

TALLINNA TEHNIKAÜLIKOOL

Majandusteaduskond

Ärikorralduse instituut

Janar Vahter

**PANKROTIMUDELI KONSTRUEERIMINE EESTI TÖÖTLEVA  
TÖÖSTUSE ETTEVÕTETE NÄITEL**

Magistritöö

Õppekava ärirahandus ja majandusarvestus, peeriala majandusarvestus

Juhendaja: külalisdotsent Margus Tinitis

Tallinn 2018

Deklareerin, et olen koostanud töö iseseisvalt ja olen viidanud kõikidele töö koostamisel kasutatud teiste autorite töödele, olulistele seisukohtadele ja andmetele, ning ei ole esitanud sama tööd varasemalt ainepunktide saamiseks. Töö pikkuseks on 12 645 sõna sissejuhatusest kuni kokkuvõtte lõpuni.

Janar Vahter .....

(allkiri, kuupäev)

Üliõpilase kood: 153045TARM

Üliõpilase e-posti aadress: janar.vahter@gmail.com

Juhendaja: külalisdotsent Margus Tinitis

Töö vastab kehtivatele nõuetele

.....

(allkiri, kuupäev)

Kaitsmiskomisjoni esimees:

Lubatud kaitsmisele

.....

(nimi, allkiri, kuupäev)

# SISUKORD

LÜHIKOKKUVÕTE .....	5
SISSEJUHATUS .....	6
1. ETTEVÕTTE PANKROTISTUMINE NING SELLE PROGNOOSIMISE TEOREETILISED ALUSED .....	9
1.1. Ettevõtte majanduslik ebaõnnestumine ning selle peamised mõjutegurid .....	9
1.2. Pankrotimodelite olulisus ning ülesehituslik funktsioon .....	12
1.3. Valimi koostamine ning muutujate valik pankroti modelleerimisel .....	14
1.4. Pankrotimodelite konstrueerimisel rakendatavad analüüsimeetodid .....	20
1.4.1. Statistilised meetodid .....	20
1.4.2. Masinõppe meetodid .....	23
2. PANKROTID EESTI TÖÖTLEVAS TÖÖSTUSES NING ANDMETE KAASAMINE MODELLEERIMISSE .....	25
2.1. Ülevaade ettevõtete pankrottide sagedusest Eesti töötlevas tööstuses .....	25
2.2. Analüüsi valitud algandmed ning muutujad .....	28
2.2.1. Algandmed ning nende kaasamisel rakendatud piirangud .....	28
2.2.2. Algalikus analüüsi kaasatud sõltumatud muutujad .....	31
2.3. Modelleerimisse kaasatava valimi ettevalmistus ning kirjeldav statistika .....	33
2.3.1. Null- ning ekstreemväärtuslike vaatluspunktide eemaldamine algalimist .....	33
2.3.2. Modelleerimisse kaasatavate andmete kirjeldav statistika ning jaotus .....	36
3. EESTI TÖÖTLEVA TÖÖSTUSE PANKROTIMODELI KONSTRUEERIMINE .....	40
3.1. Pankrotimodeli konstrueerimine logistilise regressiooni rakendamisel .....	40
3.2. Pankrotimodeli konstrueerimine otsustuspuu meetodi rakendamisel .....	47
3.2.1. Pankrotimodeli konstrueerimine CART algoritmi rakendamisel .....	49
3.2.2. Pankrotimodeli konstrueerimine CHAID algoritmi rakendamisel .....	51
3.3. Konstrueeritud pankrotimodelite tulemuste võrdlus ning järeldused .....	54
KOKKUVÕTE .....	57
SUMMARY .....	61
KASUTATUD ALLIKATE LOETELU .....	64
LISAD .....	71
Lisa 1. Ekstreemväärtuslikud vaatluspunktid – Mahalanobise ja Cook'i kaugused. ....	71
Lisa 2. Multikollinearsuse eemaldamine – VIF test .....	75

Lisa 3. Konstrueeritud pankrotimodelite hinnang treeningvalimisse kaasatud ettevõtete pankrotistumise tõenäosusele ( $P_i$ ).....	76
Lisa 4. Konstrueeritud pankrotimodelite hinnang kontrollvalimisse kaasatud ettevõtete pankrotistumise tõenäosusele ( $P_i$ ).....	79
Lisa 5. Konstrueeritud CART pankrotimodeli struktuur (blokkskeem).....	81
Lisa 6. Konstrueeritud CHAID pankrotimodeli struktuur (blokkskeem).....	82

## LÜHIKOKKUVÕTE

Käesoleva magistritöö eesmärgiks seadis autor Eesti töötleva tööstuse ettevõtete pankrotimudeli konstrueerimise. Töö eesmärgi saavutamiseks anti esmalt ülevaade ettevõtete pankrotistumisega ning selle prognoosimisega seotud teoreetilisest raamistikust. Modelleerimise aluseks olnud andmed Eesti töötleva tööstuse ettevõtete kohta koguti Registrate ja Infosüsteemide Keskusest. Andmete kogumisel rakendati mitmeid piiranguid, mis põhinesid ettevõtte tegevusharu klassifikatsioonil (EMTAK jagu C), tegevusaastate arvil (minimaalselt 3 aastat) ning töötajate keskmisel arvil (10–249 töötajat). Eeltoodud piiranguid rakendas autor andmete kogumisel eesmärgiga tagada konstrueeritud pankrotimudelite kõrgem kvaliteet ning praktiline väärtus.

Magistritöö empiirilises osas rakendas autor Eesti töötleva tööstuse pankrotimudelite konstrueerimisel statistilise analüüsimeetodina logistilist regressiooni ning masinõppel põhineva otsustuspuu meetodi algoritme CART (*Classification and Regression Tree*) ja CHAID (*Chi-square Automatic Interaction Detector*). Modelleerimisel seati eesmärgiks maksimeerida pankrotimudelite klassifitseerimisvõimekus (kontrollvalimil hinnatuna) vähima võimaliku hulga muutujate juures, toetades seeläbi mudelite võimalikult lihtsat tõlgendatavust ning ka hilisemat rakendatavust. Parima tulemuse ettevõtete pankrotiohu hindamisel, seda keskmiselt aastas (täpsemalt 6–18 kuud) ette vaatavas ajaperspektiivis, saavutas logistilisel regressioonil põhinenud pankrotimudel. Ehkki CHAID algoritmil ning logistilisel regressioonil põhinenud mudelite puhul ulatus hindamisel esinenud I tüüpi vigade määr (s.o pankrotistunud ettevõtete klassifitseerimine tegevust jätkavate ettevõtete hulka) samale tasemele, siis mudelite spetsiifilisuses (*specificity*) ning tundlikkuses (*sensitivity*) mõõdetud erinevuste põhjal võis logistilisel regressioonil põhinevat pankrotimudelit hinnata siiski mõnevõrra stabiilsemaks. Autori hinnangul võib viimast rakendada ka toetava abivahendina Eesti töötleva tööstuse ettevõtetega seotud juhtimis-, finantseerimis-, investeerimis- ja ka audiitorotsuste langetamisel. Seejuures tuleb aga tingimata arvestada kõigi töös kirjeldatud piirangutega.

Võttesõnad: Eesti töötlev tööstus, pankrotimudel, logistiline regressioon, otsustuspuu meetod, CART, CHAID

## SISSEJUHATUS

Kuigi ettevõtete pankrotistumine on majandusprotsessi igati loomulik osa, siis ometigi üritatakse seda igal võimalusel vältida. Peamiselt seetõttu, et ettevõtte pankrotistumine mõjutab negatiivselt paljusid osapooli (eelkõige omanikke, töötajaid, tarnijaid, krediidasutusi jne), millest tingituna kujuneb antud protsess tihti äärmiselt kulukaks, seda nii individuaalsel kui ka sotsiaalsel tasandil (McKee, Lensberg 2002; Balcaen, Ooghe 2004). Tingituna ettevõtte pankrotistumisega seotud suurtest finantsilistest ning sotsiaalsetest kuludest on pankrotiohu prognoosimine ühes usaldusväärsete pankrotimudelite väljatöötamisega tõusnud üha enam ja enam erinevate osapoolte tähelepanu keskmesse. Töökindlates pankrotimudelites nähakse aina olulisemat abivahendit juhtimis-, investeerimis- ning finantseerimis- (McKee, Lensberg 2002; Chen 2011a; Chen *et al.* 2011), aga ka audiitorotsuste langetamisel (McKee, Lensberg 2002; Balcaen, Ooghe 2004).

Ühelt poolt on pankrotimudelite arvukust ning arengut toetanud nõudlus prognoosimudelite järele. Teisalt aga tänaseni jätkuv progress modelleerimisel rakendatavate meetodite osas (Doumpos, Zopounidis 1999, 73) ning pankroti modelleerimiseks vajalike andmete läbipaistvuse ja kättesaadavuse paremine (Balcaen, Ooghe 2004, 3). Andmete kättesaadavuse paremine on omakorda loonud võimaluse konstrueerida ka oluliselt spetsiifilisemaid pankrotimudeleid, kaasates modelleerimisse tunduvalt homogeensemate tunnustega ettevõtteid (Blanco-Oliver *et al.* 2015). Pankrotimudelite konstrueerimisel on aegade jooksul rakendatud mitmeid erinevaid meetodeid. Need jagunevad üldjoontes kahte rühma: statistilised ning masinõppe meetodid (Chen 2011a, 4514). Statistilistest analüüsimeetoditest on kirjanduses enim kajastust leidnud mitmemõõtmeline diskriminantanalüüs ning logistiline regressioon (Onofrei, Lupu 2014, 250), masinõppe meetoditest aga tehislike närvivõrkude ning otsustuspuu meetod (Chen 2011a, 4516).

Käesoleva magistritöö eesmärk on Eesti töötleva tööstuse ettevõtete pankrotimudeli konstrueerimine. Töötlev tööstus tervikuna panustab Eesti sisemajanduse koguprodukti koondväärtusesse tegevusharudest kõige rohkem (Eesti statistikaamet, tabel RAA0042), osutudes tegevusharude lõikes muuhulgas ka suurimaks tööandjaks (Eesti statistikaamet, tabel

EM001). Töötleva tööstuse roll Eesti majanduse stabiilsuse ning jätkusuutlikkuse toetamisel on äärmiselt oluline. Seejuures selgub Krediidiinfo pankrotiuuringust, et Eesti majanduse tegevusharude keskmisega võrreldes ulatub töötleva tööstuse ettevõtete pankrottide sagedusmäär koguni poole kõrgemale (Creditinfo Eesti AS 2016, 11). Eeltoodud statistika toetab tugevalt vajadust ajakohase ning usaldusväärse prognoosimudeli järele, mis oleks võimeline andma erinevatele osapooltele täpset ning ajas ette vaatavat informatsiooni huvialuste töötleva tööstuse ettevõtete finantstervise kohta.

Olenemata teema olulisusest on autorile teadaolevalt Eesti töötleva tööstuse pankrotimudeli konstrueerimist varasemalt lõputöodes käsitletud vaid paaril korral, seda Grünbergi (2013) ning Kelementi (2015) magistritöodes. Kusjuures viimane käsitles pankrotimudeli konstrueerimisel konkreetsemalt Eesti töötleva tööstuse tegevusharu alla kuuluvat mööblitööstuse tegevusala. Kuivõrd aga pankrotimudelitesse kaasatud finantsnäitajate optimaalne tase ajas muutub, on mudelite usaldusväärse tagamiseks oluline neid perioodiliselt häälestada (Mensah 1984) või koguni täismahuliselt uuendada (Begley *et al.* 1996). Lisaks olulisuse printsiibile tulenes käesoleva magistritöö teema valik ka sellest, et töö autoril on ettevõtete jätkusuutlikkuse hindamisega tugev tööalane seos, seda nimelt krediidianalüütilisest perspektiivist.

Püstitatud eesmärgi saavutamiseks seab magistritöö autor järgmised uurimisülesanded:

1. Anda ülevaade ettevõtte pankrotistumisega ning selle prognoosimisega seotud teoreetilisest raamistikust.
2. Kirjeldada pankrotimodelite koostamisel enim kasutatust leidnud analüüsimeetodeid ning nende rakendamisega seotud piiranguid.
3. Anda ülevaade pankrotistumise trendidest Eesti töötlevas tööstuses ning koguda andmeid kõnealuse tegevusharu pankrotistunud ning jätkuvalt tegutsevate ettevõtete kohta.
4. Konstrueerida kogutud andmete põhjal usaldusväärne ning võimekas pankrotimudel, rakendades selleks erinevaid analüüsimeetodeid.
5. Võrrelda konstrueeritud pankrotimodelite klassifitseerimistäpsust ning stabiilsust.

Eesti töötleva tööstuse ettevõtete pankrotimodelite konstrueerimisel rakendatakse töös logistilist regressiooni ning otsustuspuu meetodi algoritme CART (*Classification and Regression Tree*) ja CHAID (*Chi-square Automatic Interaction Detector*). Autorile teadaolevalt ei ole otsustuspuu meetod pankroti modelleerimist käsitlevates lõputöodes Eestis seni rakendust leidnud.

Magistritöö jaguneb kolmeks peatükiks. Töö esimene peatükk annab ülevaate ettevõtte pankrotistumisega ning selle prognoosimisega seotud teoreetilisest raamistikust. Muuhulgas pankrotiolukorra definitsioonist, pankrotistumist põhjustavatest peamistest mõjuteguritest, pankrotimudelite olulisusest, pankroti prognoosimisel kasutatavatest andmetest ning pankroti modelleerimisel seni enim kasutatust leidnud analüüsimeetoditest. Töö esimene peatükk põhineb suuresti teemakohastel ingliskeelsetel teadusartiklidel. Teine peatükk annab esmalt ülevaate pankrotistumise trendidest Eesti töötlevas tööstuses. Seejärel kirjeldab autor modelleerimisse kaasatavate andmete kogumise ja muutujate valiku protsessi, muuhulgas ka kogutud andmete eelanalüüsi protsessi. Vaatlusaluseks perioodiks töötleva tööstuse pankroti modelleerimisel seab autor 2005–2016 ning uurimisaluste ettevõtete andmed kogutakse Justiitsministeeriumi poolt hallatavast Registrate ja Infosüsteemide Keskusest. Töö kolmas peatükk käsitleb pankrotimudelite konstrueerimist eelmainitud meetodite rakendamisel ning väljatöötatud pankrotimudelite tulemuste võrdlust.

Autor hindab käesolevas magistritöös käsitletavat teemat väga olulisuseks ning usub, et töö tulemus pakub nii teoreetilist kui ka praktilist väärtust mitmele erinevale osapoolle. Magistritöö valmimise osas soovib autor tänu avaldada esmajärjekorras töö juhendajale, Margus Tinitale, asjalike ja põhjalike soovitude eest, ning Registrate ja Infosüsteemide Keskuse infosüsteemi haldurile, Larissa Merkulovalle, koostöövalmiduse eest. Lisaks soovib autor tänada ka oma lähedasi töö koostamise vältel osutatud toetuse eest.



# 1. ETTEVÕTTE PANKROTISTUMINE NING SELLE PROGNOOSIMISE TEOREETILISED ALUSED

## 1.1. Ettevõtte majanduslik ebaõnnestumine ning selle peamised mõjutegurid

Rahvusvahelises teemakohases kirjanduses on ettevõtte majanduslikku ebaõnnestumist kirjeldava olukorra määratlus laialt varieeruv. Kraus ja Litzenberger (1973) täheldasid, et uurimistöodes rakendatakse ettevõtte ebaõnnestumist fikseerivate ilmingutena mitmeid faktoreid. Teiste hulgas näiteks likviidsusprobleemide ilmnemist, suutmatust teenindada laenukohustusi või püsivat maksejõuetust väljendavat olukorda. Eelkõige oleneb määratluse valik ettevõtte ebaõnnestumisele viitava olukorraraamistiku defineerimisel määratleja enda perspektiivist lähtuvatest huvidest ning uuritavast kontekstist (Balcaen, Ooghe 2004, 18–19), aga kindlasti ka andmete kogumisel esinevatest piirangutest.

Teemaalases kirjanduses enim rakendust leidnud ettevõtte ebaõnnestumist fikseerivaks ilminguks on ettevõtte pankrot, nimelt õiguslikus mõistes defineerituna. Seda eelkõige tingituna ettevõtte pankrotti kirjeldava definitsiooni selgusest, selle tõlgendatavuse objektiivsusest ning pankrotiolukorra ajalise määratluse konkreetsusest (*ibid.*, 18–19). Eesti pankrotiseaduse § 1 kohaselt on pankrot võlgniku kohtuotsusega väljakuulutatud maksejõuetus. Tegemist on olukorraga, kus võlgnik ei suuda rahuldada nõudeid võlausaldajate ees ja see suutatus ei ole võlgniku majanduslikust olukorrast tulenevalt ajutine. Võlgnikku hinnatakse pankrotiseaduse § 1 alusel maksejõuetuks ka sellisel juhul, kui selle vara ei kata tema kohustusi ja selline seisund ei ole võlgniku olukorrast tulenevalt ajutine. Pankrotimenetlus lõpeb kohtu otsusel, kas pankrotiavalduse rahuldamata jätmisega, pankroti väljakuulutamisega või pankrotimenetluse raugemise tõttu (PankrS § 27). Viimase näol on tegemist erijuhtumiga, kus pankrotimenetlus lõpetatakse pankrotti väljakuulutamata, seda olukorras, kus ettevõttel ei jätku vara pankrotimenetlusega tekkivate kulude katteks (PankrS § 29). Teisisõnu on pankroti raugemise korral sisult tegemist pankrotistunud ettevõtte definitsioonile vastava ettevõttega, mille puhul pankrotimenetluse läbiviimine pole menetluse vastaspooltele lihtsalt majanduslikult otstarbekas.

Ehkki õiguslikel alustel defineeritud pankrotiolukord on ettevõtte majanduslikku ebaõnnestumist kirjeldava ilminguna uurimustöodes laialdast kasutust ning üldist heakskiitu leidnud, on seda määratlusvormi mõningatel juhtudel ka kritiseeritud. Peamise probleemina kõnealuse määratluse juures on välja toodud see, et kohtuliku pankrotimenetluse lõppemise hetke ei peegelda tingimata ettevõtte tegelikku ehk majanduslikku ebaõnnestumise ajalist hetke. Seejuures toonitatakse, et ettevõtte majandusliku ebaõnnestumise tegeliku asetleidmise ning selle õiguslikus formaadis fikseerimise hetke vahele võib jääda võrdlemisi pikk ajaline periood (Ooghe *et al.* 1995)

Alternatiivina õiguslikult defineeritud pankrotistumisele rakendas näiteks Laitinen (1994) teostatud uurimustöös ettevõtte majanduslikku ebaõnnestumist fikseeriva ilminguna likviidsusprobleemide esinemist, kirjeldamaks ettevõtte võimetust katta tasumisele kuuluvaid lühiajalisi kohustusi õigeaegselt. Ward ja Foster (1997) leidsid aga, et suutmatuse teenindada laenukohustusi krediitiasutuste ees on otstarbekohane olukorda kirjeldav avaldumisvorm defineerimaks ettevõtte majanduslikku ebaõnnestumist. Seda esiteks põhjusel, et laenukohustuste teenindamise suutmatusele viitava olukorra esinemist on võimalik ajaliselt konkreetselt määratleda. Selleks on võimalik rakendada näiteks Euroopa Liidu poolt pankadele kehtestatud uuendatud Basel II raamistikus esitletud definitsiooni – üle 90 päevane makseviivitus (Hayden 2003). Lisaks argumenteerisid Ward ja Foster (1997), et laenuteenindamise suutmatuse puhul on tegemist majanduslikku ebaõnnestumist indikeeriva olukorraga, mitte üksnes õiguslikku sisu omava fikseeritava sündmusega, nagu seda on õiguslikel alustel defineeritud pankrotistumine. Ometi on eeltoodud määratluse puhul tegemist liialt krediidiorienteeritud lähenemisega, mis ei kajasta tingimata püsivat majanduslikku olukorda ning ei ole seega sobivaim määratlusvorm analüüsimeks ettevõtete ebaõnnestumist üldisemas majanduslikus kontekstis (Balcaen, Ooghe 2004, 20).

Küsimuses, kas ettevõtte pankrotistumine, seda eelkõige õiguslikus mõistes defineerituna, on sisukaim olukorra kirjeldus esindamaks ettevõtte majanduslikku ebaõnnestumise hetke, ei ole autorid seni konsensuslikule vastusele jõudnud. Sellegi poolest on just eelmainitud määratlus ettevõtete majandusliku ebaõnnestumise defineerimisel ning analüüsimisel, muuhulgas prognoosimudelite koostamise meetodikas, läbi aegade korduvalt ning jätkuvalt rakendust leidnud. Teiste hulgas näiteks Altmani (1968), Ohlsoni (1980), Taffleri (1983), Chen (2011a), Palinko ja Svoob (2016) uurimustöodes. Tulenevalt sellest, et ettevõtte pankrotistumine on ettevõtte majanduslikku ebaõnnestumist fikseeriva ilminguna teemakohase modelleerimise

raames sedavõrd laialdast rakendust leidnud, kutsutakse ettevõtte ebaõnnestumist prognoosivaid mudeleid sageli lihtsalt pankrotimudeliteks.

Oluline teoreetiline alus ettevõtte ebaõnnestumist kirjeldava olukorra definitsiooni kõrval on ka selle olukorra kasuaalne raamistik. Erinevalt ettevõtte ebaõnnestumisele viitava olukorra defineerimisest, on seda olukorda tingivate peamiste põhjuste raamistiku osas autorite arusaam mõnevõrra kattuvam. Ettevõtte ebaõnnestumist põhjustavaid potentsiaalseid tegureid on mõistagi arvukalt ning seega ei ole võimalik määratleda konkreetset tegurite komplekti, mille mõjul ettevõtte majanduslik ebaõnnestumine aset leiab. Üldine kirjanduses kajastuv seisukoht on see, et ettevõtte ebaõnnestumist, halvimal juhul koguni pankrotistumist põhjustavad tegurid jagunevad kahte fundamentaalsesse gruppi: endogeenseteks ning eksogeenseteks teguriteks (Korol, Korodi 2010; Palinko, Svoob 2016).

Endogeensed tegurid on ettevõtte poolt vähemal või rohkemal määral kontrollitavad faktorid, mis mõjutavad ettevõtte majanduslikku võimekust ning jätkusuutlikkust (Korol, Korodi 2010). Lizal (2002, 2) jaotab peamised ettevõtte majanduslikku ebaõnnestumist põhjustavad endogeensed tegurid ühistest tunnusoontest lähtuvalt omakorda kolme alamgruppi:

1. vara ebaefektiivne kasutus ettevõtte tegutsemisel;
2. vähene likviidsus ning ülemäärane võimendus;
3. ebapädev juhtkond ning selle majanduslikud otsused.

Eksogeensete tegurite puhul on seevastu tegu üldise majanduskeskkonnaga seotud faktoritega, nagu seda on näiteks inflatsioonimäär, fiskaal- ja monetaarpoliitika, konkurentsi tase, valuutakurss jmt. Eksogeenseid tegurid ei ole ettevõtte poolt mõjutatavad. Küll aga mõjutavad antud tegurid igapäevaselt ettevõtte tegevuse toimimist, seda muuhulgas läbi maksukoormuse, krediidikulukuse, valuutakursi jne. Eksogeensed tegurid jaotab mõjuulatuse avaldamise alusel omakorda kahte alamgruppi jaotanud (Korol, Korodi 2010, 202):

1. kõiki või valdavat osa tegutsevaid ettevõtteid samasuunaliselt mõjutavad tegurid (näiteks üldine konkurentsi kasv);
2. ettevõtteid vastassuunaliselt mõjutavad tegurid (näiteks muutus valuutakursis).

Tihti peale peitub ettevõtte majandusliku ebaõnnestumise põhjus mõlemat liiki majandustegurite koosmõju tulemusel ehk ebasoodsate muutuste samaaegses asetleidmises nii endo- kui ka eksogeensetes tegurites (Palinko, Svoob 2016, 528). Käesoleva magistritöö autor nõustub

siinkohal seisukohaga, et ettevõtte majanduslikule ebaõnnestumisele avaldavad olulist mõju mõlemad, nii endo- kui ka eksogeensed tegurid, seda tihtipeale samaaegselt. Küll aga leiab autor, et eelnimetatud endogeensetest teguritest avaldab ettevõtte ebaõnnestumisele suurimat mõju siiski ettevõtte ebapädev juhtkond ning selle majanduslikud otsused. Vara ebaefektiivne kasutus, vähene likviidsus ning ülemäärane võimendus on seejuures pelgalt juhtkonna ebapädevate otsuste majanduslikuks tagajärjeks.

## **1.2. Pankrotimudelite olulisus ning ülesehituslik funktsioon**

Majandusteooriale tuginedes on ettevõtete pankrotistumine majandusprotsessi osa, mille läbi süsteem puhastub ebaefektiivselt toimivatest ettevõtetest, pakkudes seejuures laiemat tegevusruumi efektiivsematele ning konkurentsivõimekamatele ettevõtetele (Numani 2015, 839). Praktikas on ettevõtte pankrotistumisega seotud aga mitmed olulised negatiivseid nüansid, millest peamine on ehk see, et ettevõtte pankrotistumine avaldab negatiivset mõju paljudele seotud osapooltele (eelkõige omanikele, töötajatele, tarnijatele, krediitiasutustele jne). Seetõttu kujuneb ettevõtte pankrotistumine tihti äärmiselt kulukaks ja seda koguni laiemas mõistes, nii individuaalsel kui ka sotsiaalsel tasandil (McKee, Lensberg 2002; Balcaen, Ooghe 2004). Tulenevalt ettevõtte pankrotistumisega seotud suurtest finantsilistest ning sotsiaalsetest kuludest on pankrotiohu prognoosimine ning seehulgas töökindlate pankrotimudelite väljatöötamine tõusnud aja jooksul üha rohkem ja rohkem erinevate osapoolte tähelepanu keskmesse.

Usaldusväärsetes pankrotimudelites nähakse aina olulisemat abivahendit juhtimis-, investeerimis- ning finantseerimisotsuste langetamisel (McKee, Lensberg 2002; Chen 2011a; Chen *et al.* 2011) aga ka näiteks jätkuvuse printsiibi hindamisel audiitorotsuste raames (McKee, Lensberg 2002; Balcaen, Ooghe 2004). Pankrotimudelite arengut on ühelt poolt mõjutanud nõudlus prognoosimudelite järele, teisalt ka tugevalt toetanud jätkuv progress andmetöötlustarkvara ning modelleerimisel rakendatavate meetodite võimekuses (Doupoumis, Zopounidis 1999, 73). Mitte vähem tähtsa tegurina on pankrotimudelite arvukuse kasvu stimuleerinud ka modelleerimiseks vajalike andmete läbipaistvuse ning kättesaadavuse paranemine (Balcaen, Ooghe 2004, 3).

Andmete kättesaadavuse paranemine on omakorda loonud võimaluse konstrueerida ka oluliselt spetsiifilisemaid pankrotimudeleid, kaasates modelleerimisse oluliselt homogeensemata

tunnustega ettevõtteid (Blanco-Oliver *et al.* 2015), lähtudes seejuures näiteks ettevõtte suurusest (näiteks Altman, Sabato 2007), tegevusharust (näiteks Chava, Jarrow 2004), vanusest (näiteks Wilson, Altanlar 2014) ja/või muudest määratletavatest tunnustest. Taffler (1983) ning Mensah (1984) soovivad pankroti modelleerimisel lähtuda eelkõige ettevõtete sarnasest suurusest, seda nimelt vara mahu mõistes, ning tegevusharust. Üldiselt leitakse, et mida homogensemata tunnustega ettevõtteid pankroti modelleerimisel kasutatud on, seda kõrgemaks kujuneb konstrueeritud pankrotimudeli prognoosivõimekus ning ka praktiline usaldusväärus (Blanco-Oliver *et al.* 2015).

Pankrotimudelite peamine funktsioon on hinnata olemasolevate andmete ning määratletud parameetrite põhjal ettevõtte seisundit, seda pankrotistumise ohu kujul. Laialdaselt kasutust leidnud meetodika pankrotimudelite konstrueerimisel põhineb sõltuva muutuja ehk mudeli väljundi dihhotoomsusel, kus väärtus 1 tähistab pankrotistuvat ettevõtet ning väärtus 0 jätkuvalt tegutsevat ettevõtet või vastupidi. Dihhotoomsel meetodikal põhinevad pankrotimudelid on ometigi pälvinud ka laia kriitikat. Eelkõige seepärast, et praktikas ei ole võimalik tõmmata eelkirjeldatud kahe kategooria ettevõtete vahele selget joont ning seetõttu ei ole dihhotoomsel väljundil põhinevad pankrotimudelid ehk niivõrd praktilised abivahendid ettevõtete klassifitseerimisel (Cybinski 2001, 30). Eelkirjeldatud kitsaskohast tulenevalt on kirjanduses toetust leidnud ka seisukoht, et pankrotimudel peaks lisaks pankrotistuvate ning jätkuvalt tegutsevate ettevõtete kategooriale võimaldama klassifitseerida kolmanda väljundkategooriana ka „halli ala“ ettevõtteid. „Halli ala“ ettevõtete all peetakse silmas neid ettevõtteid, mille puhul on hindamise hetkel ebaselge, kas tegu on pankrotistuvate või tegevust jätkavate ettevõtetega (Balcaen, Ooghe 2004, 18). Ometi on pankrotimudelite dihhotoomne ülesehitus pankroti modelleerimisel rakendust leidnud silmapaistvalt enamal määral kui eelkirjeldatud trihhotoomne ülesehitus.

