ОБАВЕШТЕЊЕ О ОДБРАНИ МАСТЕР РАДА

Име:	Невена
Презиме:	Денић
Број индекса:	120PH
Департман:	Рачунарске науке
Тема мастер рада:	Хибридно дифузно и ауторегресивно тренирање језичких модела за унапређено учење слабије заступљених језика
Ментор:	Проф. др Бранимир Тодоровић
Датум одбране:	29.10.2026.
Време одбране:	17h
Место одбране:	учионица 140

Датум:	**Потпис студента:**
23.10.2026.	Невена Денић

ИЗЈАВА

Студент: Невена Денић

Број индекса: 120РН

Студијски програм: Вештачка интелигенција и машинско учење

Наслов мастер рада: Хибридно дифузно и ауторегресивно тренирање језичких модела за унапређено учење слабије заступљених језика

Ментор мастер рада: Проф. др Бранимир Тодоровић

Изјављујем да без сагласности ментора резултати мастер рада неће бити публиковани у стручном или научном часопису нити саопштени на научном скупу/конференцији.

У Нишу, 23.10.2026.

Потпис

Невена Денић

Прилог 5/1

ПРИРОДНО - МАТЕМАТИЧКИ ФАКУЛТЕТ
НИШ

КЉУЧНА ДОКУМЕНТАЦИЈСКА ИНФОРМАЦИЈА

Редни број, **РБР**:	
Идентификациони број, **ИБР**:	
Тип документације, **ТД**:	монографска
Тип записа, **ТЗ**:	текстуални / графички
Врста рада, **ВР**:	Мастер рад
Аутор, **АУ**:	Невена Денић
Ментор, **МН**:	Бранимир Тодоровић
Наслов рада, **НР**:	Хибридно дифузно и ауторегресивно тренирање језичких модела за унапређено учење слабије заступљених језика
Језик публикације, **ЈП**:	српски
Језик извода, **ЈИ**:	енглески
Земља публиковања, **ЗП**:	Р. Србија
Уже географско подручје, **УГП**:	Р. Србија
Година, **ГО**:	2026.
Издавач, **ИЗ**:	ауторски репринт
Место и адреса, **МА**:	Ниш, Вишеградска 33.
Физички опис рада, **ФО**: (поглавља/страна/ цитата/табела/слика/графика/прилога)	69 стр. ; граф. прикази
Научна област, **НО**:	Рачунарске науке
Научна дисциплина, **НД**:	Машинско учење
Предметна одредница/Кључне речи, **ПО**:	Језички модели, Мултијезични језички модели, Маскирано дифузно тренирање, Ауторегресивно тренирање, Језици са ограниченим ресурсима
УДК	004.8:81'322
Чува се, **ЧУ**:	библиотека
Важна напомена, **ВН**:	

Извод, **ИЗ**:		Овај мастер рад бави се комбиновањем ауторегресивног и дифузионог приступа у тренирању језичких модела. Испитује се како различит однос ова два начина тренирања утиче на способност модела да генерализује при вишеструком понављању истог скупа података, како за појединачне језике, тако и за вишејезичке скупове података и језике са ограниченим ресурсима. Циљ рада је да се испита како различити фактори утичу на избор вредности параметра α, који одређује однос ауторегресивног и дифузионог начина тренирања. Посебна пажња биће посвећена утицају количине података, броја њихових понављања и језичке разноврсности при избору оптималне вредности α.
Датум прихватања теме, **ДП**:		01.10.2026.
Датум одбране, **ДО**:		
Чланови комисије, **КО**:	Председник:	Проф. др Марко Петковић
	Члан:	Проф. др Марко Миладиновић
	Члан, ментор:	Проф. др Бранимир Тодоровић

Прилог 5/2

ПРИРОДНО - МАТЕМАТИЧКИ ФАКУЛТЕТ
НИШ

KEY WORDS DOCUMENTATION

Accession number, **ANO**:	
Identification number, **INO**:	
Document type, **DT**:	monograph
Type of record, **TR**:	textual / graphic
Contents code, **CC**:	Master's thesis
Author, **AU**:	Nevena Denić
Mentor, **MN**:	Branimir Todorović
Title, **TI**:	Hybrid Diffusion and Autoregressive Training of Language Models for Improved Learning of Low-Resource Languages
Language of text, **LT**:	Serbian
Language of abstract, **LA**:	English
Country of publication, **CP**:	Republic of Serbia
Locality of publication, **LP**:	Serbia
Publication year, **PY**:	2026
Publisher, **PB**:	author's reprint
Publication place, **PP**:	Niš, Višegradska 33.
Physical description, **PD**: (chapters/pages/ref./tables/pictures/graphs/appendixes)	69 p. ; graphic representations
Scientific field, **SF**:	Computer Science
Scientific discipline, **SD**:	Machine learning
Subject/Key words, **S/KW**:	Language models, Multilingual language models, Diffusion training, Autoregressive training, Low-resource languages
UC	004.8:81'322
Holding data, **HD**:	library
Note, **N**:	
Abstract, **AB**:	This master's thesis deals with combining autoregressive and diffusion approaches in the training of language models. It examines how different ratios of these two training methods affect a model's ability to generalize when the same dataset is repeated multiple times, both for individual languages and for multilingual datasets and low-resource languages. The aim of the thesis is to examine how various factors influence the choice of the parameter α, which determines the ratio of autoregressive to diffusion training. Particular attention is given to the influence of data quantity, the number of data repetitions, and linguistic diversity on the choice of the optimal value of α.

Accepted by the Scientific Board on, **ASB**:		01.10.2026.
Defended on, **DE**:		
Defended Board, **DB**:	President:	Dr Marko Petković
	Member:	Dr Marko Miladinović
	Member, Mentor:	Dr Branimir Todorović