David C. Anastasiu
Curriculum Vitæ

April 2024

Address: Santa Clara University
Computer Science & Engineering
500 El Camino Real
Santa Clara, CA 95053
Office: (408) 551-1941
Email: danastasiu@scu.edu
WWW: davidanastasiu.net

Highlights

- Research in machine learning, data mining, computational genomics, and high performance computing.
- Passionate about teaching and mentoring students.

Education

- **PH.D. IN COMPUTER SCIENCE**, *University of Minnesota–Twin Cities*, Minneapolis, MN, 2011—2016.
  Advisor: Dr. George Karypis. Received a Doctoral Dissertation Fellowship award.
- **M.S. IN COMPUTER SCIENCE**, *Texas State University*, San Marcos, TX, 2009—2011.
  Advisor: Dr. Byron J. Gao. Received an Outstanding Graduate Student award and a Graduate Research Excellence award.
- **POST GRADUATE CERTIFICATE IN COMPUTER SCIENCE**, *Texas State University*, San Marcos, TX, 2008—2009.
- **B.A. IN BIBLE/THEOLOGY**, *Moody Bible Institute*, Chicago, IL, 1997—2001.
  Focused on ancient Greek studies. Graduated with honors.

Professional Experience

- **ASSISTANT PROFESSOR**, *Santa Clara University*, Santa Clara, CA, Fall 2019—Present.
- **ASSISTANT PROFESSOR**, *San José State University*, San José, CA, Fall 2016—Spring 2019.
- **MACHINE LEARNING CONSULTANT**, *In-Depth, Inc.*, San Francisco, CA, May—Nov 2014.
- **RESEARCH ASSISTANT**, *University of Minnesota—Twin Cities*, Minneapolis, MN, 2011—2016.
- **SYSTEMS DEVELOPER / INTEGRATION SPECIALIST**, *PPD, Inc.*, Austin, TX, 2005—2010.

Grants

External Research Grants

- *"Rainflow-Based Stream Runoff Prediction"*, Santa Clara Valley Water District, Santa Clara, CA, **$50,000.00**, 08/2/2022—05/31/2023. Role: **Principal Investigator**.

- *"Valley Water AI Models for Hydrologic Flow"*, Wood Rogers, Inc./Santa Clara Valley Water District, Santa Clara, CA, **$39,328.65**, 06/10/2020—06/09/2022. Role: **Principal Investigator**.

- *"CRII: III: RUI: Effective Protein Characterization via Fast Exact Open Modification Searching"*, National Science Foundation (NSF), #2002321, **$141,988**, 08/17/2019—01/31/2023, at Santa Clara University. Role: **Principal Investigator**.

- *"CRII: III: RUI: Effective Protein Characterization via Fast Exact Open Modification Searching"*, National Science Foundation (NSF), #1850557, **$33,012**, 02/15/2019—01/21/2020, at San José State University. Role: **Principal Investigator**.

- *"Anomaly Detection in Active DNS Data"*, Infoblox, Santa Clara, CA, **$240,000**, 01/28/2019—01/27/2021, (with Magdalini Eirinaki). Role: **Principal Investigator**.

- *"Flextronics Term Chair Grant"*, Flex Ltd., Milpitas, CA, **$90,000**, 02/01/2018—01/31/2019. Role: **Principal Investigator**.

- *"DARPA HIVE Project Subcontract"*, Intel Labs, Santa Clara, CA, **$25,000**, 06/01/2018—01/08/2019. Role: **Principal Investigator**.

- *"Mass Warning Study for the City of San José"*, San José Office of Emergency Management, **$42,000**, 10/01/2017—01/31/2019 (with Zeyu Gao and Subhankar Dhar). Role: **Co-Principal Investigator**.

- *"GPU Grant"*, NVIDIA, Santa Clara, CA, **$1,940**, 01/15/2017. Role: **Principal Investigator**.

Internal Research and Teaching Grants

- *"Inexpensive Intelligent Kidney Health Prediction"*, School of Engineering, SCU, **$15,000**, 2020—2022. Role: **Principal Investigator**.

- *"Use of a Competitive Learning Platform to Increase Student Engagement"*, The Faculty Collaborative for Teaching Innovation, SCU, **$6,800**, 2020—2023 (with Fatemeh Tehranipoor). Role: **Principal Investigator**.

- *"The College of Engineering High Performance Computing System"*, College of Engineering, SJSU, **$530,000**, 2017—2018. Role: **Principal Investigator**.

- *"Predictive Analytics for SJSU Admission Yield Estimation"*, Provost Office, SJSU, **$16,000**, 08/01/2018—07/31/2019 (with Gheorghi Guzun). Role: **Co-Principal Investigator**.

- *"Optimal Constrained Wireless Network Antenna Placement"*, College of Engineering Excellence Research Center on Smart Technology, Computing, and Complex System (STCCS), **$9,500**, 01/15/2017—12/31/2017. Role: **Principal Investigator**.

- *"Davidson Student Scholar Grants"*, College of Engineering, SJSU, **$11,000**, 2016—2019 (supporting graduate students in my lab). Role: **Advisor**.

- *"Professional Development Grants"*, College of Engineering & Department of Computer Engineering, SJSU, **$29,910**, 2016—2019. Role: **Principal Investigator**.

- *"University Grants Academy"*, SJSU, **20% teaching release (~$11,250)**, 01/01/2017—06/01/2017.

Awards & Honors

- **Outstanding Career Influencer** award, 2021-2022, Santa Clara University Career Center.

- **Outstanding Paper Reviewer** award, 2017, 26th ACM International Conference on Information and Knowledge Management, CIKM'17, Pan Pacific, Singapore.

- **Next Generation Data Scientist** award, 2016, 3rd IEEE International Conference on Data Science and Advanced Analytics, DSAA'16, Montreal, Canada. This is the first early career award given at DSAA, and the only one awarded in 2016.

- **Best Research Paper** award, 2016, 3rd IEEE International Conference on Data Science and Advanced Analytics, DSAA'16, Montreal, Canada.

Selected Publications

Student co-authors are denoted with an asterisk (*).

Journal Articles

Alex Whelan*, Ragwa Elsayed, Alessandro Bellofiore, and David C. Anastasiu. Selective partitioned regression for accurate kidney health monitoring. *Annals of Biomedical Engineering*, Feb 2024. **Contributions**: research design, model development, data analysis, co-advised student authors, co-wrote and reviewed the manuscript. **Impact factor**: 4.0.

Nivedha Balakrishnan, Rahul Katkar, Peter V. Pham, Taylor Downey*, Prarthna Kashyap, David C. Anastasiu, and Anand K. Ramasubramanian. Prospection of peptide inhibitors of thrombin from diverse origins using a machine learning pipeline. *Bioengineering*, 10(11), 2023. **Contributions**: research design, model development, data analysis, co-advised student authors, co-wrote and reviewed the manuscript. **Impact factor**: 4.6.

Yijia Li, Jonathan Nguyen*, David C Anastasiu, and Edgar A Arriaga. CosTaL: an accurate and scalable graph-based clustering algorithm for high-dimensional single-cell data analysis. *Briefings in Bioinformatics*, 05 2023. **Contributions**: helped develop the algorithms, co-advised student authors, co-wrote and reviewed the manuscript. **Impact factor**: 13.994.

