Jason C. Hung · Jia-Wei Chang · Yan Pei *Editors* 

# Innovative Computing Vol 1 - Emerging Topics in Artificial Intelligence

Proceedings of IC 2023



### Series Editors

Leopoldo Angrisani, Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Napoli, Italy

Marco Arteaga, Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico

Samarjit Chakraborty, Fakultät für Elektrotechnik und Informationstechnik, TU München, München, Germany Jiming Chen, Zhejiang University, Hangzhou, Zhejiang, China

Shanben Chen, School of Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China

Tan Kay Chen, Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore

Rüdiger Dillmann, University of Karlsruhe (TH) IAIM, Karlsruhe, Baden-Württemberg, Germany Haibin Duan, Beijing University of Aeronautics and Astronautics, Beijing, China

Gianluigi Ferrari, Dipartimento di Ingegneria dell'Informazione, Sede Scientifica Università degli Studi di Parma, Parma, Italy

Manuel Ferre, Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Spain

Sandra Hirche, Department of Electrical Engineering and Information Science, Technische Universität München, München, Germany

Faryar Jabbari, Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA

Limin Jia, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China Janusz Kacprzyk, Intelligent Systems Laboratory, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Alaa Khamis, Department of Mechatronics Engineering, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt

Torsten Kroeger, Intrinsic Innovation, Mountain View, CA, USA

Yong Li, College of Electrical and Information Engineering, Hunan University, Changsha, Hunan, China Qilian Liang, Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA Ferran Martín, Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain

Tan Cher Ming, College of Engineering, Nanyang Technological University, Singapore, Singapore Wolfgang Minker, Institute of Information Technology, University of Ulm, Ulm, Germany Pradeep Misra, Department of Electrical Engineering, Wright State University, Dayton, OH, USA Subhas Mukhopadhyay, School of Engineering, Macquarie University, NSW, Australia Cun-Zheng Ning, Department of Electrical Engineering, Arizona State University, Tempe, AZ, USA

Toyoaki Nishida, Department of Intelligence Science and Technology, Kyoto University, Kyoto, Japan Luca Oneto, Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genova, Genova, Genova, Italy

Bijaya Ketan Panigrahi, Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India

Federica Pascucci, Department di Ingegneria, Università degli Studi Roma Tre, Roma, Italy Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China Gan Woon Seng, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore

Joachim Speidel, Institute of Telecommunications, University of Stuttgart, Stuttgart, Germany Germano Veiga, FEUP Campus, INESC Porto, Porto, Portugal

Haitao Wu, Academy of Opto-electronics, Chinese Academy of Sciences, Haidian District Beijing, China Walter Zamboni, Department of Computer Engineering, Electrical Engineering and Applied Mathematics, DIEM—Università degli studi di Salerno, Fisciano, Salerno, Italy

Junjie James Zhang, Charlotte, NC, USA

Kay Chen Tan, Dept. of Computing, Hong Kong Polytechnic University, Kowloon Tong, Hong Kong

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering—quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

### China

Jasmine Dou, Editor (jasmine.dou@springer.com)

### India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

### Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

### USA, Canada

Michael Luby, Senior Editor (michael.luby@springer.com)

### All other Countries

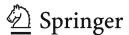
Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

\*\* This series is indexed by EI Compendex and Scopus databases. \*\*

Jason C. Hung · Jia-Wei Chang · Yan Pei Editors

# Innovative Computing Vol 1 -Emerging Topics in Artificial Intelligence

Proceedings of IC 2023



Editors
Jason C. Hung
Department of Computer Science
and Information Engineering
National Taichung University of Science
and Technology, FIET
Taichung City, Taiwan

Yan Pei Computer Science and Engineering University of Aizu Aizuwakamatsu, Fukushima, Japan Jia-Wei Chang
Department of Computer Science
and Information Engineering
National Taichung University of Science
and Technology, FIET
Taichung City, Taiwan

ISSN 1876-1100 ISSN 1876-1119 (electronic) Lecture Notes in Electrical Engineering ISBN 978-981-99-2091-4 ISBN 978-981-99-2092-1 (eBook) https://doi.org/10.1007/978-981-99-2092-1

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023, corrected publication 2023

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd. The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

## **Contents**

The International Workshop on Technique for Language and Literature Information Modeling (LIM 2023)

A Study on Data Mining for Type of Korean Painting Poetry	3
A Big Data Based Learning Model from Student Questionnaire	10
A Phonetic Investigation of Korean Monophthong Vowels by Vietnamese Female Speakers	17
Korean Causal Connective Expressions in a Cross-Linguistic and Cultural Perspective	26
ICIC2023-vol1	
Design and Implementation of Electromagnetic Field Simulation  Platform for AC Testing  Liang Zhang, Lijun Fu, Yanling Guan, Jingyuan Chi, Han Wang, and Jianquan Liang	37
Organizational Layout and Optimization Model of Agricultural Logistics Industry Based on Ant Colony Algorithm	44
Application of Image Recognition in Equipment Monitoring	54
Stability of Marine Physics Detection Sensor Based on Artificial Intelligence Technology	60
Optimization System of Microbial Test on Account of Genetic Algorithm  Mingming Shao	68
The Application of Virtual Reality Technology in Ophthalmology  Jingying Wang and Yi Zhang	76

