

Tartu Ülikool
Arvutiteaduse instituut
Informaatika õppekava

Kristjan Radsin

Masinõppe meetodid pragude tuvastamiseks

Euler-Bernoulli talades

Bakalaureusetöö (9 EAP)

Juhendaja:

Ljubov Jaanuska, PhD

Tartu 2025

Masinõppe meetodid pragude tuvastamiseks Euler-Bernoulli talades

Lühikokkuvõte:

Bakalaureusetöö eesmärk on hinnata erinevate masinõppe meetodite täpsust pragude tuvastamisel Euler-Bernoulli talades, mis on paigaldatud Pasternaki alusele. Selleks kasutatakse Pythoni teekides olevaid meetodeid. Uuritavateks masinõppemeetoditeks on lineaarregressioon, kernel ridge, Gaussi protsessi regressioon, K-lähima naabri meetod (KNN), otsustusmets, gradient boosting, ja tehisnärvivõrk. Mudeleid treeniti kolmel andmehulgal: tala omavõnkumistel põhinev, võnkemoodi kujust teisendatud Haari lainikute kordajatel põhinev, ja Fourier' teisendatud võnkumistel põhinev andmehulk. Lisaks katsetati mudeleid 5%-list valget müra sisaldaval andmestikul. Tulemused näitavad, et pragude asukoha ennustamisel saavutati parimad tulemused mudelitega, mis olid treenitud Haari lainikute kordajatel, ning parimaks meetodiks osutus Gaussi protsessi regressioon. Sügavuse ennustamisel andis parimaid tulemusi Fourier' teisendusel põhinevad võnkumised ning edukaimaks meetodiks oli otsustusmets.

Võtmesõnad: Masinõpe, Euler-Bernoulli tala Pasternaki alusel, Haari lainikud, Fourier' teisendus, Python

CERCS: P190 Matemaatiline ja üldine teoreetiline füüsika, klassikaline mehaanika, kvantmehaanika, relatiivsus, gravitatsioon, statistiline füüsika, termodünaamika, T121 Signaalitöötlus, P176 Tehisintellekt

Machine Learning Methods for Crack Identification in Euler-Bernoulli Beams

Abstract:

The objective of this bachelor's thesis is to evaluate the suitability of various machine learning methods for crack identification in Euler-Bernoulli beams placed on a Pasternak foundation. Methods available in Python libraries are used for this purpose. The machine learning methods of interest are linear regression, kernel ridge, Gaussian process regression, K-nearest neighbours, random forest, gradient boosting, and artificial neural network. The models are trained on three datasets: the one is based on the beam's natural frequencies, the

second one is based on the Haar wavelet coefficients obtained from the mode shape, and the third one is based on the Fourier transformed frequencies. In addition, the models are tested on a dataset containing white noise in the 5% range. The results show that the best results for predicting crack location can be achieved with models trained on the Haar wavelet coefficients, and Gaussian process regression turns out to be the best method for this case. Fourier transformed frequencies give the best results in predicting crack depth, and random forests are the most successful method in this case.

Keywords: Machine learning, Euler-Bernoulli beam on Pasternak foundation, Haar wavelets, Fourier transform, Python

CERCS: P190 Mathematical and general theoretical physics, classical mechanics, quantum mechanics, relativity, gravitation, statistical physics, thermodynamics, T121 Signal processing, P176 Artificial intelligence

Sisukord

Sissejuhatus.....	5
1. Teoreetiline taust.....	7
2. Masinõppe meetodid.....	10
2.1 Lineaarregressioon.....	10
2.2 Kernel ridge.....	10
2.3 Gaussi protsessi regressioon.....	10
2.4 K-lähima naabri meetod.....	11
2.5 Otsustusmets.....	11
2.6 Gradientvõimendamise masinad.....	12
2.7 Tehisnärvivõrk.....	12
3. Andmed ja metoodika.....	14
3.1 Fourier' teisendus.....	14
3.2 Andmed.....	14
3.3 Mudeldamine.....	16
4. Tulemused ja arutelu.....	18
4.1 Asukoha ennustamine.....	18
4.2 Sügavuse ennustamine.....	21
4.3 Müraga andmed.....	24
4.4 Arutelu.....	25
Kokkuvõte.....	28
Viidatud kirjandus.....	29
Litsents.....	31

Sissejuhatus

Tala on üks levinumaid konstruktsioonielemente, mida kasutatakse mitmetes valdkondades. Mõned näited taladest on ukseava või aia tala ning rongirööpad. Üheks võimalikuks defektiks talades on praod, mis vähendavad tala kandevõimet ning ohustavad konstruktsiooni terviklikkust. Selliste defektide tuvastamine konstruktsioonides on oluline edasiste kahjustuste vähendamiseks ja ohutuse tagamiseks.

Üheks viisiks konstruktsioonielementides kahjustusi tuvastada on eksperimentaalsete meetodite kasutamine, nagu magnet- ning soojusväljal põhinevad meetodid, pöörivoolu testimine, ultraheliga testimine, radioloogia meetodid jt. Need võimaldavad küll muutusi fikseerida, kuid on sageli kallid ja ressursimahukad.

Viimastel aastatel on teadlaste tähelepanu pälvinud modaalanalüüs (ingl *modal analysis*). Antud lähenemisviis põhineb eeldusel, et konstruktsiooni kahjustused muudavad selle füüsikalisi omadusi, mis omakorda mõjutab modaalseid omadusi, näiteks omavõnkumisi (ingl *natural frequency*) ja võnkemoodi kuju (ingl *mode shape*) [1].

Modaalanalüüsis kasutatakse lisaks klassikalistele meetoditele üha enam ka arvutuslikke ja statistilisi meetodeid, sealhulgas masinõppe tehnikaid, mis võimaldavad töödelda suuremahulisi mõõtmisandmeid ning ennustada modaalparameetreid keerukates või mürastes süsteemides. Samas puudub hõlmav ülevaade erinevatest meetoditest ning tunnuste valikust, mis teeb analüüsi keeruliseks [1].

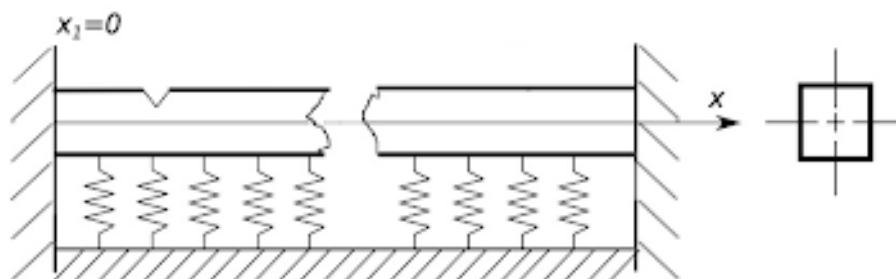
Käesoleva bakalaureusetöö eesmärk on uurida erinevate masinõppe meetodite sobivust pragude identifitseerimiseks Euler-Bernoulli talades kasutades selleks programmeerimiskeele Python teeki. Töö raames käsitletakse tala, mis asetseb Pasternaki alusel, kusjuures üks tala ots on jäigalt kinnitatud ja teine ots on rullkinnitusega. Antud tala kohta on olemas platvormil Kaggle avalik andmestik, mis sisaldab erinevate prao asukohtade ja sügavuste korral arvutatud omavõnkesagedusi ning võnkemoodi kuju iseloomustavaid Haari lainikute kordajaid.

Töö uudsus seisneb Fourier' teisenduse rakendamises andmete töötlemisel ning potentsiaalse 5% valge müra lisamises, et simuleerida mõõtmisandmetes esinevat ebatäpsust ja hinnata masinõppemeetodite ennustamistäpsust müra olemasolul.

Lõputöö on jaotatud neljaks peatükiks. Esimene peatükk annab ülevaate töö teoreetilisest taustast, kus selgitatakse täpsemalt probleemi olemust ning tutvustatakse varasemaid uurimusi. Teine peatükk tutvustab töös kasutatud masinõppe meetodeid. Kolmas peatükk kirjeldab töö protsessi, andes ülevaate algandmetest ja nende töötlemisest ning loodud mudelitest. Neljandas ehk viimases peatükis antakse ülevaade uuritud meetodite tulemustest.

1. Teoreetiline taust

Euler-Bernoulli tala (ingl *Euler-Bernoulli beam*) on üks levinumaid ja enim kasutatud konstruktsioonelemente (Joonis 1). Seda iseloomustavad eeldused, et tala käitub lineaarse elastse kehana, talal on suur pikkuse ja paksuse suhe (tüüpiliselt vähemalt 20), läbipainded on väikesed ning ristlõiked jäävad painutamisel tasaseks ja risti neutraalteljega [2]. Sellised eeldused võimaldavad lihtsustada mehaanilist analüüsi, keskendudes peamiselt paindemõjudele ning jättes nihkemõjud tähelepanuta.



Joonis 1. Praoga Euler-Bernoulli tala Pasternaki alusel.

Üheks võimalikuks defektiks Euler-Bernoulli talades on praod, mis negatiivselt mõjutavad tala kandevõimet ning võivad panna kogu konstruktsiooni terviklikkuse ohtu [1]. Seega on pragude tuvastamine tähtis, et tagada konstruktsiooni ohutus ning ennetada kahjustuste süvenemist.

Pragude tuvastamiseks on olemas mitmeid meetodeid. Eksisteerivad lokaalsed eksperimentaalsed meetodid, nagu näiteks pöörisvoolu testimine, ultraheliga testimine, radioloogia meetodid ja magnet- väljal ning soojusväljal põhinevad meetodid, kuid need on sageli ressursimahukad ning keerulistema ja raskesti ligipääsetavate struktuuride jaoks ei pruugi nende meetodite rakendamine olla mõistlik ega võimalik [1, 3].

