

Virtuaalse abilise Bürokrati treenimise juhend

Riigi Infosüsteemi Amet, 2023

Sisukord

Mõisted	2
1. Bürokrati treenimine masinõppel	3
2. Treeningliides ja selle komponendid	4
2.1 Teemad	4
2.2 Ajalugu	6
2.3 Treenimine	7

Mõisted

Bürokratt on digiriigi virtuaalne assistent. Bürokrati üks komponentidest on juturobot, mis põhineb masinõppelistel keeletehnoloogilistel mudelitel.

Keeletehnoloogia (KT) (*language technology, natural language processing* ehk NLP) on arvutilingvistika praktiline pool, kasutab arvutilingvistikas väljatöötatud teooriaid rakenduste loomiseks ning ühtlasi loob vahendeid arvutilingvistika arenguks.^[1]

Masinõppimise (ML, *machine learning*) puhul on arvuti ise võimeline ise looma algoritmi, mis suudaks piisava täpsusega näidisandmetest eraldada meid huvitava info. See on osalt tehisintellekt (AI, *artificial intelligence*), mis suudab kohaneda muutuva keskkonnaga ja omab seega õppimisvõimet.^[2]

Juturobot on arvutiprogramm või tarkvara, mis simuleerib inimsuhtlust läbi teksti- või kõnesuhtluse. Juturoboteid on loodud mitmetel erinevatel meetoditel. Tänaused juturobotid põhinevad keeletehnoloogilistel (NLP) masinõppe algoritmidel ning võimaldavad tuvastada kliendi pöördumiste teemasid. Paljud juturobotid kasutavad ka kõnetehnoloogilisi lahendusi (näiteks Alexa ja Siri), mis võimaldavad imiteerida kliendisuhtlust ka kõne kaudu.^[3]

Juturoboti treenimine on oluline, sest masinõppelised mudelid nõuavad suurt andmehulka, et tuvastada kasutajate pöördumiste teemasid võimalikult enesekindlalt ning suhelda efektiivselt. Juturobotid aitavad tõsta nii klienditeeninduse kvaliteeti kui ka kiirendada teenindajate tööd vastates lihtsamatele ning rutiinsematele küsimustele ja täites rutiinseid äriprotsesse.^[4]

Bürokratt põhineb **RASA** vabavaralisel tarkvaral, mis pakub masinõppel põhinevat tehnoloogilist lahendust juturobotite loomiseks, haldamiseks ja treenimiseks.^[5]

[1] https://courses.cs.ut.ee/MTAT.03.047/2013_fall/uploads/Main/Keeletehnoloogia.pdf

[2] <https://masinope.ee/masinoppimine/>

[3] <https://www.techtarget.com/searchcustomerexperience/definition/chatbot>

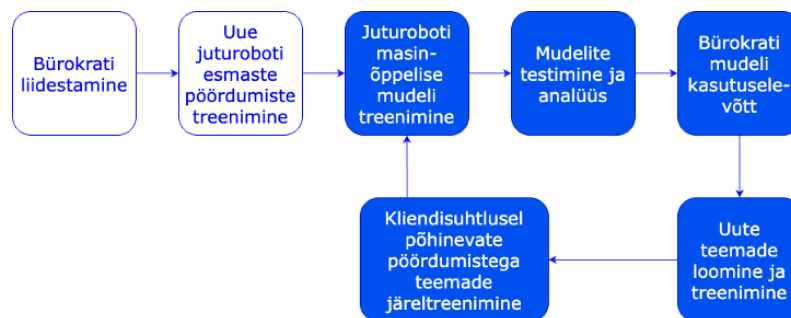
[4] <https://www.techtarget.com/searchcustomerexperience/definition/chatbot>