Mohammed Shaiqur Rahman, Archana Venkatachalapathy, Anuj Sharma, Jiyang Wang, Senem Velipasalar Gursoy, David Anastasiu, and Shuo Wang. Synthetic distracted driving (syndd1) dataset for analyzing distracted behaviors and various gaze zones of a driver. *Data in Brief*, 46:108793, Feb 2023. **Contributions**: contributed to dataset capture protocols, developed evaluation criteria, implemented evaluation software. **Impact factor**: 1.2.

Bipasa Bose, Taylor Downey*, Anand K. Ramasubramanian, and David C. Anastasiu. Identification of distinct characteristics of antibiofilm peptides and prospection of diverse sources for efficacious sequences. *Frontiers in Microbiology*, 12, 2022. **Contributions**: helped develop the algorithms, co-advised student authors, co-wrote and reviewed the manuscript. **Impact factor**: 6.064.

Andrea Tagarelli, Ester Zumpano, David C. Anastasiu, Andrea Calì, and Gottfried Vossen. Managing, mining and learning in the legal data domain. *Inf. Syst.*, 106(C), May 2022. **Contributions**: assisted with journal manuscript reviews and editing the summary paper. **Impact factor**: 7.767.

David C. Anastasiu, Jack Gaul*, Maria Vazhaeparambil*, Meha Gaba*, and Prajval Sharma*. Efficient citywide multi-class multi-movement vehicle counting: A survey. *Journal of Big Data Analytics in Transportation*, 2(3):235–250, Dec 2020. **Contributions**: co-advised student authors, co-wrote, edited, and reviewed the manuscript. **Impact factor**: N/A (new journal).

David C. Anastasiu and George Karypis. Parallel cosine nearest neighbor graph construction. *Elsevier Journal of Parallel and Distributed Computing*, 2017. **Contributions**: developed algorithms, executed experiments, wrote and edited manuscript. **Impact factor**: 1.815.

David C. Anastasiu and George Karypis. Efficient identification of tanimoto nearest neighbors; all pairs similarity search using the extended jaccard coefficient. *Springer International Journal of Data Science and Analytics*, 4(3):153–172, Nov 2017. **Contributions**: developed algorithms, executed experiments, wrote and edited manuscript. **Impact factor**: 1.241.

David C. Anastasiu and Andrea Tagarelli. Document clustering. *Wiley StatsRef: Statistics Reference Online*, pages 1–11, 2017. **Contributions**: literature survey, co-wrote and edited manuscript. **Impact factor**: N/A.

David C. Anastasiu, Evangelia Christakopoulou, Shaden Smith, Mohit Sharma, and George Karypis. Big data and recommender systems. *Novática: Journal of the Spanish Computer Scientist Association*, 1(237):39–45, October 2016. **Contributions**: literature survey, co-wrote and edited manuscript. **Impact factor**: N/A.

David C. Anastasiu, Byron J. Gao, Xing Jiang, and George Karypis. A novel two-box search paradigm for query disambiguation. *World Wide Web*, 16(1):1–29, 2013. **Contributions**: developed algorithms, executed experiments, literature survey, co-wrote and edited manuscript. **Impact Factor**: 0.653.

Byron J. Gao, David Buttler, David C. Anastasiu, Shuaiqiang Wang, Peng Zhang, and Joey Jan. User-centric organization of search results. *IEEE Internet Computing*, 17(3):52–59, May 2013. **Contributions**: developed algorithms, executed experiments, literature survey, co-wrote and edited manuscript. **Impact factor**: 0.758.

Conference Proceedings

Yanhong Li*, Jack Xu, and David C. Anastasiu. Learning from polar representation: An extreme-adaptive model for long-term time series forecasting. In *Proceedings of the Thirty-Eith AAAI Conference on Artificial*

Intelligence and Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence and Fourteenth Symposium on Educational Advances in Artificial Intelligence, AAAI'24/IAAI'24/EAAI'24. AAAI Press, 2024. **Contributions**: supervised student author, co-wrote and edited manuscript. **Acceptance rate**: 23.75%.

Yanhong Li*, Jack Xu, and David C. Anastasiu. Seed: An effective model for highly-skewed streamflow time series data forecasting. In *2023 IEEE International Conference on Big Data (Big Data)*, IEEE BigData 2023, Los Alamitos, CA, USA, Dec 2023. IEEE Computer Society. **Contributions**: supervised student author, co-wrote and edited manuscript. **Acceptance rate**: 17.5%.

Yanhong Li*, Jack Xu, and David C. Anastasiu. An extreme-adaptive time series prediction model based on probability-enhanced lstm neural networks. In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence*, AAAI'23/IAAI'23/EAAI'23. AAAI Press, 2023. **Contributions**: supervised student author, co-wrote and edited manuscript. **Acceptance rate**: 19.6%.

Arpita Vats* and David C. Anastasiu. Enhancing retail checkout through video inpainting, yolov8 detection, and deepsort tracking. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 5530–5537, 2023. **Contributions**: supervised student author, co-developed algorithms, co-wrote and edited manuscript. **Acceptance rate**: N/A.

Milind Naphade, Shuo Wang, David C. Anastasiu, Zheng Tang, Ming-Ching Chang, Yue Yao, Liang Zheng, Mohammed Shaiqur Rahman, Meenakshi S. Arya, Anuj Sharma, Qi Feng, Vitaly Ablavsky, Stan Sclaroff, Pranamesh Chakraborty, Sanjita Prajapati, Alice Li, Shangru Li, Krishna Kunadharaju, Shenxin Jiang, and Rama Chellappa. The 7th ai city challenge. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 5538–5548, June 2023. **Contributions**: developed evaluation strategy, implemented evaluation systems, helped organize challenge and workshop, co-wrote and edited summary paper. **Acceptance rate**: N/A.

Alex Whelan*, Soham Phadke*, and David C. Anastasiu. On-device prediction for chronic kidney disease. In *2022 IEEE Global Humanitarian Technology Conference (GHTC)*, GHTC 2022, pages 325–332, 2022. **Contributions**: supervised student authors, co-wrote and edited manuscript. **Acceptance rate**: N/A.

Arpita Vats, Gheorghi Guzun, and David C. Anastasiu. Clp: A platform for competitive learning. In *Educating for a New Future: Making Sense of Technology-Enhanced Learning Adoption*, page 615âĂŞ622, Berlin, Heidelberg, 2022. Springer-Verlag. **Contributions**: supervised student author, co-wrote and edited manuscript. **Acceptance rate**: 28%.

Dorian Clay* and David C. Anastasiu. Expanding neuro-symbolic artificial intelligence for strategic learning. In *The 2022 KDD Undergraduate Consortium (KDD-UC)*, July 2022. **Contributions**: supervised student author, edited manuscript. **Acceptance rate**: N/A.

Raghav Kapoor*, Casey Nguyen*, and David C. Anastasiu. Interpnet: Interpretability for autonomous driving neural networks. In *The 2022 KDD Undergraduate Consortium (KDD-UC)*, July 2022. **Contributions**: supervised student authors, edited manuscript. **Acceptance rate**: N/A.