Alternatiivseid pragude tuvastamise meetodeid võib leida modaalanalüüsist, kus pragude tuvastamiseks kasutatakse tala modaalseid omadusi, näiteks omavõnkesagedusi, võnkemoodi kujusid ja summutamist (ingl *damping*), mis on mõjutatud muutustest tala füüsikalistes omadustes [1, 4]. Näiteks Tewelde jt [3] on uurinud talade reaktsiooni võnkumistele ning saadud tulemuste põhjal tuvastanud pragusid talades. Loodi mudeleid nii tervetest taladest kui ka kahjustatud taladest, milles oli suletud pragu (ingl *closed crack*). Leiti, et tala reaktsioon

vibratsioonile on väga tundlik ja võib tuvastada vaevu nähtava prao olemasolu, mille kindlaks tegemiseks kasutati faasidiagramme (ingl *phase diagram*).

Modaalanalüüsis tihtipeale kasutatakse ka masinõpet, et tuvastada ja identifitseerida pragusid talades. Näiteks Habib jt [5] käsitlesid oma artiklis akustilise emissioonide jälgimist pragude tuvastamiseks talades, rakendades selleks masinõppemeetodeid. Vaadeldavaks objektiks oli raudbetoonist talad. Koguti andmeid akustiliste emissioonide kohta, kasutades taladel olevaid andureid, ja need emissioonid teisendati Morse lainikuteks. Nende põhjal loodi piltidel põhinev andmebaas. Andmetest pragude ennustamiseks kasutati sügavat konvolutsioonilist tehiskärvivõrku (ingl *convolutional neural network*), täpsemalt kasutati SqueezeNet mudelit¹, mida modifitseeriti vastavalt vajadusele. Mudel klassifitseeris piltide põhja talasid terveteks, mikropragusid sisaldavateks ja makropragusid sisaldavateks. Kasutatud masinõppe mudel osutus edukaks pragude tuvastamisel raudbetoonitalades, olles keskmiselt 100% klassifikatsiooni täpsusega, mis oli varasematest implementatsioonidest 20,8% suurema täpsusega.

Hein ja Jaanuska [1] on oma artiklis käsitlenud pragude identifitseerimist Euler-Bernoulli talades kasutades tehiskärvivõrke ühe vahekihiga (ingl *hidden layer*) ja otsustusmetsi. Mudelid kasutasid ennustamiseks talade võnkumise ja talade võnkemoodi kuju, millest viimane oli Haari lainikute abil muudetud kordajateks. Mudelid ennustasid pragude asukohta ja sügavust, iga ennustatava tunnuse jaoks oli kasutusel oma kärvivõrk. Uuringu tulemusena selgus, et juhumetsad suutsid ülesandega veidi paremini toime tulla ning said ka sellega kiiremini hakkama, mida põhjendati juhumetsade vähema arvu hüperparameetritega. Mõlemad mudelid kokkuvõttes andsid häid tulemusi, kusjuures pragude asukoha määramisel oli keskmine ruutviga (ingl *mean square error, MSE*) väike. Autorite sõnul osutusid talade võnkumised heaks valikuks pragude sügavuse ennustamisel, aga asukoha ennustamisel andis paremaid tulemusi Haari lainikute kordajate kasutamine.

Gidiagba jt [4] kasutas oma uuringus masinõppe meetodeid pragude tuvastamiseks konsoolis (tala, mille üks ots on jäigalt kinnitatud ja teine ots on vaba). Modelleeriti nii pragudega kui ka pragudeta talasid, mõlemalt tala tüübilt mõõdeti võnkumisi, võnkemoodi kuju ja kiirendusreaktsioone (ingl *acceleration response*). Andmeid analüüsid leiti, et võnkumissagedustes ja võnkemoodi kujus esines väheseid muutusi terve ja vigase tala vahel, mistõttu lisati nendele veel ka kiirendusreaktsioonide mõõdud. Pragude identifitseerimiseks kasutati pärilevivõrku (ingl *feedforward neural network*), mis kasutas

¹ Lisainfo mudeli kohta: <https://arxiv.org/abs/1602.07360>

kihtides sigmoid aktivatsioonifunktsiooni (ingl *activation function*) ning sisaldas üht vahekihti. Gidiagba jt arvates osutus tehismärgivõrk sobilikuks pragude tuvastamiseks ja analüüsimiseks modelleeritud talades.

Kuigi modaalanalüüsi meetodid ja modaalomadused sobivad hästi olukordadesse, kus traditsioonilisi lähenemisviise ei ole võimalik rakendada, kaasnevad nendega ka teatud piirangud. Statistiliste ja masinõppe meetodite rakendamine eeldab kriitilist ülevaadet, piisavat andmehulka ning asjakohaste tunnuste eristamist ja ettevalmistamist (ingl *feature engineering*) [1]. Sellest lähtuvalt antud bakalaureusetöö eesmärk on hinnata erinevate masinõppemeetodite ennustustäpsust pragude tuvastamisel Euler–Bernoulli talades. Oluliseks kriteeriumiks on lisaks üldisele täpsusele ka mudelite ennustusvõimekus müra sisaldavate andmete korral, et hinnata mudelite sobivust praktilisteks rakendusteks.

2. Masinõppe meetodid

Selles peatükis antakse ülevaade töös kasutatud erinevatest masinõppe meetoditest. Kirjeldatakse valitud meetodeid, kas ja kuidas neid on kasutatud ning miks osutusid need valituks.

2.1 Lineaarregressioon

Lineaarregressioon on üks lihtsamaid masinõppes kasutatavaid meetodeid. See otsib lineaarfunktsiooni andmetevahelise seose kirjeldamiseks, leides võrrandi kaalud treeningandmete põhjal. Enamasti kasutatakse funktsiooni leidmiseks vähimruutude meetodit, kus minimeeritakse iga andmepunkti ja sirge väärtuste vahede ruutude summat. Lineaarregressioon eeldab lineaarset seost andmete ja ennustatava väärtuse vahel, mistõttu on sellel raskusi mittelineaarsete seoste leidmisega [6, 7]. Sellegipoolest on meetod valitud pragude mittelineaarselt sõltuvate omaduste tuvastamiseks, peamiselt võrdluse eesmärgil teiste meetoditega.

2.2 Kernel ridge

Lineaarregressiooni puudujääkide adresseerimiseks on olemas *lasso* ja *ridge* meetodid [7]. Keskendudes viimasele nimetatud meetodile, *ridge* on lineaarregressiooni meetod, mis rakendab $L2$ regulariseerimist, et vähendada ülesobitamisest ja multikollineaarsusest tekkinud vigu. Sellest veel edasiarendus on *kernel ridge*. Tavalisest *ridge* meetodist erinevalt kasutab *kernel ridge* kernelfunktsiooni. Kerneli meetod kaardistab andmed ühest ruumist teise, mis võimaldab meetodil sobitada sirget ruumis, kus andmed sobivad lineaarfunktsiooni jaoks paremini. See vastab mittelineaarse funktsiooni õppimisele algses ruumis, seega kerneli lisamine võimaldab *ridge* meetodil lahendada suurema hulga ülesandeid [8, 9]. Seda versiooni meetodist pole varem rakendatud pragude identifitseerimiseks, mistõttu on see valitud uurimiseks.

2.3 Gaussi protsessi regressioon

Veel üheks valitud meetodiks on Gaussi protsessi regressioon (ingl *Gaussian process regression*). Rasmussen ja Williams [10] on kirjutanud, et Gaussi protsess on üldistus Gaussi tõenäosusjaotusest, täpsemalt on tegu juhusliku protsessiga, kus iga lõpliku hulga muutujate

vahel on mitmemõõtmeline normaaljaotus. Gaussi protsessi jaotus on pideva määramispiirkonnaga funktsioonide jaotus. Meetod kasutab neid protsesse, et leida andmete vahelist seost kõige paremini kirjeldav funktsioon. Selleks defineerib see eelneva funktsiooni, mis teisendatakse uueks funktsiooniks peale andmepunktide lisamiseks, mis saadakse treeningandmetest. Meetod on vähemlevinud kui teised siin töös käsitletavad, kuid seda on siiski kasutatud mitmete probleemide lahendamiseks, näiteks on Corrado jt seda kasutanud [11], et määrata pragude asukohta talalaadsetes struktuurides. Meetod on näidanud häid tulemusi oma kasutusalades, mistõttu osutus see valituks.

2.4 K-lähima naabri meetod

Artikli [12] põhjal on K-lähima naabri meetod (ingl *K-nearest neighbours*) ehk KNN lihtne juhendatud õppega mitteparameetiline meetod. Enamasti kasutatakse seda klassifikatsiooni ülesannetes, kuid seda on edukalt rakendatud ka regressiooni jaoks. Meetod arvutab antud sisendile vastava väärtuse selle K lähima naabrite väärtuste põhjal. Regressiooni puhul on ennustatud väärtus K lähima naabri keskmine väärtus. Tegemist on väga levinud meetodiga, mida on kasutatud ka pragude ennustamiseks, näiteks Omar jt [13] on seda on rakendanud pragude leviku ennustamiseks polümeer- ja metallmaterjalides. KNN meetod on näidanud häid tulemusi paljudes kasutusalades, mistõttu osutus see valituks.

2.5 Otsustusmets

Üheks valitud meetodiks on otsustusmets (ingl *random forest*). Allika [14] põhjal on tegemist ansambliõppe meetodiga. Otsustusmets koosneb mitmest otsustuspuust, kus igale puule antakse ette osa andmetest, mille põhjal õppida, ning lõpptulemus saadakse puude tulemusi koondades. Regressiooni ülesande puhul on lõppennustus otsustuspuude keskmine väärtus. Meetod on paindlik ja kaldub vähem ülesobitama, kuid võib olla ressursinõudlik, eriti suuremate andmehulkade korral. Seda on edukalt rakendatud mitmete probleemide lahendamiseks, kaasa arvatud ka pragude ennustamiseks, nagu Omar jt [13] ning Hein ja Jaanuska [1]. Kuna see on varasemalt näidanud häid tulemusi pragude uurimisel, osutus see valituks.

