



УНИВЕРЗИТЕТ У НОВОМ САДУ
ФАКУЛТЕТ ТЕХНИЧКИХ
НАУКА У НОВОМ САДУ



Марко Митошевић

**Додавање подршке за
библиотеку Keras 3 у радно
окружење TensorFlow Federated**

ЗАВРШНИ РАД

Основне академске студије

Нови Сад, 2025

	УНИВЕРЗИТЕТ У НОВОМ САДУ • ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА 21000 НОВИ САД, Трг Доситеја Обрадовића 6	Број:
	ЗАДАТАК ЗА ЗАВРШНИ РАД	Датум:

(Податке уноси предметни наставник - ментор)

Студијски програм:	Софтверско инжењерство и информационе технологије		
Студент:	Марко Митошевић	Број индекса:	SV56/2021
Степен и врста студија:	Основне академске студије		
Област:	Електротехничко и рачунарско инжењерство		
Ментор:	Игор Дејановић		
НА ОСНОВУ ПОДНЕТЕ ПРИЈАВЕ, ПРИЛОЖЕНЕ ДОКУМЕНТАЦИЈЕ И ОДРЕДБИ СТАТУТА ФАКУЛТЕТА ИЗДАЈЕ СЕ ЗАДАТАК ЗА ЗАВРШНИ РАД, СА СЛЕДЕЋИМ ЕЛЕМЕНТИМА: <ul style="list-style-type: none"> - проблем – тема рада; - начин решавања проблема и начин практичне провере резултата рада, ако је таква провера неопходна; 			

НАСЛОВ ЗАВРШНОГ РАДА:


Додавање подршке за библиотеку Keras 3 у радно окружење TensorFlow Federated
--

ТЕКСТ ЗАДАТКА:

<p>Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.</p>
--

Руководилац студијског програма:	Ментор рада:

Примерак за: <input type="checkbox"/> - Студента; <input type="checkbox"/> - Ментора
--

	УНИВЕРЗИТЕТ У НОВОМ САДУ • ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА 21000 НОВИ САД, Трг Доситеја Обрадовића 6
	КЉУЧНА ДОКУМЕНТАЦИЈСКА ИНФОРМАЦИЈА

Редни број, РБР:	
Идентификациони број, ИБР:	
Тип документације, ТД:	Монографска документација
Тип записа, ТЗ:	Текстуални штампани материјал
Врста рада, ВР:	Дипломски - бечелор рад
Аутор, АУ:	Марко Митошевић
Ментор, МН:	Др Игор Дејановић, редовни професор
Наслов рада, НР:	Додавање подршке за библиотеку Keras 3 у радно окружење TensorFlow Federated
Језик публикације, ЈП:	српски/ћирилица
Језик извода, ЈИ:	српски/енглески
Земља публиковања, ЗП:	Република Србија
Уже географско подручје, УГП:	Војводина
Година, ГО:	2025
Издавач, ИЗ:	Ауторски репринт
Место и адреса, МА:	Нови сад, трг Доситеја Обрадовића 6
Физички опис рада, ФО: <small>(поглавља/страна/ цитата/табела/слика/графика/прилога)</small>	6/26/4/1/1/0/0
Научна област, НО:	Електротехничко и рачунарско инжењерство
Научна дисциплина, НД:	Примењене рачунарске науке и информатика
Предметна одредница/Кључне речи, ПО:	Keras, TensorFlow Federated, IDE
УДК	
Чува се, ЧУ:	У библиотеци Факултета техничких наука, Нови Сад
Важна напомена, ВН:	
Извод, ИЗ:	Овај документ представља упутство за писање завршних радова на Факултету техничких наука Универзитета у Новом Саду. У исто време је и шаблон за Typst.

Датум прихватања теме, ДП:	
Датум одбране, ДО:	01.01.2025
Чланови комисије, КО:	Председник: Др Петар Петровић, ванредни професор
	Члан: Др Марко Марковић, доцент
	Члан, ментор: Др Игор Дејановић, редовни професор

Потпис ментора

	UNIVERSITY OF NOVI SAD • FACULTY OF TECHNICAL SCIENCES 21000 NOVI SAD, Trg Dositeja Obradovića 6
	KEY WORDS DOCUMENTATION

Accession number, ANO :	
Identification number, INO :	
Document type, DT :	Monographic publication
Type of record, TR :	Textual printed material
Contents code, CC :	
Author, AU :	Marko Mitosevic
Mentor, MN :	Igor Dejanović, Phd., full professor
Title, TI :	Adding support for the Keras 3 library to the TensorFlow Federated framework
Language of text, LT :	Serbian
Language of abstract, LA :	Serbian
Country of publication, CP :	Republic of Serbia
Locality of publication, LP :	Vojvodina
Publication year, PY :	2025
Publisher, PB :	Author's reprint
Publication place, PP :	Novi Sad, Dositeja Obradovica sq. 6
Physical description, PD : (chapters/pages/ref./tables/pictures/graphs/appendixes)	6/26/4/1/1/0/0
Scientific field, SF :	Electrical and Computer Engineering
Scientific discipline, SD :	Applied computer science and informatics
Subject/Key words, S/KW :	Keras, TensorFlow Federated, IDE
UC	
Holding data, HD :	The Library of Faculty of Technical Sciences, Novi Sad, Serbia
Note, N :	
Abstract, AB :	This document provides guidelines for writing final theses at the Faculty of Technical Sciences, University of Novi Sad. At the same time, it serves as a Typst template.

