

TALLINNA TEHNIKAÜLIKOOL
Majandusteaduskond
Rahanduse ja majandusteooria instituut
Rahanduse ja panganduse õppetool

Martin-Leo Kisand

**OPTIMAALSE TOOTLUSEGA LAENUPORTFELLI
KOOSTAMINE ERAISIKUST INVESTORILE –
OTSUSTUSPUUDE MEETODITE RAKENDAMINE
INIMISELT-INIMESELE LAENUKESKKONNAS BONDORA**

Magistritöö

Juhendaja: lektor Kaido Kepp

Tallinn 2015

Olen koostanud töö iseseisvalt.

Töö koostamisel kasutatud kõikidele teiste autorite töödele, olulistele seisukohtadele ja andmetele on viidatud.

Martin-Leo Kisand

(allkiri, kuupäev)

Üliõpilase kood: 132057TARMM

Üliõpilase e-posti aadress: mkisand@gmail.com

Juhendaja lektor Kaido Kepp:

Töö vastab magistritööle esitatud nõuetele

.....

(allkiri, kuupäev)

Kaitsmiskomisjoni esimees:

Lubatud kaitsmisele

.....

(ametikoht, nimi, allkiri, kuupäev)

SISUKORD

ABSTRAKT.....	3
SISSEJUHATUS	4
1. P2P PLATVORMIDE OLEMUS JA LAENUPORTFELLI KOOSTAMISE PÕHIMÕTTED.....	7
1.1. P2p laenukeskkondade mõiste ja nendes investeerimisega seostuvad riskid	7
1.2. P2p laenukeskkondade turu ülevaade ning Bondora positsioon turul	10
1.3. Laenuportfellide koostamise erinevused võrreldes väärtpaberitesse või võlakirjadesse investeerimisega.....	13
1.4. P2p laenuportfelli koostamise teoreetilised meetodid ja varasemad tulemused	16
1.5. Bondora tööpõhimõtte kirjeldus väikeinvestori vaatenurgast.....	21
2. TÖÖS KASUTATUD METOODIKA JA OTSUSTUSPUU MEETODI VALIK	24
2.1. Andmete ülevaade ja ettevalmistus.....	24
2.2. Intervjuu läbiviimine Bondora esindajaga	26
2.3. Otsustuspuu meetodi valik ja tööpõhimõte.....	27
3. PRAKTILINE LAENUPORTFELLI KOOSTAMINE BONDORAS	35
3.1. Otsustuspuude meetodite rakendamise tulemused.....	35
3.2. Bondora esindajaga läbiviidud intervjuu tulemused.....	42
3.3. Otsustuspuude rakendamise võrdlus Bondora krediidiireitingu meetodiga ning tulemuste igapäevane kasutamine.....	46
KOKKUVÕTE.....	54
VIIDATUD ALLIKAD	57
SUMMARY	62
LISAD	65
Lisa 1. USA ja Euroopa p2p laenuturgu iseloomustavad näitajad.....	65
Lisa 2. Bondora portfelli halduri seadistuste illustratsioon.....	66
Lisa 3. Bondora investeringute otsingufiltri illustratsioon.....	67
Lisa 4. Bondora laenu taotluse näide	68
Lisa 5. Bondora laenuandmetes sisalduvad laenaja tunnused ja nende selgitused	72
Lisa 6. Otsustuspuude mudelitesse valitud laenajat iseloomustavad tunnused	85
Lisa 7. Töös kasutatud valimi puuduolevad väärtused tunnuste lõikes	88
Lisa 8. Bondora esindajaga läbiviidud intervjuu küsimused	89

ABSTRAKT

Töö uuris investeerimist Bondora p2p laenukeskkonnas. Töö uurimisprobleem oli erinevate lähenemiste paljusus ja ühtse standardi puudumine p2p laenukeskkonnas investeerimisel. Töö eesmärk oli selgitada välja optimaalse laenuportfelli koostamise põhimõtted p2p laenamise keskkonnas, rakendades otsustuspuu kontseptsiooni laenuplatvormi empiirilistele andmetele.

Töö käigus Bondora kasutas autor Bondora laenuandmetel üksikutel otsustuspuudel ja *random forest* kontseptsioonidel põhinevaid mudeleid ja võrdles tulemust Bondora krediidireitingute alusel investeerimisega. Otsustuspuu on lihtsasti mõistetav mudel, mis annab täpse tulemuse ja saab hakkama suurte andmemahutudega. *Random forest* kontseptsioon hõlmab korraga mitmete otsustuspuude kasutamist, mis hääletavad parima otsuse üle. Autor kasutas mudelitel klassifitseerimiskulude algoritmi, mis pani mudeleid halva laenu heaks hindamist arvestama viis korda kulukama eksimusena võrreldes hea laenu halvaks tunnistamisega. Lisaks viidi Bondora esindajaga läbi intervjuu.

Autori otsustuspuude mudelid ületasid investori aastast tootlust krediidireitingute alusel investeerimisega võrreldes keskmiselt 12% võrra, olles selleks hetkeks investeerinud 10% kogu turu mahust. Kõige edukamaks osutusid halbade laenude vältimisel *random forest* kontseptsioonil põhinevad mudelid, millest laenuaotlustest maksimaalset infot omav mudel oskas välja pakkuda 94% kõikidest investeeringutest ilma eksimata.

Võtmesõnad: inimeselt-inimesele laenamine, p2p laenamine, sotsiaalne laenamine, laenupoliitika, riskide hindamine, krediidirisk, krediidiskoor, klassifitseerimine, otsustuspuu, *random forest*

SISSEJUHATUS

Väikelaenude rahastamine vahendaja rollis oleva internetikeskkonna kaudu on maailmas suhteliselt uus valdkond. Hulme ja Wright on tutvustanud Ühendkuningriikides 2005. aastal käivitatud Zopa nimelist portaali kui esimest, kes pakub interneti vahendusel investoritele võimalust määrata eelistatud riskitase ning laenudesse investeerida (2006, 17). Sama aastat kinnitas Klafft, kes täpsustas taoliste keskkondade kasu ühiskonnale mh tehingukulude vähendamisest, tulenevalt panga kui vahendaja asendamisest odavamaga internetiportaali ning laenuprotsessi suuremast läbipaistvusest ja paindlikkusest (2008, 1). Käesoleva töö kirjutamise ajaks täitunud kümne aasta jooksul alates esimese portaali käivitamisest on inimeselt-inimesele (ingl k *peer-to-peer*, *p2p*, edaspidi nimetatud *p2p*) laenukeskkonnad panustanud väikelaenude tehingukulude vähendamisse ja seeläbi rahastamise protsessi efektiivsusesse.

On üldteada, et investeerimisportfelli koostamise motivatsiooniks on tavaliselt sellest tegevusest tulevikus kasu saamine, mis väljub piisava intressitulu teenimises eelistatud riskitasemel. Kasu motiiv muudab tegevuse mõtestatuks, aga püstitab kindlate käitumisreeglite vajaduse eesmärgiga valida parim võimalik tee. Henry Markowitz eeldas oma 1952. aastal kirjutatud investeerimisportfelli koostamise teemalises artiklis, et investor peaks soovima oodatavat tulumäära suurendada, vältides tulu kõikumisi ehk riski (1952, 77). Eeltoodud motivatsiooniargumendiga on haakunud paljud laenuportfelli käsitlevad tööd, mis ühelt poolt selgitavad laenudesse investeerimise eripärasid (Gollinger, Morgan 1993, 40; Stevenson, Fadil 1995, 7), teiselt poolt otsivad lahendusi laenude vahel valikute tegemises. Gollinger ja Morgan on viidanud krediidiskoori kasutamisele, mis aitab laenu kvaliteeti ühe näidikuga väljendada (1993, 40).

Lisaks laenude olemuslikule erinevusele väärtpäberitega võrreldes, lisanduvad *p2p* laenukeskkonna kasutamise juures spetsiifilised iseärasused. Näiteks asjaolu, et väljastatud laenudel puudub tihti tagatis, muudab laenaja krediivõimekuse hindamise hädavajalikuks (Mild *et al.* 2015, 1). Teiselt poolt tuleb arvestada, et *p2p* platvorm on ligipääsetav igähele

ning seepärast ei saa eeldada, et investorid koosnevad ainiti finantseksperdist, mis toob kaasa lihtsamate mudelite vajaduse krediidiriski hindamiseks (Klafft 2008, 4). Puudub üldiselt aktsepteeritud parim meetod p2p laenukeskkonnas valikute tegemiseks. Pakutud lahendused jagunevad rusikareeglitest (Klafft 2008, 4) regressioonimudelite abil leitud soovituseni (Mild *et al.* 2015, 7). Hilisemates töödes on esinenud andmekaeve meetodite kasutamist, tekitades näiteks otsustuspuud (Yap *et al.* 2011, 13274; Malekipirbazari, Aksakalli 2015, 4621).

P2p laenukeskkond on alternatiiv pankade rahastamisele, pakkudes võimalust väike-laenusid läbipaistvamalt ja väiksemate tehingukuludega rahastada. Valdus on noor, puudub ühtne standard investeerimisotsuste langetamiseks. Seega on p2p laenukeskkondade arengu püsimiseks aktuaalne uurida ja arendada meetodeid, mis lubavad riske hinnata, et investorite vaatevinklist teenitaks paremat tootlust ja laenaja poolt saadaks õiglasem intressimäär.

Töö uurimisprobleem on, et p2p keskkondades investeerimisel on erinevad seisukohad parima investeerimismeetodi valiku suhtes. See tähendab, et suurema efektiivsuse saavutamiseni on ruumi ning seda lünka tuleb täita, tagamaks taoliste keskkondade arengut ning õiglasemaid intresse laenajatele ja investoritele. Töö uurimisküsimus on, milline on optimaalsem meetod erasikuna väikelaenudesse investeerimisel, kasutades Bondora laenuplatvormi ning soovides saavutada piisavat tootlust minimaalse riskitasemega.

Töö eesmärk on selgitada välja optimaalse laenuportfelli koostamise põhimõtted p2p laenamise keskkonnas, rakendades otsustuspuu kontseptsiooni Bondora laenuplatvormi empiirilistele andmetele. Eesmärgi saavutamiseks on autor püstitanud järgnevad uurimisülesanded.

