, India
14-15 December 2017

IEEE Catalog Number: CFP17M85-POD
ISBN: 978-1-5386-3447-9

Copyright © 2017 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number: CFP17M85-POD
ISBN (Print-On-Demand): 978-1-5386-3447-9
ISBN (Online): 978-1-5386-3446-2

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
E-mail: curran@proceedings.com
Web: www.proceedings.com

2017 International Conference on Machine learning and Data Science

MLDS 2017

Table of Contents

Invited Talks.....vii

Committees.....xi

MLDS 2017 Technical Papers

A Comparison of Class Imbalance Techniques for Real-World Landslide Predictions 1

Kapil Agrawal, Yashasvi Baweja, Deepti Dwivedi, Ritwik Saha, Prabhakar Prasad, Shubham Agrawal, Sunil Kapoor, Pratik Chaturvedi, Naresh Mali, Venkata Uday Kala, and Varun Dutt

A Support Vector Machine Based Approach for Code Smell Detection 9

Amandeep Kaur, Sushma Jain, and Shivani Goel

An Empirical Investigation to Overcome Class-Imbalance in Inspection Reviews 15

Maninder Singh, Gursimran S. Walia, and Anurag Goswami

Categorical Classification and Deletion of Spam Images on Smartphones Using Image Processing and Machine Learning 23

Arjit Sachdeva, Rishabh Kapoor, Amit Sharma, and Akshit Mishra

Classification of Cancerous Profiles Using Machine Learning 31

Aman Sharma and Rinkle Rani

Comprehensive Review On Supervised Machine Learning Algorithms 37

Rishabh Choudhary and Hemant Kumar Gianey

Data Driven Decision Support to Fund Graduate Studies in Abroad Universities 44

Shahriar Yazdipour and Nahid Taherian

Extractive Text Summarization Using Word Vector Embedding 51

Aditya Jain, Divij Bhatia, and Manish K Thakur

Feature Selection Using Autoencoders 56
Dhananjay Tomar, Yamuna Prasad, Manish K. Thakur, and K. K. Biswas

Influence of Different Fields of View on Decision-Making in a Search-and-Shoot Scenario 61
Akash K Rao, Utkrisht Dhankar, Chandan Satyarthi, Sushil Chandra, and Varun Dutt

INGC: Graph Clustering & Outlier Detection Algorithm Using Label Propagation 68
Vandana Bhatia, Bharti Saneja, and Rinkle Rani

Internet of Things and Decision Support System for eHealth - Applied to Cardiometabolic Diseases 75
Parag Chatterjee, Ricardo Luis Armentano, and Leandro Javier Cymberknop

JPEG Compression Detection Based on Edge-Corner Features Using SVM 80
Yongsoo Choi, Dongshik Kang, Jae Jeong Hwang, and Kang Hyeon Rhee

Learning to Filter Documents for Information Extraction Using Rapid Annotation 85
Carlos A. Aguirre, Sneha Gullapalli, Maria F. De La Torre, Alice Lam, Joshua Levi Weese, and William H. Hsu

Modeling Choice Variation in Search Strategies with Multi-Armed Bandit Problems 91
Neha Sharma and Varun Dutt

Modeling Decisions in Games Using Reinforcement Learning 98
Himanshu Singal, Palvi Aggarwal, and Varun Dutt

Patient Journey Visualizer: A Tool for Visualizing Patient Journeys 106
Kulendra Kumar Kaushal, Shruti Kaushik, Abhinav Choudhury, Krish Viswanathan, Balaji Chellappa, Sayee Natarajan, Larry Pickett, and Varun Dutt

Spotted Hyena Optimizer for Solving Engineering Design Problems 114
Gaurav Dhiman and Amandeep Kaur

Using LSTMs for Predicting Patient's Expenditure on Medications 120
Shruti Kaushik, Abhinav Choudhury, Nataraj Dasgupta, Sayee Natarajan, Larry A. Pickett, and Varun Dutt

Validation of Inspection Reviews over Variable Features Set Threshold 128
Maninder Singh, Gursimran S. Walia, and Anurag Goswami

Weighted Frequent Subgraph Mining Using Affinity Weighing in Distributed Environment 136
Vipul Kumar and Daya Gupta

Author Index 141