Dihhotoomsel väljundil põhineva pankrotimudeli rakendamisel ettevõtete klassifitseerimisel võib klassifitseerimistulemustes esineda olemuslikult kahte tüüpi vigu. Esiteks võib rakendatud mudel klassifitseerida pankrotistuva ettevõtte tegutsema jäävate ettevõtete hulka, mida kajastatakse I tüüpi vea juhtumina. Vastupidiselt, II tüüpi vea juhtumi puhul klassifitseerib pankrotimudel tegutsema jääva ettevõtte pankrotistuvate ettevõtete hulka (Gepp, Kumar 2015, 398; Korol, Korodi 2010, 210). I tüüpi vea esinemise puhul on üheks võimalikuks ning seejuures väga relevantseks negatiivseks tagajärjeks olukord, kus pankrotimudeli ebakorrektsel väljundi põhjal ärilise otsuse teinud osapooled kannavad finantsilist kahju. II tüüpi vea puhul on aga peamiseks

negatiivseks resultaadiks alternatiivkulu tekkimine, mis väljendub eelkõige tulu näol (dividenditulu, intressitulu jmt), mis pankrotimudeli ebakorrektsel väljundi põhjal langetatud otsusest tingituna osapooltel saamata jääb (Gepp, Kumar 2015, 398). Altman ja Narayanan (1997, 3) rõhutasid seejuures, et mudeli klassifitseerimisel esinevate II tüüpi vea juhtumite puhul võib tegu olla siiski lähimatel tulevastel perioodidel pankrotistuvate ettevõtetega.

Selge on see, et mõlemat tüüpi vigade esinemise minimeerimisel pankrotimudeli konstrueerimisel on suur tähtsus. Kuigi võib eeldada, et kuluaspektist on mudeli I tüüpi vea esinemise tõenäosuse minimeerimine olulisem, siis konsensuslikku seisukohta pankrotimudeli I tüüpi ning II tüüpi vigade kulusuhte osas võetud ei ole. Seda seetõttu, et I tüüpi vea esinemise olulisus kulusuhtes väljendatuna võrreldes II tüüpi vea esinemisega oleneb igakordselt mudeli rakendaja subjektiivsest hinnangust (Gepp, Kumar 2015, 398). Altman *et al.* (1977, 46) hindasid, et näiteks kommertsbankade finantseerimistegevuses osutub I ja II tüüpi vigade esinemise vaheliseks kulusuhteks koguni 35:1. Metoodiliselt lähtutakse konstrueeritud mudelite klassifikatsioonivõimekuse hindamisel enamasti mudeli üldisest korrektsusmäärast modelleerimisse kaasatud ettevõtete klassifitseerimisel. Antud lähenemise mõistes omavad mudeli klassifitseerimisvõimekuse hindamisel I ning II tüüpi vea juhtumid võrdväärset negatiivset kaalu (Gepp, Kumar 2015, 400; Balcaen, Ooghe 2004, 15). Seejuures on metoodiliselt oluline hinnata mudelit selle konstrueerimiseks kasutatud andmetest (ehk treeningvalimist) eraldiseisvatel vaatluspunktidel (ehk kontrollvalimil), seda eelkõige objektiivsema ja usaldusväärsema hinnangu võimaldamise eesmärgil (Onofrei, Lupu 2014; Li, Wang 2014).

Magistritöö autor leiab, et eeltoodud punkt on modelleerimise seisukohast koguni esmajärgulise tähtsusega, kui eesmärk on konstrueerida usaldusväärne ning seejuures praktilise väärtusega pankrotimudel. Hinnates mudeli klassifitseerimisvõimekust selle konstrueerimiseks kasutatud andmetel, saavutatakse ehk näiliselt küll parem klassifitseerimistulemus, kuid antud metoodiline lähenemine ei paku piisavat usaldusväärsust mudeli hilisemaks rakendamiseks.

### **1.3. Valimi koostamine ning muutujate valik pankroti modelleerimisel**

Pankroti modelleerimisel kasutatav andmevalim, seejuures andmevalimi koostamisel rakendatud meetod, omab äärmiselt olulist mõju konstrueeritava pankrotimudeli ülesehitusele, nii selle

funktsionaalsusele kui ka rakendatavuse ulatusele. Pankroti modelleerimisel on üheks võimaluseks analüüsivalimi koostamisel kaasata valimisse kogu populatsioon, seda kõigi uurimisaluste ettevõtete andmete kättesaadavusel, või siis juhusliku valiku alusel läbilõige ettevõtetest, mis esindavad uurimisalust populatsiooni (Balcaen, Ooghe 2004, 22). Enamik pankroti modelleerimisel kasutatud andmevalimitest on koostatud aga hoopis paaride meetodile tuginedes, mille puhul koosneb pool moodustatud koguvalimist pankrotistunud ettevõtete ning teine pool jätkuvalt tegutsevate ettevõtete vaatluspunktidest (Scott 1981, 320). Sobitatud paaride meetod on seejuures paaride meetodi edasiarendus, mille puhul ei kaasata ettevõtteid paaridesse üksnes juhuslikul valikul. Sobitatud paaride meetodi rakendamisel lähtutakse paaride moodustamisel ettevõtete sarnastest tunnustest, seda eelnevalt defineeritud kriteeriumite alusel. Meetodi rakendamisel saavutatakse andmevalim, mis koosneb homogeensete tunnustega pankrotistunud ning jätkuvalt tegutsevate ettevõtete vaatluspunktidest. Enamasti tuginetakse paaride sobitamisel ettevõtete vara mahu, müüditulude taseme, tegevusala ja/või vanuse kriteeriumitele (Balcaen, Ooghe 2004, 24). Paaride sobitamine on laialt levinud, nimelt seetõttu, et sobitatud paaride meetod lubab mudeli koostajal kontrollida mõningaid parameetreid. Nimelt neid parameetreid, millel eeldatakse olevat oluline mõju ettevõtte finantsseisundile, kuid mida ei soovita ühel või teisel põhjusel kaasata modelleerimisse täiendavate muutujate kujul (Ooghe, Verbaere 1985, 21; Keasey, Watson 1991, 94–95).

Kriitikud toovad esile, et sobitatud paaride meetodi rakendamisel koostatud valimi puhul on tegu mittejuhusliku valimiga ning see ei kattu modelleerimise meetodiliste eeldustega. Lisaks toonitatakse, et mittejuhuslik andmete kaasamine modelleerimisse võib negatiivselt mõjutada mudeli stabiilsust ning ülekantavust ettevõtete üldkogumile (Balcaen, Ooghe 2004, 23–24). Jones (1987) aga leiab, et pankroti modelleerimise kontekstis on sobitatud paaride meetod valimi koostamisel siiski eesmärgipärasem moodus. Kaasates valimisse ettevõtteid juhuslikul valikul, võib konstrueeritud pankrotimudeli tähelepanu keskpunkt pankrotistuvate ning jätkuvalt tegutsevate ettevõtete klassifitseerimiselt eemale kalduda. Jones lisab, et modelleerimise tulemuseks võib seejuures kujuneda lausa olukord, kus konstrueeritud mudel keskendub hoopis suurte ning väikeste või erinevate tegevusvaldkondade ettevõtete üksteisest eristamisele.

Pankrotimudeli funktsiooni defineerimisel mängib olulist rolli ka valimisse kaasatavate andmete ajaline mõõde ehk andmete ajaline suhe pankroti asetleidmisesse. Klassikaliste pankrotimudelite puhul on enamasti tegu staatiliste mudelitega, mis põhinevad ühel konkreetset ajalisel läbilõikel ( $t - x$  aastat) suhtena pankrotistumise hetke ( $t$ ) (Balcaen, Ooghe 2004, 39). Enamasti kasutatakse

pankroti modelleerimisel viimast saadaolevat infot, kaasates modelleerimisse vaatlusaluste ettevõtete finantsandmed seisuga üks aasta enne pankrotistumist, mida käsitletakse ka kui  $t - 1$  andmetena (*ibid.*, 36). Teoorias prognoosib  $t - 1$  andmete põhjal konstrueeritud mudel pankrotiohtu iseenesestmõistetavalt üheaastases ette vaatavas perspektiivis.

Kuigi staatilised pankrotimudelid on kirjanduses rohkelt kajastust leidnud, siis olulise puudusena toovad kriitikud staatiliste mudelite puhul esile selle, et need käsitlevad pankrotistumist hetkena, mitte protsessina, mida ettevõtte pankrotistumine tegelikult endast kujutab (*ibid.*, 41). Lisaks on täheldatud, et konkreetsel ajalisel läbilõikel (näiteks  $t - 1$  aasta andmetel) põhineva pankrotimudeli klassifitseerimise võimekus halveneb oluliselt, kui seda rakendatakse vaatlusaluste ettevõtete mõne muu perioodi andmetel (näiteks  $t - 2$  või  $t - 3$  aastat andmetel) (Laitinen 1993).

Lisaks eeltoodule omab pankrotimudeli võimaliku ülesehituse defineerimisel olulist mõju ka modelleerimisse kaasatavate sõltumatute muutujate valik ning hulk. Kui mikroökonomikal tuginevat kirjandust pankrotiohtu prognoosivate mudelite ning nendesse kaasatud sõltumatute muutujate kohta leidub laialdaselt, siis makroökonomikal põhinev lähenemine ettevõtete pankrotistumise prognoosimisel on leidnud oluliselt vähem tähelepanu ning sellest tulenevalt ka rakendust. Seetõttu põhinebki enamik kirjanduses esitletud ettevõtete pankrotistumist prognoosivatest mudelitest just mikrotasandil mõõdetavatel näitajatel, nimelt finantsnäitajatel (Korol, Korodi 2010, 203). Pankroti modelleerimisel kasutatavad ettevõtte finantsnäitajad jagunevad olemuslikult omakorda kahte gruppi: aktsiaturupõhisteks ning raamatupidamislikeks finantsandmeteks (Chen *et al.* 2011, 135–136).

Suurem osa pankrotimudelitest on konstrueeritud puhtalt ettevõtte finantsaruandluses kajastatud informatsiooni alusel, kus pankroti modelleerimisse kaasatakse üksnes aruannetes esitatud andmete põhjal arvutatud finantsuhtarve ja/või absoluutnäitajaid (Dimitras *et al.* 1996). Selline lähenemine on palju rakendust leidnud eelkõige seetõttu, et teoorias on ettevõtte finantsandmete puhul tegemist objektiivselt mõõdetavate ning kajastatavate andmetega. Lisaks on ettevõtete finantsandmete kogumine, tulenevalt nende võrdlemisi heast kättesaadavusest, lihtne ning see võimaldab koostada modelleerimise tarbeks suuremahulisi andmekogusid (Micha 1984, 282; Laitinen 1992, 325). Raamatupidamislike finantsandmete laialdasest kättesaadavusest peitub iseenesest ka antud tüüpi andmete peamine eelis aktsiaturupõhiste andmete ees. Kuivõrd aktsiaturupõhised andmed on kättesaadavad vaid börsil noteeritud ettevõtete puhul, siis aktsiaturupõhiste andmete modelleerimisse kaasamine piirab oluliselt analüüsi kaasatavate



ettevõtete võimalikku hulka (Chen *et al.* 2011, 135–136). Kirjeldatud piirang võib aga endast kujutada olulist probleemi pankrotimudelite konstrueerimisel, seda eelkõige väiksemate majanduspiirkondade põhised, kus börsil noteeritud ettevõtete arv on liiga väike usaldusväärse pankrotimudeli konstrueerimiseks. Jardini (2009) koostatud tööst, mis koondab enda alla 190 varasemalt teostatud pankrotiteemalise uurimustöö tulemused, selgub, et raamatupidamislikel andmetel arvutatud finantsuhtarvused kasutati modelleerimisel 93% töödes. Turupõhised finantsnäitajad (muuhulgas aktsiahind, aktsiatootlus) leidsid seejuures kajastust vaid 6% töödes.

Olenemata senisest kõrgeast populäärsusest on raamatupidamislike finantsandmete kasutamisel pankroti modelleerimisel omad kitsaskohad, mida tingimata silmas tuleb pidada. Hillegeist *et al.* (2004, 6) kritiseerivad, et finantsaruanded on koostatud jätkuvuse printsiibist lähtudes ning finantsaruannetes esitatud andmed on olemuslikult ajaloolised. Kuivõrd pankrotistumise puhul on aga tegu jätkuvuse printsiibi rikkumisega, siis on ettevõtte pankrotistumise prognoosimine jätkuvuse printsiibil koostatud aruannetele ning nendes esitatud andmetele tuginedes fundamentaalselt väär. Agarwal ja Taffler (2008, 1542) leiavad, et finantsaruannete koostamise aluseks olevate raamatupidamislike standardite looval tõlgendamisel on pahatahtlikul juhtkonnal võimalik aruannetes esitatava finantsinformatsiooniga manipuleerida ning varjata kajastatud numbrite taga peituvat tegelikku majanduslikku väärtust. Seetõttu ei pruugi raamatupidamislikud finantsandmed pakkuda piisavat usaldusväärset ei pankrotimudelite väljatöötamisel ega ka ettevõtete hilisemal klassifitseerimisel.

Aktsiaturupõhiste andmete kasutamises pankroti modelleerimisel nähakse eelkirjeldatud probleemidele olulist leevendust. Esiteks, efektiivselt toimivatel turgudel peegeldub turupõhistest finantsnäitajatest investorite hinnang ettevõtte tegevusele nii ajaloolises kui ka tulevikku suunatud võtmes. Teiseks, võrreldes raamatupidamislike andmetega omab ettevõtte juhtkond aktsiaturupõhiste andmete üle oluliselt nõrgemat otsest kontrolli, mis muudab pahatahtliku juhtkonna jaoks aktsiaturupõhiste andmetega manipuleerimise märgatavalt keerukamaks (Bauer, Agarwal 2014, 433). Shumway (2001) leidis, et aktsiaturupõhiste andmete kaasamine pankroti modelleerimisse kombineeritult raamatupidamislikel andmetel põhinevate näitajatega parendab oluliselt mudeli klassifitseerimisvõimekust. Chava ning Jarrow (2004) leidsid aga, et raamatupidamislikel andmetel põhinevatel finantsnäitajatel on aktsiaturupõhiste näitajate kõrval pankrotimudelis pigem madal lisandväärtus.

Nagu Bellovary *et al.* (2007) 165 pankrotiteemalist tööd koondavast ülevaatest selgub, siis ometi on raamatupidamislikel andmetel põhinevad näitajad aktsiaturupõhistest näitajatest pankroti modelleerimisel oluliselt rohkem kasutust leidnud. Vaatlusaluste tööde lõikes varieerub prognoosimudelisse kaasatud muutujate hulk ühest kuni 57 muutujani (*ibid.*, 7). Tööde lõikes enim kasutust leidnud finantssuhtarvud on toodud tabelis 1.1.

Bellovary *et al.* (2007) töid välja, et koondatud 165 töö lõikes finantssuhtarvuna on enim kasutust leidnud puhaskasum suhtena koguvarasse, seda pea igas kolmandas uurimustöös. Kõige sagedamini kasutatud suhtarvu nimetajaks on koguvara (*ibid.*, 7). Esimese turupõhise suhtarvu võib vaatlusalusest nimekirjast leida alles 13. kohalt. Antud suhtarvuks on omakapitali turuväärtuse suhe kohustiste bilansilisse väärtusesse, mida kasutati 16 vaatlusaluses töös (*ibid.*, 42). Ühe olulise järeldusena töid Bellovary *et al.* välja, et suur arv muutujaid mudelis ei tähenda tingimata mudeli kõrgemat prognoosivõimekust (*ibid.*, 13).

Tabel 1.1. Enimlevinud kaastatud finantssuhtarvud ettevõtete pankrotiuuringutes

Suhtarv	Esinemise sagedus 165 uurimustöö lõikes
Puhaskasum / koguvara	54
Käibevara / lühiajalised kohustised	51
Käibekapital / koguvara	45
Eelmiste perioodide jaotamata kasum / koguvara	42
Ärikasum / koguvara	35
Müügitulu / koguvara	32
(Käibevara - varud) / lühiajalised kohustised	30
Kohustised / koguvara	27
Käibevara / koguvara	26
Puhaskasum / bilansiline netoväärtus	23

Allikas: Bellovary *et al.* (2007, 42)

Kaasates pankroti modelleerimisse ainult finantsnäitajaid eeldatakse vaikumisi, et kogu vajalik informatsioon klassifitseerimaks ettevõtteid pankrotistuvateks või jätkuvalt tegutsevateks sisaldub finantsaruannetes. Üksnes ettevõtte kasumiaruandest ning bilansist tuletatud näitajatele tuginedes võib aga ettevõtte majandusliku seisundi hindamine kujuneda piiratuks, kuivõrd ettevõtte finantsnäitajad ei pruugi kajastada kõiki olulisi ettevõtte jätkuvusega seotud tegureid (Maltz *et al.* 2003). Seetõttu soovitatakse lisaks finantsnäitajatele kaasata pankrotimudelite konstrueerimisse ka ettevõtte majanduslikust seisundist informatsiooni andvaid mittefinantsilisi kvantitatiivseid ning kvalitatiivseid indikaatoreid (näiteks Blanco-Oliver *et al.* 2015). Sedasorti

indikaatoritest on pankrotiuuringutes rakendust leidnud näiteks ettevõtte vanus, juhtkonna kogemus aastates, tegevuse diferentseeritus võtmeklientide ning -tarnijate arvu kujul, konkurentide arv piirkonnas jmt (Balcaen, Ooghe 2004, 30). Wilson ja Altanlar (2014) rõhutavad ka maksekäitumisega seotud probleemide ning Peel ja Peel (1989) auditiotsuse arvesse võtmise olulisust ettevõtte pankrotiohu hindamisel.

Muutujate valikul on rakendust leidnud paljud erinevad meetodid. Balcaen ja Ooghe (2004, 36) toovad välja, et enim toetatakse muutujate valikul varasemale kirjandusele, lähtudes muutujate valikul nende senisest kasutusulatuses, võttes seejuures arvesse ka muutujate osas kirjeldatud klassifitseerimisvõimekust. Seda toetab ka Jardini (2009) 190 pankrotiteemalist uurimustööd koondav statistika. Jardini vaadeldud uurimustööde lõikes oli 40% juhtudel toetunud muutujate valikul varasemas kirjanduses kajastatule. Ühemõõtmelised statistilised meetodid (muuhulgas t-test ning F-test) leidsid muutujate valikul seejuures rakendust 17% töödes. Otsustuspuu meetod ühes mitmemõõtmelise regressiooniga ning majandusteoorial põhinev valik meetodiga aga vaid 4% töödes (Jardin 2009, 8).

Tabel 1.2. Jardini (2009) vaadeldud 190 uurimustöö lõikes muutujate valikul rakendatud meetodi osatähtsus.

Pankrotimudelisse kaasatud muutujate valikul rakendatud meetod	Osatähtsus kaetud 190 töö lõikes
Populaarsus ja/või testitud prognoosivõimekus olemasolevas kirjanduses	40%
Ühemõõtmelised testid: t test, F test, korrelatsioonimaatriks, koefitsiendid	17%
Sammregressioon + Wilksi lambda	16%
Sammregressioon + tõepärasuhte kriteerium	10%
Geneetiline algoritm ja muud spetsiaalsed algoritmid (mh Relief, Tabu)	6%
Eksperti otsus	4%
Mittelineaarsed analüüsimeetodid (nagu näiteks tehislilikud närivõrgud)	3%
Muud meetodid (mitmemõõtmeline regressioon, otsustuspuu meetod, majandusteoorial põhinev valik)	4%

Allikas: Jardin (2009, 8)

Mensah (1984) toob olulise asjaoluna pankrotimudelite rakendamise juures välja selle, et olenemata mudeli koostamisel rakendatud meetodikast ning mudelisse kaasatud muutujatest on oluline arvestada sellega, et suhtarvude optimaalne tase ajas muutub. Seetõttu on vajalik pankrotimudeli praktilisuse tagamiseks perioodiliselt häälestada mudelisse kaasatud muutujate koefitsiente. Begley *et al.* (1996) väidavad aga, et pelgalt koefitsientide häälestamisest ei piisa ning pankrotimudelite usaldusväärsuse tagamiseks peab teatud aja tagant aset leidma mudelite

täismahuline uuendamine. Mis puudutab mudeli häälestamise ulatust, siis siinkohal nõustub käesoleva magistr töö autor viimasega eelkirjeldatud seisukohtadest. Pelgalt koefitsientide väärtuste häälestamine eeldab justkui vaikumisi, et muutuvast majanduskeskkonnas püsivad pankrotiohule viitavad näitajad ajast muutumatud, mis ei pruugi aga vastata tegelikkusele. Isegi kui see nii ka on, siis meetodiliselt korrektne on seda sellegipoolest läbi täismahulise mudeli uuendamise taaskordselt tõestada.

## **1.4. Pankrotimudelite konstrueerimisel rakendatavad analüüsimeetodid**

### **1.4.1. Statistilised meetodid**

Pankrotimudelite konstrueerimisel on aegade jooksul välja töötatud ning rakendatud erinevaid meetodeid. Üldjoontes liigitatakse analüüsimeetodid kahte gruppi: statistilised ning masinõppe meetodid (Chen 2011a, 4514). Statistiliste meetodite rakendamise pioneeriks loetakse pankroti modelleerimise raamistikus Beaverit (1966), kes kasutas pankrotiuuringus ühemõõtmelist analüüsi (*univariate analysis*). Paar aastat hiljem andis Altman (1968) tugeva tõuke aga mitmemõõtmelise diskriminantanalüüsi (*multiple discriminant analysis*) rakendamisele ning mõnevõrra hiljem kinnitas pankroti modelleerimisel tugevalt kanda ka logistiline regressioon (*logistic regression*), eesotsas Ohlsoni (1980) pankrotimudeliga.

Ühemõõtmelise analüüsi puhul hinnatakse ettevõtte finantsseisundit iga analüüsi kaasatud finantsnäitaja põhjal eraldi (Korol, Korodi 2010, 4). Ühemõõtmelisel analüüsil põhinevatest pankrotiuuringutest on kirjanduses enim kajastust leidnud Beaveri (1966) töö. Seda tööd esitletakse ka kui tänapäevase pankrotianalüüsi alguspunkti, tingituna selle mõju ulatusest hilisematele pankrotiuuringutele ning -analüüsidele (Serrano-Cinca, Gutiérrez-Nieto 2013, 1246).

Ühemõõtmelise analüüsi eelis peitub eelkõige selle lihtsuses, kui võrd meetodi rakendamine ei sea olulisi meetodilisi eelduseid analüüsi läbiviimiseks. Samas aga eeldab see meetod, et ettevõtte pankrotistumise ning analüüsi kaasatud finantsnäitajate vahel eksisteerib ühemõõtmeline ning lineaarne seos (Balcaen, Ooghe 2004, 6). Ometi on põhjust arvata, et ettevõtte finantsseisund on kujutatav vaid mitmemõõtmeliselt (Serrano-Cinca, Gutiérrez-Nieto 2013, 1246) ning ettevõtte finantsseisundi ning kajastavate näitajate vahel esinev seos ei ole lineaarselt kirjeldatav (Balcaen, Ooghe 2004, 6).

Altmani (1968) töö oli üks esimesi, mis käsitles ettevõtte pankrotiohu hindamisel mitut finantssuhtarvu paralleelselt, rakendades selleks mitmemõõtmelist diskriminantanalüüsi. Mitmemõõtmelise diskriminantanalüüsi rakendamisel kirjeldatakse ettevõtte finantsseisundit mudelisse kaasatud sõltumatute muutujate väärtuste ning antud muutujatele omistatud kaalude põhjal (läbi koefitsientide) kombineeritult. Tegemist on lineaarse funktsiooniga, mille väljundiks on pideva väärtusega diskriminantskoor (Ooghe *et al.* 2009, 9). Altmani (1968) tööstusettevõtete põhjal konstrueeritud pankrotimudel, mida tuntakse ka kui z-skoori mudelit, sisaldab ettevõtte finantsseisundi hindamiseks viit suhtarvu. Mudeli klassifitseerimistäpsuseks, eristamaks pankrotistuvaid jätkuvalt tegutsevatest ettevõtetest, saavutati koguni 95%. Z-skoori pankrotimudel avaldub järgnevalt (*ibid.*, 594):

$$Z_i = 0,012 \times x_1 + 0,014 \times x_2 + 0,033 \times x_3 + 0,006 \times x_4 + 0,999 \times x_5 \quad (1)$$

kus

$Z_i$  – vaatluspunkti Z-skoori väärtus,

$x_1$  – käibekapital / koguvara,

$x_2$  – eelmiste perioodide jaotamata kasum / koguvara,

$x_3$  – ärikasum / koguvara,

$x_4$  – omakapitali turuväärtus / kohustised,

$x_5$  – müügitulu / koguvara.

Kui 1980-ndate aastate lõpuni osutus pankrotiuuringutes mitmemõõtmeline diskriminantanalüüs domineerivaimaks modelleerimisel rakendatavaks meetodiks, siis pärast 1980-ndaid vähenes selle kasutus oluliselt (Dimitras *et al.* 1996, 502). Eelkõige piiravad mitmemõõtmelise diskriminantanalüüsi rakendamist võrdlemisi ranged meetodilised eeldused. Peamiseks piiranguks on teiste hulgas see, et meetod eeldab analüüsi kaasatavate sõltumatute muutujate lõikes andmete normaaljaotust, millele ettevõtete finantsandmed tegelikkuses pahatihti ei vasta. Samuti eeldab meetod analüüsi kaasatud muutujate täielikku sõltumatust üksteisest. Ometi esineb erinevate finantsnäitajate vahel nõrgemaid või tugevamaid seoseid. Lisaks seab mitmemõõtmeline diskriminantanalüüs sarnaselt ühemõõtmelisele analüüsimeetodile pankroti modelleerimisel eelduse, et ettevõtte finantsseisundi ning seda kajastavate finantsnäitajate vaheline seos on lineaarselt kirjeldatav (Balcaen, Ooghe 2004, 10–13). Rangetest eelduslikest piirangutest hoolimata on mitmemõõtmeline diskriminantanalüüs tänaseni laialdast rakendust

leidev meetod pankrotianalüüside raames, tihtipeale ka võrdlusmeetodina teiste uuringus rakendatavate peamiste meetodite kõrval (Chen 2011a, 4515).

Ületamaks mitmemõõtmelise diskriminantanalüüsi seotud eelkirjeldatud eelduslikke piiranguid, võttis Ohlson (1980) ühena esimestest pankroti modelleerimisel kasutusele logistilise regressiooni. Analüüsile seatud eelduste mõistes on logistilise regressiooni puhul tegu oluliselt vähenõudlikuma meetodiga. Logistilise regressiooni eeliseks mitmemõõtmelise diskriminantanalüüsi ees on esiteks see, et logistiline regressioon eeldab muutujate lõikes andmete normaaljaotuse asemel andmete logistilist jaotust, mis sobitub paremini pankroti modelleerimisel kasutatavate andmete jaotusega. Teiseks on logistilise regressiooni sisuliseks eeliseks see, et meetod ei eelda ettevõtte seisundi ning finantsnäitajate vahel eksisteeriva suhte lineaarsust. Lisaks on võimalik logistilise regressiooni meetodi rakendamisel modelleerimisse kaasata ka kvalitatiivseid kategoriseerivaid muutujaid (Balcaen, Ooghe 2004, 15–16). Ka tulemuste tõlgendamine on logistilise regressiooni puhul võrreldes diskriminantanalüüsiga lihtsam, kuivõrd logistilisel regressioonil põhineva mudeli väljundi näol on tegemist uuritava sündmuse esinemise tõenäosusega (Ohlson 1980, 117).