Research on the Application of BIM Technology in the Whole Process  Cost Management of Construction Project	84
Modal Parameter Identification of Bridge Structure Based on Hybrid Genetic Algorithm	92
Overload Damage Detection Method of Motor Car Axle Based on Neural Network Algorithm	102
Ground Wire Status Object Detection Based on Cloud Application  Haoran Li, Shengli Xie, and Qian Wang	110
Design of Multimedia Assisted Course Based on Immersion Japanese Teaching	120
Analysis and Research of Intelligent Distribution UAV Control System Based on Optical Flow Sensor  Shuai Sun, Yanqi Han, Yue Li, Wanting Zhao, Qifeng Gao, Yang Yang, and Heng Li	128
Method of Improving Logistics Distribution Efficiency by Combining Small Multi-rotor UAV with the Visual Recognition System Shuai Sun, Weilun Xia, Qiankun Zhu, and Zeping Han	138
Prediction Model of Popular Tourist Attractions Based on Big Data Fusion Algorithm	145
Application of Genetic Algorithm in Rail Transit Comprehensive  Monitoring System	152
Construction of Inland VHF Communication System Based on Opposite-Reason Model	159
Intelligent On-Line Monitoring System of General Electrical Equipment Based on Remote Communication  Xinhe Cui, Fei Xia, Fanbo Meng, Xiaobo Huang, and Xin An	168

Sports Video Moving Target Detection and Tracking Based on Particle Filter and Related Algorithms	176
Visual Reconstruction Design Based on Image Technology Emotion	184
System of Cross-Border E-commerce Network Pattern Evolution on Account of Bayes-BP Algorithm	191
Construction of Internet Financial Risk Early Warning Model Based on Data Mining Algorithm	199
Research on an Intelligent Logistics Blockchain Consensus Algorithm  He Yu and Jingshuang Liu	206
Improved ID3 Algorithm in Sports Tourism Service System	215
Design of Public Sports Service System in the Era of 5G + Artificial Intelligence	223
Automatic Design System of Product Shape Based on Intelligent  Algorithm	231
Sports Video Tracking Technology Based on Optimized Decision Tree Algorithm(DTA)  Zhong Wu	239
Entrepreneurship Platform Under Virtual Reality Technology (VRT)  Lili Wang	246
A Model of Diversified Fiscal and Taxation System Based on Computer  Network Technology	254
Research on Smart City Development and Construction Driven by Big Data Technology  Yongling Chu, Zhiling Luan, and Shaochun Li	262

Discipline Inspection and Supervision Big Data Information Query Platform Based on Distributed Computing	269
Changyu Liu, Jinhua Chen, Hongchang Wen, Fanjun Zeng, and Sanjiong Wang	20)
Digital Audit Platform Based on Visual Data Analysis	280
The Division of Labor and Matching System of Teaching Reform Model Based on Multi-objective Classification Algorithm	291
Domestic AI Chip Based on Internet of Things Technology and Artificial Intelligence Algorithm in Power Grid Infrastructure Construction Weihua Zhong, Jun Liu, Jianxiang Xie, and Ji Zhang	299
Data Compression Algorithm of Power System	307
Bank Credit Risk Prediction Based on MCDM and CNN	315
Application of NC Machining Technology in Aircraft Manufacturing  Weiwen Ye	322
Barriers of Mathematical Language Representation of College Students Based on Data Mining	328
Mathematical Model (MM) and Signal Processing (SP) in Computer Simulation (CS) on Account of Improved Genetic Algorithm (IGA)	334
English Translation Intelligent Recognition Model Based on Reverse  Maximum Matching Segmentation Algorithm	342
A New Retraceable Clustering Algorithm	350
Digital Transformation of the Smart Home Industry	357

