

Практикум из програмирања 1 (13Е111ПП1, 13С111ПП1)

- Пети домаћи задатак -

Напомене:

- а)** Домаћи задатак се предаје преко **Moodle** курса предмета.
- б)** Рок за предају и термин одбране домаћег задатка ће бити благовремено објављени.
- в)** Домаћи задатак оцењује се путем јавних и тајних тестова за време предаје, израдом модификације оригиналног задатка за време одбране и усменим одговарањем за време одбране.
- г)** Домаћи задатак се решава самостално. Предметни наставници задржавају право да након предаје домаћег задатка изврше проверу сличности и предузму одговарајуће дисциплинске мере. Током израде решења није дозвољена употреба алата вештачке интелигенције заснованих на великим језичким моделима (*ChatGPT, Github Copilot* и сл.).
- д)** Програм мора да буде декомпонован на функције. Потпрограми не смеју да приступају променљивама главног програма директно, већ само путем својих аргумената и повратне вредности. Главни програм треба да позове функције које учитавају све потребне податке, затим да позове функције које врше захтевану обраду и да позове функцију за испис потребних резултата.
- е)** Приликом рада програма проверити и обрадити могуће грешке коришћењем механизма изузетака. У случају грешке приликом рада са датотеком (непостојање датотеке) на стандардни излаз исписати поруку `DAT_GRESKA`. У случају некоректног формата података унетих са стандардног улаза на стандардном излазу исписати поруку `PODACI_GRESKA`. Након исписа поруке коректно прекинути извршавање програма.

На програмском језику *Python* написати програм који на основу кључних речи врши предвиђање програмског језика у коме је написан задати програмски код. Описи програмских језика се састоје из кључних речи и задати су у датотеци `keywords.csv`. Сваки ред датотеке садржи једну кључну реч и један програмски језик у коме се она користи. Могуће је да се једна кључна реч или један програмски језик више пута нађу у оквиру датотеке. Подаци у оквиру једног реда одвојени су знаком `-`. Имена датотека која садрже програмски код за који се врши предвиђање се учитавају са стандардног улаза. Најпре је задатак укупан број датотека, а затим и имена датотека у засебним редовима. Све датотеке са програмским кодом имају екстензију `.txt`, а у случају да подаци нису у одговарајућем формату потребно је сигнализирати поруку о грешци. За сваку исправно учитану датотеку програм врши предвиђање програмског језика у коме је садржај датотеке написан и резултате предвиђања записује у датотеку `predictions.csv`. Предвиђање програмског језика се врши тако што се пронађе програмски језик за који важи да се највећи број различитих кључних речи тог језика налази у датотеци са програмским кодом. Кључна реч се налази у датотеци уколико у датотеци постоји низ карактера који идентично одговара низу карактера у кључној речи, а испред и иза кога се налази бланко знак или граница датотеке. Уколико се у датотеци налази реч која као свој део садржи кључну реч то се не сматра појавом кључне речи у датотеци. У случају да више програмских језика има исти број различитих кључних речи у задатом програмском коду, исписати све програмске језике, уређено лексикографски растуће, одвојене знаком `,`. Сваки ред излазне датотеке се односи на једну улазну датотеку и садржи име улазне датотеке, предикцију програмских језика у коме је написан програмски код који се налази у улазној датотеци и број кључних речи који је пронађен за тај програмски језик. Подаци у оквиру излазне датотеке одвојени су знаком `|`.

Програм треба да:

1. Учита кључне речи програмских језика и смести их у оквиру одговарајуће структуре података.
2. Учита и провери исправност имена датотека са стандардног улаза, а затим обради грешку уколико је до ње дошло.
3. За сваку датотеку извршити предвиђање програмског језика у коме је написан њен садржај и испише резултате предвиђања у одговарајућу датотеку.

Примери:

Стандардни улаз:	3 prog1.txt prog2.txt prog3.txt
Улазне датотеке:	keywords.csv prog1.txt prog2.txt prog3.txt
Излазне датотеке:	predictions.csv
Стандардни излаз:	

Стандардни улаз:	3 prog1.txt prog2.txt prog3.txt
Улазне датотеке:	keywords.csv #ne postoji prog1.txt prog2.txt prog3.txt
Излазне датотеке:	-
Стандардни излаз:	DAT_GRESKA

Стандардни улаз:	3 prog1.txt prog2.in prog3.txt
Улазне датотеке:	keywords.csv prog1.txt prog2.in prog3.txt
Излазне датотеке:	-
Стандардни излаз:	PODACI_GRESKA

Стандардни улаз:	2 prog4.txt prog5.txt
Улазне датотеке:	keywords.csv prog4.txt prog5.txt
Излазне датотеке:	predictions.csv
Стандардни излаз:	