TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Margareth Lasn 231875IABM
Hanna-Kristella Lehtsaar 231850IABM

Eesti tuju ajas. Eestikeelsete igapäevaartiklite meelestatusanalüüsi tööriista arendus Eesti Rahvusraamatukogu digilaborile.

Magistritöö

Juhendaja: Martin Verrev
MSc

Tallinn 2026

Autorideklaratsioon

Kinnitame, et oleme koostanud antud lõputöö iseseisvalt ning seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on töös viidatud.

Autorid: Margareth Lasn ja Hanna-Kristella Lehtsaar

17.05.2026

Annotatsioon

Käesoleva magistritöö eesmärk oli koostöös Eesti Rahvusraamatukoguga (RaRa) luua eestikeelne meelestatusanalüüsi tööriist RaRa digilabori veebikeskkonda. Töö keskendus vabavaraliste keelemudelite võrdlemisele ja hindamisele tekstide klassifitseerimisel positiivseks, negatiivseks või neutraalseks. Eksperimentides kasutati erinevaid BERT-tüüpi mudeleid ning EstLLM suurt keelemudelit.

Lisaks käsitleti töös ajalooliste OCR-tekstide parandamist, kuna OCR-vead võivad mõjutada meelestatusanalüüsi täpsust. Selleks töötati välja EstLLM-il põhinev OCR-vigade parandamise komponent.

Töö sisaldab teoreetilist tausta, kasutatud metoodikate kirjeldust, keelemudelite eksperimentide ja tulemuste võrdlust ning veebipõhise kasutajaliidese arendust ja valideerimist.

Töö tulemusena valmis avalikult kasutatav meelestatusanalüüsi tööriist, mis sisaldab ka OCR-vigade parandamise komponenti. Tööriista abil viidi läbi Eesti ajakirjanduse meelestatusanalüüs aastate 1925-2025 lõikes ning tulemuste põhjal koostati blogipostitus „Eesti tuju ajas“.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 60 leheküljel, 9 peatükki, 17 joonist, 12 tabelit.

Abstract

Estonian mood over time - development of a sentiment analysis tool for Estonian-language newspaper articles for the National Library of Estonia

The aim of this master's thesis was to create an Estonian-language sentiment analysis tool in collaboration with the National Library of Estonia (RaRa) for the RaRa digital laboratory web environment. The work focused on comparing and evaluating free language models in classifying texts as positive, negative or neutral. The experiments used different BERT-type models and the EstLLM large language model.

In addition, the work dealt with the correction of historical OCR texts, as OCR errors can affect the accuracy of sentiment analysis. For this purpose, an OCR error correction component based on EstLLM was developed.

The work includes a theoretical background, a description of the methodologies used, a comparison of language model experiments and results, and the development and validation of a web-based user interface.

As a result of the work, a publicly available sentiment analysis tool was created, which also includes an OCR error correction component. The tool was used to conduct a sentiment analysis of Estonian journalism for the years 1925-2025, and a blog post “Estonian Mood in Time” was written based on the results.

The thesis is in Estonian and contains 60 pages of text, 9 chapters, 17 figures, 12 tables.

Lühendite ja mõistete sõnastik

Keelemudel	Loomulikku keelt kirjeldav tõenäosuslik mudel, mis põhineb keeleandmetele toetuvatel algoritmidel (nt statistiline mudel, närvivõrgu mudel) [1].
Modelfile	Plaan kohandatud mudelite loomiseks ja jagamiseks Ollama abil [2].
Morfoloogia	Grammatika osa, mis tegeleb sõnavormidega, nende moodustamisega ning nendest arusaamisega [3].
NLP	Loomuliku keele töötlus – informaatika, tehisintellekti ja arvutilingvistika ühisosa, mis tegeleb inimkeele analüüsi, mõistmise ja genereerimisega [4].
OCR	*Optical Character Recognition* on tehnoloogia, mis kasutab automaatset andmete eraldamist, et tekstipildid kiiresti masinloetavasse vormingusse teisendada [5].
RaRa digilabor	Eesti Rahvusraamatukogu osakond, mille eesmärk on teha raamatukogude valduses olevad andmed digitaalselt paremini kättesaadavamaks ja kasutatavamaks, edendada andmete väärindamist, teadust ja innovatsiooni [6].
Token	Teksti väikseim analüüsitav üksus NLP's. Token võib olla sõna, alamsõna, märk või sümbol, sõltuvalt kasutatavast tokeniseerimismeetodist ja mudelist [7].
Tokeniseerimine	Esimene samm loomuliku keele töötluses, kus segmenteeritakse sisendtekst tõkenditeks [7].
Transformer-arhitektuur	Süvaõppemudel, mis kasutab s*elf-attention* mehhanismi, et modelleerida seoseid sisendjärjestuse elementide vahel ning võimaldada efektiivset NLP ja tekstigeneratsiooni [8].
UML	*Unified Modeling Language* on standardiseeritud modeleerimiskeel, mis koosneb integreeritud diagrammidest, eesmärgiga aidata visualiseerida tarkvara-, äri- ja mitte-tehnilisi süsteeme [9].

Sisukord

Jooniste loetelu

Tabelite loetelu

1 Sissejuhatus

Viimase saja aasta jooksul on Eesti kogenud iseseisvumist, okupatsioone, nõukogude perioodi ning lõpuks kujunenud tänapäevaseks demokraatlikuks ja digitaalseks ühiskonnaks [10]. Kõik need sündmused on jätnud jälje ka ajakirjandusse.

Meedia ei kajasta ainult sündmusi, vaid peegeldab ka üldist õhkkonda ja meelestatust, kuid selle käsitsi analüüsimine on ajakulukas töö [11]. Sellest tulenevalt lõid autorid koostöös Eesti Rahvusraamatukoguga (edaspidi RaRa) avalikult kättesaadava meelestatusanalüüsi tööriista.

Töö eesmärgiks oli eksperimenteerida kaasaegsete keelemudelitega, et leida täpseim lahendus meelestatusanalüüsi ülesande teostamiseks. Töös kasutati masinõppel põhinevaid keelemudeleid, mis võimaldavad automaatselt tuvastada teksti meelestatust, klassifitseerides selle positiivseks, negatiivseks või neutraalseks.

Püstitati järgmised uurimisküsimused:

1. Milline vabavaraline keelemudel saavutab kõige täpseima tulemuse eesti keelses meelestatusanalüüsi ülesandes?

2. Kui täpselt suudab loodud tööriist määrata eestikeelsete tekstide meelestatust?

Töö eripäraks on teksti parandamise komponendi kaasamine meelestatusanalüüsi protsessi, mida varasematest eesti keelel põhinevatest töödest ei leia. Vanemad meedia väljaanded (näiteks 1925. aastast) võivad sisaldada OCR-vigu, mis mõjutavad meelestatusanalüüsi täpsust. Seetõttu rakendatakse töös lähenemist, kus enne meelestatuse analüüsimist viiakse vajadusel läbi teksti parandamine, et tagada usaldusväärsemad tulemused.

Tulenevalt OCR-vigade olemasolust testandmestikul püstitati lisauurimisküsimus:

3. Kui palju mõjutavad OCR-vead tööriista võimet tuvastada meelestatust?

1.1 Lähteülesanne ja alameesmärgid

Töö lähteülesandeks oli luua meelestatusanalüüsi tööriist RaRa digilaborile ning teostada seda kasutades andmeanalüüs, mille leidudest saab kirjutada blogipostituse. Lähteülesande täitmiseks püstitati järgnevad alameesmärgid:

1. tööriist peab suutma analüüsida kasutaja sisestatud teksti ning tuvastada selles väljendatud meelsust,
2. töö teostamisel tuleks eelistada vabavaralisi tehnoloogiaid (k.a. andmestikud), et tööriist oleks kasutatav RaRa digilabori keskkonnas,
3. välja töötatud tööriistal peab olema graafiline kasutajaliides.
4. tööriista toimivust tuleb demonstreerida andmeanalüüsi näite abil,
5. tööriista abil tuleb teostada meelestatusanalüüs autorite poolt valitud andmestikul ning kirjutada leidudest blogipostitus.

1.2 Töö struktuur

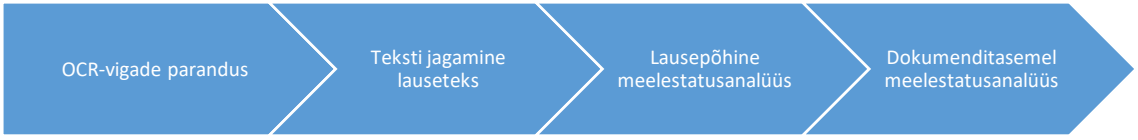
Käesolevas töös tutvustatakse esmalt autorite poolt kasutatud metoodikaid ja tööriistu (peatükis 2). Järgnevalt antakse ülevaade tööga seotud teoreetilisest taustast, mis hõlmab ka antud teemal koostatud varasemate tööde ülevaadet (peatükis 3). Töö praktilise osa alla kuuluvad teksti korrastamine (peatükk 4), mudeli valimine (peatükk 5), dokumenditaseme meelestatuse määramine (peatükk 6) ja kasutajaliidese arendus (peatükk 7). Peatükk 8 koondab kokku tööriista valideerimisviisid ning autorite poolsed järeldused saadud tulemustes.

2 Metoodika

2.1 Objekti kirjeldus

Käesoleva magistritöö objektiks on meelestatusanalüüsi tööriist, mis on loodud RaRa digilabori jaoks. Tööriista põhifunktsionaalsuste hulka kuuluvad teksti korrastamine ning meelestatusanalüüsi teostamine (vaata joonist 1). Kasutaja sisendi põhjal määrab tööriist tekstile ühe kolmest võimalikust meelestatusklassist: positiivne, negatiivne või neutraalne.

Tööriista sihtgrupiks on eesti keelt valdavad isikud, kes külastavad RaRa digilabori veebilehte. Tegemist on avalikult kättesaadava veebirakendusega, mis on mõeldud kasutamiseks laiale huviliste ringile. Töö osana loodi tööriistale ka võimekus töödelda faile, mis on suunatud RaRa siseseks kasutamiseks.



Joonis 1. Loodava tööriista töövoo joonis.

2.2 Tööprotsessi kirjeldus

Esialgsete OCR-vigade parandamise katsete läbiviimiseks edastasid RaRa esindajad väljaannete „Koit“ (23.10.1975) [12] ja „Postimees“ (23.10.1925) [13] tekstfailid. Neid kasutati OCR-vigade mõju ja parandamise võimaluste esmaseks analüüsiks.

Keelemudelitega eksperimenteerimise aluseks valiti Eesti valentsikorpuse andmestik, mis sisaldab ajalehtede „Postimees“ ja „Õhtuleht“ artiklitest ja kommentaaridest võetud lõike, millele on määratud meelestatusklass (positiivne, negatiivne, neutraalne ja vastuoluline) [14]. Andmestik valiti, kuna see oli töö koostamise ajal üks väheseid eestikeelseid märgendatud meelestatusandmestikke, mis sisaldas nii tekstisisu kui ka vastavat meelestatusklassi.

Eksperimentide läbiviimiseks jaotati andmestik kolmeks: 70% kirjetest mudeli treenimiseks, 10% mudeli valideerimiseks ning 20% testimiseks. Tegemist on andmejaotusega, mis võimaldas autoritel kasutada ühe andmestiku sisu nii treenimiseks kui ka testimiseks [15]. Antud lähenemine valiti, kuna töö koostamise ajal puudus eraldiseisev eestikeelne märgendatud andmestik mudelite sõltumatuks testimiseks.

Mudelite hindamisel kasutati peamise mõõdikuna F1-skoori ning seda täiendavaid mõõdikuid (*accuracy, precision, recall, macro avg* ja *weighted avg*). Tegemist on loomuliku keele töötluses laialdaselt kasutusel oleva mudelite hindamisraamistikuga, mistõttu valiti see ka selles töös täpseima mudeli väljaselgitamise aluseks [16].

Tööriista tervikliku kasutajakogemuse hindamiseks viidi läbi kvalitatiivne kasutajatestimine valjult mõtlemise (*think aloud*) meetodil. Meetodi eesmärk oli koguda kasutajatelt jooksvalt tagasisidet ning samaaegselt jälgida nende tegevusi ja mõtteprotsesse tööriista kasutamise ajal [17]. Tagasiside kogumisel lähtuti põhimõttest, et kui uue kasutajatestiga ei ilmnenud uusi leide, siis oli valideerimise protsess lõpule viidud [18].

Blogi postituse jaoks teostatud analüüsitulemusi võrreldi Excelis [19]. See võimaldas teha jooniseid, millest järeldusid aastavahemike vahelised trendid.

2.3 Tööriistade kirjeldus

Arenduskeskkonnana kasutati Visual Studio Code'i [20], milles teostati nii taga- kui ka esirakenduse arendus. Keelemudelitega eksperimenteerides kasutati Python programmeerimiskeelt [21], mis on laialdaselt kasutusel loomuliku keele töötlemise lahendustes [22]. Arenduse käigus kasutati ChatGPT 5.5 mudelit [23] programmeerimiskeelega seotud tehniliste küsimuste lahendamisel ja koodi analüüsimisel. Kasutades mainitud programmeerimiskeelt implementeeriti valminud tööriista tagarakenduse kood.

Kasutajaliidese loomisel kasutati veebitehnoloogiaid HTML [24], CSS [25] ja JavaScript [26]. Kuna tegemist on lihtsa veebirakenduse lahendusega piisas standardsetest veebitehnoloogiatest. Tagarakenduse koodiga suhtluse jaoks kasutati FastAPI raamistikku [27], kuna see liidestus sujuvalt Python koodil põhineva tagarakenduse lahendusega.

Keelemudelite kasutamine nõudis rohkem arvutusressursse, kui autorite isiklikud arvutid võimaldasid. Selle tõttu kasutati töö käigus GPU-toega arvutuskeskkondi. Esialgne katsetamine viidi läbi TalTech AI-Lab keskkonnas [28], mis võimaldas kasutada suurema arvutusvõimsusega GPU-ressursse. Hiljem viidi töövoog üle RunPodi pilveplatvormile [29], mis pakkus paindlikumat ja stabiilsemat töökeskkonda mudeli edasiseks kasutamiseks.

Andmeanalüüsiks kasutati Excel rakendust. See võimaldas andmeid töödelda ning visualiseerida erinevate diagrammide ja jooniste abil, et lihtsustada järelduste tegemist.

3 Teoreetiline taust

3.1 Meelestatusanalüüs (*sentiment analysis*) ja selle protsessi automatiseerimine

Meelestatusanalüüs on loomuliku keele töötluse valdkonna alamharu, mille eesmärk on tuvastada ja liigitada tekstis väljendatud hoiak või emotsioon. Tavapäraselt jaotatakse tekstid kolme põhikategooriasse: positiivne, negatiivne ja neutraalne. Meelestatusanalüüsil on oluline roll paljudes valdkondades, kuna avaliku arvamuse analüüs võimaldab teha informeeritud otsuseid [30]. Näiteks 2019 aasta UniCredit panga sotsiaalmeedia analüüs näitas, et meelestatusanalüüsi abil oli võimalik struktureerimata kasutajate tagasiside muuta kvantifitseeritavaks maineindikaatoriks. Analüüsi kaasati 953 inglise keelset *tweet*'i, mis koguti vahemikus 16.02.2016-12.07.2019, millest 499 klassifitseeriti positiivseks. Lisaks jälgiti ka millised märksõnad kordusid, et anda hinnang ettevõtte kommunikatsioonitegevuste mõjule ning tuvastada varakult võimalikud negatiivsed trendid. UniCrediti puhul kordusid positiivsetes *tweet*'ides märksõnad nende heategevuslike algatuste kohta, mis on selge kinnitus nende kommunikatsioonistrateegia toimimisele [31].

Meelestatusanalüüsi protsess koosneb mitmest järjestikusest etapist, millest olulisemad on eeltöötlus, tunnuste eraldamine ning klassifitseerimine [30].

Eeltöötluse käigus puhastatakse ja standardiseeritakse algne tekst, et vähendada andmetes esinevat müra. Selle käigus eemaldatakse ebaolulised elemendid, nagu stoppsõnad (näiteks side- ja asesõnad), erimärgid ja numbrid ning vajadusel viiakse sõnad ühtsele kujule. Samuti võib eeltöötlus hõlmata teksti normaliseerimist, näiteks suur- ja väiketähtede ühtlustamist. Nende tegevuste eesmärk on parandada andmete kvaliteeti ning suurendada analüüsi täpsust [30].

Tunnuste eraldamisel teisendatakse eeltöödeldud tekst masinõppe algoritmidele sobivale numbrilisele esituskujule. Selleks kasutatakse erinevaid meetodeid, näiteks sageduspõhiseid lähenemisi (nt *Bag-of-Words*), kaalutud meetodeid või semantilisi

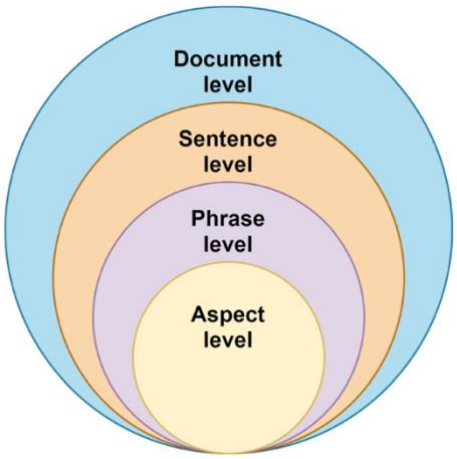
representatsioone, sealhulgas sõnavektoreid ja eelõpetatud keelemudeleid. Selle etapi eesmärk on säilitada teksti oluline informatsioon viisil, mis võimaldab mudelitel tuvastada mustreid ja seoseid [30].

Klassifitseerimisel määratakse tekstile vastav meelsuskategooria, näiteks positiivne, negatiivne või neutraalne. Protsessi tulemusena saadakse lõplik hinnang tekstis väljendatud hoiakule [30].

Meelestatusanalüüsis esineb ka probleeme. Näiteks on analüüsis raskusi tekitavaks inimeste mitteametlik ja varieeruv kirjutamisstiil, mis võib sisaldada slängi, lühendeid ja grammatilisi vigu. Samuti on raske tuvastada sarkasmi ja irooniat, mille puhul teksti otsene tähendus ei pruugi vastata tegelikule hoiakule [32].

Lisaks sõltuvad sageli tähendus ja emotsionaalne hoiak tugevalt kontekstist, mistõttu ei ole nende automaatne tõlgendamine alati üheselt määratav. Keelepõhised eripärad muudavad olukorra veelgi keerulisemaks, kuna kõigi keelte jaoks ei ole piisavalt ressursse ega tööriistu [32].

Meelestatusanalüüsi saab käsitleda neljal erineval tasemel: dokumendi, lause, fraasi ja aspekti (vt joonis 2) [32].



Joonis 2. Meelestatusanalüüsi tasemed [32].

Dokumenditasemel hinnatakse analüüsis kogu teksti tervikuna ning sellele omistatakse üks üldine meelestatus. Sellist lähenemist kasutatakse näiteks pikemate tekstide, nagu artiklite või raamatu peatükkide, üldise meelestatuse määramiseks. Meetodi eelis seisneb

selle lihtsuses ja võimes anda kiire ülevaade teksti üldisest hoiakust, kuid samas on piiratud, sest ei arvesta teksti sees esinevaid vastandlikke seisukohti või nüansse [32].

Lausetasemel käsitletakse igat lauset eraldi, määrates igale neist vastav meelestatus. Selline lähenemisviis on kasulik tekstide puhul, kus esineb mitmekesine ja vastandlik meelestatus, mida dokumenditasemel analüüs ei suuda adekvaatselt kajastada. Lausetasemel saadud tulemusi saab kasutada nii eraldiseisvalt kui ka agregeerida dokumendi üldise meelestatuse leidmiseks [32].

Fraasitasemel keskendutakse üksikutele väljenditele või sõnaühenditele, kus tuvastatakse konkreetset hinnangut kandvad sõnad või fraasid ning määratakse nende meelestatus. Erinevalt lausetaseme lähenemisest võimaldab fraasitase tuvastada, millise aspekti suhtes meelestatus väljendub. Fraasitasemel analüüs põhineb eeldusel, et sõnad ja nende kombinatsioonid kannavad meelestatust ning moodustavad teksti meelestatuse aluse. Selline lähenemine võimaldab tuvastada detailseid ja kontekstispetsiifilisi hinnanguid, mida ei ole võimalik dokumendi- ega lausetasemel piisava täpsusega eristada. Samas muudab see analüüsi keerukamaks, kuna nõuab täpsemat keelelist ja semantilist mõistmist, sealhulgas konteksti, sõnakasutuse ja väljendusviiside varieeruvuse arvestamist [32].

Kõige detailsem lähenemine on aspektitaseme meelestatusanalüüs, kus analüüsitakse tekstis esinevaid omadusi või aspekte eraldi, arvestades, et üks lause võib sisaldada mitut erinevat hinnangut erinevate aspektide kohta. Selline lähenemine on eriti kasulik olukordades, kus sama objekti mitmeid omadusi hinnatakse erinevalt. Näiteks võib inimene väljendada ühe toote puhul positiivset arvamust selle hinna, kuid negatiivset arvamust kvaliteedi kohta. Aspektitasemel analüüs võimaldab neid hinnanguid eristada. Seda ei ole võimalik dokumendi- ega lausetasemel piisava täpsusega saavutada. Aspektitasemel meelestatusanalüüs pakub kõige detailsemat analüüsitaset, aga see on kõige keerukam, kuna nõuab nii aspektide automaatset tuvastamist kui ka nende kontekstipõhist tõlgendamist [32].

3.2 Varasemad tööd

Eesti keele peal on meelestatusanalüüsi automatiseerimist käsitletud mitmetes varasemates töödes, keskendudes erinevatele ülesannetele ja tehnoloogiatele. Näiteks

kasutas Gerth Jaanimäe (2018) oma lõputöös leksikonipõhist lähenemist eesti keelsete tekstide meelestatuse hindamiseks. See lähenemine põhineb eeldefineeritud sõnastikul, kus sõnadele on määratud kindel meelestatus ning teksti üldine meelestatus tuletatakse nende sõnade esinemise põhjal. Töö tulemustest selgus, et selline meetod võimaldab luua toimiva baassüsteemi, kuid selle täpsus sõltub tugevalt kasutatava leksikoni kvaliteedist ja katvusest. Jaanimäe töös jäi leksikaalse ja masinõppepõhise lähenemise täpsus 75% lähedale, kuid erinevad lähenemised saavutasid varieeruvaid tulemusi erinevat tüüpi tekstide peal [33].

Siim Kaspar Uustalu (2019) viis läbi meelestatusanalüüsi katsetuse rakendades klassifikatsioonipõhist lähenemist, mille eesmärk oli määrata tekstidele vastav meelestatusklass. Selline lähenemine eeldab märgendatud andmestiku olemasolu, mille põhjal mudel õpib eristama positiivset, negatiivset ja neutraalset sisu. Uustalu saavutas oma katsetusega mudeli täpsuseks 0.66 ehk 66%. Täpsema tulemuse saavutamiseks soovitas autor järgmisi katsetusi teostada mõne keelemudeliga [34].

Regita Luukas (2023) töötas oma bakalaureusetöö raames välja algoritmi õppeainete tagasiside meelestatuse analüüsimiseks. Luukas arendas leksikonipõhise algoritmi, mis kategoriseerib tagasiside kommentaarid järgnevalt: positiivne, negatiivne, neutraalne või vastuoluline. Tema arendatud lahendus saavutas 67% täpsuse [35].

Katsetusi on tehtud ka kasutades keelemudeleid. M. Mets jt [36] peenhäälestasid oma 2024 aasta töös mitmeid BERT-tüüpi mudeleid eesti keele meelestatusanalüüsi ülesande lahendamiseks ning arvutasid tulemustele F1-skoori, mis võtab arvesse kõik klassid (*against*, *neutral*, *supportive*) võrdselt, näitamaks iga katsetuse täpsust (vt joonis 3) [36].

