

# Систем за утврђивање знања заснован на препознавању природног језика

---

Савремене образовне технологије и стандарди

## 1. Мотивација

Развој система за утврђивање знања заснованог на препознавању природног језика српског говорног подручја, представља значајан корак ка унапређењу образовања у дигиталном добу. Овај пројекат доноси са собом низ идеја које имају потенцијал да трансформишу традиционалне методе креирања, чувања и оцењивања тестова, обезбеђујући ефикаснији и прилагодљивији приступ едукацији.

Наставно особље се суочава са изузетним теретом који произлази из опсежне административне папирологије коју морају обрађивати. Поред обавеза креирања и одржавања наставе, присутно је и оцењивање великог броја тестова, које пажљиво треба прегледати и анализирати како нико од ученика не би био оштећен. Такође, одговорни су и за њихово чување на одређени временски период. Када се у обзир узме велики број ученика по разреду, то представља преплављујуће количине папира. Са друге стране, ученици често чекају дужи временски период на резултате и ограничен им је број увида урађених тестова.

Овим радом описан је систем који би аутоматизовао оцењивања тестова, олакшао њихово креирање и чувања, при чему би донео значајне административне олакшице и убрзао пружање повратних информација, ослобађајући време наставницима и ученицима да се фокусирају на суштинске аспекте учења и унапређивања. Осим свега наведеног, обезбедио би добар увид у обавезе и урађене провере знања.

Имплементација система на српском језику доприноси локализованој подршци корисницима у Србији и српскоговорним регионима, док могућност креирања тестова са више унапред понуђених одговора или слободним текстом, омогућава персонализацију оцењивања према специфичностима различитих предмета и разреда, како основних тако и средњих школа.

Кроз идеју овог пројекта, желимо да подстакнемо напредак образовних пракси у Србији, пружајући модерна технолошка решења која подржавају ефикасност, ажурност и приступачност у утрђивању знања.

## 2. Преглед стања области

У овом одељку биће предочена 3 рада тернутног стања области

Рад [1] описује истраживање, спроведено на Универзитету Хертфордшире у Великој Британији, које је анализирано употребу *ChatGPT* технологије у одговарању на питања из хемије у оквиру програма фармацеутске науке. Скуп података обухвата питања из два обавезна хемијска модула прве и друге године овог програма. Питања су категоризована помоћу Блумове таксономије како би се проценио ниво интелектуалног захтева. Додатне помоћи, као што је референца на одговарајуће изворе, коришћени су како би се побољшало генерисање одговора. Одговори су затим оцењивани према шеми оцењивања испита.

Испитивање показује да су одговори генерисани од стране *ChatGPT*-а били добро написани, али постоје одређена ограничења, посебно у вези са сложенијим питањима која захтевају дубљу анализу и примену знања. Аутори предлажу редизајнирање процене, посебно кроз коришћење проблема решавања интерпретације података или питања заснованих на студијама случаја. Ова технологија може бити корисна у образовању, али је важно разумети њена ограничења и прилагодити је потребама наставног плана.

Други рад [2] је за циљ поставио решавање недостатака постојећих метрика ослоњених на лексичку сличност, као што су ЕМ (енгл. *Exact Match*), Ф1-мера и тачност на првих  $n$  примера у евалуацији модела заснованих на питањима и одговорима (енгл. *Question-Answering model*) предстаљањем новог приступа, SAS (енгл. *Semantic Answer Similarity*). Коришћени су постојећи скупови података: *SQuAD*, *GermanQuAD* и *NQ-open*, анотирани су ручно, садржећи парове одговора са означеном семантичком сличношћу. Евалуација се вршила упоређивањем са људским судом о сличности датих одговора, а закључак је био да SAS метрика и *Cross-encoder* модел дају знатно боље резултате у односу на друге метрике у оцени семантичке сличности и да се могу применити за овај конкретан проблем.

Рад [3] се бави анализом и евалуацијом семантичких метрика сличности одговора, користећи три подскупа великих скупова података анотираних од стране три људска оцењивача, посебно се фокусирајући на теже случајеве, као што су

имена јавних личности, бројеви, датуми, итд. Коришћени скупови података су *SQuAD*, *GermanQuAD* и *NQ-open*, а затим је представљен нови скуп података имена, акцентован на јавне личности, како би се побољшала евалуација на *NQ-open* скупу. Студија поставља две главне хипотезе. Прво, тврди да метрике засноване на лексици нису погодне за аутоматску евалуацију модела заснованих на питањима и одговорима јер им недостаје појам контекста и семантике. Друго, тврди да већина метрика, као што су *BERTScore* и посебно *SAS*, иако се у раду [2] показала као одлична примена сијамске неуронске мреже, имају потешкоће у процени сличности за одређене врсте података попут имена, бројева, датума и акронима, што би значило да се анализа сличности тих података мора решити неким другим приступом иако за све остало *SAS* показује изванредне резултате.

### 3. Имплементација

У наставку је описан скуп података и архитектуре апликација које сачињавају имплементирано решење. Такође, предочена је и презентација начина рада система, при чему је акценат постављен на функционалности којима се утврђује знање ученику.

