CSE 302LR: Introduction to Experiential Learning/Research (Section B) - Fall 2025

Lecture 4: Plotting graphs using matplotlib

9/16/2025

University at Buffalo
Department of Computer Science and Engineering
School of Engineering and Applied Sciences

Clemen 217, 3:30 pm - 4:50 pm. In-person attendance required.

Find more on course website:
https://cse.buffalo.edu/adblab/elrr/

Plotting graphs

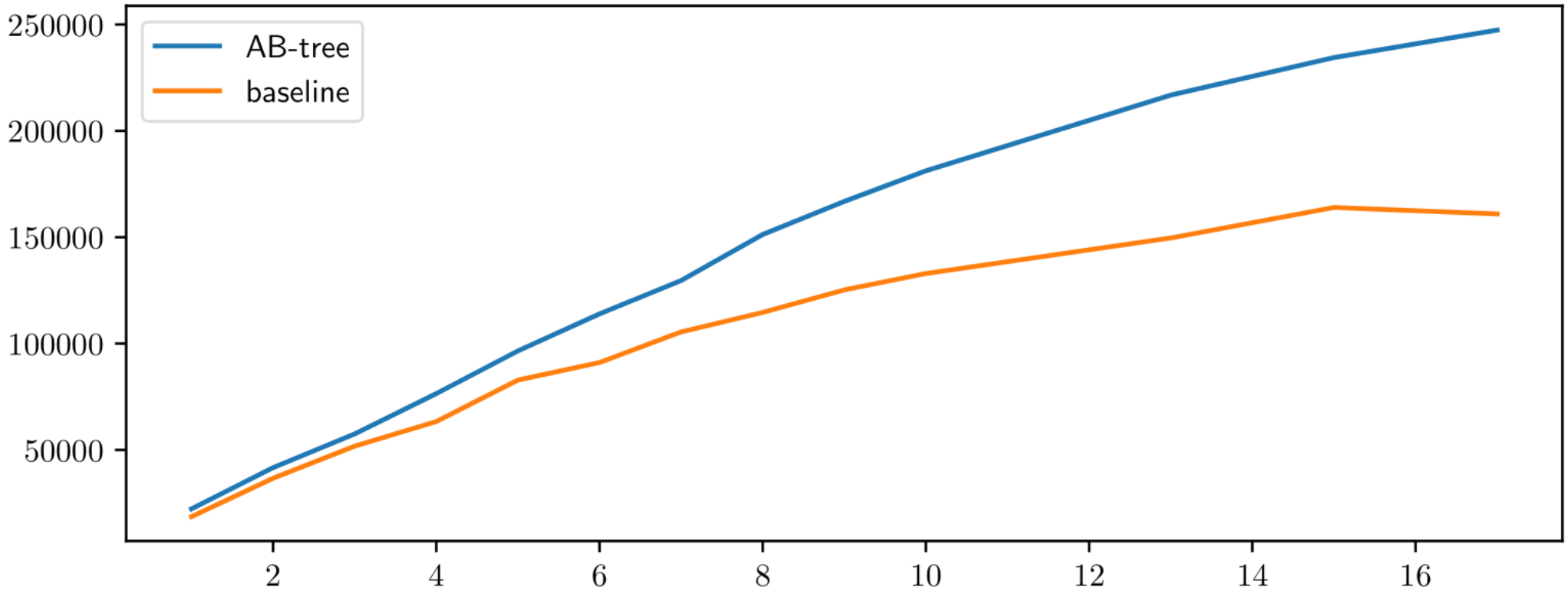
- Experiment graphs
  - visualizes data from experimental results
  - helps convey key information

Example: system scalability under concurrency

- To demonstrate the scalability of a system in terms of data ingestion
  - Ab-tree vs baseline
  - Download: https://cse.buffalo.edu/adblab/elrr/fa25/plot/
    - Starter code: `throughput.py`
    - Data: `abt_ingestion_throughput.csv`
    - Your task: load data and generate `throughput.pdf`
    - Hint: to run the script

```
chmod +x throughput
./throughput.py
```

