



KING MONGKUT'S UNIVERSITY OF TECHNOLOGY NORTH BANGKOK  
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ



To Encourage Innovation in Science and  
Technology Through the Development of People

# JCSSE 2020

## Book of Abstracts



## 2020 17th International Joint Conference on Computer Science and Software Engineering (JCSSE)

November 5, 2020	
09.00 - 09.30	Opening Ceremony
09.30 - 10.15	Keynote Speaker I [Professor Dr. Yuefeng Li, Queensland University of Technology] Title: "AI-based Data Analyses for Text Classification and Document Summarization"
10.15 - 10.30	Refreshment Break
10.30 - 11.15	Keynote Speaker II [Associate Prof. Dr. Chih-Lin Hu, National Central University] Title: "Smart Home with Software Intelligence, IoT and AI"



Professor Dr. Yuefeng Li

Professor Yuefeng Li is the leader of AI-Based Data Analysis Group in the School of Electrical Engineering and Computer Science, QUT. He has published over 190 refereed papers (including 65 journal papers). He has demonstrable experience in leading large-scale research projects and has achieved many established research outcomes that have been published and highly cited in many significant Journals and Conferences. He has been a program chair of several International Conferences and workshops. He is the Editor-in-Chief of Web Intelligence Journal.

**Research Areas:**

- Text mining and analysis


His contributions comprise pattern based models and relevance feature discovery to improve the effectiveness of information filtering. Another major contribution has been introducing theory and model into text classification research and document summarization from the field of three-way decision theory.

- Topic Modelling and Feature Selection

His contributions include models for integrating LDA (Latent Dirichlet Allocation, a famous topic modelling method) and pattern mining, and using extended random sets to interpret latent LDA topics to enhance the performance of topic modelling and feature selection.

- Ontology Learning and Web Intelligence

Ontology learning is an automatic or semi-automatic process for building of ontologies (knowledge bases) to enhance text analysis and support conversation between machine and human being. His contributions comprise ontology learning frameworks for information gathering and a new methodology for personalized ontology construction and its application for the interpretation of text mining results.

	<p>- Foundations of Data mining</p> <p>His contributions include algorithms for mining non-redundant rules and a theory for extending association rule mining and clustering into granule mining.</p>
 <p>Associate Prof. Dr. Chih-Lin Hu</p>	<p>Chih-Lin Hu received the BS degree in computer science from the National Cheng-Chi University in 1997, the MS degree in computer science from the National Chung-Hsing University in 1999, and the PhD degree in electrical engineering from the National Taiwan University in 2003. He was a researcher at BenQ and Qisda Advanced Technology Centers, Taipei City, Taiwan, from 2003 to 2007. Since 2008, he joined in the Department of Communication Engineering, National Central University, Taoyuan, Taiwan, R.O.C., where he was an assistant professor from 2008 to July 2012 and has been an associate professor since August 2012. He was a visiting researcher in Academia Sinica, Taiwan, from February to July 2018, and also a visiting researcher in Smart House Research Center, Kanagawa Institute of Technology, Japan, in 2018. He served as the Director, Career Center, NCU Office of Student Affairs during August 2013 to January 2016. He had the honor to get the BenQ innovation awards in 2006 and 2007, the best paper awards in IEEE ICPADS 2000 and IET FC 2012, the student best paper award in IEEE VTS APWCS 2013, TIEEE Best Thesis Award 2014, the First Prize Paper Award in IEEE ICASI 2017, the Excellent Paper Award in Taiwan Internet Workshop 2018, the new scholar outstanding research award in College of Electrical Engineering &amp; Computer Science, National Central University (NCU) in 2011, NCU Service Excellence Award in 2012, NCU Excellent Career Tutor Awards in 2012 and 2013, NCU Distinguished Patent Awards in 2013, 2015 and 2016, NCU Excellent Innovation Team Award in 2018, The 10th Taipei International Invention Show and Technomark - Bronze Medal Award in 2014, TIEEE Best Thesis Award in 2014, as well as Hewlett Packard (HP) Enterprise AI Hackathon (Taiwan Region) The First Prize Award in 2018. He had coorganized MDM'09 Workshop on Mobile Peer-to-Peer Information Services (MP2PIS), IEEE PerCom'10, '11, '12 and '13 Workshops on Mobile Peer-to-Peer Computing (MP2P), IEEE ICPADS'14 Workshop on Advanced Data and Security Technologies in Wireless and Mobile Networks (ADST-WMN'14), Ubi-Media'15, '16, '17, '18, and '19, as well as The 1st, 2nd and 3rd Workshops on Multimedia and Internet of Thing (MIOT) in 2017, 2018 and 2019. He had editorial services in International Journal of Mobile Network Communications &amp; Telematics, International Journal of Wireless and Mobile Networks, International Journal of Mobile Device Engineering, and a special issue of Internet Journal of Internet Protocol Technology. His research interests include mobile and pervasive computing systems, consumer communications and networking (smart home), Internet of Things (IoT), and broadcast information system. Dr. Hu is a senior member of the IEEE.</p>

