

# 2019 the 2nd International Conference on Computing and Big Data (ICCBD 2019)

Workshop

## 2019 International Conference on Computer, Software Engineering and Applications (CSEA 2019)

Taichung Software Park, Taichung, Taiwan

October 18-20, 2019

Sponsored by



東海大學  
TUNG HAI UNIVERSITY



Co-Sponsored by



Technical Supported by



Published by



# Table of Contents

<b>Welcome Remarks</b> .....	3
<b>Conference Venue</b> .....	4
<b>Instructions for Presenter</b> .....	5
<b>Information on Publication</b> .....	6
<b>Keynote Speakers Overview</b> .....	7
Keynote Speech I - Prof. Yongsheng Ma .....	7
Keynote Speech II - Prof. Wenbing Zhao .....	8
<b>Agenda Overview</b> .....	9
Day 1 (October 18, 2019): On-site Registration Only and Materials Collection .....	9
Day 2 (October 19, 2019): Keynote Speeches and Authors' Presentations.....	10
Day 3 (October 20, 2019): Academic Visit.....	10
<b>Author Presentations Overview on October 19</b> .....	11
Session I –Machine Learning and Data Mining .....	13
Session II –Big Data and Network Security .....	16
Session III –Computer Technology and Software Engineering.....	20
Session IV –Computer Vision and Image Processing.....	23
<b>Poster Presentation</b> .....	26
<b>Listeners' List</b> .....	27
<b>Upcoming Conferences</b> .....	28
ICBCT 2020 .....	28
ICBDE 2020 .....	29

## **Welcome Remarks**

**On behalf of organizing Committee, we welcome you to Taiwan to attend 2019 the 2nd International Conference on Computing and Big Data (ICCBD 2019) and its workshop 2019 International Conference on Computer, Software Engineering and Applications (CSEA 2019). We're confident that over the three days you'll get theoretical grounding, practical knowledge, and personal contacts that will help you build long-term, profitable and sustainable communication among researchers and practitioners working in a wide variety of scientific areas with a common interest in Computing, Big Data, Computer, Software Engineering and Applications.**

**We hope that your work and that of your institution or company will be enhanced both by what you learn and by those with whom you connect over the next 3 days. Our field is enriched by the dialogue among colleagues from around the world which occurs during presentation sessions as well as informal conversations. We hope this is a memorable, valuable, and enjoyable experience!**

**On behalf of conference chair and all the conference committee, we would like to thank all the authors as well as the program Committee members and reviewers. Their high competence, their enthusiasm, their time and expertise knowledge, enabled us to prepare the high-quality final program and helped to make the conference a successful event. We hope that all participants and other interested readers benefit scientifically from the proceedings and also find it stimulating in this process. Finally, we would like to wish you success in your technical presentations and social networking.**

**Once again, thanks for coming to this conference. We are delegate to higher and better international conference experiences. We will sincerely listen to any suggestion for improvement and we are looking forward to meeting you next time.**

**ICCBD 2019 & CSEA 2019  
Conference Committee**

# Conference Venue

## Taichung Software Park – Yessoon Super One

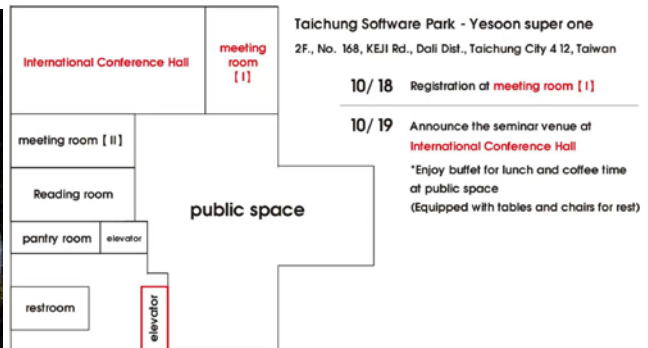
Address: International Conference Hall, S1 zhihui center, No. 1 Science and Technology Road, Dali district, Taichung, Taiwan

(Yessoon Super One: 2F., No. 168, KEJI Rd., Dali Dist., Taichung City 4 12, Taiwan)

Webpage: <https://www.daliartplaza.com.tw/>

<https://www.epza.gov.tw/english/page.aspx?Pageid=ecc3eefec73b3559>

[https://www.ideasoft3d.net/base\\_intro.aspx](https://www.ideasoft3d.net/base_intro.aspx)



The Location and Expected Performance of Taichung Software Park

(1) Location and Area of Taichung Software Park: Taichung Software Park is located at Sec. 1, Zhongxing Rd., Dali Dist., Taichung City. It was originally the land to test new tobacco and liquor from Taiwan Tobacco and Wine Monopoly Bureau and has an area of around 4.96 hectares. It is now currently under the affiliation of Export Processing Zone Administration, MOEA.

(2) Orientation of the Park: MOEA has planned it to become the R&D base for digital contents, cultural creativity and innovation.

(3) Expected Performance: The establishment plan of Taichung Software Park was approved by Executive Yuan on December 12th, 2011 and was organized to primarily introduce information service industry, digital content industry, cloud computing industry, and Mandarin e-commerce industry. After its establishment, it is expected to create 5000 employment opportunities to Taichung region, introduce around NT\$8 billion of investment and the annual product value can reach NT\$15 billion.

**Notice:** The restaurant of Dinner: LIN JI clay pot roasted chicken (Address: No. 686, Xinxing Rd., Wuri Dist., Taichung City 4 14, Taiwan)

### LIN JI clay pot roasted chicken



No. 686, Xinxing Rd., Wuri Dist., Taichung City 4 14, Taiwan

# Instructions for Presenters

## Onsite Registration

- 1) You can also register at any time during the conference.
- 2) Certificate of Participation will be awarded at the end of your session.
- 3) Your paper ID will be required for the registration.
- 4) The organizer won't provide accommodation, and we suggest you make an early reservation.

## Oral Presentation

- 1) Devices Provided by the Conference Organizer:
- 2) Laptops (with MS-Office & Adobe Reader)
- 3) Projectors & Screens
- 4) Laser Sticks
- 5) **Materials Provided by the Presenters:**
- 6) Power Point or PDF Files (Files should be copied to the conference laptop at the beginning of each session)
- 7) **Duration of each Presentation (Tentatively):**
- 8) Regular Oral Presentation: 15 Minutes of Presentation, including Q&A
- 9) Keynote Speech: 45 Minutes of Presentation, including Q&A

## Poster Presentation

### **Materials Provided by the Conference Organizer:**

The place to put poster

### **Materials Provided by the Presenters:**

Home-made Posters

Maximum poster size is A1

Load Capacity: Holds up to 0.5 kg

## Best Presentation Award

One Best Oral Presentation will be selected from each presentation session, and it will be awarded after each session.

## Dress Code

Please wear formal clothes or national representative clothing.

## Important Note:

**The time slots assigned in the schedule are only tentative.**

## Legal Disclaimer Notice:

To ensure that the meeting is carried out smoothly and efficiently in a safe environment, we make the following disclaimers:

First, the participants must wear the conference name tag to enter the conference areas during the entire conference. Individuals not wearing it will not be denied entry. Participants are not allowed to bring their associates or families into the conference areas. If you wish to bring someone, please inform the conference person in charge in advance to obtain a special name tag for whatever event like presentations, breaks or meals.

Second, participants should confirm that they have good physical, mental and psychological preparation to participate in the event. If the participant is under 18 years old, he or she must be accompanied by his or her legal guardian. The conference organizer and associations will not responsible for their safety.

Third, all legal responsibilities and liabilities arising directly or indirectly from participants' activities during the event shall be borne by the participants themselves.

Fourth, the conference area is public, so participants should take care of their belongings at all times.. All losses related to personal property will not be the responsibility of the event organizers and associations.

## Introductions for Publication



All accepted papers for **the 2019 the 2nd International Conference on Computing and Big Data (ICCBD 2019)** and its workshop **2019 International Conference on Computer, Software Engineering and Applications (CSEA 2019)** will be published in the the International Conference Proceedings Series by ACM, which will be archived in the ACM Digital Library, and sent to be indexed by EI Compendex and Scopus and

submitted to be reviewed by Thomson Reuters Conference Proceedings Citation Index (ISI Web of Science).