2.6 Gradientvõimendamise masinad

Gradientvõimendamise masinad (ingl *gradient boosting machine*) on üks valitud meetoditest. Artikli [15] põhjal on tegemist ansambelõppe meetodiga, mille lõplik mudel saadakse mitme nõrga õppealgoritmi järjestikusel kombineerimisel. Enamasti kasutatakse selleks madala sügavusega otsustuspuud. Meetod töötab iteratiivselt. See üritab järjest parandada oma ennustuste tulemust, püüdes vähendada etteantud kaofunktsiooni (ingl *loss function*). Regressiooni ülesannete puhul on tavaliselt kaofunktsiooniks keskmine ruutviga. Iga sammuga arvutatakse kaofunktsiooni negatiivne gradient, mis näitab, kuhu suunas tuleb ennustust parandada. Sellele gradiendile sobitatakse järgmine nõrk otsustuspuu, mille ennustust kasutatakse olemasoleva mudeli täiendamiseks. Sellist täiendamist korratakse määratud arv kordi ning iga uue mudeli mõju lõppennustusele määrab õppekiirus (ingl *learning rate*).

Gradientvõimendamise masinad on tuntud oma kõrge täpsuse ja suutlikkuse poolest käsitleda keerukaid andmeid, kuid selle peamised puudused on suur arvutuslik koormus ja ülesobitamise risk [13]. Meetod osutus valituks tänu oma täpsusele ning sellele, et see on andnud paljulubavaid tulemusi pragude tuvastamisel, nagu on näidatud ka Omar jt [13] uuringus.

2.7 Tehisnärvivõrk

Uurimise all on ka tehisnärvivõrk. Tehisnärvivõrk (ingl *artificial neural network*) on masinõppe mudel, mis oma tegevuselt sarnaneb inimaju toimimisele [16]. Tehisnärvivõrk koosneb neuronitest (ingl *neurons*), mis jagunevad mitmetesse kihtidesse. Need kihid omakorda jagunevad sisend-, väljund-, ja vahekihtideks.

Tuginedes allikatele [16, 17], tehisnärvivõrgus olevad neuronid on omavahel ühendatud. Neuronitele on sisendiks neile ühendatud eelnevate neuronite väljund, mis korrutatakse läbi kaaluga, seejärel liidetakse kõik sisendid kokku. Kui tulemus on suurem kui lävend, siis neuron annab edasi väljundi, vastasel juhul mitte. Neuroni lävendid määrab selle aktivatsioonifunktsioon. Algselt on kaalude ja lävendite väärtused suvalised, andes närvivõrgule treeningandmeid muudetakse kaalude väärtusi, et vähendada väljundi viga. Populaarsed närvivõrgud on pärilevivõrgud (ingl *feedforward*), see tähendab, et nendes õppimiseks kasutatavad andmed liiguvad ühes suunas ehk sisendist väljundi. Nii pärilevivõrgud kui ka muud tüüpi võrgud võivad kasutada lisaks tagasilevi (ingl

backpropagation) õppimisalgoritmi, kus info saab veel liikuda väljundist tagasi sisendini, et paremini häälestada kaalude ja lävendite väärtusi.

Sügavaks närvivõrguks (ingl *deep neural network*) nimetatakse võrku, mis sisaldab vähemalt kolme vahekihti [16]. Selle töö raames piirduakse tavaliste närvivõrkudega, kuna töös kirjeldatud ülesande kontekstis on need piisavalt täpsed [18].

Erinevaid tehisnärvivõrke on kasutatud pragude tuvastamiseks, näiteks on seda rakendanud Hein ja Jaanuska [1], Gidiagba jt [4], Habib jt [5]. Oma paindlikkuse ja populaarsuse tõttu on see üks valitud meetodites selles töös.

3. Andmed ja metoodika

Selles peatükis kirjeldatakse kasutatud andmete töötlemist ja pragude identifitseerimismudelite ülesehitust. Praktilises osas kasutatakse Jupyter Notebooki, mis asub Google Colabi keskkonnas. Antud töö raames kasutati järgmisi Python teeke: scikit-learn ja Keras masinõppe jaoks, signaalitöötlemiseks teek SciPy, andmetöötlemise ja -analüüsi jaoks pandas ja NumPy ning visualiseerimiseks teek Matplotlib.

3.1 Fourier' teisendus

Tihti peale on vaja kogutud andmeid edasi töödelda, et saaks neid paremini uurida ja kasutada. Fourier' analüüs on üks selline meetod andmete töötlemiseks. Fourier' teisendus on operatsioon, mis teisendab funktsiooni selle sagedusvahemikke iseloomustavaks funktsiooniks. Seda kasutatakse lainekujude ja funktsioonide uurimiseks, dekonstrueerides need perioodilisteks komponentideks. Kui uuritavad andmed on diskreetsed, siis on tegu diskreetse Fourier' teisendusega [19].

Fourier' kiirteisendus on algoritm, mis võimaldab kiiresti arvutada diskreetset Fourier' teisendust. Seda kasutatakse mitmetes valdkondades, kuna see võimaldab efektiivselt analüüsida vibratsioone ja signaale. Enamjaolt kasutatakse seda perioodiliste andmete uurimiseks, kuid seda on rakendatud ka mitteperioodiliste andmete jaoks. Fourier' teisendus eraldab andmed üksikuteks komponentideks, mis lubab uurida neid uuel viisil. Komponente saab eraldi töödelda, mis on kasulik andmete uurimiseks erikriteeriumite põhjal ja nendest mustrite leidmiseks [19].

Pragude uurimises on Fourier' teisendust rakendatud näiteks akustilistele emissioonidele ja mitmetele vibratsioonidele [20]. Käesoleva töö raames kasutatakse Fourier' teisendust andmete töötlemisel, rakendatakse Fourier' kiirteisendust tala omavõnkumistele.

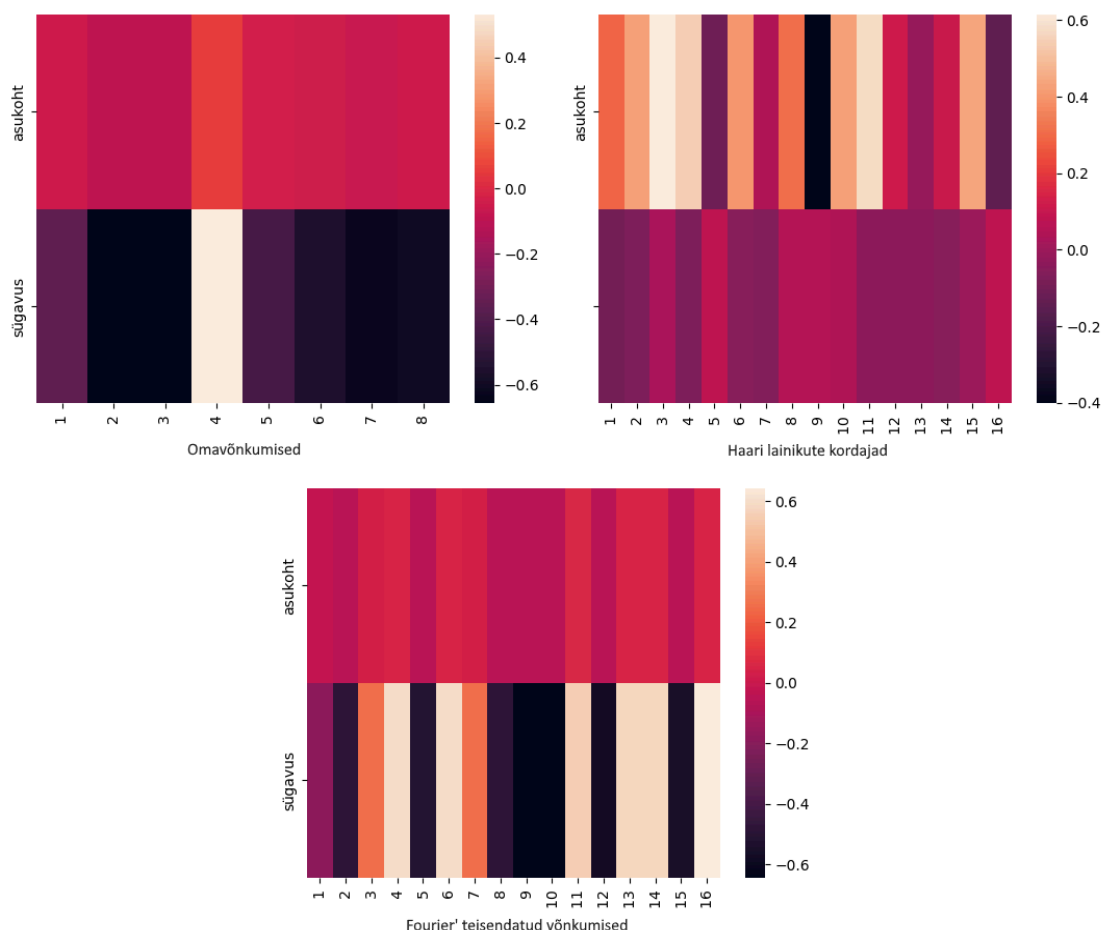
3.2 Andmed

Töös vaadeldav element on Euler-Bernoulli tala, mis on paigaldatud Pasternaki alusele. Tala üks ots on jäigalt kinnitatud ja teine ots on rullkinnitusega. Tala kõrguse ja pikkuse suhe on 0.1. Elemendi pikkus on skaleeritud üheni. Andmed pragude kohta talas jagunesid kaheks andmekoguks. Mõlemas oli toodud prao asukoht paremalt poolt ja prao sügavus. Asukoht oli vahemikus 0.01-0.99 ja sügavus vahemikus 0.01-0.50. Esimeses andmekogus oli toodud prao

asukohale ja sügavusele vastavad kaheksa esimest omavõnkumise parameetrit. Teises andmekogus oli tala esimesest võnkemoodi kujust Haari teisendusega saadud 16 Haari lainiku kordajat. Kokku oli igas andmestikus 4950 rida. Andmebaasides olevad andmed olid saadud arvutuslikul meetodil [18] ning on avalikult kättesaadavad Kaggle platvormil².

