

CYRIL WEERASOORIYA

cyrilw.com — me@cyrilw.com — Rochester, NY

Education

Ph.D., in Computer Science (GPA 3.9) at *Rochester Institute of Technology, NY*. Present
Related Courses in Data Science, Machine Learning, Deep Learning, and Quantitative Foundations

Bachelor of Science (Honors), in Computer Science at *University of Kelaniya, Sri Lanka* March 2017

Recent Research Experience

Population level label distribution learning for human disagreement prediction August 2018 - Present
Rochester Institute of Technology, NY

Tools used - Python (Scikit-learn, Keras, and TensorFlow), MongoDB for data storage, and Google Cloud.

- Predicting human disagreements for supervised machine learning algorithms using unsupervised methods.
- Working with clustering methods (K-Means, LDA, and GMM), and supervised learning methods (CNN and LSTM).
- Research work accepted at ECAI 2020* with new methods;
 - Semi-supervised (neighborhood based pooling) learning method for label pooling.
 - Statistical measures for the hyper-parameter selection for clustering and label pooling methods.
- Presently working on predicting human reactions from language based Facebook posts.

LKEMR - Cloud-based Electronic Medical Record System, Sri Lanka March 2018 - Present
Collaborative project with the Faculty of Medicine, and Colombo North Teaching Hospital, Sri Lanka. Tools used - PHP, MySQL, HTML, CSS, and UNIX. Python for data mining. Hosted on DigitalOcean and AWS.

- Web based electronic patient record management system, tailored for Sri Lanka using PHP backend, and MySQL database.
- EMR currently holds records for over 1,000 patients across several cities in Sri Lanka.
- Currently been used to identify potential patients with COVID-19, as it is the only EMR widely used in Sri Lanka.
- The work also contributes to open sourced EMR Project, Open-EMR.
- Research work published in NITC 2019 and workshop organizing committee for WONCA 2020.

Food Data Aggregation - Project to aggregate nutritional information for crowdsourced recipes from FDA Summer 2019
HEALS Project (part of IBM AI Horizon) at *Rensselaer Polytechnic Institute, NY*.

Tools used - Python for data scrapping and SpaCy for NLP aggregation.

- Research based on natural language processing and information retrieval.
- Conducted research on aggregating food related data sources for food knowledge graph, which is used with IBM Watson.
- The tool was able to generate the nutritional content per FDA guidelines from a crowdsourced recipe.
- Evaluated the method with the 1M Recipe (1 million) crowdsourced dataset.

KeyXtract:Twitter Model - Keyword Extraction Method using syntactic NLP rules. March 2016 - March 2017
University of Kelaniya, Sri Lanka

Tools used - Stanford CoreNLP (Java) & NLTK (Python) for the keyword extraction & used Python for Data Mining.

- Collaborated with a domain expert from Department of English on gathering & building a model for analyzing Tweets.
- Evaluated measures using standard measures (F1 Score) and human based evaluation (A/B Test).
- Published research on Top-Tier Computer Science conferences in Sri Lanka*,
 - NITC 2016 (National Information Technology Conference). Only 16 papers selected for publication annually.
 - ICTer 2016-2017 (Information Communication Technology for Emerging Regions).

Recent Leadership Experience

Vice-President of the **Doctoral Students Association**, in *Rochester Institute of Technology, NY*. Present
• Member of the Graduate Student Communication COVID-19 Task Force

| Skills | Awards and Certifications | Research Interests |
|--|--|---|
| Data Exploration: Tableau, Jupyter Notebooks (Colab), and Excel | Competent Gavelier Leader , awarded by <i>Toastmasters International USA</i> 2017 | Machine learning, Label distribution learning, Deep learning, Data science, Algorithms, Statistics, Probabilistic learning, Human computation, Crowdsourcing, Information retrieval, Natural language processing, Sociolinguistics, Computational chemistry, and Analytical chemistry |
| Programming: Python, R, Java, C, SQL, and PHP | Competent Gavelier Communicator , awarded by <i>Toastmasters International USA</i> 2017 | |
| Scientific Programming: Matlab, Keras, TensorFlow, Genism, and scikit-learn | Social and Behavioral Research Certificate , by <i>Citi Program</i> 2018 | |
| Cloud Management: UNIX, Google Cloud, AWS, and DigitalOcean. | | |