[5] RASA link

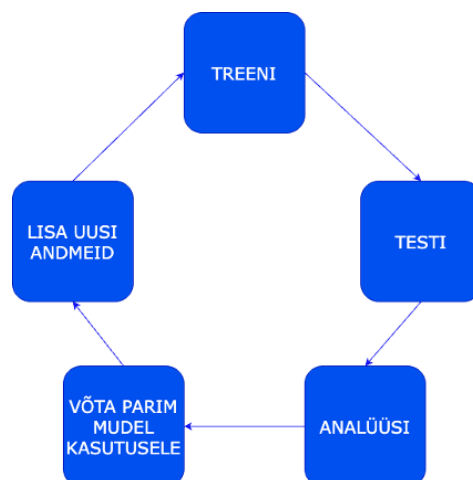
1. Bürokrati treenimine masinõppel

Bürokrati esmasel treenimisel treenitakse kõige korduvamad teemad olemasolevate asutuse ja klienditeenindaja vaheliste vestluste põhjal. Vestlusteks kasutame nii e-maile, juturoboti vestlusi kui ka muude kasutusele võetud sõnumivahetuse lahenduste vestlusi (näiteks Facebook Messenger). Oluline on treenida mudeleid päris andmetel ehk imiteerida kasutajate pöördumisi sellisel kujul nagu neid edastatakse.

Alloleval joonisel on kirjeldatud üldine ülevaade sellest, kuidas toimub uute mudelite ja andmete treenimine.



Bürokrati masinõppeliste mudelite treenimisel järgime printsiipe, et esmalt täiendatakse kasutusel olevat mudelit uute andmetega (kasutajate pöördumised ja nendel põhinevad tegevused). Kui juturoboti treenija on lisanud piisavas mahus uusi andmeid, siis mudel treenitakse ja testitakse. Mudelite treenimise ja testimise järel vaadatakse sisse mudelite testimise tulemusena loodud statistilistele mõõdikutele, mis viitavad mudeli ja teemade tuvastamise kvaliteeti. Kui kvaliteedimõõdikud on oodatud tasemel (kirjeldatud peatükis „Juturoboti mudelite treenimine ja analüüsimine“) võetakse kasutusele viimane treenitud mudel. Juhul kui mudelite mõõdikud on olulisel määral langenud, proovitakse leida mudeli kvaliteedi langemise põhjused, need eemaldada ning treenida uus mudel. Alljärgnev joonis kirjeldab üldist mudelite treenimise põhimõtet Bürokratis.



2. Treeningliides ja selle komponendid

Treeningmoodulit saab kasutada, et õpetada vestlusrobotit täpsemalt ning rohkematele pöördumistele vastama.

Vestlusroboti treenimise loogika on järgnev:

Kui klient kirjutab pöördumise, siis vestlusrobot klassifitseerib pöördumise üheks konkreetseks teemaks. Teema on kliendi pöördumise eesmärk. Näiteks klient kirjutab pöördumise: "Soovin väga teada, mis ilm on, sest mul on täna plaan pikalt väljas viibida". Selle pöördumise teema oleks "ilmateade". Kui klient aga kirjutab "mis ilm on", siis selle pöördumise teema oleks samuti "ilmateade". Kuna iga kliendi sõnavara, keeleoskus ja kirjutamise viis võib olla erinev, siis tuleb vestlusrobotit õpetada, et ta suudaks tuvastada kõikide erinevalt sõnastatud pöördumiste puhul üldist pöördumise eesmärki ehk teemat.

Selleks, et vestlusrobot oskaks teemat tuvastada, peab teema kohta lisama näiteid, mis käivad selle teema kohta. Võttes näidiseks eelnevalt toodud teema "ilmateade", võime õpetada vestlusrobotit järgmiste näidetega:

- mis ilm on
- kas väljas on soe
- kas täna hakkab sadama
- mis temperatuur on

Nüüd kui klient pöördub vestlusroboti poole, siis vestlusrobot analüüsib, milline temale õpetatud näide sobib kõige paremini kokku kliendi pöördumisega. Kui ta on otsuse teinud, siis ta leiab, mis teema alla see konkreetne näide kuulus. Seejärel suudab vestlusrobot anda selle konkreetse teema kohta õpetatud vastuse.

