DAY 1

Scientific Tracks & Abstracts

8th Edition of International Conference on

Big Data & Data Science

March 04-05, 2019 | Barcelona, Spain

DAY 1

March 04, 2019

Sessions

Hadoop Map-Reduce For Analyzing Information | Big Data In Neural Network For Deep Learning | Machine Learning In Data Science | Data Mining With Big Data | Big Data Analytics | Big Data Applications

Session Chair
Plyushchenko Andrey N
Eastwind, Russia

Session Introduction

Title: **Deep neural network system for fire emergencies: Detection, analysis and evacuation**
Jivitesh Sharma, University of Agder, Norway

Title: **Quality assessment of user generated content on twitter-A deep learning based approach**
Irfan Mohiuddin, King Saud University Riyadh, Saudi Arabia

Title: **Prediction of hourly floating population based on mobile phone data in Korea**
Ho-Chan Kwak, Korea Railroad Research Instutute, South Korea

Title: **Crime prediction using social big data and machine learning**
Juyoung Song, Pennsylvania State University, USA

Title: **Recursive estimators of the parameters of a regression**
Khedidja Djaballah, University of Algiers, Algeria

Title: **Investigating the application of machine learning techniques to detect data quality issues in big health data sets**
Suraj Juddoo, Middlesex University Mauritius, Mauritius

Title: **Can you tell the author? Natural language processing for literary studies**
Inez Okulska, NASK National Research Institute, Poland

Big Data &
Data Science

Quality assessment of user generated content on twitter-A deep learning based approach

Irfan Mohiuddin, Hassan Mathkour, Muhammad Al-Qurishi and **Majed Al-Rubaian**
King Saud University Riyadh, Saudi Arabia

Social Media today is a platform for millions of active users globally to share their content. Each second, there are thousands of messages or comments posted on different social networks. With these staggering numbers of user generated content (UGC), challenges are bound to surface. One such challenge is to assess the quality of UGC in social media because the content generated in social media could have positive or negative impact on fellow users and common people too. Low quality content not only impacts the user's content browsing experience, but also deteriorates the aesthetic value of social media. Therefore, our aim is to assess the quality of content accurately to promote the propagation of high quality content. Successful assessment of quality of UGC in social media fosters the growth of high utility UGC, which could be used by other applications and organizations for societal or organizational benefits. In this paper, we propose a deep learning based model, that leverages the quality assessment of UGC. The experimental results demonstrate that our proposed model results in high accuracy and low loss.

Recent Publications

1. "Secure distributed adaptive bin packing algorithm for cloud storage" in Future Generation Computer Systems (Q1, IF:3.99), (2018).

2. "Cloud computing services for iot −analyzing the security challenges and strategies" in international conference on industrial internet of things and smart manufacturing [(isbn: 978-1-912532-06-3)] (2018)

3. "Workload aware vm consolidation method (wavmcm) in cloud computing environment" in Journal of Parallel and Distributed Computing (Oct, 2018)

4. Contributor in a book titled "multimedia and cloud computing-architecture and applications", College of Computer and Information Sciences, King Saud University (2018).

5. Authored a chapter in book titled "industrial internet of things and smart manufacturing", Springer Publications. (Due for release).

Biography

Irfan Mohiuddin received his M.Sc. in Computer Science from King Saud University, Riyadh-Saudi Arabia, where he is currently working as a Researcher while pursuing his Ph.D. degree in Computer Science. His research interests include Data Science, Social Media Data Analysis, Cloud Computing, Virtualization and Social Internet of Things.

irfanm@ksu.edu.sa

Big Data &
Data Science

Prediction of hourly floating population based on mobile phone data in Korea

Ho-Chan Kwak and **Jiyoung Song**
Korea Railroad Research Institute, South Korea

The floating population is useful to figure out dynamic activities in urban area. Therefore, the prediction of floating population is required to practical use in urban and transportation planning. In Korea, the hourly floating population is estimated based on communication log of mobile phone. The communication log contains contents such as communication location, origin of mobile phone user. This paper is aimed to predict hourly floating population using mobile phone data, macroscopic data such as socioeconomic index, and several data mining techniques. The data collected in Seoul, the capital of South Korea, is used in this study. Also, the prediction accuracy by data mining technique is compared with each other, and the best model to predict hourly floating population is proposed in this study.