November 5, 2020

Parallel Track 1: IOT and System Track

<b>Time</b>	<b>Page</b>	<b>Paper ID</b>	<b>Paper Title</b>	<b>Author(s)</b>
13.00 - 13.20	9	1570620007	Personalized Walking Exercise Support System for Elderly based on Machine Learning	Watsawe Sansrimahachai
13.20 - 13.40	2	1570628743	Smart Greenhouse Management System based on NB-IoT and Smartphone	Fan Zhang, Xuefen Wan, Tao Zheng, Jian Cui, Xuedong Li, and Yi Yang
13.40 - 14.00	4	1570633850	Smart Collar Design to Predict Cow Behavior	Supakit Nootyaskool and Pimolrat Ounsrimuang
14.00 - 14.20	5	1570634054	Improving Accuracy of an AoA-based Wi-Fi Indoor Localization using Kalman Filter	Boonsit Yimwadsana and Vichhaiy Serey
14.20 - 14.40	6	1570629031	Reconfigurable Packages for Low Cost Sensor Prototyping	Matthew Nakamura, Noah Hafner, Wansuree Massagram, and Joseph Brown
14.40 - 15.00	7	1570635581	Data Integration Patterns for IIoT Systems	Roland Petrasch
15.00 - 15.20	1	1570634039	Building Two Communication Towers over a Terrain	Chawin Aiemvaravutigul, Jittat Fakcharoenphol, and Nattawut Phetmak
15.20 - 15.40	3	1570628957	Overhead Study of Telegraf as a Real-time Monitoring Agent	Prapaporn Rattanathamrong, Yootana Boonpalit, Siwakorn Suwanjinda, Ayuth Mangmeesap, Ken Subraties, Vahid Daneshmand, Shava Smallen, and Jason Haga
15.40 - 16.00	30	1570619525	Initial Study of Low Cost Crepitus Joint Degradation Finding using Acoustic Localization	Leong Yeng Weng, Hiroaki Seki, Adzly Anuar, Chandran Nadarajan, Mohd Hanifah Jusoh, Muhammad Fairuz, Abdul Jalal, Hang See Ong, and Khairul Salleh Mohamed Sahari

November 5, 2020

Parallel Track 2: Data Science Track

<b>Time</b>	<b>Page</b>	<b>Paper ID</b>	<b>Paper Title</b>	<b>Author(s)</b>
13.00 - 13.20	11	1570629279	TAE-ML:A Random Forest Model for Detecting RNA Editing Site	Pariwat Tongnueasuk and Duangdao Wichadakul
13.20 - 13.40	12	1570633755	Getting to Know One's Role in the Team through Personality-based Clustering	Jidapa Kraisangka, Wudhichart Sawangphol, PunnawichRojcharoenpreeda, Chosita Tangchadakorn, Mathas Vechjatuporn, Chanya Limpasitiponm, Panumas Itthisaeng, and Sirapat Boonwan
13.40 - 14.00	13	1570633869	Globular Cluster Classification in Galaxy M81 using Machine Learning Techniques	Phatcharapon Sookmee, Chutipong Suwannajak, Prapaporn Techa-Angkoon, Benjamas Panyangam, and Nahathai Tanakul
14.00 - 14.20	14	1570627170	Feature Distillation of High Dimension Datasets: Dimension Contraction and Component Dilation	Sabari Vishnu, Jayanthan Jaikrishnan, and Kanchana Viriyapant
14.20 - 14.40	28	1570633874	A Classification of Biochemical Compounds based on their Primitive Structures and Graph Kernels	Peemapat Wongsriphisant, Chidchanok Lursinsap, Apichat Suratane, and Kitiporn Plaimas
14.40 - 15.00	32	1570662412	A Collaborative Filtering Model based on Matrix Factorization and Trust Information	Dussadee Praserttipong and Wijak Srisujalertwaja
15.00 - 15.20	23	1570661606	GPU-Accelerated Method for Simulating Efficient Portfolios in the Mean-Variance Analysis	Paradorn Charoenphaibul and Nopadon Juneam
15.20 - 15.40	27	1570634020	Primary Diagnosis Prediction from Chief Complaints	Somporn Pettakorn, Vipas Sutantayawalee, and Ratchainant Thammasudjarit

November 6, 2020

09.00 - 09.45

Keynote Speaker III [Professor Dr. David Abramson, University of Queensland]  
Title: "Data intensive infrastructure"



Professor Dr. David Abramson

Professor Abramson has been involved in computer architecture and high performance computing research since 1979. He has held appointments at Griffith University, CSIRO, RMIT and Monash University. Prior to joining UQ, he was the Director of the Monash e-Education Centre, Science Director of the Monash e-Research Centre, and a Professor of Computer Science in the Faculty of Information Technology at Monash. From 2007 to 2011 he was an Australian Research Council Professorial Fellow. Abramson has expertise in High Performance Computing, distributed and parallel computing, computer architecture and software engineering. He has produced in excess of 200 research publications, and some of his work has also been integrated in commercial products. One of these, Nimrod, has been used widely in research and academia globally, and is also available as a commercial product, called EnFuzion, from Axceleon. His world-leading work in parallel debugging is sold and marketed by Cray Inc, one of the world's leading supercomputing vendors, as a product called ccbd. Abramson is a Fellow of the Association for Computing Machinery (ACM), the Institute of Electrical and Electronic Engineers (IEEE), the Australian Academy of Technology and Engineering (ATSE), and the Australian Computer Society (ACS). He is currently a visiting Professor in the Oxford e-Research Centre at the University of Oxford.