Arpita Vats* and David C. Anastasiu. Key point-based driver activity recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 3274–3281, June 2022. **Contributions**: supervised student author, co-developed algorithms, co-wrote and edited manuscript. **Acceptance rate**: N/A.

Milind Naphade, Shuo Wang, David C. Anastasiu, Zheng Tang, Ming-Ching Chang, Xiaodong Yang, Yue Yao, Liang Zheng, Pranamesh Chakraborty, Christian E. Lopez, Anuj Sharma, Qi Feng, Vitaly Ablavsky, and Stan Sclaroff. The 5th ai city challenge. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern*

Recognition (CVPR) Workshops, CVPRW'21, pages 4263–4273, June 2021. **Contributions**: developed evaluation strategy, implemented evaluation systems, helped organize challenge and workshop, co-wrote and edited summary paper. **Acceptance rate**: N/A.

Milind Naphade, Shuo Wang, David C. Anastasiu, Zheng Tang, Ming-Ching Chang, Xiaodong Yang, Liang Zheng, Anuj Sharma, Rama Chellappa, and Pranamesh Chakraborty. The 4th ai city challenge. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, volume 1 of *CVPRW'20*, pages 2665–2674, June 2020. **Contributions**: developed evaluation strategy, implemented evaluation systems, helped organize challenge and workshop, co-wrote and edited summary paper. **Acceptance rate**: N/A.

Zheng Tang, Milind Naphade, Ming-Yu Liu, Xiaodong Yang, Stan Birchfield, Shuo Wang, Ratnesh Kumar, David C. Anastasiu, and Jenq-Neng Hwang. Cityflow: A city-scale benchmark for multi-target multi-camera vehicle tracking and re-identification. In *CVPR 2019: IEEE Conference on Computer Vision and Pattern Recognition*, CVPR 2019, 2019. **Contributions**: developed evaluation strategy, implemented evaluation system, co-wrote and edited manuscript. **Acceptance rate**: 25.2%.

Manika Kapoor* and David C. Anastasiu. A data-driven approach for detecting autism spectrum disorders. In Peter Haber, Thomas Lampoltshammer, and Manfred Mayr, editors, *Data Science – Analytics and Applications*, iDSC 2019, Wiesbaden, 2019. Springer Fachmedien Wiesbaden. **Contributions**: supervised student author, developed algorithms, co-wrote and edited manuscript. **Acceptance rate**: N/A.

Saloni Mohan*, Sahitya Mullapudi*, Sudheer Sammeta*, Parag Vijayvergia*, and David C. Anastasiu. Stock price prediction using news sentiment analysis. In *2019 IEEE Fourth International Conference on Big Data Computing Service and Applications (BigDataService)*, BDS 2019. IEEE, April 2019. **Contributions**: supervised student authors, developed algorithms, co-wrote and edited manuscript. **Acceptance rate**: 19.5%.

Shuai Hua* and David C. Anastasiu. Effective vehicle tracking algorithm for smart traffic networks. In *Thirteenth IEEE International Conference on Service-Oriented System Engineering (SOSE)*, SOSE 2019. IEEE, April 2019. **Contributions**: supervised student author, co-developed algorithms, co-wrote and edited manuscript. **Acceptance rate**: 28%.

Anupama Upadhayula*, Avinash Ravilla*, Ishwarya Varadarajan*, Sowmya Viswanathan*, and David C. Anastasiu. Study area recommendation via network log analytics. In *The Seventh IEEE International Conference on Mobile Cloud Computing, Services, and Engineering*, MCCSE 2019. IEEE, April 2019. **Contributions**: supervised student authors, developed algorithms, co-wrote and edited manuscript. **Acceptance rate**: N/A.

Milind Naphade, Zheng Tang, Ming-Ching Chang, David C. Anastasiu, Anuj Sharma, Rama Chellappa, Shuo Wang, Pranamesh Chakraborty, Tingting Huang, Jenq-Neng Hwang, and Siwei Lyu. The 2019 ai city challenge. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, volume 1 of *CVPRW'19*, pages 452–460, June 2019. **Contributions**: developed evaluation strategy, implemented evaluation systems, helped organize challenge and workshop, co-wrote and edited summary paper. **Acceptance rate**: N/A.

Swapnil Gaikwad*, Melody Moh, and David C. Anastasiu. Data structure for efficient line of sight queries. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, CIKM'18, New York, NY, USA, 2018. ACM. **Contributions**: supervised student author, co-developed algorithms, co-wrote and edited manuscript. **Acceptance rate**: 18%.

Manika Kapoor*, Shuai Hua*, and David C. Anastasiu. Improving student motivation through competitive active learning. In *2018 IEEE Frontiers in Education Conference*, FIE 2018. IEEE, October 2018. **Contributions**: supervised student authors, co-developed algorithms, co-wrote and edited manuscript. **Acceptance rate**: 20%.

Shuai Hua*, Manika Kapoor*, and David C. Anastasiu. Vehicle tracking and speed estimation from traffic videos. In *2018 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, volume 1 of *CVPRW'18*, pages 153–1537, July 2018. **Contributions**: supervised student authors, co-developed algorithms, co-wrote and edited manuscript. **Acceptance rate**: N/A.

Milind Naphade, Ming-Ching Chang, Anuj Sharma, David C. Anastasiu, Vamsi Jagarlamudi, Pranamesh Chakraborty, Tingting Huang, Shuo Wang, Ming-Yu Liu, Rama Chellappa, Jenq-Neng Hwang, and Siwei Lyu. The 2018 nvidia ai city challenge. In *2018 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, volume 1 of *CVPRW'18*, pages 53–60, July 2018. **Contributions**: developed evaluation strategy, implemented evaluation systems, helped organize challenge and workshop, co-wrote and edited summary paper. **Acceptance rate**: N/A.

Swapnil Gaikwad* and David C. Anastasiu. Optimal constrained wireless emergency network antenna placement. In *Proceedings of the IEEE Smart City Innovations 2017 Conference*, IEEE SCI 2017, 2017. **Contributions**: supervised student author, co-developed algorithms, co-wrote and edited manuscript. **Acceptance rate**: N/A.

David C. Anastasiu. Cosine approximate nearest neighbors. In Peter Haber, Thomas Lampoltshammer, and Manfred Mayr, editors, *Data Science – Analytics and Applications*, iDSC 2017, pages 45–50, Wiesbaden, 2017. Springer Fachmedien Wiesbaden. **Contributions**: literature survey, developed and implemented algorithms, performed experiments, wrote and edited manuscript. **Acceptance rate**: N/A.

Niveditha Bhandary*, Charles MacKay*, Alex Richards*, Ji Tong*, and David C. Anastasiu. Robust classification of city roadway objects for traffic related applications. In *2017 IEEE Smart World NVIDIA AI City Challenge*, SmartWorld'17, Piscataway, NJ, USA, 2017. IEEE. **Contributions**: supervised student authors, co-developed algorithms, co-wrote and edited manuscript. **Acceptance rate**: N/A.