Accepted by the Scientific Board on, ASB :				
Defended on, DE :	01.01.2025			
Defended Board, DB :	President:	Petar Petrović, Phd., assoc. professor		
	Member:	Marko Marković, Phd., asist. professor		
	Member, Mentor:	Igor Dejanović, Phd., full professor		
		<table><tr><td>Menthor's sign</td></tr><tr><td></td></tr></table>	Menthor's sign	
Menthor's sign				



УНИВЕРЗИТЕТ У НОВОМ САДУ • ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА
21000 НОВИ САД, Трг Доситеја Обрадовића 6

ИЗЈАВА О НЕПОСТОЈАЊУ СУКОБА ИНТЕРЕСА

Изјављујем да нисам у сукобу интереса у односу ментор – кандидат и да нисам члан породице (супружник или ванбрачни партнер, родитељ или усвојитељ, дете или усвојеник), повезано лице (крвни сродник ментора/кандидата у правој линији, односно у побочној линији закључно са другим степеном сродства, као ни физичко лице које се према другим основама и околностима може оправдано сматрати интересно повезаним са ментором или кандидатом), односно да нисам зависан/на од ментора/кандидата, да не постоје околности које би могле да утичу на моју непристрасност, нити да стичем било какве користи или погодности за себе или друго лице било позитивним или негативним исходом, као и да немам приватни интерес који утиче, може да утиче или изгледа као да утиче на однос ментор-кандидат.

У Новом Саду, дана _____

Ментор

Кандидат

Садржај

1	Увод	1
1.1	Мотивација	1
1.2	Циљеви рада	2
1.3	Организација рада	2
2	Теоријске основе	3
2.1	Федеративно учење	3
2.2	Оптимизована диференцијална приватност	4
2.3	Радно окружење <i>TensorFlow Federated</i> (TFF)	6
2.4	Библиотека <i>Keras</i>	7
3	Имплементација подршке за <i>Keras</i> 3	9
3.1	Анализа имплементације проблема компатибилности	9
3.2	Архитектура рефакторисања	10
3.3	Апстракција логике руковања моделом	10
4	Закључак	15
	Биографија	23
	Литература	25

1.1 Мотивација

Модерна интегрисана развојна окружења (*Integrated Development Environment*, IDE) имају велики број функционалности које помажу програмерима да ефикасно рукују пројектима великог обима. Једна од кључних функционалности у IDE-овима, попут *IntelliJ IDEA*, је претрага *Search Everywhere* (SE), која кориснику на једном месту омогућава да претражи све функционалности окружења, као и све датотеке и њихов садржај унутар пројектне структуре [1].

Корисници често не знају тачно име функционалности коју желе да пронађу, што отежава претрагу. Како би се овај проблем ублажио, истражују се решења модела машинског учења (ML) која могу да предвиде име жељене функционалности. Коришћењем ових модела могуће је направити препоруку претраге, чиме би се олакшало коришћење функционалности SE. Препорука би била најефикаснија уколико би модел био трениран над претходним претрагама корисника и других корисника, као и над контекстом под којим је претрага позвана.

Тренирање модела над осетљивим корисничким подацима представља изазов у погледу приватности и поверљивости. Многе компаније морају чувати поверљивост свог кода и информационих система. Уредбе о заштити података корисника попут Опште уредбе о заштити података (*General Data Protection Directive*, GDPR), која је ступила на снагу у Европској Унији, забрањују обраду података без експлицитне сагласности корисника или адекватног правног основа [2]. Наведена ограничења онемогућавају директно слање података на централни сервер ради тренирања модела.

Постоје технике ML које имају способност да чувају поверљивост тренинг података. Најпознатије су федеративно учење (*Federated Learning*, FL) [3] и диференцијална приватност (*Differential Privacy*, ODP) [4]. Ове технике се базирају на дистрибутивном ML којим се гарантује поверљивост података. Мана код тренутне имплементације FL је недостатак подршке за учитавање најновијих модела, који су потребни да тачно предвиде претрагу SE. Елиминисањем ове мане би се отворила могућност за шире коришћење FL у индустрији.

Циљ овог рада је додавање подршке за новије моделе у радно окружење отвореног кода *TensorFlow Federated* (TFF), које имплементира технику FL [5].