November 6, 2020

Parallel Track 1: Data Mining Track

Time	Page	Paper ID	Paper Title	Author(s)
11.00 - 11.20	15	1570633835	Detection of Drug-Associated Rhabdomyolysis Through Data Mining Techniques	Patratorn Kunakorntham, Oraluck Pattanaprteep, Charungthai Dejthevaporn, Ratchainant Thammasudjarit, and Ammarin Thakkinstian
11.20 - 11.40	33	1570665462	Hybrid Multi-population Evolution based on Genetic Algorithm and Regularized Evolution for Neural Architecture Search	Yutana Jewajinda and Phanomphon Yotchon
11.40 - 12.00	10	1570633982	Workload Prediction with Regression for over and under Provisioning Problems in Multi-agent Dynamic	Noppanut Suksriupatham and Apirak Hoonlor

Parallel Track 1: Image Processing and Computer Vision

13.00 - 13.20	8	1570629615	Facial Image Verification for Government Kiosk System	Srirath Sawetsutipun, Thananop Kobchaisawat, and Thanarat H. Chalidabhongse
13.20 - 13.40	16	1570629012	Multiscale and Nonlinearity Convolutional Regression for Locating the Eye's Pupil Center	Phitchapha Lertsiravarameth and Pinyo Taeprasartsit
13.40 - 14.00	20	1570634001	Accelerating Cylinder Detection in Point Clouds using GPU	Krit Chaiso and Paruj Ratanaworabhan
14.00 - 14.20	29	1570623206	Automatic White Matter Hyperintensities Segmentation from Brain Magnetic Resonance Images using Polar Transform	Sirikan Chucherd, Samart Moodleah, and Annupan Rodtook
14.20 - 14.40	31	1570633978	Hand Writing Recognition Model Enhancement Exploration based on Skipped Connections	Phattharaphon Romphet, Supasit Kajkamhaeng, and Chantana Chantrapornchai

November 6, 2020

Parallel Track 2: Blockchain Track

<b>Time</b>	<b>Page</b>	<b>Paper ID</b>	<b>Paper Title</b>	<b>Author(s)</b>
11.00 - 11.20	19	1570633691	An Architecture for Game to Game Data Transfer using Blockchain	Chanon Yaklai and Vishnu Kotrajaras
11.20 - 11.40	17	1570633972	Fault-Tolerant Architecture Design for Blockchain-based Electronics Voting System	Siriboon Chaisawat and Chalee Vorakulpipat
11.40 - 12.00	22	1570633678	Bitcoin Trading Using Deep Reinforcement and Supervised Learning	Daniel Ekwuazi and Chaiwat Nuthong

Parallel Track 2: IT and Security Track

13.00 - 13.20	24	1570628631	Enhancing Visualization Applications using Open Data Sources	Ponlakit Suwanworaboon, Steven Lynden, and Suppawong Tuarob
13.20 - 13.40	18	1570628406	Identification of Critical Factors on Structure and Conduct of ICT Industry using SCP Framework: Evidence Indonesian ICT Industry	Putri Nastiti and Djoko Budiyanto Setyohadi
13.40 - 14.00	25	1570634302	Database System for Royal Thai Orchid Plantation in Chiang Mai Province	Parot Ratnapinda, Watcharin Sarachai, and Pitchayanida Khumwichai
14.00 - 14.20	26	1570634010	Web Encryption Analysis of Internet Banking Websites in Thailand	Suphannee Sivakorn, Patsita Sirawongphatsara, and Nuttaya Rujiratanapat
14.20 - 14.40	21	1570633926	An IDS Rule Redundancy Verification	Piyawat Noiprasong and Assadarat Khurat





## Building Two Communication Towers over a Terrain

Chawin Aiemvaravutigul, Jittat Fakcharoenphol, and Nattawut Phetmak

We consider visibility problems on a terrain. Given two stations located at different points on the terrain, we would like to find the minimum heights of communication towers at the two locations so that they can communicate freely. Formally, given two points  $q_1$  and  $q_2$  on the terrain, we would to find the height  $h_1$  and  $h_2$  such that if one lift  $q_i$  up by  $h_i$ , for  $i=1, 2$ , the lifted points see each other with the minimum cost that depends linearly on  $h_1$  and  $h_2$ . We consider two versions of the problem. When given two points  $q_1$  and  $q_2$  on a terrain with  $n$  vertices, with the preprocessing time of  $O(n \log n)$ , we can answer the query in time  $O(\log n + d)$  where  $d$  is the number of terrain edges between  $q_1$  and  $q_2$ . When one point  $q_1$  is fixed, given  $q_2$  and  $\epsilon > 0$ , we can give an approximate answer with additive error at most  $\epsilon$  in time  $O(\log^2(H_{\max}/\epsilon) \cdot \log n)$  where  $H_{\max}$  is the maximum height in terrain. In this case, the preprocessing time is  $O(n^2)$ .



## **Smart Greenhouse Management System based on NB-IoT and Smartphone**

Fan Zhang, Xuefen Wan, Tao Zheng, Jian Cui, Xuedong Li, and Yi Yang

In order to realize intelligent agricultural greenhouse management and improve the density and refinement of the sensing and control terminal, a smart greenhouse management system based on NB-IoT (Narrow Band Internet of Things) and smartphone is designed. The terminal node supports NB-IoT network and Bluetooth communication, periodically collect environmental data such as temperature, humidity, light intensity, wind speed and wind direction, and uploads it to the cloud platform in real time through the NB-IoT network. The cloud platform that suit for NB-IoT is designed, and includes data storage, server-side and web applications for centralized management. The web application implements functions such as user management, multi-node information management and real-time monitoring, alarm recording, and historical query. The Android APP is designed for distributed on-sited management of greenhouses, implements real-time monitoring and temporary data storage of a single node. The test results show that the system has stable data transmission, low packet loss rate in the greenhouse environment which meets the real-time management requirements for the greenhouse.



## **Overhead Study of Telegraf as a Real-time Monitoring Agent**

Prapaporn Rattanathamrong, Yoottana Boonpalit, Siwakorn Suwanjinda,  
Ayuth Mangmeesap, Ken Subraties, Vahid Daneshmand, Shava Smallen, and  
Jason Haga

Large-scale distributed systems have become an essential part of our everyday life. These systems have a large number of hardware and software components, often cooperating in complex and unpredictable ways. Operating these kinds of systems requires centralized monitoring to understand their overall states. While running software to collect metrics in a server is considered common nowadays, it often goes unstudied the impact metric collection software have on the base system. This is especially important in low-power, IoT applications. According to our review, one particular software, Telegraf, has never been formally studied before in terms of how much overhead Telegraf adds to the base system. In this work, we conducted several experiments to study how the base system is affected by Telegraf in two scenarios: a datacenter server and an IoT node. The results show that Telegraf is lightweight and suitable to serve as a real-time monitoring agent in both scenarios.