Kolmanda andmehulga saamiseks jaoks rakendati tala omavõnkumistele Fourier' kiirteisendust, mille tulemusena saadi 16 tunnust iga kirje jaoks. Andmete töötlemisel loodi ka neljas andmehulk, kus tunnuste väärtuseid oli juhuslikult muudetud 5% ulatuses, et jäljendada valget müra, mis võib esineda modaalsete omaduste mõõtmisel.

Ennustatavateks väärtusteks on asukoht ja sügavus. Ülejäänud tunnustest moodustati kolm tunnusvektorit (ingl *feature vector*): üheks vektoriks on kõik omavõnkumised, teiseks Haari lainikute kordajad, kolmandaks Fourier' teisendatud võnkumised. Tunnuste ja otsitavate väärtuste vahel arvutati nende korrelatsioon, mis on toodud joonisel 2.



Joonis 2. Omavõnkumiste, Haari lainikute kordajate, ja Fourier' teisendatud võnkumiste korrelatsioonigraafikud pragude asukoha ja sügavusega.

² Andmete link:

<https://www.kaggle.com/datasets/ljubovjaanuska/open-crack-in-euler-bernoulli-beam-on-pasternak-fn>

Korrelatsioonigraafikutest selgub, et võnkumised ning Fourier' teisenduse abil saadud sagedused on tundlikud prao sügavuse suhtes, samas kui Haari lainikute kordajad näitavad suuremat tundlikkust prao asukoha suhtes. See on ootuspärane, kuna Haari kordajad on seotud võnkemoodi kuju muutustega, samas kui sagedused sõltuvad peamiselt konstruktsiooni jäikusest, kinnitustest ja terviklikkusest.

3.3 Mudeldamine

Enamik koostatud mudeleid³ põhines teegi scikit-learn vastaval meetodi implementatsioonil, tehisnärvivõrk loodi Kerase abil. Iga meetodi jaoks koostati kuus erinevat mudelit, mis erinesid ennustatava tunnuse poolest (prao asukoht või sügavus) ning sisendandmete ehk tunnusvektori põhjal (omavõnkumised, Haari lainikute kordajad, Fourier' teisendatud võnkumised). Mudelite jaotuse ülevaade on esitatud tabelis 1.

Tabel 1. Loodud mudelite jagunemine ennustatava tunnuse ja ennustamiseks kasutatavate andmete põhjal

Ennustatav tunnus	Ennustamiseks kasutatav tunnusvektor		
Asukoht	Omavõnkumised	Haari lainikute kordajad	Fourier' teisendatud võnkumised
Sügavus	Omavõnkumised	Haari lainikute kordajad	Fourier' teisendatud võnkumised

Kasutatud andmestikud jagati 80% treenimisandmestikuks ja 20% testandmestikuks. Mudelite hindamiseks kasutati mõõdikutena keskmist ruutviga MSE ja determinatsioonikordajat R^2 . Mõlemad on levinud mõõdikud regressiooni ülesannetes. MSE on tundlik suuremate vigade osas, R^2 võrdleb ruutvigade keskmist suurust märgendite dispersiooniga ja näitab, palju on loodud mudel parem nullmudelist. R^2 väärtus 1 tähendab, et mudel selgitab täielikult andmete varieeruvuse, väärtuse 0 või negatiivne väärtus näitab, et mudel ei ole andmete kirjeldamisel parem kui triviaalne mudel [21].

Mudelite kasutaja poolt määratavad parameetrid valiti kas lähtuvalt ülesande püstitusest või leiti parim väärtus katsetamise teel. Need katsetused viidi läbi korduvalt, et vähendada juhuslikkuse mõju parameetrite valikule. Mudelite kasutajapoolselt seatud parameetrid ja

³ Link Colab'is loodud mudelitele:

https://colab.research.google.com/drive/1ze4MNRf8OB__FFCDExaRbq07jBdVLYPk?usp=sharing

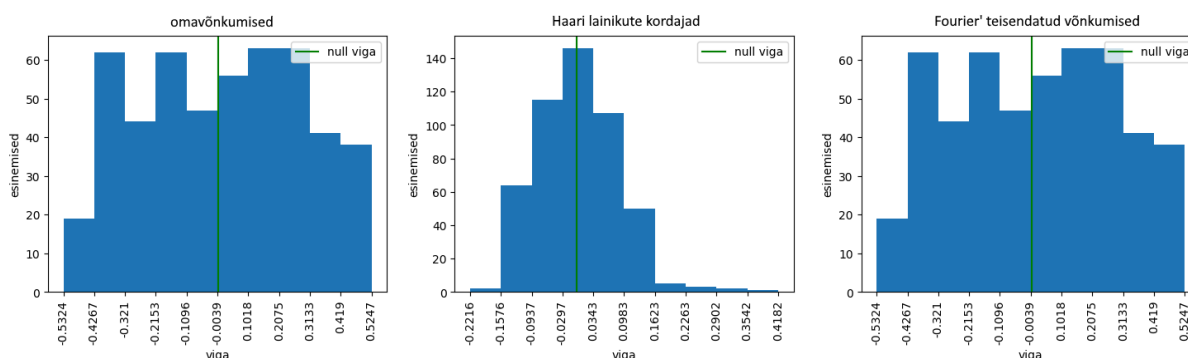
ülesehitis ei erinenud asukoha ja sügavuse ennustamise vahel, kuid osade meetodite puhul oli erinevate sisendandmete korral erinevad parameetrid kasutusel. Mudelite seadistused, mis andsid parima tulemuse, on välja toodud järgmises peatükis koos nende saavutatud tulemustega.

4. Tulemused ja arutelu

Selles peatükis tuuakse välja loodud mudelite tulemused pragude kirjeldamisel, kus asukoha ja sügavuse ennustamise tulemused on kirjeldatud vastavates alapeatükkides. Peatüki lõpus võrreldakse erinevate mudelite toimivust ka müraga andmestiku korral.

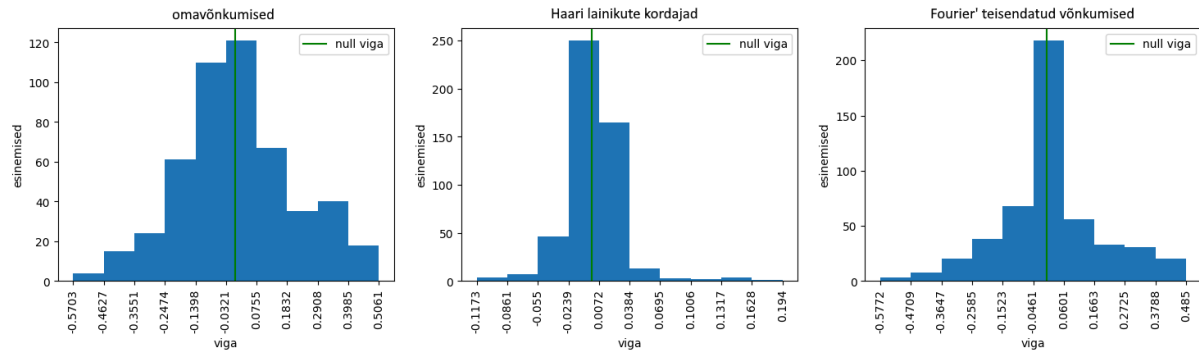
4.1 Asukoha ennustamine

Lineaarregressiooni puhul kasutati selle vaikeimplementatsiooni scikit-learn teegis. Joonisel 3 on toodud lineaarregressiooni tulemused vigade histogrammi kujul. Kasutades ennustamiseks omavõnkumisi oli selle meetodi MSE $7,613 \cdot 10^{-2}$, Haari lainikute kordajate korral oli see $7,158 \cdot 10^{-3}$, Fourier' teisendusega võnkumiste korral $7,613 \cdot 10^{-2}$. Mudelite R^2 väärtused olid vastavalt $-0,006$, $0,905$, ja $-0,006$. Kolme mudeli treenimiseks kulus kokku $0,023$ sekundit.



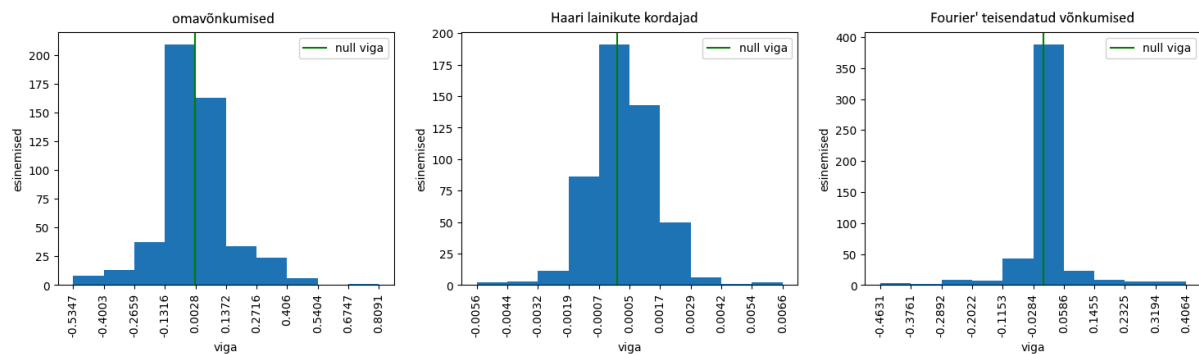
Joonis 3. Lineaarregressiooni vigade histogrammid asukoha ennustamisel.