Kuigi logistiline regressioon on võrreldes diskriminantanalüüsiga oluliselt paindlikum analüüsile seatud eelduslike tingimuste osas (Chen 2011a, 4515), siis on sellel meetodil ka mõningaid nõrkusi. Esiteks on logistiline regressioon äärmiselt tundlik valimis esinevate ekstreemsete ja ka puuduolevate väärtuste suhtes. Taoliste väärtuste esinemine analüüsi valimis võib logistilise regressiooni mudeli klassifikatsioonivõimekust olulisel määral vähendada (Balcaen, Ooghe 2004, 15–16). Lisaks on logistiline regressioon väga tundlik sõltumatute muutujate vahel esineva multikollineaarsuse suhtes (Doumpos, Zopounidis 1999, 81).

Käesoleva magistr töö autor leiab, et eelkirjeldatud statistilistest analüüsimeetoditest sobib logistiline regressioon kõige paremini pankroti modelleerimiseks. Seda eelkõige põhjusel, et logistiline regressioon ei eelda ettevõtte pankrotiohu ning finantsnäitajate vahelist lineaarsust. Lisaks räägib eelmainitud meetodi kasuks ka see, et analüüsi teostajal on võimalik logistilist regressioonanalüüsi negatiivselt mõjutavaid tegureid (ekstreemsed ja puuduolevad väärtused, multikollineaarsus) teatud ulatuseni kontrollida, vähendades antud tegurite poolt modelleerimise tulemustele avalduvat negatiivset mõju.

### 1.4.2. Masinõppe meetodid

Statistiliste meetodite rakendamisega seotud ranged eelduslikud piirangud ning ka kategoriseerivate muutujate modelleerimisse kaasamise kasvav vajadus andsid omalt poolt tõuke aga võimekate masinõppel põhinevate meetodite arengule. Masinõppe meetoditest on pankroti modelleerimisel seni kõige laialdasemat kasutust leidnud tehislake närvivõrkude (*artificial neural networks*) ning otsustuspuu (*decision tree*) meetod (Chen 2011a, 4516).

Tehislake närvivõrkude meetod on klassifitseerimismudelite väljatöötamisel seni palju rakendust leidnud. Pankroti ennustamisel on tehislake närvivõrkude meetodit rakendatud teadaolevalt juba 1990-ndate aastate algusest saadik (Chen *et al.* 2011, 136). Närvivõrkude meetodi näol on tegu bioloogiliste närvivõrkude struktuurist ning funktsioonist inspireeritud analüüsimeetodiga. Närvivõrkude meetod on võimeline modelleerima väga keerulisi mittelineaarseid seoseid sõltuva ning sõltumatute muutujate vahel, mis on ka antud meetodi üheks peamiseks tugevuseks võrreldes paljude teiste analüüsimeetoditega (Olson *et al.* 2012, 466).

Lisaks sellele, et tehislake närvivõrkude meetod suudab edukalt modelleerida muutujate vahel esinevaid seoseid, ei eelda see ka kirjeldatava suhte lineaarsust ega ka muutujate vahelist sõltumatust (Chen *et al.* 2011). Boritz *et al.* 1995 võrdles närvivõrkude meetodit mitme erineva statistilise meetodiga (mh diskriminantanalüüsiga, logistilise regressiooniga) ning leidis, et närvivõrkude meetodil põhinev mudel ületas klassifitseerimisvõimekuselt kõiki võrdlusaluseid statistilisi mudeleid.

Tehislake närvivõrkude meetodit puudutavate kitsaskohtadena on kirjanduses esile tõstetud närvivõrkude mudeli tõlgendamise keerukus ning aeganõudvus optimaalse mudeli konstrueerimisel. Optimaalse mudeli väljatöötamine on aeganõudev ning keeruline eelkõige seetõttu, et närvivõrkude meetodi rakendamisel tuleb analüüsi läbiviijal määratleda hulga parameetreid (nt peidetud kihtide arv, peidetud neuronite arv, õppimismäär, ülekande funktsioon jm), millest sõltub nii mudeli ülesehitus kui ka klassifitseerimisvõimekus (Shin *et al.* 2005, 130). Mudeli tõlgendamise keerukus tuleneb sellest, et närvivõrkude meetod on ülesehituslikult „musta kasti“ meetod, mille analüüsiprotsessid ei ole lihtsasti kirjeldatavad (Li, Wang 2014, 78).

Otsustuspuu meetodi puhul on tegu masinõppe meetodiga, mis põhineb rekursiivsel andmete klassifitseerimise protsessil (Gepp, Kumar 2015, 398). Struktuurilt meenutab otsustuspuu

meetodil põhinev mudel tagurpidi puud. Sõltumatutele muutujatele tuginedes jaotab otsustuspuu meetod analüüsi kaasatud koguvalimi sarnaste tunnustega väiksemateks gruppideks (Berry, Linoff 1997, 166). Otsustuspuu meetodit rakendati pankrotianalüüsis juba 1980-ndatel aastatel (Gepp, Kumar 2015, 397), esimeste hulgas Frydman *et al.* (1985) poolt, kes töö tulemusena väitsid, et otsustuspuu meetodil põhinev pankrotimudel ületab klassifitseerimisvõimekuselt diskriminantanalüüsil põhinevat mudelit. Seda väidet toetas ka aastakümneid hiljem Gepp *et al.* (2010) teostatud uurimustöö tulemus. Joos *et al.* (1998) võrdlesid otsustuspuu meetodit logistilise regressiooniga ning leidsid, et hoopis viimane demonstreeris mudeli üldise klassifitseerimisvõimekuse osas mõnevõrra paremaid tulemusi.

Võrreldes närvivõrkude meetodiga on otsustuspuu meetodi ühe eelisena kirjanduses ära märgitud see, et otsustuspuu meetodil põhinevate mudelite skaleeritavus on mõnevõrra kõrgem. Lisaks suudab otsustuspuu meetod modelleerimisel edukalt toime tulla väiksemamahuliste valimitega, mis muudab antud meetodi rakendamise pankroti modelleerimisel närvivõrkude meetodiga võrreldes praktilisemaks. Seda eelkõige spetsiifilisemate pankrotimudelite konstrueerimisel. (Markham *et al.* 2010, 244–245)

Otsustuspuu meetodil põhineva mudeli suurim tugevus peitub aga selle lihtsuses. Mudel on struktuurilt selgesti mõistetav ning klassifitseerimisprotsess mudeli rakendamisel lihtsasti tõlgendatav. Võrreldes statistiliste meetoditega ei sea otsustuspuu meetod analüüsitavaatele andmetele statistilisi eelduseid ning muuhulgas suudab meetod hästi toime tulla ka valimis esinevate ekstreemsete ning puuduolevate väärtustega (Gepp, Kumar 2015, 398). Ometi esineb meetodil siiski ka mõningaid puuduseid. Suurimaks ohuks peetakse otsustuspuu meetodi rakendamisel mudeli ülekohandamist (*over fitting*). Mudeli ülekohandamine viitab olukorrale, kus mudel on ülimalt täpne klassifitseerima vaatluspunkte treeningvalimi lõikes, kuid rakendades mudelit kontrollvalimil, osutub mudeli klassifitseerimisvõimekus oluliselt madalamaks. Seetõttu on otstarbekas otsustuspuu mudeli treenimisel lähtuda klassifitseerimistäpsuse maksimeerimisest kontrollvalimil, mitte treeningvalimil (Serrano-Cinca, Gutiérrez-Nieto 2013, 1246).

Nagu kirjandusest selgub, on nii tehnilike närvivõrkude kui ka otsustuspuu meetod võimekad abivahendid keeruliste klassifitseerimisprobleemide lahendamisel. Ometi on pankrotimudelite juures äärmiselt oluliseks väärtuseks ka nende tõlgendatavuse ning rakendatavuse lihtsus. Sellest tulenevalt osutub magistritöö autori hinnangul otsustuspuu meetod eelkirjeldatud kahest masinõppe meetodist praktilisemaks abivahendiks pankroti modelleerimisel.



## **2. PANKROTID EESTI TÖÖTLEVAS TÖÖSTUSES NING ANDMETE KAASAMINE MODELLEERIMISSE**

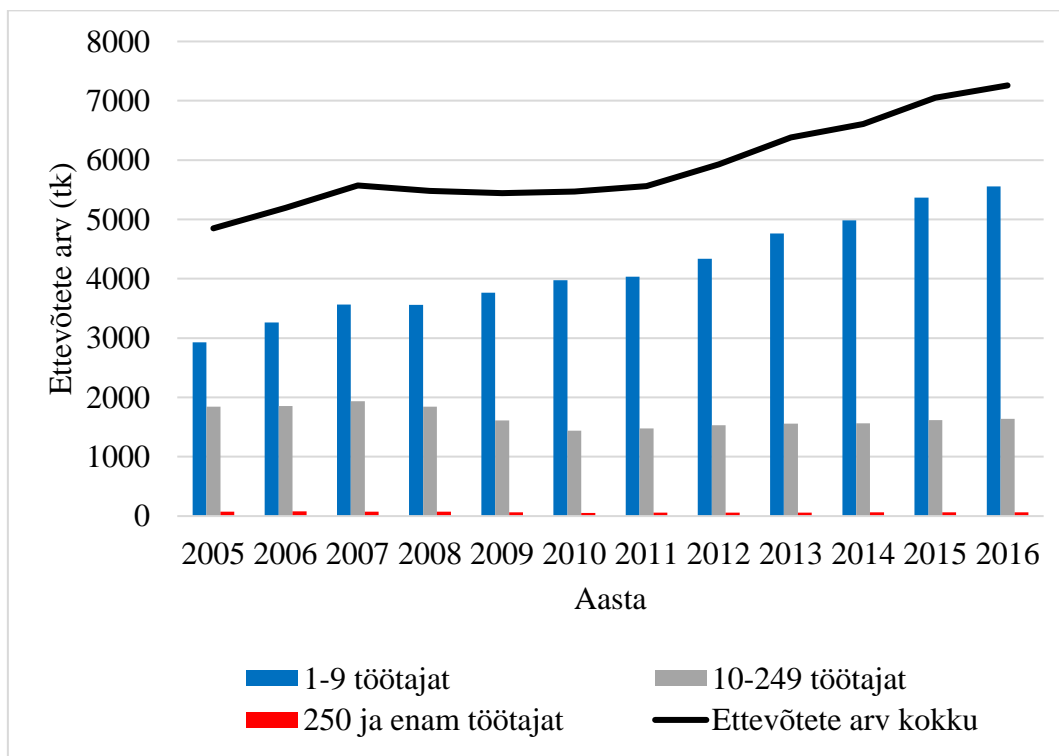
### **2.1. Ülevaade ettevõtete pankrottide sagedusest Eesti töötlevas tööstuses**

Eesti töötlev tööstus on tegevusharu, mis tegeleb materjalide, ainete või komponentide töötlemisega uuteks toodeteks. Töötleva tööstuse EMTAK tegevusharu klassifikatsiooniks on C, mille alla kuulub omakorda 24 tegevusvalkonda (C10–C33) (E-äriregister 2008). 2016. aastal ulatus töötlevas tööstuses tegutsevate ettevõtete arv 7259 ettevõteteni (kasv võrreldes eelneva aastaga 2,9%) (Eesti statistikaamet, tabel EM001). Ühtekokku moodustas tegevusharu toodangumaht 2016. aastal jooksevhindades 11,24 miljardit eurot, millesse tegevusvaldkondadest suurima panuse andsid puidutöötlemine ja puittoodete tootmine (C16), arvutite, elektroonika- ja optikaseadmete tootmine (C26) ning toiduainete tootmine (C10) (Eesti statistikaamet, tabel EM008).

Töötlev tööstus tervikuna panustab Eesti sisemajanduse koguprodukti koondväärtusesse tegevusharudest kõige rohkem (2016. a 15,7%), osutudes tegevusharude lõikes muuhulgas ka suurimaks tööandjaks (Eesti statistikaamet, tabel RAA0042). Töötajate keskmine arv (taandatud täistööajale) 2016. aastal ulatus kõnealuses tegevusharus 107 052 töötajani, mis moodustas koguni 24,7% kogu majanduse tööhõivest (Eesti statistikaamet, tabel EM001). Eesti töötleva tööstuse näol on Eesti majanduse stabiilsuse ning jätkusuutlikkuse toetamise kontekstis seega tegemist väga olulise tegevusharuga

Eesti töötleva tööstuse ettevõtete arv perioodi 2005–2016 lõikes on illustreeritud joonisel 2.1, millel on kajastatud tegevusharu ettevõtete arv ka töötajate arvu järgi jaotatuna (statistika vastava aasta lõpuseisuga). Joonisel toodud ettevõtete jaotuses on lähtutud Euroopa Komisjoni (Komisjoni Määrus 651/2014 I lisa art 2) sätestatud töötajate arvu põhisest kriteeriumist, eristamaks mikroettevõtteid, väikeseid ning keskmiseid ettevõtteid ja suurettevõtteid (sätestatud müügitulu ning vara mahuga seotud kriteeriume seejuures joonisel esitletud jaotuses rakendatud ei ole).

Jättes kõrvale möödunud kriisist tingitud ettevõtete arvu mõningase languse (2007–2009 löikes muutus -2,4%), on vaatlusaluse perioodil Eesti töötleva tööstuse tegutsevate ettevõtete arv kasvanud 2409 võrra (kogumuutus 49,7%). Ootuspäraselt moodustavad keskmiselt 1–9 (lepingulist) töötajaga ettevõtted tegevusharust arvuliselt olulisima osa, kuivõrd 2016. aasta lõpuseisuga moodustas 1–9 (lepingulist) töötajaga ettevõtete arv kogu Eesti töötleva tööstuse tegutsevatest ettevõtetest arvuliselt 76,6% (5558 ettevõtet). Seejuures 10–249 töötajaga ettevõtted moodustasid 2016. aasta lõpuseisuga 22,6% (1640 ettevõtet) ning 250 ning enama töötajaga ettevõtted pelgalt 0,8% (61 ettevõtet) tegevusharu tegutsevatest ettevõtetest. Silmatorkava fenomenina võib joonisel 2.1 illustreeritud statistika põhjal täheldada, et kuigi vaatlusperioodi löikes on majandusharu ettevõtete arv jõudsalt kasvanud, siis seda kasvu on toetanud nimelt 1–9 töötajaga ettevõtted, kuivõrd 10–249 töötajaga ja 250 ning enama töötajaga ettevõtete arv on 2005–2016 perioodi löikes koguni kahanenud, muutused koondperioodi löikes vastavalt -11,1% ning -18,7%.



Joonis 2.1. Ettevõtete arv Eesti töötlevas tööstuses 2005–2016 (aastalõpu seisuga)  
Allikas: Eesti Statistikaamet (tabel EM001); autori koostatud

Vaadeldes pankrotide statistikat tegevusharude löikes, selgub Kredidiinfo 2016. aasta pankrotiuuringu tulemustest, et Eesti töötlev tööstus kvalifitseerub pankrotide sageduse alusel

kolmandale kohale, seda vahetult majutuse ja toitlustuse ning ehituse tegevusharude järel. Kui keskmiselt pankrotistus 2016. aastal tuhande Eestis registreeritud ettevõtte kohta 1,58 ettevõtet, siis kõnealuse uuringu andmetel ulatus töötlevasse tööstusesse registreeritud ettevõtete puhul antud määr koguni 3,19 tasemele (Creditinfo Eesti AS 2016, 11).

Järgnevas tabelis (vt tabel 2.1) esitatud statistika põhjal võib täheldada, et pankrotijuhtude arv nii töötlevas tööstuses kui ka tegevusharude lõikes üldiselt kulmineerus 2009. aastal, seda tingituna ülemaailmsest majandussurumisest. Töötlevas tööstuses ulatus pankrotistunud ettevõtete arv 2009. aastal koguni 149 ettevõteteni, pankrottide arvu kasv võrreldes eelneva aastaga ulatus seejuures märkimisväärsele 116% tasemele. Pärast 2009. aastat on pankrottide arv töötlevas tööstuses järjepidevalt vähenenud. 2015. ning ka 2016. aastal pankrotistus töötlevas tööstuses kokku 42 ettevõtet, mis võrreldes 2009. aasta tipuga (149 ettevõtet) oli koguni 72% võrra vähem. Sellegipoolest on Eesti töötleva tööstuse näol tegemist tegevusharuga, kus pankrottide sagedus ja üldistatult ka pankrotistumise oht osutub võrreldes tegevusharude keskmisega poole kõrgemaks.

Tabel 2.1. Pankrotistunud ettevõtete arv 2008–2016 suurima pankrotisagedusega tegevusharude lõikes (2016. a seisuga)

Tegevusharu	2008	2009	2010	2011	2012	2013	2014	2015	2016	Pankrote per 1000 (2016. a)
Majutus ja toitlustus	15	58	59	42	27	36	31	27	20	3,85
Ehitus	68	269	222	121	86	79	64	71	70	3,63
Töötlev tööstus	69	149	126	63	46	50	55	42	42	3,19
Tegevusharud kokku	423	1055	1029	623	495	459	428	376	335	1,58

Allikas: Creditinfo Eesti AS (2016, 8)

Töötleval tööstusel on Eesti majanduse stabiilsuse ning jätkusuutlikkuse toetamisel äärmiselt oluline roll. Seejuures on Eesti töötleva tööstuse näol tegu majandusharuga, kus viimaste andmete kohaselt osutub ettevõtete pankrottide sagedusmäär võrreldes majandusharude keskmisega poole kõrgemaks, mis muudab seda olulisemaks kõnealuse tegevusharu alla kuuluvate ettevõtete majandusliku tervise perioodilise monitoorimise. Eeltoodud statistika toetab tugevalt vajadust usaldusväärse ning ajakohase prognoosimudeli järele, mis oleks võimeline andma erinevatele osapooltele võimalikult täpset ning ajas ette vaatavat informatsiooni huvialuste Eesti töötleva tööstuse ettevõtete finantstervise kohta.

## 2.2. Analüüsi valitud algandmed ning muutujad

### 2.2.1. Algandmed ning nende kaasamisel rakendatud piirangud

Esimene samm Eesti töötleva tööstuse pankrotimudeli konstrueerimisel on algandmete kogumine. Seda nii pankrotistunud kui ka jätkuvalt tegutsevate töötleva tööstuse ettevõtete kohta. Pankrotimudeli konstrueerimisse kaasatud algandmed on kogutud Justiitsministeeriumi poolt hallatavast Registrite ja Infosüsteemide Keskusest (RIK). Teiste hulgas haldab RIK ka Äriregistri andmesüsteemi, mis sisaldab kõigi Eestis registreeritud ettevõtete tegevusandmeid ning majandusaasta aruandeid.

Esimeses järjekorras kogus töö autor andmeid töötleva tööstuse pankrotistunud ettevõtete kohta, mille osas perioodil 2005–2016 kuulutati kohtulikult välja pankrot või lõpetati pankrotimenetlus pankroti raugemisega. Modelleerimisse kaasati aga vaid need perioodil 2005–2016 pankrotistunud ettevõtted, mis rahuldasiid kõiki järgnevaid piiranguid:

1. Ettevõtte on enne pankrotimenetluse lõppemist tegutsenud minimaalselt 3 aastat. Piirangu eesmärgiks on vähendada tõenäosust käivitusfaasis olevate ettevõtete andmete kasutamist pankrotimudeli koostamisel. Autori hinnangul on käivitusfaasis oleva ettevõtte riskiprofiil sootuks erinev ning säärase ettevõtte puhul ei kajastu peamised pankrotiohtu väljendavad riskid tingimata kvantitatiivselt mõõdetavates finantsnäitajates. Taolise profiiliga ettevõtete andmete kasutamine modelleerimisel võib aga negatiivselt kallutada finantsnäitajatel põhineva pankrotimudeli võimekust.
2. Ettevõtte on Äriregistrile esitanud tervikliku majandusaasta aruande, milles kajastatud majandusaasta lõpuseisu ning pankrotimenetluse lõppemise vaheline periood jääb ajavahemikku 6–18 kuud. Antud piirangu eesmärgiks on konkreetselt määratleda modelleerimisel kasutatavate andmete ajaline suhe pankrotistumise hetke (antud töö analüüsis käsitletud kui pankrotimenetluse lõppemise kuupäev). Piirangu toel tehakse käesoleva töö raames üldistus, et pankrotimudeli koostamisel kasutatakse (keskmiselt)  $t - 1$  aasta finantsandmeid, kus  $t$  märgib ettevõtte pankrotimenetluse lõppemist (pankroti väljakuulutamise või raugemine).
3. Eeltoodud punktis viidatud ettevõtte majandusaasta aruande kohaselt ulatus ettevõtte keskmine töötajate arv (taandatuna täistööajale) vahemikku 10–249 ning ettevõtte tegutses antud majandusaastal töötleva tööstuse alla klassifitseerivas (EMTAK tegevusharu jagu C) tegevusvaldkonnas. Piirangu eesmärk on kaasata modelleerimisse vaid töötleva tööstuse ettevõtted, mis Euroopa Komisjoni (Komisjoni Määrus 651/2014 I lisa art 2)

poolt sätestatud töötajate arvu põhisele kriteeriumile toetudes kvalifitseeruvad seejuures väikeste ja keskmiste ettevõtete alla (vara ja müügitulu mahu osas sätestatud kriteeriumid ettevõtete klassifitseerimisel suuruse järgi jäetakse antud piirangu juures kõrvale). Kõnealuse piirangu läbi eemaldatakse modelleerimisse kaasatavast valimist töötajate arvu kriteeriumi põhiselt mikro- ja suurettevõteteks klassifitseeruvad ettevõtted (Euroopa Komisjoni definitsiooni kohaselt), ühtlustades seeläbi valimi homogeensust ning võrdlusbaasi. Kui mikroettevõtete finantsnäitajad on autori hinnangul ajas keskmisest oluliselt volatiilsemad (üksikutest tehingutest kergesti mõjutatavad) ning antud ettevõtete aruandluses kajastuv finantsseis ei pruugi osutada kõige usaldusväärsemaks, siis suurettevõtete puhul ei ole tulemused lihtsalt valdavale enamusele Eesti töötleva tööstuse ettevõtetest ülekantavad.

Kuivõrd modelleerimisse kaasati ettevõtete  $t - 1$  aasta finantsandmed, siis on käesoleva magistritöö fookusesse võetud ettevõtete finantsseisundi ning pankrotistumise ajaline suhe (keskmiselt) üheaastases ette vaatavas perspektiivis. Kaugema tuleviku prognoosimisel osutub pankrotimudelite klassifitseerimistäpsus kehvemaks (Onofrei, Lupu 2014), seda eelkõige tingituna pikema ajalise perioodiga seotud kõrgemast ebamäärasusest (suhtena pankrotistumisse), ning mudelite usaldusväärsus kujuneb seetõttu madalamaks.

Rakendatud piirangud töötleva tööstuse pankrotistunud ettevõtete algvalimisse kaasamisel küll vähendavad modelleerimisel kasutatavat andmehulka, kuid autori hinnangul kokkuvõttes parendavad töö raames konstrueeritud pankrotimudelite usaldusväärsust ning seega ka praktilisust. Kokku rahuldab kõiki andmete kogumisel seatud piiranguid 129 perioodil 2005–2016 pankrotistunud töötleva tööstuse ettevõtet.

Teises järjekorras kaasati modelleerimisse koguvalimi moodustamiseks juhuslikul valikul jätkuvalt tegutsevate töötleva tööstuse ettevõtete finantsandmed. Selleks rakendati käesolevas töös sobitatud paaride meetodit (*matched-pairs method*), millest lähtuvalt iga eelnevas etapis valimisse kaasatud pankrotistunud ettevõtte finantsandmed sobitati paari jätkuvalt tegutseva töötleva tööstuse ettevõtte finantsandmetega, seda kattuva majandusaasta põhjal. Oluline on siinkohal ära märkida, et käesoleva töö analüüsi raames defineeritakse ettevõtte jätkuvalt tegutsevaks ainult juhul, kui ettevõtte on esitanud Äriregistrile tervikliku 2016. majandusaasta aruande. Juhul, kui ettevõtte tegevusstaatatus on Äriregistri andmeil küll aktiivne, kuid ettevõtte pole 2016. majandusaasta aruannet Äriregistrile esitanud (seda põhjusest olenemata), välistatakse

selle ettevõtte finantsandmete võimalik algvalimisse sattumine. Muus osas rakendati jätkuvalt tegutsevate ettevõtete andmete kaasamisel samasisulisi kriteeriume võrreldes pankrotistunud ettevõtete andmete kaasamisel rakendatud piirangutega – minimaalselt 3 aastat tegevust, 10–249 töötajat ning ettevõtte EMTAK tegevusharu jagu C (töötlev tööstus) klassifikatsioon.

Kokku moodustati kirjeldatud metoodikale põhinedes Eesti töötleva tööstuse pankroti modelleerimise tarbeks algvalim 258 eraldiseisva vaatluspunktiga, mis koosnes 129 pankrotistunud ning jätkuvalt tegutseva ettevõtte finantsandmete paarist, seda kattuvate majandusaastate põhjal (Registrite ja Infosüsteemide Keskus, tabel Algandmed).

Tabel 2.2. Algvalimisse kaasatud ettevõtete arv töötleva tööstuse tegevusvaldkondade lõikes

Töötleva tööstuse tegevusvaldkonnad		Ettevõtteid algvalimis	
		Pankrotis (n=129)	Jätakuvalt tegutsev (n=129)
Toiduainete tootmine	C10	13	10
Joogitootmine	C11	3	0
Tubakatoodete tootmine	C12	0	0
Tekstiilitootmine	C13	4	7
Rõivatootmine	C14	14	13
Nahatöötlemine ja nahktoodete tootmine	C15	2	3
Puidutöötlemine ning puit- ja korktoodete tootmine, v.a mööbel	C16	22	14
Paberi ja pabertoodete tootmine	C17	1	4
Trükindus ja salvestiste paljundus	C18	2	8
Koksi ja puhastatud naftatoodete tootmine	C19	0	0
Kemikaalide ja keemiatoodete tootmine	C20	3	1
Põhifarmaatsiatoodete ja ravimpreparaatide tootmine	C21	0	0
Kummi- ja plasttoodete tootmine	C22	7	8
Muude mittemetalletest mineraalidest toodete tootmine	C23	5	5
Metallitootmine	C24	1	0
Metalltoodete tootmine, v.a masinad ja seadmed	C25	24	20
Arvutite, elektroonika- ja optikaseadmete tootmine	C26	0	5
Elektriseadmete tootmine	C27	1	4
Mujal liigitamata masinate ja seadmete tootmine	C28	2	8
Mootorsõidukite, haagiste ja poolhaagiste tootmine	C29	1	1
Muude transpordivahendite tootmine	C30	1	1
Mööblitootmine	C31	18	10
Muu tootmine	C32	2	5
Masinate ja seadmete remont ja paigaldus	C33	3	2

Allikas: autori arvutused

Algvalimisse kaasati nii osäühinguid kui ka aktsiaseltse, kuivõrd konkreetset ettevõtlusvormi piiramist äriühingute raames ei pidanud autor analüüsivalimi koostamisel vajalikuks ega ka otstarbekaks. Samuti ei pidanud autor vajalikuks analüüsivalimi koostamisel ettevõtteid paaridesse sobitada konkreetse töötleva tööstuse tegevusvaldkonna alusel. Sellise parameetri kontrollimine oleks autori hinnangul valimi koostamisel esinevat juhuslikkust tuntavalt vähendanud, lisamata seejuures olulisel määral lisandväärtust konstrueeritud pankrotimudelite üldisele kvaliteedile.

Enim pärineb algvalimisse kaasatud töötleva tööstuse ettevõtteid (vt tabel 2.2) tegevusvaldkondadest metalltoodete tootmine, v.a masinad ja seadmed (kokku 44 vaatluspunkti); puidutöötlemine ning puit- ja korktoodete tootmine, v.a mööbel (36 vaatluspunkti) ning mööblitootmine (28 vaatluspunkti). Seevastu ei ole algvalimis esindatud ühtegi tubakatoodete, koksi ja puhastatud naftatoodete ning põhifarmaatsiatoodete ja ravimipreparaatide tootmise ettevõtet.

### **2.2.2. Algvalikus analüüsi kaasatud sõltumatud muutujad**

Modelleerimisse kaasatavate sõltumatute muutujate valikul lähtus autor järgnevatest printsiipidest. Esiteks sellest, et valitud muutujate näol oleksid kaetud mitmesugused majandustegevust kajastavad mõõdetavad valdkonnad, nii ettevõtte rentaablus, likviidsus ja maksevõime, kapitali struktuur, laenu teenindamise võimekus ning tegevuse efektiivsus. Teiseks arvestati muutujate valikul kaasatavate finantsnäitajate senist kajastust teemakohases kirjanduses (põhinedes eelkõige tabelis 1.1 esitatud teabele). Kolmandaks toetus autor muutujate valikul ka enda senisele krediidianalüütilisele oskusteabele.