## **Smart Collar Design to Predict Cow Behavior**

Supakit Nootyaskool and Pimolrat Ounsrimuang

Productivity from agriculture and farming is one major to drive Thailand's economy. Low price sensors and easy to make the Internet of Thing devices for making the data collector applied in various fields such as industry, communication, and transportation. However, smart application in agriculture does not apply widespread usage. This research proposed the detail design of the smart cow collar to use monitoring cow healthy. Data of walking and scraping of the cow sent to collect on the server and report to the user known the cow healthy by relating to the behavior. The purpose of the device helps the cow keeper reducing the monitoring of cow behaviors and it can predict illness and vital of the cow occurring in the future. For predicting part, design by the data from the gyroscopes after preparing send to Hidden Markov Models to predict cow behavior, however, this paper mainly focuses on the detail building the cow collar and creating the prototype device, then testing on the real environment.



## **Improving Accuracy of an AoA-based Wi-Fi Indoor Localization using Kalman Filter**

Boonsit Yimwadsana and Vichhaiy Serey

Indoor location-based system (Indoor LBS) has increasingly attracted attentions in research and industrial community in the recent years. However, the adoption of indoor LBS is still slow due to many obstacles, in particular, its low accuracy performance. Many real-world applications require challenging performance targets such as real-time operation, high accuracy, and energy efficiency. In order to meet the requirements fast and accurate positioning methods are necessary. Noise from interference and multipath in the indoor environment is one of the most important factors preventing accurate positioning. Since it is difficult to make changes to the sensor technologies especially in the hardware, improving the accuracy of indoor positioning by removing noise from the positioning measurement offers an effective solution to the accuracy problem for indoor LBS. We propose a Kalman Filter method which could be applied to the measurements of the indoor LBS. The result from the experiment shows that the positioning accuracy has improved for over 30 percent.



## **Reconfigurable Packages for Low Cost Sensor Prototyping**

Matthew Nakamura, Noah Hafner, Wansuree Massagram, and Joseph Brown

Presented are sensor packages for prototyping sensors and network configurations to support humanitarian assistance and disaster relief efforts. Hardware prototypes were assembled using commercially available modules in 3d printed enclosures. The cost of Components for a single unit is approximately US\$110. Hardware and software implementation for mobile sensors and network gateways is described with cost vs. capability design decisions highlighted. Results are presented from exercise in Hawaii and Thailand with 15 sensor nodes collecting environmental data and position while worn by people or attached to vehicles. Personnel tracking capability is demonstrated over Long Range (LoRa) data link. 15 sensor nodes each transmitted four data updates messages per minute to three gateway nodes with single channel radios. The aggregate message rate of 1 message/second was above the capacity limit of a single gateway, and sensor nodes automatically switched to use all the gateways.



## **Data Integration Patterns for IIoT Systems**

Roland Petrasch

IIoT (Industrial Internet of Things) systems are often part of a larger enterprise application system, e.g. in the context of manufacturing machines produce sensor data that need to be merged with ERP data (customer orders, machine maintenance plans etc.). Before this data merger can be performed, a concept for IoT data integration into the enterprise data needs to be provided in order to avoid data inconsistencies and an overload of unused data. Another aspect is the right layer for data conversion, down- or upsampling, merging, aggregation etc.: Not all IoT data points need to be transferred and merged with the enterprise data. Different layers like sensor device, gateway, edge/fog, cloud must be considered. This paper presents a conceptual model based on data integration patterns that are specialized on IIoT data. The data structures of the IoT data and the data flow across the layers are analyzed and patterns for data integration can be applied. This IoT data integration concept that also uses enterprise integration patterns is a robust basis for an implementation or code generation. The approach is exemplified with a simple IoT sensor data flow that needs to be merged with ERP production data in order to monitor the manufacturing process in real-time.



## **Facial Image Verification for Government Kiosk System**

Srirath Sawetsutipun, Thananop Kobchaisawat, and Thanarat H. Chalidabhongse

Some government organizations have been started mega-projects from which Thai citizens can access their welfare data and receive services from the government via kiosk machines. The kiosk machines are typically installed at the government offices and public places. To access any data or services, the user authentication is done by the Thai citizen smart ID card. However, the weakness of the current system is person verification; it is possible that someone uses the other's ID card. In this paper, we propose a facial image verification system prototype that can integrate to the existing well-fare kiosks. The face verification is done by comparing the face image retrieved from the smart ID card and the face image taken from a camera on the kiosk. The prototype system aims to cross-check the ownership of the smart card holder and to protect personal data privacy. We implemented Face Verification APIs based on the ResNet and MobileNet. From the experimental results, when tested with Cross-Age LFW face dataset, the accuracies are 99.996% and 99.988% on the ResNet50 and the MobileNet, respectively. However, the accuracies reduce to 96.296% and 95.556% when tested on the low-quality face images that embedded in the national ID cards. The system is designed as a client-server model using Microservices Architecture (MSA) style that complements "loosely coupled" multiple services which can be leveraged to share important functionality throughout other government agencies.