Model	Against	Neutral	Supportive	F1 macro
Naive Bayes	0.59	0.58	0.38	0.52
EstBert class	0.69	0.70	0.53	0.64
Est-RoBERTa	**0.74**	**0.69**	**0.55**	**0.66**
XLM-RoBERTa	0.73	0.65	0.54	0.64
mBERT (cased)	0.66	0.64	0.40	0.56
mBERT (uncased)	0.64	0.58	0.38	0.54
GPT 3.5	0.74	0.64	0.57	0.65
Est-RoBERTa Sentiment	0.63	0.42	0.42	0.49

Joonis 3. M. Mets, A. Karjus, I. Ibrus ja M. Schich töö tulemused [36].

M. Mets jt. katsetused näitasid, et eesti keelele spetsialiseeritud Est-RoBERTa mudel andis parima tulemuse meelestatusanalüüsi raames, saavutades F1 makroskoori 0.66. Töös katsetati ka GPT 3.5 mudelit mitte-peenhäälestatud kujul. F1 makroskoori tulemus

osutus ainult 0.01 võrra väiksemaks Est-RoBERTa omast. Tulemus viitab vabavaralise mudeli (Est-RoBERTa) kõrgele võimekusele, kuid lahendus on ressursinõudlikum, sest nõuab eelnevat peenhäälestamist [36].

3.2.1 OCR-vigade parandamine ajaloolistes eestikeelsetes tekstides

Eestikeelsete ajalooliste OCR-tekstide automaatset parandamist on seni uuritud vähe. Töö koostamise ajal oli Meimeri ja Lehtmetsa (2025) bakalaureusetöö „Ajalooliste eestikeelsete OCR tekstide järeltöötluse ja hindamise automatiseerimine Eesti Rahvusraamatukogu jaoks“ ainus töö, mis käsitleb suurte keelemudelite kasutamist eestikeelsete ajalooliste OCR-tekstide parandamisel [37].

Töös uuriti erinevate keelemudelite kasutamist OCR-vigade parandamiseks ning paranduste kvaliteedi hindamiseks. Meimeri ja Lehtmetsa tõid oma töös välja, et ajalooliste OCR-tekstide töötlemine on keeruline probleem, kuna tekstides esineb arhailist keelekasutust, OCR-moonutusi ning varieeruvat tekstistruktuuri. Samuti leiti, et isegi kaasaegsete keelemudelite kasutamisel ei olnud võimalik saavutada täielikult stabiilseid ega vigadeta OCR-järeltöötlust [37].

3.3 BERT-tüüpi keelemudelid

BERT (*Bidirectional Encoder Representations from Transformers*) on Google poolt välja töötatud NLP mudel, mis põhineb *Transformer*-arhitektuuril. Mudel on disainitud õppima sügavaid kahepoolselt kontekstitundlikke tekstiesitusi märgendamata andmetest, võttes kõigis mudeli kihtides korraga arvesse nii eelnevat kui järgnevat konteksti [38].

Käesoleva töö fookuses on BERT-tüüpi mudelid, sest BERT kasutab *pre-training* ja *fine-tuning* lähenemist, mis võimaldab mudelit kohandada erinevatele ülesannetele minimaalse arhitektuurilise muutusega, tehes selle sobivaks ka piiratud andmestikku kasutades. Eelõppes kasutab mudel *masked language modeling* (MLM) ja *next sentence prediction* (NSP) ülesandeid. MLM võimaldab mudelil mõista sõnade tähendust kontekstis, NSP seejuures aitab mõista teksti struktuuri. Peenhäälestamise käigus ei treenita ainult juurde lisatud klassifikatsiooni väljundkihti, vaid uuendatakse kogu mudeli parameetreid. See võimaldab BERT mudelil kohanduda konkreetse ülesande eripäradega ka piiratud andmestiku korral [38].

J. Devlin jt [1] töös testiti BERT mudelit *General Language Understanding Evaluation* (GLUE) loomuliku keele ülesannetel [39]. BERTi koondtulemuseks GLUE raamistiku ülesannetes oli 80.5% [38].

3.3.1 EstBERT

EstBERT on eesti keelele spetsialiseeritud BERT-tüüpi keelemudel. See on treenitud 2017. aasta eesti keele ühendkorpuse põhjal, mis oli mudeli treenimise ajal kõige mahukam andmekogu [40].

Mudeli eeliseks on selle kohandatus käänete rohkele eesti keelele, mis võimaldab saavutada täpsemaid tulemusi võrreldes mitmekeelsete mudelitega eesti keele spetsiifilistes ülesannetes. K. Sirts jt töös võrreldi EstBERTi mitmekeelsete, mBERT ja XLM-RoBERTa, mudelitega kuues ülesandes, millest üks oli meelestatusanalüüs. EstBERT saavutas kõikides ülesannetes kõige täpsema tulemuse, välja arvatud meelestatusanalüüsi ülesandes, kus XLM-RoBERTa oli kõige täpsem [40].

3.3.2 XLM-RoBERTa ja EstRoBERTa

XLM-RoBERTa põhineb RoBERTa arhitektuuril ning on optimeeritud suurte ja mitmekeelsete tekstikorpuste peal treenimiseks, võimaldades saavutada häid tulemusi ka madalama ressursiga keeltes [41]. Varasemates töödes on näidatud, et XLM-RoBERTa võib peenhäälestatud kujul ületada eesti keelele spetsialiseeritud mudeleid meelestatusanalüüsi ülesandes. Peatükis 3.3.1 toodi esile, et XLM-RoBERTa mudel peenhäälestatud kujul annab paremaid tulemusi meelestatusanalüüsi ülesandes kui EstBERT. Selle tulemuse saavutamiseks treeniti XLM-RoBERTa mudelit Eesti valentsikorpusel [14].

EstRoBERTa on eesti keelele kohandatud keelemudel, mis põhineb samuti RoBERTa arhitektuuril, kuid on treenitud ükskeelselt suuremahulisel eestikeelsel korpusel. Treeningandmestik sisaldab uudisartikleid ning Eesti keele osa *Universal Dependencies* CoNLL 2017 korpusest [42].

Keelespetsiifiliste ja mitmekeelsete mudelite võrdlus on oluline uurimissuund loomuliku keele töötluses. Uuringud on näidanud, et kuigi mitmekeelsed mudelid suudavad kasutada suuremahulist mitmekeelset teadmust, võivad ükskeelsed mudelid saavutada paremaid

tulemusi ülesannetes, kus on oluline keele spetsiifiliste nüansside täpne mõistmine [43] [44].

3.4 EstLLM

Lisaks BERT-tüüpi mudelitele kasutati käesolevas töös ka Llama 3.1 arhitektuuril põhinevat EstLLM mudelit, mis on eesti keele jaoks kohandatud suur keelemudel. Tegemist on generatiivse keelemudeliga, mis toetab nii tekstiloome- kui ka tekstimõistmisülesandeid [44].

EstLLM põhineb avatud lähtekoodiga baasmudelitel, mida on jätkuva eelõpetamise (*continued pretraining)* abil täiendavalt treenitud eestikeelsetel andmetel. Sellise lähenemise eesmärk on parandada mudeli võimekust väikese ressursiga keelte töötlemisel, ilma et kaoksid baasmudeli üldised teadmised ja arutlusvõime. Väikese ressursiga keelte puhul on see oluline, kuna mitmekeelsed mudelid ei pruugi saavutada piisavalt häid tulemusi keelespetsiifiliste nüansside mõistmisel [44].

Käesolevas töös kasutati mudeli varianti llama-estllm-prototype-0825-Q8_0-GGUF [20], mida käitati lokaalselt Ollama keskkonnas. EstLLM valiti töö jaoks, kuna see võimaldas töödelda eestikeelseid tekste lokaalselt ning kasutada vabavaralist lahendust, mis vastas töö lähteülesandes seatud nõudele kasutada võimalikult avatud tehnoloogiaid. Lisaks toetab EstLLM loomulikus keeles antud juhiste järgimist, võimaldades kasutada juhenduspõhiseid lähenemisi ilma mudelit täiendavalt peenhäälestamata [44], [45].

Samuti sobib generatiivne keelemudel OCR-vigade parandamise ülesandeks paremini kui klassikalised sõnastikupõhised õigekirjakontrolli süsteemid, kuna mudel arvestab sõnade laiemat konteksti ja taastab semantiliselt tõenäolisi sõnavorme ka tugevalt moonutatud sisendi korral [44], [37].

3.5 Mudeli juhendamine (*prompting*)

Mudeli juhendamine (*prompting*) tähendab, et keelemudelile antakse loomulikus keeles juhised konkreetse ülesande lahendamiseks [45]. Selle lähenemise puhul ei ole vaja mudelit eraldi peenhäälestada, vaid ülesanne kirjeldatakse mudelile tekstilise juhise kaudu.

Käesolevas töös kasutati *zero-shot* ja *few-shot* juhendamise lähenemisi. *Zero-shot* lähenemise korral antakse mudelile ainult ülesande kirjeldus ilma näideteta. *Few-shot* lähenemise korral lisatakse sisendisse ka näiteid soovitud väljunditest, et aidata mudelil ülesande vormi paremini mõista [45].

Varasemates töödes on näidatud, et juhiste sõnastus võib mõjutada mudeli väljundi kvaliteeti ja täpsust [46]. Seetõttu pöörati ka käesolevas töös tähelepanu juhiste sõnastusele ja ülesehitusele.

Käesolevas töös kasutati juhenduspõhist lähenemist EstLLM mudeli rakendamisel nii OCR-vigade parandamiseks kui ka meelestatusanalüüsi ülesannete lahendamiseks. Eksperimentides katsetati erinevaid juhiste ülesehitusi, sealhulgas *zero-shot*, struktureeritud *zero-shot* ja *few-shot* lähenemisi.

3.6 EstNLTK raamistik

EstNLTK on avatud lähtekoodiga eesti keele loomuliku keele töötluse raamistik, mis pakub erinevaid tekstianalüüsi tööriistu, sealhulgas tokeniseerimist, lemmatiseerimist ja morfoloogilist analüüsi [47]. Käesolevas töös kasutati EstNLTK raamistikku peamiselt tekstide morfoloogiliseks analüüsiks.

Morfoloogilise analüüsi teostamiseks kasutab EstNLTK Vabamorfi analüsaatorit, mis võimaldab määrata sõnade võimalikke morfoloogilisi tunnuseid ja sõnaliike. Selliseid tööriistu kasutatakse loomuliku keele töötluses tekstide grammatiliseks ja struktuurseks analüüsimiseks [48].

Käesoleva töö sisendtekstid sisaldasid OCR-töötlusest tulenevaid moonutusi, arhailist kirjaviisi ning ebastandardseid sõnavorme. Sellest tulenevalt kasutati EstNLTK raamistikku tekstide morfoloogilise analüüsitavuse hindamiseks enne meelestatusanalüüsi rakendamist

4 Teksti korrastamine

Käesolevas peatükis käsitletakse tekstide korrastamist enne meelestatusanalüüsi rakendamist. Analüüsitav tekstikorpus pärineb DIGARi arhiivist [49] ning OCR-vigade parandamise meetodeid eksperimenteeriti ajalehtede „Koit“ (23. oktoober 1975) ja „Postimees“ (23. oktoober 1925) tekstidel.

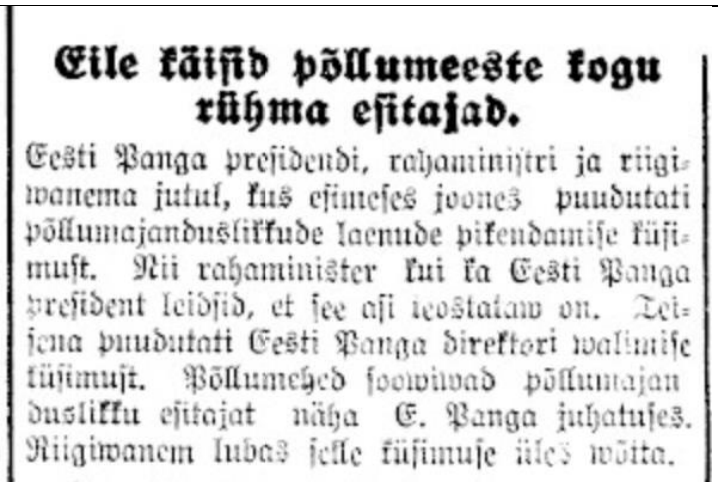
Tekstid on saadud OCR-töötluse abil ning sisaldavad seetõttu mitmesuguseid vigu, nagu katkiseid sõnu, valesid sümboleid ning ebakorrektsusi tühikute kasutuses. Näiteks „*ciiresti*”, „*ämblikxdiikide*” või „*kaub.°tõõst"sministri*”. Lisaks sisaldasid tekstid ajaloolisele kirjaviisile omaseid vorme, näiteks tähe „*w*” kasutamist tänapäevase „*v*” asemel. Sellised vead võivad raskendada tekstide automaatset analüüsi [50].

Tekstide korrastamiseks eksperimenteeriti kolme erineva lähenemisega:

1. morfoloogilisel analüüsil põhinev kvaliteedihindamine EstNLTK abil,
2. klassikaliste tekstiparandustööriistade katsetamine,
3. juhenduspõhine tekstiparandus EstLLM mudeli abil.

Tabelis 1 on esitatud tekstide korrastamise töövoog ajalehe „Postimees“ 23. oktoobri 1925. aasta artikli „Eile käisid põllumeeste kogu rühma esitajad“ näitel [51]. Tabelis on näidatud kogu töövoog alates algsest ajalehepildist kuni parandatud lõpptulemuseni. „Postimees“ väljaande artikkel valiti, kuna see on vanem kui „Koit“ väljaanne ning sisaldab rohkem OCR-vigu ja ajaloolisele kirjaviisile omaseid vorme.

Tabel 1. Tekstide korrastamise töövoog.

Töövoo samm	**Väljund**
Algne artikkel ajalehes	 **Eile käisid põllumeeste kogu rühma esitajad.** Eesti Panga presidendi, rahaministri ja riigiwanema jutul, kus esimeses joones puudutati põllumajanduslikkude laenude pikendamise küsimust. Nii rahaminister kui ka Eesti Panga president leidsid, et see asi teostataw on. Teisena puudutati Eesti Panga direktori walimise küsimust. Põllumehed soowiwad põllumajan duslikku esitajat näha E. Panga juhatuses. Riigiwanem lubas selle küsimuse üles wõtta.
OCR-väljund HTML-kujul	*postimeesew19251023.2.61 Eile käisid põllumeeste kogu rühma esitajad. Eesti Panga presidendi, rahaministri ja riigiwanema jutul, kus esimeses joones puudutati põllumajanduslikkude laenude pikendamise küsimust. Nii rahaminister kui ka Eesti Panga president leidsid, et see asi teostataw on. Teisena puudutati Eesti Panga direktori walinnse küsimust. Põllumehed soowiwad põllumajan duslikku esitajat näha E. Panga juhawses. Riigiwanem lubas selle küsimuse üles wõtta. *
HTML-ist puhastatud tekst	*postimeesew19251023.2.61* *Eile käisid põllumeeste kogu rühma esitajad.* *Eesti Panga presidendi, rahaministri ja riigiwanema jutul, kus esimeses joones puudutati põllumajanduslikkude laenude pikendamise küsimust. Nii rahaminister kui ka Eesti Panga president leidsid, et see asi teostataw on. Teisena puudutati Eesti Panga direktori walinnse küsimust. Põllumehed soowiwad põllumajan duslikku esitajat näha E. Panga juhawses. Riigiwanem lubas selle küsimuse üles wõtta.*
EstLLM parandatud tekst	*postimees 23.05.1925.2.61 Eile käisid põllumeeste kogu rühma esitajad Eesti Panga presidendi, rahaministri ja riigivanema jutul, kus esimeses joones puudutati põllumajanduslikkude laenude pikendamise küsimust. Nii rahaminister kui ka Eesti Panga president leidsid, et see asi teostataw on. Teisena*

Töövoo samm	Väljund
	puudutati Eesti Panga direktori valimise küsimust. Põllumehed soovivad põllumajanduslikku esitajat näha E.Panga juhatuses. Riigivanem lubas selle küsimuse üles võtta.

4.1 Morfoloogilisel analüüsil põhinev kvaliteedihindamine

Esialgse lähenemisena rakendati morfoloogilist analüüsi ehk Vabamorfi, et hinnata tekstide kvaliteeti sõnavormide analüüsitavuse kaudu [48]. Selleks arvutati iga teksti kohta sõnade osakaal, mille puhul morfoloogiline analüüs andis ebakindlaid või mittestandardseid tulemusi.

Analüüsiprotsess koosnes mitmest etapist. Esmalt eraldati artiklite tekstiline sisu ning eemaldati HTML-märgistus, et tagada ühtne ja analüüsitav sisend. Seejärel viidi läbi iga teksti morfoloogiline analüüs, kasutades EstNLTK teeki ja selle Vabamorfi analüsaatorit. Iga sõna kohta koguti kõik võimalikud morfoloogilised analüüsid. Sõna loeti vigaseks juhul, kui kõik sellele vastavad analüüsid vastasid vähemalt ühele järgmistest tingimustest:

- analüüs märgiti ebakindlaks (tunnus *form='?'*), või
- tegemist oli mittesõnalise üksusega (tunnus *partofspeech='Y'*).

Lisaks rakendati filtreerimist, mille käigus jäeti analüüsist välja numbreid sisaldavad üksused, mitte-tähelised üksused ning isikunimede initsiaalid (nt „*K.*"). Sõnad, millele ei leitud ühtegi morfoloogilist analüüsi, jäeti vigade loendist välja ning neid ei arvestatud analüüsitud sõnade koguarvu hulka. Iga teksti kohta arvutati vigade osakaal, jagades vigaseks loetud sõnade arvu analüüsitud sõnade koguarvuga. Saadud vigade osakaalu kasutati teksti morfoloogilise kvaliteedi ligikaudseks hindamiseks. Rakendatud lähenemisel ilmnesid mitmed olulised piirangud, mis mõjutavad saadud tulemuste tõlgendamist (nendest lähemalt peatükis 4.2).

4.2 Morfoloogilise analüüsi piirangud

Morfoloogiline analüüs on eelkõige mõeldud sõnavormide ja nende grammatiliste tunnuste analüüsimiseks, mitte õigekirjavigade tuvastamiseks [47] Vabamorfi analüsaator suudab sageli leida võimaliku analüüsi ka ebastandardsetele või vigastele sõnavormidele, mistõttu ei erista see alati selgelt korrektseid ja vigaseid kirjapilte. Lisaks käsitleb analüsaator ebamäärastena mitmeid keeleliselt korrektseid üksusi, sealhulgas pärisnimesid, lühendeid ja võõrsõnu [48].

Empiiriline analüüs näitas, et vaatamata rakendatud filtreerimisele märgiti probleemsetena ka mõningaid lühendeid (nt „*tel*"). Samal ajal jäid paljud OCR-vead tuvastamata juhtudel, kus analüsaator suutis sõnale määrata vähemalt ühe võimaliku morfoloogilise analüüsi. Seetõttu ei peegelda arvutatud vigade osakaal otseselt õigekirjavigade hulka, vaid kirjeldab pigem teksti üldist morfoloogilist analüüsitavust.

4.2.1 Morfoloogilise analüüsi piirangud tekstinäitel

Meetodi piirangute illustreerimiseks analüüsiti artiklit „Ungari teaduste akadeemia 100-aasta juubel." [52] „Postimees" 23. oktoobri 1925. aasta väljaandest.

Ungari teaduste akadeemia 100aasta juubel. Ka Eesti wõtab Pidustusest osa. Ungari teaduste akadeemia pühitseb 3. now. oma 100-a. juubelit. Silmas pidades neid sõbralikke suhteid, mis Eesti ja Ungari Vahel loa litsemad, otsustas Valitsus kõmandeerida Eesti esitajaks Ungari piduswstele haridusmin. prof. RahamägiHaridusministri reis kestab 30. okt. kuni 8. now. Tema äraolekul täidab haridusministri kohuseid Põllutöömin. A- Keren:.

Lõik sisaldab ajaloolise kirjaviisi vormi nagu „*wõtab*" ning OCR-moonutusi nagu „piduswstele", „loa litsemad", „kõmandeerida" ja „RahamägiHaridusministri". Vabamorfi ei tuvastanud neid vigadena, kuna leidis neile vähemalt ühe võimaliku morfoloogilise analüüsi.

Samas ilmnes vastupidine probleem. Vigasena märgiti sellised üksused nagu „*okt.*", „*prof.*" ja „*A-*", kuigi tegemist ei ole meelestatusanalüüsi seisukohalt oluliste vigadega. „*okt.*" ja „*prof.*" on tavapärased lühendid ning „*A-*" on OCR-moonutusest tekkinud eraldiseisev üksus.

Näide illustreerib meetodi mõlemat põhipiirangut korraga: tegelikud OCR-vead ja ajaloolise kirjaviisi vormid võivad jääda tuvastamata, samal ajal märgitakse vigasena ka keeleliselt korrektseid või vähese sisulise tähendusega üksusi. Tulemuste põhjal järeldati, et morfoloogilisel analüüsil põhinev lähenemine ei ole piisavalt täpne OCR-vigade tuvastamiseks ning otsustati uurida automaatseid tekstiparanduse meetodeid.

4.3 Klassikaliste tekstiparandustööriistade sobivuse hindamine

Eesti keele tekstiparanduseks on olemas mitmeid tööriistu, kuid käesoleva töö kontekstis ei osutunud need praktiliselt sobivaks. Eesmärk oli leida lahendus, mis toetaks eesti keelt, võimaldaks töödelda OCR-vigadega tekste ning sobituks olemasoleva Python-põhise töövooga.

Esialgu uuriti EstNLTK tekstiparanduse võimalusi, katsetades SpellCheckTagger komponenti [47]. Katsetuste käigus selgus, et kasutatud EstNLTK versioonis ei olnud õigekirjaparanduse funktsionaalsus rakendatav. Lisaks põhineb EstNLTK lahendus sõnastikupõhisel lähenemisel, mis käsitleb sõnu eraldiseisvate üksustena ega arvesta nende konteksti. OCR-tekstide puhul on see oluline piirang, kuna moonutatud sõnavormide korrektne taastamine sõltub sageli laiemast kontekstist [53].

Hunspell [54] on laialt kasutatav sõnastikupõhine õigekirjakontrolli süsteem, mida katsetati Pythonis *pyhunspell* ja *spylls* teekide kaudu. Eksperimentide käigus ilmnesid probleemid eesti keele sõnastikufailide kodeeringutega. Failid olid ISO-8859-15 formaadis, mis põhjustas ebausaldusväärseid tulemusi täpitähtede töötlemisel. Sõnastikke prooviti konverteerida UTF-8 kodeeringusse, kuid probleemid jäid püsima.

LanguageTool [55] on grammatika- ja õigekirjakontrolli süsteem, millel on olemas eesti keele tugi. Lokaalne kasutamine eeldas eraldiseisva serveri käivitamist, mis muutis lahenduse olemasolevas Python-põhises töövoos keerukaks. Avaliku API kasutamisel esinesid päringupiirangud, mis muutsid suuremahulise tekstikoguse automatiseeritud töötlemise ebaotstarbekaks.

Kokkuvõttes ei olnud käesoleva töö jaoks saadaval klassikalist tekstiparanduslahendust, mis oleks samaaegselt usaldusväärselt toetanud eesti keelt ja arvestanud OCR-tekstide eripärasid. Lisaks ei sobitunud olemasolevad lahendused praktiliselt kasutatud töövoogu.

Seetõttu otsustati jätkata generatiivsel keelemudelil põhineva lähenemisega, kasutades EstLLM mudelit.

4.4 EstLLM-il põhinev OCR-vigade parandamine

OCR-vigade parandamiseks kasutati EstLLM mudelit. Käesolevas töös kasutati mudeli varianti llama-estllm-prototype-0825-Q8_0-GGUF [56]. Mudelit käitati lokaalselt Ollama keskkonnas, mille kaudu teostati päringud Pythoni programmliidese abil.

Sisendtekstid jaotati automaatselt väiksemateks osadeks, et vältida mudeli kontekstipiirangute ületamist. Jaotamine toimus lausete kaupa ning tekstiosade maksimaalne pikkus oli ligikaudu 800 märki. Iga tekstiosa töödeldi mudeliga eraldi ning tulemused liideti hiljem terviktekstiks. Tekstiparanduse ülesandes suunati mudeli käitumist süsteemijuhise (*system prompt*) ning väljundi piirangute ja kontrollimehhanismide abil.