Kuigi paljudes ettevõtete pankrotistumist käsitlevates uurimustöodes on modelleerimisse kaasatud ka kvantitatiivselt mõõdetavad ettevõtte suurus- ning küpsusnäitajaid (näiteks varade või müügitulu maht, töötajate arv ning ettevõtte vanus), siis käesoleva töö modelleerimisse kirjeldatud näitajaid ei kaasata. Seda põhjusel, et käesolevas töös on algvalimi koostamisel rakendatud erinevaid piiranguid (vaata peatükki 2.2.1), mis autori hinnangul vähendavad teatavas suuruses (väga väikesed ning suured) või vanuses (kävitusfaasis) ettevõtete modelleerimisse sattumise tõenäosust oluliselt. Seetõttu ei pea autor otstarbekaks kaasata ettevõtte suurust ning tegevusküpsust kajastavaid näitajaid modelleerimisse veel eraldi sõltumatute muutujatena.

Kirjanduses esineb ka näiteid, kus pankroti modelleerimisse on kaasatud muuhulgas ka kvalitatiivsed muutujad. Autor tõdeb, et kvalitatiivsed näitajad kätkevad endis ettevõtete finantstervise määramisel olulist lisandväärtust. Ometi on ettevõtteid puudutava kvalitatiivse informatsiooni kättesaadavus praktikas liigselt piiratud. Sellest tulenevalt on käesolevas töös piiratud Eesti töötleva tööstuse pankrotimudeli konstrueerimisel finantssuhtarvudel põhinevate kvantitatiivsete muutujatega. Lähtudes eelkirjeldatud printsiipidest, valiti esialgselt välja 19 modelleerimisse kaasatavat finantssuhtarvu põhist muutujat (vaata tabel 2.3).

Tabel 2.3. Esialgne valik modelleerimisesse kaasatavatest sõltumatutest muutujatest

Valdkond	Muutuja
Rentaablus	puhaskasum / koguvara
	ärikasum / koguvara
	puhaskasum / müügitulu
	ärikasum / müügitulu
	ärikasum / keskmine töötajate arv
Likviidsus ja maksevõime	käibevara / lühiajalised kohustised
	(käibevara - varud) / lühiajalised kohustised
	käibevara / koguvara
	(käibevara - lühiajalised kohustised) / koguvara
Kapitali struktuur	kohustised / omakapital
	laenukohustused / omakapital
	omakapital / koguvara
	laenukohustused / koguvara
Laenu teenindamine	lühiajalised laenukohustused / EBITDA
	laenukohustused / EBITDA
	intressikulu / ärikasum
Efektiivsus	müügitulu / koguvara
	varud / müügitulu
	koguvara / keskmine töötajate arv

Allikas: autori koostatud

Rentaablusnäitajate kaasamine pankrotianalüüsi on oluline seetõttu, et ettevõtte rentaablus ehk tegevuse kasumlikkus enamasti kehveneb, kui ettevõtte liigub pankrotistumise suunas (Ohlson 1980, 119–120). Seetõttu on rentaablusnäitajad osutunud väga efektiivseteks indikaatoriteks pankroti prognoosimisel (Chen *et al.* 2011, 141). Vähem olulist rolli ei oma ettevõtte pankrotiohule viitamisel likviidsusnäitajad ning kapitali struktuuri näitajad (Zmijewski 1984, 76; Ni *et al.* 2014, 8–9). Likviidsusnäitajad kajastavad eelkõige ettevõtte lühiajalist maksevõimekust, kapitali struktuuri näitajad indikeerivad seevastu ettevõtte pikaajalist finantstugevust ning -stabiilsust. Kui pankrotimudeli konstrueerimisel tuginetakse meetodiliselt pankrotistumise



õiguslikule definitsioonile, omavad nii ettevõtte likviidsus kui ka kapitali struktuuri kirjeldavad näitajad pankrotiohule viitamisel veelgi ilmselgemat rolli. Seda põhjusel, et ettevõtte pankrotistumist indikeeriva olukorra kirjeldus õiguslikus mõistes põhineb enamikel juhtudel justnimelt antud valdkonna näitajatel (Balcaen, Ooghe 2004, 21). Senisele krediidianalüütilisele kogemusele põhinedes, kaasab autor analüüsi ka erinevad laenu teenindamist indikeerivad finantsnäitajad, eesmärgiga lisada pankroti modelleerimisse ka mõnevõrra krediidiorienteeritust. Kuivõrd ettevõtte ressursside (seejuures ka inimressursside) ebaefektiivne kasutamine suurendab ettevõtte pankrotistumise ohtu (Ni *et al.* 2014, 9), siis kaasab autor modelleerimisse viimaste hulgas ka ettevõtte tegevuse efektiivsust kajastavad näitajad.

## **2.3. Modelleerimisse kaasatava valimi ettevalmistus ning kirjeldav statistika**

### **2.3.1. Null- ning ekstreemväärtuslike vaatluspunktide eemaldamine algvalimist**

Modelleerimisse kaasatava andmevalimi puhul on oluline selle terviklikkus ning puhtus, kuivõrd teatud analüüsimeetodid, eelkõige statistilised, ei tule hästi toime valimis sisalduvate null- ning ekstreemväärtuslike vaatluspunktidega. Esimeses järgus käsitleti koostatud algvalimis sisalduvaid nullväärtuslikke ehk sisult puuduolevate väärtustega vaatluspunkte, tagamaks modelleerimisse kaasatava andmevalimi terviklikkuse. Arvutuslikus mõistes on kaasatud finantssuhtarvupõhiste muutujate lõikes nullväärtuslikud need vaatluspunktid, mille arvutamisel esineb lugejas või nimetajas väärtus 0. Kuigi väärtuse 0 esinemisel suhtarvu lugejas on suhtarvu väärtus küll matemaatiliselt tuletatav, seejuures igakordselt tulemuseks 0, siis autori hinnangul on pankroti modelleerimise raames nullväärtusliku vaatluspunkti puhul tegu pseudoväärtusega, mis omab modelleerimisse kaasamisel ometi olulist soovimatut mõju. Olulises mahus nullväärtusi sisaldava sõltumatu muutuja kaasamine modelleerimisse kallutab konstrueeritava pankrotimudeli tulemusi ning eeldatavalt vähendab selle kvaliteeti prognoosivõimekuse näol.

Käesoleva töös toimiti algvalimis sisalduvate nullväärtustega selliselt, et vähemal määral kui 5% nullväärtuslike vaatluspunktide esinemise korral konkreetse muutuja lõikes, otsustati asendada nullväärtused antud muutuja keskvärtusega. Seda nimelt vastava kategooria ettevõtete (pankrotistunud või jätkuvalt tegutsevad) vaatluspunktide väärtuste põhjal tuletatud keskvärtusega. Enam kui 5% nullväärtuste esinemise korral ei pidanud autor konkreetse muutuja lõikes nullväärtuste asendamist keskvärtusega otstarbekaks, kuivõrd sellisel juhul oleks tegemist olnud olulises mahus kunstlike väärtuste genereerimisega. Antud toiming oleks

kallutanud konstrueeritud prognoosimudeli tulemusi tegelikkusele mittevastavas suunas ning vähendanud seeläbi mudeli praktilisust.

Kuivõrd kõigi nullväärtusi sisaldavate muutujate löikes ulatus nullväärtuste osatähtsus üle 5% (vaata tabel 2.4), ei leidnud eelmäaratletud metoodikast lähtuvalt ühegi muutuja löikes nullväärtuste asendamine keskväertusega aset ning nullväärtused otsustati algvalimist eemaldada. Efektiivsemaks lahenduseks nullväärtuste algvalimist eemaldamisel, minimeerimaks seejuures tekkivat andmekadu, hindas autor aga nullväärtusi sisaldavate sõltumatute muutujate kogu mahus algvalimist eemaldamist. Seda konkreetsete nullväärtusi sisaldavate vaatluspunktide eemaldamise asemel. Arvutuslikus mõistes eemaldati kõnealuses etapis muutujad, mille arvutamisel esines enam kui 5% juhul vaatluspunktidel lugejas või nimetajas väärtus 0. Ühtekokku eemaldati kuus muutujat, seehulgas kaks kapitali struktuuri, üks ettevõtte tegevuse efektiivsust ning kõik kolm laenu teenindamise võimekust kajastavad muutujad.

Tabel 2.4. Esiialgselt valikust eemaldatud muutujad

Valdkond	Muutuja	Nullväärtuste osakaal algvalimist ( $n = 258$ )
Kapitali struktuur	laenukohustused / omakapital	21,3%
	laenukohustused / koguvara	21,3%
Laenu teenindamine	lühiajalised laenukohustused / EBITDA	23,6%
	laenukohustused / EBITDA	21,3%
	intressikulu / ärikasum	22,9%
Efektiivsus	varud / müügitulu	7,0%

Allikas: autori arvutused

Alternatiivse võimalusena andmevalimi terviklikkuse saavutamisel oleks võinud olulises mahus nullväärtusi sisaldavate muutujate asemel ka nullväärtusi sisaldavad konkreetsete vaatluspunktide algvalimist eemaldada. Sellise lähenemise korral oleks andmete kadu modelleerimisse kaasatud 258 vaatluspunkti löikes ulatunud aga 92 vaatluspunktini (35,7% koguvõimist), mida autor hindas palju olulisemaks andmekaoks. Lisaks toetab valitud lähenemist nullväärtustega toimimisel see, et kuivõrd konstrueeritava pankrotimudeli praktilisus oleneb muuhulgas ka mudeli rakendatavuse ulatusest, siis olulises mahus nullväärtuslike vaatluspunkte sisaldavate muutujate pankrotimudelisse jõudmine vähendaks märgatavalt konstrueeritud pankrotimudeli rakendatavust. Seda põhjusel, et taoline pankrotimudel ei oleks rakendatav sisuliselt ühelgi sellisel huvialusel vaatluspunktis (ehk ettevõtte andmetel), mis antud muutujate löikes sisaldab

nullväärtust ehk majanduslikku pseudoväärtust. Ainuüksi kõnealuse töö raames koostatud algvalimis esineb nullväärtuslikke vaatluspunkte vastavate muutujate lõikes 7,0–23,6% juhtudel.

Teises järgus eemaldati algvalimist mitmemõõtmelised ekstreemväärtuslikud vaatluspunktid (*multivariate outliers*) ehk ekstreemväärtused, tagamaks seejuures modelleerimisse kaasatava andmevalimi puhtuse. Mitmemõõtmeliste ekstreemväärtuste tuvastamist võimaldab teiste seas Mahalanobise kauguse ning Cook'i kauguse mõõdikute rakendamine. Mahalanobise kaugust võib tõlgendada kui vaatluspunkti kaugust konkreetse muutuja jaotuse keskvärtuse suhtes, mille väärtus leitakse sõltumatute muutujate normaliseeritud väärtuste ning nende muutujate korrelatsioonikordajate põhjal (Kumar *et al.* 2010, 2056–2057). Otsustamaks Mahalanobise kauguse väärtuse põhjal, kas konkreetse vaatluspunkti näol on tegu ekstreemväärtusega, võrreldakse konkreetse vaatluspunkti Mahalanobise kauguse väärtust anaüüsi alimisi parameetrite põhjal tuletatud kriitilise väärtusega (*ibid.*, 2057). Mahalanobise kauguse kriitiliseks väärtuseks on *Hii*-ruut jaotuse väärtus, kus olulisuse nivool ( $\alpha$ ) 0,001 vabadusastmete arv (*df*) võrdub analüüsi kaasatud sõltumatute muutujate arvuga. Juhul kui vaatluspunkti Mahalanobise kauguse väärtus ületab tuletatud kriitilist väärtust, on tegu potentsiaalse ekstreemväärtusliku vaatluspunktiga (Tabachnick, Fidell 2007, 74).

Cook'i kauguse mõõdikut rakendades tuvastatakse ekstreemväärtuslikud vaatluspunktid seevastu läbi vähimruutude (*Ordinary Least Squares*) regressiooni, hinnates kaasatud vaatluspunktide eraldiseisvat mõju analüüsi väljundile. Peamine erinevus Cook'i ning Mahalanobise kauguse mõõdikus seisnebki selles, et kui Mahalanobise kauguse mõõtmine vaatluspunktide osas toimub vaid sõltumatute muutujate keskselt, siis Cook'i kaugus kaasab mõõtmisesse lisaks sõltumatutele muutujatele ka vaatlusaluse sõltuva muutuja. Cook'i kaugus hindab nimelt vaatluspunktide eraldiseisvat mõju huvisaluse sõltuva muutuja suhtes (Miles, Shevlin 2000, 96). Mida mõjukam konkreetne vaatluspunkt sõltuva muutuja suhtes on, seda kõrgemaks kujuneb antud vaatluspunkti Cook'i kauguse väärtus. Ka Cook'i kauguse mõõdiku põhjal defineeritakse ekstreemväärtuslikeks need vaatluspunktid, mille arvatud Cook'i väärtus ületab määratud kriitilist piirväärtust. Tuvastamaks ekstreemväärtuslikke vaatluspunkte rakendatakse Cook'i kauguse kriitilise piirväärtusena tihtipeale 1,0 taset (Chatterjee, Hadi 2006, 104). Bollen ja Jackman (1990, 265–266) leidsid aga, et konservatiivsemaks lähenemiseks ekstreemväärtuslike vaatluspunktide tuvastamisel võib rakendada Cook'i kauguse piirväärtusena  $4/n$ , kus  $n$  tähistab analüüsi kaasatud vaatluspunktide arvu.

Kuna analüüsi kaasatud pankrotistunud ning tegutsevate ettevõtete puhul on tegu kahe ettevõttekategoriaga, siis ekstreemväärtuslike vaatluspunktide tuvastamiseks hinnati Mahalanobise ning Cook'i kauguseid pankrotistunud ning jätkuvalt tegutsevate ettevõtete alamvalimite lõikes eraldiseisvalt. Mahalanobise kauguse puhul osutus kriitiliseks väärtuseks 34,528 (tuletatud  $H_{ii}$ -ruut jaotuse põhjal, kus  $\alpha = 0,001$  ning  $df = 13$ ) ning Cook'i kauguse puhul 0,031 (lähtunud konservatiivsemast  $4 / n$  piirväärtusest, kus  $n = 129$  ehk vaatluspunktide arvu eraldiseisvalt pankrotistunud ja jätkuvalt tegutsevate ettevõtete lõikes). Antud töö analüüsi raames defineeriti vaatluspunkt ekstreemväärtuslikuks ainult sellisel juhul, kui antud vaatluspunkti puhul mõõdetud Mahalanobise ning Cook'i kaugus ületas rakendatud mõõdikute osas vastavaid tuletatud kriitilisi piirväärtuseid paralleelselt (samaaegselt).

Rakendades SPSS andmetöötlustarkvara arvutati kõigile algvalimisse kaasatud vaatluspunktile Mahalanobise ning Cook'i kaugus (vt lisa 1), mida võrreldi seejärel tuletatud piirväärtustega. Ühtekokku tuvastati vaatluspunktide lõikes 15 ekstreemväärtuslikku vaatluspunkti (seehulgas 8 pankrotistunud ning 7 jätkuvalt tegutsevat ettevõtet), mis seejärel ka algvalimist eemaldati, toetamaks modelleerimisse kaasatava andmevalimi kvaliteeti võimalikult kõrge puhtuse näol.

### **2.3.2. Modelleerimisse kaasatavate andmete kirjeldav statistika ning jaotus**

Eesti töötleva tööstuse pankrotimudeli konstrueerimisse kaasati olulises mahus nullväärtusi sisaldavate muutujate ning ekstreemväärtuslike vaatluspunktide eemaldamise järel ühtekokku 121 pankrotistunud ning 122 jätkuvalt tegutseva ettevõtte finantsandmed. Need 13 suhtarvupõhise sõltumatu muutuja lõikes. Järgnevas tabelis (vt tabel 2.5) on esitatud modelleerimisse kaasatud pankrotistunud ning järkuvalt tegutsevate töötleva tööstuse ettevõtete (kokku 243 vaatluspunkti) finantsandmete võrdlev kirjeldav statistika muutujate lõikes.

Tabelis 2.5 esitatud andmete kirjeldava statistika põhjal joonistuvad välja teatavad kergesti hoomatavad erinevused modelleerimisse kaasatud pankrotistunud ning jätkuvalt tegutsevate töötleva tööstuse ettevõtete vahel. Esiteks võib tabelis esitatud keskväärtuste põhjal täheldada, et kõigi rentablust kajastavate muutujate ( $x_1, x_2, x_3, x_4$  ning  $x_5$ )  $t - 1$  keskväärtused on pankrotistunud ettevõtete puhul negatiivsed, mis tõlgendatuna tähendab seda, et (keskmiselt) aasta enne pankrotistumist tegutses keskmine pankrotistunud töötleva tööstuse ettevõte (viitab valimi andmete aritmeetilisele keskväärtusele) kahjumlikult. Võrdluseks, tegutsevate ettevõtete puhul osutuvad kõigi kaasatud rentablusnäitajate keskväärtused positiivseks.

Tabel 2.5. Modelleerimisse kaasatud andmete kirjeldav statistika pankrotistunud ning tegutsevate ettevõtete lõikes

Muutujad		Aritmeetiline keskmine		Shapiro-Wilk testi olulisuse tõenäosus		U-testi olulisuse tõenäosus
		Pankrotis ( <i>n</i> = 121)	Tegutsev ( <i>n</i> = 122)	Pankrotis	Tegutsev	
puhaskasum / koguvara	x <sub>1</sub>	-0,306	0,093	0,000	0,280	0,000
ärikasum / koguvara	x <sub>2</sub>	-0,278	0,101	0,000	0,315	0,000
puhaskasum / müügitulu	x <sub>3</sub>	-0,170	0,055	0,000	0,001	0,000
ärikasum / müügitulu	x <sub>4</sub>	-0,150	0,060	0,000	0,004	0,000
ärikasum / keskmine töötajate arv	x <sub>5</sub>	-4701,4	3003,5	0,000	0,000	0,000
käibevara / lühiajalised kohustised	x <sub>6</sub>	0,757	2,679	0,000	0,000	0,000
(käibevara - varud) / lühiajalised kohustised	x <sub>7</sub>	0,406	1,850	0,000	0,000	0,000
käibevara / koguvara	x <sub>8</sub>	0,554	0,523	0,000	0,011	0,471
(käibevara - lühiajalised kohustised) / koguvara	x <sub>9</sub>	-0,392	0,230	0,000	0,047	0,000
kohustised / omakapital	x <sub>10</sub>	3,776	1,258	0,000	0,000	0,010
omakapital / koguvara	x <sub>11</sub>	-0,107	0,600	0,000	0,000	0,000
müügitulu / koguvara	x <sub>12</sub>	2,603	1,813	0,000	0,000	0,038
koguvara / keskmine töötajate arv	x <sub>13</sub>	33182,3	35213,9	0,000	0,000	0,016

Allikas: autori arvutused

Teiseks võib silmatorkava erinevusena välja tuua pankrotistunud ettevõtete oluliselt madalamad *t* - 1 keskväärtused likviidsust ning maksevõimet kajastavate muutujate lõikes (x<sub>6</sub>, x<sub>7</sub> ning x<sub>9</sub>). Ainuke antud valdkonna muutuja, mille keskväärtustes silmatorkav vahe ei esine, on käibevara / koguvara (x<sub>8</sub>). Toetudes kapitali struktuuri muutujate (x<sub>10</sub> ning x<sub>11</sub>) lõikes arvatud *t* - 1 keskväärtustele, võib väita, et keskmise pankrotistunud ettevõtte omakapitali tase nii varade kui ka kohustiste suhtes on kordades madalam. Kaasatud efektiivsust kajastavate muutujate (x<sub>12</sub> ning x<sub>13</sub>) juures võib huvitekitava aspektina välja tuua, et pankrotistunud töötleva tööstuse ettevõtte müügitulu / koguvara (x<sub>12</sub>) keskväärtuse tase ulatub jätkuvalt tegutseva omast kõrgemale, mille näol on tegu ka ainsa üllatava tulemusega modelleerimisse kaasatud muutujate keskväärtuste lõikes.

Kahe eraldiseisva andmekategooria muutujate keskväärtustes esineva erinevuse statistilise olulisuse hindamisel on üheks rakendatavaks vahendiks t-test. Samas on t-testi rakendamise eelduseks andmete normaaljaotus vastava kategooria sõltumatute muutujate lõikes (Urdan 2005,

161). Seetõttu tuli esmalt hinnata statistilist tõenäosust selle osas, et modelleerimisse kaasatud andmed jaotuvad muutujate lõikes vastavalt normaaljaotusele, seda eraldiseisvalt nii pankrotistunud kui ka tegutsevate ettevõtete lõikes.

Andmete normaaljaotuse testimiseks kasutatakse Shapiro-Wilk testi (Goos, Meintrup 2016, 192). SPSS andmetöötlustarkvara rakendamisel hindas autor analüüsi kaasatud andmetel Shapiro-Wilk testi olulisuse tõenäosust. Kuivõrd testi tulemus (vaata tabel 2.5) osutus kõigi muutujate lõikes vähemalt ühe eraldiseisva kategooria (pankrotistunud või tegutsevad ettevõtted) andmete puhul oluliseks (olulisuse nivool  $\alpha = 0,05$ ), siis ei saa modelleerimisse kaasatud andmete jaotuvust vastavalt normaaljaotusele hinnata statistiliselt rahuldatuks. See tähendab, et t-testi rakendamiseks seatud peamine statistiline eeldus pole käesolevate andmete puhul täidetud.

Hindamaks kaasatud muutujate keskväärtustes esineva erinevuse statistilist olulisust rakendatakse SPSS andmetöötlustarkvara läbi t-testi alternatiivina mitteparameetrilist Mann-Whitney U-skoori testi, mis erinevalt t-testist ei eelda andmete jaotuvust vastavalt normaaljaotusele (Urdan 2005, 161). Testi tulemustele põhinedes (vaata tabel 2.5) on pankrotistunud ning jätkuvalt tegutsevate ettevõtete muutujate keskväärtustes esinev erinevus oluline kõigi muutujate puhul peale ühe, milleks on käibevara / koguvara ( $x_8$ ). Ülejäänud 12 muutuja osas võib arvutatud U-testi tulemustele põhinedes väita, et antud suhtarvud sisaldavad teatavat olulist informatsiooni eristamiseks aasta pärast pankrotistuvaid ettevõtteid jätkuvalt tegutsevatest. Siinkohal täidavad aga esitletud U-testi tulemused ning antud tulemuste põhjal tehtavad järeldused siiski vaid informatiivset rolli, kuivõrd modelleerimisel rakendatavatesse analüüsimeetoditesse kaasatakse olenemata U-testi tulemustest kõik eeltoodud 13 muutujat. Seda eesmärgiga tagada modelleerimisel vabadus konstrueerida optimaalne pankrotimudel lähtuvalt kasutatavate analüüsimeetodite meetoodilisest spetsiifikast.

Objektiivsuse tagamiseks mudelite hindamisel, on modelleerimisse kaasatud koguvalim (kokku 243 vaatluspunkti) jagatud treeningvalimiks, mille põhjal toimub analüüsimeetodeid rakendades pankrotimudelite konstrueerimine, ning kontrollvalimiks (vt tabel 2.6), mille peal toimub konstrueeritud pankrotimudelite valideerimine, mis antud töö kontekstis tähendab mudelite prognoosivõimekuse ning stabiilsuse hindamist. Treeningvalimisse kaasati vahemikus 2005–2010, kontrollvalimisse vahemikus 2011–2016 pankrotistunud ettevõtetele vastavad vaatluspunktid ( $t - 1$  andmetega) ning antud vaatluspunktidele eelnevalt paariliseks sobitatud jätkuvalt tegutsevate ettevõtete vaatluspunktid ( $t - 1$  andmetega).

Tabel 2.6. Modelleerimisse kaasatud vaatluspunktide jaotus treening- ning kontrollvalimiks

Parameetrid	Pankrotistunud ettevõtted		Tegutsevad ettevõtted		Kokku
	treening- valim	kontroll- valim	treening- valim	kontroll- valim	
Vaatluspunktide arv	82	39	83	39	243
Pankrotistumise aasta	2005–2010	2011–2016	-		2005–2016
Finantsnäitajate ( $t - 1$ ) aasta modelleerimises	2003–2009	2009–2014	2003–2009	2009–2014	2003–2014

Allikas: autori arvutused

Ajavahemikud treening- ning kontrollvalimi koostamiseks määrati eelkirjeldatult seetõttu, et suurusjärg 2/3 vaatluspunktidest leiaksid rakendust mudelite treenimisel ning 1/3 vaatluspunktidest koostatud mudelite valideerimisel. Sarnast meetodilist ülesehitust pankroti modelleerimisel on rakendatud ka varasemates uurimustöodes (Chen 2011b; Gepp, Kumar 2015). Seeläbi moodustati pankrotimudelite konstrueerimiseks treeningvalim 165 vaatluspunktiga (67,9% koguvalimist) ning mudelite valideerimiseks kontrollvalim 78 vaatluspunktiga (32,1% koguvalimist).

Valimi jaotusel rakendatud meetodika, kus kaasatav koguvalim jaotatakse treening- ning kontrollvalimiks, millesse kaasatud vaatluspunktid pärinevad ajaliselt iseseisvatest ehk mittekattuvatest perioodidest, on kasutust leidnud ka varasemates pankroti modelleerimist puudutavates uurimustöodes (Barniv, Raveh 1989; Lin 2009; Chen *et al.* 2011). Kirjandusest leidub teisalt ka näiteid, kus nii treening- kui ka kontrollvalimisse kaasatakse vaatluspunktid kattuvast perioodist (näiteks Mckee 2003; Tseng ja Hu 2010). Autori hinnangul ühtib käesolevas töös rakendatud valimi jaotuse disain aga oluliselt tugevamalt töö eesmärgiga, milleks on nimelt prognoosimudelite konstrueerimine ning tagada seejuures ka konstrueeritud mudelite võimekuse võimalikult objektiivne hindamine. Kirjeldatud disain pakub vaikumisi ka tugevamat kindlustunnet konstrueeritud pankrotimudelite rakendamisel hilisemas praktikas.

### 3. EESTI TÖÖTLEVA TÖÖSTUSE PANKROTIMUDELI KONSTRUEERIMINE

#### 3.1. Pankrotimudeli konstrueerimine logistilise regressiooni rakendamisel

Logistiline regressioon jaguneb olemuselt binaarseks ning mitmemõõtmeliseks. Binaarset logistilist regressiooni rakendatakse klassifitseerimaks kaheväärtuselisi ehk dihhotoomseid vaatluspunkte, mistõttu sobib meetod ülesehituslikult ka pankrotistuvate ning jätkuvalt toimivate ettevõtete klassifitseerimiseks. Kui sõltumatu muutuja osas eeldab logistiline regressioon kategeoriseerivat väärtust, siis regressiooni kaasatavad sõltumatud muutujad võivad omada pidevat või diskreetset väärtust (Chen 2011a, 4515). Logistiline regressioon on mittelineaarne analüüsimeetod, mille rakendamisel kombineeritakse olulisimad analüüsi kaasatud sõltumatud muutujad tõenäosust kajastavaks logit-skooriks (Balcaen, Ooghe 2004, 14). Mudeli väljundi näol on tegemist uuritava sündmuse esinemise tõenäosusega, mis muudab tulemuste tõlgendamise võrdlemisi lihtsaks (Ohlson, 1980). Logistilise regressiooni pankrotimudeli väljundiks olev väärtus jääb seega alati vahemikku 0–1. Logistilise regressioonimudeli üks võimalikest avaldistest matemaatilisel kujul on järgmine (Balcaen, Ooghe 2004, 14):

$$P(x_i) = \frac{1}{1 + \exp -(B_0 + B_1 \times x_1 + B_2 \times x_2 + \dots + B_n \times x_n)} \quad (2)$$

kus

$P(x_i)$  – vaatluspunkti  $i$  tõenäosus pankrotistuda,

$B_0$  – konstant,

$B_1 \dots B_n$  – muutujate koefitsiendid,

$x_1 \dots x_n$  – sõltumatud muutujad

ehk teisiti avaldatuna:

$$P(x_i) = \frac{1}{1 + \exp -(Y_i)} \quad (3)$$

kus  $Y_i$  on vaatluspunkti logit-skoor.



Logistilise regressiooni mudeli klassifikatsioonivõimekust saab hinnata nii I tüüpi ning II tüüpi vigade esinemise ning üldise klassifikatsioonitäpsuse, samuti ka ROC kõvera (*Receiver Operating Curve*) põhjal. Lisaks võib logistilise regressiooni mudeli võimekuse hindamisel rakendada statistilisi  $R^2$  mõõdikuid (nt Nagelkerke  $R^2$ ), aga ka entroopiaal põhinevaid hinnanguid (Balcaen, Ooghe 2004).