1. P2p laenukeskkondade olemuse selgitamine ja võrdlemine pangandussektori rahastamismudeliga.
2. Bondora positsiooni iseloomustamine p2p laenukeskkondade kontekstis.
3. Laenuportfellide koostamise iseärasuste väljatoomine, võrreldes traditsiooniliste rahapaigutusviisidega.
4. P2p laenukeskkonnas väljastatavate laenude hindamise meetodite ja väikelaenaja krediidikäitumist mõjutavate tunnuste lahtimõtestamine.
5. Optimaalse investeerimisotsuste tegemise meetodini jõudmine läbi otsustuspuu meetodi rakendamise ja Bondoraga läbiviidud intervjuu ning saadud empiiriliste tulemuste võrdlemise Bondora krediidireitingute süsteemiga.

Esiteks püstitab autor hüpoteesi, et investeerimisel on võimalik otsustuspuuid kasutades tootlust parandada, võrreldes keskkonna krediidireitingute lähenemisega. Teiseks püstitab autor hüpoteesi, et Bondora laenukeskkond pakub investoritele teiste platvormidega võrreldes kõrgeimat keskmist tootlust, arvestades riski.

Töö uurimisobjektiks on Bondora laenukeskkond, mille kaudu on võimalik Eesti eraisikust väikeinvestoril laenuaotlustesse investeerida. Töös on kasutatud peamiselt kahte uurimismeetodit. Bondora laenude ajaloolistel andmetel on kasutatud otsustuspuu lähenemist, mille alusel investeerimisotsuseid langetada. Autor on sisse toonud klassikalise otsustuspuu mudeli ja selle edasiarenduse, nn *random forest* meetodi, mis kasutab lõpptulemuse saamisel korraga mitut otsustuspuud. Autor on tulemusi võrrelnud Bondora krediidireitingute strateegiaga. Töös kasutati Bondora poolt kogutud ja avalikustatud ajalooliseid laenuandmeid, mida allalaadimise hetkeks oli kogunenud ligi 30 tuhande taotluse kohta. Andmed olid Exceli kujul ning sisaldasid erinevaid laenuaotleja tunnuseid, mis lubab neid pidada usaldusväärseks. Lisaks on kvalitatiivse meetodina teostatud intervjuu Bondora esindajaga, eesmärgiga välja tuua Bondora seisukohad investeerimispõhimõtete valikul ning tagasiside otsustuspuu meetodi atraktiivsusest p2p keskkonnas.

Töö on jaotatud kolme peatükki. Esimeses peatükis selgitatakse p2p laenukeskkondade olemust ja riske. Seejärel iseloomustatakse p2p laenude turgu. Järgnevalt käsitleb autor laenuportfellide koostamise iseärasusi võrreldes klassikaliste väärtpaberiportfellidega ning esitab varem kasutatud teoreetilised meetodid ja tulemused. Viimaks seletab autor lahti Bondora keskkonna tööpõhimõtte. Teine peatükk käsitleb töös kasutatud metoodikat. Kolmandas peatükis on esitatud otsustuspuu meetodi ja intervjuu rakendamise tulemused ning järeldused investeerimisel parimaks osutunud otsustuspuude rakendamise kohta.

Töö sihtgrupiks võib pidada väikeinvestorit, kes võib tulemusi kasutada investeerimisstrateegia koostamisel. Töö annab panuse p2p valdkonda uurivale akadeemilisele poolele, sest teemakohast teaduslikku tööd on tehtud suhteliselt vähe. Töö annab üldise ülevaate p2p laenuurust, selgitades p2p laenuäri valdkonna kui alternatiivse rahastamismudeli tausta.

Töö autor avaldab suurt tänu töö juhendajale, Kaido Kepp'ile, igakülgse abi ja heade soovitude eest ning Bondorale ja ettevõtte töötajale Taavi Pertmanile intervjuusoovile vastutulemise ning hilisema tagasiside eest.

1. P2P PLATVORMIDE OLEMUS JA LAENUPORTFELLI KOOSTAMISE PÕHIMÕTTED

1.1. P2p laenukeskkondade mõiste ja nendes investeerimisega seostuvad riskid

P2p laenukeskkonna mõiste võib kokku võtta kui turu, mis ühendab eraisikutest laenuandjad ja laenajad. Viimased peavad esitama laenuaotlused ning laenuandjad ehk investorid otsustavad, millistel tingimustel (ehk intressiga) nad laenu väljastavad. (Mild *et al.* 2015, 1) Käesoleva töö autor lisab, et mõned portaaliid, näiteks Eestis ligipääsetavad Bondora ja Omaraha pakuvad investeerimisvõimalust ka juriidilistele isikutele (Korduma kippuvad ... 2015; Samm-sammult 2015). Samamoodi on juriidilistel isikutel varasemalt Bondora portaalis võimaldatud laenu võtta, mida tõestavad portaali ajaloolised laenuandmed, mida on lähemalt tutvustatud töö teises peatükis (vt lk 24). Seega tuleks p2p laenukeskkonna liikmete (ingl *k peers*) alla arvestada ka juriidilised isikud, kes omale investeerimisportfelli koostavad või laenuvõtjana võõrkapitali kaasavad.

Üheks olulise probleemina p2p laenuplatvormide juures on paljud autorid välja toonud informatsiooni asümmeetria (Emekter *et al.* 2014, 55; Yum *et al.* 2012, 469; Bachmann *et al.* 2011, 7). See tähendab, et laenuandja ei oma laenuvõtja kohta nii palju infot kui viimane ise omab, mis võib viia kehvade investeerimisotsusteni. P2p laenukeskkonnas muudab probleemi tähtsamaks (võrreldes näiteks pankade poolt antavate laenudega) see, et osapooled ei kohtu omavahel füüsiliselt (Emekter *et al.* 2014, 55). Informatsiooni asümmeetria probleemi lahendamine seisab käesoleva töö autori arvates vahendajate õlgadel (kellel on võimalik laenaja kohta võimalikult palju taustainformatsiooni uurida). Võib selgelt väita, et juhul kui vahendaja täidab enda rolli informatsiooni asümmeetria vähendamisel kehvasti (või ei sea seda eesmärgiks), võivad investorid langeda pettuse ohvriks ka siis, kui nad kasutavad laenuaotluses esitatud informatsiooni laenaja kohta maksimaalselt ära. Teiselt poolt ei ole

investoritel otstarbekas ega sageli võimalik laenaja kohta iseseisvalt piisavas mahu taustainformatsiooni välja uurida.

Investori seisukohalt laenuvõtja riski mõjutava aspektina on p2p laenuplatvormide puhul välja toodud, et väljastatavad laenud ei ole tavaliselt tagatisega kaetud (Mild *et al.* 2015, 1; Yum *et al.* 2012, 470). See tõstab riskitaset laenuandja silmis ning seetõttu võib suurendada intressi laenuvõtjale. Lisaks kirjutavad Yum *et al.* oma töös, et iga väikese laenu eraldi üle vaatamine võib olla ajamahukas (2012, 470). Tõepoolest, p2p laenuportaalides võivad laenusummad jääda suhteliselt väikeseks, näiteks Bondora portaali keskmise laenuaotluse summa on ligikaudu 2500 eurot (Üldine statistika 2015), USA-s baseeruva Prosper.com portaali keskmise laenuaotluse summa on ligikaudu 12 800 USA dollarit (sõltub mõõdetava perioodi pikkusest) (Seasoned Returns ... 2015). Veel enam, laenuandja, kes reeglina kogu laenu üksi ei rahasta, teeb Bondora andmetel pakkumuse keskmises summas ca 25 eurot (Üldine statistika 2015). Seevastu panganduses, näiteks eraisikute kodulaenu puhul, oli Eesti Panga andmetel keskmine eluasemelaenu jääk 2015. aasta veebruaris 37 000 eurot (3.3.3 Kodumajapidamistele antud ... 2015). On arusaadav, et pangad peaksid ühe laenu riskide hindamisse panustama rohkem kui eraisikust laenuandja p2p laenukeskkonna vahendusel, sest kaalul olevad rahasummad erinevad kordades. Samas ei ole portfelli hajutava eraisikust investori kohta õige arvata, et tema peaks iga laenuaotluse peale kulutama võimalikult vähe aega pelgalt seetõttu, et laenusumma on üksikult vaadates suhteliselt väike. Pigem on laenuandjal p2p portaalil otstarbekas kasutada teatud standardiseeritud lahendust, mille tulemusel oleks ajakulu ühe laenu ülevaatamiseks minimaalne (arvestades, et ühte laenu panustatud rahasumma on suhteliselt väiksem), aga saadav tulemus piisav, et tervele portfelliga küllaldane tootlus tagada.

Mõned autorid on p2p laenukeskkondade puhul esile toonud nn valikunihke eksisteerimise (ingl k *selection bias*), mis tähendab seda, et p2p laenukeskkondadesse pöörduvad tihti laenuaotlejad, kellele on pangast laenu saamine raskendatud ehk juba eos on tegemist kõrgema krediidiriskiga isikutega (Malekipirbazari, Aksakalli 2015, 4621; Emekter *et al.* 2014, 63). Valikunihke esinemine võib tõepoolest p2p laenukeskkondade puhul olla riski kergitavaks asjaoluks, aga seda enam on tähtis teha igapäevaseid investeerimisotsuseid kindlatele reeglitele toetudes, eesmärgiga valida väiksema riskiga laenusid.

Yum *et al.* on lühidalt ära maininud veel ühe olulise riskiteguri, mis p2p laenamisega kaasneb – nimelt keskkonnast enesest tuleneva riski (Yum *et al.* 2012, 472). Kuigi viidatud artiklis ei ole teemat põhjalikumalt käsitletud, ei maksa käesoleva töö autori hinnangul platvormi riski alahinnata. Arvestades, et keskmise laenupakkumise summa on Bondora platvormis ligikaudu 25 eurot (vt lk 8), võiks näiteks 10 000 euro suuruse investeerimisportfelliga investoril eksisteerida 400 laenulepingut. Juhul kui Bondora peaks otsustama oma tegevuse lõpetada vahendaja rolli teisele ettevõttele üle andmata (või tehes seda ebapiisaval kvaliteeditasemel), jääks väikeinvestor laenulepingutega üksinda. Investori portfelli summaarne väärtus oleks tähelepanuväärne, kuid mitte ühegi lepinguga eraldi ei oleks otstarbekas (näiteks kohtu kaudu võlga sisse nõudes) tegeleda, sest asjaajamise kulud tõenäoliselt ületaksid iga lepingu puhul välja nõutava summa.