Milind Naphade, David C. Anastasiu, Anuj Sharma, Vamsi Jagrlamudi, Hyeron Jeon, Kaikai Liu, Ming-Ching Chang, Siwei Lyu, and Zeyu Gao. The nvidia ai city challenge. In *2017 IEEE SmartWorld Conference*, Smart-World'17, Piscataway, NJ, USA, 2017. IEEE. **Contributions**: developed evaluation strategy, implemented evaluation systems, helped organize challenge and workshop, co-wrote and edited summary paper. **Acceptance rate**: N/A.

David C. Anastasiu and George Karypis. Efficient identification of tanimoto nearest neighbors. In *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, DSAA'16, pages 156–165, 2016. **Contributions**: literature survey, developed and implemented algorithms, performed experiments, wrote and edited manuscript. **Acceptance rate**: 20.2%. *Best Research Paper Award*.

David C. Anastasiu and George Karypis. Fast parallel cosine k-nearest neighbor graph construction. In *2016 6th Workshop on Irregular Applications: Architecture and Algorithms (IA3)*, IA3 2016, pages 50–53, Nov 2016. **Contributions**: literature survey, developed and implemented algorithms, performed experiments, wrote and edited manuscript. **Acceptance rate**: 48.3%.

David C. Anastasiu and George Karypis. L2knng: Fast exact k-nearest neighbor graph construction with l2-norm pruning. In *Proceedings of the 24th ACM International Conference on Information and Knowledge Management*, CIKM'15, pages 791–800, New York, NY, USA, 2015. ACM. **Contributions**: literature survey, developed and implemented algorithms, performed experiments, wrote and edited manuscript. **Acceptance rate**: 26%.

David C. Anastasiu, Al M. Rashid, Andrea Tagarelli, and George Karypis. Understanding computer usage evolution. In *31st IEEE International Conference on Data Engineering*, ICDE 2015, pages 1549–1560, 2015. **Contributions**: literature survey, developed and implemented algorithms, performed experiments, co-wrote and edited manuscript. **Acceptance rate**: 25%.

David C. Anastasiu and George Karypis. Pl2ap: Fast parallel cosine similarity search. In *Proceedings of the 5th Workshop on Irregular Applications: Architectures and Algorithms, in conjunction with SC'15*, IA3 2015, pages 1–8, New York, NY, USA, 2015. ACM. **Contributions**: literature survey, developed and implemented algorithms, performed experiments, wrote and edited manuscript. **Acceptance rate**: 58.3%.

David C. Anastasiu and George Karypis. L2ap: Fast cosine similarity search with prefix l-2 norm bounds. In *The 30th IEEE International Conference on Data Engineering*, ICDE 2014, pages 784–795, 2014. **Contributions**: literature survey, developed and implemented algorithms, performed experiments, wrote and edited manuscript. **Acceptance rate**: 20%.

David C. Anastasiu, Byron J. Gao, and David Buttler. A framework for personalized and collaborative clustering of search results. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*, CIKM'11, pages 573–582, New York, NY, USA, 2011. ACM. **Contributions**: literature survey, developed and implemented algorithms, performed experiments, co-wrote and edited manuscript. **Acceptance rate**: 25%.

David C. Anastasiu, Byron J. Gao, and David Buttler. Clusteringwiki: personalized and collaborative clustering of search results. In *The 34th ACM SIGIR International Conference on Research and Development in Information Retrieval*, SIGIR 2011, pages 1263–1264, 2011. **Contributions**: co-developed and implemented algorithms, performed experiments, co-wrote and edited manuscript. **Acceptance rate**: 20%.

Byron J. Gao, David C. Anastasiu, and Xing Jiang. Utilizing user-input contextual terms for query disambiguation. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, COLING'10, pages 329–337, Stroudsburg, PA, USA, 2010. Association for Computational Linguistics. **Contributions**: co-developed and implemented algorithms, performed experiments, co-wrote and edited manuscript. **Acceptance rate**: 41%.

Byron J. Gao, Mingji Xia, Walter Cai, and David C. Anastasiu. The gardener's problem for web information monitoring. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management*, CIKM'09, pages 1525–1528, New York, NY, USA, 2009. ACM. **Contributions**: co-developed and implemented algorithms, performed experiments, co-wrote and edited manuscript. **Acceptance rate**: 15%.

Walter Cai, David C. Anastasiu, Mingji Xia, and Byron J. Gao. Olap for multicriteria maintenance scheduling. In *The 5th International Conference on Data Mining*, DMIN'09, pages 35–41. CSREA Press, 2009. **Contributions**: co-developed and implemented algorithms, performed experiments, co-wrote and edited manuscript. **Acceptance rate**: N/A.

Book Chapters

Evangelia Christakopoulou, Shaden Smith, Mohit Sharma, Alex Richards*, David C. Anastasiu, and George Karypis. Scalability and distribution of collaborative recommenders. In *Collaborative Recommendations: Algorithms, Practical Challenges and Applications*, CR2019. World Scientific Publishing, Singapore, April 2019. **Contributions**: organized project, co-developed literature survey, co-wrote and edited manuscript.

David C. Anastasiu, Jeremy Iverson, Shaden Smith, and George Karypis. Big data frequent pattern mining. In *Frequent Pattern Mining*, pages 225–260. Springer International Publishing, Switzerland, 2014. **Contributions**: organized project, co-developed literature survey, co-wrote and edited manuscript.

David C. Anastasiu, Andrea Tagarelli, and George Karypis. Document clustering: The next frontier. In *Data Clustering: Algorithms and Applications*, pages 305–338. CRC Press, Boca Raton, FL, USA, 2013. **Contributions**: organized project, co-developed literature survey, co-wrote and edited manuscript.

Technical Reports

Yanhong Li*, Jack Xu, and David C. Anastasiu. Learning from polar representation: An extreme-adaptive model for long-term time series forecasting. Technical Report 2312.08763, arXiv, 2023. **Contributions**: supervised student author, co-wrote and edited manuscript.

Yanhong Li*, Jack Xu, and David C. Anastasiu. An extreme-adaptive time series prediction model based on probability-enhanced lstm neural networks. Technical Report 2211.15891, arXiv, 2022. **Contributions**: supervised student author, co-wrote and edited manuscript.

Yijia Li, Jonathan Nguyen*, David C. Anastasiu, and Edgar A. Arriaga. Costal: An accurate and scalable graph-based clustering algorithm for high-dimensional single-cell data analysis. *bioRxiv*, 2022. **Contributions**: helped develop the algorithms, co-advised student authors, co-wrote and reviewed the manuscript.

Bipasa Bose, Taylor Downey*, Anand K. Ramasubramanian, and David C. Anastasiu. Identification of distinct characteristics of antibiofilm peptides and prospection of diverse sources for efficacious sequences. *bioRxiv*, 2021. **Contributions**: helped develop the algorithms, co-advised student authors, co-wrote and reviewed the manuscript.

Manika Kapoor* and David C. Anastasiu. A data-driven approach for detecting autism spectrum disorders. Technical Report 2019-1, San José State University, San José, CA, USA, 2019. **Contributions**: supervised student author, developed algorithms, co-wrote and edited manuscript.

David. J. Buttler, David Andrzejewski, Keith D. Stevens, David C. Anastasiu, and Byron J. Gao. Rapid exploitation and analysis of documents. Technical Report LLNL-TR-517731, Lawrence Livermore National Laboratory, Livermore, CA, USA, 2011. **Contributions**: co-wrote and edited manuscript.