Kernel ridge puhul andis kõige paremaid tulemusi *laplacian* kernel ning regulatsiooni määr $0,1$. Asukoha ennustamisel kasutades selleks omavõnkumisi oli selle meetodi MSE $4,089 \cdot 10^{-2}$, Haari lainikute kordajate kasutamisel $8,499 \cdot 10^{-4}$, Fourier' teisendusega võnkumiste kasutamisel $3,219 \cdot 10^{-2}$. Mudelite R^2 väärtused olid vastavalt $0,459$, $0,988$, ja $0,574$. Kolme mudeli treenimiseks kulus kokku $8,46$ sekundit. Joonisel 4 on toodud meetodi tulemused vigade histogrammi kujul.



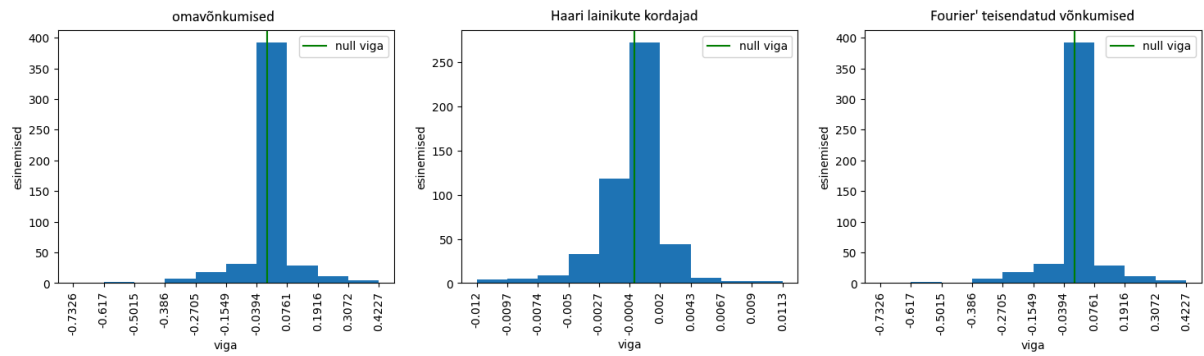
Joonis 4. *Kernel ridge* vigade histogrammid asukoha ennustamisel.

Gaussi protsessi regressiooni rakendavate mudelitel seati α parameetrit parimate tulemuste jaoks vastavalt sisendandmetele: omavõnkumiste korral 10^{-6} , Haari lainikute kordajate korral 10^{-13} , Fourier' teisendatud võnkumiste korral 10^{-9} . Kolme mudeli MSE olid vastavalt $2.258 \cdot 10^{-2}$, $1.865 \cdot 10^{-6}$, ja $6.810 \cdot 10^{-3}$. Mudelite R^2 väärtused olid vastavalt 0,702, 0,999, ja 0,910. Nende mudelite treenimiseks kulus kokku 17,94 sekundit. Joonisel 5 on toodud meetodi tulemused.



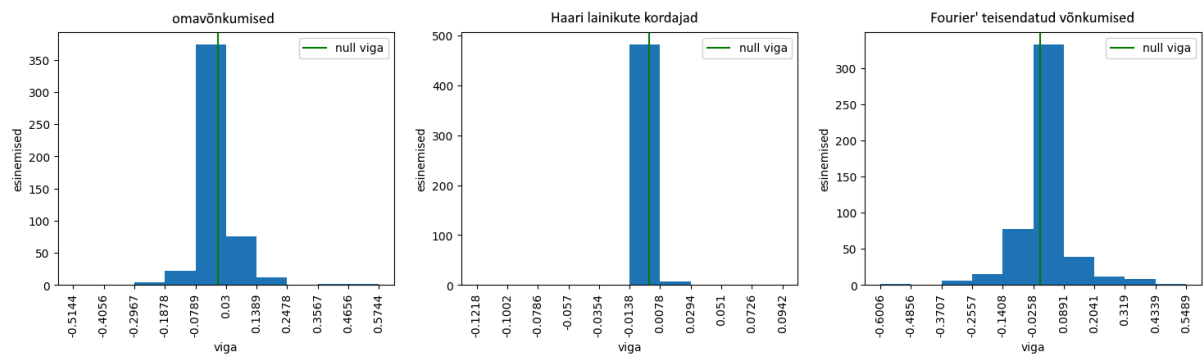
Joonis 5. Gaussi protsessi regressiooni vigade histogrammid asukoha ennustamisel.

KNN-i kasutatavate mudelite korral leiti, et parimad tulemused saadi, kui naabrite arv oli 4. Mudelite tulemused on toodud joonisel 6. Täpsemalt omavõnkumiste kasutamisel oli selle MSE $9,733 \cdot 10^{-3}$, Haari lainikute kordajate kasutamisel $5.705 \cdot 10^{-6}$, Fourier' teisendatud sageduste kasutamisel $9,733 \cdot 10^{-3}$. Mudelite R^2 väärtused olid vastavalt 0,881, 0,999, ja 0,881. Kolme mudeli treenimiseks kulus 0,0176 sekundit.



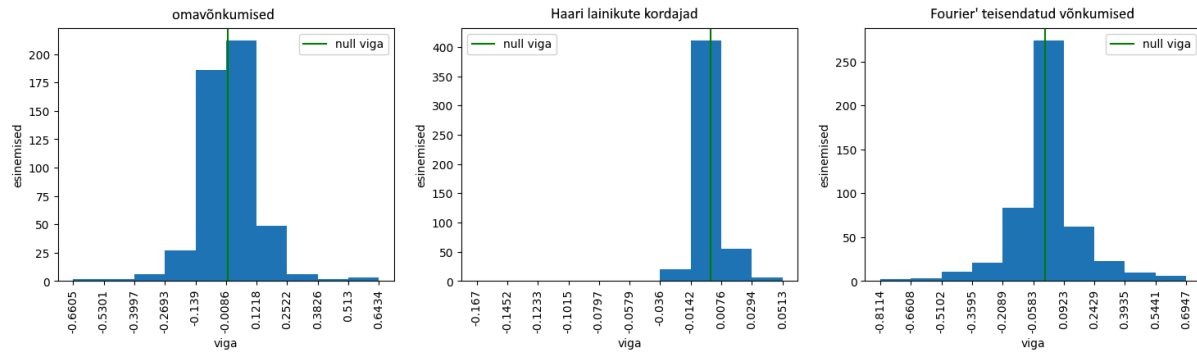
Joonis 6. KNN vigade histogrammid asukoha ennustamisel.

Otsustusmetsa puhul leiti, et parimate tulemuseks jaoks oli optimaalseim puude arv 90. Selliselt üles seatud mudelite tulemused on toodud joonisel 7. Tulemustest täpsemalt oli omavõnkumiste kasutamisel asukoha ennustamiseks $MSE\ 6,009 \cdot 10^{-3}$, Haari lainikute kordajate kasutamisel $6,159 \cdot 10^{-5}$, Fourier' teisendatud sageduste kasutamisel $1,226 \cdot 10^{-2}$. Mudelite R^2 väärtused olid vastavalt 0,920, 0,999, ja 0,838. Kolme mudeli trenimiseks kulus kokku 27,53 sekundit.



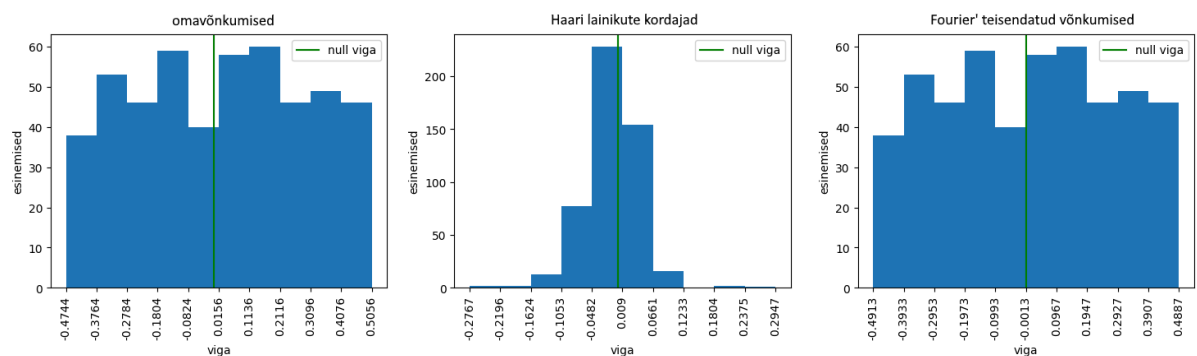
Joonis 7. Otsustusmetsa vigade histogrammid asukoha ennustamisel.

Gradientvõimendamise masinate puhul saadi parimad tulemused, kui õppesamm oli 0,4 ning korduste arv oli 500. Selliste parameetritega mudelite tulemused asukoha ennustamisel on toodud joonisel 8. Täpsemalt oli omavõnkumiste kasutamise korral $MSE\ 1,509 \cdot 10^{-2}$, Haari lainikute kordajate korral $1,429 \cdot 10^{-4}$, Fourier' teisendatud võnkumiste korral $3,180 \cdot 10^{-2}$. Mudelite R^2 väärtused olid vastavalt 0,800, 0,998, ja 0,579. Kolme mudeli trenimiseks kulus kokku 46,46 sekundit.



Joonis 8. Gradientvõimendamise masinate vigade histogrammid asukoha ennustamisel.

Töös kasutatud tehisnärvivõrk sisaldas kahte vahekihti, kokku oli neli kihti. Vahekihtide aktivatsioonifunktsioon oli sigmoid, sisend- ja väljundkihi aktivatsioonifunktsioon oli lineaarne. Neuronite arv mõlemas vahekihis oli 10. Kasutatav optimeerija oli *adam*. Mudelite treenimine kestis kuni 200 iteratsiooni või kuni mudeli kadu ei olnud 10 iteratsiooni järjest muutunud. Sellise ülesehitusega närvivõrgu tulemused on toodud joonisel 9.

