

**UNIVERSITETI I PRISHTINËS “HASAN PRISHTINA”
FAKULTETI I SHKENCAVE MATEMATIKE-NATYRORE**

**DEPARTAMENTI I MATEMATIKËS
PROGRAMI SHKENCË KOMPJUTERIKE**



TEZA MASTER

**DETEKTIMI I PARAGJYKIMEVE DHE STEREOTIPEVE GJINORE NË
RAPORTIMIN E MEDIAVE PËR DHUNËN NDAJ GRAVE NË KOSOVË
PËRMES PËRPUNIMIT TË GJUHËS NATYRORE (NLP)**

Mentori:

Prof. Asoc. Dr. Eliot Bytyci

Kandidati:

BSc. Trime Ismajli

**Prishtinë
Maj, 2026**

ABSTRAKT I ZGJERUAR

Dhuna ndaj grave mbetet një nga problemet më serioze shoqërore, institucionale dhe gjinore, jo vetëm për shkak të pasojave të drejtpërdrejta që ka mbi viktimat, familjet dhe komunitetin, por edhe për shkak të mënyrës se si ajo kuptohet, interpretohet dhe diskutohet në hapësirën publike. Në këtë proces, mediat kanë rol të veçantë sepse ato nuk shërbejnë vetëm si burim informacioni për ngjarjet, por edhe si hapësirë ku ndërtohet kuptimi shoqëror i dhunës. Përmes fjalëve që zgjedh, detajeve që thekson dhe atyre që i lë në heshtje, raportimi mediatik mund ta vendosë përgjegjësinë te autori ose, përkundrazi, ta zhvendosë drejt viktimës; mund ta paraqesë dhunën si problem strukturor ose ta zbehë në një konflikt të rastit.

Qëllimi kryesor i këtij punimi është të analizojë mënyrën se si mediat online në Kosovë raportojnë për dhunën ndaj grave dhe të vlerësojë mundësitë e përdorimit të metodave të përpunimit të gjuhës natyrore (ang. natural language processing – NLP) për identifikimin e tematikave dominuese, paragjyqimeve dhe stereotipeve gjinore në tekstet mediatike në gjuhën shqipe.

Për këtë qëllim është ndërtuar një dataset origjinal me 3,078 artikuj në gjuhën shqipe, të mbledhur nga pesë portale online në Kosovë që, gjatë periudhës së mbledhjes së të dhënave, figuronin ndër mediat me audiencën më të lartë në vend: Telegrafi, Klan Kosova, Koha, Indeksonline dhe Kallxo. Të dhënat mbulojnë periudhën nga viti 2020 deri në fillim të vitit 2026. Çdo artikull u pastrua, u strukturua dhe u etiketua manualisht sipas llojit të dhunës (fizike, psikologjike ose vrasje/femicid) dhe sipas pranisë apo mungesës së paragjykimit gjinor. Ndërtimi i këtij dataseti të etiketuar në shqip përbën një ndër kontributet kryesore të punimit, pasi i përgjigjet drejtpërdrejt boshllëkut të identifikuar në literaturë për mungesën e të dhënave të strukturuar dhe të etiketuara për këtë temë.

Analiza përshkruese e të dhënave nxori në pah disa modele domethënëse në mënyrën se si mediat online në Kosovë raportojnë për dhunën ndaj grave. Raportimi rezultoi i përqendruar kryesisht te portali Kallxo, i cili përbën rreth 61% të artikujve të përfshirë në dataset. Sa i përket llojit të dhunës, kategoria më e raportuar ishte dhuna fizike (67%), e ndjekur nga dhuna psikologjike (25%) dhe vrasja/femicidi (8%). Paragjykimi i hapur gjinor rezultoi relativisht i rrallë, i pranishëm në vetëm 3.8% të artikujve; ndonëse në rastet ku shfaqej, materializohej kryesisht përmes formulimeve që mund ta zbusin përgjegjësinë e autorit. Grupi më i shpeshtë i termave justifikues ishte “alkool/dehje”, i pranishëm në pjesën më të madhe të rasteve të justifikimit, duke e zhvendosur vëmendjen nga përgjegjësia e kryerësit drejt rrethanave shoqërore.

