

# Lab I

---

- Find a motif from a set of strings using EM and gibbs sampling
  - input in a text file
    - k: the length of motifs
    - t: the number of string in a input file
    - m: max length of each string
    - t strings
  - output in a text file
    - t motifs
    - a profile with t motifs found
- Due on 2018.9.23 11:59 PM
- You should submit a code and a readme file including usages to portal

# Lab I: an input example

---

7

10

100

atgaccgggatactgataaaaaaaagggggggggcggtacacattagataaacgtatgaagtacgttagactcggcgccgccg  
acccctatTTTTgagcagatttagtgacctggaaaaaaaatttgagtacaaaactttccgaataaaaaaaaaaggggggga  
tgagtatccctgggatgacttaaaaaaaaggggggggtgctctcccgattttgaatatgtaggatcattcgccaggggtccga  
gctgagaattggatgaaaaaaaaggggggggtccacgcaatcgcgaaaccaacgcggacccaaaggcaagaccgataaaggaga  
tccctttgcggtaatgtgccgggaggctggttacgtagggaagccctaacggacttaataaaaaaaaggggggggcttatag  
gtcaatcatgttcttgtgaatggatttaaaaaaaagggggggggaccgcttggcgcacccaaattcagtggtggcgagcgcaa  
cggttttggcccttgtagaggcccccgtaaaaaaaaggggggggcaattatgagagagctaatactatcgcggtgcgtgttcat  
aacttgagttaaaaaaaaggggggggctggggcacatacaagaggagtcttccttatcagttaatgctgtatgacactatgta  
ttggcccatggctaaaagcccaacttgacaaatggaagatagaatccttgcataaaaaaaaggggggggaccgaaagggaag  
ctggtgagcaacgacagattcttacgtgcattagctcgcttccggggatctaatagcacgaagcttaaaaaaaagggggggga

## Lab 1: an output example

aggtcgg

acgtcgg

tcgtcgg

tcgtcgg

acgccgg

acgtctg

acgtcgt

acgtcgc

acttcgg

aggtcgg

0.2222 0.1122 0.2211 0.1122 0.1122 0.1122 0.1122

0.2222 0.1122 0.2211 0.1122 0.1122 0.1122 0.1122

0.2222 0.1122 0.2211 0.1122 0.1122 0.1122 0.1122

0.2222 0.1122 0.2211 0.1122 0.1122 0.1122 0.1122

0.2222 0.1122 0.2211 0.1122 0.1122 0.1122 0.1122

0.2222 0.1122 0.2211 0.1122 0.1122 0.1122 0.1122

0.2222 0.1122 0.2211 0.1122 0.1122 0.1122 0.1122

0.2222 0.1122 0.2211 0.1122 0.1122 0.1122 0.1122

0.2222 0.1122 0.2211 0.1122 0.1122 0.1122 0.1122

0.2222 0.1122 0.2211 0.1122 0.1122 0.1122 0.1122