Jose Lugo-Martinez, Ph.D.

Assistant Professor
Computational Biology Department
School of Computer Science
Carnegie Mellon University

5000 Forbes Avenue
Pittsburgh, PA 15217
Email: jlugomar@cs.cmu.edu
Phone: +1 (412) 268-8693
Web: https://jlugomar.github.io

EDUCATION

Ph.D. in Computer Science — August 2009 – December 2016
Department of Computer Science
Indiana University, Bloomington, IN

- *Minor:* Bioinformatics
- *Thesis:* Flexible kernel functions for learning on graphs and hypergraphs
- *Advisor:* Predrag Radivojac
- *Thesis Committee:* Esfandiar Haghverdi, Matthew W. Hahn, Haixu Tang, Yuzhen Ye

M.S. in Computer Science — August 2007 – January 2010
Department of Computer Science and Engineering
University of California, San Diego, CA

- *Concentration:* Computer Architecture and Compilers
- *Thesis:* Strategies for sharing the floating-point unit between specialized processing elements

Dual B.S. in Computer Science and Mathematics — August 1999 – May 2003
Department of Mathematics and Computer Science
University of Puerto Rico, Rio Piedras, PR

- *Magna cum laude*

RESEARCH INTERESTS

Computational biology, Automated science, Machine learning, Biomedical data science

PROFESSIONAL EXPERIENCE

Assistant Professor — June 2022 – present
Computational Biology
School of Computer Science
Carnegie Mellon University, Pittsburgh, PA

Assistant Professor — August 2021 – May 2022
Department of Computer Science
University of Puerto Rico, San Juan, PR

Lane Fellow September 2017 – July 2021
Computational Biology Department
School of Computer Science
Carnegie Mellon University, Pittsburgh, PA

- Development of computational methods for integration and analysis of longitudinal multi-omics microbiome data.
- Development of computational approaches for integration of heterogeneous data towards prediction of novel protein complexes.

Post-Doctoral Fellow January 2017 – August 2017
Precision Health Initiative
Department of Computer Science
Indiana University, Bloomington, IN

- Development of computational approaches towards understanding protein function and how disruption of protein functions leads to disease.

Research Assistant August 2009 – December 2016
Department of Computer Science
Indiana University, Bloomington, IN

- Development of robust graph-based kernel methods for learning on graphs and hypergraphs
- Development of computational models for understanding and predicting molecular mechanisms of disease upon mutation

Research Intern May – August 2014
SystemML Research Group
IBM Research, Almaden, CA

- Design and implementation of a basic debugger for SystemML. SystemML enables declarative machine learning on Big Data in a MapReduce environment. **Debugger prototype approved for production.**

Research Assistant September 2007 – July 2009
Department of Computer Science and Engineering
University of California, San Diego, CA

- Development of a fully automated toolchain that generates specialized circuits (c-cores) from source code extracted from applications.

Project and Services Coordinator August 2006 – July 2007
High Performance Computing Facility
University of Puerto Rico, Rio Piedras, PR

Co-founder and Chief Executive Officer June 2003 – June 2006
Interbase Solutions Corp
San Juan, PR

TEACHING EXPERIENCE

Assistant Professor
Computational Biology Department
Carnegie Mellon University, Pittsburgh, PA

- *Automation of Scientific Research* (graduate course) –SPRING 2023

Assistant Professor
Department of Computer Science
University of Puerto Rico, San Juan, PR

- *Introduction to Computer Science* (undergraduate course) – FALL 2021, SPRING 2022
- *Undergraduate Seminar I* (undergraduate course) – FALL 2021

Associate Instructor/Teaching Assistant
Department of Computer Science
Indiana University, Bloomington, IN

- *Data mining* (graduate course) – FALL 2013, SPRING 2016
- *Machine learning* (graduate course) – SPRING 2014
- *Introduction to bioinformatics* (graduate course) – FALL 2014

Instructor
Department of Linguistics
University of California, San Diego, CA

- *Conversation Spanish* (undergraduate course) – FALL 2007 & 2008, WINTER 2007 & 2008, SPRING 2008 & 2009

Adjunct Lecturer
Department of Computer Science
University of Puerto Rico, Rio Piedras, PR

- *Compilers design* (upper-level undergraduate course) – SPRING 2004 & 2007
- *Introduction to computer science* (undergraduate course) – FALL 2005