Design Yu Gui, Yangyang Fan, Jian Tian, Zonghui Bo, Yuxin Miao, Yuxiao Fan, Shunzi Ren, and Xi Chen	4
Seismic Stability Assessment of Civil Building Projects Based on BIM	
Technology	4
Design and Implementation of Virtual Simulation Teaching Platform of Sports Action Technology Based on Maya	4
Industrial Economy and Regional Growth Strategy Based on Particle Swarm Optimization Algorithm Hengran Bian and Yi Liu	4
Design and Implementation of Online Adaptive Teaching System Based on Data Clustering Algorithm	4
Design and Implementation of Online Teaching System Based on J2EE  Qingfeng Guo	4
Design and Implementation of Classroom Teaching Line Analysis Platform Based on Pose Recognition Algorithm (MCG-CPR)	4
Double Attention Mechanism Text Detection and Recognition Based on Neural Network Algorithm	5
A Text Detection and Recognition System Based on Dual-Attention  Mechanism with Artificial Intelligence Technology	5
Data Transmission and Isolation Method for Power Bus Communication  Yuxuan Ye, Rongtao Liao, Lei Li, Shengwei Wang, and Jian Zhang	5
Configuration Method of Safety Link of Electric Power Communication Network	5
Huanjun Hu, Shengwei Wang, Yixi Wang, Lei Li, and Hao Ning	
Method for Locating Communication Device Faults	5

Contents	X1
Digital Twin Simulation System for Computer Room Equipment	547
Fault Detection Methods for Communication Equipment	555
Design and Implementation of Teaching Assistant System Based on Intelligent Recommendation Algorithm	564
Construction of User Profile for Teaching Archive Knowledge Service	
Domain Wenyan Guo, Liming Du, Fengying Wang, Lei Cui, and Jie Dong	571
Artificial Intelligence Technology in Computer Network Security  Cuijin Lao and Shen Qin	579
Dynamic Capacity Expansion System and Control Method for User-Side Power Rongtao Liao, Huanjun Hu, Yixi Wang, Yuxuan Ye, and Jian Zhang	587
Secure and Non-secure Data Processing of Power Bus Communication	597
Dynamic Monitoring Method for Power Big Data Quality	606
Detection of Load Loss of Large Capacity Transformer  Feifeng Wang, Shengchao Jiang, Haoze Zhuo, Tailin Li, Bin Wang, and Yunqing Pei	614
Intelligent Power Measurement System  Chengwu Zou, Dan Ye, Qing Shen, Naiqing Wang, Jiaqi Lu,  Mingquan Zeng, and Hangkang Li	622
Routing Algorithm for Connected Vehicles Based on Bus Route Information	633
Ning Tong, Man Jin, and Kunrui Wang	
Processing Method of Computer Data	642
Storage System of Computer Data	651

Design of Human-Computer Interaction Experience Platform in Virtual Space Under Scientific Information Technology	659
Design of Intelligent Color Matching System for Cultural and Creative Products Based on Data Analysis Algorithm	667
Design of an Intelligent Educational Virtual Classroom Platform Based on the Mobile Terminal APP of the Internet of Things	675
Design and Development of English Autonomous Learning Database Framework Based on Artificial Intelligence Algorithm	684
Devise and Application of Intelligent Guidance System in Exhibition Space  Heng Shen	692
Configuration of Crop Supply Chain Based on Gale-Shapley Algorithm  Wenting Kan and Bin Xiao	700
Application of Particle Swarm Optimization in BIM Building Modeling Guang Yang and Xin Guo	707
Intelligent Prediction Model of Agricultural Environment Based on Intelligent Algorithm	716
Mathematical Modeling and Optimal Design Based on Information Technology Algorithms	721
Internet of Things Data Mining Technology in Sports Physical Health Analysis  Qian Zhang	729
Application of Artificial Intelligence Technology in Big Data Nining  Xueyun Zhou, Yongjun Qi, and Hailin Tang	737
Enterprise Credit Rating Prediction Model Based on Data Mining Algorithm	745

Contents	xiii
Application of Machine Learning Algorithm in Financial Market Risk Prediction	752
Credit Risk Assessment Model of Supply Chain Finance Under the Background of Big Data	759
Large Screen for 3D Data Visualization Based on RFG-SVM Algorithm  Guojuan Wang, Lijun Wang, Heqi Zhao, Qi Mu, and Xiangchun Ji	767
Construction of Inspection Work Management System Based on Data Mining Technology	775
Constructing and Evaluating Service Quality Evaluation Index System of Airline Using PSM and Entropy Weight  Liying Wang, Hongtao Li, Jing Bai, Xingming Lu, Pengpeng Zhou, and Zhijie Liu	784
Improving Service Quality Measurement with Servqual: Evidence from Airport Industry	792
Research on Visualization of Student Management System Based on Distributed Storage Technology	800
Banking Fintech Capability Evaluation System Based on Data Mining  Jie Fan	805
Technology Financial Service Platform Based on Data Mining Algorithm  Shuyu Hu	815
Overlapping Region Extraction Method of 3D Laser Point Cloud Registration on Account of Artificial Intelligence Algorithm	823
Smart Urban Carbon Emission Management Platform Based on Energy Big Data	831
Xue Li, Zheyong Piao, Yushan Zheng, Jieping Han, and Rijie Cong	501

