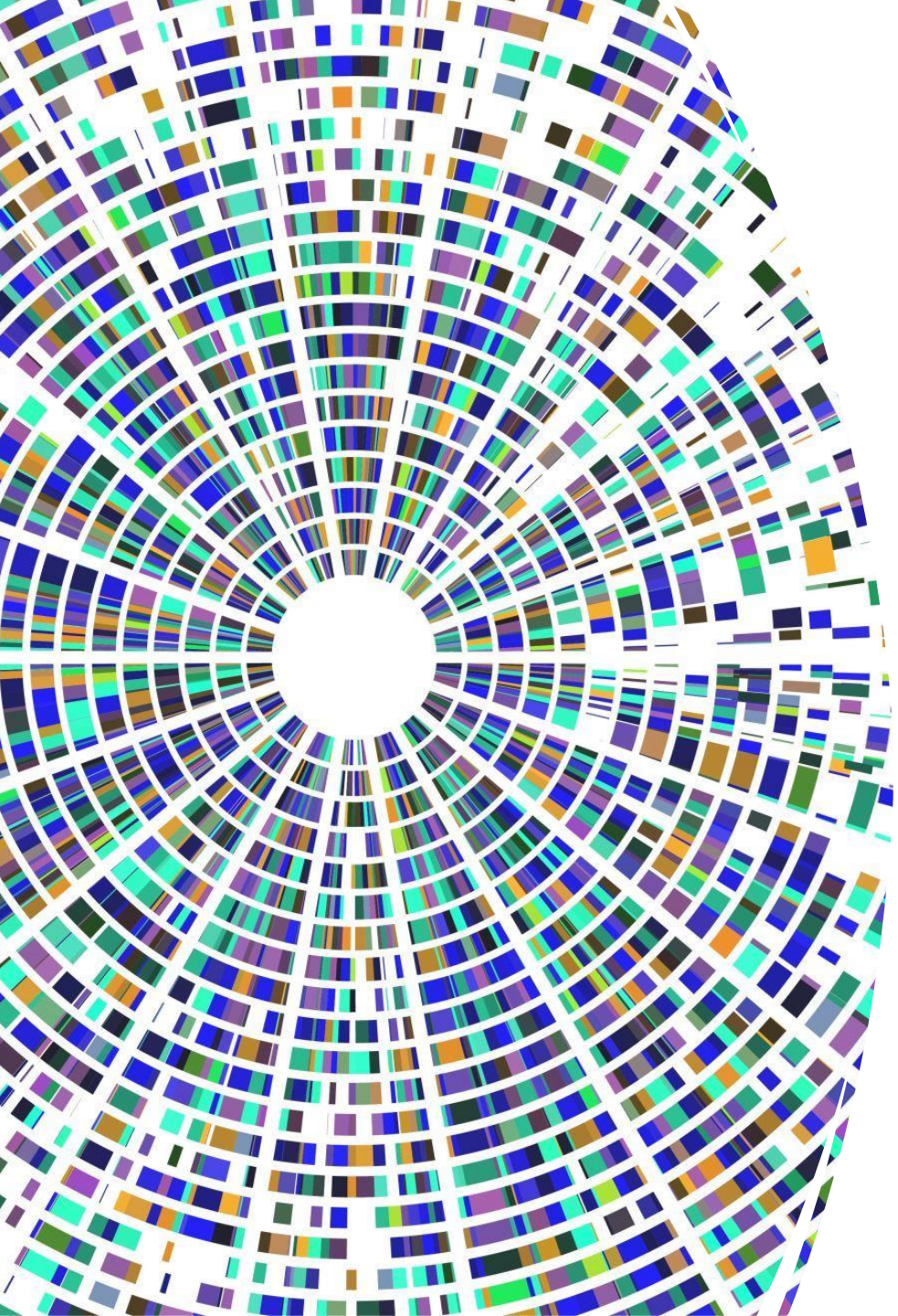


Visualization and Quantitative Comparisons of Sorting Algorithms for Large Data Sets

Shriya Dhaundiyal
Mariah Maynard
Aditya Puttigampala
Dmitrii Troitskii

Northeastern University

Problem

What are the tradeoffs between different sorting algorithms in an industry setting?