2.1 Teemad

„Teemad” saki alt saab ülevaate kõigist olemasolevatest teemadest ning nende konkreetsete teemade näidetest. Vasakpoolses nimekirjas on kõik **teemad** (joonis 1), mida on eraldi võimalik avada, et näha konkreetse teema kohta käivaid **näiteid** (joonis 1). Neid näiteid saab **muuta, kustutada** (joonis 1) või soovi korral teemale näiteid juurde **lisada** (joonis 1). Võimalik on ka vastavalt **redigeerida teema vastust** (joonis 1).

Lisa teema (joonis 1) nupule vajutades saab luua täiesti uue teema. Sellisel juhul on vajalik lisada ka vastavale teemale vastus ning vähemalt üks näide, kuid alati võib lisada rohkem näiteid (joonis 2). Teema nimetamisel on rangelt soovituslik kasutada struktuuri, kus tühiku asemel kirjutatakse alakriips. Näiteks soovi korral lisada teema "info asutusest" peaks nimetama teema "info_asutusest".

The screenshot shows the 'Teemad' management interface. On the left, there is a list of topics with their names and 'ava teema' buttons. The main area displays a table of topics with columns for the topic name and actions like 'kustuta' and 'muuda'. The 'lisa näide' button in the sidebar is highlighted with a blue box.

joonis 1 Teemad

loo uus teema



teema nimetus

pöördumise vastus

lisa näide/näited

[lisa näide](#)

[tühista](#)

[salvesta](#)

joonis 2 Uue teema loomine

Oluline on teada, et uue teema loomisel ei lisata seda automaatselt mudelisse. See tähendab, et kuigi klienditeenindaja jaoks on teema nimekirjas, siis vestlusrobot ei õpi sellest ning kliendile selle teema puhul ei vastata. Seda, kas teema on mudelis, on näha teema nimetusega samal real (joonis 1). Et vestlusrobot õpiks lisatud teemal vastama, on pärast teema loomist tarvis vajutada nupule **lisa teema mudelisse** (joonis 3). Sarnaselt on võimalik olemasolev teema vestlusrobotist eemaldada, kasutades nuppu **eemalda teema mudelist** (joonis 3).

teema tegevused

[lisa näide](#)

[kustuta teema](#)

[lisa teema mudelisse](#)

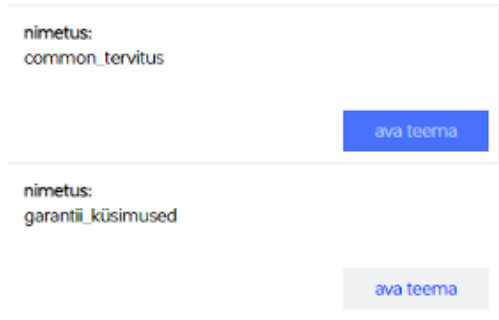
teema tegevused

[lisa näide](#)

[kustuta teema](#)

[eemalda teema mudelist](#)

Osa teemadest on piiratud õigustega. See tähendab, et need on loodud erireeglitega ning neid ei saa kustutada ega mudelist eemaldada. **Piiratud õigustega teemad** tunneb nimekirjast ära värvilise märgistuse järgi (joonis 4). Juhul kui teema on piiramatute õigustega, siis saab selle täielikult kustutada nii kliendi kui klienditeenindaja jaoks, kasutades nuppu **kustuta teema** (joonis 5).



