Kristina P. Sinaga

Address
No. 17A Naga Terbang St.
P. Siantar City
North Sumatera Province
21132, Indonesia

Contact
kristinasinaga41@gmail.com
https://kristinap09.github.io
ORCID 0009-0000-6184-829X
Scopus Author ID 57202362864

SUMMARY

I holds a Ph.D. in applied mathematics from Chung Yuan Christian University, Taiwan. I finished my PhD supervised by Prof. Miin-Shen Yang, working on multi-view clustering to address multi-view problems with (out) feature reduction in a collaborative manner. Major lines of my research, professional work and expertise are in clustering and pattern recognition, especially within areas of single, multi-view, and multiple users learning. I have a proven publication track record of success in these areas and am proficient in designing and analyzing algorithms for mathematical optimization of unsupervised machine learning such as constructing/modifying objective functions for complex design (i.e. with curse of dimensionality problems, finding optimal number of clusters, etc.) in non-federated and federated environments. To test my proposed algorithms on a large dataset, mostly I used programming tools like Matlab and Python. During the past four years I have been a lecturer specialist - S3 in the information systems management department at BINUS university, Indonesia (2020-2022) and a postdoc in the Department of applied mathematics at Chung Yuan Christian University, Taiwan (2023-2024).

PROFESSIONAL EXPERIENCE

Independent Researcher
Self-employed **Mar. 2024 – Present**

- I am working in the development of innovative algorithms capable of extracting meaningful insights from complex data without the need for labeled data. The main goal is to continue advancing state-of-the-art in unsupervised learning algorithms for pattern recognition tasks, with a focus on scalability, interpretability, and real-world applicability using publicly available datasets.

Post-doctorate Fellow
Department of Applied Mathematics, CYCU, Taiwan **Mar. 2023 – Mar. 2024**

- Primarily works in an office environment and working from home using personal PC for much of the day.
- Work independently and weekly meeting with PI to discuss a new idea or new accomplishment related to the research works.
- Proposed a new objective function of soft and hard clustering to address multiple resources, clients or users data.
- Designed new algorithms of multi-view k-means (MVKM) and multi-view fuzzy c-means (MVFCM) in non-federated and federated environments.
- Provided (Matlab) codes for the problems of multiple resources and multiple clients or users data.
- Conducted experiment/simulation on different publicly available multi-view data sets and interpret the results.
- Wrote academic papers that implemented a soft or hard clustering algorithm to assure efficiency, repeatability, and standardization in the use of multiple-resources data over multiple clients or users.
- Served as a reviewer of IEEE Access.

Lecturer Specialist - S3
Information Systems Management Department, **Nov. 2020 – Mar. 2022**
BINUS Graduate Program, Indonesia

- As a lecturer my role, is to teach, do research, serve communities in Indonesian society, provide and grading students examinations/their other assessment items. I also got the opportunity of contributing to and developing the course content, homeworks, exams and programming assignments.

Staff (Badge: Supervisor)
BINUS University, Indonesia **Nov. 2020 – Mar. 2022**

- As a staff my role, is to advance the university by participating and contributing in any tasks related to the binus graduate program of master of management of system information.

EDUCATION

Doctor of Philosophy, Applied Mathematics
Chung Yuan Christian University (CYCU), Taiwan 2020
Thesis title: *Multi-view fuzzy clustering algorithms for multi-view data*
Thesis' PPT: *Click here*
CGPA: 3.842 out of 4.000

Master of Science, Mathematics in Operation Research
University of Sumatera Utara (USU), Indonesia 2015
CGPA: 3.78 out of 4.000

Bachelor of Science, Mathematics in Statistics
University of Sumatera Utara (USU), Indonesia 2013
CGPA: 3.30 out of 4.000

RESEARCH SUMMARY

Research Interests

- Clustering: I work on developing k-means and fuzzy c-means (FCM) algorithms for addressing single and multi-view data. I occasionally build a new developed clustering algorithm based on the new objectives of mathematics formulation. Prior to that, I also provided and publicly shared the codes of my proposed algorithms on my GitHub page. Most recently, I am leveraging my research interests into graph clustering, manifold regularizations, and kernel-based approaches to separate data points into different clusters.
- Pattern Recognition: I work on clustering-based algorithms such as k-means and FCM for dimensionality reduction. I facilitate the principal analysis of un-wanted, less likely, and relevant features phenomena on single and multi-view data. Specifically, I create a collaborative approach to select informative features with single/multi-view features representation and unsupervised learning. In such a way, the implementation of this feature selection-based dimension reduction technique can effectively provide the optimal number of clusters k but also significantly improve the accuracies.
- Federated Learning: Currently, I work on federated learning (FL), developed and designed conventional multi-view clustering algorithms into parallel algorithms for mathematical optimization of recognizing data pattern from multiple clients' multi-view data. Unlike my previous works on non-federated unsupervised machine learning techniques, in this topic, I organized, brought creative/innovative perspectives by proposing new algorithmic approaches to address multiple clients' multi-view data with privacy and effective communication concerns.