ISBN: 978-1-4503-7290-9

# Introductions for Speakers

*Keynote Speaker I*

*Prof. Yongsheng Ma*  
*University of Alberta, Canada*



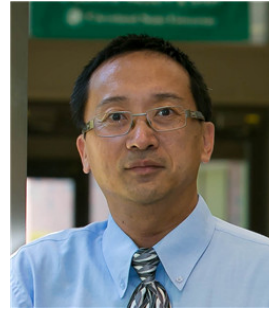
Dr. Yongsheng Ma joined the faculty of engineering, University of Alberta in 2007, and has been a full professor since 2013. Dr. Ma is a member of ASEE, SME, SPE and an Alberta registered Professional Engineer. He teaches Capstone Design Projects, Engineering Informatics, Engineering Economics and Manufacturing Processes. His main research areas include e-Manufacturing, feature-based design and manufacturing modeling, CAD/CAM, ERP process modeling and product lifecycle management. Dr. Ma received his B.Eng. from Tsinghua University, Beijing (1986), both M.Sc. (1990) and Ph.D. (1994) from UMIST, UK. In 2000-2007, he was a faculty with Nanyang Technological University, Singapore. Dr. Ma publishes actively in top international journals and conferences, especially in the field of Semantic Modeling for Product and Process Engineering. Dr. Ma had been an associate editor of IEEE Transaction of Automation Science and Engineering (2009-2013). Since 2012, he has served as an editor of Advanced Engineering Informatics. Due to his fruitful collaboration with a Canadian manufacturer, in 2012, he won the prestigious ASTech award sponsored by Alberta Science and Technology Leadership Foundation. Dr. Ma started his career as a polytechnic lecturer in Singapore (1993); and then a senior research fellow and group manager (1996-2000) at Singapore Institute of Manufacturing Technology. Dr. Ma also served in department selection committee (2012), faculty nominating (2009-2011) and evaluation committees (2011). Dr. Ma had been a member of General Faculty Council (GFC) and a university senator from July, 2013 to July, 2016.

***Speech Title: Gamification and Predictive Analytics for the Next Generation of Workers***

***Abstract:*** In this era of new industrial revolution, Industry 4 framework championed in Europe is important for the management to determine ways to satisfy employees' needs, and the companies' needs simultaneously. It has never been so dynamic and pressurized to keep employees focused, engaged, motivated and involved with their jobs. This research proposes a network based data driven real-time solution which will help companies stay competitive and as well as recruit and retain talents. The idea is that the future of employee motivation is to take advantage of gamification elements using machine learning and statistical model, that have dominated social media and gaming applications. A review of psychological theories that induce extrinsic and intrinsic motivations in the workplace, as well as the current applications of game elements in a non-gamified environment, is done in this research. Suggested approach is discussed for the implementation of psychological theories in the study of gamification has played an important role in our digital lives as multiple social media and gaming apps compete to harness user engagement to stay popular and relevant. Following that, a modified gamified model is proposed that uses real time data, Weibull statistical distribution and a K-means clustering algorithm, machine learning along with a reward system that would help increase skill development and retention along with employee satisfaction.

*Keynote Speaker II*

*Prof. Wenbing Zhao*  
*Cleveland State University, USA*



Wenbing Zhao received his Ph.D. in Electrical and Computer Engineering at University of California, Santa Barbara, in 2002. Dr. Zhao has a Bachelor of Science degree in Physics in 1990, and a Master of Science degree in Physics in 1993, both at Peking University, Beijing, China. Dr. Zhao also received a Master of Science degree in Electrical and Computer Engineering in 1998 at University of California, Santa Barbara. Dr. Zhao joined Cleveland State University (CSU) faculty in 2004 and is currently a Professor in the Department of Electrical Engineering and Computer Science (EECS) at CSU. Dr. Zhao has authored a research monograph titled: “Building Dependable Distributed Systems” published by Scrivener Publishing, an imprint of John Wiley and Sons. Furthermore, Dr. Zhao published over 200 peer-reviewed papers in the area of distributed systems (three of them won the best paper award), smart and connected health, physics, and education. Dr. Zhao’s research is supported in part by the US National Science Foundation, the US Department of Energy, the US Department of Transportation, Ohio State Bureau of Workers’ Compensation, Ohio Department of Higher Education, and by Cleveland State University. Dr. Zhao is currently serving on the organizing committee and the technical program committee for numerous international conferences. He is an Associate Editor for IEEE Access and for MDPI Computers. Dr. Zhao is a senior member of IEEE and a senior member of International Economics Development and Research Center (IEDRC).

***Speech Title: Data Science in Professional Baseball: A Preliminary Study***

***Abstract:*** Unlike many other sectors in the society, the professional sports industry is driving technical innovations. Team owners, managers, and players recognize the huge value of technology for improving athlete performance, game officiating, as well as helping fans to better enjoy the games. A good example is the US major league baseball (MLB). In this talk, I report a preliminary study on video-based pitch type recognition using deep learning. To facilitate the study, we first developed a semi-automated way of building datasets based on publicly available video and pitch information for MLB games. For pitch type recognition, we used the two-stream inflated 3D convolutional neural network (I3D). To improve the state-of-the-art of research, we trained and tuned the I3D model extensively, primarily combating the problem of overfitting while still trying to improve final validation accuracy. We are able to achieve an accuracy of 53.43% +/- 3.04% when oversampling and 57.10% +/- 2.99% when not oversampling, which is a significant improvement over the published best result of an accuracy of 36.4% on the same six pitch type classes.



# Conference Agenda

## Day 1(October 18, 2019): Onsite Registration & Conference Materials Collection

(Note: October 18, 2019 is only for onsite Registration, but on October 19, 2019, the registration is also open outside the conference room)

Meeting Room I	10:00-17:00	Registration & Conference Materials Collection
----------------	-------------	--

## Day 2(October 19, 2019) (Morning) Keynote Speeches

<p><b>Opening Remark</b> 09:00-09:10</p>		<p><b>Prof. Wenbing Zhao</b> Cleveland State University, USA</p>
<p><b>Keynote Speech 1</b> 09:10-09:55</p> <p>International Conference Hall</p>		<p><b>Prof. Yongsheng Ma</b> University of Alberta, Canada</p> <p><b>Speech Title: Gamification and Predictive Analytics for the Next Generation of Workers</b></p>
<p>09:55-10:20</p>	<p><b>Coffee Break &amp; Group Photo</b></p>	
<p><b>Video Presentation</b> 10:20-10:40</p> <p>International Conference Hall</p>	<p><b>The Introduction of Taichung Software Park</b></p>	
<p><b>Keynote Speech 3</b> 10:40-11:25</p> <p>International Conference Hall</p>		<p><b>Prof. Wenbing Zhao</b> Cleveland State University, USA</p> <p><b>Speech Title: Data Science in Professional Baseball: A Preliminary Study</b></p>
<p>12:00-13:30</p>	<p><b>Lunch</b></p>	

**Day 2(October 19, 2019) (Afternoon): Authors' Presentation**

<b>Time</b>	<b>Venue</b>	<b>Session</b>
<b>13:30- 15:30</b>	<b>International Conference Hall</b>	<b>Session 1</b> <b>Topic: Machine Learning and Data Mining</b> 8 papers
<b>13:30-15:30</b>	<b>Meeting Room I</b>	<b>Session 2</b> <b>Topic: Big Data and Network Security</b> 8 papers
<b>15:30-15:45</b>	<b>Coffee Break</b>	
<b>15:45-17:30</b>	<b>International Conference Hall</b>	<b>Session 3</b> <b>Topic: Computer Technology and Software Engineering</b> 7 papers
<b>15:45-17:45</b>	<b>Meeting Room I</b>	<b>Session 4</b> <b>Topic: Computer Vision and Image Processing</b> 8 papers
<b>18:00-20:00</b>	<b>Dinner</b> <b>Venue—LIN JI clay pot roasted chicken</b>	

**Day 3(October 20, 2019): Academic Visit**

# Presentations Quick Review

**Session 1**      **Theme: Machine Learning and Data Mining**  
**Times: 13:30-15:30** | **Conference Room: International Conference Hall**  
**Session Chair:**      **Assoc. Prof. Wendy Hui, Lingnan University, Hong Kong, China**

<b>CB0015</b>	Application of Literature-Based Discovery in Nonmedical Disciplines: A Survey <i>Wendy Hui and Wai Kwong Lau</i>
<b>CB0006</b>	Filter Selection Methods for Multiclass Classification <i>Rhodessa Jurilla Cascaro, Bobby D Gerardo and Ruji P Medina</i>
<b>CB3002</b>	Regression and Classification in Supervised Learning <i>Jiachong Li</i>
<b>CB0016</b>	Multi-class Document Classification Using Improved Word Embeddings <i>Benedict A. Rabut, Arnel C. Fajardo and Ruji P. Medina</i>
<b>CB3004</b>	Utilizing Tandem Features for Text-Independent Speaker Recognition on Short Utterances <i>Alvarez Arvin Kenneth, Pelipas Mary Tricia Ann, Rayos del Sol Carl Ivan and Tomas John Paul</i>
<b>CB1003</b>	Handling Imbalanced Data through Affinity Propagation and SMOTE <i>Lanie B. Laureano, Ariel M. Sison and Ruji P. Medina</i>
<b>CB0005</b>	Abstract Generation System for Chinese Articles and Reviews of 3C Products <i>Yi-Ting Chen and Tzone-I Wang</i>
<b>CB1001</b>	Relocating Local Outliers Produced by Partitioning Methods <i>Rogelio O. Badiang Jr., Bobby D. Gerardo and Ruji P. Medina</i>