Tutorials, Keynotes, and Invited Talks

David C. Anastasiu. The ai revolution of traffic analytics. Talk at the the 4th Silicon Valley Cybersecurity Conference (SVCC), San Jose, CA. May 18, 2023, 2023.

David C. Anastasiu. The joys of working with data. Talk at the Breaking Into Research event, part of the Big Data at Berkeley Speaker Series, at UC Berkeley, Berkeley, CA. Apr 6, 2023, 2023.

David C. Anastasiu, Huzefa Rangwala, and Andrea Tagarelli. Are you my neighbor? bringing order to neighbor computing problems. In *The 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD'19, New York, NY, USA, 2019. ACM. **Contributions**: Co-developed tutorial, co-wrote tutorial paper. Tutorial presented at KDD'19, Aug 04, 2019.

David C. Anastasiu. The ai data revolution: Doing more with less data labeling. Keynote at the 2nd International Data Science Conference 2019, Salzburg, Austria. May 22, 2019, 2019.

David C. Anastasiu. The biomedical ai revolution. Keynote at the 10th Annual Bay Area Biomedical Device Conference, San José, CA. Apr 03, 2019, 2019.

David C. Anastasiu. Fireside chat on current trends in machine learning. Talk at IBM Machine Learning Hub, San José, CA. Sep 25, 2018, 2018.

Alessandro Bellofiore, Ragwa Elsayed*, David C. Anastasiu, and Rathna Ramesh*. A novel image-based creatinine monitor for chronic kidney disease. Talk by Alessandro Bellofiore at the 8th World Congress of Biomechanics, held in Dublin, Ireland., 2018.

David C. Anastasiu. Efficient neighborhood graph construction for sparse high dimensional data. Talk at Lawrence Livermore National Laboratory, Livermore, CA, Feb 08, 2017, 2017.

Posters

Arpita Vats*, Gheorghe Guzun, and David C. Anastasiu. Clp: A platform for competitive learning, 2023. Poster presented by Arpita Vats at the 2023 School of Engineering Research Showcase, Santa Clara University, Santa Clara, CA. ***Third-place poster presentations winner in the Ph.D./Post-doc category***.

Yanhong Li*, Jack Xu, and David C. Anastasiu. An extreme-adaptive time series prediction model based on probability-enhanced lstm neural networks, 2023. Poster presented by Yanhong Li at the 2023 School of Engineering Research Showcase, Santa Clara University, Santa Clara, CA.

Alex Whelan*, Soham Phadke*, Alessandro Bellofiore, and David C. Anastasiu. On-device prediction for chronic kidney disease, 2023. Poster presented by Alex Whelan at the 2023 School of Engineering Research Showcase, Santa Clara University, Santa Clara, CA.

Yijia Li, David C. Anastasiu, and Edgar Arriaga. Costal: A network-based algorithm for clustering high-dimensional single-cell datasets, 2022. Abstract and poster presented by Yijia Li at the 30th annual conference on Intelligent Systems for Molecular Biology (ISMB), Madison, WI. Acceptance rate: 19.8%.

David C. Anastasiu and Gheorghi Guzun. Using competitive learning to increase student engagement, 2021. Poster presented by David C. Anastasiu at the 8th Academic Technology Expo (ATXpo 2021), Berkeley, CA.

Manika Kapoor* and David C. Anastasiu. A data-driven approach for detecting autism spectrum disorders, 2018. Poster presented by Manika Kapoor at the 2018 Grace Hopper Celebration conference, Houston, TX.

David C. Anastasiu. Teaching with jupyter in-class activities: Lessons learned and next steps, 2018. Poster presented at the 20th CSU Symposium on University Teaching, held at Cal-Poly Pomona, Pomona, CA.

Ragwa M. El Sayed*, Rathna Ramesh*, Alessandro Bellofiore, David C. Anastasiu, and Melinda Simon. Patient friendly kidney function screening, 2018. Poster presented by Ragwa M. El Sayed at the National Kidney Foundation 2018 Spring Clinical Meeting, Austin, TX.

Software

Alex Whelan* and David C. Anastasiu. Spr. `https://github.com/davidanastasiu/spr`, 2024.

Yanhong Li* and David C. Anastasiu. Dan. `https://github.com/davidanastasiu/dan`, 2024.

Yanhong Li* and David C. Anastasiu. Seed. `https://github.com/davidanastasiu/seed`, 2023.

Marcus Chavez*, Sean Leininger*, Joseph Pham Nguyen*, and David C. Anastasiu. Flowview: Web application for model training, inference, and visualization of stream flow and reservoir water level predicted levels in santa clara county. `https://flowview.info/`, 2023.

Arpita Vats* and David C. Anastasiu. Retailcounter: Enhancing retail checkout through video inpainting, yolov8 detection, and deepsort tracking. `https://github.com/davidanastasiu/retailcounter`, 2023.

Arpita Vats* and David C. Anastasiu. Key point-based naturalistic driving action recognition (kndar). `https://github.com/davidanastasiu/kndar`, 2022.

Yanhong Li* and David C. Anastasiu. Necplus. `https://github.com/davidanastasiu/NECPlus`, 2022.

David C. Anastasiu and George Karypis. Tapnn: Efficient identification of tanimoto nearest neighbors. `http://davidanastasiu.net/software/tapnn/`, 2016.

David C. Anastasiu and George Karypis. L2knng: Fast exact k-nearest neighbor graph construction with l2-norm pruning. `http://davidanastasiu.net/software/l2knng/`, 2015.

David C. Anastasiu and George Karypis. Orion: Multivariate resource utilization time series evolution analysis. http://davidanastasiu.net/software/orion/, 2014.

David C. Anastasiu and George Karypis. L2ap: Fast cosine similarity search with prefix l-2 norm bounds. http://davidanastasiu.net/software/l2ap/, 2014.

Other

Mohammed Shaiqur Rahman, Archana Venkatachalapathy, Anuj Sharma, Jiyang Wang, Senem Velipasalar Gursoy, David Anastasiu, and Shuo Wang. Synthetic distracted driving (syndd1) dataset for analyzing distracted behaviors and various gaze zones of a driver, 2022.

Under Submission

David C. Anastasiu and Gheorghi Guzun. Active learning through competition on continuous improvement problems, 2023. *Manuscript submitted for publication.*

Alex Whelan*, Ragwa Elsayed, Alessandro Bellofiore, and David C. Anastasiu. Selective partitioned regression for accurate kidney health monitoring.