A not very readable graph

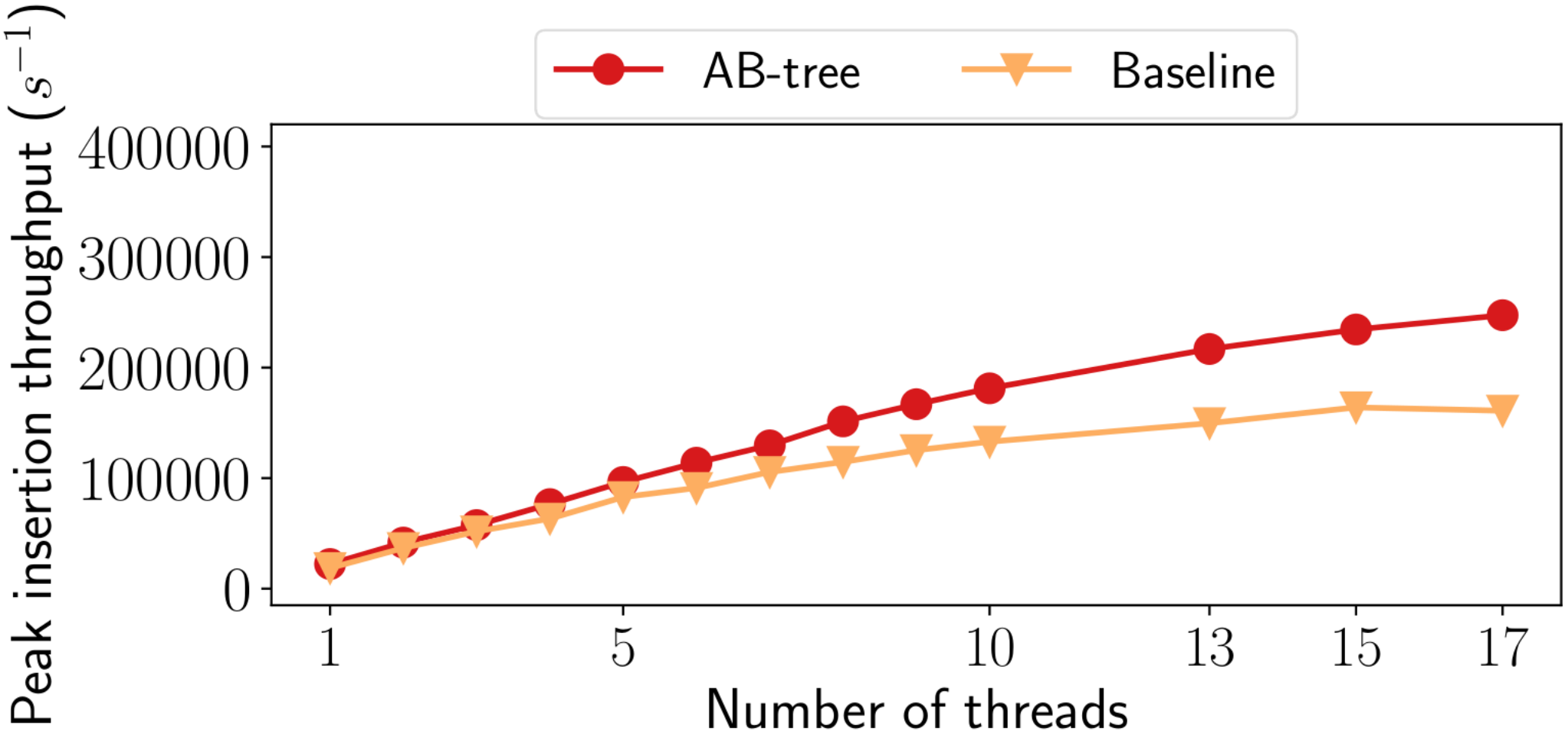


Things to consider when plotting graphs

- What're the key messages we'd like to send?
- Are we (accidentally) hiding any information?
  - Often time they are scrutinized by your peer reviewers.
- Is the graph misleading?
- Is the graph accessible?
  - Font size, color blindness friendly colors, color contrast in grayscale printing

How to make the graph readable?

- How to make the following graph instead?



Is the graph good enough?

- There might be other hidden facts we fail to show in the graph, for instance,
  - Peak throughput vs average throughput?
    - You might want to throw out the warm up and cool down phases
  - Does the trend change when we have more ingestion threads?
    - The test machine had 18 cores but 36 threads (hyperthreading)
  - Update the graph following one of the above two directions
    - More data in `rawdata/`
  - Demo your graph at 4:30 pm