Kuigi logistilise regressioon ei sea kasutatavatele andmetele rangeid, siis nagu eelnevalt mainitud, on üheks logistilist regressiooni negatiivselt mõjutavaks teguriks multikollineaarsus, mille esinemist modelleerimisse kaasatavas andmevalimis tuleb vältida. Multikollineaarsuse esinemine tähendab sisuliselt seda, et analüüsi kaasatud sõltumatud muutujad korreleeruvad teineteisega tugevalt ning kirjeldavad seetõttu sõltuva muutuja varieeruvuses sarnaseid omadusi, loomata seejuures olulist lisandväärtust (Hinton 2014, 271).

Tabel 3.1. Modelleerimisse kaasatud andmete korrelatsioonimaatriks

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
X1	1	-	-	-	-	-	-	-	-	-	-	-	-
X2	0,99	1	-	-	-	-	-	-	-	-	-	-	-
X3	0,85	0,81	1	-	-	-	-	-	-	-	-	-	-
X4	0,85	0,82	0,99	1	-	-	-	-	-	-	-	-	-
X5	0,55	0,57	0,57	0,62	1	-	-	-	-	-	-	-	-
X6	0,29	0,28	0,24	0,25	0,37	1	-	-	-	-	-	-	-
X7	0,27	0,26	0,22	0,22	0,33	0,92	1	-	-	-	-	-	-
X8	-0,10	-0,11	-0,08	-0,08	0,10	0,36	0,29	1	-	-	-	-	-
X9	0,80	0,76	0,76	0,74	0,38	0,51	0,44	0,21	1	-	-	-	-
X10	0,06	0,06	0,15	0,17	0,11	-0,06	-0,06	-0,08	-0,02	1	-	-	-
X11	0,83	0,79	0,80	0,78	0,37	0,44	0,39	-0,09	0,93	-0,01	1	-	-
X12	-0,06	-0,07	0,11	0,11	0,07	-0,10	-0,07	0,38	-0,04	0,11	-0,12	1	-
X13	0,09	0,08	-0,10	-0,12	-0,21	0,05	-0,01	-0,22	0,06	-0,07	0,09	-0,46	1

Allikas: autori arvutused

Märkus: muutujate nimetused esitatud tabelis 2.5

Mida lähemal on muutujate vahel mõõdetud korrelatsioonikordaja absoluutväärtus ühele, seda tugevam on konkreetsete muutujate vaheline seos. Muutujate vahelisse korrelatsioonikordaja väärtusesse  $> 0,7$  tuleb suhtuda kriitiliselt, väärtus  $> 0,9$  on regressioonanalüüsi seisukohast aga lubamatu (Tufféry 2011, 88). Esitatud korrelatsioonimaatriksist (vt tabel 3.1) avaldub, et mõningate analüüsi kaasatud muutujate vaheline korrelatsioonikordaja ületab koguni väärtust 0,9, mis viitab ilmselgele multikollineaarsuse esinemise ohule.

Multikollinearsuse eemaldamiseks on mitmemõõtmelise analüüsi korral rakendatav VIF (*Variable Inflation Factor*) ehk dispersiooni inflatsiooniteguri analüüs. VIF analüüsi läbi leitakse kõigile analüüsi kaasatud muutujatele vastav inflatsioonitegur, mis indikeerib konkreetse muutuja korreleeruvust kõigi teiste analüüsi kaasatud muutujate suhtes. Mida suurem on VIF teguri väärtus, seda tugevamalt antud muutuja teiste muutujatega korreleerub (Bager *et al.* 2017, 35). Multikollinearsuse esinemist tähistava VIF piirmäära osas on kirjanduses erinevaid seisukohti. Ühe rusikareeglina märgitakse kirjanduses  $VIF > 10,0$  kui multikollinearsuse esinemise ohtu (Freund, Littell 2000, 98). O'Brien (2007, 679) kasutab uurimustöös aga mõnevõrra konservatiivsemat  $VIF > 4,0$  piirmäära multikollinearsuse esinemise määratlemisel.

Käesoleva töö logistilise regressiooni analüüsis on multikollinearsuse eemaldamisel lähtunud nimelt O'Brieni (2007) poolt välja pakutud VIF piirmäärast 4,0 (ehk tolerants 0,25). Järgnevalt arvutati modelleerimisse kaasatud treeningvalimi andmete põhjal sõltumatutele muutujatele VIF väärtus ning modelleerimisest eemaldati ükshaaval järk-järgult suurima VIF väärtusega muutuja. Antud protsessi korrati kuni ühegi analüüsivalimisse järele jäänud sõltumatu muutuja VIF väärtus ei ületanud autori poolt määratud 4,0 piirmäära (vt lisa 2). Kokku eemaldati analüüsi kaasatud 13 muutujast ükshaaval ühtekokku 5 muutujat, nende hulgas puhaskasum / koguvara ( $x_1$ ), puhaskasum / müügitulu ( $x_3$ ), ärikasum / müügitulu ( $x_4$ ), käibevara / lühiajalised kohustised ( $x_6$ ), (käibevara - lühiajalised kohustised) / koguvara ( $x_9$ ).

Tabel 3.2. VIF väärtused modelleerimisse kaasatavate muutujate lõikes ( $VIF < 4$  muutujad)

Muutujad		VIF väärtus	Tolerants
ärikasum / koguvara	$x_2$	3,76	0,27
ärikasum / keskmine töötajate arv	$x_5$	1,85	0,54
(käibevara - varud) / lühiajalised kohustised	$x_7$	1,48	0,68
käibevara / koguvara	$x_8$	1,41	0,71
kohustised / omakapital	$x_{10}$	1,05	0,95
omakapital / koguvara	$x_{11}$	3,13	0,32
müügitulu / koguvara	$x_{12}$	1,50	0,67
koguvara / keskmine töötajate arv	$x_{13}$	1,39	0,72

Allikas: autori arvutused

Logistilise regressiooni analüüsimeetodit rakendades seadis autor eesmärgiks konstrueerida kõrgeima klassifitseerimisvõimekusega pankrotimudel (hindamine kontrollvalimi alusel), millesse kaasatud muutujad oleksid kõik statistiliselt olulised (olulisuse nivool 0,05).

Modelleerimisse kaasati sõltumatute muutujatena kõik VIF analüüsi järgselt järele jäänud 8 muutujat (vt tabel 3.2). Sõltuva muutujana kaasati analüüsi kategoriseeriv muutuja tähistamiseks valimisse kaasatud vaatluspunktidel vastavate ettevõtete resultaati ajahetkel  $t$ , kus väärtusega 1 märgiti pankrotistunud ning väärtusega 0 jätkuvalt tegutsevad ettevõtted.

Logistiline regressioonanalüüs viidi läbi SPSS andmetöötlustarkvara rakendamisel. Logistilise regressiooni kasutamisel saavutati parim pankrotimudel rakendades tõepärasuhte (likelihood ratio) põhinevat muutujate tagasisuunas valikut (*backward elimination*). Seejuures konstanti parim konstrueeritud logistilise regressiooni pankrotimudel ei sisaldanud. Autor hindab ilma konstantita pankrotimudelit iseenesest ka praktilisemaks, kui võrd konstanti sisu ei ole pankrotimudelis majanduslikult tõlgendatav. Lisaks ei pruugi määratud konstandi efektiivne väärtus pankrotimudelis püsida ajas kuigivõrd stabiilne, mis vaikimisi vähendab ka pankrotimudeli usaldusväärsust aja möödudes.

Kokku kaasati logistilise regressiooni mudelisse kaheksast võimalikust ( $VIF < 4$  piirmäära) muutujast viis (vt tabel 3.3), millest kõik osutusid statistiliselt oluliseks (olulisuse nivool 0,05). Logistilise regressiooni rakendamisel konstrueeritud kõrgeima klassifitseerimisvõimega pankrotimudeli üldkuju, hindamiseks töötleva tööstuse ettevõtete pankrotiohtu, on võrrandis avaldatuna järgmine:

$$Y_i = -8,19474 \times \frac{\text{ärikasum}}{\text{koguvara}} - 1,27174 \times \frac{(\text{käibevara} - \text{varud})}{\text{lühiajalised kohustised}} - 5,21146 \times \frac{\text{omakapital}}{\text{koguvara}} + 1,00375 \times \frac{\text{müügitulu}}{\text{koguvara}} + 0,00002 \times \frac{\text{koguvara}}{\text{keskmise töötajate arv}} \quad (4)$$

kus  $Y_i$  tähistab vaatluspunkti logit-skoori ehk pankrotistumise ning tegevuse jätkumise šansside suhet vaatlusala ettevõtte pankrotistumise tõenäosuse ( $P(x_i)$ ) hindamiseks avaldises:

$$P(x_i) = \frac{1}{1 + \exp(-Y_i)} \quad (5)$$

Mida suuremaks kujuneb mudeli üldkujulises avaldises  $Y_i$  väärtus, seda suuremaks osutub ka vaatlusala ettevõtte pankrotistumise tõenäosus (keskmiselt) üheaastases (6–18 kuud) perspektiivis.

Tabel 3.3. Logistilise regressiooni pankrotimudelisse kaasatud muutujate kirjeldav statistika

Muutujad		Koefitsient ( <i>B</i> )	Olulisuse tõenäosus ( <i>p</i> )	Šansside suhe <i>Exp(B)</i>	95% usaldusintervall <i>Exp(B)</i> osas	
					Alumine piir	Ülemine piir
ärikasum / koguvara	$x_2$	-8,19474	0,000	0,00028	0,00001	0,01334
(käibevara – varud) / lühiajalised kohustised	$x_7$	-1,27174	0,045	0,28034	0,08068	0,97415
omakapital / koguvara	$x_{11}$	-5,21146	0,000	0,00545	0,00039	0,07680
müügitulu / koguvara	$x_{12}$	1,00375	0,000	2,72850	1,79672	4,14351
koguvara / keskmine töötajate arv	$x_{13}$	0,00002	0,011	1,00002	1,00001	1,00004

Allikas: autori arvutused

Mudeli praktilise rakendatavuse hindamisel on lisaks klassifikatsioonivõimekusele oluline hinnata ka seda, kas konstrueeritud mudel on ka majanduslikult selgitatav. Erinevalt lineaarsest regressioonist ei ole logistilise regressiooni puhul koefitsientide (*B*) märgidele (+/-) tuginedes sõltumatute muutujate mõjusuund ettevõtte pankrotistumise tõenäosuse suhtes otse tõlgendatav. Seda tõlgendust pakub logistilise regressiooni puhul aga muutujate löikes kirjeldatud šansside suhe (*Exp(B)*).

Kuivõrd muutujate ärikasum / koguvara ( $x_2$ ), (käibevara - varud) / lühiajalised kohustised ( $x_7$ ) ning omakapital / koguvara ( $x_{11}$ ) šansside suhte väärtus on väiksem kui 1,0, siis kõnealuste muutujate kõrgem väärtus vähendab pankrotistumise ohtu. Vastupidiselt, muutujate müügitulu / koguvara ( $x_{12}$ ) ning koguvara / keskmine töötajate arv ( $x_{13}$ ) šansside suhte väärtus on suurem kui 1,0 ning kõrgem väärtus antud muutujate puhul suurendab pankrotistumise ohtu. Kuivõrd väärtus 1,0 (niinimetatud šansside suhte nullväärtus) ei jää ühegi mudelisse kaasatud muutuja šansside suhte 95% usaldusintervalli sisse, siis võib šansside suhet kirjeldavaid tulemusi hinnata mudelisse kaasatud muutujate löikes statistiliselt oluliseks (olulisuse nivool 0,05).

Mõnevõrra üllatavaks hindab autor esitatud šansside suhte statistiku tulemustes muutuja müügitulu / koguvara ( $x_{12}$ ) suhet mudeli poolt kirjeldatava pankrotistumise ohuga, mille kohaselt kõnealuse efektiivsusnäitaja kõrgem väärtus suurendab ettevõtte pankrotistumise tõenäosust. Vaadeldes konkreetset konteksti ei ole ometigi tegu majandusloogika vastase tulemusega. Võimalikud majanduslikud selgitused kõnealuse muutuja dünaamika osas konstrueeritud pankrotimudelil on järgnevad:

- Keskmise (valimisse kaasatud) tegevust jätkav ettevõtte investeerib ning kasvatab käibevarade mahtu (suhtena müügitulusse) kõrgemal määral. Seda eesmärgil olla suuteline ära teenindama tuleviku osas oodatavat nõudluse kasvu. Vara mahu kasvatamise tulemusel (suhtena müügitulusse) osutubki keskmise tegutseva ettevõtte müügitulu / koguvara väärtus madalamaks kui seda keskmise pankrotistunud ettevõtte puhul.
- Keskmise (valimisse kaasatud) pankrotistuva ettevõtte varade maht väheneb müügitulu suhtes (eelkõige kahjumliku majandustegevuse tõttu), millest tingituna müügitulu / koguvara osas tuletatav väärtus aga kasvab. Ometi ei väljenda antud olukord efektiivsuse kasvu.

Kui konstrueeritud mudel klassifitseeris vaikumisi määratud löikepunkti (s.o pankrotistumise tõenäosuse piirmäär klassifitseerimaks vaatluspunkte pankrotistunud või jätkuvalt tegutsevate ettevõtete hulka) 0,5 juures korrektselt 87,9% kõigist treeningvalimisse kaasatud 165-st ettevõttest, siis kontrollvalimisse kaasatud 78 ettevõtte löikes ulatus loodud mudeli klassifitseerimisvõimekus lausa 91,0% tasemele (vt tabel 3.4), mida võib hinnata väga kõrgeks kirjeldusvõimeks. I tüüpi viga leidis kontrollvalimisse kaasatud vaatluspunktide klassifitseerimisel aset 5,1% ning II tüüpi viga 12,8% juhtudel.

Alternatiivina võib kirjeldada mudeli klassifitseerimisvõimekust eraldiseisvalt pankrotistunud ning jätkuvalt tegutsevate ettevõtete löikes läbi mudeli spetsiifilisuse (*specificity*) ning tundlikkuse (*sensitivity*). Spetsiifilisus märgib mudeli poolt korrektselt klassifitseeritud jätkuvalt tegutsevaid ettevõtteid ning tundlikkus korrektselt klassifitseeritud pankrotistunud ettevõtteid (Sinha, May 2004). Konstrueeritud logistilise regressiooni pankrotimudeli spetsiifilisus kontrollvalimil hinnatuna ulatus seega 87,2% ja tundlikkus 94,9% tasemele (vaikumisi löikepunkt 0,5).

Vaikumisi määratud löikepunkt 0,5 on pankrotimudelitel puhul teoreetiliseks ideaaliks klassifitseerimaks ettevõtteid pankrotistuvate või jätkuvalt tegutsevate ettevõtete hulka. Ometi ei pruugi antud väärtus praktikas osutada optimaalseks löikepunktiks (löikepunktiks millel maksimeeritakse mudeli klassifitseerimisvõimekus). Käesolevas töös konstrueeritud logistilise regressiooni mudeli puhul saavutati kontrollvalimil hinnatuna kõrgeim klassifitseerimisvõimekus löikepunkti intervallis 0,63–0,72 (keskmise optimaalne murdepunkt 0,675). Optimaalsel löikepunkt 0,5 ulatus mudeli üldine prognoosikorrektus koguni 93,6% tasemele, mida võib hinnata suurepäraseks tulemuseks. Võrreldes löikepunkt 0,5 saavutatud klassifitseerimis-

tulemustega, püsis optimaalsel murdepunktil mudeli I tüüpi vea määr samal tasemel (5,1%), kuid II tüüpi vea määr vähenes 7,1% tasemele (0,5 murdepunktil 12,8%), mida võib hinnata oluliseks paranemiseks.

Tabel 3.4. Logistilise regressiooni rakendamisel koostatud pankrotimudeli klassifitseerimisvõimekus

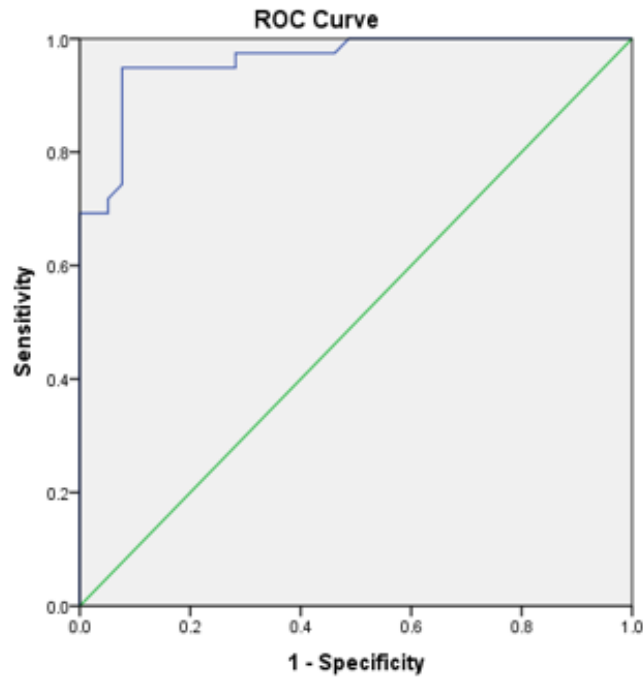
Tegelik tulemus		Prognoositud tulemus treeningvalimis (lõikepunktil 0,5)			Prognoositud tulemus kontrollvalimis (lõikepunktil 0,5)			Prognoositud tulemus kontrollvalimis (lõikepunktil 0,675)		
		Y		Korrektus (%)	Y		Korrektus (%)	Y		Korrektus (%)
		0	1		0	1		0	1	
Y	0	73	10	88,0	34	5	87,2	36	3	92,3
	1	10	72	87,8	2	37	94,9	2	37	94,9
Üldine mudeli korrektsus		87,9			91,0			93,6		

Allikas: autori arvutused

Märkus: väärtus Y = 1 esindab pankrotistunud ettevõtet, Y = 0 jätkuvalt tegutsevat ettevõtet

Mudeli klassifitseerimisvõimekust on võimalik hinnata ka ROC-kõver abil (*receiver operating characteristic curve*), mida on teiste hulgas rakendanud ka näiteks Chava ja Jarrow (2004) ning Agarwal ja Taffler (2008). ROC kõver kajastab mudeli spetsiifilisust ning tundlikkust ehk mudeli võimekust klassifitseerida korrektselt tegevust jätkavaid ning pankrotistunud ettevõtteid igal võimalikul lõikepunktil. Mudeli üldise klassifikatsioonivõimekuse hindamisel tuginetakse eelkõige aga ROC-kõvera alusele pinna suurusele ehk AUC (*area under the curve*) väärtusele (Hosmer, Lemeshow 2004, 160–162).

Maksimaalne AUC väärtus 1 märgib mudeli perfektset klassifitseerimisvõimekust, väärtus 0,5 (mida graafiliselt illustreerib 45° diagonaal ROC-kõvera joonisel) seejuures aga mudeli klassifitseerimisvõimekuse täielikku puudumist. Seega, mida rohkem erineb mudeli AUC väärtus 0,5-st, seda kõrgem on antud mudeli üldine klassifitseerimisvõimekus. Vastavalt Hosmer ja Lemeshow (2004, 160–162) seisukohale märgib AUC väärtus vahemikus 0,7–0,8 aktsepteeritavat, vahemikus 0,8–0,9 head ning väärtus üle 0,9 suurepärase mudeli klassifitseerimisvõimekust. Rakendades SPSS andmetöötlustarkvara sai autor konstrueeritud logistilise regressiooni mudeli ROC-kõvera AUC väärtuseks 0,962, mille põhjal võib hinnata mudeli klassifitseerimisvõimekust suurepäraseks, seda esitletud AUC piirmäärdest lähtuvalt.



Joonis 3.1. Logistilise regressiooni rakendamisel konstrueeritud mudeli ROC-kõver (hinnatud kontrollvalimil)

Allikas: autori koostatud SPSS rakendamisel

Statistilisest vaatepunktist väärrib konstrueeritud mudeli juures äramärkimist SPSS andmetöötlustarkvara poolt analüüsi väljundina esitatud Hosmer ja Lemeshow testi olulisuse tõenäosus 0,903. Hosmer-Lemeshow testi tulemuse põhjal võib väita, et statistiliselt (olulisuse nivool 0,05) sobitub konstrueeritud mudel hästi analüüsi kaasatud andmetega (*ibid.*, 160-162), mida toetab iseenesest ka eelnev mudeli klassifikatsioonivõimekuse hinnang.

### 3.2. Pankrotimudeli konstrueerimine otsustuspuu meetodi rakendamisel

Otsustuspuu meetodi puhul on tegu masinõppemeetodiga, millel on mitmeid variatsioone erinevate algoritmide näol, muuhulgas CART, CHAID. Algoritmid erinevad üksteisest eelkõige optimaalse valiku tegemise meetodika ja/või mudeli võimaliku struktuuri poolest (Chen 2011a; Chen 2011b; Gepp *et al.* 2010). Sõltumata rakendatavast algoritmist meenutab otsustuspuu meetodi läbi konstrueeritud mudel struktuurilt tagurpidi puud. Mudeli struktuuriosadeks on seejuures juurepunkt, oksad, sõlmpunktid (otsustussõlmed) ning lehed (terminaalsed sõlmed). Klassifitseerimine sellise mudeli põhjal toimub liikudes sõlmpunkti puu okste läbi seatud kriteeriumite alusel puu lehtedeni ehk terminaalsete sõlmedeni. Terminaalsetes sõlmedes toimub

eeldefineeritud kriteeriumite alusel vaatluspunktide lõplik klassifitseerimine (Gepp *et al.* 2010, 539).

Otsustuspuu meetod on mittelineaarne meetod, mis jaotab sõltumatute muutujate abil koguvalimi väiksemamahulisteks gruppideks. Otsustuspuu mudeli klassifitseerimisprotsessis liigutakse iteratiivselt sõlmest sõlme. See tähendab, et vaatluspunktide klassifitseerimise protsess toimub korduvalt ning eraldiseisvalt igal mudeli otsustustasandil, seda vastaval tasandil (ehk oksal) parimaks klassifikaatoriks osutuva sõltumatu muutuja põhjal (Berry, Linoff 1997, 165–167; Kayri, Kayri 2015, 21).

Käesolevas töös rakendati otsustuspuu konstrueerimisel nii CART (*Classification and Regression Tree*) kui ka CHAID (*Chi-square Automatic Interaction Detector*) algoritmi. Modelleerimisel rakendati andmetöötlustarkvara SPSS. Autorile teadaolevalt ei ole neid algoritme varasemalt Eesti ettevõtete pankrotimudelite väljatöötamisel kasutatud.

CART on otsustuspuu algoritm, mis pakub igas mudeli otsustussõlmes vaatluspunkti binaarset klassifitseerimisvõimalust (pankrotimudeli puhul pankrotistunud või jätkuvalt tegutsev ettevõtte) (Kayri, Kayri 2015, 21). CART otsustuspuu on seega juurest kuni lehtedeni kaheharuline. Kaks haru väljendavad sisuliselt igal klassifitseerimise sammul esinevat kahte võimalikku edasiliikumise suunda – väiksem (või võrdne) või suurem kui algoritmi poolt konkreetsele oksale seatud sõltumatu muutuja osas määratud optimaalne murdeväärtus. CHAID on mõnevõrra keerukam algoritm, kus klassifitseerimisvõimaluste hulk võib osutada igal mudeli tasandil kahest suuremaks (Berry, Linoff 1997, 182–183).

Mudeli konstrueerimisel lähtub CART algoritm vastavalt analüüsi koostaja valikule informatsiooni entroopia meetrikast või Gini meetrikast, millele tuginedes toimub muutujate valik ning optimaalsete murdepunktide määramine mudeli klassifikatsioonivõimekuse suurendamiseks (Kayri, Kayri 2015, 21–22; Topi, Tucker 2014, 19–6). Olulisel määral entroopia ning Gini meetrikate rakendamisel saadavad tulemused üksteisest aga praktikas ei erine (Murthy 1998, 354). CHAID algoritmi puhul rakendatakse mudeli konstrueerimisel *Hii*-ruut testi, genereerimaks mitmeharulisi klassifikatsioonivõimalusi (Berry, Linoff 1997, 182).

Otsustuspuu meetodile tugineva modelleerimise eesmärgiks seati kummagi CART ning CHAID algoritmi rakendamise läbi konstrueerida kõrgeima klassifitseerimisvõimekusega pankrotimudel



(kontrollvalimil hinnatuna), millesse kaasatud erinevate muutujate arv ei ületaks logistilise regressiooni pankrotimudelisse kaasatud muutujate arvu. Üheltpoolt üritati muutujate arvulise piirangu seadmisega vähendada mudeli üle kohandamise ohtu otsustuspuu mudelite treenimisel, teisalt aga tagada ka mudelite lihtne tõlgendatavus ning hilisem rakendatavus.

### 3.2.1. Pankrotimudeli konstrueerimine CART algoritmi rakendamisel

Erinevalt logistilisest regressioonist ei mõjuta CART meetodi rakendamist multikollineaarsuse esinemine modelleerimisse kaasatavate sõltumatute muutujate vahel. Seega kaasati CART analüüsi kõik varasemalt esitletud 13 modelleerimisse jõudnud muutujat (vt tabel 2.5). CART mudeli konstrueerimisel rakendati SPSS andmetöõtlustarkvara. Kõrgeima kirjeldusvõimega CART pankrotimudel konstrueeriti rakendades mudeli konstrueerimisel järgmisi parameetreid – mudeli ebapuhtuse hindamine Gini indeksi funktsiooni põhjal, mudeli treenimise peatamine puhtuse kasvumäära lävel 0,12.

Tabel 3.5. CART rakendamisel koostatud pankrotimudeli struktuur

Mudeli sõlmed		Proгноositav pankrotistumise tõenäosus	Proгноositav kategooria	Sõltumatud muutujad mudelis		
Sõlm (Jätk)	Sõlme tüüp			Muutuja	Puhtuse kasvumäär	Murdeväärtus
0	Juur	-	-	-	-	-
1 (3, 4)	Otsustus	76,2%	Pankrotistuv	$x_{11}$	0,222	$\leq 0,449$
2	Terminaalne	7,8%	Tegutsev	$x_{11}$	0,222	$> 0,449$
3	Terminaalne	86,9%	Pankrotistuv	$x_4$	0,069	$\leq 0,041$
4 (5, 6)	Otsustus	23,5%	Tegutsev	$x_4$	0,069	$> 0,041$
5 (7, 8)	Otsustus	13,3%	Tegutsev	$x_4$	0,016	$\leq 0,149$
6	Terminaalne	100,0%	Pankrotistuv	$x_4$	0,016	$> 0,149$
7	Terminaalne	66,7%	Pankrotistuv	$x_7$	0,013	$\leq 0,318$
8	Terminaalne	0,0%	Tegutsev	$x_7$	0,013	$> 0,318$

Allikas: autori koostatud

Struktuurilt osutus konstrueeritud CART mudel (vt tabel 3.5) ühe haruga puuks, koosnedes kolmest otsustussõlmest ning viiest terminaalsest sõlmest (vt CART mudeli illustratiivset avaldist blokk skeemi kujul lisas 5). Mudel hõlmas ühtekokku kolme muutujat, milleks olid omakapital / koguvara ( $x_{11}$ ) esimesel oksal, ärikasum / müügitulu ( $x_4$ ) teisel ning kolmandal oksal ning (käibevara - varud) / lühiajalised kohustised ( $x_7$ ) mudeli viimasel neljandal oksal. CART algoritmil põhineval konstrueeritud töötleva tööstuse pankrotimudelis olid seega esindatud nii ettevõtte kapitali struktuuri ( $x_{11}$ ), rentaablust ( $x_4$ ) kui ka likviidsust ( $x_7$ ) väljendavad muutujad. Olulisimaks osutus mudelisse kaasatud muutujatest ärikasum / müügitulu ( $x_4$ ).

Hinnates mudeli majanduslikku tõlgendatavust kerkib CART mudeli puhul esile aga majanduslik anomaalia, seda mudeli kuuendas sõlmes (terminaalne sõlm). Konstrueeritud mudeli alusel klassifitseeritakse ettevõtte, mille ärikasum / müügitulu ( $x_4$ ) väärtus ületab 14,9%, pankrotistuvate ettevõtete hulka. Analüüsidest modelleerimisse kaasatud andmeid, selgub, et treeningvalimis ulatub kahel (mudeli teisele ja kolmandale oksale jõudnud) pankrotistunud ettevõtte vaatluspunktil ( $t - 1$ ) ärikasum / müügitulu ( $x_4$ ) väärtus üle 14,9%. Teisele ning kolmandale oksale jõudnud tegevust jätkavate ettevõtete vaatluspunktidel (seda nii treening- kui ka kontrollvalimi puhul) jääb aga kõigil ( $t - 1$ ) ärikasum / müügitulu ( $x_4$ ) väärtus alla 14,9%. Sellest tulenevalt hindas rakendatud CART algoritm analüüsi kaasatud andmete põhjal otstarbekaks klassifitseerida kolmandal oksal vaatluspunkti ärikasum / müügitulu ( $x_4$ ) väärtusega üle 14,9% pankrotistuvate ettevõtete hulka.

