

Tehisintellekt Sälli otsustusmudeli loomine

III etapi kokkuvõte

Koostasid:

Meri Liis Treimann
Kurmo Konsa
Kristiina Piirisild

Kuupäev:

06.07.2021

STACC OÜ

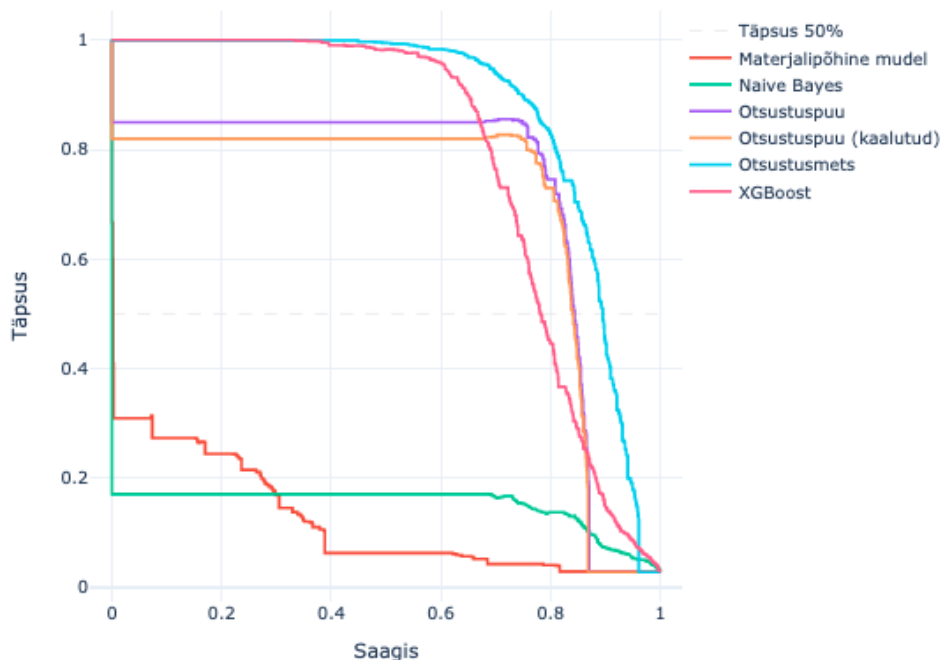
Narva mnt 20, 51009 Tartu, Estonia

II etapi käigus leidsime, et masinõppemudeliga on teoreetiliselt võimalik tuvastada vähemalt 84% halvnevatest museaalidest (saagis), kui lubame mudelil anda kuni 50% valepositiivseid (kus seisund tegelikult ei halvne). Vähendades valepositiivsete lubatud hulka 25%-ni on masinõppemudeliga võimalik tuvastada 78% halvnevatest museaalidest.

III etapi kaks peamist eesmärki olid **mudeli edasiarendus** (algoritmide ja sisendandmetega eksperimenteerimine) ning **prototüübi arendus ja selle testimine** pilootmuseumides.

Eksperimendid

Lisaks II etapis katsetatud otsustuspuule ja XGBoostile testisime veel kolme algoritmi: materjalipõhine mudel, Naive Bayes ja otsustusmets. Neist kaks esimest on lihtsad tõenäosuslikud mudelid, mis sobivad pigem kontrollmudeli rolli ja asendavad II etapis kasutatud juhuslikku mudelit. Otsustusmets on ansambelõppe versioon ühest otsustuspuust, kus luuakse korraga mitu väiksemat puud ja tehakse lõplik otsus enamushääletuse põhjal. Joonisel 1 on toodud katsetatud algoritmide saagise-täpsuse kõverad testandmetel. **Kõige paremaid tulemusi annab otsustusmets**, mis 50% täpsuse juures tuvastab testandmetel 90% halvnevatest museaalidest.