Süsteemijuhises defineeriti mudeli rolliks eesti keele OCR-tekstide parandamine (vt Lisa 5). Mudelit suunati parandama ainult OCR- ja õigekirjavigu ning vältima teksti tähenduse muutmist või lisasisu genereerimist.

Juhiseid ja mudeli konfiguratsiooni kohandati iteratiivselt empiiriliste katsetuste põhjal. Kui parandatud tekst muutus semantiliselt loetavamaks ning säilitas algse tähenduse, ei tehtud juhistes enam olulisi muudatusi.

Sarnaseid probleeme ajalooliste eestikeelsete OCR-tekstide parandamisel on täheldatud ka Meimeri ja Lehtmetsa (2025) töös [37], kus toodi välja, et suured keelemudelid võivad OCR-järeltöötluse käigus genereerida uut teksti, lisada sisendiga mitteseotud sisu või muuta algset tähendust.

4.4.1 Mudeli parameetrid ja väljundi kontroll

Mudeli tööparameetrid valiti eesmärgiga saavutada võimalikult kontrollitud ja stabiilne väljund. Temperatuuriks määrati 0, mis vähendab juhuslikkust tekstigenereerimisel, kuid ei taga täielikult deterministlikku käitumist generatiivsete keelemudelite puhul [57].

Genereeritava teksti maksimaalne pikkus määrati dünaamiliselt vastavalt sisendi pikkusele, ligikaudu 1.2-kordseks sisendi pikkuseks. See piirang vähendas olukordi, kus mudel genereeris liigset või sisendist sõltumatut teksti.

Lisaks rakendati stopptokeneid (*stop tokens*), mille abil katkestati genereerimine soovimatute mustrite ilmnemisel. Stop-sümbolite hulka lisati nii tehnilised markerid kui ka vestluslikud fraasid. Stopptokeneid kasutati nii mudeli konfiguratsioonis (*modelfile*) kui ka päringutasemel.

Täiendavalt rakendati väljundi puhastamist, mille käigus eemaldati võimalikud genereerimise markerid. Lisaks kasutati kontrollreeglit: kui väljundi pikkus ületas sisendi pikkust rohkem kui 20%, käsitleti väljundit ebausaldusväärsena. Sellisel juhul säilitati parandatud versiooni asemel algne tekst.

Sellised järeltöötlus- ja kontrollimehhanismid on levinud keelemudelite praktilises rakendamises, kus heuristikate abil püütakse vähendada hallutsinatsioonide ja soovimatu genereerimise mõju [58].

4.4.2 Näide tekstiparandusest

Meetodi toimimist illustreerib järgmine tekstilõik artiklist „Kes pääsewad kaubandus-tööstuskotta?“ ajalehe „Postimees" 23. oktoobri 1925. aasta väljaandest [13], kus on näha erinevat tüüpi OCR-vigade parandamist (vt tabel 2).

Tabel 2. EstLLM OCR-vigade paranduse näide.

Algne tekst	**Parandatud tekst**
Kes pääsewad kaubandus-tööstuskotta? üksmeele tagajärjed toowad wähemad linnad wäljapaist" wale kohale. Unnese kaubandus-tööstuskoja walimiste tagajärjed hakkawad selguma. Walimismaterzalide läbitöötamine kestub, kuid suur osa tööst on juba tehtud. Seniste andmete järele pääsewad kau-bandus-tööstuskotta järgmised esitajad.	*Kes pääsevad kaubandus-tööstuskotta? Üksmeele tagajärjed toovad väiksemad linnad väljapaistvale kohale. Unnese kaubandus-tööstuskoja valimiste tagajärjed hakkavad selguma. Valimismaterjali läbitöötamine kestab, kuid suur osa tööst on juba tehtud. Seniste andmete järgi pääsevad kaubandus-tööstuskotta järgmised esitajad.*

Näide illustreerib, et mudel suutis parandada erinevat tüüpi OCR-vigu, sealhulgas ajaloolise kirjaviisi vorme (nt „*pääsewad*" → „*pääsevad*", „*toowad*" → „*toovad*"), katkiseid sõnavorme (nt „*kau-bandus*" → „*kaubandus*", „*Walimismaterzalide*" → „*Valimismaterjali*") ning vigaseid sõnaühendeid (nt „*wäljapaist" wale kohale*" →

„*väljapaistvale kohale*"). Samas ilmnes, et mudel tõlgendas mõnel juhul ajaloolisi sõnavorme konteksti põhjal ümber. Näiteks „*wähemad*" parandati „*väiksemad*", mis muutis algset tähendust.

4.4.3 Mudeli käitumise piirangud

Rakenduse käigus täheldati juhtumeid, kus mudel ei järginud etteantud juhiseid ning genereeris lisaks parandatud tekstile ka vestluslikku sisu. Näiteks lisas mudel mõnel juhul väljundi lõppu järgmise teksti:

„Tere! Kas sa soovid, et ma parandan OCR-i vigu või õigekirjavigu?“

Selline käitumine ei vasta etteantud ülesandele ning viitab mudeli rollist väljumisele. Sarnaseid nähtusi on kirjeldatud ka teaduskirjanduses, kus keelemudelid ei pruugi alati täpsel t järgida juhiseid, eriti juhul kui sisend ei sisalda selgesõnalist ülesandekirjeldust [59], [60].

Stopptokenite ja teiste kontrollimehhanismide valik põhines praktilistel katsetustel ning jooksval väljundite hindamisel. See vähendas oluliselt soovimatute väljundite hulka. Generatiivsete keelemudelite väljund põhineb tõenäosuslikul järjestikgeneratsioonil, mistõttu ei ole võimalik kõiki kõrvalekaldeid täielikult vältida [57]. Lisaks ilmnes kompromiss parandamise agressiivsuse ja väljundi stabiilsuse vahel. Rangemad piirangud vähendasid soovimatuid muudatusi, kuid suurendasid olukordade arvu, kus osa vigadest jäi parandamata.

4.4.4 EstLLM-il põhineva lähenemise hinnang

Rakendatud lähenemine võimaldas parandada OCR-vigu kontekstitundlikult ning säilitada enamikel juhtudel teksti algse tähenduse ja struktuuri. Võrreldes klassikaliste sõnastikupõhiste lahendustega suutis mudel taastada ka selliseid sõnavorme, mille korrektne tõlgendamine sõltus laiemast kontekstist.

Paranduste kvaliteeti hinnati kvalitatiivselt, jälgides, kas tekst muutus semantiliselt loetavamaks ning sobivamaks edasiseks meelestatusanalüüsiks. Eksperimenti ei korratud fikseeritud arv kordi, kuna OCR-parandus oli töö lisakomponent ning eesmärk oli saavutada meelestatusanalüüsi jaoks piisav tekstikvaliteet, mitte maksimeerida OCR-

paranduse täpsust eraldiseisva ülesandena. Mudeli väljund ei olnud alati täielikult stabiilne ega vigadeta, aga saavutatud tekstikvaliteet oli meelestatusanalüüsi jaoks piisav.

4.4.5 OCR-paranduse mõju meelestatusanalüüsile

Tekstide korrastamise praktilise mõju hindamiseks võrreldi meelestatusanalüüsi tulemusi parandamata ja parandatud tekstidel, kasutades ajalehe „Postimees" 23. oktoobri 1925. aasta väljaande 74 artiklit. See väljaanne valiti hindamise aluseks, kuna 1925. aasta tekstid sisaldavad ajaloolisele kirjaviisile omaseid vorme ning ulatuslikumaid OCR-moonutusi võrreldes hilisema „Koit" väljaandega.

OCR-vead, katkised sõnavormid ja ajaloolise kirjaviisi moonutused raskendasid mudelil teksti semantilist tõlgendamist. Mitmel juhul mõjutas see nii mudeli määratud meelestatuse klassi kui ka kindlusmäära. 74 analüüsitud artiklist 50 puhul muutus OCR-paranduse järel dokumenditaseme kindlusmäär, säilitades sama meelestatuse klassi, ning 17 artikli puhul muutus ka määratud meelestatuse klass. Muutunud klassifikatsioonidega juhtumid valideeriti autorite poolt kvalitatiivse käsitsi hindamise teel. Hindamise põhjal ei tuvastatud juhtumeid, kus OCR-parandus oleks muutnud meelestatusklassifikatsiooni sisuliselt ebatäpsemaks. Enamikel juhtudel võimaldas parandatud tekst mudelil paremini tõlgendada artikli semantilist sisu.

Järgnevad näited illustreerivad tekstiparanduse mõju meelestatuse määramise tulemustele juhtudel, kus OCR-parandus muutis ka määratud meelestatuse klassi. Dokumenditaseme meelestatuse muutus tuleneb lausetaseme hinnangutest. Kui parandatud teksti lausetes muutus mudeli tõlgendus, kajastus see ka dokumendi üldhinnangus.

Näide 1

Näide 1 põhineb ajalehes „Postimees“ ilmunud artiklil „Kaitseliidu aastapäewa pidustused” [61].

Parandamata tekst - negatiivne meelestatus, kindlusmäär 0.26

Kaitseliidu aastapäewa pidustused.

Llpn üleandmine.

Kaitseliidu keskjuhatuse koosolekul 20. okt. otsustati kaitseliidu aastapäetov Puhul kaitseliidult lippl iile 1 unda. Sel päewal autakse ka Tallinna maletoale lipp kätte. Lipu üleandmisele kutsutakse igast malewast J esitajas pealei selle; weel külalised Lätist ja Soouiest. kadMsh on weel toöetud: muusikakoorida mäng Hommikul kirikutornides, päeval paraad Wabudnsplatsil, pärast lõnnat linnvs mitmel poot aktused päetoakohase sisuga.' Õhtul üldine! eine, millest palju pealikuid km ka lihtmalewlaft osa tootma kntsuchkfes Kawatftlftll on tveeh õhtul! tõrtoikutega rongikäik. korraldada" riigitoanema maja ette.'

Parandatud tekst - neutraalne meelestatus, kindlusmäär 0.11

Kaitseliidu aastapäeva pidustused.

Lipu üleandmine. Kaitseliidu keskjuhatuse koosolekul 20. okt otsustati kaitseliidult lippliile anda lipp. Sel päeval antakse Tallinna malevale lipp kätte. Lipu üleandmisele kutsutakse igast malevast esitaja peale selle veel külalised Lätist ja Soomest. Kadetsh on weel toetud: muusikakoorid mängivad hommikul kirikutornides, päeval paraad Vabaduseplatsil, pärast lõunat linnas mitmel pool aktused päevakohase sisuga. Õhtul üldine eine, millest palju pealikke ja ka lihtmalevlasi osa võtma kutsub keskjuhatus. Kavatsetud on teha õhtul tõrvikutega rongikäik. Korraldada riigitoanema maja ette.

OCR-paranduse järel muutus dokumendi meelestatus negatiivsest neutraalseks. Parandatud tekst võimaldas mudelil sündmuse kirjeldavat ja informatiivset iseloomu täpsemalt tõlgendada, mistõttu vähenes negatiivsete tõlgenduste mõju dokumendi üldhinnangule. *Confidence*-väärtuse langus 0.26-lt 0.11-le peegeldab asjaolu, et parandatud tekstis esines samaaegselt erineva meelestatusega lauseid, mille mõjud tasakaalustusid. Käesolevas näites tulenes madalam *confidence*-väärtus sellest, et parandatud tekst sisaldas nii neutraalseid kui ka erineva meelestatusega lauseid, mistõttu ei domineerinud ükski meelestatus selgelt.

Näide 2 põhineb ajalehes „Postimees“ ilmunud artiklil „Põhiwöimlemine” [62].

Parandamata tekst - neutraalne meelestatus, kindlusmäär 0.07

ISudZs müügile laua oodatud Niels B"kh: „Põhiwöimlemine". Hind 320 marka. Tekstis 175 kujutust. Tõlkija E. Rein,vald. Käesole"v raamat on sõna tõsises mõttes oodatud

‚vahend, sest kõik seni ilmunud raamatud tawad ainult üksikuid wõimlemise harusid. Kõnesolew raamat on kuulsa Taann wõunleja ja wõunlennsülikooli juhataja uuesti käsitatud wõimlemise tööwiisi monumentaalne ehitus. Raamat sisaldab kohalise kaswatuse mõtteid, päeivade kaupa koondatud kamade,"". Raamat on paljudesse kultuur-keeltesse tõlgitud ja igal pool oma uudsuse peale waatamata paari aasta jookjul korduwates trükkides ilmunud. Loodetawasti leiab ta ka meie õpetajate ja spordiringides omale waa" riwat wastuwõttu. K.-ü. „Loodus", Tartus.

Parandatud tekst - positiivne meelestatus, kindlusmäär 0.36

ISudZs müügile laua oodatud Niels Böhki:

„Põhiwöimlemine". Hind 320 marka. Tekstis 175 kujutust. Tõlkija E. Reinvald. Käesolev raamat on sõna tõsises mõttes oodatud vahend, sest kõik seni ilmunud raamatud käsitlevad ainult üksikuid wõimlemise harusid. Kõnesolev raamat on kuulsa Taani võunleja ja võunlennsülikooli juhataja uuesti käsitatud wõimlemise tööwiisi monumentaalne ehitus. Raamat sisaldab kohalise kasvatuse mõtteid, päevade kaupa koondatud harjutusi. Raamat on paljudesse kultuurkeeltesse tõlgitud ja igal pool oma uudsuse peale vaatamata paari aasta jooksul korduvates trükkides ilmunud. Loodetavasti leiab ta ka meie õpetajate ja spordiringides omale vastuvõttu. K.-ü. „Loodus", Tartus.

OCR-vead moonutasid mitmeid positiivset tähendust kandvaid lauseid ning raskendasid mudelil teksti soovitusliku iseloomu tuvastamist. Parandatud tekst võimaldas mudelil positiivset meelestatust selgemini tuvastada, mille tulemusena muutus nii dokumendi meelestatuse klass kui ka *confidence*-väärtus. *Confidence*-väärtuse tõus 0.07-lt 0.36-le viitab sellele, et parandatud tekstis tuvastati positiivse meelestatuse väljendusi ühtlasemalt ja selgemini. Parandamata tekstis olid OCR-vead moonutanud lauseid määral, mis takistas mudelil meelestatuse usaldusväärset tõlgendamist.

Lisaks EstLLM-il põhinevale lähenemisele katsetati töö käigus ka spetsialiseeritud OCR-järeltöötluseks loodud mudelit Llammas-OCR-FT13k [63]. Mudelit siiski käesoleva töö töövoogu ei liidestatud, kuna eesmärgiks oli kasutada olemasolevasse Python-põhisesse lahendusse sobituvat ühtset töövoogu. Seetõttu otsustati OCR-paranduse ülesandes jätkata EstLLM-il põhineva lähenemisega.

Tulemused viitavad sellele, et OCR-vead ei mõjuta ainult teksti loetavust, vaid võivad otseselt muuta ka mudeli semantilist tõlgendust ja meelestatuse klassifikatsiooni. Kuigi OCR-parandus ei kõrvaldanud kõiki vigu täielikult, muutusid tekstid enamikel juhtudel semantiliselt loetavamaks ning meelestatusanalüüsi tulemused olid paremini kooskõlas teksti sisulise tähendusega

5 Mudeli valimine

Selles peatükis keskendutakse erinevate keelemudelite võrdlevale hindamisele eesti keele meelestatusanalüüsi ülesandes. Mudelite valikul lähtuti varasematest teadustöödest ja olemasolevatest lahendustest, mille puhul on juba esitatud häid tulemusi just eesti keele kontekstis.

Katsetustesse kaasati nii eelnevalt peenhäälestatud kui ka üldotstarbelisi mudeleid, millele viidi läbi täiendav treenimine. See võimaldab võrrelda erinevaid lähenemisi: ühelt poolt valmis kujul kasutatavaid lahendusi ning teiselt mudeleid, mis vajavad kohandamist konkreetse ülesande jaoks.

Käesolevas töös kasutati EstBERT128_sentiment, EstRoBERTa, XLM-RoBERTa ja EstLLM mudeleid. Antud valik sai tehtud teaduskirjanduse, nende kättesaadavuse ning töös kasutamise võimaluste põhjal. Kuna töö üheks kriteeriumiks oli vabavaraliste lahenduste kasutamine lõpptulemi saavutamiseks, siis oli võrdlusesse kaasatavate mudelite valik piiratud.

EstBERT128_sentiment on eesti keelele ja meelestatusanalüüsi ülesandele kohandatud mudel, mis vastas käesoleva töö eesmärkidele. XLM-RoBERTa kaasati võrdlusesse, et hinnata, kas eesti keelele kohandatud mudelid saavutavad paremaid tulemusi eestikeelses meelestatusanalüüsi ülesandes kui mitmekeelne mudel. Lisaks selgus teaduskirjandust läbi töötades, et XLM-RoBERTa on katsetuste käigus saavutanud meelestatusanalüüsi ülesandes paremaid tulemusi kui eesti keelele kohandatud mudelid (vt peatükki 3.3.1). EstRoBERTa kaasati töösse, et võrrelda, kas peenhäälestatud kujul RoBERTa arhitektuuril põhinev eesti keelele kohandatud mudel sobib ka meelestatusanalüüsi ülesande jaoks. Täiendava võrdluspunkti lisab XLM-RoBERTa ja EstRoBERTa tulemuste analüüs, kuna varasemas uuringus (vt peatükki 3.3.1) leiti, et XLM-RoBERTa on meelestatusanalüüsi ülesandes parem kui EstRoBERTa. EstLLM kaasati võrdlusesse, et hinnata, kuidas generatiivne keelemudel sobib traditsiooniliste klassifitseerimismudelite kõrval meelestatusanalüüsi ülesande lahendamiseks.

Mudelite võrdlemiseks kasutati ühtset hindamisraamistikku, kus kõik mudelid hinnati sama andmestiku (Eesti valentsikorpus) ning samade hindamismeetrikate (*accuracy*, *precision*, *recall* ja F1-skoor) alusel. F1-skoor näitab, kui palju kattuvad mudeli vastused ja tegelik õige vastus [64]. Lisaks arvutati ka makro- ja kaalutud keskmised näitajad, mis võimaldavad hinnata mudeli sooritust nii klasside lõikes kui ka tervikuna.

Kirjed jaotati kasutades 70/10/20 jaotust [15], et andmestikust eraldada treenimis-, valideerimis- ja testandmed. Nendest kirjetest 2475 kasutati EstRoBERTa ja XLM-RoBERTa treenimiseks meelestatusanalüüsiülesande jaoks, 353 kirjet kasutati mudeli valideerimiseks ning 708 kirjet testimiseks. Testimistulemused võeti aluseks kõige parema tulemuse saavutanud mudeli väljaselgitamiseks.

Järgnevalt kirjeldatakse eksperimentide ülesehitust, kasutatud andmestikke ning treenimisprotsessi, millele järgnev mudelite tulemuste analüüs ja võrdlus.

5.1 EstBERT128_sentiment eksperiment

EstBERT-i eksperimentides kasutati eelnevalt peenhäälestatud keelemudelit EstBERT128_sentiment, mis on treenitud eesti keele meelestatusanalüüsi ülesande jaoks ning klassifitseerib teksti kolme kategooriasse: positiivne, negatiivne ja neutraalne. Mudel on kättesaadav HuggingFace platvormil ning seda rakendati *Transformers* teegi abil, mis võimaldab eelõpetatud keelemudeleid standardiseeritult kasutada ja integreerida NLP ülesannetes [65].

Mudelile anti ülesandeks hinnata Eesti valentsikorpuses olevate tekstide meelestatust. Korpusest eemaldati vastuoluline meelestatusklass, sest selle määramise võimekus puudub mudeli klassifitseerimiskihis.

EstBERT128_sentiment tulemused on esitatud tabelis 3. Üldiseks mudeli täpsuseks (*accuracy*) kujunes 95%, mis tähendab, et mudel klassifitseeris enamiku tekstidest korrektselt.

Tabel 3. EstBERT128_sentiment eksperimendi tulemused.

	Precision	***Recall***	***F1-score***	***Support***
Negatiivne	0.95	0.99	0.97	386

	Precision	*Recall*	*F1-score*	*Support*
Neutraalne	0.93	0.92	0.93	145
Positiivne	0.96	0.88	0.92	177
Accuracy			0.95	708
Macro avg	0.95	0.93	0.94	708
Weighted avg	0.95	0.95	0.95	708

Klassipõhine analüüs näitab, et mudel saavutas parima tulemuse negatiivse klassi tuvastamisel. Selles klassis saavutas mudel F1-skooriks 97%. Neutraalse ja positiivse klassi tulemused on mõnevõrra madalamad võrreldes negatiivse klassiga. Neutraalse klassi F1-skooriks kujunes 93%, samas kui positiivse klassi F1-skoor oli 92%. Need tulemused viitavad mudeli kõrgele võimekusele klassifitseerida tekstide meelestatust korrektselt.

5.1.1 EstBERT128_sentiment 4-meelestatusklassi eksperiment

EstBERT128_sentiment mudeli algne klassifitseerimiskiht ei sisaldanud meelestatusklassi vastuoluline, mis Eesti valentsikorpuses on esitatud, otsustati teha katsetus mudeli peenhäälestamisega.

Eksperimendis kasutati Eesti valentsikorpuse andmestikku peenhäälestamiseks ja testimiseks sama andmestiku jaotuse põhimõtte alusel, mis eelnevalt mainitud peatükis 5. Selle jaotuse järgi oli treeningandmete ridu 2861, valideerimise ridu 409 ja testandmeid 818. Testimistulemused on esitatud tabelis 4.

Tabel 4. EstBERT128_sentiment 4-klassi peenhäälestatud mudeli eksperimendi tulemused.

	Precision	*Recall*	*F1-score*	*Support*
Negatiivne	0.86	0.93	0.90	385
Neutraalne	0.90	0.90	0.90	145
Positiivne	0.89	0.86	0.88	177
Vastuoluline	0.54	0.42	0.47	111
Accuracy			0.84	818
Macro avg	0.80	0.78	0.79	818

	Precision	***Recall***	***F1-score***	***Support***
Weighted avg	0.83	0.84	0.83	818

Eksperimendi tulemused näitavad, et kõikide meelestatusklasside F1-skoori tulemused langesid. Lisaks osutus ka vastuolulise klassi tulemus pigem madalaks, täpsemalt 47%, mis on ligikaudu 40 protsendipunkti võrra väiksem võrreldes järgmise kõrgeima tulemusega (positiivne klass). Makro keskmise F1-skoori tulemus ka langes, 94% pealt 79% peale. Sellest tulemusest saab järeldada, et ühe klassi juurde lisamine vähendas mudeli võimet erinevaid klasse tasakaalustatult eristada.

Tulemuste põhjal otsustati järgmistes eksperimentides kasutada kolme meelestatusklassi lähenemist.

5.2 EstRoBERTa eksperiment

Käesolevas eksperimendis kasutati eesti keelele spetsialiseeritud keelemudelit EstRoBERTa (tartuNLP/EstRoBERTa), eesmärgiga hinnata selle sobivust meelestatusanalüüsi ülesandes ning võrrelda sooritust teiste mudelitega.

EstRoBERTa mudel ei ole treenitud meelestatusanalüüsi ülesande jaoks, seega enne eksperimendi sooritamist viidi läbi mudeli peenhäälestamine.

EstRoBERTa eksperimendi tulemused on esitatud tabelis 5. Mudeli täpsuseks (*accuracy*) kujunes 77%, mis viitab mudeli mõõdukale võimekusele klassifitseerida tekstide meelestatust.

Tabel 5. EstRoBERTa peenhäälestatud mudeli eksperimendi tulemused.

	Precision	***Recall***	***F1-score***	***Support***
Negatiivne	0.86	0.88	0.87	386
Neutraalne	0.65	0.53	0.59	145
Positiivne	0.65	0.71	0.68	177
Accuracy			0.77	708
Macro avg	0.72	0.71	0.71	708

	Precision	***Recall***	***F1-score***	***Support***
Weighted avg	0.76	0.77	0.76	708

Klassipõhisest analüüsist saab välja lugeda, et mudel saavutas parima tulemuse negatiivse klassi määramisel, kus F1-skooriks kujunes 87%. Positiivse klassi määramisel saavutas mudel madalama tulemuse, saavutades F1-skooriks 68%. Mudeli kõige nõrgemaks klassiks osutus neutraalne klass, saavutades F1-skooriks 59%.