A Distribution Network Reconfiguration Method Considering Optimal  Distribution of DG Output	841
Chao Feng, Yuanhong Hou, Shijun Li, Hui Xue, Yulai Yang, and Shunhai Xue	
A Study of Intersubjectivity in Translation Based on Communicative Behavior Theory and Computer Technology	848
Analysis and Research on Children's Cognitive Psychological System  Based on Digital Learning Terminal	855
Analysis of E-Commerce Live Broadcast Marketing Strategy Based on Big Data Algorithm	862
Analysis of Flow Data of E-Commerce Project Incubation Park Based on Computer Algorithm	870
Analysis of Teaching Management Data Based on Decision Tree  Algorithm  Dan Yu	878
Analysis on Reasonable Planning and Design of Port Waters and Channels  Wenting Liu and Zhiwei Liu	884
Ant Colony Algorithm Theory and Its Application in Control Engineering Chunxiang Huang, Nenjun Ben, and Guojun Yan	890
Application and Optimization of Style Transfer Algorithm in Contemporary Cultural and Creative Products  Jinxia Chen and Gui Wang	896
Application of Ant Colony Algorithm in Chinese Language and Literature Education	902
Application of BP Neural Network Algorithm in Teaching Evaluation and Learning Prediction	908

Contents	XV
Application of Computer Aided Technology in Kindergarten Interior  Design Under the "New Crown" Epidemic	914
Application of Data Mining Algorithm in Computer Mathematical  Modeling	922
Application of Hybrid Recommendation Algorithm in College Life Safety Education Management System Shuchao Liu and Yuanyuan Xu	928
Application of Improved Genetic Algorithm in Distribution Network Planning with Distributed Generation Yuanhong Hou, Chao Feng, Yougui Aer, Shunhai Xue, Yuqiang Zhou, and Hongchao Wang	937
Application of Mathematical Simulation Method in Phosphate Ore  Dressing  Lingpan Du and Yongjie Guo	944
Application of New Light and High Strength Materials in Prefabricated Buildings	951
Application of Virtual Simulation Technology in Microbial Image Segmentation in Sewage	958
Application of Virtual Simulation Technology in Traditional Handicraft Protection	964
Big Data Visual Accounting Model for Ecotourism Carrying Capacity  Yang Ye	971
Biomethanation System for Carbon Dioxide Capture in Renewable Energy Hydrogen Production Coupling Power Plant	979

Children's Safety in Smart Home Based on Intelligent Optimization

Algorithm .....

986

Changchun Xu

Wenxin Yang

### xvi Contents

Clustering Analysis of English Scores Based on K-means Clustering  Algorithm	993
Yueying Shen	
College Students' Learning Engagement Based on Engineering Thinking  Mingyuan Ma	1001
Comparison and Research of Roof Waterproof Materials Based on Computer Algorithm	1010
Comprehensive Evaluation System of Cultural Industry Development Based on Computer-Aided	1016
Correction to: Innovative Computing Vol 1 - Emerging Topics in Artificial Intelligence	C1
Author Index	1023

# The International Workshop on Technique for Language and Literature Information Modeling (LIM 2023)



# A Study on Data Mining for Type of Korean Painting Poetry

Haeyoung Park  $^1$   $^{\text{(b)}}$ , Younghoon An $^{2(\boxtimes)}$   $^{\text{(b)}}$ , and Hwayoung Jeong  $^{1(\boxtimes)}$   $^{\text{(b)}}$ 

Humanitas College, Kyung Hee University, 26, Kyungheedae-ro, Dongdaemun-gu, Seoul, Republic of Korea

{hy000p,hyjeong}@khu.ac.kr

Department of Korean Language and Literature, Kyung Hee University, 26, Kyungheedae-ro, Dongdaemun-gu, Seoul, Republic of Korea

yhnahn@khu.ac.kr

**Abstract.** This study introduced the work of analyzing the meaning by using the computer information processing method. We extracted the painting poetry of Mukjukdo (Bamboo Paintings) in the early Joseon Dynasty from the *Hanguk Munjip Chonggan* (Korean Literary Collections in Classical Chinese) of *Hanguk Gojeon Jonghap DB* (Korea Classics DB). Through the data mining method, we divided the types by extracting and cataloging the painting poetry from the Korean classical literature. Then, the scope was narrowed down to painting poetry of Mukjukdo in the early Joseon Dynasty, and the text was analyzed in units of syllable corpus.