**Session 2**      **Theme: Big Data and Network Security**  
**Times: 13:30-15:45** | **Conference Room: Meeting Room I**  
**Session Chair:**      **Assoc. Prof. Hui Na Chua, Sunway University, Malaysia**

<b>CB0004</b>	An Application for Classifying Depression in Tweets <i>Bernice Ziwei Yeow and Hui Na Chua</i>
<b>CB0014</b>	A Secure Key Scheduling Operation for International Data Encryption Algorithm using Serpent Key Schedule Operation <i>Gerald Tomelden Cayabyab, Ariel M Sison and Ruji P Medina</i>
<b>CB0008</b>	False Alert Buster: an Adaptive Approach for NIDS False Alert Filtering <i>Yuan-Hsiang Su, Michael Cheng Yi Cho and Hsiu-Chuan Huang</i>
<b>CB1002</b>	A Secured RFID Lightweight Authentication Protocol based on Elliptic Curve Cryptography Using Modified PRNG Algorithm <i>Domingo V. Origines JR, Ariel M. Sison and Ruji P. Medina</i>
<b>CB2001</b>	NetFlowTotal: A Cloud Service Integration Platform for Malicious Traffic Analysis and Collaboration <i>Tzung-Han Jeng, Wei-Ming Chan, Wen-Yang Luo, Chuan-Chiang Huang, Chien-Chih Chen and Yi-Ming Chen</i>
<b>CB0017</b>	A Modified Least Significant Bit Randomized Embedding Method based on Image Partitioning and Columnar Transposition with Encryption <i>Fredilyn Balayan Calanda, Ariel Mora Sison, Mark Rennel D Molato and Ruji P Medina</i>
<b>CB0003-A</b>	Big Data Management: Synchronising Complexities of the South African Higher Education Admissions Processes <i>Salamina Manoge Dzingwa and Fayth Ruffin</i>
<b>CB0013</b>	GISKOP: A Modified Key Scheduling Operation of International Data Encryption Algorithm using Serpent Key Scheduling <i>Gerald Tomelden Cayabyab, Ariel M Sison and Alexander A Hernandez</i>

**Session 3**      **Theme: Computer Technology and Software Engineering**  
**Times: 15:45-17:30** | **Conference Room: International Conference Hall**  
**Session Chair:**      **Prof. Wenbing Zhao, Cleveland State University, USA**

<b>CB2005</b>	Effect of Preservation Technology for Sustainable Cold Chain <i>Yung-Fu Huang, Ming-Wei Weng, I-Sung Lai and Lin Huang</i>
---------------	---

<b>CB3001</b>	Investigating the Effect of the Cloud Computing On Education Internet of Things (Eiot) <i>Yu-Hsin Hung and Shen-Hong Wu</i>
<b>CB2006</b>	Wavelet Analysis on the Relevance of Stock Market and Exchange Rate between China and India <i>Hsin-Pei Hsueh and Fangjhy Li</i>
<b>CB2002</b>	Applying ISO/IEC 29110 to ISO/IEC 62304 for Medical Device Software SME <i>Natsuda Kasisopha and Panita Meananeatra</i>
<b>CB0019-A</b>	Data Analytics & Big Data Impact on the Auditing Profession: Sea Change or Just Another Fad? <i>Ramesh Narasimhan</i>
<b>CB0007</b>	Ballot Mechanism Design Based on Blockchain Methodologies <i>Meng-Hsuan Fu</i>
<b>CB0002</b>	Intelligent Maintenance Diagnosis System <i>Chu-Chai Henry Chan and Gino Liu</i>
<b>Session 4            Theme: Computer Vision and Image Processing</b> <b>Times: 15:45-17:45 Conference Room: Meeting Room I</b> <b>Session Chair:      Prof. Joel C. De Goma, Mapua University, Philippines</b>	
<b>CB3003</b>	Disease Detection of Asian Rice (Oryza Sativa) in the Philippines Using Image Processing <i>Elysse Joy Angelica V. Pascual, Joe Mhar J. Plaza ,Jose Lorenzo L. Tesorero and Joel C. De Goma</i>
<b>CB2004</b>	Using Virtual Reality Technology to Enhance the Experience of Immersive Sensation Training Color Applied to the Environment by Means Of EEG <i>Wei-Te Tsai and Chien-Hsu Chen</i>
<b>CB3008</b>	Detection of Proper Form on Upper Limb Strength Training using Extremely Randomized Trees for Joint Positions <i>Aldrin John Torres, Clowie Silubrico, Daniel Torralba and John Paul Tomas</i>
<b>CB3007</b>	Producing 3D Animation Teaching Materials – A Case Study of the Huge Differences Between Male and Female Wealth Management <i>Ting-sheng Weng</i>
<b>CB3006</b>	Classification of Varietal Type of Philippine Rice Grains Using Image Processing Through Multi-View 3d Reconstruction <i>Jan Rei S. Buenaventura, Jun T. Kobayashi, Larah Mae P. Valles, Joel C. De Goma and Ariel Kelly D. Balan</i>
<b>CB3009</b>	Detection of Overall Fruit Maturity of Local Fruits using Convolutional Neural Networks Through Image Processing <i>Mark Anthony Ayllon, Melwin James Cruz, Justin Jason Mendoza and Mary Christine Tomas</i>
<b>CB0001</b>	Enhancing Feed-Forward Neural Network in Image Classification <i>Mark Jovic A Daday, Arnel C Fajardo and Ruji P Medina</i>
<b>CB3005</b>	Real-Time Multi-Person Smoking Event Detection <i>Waynebert Jan D. Cabanto, Aira Danielle B. Jocson, Renzel Laurence T. Lateo and Joel C. De Goma</i>

# Authors' Presentations

## Session 1

**Topic: Machine Learning and Data Mining**

**Time: 13:30-15:30**

**Venue: International Conference Hall**

**Session Chair: Assoc. Prof. Wendy Hui**

**Lingnan University, Hong Kong, China**

\*The time slots assigned here are only tentative. Presenters are recommended to stay for the whole session in case of any absence.

\*\*After the session, there will be a group photo for all presenters in this session.

<p>13:30-13:45 <b>CB0015</b></p>	<p>Application of Literature-Based Discovery in Nonmedical Disciplines: A Survey <b>Wendy Hui</b> Lingnan University, HongKong, China</p> <p><b>Abstract:</b> The rapid increase in the academic literature has resulted in overspecialization of knowledge. To overcome this overspecialization, literature-based discovery (LBD) has been proposed to help discover promising and under-investigated hypotheses that connect two bodies of knowledge. The method has been around for more than 30 years, but its application has mostly been confined to the medical sciences. In this study, we perform a comprehensive survey of the literature on the application of LBD in nonmedical disciplines, summarize the approaches used, and discuss ways to promote the use of LBD in nonmedical disciplines.</p>
<p>13:45-14:00 <b>CB0006</b></p>	<p>Filter Selection Methods for Multiclass Classification <b>Rhodessa Jurilla Cascaro, Bobby D Gerardo and Ruji P Medina</b> Technological Institute of the Philippines, Philippines</p> <p><b>Abstract:</b> Feature selection is used in choosing relevant features that contribute to the predictive power of a machine learning model. Irrelevant features tend to decrease model accuracy and cause overfitting. Feature selection is the solution to dimensionality problems, especially that data nowadays are unstructured. There are three types of feature selection techniques; filter, wrapper and embedded. Filter methods uses statistical scoring and ranks features by the score. Wrapper methods uses a model to select features and evaluates according to model accuracy. Embedded methods combine the properties of both previous algorithms and selects features in the process of training. This study is focused on the filter types, specifically Chi-square, Information Gain and Relief. On the other hand, multiclass classification is a task that involves classifying instances into three or more classes. This paper aims to compare the performance of the Support Vector Machine (SVM) multiclass classifier when entered with feature subsets generated from three different filter feature selection methods. The dataset is a clothing review text data taken from Kaggle. It contains multiple classes with 23486 review instances. Since filter selection methods utilize ranking, CHI, IG, and Relief were able to select and rank almost the same number of features. When the data was fed into SVM, CHI garnered 66.84% accuracy, IG got 32.90% accuracy while Relief obtained 29.69% accuracy. Experiments on SVM showed that among the generated data subset from the three filter selection methods, the subset using CHI yielded higher accuracy, precision, recall and F1 score compared to others.</p>
<p>14:00-14:15 <b>CB3002</b></p>	<p>Regression and Classification in Supervised Learning <b>Jiachong Li</b></p> <p><b>Abstract:</b> The problem of recognizing patterns from big data has attracted a lot of attention these days, especially in artificial intelligence and machine learning fields. People are interested in training computers to make predictions or classifications on their own based on past experience, i.e., data. In this paper, we review three fundamental supervised learning models (linear regression, logistic regression, and perceptron) for both regression and classification tasks, including their theoretical background, algorithmic solutions, and application scenarios. We also conduct synthetic experiments to demonstrate their performance.</p>