Kui portaalist endast tulenevat riski võib autori hinnangul vähendada vaid mitme portaali vahendusel korraga investeerides, siis laenajate erineva krediitkvaliteediga seostuvaid riske on võimalik juhtida, võttes arvesse laenaja kohta kättesaadavat informatsiooni laenuaotluselt või portaali enda poolt välja töötatud hinnangutelt.

Kokkuvõttes saab järeldada, et p2p laenukeskkond on vahendaja, mida kasutavad nii laenuandjad kui –võtjad. P2p keskkonna tüüpiline probleem on info asümmeetria, mis väljendub mittetäielikus infomahus laenuvõtja kohta. Probleemi saab leevendada p2p keskkond, nõudes vajaliku informatsiooni edastamist. P2p keskkonna kaudu laenu andmise riski kasvatab asjaolu, et laenud on sageli tagatiseta. Veel enam, kuna p2p platvormil korraga investeeritavad summad on tihtipeale väikesed, tekib vajadus standardiseeritud otsustusprotsessi järele, sest iga laenu põhjalikuks analüüsimiseks ei ole otstarbekas panustada liialt aega. Lisaks esineb p2p laenukeskkonnas valikunihe, mis tähendab, et laenajateks on tihtipeale isikud, kellele mujalt laenu saamine on keeruline. Viimaks lisandub riskide hulka p2p portaalist endast tulenev risk, mis realiseerudes jätaks investorid suure tõenäosusega laenudega üksinda, sest iga laenuga eraldi asjaajamine oleks liialt kulukas.

1.2. P2p laenukeskkondade turu ülevaade ning Bondora positsioon turul

Bachmann on p2p laenuplatvormid kaheks jaganud: ärilised ja mitte-ärilised (ingl *commercial and non-commercial*). Mitte-ärilised platvormid ei paku kasutajatele intressitulu, vaid võimalust teha annetusi maailma vaesematesse piirkondadesse (Bachmann *et al.* 2011, 5). Käesolev töö käsitleb investeerimist eesmärgiga sellest tulevikus tulu teenida, seega jätab autor järgnevas turuülevaates kõrvale p2p platvormid, mille eesmärk ei ole investoritele seda võimalust pakkuda.

P2p laenuplatvormide mahu hüppelist kasvu on käsitlenud Yum *et al.* töö, mille kohaselt oli 2008. aasta märtsiks üle 20 p2p laenuplatvormi poolt välja antud laene 500 mln USA dollari mahus, lisaks on autorid viidanud 2010. aasta prognoosile, mille kohaselt pidi 2013. aastaks p2p laenuturu maht jõudma 5 mld USA dollarini. Autorid on lisanud, et lähiaastatel võib sotsiaalse laenamise platvormide maht omada 10% suurust osa kogu maailma eraisiku väikelaenude turust. (Yum *et al.* 2012, 469)

Majandusajakirja The Economist andmetel omavad USA p2p laenuturust 98% platvormid Lending Club ja Prosper, kes väljastasid 2013. aastal laene mahus 2,4 mld USA dollarit. Suurbritannia p2p laenuturg ületas mahu 1 mln Suurbritannia naela (1,7 mld USA dollarit) 2014. aasta alguses. (Banking without ... 2014)

Käesolevas töös lähtub autor p2p laenuturu mahtude täpsemal hindamisel USA platvormide Lending Club ja Prosper poolt väljastatud andmetest ning Euroopa alternatiivsete finantsteenuste andmeid koguva keskkonna Altfi.com infost (Altfi.com 2015). Seega jäävad täpsemate andmete puudumisel arvestamata p2p laenukeskkonnad mujal kui USAs ja Euroopas.

Põhjalik visuaalne ülevaade p2p laenuturu hetkeseisust on esitatud lisa (vt lisa 1 lk 65). Autor töötles algandmeid, arvutades viimase 12 kuu kasvust (kuni 2015. aasta märtsini) ja kumulatiivsest kogumahust lähtudes välja 12 kuu jooksul väljastatud laenude mahu ning p2p ettevõtete turuosad selles. Ettevõtted on töö lisa paikneval joonisel järjestatud nende kumulatiivse rahastatud laenude mahu alusel.

Olulised tulemused eelmainitud kokkuvõttest on järgmised. P2p laenuturu mahud on kasvanud kiiresti. Viimase 12 kuu jooksul (kuni märts 2015) rahastati p2p laenuturul 9,1 mld

euro väärtuses laene, võrreldes eelmise perioodiga kasvas maht 145% (see ületab eelnevalt mainitud prognoosi Yum *et al.* poolt). Euroopa kontekstis rahastati viimase 12 kuu jooksul 2,6 mld euro ulatuses laene. Autori saadud tulemused on kooskõlas 2015. aasta veebruaris ilmunud uuringuga Euroopa alternatiivrahastuse mahtude kohta. Uuringu kohaselt tehti kogu Euroopa alternatiivrahastuse turul 2014. aastal 2,96 mld euro ulatuses tehinguid, mis on 144% enam kui varasemal aastal (Wardrop *et al.* 2015, 12). Aastaks 2015 on sama uuring prognoosinud rahastamise mahuks Euroopas 7 mld eurot, mis on üle kahe korra rohkem kui 2014. aastal (*Ibid.*, 40).

Suurima osa p2p laenude turust moodustasid viimase 12 kuu jooksul (kuni 2015. aasta märtsini) USA ettevõtted Lending Club ja Prosper, kes kokku väljastasid ligikaudu 72% laenudest USAs ja Euroopas, sh moodustas Lending Clubi osakaal 53% (suurim turuosaline). Euroopa kontekstis olid turuosad rohkem jaotunud, aga vaatamata sellele oli suurima viie ettevõtte käes 73% turust.

Käesoleva töö lisas paiknevalt graafikult (vt lisa 1 lk 65) selgub veel kaks seaduspära. Vaatamata sellele, et p2p ettevõtete laenumahtude kasvud on olnud kiired, ei ole need olnud suurimate turuosadega ettevõtete puhul (Lending Club, Zopa, Funding Circle) ühtlased, vaid on osaliselt isegi keskmisele turu kasvutempole alla jäänud (keskmine kasvutempo oli 145% viimase 12 kuu jooksul). See võib tähendada, et suurema turuosaga ettevõtete kasvukiirus on hakanud aeglustuma ja nad on jõudmas nõ küpse faasi.

Teiseks on eelnevalt viidatud intresside graafikult näha, et p2p laenuäri suurimad tegijad on pakkunud investoritele pigem keskmisest väiksemaid tootluseid (keskmine intressitulu joonisel toodud ettevõtete hulgas oli 9% aastas), seevastu väiksema rahastatud laenude mahuga ettevõtted (ehk reeglina lühemat aega tegutsenud turuosalised) on pakkunud kohati isegi kaks korda kõrgemat tootlust. Võimalik, et tootlust mõjutab nn valikunihke aspekt, mida käsitleti eespool (vt lk 8). Juhul kui suuremad portaalid on kaasanud rohkem laenajaid (mis on vajalik suuremate laenumahtude tekitamiseks) võib arvata, et nende hulgas on suurem madalama riskiga ehk kvaliteetsete laenusoovijate hulk, mis kokkuvõttes võib keskmise tootluse allapoole tuua (samamoodi väheneb ka risk).

Viimaks on p2p laenuturu käsitlemisel oluline mainida, et erinevalt näiteks väärtpaperiturgudest, mis on üldiselt ligipääsetavad üle kogu maailma, ei ole p2p

laenuplatvormide puhul sageli lihtne investeerida p2p ettevõtte vahendusel, kes ei asu investoriga samas riigis. P2p infoportaal p2p-banking.com on välja toonud vaid kolm ettevõtet, kes pakuvad võimalust investoril nende kaudu investeerida nii, et viimane ei pea omama ettevõttega samas riigis pangakontot ega olema kohalik resident. Nendeks olid töö kirjutamise seisuga Bondora, Ablrate ja Mintos (viimased kaks ei ole töö lisas 1 olevas ülevaates käsitletud). Samas on viidatud artiklis mainitud, et mitmed teised p2p ettevõtted lubavad samuti rahvusvahelisi investoreid, aga sageli rakenduvad lisatingimused nagu näiteks vajalikud suured investeerimismahud. (P2P Lending Services ... 2015) Taaskord on näha, et p2p laenamine kui uus võimalus investeerimisvarade turul hõlmab endas kitsaskohti, mis tõenäoliselt lahenevad valdkonna populaarsuse suurenedes.

Kui liikuda turu üldiselt sentimentidelt Bondora platvormile, siis vaatamata sellele, et Bondora turuosa Euroopa p2p platvormide hulgas oli viimase 12 kuu käsitletud andmete alusel suhteliselt väike (~1%), ületas Bondora laenumahude aastane kasv (161%) turu keskmist (145%) (vt lisa 1 lk 65). See tähendab, et Bondora võib lähiaastatel suuta turuosa oluliselt suurendada, kasvades paljudest teistest p2p ettevõtetest kiiremini (*ceteris paribus*).

Bondora on enda kodulehel esitlenud ajaloolist teenitud netotootlust (arvestab pankrotistunud laenude mõju) suuruses 19,8% aastas (Teenitud tootlus 2015), see on kõrgeim kui teiste vaadeldud turuosaliste poolt pakutav. Kui aga uurida lähemalt Bondora platvormis kaasnevaid riske, siis selgub, et tegelik laenukadude määr kvaliteetsema krediidiireitinguga laenude hulgas on olnud 11,9%, kõrge riskiga laenude hulgas aga 40%. Määramata krediidiireitinguga laenude hulgas oli sama näitaja 16,8%. (Eeldatav tootlus 2015) See tähendab, et p2p laenuturu üldises kontekstis on Bondora pakutav keskmine tootlus atraktiivne, aga Bondoras investeerimisotsuseid tehes võib see oluliselt erineda, sõltuvalt valitud laenudest.