AWARDS AND HONORS

NIH Diversity Supplement	December 2019 – July 2021
Lane Fellowship, Carnegie Mellon University	September 2017 – July 2021
LatinX in AI Workshop Travel Grant for International Conference in Machine Learning	June 2019
Graduate student speaker, Latino Congratulatory Ceremony	May 2016
Nominated for Associate Instructor (Teaching Assistant) of the year, Computer Science Program	SPRING 2016
Best poster award, 2nd Symposium of Health Informatics in Latin America and the Caribbean	November 2015
Best oral graduate student presentation, American Association for the Advancement of Science (AAAS) Emerging Researchers National Conference	February 2013
ACM SIG Travel Award for Richard Tapia Celebration of Diversity in Computing	February 2013
FASEB-MARC Program Travel Award for International Society for Computational Biology (ISCB) Rocky Mountain Bioinformatics Conference	December 2012

Best oral graduate student presentation in Computer Science, Society for the Advancement of Chicanos and Native Americans in Science (SACNAS) National Conference	October 2012
Ford Foundation Pre-doctoral Diversity Fellowship	August 2010 – May 2013
Computer Packages Inc. Hispanic College Fund (HCF) Scholarship	August 2011 – May 2012
Computer Science Corporation HCF Scholarship	August 2010 – May 2011
Google Society of Hispanic Professional Engineers Travel Scholarship	October 2009
Google HCF Scholarship	August 2009 – May 2010
Indiana University Graduate Scholars Fellowship	August 2009 – May 2010
University of Puerto Rico; Computer Science, Engineering and Mathematics Undergraduate Scholarship (CSEMS-NSF)	August 2001 – May 2003

PUBLICATIONS

P Jiang, R. Cai, **J. Lugo-Martinez**[+], Y. Guo[+], *A hybrid positive unlabeled learning framework for uncovering scaffolds across human proteome by measuring the propensity to drive phase separation*, Briefings in Bioinformatics, bbad009, 2023 ([+]Indicates corresponding authors).

J. Lugo-Martinez*, S. Xu*, J. Levesque, D. Gallagher, L.A. Parker, J. Neu, C.J. Stewart, J.E. Berrington, N.D. Embleton, G. Young, K.E. Gregory, M. Good, A. Tandon, D. Genetti, T. Warren and Z. Bar-Joseph, *Integrating longitudinal clinical and microbiome data to predict growth faltering in preterm infants*, Journal of Biomedical Informatics, 128:104031, 2022 (*Indicates equal contribution).

D. Ruiz-Perez*, **J. Lugo-Martinez***, N. Bourguignon, K. Mathee, B. Lerner, Z. Bar-Joseph and G. Narasimhan, *Dynamic Bayesian networks for integrating muti-omics time-series microbiome data*, mSystems, 6(2):e01105-20, 2021 (*Indicates equal contribution).

V. Pejaver, J. Urresti, **J. Lugo-Martinez**, K.A. Pagel, G.N. Lin, H.J. Nam, M. Mort, D.N. Cooper, J. Sebat, L.M. Iakoucheva, S.D. Mooney and P. Radivojac, *Inferring the molecular and phenotypic impact of amino acid variants with MutPred2*, Nature Communications 11(1):5918, 2020.

J. Lugo-Martinez, D. Zeiberg, T. Gaudelet, N. Malod-Dognin, N. Pržulj and P. Radivojac, *Classification in biological networks with hypergraphlet kernels*, Bioinformatics 37(7):1000-1007, 2020.

J. Lugo-Martinez, J. Dengjel, Z. Bar-Joseph and R.F. Murphy, *Integration of heterogeneous experimental data improves global map of human protein complexes*, In the Proceedings of the 10th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM-BCB), 2019.

J. Lugo-Martinez*, D. Ruiz-Perez*, G. Narasimhan and Z. Bar-Joseph, *Dynamic interaction network inference from longitudinal microbiome data*, Microbiome, 7(1):54, 2019 (*Indicates equal contribution).

M. Yamada, J. Tang, **J. Lugo-Martinez**, E. Hodzic, R. Shrestha, A. Saha, H. Ouyang, D. Yin, H. Mamitsuka, C. Sahinalp, P. Radivojac, F. Menczer and Y. Chang, *Ultra high-dimensional nonlinear feature selection for big biological data*, IEEE Transactions on Knowledge and Data Engineering, 30(7):1352-1365, 2018.

T. Gaudelet, N. Malod-Dognin, **J. Lugo-Martinez**, P. Radivojac and N. Pržulj, *Hypergraphlets give insight into multi-scale organization of molecular networks*, The 6th International Conference on Complex Networks and Their Applications, 41-43, 2017.

J. Lugo-Martinez, V. Pejaver, K.A. Pagel, S. Jain, M. Mort, D.N. Cooper, S.D. Mooney and P. Radivojac, *The loss and gain of functional residues is a frequent mechanism of human genetic disease*, PLoS Computational Biology, 12(8):e1005091, 2016.