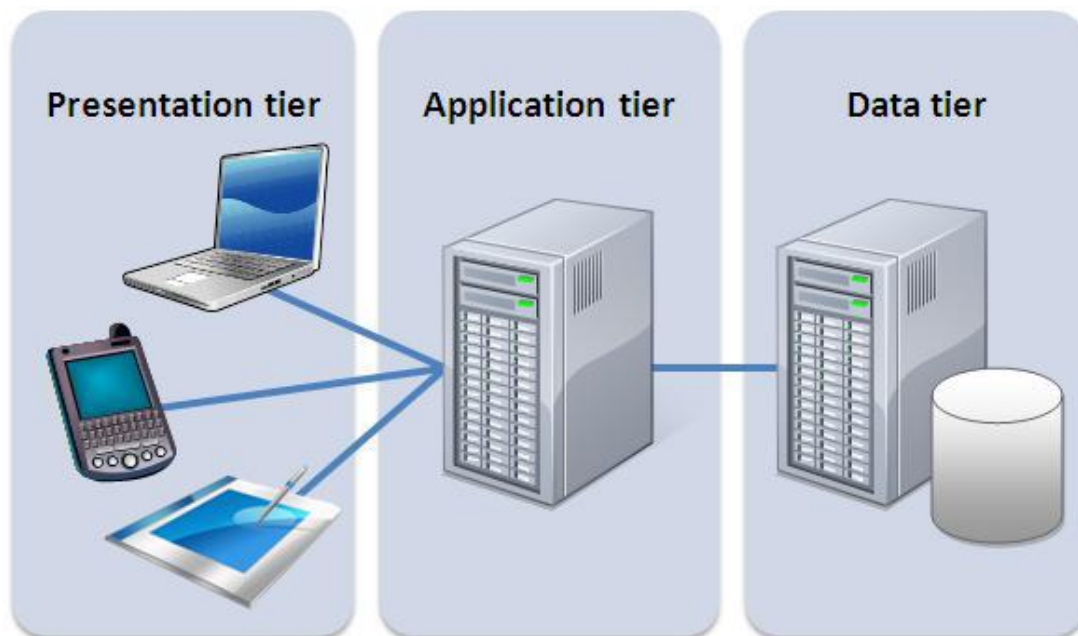
#### 3.1. Скуп података

Скуп података чине тестови из биологије Основне школе „Вук Караџић“ у Бачком Добром Пољу. Аутор поседује тест са тачним одговорима, датим од стране наставника, али такође и оцењене тестове ученика. Тестови су за узраст петог, шестог и седмог разреда, сва питања која нису била у форми слободног одговора или више унапред понуђених одговора су ручно пребачена, док је број бодова остао исти.

#### 3.2. Архитектура апликације

Апликација је имплементирана према multi-tier архитектури, прецизније *three-tier* архитектури. Таква архитектура подразумева постојање презентационог, апликативног и слоја за управљање подацима (слика 3.1.). Свака од компоненти покреће се на одвојеном серверу, чиме се постиже виши ниво флексибилности

решења приликом измена унутар слојева или додавања нових слојева. Директна комуникација постоји између презентационог и апликативног слоја, као и између апликативног и слоја за управљање подацима и обавља се према стилу клијент-сервер. *Hypertext Transfer Protocol (HTTP)* задужен је за размену порука између презентационог и апликативног слоја. Слој за управљање подацима чини *PostgreSQL* систем за управљање базама података (*SUBP*).



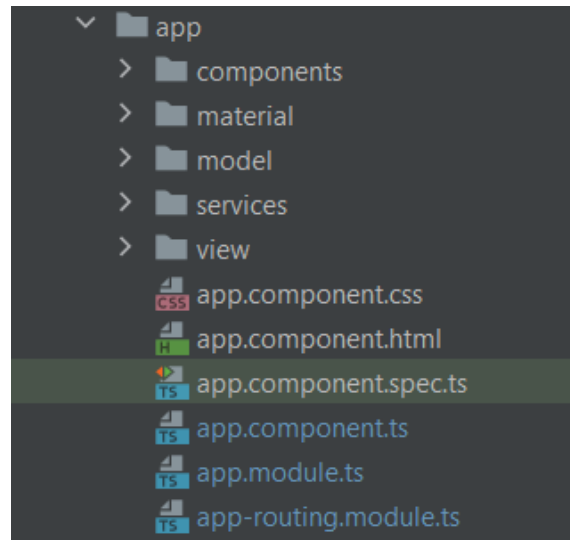
Слика 3.1. Three-tier архитектура

### 3.2.1. Презентациони слој

Клијентски или презентациони слој имплементиран је употребом *TypeScript* језика и *Angular* радног оквира. Изглед корисничког интерфејса креиран је употребом *HTML*-а, *CSS*-а и уз помоћ *Angular Material* библиотеке.

Компоненте су хијерархијски организоване у стабло и унутар пројекта су подељене у два директоријума: *components* који садржи атомичне компоненте даље коришћене у компонентама *view* директоријума, *view* директоријум, сачињен од листе компоненти које представљају странице, енкапсулирајући кориснички приказ.

Преостали директоријуми су: *service* – сачињен од сервиса за комуникацију са серверском апликацијом, *model* – сачињен од класа којима се енкапсулирају подаци за размену порука са апликативним слојем.