Punimi ndjek dy drejtime kryesore analitike. Drejtimi i parë lidhet me analizën tematike të artikujve, ku janë përdorur dy metoda të modelimit të temave: modeli klasik Latent Dirichlet Allocation (LDA) dhe modeli modern BERTopic. Qëllimi ishte të vlerësohej nëse artikujt grupohen natyrshëm sipas llojit të dhunës; mirëpo rezultatet treguan se kjo nuk ndodh. Artikujt priren të grupohen sipas burimit institucional dhe fazës procedurale të raportimit, si policia, prokuroria dhe gjykata. Të dyja metodat shfaqën përputhje shumë të ulët me etiketimin manual; megjithatë, LDA u afrua pak më shumë me kategoritë e vendosura nga njeriu, ndërsa BERTopic prodhoi tema disi më koherente nga ana kuptimore.

Drejtimi i dytë i punimit lidhet me detektimin automatik të paragjykimeve gjinore. Fillimisht u testua qasja zero-shot me modelin shumëgjuhësor mDeBERTa-v3, ku modeli u përdor për të klasifikuar artikujt pa u trajnuar paraprakisht në shembuj të etiketuar në gjuhën shqipe. Rezultatet treguan performancë të kufizuar: edhe pas optimizimit të etiketave dhe pragut të vendimit (ang. decision threshold), modeli nuk arriti aftësi të qëndrueshme dalluese, me ROC-AUC rreth 0.52–0.55. Kjo tregon se sjellja e tij ishte afër një klasifikuesi të rastësishëm, me tendencë për të etiketuar gabimisht një numër të madh artikujsh neutralë si paragjykes.

Në fazën tjetër, i njëjti model u rafinua (ang. fine-tuning) mbi datasetin e ndërtuar në këtë punim. Rafinimi i thjeshtë e bëri modelin tepër agresiv: kapte të gjitha rastet paragjyuese, por sinjalizonte gabimisht pothuajse çdo artikull neutral. Rafinimi me funksion humbjeje të peshuar (ang. weighted loss) ra në ekstremin e kundërt: arriti saktësi mbi 96%, por nuk identifikoi asnjë rast të vetëm paragjykimi, një kujtesë se në të dhëna kaq të pabalancuara saktësia e lartë mund të jetë mashtruese. Konfigurimi i tretë me mbikampionim të klasës minoritare (ang. minority-class oversampling) dha performancën më të mirë, me ROC-AUC = 0.741 dhe F1-score = 0.207, duke ofruar balancën më të dobishme ndërmjet kapjes së rasteve reale dhe shmangies së alarmeve të rreme.

Këto rezultate e konfirmojnë hipotezën themelore të punimit: adaptimi në domen prodhon sinjal dukshëm më të dobishëm sesa transferimi i drejtpërdrejtë nga gjuhë të tjera për detyra kaq specifike dhe të ndjeshme në gjuhën shqipe. Megjithatë, performanca absolute mbetet modeste, gjë që lidhet me rrallësinë e klasës me paragjykim, subjektivitetin e vetë detyrës dhe burimet ende të kufizuara të NLP-së për shqipen. Për këtë arsye, detektimi plotësisht automatik nuk mund të konsiderohet ende i mjaftueshëm, dhe vlerësimi njerëzor mbetet thelbësor si plotësues i analizës kompjuterike.

Kontributi kryesor i punimit është i dyfishtë. Së pari, ai ofron një analizë sistematike të mënyrës se si mediat online në Kosovë raportojnë për dhunën ndaj grave, duke identifikuar modele të dukshmërisë, mungesës së kontekstit dhe formave të mundshme të paragjykimit gjinor. Së dyti, ai ndërton dhe teston një metodologji kompjuterike të replikueshme për analizën e teksteve shqipe, duke krahasuar qasje klasike dhe moderne të NLP-së. Kështu, punimi kontribuon si në fushën e studimeve gjinore dhe mediatike, ashtu edhe në atë të shkencës kompjuterike dhe përpunimit të gjuhës natyrore.

Fjalë kyçe: dhuna ndaj grave, paragjykime gjinore, media online, Kosovë, përpunim i gjuhës natyrore (NLP), BERTopic, mDeBERTa-v3.