Kokkuvõttes selgub, et p2p laenuturg on kümne aasta vanuselt alates p2p keskkonna Zopa asutamisest (vt lk 4) suhteliselt uus, võrreldes traditsioonilise pangandussektori vanusega ning laenumahud kasvavad aasta-aastalt kiiresti. Turu suurimad tegijad omavad turust kokkuvõttes suurt osa, aga pakutavad intressitulud jäävad mõnevõrra madalamateks väiksematest (ja noorematest) p2p ettevõtetest. See võib samas tähendada, et investor võtab väiksema tulu juures väiksemat riski. Siiski ei saa eraisikust väikeinvestor platvormi riski edukalt maandada, sest vaatamata paljudele tegutsevatele p2p laenukeskkondadele pakuvad

rahvusvahelist ligipääsu vähesed. Bondora p2p platvorm eristub teistest mõnevõrra kõrgema keskmise aastase kasvukiirusega, kõrgeima pakutava intressituluga, aga ka asjaoluga, et Bondora kaudu saavad investeerida mujalt kui vaid Eestist pärit väikeinvestorid.

1.3. Laenuportfellide koostamise erinevused võrreldes väärtpaberitesse või võlakirjadesse investeerimisega

Äriahanduse valdkonnas on üldtuntud teadmine, et tänapäevasele portfelli teooriale pani aluse Henry Markowitz oma 1952. aastal kirjutatud tööga investeerimisportfelli koostamisest (1952, 82). Markowitz tutvustas oodatud tulumäära ja kaasneva riski maksiimi ehk seda, et investor peab teadma oma eelistusi nii tulumäära kui riski osas ja koostama investeerimisportfelli vastavalt. Nii saab valida paljudest võimalikest kõige efektiivsema portfelli, mille puhul on eeldatud riskimäära juures oodatav tulu kõige suurem või soovitud tulumäära juures risk kõige väiksem. Investor peab vältima olukorda, kus võrreldes valituga eksisteerib parem portfelli (suurema oodatava tuluga sama riski juures või väiksema riskiga sama oodatava tulu juures). (*Ibid.*, 82)

Markowitz küll nentis, et võib eksisteerida efektiivne portfelli, mis koosneb vaid ühest väärtpaberist (mis juhtumisi omab väga kõrget oodatavat tulu ja teistest väärtpaberitest madalamat riski), aga tõi välja, et üldiselt on efektiivsed portfelli diversifitseeritud ehk koosnevad väärtpaberitest, mille tootlused ei käitu samadel ajaperioodidel ühtemoodi (*Ibid.*, 82). Käesoleva töö autor aktsepteerib seda tõekspidamist ja eeldab töös kasutatud mudelite puhul, et investor peab otsustamisel arvesse võtma nii oodatavat tulumäära kui kaasnevat riski.

Kui tuua eelkirjeldatud teooria käesoleva töö uurimisobjekti konteksti, siis on Bondora reitingu puhul (lähemalt selgitatud lk 22) tegu informatsiooniga, mille alusel saab investeerimisportfelli koostada tuginedes osaliselt Markowitzi teooriale, pannes eelnevalt paika eelistatud tootluse ja riskitaseme ning valides vastavad portfelli kaalud. Samas peab Bondora reitingutest lähtudes investeerides arvestama, et klassifikatsiooni adekvaatsust ja meetodikat ei ole võimalik iseseisvalt täielikult kontrollida, sest osa tulemuseni viinud informatsioonist ei ole avalikult kättesaadav – ehk tuleb täielikult usaldada vahendajat. Lisaks peab kontrolli aspekti puudumise tõttu arvestama, et riski korrelatsioone laenuaotluste vahel

ei ole võimalik leida vaid tuleks lähtuda infost, mis on parajasti Bondora veebilehe vahendusel iga riskiklassi kohta saadaval (eeldatav tulu ja kaasnev risk).

Liikudes üldisest investeerimisportfelli kontekstist kitsamasse, laenuportfelli käsitusse, rakenduvad portfelli optimaalse koostise leidmisel teatavad piirangud.

Gollinger ja Morgan on enda 1993. aastal kirjutatud artiklis välja toonud järgmised probleemid Markowitzi teooria rakendamisel laenuportfellile (1993, 40):

- Laenude tootluse info ei ole nii hästi kättesaadav kui aktsiate puhul (seda ei avalikustata võrdväärse mahus).
- Laenude tootlused ei ole sümmeetrilised, maksimaalne tootlus on piiratud (võrreldes aktsiatega).
- Kui laenu krediitkvaliteet halveneb tagasimakseperioodi kestel, siis reeglina ei muutu tootlus, sest laenuintress on fikseeritud.
- Laenuportfell ei ole nii likviidne kui väärtpaberiportfell. Rohkem kaubeldakse vaid kvaliteetsemate laenulepingutega.

Autorid on lisanud positiivse tooniga aspekti, nimelt pakub laenuportfelli krediitkvaliteedi informatsioon riski osas põhjalikumat teavet kui seda suudab pakkuda tootluste muutlikkus ajas, mida aktsiate puhul riski mõõdikuna tavapäraselt kasutatakse (Gollinger, Morgan 1993, 40). Tuleb lisada, et krediitkvaliteedi hindamine on laenude puhul jällegi keerukam kui aktsiate tootluse arvutamine.

Oma töös on Gollinger ja Morgan riski mõõtmisel kasutanud enese arvutatud Edward Altmani Z-skoori mudelil põhinevaid ZETA¹ krediitdiskoore, mis baseeruvad ettevõtte finantsnäitajatel. Markowitzi mudeli rakendamiseks on tööstusharusid arvestatud kui ühte (võla) väärtpaberit ning riskina vaadeldud krediitdiskooride varieerumist haru sees. (*Ibid.*, 40) Töö peamine tulemus on, et Markowitzi teooriat saab edukalt laenuportfellile üle tuua (*Ibid.*, 46), samas on mh välja toodud asjaolu, et krediitdiskooride minevikuandmed ei pruugi peegeldada täpselt tulevikku ning seepärast oleks soovitatav, kui pangad arvestaksid krediitdiskooride prognoose (*Ibid.*, 44).

¹ Edward Altmani 1977. aastal loodud mudeli järgi, mis prognoosis *ex ante* ettevõtete pankrotistumist kuni viis aastat ette 70% täpsusega (Altman *et al.* 1977, 31)

Eelnevalt mainitud tööd arendasid edasi Stevenson ja Fadil, kes jõudsid sarnase tulemuseni, et laenuportfelle saab optimeerida ja see on hea lahendus vähendamaks riski, sealjuures piisavat tootlust säilitades (1995, 5). Küll aga tõid nad välja ühe aspekti lisaks Gollingeri ja Morgani tööle – nimelt laenuklasside vaheliste korrelatsioonide leidmise. Kuna pangad hoiavad laene reeglina algusest lõpuni, on laenude tootluste korreleerumist raskem võrrelda. Autorid lahendasid probleemi, võrreldes omavahel vastavate tööstuste pankrotistumise riskide korrelatsioone. (*Ibid.*, 5)

Käesoleva töö eesmärki silmas pidades tuleb lisada, et p2p väikelaenude puhul oleks sama lähenemise rakendamiseks tarvis reeglit laenude gruppidesse jagamiseks (et asendada eeltoodud töödes kasutatud tööstusharusid) ning mõõta riskina näiteks laenude krediitiskooride volatiilsust gruppide sees. Esiteks aga väikelaenud reeglina eraldi tööstusharusid ei esinda (vaid seostuvad eraisikute rahavajadusega), teiseks ei ole võimalik arvutada krediitiskoori ettevõtetega analoogsete näitajate alusel – järelikult on tarvis nii laenude liigitamise kui riski hindamise osas siiski teistsugust lähenemist, mis arvestab p2p laenukeskkonnaga lisanduvaid aspekte.

Küll saab eeltoodud töödest välja lugeda, et laenuportfelli on otstarbekas hinnata riski ja tulumäära omavahelist seost silmas pidades, mitte keskendudes näiteks vaid tootlusele või riskile eraldi (samamoodi nagu Markowitzi lähenemises väärtpaberiportfelli kohta).

Kontseptsiooni riskide hindamisest laenuportfellis täiendasid Jacobson ja Roszbach (2003, 627). Nad leidsid, et portfelli koostamisel ei peaks vaatlema ainult individuaalseid laenuriske (näiteks krediitiskooride alusel), vaid silmas pidama terve portfelli riski, kaaludes laenu riski läbi mahuga. Ehk siis soovitati järgida VaR (ingl k *value-at-risk*) põhimõtet, mis antud kontekstis tähendab kaotust suurusega X laenuportfellis, mis võib Y tõenäosusega kujuneda Z pikkusega perioodis suuremaks (kui X). (*Ibid.*, 627) Lisaks tehti kasutatud mudeli alusel järeldus, et laenude maht ei mõjuta laenuga seotud pankrotiriski väikelaenude puhul (*Ibid.*, 625). Seega tuleneb viidatud artiklist, et ainult krediitiskooride järgi portfelli juhtides võib tekkida riskide kuhjumine (portfell muutub üheülbaliseks) ning, et näiteks suuremahulisemate laenude eelistamine (juhul kui need pakuvad kõrgemat tootlust) ei pruugi hõlmata suuremat riski kui väiksemate laenude puhul ehk siis mahu ja riski vahel ei pruugi seisneda valikukriteeriumit. Viimast teadmist võiks saada otseselt üle viia p2p laenude konteksti hindamiseks seda, kas laenuaotluse mahu ja kaasneva riski vahel peaks esinema seos.

Eelnevalt toodud alapeatüki kokkuvõtteks saab öelda, et laenudest koosneva investeerimisportfelli koostamisel on hea silmas pidada oodatava tulu ja kaasneva riski lähenemist, et mitte valida ebaefektiivset portfelli (saadava tulu juures risk liialt kõrge või võetava riski juures tulu liialt väike). Tuleks käituda sarnaselt nagu näiteks aktsiatest koosnevat väärtpaberiportfelli koostades. Samas kehtivad laenude puhul väärtpaberiportfelliga võrreldes piirangud, mis tulenevad laenulepingute eripäradest börsil kaubeldavate aktsiatega võrreldes ja raskendavad klassikalise Markowitzi lähenemise kasutamist. Laenudega rakenduvad eripärad, mis muudavad protsessi keerukamaks peamiselt tootluse aspektist (tootlus on raskemini mõõdetav, ebasümmeetriline, fikseeritud). Lisaks mõjutab riski laenuportfelli ebalikviidsus. Vaatamata eelnevale on varasemates töödes leitud, et laenuportfelle saab optimeerida riski ja tulumäära seost silmas pidades. Sealjuures tuleb kasutada erinevat lähenemist riski mõõtmisel (nt krediidiskooride näol) ning laenude korreleerumise arvestamisel (nt erinevate tööstusharude vaatlemine). P2p laenukeskkondades portfelli koostamise puhul lisanduvad veel omakorda teatud aspektid, mida käsitleb järgnev alapeatükk.