J.F. Denton, **J. Lugo-Martinez**, A.E. Tucker, D.R. Schrider, W.C. Warren and M.W. Hahn, *Extensive error in the number of genes inferred from draft genome assemblies*, PLoS Computational Biology, 10(12):e1003998, 2014.

J. Lugo-Martinez and P. Radivojac, *Generalized graphlet kernels for probabilistic inference in sparse graphs*, Network Science, 2(2):254-276, 2014.

J. Lugo-Martinez and P. Radivojac, *Vertex classification in graphs*, Biomedical Computation Review, Summer 2013.

M.V. Han, G.W.C. Thomas, **J. Lugo-Martinez** and M.W. Hahn, *Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3*, Molecular Biology and Evolution, 30:1987-1997, 2013.

G. Venkatesh, J. Sampson, N. Goulding, S. Garcia, S. Bryksin, **J. Lugo-Martinez**, S. Swanson and M. Taylor, *Conservation cores: Reducing the energy of mature computations*, In the Proceedings of the 15th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS 2010), 38:205-218, 2010.

PREPRINTS

J. Ding*, **J. Lugo-Martinez***, Y. Yuan*, J. Huang, A.J. Hume, E.L. Suder, E. Mühlberger, D.N. Kotton and Z. Bar-Joseph, *Reconstructed signaling and regulatory networks identify potential drugs for SARS-CoV-2 infection*, 2021 (*Indicates equal contribution). An older preprint version is available at https://doi.org/10.1101/2020.06.01.127589

TECHNICAL REPORTS

J. Lugo-Martinez, *Flexible kernel functions for learning on graphs and hypergraphs*, PhD Thesis, Indiana University, 2016.

J. Lugo-Martinez, *Strategies for sharing the floating point unit between specialized processing elements*, Master's Thesis, University of California-San Diego, 2010.

K. Acosta*, M. Hernandez* and **J. Lugo-Martinez***, *Lexicographic and non-lexicographic greedy codes*, Summer Institute in Mathematics for Undergraduates: Technical Reports, 2001 (*Indicates equal contribution).

INVITED PRESENTATIONS

Computational approaches for integrating and analyzing biological data

- University of Pittsburgh, Department of Computational & Systems Biology, January 2022
- Carnegie Mellon University, School of Computer Science, December 2021
- University of Puerto Rico, Department of Computer Science, November 2020

Computational approaches for integrating and analyzing longitudinal multi-omics microbiome data

- Southwest Regional IDeA Conference, November 2021

Dynamic interaction network inference from longitudinal microbiome data

- Graduate Student Association Seminar, Carnegie Mellon University, March 2020
- Symposium on Evolutionary Ecology and Host-Virus Dynamics, SMBE 2019, July 2019

Flexible kernel functions for leaning on graphs and hypergraphs

- Carnegie Mellon University, Department of Computational Biology, April 2017
- ETH Zürich, Department of Biosystems Science and Engineering, October 2016

Computational approaches for understanding molecular mechanisms of disease

- Indiana University, Bioinformatics for precision medicine course, April 2017
- Rutgers University, Department of Biochemistry and Microbiology, July 2016
- University of Puerto Rico, Department of Computer Science, May 2016

Graphlets for learning and inference on real-world networks

- University of Puerto Rico, Department of Computer Science, May 2016

Graphlet kernels for vertex classification

- Indiana University, Data mining course, April 2016
- University of Puerto Rico, Department of Computer Science, November 2015
- 45th Symposium of the Interface on Computing Science and Statistics, June 2015

REFEREED CONFERENCE PRESENTATIONS

Integration of heterogeneous experimental data improves global map of human protein complexes

- (Oral) 10th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics, September 2019

Dynamic interaction network inference from longitudinal microbiome data

- (Poster) 1st Rust Belt Microbiome Conference, November 2019
- (Poster) 10th ACM Conference on Bioinformatic, Computational Biology, and Health Informatics, September 2019
- (Poster) 36th International Conference on Machine Learning, June 2019
- (Oral) Great Lakes Bioinformatics Conference, May 2019
- (Poster) 23rd Research in Computational Molecular Biology, May 2019

Classification in biological networks with hypergraphlet kernels

- (Poster) 36th International Conference on Machine Learning, June 2019
- (Poster) 23rd Research in Computational Molecular Biology, May 2019
- (Oral) Great Lakes Bioinformatics Conference, May 2017

In Silico analysis from protein structures reveals loss and gain of functional sites as active mechanism of disease

