

Assignment 6: Suffix Index

Computational Genomics

Objectives: Become familiar with using different suffix index data structures to search large strings

Tasks:

1. Accept the assignment at <https://classroom.github.com/a/IXHCJRSn>
2. Clone the repository
3. In `src/suffix_trie.py` implement
 - a. `build_suffix_trie`
 - b. `search_trie` which returns the prefix overlap length
4. In `src/suffix_tree.py` implement
 - a. `search_tree` which returns the prefix overlap length
5. In `src/suffix_array.py` implement
 - a. `Build_suffix_array` (HINT: modify BFS)
 - b. `search_array` which returns the prefix overlap length (HINT: modify binary search)
6. Experiment with the size and speed of these different data structures.
7. Create figures that demonstrate the results of your experiments.
8. Update `README.md` and create a `doc/suffix_index.tex` (or similar) to include your new experiments.
9. Push your final code to GitHub.
10. Submit your final PDF to Canvas.