

Upute za izradu projekta na predmetu Bioinformatika 2018./2019.

Projekt se sastoji od pismenoga dijela i implementacije/testiranja.

Pismeni dio

U pismenom dijelu je potrebno:

- opisati zadani algoritam/problem
- objasniti isti kroz korake na jednostavnom primjeru
- prikazati rezultati testiranja (vrijeme izvođenja i količina zauzete memorije - ako nije drugačije navedeno)
- napisati zaključak
- obavezno navesti korištenu literaturu
- koristiti standardni FER obrazac za seminare

Implementacija

Zahtjeva se:

- uz sam kod potrebno je predati i upute za instalaciju i izvođenje, te testni primjer koji radi
- Kod treba biti pohranjen na github repozitoriju.

Svi članovi rade na istom repozitoriju. Naslovna strana je na engleskom, staviti link na stranice predmeta. Svako mora imati barem 10 commita na repozitorij u različitim fazama projekta. Uvjet od 10 commita se jedino ne odnosi na projekte gdje se radi testiranje različitih softvera, ali i tu treba podatke prebaciti na github.

- koristiti MIT licencu
- svaka funkcija i klasa moraju biti komentirani

Dobro pisan kod ne treba pretjerane komentare. Ako se implementiraju poznati algoritmi dovoljno je samo navesti koji se algoritam implementira. Ako je neki komad koda nerazumljiv, onda se u komentaru prije njega treba jasno objasniti što će se raditi.

- treba naznačiti tko je napisao koji dio koda
- imena varijabli i komentari su na engleskom
- kod treba biti čitljiv - rađe napisati nekoliko linija koda nego u jednoj teško razumljivoj
- projekt treba raditi na linux OS - instalirati kao dual boot ili virtualnu mashinu npr. VirtualBox

<https://www.virtualbox.org/>

- sve dodatne vanjske biblioteke moraju biti uključene u izvorni kod (ako ih nema na bio-linuxu) i moraju se instalirati automatski sa kodom te ih je potrebno navesti i referencirati u pismenom dijelu.

Alternativno, dodatne biblioteke automatski downloadati i instalirati.

- instalacija i pokretanje iz naredbenoga retka.

Ne smije se koristiti IDE (npr. Eclipse).

- svi parametri se unose jedino prilikom pokretanja programa u naredbenom retku ili iz konfiguracijske datoteke

- izlaz programa mora biti u datoteku, a ukoliko se radi o poravnanju koristiti MAF format
<http://genome.ucsc.edu/FAQ/FAQformat.html#format5>

- koristiti standardne konvenciju za programski jezik: Python PEP8 (poželjno u kombinaciji s Pylintom)
 - C++ Google styleguide
 - Java - Oracle standards
 - C - <http://cs.brown.edu/courses/cs033/docs/guides/stylePSA.pdf>
 - ...

Preporuka

- preporuča se, ali nije uvjet da sve javne klase i funkcije moraju imati Unit testove
Privatne/interne klase mogu, ali ne moraju. http://en.wikipedia.org/wiki/Unit_testing.

Testiranje performansi

- Potrebno je za svaki algoritam napraviti analizu točnosti, vremena izvođenja i utroška memorije za različite testne slučaje i programske jezike u kojima su algoritmi implementirani.
- Generirati barem 10 različitih sintetskih testnih podataka s različitim duljinama znakova. Testni slučajevi moraju obuhvaćati različite duljine nizova (**100 – 1000000 znakova**).
- Za podatke i testiranje koristite genome bakterija sa stranice <http://bacteria.ensembl.org/index.html> (npr. Escherichia coli) i wgsim alat (<https://github.com/lh3/wgsim>)
- Testovi ne smiju biti uključeni u kod.

Predavanje koda i dokumentacije

- predaja se obavlja slanjem maila nastavnicima. U tome trenutku treba biti sve gotovo. Naknadni commitovi neće biti uvaženi.

Ocjenjivanje

Kod ocjenjivanja vodit će se računa o:

- tome da li su poštovani zahtjevi iz prethodnoga poglavlja
- točnosti rješenja
- poznavanju koda (svatko treba znati objasniti svoj dio koda)
- prepisivanje gotovih rješenja automatski nije dozvoljeno i u tome slučaju projekt dobiva 0 bodova

Rok predaje

- **14. 1. 2019. u 23:59**

- Predaja se obavlja slanjem e-maila nastavnicima. U tome trenutku treba biti sve gotovo. Naknadni commitovi neće biti uvaženi.
- U dogovorenom terminu pred nastavnicima na vlastitom računalu napraviti git clone, instalirati i pokrenuti aplikaciju na testnom primjeru.
- Prezentacije projekta:
 - 2 -3 člana: 5 minuta
 - 4 - 6 članova: 7 minuta