<p>14:15-14:30 <b>CB0016</b></p>	<p>Multi-class Document Classification Using Improved Word Embeddings <b>Benedict A. Rabut</b>, Arnel C. Fajardo and Ruji P. Medina Technological Institute of the Philippines, Philippines</p> <p><b>Abstract:</b> In this paper, we conducted an experiment to build a classification model that combines different techniques in most of the Natural Language Processing Tasks. We used the word embedding method to transform every word in the dataset and to obtain the custom-built word embedding vectors. This is in contrast to the approaches in the previous literature that implement word embedding using the pre-trained word embedding vectors. We enriched the custom-built word embedding vectors by incorporating Part-of-Speech (POS) tag vectors to provide additional semantic information about the word to be used in training our proposed classification model. The proposed model was built using the neural network approach, which is considered to be more efficient and reliable in solving real problems for document classification tasks. We fine-tuned the parameters during the training of our neural network classification model with our aim to increase the performance in terms of classification accuracy. The experimental result demonstrates that our model performs remarkably well and increase the percentage accuracy up to 1.7% compared to the accuracy results obtained by the previous baseline word embedding methods using the same dataset. It was also observed that our model outperforms some other traditional classification models implemented using different techniques and machine learning algorithms.</p>
<p>14:30-14:45 <b>CB3004</b></p>	<p>Utilizing Tandem Features for Text-Independent Speaker Recognition on Short Utterances Alvarez Arvin Kenneth, Pelipas Mary Tricia Ann, Rayos del Sol Carl Ivan and <b>Tomas John Paul</b> Mapua University, Philippines</p> <p><b>Abstract:</b> This study focuses on the application of Deep Neural Networks trained to discriminate amongst senones to improve speaker recognition performance when dealing with text-independent Short Utterance Speaker Recognition (SUSR). The features derived from the said network are theorized to be more robust given that they can eliminate any unnecessary information in the final representation of the speaker. The efficacy of these features is evaluated using the test subset of the LibriSpeech speech corpus. It is found that the system's performance, especially when dealing with SUSR is greatly improved when these above-mentioned features are concatenated with traditional, more widely used Mel Frequency Cepstral Coefficients (MFCC) as measured in terms of Equal Error Rate (EER).</p>
<p>14:45-15:00 <b>CB1003</b></p>	<p>Handling Imbalanced Data through Affinity Propagation and SMOTE <b>Lanie B. Laureano</b>, Ariel M. Sison and Ruji P. Medina Technological Institute of the Philippines, Philippines</p> <p><b>Abstract:</b> Imbalanced datasets occur in real-world scenarios and affect the performance of the classifiers in supervised learning. To combat the imbalance between majority and minority instances, several techniques including SMOTE which generates artificial data along the line between minority samples and its selected neighbors are used in preprocessing. However, SMOTE suffers from generating synthetic data in noise regions and imbalanced distribution of samples within the minority class. Using clustering techniques prior to the oversampling step to determine the target areas of the input space where the generation of synthetic samples is effective, is known to be an extension of SMOTE. The proposed method uses affinity propagation as a clustering technique to generate clusters and its cluster exemplar without requiring the specification of the number of clusters a priori. The cluster exemplar becomes the basis of which samples are to be oversampled in each cluster. Simulation results on some publicly available datasets show the effectiveness of the proposed method in terms of F-measure, G-mean and AUC with a mean value across datasets of .685,.756 and .765 respectively.</p>
<p>15:00-15:15 <b>CB0005</b></p>	<p>Abstract Generation System for Chinese Articles and Reviews of 3C Products <b>Yi-Ting Chen</b> and Tzone-I Wang National Cheng Kung University, Taiwan</p> <p><b>Abstract:</b> The Internet has become popular and convenient. Product articles and reviews are written by people on digital media platforms such as Facebook, PTT, Mobile01, and Apple Daily News. Most people read many articles and reviews on digital media when they want to buy a</p>

	<p>product. However, an overwhelming number of articles and reviews of products is available on the Internet, and a prospective buyer can become confused. People should organize these articles and reviews before deciding on whether to buy a product or not. Therefore, it is essential to summarize the available data and provide customers useful information. Many researchers have investigated this task, but most studies have been focused on English reviews. This work focuses on Chinese articles and reviews in digital media, and propose a system designed to summarize data from digital media. When a user is interested in a product, the system extracts features and opinion words of the product from review articles and uses these features to identify sentences highly related to the product. After obtaining these sentences, the approach in this work selects top 20 important sentences to form the summary of the product, which is presented to the user. This work conducts several experiments to compare the effectiveness of TextRank, Luhn's method, and the proposed approach. Among them, the approach proposed in this work exhibits the best performance.</p>
15:15-15:30 <b>CB1001</b>	<p>Relocating Local Outliers Produced by Partitioning Methods  <b>Rogelio O. Badiang Jr.</b>, Bobby D. Gerardo and Ruji P. Medina  Technological Institute of the Philippines, Philippines</p> <p><b>Abstract:</b> Partitioning methods is one of the most known and most used data clustering methods. The cluster center, known as centroid or medoid, represents each cluster that these methods produced. When the cluster center cannot describe all the data in the group, the undescribed data are mistakenly classified to their nearest cluster. These data are considered as local outliers. With this, the study seeks to address the problem of local outliers produced by partitioning methods specifically k-means and k-medoids. The Local Outlier Rectifier (LOR) is a proposed method used to relocate local outliers to its appropriate clusters. The results of the simulation show that several outliers are successfully transferred and the accuracy of the clustering performance increases.</p>
15:30-15:45	<b>Coffee Break</b>

**Session 2**  
**Topic: Big Data and Network Security**  
**Time: 13:30-15:30**  
**Venue: Meeting Room I**  
**Session Chair: Assoc. Prof. Hui Na Chua**  
**Sunway University, Malaysia**

\*The time slots assigned here are only tentative. Presenters are recommended to stay for the whole session in case of any absence.

\*\*After the session, there will be a group photo for all presenters in this session.

13:30-13:45 <b>CB0004</b>	<p>An Application for Classifying Depression in Tweets  Bernice Ziwei Yeow and <b>Hui Na Chua</b>  Sunway University, Malaysia</p> <p><b>Abstract:</b> A depressed person feels sad, hopeless on a regular basis, lose interest in activities and experiences physical symptoms such as chronic pain or digestive issues [1]. Each day, millions of people suffered from depression and only a small fraction of them undergo proper treatment. In the past, doctors analyzed depressed people via a face to face session and apply the diagnostic criteria written by a professional psychologist to determine depression. However, past research revealed that most patients would not seek help from doctors at the early stage of depression which results in a declination of their mental health condition. On the other hand, many people are using the social media platform to share their feelings daily. Since then, there have been many studies on using social media to predict mental and physical diseases such as cardiac arrest, zika virus, Muppalla, and prescription drug abuse. However, currently there is a lack of work on an application that leverages social media data for detecting depressive thoughts to determine the depression state of a Twitter user. To achieve this aim, we developed a web application that performs sentiment analysis through a classification function that discerns the proportion of depressive and non-depressive thoughts of an individual. This application further provides individuals to visualize their depression state via a web interface.</p>
13:45-14:00 <b>CB0014</b>	<p>A Secure Key Scheduling Operation for International Data Encryption Algorithm using Serpent Key Schedule Operation  <b>Gerald Tomelden Cayabyab</b>, Ariel M Sison and Ruji P Medina  Technological Institute of the Philippines Quezon City, Philippines</p> <p><b>Abstract:</b> The International Data Encryption Algorithm (IDEA) is one of an encryption algorithm that uses a cyclic process in producing sub keys from secret key using its key schedule but generates large weak keys. GISKOP is a modified IDEA that uses the concept of serpent key scheduling operation to address the weak keys found on the original IDEA. Randomness test was used to identify the dependence and correlations between generated sub-keys of a key scheduling operation. Statistical evaluation was done to analyze and compare key distribution and exchange operations of the modified key scheduling operation. The GISKOP key scheduling operation obtained an increase of 8% randomness using frequency test in the sub-key independence level. Also, it was able to maintain its 100% randomness under the sub-key level independence. Thus, proving that each generated sub keys are not correlated with each other. With this, many attacks on systems that implement IDEA can be mitigated. Furthermore, the findings will take advantage of independency or nonlinear relationship between sub-keys and prevent breaking the whole cipher.</p>
14:00-14:15 <b>CB0008</b>	<p>False Alert Buster: an Adaptive Approach for NIDS False Alert Filtering  <b>Yuan-Hsiang Su</b>, Michael Cheng Yi Cho and Hsiu-Chuan Huang  Information &amp; Communication Security Laboratory, Chunghwa Telecom Laboratories, Taiwan</p> <p><b>Abstract:</b> Detecting cyber threats is one of the most important security issues today, and network based intrusion detection systems (NIDS) play a vital role. Many studies have adopted Artificial Intelligence/Machine Learning (AI/ML) technology to create powerful NIDS to detect cyber threats. While most of NIDS studies focus on improving classification/detection accuracy by proposing new AI/ML models, all models produce a ratio of false positive alerts in the field. Very few studies discuss the methodology to handle false positive alerts. Giving a busy network, handling the number of false positive alerts become a time consuming task for security personnel. For this reason, false</p>