Which algorithms are best applied to large datasets?

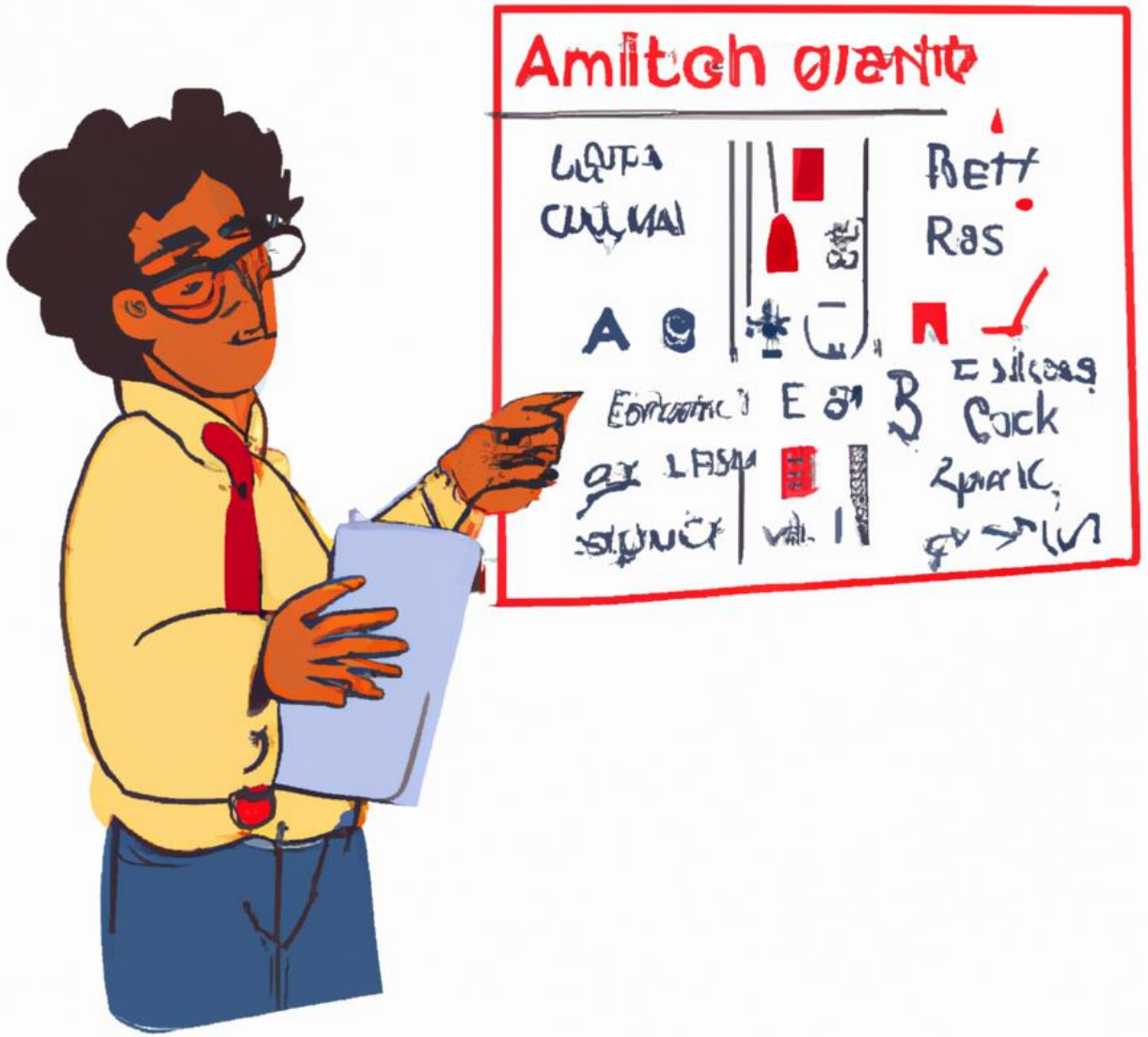
How to improve students' understanding of **sorting** algorithms?