1.4. P2p laenuportfelli koostamise teoreetilised meetodid ja varasemad tulemused

P2p laenukeskkonna kaudu investeerimisel parimate valikute tegemiseks on tehtud mitmeid empiirilisi töid. Käesolevas töös on varasemalt lähenemised jagatud järgnevas loetelus kolmeks.

- Laenuportfelli koostamine lähtudes lihtsamatest rusikareeglitest või laenuportaali enda poolt pakutavatest soovitustest nagu krediidiskoori alusel investeerimine.
- Laenuportfelli koostamine lähtudes õppimisvõimelistest andmekaeve mudelitest, mis suudavad investori suunata konkreetse otsuse suunas.
- Laenuportfelli koostamine arvestades teiste investorite käitumist.

Nn rusikareglite ja portaali soovitude alusel investeerimise kohta võib esiteks tuua Klaffti nägemuse tüüpilistest investoritest p2p laenuportaalid. Nimelt leidis ta, et harilikud investorid p2p laenuportaalid on eraisikud ilma panganduse ja finantsalase kogemusega –

seepärast on tarvis arusaadavaid ja lihtsasti rakendatavaid reegleid, et investeeringute kasumlikkust parandada. Klafft pakkus välja, et investeerida tuleb laenajatesse kellel ei ole varasemaid võlgnevusi, kelle võlakohustuste suhe sissetulekusse jääb alla 20% ning kes ei ole viimase kuue kuu jooksul laenu taotlenud. Nii saavutatakse investeeringutele kasumlikkus, mis ületab portaali krediidireitingute süsteemi poolt esialgselt pakutavat. (Klafft 2008, 4) Käesoleva töö autor nõustub Klaffti klassifikatsiooniga tüüpilisest investorist, aga on veendunud, et ka keerukamate võtete kasutamine investeerimisel on õigustatud juhul, kui tulemus ületab nõ rusikareegli lähenemist piisavalt. Kasumit maksimeeriva investori jaoks peab ajalise panuse suuruse küsimus taanduma sellele, kas saadav kasu ületab lisaks panustatava ressursi kulusid (mõõdetuna näiteks ajas). Sarnasele tulemusele jõudis autor ka intervjuu kaudu, kus Bondora esindaja selgitas, et kokkuvõttes taandub probleem mh saadavale lisakasule võrreldes krediidireitingute alusel investeerimisega (vt lk 44).

Iyer *et al.* uurisid prosper.com poolt pakutavat ajaloolist infot sealsete laenude kohta ning jõudsid järeldusele, et pankroti tõenäosust ennustab hästi krediidiskoor (2009, 16). Krediidiskoori adekvaatsust hindasid ka Emekter *et al.*, kes viisid läbi regressioonianalüüsi USA p2p laenukeskkonna Lending Club andmetel (2014, 55). Autorid leidsid, et laenu pankroti tõenäosust mõjutavad krediidiskoorid, võla suhe sissetulekusse, FICO² skoorid ja olemasoleva võla suhe maksimaalsesse võlalimiiti (ingl k *revolving line utilization*) (Emekter *et al.* 2014, 69). Juhul kui krediidiskoor on info, mida p2p laenude vahendajast portaal laenude kohta arvutab ja avalikustab, võib seda võtta kui portaalipoolset soovitusi või hinnangut sellele, kas antud laenu taotlusesse tasub investeerida või mitte.

Lihtsustatud lähenemise tõid välja Mild *et al.*, kes uurisid laenude tagasimaksemäära sõltuvust erinevatest parameetritest Taani p2p laenuplatvormil myc4.com. Autorid jõudsid järeldusele, et tagasimaksemäära mõjutavad kohalik krediiti jaotav asutus³, tagasimakseperioodi pikkus (mida pikem, seda kehvem oodatav tulemus) ning tagatise olemasolu (võimalus eksisteeris antud keskkonnas). (Mild *et al.* 2015, 6) Viimase näite puhul on oluline, et myc4.com keskkond annab laenu väikestele ja keskmistele Aafrika ettevõtetele (*Ibid.*, 2),

² Akronüüm Fair Isaac Corporationist (FICO), kes töötab välja valemi, mida kasutavad paljud krediidasutused USAs. FICO skoor arvestab inimese varasemat maksekäitumist, kohustuste mahtu, sissetulekut jmt (Arya *et al.* 2013, 175)

³ Laenusaja ettevõtte riigis paiknev kohalik asutus, kes laenu käsitles (Mild *et al.* 2015, 11)

pigem saab üldistamise eesmärgil esile tuua tagasimakseperioodi pikkuse kui võimaliku laenu kvaliteedi mõjutaja.

Suures osas p2p laenukeskkonnas investeerimist uurivatest töödest on empiirilise meetodina kasutatud klassikalist vähimruutude meetodil põhinevat regressioonianalüüsi, selle tulemusena on leitud konkreetsete laenaja parameetrite (nt eelmainitud laenuaotluse pikkus) võimalik mõju laenu edukusele ehk sellele, kas laen makstakse tagasi või see pankrotistub. Tuginedes kättesaadavale kirjandusele, on käesoleva töö autori hinnangul vähe rakendatud keerukamaid algoritme, mis kasutavad masinõppimist (ingl k *machine learning*) ja andmekaevet (ingl k *data mining*). Siiski ei ole nimetatud valdkond p2p laenukeskkondadel täiesti kaardistamata.

Malekipiribazari ja Aksakalli poolt 2015. aastal avaldatud artiklis uuriti otsustuspuude meetodi rakendamist Lending Club portaalis. Autorid genereerisid juhuslikkuse alusel palju otsustuspuuid, saadud tulemused kombineeriti üheks parimaks otsustuspuuks (lähenemise täpsem nimetus ingl k *random forest*) (Malekipiribazari, Aksakalli 2015, 4628). Meetodi rakendamisel jõuti järeldusele, et see on töötav lahendus parimate laenajate väljaselgitamiseks, ületades näiteks FICO skooride ja Lending Clubi krediidiskooride tulemust (*Ibid.*, 4630). Samas töös katsetati teisigi masinõppimise meetodeid nagu tugivektor-masinaid (ingl k *support vector machines*), regressioon ja k-lähima naabri meetod (ingl k *k-nearest neighbor*) (Malekipiribazari, Aksakalli 2015, 4622). Ühtlasi põhjendati otsustuspuu meetodit kui sellist, mille tulemusi on lihtne tõlgendada (*Ibid.*, 4628). Seepärast aitab mainitud lähenemine lahendada Klaffti poolt esitatud probleemi (vt lk 16) meetodi keerukusest kui võimalikust takistusest keskmisel investoril seda kasutamast seoses finantshariduse puudumisega.

Eelmainitud otsustuspuuid käsitleva tööga vastuolulise tulemuse said Yap *et al.*, kes uurisid krediidiskoori kaardi (ingl k *credit scorecard*), regressioonimudeli ja otsustuspuu rakendamist terviseklubi liikmemaksude andmete peal (tulemusi üldistati finantssektorile, sest paraku ei olnud autoritele laenude empiirilisi andmeid kättesaadaval) (2011, 13784). Töös võrreldi nimetatud kolme lähenemist ning leiti, et tulemused ei erine üksteisest eriti ja seetõttu nõ parimat meetodit ei leidu. Samas mõõnsid autorid, et tulemus sõltub mh kasutatud andmetest (*Ibid.*, 13283). Seega tuleb arvestada, et autorite tulemus oli andmete suhtes tundlik

ning oleks võinud erineda kui kasutatud oleks pankade laenuportfelli või p2p keskkonna ajalooliseid andmeid.

Otsustuspuu meetodit on varem Bondora keskkonnal (mis on käesoleva töö uurimisobjekt) katsetatud. P2p laenukeskkondade uudiseid koondavas portaalis p2p-banking.com on kirjeldatud otsustuspuude rakendamist Bondoras ühe investori poolt. Kahjuks ei selgu artiklist täpset kvantitatiivset hinnangut tootluse paranemise kohta (otsustuspuu alusel alustati investeerimist liialt vähe aega tagasi vastava artikli kirjutamise hetkest), aga esialgsed tulemused olid väidetavalt positiivsed ning eelkirjeldatud teoreetiliste töödega sarnaselt toodi välja asjaolu, et laenude krediitkvaliteedi kujunemine on otsustuspuu mudelis lihtsasti arusaadav. (Decision Trees – Using ... 2015) See annab positiivse signaali kasutamaks sama meetodit käesolevas töös. Lisaks möönab käesoleva töö autor, et mõistlik on investeerimiseks kasutatavat mudelit esmalt testida olemasolevate andmete peal, mitte läbi nõ realse raha kasutamise ja lõpptulemuse ära ootamise (mis lähtus eelmainitud artiklist).

Viimaks on leitud lähenemisi, mis pakuvad välja, et laenuportfelli koostamisel p2p laenukeskkonnas võib olla otstarbekas arvestada teiste investorite arvamusega laenu kohta. Yum *et al.* viisid läbi regressioonimudelil põhineva analüüsi Lõuna-Korea p2p laenukeskkonnal popfunding.com. Nimetatud keskkond pakub unikaalset hääletusmehhanismi, mis seisneb selles, et investorid annavad oma hääle, kas nende arust saab parajasti avatud olev laenu taotlus tulevikus edukas olema või mitte (kas laen makstakse õigeaegselt tagasi) ja portaal avalikustab hääletustulemuse taotluse avatuna püsimise perioodi kestel (Yum *et al.* 2012, 471). Seega saavad üksikud investorid hääletuse tulemust otsustamisel arvesse võtta (kas teha taotlusesse pakkumist). Yum *et al.* leidsid, et hääletustulemused olid seotud laenu hilisema käitumisega (laen kujunes edukaks) olukorras, kus laenaja ajaloo kohta oli vähem teada (varasemalt taotletud laenu kohta ei olnud infot) (*Ibid.*, 471) Autorite põhiline järeldus oli, et kollektiivsest mõtlemisest p2p laenukeskkonnas on kasu ja seda arvestatakse investeerimisotsuste tegemisel (*Ibid.*, 481). Antud lahendus hääletuse mudeli näol seega suunab investoreid enam nõ üksteist kuulama, sest muu lisainfo maht on piiratud (millest lähtuvalt saaks teha kindlamaid iseseisvaid otsuseid).