	<p>positive alert filtering automation is an important issue. In this paper, we propose a scheme to automate false positive alert filtering by leveraging kernel density estimation. Disregard of the deployed NIDS, our proposed scheme can assist security personnel in alert verification task. Our experiments show that our proposed scheme is 34% to 62% better in performance (in terms of error ratio) compared to other algorithms. Our proposed scheme also reduces 75% of the time in alert verification process.</p>
14:15-14:30 <b>CB1002</b>	<p>A Secured RFID Lightweight Authentication Protocol based on Elliptic Curve Cryptography using modified PRNG Algorithm  <b>Domingo V. Orignes JR</b>, Ariel M. Sison and Ruji P. Medina  Technological Institute of the Philippines, Philippines</p> <p><b>Abstract:</b> In this paper, the goal is to enhance the lightweight Authentication Protocol using Epoch Timestamp as Seed for PRNG Protocol based on Elliptic Curve Cryptography. The proposed enhancement uses PRNG Algorithm where randomness is generated using real-time system clock converted into running epoch timestamp as Seed for PRNG on LCG Algorithm that will do Bootstrap Iteration in generating a Random Number as keys for ECC Mutual Authentication Protocol. The result shows that the enhanced ECC Protocol utilizes smaller storage requirements of 2560bits + 640w bits, communication cost of 1,120bits faster response time and throughput, and lesser computational cost of 207ms for the Tag and the Server was achieved compared to Zheng Protocol. Also, the performance of this proposed Protocol has many advantages in terms of confidentiality, mutual authentication, availability, forward security, anonymity, spoofing/cloning, scalability and so on, which can resist tracking attacks, denial of service attacks, and system internal attack. It is therefore suggested that some practical applications using a RFID system that requires security for the <u>lightweight authentication protocol</u> should be introduced and more appropriate.</p>
14:30-14:45 <b>CB2001</b>	<p>NetFlowTotal: A Cloud Service Integration Platform for Malicious Traffic Analysis and Collaboration  <b>Tzung-Han Jeng</b>, Wei-Ming Chan, Wen-Yang Luo, Chuan-Chiang Huang, Chien-Chih Chen and Yi-Ming Chen  Chunghwa Telecommunication Labs, Taiwan</p> <p><b>Abstract:</b> Network security lacks the verification of real world net flow data, and lacks a platform to collect and integrate net flow data and threat intelligence, so as to generate an evaluation benchmark for machine learning on cybersecurity. NetFlowTotal develop many net flow analysis tools to detect malicious threats in the net flow data. Through the two-side market strategies, NetFlowTotal platform tie together two distinct groups of users in a network. One kind of user can upload net flow data to the NetFlowTotal platform to obtain security incidents reports; the other kind of user can share threat intelligence to the NetFlowTotal platform to obtain more associate threat intelligence according to global net flow data. The goal of this paper is to establish a net flow evaluation platform to provide real world dataset with security incidents reports for machine learning evaluation.</p>
14:45-15:00 <b>CB0017</b>	<p>A Modified Least Significant Bit Randomized Embedding Method based on Image Partitioning and Columnar Transposition with Encryption  <b>Fredilyn Balayan Calanda</b>, Ariel Mora Sison, Mark Rennel D Molato and Ruji P Medina  Technological Institute of the Philippines, Philippines</p> <p><b>Abstract:</b> While traditional embedding methods which are usually based on LSBs provide security to hidden messages in image-based Steganography, steganalyses have shown these methods to be vulnerable to detection and extraction, thus compromising the security of embedded messages. In this study, a new embedding method based on LSB is introduced to address this vulnerability. By using a one-time pad encryption method, the security of secret messages before embedding is ensured. A partitioning method based on columnar transposition and RGB color plane scattering technique is employed to improve scattering and ensure randomization during embedding. The experimental result shows an excellent outcome of the Peak-signal-to-noise Ratio (PSNR), Mean square Error (MSE), and Structural Similarity (SSIM) index. The PSNR value of the 5 test images results ranges from 53.2122 to 60.8716 all are higher than the acceptable value of PSNR, which is 30db. The average MSE result is 0.2766 for all color planes, which also provide an outstanding result</p>

	<p>since the lower the MSE means the accumulative squared error between the stego image and the original one is lower. The average SSIM obtained in the testing is 0.9792 for the 5 test images, which is very close to 1. The result supports the claim that the resulting stego images have low distortion which means that the cover image is almost the same with the stego image and that the modified LSB embedding algorithm provides excellent image quality with high imperceptibility and provides added security layers.</p>
<p>15:00-15:15 CB0003-A</p>	<p>Big Data Management: Synchronising Complexities of the South African Higher Education Admissions Processes <b>Salamina Manoge Dzingwa and Fayth Ruffin</b> Vaal University of Technology, South Africa; University of KwaZulu- Natal, South Africa</p> <p><b>Abstract:</b> Higher education institutions (HEIs) are a source of big data and subject to a variety of forms of analytics for different purposes. This empirical study is concerned with management of big data for admissions processes with an eye toward student success. In South Africa, undergraduate admission processes have been roblematizes by numerous challenges and a series of negative consequences. Institutional management of admissions processes has been termed inefficient and ineffective. Massification of higher education in post-apartheid South Africa exacerbates the situation. The national government proposes the nation-wide use of the Post School Education and Training Central Applications Service (PSET-CAS) for HEIs as a mandatory approach to ameliorating admissions related problems. The Centralised Applications Systems, within which the PSET-CAS is embedded, is expected to facilitate access to post-school education and training by making choices and placement of students across the sector more efficient and effective. Of the nine provinces in South Africa, the province of KwaZulu-Natal was selected by government to model the CAS. In KwaZulu-Natal the Central Applications Office (CAO) is the hub of the CAS. This study appears to be the first of its kind to test big data management by the CAS. However, there are diverse information technology (IT) systems employed by the various educational institutions. In this study we questioned: to what extent can the use of different IT systems employed by varied HEIs and the CAO affect implementation of the CAS and the nature of big data extracted for admissions decision-making? To begin to answer this question we set out to determine if the CAS is a fit system for addressing admissions challenges and whether it is sufficiently efficient and effective to be worthy of extension to the rest of the country. This mixed method study was driven by a case study strategy, underpinned by post-positivist, pragmatism and constructivist worldviews. It was guided by an input-output-outcome meta-theoretical framework □roblematizes□ by network governance theories. Data were collected from two universities in South Africa that currently use the CAS model, namely the Durban University of Technology (DUT), and the University of KwaZulu-Natal (UKZN). Data were also collected from the CAO. There were 312 study participants consisting of staff and students. Semi-structured interviews, focus groups and a group interview were conducted of staff. Students responded to a survey. Emerging themes from combined application of content, matrix and thematic analyses provided guidance on how to improve the implementation of the PSET-CAS before it becomes operational nationwide. Stata 11.0 results from the surveys revealed a number of statistically significant relationships between relevant variables. We found that the CAS is a fit and preferred system for applications processing and applicants' information (school grades, background etc.) can significantly contribute to informed decision making by HEIs. Findings further suggest that the use of different IT systems in the admissions processes present challenges which need careful scrutiny. For example, both DUT and UKZN use different Enterprise Resource Programs while the CAO uses its own system. The use of these different systems could affect the efficiency and effectiveness of the application processes, and subsequently the implementation of the national CAS. This in turn, □roblematizes the utility of big data management. Based on empirical findings, the study makes recommendations on how to handle the lack of synchronicity between IT systems for a nationwide CAS. Efficiency and effectiveness of the CAS is particularly central for Global South countries seeking enriched human capital development to contribute to the global knowledge economy. Therefore, lessons learnt could be transferable to other jurisdictions.</p>

<p>15:15-15:30 <b>CB0013</b></p>	<p><b>GISKOP: A Modified Key Scheduling Operation of International Data Encryption Algorithm using Serpent Key Scheduling</b>  <b>Gerald Tomelden Cayabyab</b>, Ariel M Sison and Alexander A Hernandez          Technological Institute of the Philippines Quezon City, Philippines</p> <p><b>Abstract:</b> Cryptography is a method of storing and transmitting data in a particular form. Only those for whom it is intended can read, use it, and return it back to the original data by using various techniques. The International Data Encryption Algorithm “IDEA” is a block cipher that works with 64-bit plaintext block and ciphertext blocks and it has a 128-bit input key. This paper describe the designing and implementation of a modified key schedule operation of IDEA called GISKOP. It uses the same number of rounds and output transformation that operates using 128 bit user input plaintext and a modified way of key scheduling operation of 256 bit keys. The modified algorithm uses Serpent key scheduling operation to derive the different sub keys to be used in each rounds. The algorithm was implemented to provide better security on user’s password within the Document Management System to protect user’s data within the cloud database. It has gone through initial testing and evaluations with very encouraging results.</p>
<p>15:30-15:45</p>	<p>Coffee Break</p>

**Session 3**  
**Topic: Computer Technology and Software Engineering**  
**Time: 15:45-17:30**  
**Venue: International Conference Hall**  
**Session Chair: Prof. Wenbing Zhao**  
**Cleveland State University, USA**

\*The time slots assigned here are only tentative. Presenters are recommended to stay for the whole session in case of any absence.