Makro keskmine F1-skoor, milleks osutus 71%, mis peegeldab mudeli üldist tasakaalustatud võimekust erinevate klasside eristamisel. Kaalutud keskmise F1-skooriks tuli 76%, mis viitab sellele, et mudel saavutas paremaid tulemusi klassides, mis olid andmestikus rohkem esindatud.

5.3 XLM-RoBERTa eksperiment

XLM-RoBERTa puhul on samuti tegemist mudeliga, mis ei ole kohandatud meelestatusanalüüsi ülesande teostamiseks.

XLM-RoBERTa mudeli täpsuse tulemused on esitatud tabelis 6. Mudel saavutas eksperimendi tulemusena täpsuseks (*accuracy*) 55%. Tulemus näitab, et mudel ei osutunud võimekaks meelestatusanalüüsi ülesandes eesti keeles.

Tabel 6. XLM-RoBERTa peenhäälestatud mudeli eksperimendi tulemused.

	Precision	***Recall***	***F1-score***	***Support***
Negatiivne	0.55	1.00	0.71	386
Neutraalne	0.00	0.00	0.00	145
Positiivne	0.00	0.00	0.00	177
Accuracy			0.55	708
Macro avg	0.18	0.33	0.24	708
Weighted avg	0.30	0.55	0.38	708

Klassipõhine tulemuste analüüs näitab, et mudel ei suutnud erinevaid meelestatuseklasse tasakaalustatult eristada. Negatiivse klassi F1-skooriks saavutas mudel 71%, teise kahe klassi puhul tuli tulemuseks 0%. Negatiivse klassi *Recall* tulemus 100% saame

tõlgendada kahte moodi: kas mudel määras kõik tekstid negatiivseks või peaaegu kõik tekstid negatiivseks. *Recall* tulemus viitab sellele, et mudel tuvastas kõik tegelikult negatiivsed tekstid.

Selle katse puhul ei peegelda *accuracy* tulemus 55%, et mudel pooltel juhtudel korrektse tulemuse annab. Kuna andmestikus esines negatiivseid kirjeid rohkem ja seda klassi mudel suutis määrata, siis sellest keskmine täpsusega tulemus.

Antud eksperimendi puhul osutub oluliseks makro keskmine F1-skoor, sest see arvestab kõiki kolme klassi võrdselt. F1-skooriks selle mõõdiku alusel osutus 24%, mis on madal tulemus.

5.4 EstLLM eksperiment

Käesolevas eksperimendis kasutati EstLLM mudeli kvantifitseeritud varianti llama-estllm-prototype-0825-Q8_0-GGUF [56]. Mudeli käivitamiseks seadistati vajalik tarkvarakeskkond ning käivitati Ollama server, mille kaudu teostati päringud programmliidese abil. Mudelit kasutati kohandatud konfiguratsiooniga, mis oli defineeritud eraldi *modelfile*'i abil. *Modelfile* sisaldas süsteemijuhiseid, mille kaudu määrati mudeli käitumine katsete läbiviimisel. Erinevate eksperimentide jaoks muudeti *modelfile*'is defineeritud juhiseid ning mudel loodi vastava konfiguratsiooniga uuesti.

Eksperimentide automatiseerimiseks kasutati skripte, mis saatsid mudelile sisendtekstid ning töötlesid saadud vastused edasiseks analüüsiks. Eksperimendis kasutati sama andmestikku ning selle jaotamise põhimõtet, mis on välja toodud peatükis 5.

Mudelit kasutati generatiivse keelemudelina, mille käitumist suunati loomulikus keeles sõnastatud juhiste abil meelestatusanalüüsi teostamiseks. Käesolevas töös käsitletakse juhiseid kõrgtaseme strateegiatena (*zero-shot*, struktureeritud *zero-shot* ja *few-shot*), keskendudes juhiste sisule ja ülesande sõnastusele. Mudel suunati juhiste abil tagastama väljundit etteantud klassifikatsioonisiltide kujul (positiivne, negatiivne või neutraalne), mis võimaldas käsitleda selle vastuseid otseselt klassifitseerimistulemustena. Mõnel juhul ei piirdunud mudel ainult klassifikatsioonisiltidega ning need tulemused arvestati katsest välja.

Eksperimentide eesmärgiks oli hinnata, kuidas erinevad juhiste strateegiad mõjutavad mudeli sooritust. Selleks võrreldi *zero-shot* ja *few-shot* lähenemisi ning analüüsiti, milline juhiste struktuur võimaldab saavutada parima tasakaalu erinevate meelestatusklasside vahel.

Tuleb arvestada, et generatiivse mudeli väljund ei ole täielikult deterministlik ning võib varieeruda sõltuvalt juhiste sõnastusest. Seetõttu pöörati tähelepanu juhiste ühtsele ja kontrollitud kasutamisele kõigis katsetes.

5.4.1 *Zero-shot* lähenemise eksperiment

Zero-shot lähenemise korral esitati mudelile juhised ilma näideteta. Antud lähenemine on laialdaselt kasutatav suurte keelemudelite hindamisel ning toimib sageli baasvõrdlusena, kuna see võimaldab hinnata mudeli võimekust ilma ülesandespetsiifiliste näideteta [45].

Esmalt katsetati *zero-shot* juhist, mis sisaldas ainult ülesande kirjeldust ilma täiendavate näidete või otsustusreegliteta. Kasutatud juhis on esitatud joonisel 4.

```
Oled tekstide meelestatuse analüüsija.
Määra teksti üldine meelestatus:
- positiivne
- negatiivne
- neutraalne
Vali täpselt üks klass vastavalt teksti üldisele toonile.
Ära lisa muud teksti.
```

Joonis 4. Meelestatusanalüüsi eksperimendi *zero-shot* juhis.

Zero-shot eksperimendi puhul andis mudel 2 kirje puhul tulemuse, mis ei vastanud klassifikatsioonisiltidele. Eksperimendi tulemused on esitatud tabelis 7. Antud eksperiment saavutas üldise täpsuse tulemuseks 57%. Tegemist on üsna madala tulemusega, millest võib järeldada, et *zero-shot* strateegia generatiivse keelemudeli puhul ei anna paremat tulemust kui klassifitseerimiseks kohandatud mudel.

Tabel 7. EstLLM *zero-shot* lähenemise eksperimendi tulemused.

	Precision	***Recall***	***F1-score***	***Support***
Negatiivne	0.93	0.59	0.73	385
Neutraalne	0.30	0.72	0.42	144
Positiivne	0.62	0.38	0.48	177

	Precision	***Recall***	***F1-score***	***Support***
Accuracy			0.57	706
Macro avg	0.62	0.57	0.54	706
Weighted avg	0.72	0.57	0.60	706

Makro keskmine F1-skoor 54% näitab, et lähenemise üldine tasakaalustatud võimekus oli madal. Negatiivse klassi määramisel oli EstLLM *zero-shot* lähenemist kasutades kõige täpsem, saavutades seal F1-skooriks 73%. Teised meelestatuseklassid saavutasid tulemuse, mis jäid alla 50%.

Neutraalse klassi puhul saavutas EstLLM antud strateegiat kasutades *precision* väärtuseks 30% ning *recall* 72%. See tulemus viitab sellele, et mudel kaldus neutraalset klassi üle klassifitseerima.

5.4.2 Struktureeritud *zero-shot* lähenemise eksperiment

Struktureeritud *zero-shot* lähenemise korral täiendati peatükis 5.4.1 välja toodud juhist samm-sammulise otsustusloogikaga. Selle eesmärk oli suunata mudelit analüüsima teksti süsteemsemalt. Erinevalt eelmisest katsest ei piirdutud ainult ülesande kirjeldusega, vaid mudelile anti reeglid, mille alusel klassifikatsioon teostada.

Juhis suunas mudelit esmalt tuvastama, kas tekst sisaldab hinnangulist või emotsionaalset keelt. Kui hinnang puudus, klassifitseeriti tekst neutraalseks. Hinnangu olemasolul suunati mudelit eristama positiivset ja negatiivset meelestatust. Juhtudel, kus tekst sisaldas nii positiivseid kui negatiivseid elemente, rakendati domineeriva hinnangu põhimõtet.

Selline lähenemine põhineb struktureeritud juhendamise ja samm-sammulise arutluse põhimõtetel, mis võivad parandada keelemudelite võimet keerukamaid ülesandeid lahendada [66]. Samuti on näidatud, et selgelt formuleeritud juhised aitavad suunata mudeli käitumist ning parandada ülesandespetsiifilist sooritust [59].

Kasutatud struktureeritud *zero-shot* juhis on esitatud joonisel 5.

```
Oled tekstide meelestatuse analüüsija.
Sinu ülesanne on määrata tekstile TÄPSELT üks klass:
```

```
- positiivne
- negatiivne
- neutraalne
Järgi seda otsustusloogikat:
1. Kas tekst sisaldab hinnangut või emotsiooni?
  - Kui EI -> neutraalne
  - Kui JAH -> mine edasi
2. Kas hinnang on:
  - ainult positiivne -> positiivne
  - ainult negatiivne -> negatiivne
  - nii positiivne kui negatiivne -> mine edasi
3. Kui tekst sisaldab nii positiivseid kui negatiivseid elemente:
  - vali domineeriv hinnang (tugevam, rõhutatum või detailsem)
TÄPSUSTUSED:
- neutraalne = faktid või kirjeldus ilma hinnanguta
- ära eelda emotsiooni kui seda pole selgelt väljendatud
- ära liigita neutraalseks, kui tekstis on hinnang olemas
REEGLID:
- vali ainult üks klass
- ära lisa muud teksti
```

Joonis 5. Meelestatusanalüüsi eksperimendi struktureeritud *zero-shot* juhis.

Struktureeritud *zero-shot* lähenemisel andis mudel 6 kirje puhul tulemuse, mis ei vastanud klassifikatsioonisiltidele. Eksperimendi tulemused on esitatud tabelis 8. Antud juhise strateegiat kasutades kujunes mudeli üldiseks täpsuseks 63%.

Tabel 8. EstLLM struktureeritud *zero-shot* lähenemise eksperimendi tulemused.

	Precision	***Recall***	***F1-score***	***Support***
Negatiivne	0.69	0.94	0.79	383
Neutraalne	0.42	0.48	0.45	142
Positiivne	0.89	0.10	0.17	177
Accuracy			0.63	702
Macro avg	0.67	0.50	0.47	702
Weighted avg	0.69	0.63	0.57	702

Ainult täpsuse põhjal ei saa öelda, et tegemist on parema tulemusega kui *zero-shot* lähenemisega. Struktureeritud *zero-shot* lähenemisel saavutas mudel makro keskmise F1-skooriks 47%. See tulemus on madalam kui *zero-shot* eksperimendi oma, milleks oli 54%. Makro keskmine F1-skoor näitab kõikide meelestatusklasside tuvastamise võimekust, mis on oluline indikaator meelestatusanalüüsi ülesandes. Positiivse klassi puhul näeme, et F1-skoor on 17%, mis on madal tulemus.

Järgneva eksperimendi puhul kasutatakse võrdlusbaasina *zero-shot* lähenemist.

5.4.3 *Few-shot* lähenemise eksperiment

Few-shot lähenemise korral anti juhis konkreetsete näidetega, mille eesmärk oli suunata mudelit paremini mõistma erinevate meelestatuskategooriate tähendust. Erinevalt *zero-shot* lähenemisest, kus mudel peab ülesande lahendama üksnes juhiste põhjal, võimaldab *few-shot* lähenemine mudelil tugineda näidete kaudu omandatud mustritele.

Käesolevas eksperimendis esitati mudelile kolm näidet, mis esindasid vastavalt positiivset, negatiivset ja neutraalset meelestatust. Lisaks säilitati juhis valida segatud hinnanguga tekstide puhul domineeriv meelestatus, mis esitati peatükis 5.4.2.

Selline lähenemine on kooskõlas *in-context learning* põhimõttega [45]. Samuti on näidatud, et sobivalt valitud näited võivad mõjutada mudeli otsustuspiire ning parandada keerukamate või ebamäärasemate klasside tuvastamist [67] [68].

Few-shot eksperimendis kasutatud juhis on esitatud joonisel 6.

```
Oled tekstide meelestatuse analüüsija.
Sinu ülesanne on määrata teksti üldine meelestatus.
Võimalikud vastused:
positiivne
negatiivne
neutraalne
Näited:
Tekst: "See film oli täiesti suurepärane!"
Vastus: positiivne
Tekst: "Ma ei jäänud selle teenusega rahule."
Vastus: negatiivne
Tekst: "Täna on kolmapäev."
Vastus: neutraalne
Kui tekst sisaldab nii positiivseid kui negatiivseid elemente,
vali domineeriv üldtoon.
Vasta ainult ühe sõnaga (positiivne, negatiivne või neutraalne).
Ära lisa muud teksti.
```

Joonis 6. Meelestatusanalüüsi eksperimendi *few-shot* juhis.

Few-shot lähenemise puhul esines märkimisväärselt rohkem mittesobivaid väljundeid (59 kirjet), kuna mudel kaldus näidete põhjal genereerima pikemat seletust ühe klassifitseerimissõna asemel. Mudeli täpsuseks kujunes 62% (vt tabel 9). Tulemus on kõrgem, kui *zero-shot* eksperimendi oma, kuid selle tulemuse pealt parimat lähenemisviisi selgitada ei saa.

Tabel 9. EstLLM *few-shot* lähenemise eksperimendi tulemused.

	Precision	***Recall***	***F1-score***	***Support***
Negatiivne	0.85	0.75	0.79	364
Neutraalne	0.31	0.60	0.41	126
Positiivne	0.63	0.33	0.43	159
Accuracy			0.62	649
Macro avg	0.60	0.56	0.55	649
Weighted avg	0.69	0.62	0.63	649

Few-shot lähenemise makro keskmine tulemuseks oli 55%. See on parem kui *zero-shot* lähenemise tulemus, kuid mitte märkimisväärselt. Kaalutud keskmise F1-skoori tulemus tõusis 60%-lt 63%-le.

Klassipõhine analüüs näitab, et paranes negatiivse klassi määramise võimekus, kuid selle arvelt langesid positiivse ja neutraalse klassi F1-skooride tulemused.

Few-shot lähenemise puhul on ka oluline märkida, et suurenes mittesobivate tulemuste arv. Kuna täpsuse ja klassipõhised F1-skoorid ei paranenud märkimisväärselt, siis on klassifikatsioonisiltidele vastavate väljundite arv oluliseks indikaatoriks parima lähenemise otsustamisel. Kuna *zero-shot* lähenemine saavutas võrreldava makro keskmise F1-skoori väiksema arvu mittesobivate väljunditega, kasutatakse järgnevas võrdluses BERT-tüüpi mudelitega *zero-shot* lähenemise tulemusi.

5.5 Mudelite tulemuste võrdlus ja järeldused

Käesolevas peatükis analüüsitakse kõikide mudelite tulemusi ning tehakse nende põhjal otsus, millise mudeliga jätkatakse meelestatusanalüüsi tööriista arendamist. Mudeleid võrreldakse nende üldise täpsuse, makro keskmise- ja kaalutud keskmise F1-skoori alusel. Tulemused on esitatud tabelis 10.

Tabel 10. Eksperimentidesse kaasatud mudelite võrdlus.

Mudel	***Accuracy***	***Macro avg F1-score***	***Weighted avg F1-score***
EstBERT128_sentiment	95%	94%	95%

Mudel	***Accuracy***	***Macro avg F1-score***	***Weighted avg F1-score***
EstRoBERTa	77%	71%	76%
XLM-RoBERTa	55%	24%	39%
EstLLM *zero-shot*	57%	54%	60%

Tabelist saab välja lugeda, et kõige nõrgema sooritusega meelestatusanalüüsi ülesandes tegi XLM-RoBERTa peenhäälestatud mudel. XLM-RoBERTa üldiseks täpsuseks kujunes 55%, kuid kui vaatame makro keskmist F1-skoori, siis see on märkimisväärselt madalam (24%). Madal tulemus on tingitud sellest, et mudel oli suuteline ainult tuvastama negatiivseid tekste korrektselt. Ühe võimaliku põhjusena, miks mudel saavutas madalad tulemused, võib olla, et treeningandmeid ei olnud mudelil piisavalt, et see kohaneks meelestatusanalüüsi ülesandega (2475 kirjet). Arvestades teiste mudelite paremaid tulemusi ning täiendava treeningandmestiku puudumist, ei peetud mudeli edasist peenhäälestamist käesoleva töö raames põhjendatuks.

EstLLM *zero-shot* lähenemine saavutas samuti üldise täpsuse tulemuse 57%, kuid ka selle mudeli ja strateegia klassipõhine tasakaal oli võrreldes teiste mudelitega nõrk. Eksperimendis saavutas mudel makro keskmise F1-skoori 54%, mis võrreldes teiste BERT-tüüpi mudelitega on märkimisväärselt nõrgem tulemus. Lisaks osutus EstLLM miinuseks väljundite stabiilsus. EstLLM *zero-shot* lähenemisel jäi 2 teksti lõplikust hindamisest välja.

EstRoBERTa peenhäälestatud mudeli puhul on näha paremaid tulemusi kui XLM-RoBERTa ja EstLLM *zero-shot* mudelite juures. EstRoBERTa üldine täpsus ja ka makro kesmine- ning kaalutud keskmine F1-skoor olid üle 70%. Need numbrid näitavad, et mudel oli enamikel juhtudel võimeline määrama meelestatusklassi korrektselt ning oli erinevaid klasse määrates pigem tasakaalustatum. Nagu ka XLM-RoBERTa mudeli puhul, oli ka EstRoBERTa treeningandmeid pigem vähe. Kuna üks eksperimentidest osutus edukamaks ning lisaandmed puudusid, siis ei liigutud edasi mudeli täiendava treenimisega.

Kõige paremaid tulemusi näitas EstBERT128_sentiment mudel. Kõikide tabelis X esitatud mõõdikute tulemused ületasid 90% piiri. Nendest tulemustest saame järeldada,

et mudel suutis peaaegu igal juhul määrata tekstide meelestatust korrektselt ning oli selles ka tasakaalukas.

Huvitava leiuna suutsid kõik mudelid, väljaarvatud EstBERT128_sentiment 4-klassi treenitud versioon, kõige paremini määrata negatiivset meelestatuseklassi. Sellist tulemust võis mõjutada asjaolu, et negatiivseid tekste oli Eesti valentsikorpuses kõige rohkem esindatud.

Saadud tulemuste põhjal otsustati tööriista arendamisel liikuda edasi EstBERT128_sentiment mudeliga. Mudel saavutad võrreldud lahendustest kõige kõrgemad tulemused ning oli erinevate mõõdikute lõikes kõige stabiilsem. Tulemuste põhjal võib järeldada, et eestikeelse meelestatusanalüüsi ülesandes osutus kõige sobivamaks vabavaraliseks keelemudeliks EstBERT128_sentiment, millega vastati töö esimesele uurimisküsimusele.

6 Dokumenditaseme meelestatuse määramine

Dokumenditaseme agregeerimismeetodi valikul katsetati kolme erinevat lähenemist ning võrreldi nende tulemusi konkreetsetel tekstinäidetel.

Esimese lähenemisena kasutati mitte-neutraalsete lausete kindlusmäärade põhjal arvutatud kaalutud keskmist. See meetod oli arvutuslikult lihtne, kuid andis lõpptulemuseks ainult dokumendi meelestatuse skoori ega eristanud eraldi seda, kui kindel või laialdaselt esindatud meelestatuse signaal tekstis oli. Näiteks ChatGPT abil loodud illustratiivsel näitetekstil:

Tallinna börs sulges eile langusega, kaotades päeva jooksul kaks protsenti. Investorid olid murelikud intressimäärade tõusu pärast. Samas teatas tehnoloogiaettevõte Nortal rekordkasumist, mis ületas analüütikute ootusi. Ettevõtte aktsiad tõusid järsult ja meeleolu turgudel paranes õhtu poole.

klassifitseeris mudel kaks lauset negatiivseks ja kaks positiivseks, kõik kindlusmääraga *(confidence*) ligikaudu 1.0. Kuna positiivsete ja negatiivsete lausete kaalud olid praktiliselt võrdsed, kujunes dokumendi skooriks nullilähedane positiivne väärtus, mis klassifitseeriti positiivseks. Meetod võimaldas määrata dokumendi domineeriva meelestatuse, kuid ei kirjeldanud eraldi seda, kui tugev või ühtlane meelestatuse signaal tekstis oli.

Teise lähenemisena katsetati kombineeritud meetodit, mille väljatöötamisel lähtuti ideest, et kõik laused ei panusta dokumendi üldisesse meelestatusse võrdselt [69]. Selleks ühendati kogu teksti põhjal arvutatud kaalutud keskmine ning tugevaima meelestatusega lausete põhjal arvutatud täiendav skoor. Täiendava skoori arvutamisel kasutati kõrgeimate kindlusmääradega mitte-neutraalseid lauseid, mille eesmärk oli rõhutada tugevaima meelestatuse signaaliga tekstiosi. Katsetamisel ilmnes aga, et kui mudel andis enamikele lausetele sarnased ja kõrged kindlusmäärad, sõltus lausete valik suurel määral lausete järjestusest ega võimaldanud tugevaima meelestatusega lauseid usaldusväärselt

eristada. See muutis tulemused tundlikuks väikestele kindlusmäärade erinevustele ning põhjustas ebastabiilseid klassifikatsioone, mistõttu otsustati kombineeritud meetodist loobuda.

Seetõttu kasutati lõplikus lahenduses kaalutud keskmist koos eraldi dokumenditaseme kindlusmääraga. Sama tekstinäite puhul andis lõplik meetod dokumendi skooriks 0.0 ja klassifitseeris dokumendi kindlusmääraga 0.30 neutraalseks, mis peegeldab täpsemalt nii meelestatuse tasakaalu kui ka meelestatuse signaali vähest domineerivust dokumendis.

Esmalt määrati iga lause meelestatus eraldi valitud mudeli abil, mis tagastas iga lause jaoks meelestatuse (positiivne, negatiivne või neutraalne) ning kindlusmäära. Pikemate tekstide korral ei olnud võimalik kogu dokumenti korraga mudelisse sisestada, kuna see ületas mudeli sisendpiirangu. Seetõttu jagati tekstid esmalt lauseteks EstNLTK lausetuvastuse abil. Kui üksik lause oli liiga pikk, jagati see täiendavalt väiksemateks osadeks. Dokumendi üldise meelestatuse leidmiseks kasutati kaalutud keskmist, kus suurema kindlusmääraga klassifitseeritud laused mõjutasid lõpptulemust rohkem kui väiksema kindlusmääraga laused.

Iga lause meelestatus teisendati arvuliseks väärtuseks: positiivne = +1, neutraalne = 0, negatiivne = −1. Neutraalsed laused jäeti dokumendi meelestatuse skoori arvutamisest välja, kuna need ei kanna otsest meelestatuse signaali. Samas võeti nende kindlusmäärad arvesse katvuse komponendi arvutamisel, et hinnata mitte-neutraalse meelestatuse osakaalu kogu tekstis.

Dokumendi meelestatus arvutati mitte-neutraalsete lausete kaalutud keskmisena:

$$\mathrm{Sdok} = \frac{\sum_{i=1}^{n} \mathrm{SiCi}}{\sum_{i=1}^{n} \mathrm{Ci}}$$

Kus s_i tähistab lause meelestatuse väärtust (+1 positiivne, -1 negatiivne), c_i tähistab lause kindlusmäära ja n on mitte-neutraalsete lausete arv. Saadud väärtus jääb vahemikku [−1, 1].

Lõplik dokumendiklass määrati lävepõhiselt, kasutades läve väärtust 0.2: kui $S_{dok} > 0.2$, klassifitseeriti dokument positiivseks. Kui $S_{dok} < -0.2$, klassifitseeriti dokument negatiivseks ja muul juhul neutraalseks. Läve kasutamise eesmärk oli vältida olukorda,

kus väga nõrk kõrvalekalle tõlgendatakse ekslikult selgelt positiivse või negatiivse meelestatusena. Läve väärtus 0.2 valiti empiiriliste katsetuste põhjal. Väiksema läve korral klassifitseeriti liiga paljud tasakaalustunud tekstid positiivseks või negatiivseks, suurema läve korral suurenes neutraalsete dokumentide osakaal ebaproportsionaalselt.