How does the choice of an algorithm affect performance in the real world?

Team Motivations

Aditya

Understand tradeoffs during cost/benefit analysis phase of product development



Shriya

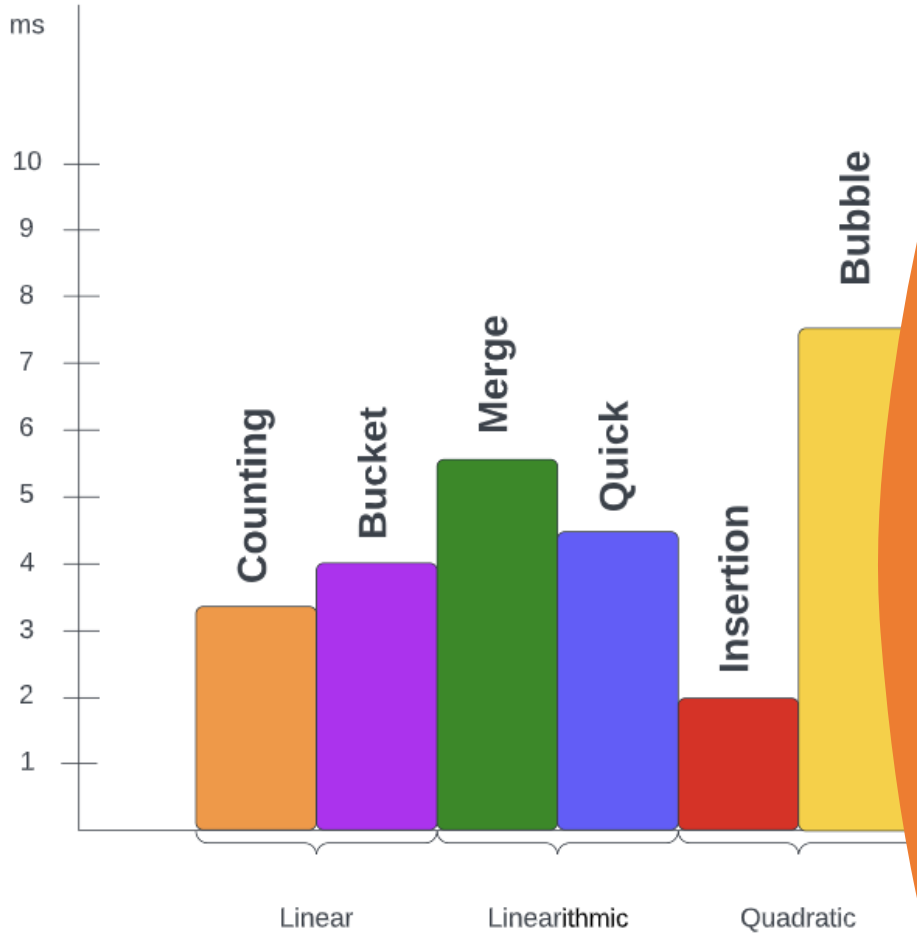
Stemming from the motivation for better categorization of patient information in the medical field which can be extrapolated for sorting all sorts information for all industries.

