Data Visualization

Course Rationale

Today quantitative and symbolic data are easily collected in computer for- mat, from databases, websites, smart devices, and anything that has interconnect capabilities. When such large amounts of data are put in spreadsheets or tabular reports, it becomes difficult to see the patterns, structure, trends, or relationships inherent in the data. Effective data visualization exposes these inherent relationships, consolidating and illustrating them in graphics.

Course Description

Visualization organizes data in a way that the structure and relationships in the data that may not be so easily understood becomes easily understood and interpreted with the visualization. Visualizations of a data set give the reader a narrative that tells the story of the data.

The purpose of data visualization is to convey information contained in data to clearly and efficiently communicate an accurate picture of what the data says through understandable and context appropriate visualizations.

To do a visualization can be just exploratory or entails using Machine Learning techniques that determine the structure of the data. The visualization is then matched to the data structure.

The course will explore how principles of information graphics and design and how principles of visual perception, can be used with machine learning techniques to make effective data visualizations.

Each student will make a presentation of some principles of data visualization or do a visualization project.

The course is open to PhD students in all programs. Non-computer science students will be paired with computer science students for the visualization project.

Topic List

The topic list may include but is not limited to:

- – Why visualizations
- – Basics about visualizations (marks and channels)

- Visualization Techniques

  - – Pie and Donut Charts
  - – Histograms
  - – Scatter Plots
  - – Heat Maps
  - – Matrix Diagrams
  - – Candlestick Charts
  - – Bubble Charts
  - – Graphs and Networks
  - – Alluvial Diagrams
  - – Dendrograms
  - – Ring Charts
  - – Tree Diagrams
  - – Treemaps
  - – Scatterplot matrices
  - – Parallel Coordinate Displays
  - – Multidimensional data visualizations
  - – Time Series
  - – Line Charts
  - – Cartograms and Choropleths
  - – Dot Distribution Maps

- Visualization Issues

  - – Visualization Tools: Vega lite, D3.js, R Shiny

Learning Goals

- Be able to describe the key design guidelines and techniques used for the visual display of information

- Understand how to best use the capabilities of visual perception in a graphic display

- Understand the principles of interactive visualizations

- Understand how Machine Learning techniques can determine data structure and pattern

- Explore and critically evaluate a wide range of visualization techniques and applications

Assessment

Every student will do a project involving a presentation of the project at the end of the course.