- (Poster) 24th International Conference on Intelligent Systems for Molecular Biology, July 2016
- (Poster) 2nd Symposium of Health Informatics in Latin America and the Caribbean, November 2015. **Awarded Best Poster**.
- (Oral) Great Lakes Bioinformatics Conference, May 2015

Flexible graphlet kernels for prediction of functional residue in protein structures

- (Oral) AAAS Emerging Researchers National Conference, February 2013. **Awarded Best Oral Graduate Student Presentation**.
- (Oral) Ford Foundation Fellows Conference, September 2010

Flexible kernels for learning on large sparse graphs and hypergraphs

- (Oral) ACM Richard Tapia Celebration of Diversity in Computing Conference, February 2013

Hypergraph kernels for protein function prediction using protein complexes

- (Oral) 10th Annual Rocky Mountain Bioinformatics Conference, December 2012

Flexible graphlet kernels for vertex classification in sparse graphs

- (Oral) SACNAS National Conference, October 2012. **Awarded Best Oral Graduate Student Presentation in Computer Science**.
- (Oral) 20th International Conference on Intelligent Systems for Molecular Biology, SIG Automated Function Prediction, July 2012.

PROFESSIONAL SERVICE

Current Trainees

Post-doctoral fellows

- Duc Nam Nguyen, 2022 – present, Lane Fellow (CMU)
- Hao Chen, 2022 – present, Post-doctoral Fellow (CMU)

PhD students

- Young Je Lee, 2023 – present, CPCB Program (CMU)
- Peiran Jiang, 2022 – present, CPCB Program (CMU)

Master students

- Mohammadreza Moradi, 2023 – present, MSBME (CMU)
- Daniel Jimenez, 2022 – present, MSAS Program (CMU)
- Yifan Wu, 2022 – present, MSCB Program (CMU)

Undergraduate students

- Peter Sauer, 2022 – present, Biology and Statistics (CMU)

Dissertation committee

- Steven W. Smeal, 2023 – present, CPCB Program (Pitt)
- Mustafa Guler, 2023 – present, CPCB Program (CMU)
- Haotian (Frank) Zhang, 2022 – present, CPCB Program (Pitt)

Prior Trainees

Graduate mentees

- Sofia Melendez-Cartagena, Summer 2020, Biology (UPR-RP)
- Daniel Ruiz-Perez, 2017 – 2020, Computer Science (FIU)

Undergraduate students

- Michael Figueroa-Muñiz, Summer 2019, Mathematics (UPR-RP)
- Matthew Carey, 2015 – 2016, Computer Science (IU)
- Jacob Weimer, 2013 – 2014, Computer Science (IU). **Awarded Best Poster Award**, Research Experience for Undergraduates, School of Informatics and Computing, Indiana University

Referee/Subreferee

- *Journals* – Nature Methods, Bioinformatics, PLoS Computational Biology, BMC Bioinformatics, IEEE/ACM Transactions in Computational Biology and Bioinformatics, IEEE Transactions on Knowledge and Data Engineering, and Journal of Machine Learning Research
- *Conferences* – RECOMB, PSB, ISMB, ACM BCB

OUTREACH ACTIVITIES

Organizing Committee, 8th Workshop on Computational Network Biology: Modeling, Analysis, and Control (CNB-MAC), 2023

Organizing Committee, 3rd Nobel Turing Challenge Initiative Workshop, 2023

Organizing Committee, 7th Workshop on Computational Network Biology: Modeling, Analysis, and Control (CNB-MAC), 2022

Program Committee, Latinx in Artificial Intelligence Research Workshop, 36th International Conference on Machine Learning (ICML), 2019

Program Committee, Latinx in Artificial Intelligence Research Workshop, 32nd Conference on Neural Information Processing Systems (NeurIPS), 2018

Organizing Committee, Broadening Participation in Data Mining Workshop, ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2017

Mentor, Latino Male Initiative, Indiana University, 2015 – 2017

- **Awarded Program/Group of the Year** by the Latino Faculty and Staff Council, 2016

Co-founder, Bioinformatics Club, Indiana University, 2011 – 2017

Member, SACNAS Chapter, Indiana University, 2009 – 2017

- President, June 2012 – June 2013, **Awarded Graduate Chapter of the Year**, 2013
- Executive board, June 2011 – June 2012, **Awarded Role Chapter of the Year**, 2012

PROFESSIONAL MEMBERSHIPS

Association for Computing Machinery (ACM)
International Society for Computational Biology (ISCB)
Society for the Advancement of Chicanos and Native Americans in Science (SACNAS)

LANGUAGES

Spanish: native language; English: full professional proficiency; French: limited working proficiency