

Predicting RNAcompete binding
based on an amino acid sequence

Project in
Deep Learning in
Computational Biology

Administrative announcements

- The submission deadline: 2nd July, 10:00

Implementation

- Input: the 1st argument is the RBP filename, and the 2nd is the RNAcompete filename
- `python main.py <ofile> <RBP> <RNCMPT>`
- Output file to include: scores for sequences in RNAcompete file in the same order (textual file, each line one number)
- Training time performance will be measured
- Reasonable documentation

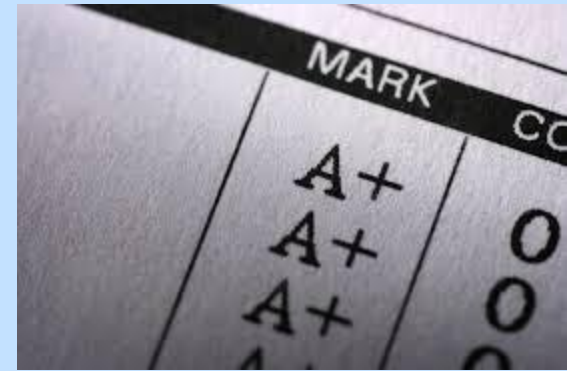
Submission

- 1. Electronic design document
- 2. Electronic code submission
- Submit by deadline
- 44 scored RNAcompete files, e.g., RBP201.txt
- 3. All score files zipped in one zip file
- Python code for runtime test
- Submit via moodle

Design document

- 3-5 pages (pdf), Hebrew/English
- Briefly describe main goal, input and output of program
- Describe neural net architecture, chosen parameters, and pre- and post-processing of the data, parameter search and training, results on training data, performance (time, memory, CPU)

Grades



MARK	CO
A+	0
A+	0
A+	0

- 10% HW assignments submission (no grading)
- ~~20% final exam (must pass)~~
- 90% final project:
 - 10% for implementation (modularity, clarity, documentation)
 - 10% for efficiency: $\max(\min(1, 2 - \text{time}[s]/3600), 0)$
 - 40% for accuracy: $10\% * f(r, k) + 30\% * (\text{score} / \text{max_score})$
 - $f(r, k) = r/k$, $k = \# \text{projects}$, $r = \text{rank (k at the top)}$
 - 10% for final report
- HW assignments and final exam are individual
- Project submission is be in pairs

Grading

- 30% accuracy = $\text{Corr} / \max(\text{Corr})$
 - Average Corr over 44 submitted RNCMPT files
 - $\max(\text{Corr})$ = the best team's average Corr
- 10% rank among projects / total #projects
- 10% run-time = $\max(\min(1, 2\text{-time}/36000), 0)$ in total
 - Training time on 200 training RBPs in seconds
- 10% final report
- 10% for implementation (modularity, clarity, documentation)