1.2 Циљеви рада

Главни циљ овог рада је да се имплементира функционална подршка за библиотеку *Keras* 3 унутар радног окружења TFF. Да би се то постигло, потребно је испунити следеће циљеве:

- Анализирати постојећу архитектуру TFF-а и њену зависност од библиотеке *Keras* 2.
- Рефакторисати компоненте TFF-а тако да се постигла компатибилност са верзијом 3 библиотеком *Keras*, уз очување компатибилности са верзијом 2.
- Имплементирати компоненту која омогућава јединствен начин руковања моделима и компонентама обе верзије библиотеке *Keras*.
- Имплементирати тестове за тестирање компатибилности библиотеке *Keras* 3 са рефакторисаним компонентама TFF.
- Евалуирати функционалну исправност имплементиране подршке проласком свих тестова.
- Демонстрирати предности интеграције *Keras* 3 коришћењем модела са знатно бољим перформансама (нпр. *Gemma* 3) у односу на *Keras* 2 моделе (нпр. GPT-2).

1.3 Организација рада

Рад је организован у пет поглавља. Прво поглавље, Увод, дефинише мотивацију и циљеве рада, уводећи читаоца у проблематику приватности података у машинском учењу и циљ рада. Друго поглавље, Теоријске основе, описује концепт федеративног учења, архитектуру и примену радног окружења TFF, као и библиотеке *Keras*. Треће поглавље, Имплементација подршке за *Keras* 3, описује детаље рефакторисања кода и увођење нових компатибилних функционалности. Четврто поглавље, Резултати и дискусија, представља евалуацију решења и поређење перформанси нових и старих модела. Пето поглавље, Закључак, сумира постигнуте резултате и предлаже правце за даља унапређења.

Глава 2

Теоријске основе

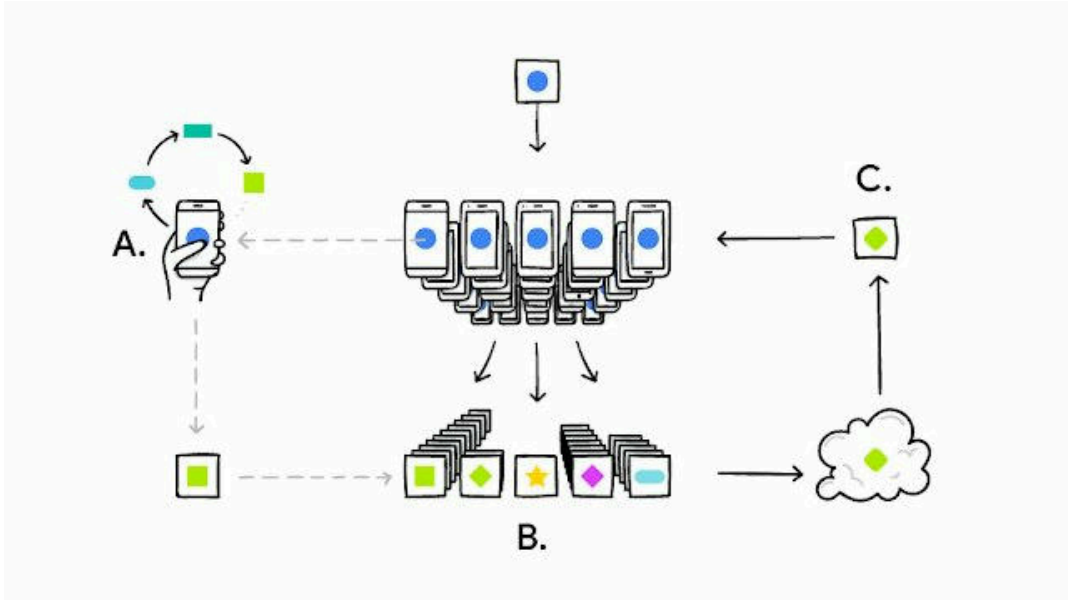
2.1 Федеративно учење

Федеративно учење (*Federated Learning*, FL) је техника машинског учења која омогућава децентрализовано тренирање модела над поверљивим подацима. За разлику од традиционалног централизованог приступа, где се сви подаци прикупљају на једном серверу, FL гарантује да подаци никада не напуштају оригинални уређај. Овај приступ је кључан за очување поверљивости и усклађеност са регулативама за заштиту података [3].

Систем федеративног учења функционише на принципу децентрализоване обуке. Сваки уређај (клијент) у систему поседује локалну копију дељеног модела. Локални модел се тренира искључиво коришћењем података који се налазе на том уређају. Након завршетка локалног тренинга, уређај не шаље своје приватне податке, већ само ажуриране тежине модела централном серверу.

Централни сервер има улогу агрегатора. Он прикупља ажуриране тежине од великог броја уређаја и агрегира резултате. Израчунати, побољшани глобални модел се затим шаље назад свим уређајима у систему, и циклус обуке се понавља. Процес се изводи итеративно, чиме се временом побољшавају перформансе дељеног модела, без употребе приватних података.