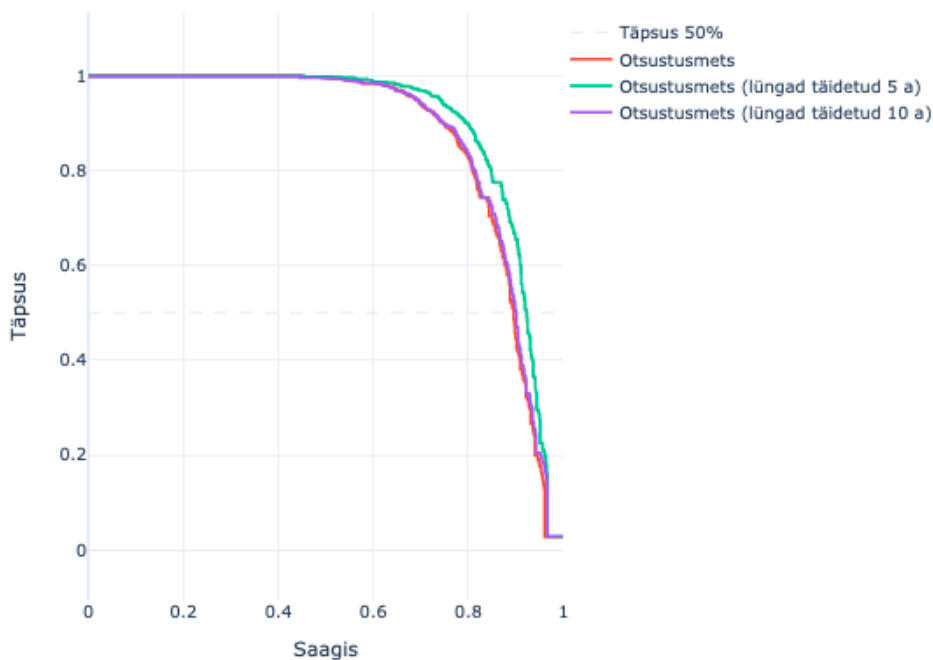


Joonis 1. Katsetatud algoritmide saagise-täpsuse kõverad, mis on saavutatud erinevaid otsustuslängendeid kasutades. Mida suurem on kõveraalune pindala ehk mida kaugemal paremal ülal nurgas kõver asub, seda paremaks üldiselt mudel loetakse (nii saagis kui täpsus on keskmiselt kõrgemad).

Katsetasime otsustusmetsa algoritmi erinevate andmete ettevalmistamise viisidega:

- Sarnaselt materjalirühmadele lisasime ka olemuste rühmad (kokku 97 rühma). Olemuste rühmade lisamine mudeli tulemusi ei muutnud.
- Lisasime kunstlikult museaalidele seisundihinnanguid juurde, nii et kahe seisundihinnangu vahele ei jääks rohkem kui n aastat. Lisatud andmepunkt sai viimase seisundihinnangu väärtuse. Nii jäid treeningandmetesse ka need andmepunktid, millele originaalis ei saanud leida silti, kas seisund

halvenes või mitte, kuna järgmine seisundihinnang oli rohkem kui 10 aastat hiljem. Kuna selliselt andmete ettevalmistamise protsess oli küllaltki ajakulukas ja mudeli treenimine võttis lisatud andmete hulga tõttu samuti kauem aega, siis testisime kahe n väärtusega: 5 ja 10 aastat (Joonis 2). **Kõige paremad tulemused saime, kui lisasime andmetesse seisundihinnanguid juurde, nii et kahe seisundihinnangu vahele ei jääks rohkem kui 5 aastat.** Nii oli 50% täpsuse juures võimalik testandmetes tuvastada 92% halvenevatest museaalidest (2% parem, kui ilma andmeid juurde lisamata).