Juhul kui kõik dokumendi laused klassifitseeriti neutraalseks, tagastati dokumendi meelestatuseks neutraalne, meelestatuse skooriks 0.0 ning kindlusmääraks 1.0, kuna kõik laused olid omavahel kooskõlas ega sisaldanud mitte-neutraalset meelestatuse signaali.

Lisaks lõplikule klassile arvutati dokumenditaseme kindlusmäär. Tegemist ei ole kasutatud mudeli otsese väljundiga, vaid käesoleva töö jaoks defineeritud heuristilise hinnanguga dokumenditaseme otsuse kindlusele. Dokumenditaseme kindlusmäära eesmärk oli anda kasutajale ligikaudne hinnang selle kohta, kui tugev, laialdaselt esindatud ja omavahel kooskõlaline on tekstis leiduv meelestatuse signaal.

Dokumendi kindlusmäär arvutati kolme komponendi põhjal:

- tugevus - dokumendi meelestatuse skoori absoluutväärtus,
- katvus - kui suur osa kogu teksti kindlusmäärade summast pärineb mitte-neutraalsetest lausetest,
- kooskõla - positiivsete ja negatiivsete meelestatuse signaalide omavaheline kooskõla.

$$kindlusmäär = 0.5 \cdot tugevus + 0.3 \cdot katvus + 0.2 \cdot kooskõla$$

kus:

$$tugevus = \mid S_{dok} \mid,$$

$$katvus = \frac{\sum C_{mitte-neutraalne}}{\sum C_{kõik}},$$

$$kooskõla = \frac{\mid pos - neg \mid}{pos + neg}$$

ning pos ja neg tähistavad vastavalt positiivsete ja negatiivsete lausete kindlusmäärade summasid.

Kaalud 0.5, 0.3 ja 0.2 määrati empiiriliste katsetuste põhjal, lähtudes ajaleheartiklite sisulistest omadustest. Meelestatuse tugevusele omistati suurim kaal, kuna laused panustavad dokumendi üldise meelestatuse kujunemisse erineval määral [69]. Katvuse komponendi eesmärk oli hinnata, kui suur osa tekstist sisaldab mitte-neutraalset meelestatuse infot, sest ajalehetekstid sisaldavad sageli neutraalset ja informatiivset sisu, mis ei kanna meelestatuse signaali [70]. Lausetevahelisele kooskõlale anti väiksem kaal, kuna ajaleheartiklid võivad sisaldada samaaegselt nii positiivseid kui ka negatiivseid hinnanguid, mistõttu võivad dokumendis esineda vastandlikud meelestatuse signaalid.

Agregeerimisloogika lõplikku toimivust hinnati peatükis 8.1 kirjeldatud kontrolltesti käigus.

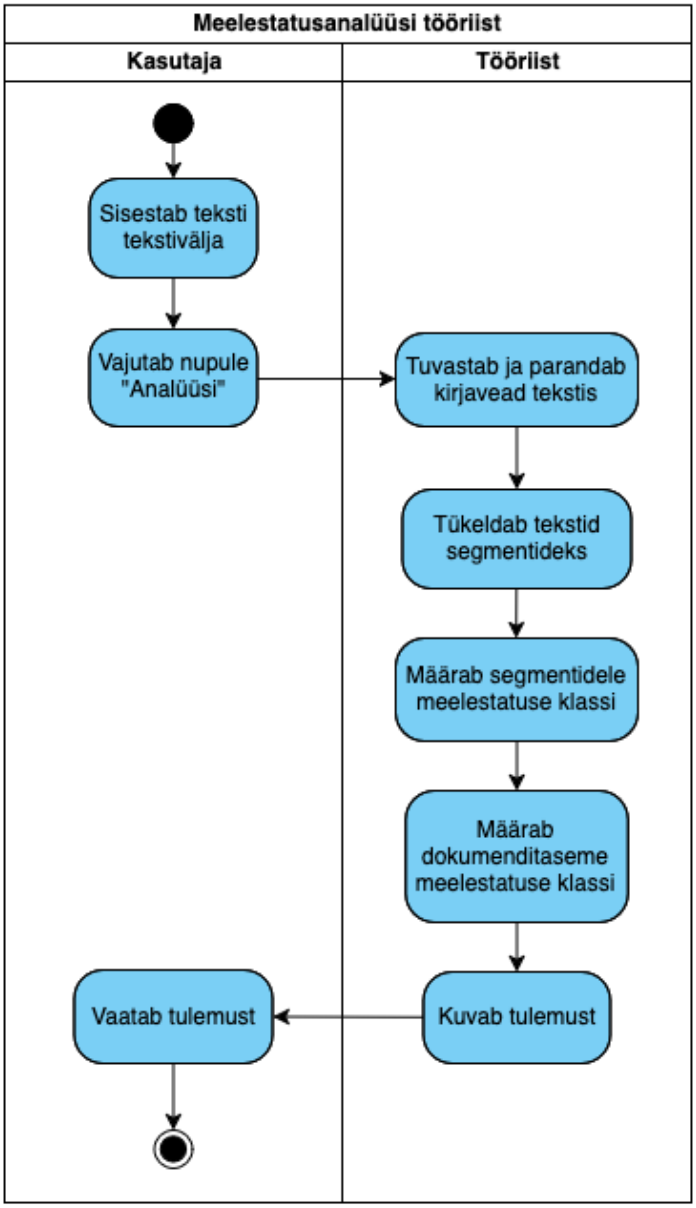
7 Kasutajaliides

Tööriista üheks osaks oli kasutajaliidese arendamine, mille eesmärk oli muuta loodud meelestatusanalüüsi teostamine kasutajale visuaalselt selgeks, kasutades RaRa visuaalset identiteeti [71]. Käesolev töö ei keskendunud ainult mudeli valimisele ja meelsuse määramisele, vaid ka reaalse tööriista loomisele, mille oluliseks rolliks oli kasutajaliidese arendus.

Kasutajaliidese eesmärk on võimaldada kasutajal sisestada eestikeelset teksti, käivitada automaatne meelestatusanalüüs ning kuvada saadud tulemused visuaalselt ja selgitaval kujul. Lisaks lõpptulemuse kuvamisele, esitab tööriist ka sisendisse kuuluvate erinevate segmentide eraldiseisvaid meelestatushinnanguid, et kasutajale näidata tööriista otsustusprotsessi.

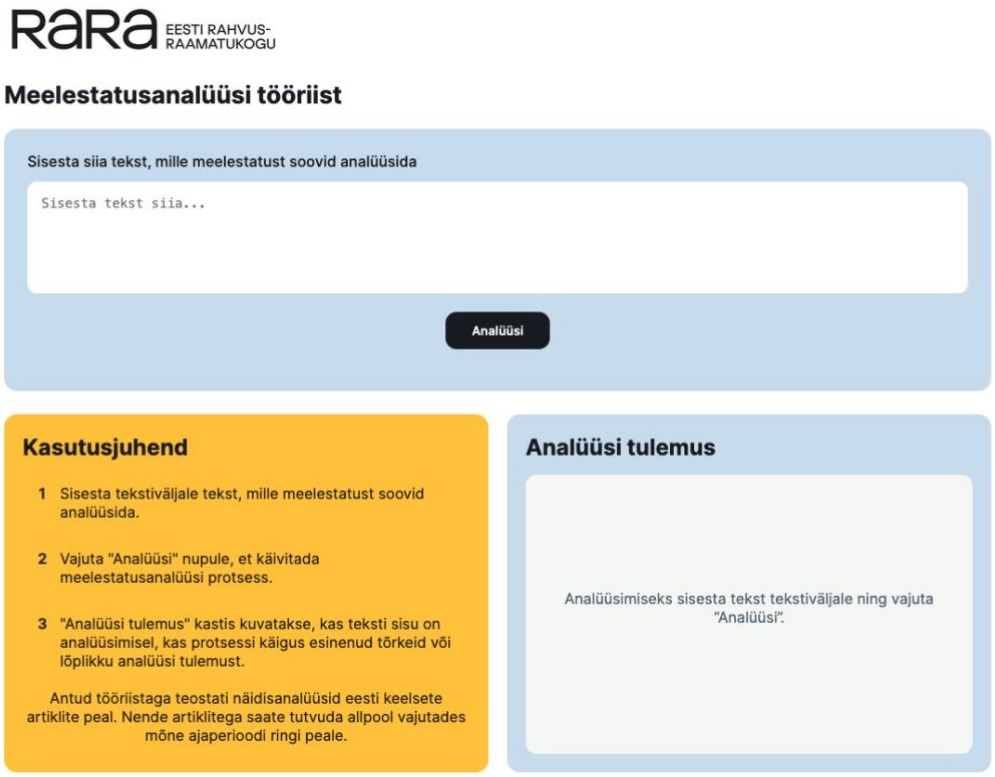
Arenduse käigus lähtuti põhimõttest, et süsteem peab olema kasutatav ka inimese poolt, kellel puudub tehniline taust masinõppe või loomuliku keele töötluse valdkonnas. Selle valideerimiseks viidi läbi kasutajatestid (tulemusi vaata peatükist 8.1.2).

Enne kasutajaliidese realiseerimist modelleeriti tööriista töövoog UML tegevusdiagrammina, kus kirjeldati nii kasutaja kui ka tööriista tegevusi sisendi töötlemise protsessis (vt joonist 7). Diagrammi eesmärk oli visualiseerida kasutaja ja tööriista vahelist suhtlust ning määratleda tegevuste loogiline järjestus alates teksti sisestamisest kuni tulemuse kuvamiseni. See diagramm aitas planeerida kasutajaliidese struktuuri ja komponentide paigutust, mis toetaks tööriista arusaadavust.



Joonis 7. Meelestatusanalüüsi tööriista tegevusdiagramm.

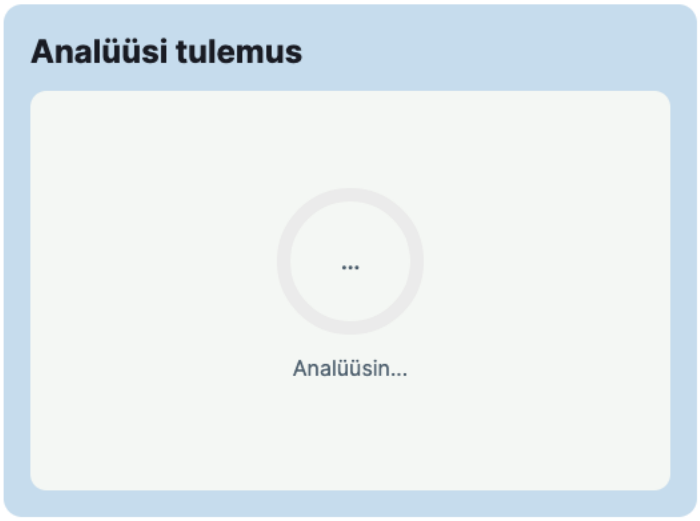
Joonisel 8 on kuvatud valminud tööriista kasutajaliides, kus kuvatakse kasutajale tekstisisestusvälja, tulemuse kuvamise komponenti (mille sisu muutub vastavalt tegevusele) ja kasutusjuhendit.



Joonis 8. Loodud meelestatusanalüüsi tööriista kasutajaliides.

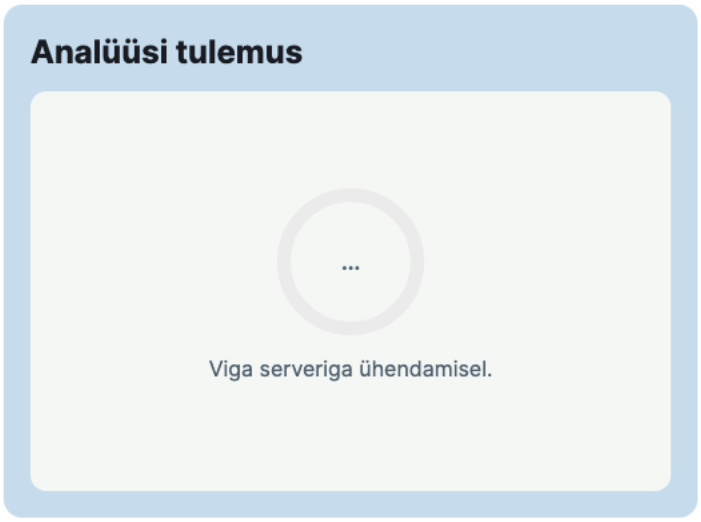
Tekstisisestusvälja kasuks otsustati, et vältida pahaloomuliste failide sisestamist süsteemi ning võimaldada kasutajal analüüsida tekste nii, et neid ei peaks eraldi faili salvestama. RaRa enda kasutuseks loodi skript, millega saab faile analüüsida.

Kui kasutaja käivitab analüüsimisprotsessi, siis kommunikeeritakse talle, mis taustal toimub. Kui sisendit veel töödeldakse kuvatakse kasutajale kirja „Analüüsin“ (joonis 9).



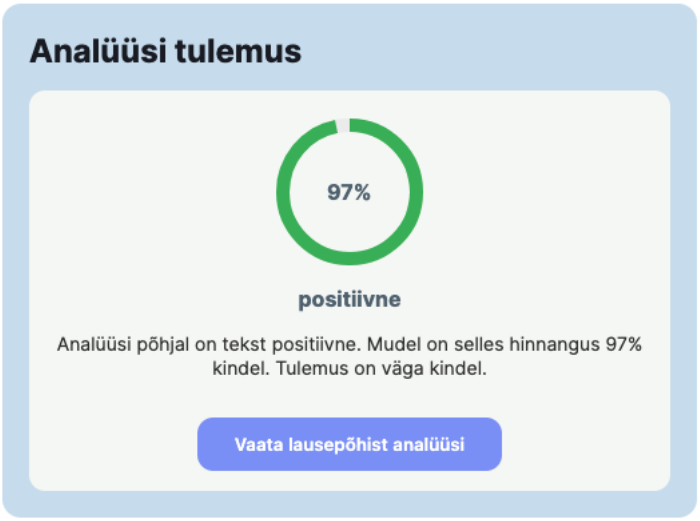
Joonis 9. Analüüsi tulemuse komponent, kui tööriist tegeleb sisendi analüüsimisega.

Kui serveriga ühendamisel peaks esinema viga, siis antakse sellest teada kasutajale, kuvades teksti „Viga serveriga ühendamisel.“ (joonis 10).



Joonis 10. Tulemuse kuvamise komponent, kui serveriga ühendamisel tekkis tõrge.

Töötlemise õnnestumisel saab kasutaja näha tulemust samast komponendist (vaata näidistulemust jooniselt 11)



Joonis 11. Tulemuse kuvamise komponent, kui sisendi töötlemine õnnestus.

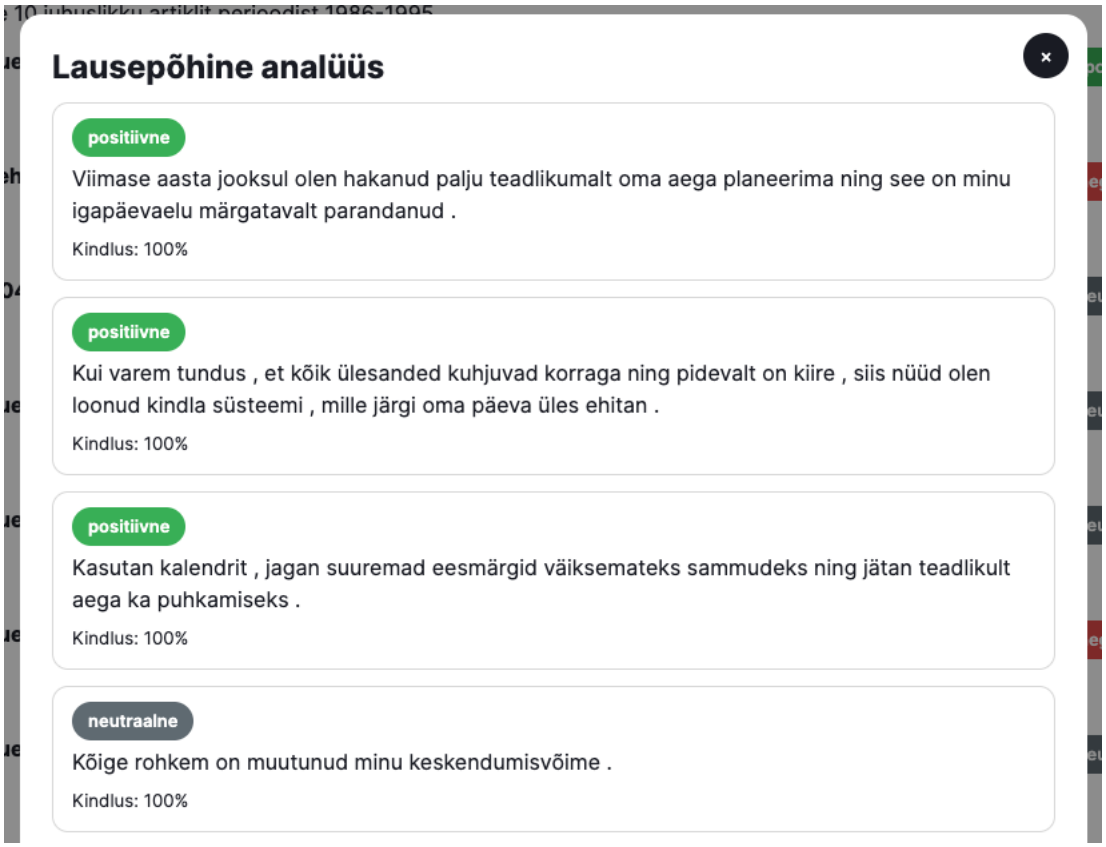
Õnnestunud sisendi töötlemise puhul kuvatakse kasutajale:

- sõõrikdiagrammi, mis omandab värvi vastavalt analüüsi tulemusele (positiivne on roheline, negatiivne on punane ja hall on neutraalne),
- tööriista tulemuse kindluse protsenti,
- sõnaline tulemus (positiivne, negatiivne või neutraalne):
- tulemuse selgitus:
    - kui mudeli kindlus on üle 80% kuvatakse teksti „Tulemus on väga kindel.“
    - Kui mudeli kindlus on üle 60%, aga alla 81%, kuvatakse teksti „Tulemus on üsna kindel.“
    - Kui mudeli kindlus jääb alla 61% kuvatakse teksti „Tulemus on ebakindlam ja võib vajada täiendavat tõlgendamist.“

Kindluse vahemikud ja nende kirjeldused määrati heuristiliselt eesmärgiga selgitada kasutajale arusaadavalt tööriista tulemust ja selle usaldusväärsust. Vahemikke ei käsitleta töö analüüsiprotsessis järelduste tegemiseks, vaid on kasutajale suunatud indikaatoritena.

Sõõrikdiagramm toodi rakendusse tööriista kindluse visualiseerimiseks, et kasutajale anda juurde veel üks väljund, lisaks protsendile ja kirjeldusele, mis aitaks kindlusetulemust tõlgendada.

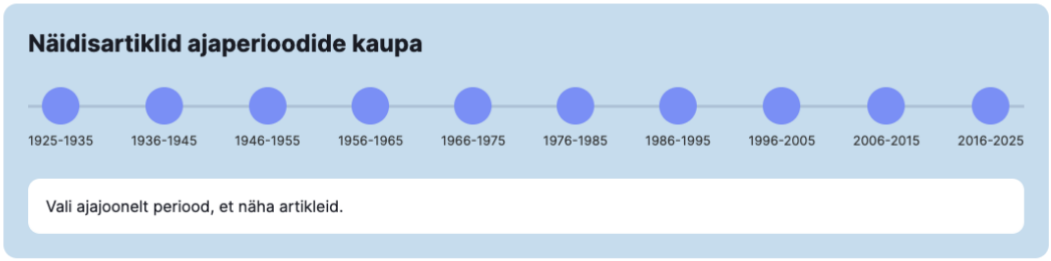
Joonsel 12 on kujutatud lausepõhise analüüsi detailvaadet, mis avaneb hüpikaknana vajutades „Vaata lausepõhist analüüsi“ nupule tulemuste kuvamise komponendi juures. Nupp ilmub vaatele õnnestunud töötlemise puhul.



Joonis 12. Lausepõhise analüüsi detailvaade.

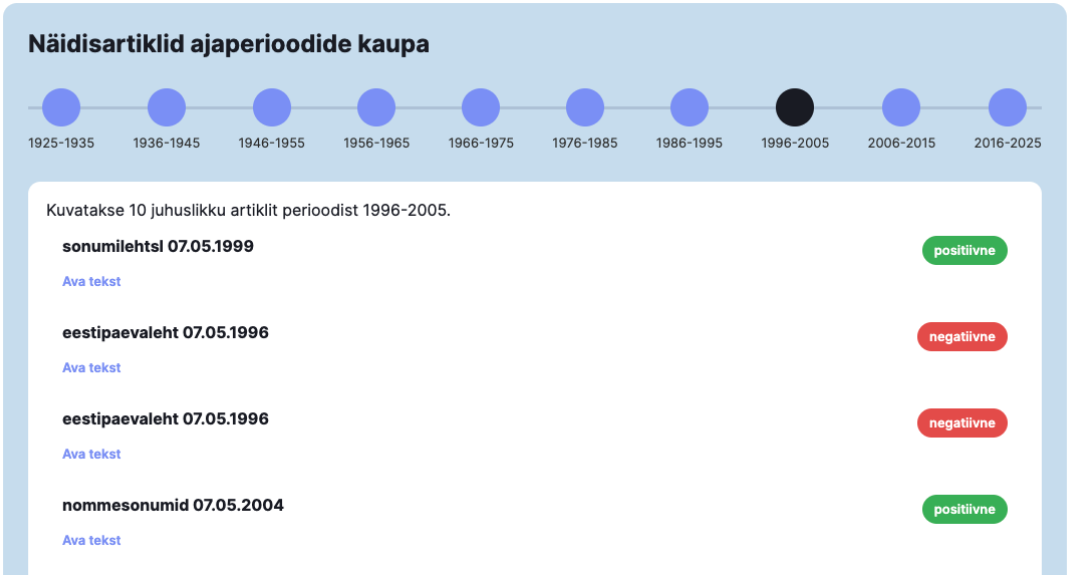
Antud modaal sai kasutajaliidesesse toodud sisendi töötlemise eripärast. Pikkade sisendtekstide puhul jagatakse tekst väiksemateks segmentideks (lauseteks), millele määratakse eraldiseisvad meelestatushinnangud. Detailvaates kuvatakse neid hinnanguid kasutajale eraldi, võimaldades analüüsida ka teksti erinevate osade lõikes.

Töö käigus teostati ka tööriista kasutades eesti artiklite meelestatusanalüüs, et saaks kasutajatele esitada näidiseid tööriista toimimisest. Artiklid pärinesid vahemikust 1925-2025, mis jaotati kümnendikeks. Kümnenditest koostati ajatelg, mis on kujutatud joonisel 13.



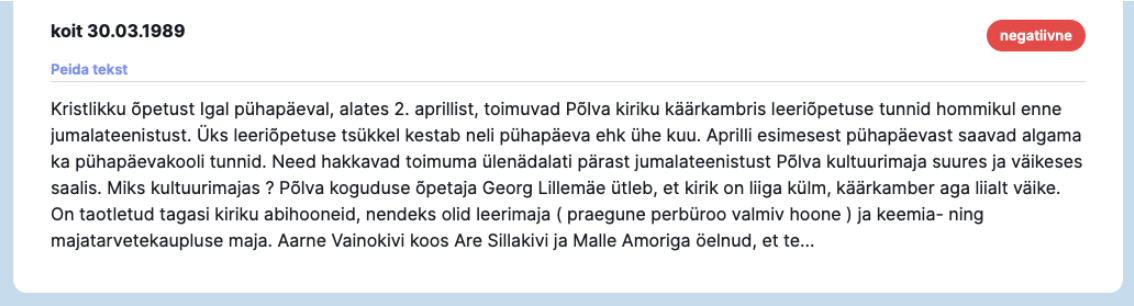
Joonis 13. Näidisartiklite ajajoon.