Потенцијалне области примене федеративног учења су индустрије које захтевају учење модела над поверљивим скуповима података великог броја корисника, као што су медицина, роботика и софтверско инжењерство. Најутицајнија примена FL-а над великим скупом података је предвиђање корисничког уноса на Gboard Android тастатури. На слици 1 приказан је примена FL за предвиђање корисничког уноса на Gboard Android тастатури [6].



Слика 1: Примена FL за предвиђање корисничког уноса на *Gboard Android* тастатури. Мобилни уређај локално тренира над претходним корисничким уносима и ажуриране тежине шаље централном серверу (А). Централни сервер агрегира примљене тежине (В) и шаље свим уређајима побољшан дељени модел (С) [6].

Кључна предност Федеративног учења је гарантована поверљивост података, јер подаци остају искључиво на оригиналним уређајима, што омогућава децентрализовано тренирање. Ова техника, заједно са оптимизованом диференцијалном приватношћу, базира се на дистрибутивном машинском учењу којим се ефикасно чува поверљивост тренинг података.

2.2 Оптимизована диференцијална приватност

Диференцијална приватност (*Differential Privacy*, DP) је математички дефинисан оквир који квантификује и гарантује приватност појединца у оквиру базе података. Основна идеја је увођење контролисаног шума у процес обраде података, тако да присуство или одсуство записа једног појединца у скупу података не може значајно утицати на коначни излаз анализе или модела. Овај концепт пружа гаранцију да нападач не може са сигурношћу закључити да ли је одређени појединац био део скупа података [7].

Математичка дефиниција DP се најчешће изражава параметрима ϵ (епсилон) и δ (делта). Алгоритам A је (ϵ, δ) диференцијално приватан ако за свака два скупа података која се разликују за само један узорак D и D' и за сваки излазни скуп резултата S важи следећа неједнакост [8]:

$$P[A(D) \in S] \leq e^\epsilon * P[A(D') \in S] + \delta$$

, где је P вероватноћа. Загарантовано је да се вероватноћа добијања било ког излаза не може променити за фактор већи од e^ϵ ако се један узорак дода или уклони.

Параметар ϵ одређује горњу границу промене вероватноће излаза приликом укључивања или изостављања једног узорака података. Нижа вредност ϵ означава бољу приватност, али потенцијално и мању тачност модела. За ϵ једнак нули излази су индентични и добија се загарантована приватност, што је у пракси немогуће достићи.

Параметар δ представља вероватноћу да горња граница ϵ неће бити задовољена. У идеалном теоријском случају δ тежи нули, чиме се добија диференцијална приватност која зависи само од параметра ϵ , док се у практичним применама поставља на малу вредност (нпр. 10^{-5}), која би требало да буде мања од инверзне вредности величине скупа података ($\frac{1}{\text{величина скупа података}}$), чиме се смањује ризик угрожавања приватности појединог узорака у скупу података.

Оптимизована диференцијална приватност (*Optimized Differential Privacy*, ODP) представља варијанту примене DP, која је специјализована за контекст машинског учења и великих, дистрибуираних система. Главни изазов код стандардне DP је постизање оптималног компромиса између приватности и тачности модела. ODP користи методе попут клиповања градијента (*Gradient Clipping*) [9] пре додавања шума, како би се смањио потребан ниво шума и одржала тачност модела, уз истовремено задовољавање дефинисане ϵ границе приватности.

У контексту дистрибутивног машинског учења, DP се може применити на два начина [10]:

- Локална диференцијална приватност (*Local Differential Privacy*, LDP): Шум се додаје директно на сирове податке или ажурирања модела на самом уређају (клијенту), чиме се добија заштита од злонамерног централног сервера.
- Централна диференцијална приватност (*Central Differential Privacy*, CDP): Шум се додаје на централном серверу, након што су прикупљени сви доприноси клијената. У FL, шум се додаје на тежине модела приликом агрегације.

Технике FL и DP су комплементарне. FL пружа технички механизам за децентрализацију тренинга и спречава прикупљање сирових података на серверу, док DP/ODP пружа математичку гаранцију да чак ни ажурирања модела не могу открити податке појединаца. Применом ODP на ажуриране тежине модела током процеса агрегације, спречава се да сервер може из модела закључити осетљиве информације о појединачним тренинг подацима. Ова комбинација је неопходна за изградњу модерних система машинског учења који су у потпуности усклађени са регулативама за заштиту података.

2.3 Радно окружење *TensorFlow Federated* (TFF)

TensorFlow Federated (TFF) је радно окружење отвореног кода које је развила компанија *Google* са циљем да олакша истраживање и примену FL у реалним окружењима [5]. Поред имплементације постојећих FL алгоритама, радно окружење TFF пружа и модуларни оквир за истраживање и дизајн нових дистрибуираних алгоритама. Главна предност је подршка за симулирање FL алгоритама користећи *TensorFlow* (TF) моделе над симулираним уређајима.