\*\*After the session, there will be a group photo for all presenters in this session.

15:45-16:00 <b>CB2005</b>	<p>Effect of Preservation Technology for Sustainable Cold Chain  Yung-Fu Huang, <b>Ming-Wei Weng</b>, I-Sung Lai and Lin Huang  Chaoyang University of Technology, Taiwan</p> <p><b>Abstract:</b> The aim of this paper was to analyze food preservation technology in selected food manufacturer in Taiwan. The increasingly complex and extended food supply chain increases the risk of food quality. A generalized model is presented to determine the optimal temperature, preservation technology investment and replenishment strategies that minimize the food manufacturer's total cost. Advances in science and technology such as RFID technology, information computing technology to ensure food quality and safety. Numerical examples and sensitivity analysis are then provided by the collecting real data from Taiwan. Finally, concluding remarks are offered.</p>
16:00-16:15 <b>CB3001</b>	<p>Investigating the Effect of the Cloud Computing On Education Internet of Things (Eiot)  <b>Yu-Hsin Hung</b>, Shen-Hong Wu  National Yunlin University of Science and Technology, Taiwan</p> <p><b>Abstract:</b> Internet of things (IoT) has been applied widely in various industry, especially for manufacture industry, energy industry, transportation, and education. Its technology has brought transformations in various fields, educational technology. Education IoT (EIoT) is the innovation framework in the educational field, EIoT can be described as the interactive framework where numerous information is connected and synchronized by applying cloud computing, third-party technologies, gateway, and in a data communication. This study surveyed the EIoT-related studies from 2012 to 2019. It investigates the disciplinary distribution of emerging the EIoT-related topics. The findings of this study show that cloud computing, big data, e-learning, online services, and machine learning are the key issues driving researches, which are associated with EIoT. Recently, the application of cloud computing in the EIoT has attracted considerable attention from various researchers and professionals. Education as a Service (EaaS) is considered as one of the cloud-computing type. Additionally, this study investigates the present EaaS for achieving the goal of EIoT using the cloud machine learning module. The results demonstrated that the cloud-computing solution offers the appropriate tools needed for the users to analyze their learning profile data and build the intelligent educational strategy.</p>
16:15-16:30 <b>CB2006</b>	<p>Wavelet Analysis on the Relevance of Stock Market and Exchange Rate between China and India  Hsin-Pei Hsueh and <b>Fangjhy Li</b>  Hubei University of Economics, China</p> <p><b>Abstract:</b> China and India are the two countries with the highest population in the world, and their GDP is also among the top 10 in the world and continues to increase. This study is intended to explore the correlation between the exchange rate of China and India and the stock price index and to understand its leading backward relationship. Wavelet Analysis (Wavelet Analyses) was used as the research method and the interest rate was included as the control variable. The research period was from October 1996 to December 2016. We use wavelet analysis as the main research method. There is a significant correlation between the exchange rate of China and India and the stock price index (<math>&gt; 0.7</math> and <math>&gt; 0.8</math>). The empirical results show that first China and India have significant linkages. Except that China has no correlation between 2000 and 2004, there are significant correlation periods in the short-term (1-4 years) or long-term (4-8 years) correlation coefficient <math>&gt; 0.8</math>. Second, Short-term (1-4 years) linkage: After incorporating control variables, in the short-term (1-4 years) phase difference, China has a constant positive correlation between short-term (1-4 years) exchange rate and stock price index during the study period. In the short-term (1-4 years), India has a positive correlation with</p>

	<p>the exchange rate and the leading backward relationship is different and mutually causal. Third, Long-term (4-8 years) linkage relationship: After joining the control variables, the stock price index and the exchange rate between China and India have a two-way causal relationship. Forth, Long-term (4-8 years), regardless of whether or not the control variables are included, the correlation between China and India is significant (<math>&gt; 0.8</math>). When the fluctuations of the stock price index suddenly change, the correlation between the stock price index and the exchange rate is reversed. When the change, from the positive (negative) relevant region to the negative (positive) related region, is the large fluctuation of the stock price index, the structural change on behalf of the stock market does affect the linkage between the stock price index and the exchange rate.</p>
16:30-16:45 CB2002	<p>Applying ISO/IEC 29110 to ISO/IEC 62304 for Medical Device Software SME  <b>Natsuda Kasisopha</b> and Panita Meananeatra  National Electronics and Computer Technology Center, Thailand</p> <p><b>Abstract:</b> Medical Device (MD) and Medical Device Software (MDS) imperatively precise as they are dealing with human life and human health. Nowadays, medical device product manufactured by small and medium companies/entities/organization, which diverge from the past. To control the quality and security of those products, the Food and Drug Administration (FDA) regulation and IEC60601 are compulsory. Moreover, IEC60601 are referring to ISO/IEC 62304 on MDS development. The burden of small entities is an unawareness of the initiation point to adopting the ISO/IEC 62304 processes. This paper, we present a directive for a small entity who employed the ISO/IEC 29110 and aspire to apply the ISO/IEC 62304 processes. This is to prepare the small entity in planning to apply ISO/IEC 62304 before further testing and certification. Hence, the reviewed and aligned context of ISO/IEC 62304, ISO/IEC 29110, and ISO/IEC 12207 standards on the System Life Cycle processes are necessary.</p>
16:45-17:00 CB0019-A	<p>Data Analytics &amp; Big Data Impact on the Auditing Profession: Sea Change or Just Another Fad?  <b>Ramesh Narasimhan</b>  Montclair State University, USA</p> <p><b>Abstract:</b> The auditing profession has evidenced many technological and other innovations in the past fifty years. Statistical sampling, use of personal computers, expert systems, ERPs, and other changes that have impacted businesses were supposed to change the auditing profession dramatically. However the techniques used by auditors have not changed very radically from what was being used even a century ago. The latest developments that are impacting the profession are data analytics and big data. In this paper we trace the advances in the profession to date and analyze the reasons the auditing techniques has not changed considerably with these developments and the possibility that data analytics and big data may impact the profession more than the previous developments.</p>
17:00-17:15 CB0007	<p>Ballot Mechanism Design Based on Blockchain Methodologies  <b>Meng-Hsuan Fu</b>  Shih Hsin University, Taiwan</p> <p><b>Abstract:</b> Ballot mechanism design is proposed to improve existing ballot procedures. Presently, direct election of government officials at various levels is still under the mechanism of hand counting. Occasionally, the outcome from a hand count could still save the day because of the mistakes or inconsistent standards. Nowadays, network architecture and IoT devices are well developed. Based on the infrastructure, the ballot mechanism is designed by adapting blockchain methodologies to solve the problems of existing election procedures. In ballot mechanism, blockchain methodologies are adopted including cryptography, peer-to-peer transaction, data storage. In which, the complete cryptography is necessary for protecting personal information and defending outside destruction. Secure hash algorithm is a methodology to convert meaningful text into a meaningless string. Digital signature is adopted to ensure the vote was made by the person who has the related private key. Peer-to-peer transmission between voters and election center server to reduce the exchange errors. Each vote will be stored in the blockchain, where is under the secure mechanism, data is hardly altered and records are stored permanently. The goal of ballot mechanism design is aim to raise the election efficiency and accuracy, improve the manual counting procedures, and reduce cost of human resource and misjudgment through the blockchain methodologies.</p>

17:15-17:30 <b>CB002</b>	<p>Intelligent Maintenance Diagnosis System <b>Chu-Chai Henry Chan</b> and Gino Liu Chaoyang University of Technology, Taiwan</p> <p><b>Abstract:</b> Traditionally, the process of trouble shooting is time-consuming done by expert's personal experiences. Pre-automatic diagnosis seems a good solution to increase the service quality of trouble shooting for information technology outsourcing. Developing an expert system for pre-automatic diagnosis should be the better approach to speed up the service. The main purpose of this study is to explore how IT outsourcing service providers build a case-based reasoning expert system for trouble shooting. To solve problems, a case-based reasoning system retrieves information from one or more previous cases for reference. Then this developed system modifies the cases to adapt to the new situation. By this way, cases can be reused, and previous experiences can be retained. Finally, current cases will be combined into existing knowledge base After searching solutions for references, the study chooses the method of K-Nearest Neighbor (KNN) to classify clients' problems. This will increase the efficiency and accuracy in searching for the best solution. A pilot case study is implemented and the conclusion shows that the developed system really assists the customers effectively.</p>
18:00-20:00	Dinner

**Session 4**  
**Topic: Computer Vision and Image Processing**  
**Time: 15:45-17:45**  
**Venue: Meeting Room I**  
**Session Chair: Prof. Joel C. De Goma**  
**Mapua University, Philippines**

\*The time slots assigned here are only tentative. Presenters are recommended to stay for the whole session in case of any absence.