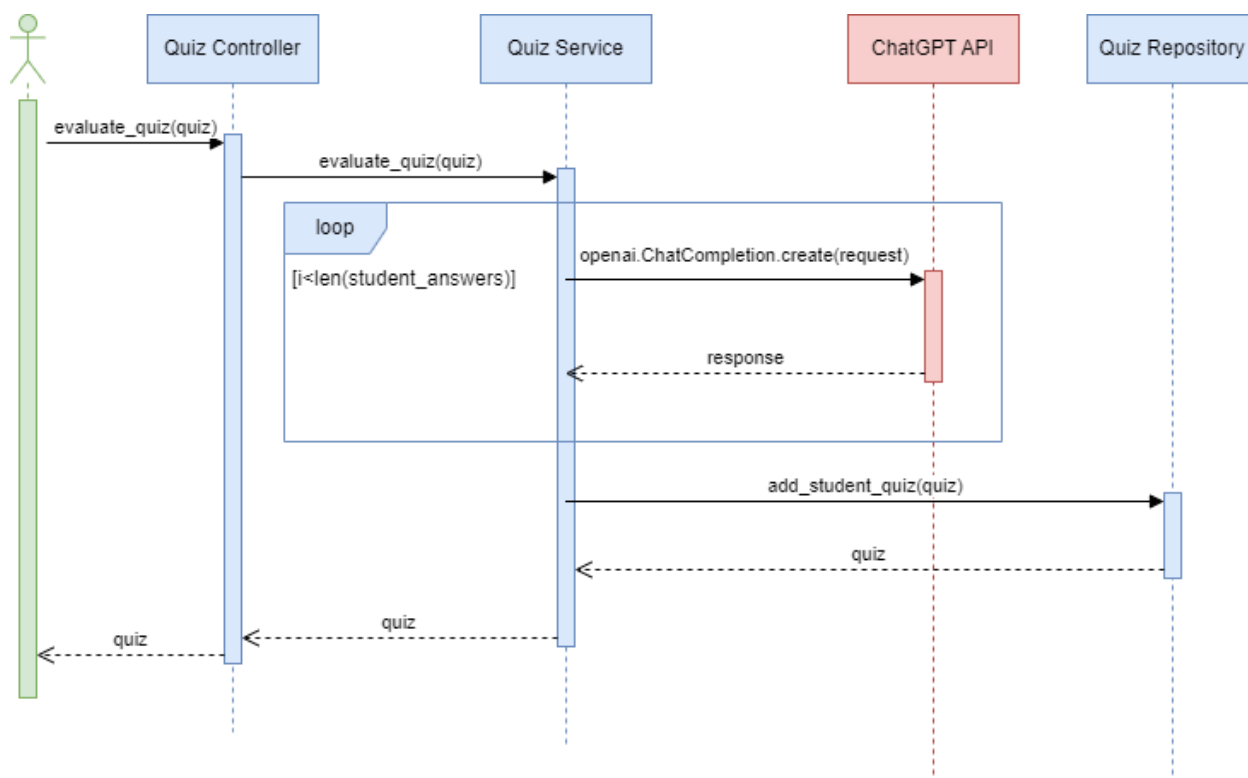


### 3.2.2. Апликативни слој

Апликативни слој је имплементиран као *Representational State Transfer (REST)* сервис у програмском језику *Python*, користећи *Flask* радни оквир. Поред *Flask*-а такође се користи и *SQLAlchemy* за рад са базом података.

Структура овог слоја поштује слојни стил архитектуре (енгл. *layered architecture*), поседује пакет *controllers* који садржи *Flask REST* контролере који прихватају захтеве клијената, обрађују их и у складу са тим враћају одговор. Пакет *dto* поседује класе намењене преносу података (енгл. *Data Transfer Object*) са клијентском страном. Оне доприносе смањењу саобраћаја на мрежи јер се врши транспорт само неопходних података. Пакет *model* садржи класе које чине доменски модел система. Садржај *repository* пакета чине репозиторијуми који обављају комуникацију са базом података, док пакет *service* обухвата логику апликације кроз сервисне интерфејсе и њихове имплементације.

Апликативни слој комуницира са интерфејсом ChatGPT-ја путем одговарајућих *HTTP* захтева. Приликом постављања питања, захтев се конструише са неопходним подацима, укључујући и АПИ кључ. Након обраде захтева, одговор се прима и даље обрађује у апликативном слоју. У наставку је приказ дијаграма



Слика 3.3. Дијаграм секвенце оцењивања теста

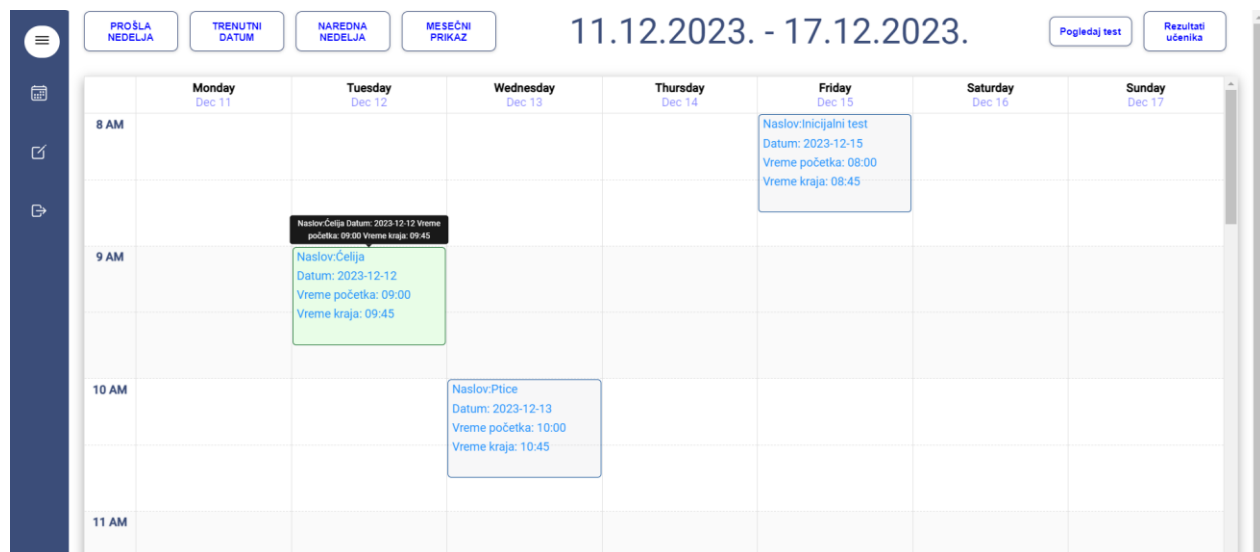
### 3.3. Приказ имплементираног решења

Систем пружа функционалности утврђивања градива за наставнике и ученике. Обе улоге имају могућност прегледа календара тестова, наставнику се пружа могућност увида у изглед теста, резултате тестирања, као и увид у појединачне одговоре ученика. С друге стране, ученици, поред прегледа календара, имају могућност полагања теста.