Архитектура TFF-а је подељена на два главна слоја [5]:

- Федеративни модел API (*Federated Model API*): слој намењен истраживачима и инжењерима машинског учења. Пружа скуп модула високог нивоа који омогућавају кориснику да постојеће моделе изграђене у TF-у, укључујући и *Keras* моделе, лако прилагоди за FL окружење. Главна функција овог слоја је аутоматско генерисање FL алгоритама попут федеративног просека (*FedAvg*) из стандардног модела машинског учења.
- Федеративни основни API (*Federated Core API*): темељ TFF-а, намењен за имплементацију нових FL алгоритама. Федеративни основни API је декларативни језик који омогућава експлицитно дефинисање дистрибуираних рачунарских операција. Овај језик је дизајниран за рад са федеративним типовима података, као што су подаци на клијентима (CLIENTS) и подаци на серверу (SERVER). Све комуникационе операције, као што су агрегација и дифузија (слање модела), дефинишу се експлицитно, омогућавајући потпуну контролу над током података и комуникационом ефикасношћу.

Кључна карактеристика TFF-а је његов систем федеративних типова. За разлику од библиотеке *TensorFlow* која ради само са тензорима (низовима података), TFF уводи концепт локације података. На пример, подаци који се налазе на свим клијентима дефинисани су као {тип}@CLIENTS, док су подаци који се налазе на централном серверу дефинисани као {тип}@SERVER. Јасно раздвајање локације података осигурава да се подаци не могу случајно преместити са приватне локације на централни сервер, чиме се пружа основна гаранција приватности.

Радно окружење TFF-а нуди подршку за учитавање *Keras* модела због њихове једноставности и популарности. TFF користи функцију `tff.learning.from_keras_model` за претварање *Keras* модела у `tff.learning.Model` интерфејс. Овај интерфејс омогућава коришћење традиционалног централизованог тренирања са дистрибуираним FL протоколима. Претворени модел се може користити за креирање комплетног федеративног итеративног процеса, који дефинише иницијализацију (`initialize`) и један корак тренирања (`next`) у FL циклусу.

Механизам конверзије и адаптације *Keras* модела је кључан за додавање подршке за верзију 3 библиотеке *Keras*. Постојећа архитектура `from_keras_model` је чврсто везана за архитектуру *Keras* 2 библиотеке. Због тога је за подршку новој верзији библиотеке *Keras* неопходно рефакторисање компоненти TFF-а које управљају стањем модела, оптимизаторима и метрикама, како би TFF могао да искористи предности новијих модела за машинско учење.

2.4 Библиотека *Keras*

Keras је библиотека отвореног кода за машинско учење која служи као интерфејс високог нивоа, омогућавајући једноставно коришћење сложених модела. Примарна сврха је једноставно коришћење популарних серверских радних окружења (*backend*) за машинско учење, као што су *TensorFlow*, *JAX* и *PyTorch*. *Keras* пружа конзистентан и једноставан апликативни програмски интерфејс (*Application Programming Interface*, API), који апстрахује сложеност *backend* имплементације, што омогућава истраживачима и инжењерима да се фокусирају на дизајн модела [11]. Такође, *Keras* подржава учитавање и тренирање модела са екстерних платформи, попут попут платформе *Kaggle*-а [12].

Док је верзија 2 библиотеке *Keras* било чврсто везан за *TensorFlow* као свој примарни *backend*, верзија 3 библиотеке *Keras* донела је темељну архитектонску промену. Најновија верзија имплементације библиотеке *Keras* је у потпуности редизајнирана да постане библиотека која подржава више *backend* окружења (*multi-backend*), чиме се пружа флексибилност у избору најбољег окружења за специфичан хардвер или задатак. Због ових промена, *Keras* 3 је морао да редефинише како се интерно рукује компонентама попут оптимизатора и метрика.

Главни значај библиотеке *Keras* 3 за федеративно учење и овај рад лежи у његовим побољшањима перформанси и скалабилности. *Keras* 3 је омогућио подршку за учитавање и тренирање модела са више милијарди параметара, као што је модел *Llama* 3 [13]. Нови модели имају знатно боље перформансе у поређењу са претходним моделима са мањим бројем параметара. Немогућност коришћења ових напредних модела у TFF-у била је главна мотивација за рефакторисање радног окружења.

Архитектонска промена на *multi-backend* у библиотеци *Keras* 3 има директан утицај на имплементацију TFF-а јер је библиотека *Keras* 3 увела *stateless* парадигму за кључне компоненте. *Stateless* парадигма представља систем који не чува податке о претходним интеракцијама са корисником, док *statefull* парадигма чува претходне интеракције са корисником. У библиотеци *Keras* 3, оптимизатори више нису директно *statefull*, већ користе методе попут `stateless_apply()`. Ово представља проблем код FL, јер TFF мора експлицитно да рукује променљивим стањима оптимизатора (*optimizer slots*) приликом агрегације и дистрибуције.

