Research Experience for Undergraduates
Summer School on Mathematical Foundation of Data Science

June 2, 2025 --- July 11, 2025

LeConte 101
Department of Mathematics
University of South Carolina
or
Join Virtual Zoom Program
https://sc-edu.zoom.us/j/89463183663?pwd=WgVNqseRvUDYa3zyPCd15a79IgfNeb.1

Meeting ID: 894 6318 3663
Passcode: 024115

Sponsored by:
Department of Mathematics
University of South Carolina

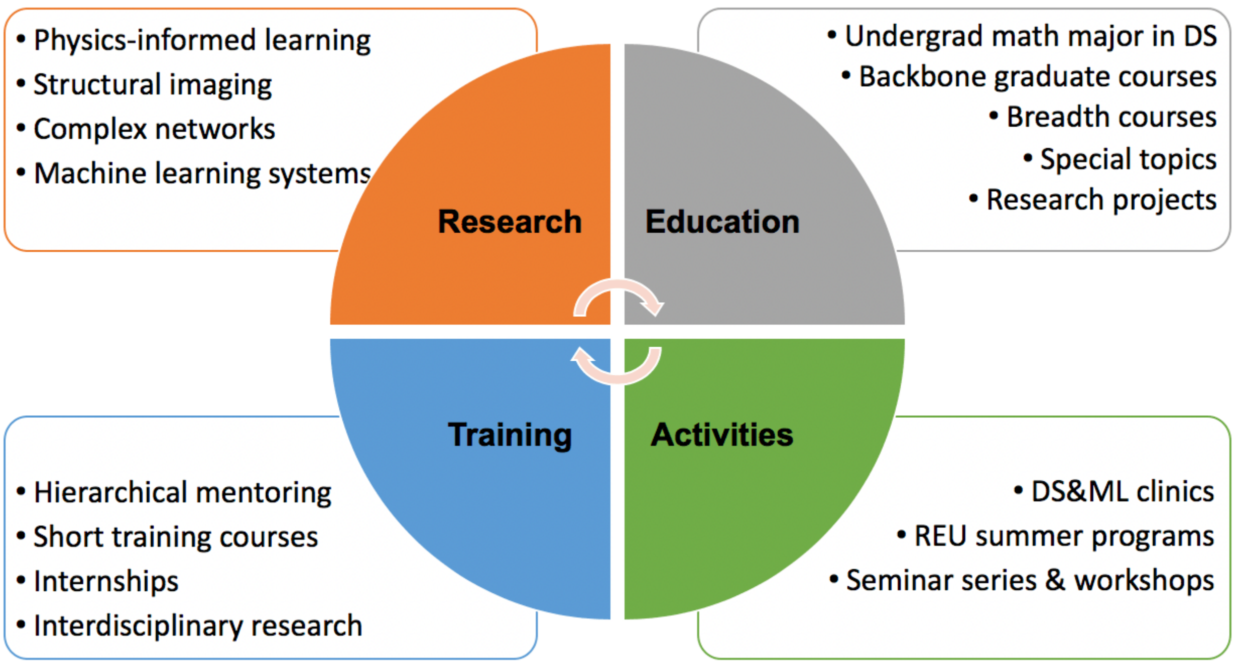
National Sciences Foundation
RTG award DMS 2038080

Organized by:

Prof. Linyuan Lu,
Prof. Qi Wang,
Prof. Peter Binev,
Prof. Pooyan Jamshidi,
Prof. Ray Bai
Prof. Changhui Tan
Prof. Siming He

Section 1: Program Overview

This REU summer program is part of the NSF RTG project **"RTG: Mathematical Foundation of Data Science at University of South Carolina"**, which aims to develop a multi-tier Research Training Program at the University of South Carolina (UofSC) designed to prepare the future workforce in a multidisciplinary paradigm of modern data science. The education and training models will leverage knowledge and experience already existing among the faculty and bring in new talent to foster mathematical data science expertise and research portfolios through a vertical integration of post-doctoral research associates, graduate students, undergraduate students, and advanced high school students. A primary focus of this project is to recruit and train U.S. Citizens, females, and underrepresented minority (URM) among undergraduate and graduate students, and postdocs through research led training in Data Science.



For more information on the NSF RTG project, please visit us at the following URL: https://sc.edu/study/colleges_schools/artsandsciences/mathematics/my_mathematics/rtg/index.php. The REU summer program of this year runs in-person and virtually from June 6 to July 15. Starting from the first week, students will be divided into several groups to work on research projects. Some guest speakers are invited to give talks on the latest development in the Mathematical Foundation of Data Science. On the last day of the program, students will present their research findings.