**Keywords:** Data mining · Painting poetry · Bamboo paintings · Extracting · Cataloguing · Type Classification

### 1 Introduction

This study introduced the work of analyzing the meaning by using the computer information processing method. We extracted the painting poetry of Mukjukdo (Bamboo Paintings) in the early Joseon Dynasty from the *Hanguk Munjip Chonggan* (Korean Literary Collections in Classical Chinese) of *Hanguk Gojeon Jonghap DB* (Korea Classics DB). This is one of the basic tasks of research on Hansi (poems in the Chinese style), which has subdivided types of Hansi according to their material. This study presents expression techniques and meanings objectively by converting them into numerical values by using a computer information processing method. Traditional research on Hansi has mainly relied on researchers' intuition, however, this digital data analysis, which is a quantitative method, can complement the traditional method.

### 2 Research Methods

### 2.1 Data

The basic data for this study is *Hanguk Munjip Chonggan*; the Database of Hanguk Gojeon Beonyeokwon (Institute for the Translation of Korean Classics). It is called

<sup>©</sup> The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023 J. C. Hung et al. (Eds.): IC 2023, LNEE 1044, pp. 3–9, 2023. https://doi.org/10.1007/978-981-99-2092-1\_1

### H. Park et al.

"Hanguk Gojeon Jonghap DB". Hanguk Gojeon Beonyeokwon organizes and translates Korean classics into this database and then digitally converts them and discloses them to the public. This data is provided in XML format in the form of openAPI [1]. This data is categorized by author, book name, style, title, original text, year of publication, etc. Based on this, the painting data from the early Joseon Dynasty were extracted and analyzed.

### 2.2 Data Mining Techniques

Data mining is a technique of extracting useful information by analyzing statistical patterns, rules, and relationships in large amounts of data [2]. Today, it is used in various fields such as computerization, statistics, and management. We also intend to apply this technology to the study of Korean classical liteture. Currently, 142 anthologies from the early Joseon Dynasty are included in the *Hanguk Munjip Chonggan*. The primary data mining is to select the painting poetry based on this database and to classify them by their type. The secondary mining is to analyze the text by narrowing the scope [3] to painting poetry of Mukjukdo in the early Joseon Dynasty.

### 3 Result

### 3.1 Extracting and Cataloguing

We chose the poetry from the style category at Hanguk Gojeon Jonghap DB, and selected poems with characters in the title, such as 'do (picture)', 'hwa (painting)', 'je (mention)', 'muk (ink)', 'hoe (drawing)', 'sa (drawing)', 'byeong (folding screen), 'jok (hanging scroll)', 'cheop (album)', and 'chuk (scroll)'. After that, we compare the title and the content of each poem to determine whether the poem was painting poetry or not. Through this, it was found that there are a total of 842 poetry poems currently included in *Hanguk Munjip Chonggan*.

### 3.2 Type Classification

Based on the previously extracted lists, the painting poetry of the early Joseon Dynasty was classified by type based on the subject matter of the painting [4]. This classification is possible because the subject matter of the picture is presented as the name of each object before and after 'do' or 'hwa', which usually means a picture. For example, the titles of most painting poems of Mukjukdo appear in expressions such as "Mukjukdo" or "Jejuk (reciting a bamboo)". The figure below is the result of categorizing Mukjukdo painting poems using the TOPIC MAP based on the list of painting poems in the early Joseon Dynasty and the types of paintings (Fig. 1).

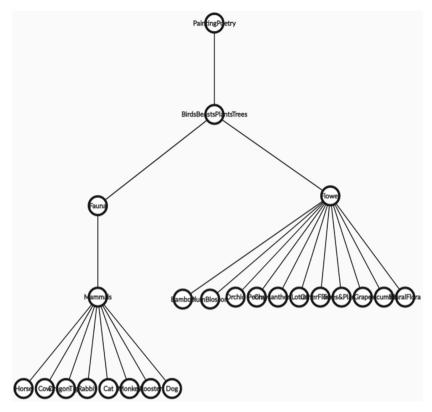


Fig. 1. A Type classification of Mukjukdo painting poetry using TOPIC MAP

### 3.3 Syllable Unit Segments and Statistics

Next, we analyzed the text data of the painting poetry of Mukjukdo in the early Joseon Dynasty. In order to interpret the poem [5], it is convenient to divide it into morphemes, the smallest unit of meaning. Since Hansi is composed of Chinese characters, it can be divided into morphemes up to one syllable unit. In addition, the units of these syllables are basically combined into five and seven words to form a row. Therefore, we set one row as the basic unit. Among the original text of painting poetry, o'eon (five-character) was divided from 1 to 5, and chil'eon (seven-character) was divided from 1 to 7 syllables. In succession, per line, o'eon generated five corpus of one syllable, two syllables, three syllables, and chil'eon generated seven corpus of one syllable, two syllables, three syllables, four syllables, five syllables, six syllables, and seven syllables (Table 1).