Слично променама у оптимизаторима, и метрике (Metrics) су прешле на *stateless* приступ (нпр. функција `stateless_update_state()`), што је променило начин на који TFF издваја и агрегира нефинализоване метрике са клијената. Постојећа функција `from_keras_model` у TFF-у ослањала се на имплементацију *Keras 2* модела за трансформацију модела, које су у библиотеци *Keras 3* комплетно промењене. Функција би морала бити модификована, да би подржала нове структуре и понашања нове верзије библиотеке *Keras*.

Тренутна верзија TFF-а подржава само *Keras 2*, што значајно ограничава коришћење новијих, напреднијих модела. Додавањем подршке за *Keras 3*, омогућило би се тренирање над најновијим моделима, чиме би се знатно побољшало предвиђање претраге (SE) унутар IDE-а, што је крајњи циљ примене федеративног учења у овом домену.

Глава 3

Имплементација подршке за *Keras 3*

3.1 Анализа имплементације проблема компатибилности

Архитектонска промена библиотеке *Keras* из верзије 2 у верзију 3, описана у поглављу Библиотека *Keras*, је разлог проблема компатибилности библиотеке унутар радног окружења TFF. Функција `tff.learning.from_keras_model`, била је дизајнирана да очекује и користи унутрашње API позиве специфичне за *Keras 2* (нпр. за директно издавање *statefull* променљивих оптимизатора). Пошто *Keras 3* не подржава ове позиве, директна интеграција је постала немогућа.

Анализа TFF кода показала је да би потпуна миграција целог радног окружења TFF-а на *Keras 3* архитектуру била претерано сложена и ризиковала би нарушавање функционалности постојећег, верификованог кода који се ослања на библиотеку *Keras 2*. Стога је донета одлука да се имплементира паралелна подршка за обе верзије библиотеке *Keras*. Овај приступ је захтевао да се TFF функције модификују тако да могу да рукују и са *Keras 2* и са *Keras 3* компонентама истовремено.

Паралелна подршка за обе верзије библиотеке *Keras* подразумева два примарна техничка захтева који су формирали оквир рефакторисања:

- Подршка за унију типова: Функције у TFF-у морају бити у стању да прихвате објекте који су или *Keras 2* или *Keras 3* типови. Ово је решено имплементацијом уније типова за улазне параметре модела, метрика и функција грешке.
- Апстракција логике руковања моделом Креирање засебног механизма који би, на основу верзије *Keras* објекта, вратио одговарајуће променљиве (нпр. из *stateful Keras 2* или *stateless Keras 3* API-ја). Ова потреба је директно довела до развоја посебне компоненте `keras_compat`.

Рефакторисање се фокусирало на додавање логике компатибилности у кључне TFF модуле (*models*, *optimizers*, *metrics*), омогућавајући рад са обе верзије библиотеке *Keras*, чиме је изазов миграције претворен у изазов апстракције.

3.2 Архитектура рефакторисања

Већина рефакторисаног кода се налази унутар сегмента радног окружења TFF који је имплементиран у програмском језику *Python* [14]. Унутар *Python* имплементације се налази секција *learning* у којој се налази код одговоран за алгоритме за федеративно учење, као и модули *models*, *optimizers* и *metrics*, који рукују понашањем модела током обучавања [5].

Модул *models* представља најважнију тачку за интеграцију *Keras* модела у TFF. Његова примарна улога енкапсулација логики *Keras* модела у TFF-ов интерфејс *tff.learning.Model*. TFF интерфејс дефинише потребне атрибуте и методе које су кључне за FL процес, укључујући: обучиве променљиве (*trainable variables*), необучиве променљиве (*non-trainable variables*), променљиве стања модела (*model state variables*), методу *forward_pass* за прослеђивање података кроз модел и методу *report_local_unfinalized_metrics* за прикупљање локалних статистика [5].

Модул *optimizers* садржи логику неопходну за дефинисање и управљање оптимизаторима клијената (*client optimizers*). У федеративном учењу, оптимизатори имају клијентску и серверску улогу. На клијентској страни се користе за израчунавање локалних ажурирања модела на основу локалних података. Њихово стање, које укључује променљиве попут импулса (*momentum*) или адаптивних стопа учења (*adaptive learning rates*) мора бити укључено у стање сервера (*server state*). Овај модул обезбеђује да TFF може исправно да серијелизује и десеријелизује стање оптимизатора приликом дистрибуције и агрегације, омогућавајући континуирано учење током FL рунди [5]. Промене у понашању *Keras* 3 оптимизатора директно утичу на то како овај модул приступа њиховим променљивим стањима.

Модул *metrics* је одговоран за израчунавање, прикупљање и агрегацију статистичких мера перформанси модела, као што су тачност (*accuracy*) и губитак (*loss*) током процеса федеративног учења. У радном окружењу TFF, метрике се деле на две фазе: нефинализоване метрике и финализоване метрике. Клијенти израчунавају нефинализоване метрике (сирове вредности, као што су укупна сума губитака и број обрађених примера) и враћају их серверу. Овај модул садржи логику за агрегацију нефинализованих метрика преко свих клијената, као и логику за њихову финализацију (нпр. израчунавање просечне вредности) [5]. Прелазак *Keras* 3 метрика на *stateless* парадигму је захтевао рефакторисање модула, како би се обезбедило исправно извлачење и враћање нефинализованих тензора неопходних за централизовану агрегацију на серверу.