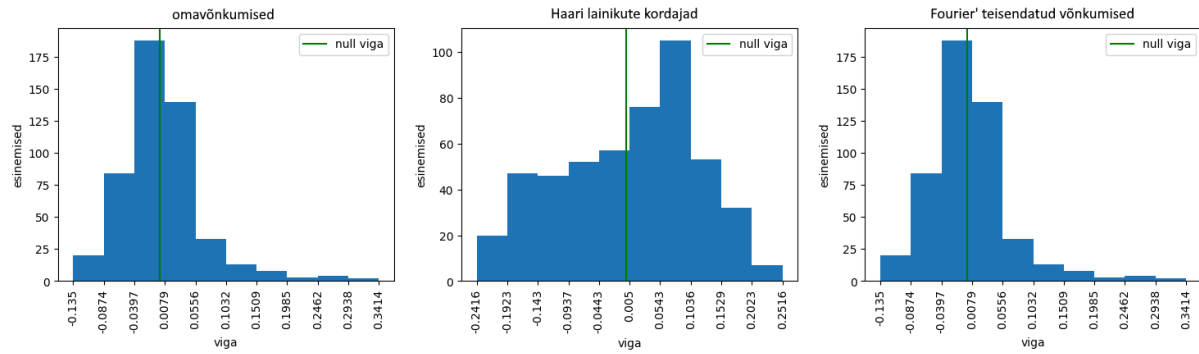


Joonis 9. Tehisnärvivõrgu vigade histogrammid asukoha ennustamisel.

Kirjeldatud närvivõrgu *MSE* oli omavõnkumiste kasutamisel $7,617 \cdot 10^{-2}$, Haari lainikute kordajate kasutamisel $2,999 \cdot 10^{-3}$, Fourier' teisendatud saguste kasutamisel $7,570 \cdot 10^{-2}$. Mudelite R^2 väärtused olid vastavalt -0,006, 0,960, ja 0,004. Kolme meetodi treenimiseks kulus 67,56 sekundit.

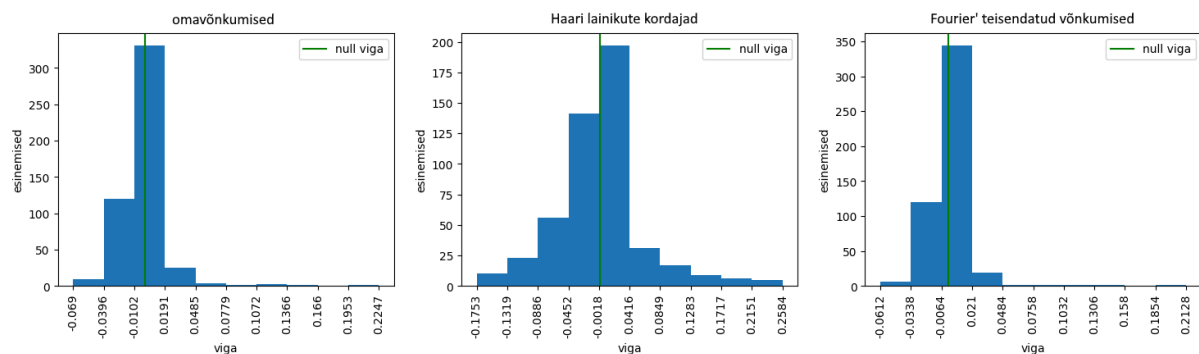
4.2 Sügavuse ennustamine

Lineaarregressiooni tulemused sügavuse ennustamisel on toodud joonisel 10. *MSE* oli omavõnkumisi kasutades $3,985 \cdot 10^{-3}$, Haari lainikute kordajaid kasutades $1,242 \cdot 10^{-2}$, ja Fourier' teisendatud võnkumisi kasutades $3,985 \cdot 10^{-3}$. Mudelite R^2 väärtused olid vastavalt 0,809, 0,406, ja 0,809.



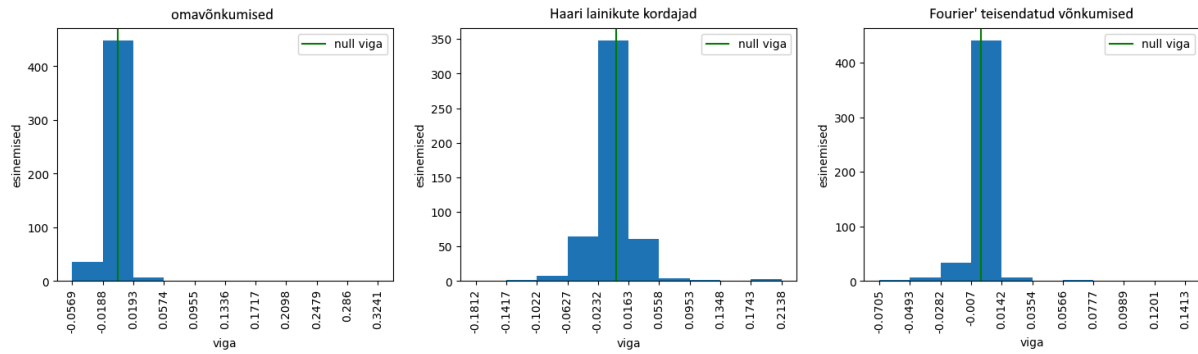
Joonis 10. Linearregressiooni vigade histogrammid sügavuse ennustamisel.

Kernel ridge meetodi *MSE* sügavuse ennustamisel oli omavõnkumisi kasutades $5,128 \cdot 10^{-4}$, Haari lainikute kordajaid kasutades $3,839 \cdot 10^{-3}$, Fourier' teisendatud võnkumisi kasutades $3,388 \cdot 10^{-4}$. R^2 väärtused olid vastavalt 0,975, 0,816, ja 0,983. *Kernel ridge* tulemused vigade histogrammi kujul on toodud joonisel 11.



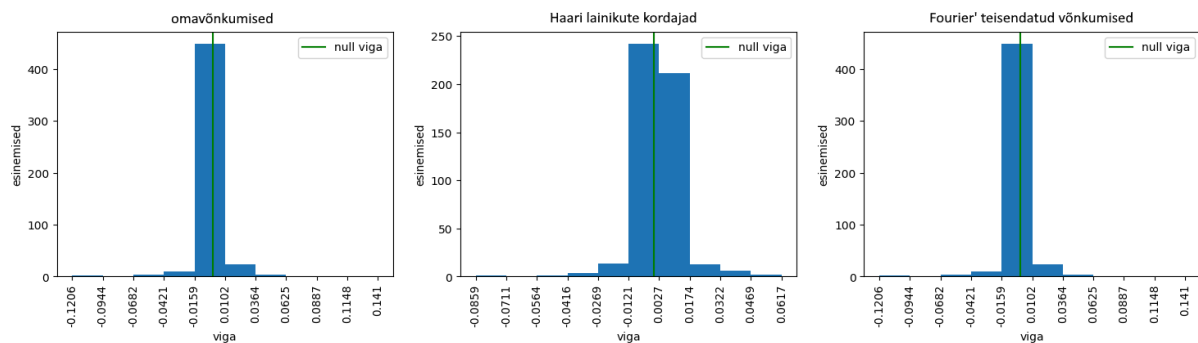
Joonis 11. *Kernel ridge* vigade histogrammid sügavuse ennustamisel.

Gaussi protsessi regressiooni oli *MSE* sügavuse ennustamiseks omavõnkumisi kasutades $5,155 \cdot 10^{-4}$, Haari lainikute kordajaid kasutades $9,949 \cdot 10^{-4}$, Fourier' teisendatud võnkumisi kasutades $1,627 \cdot 10^{-4}$. R^2 väärtused olid vastavalt 0,975, 0,952, ja 0,992. Nende vigade histogrammid on toodud joonisel 12.



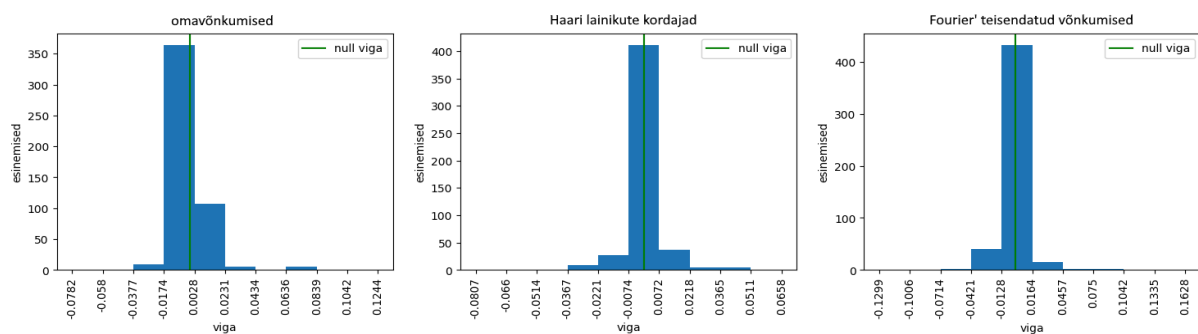
Joonis 12. Gaussi protsessi regressiooni vigade histogrammid sügavuse ennustamisel.

KNN tulemused sügavuse ennustamisel on toodud joonisel 13. Täpsemalt oli MSE omavõnkumisi kasutades $2,053 \cdot 10^{-4}$, Haari lainikute kordajaid kasutades $9,920 \cdot 10^{-5}$, ja Fourier' teisendatud sagedusi kasutades $2,053 \cdot 10^{-4}$. R^2 väärtused olid vastavalt 0,990, 0,995, ja 0,990.