Ajateljel olevatele ringidele saavad kasutajad vajutada, et avada nimekirja artiklitest, mis kuuluvad antud vahemikku. Joonisel 14 on väljatoodud artiklite nimekiri, mis peale ringile vajutamist avaneb.



Joonis 14. Näidisartiklite loetelu 1996-2005 ajavahemiku kohta.

Kuna artikleid artiklite arv, mida analüüsiti oli liiga suur kuvamiseks (4557 kõikide ajavahemike peale kokku), siis otsustati kasutajale kuvada juhuslikku 10 artiklit valitud perioodist. See otsus aitas kaasa ka tööriista jõudlusele, tehes artiklite kuvamise sujuvamaks. Näidisartiklite puhul on näha ajalehte, millest artikkel pärineb, artikli avaldamise kuupäeva, artikli meelestatushinnangut ning valikut „Ava tekst“. Joonisel 15 on kujutatud vaade, mis avaneb antud valikule vajutades.



Joonis 15. Näidisartikli osa avatud kujul.

Näidisartikli avatud olekus kuvatakse ainult osa artiklist. Kuna paljud artiklid osutusid pikaks, siis jõudlusprobleemide vältimiseks kaasati vaatesse ainult 600 tähemärk. Antud arv sai valitud eksperimenteerides, mis on suurim hulk, mida suudab tööriist sujuvalt kuvada ning mis visuaalselt sobitub.

7.1 Kasutajaliidese valideerimine

Käesoleva töö raames viidi läbi viis kasutajatesti. Testkasutajad leiti autorite suhtlusvõrgustikust, sealjuures püüti valida võimalikult erineva taustaga kasutajad, eesmärgiga saada mitmekesisemat tagasisidet tööriista kasutatavuse kohta.

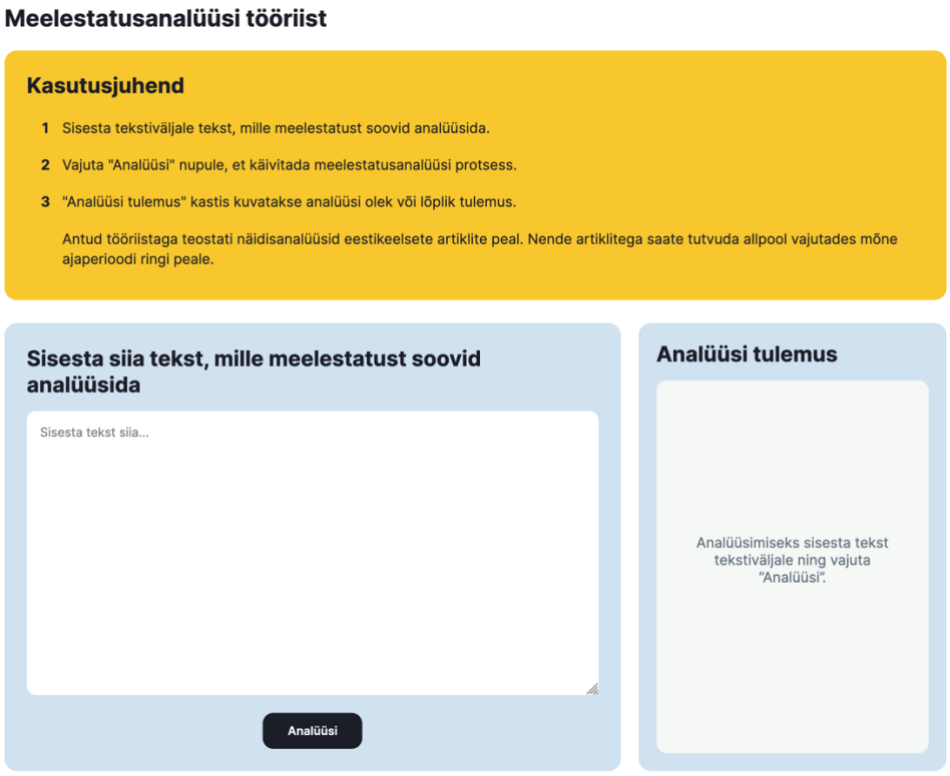
Testkasutajatele anti kolm ülesannet: sisestada tekst ning analüüsida selle meelestatust kasutades tööriista, vaadata tulemuse lausepõhist analüüsi, tutvuda näidisartiklitega. Testimine viidi läbi kasutades valjusti mõtlemise meetodit, mille käigus paluti kasutajatel tööriista kasutamise ajal oma mõtteid, tähelepanekuid ja tegevusi jooksvalt kirjeldada. Selline lähenemine võimaldab tuvastada kasutajaliidese elemente, mis võivad kasutajates segadust tekitada.

Pärast kasutajatestide läbiviimist analüüsisid autorid kogutud tagasisidet. Testimise tulemuste põhjal leiti, et kasutajaliides on arusaadav ning tööriista põhifunktsionaalsuse kasutamine ei tekitanud kasutajatele raskusi. Kuna tööriist keskendub ainult ühele põhifunktsionaalsusele, milleks on meelestatusanalüüsi teostamine, siis kasutusloogika oli kiiresti mõistetav.

Ainsa parandusettepanekuna toodi välja kasutusjuhendi paigutus kasutajaliideses. Testimise hetkel paiknes tekstisisestusväli enne kasutusjuhendit (vt joonis 8), mistõttu alustasid kasutajad tööriistaga meelestatusanalüüsi teostama enne juhendiga tutvumist. Kasutajatelt kasutajaliidese ülesehituse kohta tagasisidet küsides nõustusid nad, et

kasutusjuhendi paigutamine esimeseks elemendiks oleks muutnud esmase kasutamiskogemuse sujuvamaks.

Saadud tagasiside põhjal viidi tööriista lõplikus versioonis sisse muudatus, mille käigus paigutati kasutusjuhend enne põhifunktsionaalsuse kasutamise komponente (vt joonis 16).



Joonis 16. Kasutajaliidese valideerimisest saadud tagasiside muudatus kasutajaliideses.

8 Tulemused ja järeldused

8.1 Tööriista meelestatuse klassifitseerimistäpsus

Valminud tööriista toimimist testiti kahel viisil: analüüsides kogu Eesti valentsikorpuse andmestikku (varasema eksperimendi käigus testiti 20% andmestiku peal) ja luues tehisintellektiga näidistekstid, mille meelestatust analüüsida. Esimese testimismeetodi tulemused on esitatud tabelis 11.

Tabel 11. Arendatud meelestatusanalüüsi tööriista tulemused Eesti valentsikorpuse peal.

	Precision	***Recall***	***F1-score***	***Support***
Negatiivne	0.93	0.98	0.95	1927
Neutraalne	0.90	0.84	0.87	727
Positiivne	0.93	0.87	0.90	882
Accuracy			0.92	3536
Macro avg	0.92	0.90	0.91	3536
Weighted avg	0.92	0.92	0.92	3536

Tööriista üldine klassifitseerimistäpsus Eesti valentsikorpuse põhjal on 92%, mis on veidi madalam kui EstBERT128_sentiment mudeli eksperimendis. Selline erinevus on ootuspärane, sest tööriista hinnati suuremal ja mitmekesisemal andmekogul kui esialgses eksperimendis. Tööriista makro keskmine F1-skoor oli 91%. Sellest tulemusest saame järeldada, et tööriist suudab enamikel juhtudel määrata korrektselt meelestatusklassi ning on erinevaid klasse määrates tasakaalustatud. Tööriista klassifitseerimistulemus on kõige nõrgem neutraalse klassi puhul, kuid 87% F1-skoor on siiski kõrge tulemus.

Eesti valentsikorpuses ei ole esindatud pikemaid tekste, seega teise testimismeetodina otsustati analüüsida tehisintellekti poolt genereeritud tekste. See test hõlmas ka dokumenditaseme agregeerimisloogika toimivuse hindamist peatükis 6 kirjeldatud

meetodil. Tööriista toimimise hindamiseks pikkade tekstide korral genereeriti ChatGPT abil kontrollitud meelestatusega näidistekstid (tekstid välja toodud lisas 4), et läbi viia väikese mahuga kontrolltest. Antud testi oli kaasatud 9 teksti, nendest 3 positiivsed, 3 negatiivsed ja 3 neutraalsed. Loodud tööriist märkis korrektselt 8/9 tekstidest. Raskusi tekkis pika neutraalse teksti meelestatuse määramisega, mille klassifitseeris tööriist positiivseks. Tekst, kus tööriist andis erineva meelestatuse oodatud klassist, esines kalduvust positiivsuse poole. Tekstis kirjeldati erinevaid õppimisvõimalusi õpilastele, mis võisid anda positiivseid signaale.

Saadud tulemuste põhjal võib järeldada, et tööriist toimib nii lühikeste kui ka pikkade tekstide korral. Suurimaks väljakutseks osutus neutraalsete tekstide klassifitseerimine, kus tööriist kaldus mõnel juhul määrama teksti positiivseks või negatiivseks. Tööriista üldiseks klassifitseerimistäpsuseks kujunes 92%, mille põhjal leiti vastus ka töö teisele uurimisküsimusele loodud lahenduse täpsuse kohta.

8.2 Manuaalse ja automaatse meelestatusanalüüsi ajaline võrdlus

Testimaks, kui kaua kulub inimesel ühe artikli meelestatuse määramiseks võrreldes valminud automaatse tööriistaga, viisid autorid läbi illustratiivse eksperimendi ühe ERR-i artikli põhjal. Võrdluse tulemused on esitatud tabelis 12. Artikkel käsitles põhikooli lõpueksamite uut korda. Analüüsitud artikkel sisaldas 1155 sõna [72].

Tabel 12. Manuaalse- ja automaatse meelestatusanalüüsi tulemuste võrdlus.

	Inimene	**Tööriist**
Artikli meelestatuse hinnang	Negatiivne	Negatiivne 53% kindlusega
Aeg	6 minutit ja 27 sekundit	4.27 sekundit

Tööriist klassifitseeris artikli negatiivseks 53% kindlusega ning kogu analüüsiprotsess kestis 4,27 sekundit. Artikli üldise meelestatuse hindas tööriist negatiivseks, kuid tuvastas tekstis ka neutraalseid ning positiivseid lauseid.

Sama artiklit hindas seejärel inimene, kellel puudus eelnev kokkupuude töö ja loodud rakendusega. Inimesel kulus artikli lugemiseks ja üldise meelestatuse määramiseks 6 minutit ja 27 sekundit. Inimese hinnangul oli artikkel üldpildis samuti negatiivne, kuid

sisaldas suurel hulgal neutraalset sisu ning mõningaid positiivseid hoiakuid, mis muutis lõpphinnangu määramise mõnevõrra keerukamaks. Lõpphinnangu kohaselt jäi artikli meelestatus siiski negatiivseks.

Eksperimendi tulemused näitavad, et automaatne meelestatusanalüüsi tööriist suutis jõuda inimese hinnanguga sarnasele lõpptulemusele märgatavalt kiiremini. Kui inimese hinnangu andmine võttis üle kuue minuti, siis mudel suutis sama ülesande täita mõne sekundiga. See viitab, et automaatne meelestatusanalüüs võimaldab suurte tekstikorpuste töötlemist oluliselt efektiivsemalt kui täielikult manuaalne analüüs. Kuigi tegemist on ühe näitega ega võimalda laiemaid üldistusi, illustreerib see selgelt automaatse tööriista ajalist eelist manuaalse analüüsi ees.

8.3 RaRa tagasiside meelestatusanalüüsi tööriistale

RaRa-lt valminud tööriistale tagasiside saamiseks lähtuti samuti kvalitatiivsest andmete kogumise meetodist (kirjeldatud peatükis 2.2). Autorid organiseerisid RaRa esindajatega koosoleku, mille ülesehitus oli järgmine:

- kasutajateekonna demonstreerimine,
- erinevate tulemuste kuvamise viiside (dokumendi- ja lausetaseme) tutvustamine,
- näidisartiklite ajajoone kasutamise tutvustus,
- lõpliku tagasiside kogumine.

Kõikide sammude vältel paluti jagada tagasisidet lahendusele.

RaRa tagasiside meelestatusanalüüsi tööriistale oli positiivne. Tugevustena toodi välja intuitiivne kasutajaliides, mis samaaegselt on visuaalselt meeldiv ning ei vaja tehnilist ega masinõppealaseid teadmisi. Lisaväärtusena nähti lausepõhise analüüsi hüpikakent, mis muudab tööprotsessi kasutaja jaoks läbipaistvamaks ning võimaldab sooritada detailsemat analüüsi sisendi erinevate lausete lõikes.

Aruteludes toodi välja, et OCR-vigade parandamine ei olnud alati täpne. Samas märgiti, et ka varasemad OCR-vigade parandamise lahendused ei ole saavutanud täielikku täpsust, mistõttu nähti OCR-parandust kasuliku lisana.

Saadud tagasiside põhjal järeldati, et loodud tööriist vastas RaRa poolt seatud ootustele ning täitis eesmärgi pakkuda arusaadavat ja intuitiivselt kasutatavat meelestatusanalüüsi tööriista.

8.4 Blogipostitus „Eesti tuju ajas“

8.4.1 Andmete valimine

Sobilikud artiklid otsisid autorid DIGAR keskkonnast. Artiklid valiti järgmiste kriteeriumite alusel:

- ilma autoriõiguste piiranguteta,
- igapäeva artiklid, mitte spetsiifilise temaatikaga väljaanded,
- väljaannete ilmumisaasta jäi vahemikku 1925-2025.

Valitud väljaannete loetelu edastati RaRa töötajatele, kuna autoritel ei olnud võimalik DIGAR keskkonnast artikleid tekstifailidena alla laadida. RaRa töötajad koostasid nende põhjal ühe tekstifaili, mis sisaldas kõigi valitud artiklite tekste. Enne faili edastamist eemaldasid nad andmestikust reklaamartiklid ja muu mittesisulise materjali (näiteks tühjad artiklid).

Artiklite tekstifailina kättesaamine oli vajalik, sest käsitsi töö käigus loodud kasutajaliidese kaudu artiklite töötlemine oleks olnud ajakulukas. Andmete töötlemine automatiseeriti skripti abil, mida on võimalik edaspidi kasutada ka RaRa töötajatel sarnaste analüüside läbiviimiseks.

8.4.2 Andmete puhastus ja eeltöötlus

Andmete töötlemine toimus mitmeetapilise töötlusprotsessi abil, mille eesmärk oli filtreerida välja ebakvaliteetsed kirjed, puhastada tekst ning valmistada see ette OCR-vigade paranduseks ja meelestatusanalüüsiks.

Sisendandmed olid TSV-formaadis (*Tab-Separated Values*), kus iga kirje koosnes artikli identifikaatorist, artikli tüübist ja tekstisisust. Väljad olid omavahel eraldatud tabulaatori

märgiga. Töötlemisse jäeti ainult artiklid tüüpidega *article* ja *article+illustration*, sest ülejäänud kirjed ei sisaldanud meelestatusanalüüsi jaoks sobivat tekstilist sisu.

Esimeses etapis kontrolliti, kas artiklis leidub piisavalt tekstilist sisu. Selleks eemaldati HTML-märgistus ning HTML-i erimärkide kodeering. Kui pärast puhastust sisaldas tekst vähem kui 60 märki, jäeti artikkel edasisest töötlusest välja. Alla 60 märgi pikkused tekstid jäeti välja, et eemaldada väga lühikesed ja vähese sisulise väärtusega kirjed, mis ei olnud meelestatusanalüüsi jaoks piisavalt informatiivsed. Sellised tekstid sisaldasid sageli ainult pealkirju, katkist sisu, metaandmeid või muud müra, mis oleks võinud analüüsi tulemusi moonutada.

Järgmises etapis viidi läbi põhjalikum teksti puhastus. Korrastati reavahed, normaliseeriti erinevad sümbolid ja jutumärgid ning eemaldati tühjad read, ainult numbreid või sümboleid sisaldavad read, väga lühikesed read ning uudisteagentuuride märgendid nagu BNS, AFP, AP, ETA ja REUTER. Samuti eemaldati read, kus leidus vähem kui kaks tähte, et vähendada müra sattumist analüüsi. Kui pärast puhastust ei jäänud faili enam sisulist teksti alles, jäeti see edasisest töötlusest välja.

Seejärel rakendati tekstidele õigekirjaparandus ning pikemad tekstid jagati lausete põhjal väiksemateks osadeks, et võimaldada nende töötlemist.

Viimases etapis viidi läbi meelestatusanalüüs nii lause kui ka kogu dokumendi tasandil. Analüüsi tulemused salvestati struktureeritud kujul edasiseks statistiliseks analüüsiks ja visualiseerimiseks.

8.4.3 Tulemuste analüüs

Tulemuste analüüsimisel jaotati artiklid kümne aasta kaupa perioodidesse vahemikus 1925-2025. Selline lähenemine võimaldas võrrelda erinevate ajastute üldist meelestatust ning seostada tulemusi Eesti ühiskonnas toimunud ajalooliste sündmustega.

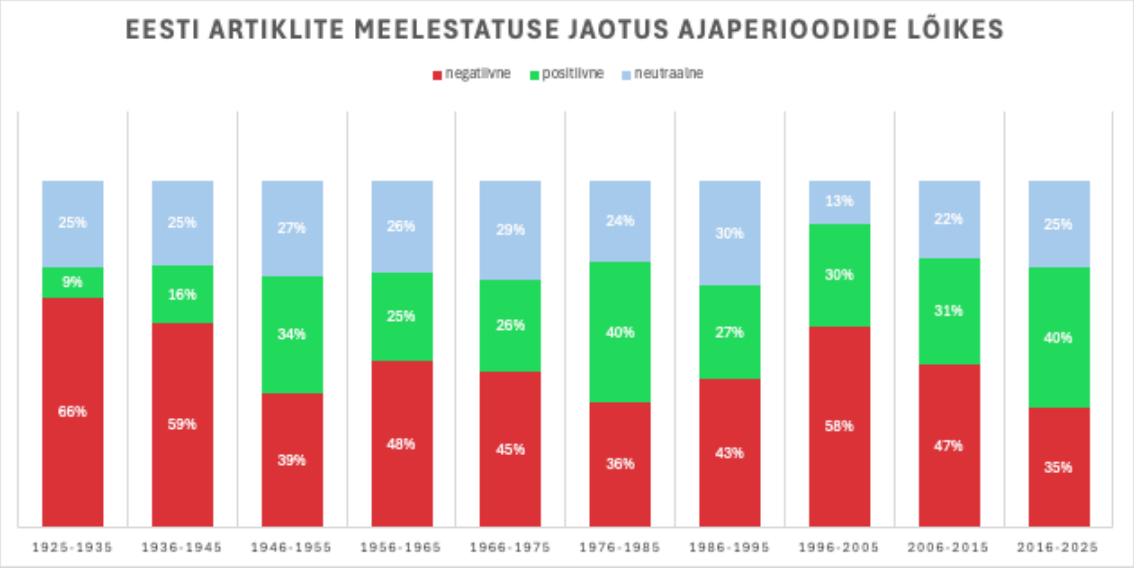
Iga perioodi puhul vaadeldi esmalt meelestatusanalüüsi tulemusi ehk negatiivsete, positiivsete ja neutraalsete artiklite osakaalu. Seejärel loeti ja võrreldi ka perioodidele iseloomulikke artikleid, et saada parem ülevaade, millest ajakirjandus kirjutas ning milline oli tekstide üldine toon ja käsitlusviis.

Tulemuste tõlgendamisel arvestati ka erinevate ajastute ajaloolist ja ühiskondlikku tausta. Analüüsi käigus vaadeldi, kuidas võisid suuremad sündmused ja muutused mõjutada artiklite toonivalikut ja käsitletavaid teemasid.

Analüüsi eesmärk ei olnud ainult hinnata, kas artiklid olid positiivsed või negatiivsed, vaid saada parem ülevaade, milline oli eri ajastute meediaruum ja kuidas ajakirjandus ühiskonnas toimuvat kajastas. Tuleb arvestada, et kõik artiklid ei ole ühiskondliku tähenduse poolest võrdsed. Mõni oluline kriis või ajalooline sündmus võib mõjutada ühiskonda rohkem kui väiksemad positiivsed uudised. Seetõttu kirjeldab meelestatusanalüüs eelkõige artiklite tonaalsust, mitte sündmuste tegelikku mõju ühiskonnas.

Analüüsi tulemuste paremaks tõlgendamiseks koostati joonised, mis näitavad artiklite meelestatuse muutumist ajaperioodide lõikes. Joonised aitavad võrrelda positiivsete, negatiivsete ja neutraalsete artiklite osakaalu ning tuua esile suuremaid muutusi Eesti meediaruumis.

Joonisel 17 on kujutatud positiivsete, negatiivsete ja neutraalsete artiklite protsentuaalne jaotus erinevatel ajaperioodidel.



Joonis 17. Eesti artiklite meelestatuse jaotus ajaperioodide lõikes.

Analüüsi käigus ilmnes mitmeid huvitavaid mustreid, mis üksnes numbriliste tulemuste põhjal ei olnud nähtavad.

Esiteks selgus, et positiivsem ajakirjanduse toon ei tähendanud alati ühiskonnas paremat olukorda. Näiteks nõukogude perioodil kasvas positiivsete artiklite osakaal, aga tegelikkuses toimusid repressioonid, küüditamised ja tugev ideoloogiline kontroll. See viitab sellele, et propaganda ja tsensuur mõjutasid tugevalt ajakirjanduse toonivalikut.

Teiseks tuli välja, et sama meelestatuse tulemus võib eri ajastutel tähendada täiesti erinevaid asju. Kui nõukogude perioodil võis positiivne toon viidata propagandale, siis taasiseseisvunud Eestis tähendas suurem negatiivsete artiklite osakaal hoopis vabamat ja kriitilisemat ajakirjandust. Probleemidest kirjutati avalikumalt ning rohkem käsitleti erinevaid arvamusi ja ühiskondlikke muresid.

Kolmandaks mõjutas tulemusi ka väljaande tüüp. Analüüsi viimases perioodis sisaldas andmestik rohkem maakonnalehti, kus lisaks kriisidele kajastati palju kohalikke kultuuri-, spordi- ja kogukonnasündmusi. See tõi artiklitesse rohkem positiivset ja neutraalset tooni.

Kogu sajandi lõikes oli näha, et ajakirjanduse keel ja toon muutusid ajas märgatavalt. Kui varasematel perioodidel domineeris ametlik ja ideoloogiline keelekasutus, siis hilisemates perioodides suurenes isiklike lugude, arvamuslugude ja avalike arutelude osakaal.

8.5 Edasine arendus

Töö käigus saavutatud tulemused näitasid, et kasutatud lähenemised olid töö eesmärkide täitmiseks sobivad. Sellegipoolest eksperimenteerimise käigus ilmnes ka mitmeid võimalikke edasiarendussuundi, mis võimaldaksid süsteemi tulevikus täiustada, kuid ei mahtunud käesoleva töö skoopi.

Ühe võimaliku edasiarendusena võiks täiustada OCR-vigade parandamise kvaliteeti. Käesolevas töös kasutatud EstLLM mudel võimaldas parandada mitmesuguseid OCR-vigu ning muutis tekstid enamasti semantiliselt paremini tõlgendatavaks. Samas ilmnes, et väga tugevalt moonutatud ajalooliste tekstide puhul jäi osa vigadest parandamata. OCR-vigade parandamist võiks tulevikus edasi arendada, kuna see ei olnud käesoleva töö põhifookus.

Samuti võiks edasi uurida kasutatud keelemudeli täiendavat peenhäälestamist. Käesolevas töös kasutati peamiselt olemasolevat eelõpetatud mudelit, mis saavutas juba head tulemused eestikeelses meelestatusanalüüsi ülesandes. Suuremahulise ja mitmekesisema eestikeelse märgendatud andmestiku olemasolul oleks võimalik mudelit täiendavalt kohandada ajalooliste tekstide ja ajakirjanduskeele eripärade jaoks.