3.3 Апстракција логики руковања моделом

Апстракција логики руковањем моделом је неопходна за несметано руковање различитим верзијама библиотеке *Keras*. Будући да постоје измене у руковању своства

модела између верзија библиотеке, потребно је увести компоненту која ће на основу библиотеке прикладно руковати са имплементацијом модела. Креирана је компонента `keras_compat` која садржи функције наведене у табели 1.

Табела 1: Функције компоненте `keras_compat` за апстракцију руковања моделом

Назив функције	Сврха функције
<code>is_compiled</code>	Проверава да ли је модел претходно компајлиран.
<code>keras_dtype_to_tf</code>	Претвара Keras типове података у TensorFlow типове података.
<code>is_keras3</code>	Проверава да ли објекат потиче из имплементације библиотеке Keras 3.
<code>get_optimizer_variables</code>	Добавља променљиве асоциране са оптимизатором (<i>optimizer</i>).
<code>ref</code>	Враћа референцу на променљиву која може бити Keras или TensorFlow типа.
<code>get_variable</code> и <code>get_variables</code>	Добавља променљиве Keras или TensorFlow модела.
<code>clone_model</code>	Враћа копију прослеђеног модела.
<code>int_shape</code>	Враћа облик модела као тип <i>int</i> .

Функција `is_compiled` служи да провери да ли је прослеђени модел компајлиран, што је неопходан предуслов за започињање процеса машинског учења. Због разлика у интерним API позивима, функција мора експлицитно да разликује верзије, где се + за објекте типа `tf.keras.Model` (имплементација Keras 2) провера врши преко приватног атрибута `model._is_compiled`, док се за објекте типа `keras.Model` (имплементација Keras 3) користи директан атрибут `model.compiled`. Функција осигурава да радно окружење TFF увек може да потврди спремност модела за рад, без обзира на верзију Keras библиотеке.

Функција `keras_dtype_to_tf` осигурава конзистентности типова података (*data types*) у радном окружењу TFF. TFF је изграђен на библиотеци *TensorFlow*, која захтева специфичне `tf.dtype` објекте. Пошто Keras може да излаже типове података као једноставне *string* репрезентације (нпр. “float32”), функција мапира *string* вредности у њихове одговарајуће `tf.dtype` објекте. Поред стандардних типова, функција такође подржава и новије типове података уведене у библиотеци *TensorFlow*, као што су `float8_e4m3fn` и `float8_e5m2`, осигуравајући компатибилност са хардверски оптимизованим типовима.

Функција `is_keras3` утврђује да ли прослеђени објекат (нпр. модел, променљива, метрика) потиче из Keras 3 библиотеке. Провера се врши коришћењем `isinstance` методе уграђене у програмски језик *Python* над унијом различитих Keras 3 класа као што су `keras.Model`, `keras.Variable`, `keras.Metric`, `keras.Layer`. Провера омогућава осталим функцијама компоненте да изаберу исправан API позив.

Функција `get_optimizer_variables` решава проблем некомпатибилности у приступу променљивама оптимизатора између различитих верзија библиотеке *Keras*. У верзији 2, где су оптимизатори често имплементирани као `OptimizerV2` или сличне класе, атрибут `optimizer.variables` је функција (*callable*), која се мора позвати ради добијања листа променљивих. Насупрот томе, у верзији 3, `optimizer.variables` је директни атрибут. Функција апстрахује разлику и ако је `variables` функција, она је позива, у супротном, враћа атрибут директно, чиме се осигурава пренос стања оптимизатора током FL рунди.

Функција `ref` служи за добијање јединствене референце (ID) за променљиву, што је важно за интерно праћење стања у радном окружењу TFF. За `tf.Variable` (који се типично користи у имплементацији Keras 2), функција користи `variable.ref()` да би добила *TensorFlow* референцу. Међутим, за имплементацију Keras 3 променљивих (`keras.Variable` или `keras_common.KerasVariable`), које су апстрахованије, функција враћа Python ID објекта користећи `id(variable)`. Овај механизам омогућава TFF-у да поуздано идентификује и мапира променљиве без обзира на њихов основни тип.

Функције `get_variable` и `get_variables` служе за издвајање стварне вредности променљиве. У Keras 3, променљиве су представљене класама које имају атрибут `variable.value`, који садржи сам тензор. Функција `get_variable` проверава да ли је у питању имплементација Keras 3 променљиве и, уколико јесте, враћа њену вредност. У супротном, враћа саму променљиву (што важи за `tf.Variable` у имплементацији Keras 2). Функција `get_variables` је помоћна функција која примењује `get_variable` на колекцију променљивих (листу или торку), обезбеђујући да радно окружење TFF увек рукује само са сировим тензорским вредностима.