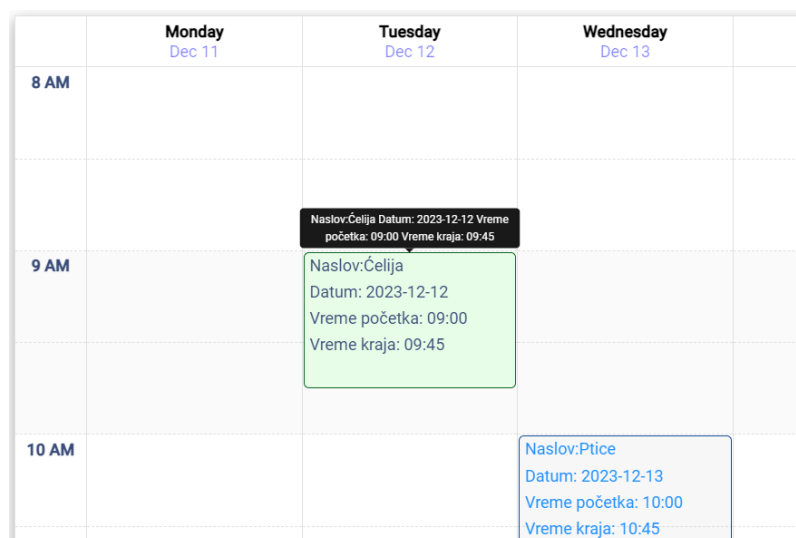
### 3.3.1. Приказ календара тестова

Након што се улогују на систем, корисницима се приказује календар са тестовима. Календар представља корисну функционалност у нашем систему обезбеђујући наставницима и ученицима једноставан и прегледан начин организације и планирања предстојећих обавеза. У горњем левом углу се налази навигациона трака која омогућава прелазак на жељени датум, па и промену на месечни режим.

Дођађаји су представљени картицом, чија величина зависи од дужине самог теста. На картици се налазе основне информације о тесту попут назива, предмета, датума, временског почетка и краја. Кликом на ту картицу у десном горњем углу, корисник може да полаже тест, уколико је ученик или да погледа тест и резултате теста уколико је наставник.



Слика 3.4. Приказ календара за наставнике



Слика 3.5. Приказ картице заказаног теста

### 3.3.2. Креирање теста

Како би се извршило утврђивање знања ученика, потребно је да наставник креира тестове. Уношење основних информација теста подразумева уписивање наслова теста, предмета, напомене за ученике, избора датума, као и временског избора почетка и краја теста. Форма основних информација теста представљена је сликом 3.6.

The screenshot shows a web interface for creating a test. On the left is a dark blue sidebar with icons for a menu, calendar, edit, and share. At the top, a progress bar has three steps: '1. Opšte informacije' (active), '2. Test pitanja', and '3. Gotovo'. The main heading is 'Osnovne informacije'. The form contains several input fields: 'Naslov: \*', 'Predmet: \*', 'Napomena za učenike: \*', 'Izaberi datum' with a calendar icon and a placeholder 'MM/DD/YYYY', and two time selection fields 'Izaberi vreme' and 'Izaberi vreme završetka'. A blue 'Next' button is on the right.

Слика 3.6. Приказ форме основних информација теста

Након што је корисник унео основне информације, потребно је да унесе питања. Питање садржи текст, број бодова, тачан одговор и тип. Тип питања може бити слободан одговор или више понуђених одговора. Уколико се корисник одлучи за други избор, форма ће се проширити са још 3 додатна поља која представљају нетачне одговоре који ће бити понуђени ученику.

The screenshot shows the 'Dodaj pitanja' form. The progress bar now shows '1. Opšte informacije' as completed with a checkmark, and '2. Test pitanja' as the active step. The main heading is 'Dodaj pitanja'. The form has four input fields: 'Pitanje:', 'Tačan odgovor:', 'Bodovi:', and 'Izbaeri tip pitanja:' with a dropdown arrow. Below these is a wide blue button labeled 'Dodaj'. A blue 'Dalje' button is on the right. At the bottom center, there is a small 'Discord' link.

