

TALLINNA TEHNIKAÜLIKOOL

Infotehnoloogia teaduskond

Markus Järavere 232719IABB

Sander Talisoo 223242IABB

Oliver Villomann 233385IABB

**Konkurentide monitoorimise rakenduse
realiseerimine finantsasutuse töö efektiivsuse
tõstmiseks**

Bakalaureusetöö

Juhendaja: Viljam Puusep

MSc

Tallinn 2026

Kokkuvõte

Eesti pangandusturul on finantsteenuste hinnakujundus äärmiselt dünaamiline valdkond, kus intressimäärad, teenusetasud ja kampaaniatingimused muutuvad pidevalt. Tarbijate ja finantsasutuste jaoks tähendab see suurt vajadust pidevalt jälgida ning võrrelda erinevate krediidasutuste pakkumisi. Kuna Eesti turul puudub ühtne, standardiseeritud ja sõltumatu liides või API-platvorm (välja arvatud PSD2 direktiivi spetsiifilised kasutusjuhud), mis koondaks ajakohased ja omavahel võrreldavad andmed kõigi suuremate pankade hinnakirjadest, nõuab konkurentide analüüs märkimisväärset manuaalset tööd. Eelnevate projektide käigus on selgunud, et traditsioonilised veebikraapimise tehnoloogiad, mis põhinevad CSS-selektoritel ja XPath-päringutel, on sellise andmemahu haldamiseks liiga haprad ning lakkavad töötamast isegi kõige väiksemate pangaveebide struktuurimuudatuste peale.

Lõputöö peamiseks eesmärgiks on välja arendada terviklik, automatiseeritud ja vastupidav veebirakendus, mille abil saab finantsasutus monitoorida suurimate Eestis tegutsevate konkurentide finantsteenuste hinnakirju ja sotsiaalmeedia turundustegevusi. Seatud ülesannete hulka kuulusid uue keelemudelitel (LLM) põhineva kraapimisarhitektuuri projekteerimine, laiendatud tootevaliku analüüsimine, andmete visualiseerimine, automaatsete hindade muutmise teavituste saatmine ning tehisintellektil põhineva turundusstrateegiate kokkuvõtte genereerimine.

Loodud lahendus toetub klassikalisele kolmetasandilisele arhitektuurile, mis koosneb esitluskihist, äriloogikakihist ja andmekihist.

- Esitluskiht (klientrakendus) on arendatud Nuxt 3 raamistikuga, mis tagab serveripoolse renderdamise ning komponendipõhise ülesehituse. Klientrakendus koosneb avalikust võrdluslehest, kus kasutaja saab dünaamilistes tabelites kõrvutada toodete parameetreid, ning turvatud administraatoripaneelist.
- Äriloogikakiht (serverirakendus) tugineb ASP.NET Core 8 platvormile, mis haldab REST API liidest ning taustateenuseid. Autentimine toimub turvaliselt JWT-tokenite abil, mis piirab kraapimisprotsesside ja teavituste käsitsi käivitamist vaid administraatoriõigustega kasutajatele.

- Andmekiht kasutab PostgreSQL relatsioonilist andmebaasi, integreerides Entity Framework Core'i. Andmebaas on üles ehitatud paindlikule EAV (Entity-Attribute-Value) disainimustrile, mis lubab süsteemi tulevikus sujuvalt laiendada uute tootekategooriatega ilma andmebaasi skeemi muutmata.
- Kogu rakendus on töökindluse ja juurutamise lihtsuse tagamiseks konteineriseeritud Dockeri ja Docker Compose'i abil.

Süsteemi tehniline tuum ja peamine innovatsioon seisneb kolmekihilises kraapimisarhitektuuris, mis asendas varasema ebastabiilse lahenduse. Süsteem käivitub igapäevaselt 24-tunnise intervalliga automatiseeritud taustateenuse vahendusel. Esmalt avab Playwrighti brauseriautomatiseerimise raamistik pangaveebi (või klõpsab lahti JavaScripti renderdatavad alamlehed) ja eemaldab vajadusel botitõrjed. Seejärel edastatakse sisu Jina.ai dokumenditeisendusteenusele, mis puhastab veebilehe HTML-koodi või PDF-dokumendid visuaalsest müra, jättes järele LLM-idele kergesti töödeldava Markdown-formaadis teksti. Kolmanda sammuna edastatakse puhas tekst OpenAI GPT-4o mini API-le, mis eraldab sealt semantiliselt struktureeritud JSON-andmed koos spetsiifiliste usaldusväärsusskooridega. See meetoodika katab edukalt koguni 11 finantstoodete kategooriat enam kui 80 erineva parameetri ulatuses.

Lõputöö käigus realiseeriti kaks suurt lisaväärtust pakkuvat moodulit:

- 1) Hindade muutumise teavitussüsteem: Pärast igat kraapimisseansi võrdleb süsteem uusi väärtusi ajaloolistega ning koostab leitud erinevustest koondraporti. Tellija postkasti säästmiseks ja LLM-i mittedeterministlikust loomusest tuleneva müra vähendamiseks arendati välja dedubleerimisloogika, mis tagab, et iga muudatuse kohta saadetakse vaid üks e-kiri.
- 2) Meta reklaamide jälgimise moodul: Täiesti uude funktsionaalsusena lisandus süsteemi Meta Ad Libraryst reklaamide kogumise loogika, mis kasutab Playwrighti juhuslike kerimisliigutustega, et jäljendada inimese käitumist ja koguda konkurentide sotsiaalmeediareklaame. Kõik andmed filtreeritakse, eemaldatakse korduvused ning saadetakse OpenAI API-le, mis astub turundusanalüütiku rolli. Tehisintellekt sünteesib igale pangale kolmeosalise strateegilise kokkuvõtte: aktiivsete kampaaniate fookus,

ajalooliste kampaaniate sisu ning ennustus tulevaste turundussammude kohta, lisades sinna ka mudeli enda määratud analüütilise kindluse skoori (0.0 - 1.0).

Töö tulemusena anti tellijale üle täielikult automatiseeritud, iseseisev ja tootmiskeskonna valmiduses lahendus, mis automatiseerib varem manuaalselt tehtud turutrendide ja konkurentide jälgimise. Kõik püstitatud põhieesmärgid said realiseeritud.

Tehtud arhitektuuri ja metoodika analüüs tõestas, et puhtalt LLM-põhine andmekorje ei ole kulude ja kohatise ebatäpsuse tõttu alati optimaalne. Seetõttu osutuski parimaks kompromissiks just antud töös rakendatud hübriidlahendus, mis kombineerib reeglipõhised koodilahendused nutika semantilise infoeraldamisega. Väljatöötatud süsteemi on võimalik ka tulevikus lihtsasti skaleerida, olgu selleks uute riikide pankade lisamine, ajaridade graafiline visualiseerimine esitluskihis, andmete mitmekordne ristvalideerimine teiste mudelitega või kogutud info jagamine avaliku API kaudu.