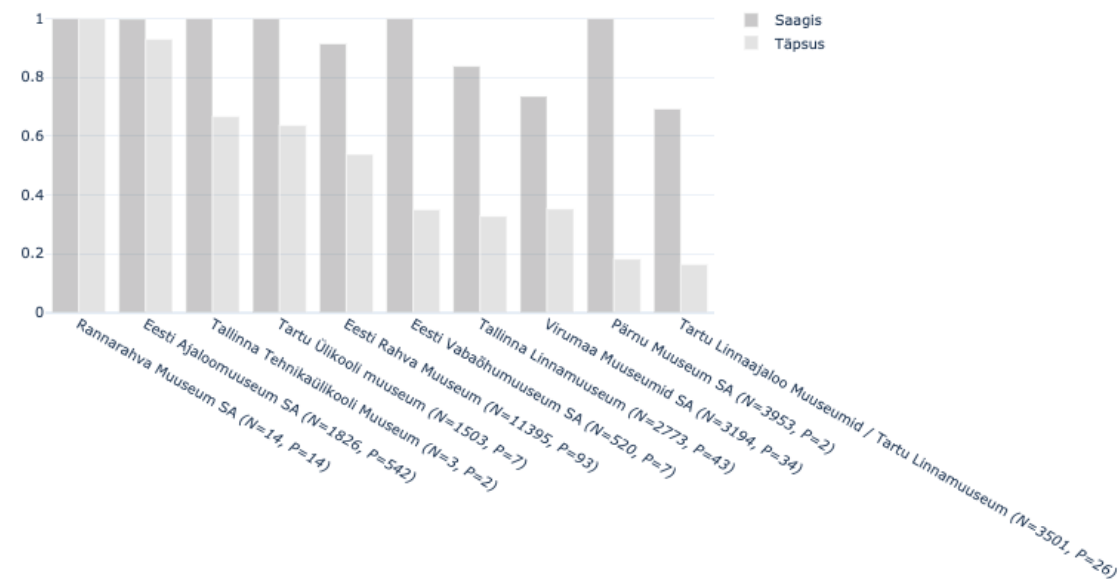


Joonis 2. Otsustusmetsa tulemused erinevatel andmete ettevalmistamise viisidel.

Eksperimentides kasutasime MuSi andmeid seisuga 27.01.2021. Et mudelist maksimumi saada, treenisime prototüübi jaoks parima mudeli (seisundihinnangute lisamine + otsustusmets) üle, kasutades MuSi andmeid seisuga 03.06.2021. **Lõpliku mudeli võimekus 50% täpsuse juures jäi 92% saagise juurde.**

Joonisel 3 on toodud viimase mudeli tulemused pilootmuuseumides testandmetel, kui jääme otsustuslävendi juurde, mille korral on mudelil üle kõigi testandmete 50% täpsus (otsustuslävendi väärtus 6%). On näha, et mudel annab suhteliselt häid tulemusi näiteks Eesti Ajaloomuuseumis, kus see tuvastas testandmetest korrektselt 541 halvenevat museaali 542st (saagis peaaegu 100%), andes seejuures vaid 41 ehk 7% valepositiivset (täpsus 93%). Mõõdikute põhjal saavutas mudel kehvad tulemused mitmes suure museaalide hulgaga muuseumis, sh Tallinna Linnamuuseum, Virumaa Muuseumid, Pärnu Muuseum ja Tartu Linnamuuseum, kus üle 60% ja isegi üle 80% positiivsetest prognoosidest olid valepositiivsed. Samas ei ole see nii suur probleem, kui halvenevate museaalide hulk on väike. Näiteks Pärnu Muuseumist oli testandmetes 2 halvenevat museaali ja mudel leidis need mõlemad üles, lisaks prognoosis mudel seisundi halvenemist veel 9 museaalile (muidugi kokku on Pärnu Muuseumis umbes 40 korda rohkem museaale ja 400 valepositiivse museaali kontrollimine võtab juba kauem aega).