Слика 3.7. Приказ форме за додавање питања

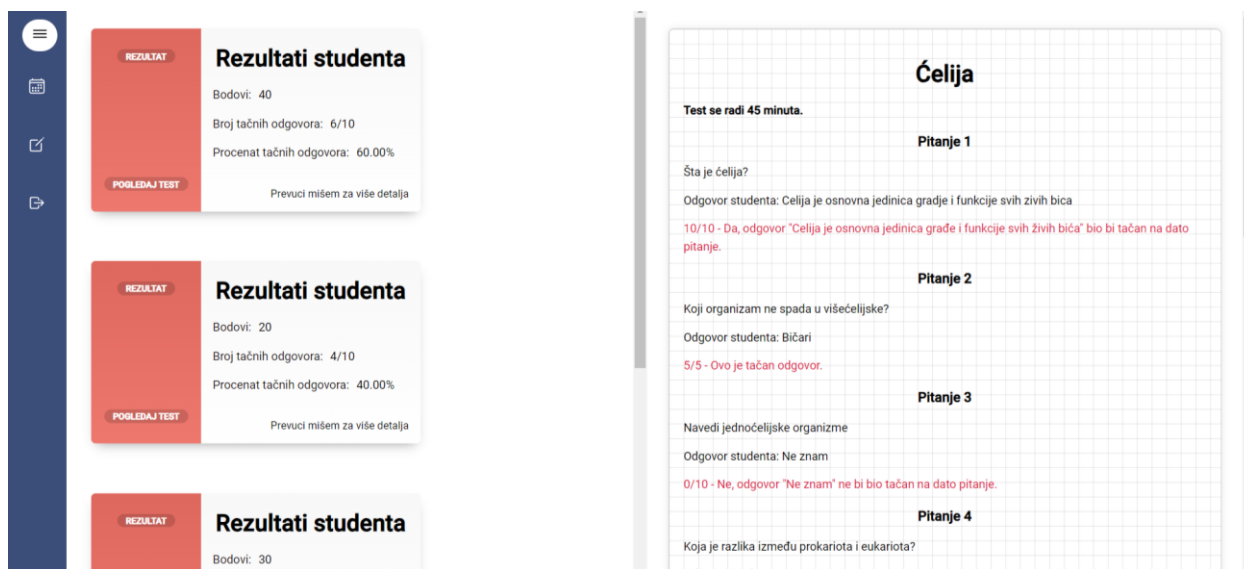


Након уношења свих информација, корисник додаје питање у тест кликом на дугме „Додај“, при чему може да уноси нова питања. Кад заврши уношење свих питања кликом на дугме „Даље“ завршава процес креирања теста.

### 3.3.3. Приказ креираног теста и резултата тестирања

Наставник може да погледа како изгледа било који тест у његовом календару, независно од тога да ли је био у прошлости или тек треба да буде реализован. Осим тога, може да погледа и какви су резултати ученика. У левом делу екрана, кориснику се приказује листа картица са информацијама урађених тестова.

Картица се садржи податке о томе колико бодова је ученик освојио, који је број и проценат тачних одговора, након што превуче мишем преко картице приказују се подаци о имену, презимену, кад је почео да ради тест и кад је завршио. Уколико корисник кликне на дугме „Погледај тест“, на десној страни ће имати увид у све ученикове одговоре.



Слика 3.8. Приказ ученичких резултата

### 3.3.4. Полагање теста

Сликом 3.9. приказан је пример полагања теста. Како би се избегло могуће преписивање, распоред одговора, у питањима са понуђеним одговорима, је насумичан. Након што ученик унесе одговоре на дата питања, тест се аутоматски оцењује. За свако питање, приказује се број бодова које је освојио и коментар система. Коментар је у виду тачног одговора, образложења зашто је учеников одговор нетачан или поруке похвале уколико је одговор тачан, поред тога, на крају теста налази се и укупан број бодова остварених на тесту.

Ćelija

Test se radi 45 minuta.

Pitanje 1

Šta je ćelija?

Type here

Pitanje 2

Koji organizam ne spada u višećelijske?

☐ Gljive

☐ Biljke

☐ Bičari

☐ Životinje

Pitanje 3

Navedi jednoćelijske organizme

Type here

Pitanje 4

Koja je razlika između prokariota i eukariota?

Type here

Слика 3.9. Приказ полагања теста

0/10 - Tačan odgovor je: Glavna funkcija mitohondrija u ćeliji je proizvodnja energije putem procesa respiracije.

Pitanje 8

Kako se zove deoba somatskih ćelija, a kako polnih?

Mejoza je deoba somatskih ćelija, mitozu polnih

0/10 - Ne. Odgovor "Mejoza je deoba somatskih ćelija, mitozu polnih" je netačan. Mejoza je deoba polnih ćelija (gameta), a mitozu je deoba somatskih (telo) ćelija. Pravilan odgovor je: "Deoba somatskih ćelija se zove mitozu, a polnih mejoza."

Pitanje 9

Navedi faze mitoze.

Type here

0/10 - Tačan odgovor je: Profaza, metafaza, anafaza, telofaza

Pitanje 10

Koliko hromozoma sadrži jedna somatska ćelija?

☐ 23

☒ 46

☐ 44

☐ 1

5/5 - Ovo je tačan odgovor.

Submit

Your points: 40

Слика 3.10. Приказ оцењеног теста

## 4. Експериментални резултати и дискусија

У овом поглављу биће представљени експериментални резултати, који су добијени током истраживања, као и дискусија базирана на њима. Избору *ChatGPT*-а као средства за процену тачности ученичког одговора, претходила је упоредна анализа других модела и приступа.

Првенствено је идеја била да модел обезбеди векторе репрезентације (енгл. *embeddings*) одговора, ученичких и тачног, наставничког, како би се, помоћу метрике косинусне сличности, проценила сличност парова ученичког и тачног одговора. Модел би поседовао задати праг, уколико је сличност преко тог прага, сматра се да је ученик добро одговорио на питање.

Након анализе резултата, изабрани су модели који су се показали довољно добро да би се на њима даље вршили експерименти, попут превођења оба одговора на енглески језик у покушају бољег генерисања вектора репрезентације, а самим тим и упоређивања сличности.

Уместо да модел служи за генерисање вектора репрезентације, а проста метрика изврши процену сличности, испробан је и на крају и изабран приступ коришћења *ChatGPT*-а као директног средства за процену тачности ученичког одговора. У наставку су приказани и детаљније описани експериментални резултати поменутих фаза истраживања.

### 4.1. Вектори репрезентације одговора и косинусна сличност

Изабрано је неколико језичких модела: ***Bertić***, модел трансформатора трениран над подацима српског, хрватског и босанског језика али за проблем препознавања именованих ентитета (енгл. *Named Entity Recognition*), ***paraphrase-multilingual-MiniLM-L12-v2*** и ***stsb-xlm-r-multilingual*** два вишејезична модела *BERT*-а коришћених и тренираних у архитектури сијамске неуронске мреже за проблеме препознавања парафраза и процене семантичке текстуалне сличности, ***all-MiniLM-L6-v2*** и ***laBSE*** модели избарани због количине података над којима су тренирани и ***OpenAI GPT***.