*Publications available overleaf. *Résumé* as of February 26, 2021

Selected Recent Publications

h-index = 2, Citations (February 2, 2021) = 23

Weerasooriya, T. C., Liu, T., & Homan, C.H. (2020). *Neighborhood-based Pooling for Population-level Label Distribution Learning*. 24th European Conference on Artificial Intelligence (ECAI 2020).Mendis, K., **Weerasooriya, T. C.**, Withana, S., Liyanage, P., Silva, A. W., Wickramasinghe, R., & Weerabaddna, C. (2019). *Cloud-Based Open Source Primary Care Electronic Patient Record System for Sri Lankan Citizens*. 2019 National Information Technology Conference (NITC), pp. 41–46.**Weerasooriya, T.**, Perera, N., & Liyanage, S. R. (2017). *A Framework for Automated Corpus Compilation for KeyXtract : Twitter Model*. 2017 International Conference on Advances in ICT for Emerging Regions (ICTer), pp. 43–48.**Weerasooriya, T.**, Perera, N. and Liyanage, S.R. (2016). *A Method to Extract Essential Keywords from a Tweet Using NLP Tools*. 2016 International Conference in Advances in ICT for Emerging Regions (ICTer). pp. 29-34.Perera, N., & **Weerasooriya, T.** (2016). *The Racecourse of Then and Now: Evolution of the Sri Lankan English Vocabulary Over Two Generations of SLE Speakers*. VISTAS Journal, 10, pp. 1–23..**Weerasooriya, T.** (2016). *Automated Railway Ticketing System: Replacing the paper based ticket with the Electronic National Identity*. ICT for Development Working Paper Series, 6(1–2), pp. 67–77.**Recent Work Experience**

Tech Lead at *LKEMR (Cloud-based Electronic Medical Record System)* March 2018 - Present

- Leading a team for building a web based electronic patient record management system, tailored for Sri Lanka.
- Created the pipeline to manage an instance based EMR implementations for clinics on AWS using CI/CD.
- EMR currently holds records for over 1,000 patients across several cities in Sri Lanka and expanding nationwide.

Adjunct at *Rochester Institute of Technology, NY*. Summer 2020

- Taught CSCI 635 - Introduction to Machine Learning.

Research Assistant for HEALS Project (part of IBM AI Horizon) at *Rensselaer Polytechnic Institute, NY*. Summer 2019

- Conducted research on NLP methods to aggregating food related data sources for food knowledge graph.
- Worked on a way to extract information from ADA guidelines based on the queries using doc2vec.

Intern at *Research and Development Team, Lesova Holdings Private Limited, Sri Lanka* February 2016 - December 2016

- Implemented and maintained customer relations management (CRM) System using Zendesk and Freshdesk.
- Designed and Developed Official Website for T-Shirts.LK Maldives Website.

Teaching Assistant at *Department of Computer Science, University of Kelaniya, Sri Lanka* March 2017 - August 2018

- Assisted in teaching Programming in C, Object Oriented Programming and Data Structures and Algorithms for approximately 150 students.

Research Assistant (Analytical Chemistry) at *University of Peradeniya, Sri Lanka* December 2014 - March 2015Science and Technology Research Partnership for Sustainable Development Laboratory at *University of Peradeniya, Sri Lanka*

- Experienced in the analysis of TOC, BOD and COD in water samples collected from different lakes around Kandy.

Leadership Experience

HackLN 2017, Hackathon organized by the *Department of Computer Science, University of Kelaniya, Sri Lanka*. 2017

- Guided the students as the teaching assistant in charge for the Department for Vidya 50, the 50th Celebratory Exhibition of Faculty of Science.

Gavel Club affiliated with *Toastmasters International, USA*. November 2014 - November 2017

- Organizing Committee Member of Inter University Best Speaker Contest 2014 - 2017.
- Designed, developed and maintained the official Club Website, Contest Website and the online account management system for the executive committee.

Career Fair organized by the *Career Guidance Unit, University of Kelaniya, Sri Lanka*. January 2015 - December 2016

- Designed, developed and maintained the event website and coordinated the registrations for 600+ undergraduates annually.