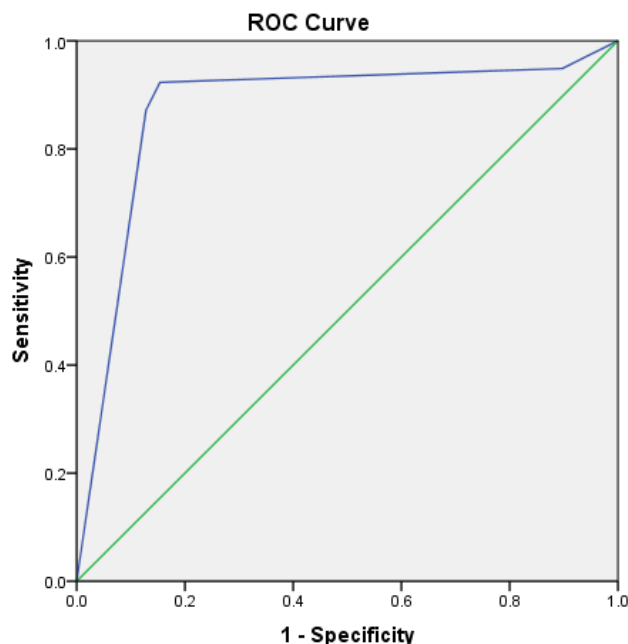
Tabel 3.6. CART rakendamisel koostatud pankrotimudeli klassifitseerimisvõimekus (hinnang kontrollvalimil)

Tegelik tulemus		Proгноositud tulemus treeningvalimis (lõikepunkt 0,5)			Proгноositud tulemus kontrollvalimis (lõikepunkt 0,5)		
		Y		Korrektus (%)	Y		Korrektus (%)
		0	1		0	1	
Y	0	71	12	85,5	33	6	84,6
	1	5	77	93,9	3	36	92,3
Üldine mudeli korrektus		89,7			88,5		

Allikas: autori arvutused

Märkus: väärtus 1 esindab pankrotistunud ettevõtet ning 0 jätkuvalt tegutsevat ettevõtet.

Klassifitseerimisel vaikimisi määratud lõikepunkt 0,5 langes ühtlasi ka konstrueeritud CART mudeli optimaalse lõikepunkti intervalli (0,08–0,66) piiridesse. Kui lõikepunkt 0,5 klassifitseeris konstrueeritud CART pankrotimudel korrektselt 89,7% kõigist treeningvalimisse kaasatud 165-st ettevõttest, siis kontrollvalimisse kaasatud 78 ettevõtte puhul langes mudeli klassifitseerimisvõimekus 88,5%-le. I tüüpi viga leidis kontrollvalimisse kaasatud vaatluspunktide klassifitseerimisel aset 7,7% ning II tüüpi viga 15,4% juhtudel. CART mudeli spetsiifilisus ulatus kontrollvalimi tulemuste põhjal hinnates seega 84,6%-ni ning tundlikkus seejuures 92,3%-ni.



Joonis 3.2. CART rakendamisel konstrueeritud mudeli ROC-kõver (hinnatud kontrollvalimil)  
Allikas: autori koostatud SPSS rakendamisel

Konstrueeritud CART mudeli ROC-kõver indikeerib samuti pankrotimudeli head klassifitseerimisvõimekust, kuivõrd mudeli ROC-kõvera aluse pinna väärtus AUC küündib 0,875 tasemele. Kuigi klassifitseerimistulemuste põhjal, arvestades seejuures ka CART mudeli AUC väärtust, võib konstrueeritud mudeli võimekust hinnata justkui heaks, siis ometi on mudeli ülesehitus majandusloogika vastane. Autori hinnangul ei ole käesolevas magistritöös konstrueeritud CART pankrotimudeli näol tegemist usaldusväärse pankrotimudeliga, mistõttu puudub sellel ka praktiline väärtus.

### 3.2.2. Pankrotimudeli konstrueerimine CHAID algoritmi rakendamisel

Sarnaselt CART algortimiga ei ole ka CHAID algoritmil rakendamise eelduseks multikollineaarsuse eemaldamine analüüsi kaasatavate sõltumatute muutujate hulgast. Seetõttu kaasatakse ka CHAID analüüsi kõik 13 modelleerimisse kaasatud muutujat (vaata tabel 2.5). Mudeli konstrueerimisel rakendati taaskord SPSS andmetöötlustarkvara. Kõrgeima klassifitseerimisvõimega CHAID pankrotimudel konstrueeriti järgmisi parameetreid rakendades – mudeli treenimine tõepärasuhte *Hii*-ruut statistiku (olulisuse nivool 0,01) põhjal, mudeli kasvatamise piiramine väärtusega 2 ning sõltumatute muutujate intervallskaala piiramine väärtusega 20.

Konstrueeritud mudel (vt tabel 3.7) osutus kaheharuliseks, koosnedes ühtekokku kahest otsustussõlmest ning seitsmest terminaalsest sõlmest (vt CHAID mudeli illustratiivset avaldist blokk skeemi kujul lisas 6). Mudel koondas enda alla kolm muutujat, kõik statistiliselt olulised (olulisuse nivool 0,01), millest esimesele oksal asetseb omakapital / koguvara ( $x_{11}$ ), vasakpoolse haru teisel oksal ärikasum / keskmine töötajate arv ( $x_5$ ) ning puu parempoolse haru teisel oksal muutuja puhaskasum / koguvara ( $x_1$ ). Seega on CHAID algoritmi läbi konstrueeritud pankrotimudelis esindatud vaid ettevõtte kapitali struktuuri ( $x_{11}$ ) ning rentaablust kajastavad näitajad ( $x_5$  ja  $x_1$ ). Analüüsi väljundile tuginedes osutus olulisimaks muutujaks konstrueeritud CHAID mudelis omakapital / koguvara ( $x_{11}$ ).

Tabel 3.7. CHAID rakendamisel koostatud pankrotimudeli struktuur

Mudeli sõlmed		Proгноositav pankrotistumise tõenäosus	Proгноositav kategooria	Sõltumatud muutujad mudelis		
Sõlm (Jätk)	Sõlme tüüp			(x)	(p)	Murdeväärtus
0	Juur	-	-	-	-	
1	Terminaalne	93,9%	Pankrotistuv	$x_{11}$	0,000	$\leq 0,191$
2 (5, 6)	Otsustus	61,9%	Pankrotistuv	$x_{11}$	0,000	$0,191 - 0,399$
3 (7, 8, 9)	Otsustus	30,3%	Tegutsev	$x_{11}$	0,000	$0,399 - 0,629$
4	Terminaalne	0,0%	Tegutsev	$x_{11}$	0,000	$> 0,629$
5	Terminaalne	89,5%	Pankrotistuv	$x_5$	0,007	$\leq -297,967$
6	Terminaalne	39,1%	Tegutsev	$x_5$	0,007	$> -297,967$
7	Terminaalne	100,0%	Pankrotistuv	$x_1$	0,000	$\leq -0,071$
8	Terminaalne	33,3%	Tegutsev	$x_1$	0,000	$-0,071 - 0,087$
9	Terminaalne	0,0%	Tegutsev	$x_1$	0,000	$> 0,087$

Allikas: autori koostatud

Mudeli struktuurist nähtub, et CHAID algoritm on rakendatud parameetritele tuginedes moodustanud esimese oksa muutuja omakapital / koguvara ( $x_{11}$ ) osas neli võimalikku sõlme, millest esimene ( $x_{11} \leq 0,191$ ) ja neljas ( $x_{11} > 0,629$ ) on terminaalsed sõlmed, tähistamaks lõikepunkt 0,5 vastavalt (keskmiselt) aasta pärast pankrotistuvaid (pankrotistumise tõenäosus 93,9%) ning jätkuvalt tegutsevaid ettevõtteid (pankrotistumise tõenäosus 0%). Keskmiste omakapital / koguvara ( $x_{11}$ ) sõlmede puhul on aga tegu otsustussõlmedega, kus vaatluspunktide lõplik prognoositav klassifikatsioon selgub vastavalt muutujate ärikasum / keskmine töötajate arv ( $x_5$ ; kaks võimalikku terminaalselt sõlme) ning puhaskasum / koguvara ( $x_1$ ; kolm võimalikku terminaalselt sõlme) väärtuste hindamisel.

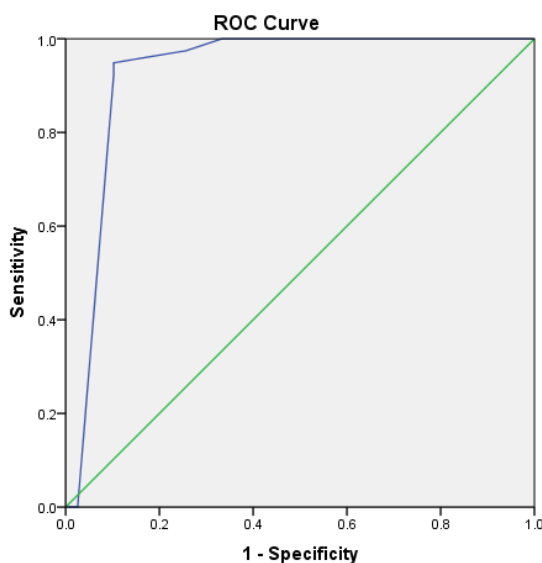
Tabel 3.8. CHAID rakendamisel koostatud pankrotimudeli klassifitseerimisvõimekus (hinnang kontrollvalimil)

Tegelik tulemus		Proгноositud tulemus treeningvalimis (lõikepunkt 0,5)			Proгноositud tulemus kontrollvalimis (lõikepunkt 0,5)		
		Y		Korrektus (%)	Y		Korrektus (%)
		0	1		0	1	
Y	0	78	5	94,0	35	4	89,7
	1	13	69	84,1	2	3	94,9
Üldine mudeli korrektsus		89,1			92,3		

Allikas: autori arvutused

Märkus: väärtus 1 esindab pankrotistunud ettevõtet ning 0 jätkuvalt tegutsevat ettevõtet.

Klassifitseerimisel vaikumisi määratud lõikepunkt 0,5 langes ühtlasi ka konstrueeritud CART mudeli optimaalse lõikepunkti intervalli (0,39–0,8) raamidesse. Esitletud CHAID pankrotimudel klassifitseeris korrektselt 89,1% kõigist treeningvalimisse kaasatud 165-st ettevõttest, kontrollvalimisse kaasatud 78 vaatluspunkti puhul ulatus mudeli klassifitseerimisvõimekus aga 92,3%-le, mida võib hinnata suurepäraseks kirjeldusvõimeks. I tüüpi viga leidis CHAID mudelit rakendades aset 5,1% juhtudest ning II tüüpi viga 10,3% juhtudest kontrollvalimisse kaasatud vaatluspunktide klassifitseerimisel. CHAID mudeli spetsiifilisus ulatus kontrollvalimi klassifitseerimistulemuste põhjal hinnates seega 89,7%-ni ning tundlikkus seekõrval 94,9%-ni.



Joonis 3.2. CHAID rakendamisel konstrueeritud mudeli ROC-kõver (hinnatud kontrollvalimil)  
Allikas: autori koostatud SPSS rakendamisel

Konstrueeritud CHAID mudeli ROC-kõvera statistikule tuginedes võib väita, et mudeli kirjeldusvõimekus on suurepärase, kuivõrd mudeli ROC-kõvera aluse pinna väärtus AUC küündib 0,926 tasemele. CHAID mudeli AUC väärtus jääb küll mõnevõrra alla logistilise regressiooni pankrotimudeli AUC väärtusele (0,962), kuid võrreldes CART pankrotimudeli AUC väärtusega (0,875) on tegemist arvestatavalt parema tulemusega. Olulise märkusena tuleb siinkohal välja tuua, et võrreldes konstrueeritud CART mudeliga on CHAID mudel ülesehituselt majanduslikult tõlgendatav, millest tulenevalt on viimasel ka praktiline väärtus.

### 3.3. Konstrueeritud pankrotimodelite tulemuste võrdlus ning järeldused

Modelleerimisel seadis autor käesolevas töös eesmärgiks pankrotimodelite klassifitseerimisvõimekuse maksimeerimise (nimelt kontrollvalimil hinnatuna) ning seda vähima võimaliku hulga muutujate juures, tagades lihtsuse mudelite tõlgendatavuses ning praktilisuse konstrueeritud mudelite hilisemal rakendamisel. Modelleerimisse kaasatud 13 muutujast osutusid logistilise regressiooni mudelisse valituks viis, CART ning CHAID mudelisse aga kolm (vt tabel 3.9).

Tabel 3.9. Konstrueeritud pankrotimudelitesse kaasatud muutujad

Mudel	Muutujad	
Logit	X <sub>2</sub>	ärikasum / koguvara
	X <sub>7</sub>	(käibevara - varud) / lühiajalised kohustised
	X <sub>11</sub>	omakapital / koguvara
	X <sub>12</sub>	müügitulu / koguvara
	X <sub>13</sub>	koguvara / keskmine töötajate arv
CART	X <sub>4</sub>	ärikasum / müügitulu
	X <sub>7</sub>	(käibevara - varud) / lühiajalised kohustised
	X <sub>11</sub>	omakapital / koguvara
CHAID	X <sub>1</sub>	puhaskasum / koguvara
	X <sub>5</sub>	ärikasum / keskmine töötajate arv
	X <sub>11</sub>	omakapital / koguvara

Allikas: autori koostatud

Konstrueeritud pankrotimudelites on läbivaks ühisosaks ettevõtte kapitali struktuuri kajastav muutuja omakapital/koguvara (x<sub>11</sub>), mis on teostatud analüüsi tulemustele tuginedes äärmiselt oluline näitaja ettevõtte pankrotistumise tõenäosuse prognoosimisel keskmiselt üheaastases (6–18 kuud) ette vaatavas perspektiivis. Lisaks on kõigis mudelites esindatud ka rentaablusnäitajad

( $x_1$ ,  $x_2$ ,  $x_4$  ning  $x_5$ ). Logistilise regressiooni ning CART rakendamisel konstrueeritud kõrgeima klassifikatsioonivõimekusega mudelites on esindatud ka likviidsust ja maksevõimet kajastav muutuja ( $x_7$ ). Logistilise regressiooni pankrotimudel on kolmest konstrueeritud mudelist aga ainuke, kus muuhulgas ka ettevõtte tegevuse efektiivsust kajastavad muutujad esindatud on ( $x_{12}$ ,  $x_{13}$ ).

Mudelitest osutus ettevõtete klassifitseerimisel võimekaimaks logistilise regressiooni analüüsimeetodil põhinev pankrotimudel, mis suutis optimaalsel löikepunktil (0,675) korrektselt klassifitseerida 93,6% kõigist kontrollvalimisse kaasatud vaatluspunktidest. I tüüpi viga leidis logistilise regressiooni pankrotimudeli rakendamisel seejuures aset 5,1% juhtudel ning II tüüpi viga 7,7% klassifikatsioonijuhtudel. Klassifitseerimisvõimekuselt teisele kohale jäi otsustuspuu meetodi CHAID algoritmi, mille klassifitseerimisvõimekus ulatus löikepunkt 0,5 (langes ka mudeli optimaalse löikepunkti intervalli piiridesse) vastavalt 92,3% tasemele. CHAID mudeli puhul ühtis klassifitseerimisel aset leidnud I tüüpi vea määr logistilise regressiooni mudeli omaga (5,1%). Ka parima CART pankrotimudeli klassifitseerimisvõimekus ulatus heale tasemele (88,5%), kuid mudeli ülesehitus osutus seejuures majandusloogika vastaseks, mistõttu puudub sellel mudelil sügavam praktiline väärtus. Eeltoodud järjestust pankrotimudelite võimekuse osas toetab ka mudelite ROC-kõvera alused AUC väärtused – logistilise regressiooni mudeli AUC 0,962, CHAID mudeli AUC 0,926 ning CART mudeli AUC 0,875 (vt tabel 3.10).

Tabel 3.10. Konstrueeritud pankrotimudelite üldine klassifitseerimismäär, spetsiifilisus ning tundlikkus

Kontrollvalim	Logit mudel (löikepunkt 0,675)	CART mudel (löikepunkt 0,5)	CHAID mudel (löikepunkt 0,5)
Spetsiifilisus	92,3%	84,6%	89,7%
Tundlikkus	94,9%	92,3%	94,9%
Üldine klassifikatsioonimäär	93,6%	88,5%	92,3%
Spetsiifilisuse ning tundlikkuse erinevus absoluutväärtuses	2,6%	7,7%	5,2%

Allikas: autori arvutused

Hinnates konstrueeritud mudelite klassifitseerimistulemuste stabiilsust kontrollvalimil saavutatud spetsiifilisuse ning tundlikkuse löikes, võib täheldada, et enim volatiilsust esineb mudelite löikes CART mudeli puhul (spetsiifilisuse ning tundlikkuse erinevus absoluutväärtuses 7,7%). Enim demonstreerib tulemustes stabiilsust logistilise regressiooni mudel (erinevuse absoluutväärtus 2,6%), mille üldine klassifitseerimisvõimekus kontrollvalimi põhjal hinnatuna

ulatus samuti käsitletud mudelitest kõrgeimale tasemele. Kuigi logistilise regressiooni ning CHAID algoritmi pankrotimodelite puhul esinenud I tüüpi vigade määr ei erinenud üksteisest (5,1% juhtudest), siis mõõdetud erinevuste põhjal mudelite spetsiifilisuses ning tundlikkuses võib logistilise regressiooni mudelit hinnata mõnevõrra stabiilsemaks.

Käesoleva magistritöö raames konstrueeritud kolmest pankrotimodelist on võimekaim ning stabiilsem logistilisel regressioonil põhinev mudel. Autori hinnangul pakub konstrueeritud logistilisel regressioonil põhinev pankrotimudel täpset ning usaldusväärset väljundit Eesti töötleva tööstuse ettevõtete pankrotiohu hindamisel, seda keskmiselt üheaastases (6–18 kuud) ette vaatavas perspektiivis. Seejuures tuleb arvestada võimalike ajas aset leidvate muutustega mudelisse kaasatud näitajate optimaalse taseme dünaamikas. Tingimata on oluline konstrueeritud pankrotimudeleid perioodiliselt testida, hinnata ning vajadusel häälestada vastavalt majanduskeskkonnas toimunud muutustele.

Magistritöös konstrueeritud töötleva tööstuse pankrotimodelite rakendamisel on oluline silmas pidada modelleerimisel seatud piiranguid, eelkõige seda, et valimisse kaasati vaid minimaalselt kolmeaastase tegevusajalooga töötleva tööstuse ettevõtteid, milles tegutsevate töötajate keskmine arv konkreetsel majandusaastal jäi vahemikku 10–249. Töö edasiarendusena oleks mõistlik testida konstrueeritud mudelite klassifikatsioonivõimekust ka töötleva tööstuse ettevõtete peal, milles tegutsevate töötajate keskmine arv jääb konkreetsel majandusaastal vahemikku 1–9 ja/või ületab 249 piiri. Eeldusel, et mudeli kasutamisel esineva I tüüpi ning II tüüpi vea tagajärg ei ole praktikas võrdväärne, on töö edasiarendamise võimaluseks hinnata konstrueeritud pankrotimodelite kvaliteeti ka tulu-kulu printsiibist, mõõtes mudelite põhjal tehtud finantsotsuste tulemusi monetaarsetel alustel. Eeldades, et finantsnäitajad ei pruugi katta kõiki olulisi pankrotiohule viitavaid ilminguid, on käesoleva töö võimalik edasiarendus ka kvalitatiivsete muutujate kaasamine pankroti modelleerimisse, parendamaks pankrotimudeli prognoosivõimekust ning mitmekülgset veelgi enam.



## KOKKUVÕTE

Käesoleva magistritöö eesmärk oli Eesti töötleva tööstuse ettevõtete pankrotimudeli konstrueerimine. Töötleval tööstusel on tegevusharuna äärmiselt oluline roll Eesti majanduse stabiilsuse ning jätkusuutlikkuse toetamisel. Seejuures ulatub aga ettevõtete pankrottide sagedusmäär töötlevas tööstuses võrreldes majandusharude keskmisega koguni poole kõrgemale. Kirjeldatud statistika toetab tugevalt vajadust ajakohase ning usaldusväärse töötleva tööstuse ettevõtete pankrotimudeli järele, mis pakuks usaldusväärset toetust juhtimis-, investeerimis-, finantseerimis- aga ka audiitorotsuste langetamisele.

Töö eesmärgi saavutamiseks andis autor esmalt ülevaate ettevõtte pankrotistumise ning selle prognoosimisega seotud teoreetilisest raamistikust. Enamik olemasolevaid pankrotimudeleid põhinevad väljundi dihhotoomsusel ehk kaheväärtuslikkusel. Dihhotoomsel väljundil põhineva pankrotimudeli rakendamisel ettevõtete klassifitseerimisel võib klassifitseerimistulemustes esineda olemuslikult kahte tüüpi vigu. Esiteks võib rakendatud mudel klassifitseerida pankrotistuva ettevõtte tegutsema jäävate ettevõtete hulka, mida kajastatakse I tüüpi vea juhtumina. II tüüpi vea juhtumi puhul klassifitseerib pankrotimudel tegutsema jääva ettevõtte pankrotistuvate ettevõtete hulka.

Enamik kirjanduses kajastust leidnud pankrotimudelitest on konstrueeritud puhtalt ettevõtte finantsaruandluses kajastatud informatsiooni põhjal, kus pankroti modelleerimisse kaasatakse üksnes aruannetes esitatud andmete põhjal arvutatud finantsuhtarve ja/või absoluutnäitajaid. Seda põhjusel, et ettevõtete finantsandmete kogumine tulenevalt nende võrdlemisi heast kättesaadavusest on lihtne. Analüüsi kaasatavate finantsnäitajate ehk sõltumatute muutujate valikul tuginetakse tihti nende esinemise populaarusele kirjanduses, võttes seejuures arvesse ka muutujate osas kirjeldatud klassifitseerimisvõimekust.

Üldjoontes jagunevad pankroti modelleerimisel rakendatavad analüüsimeetodid statistiliste ning masinõppe meetodite vahel. Statistiliste meetodite hulka kuuluvad näiteks ühemõõtmeline analüüs ja rohkelt kajastust leidnud mitmemõõtmeline diskriminantanalüüs ning logistiline

regressioon. Logistilise regressiooni peamiseks eeliseks peetakse selle paindlikkust analüüsile seatud eelduslike tingimuste osas ning seda, et antud meetod ei eelda ettevõtte seisundi ning finantsnäitajate vahel esineva suhte lineaarsust. Logistilise regressiooni peamiseks puudustena nähakse selle tundlikkust valimis esinevate ekstreemsete ja ka puuduolevate väärtuste, aga ka muutujate vahelise multikollineaarsuse suhtes. Masinõppe meetoditest on pankroti modelleerimisel seni kõige laialdasemat kasutust leidnud tehnilike närvivõrkude ning otsustuspuu meetod. Otsustuspuu meetod suudab teoorias hästi toime tulla ka valimis esinevate ekstreemsete ning puuduolevate väärtustega. Suurimaks ohuks peetakse otsustuspuu meetodi rakendamisel aga mudeli ülekohandamist, mille vältimiseks soovitatakse otsustuspuu mudeli treenimisel lähtuda mudeli klassifitseerimistäpsuse maksimeerimisest kontrollvalimil, mitte treeningvalimil.

Magistritöös teostatud modelleerimise aluseks olnud andmed Eesti töötleva tööstuse ettevõtete kohta kogus autor Registrate ja Infosüsteemide Keskusest. Andmete kogumisel rakendati tegevusharu klassifikatsioonile (EMTAK klassifikatsioon C), ettevõtte töötajate arvule (10–249 töötajat) ning tegevusaastate arvule (minimaalselt 3 aastat) tuginevaid piiranguid. Eeltoodud piiranguid rakendati andmete kogumisel eesmärgiga tagada konstrueeritud pankrotimudelite kõrgem kvaliteet ning praktilisus. Tuginedes seatud piirangutele kaasati töö analüüsi perioodil 2005–2016 pankrotistunud ettevõtete ( $t - 1$  aasta) andmed. Kokku kaasati analüüsi 129 pankrotistunud töötleva tööstuse ettevõtet. Tuginedes sobitatud paaride meetodile kaasati koguvalimi moodustamiseks analüüsi ka 129 jätkuvalt tegutseva töötleva tööstuse ettevõtte finantsandmed.

Analüüsi kaasas autor kirjandusele ning enda kogemuspõhisele oskusteabele tuginedes sõltumatute muutujatena algselt 19 finantsnäitajat. Valitud muutujate näol olid esindatud mitmesugused erinevad majandustegevust kajastavad valdkonnad – ettevõtte rentaablus, likviidsus ja maksevõime, kapitali struktuur, laenu teenindamise võimekus ning tegevuse efektiivsus. Modelleerimisse kaasatava andmevalimi terviklikkuse ning puhtuse tagamiseks eemaldati valimist kokku aga 6 algselt kaasatud muutujat (olulises mahus nullväärtuslikke vaatluspunkte sisaldanud muutujad), muuhulgas ka kõik algselt kaasatud laenu teenindamise näitajad. Lisaks eemaldati ka 15 ekstreemväärtuslikku konkreetset vaatluspunkti (seehulgas 8 pankrotistunud ning 7 jätkuvalt tegutsevat ettevõtet). Ekstreemväärtuslike vaatluspunktide tuvastamiseks rakendati käesolevas töös Mahalanobise ning Cook'i kauguse mõõdikuid. Eesti töötleva tööstuse pankrotimudeli konstrueerimisse kaasati pärast olulises mahus nullväärtusi

sisaldavate muutujate ning ekstreemväärtuslike vaatluspunktide eemaldamist seega kokku 121 pankrotistunud ning 122 jätkuvalt tegutseva ettevõtte finantsandmed 13 suhtarvupõhise sõltumatu muutuja lõikes.

Magistritöö empiirilises osas rakendati Eesti töötleva tööstuse pankrotimudelite konstrueerimisel logistilist regressiooni ning otsustuspuu meetodi algoritme CART (*Classification and Regression Tree*) ja CHAID (*Chi-square Automatic Interaction Detector*). Modelleerimisel seati eesmärgiks maksimeerida pankrotimudelite klassifitseerimisvõimekus (kontrollvalimil hinnatuna) ning seda vähima võimaliku hulga muutujate juures, tagades mudelite võimalikult lihtsa tõlgendatavuse.

Logistilise regressiooni rakendamisel konstrueeritud kõrgeima klassifitseerimisvõimekusega pankrotimudeli üldkuju, hindamaks pankrotistumise tõenäosust (keskmiselt) üheaastases ette vaatavas perspektiivis (6–18 kuud), avaldus järgneval kujul ( $Y_i$  on vaatluspunkti logistiline skoor pankrotistumise tõenäosuse avaldises  $P(x_i) = \frac{1}{1+\exp-(Y_i)}$ ):

$$Y_i = -8,19474 \times \frac{\text{ärikasum}}{\text{koguvara}} - 1,27174 \times \frac{(\text{käibevara} - \text{varud})}{\text{lühiajalised kohustised}} - 5,21146 \times \frac{\text{omakapital}}{\text{koguvara}} + 1,00375 \times \frac{\text{müügitulu}}{\text{koguvara}} + 0,00002 \times \frac{\text{koguvara}}{\text{keskmine töötajate arv}} \quad (7)$$

Logistilisel regressioonil põhinev konstrueeritud mudel suutis optimaalsel lõikepunktil (0,675) korrektselt klassifitseerida 93,6% kõigist kontrollvalimisse kaasatud vaatluspunktidest. I tüüpi viga leidis mudeli klassifitseerimise juures aset 5,1% juhtudel ning II tüüpi viga 7,7% klassifikatsiooni juhtudel.

Otsustuspuu meetodi CART algoritmi rakendamisel koostatud kõrgeima kirjeldusvõimega mudelis olid muutujatest esindatud omakapital / koguvara, ärikasum / müügitulu ning (käibevara - varud) / lühiajalised kohustised. Mudeli klassifitseerimisvõimekus ulatus kontrollvalimisse kaasatud ettevõtte puhul 88,5%-le. Ometi osutus CART mudeli ülesehitus majandusloogika vastaseks ning sellel mudelil puudub praktiline väärtus.

Otsustuspuu meetodi CHAID algoritmi rakendamisel konstrueeritud kõrgeima klassifitseerimisvõimekusega mudel koondas enda alla kolm muutujat, milleks olid omakapital / koguvara, ärikasum / keskmine töötajate arv ning puhaskasum / koguvara. Konstrueeritud

CHAID pankrotimudel klassifitseerimisvõimekus kontrollvalimisse kaasatud vaatluspunktide puhul ulatus mudeli 92,5%-le. I tüüpi viga leidis CHAID mudelit rakendades aset 5,1% juhtudest ning II tüüpi viga 10,3% juhtudel kontrollvalimisse kaasatud vaatluspunktide klassifitseerimisel.