## **Personalized Walking Exercise Support System for Elderly Based on Machine Learning**

Watsawe Sansrimahachai

Over the last decade, the world's population structure has significantly changed where the proportion of elderly people in the population increases and that of young people declines considerably. This change of demographic landscape has caused a severe manpower shortage. It reduces the pool of people in their productive years. Several countries have conducted plans to extend the retirement age to keep elderly workers in the labor market longer. However, maintaining health and physical ability among elderly workers is a primary issue. To address the problem, regular physical exercise is crucial. In this paper, we propose an innovative software system that support the elderly to participate in walking exercise regularly. Our system uses a mobile phone together with a fitness tracker as data collection devices. It applies an artificial neural network learning algorithm to construct a regression-based predictive model that can predict exercise minutes based on exercise history and walking statistics. The exercise minutes predicted are then used to generate a personalized walking exercise plan designed for each individual elder. The experimental result indicates that our predictive model is practical to be used in real-world settings with reasonable prediction accuracy, about +/- 17 weekly exercise minutes observed.



## **Workload Prediction with Regression for over and under Provisioning Problems in Multi-agent Dynamic**

Noppanut Suksriupatham and Apirak Hoonlor

The infrastructure of cloud computing is distributed environment, where resources, such as CPU, memory, storage, and network bandwidth, are shared. The cloud computing system can experience the over and under provision problems. Recently, the problems were handled using the multi-agent dynamic resource provisioning framework, where Machine Learning algorithm is used for workload prediction. We empirically investigated the effectiveness of regression models for the over and under provisioning on cloud system. We found that while linear regression, polynomial regression, and support vector machine (SVM) performed equally well in term of workload prediction. However, polynomial regression was able to distribute the load among the hosts better. This resulting in the lower total energy consumption. Overall performance wise, all regression model requires little time for parameter tuning. As such, we suggest that the model's parameter should be readjusted as needed.



## **TAE-ML: A Random Forest Model for Detecting RNA Editing Site**

Pariwat Tongnueasuk and Duangdao Wichadakul

RNA editing is a post-transcriptional alteration that changes the RNA sequence by insertions, deletions and base substitutions. These changes can directly affect protein expression and structure. In humans, the most common RNA editing is adenosine to inosine (A-to-I editing) but in next-generation sequencing, inosine is read as guanine (A-to-G) and mostly found in Alu regions. With this information, several algorithms and software tools have been developed for identifying RNA editing sites. In this paper, we present TAE-ML, a random forest model for detecting RNA editing sites from SNVs in a VCF file called from an RNA-Seq experiment. TAE-ML was integrated with multiple filters to remove artificial variants and incorporated the surrounding sequence of SNVs to accurately identify RNA editing sites. The YH dataset from previous research was used for training the model. The HeLa, CH24T, and CH62T datasets were used for testing. Compared with RED-ML and RDDpred with the same candidate sites, TAE-ML got a higher accuracy and F1 score than both tools. These results suggested that our model is capable of predicting true RNA editing sites. We designed TAE-ML to be a simple python script to make it easy to run on various operating systems that support python. The model is available at <https://github.com/cucpbioinfo/TAE-ML>.



## **Getting to Know One's Role in the Team through Personality-based Clustering**

Jidapa Krajangka, Wudhichart Sawangphol, PunnawichRojcharoenpreeda,  
Chosita Tangchadakorn, Mathas Vechjatuporn, Chanya Limpasitiponm,  
Panumas Itthisaeng, and Sirapat Boonwan

Matching people to the right job with the right role has been a challenge for any organizations. In addition to the general criteria from the resume, we can use personality data from the personality test to identify the suitable role of an employee in the team. In this paper, we explore the characteristic of roles in the team through the personality-based clustering. Data from the personality assessment of the 573 anonymous employees were used in the analysis. After the series of experiments, we obtained the optimal model with k-mean algorithm ( $k=3$ ). We further verified the clusters with the psychologists to identify the role in the team characterized by each cluster. As a result, together with the psychologist research team, we identified three main roles in a team, including logicians, supporters, and adapters.



## **Globular Cluster Classification in Galaxy M81 using Machine Learning Techniques**

Phatcharapon Sookmee, Chutipong Suwannajak, Prapaporn Techa-Angkoon,  
Benjamas Panyangam, and Nahathai Tanakul

Globular clusters are very important in astronomy since they can be used to study the process of galaxy formation and evolution. With the exponential growth of data in astronomy, it is currently inefficient to classify globular clusters from the other types of astronomical objects by humans. In this study, we explore the possibility of using machine learning in globular cluster classification to replace the classification by humans. We selected five standard classification methods including k-NN, Random Forest, SVM, Neural Network, and Decision Tree. All models were built and tested by using the Weka software with datasets from a galaxy M81. Our experiments showed that k-NN, Random Forest, and SVM are the best approaches for globular cluster classification, with 97.7% accuracy, 97.8% precision, 97.7% recall, and 97.7% F-measure. Finally, when we applied these models to an unseen dataset to predict new globular cluster candidates, we acquired 6.32% succession rate compared to 30% succession rate by humans. This suggests that machine learning techniques can be applied to globular clusters classification. However, our models need to be improved to achieve a higher success rate to replace the classification by humans.



## **Feature Distillation of High Dimension Datasets: Dimension Contraction and Component Dilation**

Sabari Vishnu, Jayanthan Jaikrishnan, and Kanchana Viriyapant

Handling high dimensional datasets is always a challenge in machine learning. The operations that deal with datasets containing vast amounts of features are extremely time-consuming and computationally expensive. In an even more challenging scenario, a data scientist may be given a dataset without any information about the target feature and he/she needs to find out the features that can be possibly predicted from the dataset without loss of interpretation. To satisfy this requirement we introduce a new technique, "Feature Distillation". In addition to that a new algorithm, Dimension Contraction and Component Dilation (DCCD) is proposed. The idea is to return features that can be predicted with high accuracies from a dataset using a combination of information measurement, discretization, dimensionality reduction, supervised and unsupervised techniques. The DCCD algorithm is tested with various synthetic datasets and proven to be faster than traditional methods.