Investorite käitumise uurimisele pühendusid Herzensteini *et al.* ja Lee *et al.* tööd, milles uuriti karjaefekti (ingl k *herding behaviour*) investorite käitumises p2p laenuportaalides. Herzenstein *et al.* selgitas karjaefekti p2p laenukeskkonnas kui suuremat

tõenäosust, et tehakse pakkumine laenuoksjonile, kuhu juba on teiste taotlustega võrreldes rohkem pakkumisi tehtud (Herzenstein *et al.* 2011, 27). USAs paikneva p2p laenukeskkonna prosper.com andmeid uurinud autorid leidsid, et oluline karjaefekt esineb kuni laen on täielikult rahastatud ning seejärel väheneb⁴. Investoriid saavad karjaefektist kasu esiteks tänu kõrgemale intressimäärale, mida soodustab aktiivsem pakkumistegevus ning teiseks tänu sellele, et karjaefektiga laene seostati kui paremini toimivaid (ehk pankrotirisk oli madalam). (Herzenstein *et al.* 2011, 35) Lisaks võib käesoleva töö autori hinnangul võimaliku kasuna välja tuua asjaolu, et karjaefekti käivitumisel võib laenuaotlus saada kiiremini rahastatud.

Karjaefekti olemasolu leidsid ka E. Lee B. Lee, kes uurisid seda popfunding.com laenukeskkonnal. Autorid leidsid, et karjaefekt on investorite käitumises olemas ning samuti kipub see vähenema siis kui laen on nõ täis saanud. (E. Lee, B.Lee 2012, 512)

Erinevate autorite tööde eelnevalt selgitatud põhitulemused ja probleemid p2p laenukeskkondade uurimisel on ülevaatlikult esitatud järgnevas tabelis.

Tabel 1. Varasemate p2p laenuportaalides tehtud uuringute ja järelduste kokkuvõte

Autorid	Uurimismeetod või -küsimus	Tulemused ja järeldused
(Malekipirbazari, Aksakalli 2015)	Peamiselt masinõppe meetodite katsetamine laenuportfelli koostamise otsuste tegemiseks (k-lähima naabri meetod, regressioonimudel, tugivektor masinad, <i>random forest</i>).	Krediidiskooridega võrreldes (FICO skoorid ja uuritud p2p platvormi hinnangud) valis otsustuspuu teatud tingimustel vähem pankrotistunud laene.
(Yap <i>et al.</i> 2011)	Katsetati erinevaid meetodeid krediidiotsuste tegemiseks: krediidskoori kaart (<i>credit scorecard</i>), regressioonimudel ja otsustuspuu mudel.	Mudelite eksimismäärad (<i>missclassification rate</i>) olid sarnased. Parim mudel sõltub olukorrast (andmed, eesmärk).
(Mild <i>et al.</i> 2015)	Regressioonimudel (sõltuv muutuja oli laenu tagasimaksemäär, sõltumatud muutujad kirjeldasid laenuvõtja tausta).	Teatud tunnused mõjutasid oluliselt tagasimaksemäära. Olulisemad: kohalik krediidi korraldaja, tagasimakseperiood, tagatise olemasolu.
(Emekter <i>et al.</i> 2014)	Regressioonimudel (sõltuv muutuja oli laenu pankroti tõenäosus).	Madala pankrotiriski tõid laenuvõtjad kõrge FICO skoori ja krediidireitinguga ning madala võla/sissetuleku suhtega ja võetud kohustuste arvuga.
(Yum <i>et al.</i> 2012)	Regressioonimudel. Uuriti p2p platvormi hääletusmehhanismi ja küsiti, kas investorite arvamust laenu oodatavast tulemusest tuleks arvestada	Eelneva krediidiajaloota laenajate puhul mõjutasid hääletustulemused laenu edukust. Eelneva ajaloo laenajate puhul (minevikus edukad laenajad), ei

⁴ Karjaefekt ei pruugi täielikult kaduda seetõttu, et mitmed p2p laenuaotlused põhinevad oksjoni meetodil, mis mh tähendab, et kui laen on juba täielikult rahastatud, on endiselt uutel investoritel võimalik pakkumine teha juhul kui nad on nõus eelmisest pakkumisest väiksema intressiga

Autorid	Uurimismeetod või -küsimus	Tulemused ja järeldused
	laenuandmise otsuse tegemisel.	olnud hääletus enam efektiivne.
(E. Lee, B. Lee 2012)	Regressioonimudeli abil uuriti karjaefekti olemasolu laenupakkumiste tegemisel.	Karjaefekt eksisteeris (aga hääbuv kui laen on täis saamas). Toimus ka „viimase-minuti pakkumine“.
(Herzenstein <i>et al.</i> 2011)	Regressioonimudeli abil uuriti karjaefekti olemasolu laenupakkumiste tegemisel.	Karjaefekt eksisteeris (aga hääbuv). Leiti positiivne seos karjaefekti ja laenu hilisema edukuse vahel ehk karjal oli üksteisest kasu. Kasu laenajale seisnes laenu täitumises, intress ei langenud oluliselt.
(Klafft 2008)	Uuriti laenu oodatavat kahjumäära ja selle vastavust oodatavale laenuintressile ning samale intressile siis kui on arvestatud kolme lihtsa reegluga. Tehti arvutused ajaloolisest laenude statistikast lähtudes.	Investeerides vaid reitingute alusel olid investorid osade reitingute puhul kahjumis (intress ei katnud oodatavat kahju). Tulemus paranes oluliselt järgides kolme reeglit korraga: ei investeeeri eelnevate võlgnevustega laenajasse, võla/sissetuleku suhe väiksem kui 20%, viimase 6 kuu jooksul polnud laene võetud.
(Iyer <i>et al.</i> 2009)	Regressioonimudelite abil uuriti erinevate laenu parameetrite mõju intressimääradele.	Krediidiskoor prognoosis pankroti tõenäosust. Ühtlasi kasutasid laenuandjad taotluses olevat muud infot.

Allikas: autori koostatud tuginedes tabelis viidatud allikatele

Kokkuvõttes selgus, et suur osa p2p laenukeskkondi uurivatest töödest on keskendunud lähenemisele, mille eesmärk on välja pakkuda mõned konkreetsed parameetrid laenukvaliteedi hindamiseks. Tulemused on reeglina saadud klassikalist regressioonimudelit kasutades. Uuringuid masinõppel põhinevatel mudelitel on p2p laenukeskkonna puhul vähe, samas on saadud tulemused olnud kohati paljulubavad, mis õigustab selle lähenemise edasist uurimist. Viimaks on välja pakutud alternatiivne meetod investeerimisel, mis kasutab ära p2p laenukeskkonna omapära – nimelt arvestada rahvahulkade teadmise ja teiste investorite huviga või hinnanguga konkreetse laenu taotluse osas.

1.5. Bondora tööpõhimõtte kirjeldus väikeinvestori vaatenurgast

Bondora platvormil investeerimine on sarnane paljude teiste p2p laenukeskkondadega. Investor peab olema täisealine Euroopa Liidu või Šveitsi kodanik või Euroopa Liidu registreeritud juriidiline isik. Bondora on investorid jaganud aktiivseteks ja passiivseteks.

Aktiivsed investorid võivad investeerida laenukeskkonnas nõ käsitsi, valides eelistatud kriteeriumite alusel laenuaotlustest sobivamad. Passiivsed investorid võivad investeringud portfelli halduri nimelise teenuse abil automatiseerida. (Korduma kippuvad ... 2015)

Investor, kes soovib kasutada portfelli halduri teenust, saab sätestada standardsed investimise eelistused ehk portfelli halduri kriteeriumid. Nendeks on investeeritav rahasumma, portfelli soovitud jaotus Bondora laenureitingu alusel ning investeeritav maksimaalne summa ühte laenu, mis kuulub vastavasse reitingugruppi. Lisaks saab valida kolme eeltäidetud eelistustega portfelli halduri vahel, mis erinevad oodatava tootluse ja kahjumäära poolest (Automaatne investimine 2015) Bondora portfelli halduri valikute illustatsioon on toodud lisa 2 (vt lk 66).

Bondora reitingugrupid koosnevad seitsmest riskiklassist, milles erinevad laenu oodatavad tootlused ja kahjumäärad. Bondora veebilehe alusel hinnatakse iga taotluse puhul esiteks riski laenu mikrotasandil (taotluse tasemel), arvestades laenu kriteeriume, teiseks hindab Bondora makrotasandi riske, näiteks riigi riskifaktorit.

Bondora riskigrupi arvutused põhinevad kolme tüüpi andmeallikatel, mis on toodud järgnevas loetelus (Bondora rating 2015):

- kõikidele investoritele avalikult kättesaadavad laenuaotluse andmed;
- välised andmed, mida kontrollitakse teiste andmebaaside abil;
- Bondora veebilehel tehtud toimingute alusel kogutud laenu käitumise iseloomustavad andmed.

Bondora liigituse järgi aktiivne investor saab sobivaid investeringuid valida nõ käsitsi, kasutades isiklikult eelistatud kriteeriumeid. Valida on võimalik veebilehel avatud laenuaotluste vahel ehk nende taotluste hulgast, mis ei ole automaatsete portfelli haldurite poolt täielikult rahastatud. Bondora soovib aktiivseks investoriks hakata nendel, kes omavad suuremat kogemust ning soovivad investeerida suuremaid summasid. (Korduma kippuvad ... 2015)

Kuna korraga võib eksisteerida väga palju erinevaid laenuaotluseid (kuupäeval 30.03.2015 Bondora veebilehte külastades oli avatud laenuaotluseid üleval üle 400), on võimalik otsingufiltri kaudu teha laenu eelselektioon, filtreerida saab kriteeriumite alusel

nagu laenaja asukohariik, laenu pikkus, summa, intress, maksekäitumise ajalugu, haridustase, vanus jne (põhjalikum otsingufiltri kriteeriumite ülevaade on esitatud lisas 3, vt lk 67).