If 1 line of o'eon poem	If 1 line of chil'eon poem
is assumed to "1 2 3 4 5",	is assumed to "1 2 3 4 5 6 7",
the corpus unit	the corpus unit
1 syllable corpus(5): 1/2/3/4/5	1 syllable corpus (7): 1/2/3/4/5/6/7
2 syllable corpus(4): 12/23/34/45	2 syllable corpus (6): 12/23/34/45/56/67
3 syllable corpus(3): 123/234/345	3 syllable corpus (5): 123/234/345/456/567
4 syllable corpus(2): 1234/2345	4 syllable corpus (4): 1234/2345/3456/4567
5 syllable corpus(1): 12345	5 syllable corpus (3): 12345/23456/34567

Table 1. The corpus unit of o'eon and chil'eon poem

The following is the result of dividing the original data of Mukjukdo painting poetry by syllables and calculating the frequency. This is the figure measured using the JAVA program to see how many syllable corpus units are repeated in the actual painting poetry text data. The numerical value of 1–5 syllable corpus is the sum of the data for each syllable of o'eon and chil'eon, and the numerical values of 6–7 syllables is for chil'eon only. In these values, only words from 1 to 3 syllables were actually valid, and all 4–7 syllables were meaningless because the word combination was under twice (Figs. 2, 3 and 4).

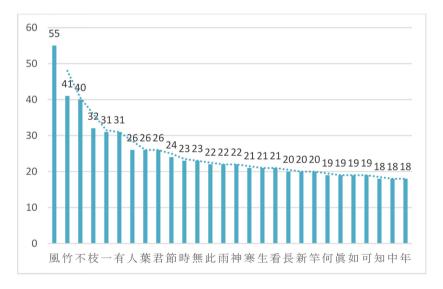


Fig. 2. The frequency of one syllable corpus in Mukjukdo painting poetry

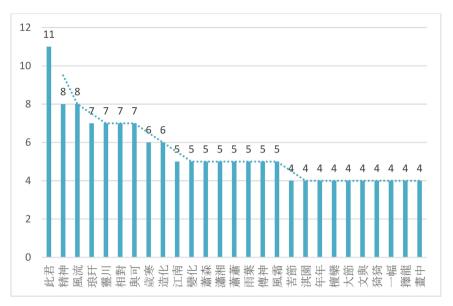


Fig. 3. The frequency of two syllable corpus in Mukjukdo painting poetry

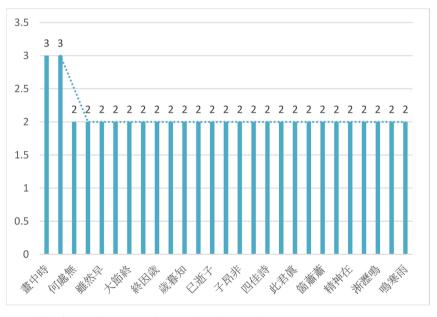


Fig. 4. The frequency of three syllable corpus in Mukjukdo painting poetry

### 3.4 Analysis of the Meaning of Corpus

The most commonly used one-syllable word in painting poetry of Mukjukdo is "poong (wind: 55 times)". "poong" is a combination of words in the order of "poong ryu (a taste for the arts: 8 times)", "poong sang (wind and frost: 5 times)", "Chun poong(the spring breeze: 3 times)" etc. in 2 syllables. Poong ryu is presented when referring to the wonderful scenery of bamboo. Pung sang is wind and frost, which gives bamboo trials and tribulations. Chun poong appears when describing a situation in which bamboo shoots sprout.

Next is "juk (a bamboo: 41 times)", but "juk" refers to bamboo itself, so it is not a discriminating result. And "bul (no: 40 times)", which belongs to an adverb that represents negativity in Chinese characters, is not very meaningful in single syllables, but further emphasizes its meaning when combined with other words. And "ji (Branch: 32 times)", which is connected in two syllables: "poong ji (a branch swaying in the wind: 2 times)", "so ji(a slender branch: 2 times). It excludes the numerical value that does not form a special meaning, such as juk, bul, ji.

The next significant figure is "goon (a man of virtue: 26 times)". "goon" is a combination of words in the order of "cha goon (this man of virtue: 11 times)", "goon jin (a man of virtue is truly ~: 3 times)", "goon dok (a man of virtue alone ~: 2 times)" etc.in 2 syllables. And "jeol" is 24 times, it means principles. "jeol" is a combination of words in the order of "go jeol(a distressed principles: 4 times)", "no jeol (mature principles: 2 times)", "jik jeol (a straight principles: 2 times)", etc., in 2 syllables. The jeol represents the principles of bamboo, and the attitude of keeping faith firmly is matched with "go" and "no", to reinforce the unchanging properties.

### 4 Conclusion

The data figures of corpus by syllable in Mukjukdo painting poetry are in line the symbolic meanings of bamboo in the early Joseon Dynasty. The symbolic meaning of Mukjukdo [6] is as follows. As an object, a bamboo was personified as the ideal existence of a god dragon and a Junzi. As an event, the virtual space of bamboo about the bamboo of Two queens and Qu Yuan was realized as the sorrowful and unworldly image.