Research Activity

Total refereed papers:	7
Total books / book chapters:	0 / 2
Journals reviewed for:	0
Conference / workshop chairs:	4 / 0

PUBLICATIONS

According to Google Scholar Citations, my h-index is **5** and I have over **1,600** citations.

My top cited publications (with over 100 citations) are: "Unsupervised k-means clustering algorithm" (with over 1,212 citations); "A feature-reduction multi-view k-means clustering algorithm" (cited by 112). My top second cited publications (with over 20 citations) are: "Collaborative feature-weighted multi-view fuzzy c-means clustering" (cited by 39); "Entropy k-means clustering with feature reduction under unknown number of clusters" (cited by 31). My top third cited publications are: "Poverty data modeling in North Sumatera Province using geographically weighted regression (GWR) method (cited by 7)"; "Modified relational mountain clustering method (cited by 3)", "Spatial variation in infant mortality with geographically weighted poisson regression (GWPR) approach (cited by 3)"; "Machine learning approaches for marketing campaign in Portuguese banks (cited by 2)"; "Unsupervised multi-view fuzzy c-means clustering algorithm (cited by 1)".

Journal Paper

Hussain, Ishtiaq, **Sinaga, Kristina P**, and Yang, Miin-Shen (2023). Unsupervised multi-view fuzzy c-means clustering algorithm. *Electronics*, *12*, 4467. (link)

Yang, Miin-Shen and **Sinaga, Kristina P** (2021). Collaborative feature-weighted multi-view fuzzy c-means clustering. *Pattern Recognition*, *119*, 108064. (link, data and materials)

Sinaga, Kristina P, Hussain, Ishtiaq, and Yang, Miin-Shen (2021). Entropy k-means clustering with feature reduction under unknown number of clusters. *IEEE Access*, *9*, 67736–67751. (link, pdf, data and materials)

Sinaga, Kristina P and and Yang, Miin-Shen (2020). Unsupervised k-means clustering algorithm. *IEEE Access*, *8*, 80716–80727. (link, pdf, data and materials)

Yang, Miin-Shen and **Sinaga, Kristina P** (2019). A feature-reduction multi-view k-means clustering algorithm. *IEEE Access*, *7*, 114472–114486. (link, pdf, data and materials)

Sinaga, Kristina P and Hutahaean, Manuntun and Gea, Petrus (2016). Spatial Variation in Infant Mortality with Geographically Weighted Poisson Regression (GWPR) Approach. *International Journal of Science and Research*, *5*(3), 96–100.

Sinaga, Kristina P(2015). Poverty Data Modeling in North Sumatera Province Using Geographically Weighted Regression (GWR) Method. *International Journal of Science and Research*, *4*(2), 1738–1742.

Preprint Paper

Sinaga, Kristina P (2024). Rectified Gaussian kernel multi-view k-means clustering. *arXiv preprint arXiv:2405.05619*. (link, pdf, data and materials)

Book Chapter

Sinaga, Kristina P, Benjamin, J.B.M., and Yang, Miin-Shen (2018). Modified relational mountain clustering method. *Artificial Intelligence and Soft Computing: 17th International Conference, ICAISC 2018*, Zakopane, Poland, June 3-7, *Part I 17*, 690–701.

D. Yuniati and **Sinaga, Kristina P**(2021). Analytics-based on classification and clustering methods for local community empowerment in Indonesia. *(eds) Soft Computing in Data Science, SCDS 2021*, *Communication in Computer and Information Science*, *vol. 1489*, Springer, Singapore.

Conference Paper

A. Jennifer and **Sinaga, Kristina P**(2021). Machine learning approaches for marketing campaign in Portuguese banks. *2021 3rd International Conference on Cybernetics and Intelligent System (ICORIS)*, Makasar, Indonesia, 1–6.