Ради лакшег приказа резултата, моделима су додељене ознаке којима ће бити референцирани у даљим табелама, такође приказани су и прагови сличности (табела 4.1.) изнад којих се одговор сматра тачним.

Модел	Ознака	Праг
Bertić	A	0.93
paraphrase-multilingual-MiniLM-L12-v2	B	0.70
stsb-xlm-r-multilingual	C	0.77
all-MiniLM-L6-v2	D	0.60
laBSE	E	0.65
OpenAI GPT	F	0.89

Табела 4.1. Модели са својим ознакама и праговима

У заглављу табела 4.2., 4.3., 4.4., 4.5., прва колона представља тачан одговор, а онда следе ознаке модела којима резултати у њиховој колони припадају. Уколико се поред ученичких одговора налази ознака **T**, то означава да је тај одговор процењен као тачан. Резултати у колонама модела представљају резултат косинусне сличности. Одговори ученика на питања у наредним примерима су дати од стране аутора како би се испитале и тежи облици провере тачности одговора. Црвеном бојом означене су погрешне процене.

#### Питање 1: Која је функција митохондрија?

Главна функција митохондрија у ћелији је производња енергије путем процеса респирације		A	B	C	D	E	F
<b>T</b>	Митохондрије производе АТП, главни извор енергије у ћелији	0.97	0.86	0.77	0.67	0.71	0.92
<b>T</b>	Митохондрије су одговорне за стварање енергије у ћелији путем оксидативне фосфорилације	0.98	0.88	0.83	0.72	0.71	0.89
<b>T</b>	Функција митохондрија је производња енергије	0.82	0.82	0.82	0.76	0.83	0.93
	Митохондрије су одговорне за фотосинтезу	0.94	0.69	0.68	0.54	0.57	0.85
	Митохондрије чувају генетску информацију	0.82	0.60	0.55	0.49	0.67	0.83
	Митохондрије обављају функцију стварања нових ћелија у организму	0.90	0.75	0.64	0.73	0.79	0.89

Табела 4.2. Приказ резултата одговора за питање 1

Овим примерима желели смо да проверимо како ће се модели снаћи са апозицијом, синонимима попут речи „стварање“ и „производња“, како у позитивном примеру (други ред), тако и у негативном (последњи ред).

#### Питање 2: Која је улога еритроцита?

Основна функција еритроцита је преношење кисеоника из плућа до ткива и ћелија у телу		A	B	C	D	E	F
<b>T</b>	Еритроцити транспортују кисеоник до ћелија у организму	0.93	0.92	0.95	0.76	0.49	0.90
<b>T</b>	Главна улога еритроцита је преношење кисеоника до свих делова	0.93	0.85	0.92	0.83	0.60	0.94
	Функција митохондрија је производња енергије	0.79	0.46	0.60	0.43	0.57	0.85
	Еритроцити су одговорни за варење хране у телу	0.89	0.53	0.58	0.62	0.72	0.85
	Еритроцити су црвене крвне ћелије које луче хормоне	0.90	0.56	0.43	0.62	0.44	0.87
	Еритроцити помажу у филтрацији крви у бубрезима	0.51	0.64	0.74	0.61	0.51	0.87

Табела 4.3. Приказ резултата одговора за питање 1

### Питање 3: Која је функција јетре у организму?

Функција јетре је детоксикација организма		A	B	C	D	E	F
T	Јетра има улогу у чишћењу организма од токсина	0.95	0.70	0.83	0.80	0.53	0.89
	Јетра је одговорна за производњу хормона	0.96	0.36	0.61	0.53	0.52	0.82
	Плућа су одговорна за филтрацију крви у организму	0.97	0.43	0.72	0.61	0.50	0.84
	Јетра има улогу у варењу хране	0.92	0.35	0.73	0.65	0.46	0.83
	Јетра производи црвена крвна зрнца	0.90	0.20	0.67	0.43	0.44	0.80
	Еритроцити помажу у филтрацији крви у бубрезима	0.52	0.37	0.71	0.55	0.38	0.82

Табела 4.4. Приказ резултата одговора за питање 3

Коришћењем израза „чишћење организма од токсина“ за реч детоксикација, направили смо велику варијацију описа појма.

### Питање 4: Која је функција бубрега?

Функција бубрега је филтрација крви и регулација равнотеже електролита		A	B	C	D	E	F
T	Бубрези чисте крв и одржавају електролитску равнотежу	0.97	0.80	0.78	0.80	0.71	0.89
	Бубрези су задужени за баланс електролита	0.95	0.80	0.87	0.65	0.70	0.87
	Срце је одговорно за варење хране	0.87	0.35	0.50	0.44	0.50	0.79
	Бубрези помажу у транспорту кисеоника у тело	0.98	0.54	0.74	0.43	0.40	0.84
	Бубрези луче инсулин	0.89	0.40	0.58	0.40	0.31	0.82
	Функција бубрега је филтрација електролита и регулација равнотеже крви	0.99	0.99	0.99	0.99	0.99	0.99

Табела 4.5. Приказ резултата одговора за питање 4

Анализом резултата увидели смо да модели А, D, Е нису добри за проблем овог рада, те их у даљем наставку текста нећемо коментарисати. Закључак прве фазе јесте да тренирање над сијамском мрежом показује одличне резултате, као и модел GPT-а.