Section 2: Research Projects

2.1 Developing AI tools for Scanning Tunneling Microscopy (STM) (Advisor: Peter Binev)

Scanning Tunneling Microscopy (STM) is used to observe the atomic structure of the surface of the materials. Data acquisition tasks in STM usually require a human-in-the-loop component to constantly evaluate the structural phenomena on the surface and navigate towards the ones of interests. Using AI to take care of most of these evaluations would allow a rapid increase in materials research. The first step in achieving this is the development of data processing tools and deep learning routines to recognize the usual structures and classify them. The research project involves investigating and training deep learning autoencoders to extract pattern information from artificial and real data, as well as using this information to advance different image processing tasks.
Further information about the research area can be found in last year's project report https://arxiv.org/pdf/2501.13283.

2.2 AI Tools for Digital Twins for Health (Advisor: Qi Wang)

Biomedical applications are ideal implementation venues for AI models and tools. With AI integration, biomedical research and healthcare deliveries could experience a transformational change.The project goal is to develop biomedical AI models, data processing and computational infrastructures to support the development and implementation of AI-enabled digital twins for health in the form of software. Several fundamental theories/foundation models, methods, and tools are developed to facilitate research projects in this area. The REU component will include adding front-end user-interfaces for some basic models for specific diseases so that the models can be used in the clinical setting. The three major tasks in the current digital twin project consist of data imputation and synthetic data generation, multiclass classification, and time-series forecasting. The student will work with the existing research group to implement the already developed models and modules. New models may be developed for enhance the existing digital twin platform.

2.3 From Language to Action: Distributed LLM-Based Robot Control and Planning (Advisor: Pooyan Jamshidi, Abir Hossen)

The topic I would like you to explore would be the synergy between LLM and Robotics. Specifically, task planning and execution by:

- Translating natural language commands into robot actions,
- Multi-modal task planning using visual and textual information.

To achieve the above, we would explore distributed LLM strategy. This distributed framework would be the outcome of your internship by demonstrating: (i) a robot will go to a location, (ii) a manipulator will pick and place an object, and (iii) combination of (i) and (ii), by taking natural language command. Example natural language command: "Robot, go to the kitchen and bring the coffee mug from the coffee machine to the living room table if there is coffee in the mug."

The project would be mostly developing a unified software. We may use technologies such as Ollama, Cake, ROS, Python. The goal would be to make the framework software and hardware agnostic. For this, we may implement APIs using Flask.

2.4 Spatiotemporal Modeling for Disease Mapping and Causal Inference

(Advisor: Ray Bai)

A critical task in public health is disease mapping, which aims to analyze the distribution and spread of diseases across time and space. This can help epidemiologists and policymakers to identify disease patterns and predict areas that are at elevated risk of future outbreaks. Bayesian spatiotemporal modeling is widely used in practice for disease mapping. In addition to predicting disease prevalence, these models also help policymakers to identify county-level characteristics (e.g. percentage of residents who are uninsured) that are highly associated with disease prevalence. However, association does not imply causation. If we want to study whether a particular county-level characteristic *causes* disease prevalence to either increase or decrease, then we need to bring in causal inference.

In this project, participants will use Bayesian spatiotemporal models to predict disease prevalence of a disease of their choice (e.g. an infectious disease or a chronic disease) at the county level in the United States. If time allows, we will also extend the model to infer the causal effect of a treatment on the disease prevalence.

2.5 Simulating Transport and Diffusion in Physical and Biological Systems

(Advisor: Changhui Tan/Siming He)

This REU project explores how transport (directed movement) and diffusion (random spreading) interact in systems governed by partial differential equations (PDEs). Such models arise in many contexts. For example, to describe the spread of biological populations (Fisher-KPP equations) or the phase separation of materials (Allen-Cahn equation).
However, when external transport effects are incorporated, the qualitative behavior of solutions can change substantially, influencing the speed of propagation, pattern formation, and solution stability. Students will develop MATLAB code to simulate the PDEs numerically. The project involves:

- Learning basic PDE knowledge about the transport and diffusion phenomena;
- Developing numerical schemes to solve these equations;
- Running numerical experiments to explore how transport modifies diffusion-driven behavior.

This project suits students interested in scientific computing, numerical modeling, and applying data-driven approaches to real-world systems.

2.6 - Ricci Curvature on Graphs

(Advisor: Linyuan Lu)

Throughout the last thirty years, extensive efforts have been made to adapt the concept of Ricci curvature from Riemannian geometry to a discrete framework. This adaptation enables the treatment of graphs similarly to manifolds. Curvature has been applied in various research areas such as network analysis, quantum computation, dynamic Networks, etc. In this project, we aim to explore the characteristics of graphs exhibiting positive Lin-Lu-Yau Ricci curvature.

