ISTeC
The Information Science and Technology Center

COLORADO STATE UNIVERSITY

Distinguished Lectures
Fall 2018

Colorado State University's Information Science and Technology Center (ISTeC) presents two lectures by

Dr. Vipin Kumar
Regents Professor and William Norris Endowed Chair
Department of Computer Science and Engineering
University of Minnesota

ISTeC Distinguished Lecture

In conjunction with the Department of Electrical and Computer Engineering, and Department of Computer Science Seminar Series

"Big Data in Climate and Earth Sciences: Challenges and Opportunities for Machine Learning"

Wednesday, September 19, 2018
Reception with refreshments: 10:30 a.m.
Lecture: 11:00 a.m.-12:00 noon
Morgan Library Event Hall

Department of Electrical and Computer Engineering and Department of Computer Science Seminar Series Sponsored by ISTeC

"Physics Guided Machine Learning: A New Paradigm for Modeling Science and Engineering Problems"

Wednesday, September 19, 2018
Lecture: 2-3 p.m.
Morgan Library Event Hall

Abstracts

Big Data in Climate and Earth Sciences: Challenges and Opportunities for Machine Learning

The climate and earth sciences have recently undergone a rapid transformation from a data-poor to a data-rich environment. In particular, massive amount of data about Earth and its environment is now continuously being generated by a large number of Earth observing satellites as well as physics-based earth system models running on large-scale computational platforms. These massive and information-rich datasets offer huge potential for understanding how the Earth's climate and ecosystem have been changing and how they are being impacted by humans actions. This talk will discuss various challenges involved in analyzing these massive data sets as well as opportunities they present for both advancing machine learning as well as the science of climate change in the context of monitoring the state of the tropical forests and surface water on a global scale.

Physics Guided Machine Learning: A New Paradigm for Modeling Science and Engineering Problems

Physics-based models of dynamical systems are often used to study engineering and environmental systems. Despite their extensive use, these models have several well-known limitations due to incomplete or inaccurate representations of the physical processes being modeled. Given rapid data growth due to advances in sensor technologies, there is a tremendous opportunity to systematically advance modeling in these domains by using machine learning (ML) methods. However, capturing this opportunity is contingent on a paradigm shift in data-intensive scientific discovery since the "black box" use of ML often leads to serious false discoveries in scientific applications. Because the hypothesis space of scientific applications is often complex and exponentially large, an uninformed data-driven search can easily select a highly complex model that is neither generalizable nor physically interpretable, resulting in the discovery of spurious relationships, predictors, and patterns. This problem becomes worse when there is a scarcity of labeled samples, which is quite common in science and engineering domains.

This talk makes a case that in real-world systems that are governed by physical processes, there is an opportunity to take advantage of fundamental physical principles to inform the search of a physically meaningful and accurate ML model. Even though this will be illustrated in the context of two problems in modeling water quality, the paradigm has the potential to greatly advance the pace of discovery in a number of scientific and engineering disciplines where physics-based models are used, e.g., power engineering, climate science, weather forecasting, materials science, computational chemistry, and biomedicine.

Speaker Biography

Vipin Kumar is a Regents Professor and hold the William Norris Chair in the Department of Computer Science and Engineering at the University of Minnesota. His research interests include data mining, high-performance computing, and their applications in Climate/Ecosystems and health care. He is currently leading an NSF Expedition project on understanding climate change using data science approaches. He has authored over 300 research articles, and co-edited or coauthored 10 books including the widely used text book "Introduction to Parallel Computing", and "Introduction to Data Mining". Kumar has served as chair/co-chair for many international conferences and workshops in the area of data mining and parallel computing, including 2015 IEEE International Conference on Big Data, IEEE International Conference on Data Mining (2002), and International Parallel and Distributed Processing Symposium (2001). Kumar is a Fellow of the ACM, IEEE, AAAS, and SIAM. Kumar's research has been honored by the ACM SIGKDD 2012 Innovation Award, which is the highest award for technical excellence in the field of Knowledge Discovery and Data Mining (KDD), and the 2016 IEEE Computer Society Sidney Fernbach Award, one of IEEE Computer Society's highest awards in high performance computing.

To arrange a meeting with the speaker, please contact Dr. Imme Ebert-Uphoff, iebert@colostate.edu.

Upcoming Distinguished Lectures

October 15

"How 2G Computational Social Science Can Revolutionize the Study of 'Dark' Networks"

11 a.m.-12 noon

Morgan Library Event Hall
Dr. Jytte Klausen

October 22

"Laws for Cybersecurity?"

11 a.m.-12 noon

Morgan Library Event Hall
Dr. Fred B. Schneider

ISTeC (Information Science and Technology Center) is a university-wide organization for promoting, facilitating, and enhancing CSU's research, education, and outreach activities pertaining to the design and innovative application of computer, communication, and information systems. For more information please see ISTeC.ColoState.edu.