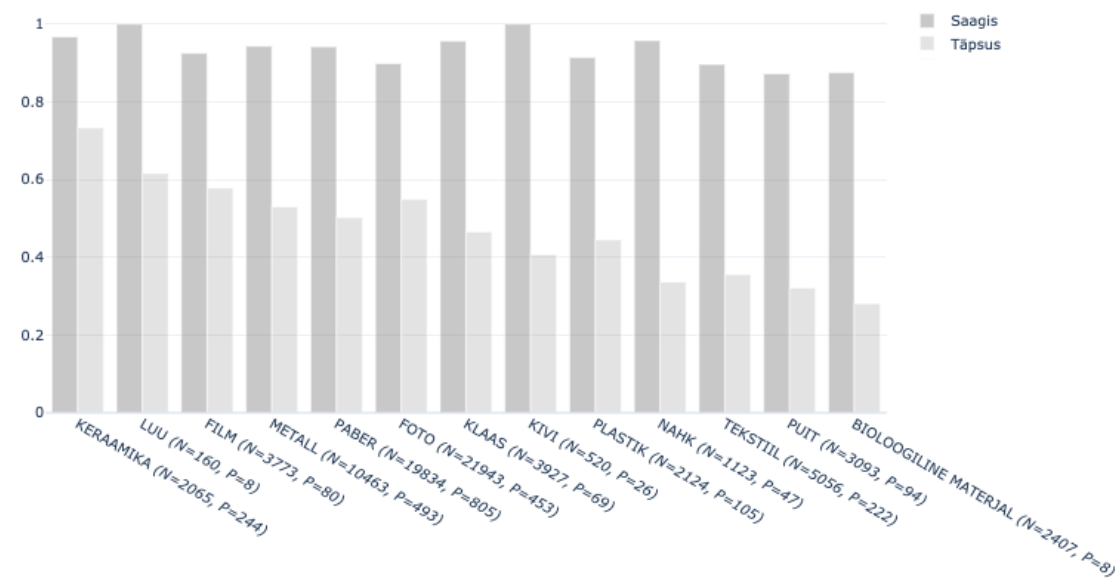
Mudeli tulemused testmuseumide lõikes



Joonis 3. Lõpliku mudeli tulemused pilootmuseumides testandmetel. N – museaalide hulk testandmetes, P – halvenevate museaalide hulk testandmetes.

Joonisel 4 on toodud viimase mudeli tulemused materjalirühmade lõikes testandmetel. On näha, et teatud materjalirühma museaalidega saab mudel paremini hakkama kui teistega.

Mudeli tulemused materjalirühmade lõikes



Joonis 4. Lõpliku mudeli tulemused materjalirühmade lõikes testandmetel. N – museaalide hulk testandmetes, P – halvenevate museaalide hulk testandmetes.

Prototüüp

Teame, et märgitud seisundihinnangud on subjektiivsed ja ei pruugi alati vastata tegelikkusele. Näiteks võib museaali seisundit esmalt hinnata koguhoidja, kes ei oska seda alati adekvaatselt teha ning märgib sel juhul seisundiks "rahuldav". Hiljem võib konservaator selle ümber hinnata näiteks "halvaks", kuid andmetest ei tule välja, kas museaali olukord nende kahe seisundihinnangu ka päriselt halvenes või oli see juba alguses "halb", kuid esmane hindaja märkis teisiti.

Eksperimentide käigus saame mudeleid valideerida vaid testandmete peal, mis on väiksem osa kõikidest kasutada olevatest andmetest. Kui nendes andmetes esineb ebamäärasusi või puudujääke, siis esinevad need ka testandmetes ja mudeli *offline* valideerimisel hindame osaliselt näiteks ka seda, kui hästi see oskab ennustada, kas konservaator annab madalama hinnangu kui koguhoidja.

Masinõppemudeli prognooside kasulikkuse ja kasutatavuse lõplikuks hindamiseks lõime veebirakenduse (<https://salli.stacc.cloud/et/dashboard>), kust muuseumitöötajad näevad oma muuseumi tuhandet kõige kõrgema riskiskooriga museaali, mille olukorra võiks eelisjärjekorras üle vaadata. Nimekirjast jätsime välja 2021. aasta jooksul seisundihinnangu saanud museaalid. Museaalide kontrollimisel kogume kasutajatelt tagasisidet, küsides nende hinnangut museaali halvenemise kohta. Projekti viimases etapis võrdleme mudeli antud riskiskoore muuseumitöötajate tagasisidega, et anda prognoosidele lõplik hinnang.

Pilootmuuseumidest ja Muinsuskaitseametist liitunud testkasutajatele andsime ligipääsu rakendusele 18.06.2021. Testperiood kestab kuni augusti keskpaigani, mil fikseerime saadud tagasiside ja teeme lõplikud järeldused.

Järgmised sammud

IV etapi käigus:

- Teeme projektist viimase kokkuvõtte.
- Toimub prototüübi ja andmete töötlemise lähtekoodi üleandmine RIKile.