Lisaks kolmele põhimeelestatuse klassile (positiivne, negatiivne ja neutraalne) võiks tulevikus rakendada ka emotsioonide klassifitseerimist, mille eesmärk on tuvastada tekstis konkreetseid emotsioone, näiteks rõõmu, kurbust, viha või hirmu. Selline lähenemine võimaldaks analüüsida tekste detailsemal tasemel ning anda täpsemat infot tekstis väljendatud emotsioonide kohta.

9 Kokkuvõte

Käesoleva magistritöö eesmärgiks oli luua koostöös Eesti Rahvusraamatukogu digilaboriga eestikeelsete tekstide meelestatusanalüüsi tööriist, mis võimaldaks automaatselt määrata tekstide üldist meelestatust. Töö keskendus vabavaraliste keelemudelite ja tehnoloogiate kasutamisele, et loodud lahendus oleks avalikult kasutatav.

Töö käigus eksperimenteeriti erinevate keelemudelitega ning võrreldi nende sobivust eestikeelse meelestatusanalüüsi ülesande lahendamiseks. Eksperimentidesse kaasati EstBERT128_sentiment, EstRoBERTa, XLM-RoBERTa ja EstLLM mudelid. Mudelite hindamisel kasutati F1-skoori, *accuracy*, *precision* ja *recall* mõõdikuid. Kõige täpsema tulemuse saavutas EstBERT128_sentiment mudel, mis osutus sobivaimaks lahenduseks loodava tööriista jaoks.

Käesoleva töö üheks oluliseks eripäraks oli OCR-vigade parandamise komponendi kaasamine meelestatusanalüüsi protsessi. Ajaloolised DIGAR portaali tekstid sisaldavad sageli OCR-töötlusest tingitud vigu. Autorid uurisid nende mõju meelestatusanalüüsi täpsusele ning katsetasid erinevaid parandamismeetodeid. Klassikalised OCR-vigade paranduse lahendused ei osutunud töö jaoks sobivaks, mistõttu rakendati EstLLM mudelil põhinevat lähenemist. Tulemused näitasid, et vigade parandamine muutis tööriista kindlusmäära ning mõnel juhul ka määratud meelestatusklassi, mis kinnitas OCR-vigade mõju meelestatuse tõlgendamisele.

Lisaks erinevate mudelitega eksperimenteerimisele arendati töö käigus ka veebipõhine kasutajaliides. See võimaldab kasutajal sisestada eestikeelset teksti, käivitada automaatne meelestatusanalüüs ning kuvada tulemust visuaalsel kujul. Tööriist toetab nii dokumendi- kui ka lausetasemel meelestatusanalüüsi ning võimaldab kasutajal tutvuda ka detailsema ülevaatega teksti erinevate osade meelestatusest. Valminud lahendust valideeriti kvalitatiivse kasutajatestimise abil, mille tulemused näitasid, et tööriist oli kasutajatele arusaadav ning vastas püstitatud eesmärkidele.

Töö praktilise väljundina viidi loodud tööriista abil läbi ka Eesti ajakirjanduse meelestatusanalüüs erinevatel ajaperioodidel ning koostati blogipostitus „Eesti tuju ajas“.

Kokkuvõttes saavutati töös püstitatud eesmärgid ning uurimisküsimustele leiti vastused. Töö tulemusena valmis praktiline meelestatusanalüüsi tööriist, mis ühendab OCR-vigade parandamise, keelemudelipõhise meelestatusanalüüsi ja kasutajaliidese.

Kasutatud kirjandus

[1] Eesti Keele Instituut, „Keelemudel,“ 19 06 2025. [Võrgumaterjal]. Available: https://sonaveeb.ee/search/unif/dlall/dsall/keelemudel/1/est. [Kasutatud 17 05 2026].

[2] Ollama, „Modelfile,“ [Võrgumaterjal]. Available: https://docs.ollama.com/modelfile. [Kasutatud 13 05 2026].

[3] Eesti Keele Instituut, „Eesti keele käsiraamat 2007,“ [Võrgumaterjal]. Available: https://arhiiv.eki.ee/books/ekk09/index.php?p=3&p1=2&id=117. [Kasutatud 13 05 2026].

[4] Eesti Keele Instituut, „Sõnaveeb,“ 23 01 2025. [Võrgumaterjal]. Available: https://sonaveeb.ee/search/unif/dlall/korpling/NLP/1/eng. [Kasutatud 13 05 2026].

[5] IBM, „What is optical character recognition (OCR)?,“ [Võrgumaterjal]. Available: https://www.ibm.com/think/topics/optical-character-recognition. [Kasutatud 14 05 2026].

[6] Eesti Rahvusraamatukogu, „RaRa digilabor,“ [Võrgumaterjal]. Available: https://digilab.rara.ee/meist/. [Kasutatud 13 05 2026].

[7] D. Jurafsky ja J. H. Martin, „Words and Tokens,“ 6 01 2026. [Võrgumaterjal]. Available: https://web.stanford.edu/~jurafsky/slp3/2.pdf. [Kasutatud 13 05 2026].

[8] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser ja I. Polosukhin, „Attention Is All You Need,“ *31st Conference on Neural Information Processing Systems*, Long Beach, CA, 2017.

[9] Visual Paradigm, „What is Unified Modeling Language (UML)?,“ [Võrgumaterjal]. Available: https://www.visual-paradigm.com/guide/uml-unified-modeling-language/what-is-uml/. [Kasutatud 14 05 2026].

[10] Eesti Entsüklopeedia, „Eesti ajaloo kronoloogia,“ [Võrgumaterjal]. Available: http://entsyklopeedia.ee/artikkel/eesti_ajaloo_kronoloogia1. [Kasutatud 14 05 2026].

[11] P. Mæhlum, D. Samuel, R. M. Norman, E. Jelin, Ø. A. Bjertnæs, L. Øvrelid ja E. Velldal, „It’s Difficult to be Neutral – Human and LLM-based Sentiment Annotation of Patient Comments,“ University of Oslo, Oslo, 2024.

[12] Koit, „Digar Eesti artiklid,“ 23 10 1975. [Võrgumaterjal]. Available: https://dea.digar.ee/?a=d&d=koit19751023. [Kasutatud 16 05 2026].

[13] Postimees (1886-1944), „DIGAR Eesti artiklid,“ 23 10 1925. [Võrgumaterjal]. Available: https://dea.digar.ee/article/postimeesew/1925/10/23/57. [Kasutatud 17 05 2026].

[14] H. Pajupuu, J. Pajupuu, R. Altrov ja K. Tamuri, „Estonian Valence Corpus / Eesti valentsikorpus,“ 2023.

[15] A. Gholamy, V. Kreinovich ja O. Kosheleva, „Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation,“ University of Texas at El Paso, El Paso, 2018.

[16] J. Opitz, „A Closer Look at Classification Evaluation Metrics and a Critical Reflection of Common Evaluation Practice,“ Transactions of the Association for Computational Linguistics, 2024.

[17] T. Vanicek ja S. Popelka, „The Think-Aloud Method for Evaluating the Usability of a Regional Atlas,“ *International Journal of Geo-Information,* kd. 12, nr 3, p. 95, 2023.

[18] W. Hwang ja G. Salvendy, „Number of people required for usability evaluation: the 10±2 rule,“ *Communications of the ACM,* kd. 53, nr 5, pp. 130-133, 2010.

[19] Microsoft, „Microsoft Excel,“ Microsoft, [Võrgumaterjal]. Available: https://www.microsoft.com/et-ee/microsoft-365/excel. [Kasutatud 16 05 2026].

[20] Microsoft, „Visual Studio Code,“ Microsoft, [Võrgumaterjal]. Available: https://code.visualstudio.com/. [Kasutatud 16 05 2026].

[21] Python Software Foundation, „Python,“ [Võrgumaterjal]. Available: https://www.python.org/. [Kasutatud 16 05 2026].

[22] S. Raschka, J. Patterson ja C. Nolet, „Machine Learning in Python: Main developments and technology trends in data science, machine learning, and artificial intelligence,“ *Information,* kd. 11, nr 4, p. 193, 2020.

[23] OpenAI, „Introducing GPT-5.5,“ 23 04 2026. [Võrgumaterjal]. Available: https://openai.com/index/introducing-gpt-5-5/. [Kasutatud 16 05 2026].

[24] Mozilla, „HTML: HyperText Markup Language,“ 22 12 2025. [Võrgumaterjal]. Available: https://developer.mozilla.org/en-US/docs/Web/HTML. [Kasutatud 16 05 2026].

[25] Mozilla, „CSS: Cascading Style Sheets,“ 29 12 2025. [Võrgumaterjal]. Available: https://developer.mozilla.org/en-US/docs/Web/CSS. [Kasutatud 16 05 2026].

[26] Mozilla, „JavaScript,“ 02 10 2025. [Võrgumaterjal]. Available: https://developer.mozilla.org/en-US/docs/Web/JavaScript. [Kasutatud 16 05 2026].

[27] FastAPI, „FastAPI,“ [Võrgumaterjal]. Available: https://fastapi.tiangolo.com/. [Kasutatud 16 05 2026].

[28] Tallinna Tehnikaülikool, „Intro,“ [Võrgumaterjal]. Available: https://ai-lab.pages.taltech.ee/en/intro/. [Kasutatud 21 04 2026].

[29] RunPod, „RunPod: AI and Cloud Infrastructure Provider,“ 2026. [Võrgumaterjal]. Available: https://www.runpod.io/. [Kasutatud 12 05 2026].

[30] K. Tan, C.-P. Lee ja K. Lim, „A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research,“ *Applied Sciences,* kd. 13, p. 4550, 2023.

[31] R. Botchway, A. B. Jibril, M. Kwarteng, M. Chovancová ja Z. Oplatková, „A review of social media posts from UniCredit bank in Europe: a sentiment analysis approach,“ *ICBIM '19: Proceedings of the 3rd International Conference on Business and Information Management,* pp. 74-79, 2019.

[32] M. Wnkhade, A. Rao ja C. Kulkarni, „A survey on sentiment analysis methods, applications, and challenges,“ *Artificial Intelligence Review,* kd. 55, pp. 5731-5780, 2022.

[33] G. Jaanimäe, *Eesti Wordnet ja meelestatuse analüüs,* Tartu: Tartu Ülikool, 2018.

[34] S. K. Uustalu, *Automated detection and sentiment analysis of registered entity mentions in Estonian language news media,* Tallinn: Tallinna Tehnikaülikool, 2019.

[35] R. Luukas, *Tartu Ülikooli õppeainete tagasiside meelsusanalüüs,* Tartu: Tartu Ülikool, 2023.

[36] M. Mets, I. Ibrus, A. Karjus ja M. Schich, „Automated stance detection in complex topics and small languages: The challenging case of immigration in polarizing news media,“ *PLoS ONE,* kd. 19, nr 4, 2024.

[37] L. Lehtmets ja M.-A. Meimer, *Ajalooliste eestikeelsete OCR tekstide järeltöötluse ja hindamise automatiseerimine Eesti Rahvusraamatukogu jaoks,* Tallinn: Tallinna Tehnikaülikool, 2025.

[38] J. Devlin, M.-W. Chang, K. Lee ja K. Toutanova, „BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,“ Google, 2018.

[39] A. Wang, A. Singh, J. Michael, F. Hill, O. Levy ja S. Bowman, „GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding,“ *ICLR 2019*, 2018.

[40] H. Tanvir, C. Kittask ja K. Sirts, „EstBERT: A Pretrained Language-Specific BERT for Estonian,“ Tartu Ülikool, Tartu, 2020.

[41] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzman, E. Grave, M. Ott, L. Zettlemoyer ja V. Stoyanov, „Unsupervised Cross-lingual Representation Learning at Scale,“ Facebook AI, 2020.

[42] A. Kotelnikova, D. Paschenko, K. Bochenina ja E. Kotelnikov, „Lexicon-Based Methods vs. BERT for Text Sentiment Analysis,“ *Analysis of Images, Social Networks and Texts: 10th International Conference, AIST 2021, Tbilisi, Georgia*, Tbilisi, Georgia, 2021.

[43] X. Qiu, T. Sun, Y. Xu, Y. Shao, N. Dai ja X. Huang, „Pre-trained Models for Natural Language Processing: A Survey,“ arXiv preprint arXiv:2003.08271v4, 2021.

[44] A. Dorkin, T. Purason, E. Kalbaliyev, H.-A. Kuulmets, M. Ojastu, M. Fišel, T. Alumäe, E. Aedmaa, K. Kruusmaa ja K. Sirts, „EstLLM: Enhancing Estonian Capabilities in Multilingual LLMs via Continued Pretraining and Post-Training,“ arXiv preprint arXiv:2603.02041, 2026.

[45] T. B. Brown, et al., „Language Models are Few-Shot Learners,“ *Conference on Neural Information Processing Systems* , Vancouver, Kanada, 2020.

[46] T. Z. Zhao, E. Wallace, S. Feng, D. Klein ja S. Singh, „Calibrate Before Use: Improving Few-Shot Performance of Language Models,“ *International Conference on Machine Learning (ICML)*, 2021.

[47] S. Laur, S. Orasmaa, D. Särg ja P. Tammo, „EstNLTK 1.6: Remastered Estonian NLP Pipeline,“ *Proceedings of The 12th Language Resources and Evaluation Conference*, Marseille, Prantsusmaa, 2020.

[48] T. Petmanson, „EstNLTK tutorial,“ 09 12 2014. [Võrgumaterjal]. Available: https://gitlab.keeleressursid.ee/timo-petmanson/estnltk/blob/8c9fbb2949156806f8fc4bfd7fcac6a93226b50c/docs/tutorial.rst. [Kasutatud 17 05 2026].

[49] Eesti Rahvusraamatukogu, „Digar,“ [Võrgumaterjal]. Available: https://testdea.nlib.ee/. [Kasutatud 17 05 2026].

[50] B. Dhingra, Z. C. Lipton ja D. Pruthi, „Combating Adversarial Misspellings with Robust Word Recognition,“ *CoRR,* kd. abs/1905.11268, pp. 5582-5591, 2019.

[51] Postimees (1886-1944), „DIGAR Eesti artiklid,“ 23 10 1925. [Võrgumaterjal]. Available: https://dea.digar.ee/article/postimeesew/1925/10/23/61. [Kasutatud 17 05 2026].

[52] Postimees (1886-1944), „DIGAR Eesti artiklid,“ 23 10 1925. [Võrgumaterjal]. Available: https://dea.digar.ee/?a=d&d=postimeesew19251023. [Kasutatud 16 05 2026].

[53] D. Del Fante ja G. M. Di Nunzio, „A Quantitative/Qualitative Approach to OCR Error Detection and Correction in Old Newspapers for Corpus-assisted Discourse Studies,“ *Proceedings of the 1st International Workshop on Computational Approaches to Historical Language Change*, 2021.

[54] Hunspell, „GitHub,“ [Võrgumaterjal]. Available: https://github.com/hunspell/hunspell. [Kasutatud 17 05 2026].

[55] D. Naber, „A Rule-Based Style and Grammar Checker,“ Universität Bielefeld, 2003.

[56] N. Kozloff, „GitHub,“ [Võrgumaterjal]. Available: https://huggingface.co/NikolayKozloff/llama-estllm-prototype-0825-Q8_0-GGUF. [Kasutatud 17 05 2026].

[57] W. X. Zhao, K. Zhou, J. Li, T. Tang, X. Wang, Y. Hou, Y. Min, B. Zhang, J. Zhang, Z. Dong, Y. Du, C. Yang, Y. Chen, Z. Chen, J. Jiang, R. Ren, Y. Li, X. Tang ja Z. Liu, „A Survey of Large Language Models,“ arXiv preprint arXiv:2303.18223, 2026.

[58] P. Liu, W. Yuan, J. Fu, Z. Jiang, H. Hayashi ja G. Neubig, „Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing,“ *ACM Computing Surveys,* kd. 55, nr 9, pp. 1-35, 2023.

[59] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Askell ja Welinder, „Training language models to follow instructions with human feedback,“ *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

[60] W. X. Z. e. al., „A Survey of Pretrained Language Models,“ Stanford University, 2022.

[61] Postimees (1886-1944), „DIGAR Eesti artiklid,“ 23 10 1925. [Võrgumaterjal]. Available: https://dea.digar.ee/article/postimeesew/1925/10/23/10. [Kasutatud 17 05 2026].

[62] Postimees (1886-1944), „DIGAR Eesti artiklid,“ 23 10 1925. [Võrgumaterjal]. Available: https://dea.digar.ee/article/postimeesew/1925/10/23/32.2. [Kasutatud 17 05 2026].

[63] M.-A. Meimer, „Hugging Face,“ 19 05 2025. [Võrgumaterjal]. Available: https://huggingface.co/mariannam/llammas-OCR-FT13k. [Kasutatud 17 05 2026].

[64] D. Jurafsky ja J. H. Martin, „Information Retrieval and Retrieval-Augmented Generation,“ %1 *Speech and Language Processing*, Stanford University, 2026.

[65] TartuNLP, „HuggingFace,“ 17 09 2024. [Võrgumaterjal]. Available: https://huggingface.co/tartuNLP/EstBERT128_sentiment. [Kasutatud 17 05 2026].

[66] J. Wei, X. Wang, D. Schuurmans, M. Bosma, E. H. Chi, Q. Le ja D. Zhou, „Chain of Thought Prompting Elicits Reasoning in Large Language Models,“ *36th Conference on Neural Information Processing Systems (NeurIPS 2022)*, 2023.

[67] J. Liu, D. Shen, Y. Zhang, B. Dolan, L. Carin ja W. Chen, „What Makes Good In-Context Examples for GPT-3?,“ 2021.

[68] S. Min, X. Lyu, A. Holtzman, M. Artetxe, M. Lewis, H. Hajishirzi ja L. Zettlemoyer, „Rethinking the Role of Demonstrations: What Makes In-Context Learning Work?,“ 2022.

[69] G. Choi, S. Oh ja H. Kim, „Improving Document-Level Sentiment Classification Using Importance of Sentences,“ *Entropy,* kd. 22, nr 12, p. 1336, 2020.

[70] R. Chandra, B. Zhu, Q. Fang ja E. Shinjikashvili, „Large language models for sentiment analysis of newspaper articles during COVID-19: The Guardian,“ 2024.

[71] Eesti Rahvusraamatukogu, „RaRa CVI,“ [Võrgumaterjal]. Available: https://www.figma.com/proto/E9b0XxYORdOObXsIEIzRvg/RaRa-CVI--client-file-?node-id=3046-2925&starting-point-node-id=3177%3A850. [Kasutatud 17 05 2026].

[72] ERR, „Koolijuhid ja õpetajad kritiseerivad põhikooli lõpueksamite uut korda,“ 11 05 2026. [Võrgumaterjal]. Available: https://www.err.ee/1610020132/koolijuhid-ja-opetajad-kritiseerivad-pohikooli-lopueksamite-uut-korda. [Kasutatud 17 05 2026].

[73] L. S. Sterling, The Art of Agent-Oriented Modeling, London: The MIT Press, 2009.

[74] Postimees (1886-1944), „DIGAR Eesti artiklid,“ 23 10 1925. [Võrgumaterjal]. Available: https://dea.digar.ee/?a=d&d=postimeesew19251023. [Kasutatud 17 05 2026].

Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks[1]

Meie, Margareth Lasn ja Hanna-Kristella Lehtsaar

1. Anname Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „Eesti tuju ajas. Eestikeelsete igapäevaartiklite meelestatusanalüüsi tööriista arendus Eesti Rahvusraamatukogu digilaborile.“, mille juhendaja on Martin Verrev
   1.1. reprodutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
   1.2. üldsusele kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. Oleme teadlikud, et käesoleva lihtlitsentsi punktis 1 nimetatud õigused jäävad alles ka autoritele.
3. Kinnitame, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest ning muudest õigusaktidest tulenevaid õigusi.

17.05.2026

1 Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal vastavalt üliõpilase taotlusele lõputööle juurdepääsupiirangu kehtestamiseks, mis on allkirjastatud teaduskonna dekaani poolt, välja arvatud ülikooli õigus lõputööd reprodutseerida üksnes säilitamise eesmärgil. Kui lõputöö on loonud kaks või enam isikut oma ühise loomingulise tegevusega ning lõputöö kaas- või ühisautor(id) ei ole andnud lõputööd kaitsvale üliõpilasele kindlaksmääratud tähtajaks nõusolekut lõputöö reprodutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. ja 1.2, siis lihtlitsents nimetatud tähtaja jooksul ei kehti.

Lisa 2 – Margareth Lasna isiklik panus ja eneseanalüüs

Käesolev magistritöö valmis kahe-liikmelise meeskonnatööna. Selline lähenemine töö teostamiseks sobis mulle, sest olen harjunud nii haridus- kui ka tööalaselt töötama tiimis. Kuna antud tööd tegime kursusekaaslasega kokku kaks semestrit, oli üksteise motiveerimine ja tööprotsessi järjepidavana hoidmine oluline osa töö valmimisel. Töömaht osutus suuremaks, kui projekt alguses tundus, mistõttu oli paaristöö antud juhul hea otsus. Töö käigus tuli samaaegselt tegeleda teaduskirjanduse analüüsi, mudelite katsetamise ja treenimisega, andmetöötluse, tarkvaraarenduse, kasutajaliidese disainimise ja loomisega.

Tööjaotus paarilisega tuli loomulikult. Otsustasime, et teadusartikleid ning teooriat analüüsime esmalt eraldi ning seejärel teeme üskteisele ülevaate olulistest leidudest. Töö praktilise osa suuremad komponendid jaotasime omavahel nii, et kumbki vastutaks ühe tööriista põhifunktsionaalsuse eest. Praktilist osa toetavad väiksemad tükid jaotasime vastavalt sellele, mis sobitusid meie põhifookusega töö juures. Lisaks tegime regulaarselt koosolekuid nii juhendaja kui ka RaRa esindajatega, kus analüüsisime erinevate lahenduste tulemusi ning valideerisime töö vahetulemusi. Kuna tegemist oli nii uurimis- kui ka praktilisetööga, tähendas see pidevat tehnilist ja metodoloogilist probleemide lahendamist kogu tarkvaraarendusprotsessi vältel.

Minu põhifookuseks oli läbi viia eksperimente keelemudelitega ja kirjutada meelestatusanalüüsi osa käsitlev *backend*'i kood. Selle käigus võrdlesin erinevate keelemudelite sobivust eestikeelse meelestatusanalüüsi ülesande lahendamiseks, viisin läbi mudelite *benchmark* katsed ning peenhäälestasin mudeleid, mis ei olnud meelestatusanalüüsi ülesande jaoks kohandatud. Tegemist oli tehniliselt keeruka tööga, kuna varasem kokkupuude *Python*-ga oli pigem olematu. See tähendas, et pidin paralleelselt õppima uut programmeerimiskeelt, süvenema masinõppe tööprotsessidesse ning mõistma, kuidas keelemudelite treenimine ja hindamine praktikas toimib. Alguses kasutasin tehisintellekti abi, et mõista uues keeles koodi kirjutamise eripärasid ning

kuidas korrektselt läbi viia eksperimente keelemudelitega. Töö käigus õppisin uut keelt tundma ning tegutsesin koodibaasis iseseisvalt.

Teiseks suureks ülesandeks oli mul tööriista kasutajaliidese loomine. Kuna olen varasemalt hobikorras kasutajaliideste projektidega tegelenud, siis siinkohal tundsin end kindlamalt. Samas ei olnud tegemist ainult visuaalse lahenduse loomisega, vaid eesmärgiks oli arendada kasutatav ja tehniliselt toimiv rakendus, mis suhtleks ka *backend* koodiga. Selle käigus oli huvitav õppida, kuidas API ühendusega *frontend*'i ja *backend*'i koodist tuleb kokku terviklik lahendus. Lisaks tuli tähelepanu pöörata sellele, kuidas meelestatusanalüüsi tulemust kasutajale loogiliselt kuvada, siin tulid appi arutelud nii tööpaarilise kui ka juhendajaga.

Kokkuvõttes oli antud magistritöö koostamine väga arendav kogemus mulle. Sain rakendada olemasolevaid ja õppida uusi tehnilisi oskusi arendusprojekti juures. Töö erines ka tavapärasest tarkvaraarenduse projektist, lisaks toimivale rakenduse loomisele tuli läheneda teaduslikult ning hinnata erinevate mudelite sobivust, analüüsida nende tulemusi vastavalt masinõppes kasutatavatele meetrikatele ning teha saadud tulemuste põhjal järeldusi, mis mõjutasid edasisi tehnilisi otsuseid.