Through this, the measurement of the frequency by dividing the original text of Mukjukdo painting poetry by syllables and setting it as a corpus composition unit is meaningful in revealing the symbolic meaning of bamboo in painting poetry. Conversely, the symbols of objects in various types of painting poetry can be demonstrated and shown with an objective indicator of language.

### References

- Lee, B., Min, K.: A Study on Visualization of the Analysis between the Collections of Korean Literature in Korea Classic DB: Hanguk Gojeon Beonyeokwon. Natl. Cult. 57, 5–32 (2021)
- ko.wikipedia.org/wiki/Data Mining. Accessed 1 Dec 2022
- Kim, D., Jeong, H.: A study of computational literature analysis based classification for a
  pairwise comparison by contents similarity in a section of Tokkijeon, 'Fish Tribe Conference':
  The Korea Contents Association. J. Korea Contents Assoc. 22, 15–25 (2022)

- 4. Haeyoung, P.: A study on the painting poetry of the former part of the Joseon Dynasty. Kyunghee University Graduate school, pp. 19–34 (2021)
- 5. Lee, B.: A Study on the Construction and utilization of Korean Chinese poetry corpus: geunyeoghanmunhaghoe. hanmunhaknonjib **53**, 153–177 (2019)
- 6. Haeyoung, P.: The symbol of the painting poetry about a bamboo painting in the former part of Joseon Dynasty: Society of Yol-Sang Academ. Yeol-sang J. Class. Stud. 73, 123-151 (2021)



# A Big Data Based Learning Model from Student Questionnaire

Hwa-Young Jeong<sup>(⊠)</sup>

Humanitas College, Kyung Hee University, 1 Hoegi-dong, Dongdaemun-gu, Seoul 130-701, Republic of Korea hyjeong@khu.ac.kr

**Abstract.** The results of analyzing students' requirements become very important data for teachers. Because it is a very important criterion for improving learning satisfaction while increasing students' learning effects. In this paper, students' requirements are analyzed and what subjects or contents students want are investigated. The purpose of this paper is to construct a learning model that reflects the needs of students.

Keywords: Big Data · Word Cloud · Student's Needs · Learning Model

### 1 Introduction

Until the late 21st century, general teaching methods were mainly based on frontal teaching in classrooms. The use of digital tools has the effect of increasing learning efficiency, but the method of education insisted on traditional classroom lectures [3]. After COVID-19 pandemic, university education was changed and operated from offline to online, and many students became familiar with online education [1]. Teleconferencing platforms such as Zoom have become hugely popular with the spread of online education due to COVID-19 pandemic. This has become a catalyst for many institutions to seek traditional education methods as various educational methods that can conduct synchronous or asynchronous remote education [3]. In online education, learners' activities and learning data were analyzed and processed using a LMS (Learning Management System). The large amount of learning data extracted from the LMS platform provided basic information for both teachers and students that could help improve learning satisfaction and educational goals [1]. In addition, students' learning satisfaction is generally obtained through a post-learning survey, and most of them analyze and use it when students answer multiple-choice questions written by teachers. However, the multiple-choice questions given by the teacher are not sufficient data for the student satisfaction survey, and the results of the student's data analysis may vary depending on the teacher's intention. For this purpose, it is necessary to analyze students' learning satisfaction and their needs through big data analysis when presented as a subjective question, not a multiple-choice question, and students freely describe their opinions and submit answers to the questionnaire.

NLP (Natural Language Processing) is natural human language and communication variables such as voice, text, audio, and video, interpretations and applications [2] We call NLP as non-formal data. Text mining that is one of technique in Big data is well known a process to extract meaning words from the data and analyze them.

This paper aims to make a learning model using analysis of students' needs. For this process, students get a subjective questionnaire and they write their opinion freely without any form and select items. Their answers is used to analyze what they want to study or learn. Consequently, this research shows a frequency of their opinions that can be the students' needs from their study.

### 2 Related Works

### 2.1 Big Data

Big data is a technology that can handle unstructured data, unlike databases that only dealt with structured data. Of course, this can perform both processing of collecting and analyzing unstructured and structured data.

It also guarantees both volatility, speed, volume, diversity and integrity. In big data, data is collected in various ways, such as a web browser and a mobile web with various data formats. In the existing analysis method, different formats of unstructured and structured data could not be managed, but both big data are possible. In big data, Hadoop is cost-efficient, scalable, and enables fast and flexible parallelism. It also uses the Hadoop framework for big data analysis because it provides availability, resilient properties, security, and authentication. It is also well known for its open-source software architecture, which includes processing and storage. The part to be stored is HDFS (Hadoop Distributed File System), and the part to be processed is MapReduce. D.K. Jain et al. [4] depicted the Hadoop architecture as shown in Fig. 1.