Функција `clone_model` је одговорна за стварање копије модела са новим, неиницијализованим тежинама, што је неопходно у FL окружењу (нпр. приликом иницијализације клијентских модела). Будући да Keras 3 и Keras 2 имају засебне модуле за клонирање модела (`keras.models.clone_model` и `tf_keras.models.clone_model`), функција користи функцију `is_keras3` да одреди коју верзију функције за клонирање треба позвати, чиме се осигурава да радно окружење TFF поуздано реплицира модел за сваког клијента, одржавајући исту архитектуру уз нове тежине.

Функција `int_shape` враћа облик (*shape*) прослеђеног тензора или променљиве као торку целих бројева. Функција је стандардна помоћна функција која се користи у

нижим слојевима машинског учења за ефикасно руковање димензијама, обезбеђујући да се димензије, које могу бити враћене као *TensorFlow* објекти, увек буду претворене у стандардни имплементацију торке (*tuple*) у програмском језику *Python*, чиме се поједностављује логика за проверу типова унутар радног окружења TFF.

Глава 4

Закључак

У закључку дајте кратак преглед онога шта урађено, са освртом на проблеме који су решени, предности и мане решења и правце даљег развоја.

Списак слика

Слика 1	Примена FL за предвиђање корисничког уноса на <i>Gboard Android</i> тастатури. Мобилни уређај локално тренира над претходним корисничким уносима и ажуриране тежине шаље централном серверу (А). Централни сервер агрегира примљене тежине (В) и шаље свим уређајима побољшан дељени модел (С) [6]	4
---------	--	---

Списак листинга

Списак табела

Табела 1 Функције компоненте `keras_compat` за апстракцију руковања моделом . . 11

Биографија

Овде написати своју кратку биографију.

Литература

- [1] JetBrains, „IntelliJ IDEA“. Приступљено: 12. Октобар 2025. [На Интернету]. Available at: <https://www.jetbrains.com/idea>
- [2] E. Union, „Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)“. Приступљено: 12. Октобар 2025. [На Интернету]. Available at: <https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng>
- [3] G. Cloud, „Federated learning: a guide to what it is and how it works“. Приступљено: 12. Октобар 2025. [На Интернету]. Available at: <https://cloud.google.com/discover/what-is-federated-learning?hl=en>
- [4] M. A. I. U. D. T. Q. Maria Iqbal Asadullah Tariq, „FL-ODP: An Optimized Differential Privacy Enabled Privacy Preserving Federated Learning“. Приступљено: 12. Октобар 2025. [На Интернету]. Available at: <https://ieeexplore.ieee.org/document/10287349>
- [5] Google, „TensorFlow Federated: Machine Learning on Decentralized Data“. Приступљено: 12. Октобар 2025. [На Интернету]. Available at: <https://www.tensorflow.org/federated>
- [6] R. M. S. R. F. B. S. A. H. E. C. K. D. R. Andrew Hard Kanishka Rao, „Federated Learning for Mobile Keyboard Prediction“. Приступљено: 12. Октобар 2025. [На Интернету]. Available at: <https://arxiv.org/abs/1811.03604>
- [7] C. Dwork, „Differential Privacy“. Приступљено: 16. Октобар 2025. [На Интернету]. Available at: https://doi.org/10.1007/11787006_1
- [8] A. R. Cynthia Dwork, *The Algorithmic Foundations of Differential Privacy*. University of Pennsylvania, 2014.
- [9] I. G. B. M. I. M. K. T. L. Z. Martín Abadi Andy Chu, „Deep Learning with Differential Privacy“. Приступљено: 19. Октобар 2025. [На Интернету]. Available at: https://link.springer.com/chapter/10.1007/978-3-540-79228-4_1
- [10] E. D. C. U. C. L. D. A. T. I. Mohammad Naseri Jamie Hayes, „Local and Central Differential Privacy for Robustness and Privacy in Federated Learning“. Приступљено: 12. Октобар 2025. [На Интернету]. Available at: <https://arxiv.org/abs/2508.10000>

пљено: 19. Октобар 2025. [На Интернету]. Available at: <https://arxiv.org/abs/2009.03561>

- [11] Keras, „Keras: Deep Learning for humans“. Приступљено: 21. Октобар 2025. [На Интернету]. Available at: <https://keras.io/>
- [12] Kaggle, „Kaggle: Your Machine Learning and Data Science Community“. Приступљено: 21. Октобар 2025. [На Интернету]. Available at: <https://www.kaggle.com/>
- [13] Meta, „Open-source AI Models for Any Application | Llama 3“. Приступљено: 21. Октобар 2025. [На Интернету]. Available at: <https://www.llama.com/models/llama-3>
- [14] Python, „The official home of the Python Programming Language“. Приступљено: 22. Октобар 2025. [На Интернету]. Available at: <https://www.python.org/>