

Evaluation

CMPS 7010 Research Seminar

Why Evaluation

- Understand the **strength** and **weakness** of your solution
 - Is your solution useful in **practice**?
 - Performance: worst case vs. average case
 - Resource requirements: time, space, communication, energy, ...
 - Sensitivity to model assumptions
 - **When** is your algorithm better (worse) than **other** algorithms?
 - Are the empirical results consistent with the theoretical analysis?
 - A gap may indicate an opportunity for improving algorithm and/or analysis

How to Evaluate Your Solution

- Get good datasets
 - Synthetic data: need to have a good coverage
 - Real data: public data vs. data collected by yourself
 - Mixed data
- Choose baselines
 - Previous solutions: code available?
 - Optimal solutions: e.g., exhaustive search for small problem instances
 - New heuristics

How to Evaluate Your Solution

- Obtain robust results
 - generate enough instances (via randomization) for each scenario
 - give confidence intervals
- Balance coverage and relevance
 - vary experiment settings to generate enough scenarios
 - show most relevant results
- Document everything
 - As you may need them for presentation/journal paper/dissertation

Evaluation Methods

- Numerical study
- Computer simulation/emulation
- Testbed experiment
- Real-world controlled experiment
- User study
- ...

Build a toolbox

- Programming: languages, coding environments, ...
- Data processing and analysis
- Image/figure creation
- Simulation tools and testbeds
- Large-scale evaluation platforms
- ...

Have a plan

- How many figures to plot and the content of each
 - What data are needed and where to get them
 - What baselines to compare with and how to implement them
- Identify bottlenecks
 - in your code and experiment process
- Automate the process
 - You can't work for 24×7, but machines can
 - Parallelize the job
 - Enable remote access to computing resources

Large Scale Evaluation

- [Tulane HPC](#)
- Public cloud platforms
 - Amazon Web Services
 - Microsoft Azure
 - ...
- Open testbeds
 - [Chameleon](#)
 - [COSMOS](#)
 - ...

Evaluation Section

- Setup
 - Summarize objective and method
 - Discuss data collection process, baselines, parameter settings, etc.
- Results
 - Organize results **logically**
 - Use tables and figures efficiently
 - Discuss observations and insights
- Discussion
 - Indicate relevance to reality
 - Acknowledge weakness

Teamwork

- Collaborative Paper Writing
 - Dropbox
 - Overleaf
 - GitHub
- Work with your collaborators
 - Lead your project
 - Be responsible
 - Develop good communication skills

About the Final Paper

- Use a conference template of your choice
- Complete all the sections (including the abstract)
- At least one result figure
- At least one table (e.g., a summary of symbols used)