Teaching Experience

Department of Computer Science and Engineering, Santa Clara University, Santa Clara, CA.
- COEN 11, ADVANCED PROGRAMMING   Fall 2019, Fall 2020
- COEN 140, DATA MINING/MACHINE LEARNING   Fall 2020, Spring 2022, Spring 2023
- COEN 145, INTRODUCTION TO PARALLEL COMPUTING   Spring 2020, Fall 2022, Fall 2023
- COEN 240, MACHINE LEARNING   Spring 2020
- COEN 281, PATTERN RECOGNITION AND DATA MINING   Winter 2020, Winter 2021, Spring 2022
- COEN 319, PARALLEL PROGRAMMING   Spring 2020, Fall 2021, Fall 2022
- COEN 342, DEEP LEARNING   Spring 2021, Winter 2022, Winter 2023, Winter 2024

Department of Computer Engineering, San José State University, San José, CA.
- CMPE 126, ALGORITHMS AND DATA STRUCTURE DESIGN   Spring 2019
- CMPE 138, DATABASE SYSTEMS I   Spring 2017
- CMPE 139, FUNDAMENTALS OF DATA MINING   Fall 2017
- CMPE 213, PARALLEL COMPUTING   Spring 2018
- CMPE 239, WEB AND DATA MINING   Spring 2017, Fall 2016
- CMPE 255, DATA MINING   Spring 2019, Spring 2018, Fall 2017

Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN.
- Guest Lecturer, INTRODUCTION TO DATA MINING   Fall 2015
- Teaching Assistant, PRINCIPLES OF DATABASE SYSTEMS   Fall 2013, Fall 2012
- Teaching Assistant, DATA MODELING   Spring 2013

Department of Computer Science, Texas State University, San Marcos, TX.
- Teaching Assistant, DATA MINING AND INFORMATION RETRIEVAL   Spring 2011

Providence Academy, Lostine, OR.

- High School Teacher, MATHEMATICS, CHEMISTRY 2001—2002

Membership in Professional Organizations

- Institute of Electrical and Electronics Engineers (IEEE).

- Association for Computing Machinery (ACM).

- Sigma Xi, The Scientific Research Honor Society.

Student Supervision, Santa Clara University

PhD Students Supervised

- Yinghua Qin. *Large Language Models*. 2023–Present.

- Alex Whelan. *Kidney Health Monitoring*. 2022–Present.

- Yanhong Li. *Hydrologic Flow Prediction*. 2021–Present.

- Sarah Anjum. *Medical Imaging*. 2021–Present.

- Paramdeep Singh. *Nearest Neighbor Search*. 2021–Present.

Graduate Students Supervised

- Griffin Ellis. *Protein Corona Prediction*, 2023–Present.

- Ridham Kachhadiya. *Color Constancy*, 2022–Present.

- Raghav Kapoor. *Machine Learning on Smart-NICs*. 2022–Present.

- Arpita Vats. *Competitive Active Learning, Traffic Analytics*. 2021–Present.

- Taylor Downey. *Hydrologic Flow Prediction, Antibiofilm Peptide Identification*. 2020-2021.

- Jay Janodia. *Vehicle Counting Using IoT Devices*. 2021.

- Rinkle Rani. *Hydrologic Flow Prediction*. 2020.

- Emma Lewis. *Open Modification Search*. 2019–2020.

- Juliana Shihadeh. *Kidney Health Prediction*. 2019–2020.

PhD Committees

- Yuan Wang. Computer Science and Engineering. Santa Clara University.

Thesis/Project Committees

- Antonio Fontan, Andy Xiao. *Naturalistic Driving Action Recognition*. 2023-2024.

- Amy Ha, Anthony Bryson, Vincent Zhou. *Corridor Counting*. 2023-2024.

- Arda Arkan, Jared Griffith, Ishan Sandhu. *Detecting Violation of Helmet Rule for Motorcyclists*. 2023-2024.

- Eerina Haque, Eric Huang, Sihang Li. *Multi-Camera People Tracking*. 2023-2024.

- Griffin Ellis. *Automated Retail Checkout by Computer Vision*. 2022-2023.

- Joseph Pham Nguyen, Marcus Chavez, Sean Leininger. *FlowView: A Web Application for Hydrologic Predictions*. 2022-2023. **Computer Science and Engineering Award of Technical Excellence**.

- Paul Le. *Kidney Health Prediction Android App*. 2022-2023.

- Jonathan Nguyen, Andrew Bredar. *EasyCompost: Using Augmented Reality to Improve Reality*. 2021-2022.

- Raghav Kapoor, Casey Nguyen. *Neural Network Interpretability for Autonomous Driving Neural Networks.* 2021-2022.

- Rishabh Chittaranjan. *Online Multivariate Time Series Analysis.* 2021-2022.

- Dorian Clay. *Expanding NeuroSymbolic Artificial Intelligence for Video Visual Question Answering.* 2021-2022.

- Sukruth Krishnakumar, Jacob Lucke, Audrey Hou. *Smart Refrigerator.* 2020–2021.

- Jay Ladhad, Colin Rioux, Maggie Dong, Chris Tian, Donovan Allen. *Multi-Intersection IoT-based Vehicle Counting.* 2020–2021.

- Ethan Paek, Tyler Niiyama, Justin Liu, Spencer Tsang, Kent Ngo, Jackson Tseng. *Multi-Camera Single-Intersection IoT-based Vehicle Counting.* 2020–2021.

- Claire Bushnell, Katherine Becknell, Rachel Fitzsimmons, Emily Sumner. Co-advised with Yuling Yan. *Deep Learning-Based Aneurysm Identification in MRI Scans.* 2020–2021.

- Michael Canniffe, Jack Kitts, Wei Wang. *Weight Lifting Feedback.* 2019–2020.

External Thesis Panel Member

- **Nivedha Balakrishnan**. San José State University. Department of Biomedical Engineering. *Prediction of Thrombin Inhibitor Peptides from Diverse Habitats Using Machine Learning.* Co-advised with Anand Ramasubramanian. 2021–2022.

- **Samuel Sciberras**. University of Malta. Department of Artificial Intelligence. *Land Vehicle Speed Estimation in an Augmented Reality Space using Computer Vision.* Aug 2021.

- **Bipasa Bose**. San José State University. Department of Biomedical Engineering. Bipasa Bose. *Prediction of Novel Antibiofilm Peptides from Diverse Habitats Using Machine Learning.* Co-advised with Anand Ramasubramanian. 2019–2020.

Undergraduate Students Supervised

- Jared Griffith. *Color Constancy.* 2023–Present.

- Sudhish Sewpaul. *Nearest Neighbor Search Library.* 2021–2022.

- Soham Phadke. *Kidney Health Monitoring.* 2021-2022.

- Uma Bahl. *Streaming Multivariate Time Series Analysis.* 2020-2021.

- Jonathan Nguyen. *Nearest Neighbor Search Library.* 2019-2021.

- Jack Gaul. *Single Camera-Based Vehicle Counting.* 2019-2020.

- Sukruth Krishnakumar. *Open Modification Search.* 2019-2020.

High School Students Supervised

- Apoorva Kulshreshtha, Niranjana Sankar. *Color Constancy for CKD prediction.* Summer 2023.

- Tiffany Tang, Kevin Zhou. *Corridor Vehicle Counting.* Summer 2023.

- Meha Gaba, Priyanka Ganguly, Prajval Sharma, Adit Singh, Advait Phadke. *Luminosity and Background Effects on Test Strip Prediction.* Summer 2022.

- Maria Vazhaeparambil, Jolie (Jiayi) Wang. *Vehicle Counting Using IoT Devices.* Summer 2021.

- Maria Vazhaeparambil, Meha Gaba, Prajval Sharma. Project lead to a published journal paper. *Vehicle Counting Survey.* Summer 2020.

