

CS661: BIG DATA VISUAL ANALYTICS

Course Description

The necessity of visual analytics capabilities for big data is becoming omnipresent due to its significant demand in the current age of data science and analytics. Interactive data visualization techniques enable us to comprehend and explore diverse types of complex data sets efficiently so that patterns and features from the data can be readily identified and studied in detail. As the data grows larger and become intricate, it poses significant challenges to manage, curate, and explore such large data sets. These data sets can come from various scientific simulations as well as from social media, IoT, various sensors, and many other industry and application domains. In this course, we will cover a comprehensive view of data visualization techniques with a focus on the techniques that are suitable for big data. We will discuss the theory and foundations of visualization techniques and have hands-on exercises on visualizing different types of data sets using available visualization software and libraries. We will study scientific and information visualization techniques with a focus on data compression, statistical and information theory techniques, and selected high-dimensional visualization algorithms. Next, we will discuss how modern machine learning and deep learning techniques are adapted for big data visual analytics. Finally, we will learn about exascale visual computing and state-of-the-art in situ analysis techniques and conclude by discussing the future paradigms of the big data visual analytics domain. The contents for this course will be based on research papers from top-tier journals and conferences such as IEEE TVCG, CGF, ACM CHI, IEEE/ACM Supercomputing, IEEE Visualization, EuroVis and EuroGraphics, IEEE Pacific Visualization, IEEE LDAV, EGPGV, etc.

Course Syllabus

Index	Module	Topics Covered
1	Fundamentals of Data Visualization	Introduction to Visualization and Visual Analytics
		Foundations of Data Visualization, Visual Abstractions, Visual Variables, Various types of Data
2	Scientific Visualization (SciVis)	Big Data Characteristics, Data Reduction, Various Data Models; Visualization Pipeline
		Scientific Visualization Software such as VTK, ParaView, etc.
		Linear Interpolation; Isosurface Algorithm; Volume Rendering Algorithm
3	Information Visualization (InfoVis)	Fundamentals of Information Visualization, Software for Information Visualization
		High Dimensional Data Visualization Techniques
4	Big Data Visualization Techniques	Statistical Data Modeling and Visualization
		Information Theory Techniques for Visualization
		Time-varying Data, Ensemble Data, and Uncertainty Visualization
5	ML/DL for Visual Computing of Large Data	ML/DL for Visualization and Visual Analytics
		Applications of ML/DL to Big Data Visualization, Data Modeling, Data Understanding
		Visual Analytics and Explainability of ML/DL Models
6	Advanced Topics	Extreme-scale Data Analytics, Exascale Computing, Parallel and High-performance Visualization
		In Situ Analysis & Visualization and Future Paradigms

Course Team

Instructor's Name: Prof. Soumya Dutta

Web: <https://soumyadutta-cse.github.io/>

Email: soumyad@cse.iitk.ac.in

Class Venue

RM-101, MW, 2:00pm-3:15pm

Course Grading Policy

Category	Split
Quiz	15%
Programming Assignments	30%
Mid Semester Exam	25%
Final Semester Project	30%

Textbooks and References

- Data Visualization: Principles and Practice by Alexandru C. Telea, CRC Press.
- Visualization Analysis and Design by Tamara Munzner, A K Peters Visualization Series, CRC Press.
- The Visualization Handbook edited by Charles D. Hansen and Chris R. Johnson.
- In Situ Visualization for Computational Science, Editors: Hank Childs, Janine C. Bennett, Christoph Garth, Springer publication.
- Research papers provided during the class to cover selected topics.

Course Audience

PhD, Masters, 3rd and 4th year UG Students