joonis 4 Piiratud teema



joonis 5 Teema kustutamine

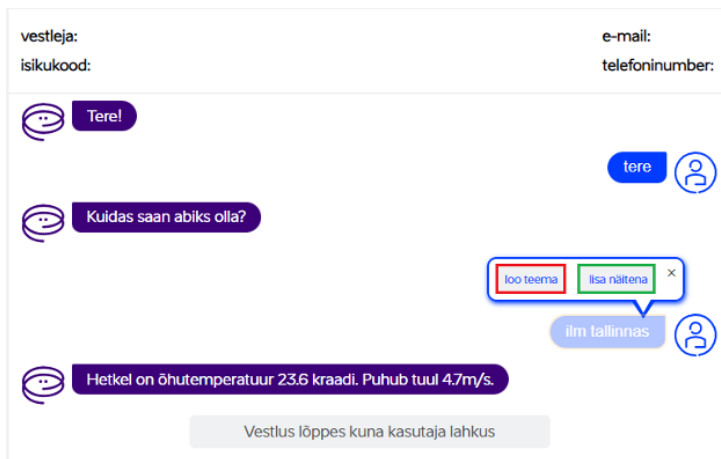
2.2 Ajalugu

Sakk „ajalugu” on visuaalselt sarnane ajaloolse vestlusmoodulis. Treeningmoodulis on lisafunktsionaalsus, mis võimaldab kliendi pöördumisi lisada näitena klienditeenindaja poolt valitud teemale või luua kliendi pöördumisest täiesti uus teema.

alguseeg	lõppaeg	nõustaja nimi	nimi	isikukood	kontaktandmed	suunatud	vestluse tulemus	ID
13. juuli 2022 12:58:11	13. juuli 2022 12:58:21	Bürokratt			ei	ei	lõpetatud	0a2b29125-3be9-405f-b4f4-d14a29eab333
13. juuli 2022 12:56:52	13. juuli 2022 12:57:07	Rasmus			ei	jah	lõpetatud	6a6c24eb-a501-453c-b168-1d4a46731070
13. juuli 2022 12:56:32	13. juuli 2022 12:56:18	Rasmus			ei	jah	lõpetatud	b386a613c-205c-4a5e-b8fa-ee813c071402
13. juuli 2022 12:55:20	13. juuli 2022 12:55:31	Bürokratt			ei	ei	lõpetatud	04a80470-c896-45fe-8669-ad650718ca4

joonis 6 Treeningmooduli ajalugu

Selle jaoks tuleb kõigepealt avada vestlus (joonis 6). Seejärel kui vajutada vastavale kliendi sõnumile, on võimalik valida **loo teema** (joonis 7). Sellisel juhul luuakse pöördumisest täiesti uus teema. Enne, kui teema luuakse, peab klienditeenindaja valima teemale nime ning vastuse (joonis 8).



joonis 7 Ajaloo kaudu uue teema loomine

Kui valida **lisa näitena** (joonis 7), siis lisatakse pöördumine juurde olemasolevale teemale näidete nimekirja. Sellisel juhul peab klienditeenindaja valima, millise teema alla pöördumine lisatakse (joonis 9). Enne salvestamist on võimalik pöördumist ka redigeerida.