Recent Publications

1. Ho-Chan K and Seungyoung K (2016) Predicting crash risk and identifying crash precursors on Korean expressways using loop detector data. Accident Analysis and Prevention 88:9-19.

2. Ho-Chan K, Dong-Kyu K, Seungyoung K and Chungwon L (2014) A crash prediction model for expressways using genetic programming. Journal of Korean Society of Transportation 32(4):369-379.

Biography

Ho-Chan Kwak has completed his PhD at Seoul National University (SNU) and Postdoctoral studies in the Institute of Construction and Environmental Engineering at SNU. He is the Senior Researcher at Korea Railroad Research Institute (KRRI), a government-funded transportation science and technology research organization. He has studied in the fields of transportation big data (such as smart card data and mobile phone data etc.) analysis and application in order to improve the quality of mobility through future transportation technology. He has published more than 10 papers in transportation planning and safety related journals and performed more than 30 projects related in transportation engineering.

kwak01@krri.re.kr

Big Data & Data Science

March 04-05, 2019
Barcelona, Spain

Crime prediction using social big data and machine learning

Juyoung Song

Pennsylvania State University, USA

The crime phenomenon of modern society is more complex and diverse than in the past. There are many ways to predict and analyze crime phenomena. The current era of the fourth industrial revolution is experiencing innovative changes as cutting-edge information and communications technology are incorporated into all areas of the economy and society; for example, artificial intelligence (AI), the Internet of Things, big data, and mobile technology. Criminologists (crime-data scientists) play a very important role in this process. They create or assemble high-quality data that can be used to train machine-learning systems, find machine-learning algorithms that are suitable for the data, and perform modeling. The discussions of politics, economy, and culture posted on social media outlets represent the opinions of the era. The method of collecting and analyzing the unstructured data from online channels, including the Social Network Service, can interpret the actual phenomenon in our society. The current study uses structured and social big data to predict crime and preemptively respond to it. The results of this study provides a detailed description of the entire research process, which consisted of gathering big data, analyzing it, and making observations to develop a crime-prediction model that uses actual big data. The study also contains an in-depth discussion of several processes: text mining, which extracts useful information from online documents; opinion mining, which analyzes the emotions contained in documents; machine learning for crime prediction and visualization analysis. Machine learning will be applied to finally suggest a prediction model. The results of the analysis and policy implication will be discussed.

Recent Publications

1. Song J, Song T M and Lee J (2018) Stay alert: Forecasting the risks of sexting in Korea using social big data. Computer in Human Behavior 81:294-301.

2. Song J, Song T M, Seo D-C, Jin D-L and Kim J S (2017) Social Big Data Analysis of Information Spread and Perceived Infection Risk During the 2015 Middle East Respiratory Syndrome Outbreak in South Korea. Cyber psychology, Behavior, and Social Networking 20(1):22-29.

3. Song J, Song T M, Seo D C and Jin J (2016) Data mining of web-based documents on social networking sites that included suicide-related words among Korean adolescents. Journal of Adolescent Health 59(6):668-673.

4. Juyoung Song and Taemin Song (2018) Crime prediction using big data. Bullsbook Publishing Co. Seoul, Korea.

5. Taemin Song and Juyoung Song (2016) Social Big Data Research Methodology with R, Hannarae Publishing Co, Seoul, Korea.

Biography

Juyoung Song is an Assistant Professor of Criminal Justice and Criminology at Pennsylvania State University. She has completed her Bachelors and Master's degrees in the College of Law at Hanyang University in Seoul, South Korea, and her Doctorate degree in Criminal Justice at Michigan State University. Her Career appointments have included an Assistant Professor at the University of West Georgia, and an Associate Researcher at the Korean Institute of Criminology. She has presented at numerous national and international conferences about "Big Data" and published several articles on big data analysis. She has recently published five books about big data analysis in Korean and is currently working on, "Crime Prediction Using Big Data in English."

jxs6190@psu.edu