Student Supervision, San José State University

Graduate Students Supervised

- Saranya Soundar Rajan. *Effect of Neighborhood Approximation on Downstream Analytics.* Masters of Science in Software Engineering. Recipient of the *Davidson Student Scholar Award* for our work. Defended Jun 21, 2019.

- Naga Boddapati. *Fast Open-Modification Spectral Library Search using Effective Exact Nearest Neighbor Search.* Masters of Science in Computer Engineering. Recipient of the *Davidson Student Scholar Award* for our work.

- Shuai Hua. *Vehicle Tracking Based on Historical Intersection Over Union.* Masters of Science in Computer Engineering. Defended Oct 26, 2018.

- Manika Kapoor. *A Data-Driven Approach for Detecting Autism Spectrum Disorders.* Masters of Science in Software Engineering. Defended Oct 18, 2018. Recipient of the *Davidson Student Scholar Award* for our work.

- Rathna Ramesh. *Machine Learning Methods for Kidney Disease Screening.* Masters of Science in Computer Engineering. Defended Jun 22, 2018. Recipient of the *Davidson Student Scholar Award* for our work.

- Swapnil Gaikwad. *Optimal Constrained Wireless Emergency Network Antennae Placement.* Masters of Science in Computer Science. Defended May 7, 2018.

Thesis/Project Committees

- Charles MacKay. *Learning for Free: Training Object Detectors with Synthetic Data.* Masters of Science in Computer Science. Defended May 17, 2019.

- Ardalan Razavi. *Combining Blockchain and Swarm Robotics to Deploy Surveillance Missions.* Masters of Science in Software Engineering. Defended Mar 29, 2019.

- Xiaohong Lui. *Data Visualization: An Exploration of Dimension Reduction.* Masters of Science in Statistics. Defended Dec 18, 2017.

- Avni Gulati. *Social Recommendation Systems.* Masters of Science in Software Engineering. Defended Oct 29, 2018.

- Ragwa Elsayed. *Patient Friendly Kidney Function Screening.* Masters of Science in Biomedical Engineering. Defended Apr 20, 2018.

Undergraduate Students Supervised

- Alexander Richards. *Nearest Neighbor Search.* Recipient of the university *Undergraduate Research Award* for our work.

- Lukasz Juraszek. *Chemical Compound Clustering.* Recipient of the university *Undergraduate Research Award* for our work.

Professional Service

National Science Foundation

- *Proposal Review Panelist.* CISE. III. 2022.

- *Ad-hoc Proposal Reviewer.* Division of Mathematical Sciences (MPS/DMS). 2022.

- *Ad-hoc Proposal Reviewer.* Division of Computing and Communication Foundations (CCF). 2021.

- *Proposal Review Panelist.* Computer & Information Science & Engineering (CISE). Division of Information Integration and Informatics (III). 2020.

- *Proposal Review Panelist.* CISE. III. 2019.

- *Proposal Review Panelist.* CISE. III. 2019.

- *Proposal Review Panelist.* CISE. III. 2019.

- *Proposal Review Panelist.* CISE. Division of Computer and Network Systems (CNS). 2018.

- *Proposal Review Panelist.* CISE. III. 2017.

- *Ad-hoc Proposal Reviewer.* CISE. III. 2017.

Department of Energy
- *Ad-hoc Proposal Reviewer.* 2021.

- *Ad-hoc Proposal Reviewer.* 2017.

Alfred P. Sloan Foundation
- *Ad-hoc Proposal Reviewer.* 2022.

Journal Guest Editor
- *Machine Learning Journal.* DSAA'2024 Journal Track. 2024.

- *Information Systems.* Special Issue on Managing, Mining and Learning in the Legal Data Domain. MML4LegalData. Elsevier. 2020-2022.

Major Professional Leadership Positions
- Co-Organizer and Student Forum Co-Chair at the 11th IEEE International Conference on Data Science and Advanced Analytics (DSAA 2024).

- Co-Organizer and Undergraduate Consortium Co-Chair at the 2024 SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2024).

- Co-Organizer and Social Events Co-Chair at the The 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2024).

- Co-Organizer and Undergraduate Consortium Co-Chair at the 2023 SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2023).

- Co-Organizer and Undergraduate Consortium Co-Chair at the 2022 SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2022).

- Founding Member and Evaluation Chair, AI City Challenge and AI City Challenge Workshops at the Conference on Computer Vision and Pattern Recognition (CVPR), 2017 – present.

Conference/Workshop Organizer or Chair
- *Session Chair.* KDD 2023. The 2023 SIGKDD Conference on Knowledge Discovery and Data Mining.

- *Publicity/Media Co-Chair.* IEEE DSAA 2023. The 10th IEEE International Conference on Data Science and Advanced Analytics.

- *Session Chair.* SDM 2021. The 2021 SIAM International Conference on Data Mining.

- *Publicity Chair.* IEEE DSAA 2020. The 7th IEEE International Conference on Data Science and Advanced Analytics.

- *Session Chair.* CIKM 2018. The 28th ACM International Conference on Information and Knowledge Management.

- *Organizer*. VOICE Big Data and Artificial Intelligence Workshop, San José State University. Fall 2017.

- *Session Chair*. iDSC 2017. The 1st International Data Science Conference.

- *Publicity Co-Chair*. IEEE SmartWorld Congress 2017.

- *Workshop/Tutorial Chair*. CBDCom 2017. The 3rd IEEE International Conference on Cloud and Big Data Computing.

- *Session Chair*. IEEE DSAA 2016. The 3rd IEEE International Conference on Data Science and Advanced Analytics.

Senior Program Committee Member

- SDM24. The 2024 SIAM International Conference on Data Mining.

- KDD 2023. The 29th ACM SIGKDD Conference of Knowledge Discovery and Data Mining.

- KDD-UC 2023. The 2nd ACM KDD Undergraduate Consortium.

- CIKM 2023. The 33rd ACM International Conference on Information and Knowledge Management.

- KDD 2022. The 28th ACM SIGKDD Conference of Knowledge Discovery and Data Mining.

- KDD-UC 2022. The 1st ACM KDD Undergraduate Consortium.

- CIKM 2022. The 32nd ACM International Conference on Information and Knowledge Management.

- DSAA 2019. The 6th IEEE International Conference on Data Science and Advanced Analytics.

Program Committee Member

- PAKDD 2024. The 28th Pacific-Asia Conference on Knowledge Discovery and Data Mining.

- WWW 2024. The Web Conference 2024 (formerly International World Wide Web Conference).

- AAAI 2023. The Thirty-Seventh AAAI Conference on Artificial Intelligence.

- SDM23. The 2023 SIAM International Conference on Data Mining.

- PAKDD 2023. The 27th Pacific-Asia Conference on Knowledge Discovery and Data Mining.

- WWW 2023. The Web Conference 2023 (formerly International World Wide Web Conference).

- WSDM 2023. The 16th ACM International Conference on Web Search and Data Mining.

- AAAI 2022. The Thirty-Sixth AAAI Conference on Artificial Intelligence.

- WWW 2022. The Web Conference 2022 (formerly International World Wide Web Conference).