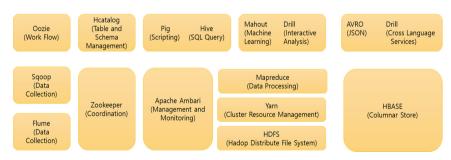


Fig. 1. Hadoop architecture by D.K. Jain et al.

Hadoop, first run by Doug Cutting and Mike Caparella in 2005, is a Java-based open-source software. Providing a distributed framework for processing and managing big data is known as Hadoop's advantage. Therefore, using Hadoop can manipulate a large amount of data. In addition, the Hadoop Distributed File System (HDFS) can store MapReduce for the purpose of the process. MapReduce is used to analyze and generate

a wide range of datasets using distributed and parallel processing methods in clusters [4].

Big data is being used in various fields. In particular, techniques for analyzing the frequency of words in unstructured data to show importance and visualize them are widely known. Figure 2 shows an example of the results analyzed by word cloud of big data for CC (Cognitive Computing) [5].



Fig. 2. A sample visualization by analyze word cloud of Big Data.

### 2.2 Learning Model

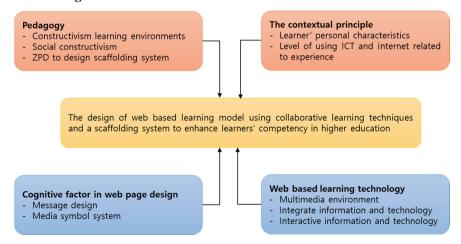


Fig. 3. The web based learning model by K. Deejring

The e-learning system promotes learning as teachers and students interact in a non-face-to-face manner, and learning materials and evaluations are also conducted online.

This approach can be a good solution to accommodating an exponentially growing number of students and their curriculum, but only popular large-scale online curriculums can cause problems such as student crowding, student performance, failure and departure [6]. K. Deejring depicted e-learning model as shown in Fig. 3 [7].

### 3 Data Analysis for Students' Needs

### 3.1 Survey of Students' Requirements for Their Class

For this study, a survey was conducted on 40 general students who did not major in IT in the liberal arts subject of K University in Seoul. The survey items are as follows.

**Question:** Feel free to describe any topic or content you want to hear in the lecture. Or if there is a subject that you want to make a lecture in general culture, please write it with the reason.

The comments as below represent one of the student's answers among the results of the survey.

Answer: In my opinion, the topic we need in the Fourth Industrial Revolution is 'human ennui among the negative aspects of liberal utopia'. In the 4th Industrial Revolution, as each of us could access information or technology so easily, it destroyed the feudal system of the past and transformed it into a nation by the public, giving everyone a chance to politics by democracy and capitalism. As educational opportunities diversified, it began to turn into a world where everyone creates opportunities through education. But we must think about whether we will be happy. I get everything I want and come across, but I feel despondent. The reason is that in the era of the Fourth Industrial Revolution, we have it whenever we want, so we work hard to achieve something and do not feel the joy, satisfaction, and fulfillment that comes from it. Therefore, we maximize convenience and efficiency, but we may not be able to enjoy our mental satisfaction and happiness.

### 3.2 Frequency from the Survey

Students' survey results are unstructured data written freely. If you analyze the frequency of words using the word cloud of big data, it is shown as shown in the Fig. 4. The results of this analysis were all processed students' opinions, and they were not able to extract actually important words from the students' answers, but only the frequency was processed from the entire content.

Figure 5 shows the results except for meaningless words. What these results show is that students want subjects related to technology and are paying a lot of attention to preparing for the future. In particular, if you look at topics such as education, development, paradigm, artificial, computer, big data, metabus, and program, it can be seen that students want lectures related to this.

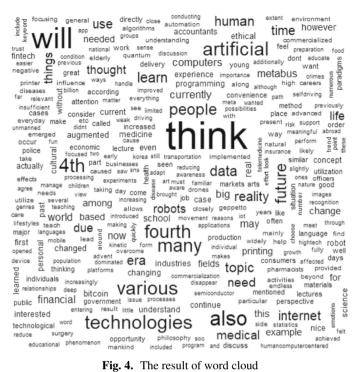


Fig. 4. The result of word cloud

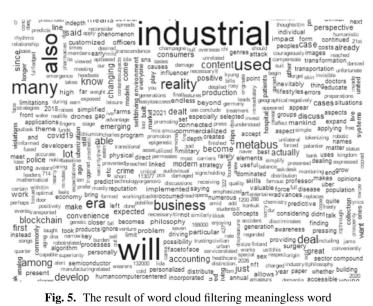


Fig. 5. The result of word cloud filtering meaningless word