Section 3: Program Calendar

Week 1

Day		Activity	Moderator
Monday June 2	9:00-12:00	Introduction to the REU program	
	12:00-2:00	Lunch break	
	2:00-5:00	Group assignments	
Tuesday June 3	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Wednesday June 4	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Thursday June 5	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Friday June 6	9:00-10:30	Parallel research sessions	
	10:30-12:00	Group presentation	
	12:00-2:00	Lunch break	
	2:00-5:00	Social/free time	

Week 2

Day		Activity	Moderator
Monday June 9	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Tuesday June 10	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Wednesday June 11	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Thursday June 12	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Friday June 13	9:00-10:30	Parallel research sessions	
	10:30-12:00	Group presentation	
	12:00-2:00	Lunch break	
	2:00-5:00	Social/free time	

Week 3

Day		Activity	Moderator
Monday June 16	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Tuesday June 17	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Wednesday June 18	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Thursday June 19	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Friday June 20	9:00-10:30	Parallel research sessions	
	10:30-12:00	Group presentation	
	12:00-2:00	Lunch break	
	2:00-5:00	Social/free time	

Week 4

Day		Activity	Moderator
Monday June 23	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Tuesday June 24	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Wednesday June 25	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Thursday June 26	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Friday June 27	9:00-10:30	Parallel research sessions	
	10:30-12:00	Group presentation	
	12:00-2:00	Lunch break	
	2:00-5:00	Social/free time	

Week 5

Day		Activity	Moderator
Monday June 30	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Tuesday July 1	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Wednesday July 2	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Thursday July 3	9:00-10:30	Parallel research sessions	
	10:30-12:00	Group presentation	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Friday July 4	Holiday	No activities	

Week 6

Day		Activity	Moderator
Monday July 7	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Tuesday July 8	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Wednesday July 9	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Thursday July 10	9:00-12:00	Parallel research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Parallel research sessions	
Friday July 11	9:00-10:30	Parallel research sessions	
	10:30-12:00	Joint research sessions	
	12:00-2:00	Lunch break	
	2:00-5:00	Final presentations and group photo	

Section 4: Contact Information

Faculty

Linyuan Lu	lu@math.sc.edu	https://people.math.sc.edu/lu/
Peter Binev	binev@math.sc.edu	https://people.math.sc.edu/binev
Qi Wang	Qwang@math.sc.edu	https://people.math.sc.edu/wangq/
Pooyan Jamshidi	PJAMSHID@cse.sc.edu	https://pooyanjamshidi.github.io/
Ray Bai	RBAI@mailbox.sc.edu	https://raybai.net/
Changhui Tan	TAN@math.sc.edu	https://people.math.sc.edu/tan/
Siming He	SIMING@mailbox.sc.edu	https://he-siming.wixsite.com/mysite

Postdoctoral Researcher

William Linz	wlinz@mailbox.sc.edu
George Brooks	ghbrooks@email.sc.edu

Graduate Assistants

Teegan Bailey	teeganb@email.sc.edu
Jasdeep Singh	jasdeep@email.sc.edu
Sehwan Kim	sehwank@email.sc.edu
Henry Simmons	has11@email.sc.edu
Nizam Hasanov	nhasanov@email.sc.edu

Undergraduate Students

Aidan	Van Voorhis	vanvoora@email.sc.edu	University of South Carolina	In-Person
Angelina	Piccolino	piccola@email.sc.edu	University of South Carolina	Virtually
Aydyn	Rohm	arohm@email.sc.edu	University of South Carolina	In-Person
Evan	Funderburg	ewf@email.sc.edu	University of South Carolina	Virtually
Ibrahim	Hussain	i.hussain2304@gmail.com	Boston University	In-Person
Lea	Frost	lkfrost@email.sc.edu	University of South Carolina	Virtually
Mark	Ritchie	RITCHIED@email.sc.edu	University of South Carolina	In-Person
Nathanael	Johnson	nathan.erikson.9701@gmail.com	The Ohio State University	In-Person
Niyathi	Kukkapalli	nk6074@princeton.edu	Princeton University	Virtually
Nolan	Jones	jonesnt@email.sc.edu	University of South Carolina	In-Person
Ronald	Smith	Ronaldrs@email.sc.edu	University of South Carolina	Virtually
Sai	Kottapali	kottapas@email.sc.edu	University of South Carolina	Virtually
Samatha	Criswell	criswes@email.sc.edu	University of South Carolina	In-Person
Sara Shree	Ramadoss Muthuselvam	muthuses@email.sc.edu	University of South Carolina	Virtually
Shahyad	Khamnei	shahyadkhamnei1@gmail.com	Boston University	Virtually
Wynton	Vaughn	wyntonv@g.clemson.edu	Clemson University	In-Person
Zachary	Valsecchi	zvv@email.sc.edu	University of South Carolina	In-Person
Emmy	Sitnik	esitnik@email.sc.edu	University of South Carolina	In-Person