- PAKDD 2022. The 26th Pacific-Asia Conference on Knowledge Discovery and Data Mining.

- SDM22. The 2022 SIAM International Conference on Data Mining.

- HiPC 2022. The 29th IEEE International Conference on High Performance Computing, Data, and Analytics.

- GHTC 2022. The 2022 IEEE Global Humanitarian Technology Conference.

- BigData 2021. The 2021 IEEE International Conference on Big Data.

- KDD21. The 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining.

- CIKM 2021. The 31st ACM International Conference on Information and Knowledge Management.

- SDM21. The 2021 SIAM International Conference on Data Mining.

- PAKDD 2021. The 25th Pacific-Asia Conference on Knowledge Discovery and Data Mining.

- BDS 2021. The 7th IEEE International Conference on Big Data Computing Service and Machine Learning Applications.
- iDSC 2021. The 4th International Data Science Conference.
- CIKM 2020. The 30th ACM International Conference on Information and Knowledge Management.
- KDD20. The 26th ACM SIGKDD Conference of Knowledge Discovery and Data Mining.
- PAKDD 2020. The 24th Pacific-Asia Conference on Knowledge Discovery and Data Mining.
- SDM20. The 2020 SIAM International Conference on Data Mining.
- BDS 2020. The 6th IEEE International Conference on Big Data Computing Service and Machine Learning Applications.
- iDSC 2020. The 3rd International Data Science Conference.
- KDD19. The 25th ACM SIGKDD Conference of Knowledge Discovery and Data Mining.
- SDM19. The 2019 SIAM International Conference on Data Mining.
- PAKDD 2019. The 23rd Pacific-Asia Conference on Knowledge Discovery and Data Mining.
- iDSC 2019. The 2nd International Data Science Conference.
- CIKM 2018. The 28th ACM International Conference on Information and Knowledge Management.
- SDM18. The 2018 SIAM International Conference on Data Mining.
- PAKDD 2018. The 22nd Pacific-Asia Conference on Knowledge Discovery and Data Mining.
- DesWeb 2018. The 9th International Workshop on Data Engineering Meets the Semantic Web. An ICDE Workshop.
- CIKM 2017. The 27th ACM International Conference on Information and Knowledge Management.
- WWW 2017. The 26th International World Wide Web Conference.
- IEEE SCI 2017. The 2017 IEEE Conference on Smart City Innovations.
- DesWeb 2017. The 8th International Workshop on Data Engineering Meets the Semantic Web. An ICDE Workshop.
- iDSC 2017. The 1st International Data Science Conference.
- IEEE BigDataService 2017. The 3rd IEEE International Conference on Big Data Service and Applications.
- CIKM 2016. The 26th ACM International Conference on Information and Knowledge Management.
- IDSC 2015. The 2nd International Conference on Data Science.
- CIKM 2015. The 25th ACM International Conference on Information and Knowledge Management.

Journal Reviewer
- ACM Transactions on Knowledge Discovery from Data (TKDD).
- ACM Transactions on Parallel Computing (TOPC).
- ACM Transactions on Intelligent Systems and Technology (TIST).
- IEEE Transactions on Knowledge and Data Engineering (TKDE).
- IEEE Transactions on Services Computing (TSC).
- IEEE Sensors.

- SIAM Journal on Scientific Computing (SISC).

- Springer Journal on Knowledge and Information Systems (KAIS).

- Springer Machine Learning Journal.

- Springer International Journal of Data Science and Analytics (JDSA).

- Springer Journal on Data Mining and Knowledge Discovery (DAMI).

- Springer Social Network Analysis and Mining (SNAM).

- Elsevier Journal of Parallel and Distributed Computing (JPDC).

- Elsevier Information Sciences (INS).

- Elsevier Information Systems (INFOSYS).

Book Manuscript Reviewer
- Jones & Bartlett Learning. Spring 2019.

University Service, Santa Clara University

Committees
- Faculty Council. School of Engineering. 2022-2025.

- Tenure Track Faculty Search committee. Department of Computer Science and Engineering. 2020–2021.

- Engineering Diversity and Inclusion Council. School of Engineering. Spring 2021 & Fall 2021.

Service to the University
- Co-Director, WAVE High Performance Computing (HPC) Center. 2023–Present.

- New Faculty Orientation Near Peer Panel. Fall 2020 & Fall 2021.

- Spearheaded an effort to enhance the WAVE HPC by upgrading its interconnect to a state-of-the-art 100 Gbps InfiniBand network.

Service to the School of Engineering
- Grad Programs Summer Workshop. Summer 2023.

- School of Engineering Open House. Fall 2022.

- Engineering Diversity and Inclusion Council member. Winter 2021 & Fall 2021.

- Virtual Open House mock course. *Introduction to Machine Learning*. Fall 2021.

- Virtual Engineering Seminar (VES). Fall 2020 & Fall 2021.

- Summer Engineering Seminar (SES). Summer 2021.

Service to the Department of Computer Science and Engineering
- Organized Tenure Track/Assistant Professor lunches. 2021–present.

- Co-wrote ABET self-study chapter. Fall 2022.

- Senior Design Coordinator. Winter 2022.

- Open House volunteer. Fall 2020.

- Proposed a new course, *COEN 342: Deep Learning*, initially offered Spring 2021.

- Revamped *COEN 145/319: Parallel Computing* in 2019, which had not been taught in some years.

University Service, San José State University

Committees

- University committee on graduate studies and research. 2018-2019.
- University curriculum committee for the Masters in Data Analytics. 2018-2019.
- University committee for establishing a Masters Degree in Data Analytics. 2017-2018.
- College graduate studies committee. 2018-2019.
- College research committee. 2017-2018.
- Department graduate studies committee. *Committee Chair*. 2018-2019.
- Department faculty recruitment committee. 2017-2018.
- Department graduate studies committee. 2017-2018.
- Department undergraduate curriculum committee. 2016-2017.

Service to the College of Engineering

- Constructed the first College of Engineering high-performance computer (HPC) from scratch as an invaluable research and teaching resource for faculty and students. The HPC is equipped with 20 compute nodes, 15 GPU nodes with NVIDIA P100 cards, 1 condo node with 2 NVIDIA P100 GPUs, 0.66 PB Lustre parallel file storage, and 56 Gbps Infiniband FDR interconnect. First students gained access to the cluster in Feb 2018.
- Led project to develop *RTrack*, an application used at the college level to track faculty research activity (grants, journal, and conference papers).

Service to the Department of Computer Engineering

- Co-advising the Society for Computer Engineering (SCE) student club with Kaikai Liu. 2018—2019.
- Department web site coordinator. 2018—2019.
- Helped organize a "research" section of *CMPE 295: Capstone Project* aimed at encouraging research-related projects and publications.
- Created a new course, *CMPE 255: Data Mining*, initially taught by me Fall 2017.
- Proposed a new course, *CMPE 258: Deep Learning*, initially taught Fall 2017.
- Revamped *CMPE 213: Parallel Computing*, which had not been taught since 2007.
- Co-authored a proposal to establish a Data Science specialization as part of the Master of Science in Software Engineering degree in the Department of Computer Engineering, presented to the graduate curriculum committee and approved in December 2016.