\*\*After the session, there will be a group photo for all presenters in this session.

<p>15:45-16:00 <b>CB3003</b></p>	<p>Disease Detection of Asian Rice (<i>Oryza Sativa</i>) in the Philippines Using Image Processing  Elysse Joy Angelica V. Pascual, Joe Mhar J. Plaza, Jose Lorenzo L. Tesorero and <b>Joel C. De Goma</b>  Mapua University, Philippines</p> <p><b>Abstract:</b> In the Philippines, rice is one of the most important food consumed for almost every meal and so, it is essential to meet the demands in rice production. However, local storms contribute to the susceptibility of rice plants to diseases debilitating the ability of farms to produce large amount of high-quality rice. Identification of diseases are done manually by farmers based on experience or an expert's advice. This paper proposes an automated detection of diseases using image processing. In the segmentation technique, the images are divided into two sets where green pixels are masked with 1) blue and 2) black pixels, along with the other features, and are then fed separately into SVM and Random Forest to compare their performance. Results show that blue pixels and SVM as its classifier have yielded better outcome having 82.41% as compared to the other model.</p>
<p>16:00-16:15 <b>CB2004</b></p>	<p>Using Virtual Reality Technology to Enhance the Experience of Immersive Sensation Training Color Applied to the Environment by Means Of EEG  <b>Wei-Te Tsai and Chien-Hsu Chen</b>  National Cheng Kung University, Taiwan</p> <p><b>Abstract:</b> The aim of this study explores whether color perception is reinforced by the immersive experience created by virtual reality technology. In an environment that emphasizes vision, the feelings of immersion stimulate people's emotions due to changes in color. Most of the digital content is now produced without systematic color matching principles for immersive sensation planning. This study will create an immersive environment through the implementation of virtual reality technology to understand the relationship between color and emotion. In the stage of implementation, a film will be created. The content of the film are sequence of images with neutral emotions that change color during continuous playback. In the detection of participants' emotions, data collection is mainly carried out for changes in EEG. The main observed values include Alpha, Beta and Gamma, which are then converted into psychological feelings. In the experimental procedure, the same image will be given different colors as an element of the operational variable to understand the emotional changes of the participants. In the results of the study, participants did show different mood swings due to changes in color. Especially in an immersive environment, emotions are strengthened. The research finds that visual perception occupies an important proportion in an immersive environment, but it is not fully summed up. We provide a concept that explores color semantics as an important visual design in an immersive environment.</p>
<p>16:15-16:30 <b>CB3008</b></p>	<p>Detection of Proper Form on Upper Limb Strength Training using Extremely Randomized Trees for Joint Positions  Aldrin John Torres, Clowie Dizon Silubrico, Daniel Torralba and <b>John Paul Tomas</b>  Mapua University, Philippines</p> <p><b>Abstract:</b> Learning how to perform a workout usually requires a little bit of research for people without experience, especially if it's in strength training or a workout that uses an equipment. Having an experienced instructor guide a trainee to perform such activity is enough, but with the advancement of computer vision, it opens up an opportunity to use technology to teach us without human intervention. In this study, Kinect V2 will be used to detect the joints of the subject and capture the starting and the end position of the workout. By tracking the position of the joints through the RGB-D sensor, the model will determine if the posture was performed correctly by the subject of these 5 (five) workouts: dumbbell lateral raise, dumbbell shoulder press, barbell front raise, dumbbell</p>

	<p>shrug, and barbell upright row. The data will be collected using the Java language and the KinectPV2 library to get the coordinates of the joints for each subject. Scikit-learn will be the chosen platform in the model creation for training and testing the datasets. Extremely Randomized Trees is a tree-based ensemble method for supervised classification and regression that will randomize. ERT will be used to the model proposed and compare it to the KNN.</p>
16:30-16:45 <b>CB3007</b>	<p>Producing 3D Animation Teaching Materials – A Case Study of the Huge Differences Between Male and Female Wealth Management <b>Ting-sheng Weng</b> National Chiayi University, Taiwan</p> <p><b>Abstract:</b> Wealth management education has invariably been a concern that cannot be neglected. The information received and risk appetite of genders in wealth management are different, resulting in different wealth management habits and investment behaviors. There are differences in the wealth management methods and investment behaviors between males and females. This study adopts three software applications, Animaker, ZEPETO, and PowerDirector, to produce three different 3D animation teaching materials on the huge differences of wealth management practices between men and women. If students have a deep understanding of the differences between men and women in wealth management and investment behavior, they can fully perceive the importance of personal and family wealth management.</p>
16:45-17:00 <b>CB3006</b>	<p>Classification of Varietal Type of Philippine Rice Grains Using Image Processing Through Multi-View 3d Reconstruction Jan Rei S. Buenaventura, Jun T. Kobayashi, Larah Mae P. Valles, Joel C. De Goma and <b>Ariel Kelly D. Balan</b> Mapua University, Philippines</p> <p><b>Abstract:</b> Number of previous works have been carried out to study the classification of rice in the flat 2D image. In computer vision, 3D reconstruction is a field that has great theoretical research value. This paper attempts to see the performance by using features extracted from the 3D reconstructed image of a rice grain in classifying different varieties of it. We identify 250 rice grains belonging to five (5) varietal groups of rice in the Philippines. The data gathered was in the form of a video from a camera then it was split into images. After applying image enhancement operations, the images were used to build the 3D model using Agisoft Photoscan software. After the reconstruction, 18 color features and 21 texture features were extracted from the image. After the feature selection, 32 features were selected for this study. The proposed method uses BPNN architecture to classify the rice grain's varietal type.</p>
17:00-17:15 <b>CB3009</b>	<p>Detection of Overall Fruit Maturity of Local Fruits using Convolutional Neural Networks Through Image Processing Mark Anthony Ayllon, Melwin James Cruz, Justin Jason Mendoza and <b>Mary Christine Tomas</b> Mapua University, Philippines</p> <p><b>Abstract:</b> Determining the ripening of a fruit is critical to a farmer, since the fresher the fruit, the better it will be priced and sold. This is also critical to the economy since the ninth (9<sup>th</sup>) most exported good in the Philippines is fruits. The researchers made use of Convolutional Neural Networks through Image Processing to determine the fruit maturity of Banana (Cavendish), Mango (Carabao), Calamansi/Calamondin, will classify said fruits into three categories for the fruit maturity: pre-matured, matured, over-matured. Of the sixty fruits used, twenty pieces of which will be used to gather data, starting from their very unripe/pre-matured stage up-to the over-matured stage. This will approximately take one (1) to two (2) weeks if stored in room temperature. The total data gathered would be 3681 pieces for Calamansi (Philippine lime); 3270 pieces for Banana (Cavendish); and 5706 pieces for Mango (Carabao). The model is written in Spyder in Anaconda Navigator, which will be applying Tensorflow-GPU and Keras. These will also be coupled with CUDA and CUDDN to process the data and determine the results. Two total experiments will be executed – one for the Red-Green-Blue (RGB) dataset, and one for the greyscale dataset.</p>