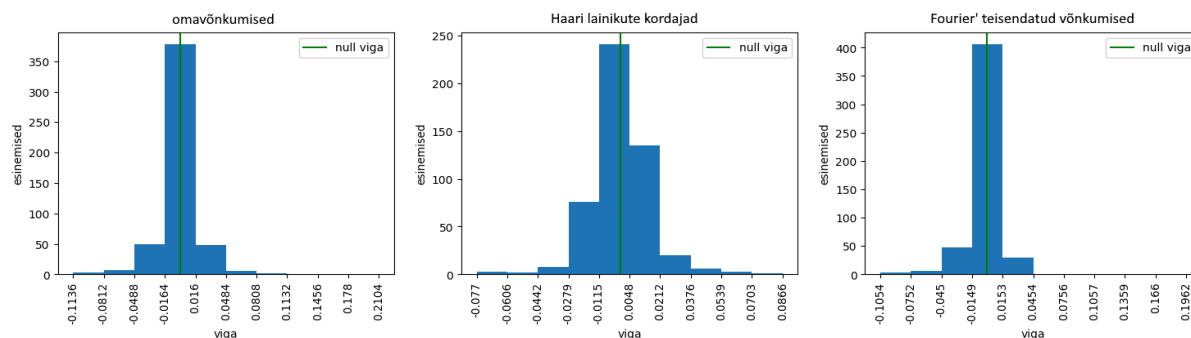
Joonis 13. KNN vigade histogrammid sügavuse ennustamisel.

Otsustumetsa puhul oli sügavuse ennustamisel MSE kasutades omavõnkumisi $1,486 \cdot 10^{-4}$, kasutades Haari lainikute kordajaid $9,204 \cdot 10^{-5}$, ja kasutades Fourier' teisendatud sagedusi $2,165 \cdot 10^{-4}$. R^2 väärtused olid vastavalt 0,993, 0,995, ja 0,990. Vigade histogrammi kujul on tulemused toodud joonisel 14.



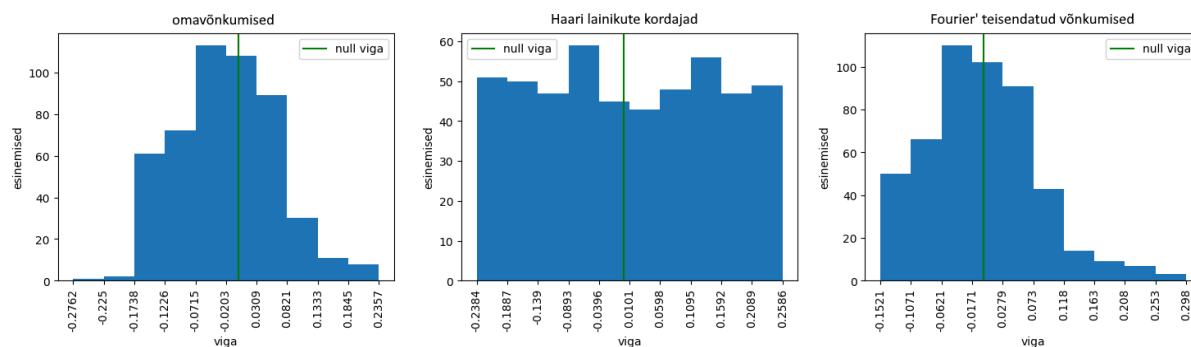
Joonis 14. Otsustumetsa vigade histogrammid sügavuse ennustamisel.

Gradientvõimendamise masinaid kasutatavate mudelite tulemused sügavuse ennustamisel on toodud joonisel 15. Täpsemalt oli MSE omavõnkumisi kasutades $4,897 \cdot 10^{-4}$, Haari lainikute kordajaid kasutades $2,523 \cdot 10^{-4}$, Fourier' teisendatud võnkumisi kasutades $3,632 \cdot 10^{-4}$. R^2 väärtused olid vastavalt 0,976, 0,987, ja 0,983.



Joonis 15. Gradientvõimendamise masinate vigade histogrammid sügavuse ennustamisel.

Tehisnärvivõrgu tulemused sügavuse ennustamisel on toodud joonisel 16. MSE omavõnkumisi kasutades oli $7,255 \cdot 10^{-3}$, Haari lainikute kordajaid kasutades $2,093 \cdot 10^{-2}$, Fourier' teisendatud sagedusi kasutades $6,708 \cdot 10^{-3}$. R^2 väärtused olid vastavalt 0,653, -0,0008, ja 0,679.



Joonis 16. Tehisnärvivõrgu vigade histogrammid sügavuse ennustamisel.

4.3 Müraga andmed

Töö raames katsetati valitud meetodeid ka mürase andmestikuga. Mudelite ülesehitusi selle jaoks ei muudetud. Iga meetodi parim tulemus müraga andmetel on toodud tabelis 2, kus on iga meetodi jaoks toodud nii asukoha kui sügavuse ennustamisel parimale tulemusele vastavad MSE ja R^2 väärtused ning kasutatud tunnused.

Tabel 2. Meetodite parimad tulemused müraga andmetel.

Asukoha ennustamise tulemused			
Meetod	Kasutatud tunnused	MSE	R²
Lineaarregressioon	Haari lainikute kordajad	1,352*10 ⁻²	0,836
<i>Kernel ridge</i>	Haari lainikute kordajad	3,579*10 ⁻³	0,956
Gaussi protsessi regressioon	Haari lainikute kordajad	2,312*10 ⁻²	0,721
KNN	Haari lainikute kordajad	3,626*10 ⁻³	0,956
Otsustusmets	Haari lainikute kordajad	1,162*10 ⁻³	0,985
Gradientvõimendamise masinad	Haari lainikute kordajad	3,124*10 ⁻³	0,962
Tehisnärvivõrk	Haari lainikute kordajad	1,470*10 ⁻²	0,822
Sügavuse ennustamise tulemused			
Meetod	Kasutatud tunnused	MSE	R²
Lineaarregressioon	Ostavõnkumised	7,936*10 ⁻³	0,632
<i>Kernel ridge</i>	Fourier' teisendatud võnkumised	7,007*10 ⁻³	0,675
Gaussi protsessi regressioon	Fourier' teisendatud võnkumised	9,901*10 ⁻³	0,541
KNN	Haari lainikute kordajad	7,703*10 ⁻³	0,643
Otsustusmets	Haari lainikute kordajad	2,184*10 ⁻³	0,898
Gradientvõimendamise masinad	Haari lainikute kordajad	4,039*10 ⁻³	0,813
Tehisnärvivõrk	Fourier' teisendatud võnkumised	1,321*10 ⁻²	0,388

4.4 Arutelu

Tabelis 3 ja tabelis 4 on toodud meetodite tulemused järjestatult. Meetodite tulemusi analüüses on näha, et need sõltusid palju algandmete valikust. Asukoha jaoks saadi paremaid tulemusi kasutades Haari lainikute kordajaid, sügavuse jaoks olid omavõnkumised ja nende Fourier' teisendatud väärtused paremad. See tulemus ühtib varasemalt arvutatud

korrelatsioonidega (joonis 2), kus võnkumised olid rohkem korreleeritud sügavusega ja Haari lainikute kordajad asukohaga. Erandiks olid KNN, otsustusmets, ja gradientvõimendamise masinad, kus Haari lainikute kordajad andsid parimaid tulemusi nii praeguse asukoha kui sügavuse ennustamisel. Ennustamiseks kasutatav tunnuste valikul tuleb seega arvestada nii ennustatava väärtusega ka valitud meetodiga.

Tabel 3. Asukoha ennustamise parimad tulemused järjestatult.

Meetod	Kasutatud tunnused	MSE	R ²
Gaussi protsessi regressioon	Haari lainikute kordajad	1,865*10 ⁻⁶	0,999
KNN	Haari lainikute kordajad	5,705*10 ⁻⁶	0,999
Otsustusmets	Haari lainikute kordajad	6,159*10 ⁻⁵	0,999
Gradientvõimendamise masinad	Haari lainikute kordajad	1,429*10 ⁻⁴	0,998
<i>Kernel ridge</i>	Haari lainikute kordajad	8,499*10 ⁻⁴	0,988
Tehisnärvivõrk	Haari lainikute kordajad	2,999*10 ⁻³	0,960
Lineaarregressioon	Haari lainikute kordajad	7,158*10 ⁻³	0,905

Tabel 4. Sügavuse ennustamise parimad tulemused järjestatult.

Meetod	Kasutatud tunnused	MSE	R ²
Otsustusmets	Haari lainikute kordajad	9,204*10 ⁻⁵	0,995
KNN	Haari lainikute kordajad	9,920*10 ⁻⁵	0,995
Gaussi protsessi regressioon	Fourier' teisendatud võnkumised	1,627*10 ⁻⁴	0,992
Gradientvõimendamise masinad	Haari lainikute kordajad	2,523*10 ⁻⁴	0,987
<i>Kernel ridge</i>	Fourier' teisendatud võnkumised	3,388*10 ⁻⁴	0,983
Lineaarregressioon	Fourier' teisendatud võnkumised	3,985*10 ⁻³	0,809
Tehisnärvivõrk	Fourier' teisendatud võnkumised	6,708*10 ⁻³	0,679

Variandiks on ka võnkumiste ja Haari lainikute kordajate kombineerimine üheks tunnusvektoriks, mida selle töö raames ei käsitletud, kuna eesmärk oli tuvastada nendest tunnustest sobivaim asukohta ja sügavuse ennustamiseks.

Meetodite tulemustest on veel näha, et prao asukohta oli lihtsam ennustada kui prao sügavust. Kõikide meetodite puhul saadi asukohta ennustamisel täpsemaid tulemusi, kui sügavuse ennustamisel. Sama järeldus on tehtud Jaanuska poolt [18].

Fourier' teisenduse rakendamine andmetele andis varieeruvaid tulemusi. Gradient-võimendamise masinate, *kernel ridge*, Gaussi protsessi regressiooni ja tehiseärvivõrgu tulemusi parandasid need märgatavalt võrreldes teisendamata võnkumiste kasutamisega, lineaarregressiooni ja KNN puhul ei olnud mingit muudatust, ning otsustusmetsa puhul tulemus hoopis halvenes.