Такође, уочени су и неки од проблема са којима модели имају потешкоћу. Уколико одговор садржи набрајања (питање 4.), уколико ученик наведе само један од два тачна готово сви модели у већини случајева то неће препознати, исто важи и за случај када ученик наведе тачне одговоре, али и неки нетачан. Пример пермутације (питање 4. -последњи пример), ниједан модел то неће схватити и сви ће сличност оценити са готово највишим процентом. Једне од идеја како да се реше ови проблеми били су н-грами, навођење честих грешака од стране наставника и тражење алтернативе која ће бити фокусиране на препознавање набрајања, са зарезом или везником и. Н-грами не би гарантовали потпуну превенцију грешака, тражење више информација од наставника би отежало и успорило његов посао што је супротно од онога што желимо да постигнемо, а алтернативе препознавања набрајања би одузеле превише времена за мали проценат побољшања.

## 4.2. Примена превођења на енглески језик

Превођење са српског језика на енглески се вршило уз помоћ бесплатног преводиоца библиотеке *googletrans*, компаније *Google*. Иако се модел D, показао лоше, како је његова предност велики скуп података над којим је трениран, проверили смо како ће се понашати са енглеским језиком. Резултати су показали да превођење не помаже.

### Питање 1: Која је функција митохондрија?

Главна функција митохондрија у ћелији је производња енергије путем процеса респирације		B	B <sub>en</sub>	C	C <sub>en</sub>	D	D <sub>en</sub>	E	E <sub>en</sub>
T	Митохондрије производе АТП, главни извор енергије у ћелији	0.86	0.85	0.77	0.80	0.71	0.79	0.92	0.91
T	Митохондрије су одговорне за стварање енергије у ћелији путем оксидативне фосфорилације	0.88	0.86	0.83	0.85	0.71	0.81	0.89	0.88
T	Функција митохондрија је производња енергије	0.82	0.77	0.82	0.81	0.83	0.85	0.93	0.92
	Митохондрије су одговорне за фотосинтезу	0.69	0.72	0.68	0.71	0.57	0.72	0.85	0.85
	Митохондрије чувају генетску информацију	0.60	0.63	0.55	0.53	0.67	0.57	0.83	0.81
	Митохондрије обављају функцију стварања нових ћелија у организму	0.75	0.75	0.64	0.68	0.79	0.77	0.89	0.88

Табела 4.6. Приказ резултата превођења на енглески језик за питање 1

### Питање 2: Која је улога еритроцита?

Основна функција еритроцита је преношење кисеоника из плућа до ткива и ћелија у телу		B	B <sub>en</sub>	C	C <sub>en</sub>	D	D <sub>en</sub>	E	E <sub>en</sub>
T	Еритроцити транспортују кисеоник до ћелија у организму	0.92	0.92	0.95	0.94	0.76	0.82	0.90	0.92
T	Главна улога еритроцита је преношење кисеоника до свих делова	0.85	0.90	0.92	0.90	0.60	0.86	0.94	0.91
	Функција митохондрија је производња енергије	0.46	0.42	0.60	0.44	0.57	0.36	0.85	0.80
	Еритроцити су одговорни за варење хране у телу	0.53	0.69	0.58	0.56	0.72	0.66	0.85	0.86
	Еритроцити су црвене крвне ћелије које луче хормоне	0.56	0.61	0.43	0.46	0.44	0.66	0.87	0.83
	Еритроцити помажу у филтрацији крви у бубрезима	0.64	0.65	0.74	0.69	0.51	0.63	0.87	0.85

Табела 4.7. Приказ резултата превођења на енглески језик за питање 2

### Питање 3: Која је функција јетре у организму?

Функција јетре је детоксикација организма		B	B <sub>en</sub>	C	C <sub>en</sub>	D	D <sub>en</sub>	E	E <sub>en</sub>
T	Јетра има улогу у чишћењу организма од токсина	0.70	0.65	0.83	0.70	0.53	0.68	0.89	0.93
	Јетра је одговорна за производњу хормона	0.36	0.71	0.61	0.59	0.52	0.57	0.82	0.85
	Плућа су одговорна за филтрацију крви у организму	0.43	0.45	0.72	0.72	0.50	0.37	0.84	0.82
	Јетра има улогу у варењу хране	0.35	0.66	0.73	0.65	0.46	0.63	0.83	0.88
	Јетра производи црвена крвна зрнца	0.20	0.61	0.67	0.61	0.44	0.55	0.80	0.83
	Еритроцити помажу у филтрацији крви у бубрезима	0.37	0.40	0.71	0.66	0.38	0.29	0.82	0.78

Табела 4.8. Приказ резултата превођења на енглески језик за питање 3

#### Питање 4: Која је функција бубрега?

Функција бубрега је филтрација крви и регулација равнотеже електролита		B	B en	C	C en	D	D en	E	E en
T	Бубрези чисте крв и одржавају електролитску равнотежу	0.80	0.86	0.78	0.83	0.71	0.81	0.89	0.91
	Бубрези су задужени за баланс електролита	0.80	0.91	0.87	0.89	0.70	0.84	0.87	0.92
	Срце је одговорно за варење хране	0.35	0.32	0.50	0.49	0.50	0.21	0.79	0.79
	Бубрези помажу у транспорту кисеоника у тело	0.54	0.68	0.74	0.72	0.40	0.63	0.84	0.84
	Бубрези луче инсулин	0.40	0.59	0.58	0.60	0.31	0.52	0.82	0.83
	Функција бубрега је филтрација електролита и регулација равнотеже крви	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99

Табела 4.9. Приказ резултата превођења на енглески језик за питање 4

### 4.3. ChatGPT

У раду је коришћена *gpt-3.5-turbo* верзија ChatGPT-а. У зависности од тога како је постављено питање, резултати су били другачији. Ради лакшег приказа резултата, структурама питања су додељене ознаке (табела 4.10.).