Olles läbi otsingufiltri eelvaliku teinud, kahaneb laenuaotluste valik piisavalt konkreetsete filtreerimiskriteeriumite korral, misjärel saab järelejäänud laenuaotluste nimekirja järjestada reitingu, intressi, eeldatava tootluse, laenuperioodi pikkuse, taotluse täituvuse osakaalu ja taotluse lõppemise aja järgi.

Olles valinud välja konkreetse laenuaotluse, saab tutvuda selles sisalduvate andmetega laenaja kohta. Andmed on suures osas sarnased filtreerimistööriistas kasutatuga, aga taotluses olev info on põhjalikum, sisaldades näiteks infot taotleja eelmiste laenuaotluste ja laenukohustuste kohta eraldi, samuti taotlejale esitatud küsimusi ja nende vastuseid. Näide Bondora lehel olnud laenuaotlusest on esitatud töö lisas 4 (vt lk 68).

Kokkuvõttes saab öelda, et Bondora pakub investorile valikute tegemiseks ühelt poolt nõ rusikareeglite meetodit, nagu seda oli eelnevalt soovitanud näiteks Klafft (vt lk 17), mis väljendub passiivse investori rolli võtmises ja investeerides portfelli halduri teenuse kaudu (ehk lähtudes Bondora panusest laenajate krediitkvaliteedi hindamisel krediidireitingu kaudu). Teiselt poolt on investorile soovi korral avatud keerukamate meetodite rakendamise võimalus, mis kätkeb endas laenuaotluste analüüsimist esiteks filtreerimise teel (lähtudes kindlatest parameetritest), aga ka individuaalselt iga laenuaotluse kaupa eraldi. Viimane meetod võib teoorias tuua veelgi paremaid investeerimistulemusi kui seda suudab pakkuda passiivne lähenemine, sest passiivne meetod hakkab tahes-tahtmata lähtuma teatud turu keskmisest tasemest, mida vastavasse krediidireitingu gruppi asetuvad laenud pakuvad. Siiski on oluline, kas nõ manuaalselt investeerimise meetodi tulemus kaalub üle investori poolt lisaks panustatud ajakulu.

2. TÖÖS KASUTATUD METOODIKA JA OTSUSTUSPUU MEETODI VALIK

2.1. Andmete ülevaade ja ettevalmistus

Käesolevas töös kasutatud andmed hõlmavad endas Bondora laenude toorandmeid, mida Bondora on avalikustanud oma veebilehel Exceli failina (Andmete eksportimine 2015). Andmebaas sisaldab infot Bondoras vahemikus 21.02.2009-10.04.2015 taotletud ja rahastamiseni jõudnud või tagasi lükatud laenude kohta. Autor on andmed alla laadinud 10. aprillil 2015. aastal. Esialgses andmebaasis on kokku 35 344 laenutaotlust, mida iseloomustab 173 kirjeldavat parameetrit mh infoga, kas laen on pankrotistunud või edukalt tagasi makstud. Kõikide parameetrite kirjelduse on autor esitanud töö lisa (vt lisa 5 lk 72).

Kuna käesoleva töö eesmärgiks on selgitada välja optimaalse laenuportfelli koostamise põhimõtted ehk seda, mille alusel investeerimisotsuseid tuleks teha, on autor esialgselt valimit vähendanud ehk ebasobivate laenutaotluste infot välja jätnud järgnevatel alustel.

- Välja on jäetud laenutaotlused, mis ei ole Bondora keskkonnas rahastatud saanud. Selliste taotluste alusel ei saa järeldusi teha tulenevalt sellest, et nende võimalikku pankrotistumist ei saa kindlalt teada.
- Välja on jäetud laenutaotlused, mis andmete laadimise hetkel ei olnud veel tagasi makstud ehk laen ei olnud maksegraafiku lõppu jõudnud. Ka nende laenude pankrotistumist ei ole analüüsi tegemise hetkel võimalik kindlalt teada.
- Välja on jäetud minevikus Bondora poolt väljastatud laenud juriidilistele isikutele. Kuna käesoleva töö kirjutamise hetkel Bondora ettevõtetele laene ei väljasta, ei pidanud autor vajalikuks vastavaid andmeid töös kasutatud mudelite treenimiseks kasutada.

Autor on vähendanud mudelisse sisenevate laenuvõtjat iseloomustavate tunnuste hulka eesmärgiga lõplikku mudelit võimalikult täpseks muuta ning mitte luua mudelit, mis võib soovitada võtta arvesse infot, millega on investeerimisotsuse langetamise hetkel liialt keeruline või lausa võimatu arvestada.

- Välja on jäetud tunnused, mida investor ei saaks laenuotsuse tegemise hetkel arvestada, sest vastavat informatsiooni ei eksisteeri veel (nt laenu rahastamise täitumine, tehtud pakkumiste arv, lõplik kujunenud intress, laenu pankrotistumine).
- Välja on jäetud tunnused, mis on mõne mudelisse sisenenud tunnuse poolt kaetud (nt tunnus taotluse vastuvõtmise kohta ei ole enam oluline kui mudelis on sees vaid kõik rahastatud taotlused).
- Välja on jäetud tunnused, mida ei saa sisuliselt tagantjärele kasutatud andmevalimiga seostada (nt mitu % laenuaotlusest on hetkel täidetud, kas investor on samasse taotlusesse juba investerinud) või millel puudub sisuline selgitus laenude toorandmete andmebaasis.

Kuna autor soovib muu hulgas teada, kas mudelid kujunevad täpsemaks siis, kui neil on kasutada rohkem informatsiooni võrreldes Bondora otsingufiltris pakutud filtreerimisvõimalustega, või mitte, siis tekitas autor kaks erineva arvu tunnustega valimit (samadest laenuaotlustest) ning suunas need ühesugustel põhimõtetel konstrueeritud mudelitesse.

Pärast osade andmete valimist kõrvaldamist eelneval leheküljel toodud põhimõtete järgi, jäi autorile kasutamiseks 4481 laenuaotluse info. Laenuvõtjat iseloomustavate tunnuste hulga korrigeerimise järgselt (nagu eelnevalt selgitatud) jäi autorile kasutamiseks kaks osaliselt kattuvate tunnustega valimit. Tunnuseid, mis võiksid investorile investeerimisotsuse hetkel lisainformatsiooni anda, kuid mida võib leida osaliselt vaid laenuaotlustest eraldi otsides, kujunes kokku 64 tk. Nendest 64-st tunnusest 21 on määratavad Bondora otsingufiltris (vt lisa 2 lk 67) ehk nende tunnustega laenude leidmine süsteemist on investorile oluliselt lihtsam. Mudelitesse suunatud tunnused ja nende sisulised tähendused on toodud töö lisas 6 (vt lk 72).

Vaid eelkirjeldatud tunnuseid silmas pidades kujunes autori jaoks lõplik valim. Selles valimis esinesid mõnede tunnuste kohta (21 tk) puuduolevad andmed, mida otsustuspuude

modelid järelikult täiel määral arvestada ei saanud. Puudolevaid andmeid omavate tunnuste kohta on autor esitanud vastava tabeli töö lisas (vt lisa 7 lk 88).

Olles andmed laadinud Rapidminer tarkvarasse, mida autor mudelite genereerimiseks kasutas, määras autor tarkvara jaoks ära tunnuste tüübid (nt eurodes väljendatud väärtused tüübiga *Numeric*, samas laenuvõtja tegevusala peegeldavad samuti numbrites väljendatud väärtused tüübiga *Polynomial*) – see kindlustas, et mudel ei hakkaks numbriliselt väljendatud väärtuseid, mis tegelikult tähendavad hoopis näiteks erinevaid vastusevariante, valesti tõlgendama.

2.2. Intervjuu läbiviimine Bondora esindajaga

Käesoleva töö raames viis autor läbi intervjuu Bondora esindajaga, kelleks oli Bondora töötaja Taavi Pertman. Taavi Pertman on töötanud Bondoras mitmetel ametikohtadel, puutudes peamiselt kokku Bondora investeerimise poolega (Pertman 2015). Tegemist on p2p platvormidel investeerimist propageeriva inimesega, kes muuseas arendab Eesti p2p maastikku ja üldiseid finantsteemasid käsitlevat portaali Rahafoorum (Rahafoorum 2015). Seega võib arvata, et tegemist on pädeva isikuga nii Bondora seisukohti esindama kui antud töö temaatikat silmas pidades kommenteerima laenuportfelli optimeerimise probleeme.

Käesoleva töö autor otsustas intervjuu meetodi kasutamise kasuks järgmistel asjaoludel. Esiteks, nagu töö 1. peatüki 4. alapeatükis selgus (vt lk 16), on teoorias esitatud mitmeid erinevaid lähenemisi ja soovitusi p2p keskkonnas laenuportfelli koostamiseks. Seega on otstarbekas aktiivselt turul tegutsevalt ettevõttelt välja uurida, milliseid meetodeid peab paremaks Bondora või milliseid kitsaskohti saab välja tuua töös käsitletud alternatiividele.

Teiseks, on eelnevalt Bondora tööpõhimõtte käsitluses selgitatud, et keskkond ise pakub kasutajatele kahte fundamentaalset alternatiivi otsuste tegemiseks – läbi aktiivse ja passiivse investori rollide (vt lk 21). Selleks, et Bondora soovitusi investeerimisportfelli koostamiseks paremini mõista, on oluline välja selgitada, miks on valitud just need meetodid ning miks peetakse neid parimaks.

Kolmandaks, on autor varem selgitanud, et Bondora arvutab krediidiskoori kasutades selleks osaliselt mitteavalikku informatsiooni (vt lk 22). Intervjuu andis võimaluse uurida, kas

ja mil määral on investoril võimalik krediidiskoori arvutustulemust ise kontrollida, eesmärgiga meetodikat paremini mõista ja sedakaudu rohkem investori kasuks tööle panna (kasvõi usalduse tõusu kaudu Bondora meetodi vastu).