joonis 8 Ajaloo kaudu uue teema loomine 2



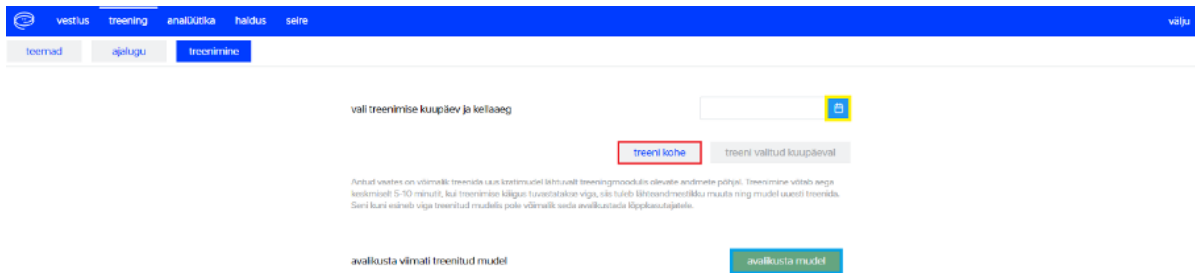
joonis 9 Ajaloo kaudu teema lisamine

2.3 Treenimine

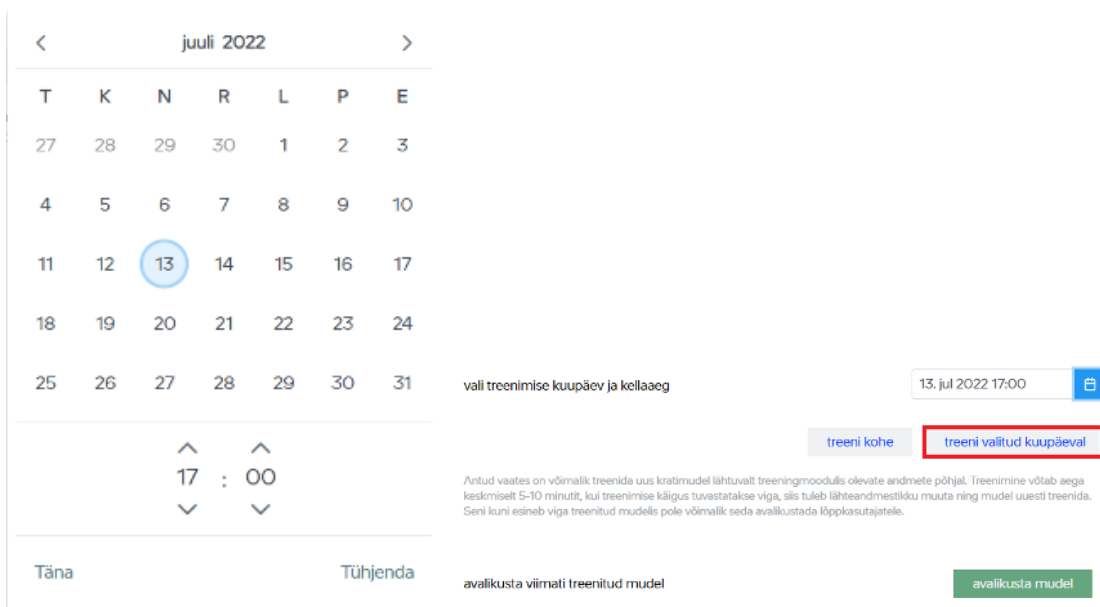
„Treenimine” saki alt on võimalik käia uusi mudeleid treenimas. See tähendab, et kui on olemas vestlusrobot, mis vastab klientidele, kuid vahepeal on klienditeenindajad lisanud uusi

teemasid, kustutanud üleliigseid teemasid või muutnud näiteid ja on soov uuendada vestlusrobotit, siis „treenimine“ saki alt on seda võimalik teha.

Treeni kohe (joonis 10) nuppu vajutades algab uue vestlusroboti treenimine. Uus vestlusrobot treenitakse kõikide teemade ja vastustega, mis on lisatud mudelisse (joonis 3). Mudeli treenimine võtab aega sõltuvalt sellest kui palju on teemasid ja näiteid. Kuni mudel treenib, ei saa „treenimine“ sakiga toimetada. Klienditeenindajal on võimalik seadistada treenimiseks **sobiv kuupäev ja kellaaeg** (joonis 10), valides endale sobiv aeg (joonis 11) ning vajutades nupule **treeni valitud kuupäeval** (joonis 12).



joonis 10 12 Mudeli treenimine



Kui mudeli treenimine on lõpetatud ning mudeli treenimisel vigu ei esinenud, siis on võimalik vahetada praegune vestlusrobot uue treenitud vestlusroboti vastu, kasutades nuppu **avalikusta mudel** (joonis 13). Juhul kui mudeli **treenimisel esines vigu** (joonis 14), siis ei luba keskkond uut mudelit avalikustada ().

avalikusta viimati treenitud mudel

avalikusta mudel

joonis 13 Mudelis ei esinenud vigu

avalikusta viimati treenitud mudel

treenitud mudelis esineb vigu

joonis 14 Mudelis esines vigu