# 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 hold a Ph.D. in Applied Mathematics from Chung Yuan Christian University, Taiwan, where my research focused on multi-view clustering under the supervision of Prof. Miin-Shen Yang, addressing complex multi-view problems with (out) feature reduction through collaborative methods. My expertise lies in clustering and pattern recognition, particularly in single-view, multi-view, and multi-user learning domains, supported by a strong publication record. I specialize in designing and analyzing algorithms for unsupervised machine learning optimization, including modifying objective functions to handle challenges like the curse of dimensionality in both non-federated and federated environments. I utilize programming tools such as Matlab and Python for testing proposed algorithms on large datasets. Previously, I served as a Lecturer Specialist - S3 at BINUS University, Indonesia (2020-2022), and as a Postdoctoral Researcher at Chung Yuan Christian University, Taiwan (2023-2024).

## PROFESSIONAL EXPERIENCE

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

- I spearhead the development of pioneering algorithms engineered to extract actionable insights from intricate datasets devoid of labeled information.
- To propel the frontier of unsupervised learning methodologies, with a specific emphasis on enhancing scalability, interpretability, and real-world applicability for pattern recognition tasks.

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

- Primarily operates within an office environment, with the flexibility to work remotely using personal PC equipment for the majority of tasks.
- Functions autonomously, collaborating with the Principal Investigator (PI) in weekly meetings to discuss new research ideas and achievements.
- Spearheaded the development of a novel objective function for soft and hard clustering, tailored to address multi-resource, client, or user data challenges.
- Conceptualized and implemented pioneering algorithms, including Multi-View K-Means (MVKM) and Multi-View Fuzzy C-Means (MVFCM), across non-federated and federated environments.
- Delivered MATLAB codes to facilitate the resolution of complex data challenges involving multiple resources and clients/users.
- Executed experiments and simulations using various publicly available multi-view datasets, meticulously interpreting results to drive insights.
- Authored academic papers showcasing the application of soft/hard clustering algorithms, ensuring efficiency, reproducibility, and standardization in multi-resource data utilization across diverse client/user landscapes.

- Contributed expertise as a peer reviewer, ensuring the rigor and quality of research within the academic community.

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

- In my capacity as a lecturer, I fulfill a multifaceted role encompassing teaching, research, community engagement, and academic assessment.
- Delivering lectures, conducting research endeavors, and actively engaging with Indonesian society.
- I am responsible for curating and refining course content, crafting assignments and examinations, and meticulously grading student assessments.
- I am privileged to contribute to the development and enhancement of curriculum materials, including homework assignments, exams, and programming tasks, thereby enriching the educational experience for students.
- Learned a valuable lesson.

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

- My primary objective is to propel the university forward by actively participating in and contributing to various initiatives within the graduate program of master of management of system information at Binus University.
- Advanced the university program's goals and objectives.
- Learned a valuable lesson.

## 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 specialize in developing k-means and fuzzy c-means (FCM) algorithms for handling single and multi-view data. Occasionally, I devise new clustering algorithms based on novel mathematical formulations. Additionally, I have made my proposed algorithm codes publicly available on my GitHub page. Currently, I am extending my research interests to include graph clustering, manifold regularization, and kernel-based approaches to effectively partition data points into distinct clusters.
- **Pattern Recognition:** My research in pattern recognition focuses on clustering-based algorithms such as k-means and FCM for dimensionality reduction. I employ principal analysis to identify irrelevant

features and emphasize informative feature selection in both single and multi-view data settings. By employing a collaborative approach to feature selection with single/multi-view feature representations and unsupervised learning, I achieve significant improvements in accuracy and efficiently determine the optimal number of clusters (k).

- **Federated Learning:** My current research centers on federated learning (FL), where I adapt conventional multi-view clustering algorithms into parallel algorithms for optimizing the recognition of data patterns from multiple clients' multi-view data. In contrast to my previous work in non-federated unsupervised machine learning techniques, I introduce innovative perspectives by proposing new algorithmic approaches to address privacy and effective communication concerns inherent in handling multiple clients' multi-view data.

## Research Activity

| | |
|---|---|
| Total refereed papers: | 7 |
| Preprint papers: | 1 |
| Total books / book chapters: | 0 / 2 |
| Journals reviewed for: | 1 |
| Conference / workshop chairs: | 1 / 0 |
| In manuscript: | >5 |

## PUBLICATIONS

According to **Google Scholar Citations**, my **h-index** is **5**, with over **1,600 citations**. Among my most highly cited publications (each with over 100 citations) are: "Unsupervised k-means clustering algorithm" (cited over 1,400 times) and "A feature-reduction multi-view k-means clustering algorithm" (cited 120 times). Additionally, my second most cited publications (each with over 20 citations) include: "Collaborative feature-weighted multi-view fuzzy c-means clustering" (cited 45 times) and "Entropy k-means clustering with feature reduction under an unknown number of clusters" (cited 34 times). Notably, my third most cited publications cover diverse topics such as poverty data modeling in North Sumatera Province using the geographically weighted regression (GWR) method (cited 7 times), an unsupervised multi-view fuzzy c-means clustering algorithm (cited 4 times), machine learning approaches for marketing campaigns in Portuguese banks (cited 4 times), modified relational mountain clustering method (cited 3 times), and spatial variation in infant mortality with geographically weighted Poisson regression (GWPR) approach (cited 3 times).

### 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 Chapters

**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.

## Under Review

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

## In Manuscripts

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 instructed over 80 undergraduate students and more than 100 graduate students across regular and online programs. In total, I have taught over 180 unique students. Additionally, I have moderated several events, including guest lecturer sessions attended by undergraduate, graduate, and doctoral students, as well as lecturers from various universities both domestically and internationally. Overall, I have moderated four events with unique attendees from Indonesia and abroad spanning the years 2020 to 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
Direct Access Certificate

*Python for Data Science, AI & Development*
IBM — Coursera Nov 24, 2022
Direct Access Certificate

*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
Direct Access Certificate

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

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

*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.
Direct Access Certificate

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

*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.