Käesoleva magistritöö raames konstrueeritud kolmest pankrotimudelist on võimekaim ning stabiilseim logistilisel regressioonil põhinev mudel. Autori hinnangul pakub logistilisel regressioonil põhinev pankrotimudel täpset ning usaldusväärset väljundit Eesti töötleva tööstuse ettevõtete pankrotiohu hindamisel, seda keskmiselt üheaastases (konkreetsemalt 6–18 kuud) ette vaatavas perspektiivis. Ehkki logistilise regressiooni ning CHAID algoritmi pankrotimudelite puhul esinenud I tüüpi vigade määr ei erinenud üksteisest (5,1% juhtudest), siis mõõdetud erinevuste põhjal mudelite spetsiifilisuses ning tundlikkuses võib logistilise regressiooni mudelit hinnata mõnevõrra stabiilsemaks.

Magistritöös konstrueeritud töötleva tööstuse pankrotimudelite rakendamisel on oluline silmas pidada modelleerimisel, eelkõige andmete kogumisel seatud piiranguid (sh tegevusharu klassifikatsioon, töötajate arv, tegevusaastad). Töö edasiarendusena oleks mõistlik testida konstrueeritud mudelite sobivust ka nende töötleva tööstuse ettevõtete klassifitseerimisel, milles konkreetsel majandusaastal tegutsevate töötajate keskmine arv jäi vahemikku 1–9 ja/või ületab 249 piiri. Eeldusel, et klassifitseerimisel esineva I tüüpi ning II tüüpi vea tagajärg ei ole praktikas võrdväärne, on töö edasiarendamise võimaliku suunana hinnata konstrueeritud pankrotimudelite võimekust ka tulu-kulu printsiibist, mõõtes mudelite põhjal tehtud finantsotsuste tulemusi monetaarsetel alustel. Kolmandaks näeb autor pankrotimudelite võimaliku edasiarendusena kvalitatiivsete muutujate modelleerimisse kaasamist, parendamaks Eesti töötleva tööstuse ettevõtete pankrotimudeli mitmekülgust ning prognoosivõimekust veelgi enam.

## **SUMMARY**

### **DEVELOPMENT OF BANKRUPTCY PREDICTION MODEL FOR ESTONIAN MANUFACTURING COMPANIES**

Janar Vahter

Although corporate bankruptcy is a natural part of economic cycle, in practice, it involves many different parties (investors, employees, suppliers, creditors etc.) and, therefore, potentially large costs. Bankruptcy does not appear costly only on an individual level, but also on a socio-economic level. With the aim of taking corrective or even preventive actions in a timely manner, it has become increasingly important to develop efficient and reliable tools, that would provide with a timely and accurate information about the financial health of the companies. The necessity for such tools has stimulated a lot of research in the area of corporate failure together with the development of bankruptcy prediction models. Reliable bankruptcy models could provide assistance to management, investors, creditors as well as auditors.

The aim of this thesis was to develop a bankruptcy prediction model for Estonian manufacturing companies. Estonian manufacturing industry holds a crucial role in the stability and sustainability of the whole economy. However, the corporate bankruptcy rate within the Estonian manufacturing industry is one of the highest across all industries. The given statistics indicates a evidential necessity for a reliable bankruptcy prediction model, which could provide efficient assistance in monitoring and decision-making processes regarding the Estonian manufacturing companies of interest. In the aim of developing an efficient bankruptcy prediction model, first, an overview regarding the theoretical framework of predicting corporate bankruptcy was given. Also an overview on the most popular methods of analysis used for the purpose of developing bankruptcy prediction models was presented.

In the theoretical part of the thesis it was concluded that most of the bankruptcy prediction models are developed dichotomous in nature (providing two classification opportunities –

bankrupt and non-bankrupt) and based solely on the information collected from the financial statements of the companies. Majority of studies employ the financial information to the analysis in form of financial ratios. The choice of specific ratios employed in the analysis is often based on the reported predictive capability of the ratios in the existing literature. Even though financial ratios have been subject to quite many criticisms, foremost regarding the reliability of such information, financial ratios continue to be the most commonly used form of information in the bankruptcy modelling. Even though market-driven, as well as non-financial variables are encouraged to be included in the analysis, the data is not available for majority of the companies, which is also the main limitation regarding the application of such data in the modelling.

It appears from the literature, that the methods of analysis available for bankruptcy modelling are commonly divide into two main groups, classical and artificial intelligent methods. The most popular classical methods used in the bankruptcy analysis are multivariate discriminant analysis and logistic regression, of which the former is argued to be more limited to strict assumptions, such as data normality, independence among the explanatory variables, linearity. Logistic regression in turn is found more flexible and also easier to interpret. Commonly employed artificial intelligent methods, that are all ultimately based on machine-learning, include artificial neural networks and decision tree method. Even though both methods have presented results in terms of classification problems, the optimal model of decision tree is found to be easier to develop as well as interpret, which are also commonly presented as the main advantages of decision tree method compared to artificial neural networks.

Firstly, the data collection and explanatory variable selection were described in the empirical part of the thesis. The financial data regarding bankrupt and non-bankrupt Estonian manufacturing companies was collected from the Centre of Registers and Information Systems. The latest available financial data on the manufacturing companies was collected for the period of 2005–2016. Several restriction, regarding the number of employees (10–249), the minimum age (3 years) and the industry (manufacturing only) of the company, were employed in the process sampling to increase the quality of the data used in the bankruptcy modelling. By employing matched-pairs sampling technique, altogether 129 pairs (258 observation points) of bankrupt and non-bankrupt manufacturing companies were initially generated for the analysis. The sample was then split into training and validation sub-samples (with a ratio of 2:1). Regarding the explanatory variables, 19 different variables were initially included in the analysis. The variables were mostly selected based on the previous literature and covered the areas of profitability,

liquidity, capital structure, debt servicing and efficiency. In the process of dealing with missing values and multivariate outliers, however, in total of 6 variables and 15 observation points (including 8 bankrupt and 7 non-bankrupt companies) were removed from the initial sample.

Secondly, the development of bankruptcy prediction model for Estonian manufacturing firms was carried out. The methods of analysis employed in the modelling were logistic regression and decision tree method, of which specifically the algorithms of CART (*Classification and Regression Tree*) ja CHAID (*Chi-square Automatic Interaction Detector*) were used. Overall, it can be concluded that the bankruptcy models based on logistic regression and CHAID demonstrated good classification accuracy, CART however, presented an economically illogical bankruptcy model. Nevertheless, as the results indicated the developed logistic regression model demonstrated the highest accuracy rate in classifying the observation points (among the validation subsample) as well as the most stability across the specificity and sensitivity of the model.

The overall accuracy rate (assessed based on validation subsample) of logistic regression model in classifying correctly bankrupt and non-bankrupt Estonian manufacturing companies appeared as high as 93.6% (at the optimal cut-off value). For CHAID model the overall accuracy rate was 92.5% and for CART model 88.5%. The explanatory variable included in all three models was total equity to total assets, which emphasizes the importance of capital structure in terms of assessing financial health of a company. Also different variables of profitability were included in all three models, demonstrating the relationship between company's sustainability and ability to generate profit.

As only the companies with 10–249 employees were included in the modelling, it would be also useful to assess the classification accuracy of the developed bankruptcy models based on the companies including less than 10 and more than 249 employees. Additionally, it would be practical to assess the developed models in terms of misclassification costs. For further development of the presented bankruptcy models the non-financial (quantitative as well as qualitative) variables could be included into the modelling.

## KASUTATUD ALLIKATE LOETELU

- Agarwal, V., Taffler, R. (2008). Comparing the Performance of Market-based and Accounting-based Bankruptcy Prediction Models. – *Journal of Banking and Finance*, Vol. 32, No. 8, 1541–1551.
- Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and The Prediction of Corporate Bankruptcy. – *The Journal of Finance*, Vol. 23, No. 4, 589–609.
- Altman, E. I., Haldeman, R., Narayanan, P. (1977). Zeta Analysis: a New Model to Identify Bankruptcy Risk of Corporations. – *Journal of Banking & Finance*, Vol. 1, No. 1, 29–54.
- Altman, E. I., Narayanan, P. (1997). An International Survey of Business Failure Classification Models. – *Financial Markets, Institutions & Instruments*, Vol. 6, No. 2, 1–57.
- Altman E. I., Sabato, G. (2007). Modeling Credit Risk for SMEs: Evidence from the US Market. – *Abacus: A Journal of Accounting, Finance and Business Studies*, Vol. 43, No. 3, 332–357.
- Bager, A., Roman, M., Algedih, M., Mohammed, B. (2017). Addressing Multicollinearity in Regression Models: a Ridge Regression Application. – *Journal of Social and Economic Statistics*, Vol. 6, No. 1, 30–45.
- Balcaen, S., Ooghe, H. (2004). 35 Years of Studies on Business Failure: an Overview of the Classic Statistical Methodologies and Their Related Problems. – Ghent University, Department of Accountancy and Corporate Finance, Working Paper D/2004/7012/34.
- Barniv, R., Raveh, A. (1989). Identifying Financial Distress: A New Nonparametric Approach. – *Journal of Business Finance & Accounting*, Vol. 16, No. 3, 361–383.
- Bauer, J., Agarwal, V. (2014). Are Hazard Models Superior to Traditional Bankruptcy Prediction Approaches? A Comprehensive Test. – *Journal of Banking & Finance*, Vol. 40, 432–442.
- Beaver, W. H. (1966). Financial Ratios as Predictors of Failure. – *Journal of Accounting Research*, Vol. 4, 71–111.
- Begley, J., Ming, J., Watts, S. (1996). Bankruptcy Classification Errors in the 1980s: An Empirical Analysis of Altman's and Ohlson's Models. – *Review of Accounting Studies*, Vol. 1, No. 4, 267–284.
- Bellovary, J. L., Giacomino, D. E., Akers, M. D. (2007). A Review of Bankruptcy Prediction Studies: 1930 to Present. – *Journal of Financial Education*, Vol. 33, 1–42.



- Berry, M. J. A., Linoff, G. S. (1997). *Data Mining Techniques for Marketing, Sales and Customer Support*. 1st ed. New York: Wiley.
- Blanco-Oliver, A., Irimia-Diequez, A., Oliver-Alfonso, M., Wilson, N. (2015). Improving Bankruptcy Prediction in Micro-Entities by Using Nonlinear Effects and Non-Financial Variables. *Czech Journal of Economics and Finance*, Vol. 65, no. 2, 144–166.
- Bollen, K. A., Jackman, R. W. (1990). Regression Diagnostics: An Expository Treatment of Outliers and Influential Cases. – *Modern Methods of Data Analysis*, 257–291. Newbury Park: Sage.
- Boritz, J. E., Kennedy, D. B., Albuquerque, M. E. (1995). Predicting corporate failure using a neural network approach. – *International Journal of Intelligent Systems in Accounting Finance & Management*, Vol. 4, No. 2, 95–111.
- Chatterjee, S., Hadi, A. S. (2006). *Regression Analysis by Example*. 4th ed. New Jersey: Wiley.
- Chava, S., Jarrow, R. A. (2004). Bankruptcy Prediction with Industry Effects. – *Review of Finance*, Vol. 8, No. 4, 537–569.
- Chen, M-Y. (2011a). Bankruptcy Prediction in Firms with Statistical and Intelligent Techniques and a Comparison of Evolutionary Computation Approaches. – *Computers and Mathematics with Applications*, Vol. 62, 4514–4524.
- Chen, M-Y. (2011b). Predicting Corporate Financial Distress Based on Integration of Decision Tree Classification and Logistic Regression. – *Expert Systems with Applications*, Vol. 38, No. 9, 11261–11272.
- Chen, S., Härdle, W. K., Moro, R. A. (2011). Modeling Default Risk With Support Vector Machines. – *Quantitative Finance*, Vol. 11, No. 1, 135–154.
- Creditinfo Eesti AS. (2016). *Pankrotid Eestis 2016 – paneeluuring*. Kättesaadav: <https://web.creditinfo.ee/pankrotid2016.pdf> , 6. Mai 2018.
- Cybinski, P. (2001). Description. Explanation, Prediction: The Evolution of Bankruptcy Studies? – *Managerial Finance*, Vol. 27, No. 4, 29–44.
- Dimitras, A., Zanakis, S., Zopounidis, C. (1996). A Survey of Business Failures with an Emphasis on Failure Prediction Methods and Industrial Applications. – *European Journal of Operational Research*, Vol. 90, No. 3, 487–513.
- Doumpos, M., Zopounidis, C. (1999). A Multicriteria Discrimination Method for the Prediction of Financial Distress: the Case of Greece. – *Multinational Finance Journal*, Vol. 3, No. 2, 71–101.
- E-äriregister. (2008) *EMTAK 2008 selgitavad märkused*. Kättesaadav: [http://www.rik.ee/sites/www.rik.ee/files/elfinder/article\\_files/EMTAK%202008%20EST%20-%20selgitavate%20m%C3%A4rkused.pdf](http://www.rik.ee/sites/www.rik.ee/files/elfinder/article_files/EMTAK%202008%20EST%20-%20selgitavate%20m%C3%A4rkused.pdf) , 6. mai 2018.

- Eesti Statistikaamet. (2018) EM001: Ettevõtete tulud, kulud ja kasum tegevusala (EMTAK 2008) ja tööga hõivatud isikute arvu järgi. — [E-andmebaas]  
<http://andmebaas.stat.ee/Index.aspx?lang=et&DataSetCode=em001> (6. mai 2018)
- Eesti Statistikaamet. (2018) EM008: Ettevõtete lisandväärtus ja tootlikkusnäitajad tegevusala (EMTAK 2008) ja tööga hõivatud isikute arvu järgi. — [E-andmebaas]  
<http://andmebaas.stat.ee/Index.aspx?lang=et&DataSetCode=EM008> (6. mai 2018)
- Eesti Statistikaamet. (2017) RAA0042: Lisandväärtus tegevusala järgi (ESA 2010) (kvartalid). — [E-andmebaas]  
<http://andmebaas.stat.ee/Index.aspx?lang=et&DataSetCode=RAA0042> (6. mai 2018)
- Eesti Vabariigi pankrotiseadus. RT 1992, 31, 403.
- Freund, R. J., Littell, R. C. (2000). *SAS System for Regression*. 3rd ed. New York: Wiley.
- Frydman, H., Altman, E. I., Kao, D. L. (1985). Introduction Recursive Partitioning for Financial Classification: The Case of Financial Distress. – *Journal of Finance*, Vol. 40, No. 1, 269–291.
- Gepp, A., Kumar, K. (2015). Predicting Financial Distress: A Comparison of Survival Analysis and Decision Tree Techniques. – *Procedia Computer Science*, Vol. 54, 396–404.
- Gepp, A., Kumar, K., Bhattacharya, S. (2010). Business Failure Prediction Using Decision Trees. – *Journal of Forecasting*, Vol. 29, No. 6, 536–555.
- Goos, P., Meintrup, D. (2016). *Statistics with JMP: Hypothesis Tests, ANOVA and Regression*. 1st ed. West Sussex: Wiley & Sons.
- Grünberg, M. (2013). Pankroti prognoosimise mudeli koostamine Eesti tööstussektori ettevõtete näitel. (Magistritöö). TÜ Ettevõtetemajanduse instituut. Tartu.
- Hayden, E. (2003). Are credit scoring models sensitive with respect to default definitions? Evidence from the Austrian market. (Dissertation Paper). University of Vienna, Department of Business Administration. Austria.
- Hillegeist, S. A., Keating, E. K., Cram, D. P., Lundstedt, K. G. (2004). Assessing the Probability of Bankruptcy. – *Review of Accounting Studies*, Vol. 9, No. 1, 5–34.
- Hinton, P. R. (2014). *Statistics Explained*. 3rd ed. East Sussex: Routledge.
- Hosmer, D. W., Lemeshow, S. (2004). *Applied Logistic Regression*. 2nd ed. New Jersey: John Wiley & Sons.
- Jardin, P. (2009). Bankruptcy Prediction Models: How to Choose the Most Relevant Variables? – *Bankers, Markets & Investors*, No. 98, 39–46.
- Jones, F. L. (1987). Current Techniques in Bankruptcy Prediction. – *Journal of Accounting Literature*, Vol. 6, 131–164.

- Joos, P., Vanhoof, K., Ooghe, H., Sierens, N. (1998). Credit Classification: A Comparison of Logit Models and Decision Trees. – *Application of Machine Learning and Data Mining in Finance*, 59–72.
- Kayri, M., Kayri, I. (2015). The Comparison of Gini and Twoing Algorithms in Terms of Predictive Ability and Misclassification Cost in Data Mining: An Empirical Study. – *International Journal of Computer Trends and Technology*, Vol. 27, No. 1, 21–30.
- Keasey, K., Watson, R. (1991). Financial Distress Models: a Review of Their Usefulness. – *British Journal of Management*, Vol. 2, No. 2, 89–102.
- Kelement, K. (2015). Eesti mööblitööstusettevõtete pankrotimudeli väljatöötamine. (Magistritöö). TTÜ Majandusarvestuse instituut. Tallinn.
- Komisjoni Määrus (EL) nr 651/2014, 17. juuni 2014, ELi aluslepingu artiklite 107 ja 108 kohaldamise kohta, millega teatavat liiki abi tunnistatakse siseturuga kokkusobivaks, 26.6.2014, I lisa, art 2, p 70.
- Korol, T., Korodi, A. (2010). Predicting Bankruptcy with the Use of Macroeconomic Variables. – *Journal of Economic Computation and Economic Cybernetics Studies and Research*, Vol. 44, No. 1, 201–219
- Kraus, A., Litzenberger, R. H. (1973). A State Preference Model of Optimal Financial Leverage. – *Journal of Finance*, Vol. 28, No. 4, 911–922.
- Kumar, S., Chow, T. W. S., Pecht, M. (2010). Approach to Fault Identification for Electronic Products Using Mahalanobis Distance. – *IEEE Transactions On Instrumentation And Measurement*, Vol. 59, No. 8, 2055–2064.
- Laitinen, E. K. (1992). Prediction of Failure of a Newly Founded Firm. – *Journal of Business Venturing*, Vol. 7, 323–340.
- Laitinen, E. K. (1993). Financial Predictors for Different Phases of the Failure Process. – *Omega: The International Journal of Management Science*, Vol. 21, No. 2, 215–228.
- Laitinen, E. K. (1994). Traditional Versus Operating Cash Flow in Failure Prediction. – *Journal of Business Finance and Accounting*, Vol. 21, No. 2, 195–217.
- Li, S., Wang, S. (2014). A Financial Early Warning Logit Model and its Efficiency Verification Approach. – *Knowledge-Based Systems*, Vol. 70, 78–87.
- Lin, T-H. (2009). A Cross Model Study of Corporate Financial Distress Prediction in Taiwan: Multiple Discriminant Analysis, Logit, Probit and Neural Networks Models. – *Neurocomputing*, Vol. 72, 3507–3516.
- Lizal, L. (2002). Determinants of Financial Distress: What Drives Bankruptcy in Transition Economy? The Czech Republic Case. – University of Michigan, William Davidson Institute, Working Paper No. 451.

- Maltz, A. C., Shenhar, A. J., Reilly, R. R. (2003). Beyond the Balanced Scorecard: Refining the Search for Organizational Success Measures. – *Long Range Planning*, Vol. 36, 187–204.
- Markham, I. S., Mathieu, R. G., Wray, B. A. (2000). Kanban Setting Through Artificial Intelligence: a Comparative Study of Artificial Neural Networks and Decision Trees. – *Integrated Manufacturing*, Vol. 11, No. 4, 239–246.
- McKee, T. (2003). Rough Sets Bankruptcy Prediction Models versus Auditor Signalling Rates. – *Journal of Forecasting*, Vol. 22, 569–586.
- McKee, T. E., Lensberg, T. (2002). Genetic Programming and Rough Sets: A Hybrid Approach to Bankruptcy Classification. *European Journal of Operational Research*, Vol. 138, No.2, 436–451.
- Mensah, Y. M. (1984). An Examination of the Stationarity of Multivariate Bankruptcy Prediction Models: a Methodological Study. – *Journal of Accounting Research*, Vol. 22, No. 1, 380–395.
- Micha, B. (1984). Analysis of Business Failures in France. – *Journal of Banking and Finance*, Vol. 8, No. 2, 281–291.
- Miles, J., Shevlin, M. (2000). *Applying Regression and Correlation: A Guide for Students and Researchers*. California: SAGE publications.
- Murthy, S. K. (1998). Automatic Construction of Decision Trees from Data: A Multidisciplinary Survey. – *Data Mining and Knowledge Discovery*, Vol. 2, No. 4, 345–389.
- Ni, J., Kwak, W., Cheng, X., Gong, G. (2014). The Determinants of Bankruptcy for Chinese Firms. – *Review of Pacific Basin Financial Markets and Policies*, Vol. 17, No. 2, 1–22.
- Numani, E. (2015). Bankruptcy And The Altman Models. Case Of Albania. – *Annals of the University of Ordea, Economic Science Series*, Vol. 24, No. 1, 839–845.
- O’Brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. – *Quality & Quantity*, Vol. 41, No. 5, 673–690.
- Ohlson, J. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. – *Journal of Accounting Research*, Vol. 18, No. 1, 109–131.
- Olson, D. L., Delen, D., Meng, Y. (2012). Comparative Analysis of Data Mining Methods for Bankruptcy Prediction. – *Decision Support Systems*, Vol. 52, 464–473.
- Onofrei, M., Lupu, D. (2014). The Modeling of Forecasting the Bankruptcy Risk in Romania. – *Economic Computation and Economic Cybernetics Studies and Research*, Vol. 48, No. 3, 249–267.
- Ooghe, H., Joos, P., De Bourdeaudhuij, C. (1995). Financial Distress Models in Belgium: The Results of a Decade of Empirical Research. – *International Journal of Accounting*, Vol. 30, 245–274.

- Ooghe, H., Spaenjers, C., Vandermoere, P. (2009). Business Failure Prediction: Simple Intuitive Models Versus Statistical Models. – *The IUP Journal of Business Strategy*, Vol. 6, No. 3/4, 7–44.
- Ooghe, H., Verbaere, E. (1985). Predicting Business Failure on the Basis of Accounting Data: The Belgian Experience. – *The International Journal of Accounting*, Vol. 9, No. 2, 19–44.
- Palinko, E., Svoob, A. (2016). Main Causes and Process of Financial Distress. An Empirical Analysis of Hungarian Firms. – *Public Finance Quarterly*, Vol. 61, No. 4, 516–532
- Peel, M. J., Peel, D. A. (1989). A Multilogit Approach to Predicting Corporate Failure – Some Evidence for the U.K. Corporate Sector. – *Omega: International Journal of Management Science*, Vol. 16, No. 4, 309–318.
- Registrite ja Infosüsteemide Keskus. (2018) Algandmed: magistritöös kasutatud Eesti töötleva tööstuse ettevõtete finantsandmed. – [E-andmebaas]  
[https://www.dropbox.com/s/n3px9vq9z9w8xc5/Janar\\_Vahter\\_Magistrit%C3%B6%C3%B6\\_Algandmed.xlsx?dl=0](https://www.dropbox.com/s/n3px9vq9z9w8xc5/Janar_Vahter_Magistrit%C3%B6%C3%B6_Algandmed.xlsx?dl=0) (21.02.2018)
- Scott, J. (1981). The Probability of Bankruptcy: a Comparison of Empirical Predictions and Theoretical Models. – *Journal of Banking and Finance*, Vol. 5, No. 3, 317–344.
- Serrano-Cinca, C., Gutiérrez-Nieto, B. (2013). Partial Least Square Discriminant Analysis for bankruptcy prediction. – *Decision Support Systems*, Vol. 54, No. 3, 1245–1255.
- Shin, K. S., Lee, T. S., Kim, H. J. (2005). An Application of Support Vector Machines in Bankruptcy Prediction Model. – *Expert Systems with Applications*, Vol. 28, No. 1, 127–135.
- Shumway, T. (2001). Forecasting Bankruptcy More Accurately: a Simple Hazard Model. – *Journal of Business*, Vol. 74, No. 1, 101–124.
- Sinha, A. P., May, J. H. (2004). Evaluating and Tuning Predictive Data Mining Models Using Receiver Operating Characteristic Curves. – *Journal of Management Information Systems*, Vol. 21, No. 3, 249–280.
- Tabachnick, B. G., Fidell, L. S. (2007). *Using Multivariate Statistics*. 5th ed. Boston: Pearson.
- Taffler, R. J. (1983). The Assessment of Company Solvency and Performance Using a Statistical Model. – *Accounting and Business Research*, Vol. 13, No. 52, 295–307.
- Topi, H., Tucker, A. (2014). *Computing Handbook, Third Edition: Information Systems and Information Technology*. 3rd ed. CRC Press.
- Tseng, F-M., Hu, Y-C. (2010). Comparing Four Bankruptcy Prediction Models: Logit, Quadratic Interval Logit, Neural and Fuzzy Neural Networks. – *Expert Systems with Applications*, Vol. 37, 1846–1853.

- Tufféry, S. (2011). *Data Mining and Statistics for Decision Making*. 1st ed. West Sussex: Wiley & Sons.
- Urdan, T. C. (2005). *Statistics in Plain English*. 2nd ed. New Jersey: Lawrence Erlbaum Associates.
- Ward, T. J., Foster, B. P. (1997). A Note on Selecting a Response Measure for Financial Distress. – *Journal of Business Finance and Accounting*, Vol. 24, No. 6, 869–879.
- Wilson, N., Altanlar, A. (2014). Company Failure Prediction with Limited Information: Newly Incorporated Companies. – *Journal of the Operational Research Society*, Vol. 65, No. 2, 252–264.
- Zmijewski, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. – *Journal of Accounting Research*, Vol. 22, 59–82.