W. Henwy and **Sinaga, Kristina P**(2021). Telecommunication analytics based on customer segmentation using unsupervised algorithms. *2021 3rd International Conference on Cybernetics and Intelligent System (ICORIS)*, Makasar, Indonesia, 1–6.

Under Review

Yang, Miin-Shen and **Sinaga, Kristina P** (2024). Federated multi-view k-means clustering. ***IEEE TPAMI***

In Manuscript

Yang, Miin-Shen and **Sinaga, Kristina P** (2024). Federated weighted multi-view fuzzy c-means.

Sinaga, Kristina P and Yang, Miin-Shen (2024). A globally collaborative multi-view k-means clustering.

Yang, Miin-Shen, Josephine. B.M. Benjamin, **Sinaga, Kristina P** (2024). A survey of soft clustering.

Sinaga, Kristina P (2024). Personalized federated learning under collaborative multi-view k-means clustering.

Sinaga, Kristina P (2024). Tensor k-means clustering algorithm.

Sinaga, Kristina P (). etc...

HONORS & AWARDS

Honorary Member
The Phi Tau Phi Scholastic Honor Society of The Republic of China, CYCU, Taiwan 2020

Recipient
Japan Science and Technology Agency (JST), Niigata University, Japan 2018

Recipient
Japan Student Service Organization (JASSO), Niigata University, Japan 2017

Recipient
CYCU International Student Scholarship, CYCU, Taiwan 2016

PROFESSIONAL ACTIVITIES

Journals Reviews

- Information Fusion, Elsevier (2022 – 2023).
- IEEE Access (2021 –2023)
- Applied Soft Computing, Elsevier (2022).
- IEEE TKDE (2022)

Conference Reviews

- IJCNN2023.
- WCCI2022.

TEACHING EXPERIENCE

In my modules, I have taught over 80 undergraduate students, and over 100 graduate students (regular and online programs). In total, I have taught over 180 unique students. I moderated some events such as guest lecturer events (participated by undergraduate, graduate, doctorate students, lecturer, etc. from different university in-and-abroad). In total, I have moderated 4 events with uniques attendances from Indonesia and abroad (2020 - 2021).

Masters in Information Systems Management

- Regular and online program of Business Intelligence and Analytics (2020 - 2022)

Bachelors in Computer Science

- Calculus I (2021)
- Discrete Mathematics (2021 - 2022)

PROFESSIONAL MEMBERSHIPS

Member, The Institute of Electrical and Electronics Engineers (IEEE) [2020 – 2021]
Member, IEEE CIS [2020 – 2021]
Member, IEEE SPS [2020 – 2021]
Scientific Committee Member, World Academy of Science, Engineering and Technology (WASET), category of Mathematical and Computational Sciences [2020 – 2021]

CERTIFICATION

The Data Scientist's Toolbox
Johns Hopkins University — Coursera Nov 30, 2022
Instructor: Jeff Leek, PhD, Roger Peng, PhD, and Brian Caffo, PhD

Python Project for Data Science
IBM — Coursera Nov 25, 2022
Instructor: Azim Hirjani & Joseph Santarcangelo

Python for Data Science, AI & Development
IBM — Coursera Nov 24, 2022
Instructor: Joseph Santarcangelo

Tools for Data Science
IBM — Coursera Nov 22, 2022
Instructor: Aije Egwaikhide, Svetlana Levitan, and Romeo Kienzler

Deep Learning.AI TensorFlow Developer
DeepLearning.AI — Coursera Nov 15, 2022
Instructor: Laurence Moroney

Neural Networks and Deep Learning
DeepLearning.AI — Coursera Nov 5, 2022
Instructor: Andrew Ng, Kian Katanforoosh, and Younes Bensouda Mourri

Machine Learning Specialization
DeepLearning.AI — Stanford University — Coursera Oct 26, 2022
Instructor: Andrew Ng

Advanced Learning Algorithms
DeepLearning.AI — Stanford University — Coursera Oct 26, 2022
Instructor: Andrew Ng

Understanding and Visualizing Data with Python
University of Michigan — Coursera Oct 18, 2022
Instructor: Brenda Gunderson, Ph.D., Kerby Shedden, Ph.D., and Brady T. West, Ph.D.

Programming for Everybody (Getting started with Python)
University of Michigan — Coursera Oct 9, 2022
Instructor: Charles Russell Severance

Learning to Teach Online
University of New South Wales (UNSW) — Coursera 2020
Instructor: Assoc. Prof. Simon McIntyre & Dr Negin Mirriah

REFERENCES

Referees are available on request.