17:15-17:30 <b>CB0001</b>	<p>Enhancing Feed-Forward Neural Network in Image Classification  <b>Mark Jovic A Daday</b>, Arnel C Fajardo and Ruji P Medina  Technological Institute of the Philippines, Philippines</p> <p><b>Abstract:</b> In this paper. The researcher use Feed Forward Neural Network for image classification. The objective of this paper is to Enhancing the structure of FFNN by adding dropout method into the input layer and hidden layer to prevent over-fitting the network and then initialized by random uniform initializer for each layer then setting a bias of zeros, also adding the following activation function ReLu&amp;Softmax to the network and the dataset that the researcher use was the MNIST-handwritten numbers. The dropout method is one of the easy and efficient regularization methods that avoids overfitting which stochastically drops out some neurons and trains the network at each weight update. After that, researcher determines the performance of the result by evaluating the FFNN with dropout on the basis of loss and accuracy to take the average of the difference between the standard FFNN and CNN. The result of FFNN with dropout method had an average of 99.86% accuracy and 0.47% of loss, the standard FFNN got 98.13% accuracy and loss of 9.15%, while the CNN had an average of 99.26% accuracy and 2.39% of loss. As the training process of neural network is an iterative process comprising forward propagation, speed improvement using dropout would also provide a significantly decreased training time.</p>
17:30-17:45 <b>CB3005</b>	<p>Real-Time Multi-Person Smoking Event Detection  Waynebert Jan D. Cabanto, Aira Danielle B. Jocson, Renzel Laurence T. Lateo and <b>Joel C. De Goma</b>  Mapua University, Philippines</p> <p><b>Abstract:</b> Smoking is a widespread problem around the globe and having an algorithm that can automatically detect smoking events is crucial in helping to public areas free of dangerous cigarette smoke. This paper presents a method of detecting smoking events through multiple human tracking, action recognition and smoke detection. In the proposed method, human objects are first detected using Histogram of Oriented Gradients (HOG). Then, the detected human objects are tracked using centroid tracking. The smoking event detection requires two factors. The first one is whether the detected human object is displaying actions that would suggest the act of smoking using CNN from Keras library and the other factor is whether there is cigarette smoke in the frame using SVM to identify the candidate areas and color features and pixel values as thresholds. The model obtained an f-score of 72.55% having one person in the frame and 64.37% having 2 or more.</p>
18:00-20:00	Dinner

## Poster Presentations

CB1004	<p>Multi-Label Classification of Employee Job Performance Prediction by DISC Personality  <b>Patompat Kamtar</b>, Duangjai Jitkongchuen and Eakasit Pacharawongsakdar  Dhurakij Pundit University, Thailand</p> <p><b>Abstract:</b> The objective of this study was to automate job performance prediction based on DISC personality test. We transformed this problem to Multi-Label Classification (MLC) by using employee's job performances as labels. In this study, three widely used MLC techniques have been employed such as Binary Relevance (BR), Label Powerset (LP) and Classifier Chains (CC) for prediction of job performances. However, these traditional techniques didn't show promising results. Therefore, we proposed another approach by building stacking MLC with model selection. The proposed method has three steps: (1) building MLC model; (2) using process from the first step and applying with a stacking model and (3) utilizing feature selection technique to select the proper models for final prediction. Using the surveys from a big financial company in Thailand, we found that the last proposed approach shows better performance, compared to the traditional MLC.</p>
CB1005	<p>Online Updating Algorithms of Statistical Methods for Big Data  <b>Yihao Li</b> and Jin Wang  Huazhong University of Science and Technology, China</p> <p><b>Abstract:</b> In this paper, we discuss online updating algorithms for Big Data. One of the main challenges of Big Data is the limitation of data storage. In the Big Data stream environment, online computation sometimes requires fast updates without the use of historical data. The focus of this research is on efficient online update algorithms for basic statistical computations, including mean, variance, covariance, skewness, kurtosis, confidence interval, test statistic, and linear regression. We demonstrate the implementation of R Language through a linear regress example.</p>
CB1006	<p>Applied Linear Algebra Methods for Data Science  <b>Chien Hsu</b> and Jin Wang  University of Nottingham, UK</p> <p><b>Abstract:</b> In this paper, we discuss efficient algorithms for using eigenvalues and eigenvectors. The main algorithm is t Principal Component Analysis (PCA), a powerful method widely used for dimensionality reduction, image compression, and face recognition. We derive the optimal <math>k</math> of the number of principal components selected. The implementation of PCA application is demonstrated by an image compression example using the MATLAB programming language. We also discuss PageRank and Network Classification algorithms.</p>
CB0009	<p>Research on the Attention Degree to Textile Apparel Information Based on Consumer Endowments  Liu Xia, Wang Guoqing, Li Jing, Chen Qianwen and Xu Kang  Quality Management Branch of China National Institute of standardization, China</p> <p><b>Abstract:</b> The consumers' attention to the basic information on the textile apparel and the purchase channels are surveyed, the relevance between the age, gender, education level, and occupation of the respondents and their attention to the information on the textile apparel, the purchase channels and the product safety indicators are analyzed, and the possibility of injury and the source of accident information under the product's usage scenarios are analyzed, for which the goodness of fit and correlation coefficient are used for analysis and test. The results show that the occupation, education level and age indicate the significant relationship with the purchase channels; the occupation is significantly correlated to the purchase channels; the occupation, education level and age are not significantly associated with the safety indicators of textile apparel.</p>

## Listeners' List

<b>Listener 1</b>	<b>Carl Ivan Trespalacioreal Rayos del Sol</b> Mapua University, Philippines
<b>Listener 2</b>	<b>Waynebert Jan Dumbrigue Cabanto</b> Mapua University, Philippines
<b>Listener 3</b>	<b>Jun Kobayashi</b> Mapua University, Philippines
<b>Listener 4</b>	<b>Elysse Joy Angelica V. Pascual</b> Mapua University, Philippines
<b>Listener 5</b>	<b>Mark Anthony V. Ayllon</b> Mapua University, Philippines
<b>Listener 6</b>	<b>Clowie Dizon Silubrico</b> Mapua University, Philippines

## Upcoming Conferences



The 2020 The 2nd International Conference on Blockchain Technology (ICBCT 2020) will be held in 'Imiloa Astronomy Center of Hawaii University at Hilo, Hilo, Hawaii, USA during March 12-14, 2020.

Recognizing the importance of Block chain and related technologies, the International Economics Development and Research Center held a conference to foster the development of an initiative. During one day event, representatives from academia, industry, government, and IEDRC volunteers and staff gathered to define that initiative. As several subject matter experts discuss the status of the technology, attendees give their own perspectives, and attendees share their research results. This diverse group of experts will generate more than great ideas; they will define the work that must be done for Blockchain to reach its potential, and will focus the energy to accomplish it

### Publication



The excellent papers accepted by ICBCT 2020 will be published in international conference proceedings, which will be sent to be indexed by EI Compendex and Scopus.

### Topics

Topics of interest for submission include, but are not limited to:

Alternative proof of work algorithms

Cyberattacks on blockchains

Economic impact on cyberattacks

Fraud and error reduction with blockchain

Inventory management with blockchain

Courier cost reduction with blockchain

Government operation with blockchain

Automobile finance management with blockchain

Global medical opinion platform with blockchain

Assets management with blockchain

System engineering with IoT and blockchain

Supply chain management with blockchain

Government identity management with blockchain

Cargo management with blockchain

IoT platform based on blockchain

Clinical trial management with blockchain

### Submission Methods

1. Email: [icbct@academic.net](mailto:icbct@academic.net)
2. Electronic Submission System: <http://confsys.iconf.org/submission/icbct2020>

### Important Dates

Submission Deadline	November 15, 2019
Notification Deadline	December 05, 2019
Registration Deadline	December 25, 2019
Conference Dates	March 12-14, 2020

[www.icbct.org](http://www.icbct.org)



2020 The 3rd International Conference on Big Data and Education (ICBDE 2020) will be held in Imperial College London, UK during April 01-03, 2020. ICBDE is expected to provide an opportunity for the researchers to meet and discuss the latest solutions, scientific results and methods in solving intriguing problems in the fields of Big Data and education.

**Publication**

Accepted excellent papers by ICBDE 2020 will be published in the International Conference Proceedings, which will be sent to be indexed by EI Compendex and Scopus.

**Topics**

Topics of interest for submission include, but are not limited to:

- |  |  |
|--|--|
| <ul style="list-style-type: none"> <li>Novel Theoretical Models for Big Data</li> <li>New Computational Models for Big Data</li> <li>Data and Information Quality for Big Data</li> <li>Algorithms and Systems for Big Data Search</li> <li>Distributed, and Peer-to-peer Search</li> <li>Big Data Search Architectures, Scalability and Efficiency</li> <li>Big Data as a Service</li> <li>Big Data Industry Standards</li> <li>Experiences with Big Data Project Deployments</li> <li>Novel Theoretical Models for Big Data</li> </ul> | <ul style="list-style-type: none"> <li>Intrusion Detection for Gigabit Networks</li> <li>Anomaly and APT Detection in Very Large Scale Systems</li> <li>High Performance Cryptography</li> <li>Visualizing Large Scale Security Data</li> <li>Portals and Virtual learning</li> <li>Partnerships in e-Learning</li> <li>Systems, Design and Technologies</li> <li>Information and Computer Education</li> <li>Language and Creative Writing</li> </ul> |
|--|--|

**Submission Methods**

3. Email: [icbde@academic.net](mailto:icbde@academic.net)
4. Electronic Submission System: <http://confsys.iconf.org/submission/icbde2020>

**Important Dates**

Submission Deadline	November 30, 2019
Notification Deadline	December 20, 2019
Registration Deadline	January 10, 2019
Conference Dates	April 01-03, 2020

[www.icbde.org](http://www.icbde.org)