## **Detection of Drug-Associated Rhabdomyolysis through Data Mining Techniques**

Patratorn Kunakorntham, Oraluck Pattanapruteep, Charungthai Dejthevaporn,  
Ratchainant Thammasudjarit, and Ammarin Thakkinstian

Rhabdomyolysis (RM) is a life-threatening adverse drug reaction (ADR). Statins are the common drugs causing RM and weakness as well as a combination with other drugs, which increase the level of statins. The estimated cost per QALY of supportive treatment for RM was 69,742.50 USD per year, but the early detection of ADRs can reduce the cost of approximately 1,400.00 USD per patient. RM has a problem of under-reporting in spontaneous reporting systems (SRSs). Detecting RM at the early stage is the crucial task by finding the relationship between drugs and RM from electronic health records (EHRs). The aim of this study is to propose a predictive model for RM analysis by predicting the probability of RM or weakness in patients who used statin alone or combined with other drugs. The proposed model can predict the probability occurrence of RM or weakness with sensitivity equal to 0.66.



## **Multiscale and Nonlinearity Convolutional Regression for Locating the Eye's Pupil Center**

Phitchapha Lertsiravarameth and Pinyo Taeprasartsit

Pupil detection has been an active research topic in recent years due to its usability in many areas including interaction design and medical diagnosis. Many existing methods are accurate in pupil localization; however, many of them were designed and tested under specialized lighting conditions. Furthermore, there are challenges to cope with images whose color of the iris and the pupil are both dark. In this work, we proposed deep regression models to locate the eye's pupil center in target groups whose iris and pupil are both dark in color and images were acquired in a typical lighting condition. The proposed neural network model concurrently utilizes features from multiple convolutional layers for final regression. In other words, features from different layers which contain different degrees of nonlinearities are concatenated so that a neural network model can explicitly employ these features together for predicting the location of the eye's pupil. Ablation analysis indicated that these model characteristics were essential for robustness. All experimented models were trained and validated against 2,500 eye images from the MPIIGaze dataset. Additional 2,515 test images were obtained from the MMU2 and BioID datasets to evaluate robustness of the method. Our experiments showed that robust and accurate localization of the pupil center could be achieved over datasets whose image were acquired under non-specialized lighting conditions. Our best model had average pixel error of 1.11 pixels, and had error within two pixels for 95.5 percent of test samples.





## **Fault-Tolerant Architecture Design for Blockchain-based Electronics Voting System**

Siriboon Chaisawat, Chalee Vorakulpipat

With the emergence of distributed ledger technology, data decentralization and data integrity in information storage are promised to be preserved. An inevitable challenge in e-voting system implementation is the trading off between system security and generality in design. Leveraging container technology along with blockchain technology, system design architecture proposed in this paper promises to deliver security in data as well as agility in system integration. The design was further analyzed and assessed by measuring utilization of computing resources and conducting performance test as to assure the aforementioned requirements can be preserved.



## **Identification of Critical Factors on Structure and Conduct of ICT Industry using SCP Framework: Evidence Indonesian ICT Industry**

Putri Nastiti and Djoko Budiyanto Setyohadi

ICT companies that focus on the field of software development are part of the era of the creative economy in Indonesia. The categories of business that are carried out include e-commerce, fin-tech, games, and so on. This study aims to identify critical factors in the industrial structure and behavior of ICT companies that affect company performance. The sample used was 64 ICT companies in the Yogyakarta area that focused on software development. Respondents consisted of top-level and middle management, who understood the real condition of the industry structure and company behavior. In this study, it was concluded that five main factors are consisting of two factors of industrial structure and three factors of corporate behavior that affect performance, including the funding factor, human resource competence, the importance of market research, clear organizational structure, and IT-business strategic alignment.



## **An Architecture for Game to Game Data Transfer using Blockchain**

Chanon Yaklai and Vishnu Kotrajaras

Blockchain has been used in several industries, especially in the financial area. In the game industry, blockchain has not been utilized much even with a large number of users who deserve to own their game data. There were attempts to allow players to own their in-game objects or data. But they were mostly crowdfunding platforms which were not yet in operation. In addition, users were required to use their ecosystem. In this paper, we propose a generic game architecture, using blockchain, that allows players to own their "spent time" in a game. The value of this spent time can be passed on to new games as ERC20 token on the Ethereum network and redeemed in such games. Our architecture can be used with any public blockchain. It can benefit many roles such as player, blockchain node, and game developer. Our experiment shows that the proposed architecture can be accepted by players, but other factors such as the type of gameplay and how the spent time is exchanged in a game can greatly affect the player acceptance.



## **Accelerating Cylinder Detection in Point Clouds using GPU**

Krit Chaiso and Paruj Ratanaworabhan

Point cloud data processing pertaining to primitive shape detection has become an active research area during the past ten years. One primitive shape that is of great interest to the research community is the cylinder. At present, there are two popular solutions for cylinder shape detection, Hough transform and RANSAC; the latter is known to be simpler and more robust than the former. This paper focuses on RANSAC cylinder detection. Specifically, it aims to parallelize the sequential RANSAC by employing GPU parallelization techniques. CUDA from NVIDIA is our choice of framework for this purpose. We evaluate our parallel RANSAC using three datasets, TUW, OSD, and WILLOW. The result shows comparable accuracy with the sequential version, but the parallel version can execute much faster on two datasets. We have released its source code under an open source license at: <https://github.com/xx-yy> (page url temporarily removed to preserve anonymity).