Lisa 3 – Hanna-Kristella Lehtsaare isiklik panus ja eneseanalüüs

Käesolev magistritöö valmis kaheliikmelise meeskonnatööna. Selline töökorraldus sobis mulle hästi, kuna projekti maht ja teemade mitmekesisus oleks olnud üksi hallata keeruline. Töö käigus tuli tegeleda nii teaduskirjanduse analüüsi, mudelite katsetamise ja andmetöötluse kui ka praktilise tarkvaraarendusega. Paaristöö võimaldas jooksvalt ideid arutada, probleeme ühiselt lahendada ning hoida tööprotsessi stabiilsena kogu projekti vältel.

Projekti alguses keskendusime peamiselt tekstide kvaliteedi hindamisele, kuid töö käigus selgus, et ainult sellest ei piisa ning vajalikuks osutus ka OCR-vigade parandamise lahenduse loomine. See muutis töö tehniliselt keerulisemaks ja suurendas oluliselt selle mahtu. Töö kujunes kombinatsiooniks uurimuslikust ja praktilisest arendusprojektist, kus tuli pidevalt katsetada erinevaid lahendusi, hinnata nende tulemusi ja teha selle põhjal edasisi tehnilisi otsuseid.

Minu peamiseks vastutusalaks kujunes EstLLM-il põhineva OCR-paranduse lahenduse arendamine. Lisaks katsetasin EstLLM-i kasutamist meelestatusanalüüsi ülesandes, eksperimenteerides erinevate juhiste sõnastamise strateegiatega, et saavutada stabiilsemaid tulemusi. Tegelesin ka tagarakenduse poolel mudelite käivitamise, andmetöötluse ning skriptide loomisega, mis võimaldasid automatiseerida artiklite töötlemist ja analüüsi. Samuti osalesin meelestatusanalüüsi tulemuste agregeerimise loogika väljatöötamises ning realiseerisin dokumenditaseme meelestatuse ja kindlusmäära arvutamise lahenduse. Osalesin ka blogipostituse „Eesti tuju ajas“ kirjutamises. Selle jaoks jagasime analüüsitavad ajaperioodid omavahel kümnendite kaupa, analüüsisime saadud tulemusi ning kirjutasime nende põhjal blogipostituse.

Praktikas osutus töö keerulisemaks, kui algselt eeldasin. Keelemudeli väljund ei olnud alati stabiilne ning mõnel juhul kaldus mudel genereerima soovimatut lisateksti või väljuma etteantud ülesande piiridest. Nende probleemide vähendamiseks tuli katsetada

erinevaid süsteemijuhiseid, stopptokeneid ja valideerimisreegleid. Lisaks tuli leida tasakaal agressiivse OCR-paranduse ja algse teksti tähenduse säilitamise vahel.

Töö käigus uurisin ka klassikaliste tekstiparandustööriistade sobivust OCR-tekstide töötlemiseks. Praktikas ilmnes mitmeid piiranguid nii eesti keele toe, kodeeringute kui ka olemasoleva töövooga integreerimise osas. Seetõttu keskendusin lõpuks EstLLM-il põhineva lahenduse edasiarendamisele ja stabiliseerimisele.

Tehniliselt oli magistritöö minu jaoks suur õppimiskogemus. Mul puudus varasem praktiline kokkupuude suurte keelemudelite, masinõppe töövoogude ja tagarakenduse arendusega. Kuigi olin Pythonit varem kasutanud, tuli töö käigus õppida uusi teeke, tööriistu ja arendusvõtteid. Projekti algfaasis kasutasin ChatGPT abi peamiselt tehniliste probleemide lahendamisel ja koodi mõistmisel, kuid töö edenedes muutusin järjest iseseisvamaks ning suutsin suure osa probleemidest juba iseseisvalt lahendada.

Kokkuvõttes oli magistritöö minu jaoks väga huvitav ja arendav kogemus. Töö käigus pidin korduvalt katsetama erinevaid lahendusi, hindama nende sobivust ning tegema tulemuste põhjal tehnilisi otsuseid. Sain väärtusliku praktilise kogemuse nii arenduses kui ka keelemudelite rakendamises.

Lisa 4 – ChatGPT poolt genereeritud tekstis meelestatusanalüüsi tööriista täpsuse testimise jaoks [23]

Positiivne – lühike tekst:

Täna oli üle pika aja tõeliselt hea päev. Sain kõik tööülesanded õigeks ajaks valmis ning meeskonnalt tuli väga positiivset tagasisidet. Õhtul käisin ujumas ja tundsin end pärast seda energilise ja motiveerituna.

Negatiivne – lühike tekst:

Olen viimastel nädalatel järjest rohkem pettunud selle teenuse kvaliteedis. Vastuseid peab ootama päevi ning probleemid jäävad lõpuks ikka lahendamata. Sellise kogemuse järel ei taha ma seda enam kasutada.

Neutraalne – lühike tekst:

Koosolek toimub esmaspäeval kell 10. Arutatakse järgmise kvartali eesmärke ning projekti ajakava muudatusi. Kõik osalejad saavad päevakorra e-posti teel.

Positiivne – keskmine tekst:

Käisin nädalavahetusel Saaremaal ning kogu reis ületas ootusi. Ilm oli soe, majutus hubane ja inimesed väga sõbralikud. Eriti meeldis mulle see, kui rahulik kogu keskkond oli võrreldes linnaga. Veetsime suure osa ajast looduses jalutades ning õhtuti grillides. Sellest reisist jäi tunne, et tahaks sinna kindlasti tagasi minna.

Negatiivne – keskmine tekst:

Tellisin uue sülearvuti vähem kui kuu aega tagasi, kuid juba praegu esineb pidevalt tehnilisi probleeme. Aku tühjeneb väga kiiresti, seade kuumeneb ning mitmed programmid jooksevad kokku. Lisaks oli klienditugi aeglane ja ebaprofessionaalne. Sellise hinna juures ootasin märksa paremat kvaliteeti.

Neutraalne – keskmine tekst:

Tallinna ühistransport koosneb bussidest, trammidest ja trollidest. Enamik liine töötab varahommikust hilisõhtuni. Piletisüsteem võimaldab kasutada nii ühiskaarti kui ka pangakaarti. Viimastel aastatel on uuendatud mitmeid tramme ja peatuseid.

Positiivne – pikk tekst:

Viimase aasta jooksul olen hakanud palju teadlikumalt oma aega planeerima ning see on minu igapäevaelu märgatavalt parandanud. Kui varem tundus, et kõik ülesanded kuhjuvad korraga ning pidevalt on kiire, siis nüüd olen loonud kindla süsteemi, mille järgi oma päeva üles ehitan. Kasutan kalendrit, jagan suuremad eesmärgid väiksemateks sammudeks ning jätan teadlikult aega ka puhkamiseks.

Kõige rohkem on muutunud minu keskendumisvõime. Tunnen, et suudan ülesannetesse süveneda palju rahulikumalt ning tööpäeva lõpuks ei ole enam tunnet, et midagi jäi tegemata. Lisaks on see aidanud vähendada stressi ja parandada und. Ka inimesed minu ümber on märganud, et olen rahulikum ja motiveeritum.

Kuigi vahel tuleb endiselt ette keerulisemaid perioode, on üldine muutus olnud väga positiivne. Tunnen, et mul on oma tegemiste üle rohkem kontrolli ning see annab kindlustunnet ka tuleviku suhtes.

Negatiivne – pikk tekst:

Viimased kuud selles ettevõttes on olnud äärmiselt kurnavad. Töökoormus suureneb pidevalt, kuid samal ajal ei ole meeskonda lisandunud uusi inimesi ega parandatud töökorraldust. Sageli tuleb teha ületunde ning paljud ülesanded antakse ette väga lühikese tähtajaga. See on tekitanud olukorra, kus inimesed on pidevalt stressis ja väsinud.

Kõige rohkem häirib mind kommunikatsioon juhtkonna poolt. Probleemidest räägitakse küll koosolekutel, kuid reaalseid lahendusi ei järgne. Samuti jääb tunne, et töötajate tagasisidet ei võeta tõsiselt. Mitmed kolleegid on viimase paari kuu jooksul lahkunud ning ausalt öeldes mõistan nende otsust täielikult.

Kuigi töö ise võiks olla huvitav ja arendav, varjutab halb töökorraldus kõik positiivse. Praegusel hetkel kaalun tõsiselt ka ise uue töökoha otsimist.

Neutraalne – pikk tekst:

Eesti kõrgharidussüsteem koosneb rakenduskõrgharidusõppest, bakalaureuseõppest, magistriõppest ja doktoriõppest. Õppetöö toimub nii eesti kui ka inglise keeles ning mitmed kõrgkoolid osalevad rahvusvahelistes koostööprogrammides. Üliõpilastel on võimalik taotleda stipendiume, õppelaenu ning osaleda vahetusprogrammides.

Viimastel aastatel on kõrgkoolid investeerinud rohkem digilahendustesse ja veebipõhisesse õppesse. See võimaldab tudengitel osaleda loengutes ka distantsilt ning kasutada erinevaid elektroonilisi õppematerjale. Lisaks tehakse järjest enam koostööd ettevõtetega, et pakkuda praktilisemaid õpivõimalusi.

Kõrghariduse eesmärk on pakkuda teadmisi ja oskusi erinevates valdkondades ning toetada teadus- ja arendustegevust. Eestis tegutseb mitu avalik-õiguslikku ülikooli ning rakenduskõrgkooli, mille õppekavad erinevad oma fookuse ja spetsialiseerumise poolest.

Lisa 5 – EstLLM OCR-paranduse modelfile põhisisu

```
FROM NikolayKozloff/llama-estllm-prototype-0825-Q8_0-GGUF

SYSTEM
Oled eesti keele OCR-tekstiparandaja.
Sisend võib olla nii vana kui tänapäevane eestikeelne ajalehetekst.
Paranda OCR-vead: valesti tuvastatud tähed, puuduvad tähed ja
üleliigsed tähed.
Kasuta konteksti vigade tuvastamiseks, kuid ära mõtle uusi sõnu välja.
Paranda ainult sõnade kirjapilti, mitte lauseid tervikuna.
Ära muuda teksti tähendust.
Ära lisa kommentaare ega küsimusi.
Tagasta ainult parandatud tekst.

PARAMETER num_ctx 8192
PARAMETER temperature
```

Lisa 6 – Blogipostitus „Eesti tuju ajas“

Sissejuhatus

Viimase saja aasta jooksul on Eestis toimunud palju suuri muutusi. Meie riik on kogenud iseseisvumist, okupatsioone, nõukogude perioodi ning lõpuks kujunenud tänapäevaseks demokraatlikuks ja digitaalseks ühiskonnaks. Kõik need sündmused on jätnud jälje ka ajakirjandusse.

Meedia ei kajasta ainult sündmusi, vaid peegeldab ka üldist õhkkonda ja meeleolu või vähemalt seda meeleolu, mida lubati väljendada. Võttes arvesse meie riigi sündmusterohket ajalugu, tekib küsimus: kuidas kujunes eestlaste meelsus ajakirjanduses sajandi jooksul? Sellele küsimusele vastuse leidmiseks analüüsisime Eesti ajakirjanduse artikleid alates 1925. aastast kuni 2025. aastani.

Analüüsi vaatluse alla võtsime 4557 Eesti igapäeva artiklit, mis olid kättesaadavad DIGAR keskkonnast. Protsessi sujuvaks kulgemiseks kasutasime meelestatuse määramiseks RaRa Digilabori meelestatusanalüüsi tööriista. Tuleb arvestada, et kõik artiklid ei ole ühiskondliku tähenduse poolest võrdsed. Mõni oluline kriis või ajalooline sündmus võib mõjutada ühiskonda rohkem kui väiksemad positiivsed uudised. Seetõttu kirjeldab meelestatusanalüüs eelkõige artiklite tonaalsust, mitte sündmuste tegelikku mõju ühiskonnas.

Eesti tuju ajas

Kui vaadata sadat aastat korraga, jääb esmapilgul mulje, et elu Eestis on sellel perioodil kulgenud pigem negatiivsel toonil. Küll aga nii pealiskaudse analüüsiga ei saa piirduda, sest peab arvestama ajalooliste sündmustega, mis mõjutasid ajakirjandus väljendusviise ja võimalusi.

Sajandi jooksul on Eesti ajakirjandus saanud tegutseda vabalt, kuid on kogenud ka tsensuuri ja propaganda dikteerimist. Seepärast ei saa sajandi pikkust meelestatust vaadata koondatult. Ühel ajastul võib kõrge positiivsus viidata propagandale. Teisel

ajastul kõrge negatiivsus, aga vabale ühiskonnale. Tulemused võivad olla kümnendite lõikes samad, kuid nende tähendus muutub sündmuste valguses.

1925-1935: Maksud, määrused ja masendus

Esimene kümnend osutus kogu analüüsi kõige süngemaks. 66% artiklitest olid negatiivse tooniga, vaid 9% positiivsed.

Artikleid lugedes torkab silma, kui palju ruumi võtsid ametlikud teadaanded, määrused ja administratiivsed arutelud. Räägiti maksudest, riigikorraldusest, hariduspoliitikast. Ühes 1929. aasta Postimehe artiklis vaidlustati näiteks haridusministri ettepanek asendada koolides inglise keel saksa keelega ja kritiseeriti läbipaistmatut otsustusprotsessi. Arutelu keskendus küll formaalselt hariduspoliitikale, kuid peegeldas ühtlasi laiemat küsimust: millise kultuurilise suuna poole Eesti peaks liikuma?

Aeg ise oli pingeline. 1929. aastal puhkes ülemaailmne majanduskriis, mis tabas Eestit valusalt. Olime tugevalt sõltuvad ekspordist ja põllumajandusest, samal ajal kasvasid poliitilised pinged. Negatiivne toon ei olnud seega juhuslik, see peegeldas tegelikku ebakindlust.

1936-1945: Kõik on kontrolli all

Võrreldes eelmise kümnendiga vähenes negatiivsete artiklite osakaal, kuid säilitas ülekaalu positiivsete ja neutraalsete üle. Sellel perioodil toimusid Eesti pinnal sõda ja repressioonid, mis võisid panna meediaväljaanded peegeldama hirmu, kaotust ning konflikte ühiskonnas.

Huvitav leid on positiivsete artiklite osakaalu kasv, 9%-lt 16%-le. Kui võtta arvesse, et tegemist oli eestlaste jaoks äärmiselt raske perioodiga, võib üheks tõlgenduseks olla ajakirjanduse tsensuur ning propaganda laialdane levik. Näiteks jäi silma lõik Harju elu ENSV 04.04.1945 väljaandes: *"Nii lööme puruks meie rahva ja kõigi vabadust armastavate rahvaste ühise vaenlase ja toome tagasi õnneliku tuleviku kõigile nõukogude rahvastele."* Lausest jääb kõlama eriti teine pool, kus ilmneb selge ideoloogiline keelekasutus.

1946-1955: Optimism vastu tahtmist

Sõjajärgsel perioodil toimus ootamatu nihe. Negatiivsete artiklite osakaal langes 39%-ni ja positiivseid artikleid moodustas lausa 34%. See on eelmise kümnendiga võrreldes ligi kahekordne kasv.

Mida see tähendab? Kindlasti mitte seda, et inimestel hästi läks. Eesti oli liidetud Nõukogude Liiduga. 1949. aastal küüditati tuhandeid eestlasi Siberisse ning talupojad kaotasid maad, mis nende peredele oli kuulunud põlvest põlve. Samas rääkis ajakirjandus hoopis teistsugust lugu.

Artikleid lugedes tekib kummaline tunne. Kirjutati tootmisplaanide ületamisest, riigilaenu tähtsusest rahva heaolule ja nõukogude noorte eredast tulevikust. Ühes 1946. aasta artiklis selgitatakse, kuidas Eesti NSV on kaasa tõmmatud võimsa arenguprotsessiga, mis annab noortele senisest kaugelt rohkem võimalusi. Kolhoosid said nimedeks "Lembitu", "Noorus" ja "Tee Kommunismile".

Numbrid ei mõõda siin ühiskonna tegelikku meeleolu, vaid süsteemi häält. Propaganda ei kadunud ajakirjandusest ka järgmistel kümnenditel.

1956-1965: Sula tuuled

Sellel perioodil oli negatiivsete artiklite osakaal 48%, kuid märkimisväärne on, et neutraalsete ja positiivsete artiklite ühine osakaal ületas negatiivsete oma.

Antud perioodil toimus Hruštšovi sula, kus süsteem NSV liidus pehmenes ning toimus kontrollitud liberaliseerimine. Ka Eestit mõjutas see aeg, kus hakkas elavnema kultuurielu. Üheks märkimisväärseks sündmuseks saab lugeda Tallinna uue laululava valmimise. Sellisele muutusele viitab ka positiivsete ja neutraalsete artiklite ühine ülekaal üle negatiivse klassi.

1966-1975: Üks samm edasi, kaks sammu tagasi

Järgmisel kümnendil astuti samm tagasi. Negatiivsete artiklite osakaal kasvas uuesti 45%-ni, positiivseid oli 26% ja neutraalseid 29%. Toon muutus jällegi formaalsemaks.

1968. aastal näitas Nõukogude Liit Praha kevade mahasurumisega, et liiga suuri vabadusi lubada ei kavatseta. Eesti ajakirjandus ei kommenteerinud Prahat otseselt, kuid

ideoloogiline surve karmistus taas. Artiklites domineerisid töökollektiivid, tootmisplaanid ja ühiskondlike eesmärkide täitmine.

Samas ei olnud tegemist enam stalinistliku hirmuajaga. Üha rohkem kirjutati spordiüritustest, teatrietendustest ja kohalikust elust, killud tavalisest igapäevaelust eksisteerisid kõrvuti ametliku ideoloogiaga.

1976-1985: (Olümpia)pidu ja (mitte)pillerkaar

Hilise nõukogude perioodi meelestatuse analüüs näitab suhteliselt tasakaalukat aega: positiivseid artikleid oli 40%, negatiivseid 36%.

Selle perioodi üheks suureks sündmuseks eestlaste jaoks oli Moskva Olümpiamängude purjeregatt, mis toimus Tallinnas. Selle raames parendati Tallinna infrastruktuuri ning tõsteti tolle aegset nõukogude Eestit mingil määral maailma areenile. Kuid selle sündmuse taustal oli ühiskonnas rahulolematust, mis avaldus Tallinnas toimunud noorte rahutustes ja meeleavaldustes.

1986-1995: Vabadus hüüab tulles

See on periood, mida on kõige raskem numbrites kinni püüda. Negatiivsed artiklid moodustavad 43%, positiivsed 27% ja neutraalsed 30%. Vaid numbrite põhjal on keeruline mõista, mis ühiskonnas tegelikult toimus.

Gorbatšovi glasnost lõdvendas tsensuuri. Eesti ajakirjandus hakkas kirjutama asjadest, millest oldi aastakümneid vaikitud, näiteks küüditamistest, MRP salajastest protokollidest ja ühiskonna tegelikest probleemidest. 1988. aastal loodi Rahvarinne, 1989. aastal moodustasid eestlased, lätlased ja leedulased Balti keti. Ka ajalehed käsitlesid neid sündmusi varasemast palju avatumalt. Kaubapuudust, majanduskriisi ja määramatust käsitleti ausalt. Avalikus ruumis hakati lõpuks rääkima teemadest, millest varem vaikiti.

Periood lõppes Eesti iseseisvuse taastamisega 1991. aastal ja sellega lõppes ka ajakirjanduse topeltelu. Enam ei pidanud kirjutama ühte, mõeldes teist. Artiklite negatiivne toon on siin vabaduse märk, mitte meeleheite väljendus. Ajakirjandus sai lõpuks tõtt rääkida.

1996-2005: Euroopa Liit, siit ma tulen!

Taasiseseisvunud Eesti algusajad olid põnevad. Käisid läbirääkimised NATO ja Euroopa Liiduga ühinemiseks, toimus esialgne digitaliseerimine, Tallinna vanalinn kanti UNESCO maailmapärandi nimistusse, Erki Noolest sai olümpiavõitja. Neid positiivseid muutusi on tunda ka meelestatuse analüüsist: positiivseid artikleid oli 30%.

Kuid tegemist ei olnud meedia poolest eestlaste jaoks täielikult positiivse ajaga. Negatiivsete artiklite osakaal antud perioodil oli 58%. See võib olla tingitud vabast ajakirjandusest, kus julgeti meedias avalikult arvamust avaldada ühiskonnas toimuva üle. Näiteks hakkasid valimis korduma artiklid, mis tõid välja iseseisvunud riigile ülemineku probleeme: “Kinnisvaraturul kiireneb hindade tõus. Eesti kinnisvaraturul on viimastel nädalatel kiirenenud hindade tõus ” (Eesti Päevaleht 07.05.1996).

Siin hakkab välja kujunema huvitav trend: **kas demokraatia märgiks võib olla kriitiline ja probleemikeskne meedia hoiak?**

2006-2015: Mis meelel, see keelel

Periood algas optimismiga. Eesti oli 2004. aastal liitunud nii NATO kui ka Euroopa Liiduga, majandus kasvas kiiresti. Ent numbrid , 47% negatiivseid, 31% positiivseid, 22% neutraalseid, peegeldavad pigem optimismi kadumist kui buumi kõrghetke.

2008. aasta sügis muutis palju. Ülemaailmne finantskriis tabas Eestit valusalt. SKP langes, töötus kasvas kiiresti ning ehitusbuumist jäid maha tühjad kortermajad. Ajakirjandus täitus artiklitega töötusest, hinnatõusust ja inimeste toimetulekuraskustest. Samal ajal valmistuti euro kasutuselevõtuks ning riik viis ellu kokkuhoiumeetmeid, mis tekitasid ühiskonnas pingeid ja palju arutelu.

Selle perioodi juures tuleb aga hästi välja üks oluline muutus. Ajakirjanduse toon oli täiesti erinev varasematest kümnenditest. See ei olnud enam formaalne. Probleemidest kirjutati ausalt ja mitmekülgselt. Kõrvuti olid päris inimeste lood, erinevad arvamused ja avalikud vaidlused. **Eesti ajakirjandus ei öelnud enam inimestele, mida mõelda, vaid lõi ruumi avalikuks aruteluks**. See on muutus, mida ainult numbrite põhjal ei ole võimalik näha.

2016-2025: Mine maale!

Viimane periood on analüüsitulemuste põhjal üllatavalt huvitav. Toimus palju negatiivseid sündmusi: inflatsioon, pandeemia, sõda Ukrainas. Kuid vaatamata sellele olid tulemused tasakaalus: negatiivseid artikleid 35%, positiivseid lausa 40%.

Perioodi kohta oli analüüsi kaasatud rohkem maakonna lehti. Avastasime, et just need ajalehed on leidnud hea tasakaalu negatiivsete ja positiivsete artiklite vahel. Kajastatakse küll päevakajalisi negatiivseid sündmusi, kuid lehe vahelt leiab ka artikleid, kus tunnustatakse maakonna parimaid sportlasi, ettevõtjaid, kultuuritegelasi.

Analüüsi leiud näitavad, et maakonna lehtedel on rohkem võimalusi kajastada tasakaalustatud sisu. Nende repertuaaris on kogukondlikud sündmused ja kohalike inimeste tunnustamine, mis toovad sisse rohkem neutraalset ja positiivset meelestatust igapäeva artiklitele.

Mida need sada aastat meile näitasid?

See saja aasta pikkune teekond ilmutab kolme suurt mustrit.

Esiteks: propaganda tõstab kunstlikult positiivsust. Nõukogude perioodil kajastus artiklites süsteemi hääl, mitte inimeste.

Teiseks: demokraatlikus ühiskonnas on loomulik, et ajakirjandus kajastab nii positiivseid kui ka negatiivseid sündmusi ning käsitleb ühiskondlikke probleeme avalikult ja kriitiliselt.

Kolmandaks: meeleolu sõltub sellest, milline meedia neid kajastab. Maakonnalehed ei keskendu ülemaailmsetele kriisidele, nagu riiklikud väljaanded.

Postituse kirjutamisel lähtuti Eesti ajaloo kronoloogia sündmustest [10].