## LISAD

### Lisa 1. Ekstreemväärtuslikud vaatluspunktid – Mahalanobise ja Cook'i kaugused.

PANKROTISTUNUD ETTEVÕTTED				JÄTKUVALT TOIMIVAD ETTEVÕTTED			
Registrikood	Y	Mahalanobise kaugus	Cook'i kaugus	Registrikood	Y	Mahalanobise kaugus	Cook'i kaugus
11086275	1	123,0837	5,8986	10340369	0	117,0281	2,6564
11519254	1	107,3166	1,4550	10210661	0	101,2304	0,5902
11259965	1	118,9181	1,2404	10584841	0	64,3032	0,2273
11141235	1	126,9339	0,3713	10045593	0	85,8910	0,1636
10464373	1	83,5249	0,2885	10500421	0	116,3373	0,0947
10163453	1	52,5874	0,1997	10857463	0	14,0513	0,0644
10962846	1	108,0010	0,1597	10014842	0	29,3445	0,0479
11092947	1	50,5098	0,0960	10114058	0	37,1888	0,0454
11541963	1	16,5593	0,0824	10823375	0	13,3997	0,0447
12149888	1	9,1886	0,0594	10038765	0	13,2066	0,0405
11914192	1	24,6227	0,0467	10657801	0	34,6587	0,0390
10188231	1	25,5817	0,0324	10803384	0	13,2516	0,0362
10719652	1	19,1363	0,0309	10732626	0	16,1121	0,0304
10965543	1	19,6710	0,0305	10796495	0	7,5890	0,0266
11983951	1	5,2758	0,0259	10726985	0	4,6869	0,0252
11705991	1	4,5530	0,0208	10207067	0	36,9981	0,0218
11736402	1	7,0528	0,0178	10378082	0	58,0422	0,0216
11804428	1	3,8225	0,0131	10468170	0	15,2470	0,0189
10935737	1	7,8704	0,0121	10012984	0	29,9879	0,0176
10057188	1	4,0459	0,0110	10059431	0	6,7505	0,0149
11261198	1	4,6455	0,0096	10052541	0	12,4340	0,0147
11371656	1	20,4003	0,0091	10724058	0	3,7303	0,0140
11054499	1	8,9179	0,0067	10732738	0	8,4200	0,0140
10193508	1	4,5505	0,0064	10327848	0	27,6933	0,0133
10008020	1	2,9579	0,0063	10395560	0	7,8597	0,0130
10055545	1	4,3131	0,0062	10102894	0	41,4427	0,0120
11164756	1	7,5504	0,0061	10842504	0	2,5826	0,0119
11416643	1	16,9412	0,0055	10739054	0	10,0260	0,0117
10014871	1	3,3152	0,0055	10095214	0	10,8228	0,0116

## Lisa 1 järg

Registrikood	Y	Mahalanobise kaugus	Cook'i kaugus	Registrikood	Y	Mahalanobise kaugus	Cook'i kaugus
10231663	1	4,7145	0,0054	10061103	0	9,1931	0,0107
10179655	1	4,8885	0,0054	10570129	0	5,6550	0,0101
11213756	1	5,2016	0,0053	10079296	0	10,4681	0,0094
10029418	1	4,8898	0,0051	10690951	0	3,7547	0,0090
10231338	1	22,1041	0,0049	10107710	0	14,9625	0,0087
10063786	1	4,4139	0,0044	10426272	0	30,2319	0,0085
10904286	1	3,6160	0,0042	10187689	0	15,8547	0,0083
10761906	1	10,5043	0,0040	10480935	0	12,3123	0,0081
11289860	1	0,7904	0,0035	10525353	0	9,3581	0,0077
10221819	1	4,5544	0,0034	10524098	0	4,0624	0,0073
10442590	1	6,8534	0,0032	10131803	0	8,0838	0,0069
10619249	1	16,4787	0,0030	10155382	0	19,5939	0,0066
10039405	1	2,7051	0,0030	10637036	0	8,4113	0,0066
10126660	1	3,1701	0,0030	10021351	0	8,1299	0,0059
10548590	1	14,8520	0,0030	10423925	0	9,7065	0,0053
10101854	1	5,8045	0,0029	10208227	0	9,2494	0,0047
10997626	1	2,8207	0,0029	10174698	0	9,9461	0,0046
10260251	1	2,4924	0,0029	10325973	0	7,0132	0,0046
10096969	1	3,1969	0,0028	10459544	0	8,3743	0,0043
10150447	1	9,9170	0,0028	10538284	0	4,3365	0,0042
10225480	1	2,9102	0,0025	10290571	0	13,3234	0,0041
10172883	1	3,9631	0,0025	10570900	0	5,9098	0,0039
10871718	1	2,5652	0,0024	10116844	0	3,8510	0,0038
10548058	1	7,0403	0,0024	10450980	0	3,4328	0,0034
10044062	1	1,3011	0,0024	10099264	0	6,4803	0,0032
11279815	1	6,7617	0,0023	10381865	0	4,2738	0,0032
10858089	1	6,6312	0,0022	10098508	0	6,1949	0,0032
10049036	1	19,2789	0,0022	10084268	0	2,4200	0,0032
10201580	1	3,6247	0,0021	10345958	0	5,7178	0,0030
10255770	1	37,2937	0,0021	10007925	0	1,6618	0,0029
10312195	1	19,8380	0,0018	10029335	0	4,8819	0,0029
10019704	1	2,5395	0,0018	10083671	0	2,1798	0,0026
10355750	1	2,9419	0,0018	10379681	0	11,6200	0,0025
10138113	1	1,7906	0,0017	10067821	0	2,0485	0,0024
10016054	1	1,8900	0,0017	10122141	0	2,8088	0,0023
11102442	1	1,2244	0,0017	10433651	0	20,5587	0,0022
10027187	1	6,7583	0,0017	10052883	0	4,2984	0,0022



## Lisa 1 järg

Registrikood	Y	Mahalanobise kaugus	Cook'i kaugus	Registrikood	Y	Mahalanobise kaugus	Cook'i kaugus
10178940	1	2,6497	0,0016	10051599	0	4,2292	0,0022
10276559	1	2,7633	0,0015	10502300	0	4,8428	0,0022
10302498	1	6,8167	0,0015	10229488	0	7,9681	0,0019
10149022	1	4,6567	0,0014	10052096	0	4,7039	0,0016
10878637	1	4,9049	0,0013	10051435	0	4,9679	0,0016
10733815	1	2,5478	0,0012	10137697	0	2,3735	0,0015
10046842	1	9,6153	0,0012	10037234	0	1,4180	0,0015
10129893	1	3,7052	0,0012	10272722	0	8,6679	0,0014
10905162	1	4,7802	0,0011	10071490	0	2,3706	0,0014
10237755	1	1,2310	0,0011	10376315	0	9,7951	0,0013
10919365	1	3,2224	0,0011	10084877	0	2,1757	0,0012
10089780	1	1,2965	0,0010	10112763	0	2,7079	0,0012
10321490	1	2,5520	0,0010	10263812	0	4,4832	0,0011
11017363	1	8,7704	0,0010	10070792	0	1,9138	0,0008
10958879	1	1,6777	0,0010	10344918	0	10,1170	0,0008
11072132	1	4,5025	0,0010	10088533	0	3,8867	0,0008
10722616	1	9,8286	0,0008	10045966	0	6,8633	0,0007
10205016	1	1,3356	0,0007	10038162	0	3,7432	0,0007
10898048	1	1,3170	0,0007	10318329	0	2,2412	0,0006
10633261	1	7,8374	0,0006	10077274	0	5,3551	0,0005
10473053	1	6,1422	0,0006	10354727	0	6,1684	0,0005
10838980	1	1,6217	0,0006	10133305	0	1,2894	0,0005
10475129	1	7,6794	0,0005	10149223	0	9,9066	0,0004
10274419	1	5,1263	0,0005	10179359	0	6,3391	0,0004
10178667	1	26,3022	0,0004	10125470	0	2,4710	0,0004
10253713	1	1,3847	0,0004	10393408	0	6,0858	0,0004
10797947	1	4,1325	0,0004	10259064	0	4,3151	0,0004
10819238	1	8,9216	0,0003	10100932	0	2,5949	0,0004
10782958	1	5,3291	0,0003	10265857	0	9,1007	0,0004
10304860	1	0,9982	0,0003	10221216	0	20,0633	0,0004
10840600	1	1,3677	0,0003	10402337	0	9,6697	0,0003
10163884	1	4,2745	0,0003	10300996	0	8,8252	0,0003
10372010	1	5,8579	0,0002	10341311	0	15,3174	0,0003
10321923	1	100,1998	0,0002	10301033	0	12,9315	0,0003
10661837	1	2,9289	0,0002	10336391	0	43,6704	0,0002
10383841	1	9,8471	0,0002	10221417	0	4,0161	0,0002
10295485	1	3,4160	0,0002	10245659	0	5,9560	0,0002

## Lisa 1 järg

Registrikood	Y	Mahalanobise kaugus	Cook'i kaugus	Registrikood	Y	Mahalanobise kaugus	Cook'i kaugus
10367049	1	2,0614	0,0001	10265900	0	3,4350	0,0002
10189010	1	4,9785	0,0001	10270388	0	2,6516	0,0002
10904903	1	3,2084	0,0001	10235454	0	2,7561	0,0002
10824713	1	3,2289	0,0001	10094410	0	8,8264	0,0001
10669253	1	4,1535	0,0001	10237022	0	3,7680	0,0001
10614128	1	2,1287	0,0001	10282360	0	3,3282	0,0001
10920397	1	3,2204	0,0001	10336592	0	3,3753	0,0001
10571896	1	5,3756	0,0001	10359328	0	4,9338	0,0001
10375341	1	3,1612	0,0001	10271378	0	4,0194	0,0001
10553154	1	5,7966	0,0001	10406163	0	6,7331	0,0001
10411483	1	3,6092	0,0001	10120018	0	3,0871	0,0001
10514326	1	2,3578	0,0000	10130136	0	5,2541	0,0001
10628886	1	1,9803	0,0000	10196642	0	10,8787	0,0001
10506551	1	1,7905	0,0000	10225875	0	1,7394	0,0000
10489120	1	34,1768	0,0000	10196978	0	8,8799	0,0000
10517098	1	2,2602	0,0000	10275070	0	4,4114	0,0000
10706646	1	4,6111	0,0000	10254173	0	4,0652	0,0000
10514421	1	4,4925	0,0000	10165162	0	3,3971	0,0000
10678418	1	3,5191	0,0000	10224203	0	1,6284	0,0000
10988283	1	6,3100	0,0000	10317956	0	8,6255	0,0000
10466426	1	2,0497	0,0000	10292529	0	16,6349	0,0000
10654518	1	2,4892	0,0000	10317139	0	21,8693	0,0000
10641378	1	2,3049	0,0000	10086445	0	7,2290	0,0000
10807531	1	4,9696	0,0000	10294184	0	1,1112	0,0000
10576155	1	6,7851	0,0000	10252659	0	4,4437	0,0000
10227213	1	7,2725	0,0000	10240403	0	2,6984	0,0000

Allikas: autori arvutused (SPSS andmetöötlastarkvara rakendamisel)

## Lisa 2. Multikollineaarsuse eemaldamine – VIF test

	Sõltumatud muutujad												
	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13
VIF I	363,3	290,3	255,2	224,9	2,8	9,5	7,2	4,8	26,3	1,1	25,2	1,8	1,7
x1 eemaldatakse													
VIF II	-	5,1	124,9	132,2	2,8	9,3	7,1	4,7	26,1	1,1	23,1	1,7	1,7
x4 eemaldatakse													
VIF III	-	4,8	5,5	-	2,5	9,0	7,0	4,5	25,7	1,1	21,9	1,7	1,5
x9 eemaldatakse													
VIF IV	-	4,3	5,4	-	2,3	9,0	7,0	1,6	-	1,1	5,4	1,7	1,5
x6 eemaldatakse													
VIF V	-	4,2	5,3	-	2,1	-	1,6	1,4	-	1,1	4,8	1,6	1,4
x3 eemaldatakse													
VIF VI	-	3,8	-	-	1,9	-	1,5	1,4	-	1,1	3,1	1,5	1,4

Allikas: autori arvutused (SPSS andmetöõtlustarkvara rakendamisel)

**Lisa 3. Konstrueeritud pankrotimudelite hinnang treeningvalimisse kaasatud ettevõtete pankrotistumise tõenäosusele ( $P_i$ )**

PANKROTISTUNUD ETTEVÕTTED					JÄTKUVALT TOIMIVAD ETTEVÕTTED				
Registri- kood	Y	Logit ( $P_i$ )	CART ( $P_i$ )	CHAID ( $P_i$ )	Registri- kood	Y	Logit ( $P_i$ )	CART ( $P_i$ )	CHAID ( $P_i$ )
10063786	1	1,000	0,870	0,940	10426272	0	0,418	0,000	0,390
10628886	1	0,989	0,870	0,940	10021351	0	0,044	0,080	0,000
10367049	1	0,517	0,870	0,890	10237022	0	0,000	0,080	0,000
10016054	1	0,945	0,870	0,940	10570900	0	0,046	0,000	0,000
10029418	1	0,975	0,870	0,390	10029335	0	0,112	0,080	0,000
10466426	1	0,524	0,870	0,890	10196978	0	0,011	0,080	0,000
10276559	1	0,908	0,870	0,940	10071490	0	0,044	0,080	0,000
10733815	1	0,619	0,870	0,940	10125470	0	0,000	0,080	0,000
10782958	1	1,000	0,870	0,940	10354727	0	0,000	0,080	0,000
10312195	1	0,375	1,000	0,940	10301033	0	0,012	0,080	0,000
10824713	1	0,713	0,870	0,940	10130136	0	0,025	0,080	0,000
10678418	1	0,893	0,870	0,940	10245659	0	0,051	0,080	0,000
10548058	1	0,479	0,870	0,940	10739054	0	0,380	0,000	0,390
10163884	1	0,532	0,870	0,390	10077274	0	0,002	0,080	0,000
10633261	1	0,974	0,870	0,940	10165162	0	0,023	0,080	0,000
10205016	1	0,616	0,870	0,890	10052096	0	0,010	0,080	0,000
10719652	1	1,000	0,870	0,940	10137697	0	0,333	0,870	0,390
10905162	1	0,697	0,870	0,890	10525353	0	0,004	0,080	0,000
10517098	1	0,996	0,870	0,940	10275070	0	0,000	0,080	0,000
10669253	1	0,896	0,870	0,940	10379681	0	0,131	0,000	0,390
10057188	1	1,000	0,870	0,890	10263812	0	0,290	0,080	0,330
10840600	1	1,000	0,870	0,940	10037234	0	0,008	0,080	0,000
10039405	1	0,976	0,870	0,890	10857463	0	0,314	0,000	0,000
10225480	1	0,823	0,870	0,890	10292529	0	0,808	0,870	0,390
10193508	1	1,000	0,870	0,940	10174698	0	0,380	0,080	0,330
10221819	1	0,582	0,080	0,330	10187689	0	0,153	0,870	0,390
10920397	1	0,983	0,870	0,940	10083671	0	0,030	0,080	0,000
10321923	1	1,000	0,870	0,940	10395560	0	0,015	0,080	0,000
10383841	1	0,997	0,870	0,940	10271378	0	0,050	0,080	0,000
10008020	1	0,988	0,870	0,940	10344918	0	0,040	0,080	0,000
11164756	1	0,054	1,000	0,390	10402337	0	0,676	0,870	0,390
10014871	1	0,968	0,870	0,940	10116844	0	0,186	0,000	0,390
10260251	1	0,190	0,870	0,390	10376315	0	0,728	0,870	0,390
10919365	1	1,000	0,870	0,940	10502300	0	0,301	0,080	0,330
10858089	1	0,959	0,870	0,940	10318329	0	0,001	0,080	0,000

### Lisa 3 järg

Registri- kood	Y	Logit (P <sub>i</sub> )	CART (P <sub>i</sub> )	CHAID (P <sub>i</sub> )	Registri- kood	Y	Logit (P <sub>i</sub> )	CART (P <sub>i</sub> )	CHAID (P <sub>i</sub> )
10661837	1	0,672	0,870	0,390	10282360	0	0,023	0,080	0,000
10878637	1	0,718	0,870	0,940	10084877	0	0,025	0,080	0,000
11017363	1	0,971	0,870	0,890	10327848	0	0,193	0,000	0,000
10089780	1	0,705	0,670	0,940	10099264	0	0,000	0,080	0,000
10411483	1	0,491	0,080	1,000	10294184	0	0,011	0,080	0,000
10904286	1	0,614	0,870	0,330	10336592	0	0,319	0,000	0,390
10049036	1	0,988	0,870	0,890	10095214	0	0,192	0,000	0,390
10654518	1	0,999	0,870	0,940	10270388	0	0,021	0,080	0,000
10150447	1	0,972	0,870	1,000	10265900	0	0,021	0,080	0,000
10237755	1	0,283	0,870	0,390	10732626	0	0,000	0,080	0,000
10253713	1	0,299	0,670	0,390	10045966	0	0,001	0,080	0,000
10988283	1	0,999	0,870	0,940	10155382	0	0,120	0,000	0,390
10149022	1	1,000	0,870	0,940	10259064	0	0,664	0,870	0,890
10571896	1	0,803	0,870	0,940	10102894	0	0,003	0,080	0,000
10641378	1	0,910	0,870	0,940	10823375	0	0,794	0,670	0,940
10475129	1	0,890	0,870	0,940	10724058	0	0,002	0,080	0,000
10189010	1	0,962	0,870	0,940	10038162	0	0,006	0,080	0,000
10055545	1	0,689	0,870	1,000	10038765	0	0,020	0,080	0,000
10871718	1	0,998	0,870	0,940	10240403	0	0,000	0,080	0,000
10027187	1	0,326	0,080	0,330	10290571	0	0,001	0,080	0,000
10958879	1	0,816	0,870	0,940	10225875	0	0,031	0,080	0,000
10231338	1	0,811	0,870	0,890	10317956	0	0,252	0,080	0,000
10553154	1	1,000	0,870	0,940	10468170	0	0,078	0,080	0,000
10355750	1	1,000	0,870	0,940	10179359	0	0,541	0,080	0,330
10838980	1	0,995	0,870	0,940	10381865	0	0,214	0,080	0,000
10101854	1	0,993	0,870	0,940	10480935	0	0,024	0,080	0,000
10321490	1	0,459	0,870	0,890	10107710	0	0,000	0,080	0,000
10576155	1	0,983	0,870	0,890	10079296	0	0,000	0,080	0,000
11102442	1	0,511	0,870	0,390	10086445	0	0,061	0,080	0,000
10227213	1	0,907	0,870	0,890	10459544	0	0,408	0,000	0,390
11213756	1	0,736	0,870	0,940	10272722	0	0,801	0,870	0,940
10096969	1	0,817	0,870	0,940	10196642	0	0,078	0,080	0,000
10614128	1	0,978	0,870	0,940	10120018	0	0,022	0,080	0,000
10178667	1	0,135	0,080	0,330	10221216	0	0,117	0,080	0,000
10019704	1	0,508	0,870	1,000	10796495	0	0,000	0,080	0,000
10489120	1	0,921	0,870	0,940	10637036	0	0,928	0,870	0,890
10178940	1	0,603	0,870	0,890	10450980	0	0,017	0,080	0,000

### Lisa 3 järg

Registri- kood	Y	Logit (P <sub>i</sub> )	CART (P <sub>i</sub> )	CHAID (P <sub>i</sub> )	Registri- kood	Y	Logit (P <sub>i</sub> )	CART (P <sub>i</sub> )	CHAID (P <sub>i</sub> )
10179655	1	0,965	0,870	0,940	10100932	0	0,117	0,080	0,330
10126660	1	1,000	0,870	0,890	10341311	0	0,009	0,080	0,000
10372010	1	0,677	0,870	0,890	10098508	0	0,159	0,080	0,000
10506551	1	0,913	0,870	0,940	10433651	0	0,891	0,870	0,940
10619249	1	0,979	0,870	0,940	10051599	0	0,038	0,080	0,000
10797947	1	0,541	0,080	1,000	10345958	0	0,130	0,080	0,330
10935737	1	0,912	0,870	1,000	10061103	0	0,176	0,870	0,330
10761906	1	0,924	0,870	0,390	10359328	0	0,482	0,870	0,390
10274419	1	0,990	0,870	0,890	10059431	0	0,006	0,080	0,000
10548590	1	0,995	0,870	0,940	10131803	0	0,316	0,000	0,000
-	-	-	-	-	10254173	0	0,572	0,080	0,330

Allikas: autori arvutused (SPSS andmetötlustarkvara rakendamisel)

**Lisa 4. Konstrueeritud pankrotimudelite hinnang kontrollvalimisse kaasatud ettevõtete pankrotistumise tõenäosusele ( $P_i$ )**

PANKROTISTUNUD ETTEVÕTTED					JÄTKUVALT TOIMIVAD ETTEVÕTTED				
Registri- kood	Y	Logit ( $P_i$ )	CART ( $P_i$ )	CHAID ( $P_i$ )	Registri- kood	Y	Logit ( $P_i$ )	CART ( $P_i$ )	CHAID ( $P_i$ )
11416643	1	1,000	0,870	0,940	10524098	0	0,000	0,080	0,000
10965543	1	1,000	0,870	0,940	10690951	0	0,556	0,000	0,390
10997626	1	0,999	0,870	0,940	10094410	0	0,057	0,080	0,000
10188231	1	0,824	0,870	0,940	10014842	0	0,849	0,080	1,000
10514326	1	1,000	0,870	0,940	10378082	0	0,000	0,080	0,000
10473053	1	0,976	0,870	0,940	10732738	0	0,413	0,000	0,940
10442590	1	1,000	0,870	0,940	10325973	0	0,052	0,080	0,000
10044062	1	0,985	0,870	0,940	10067821	0	0,012	0,080	0,000
11279815	1	0,029	0,080	0,330	10406163	0	0,476	0,000	0,390
10295485	1	0,863	0,870	0,940	10088533	0	0,071	0,080	0,000
10819238	1	1,000	0,870	0,940	10012984	0	0,196	0,080	0,000
10129893	1	0,721	0,870	0,940	10423925	0	0,436	0,080	0,000
10046842	1	0,999	0,870	0,940	10207067	0	0,871	0,870	0,390
10138113	1	0,998	0,870	0,940	10252659	0	0,000	0,080	0,000
10302498	1	0,965	0,870	0,940	10393408	0	0,001	0,080	0,000
10255770	1	0,999	0,000	0,940	10224203	0	0,006	0,080	0,000
10807531	1	0,994	0,870	0,940	10229488	0	0,001	0,080	0,000
10514421	1	0,735	0,670	0,940	10122141	0	0,011	0,080	0,000
10898048	1	0,823	0,870	0,940	10112763	0	0,002	0,080	0,000
11804428	1	1,000	0,870	0,940	10235454	0	0,120	0,080	0,000
10722616	1	0,758	0,670	0,940	10842504	0	0,446	0,000	0,390
10172883	1	0,815	0,870	0,940	10538284	0	0,088	0,080	0,330
11054499	1	0,972	0,870	0,890	10133305	0	0,035	0,080	0,000
11371656	1	1,000	0,870	0,940	10317139	0	0,871	0,870	0,940
11261198	1	0,896	0,870	0,940	10052541	0	0,627	0,670	0,940
10231663	1	1,000	0,870	0,940	10208227	0	0,000	0,080	0,000
11983951	1	0,219	0,000	0,390	10221417	0	0,002	0,080	0,000
10375341	1	0,987	0,870	0,940	10070792	0	0,000	0,080	0,000
11072132	1	0,853	0,870	0,940	10300996	0	0,001	0,080	0,000
11736402	1	1,000	0,870	0,940	10149223	0	0,008	0,080	0,000
10201580	1	0,994	0,870	0,940	10803384	0	0,000	0,080	0,000
12149888	1	1,000	0,870	0,940	10007925	0	0,158	0,080	0,330
11541963	1	0,906	0,870	0,940	10265857	0	0,004	0,080	0,000
10706646	1	0,998	0,870	0,940	10084268	0	0,376	0,870	0,390
10904903	1	1,000	0,870	0,940	10052883	0	0,006	0,080	0,000

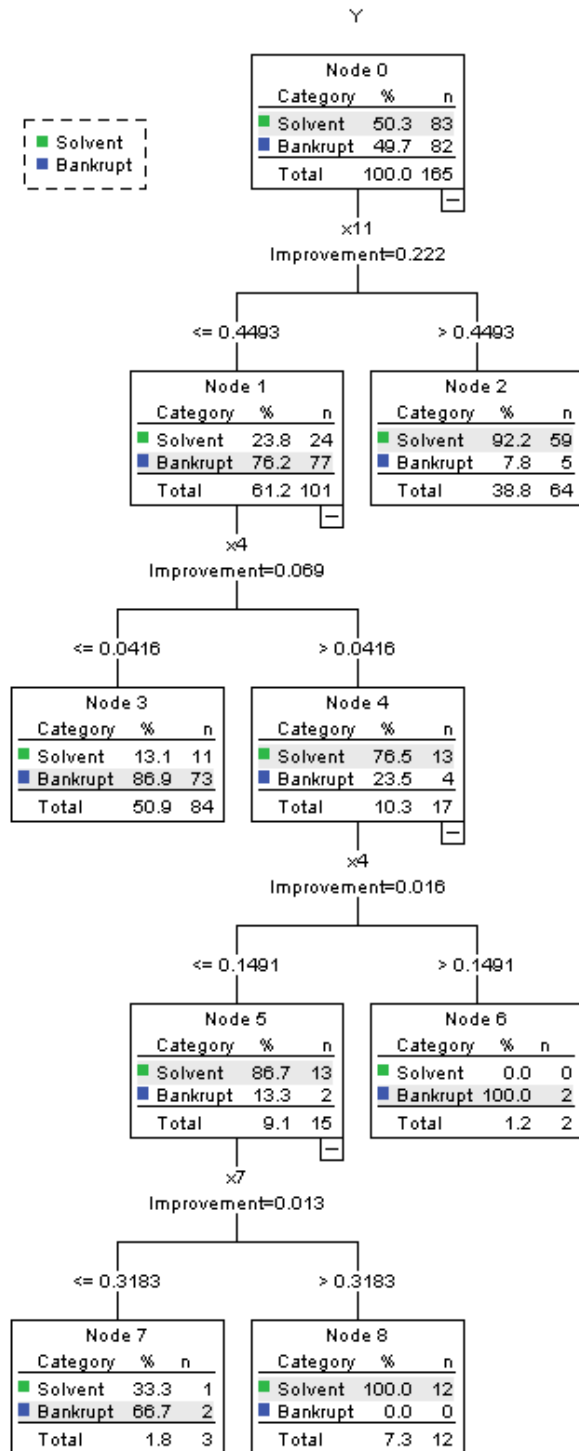
## Lisa 4 järg

Registri- kood	Y	Logit (P <sub>i</sub> )	CART (P <sub>i</sub> )	CHAID (P <sub>i</sub> )	Registri- kood	Y	Logit (P <sub>i</sub> )	CART (P <sub>i</sub> )	CHAID (P <sub>i</sub> )
11705991	1	0,998	0,870	0,940	10570129	0	0,436	0,870	0,330
11914192	1	1,000	0,870	0,940	10726985	0	0,001	0,080	0,000
11289860	1	0,837	0,870	0,940	10051435	0	0,001	0,080	0,000
10304860	1	0,815	0,870	0,940	10336391	0	0,003	0,870	0,390

Allikas: autori arvutused (SPSS andmetöõtlustarkvara rakendamisel)

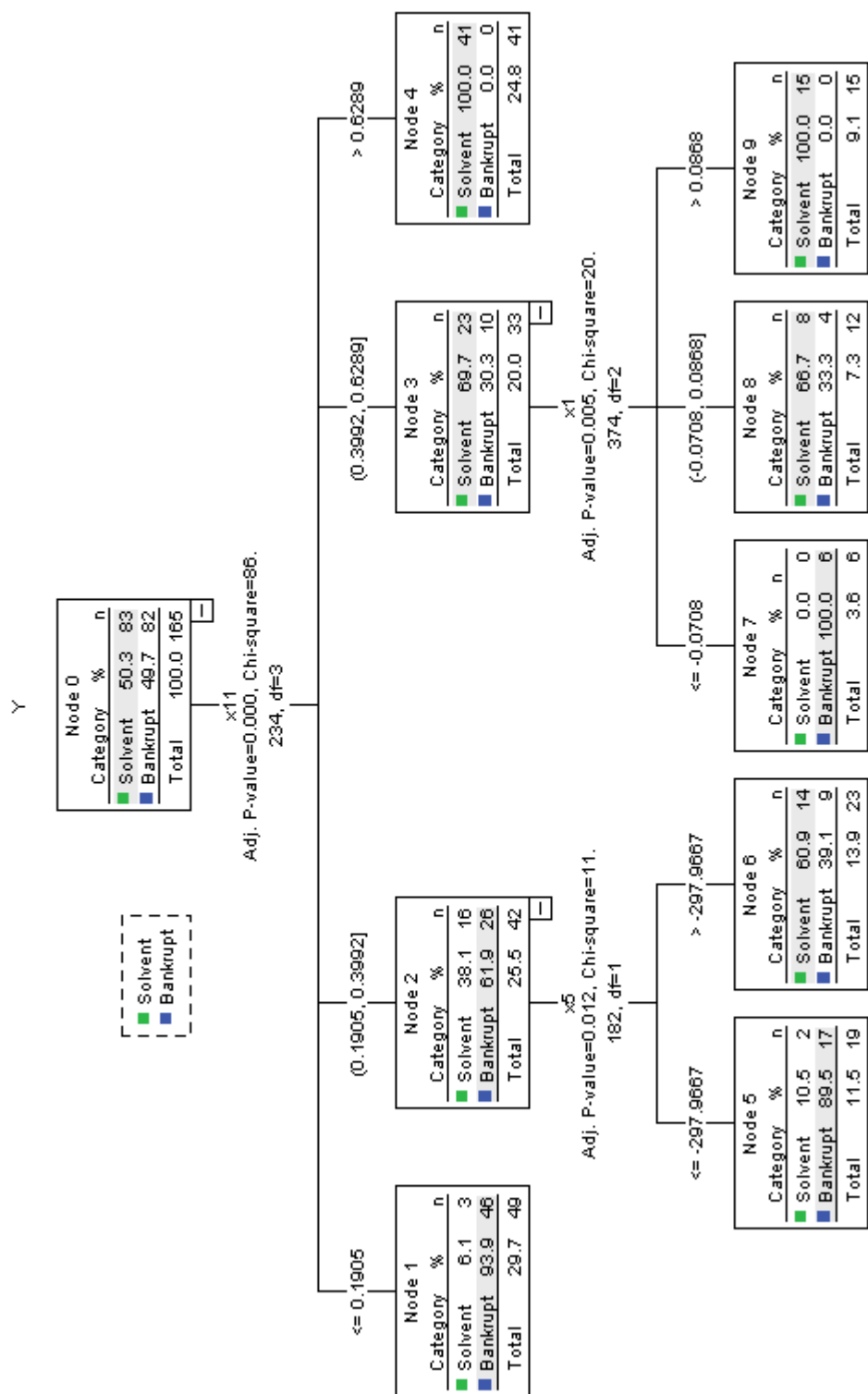


## Lisa 5. Konstrueeritud CART pankrotimudeli struktuur (blokkskeem)



Allikas: autori koostatud (SPSS andmetötlustarkvara rakendamisel)

## Lisa 6. Konstrueeritud CHAID pankrotimudeli struktuur (blokskeem)



Allikas: autori koostatud (SPSS andmetöõtlustarkvara rakendamisel)