## **An IDS Rule Redundancy Verification**

Piyawat Noiprasong and Assadarat Khurat

Intrusion Detection System (IDS) is a network security software and hardware widely used to detect anomaly network traffics by comparing the traffics against rules specified beforehand. Snort is one of the most famous open-source IDS system. To write a rule, Snort specifies structure and values in Snort manual. This specification is expressive enough to write in different way with the same meaning. If there are rule redundancy, it could distract performance. We, thus, propose a proof of semantical issues for Snort rule and found four pairs of Snort rule combinations that can cause redundancy. In addition, we create a tool to verify such redundancy between two rules on the public rulesets from Snort community and Emerging threat. As a result of our test, we found several redundancy issues in public rulesets if the user enables commented rules.



## **Bitcoin Trading using Deep Reinforcement and Supervised Learning**

Daniel Ekwuazi and Chaiwat Nuthong

The cryptocurrency market is known for its volatility; consequently, issues arise regarding when to buy and what quantity to buy, to maximize return in the long term. This paper aims to implement a deep reinforcement learning and supervised learning; Long short-term memory (LSTM) approach to cryptocurrency trading, specifically bitcoin. Experimental results showed that deep reinforcement learning outperformed the supervised learning approach.



## **GPU-Accelerated Method for Simulating Efficient Portfolios in the Mean-Variance Analysis**

Paradorn Charoenphaibul and Nopadon Juneam

This paper considers portfolio optimization whose goal consists in finding a set of efficient portfolios regarding the framework of the Mean-Variance Analysis. Our work utilizes the GPU's computing capabilities to accelerate the computation to simulate such an efficient portfolio set. In particular, we present a non-trivial GPU-accelerated method to produce a set of minimum variance portfolios for a given target range of expected returns under the basic constraints of the framework. We evaluate the experimental performance of the method by synthesizing an implementation using CUDA. The experiment results show that our implementation performs substantially faster than its implementation counterpart using CLAPACK with respect to the task of simulating the efficient frontier on large data sets with the number of assets in the range of hundreds.



## **Enhancing Visualization Applications using Open Data Sources**

Ponlakit Suwanworaboon, Steven Lynden, and Suppawong Tuarob

An increasing number of data visualization tools have started to support the automatic generation of modifications, embellishments, and natural language annotations (data facts) to aid in better understanding the data being visualized. Concurrently, many applications in data science now benefit from the use of an increasingly diverse set of open data sources to augment existing data sets to enhance their value. In this paper we present a framework for using open data-based augmentations to generate embellishments and data facts to enhance existing visualizations. Our approach is based on a semi-automated process intended to involve the user, where possible augmentations are automatically ranked based on the data facts they are capable of generating, allowing users to choose augmentations to effectively enhance existing data visualizations in an explorative manner. We show the benefit of suggesting ranked augmentations from one open data source, Wikidata, by demonstrating that a high number of data facts and embellishments can be produced utilizing the top suggested augmentations. Finally, we describe the architecture of a prototype system implementing the approach.





## **Database System for Royal Thai Orchid Plantation in Chiang Mai Province**

Parot Ratnapinda, Watcharin Sarachai, and Pitchayanida Khumwichai

An orchid plantation under the Plant Genetic Conservation Project under the royal initiation of Her Royal Highness Princess Maha Chakri Sirindhorn (RSPG) is located in Mae Rim district, Chiang Mai. The plantation hosts a wide variety of rare Thai orchids which are growth densely on the branch of the trees. Hence, it is very difficult to identify their species, locate a specific orchid, and manage orchid data. The staffs were using handwritten papers to tag an orchid. Thus, this approach had several disadvantages and limitations: vulnerability, difficult to read, and hostile to the orchid plants. We have developed an orchid database system for the plantation. The system consists of three components. First, we installed the Internet of Things weather station which are used for collecting environmental data such as temperature, humidity and wind data. Second, we developed a hand held tool for tagging using RFID technology and a Raspberry Pi. We put hundreds of RFID tags on the branches of the trees. Third, we created an internal website to manage orchid database for the staffs. We evaluated our system by conducting the user satisfactory questionnaire with the stakeholder. The overall system received the score of 4.08 out of 5.



## **Web Encryption Analysis of Internet Banking Websites in Thailand**

Suphanee Sivakorn, Patsita Sirawongphatsara, and Nuttaya Rujiratanapat

With Thailand rapidly moving to a full internet banking ecosystem, the demand for online security has never been needed more than it is today. As the security and privacy of internet users depend on HTTPS, a web encryption protocol, for securing communication between users and web servers, HTTPS is essentially the center of the web ecosystem today. Unfortunately, despite the increasing number of HTTPS adoptions, numerous studies have shown that a large number of websites have adopted HTTPS incorrectly, rendering users vulnerable to information leakages e.g., eavesdropping and man-in-the-middle attacks. The correctness of HTTPS deployment is even far greater for internet banking services due to carrying user's sensitive information and being prime targets for criminal activities. In this paper, we present WEAPONS, a novel black-box testing tool for evaluating the completeness and correctness of web encryption deployment including the deployment of HTTPS, and web encryption-related mechanisms i.e., HSTS, secure cookie, HTTPS redirect, HSTS preload. We use WEAPONS to conduct an assessment of 9 popular internet banking websites in Thailand. We demonstrate that WEAPONS is able to find HTTPS deployment incorrectness. Several of these weaknesses can expose the affected services to man-in-the-middle attacks and sensitive data exposure.