Структура питања	Ознака
„{correct}, {response}. Da li je druga rečenica tačna, ako je tačna prva rečenica, odgovori sa da/ne “	E2
„Da li je odgovor: {response} tačan na pitanje: {question}, odgovori sa da/ne“	E3
„Na pitanje {question} očekivan odgovor je {correct}. Da li bi u odnosu na očekivan odgovor, {response} bio tačan na dato pitanje, odgovori sa da/ne“	E4

Табела 4.10. Структуре питања са својом ознаком

У наредним табелама се налазе одговори ChatGPT-а на дате примере, црвеном бојом су означене грешке.

#### Питање 1: Која је функција митохондрија?

Главна функција митохондрија у ћелији је производња енергије путем процеса респирације		E2	E3	E4
T	Митохондрије производе АТП, главни извор енергије у ћелији	DA	DA	DA
T	Митохондрије су одговорне за стварање енергије у ћелији путем оксидативне фосфорилације	DA	DA	DA
T	Функција митохондрија је производња енергије	DA	DA	DA
	Митохондрије су одговорне за фотосинтезу	NE	NE	NE
	Митохондрије чувају генетску информацију	NE	NE	NE
	Митохондрије обављају функцију стварања нових ћелија у организму	NE	NE	NE

Табела 4.11. Приказ резултата ChatGPT-а за питање 1

#### Питање 2: Која је улога еритроцита?

Основна функција еритроцита је преношење кисеоника из плућа до ткива и ћелија у телу		E2	E3	E4
T	Еритроцити транспортују кисеоник до ћелија у организму	DA	DA	DA
T	Главна улога еритроцита је преношење кисеоника до свих делова	DA	DA	DA
	Функција митохондрија је производња енергије	DA	NE	NE
	Еритроцити су одговорни за варење хране у телу	NE	NE	NE
	Еритроцити су црвене крвне ћелије које луче хормоне	NE	NE	NE
	Еритроцити помажу у филтрацији крви у бубрезима	NE	NE	NE

Табела 4.12. Приказ резултата ChatGPT-а за питање 2

### Питање 3: Која је функција јетре у организму?

Функција јетре је детоксикација организма		E2	E3	E4
T	Јетра има улогу у чишћењу организма од токсина	DA	DA	DA
	Јетра је одговорна за производњу хормона	NE	NE	NE
	Плућа су одговорна за филтрацију крви у организму	NE	NE	NE
	Јетра има улогу у варењу хране	DA	NE	NE
	Јетра производи црвена крвна зрнца	NE	NE	NE
	Еритроцити помажу у филтрацији крви у бубрезима	NE	NE	NE

Табела 4.13. Приказ резултата ChatGPT-а за питање 3

### Питање 4: Која је функција бубрега?

Функција бубрега је филтрација крви и регулација равнотеже електролита		E2	E3	E4
T	Бубрези чисте крв и одржавају електролитску равнотежу	DA	DA	DA
	Бубрези су задужени за баланс електролита	DA	DA	NE
	Срце је одговорно за варење хране	NE	NE	NE
	Бубрези помажу у транспорту кисеоника у тело	NE	NE	NE
	Бубрези луче инсулин	NE	NE	NE
	Функција бубрега је филтрација електролита и регулација равнотеже крви	NE	DA	NE

Табела 4.14. Приказ резултата ChatGPT-а за питање 4

Анализом резултата закључено је да ChatGPT-а, одговори готово увек тачно уколико му се пружи питање и тачан одговор наставника који би му сузио контекст, грешке које направи структура питања E4 спада у ограничења описана у закључку. Уколико бисмо послали структуру питања E3, модел би понекад погрешно на примерима где је ученик одговорио неком последицом на тачан одговор, али не у траженом контексту. Структура питања E2, се показала као најгори приступ од поменутих три.

## 5. Закључак

Представљени систем репрезентује једно од могућих решења проблема утврђивања знања заснованог на препознавању природног језика. Корисницима система обезбеђене су функционалности прегледа тестова, њиховог креирања, полагања и аутоматског оцењивања.



Истраживали смо могућности и изазове у развоју утврђивање знања на српском језику. Анализирали смо актуелне приступе у овом домену, укључујући алате као што су *ChatGPT*, *GPT*, *BERT*, и њихове варијанте, као и њихову примену у контексту српског језика.

Истраживање је показало да постоји значајан потенцијал примене *ChatGPT*-а за задат проблем, али и да постоје одређена ограничења. Број захтева по времену је ограничен, како нисмо водили рачуна о скалабилности и брзини, у раду је овај проблем решен одложеним слањем захтева. Такође, у тренутку писања овог рада, *ChatGPT* нема приступ подацима после јануара 2022. године што би представљало једно од ограничења.

Даљи развој система би се заснивао на проширивању функционалности апликације, попут вођења анализе напретка, статистике резултата, уграђивања образовног садржаја попут лекција и многе друге. Такође, напредак би могао да се посматра и у погледу коришћења нових алата за процену утврђеног знања, неки од интересантних примера би био примена модела *Mistral*.

## 6. Литература

- [1] Evaluating Academic Answers Generated Using ChatGPT, Suzanne Fergus, Michelle Botha and Mehrnoosh Ostovar, March 2023.
- [2] Semantic Answer Similarity for Evaluating Question Answering Models Julian Risch, Timo Möller, Julian Gutsch, Malte Pietsch, August 2021.
- [3] EVALUATION OF SEMANTIC ANSWER SIMILARITY METRICS, Farida Mustafazade and Peter F. Ebbinghaus, June 2022.