Mariah

Create an educational tool to help developers in training

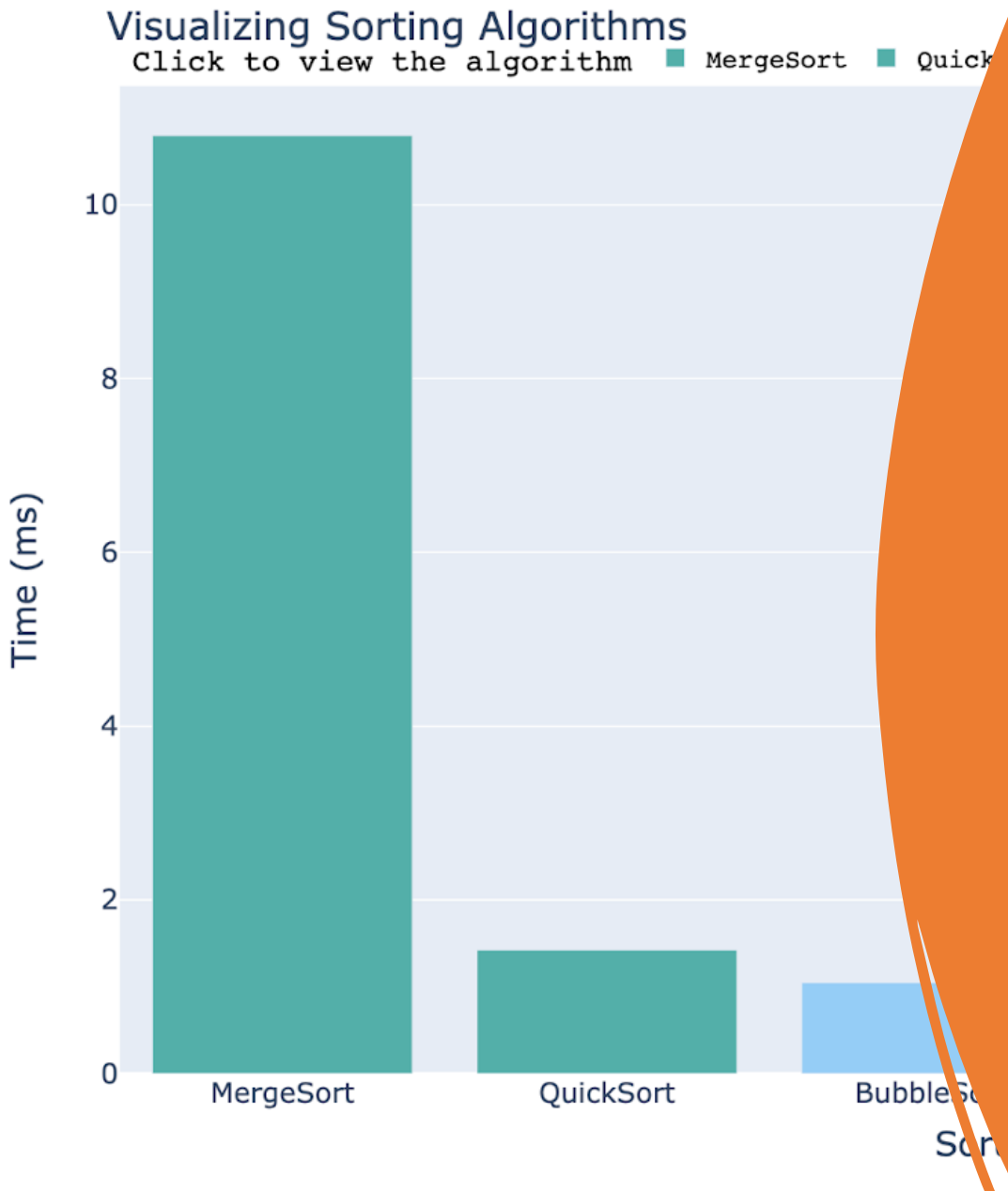
Dmitrii

Make development decisions more quickly and choose the right algorithm for the job



Idea

- Measure performance of different algorithms on the same input
- Visualize the difference in performance to easily compare algorithms
- Filter by algorithm and drill into run-time statistics for each
- Run the tool on large datasets



Results

- The tool works with 6 sorting algorithms, grouped into 3 classes by their runtime complexity
- Benchmark graph shows performance of each algorithm ran on the same data
- Large datasets can used as an input via a .csv file
- For 4 of 6 algorithms we visualized their runtime logic

Conclusion

We were able to successfully design and implement a sorting algorithm program that visual learners could use as an educational tool.

Future improvements include:

- Increasing the number of sorting algorithms
- Adding a GUI
- Creating more intricate visualizations for each sorting algorithm