## **Primary Diagnosis Prediction from Chief Complaints**

Somporn Pettakorn, Vipas Sutantayawalee, and Ratchainant Thammasudjarit

Telemedicine is one of the primary health policies in Thailand to improve life quality of people by accessing medical service from anywhere anytime. Many rural areas still have difficulties to get medical services such as primary diagnosis. We propose a web-based application for chief complaint analysis that is publicly available. Our system receives input of chief complaint text then makes predictions of primary diagnosis using machine learning algorithm. Our experiment shows that the model has promising results at 0.77 of accuracy. Our system could be used as a patient screening tool that makes impact to operational excellence for Thailand healthcare.



## **A Classification of Biochemical Compounds based on their Primitive Structures and Graph Kernels**

Peemapat Wongsriphisant, Chidchanok Lursinsap, Apichat Suratane, and  
Kitiporn Plaimas

The data encoded in the form of a graph are widely used in many fields of study. In cheminformatics, the data of biochemical compounds can be encoded as a graph where its vertices are atoms, and its edges are chemical bonds between pairs of atoms. By this representation, we can identify the molecular functions of a labeled biochemical compound with the graph classification methods which most of them depend on graph kernel. In this study, we proposed an algorithm to extract primitive substructures of the compounds and encode them in a new form of graph for feeding them a support vector machine with a graph kernel. The algorithm was tested with four data sets and yield the average accuracy of 0.847. Furthermore, the performance of the algorithm was also better than that of the solely graph kernels without primitive substructure extraction. All in all, the extraction of the primitive substructures can be a great tool in extracting compound structure features in biochemical compound classification.



## **Automatic White Matter Hyperintensities Segmentation from Brain Magnetic Resonance Images using Polar Transform**

Sirikan Chucherd, Samart Moodleah, and Annupan Rodtook

White matter hyperintensities (WMH) are areas of lost cells found in the white matter of the brain presenting hyperintensities. WMH segmentation and detection are the initial steps of many brain diseases examination. Here, we propose an automated method of WMH segmentation designed to deal with brain magnetic resonance imaging (MRI) in polar coordinate system. Moreover, the pattern of clustering used in segmentation is adjusted to achieve the desired cluster properly. Experimental results on cross-sectional images of fluid-attenuated inversion recovery (FLAIR) datasets reveal better performance of the proposed method with simplicity and robustness. The method provides good average accuracies (0.808, 0.803, 0.981 of dice similarity coefficient (DSC), sensitivity (sens) and specificity (spec) respectively) of WHH segmentation in comparison with other method.



## **Initial Study of Low Cost Crepitus Joint Degradation Finding using Acoustic Localization**

Leong Yeng Weng, Hiroaki Seki, Adzly Anuar, Chandran Nadarajan, Mohd Hanifah Jusoh, Muhammad Fairuz, Abdul Jalal, Hang See Ong, and Khairul Salleh Mohamed Sahari

Current methods of joint diagnosis depend on invasive, radiative methods, large and requiring skilled operators which are still not able to assess the joint during motion. Replacing these methods are acoustic methods which are only able to diagnose joint health based on the characteristics sound emitted by a moving joint. This paper looks into informing the medical practitioner of the locations of joint damage instead, hence the feasibility of localizing upon the damaged part of the joint which is represented by a playback of a recording of crepitus sounds emitted from a human joint in 2 dimensional settings. Three microphones, an earphone and a data acquisition device were employed to implement the testing. Using the time-difference-of-arrival (TDOA) between the 3 sensors, the Angle-of-Arrival (AOA) method was use to give the initial localization coordinates. These coordinates were then fed into a gradient descent algorithm to reduce the positioning error of the AOA method, yielding the final position results. Results show that the developed system is able to achieve localization errors of less than 0.01m.



## **Hand Writing Recognition Model Enhancement Exploration Based on Skipped Connections**

Phattharaphon Romphet, Supasit Kajkamhaeng, and Chantana Chantrapornchai

Handwriting recognition is a well-known problem utilizing a deep learning approach. Due to the variety of words and writing styles, large data set is required to build the good model. In this research, we investigate the state-of-the-art handwriting model, namely Flôr model, with our data set, and the improvement options to improve the accuracy while maintaining the network weight size. We aim to deploy the model on the mobile application. The model size is also concerned. The data set is constructed by Research Group on Computer Vision and Artificial Intelligence INF, University of Bern, containing 115,320 words. The results show that our refined model can reduce 1.5 percent in character base error and 2 percent in word base error with the unseen test data set.



## **A Collaborative Filtering Model based on Matrix Factorization and Trust Information**

Dussadee Praserttipong and Wijak Srisujjalertwaja

Collaborative filtering (CF) approach comprises of several well-known techniques, which successful in creating personalized recommendations. However, the performance of CF recommender system is dramatically degraded when dealing with the data sparsity and cold start problems. To alleviate with these problems, new collaborative filtering model based on user rating matrix factorization and trust information, named Trust-rRSVD. The empirical experiment was established. The accuracy results acquired from the proposed model depicted that Trust-rRSVD outperformed both trust-based techniques and ratings-based techniques in predictive accuracy across different testing views and across users with different trust degrees.





## **Hybrid Multi-population Evolution based on Genetic Algorithm and Regularized Evolution for Neural Architecture Search**

Yutana Jewajinda and Phanomphon Yotchon

This paper presents a hybrid multi-population evolution algorithm by combining genetic algorithms and regularized evolution for neural architecture search. The total population is divided into multiple tribes. In each tribe, there are multiple subpopulations with different crossover and mutation strategies to provide diversity. The best individual in subpopulation are migrated between subpopulations and tribes to guide search direction toward solutions. The experimental results show that the proposed algorithm performs better than traditional multi-population genetic algorithms and comparably compete with regularized evolution and other state of the art NAS algorithm on CIFAR-10 dataset.