Tunnused, millega saavutati parim tulemus prao identifitseerimisel, olid sarnased ka mürase andmestikuga ennustades. Ainukeseks erinevuseks oli lineaarregressioon, kus teisendamata võnkumised andsid nüüd täpsema tulemuse, kui teisendatud võnkumised. Müraga andmestiku katsete tulemustest on näha, et kõige vähem oli sellest mõjutatud otsustusmets, mis suutis ka mürase andmestiku korral anda täpseid tulemusi. On võimalik, et valitud meetoditest on otsustusmets kõige parema üldistamisoskusega. Kõige rohkem oli mõjutatud Gaussi protsessi regressioon, mille tulemused muutusid võrreldes puhta andmestiku tulemustega kõige rohkem. Sellele tuginedes on see meetod on väga tundlik kasutatud andmete osas, mistõttu on hoolikas andmetöötlus meetodit kasutades vajalik.

Otsustusmets on varasemalt näidanud häid tulemusi⁴ [18], mistõttu selle töö raames saavutatud tulemused ei ole üllatavad. Võib järeldada, et selle ülesande lahendamise jaoks on tegu hästi sobiva meetodiga.

Tehiseärvivõrgu tulemused olid ootuspärasest madalamad, mis võis tuleneda võrgu ülesehitusest. On võimalik, et võrgu ülesehituse muutmisel ning rohkemal uurimisel võib saavutada paremaid tulemusi, kuid tuleb arvestada, et keerukam mudel nõuab rohkem ressursse. Juba töös vaadeldud võrdlemisi lihtne ärvivõrk oli võrreldes teiste proovitud meetoditega kõige ressursinõudlikum. Lisaks ärvivõrgu ülesehituse ja treenimise parandamisele võib kaaluda juba treenitud võrgu täiendamist (ingl *fine-tuning*).

⁴Tulemused saadi MatLab keskkonnas.

Kokkuvõte

Käesoleva bakalaureusetöö eesmärk oli uurida erinevate masinõppemeetodite sobivust pragude tuvastamiseks Euler-Bernoulli talas. Tala oli paigaldatud Pasternaki alusele, selle üks ots oli jäika kinnitatud ja teises otsas oli rullkinnitus. Ennustati avatud prao asukohta ja sügavust, kasutades selleks Pythoni teekides olevaid masinõppemeetodeid.

Uuritavateks meetoditeks olid lineaarregressioon, *kernel ridge*, Gaussi protsessi regressioon, lähima naabri meetod (KNN), otsustusmets, gradientvõimendamise masinad ja tehishärvivõrk. Kasutatavad andmed olid pärit Kaggle platvormilt. Need jagunesid kaheks andmekoguks, kus esimene sisaldas tala omavõnkumisi ja teine tala esimesest võnkemoodi kujust tuletatud Haari lainikute kordajaid. Kolmas andmestik loodi Fourier' teisendust rakendades omavõnkumistele.

Tulemused näitasid, et prao asukoha ennustamisel andsid parimaid tulemusi mudelid, mis olid treenitud Haari lainikute kordajate põhjal. Sügavuse ennustamisel osutusid üldjuhul edukamateks Fourier' teisendatud võnkumistel põhinevad mudelid, kuigi ka Haari lainikute kordajate kasutamisel saavutati mitmel juhul häid tulemusi. Kõige täpsemaks meetodiks prao asukoha määramisel osutus Gaussi protsessi regressioon, mille keskmine ruutviga oli $1,865 \cdot 10^{-6}$ ja determinatsioonikordaja 0,999. Sügavuse ennustamisel saavutas parima tulemuse otsustusmets, mille vastavad näitajad olid $9,204 \cdot 10^{-5}$ ja 0,995.

Lisaks analüüsiti mudelite tundlikkust müra suhtes, kasutades selleks andmestikke, kus oli valge müra simuleerimiseks väärtusi muudetud 5% ulatuses. Tulemustest selgus, et müraga andmetel säilitas kõrge täpsuse otsustusmets, mis viitab selle meetodi heale üldistusvõimele. Müra suhtes kõige tundlikum oli Gaussi protsessi regressioon, mis näitab, et see meetod vajab hoolikat andmetöötlust ja tunnuste valikut.

Töö tulemused näitavad, et sobiva andmestiku ja masinõppemeetodi valik on pragude tuvastamisel konstruktsioonelementides määrava tähtsusega.

Viidatud kirjandus

- [1] Hein H., Jaanuska L. Quantification of cracks in beams on the Pasternak foundation using Haar wavelets and machine learning. *Proceedings of the Estonian Academy of Sciences*, 2022. https://kirj.ee/wp-content/plugins/kirj/pub/proc-1-2022-16-29_20220131231939.pdf (07.12.2024).
- [2] Shaat M., Ghavanloo E., Emam S. A Micromorphic Beam Theory for Beams with Elongated Microstructures. *Scientific Reports*, 2020, 10. <https://doi.org/10.1038/s41598-020-64542-y> (22.02.2025).
- [3] Tewelde S. A., Krawczuk M. Detection of Closing Cracks in Beams Based on Responses Induced by Harmonic Excitation. *SENSORS*, 2024. <https://www.mdpi.com/1424-8220/24/1/247> (07.12.2024).
- [4] Gidiagba J. O., Tartibu L., Okwu M.O. Crack Detection on a Structural Beam: A Simplified Analytical Method Based on Artificial Neural Network Model. *2022 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)*, 2022. <https://ieeexplore-ieee-org.ezproxy.utlib.ut.ee/document/9856177> (07.12.2024).
- [5] Habib M. A., Hasan M. J., Kim J. A Lightweight Deep Learning-Based Approach for Concrete Crack Characterization Using Acoustic Emission Signals. *IEEE Access*, 2021. <https://ieeexplore-ieee-org.ezproxy.utlib.ut.ee/document/9493229> (07.12.2024).
- [6] IBM. What is linear regression? 2021. <https://www.ibm.com/think/topics/linear-regression> (28.04.2025).
- [7] Scikit-learn developers. Linear Models. https://scikit-learn.org/stable/modules/linear_model (28.04.2025).
- [8] Scikit-learn developers. Kernel ridge regression. https://scikit-learn.org/stable/modules/kernel_ridge (28.04.2025).
- [9] Pedamkar P. Kernel Methods. 2023. <https://www.educba.com/kernel-methods/> (28.04.2025).
- [10] Rasmussen C. E., Williams C. K. I. Gaussian Processes for Machine Learning. Cambridge: The MIT Press. 2006. <https://gaussianprocess.org/gpml/chapters/RW.pdf> (27.04.2025).
- [11] Corrado N, Durrande N, Gherlone M, Mattone M, Hensman J, Surace C. Single and multiple crack localization in beam-like structures using a Gaussian process regression approach. *Journal of Vibration and Control*. 2018.

<https://journals-sagepub-com.ezproxy.utlib.ut.ee/doi/pdf/10.1177/1077546317721418>

(27.04.2025).

[12] Taunk K., De S., Verma S., Swetapadma A. A Brief Review of Nearest Neighbor Algorithm for Learning and Classification. *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, 2019.

<https://ieeexplore-ieee-org.ezproxy.utlib.ut.ee/document/9065747> (25.04.2025).

[13] Omar I., Khan M., Starr A. Comparative Analysis of Machine Learning Models for Predicting Crack Propagation under Coupled Load and Temperature. *Applied Sciences*. 2023.

<https://www.mdpi.com/2076-3417/13/12/7212> (25.04.2025).

[14] IBM. What is random forest? <https://www.ibm.com/think/topics/random-forest>

(26.04.2025).

[15] Ridgeway G. Generalized Boosted Models: A guide to the gbm package. 2024.

<https://cran.r-project.org/web/packages/gbm/vignettes/gbm.pdf> (26.04.2025).

[16] IBM. What is a neural network? 2021.

<https://www.ibm.com/think/topics/neural-networks> (28.04.2025).

[17] Hardesty L. Explained: Neural networks. 2017.

<https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414> (28.04.2025).

[18] Jaanuska L. Haar Wavelet Method for Vibration Analysis of Beams and Parameter Quantification. Tartu: University of Tartu Press. 2021.

[19] Bullard E. Fast Fourier transform (FFT). *Salem Press Encyclopedia of Science*. 2024.

<https://research-ebSCO-com.ezproxy.utlib.ut.ee/c/qlurcm/viewer/html/mjajrokv3f>

(29.04.2025).

[20] Chaudhari C. C., Gaikwad J. A., Bhanuse V. R., Kulkarni J. V. Experimental investigation of crack detection in cantilever beam using vibration analysis. *2014 First International Conference on Networks & Soft Computing (ICNSC2014)*, 2014.

<https://ieeexplore-ieee-org.ezproxy.utlib.ut.ee/document/6906685> (29.04.2025).

[21] Sügis E., Tampuu A., Aljanaki A., Fišel M., Kull M. Praktiline andmeteadus. Tartu: Tartu Ülikooli arvutiteaduse instituut. 2024.

https://courses.cs.ut.ee/t/andmeteadus/Main/HomePage?action=download&upname=Praktiline_andmeteadus_ko%CC%83rgkoolio%CC%83pik.pdf (25.04.2025).

Litsents

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Kristjan Radsin ,
(*autori nimi*)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose

Masinõppe meetodid pragude tuvastamiseks talades ,
(*lõputöö pealkiri*)

mille juhendaja(d) on Ljubov Jaanuska ,
(*juhendaja nimi*)

reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada Tartu Ülikooli digitaalarhiivi kuni autoriõiguse kehtivuse lõppemiseni;

2. annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi kaudu Creative Commons'i litsentsiga CC BY NC ND 4.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni;
3. olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile;
4. kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Kristjan Radsin
15.05.2025