Neljandaks, autor on asetanud käesoleva töö empiirilises osas suurt rõhku otsustuspuude meetodite katsetamisele Bondora andmetel ning sel viisil investori investeerimise tehnika kujundamisele, eesmärgiga optimeerida investori laenuportfelli. Töö tulemusele aitab kaasa, kui tagasisidet autori eelduste, kasutatud meetodite ja tulemuste kohta on uuritud seotud osapoolelt, kes ise on valinud teise alternatiivi investorite otsuste kujundamiseks (Bondora on valinud krediidiskooride lähenemise).

Viiendaks, on intervjuu kaudu võimalik välja selgitada esialgu käesoleva töö autori poolt käsitlemata asjaolud, mis p2p laenamise äri juures on tegelikult olulised, aga mille peale autor ise ei ole pruukinud tulla ega ole leidnud asjakohast informatsiooni muudest allikatest. Siia alla käivad küsimused teemal, milliste probleemide teadusliku uurimisega peaks p2p valdkonnas edasi tegelema, et aidata kaasa p2p laenukeskkondade aktiivsele arengule.

Autor koostas intervjuu poolstruktureeritud formaadis, mis tähendab, et ette valmistati vaid põhiküsimused ning intervjuu käigus lasti vastaspoolel rahulikult küsimustele vastata (Viires 2015). Seega jäi autorile võimalus esitada suunavaid spontaanseid küsimusi eesmärgiga katta teemat võimalikult laialt ja avatult.

Käesoleva töö autor viis intervjuu läbi Tallinnas, Bondora kontoris 14. aprillil 2015. aastal. Intervjuust on autoril olemas helisalvestis ja transkriptsioon. Intervjuu käigus esitatud küsimused on esitatud töö lisa (vt lisa 8 lk 89).

2.3. Otsustuspuu meetodi valik ja tööpõhimõte

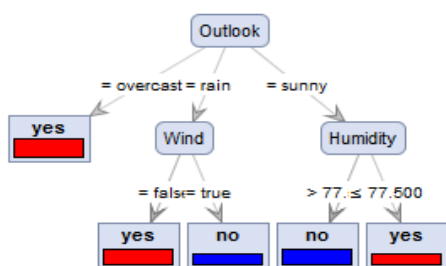
Nagu käesolevas töös varem selgitatud (vt lk 21), on laenude krediidikvaliteedi uurimiseks varasemalt sageli kasutatud regressioonimudelitel põhinevat analüüsi. Cubiles-De-La-Vega *et al.* töös toodi välja mõned tüüpilised probleemid krediidikvaliteedi andmete ja regressioonimudelite kasutamise kohta. Nimelt on krediidimudelites kasutatud andmed sageli väga mitmekesised – kõrvuti on esitatud kvalitatiivne ja kvantitatiivne informatsioon ning

sõltuva muutuja ja valitud sõltumatute muutujate (ehk laenaja parameetrite) vahel ei ole tavaliselt lineaarset seost. Selliseid probleeme aitab edukalt lahendada masinõppel põhineva algoritmi kasutamine. (Cubiles-De-La-Vega *et al.* 2013, 6911) Masinõppe meetodi alla kuulub ka otsustuspuu.

Otsustuspuu valikut põhjendavad autori hinnangul järgnevad asjaolud. Esiteks on autori hinnangul oluline mõista, et Bondora poolt avalikustatav varasemate laenude informatsioon (koos infoga selle kohta, kas laen on edukalt tagastatud või mitte) on *a priori* väärtuslik tööriist, mida oleks ebamõistlik kasutamata jätta, omandamaks põhjalikumat veendumust ja teadmist selles osas, kas teatud meetoditel investeerimine oleks varasemalt Bondora keskkonnas efektiivne olnud või mitte. Nii on võimalik hüpoteese kontrollida ilma rahaliste ressurssidega riskimata. Teiseks, nagu käesolevas töös varem viidatud, on investori igapäevaseks opereerimiseks p2p laenukeskkonnas vaja standardiseeritud meetodit (vt lk 8), sest vastasel juhul kujuneb töö ühe laenu analüüsimiseks liialt mahukaks, arvestades laenu tehtud rahalise panuse suurust. Viimaks on varasemad autorid välja toonud, et otsustamiseks on vaja lihtsamaid reegleid (vt lk 17), sest investoritel on erinevad finantsteadmised ja taust.

Otsustuspuu lähenemine on mitmete autorite poolt välja pakutud kui lihtne aga võimas klassifitseerimise mudel, mis samas annab täpse tulemuse ja töötab hästi suurte andme- mahtude peal (Basti *et al.* 2015, 15–16; Patil, Bichkar 2012, 13; Malekipirbazari, Aksakalli 2015, 4627).

Otsustuspuu väljundi võib kokku võtta järgmiselt. Tegemist on reeglite kogumiga, mis on hierarhiliselt moodustatud formaadis „kui ... siis“ (ingl k „if ... then“) (D. Zhang, Zhou 2004, 515). Olulist väärtust lisab asjaolu, et tulemus näeb visuaalselt lihtsasti mõistetav välja, meenutades puud. Näitlik joonis visualiseeritud otsustuspuust ja tema väljakirjutatud reeglistikust on esitatud järgnevalt.



Tree

```

Outlook = overcast: yes {no=0, yes=4}
Outlook = rain
| Wind = false: yes {no=0, yes=3}
| Wind = true: no {no=2, yes=0}
Outlook = sunny
| Humidity > 77.500: no {no=3, yes=0}
| Humidity ≤ 77.500: yes {no=0, yes=2}

```

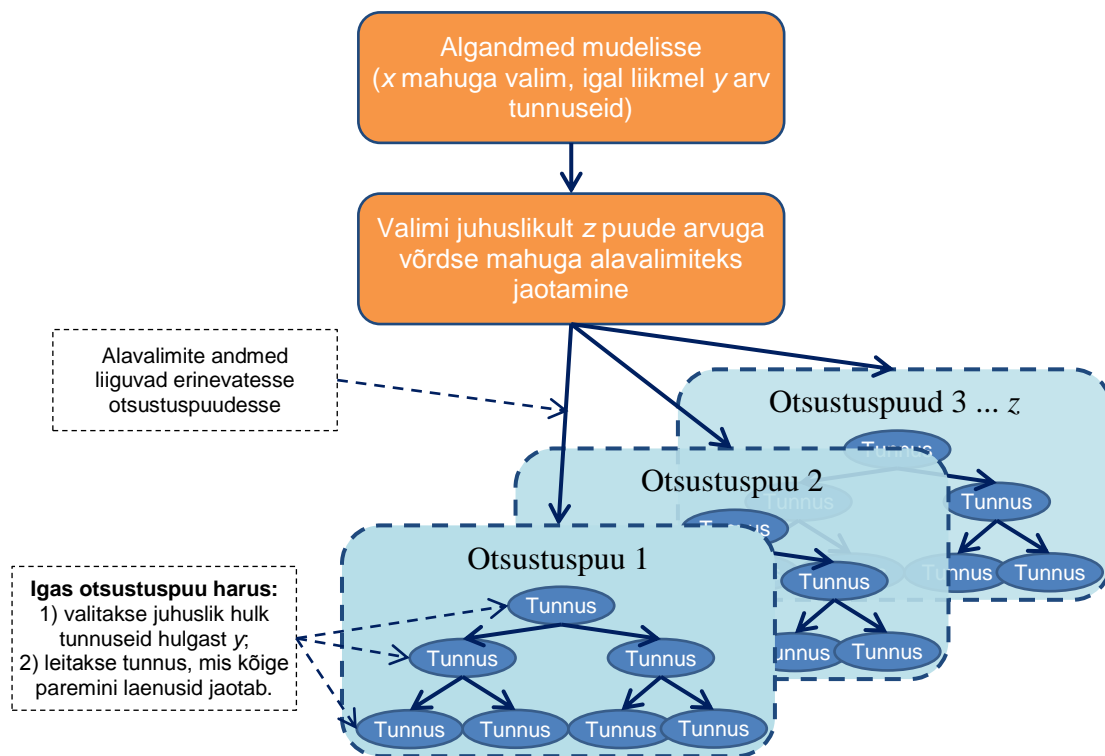
Joonis 1. Otsustuspuu visuaalne ja väljakirjutatud reeglitega vaade. Allikas: autori koostatud andmekaeveprogrammi Rapidminer näiteandmete alusel

Eeltoodud joonisel on algoritm lähtunud 14 sündmusega valimist, mis käsitleb päevi, mil golfi mängiti (või ei mängitud) ja samadele päevadele kehtinud ilmaprognoose. Kui näiteks golfiväljaku juhataja peaks otsustama, millistel päevadel väljak avada, siis oleks tema kasutada varasematel juhtumitel põhinev tulemus, et näiteks kõikidel pilvistel ilma vihmata päevadel (4 tk) on golfi mängima tulnud, samas vihmaste päevade puhul (5 tk) on tulnud vaid tuulevabadel (3 tk) jne. Eelneva lihtsustatud näite põhjal saab väita, et otsustuspuu võib pakkuda lihtsasti rakendatavat standardit otsuste kiireks langetamiseks.

Otsustuspuu miinustena on paljud autorid välja toonud, et tulemuseks saadud puu võib kujuneda liialt suureks (ehk koosneda paljudest erinevatest harudest) ja seetõttu raskesti mõistetavaks ja kasutatavaks, aga mis kõige olulisem, tema üldistusvõime väheneb ja mudel muutub üle-kohaldatuks (ingl k *over-fitting*) (Cubiles-De-La-Vega *et al.* 2013, 6912; Patil, Bichkar 2012, 13; Malekipirbazari, Aksakalli 2015, 4628). Viimase probleemi lahendamiseks kasutatakse meetodeid puu harvendamiseks (ingl k *pruning*), mis eemaldavad otsustuspuu harud, mis oluliselt ei paranda mudeli täpsust ja asendavad need lehtedega (Patil, Bichkar 2012, 15). Lehtede all mõistetakse puu lõpp-punkte, mis enam edasi ei hargne.

Optimaalse suurusega otsustuspuuks on Patil ja Bichkar nimetanud sellist, mis klassifitseerib sõltuva muutuja õigesti ja omab seejuures minimaalset harude arvu (2012, 19).

Otsustuspuude lähenemise edasiarendatud versiooniks võib nimetada nõ juhusliku metsa lähenemist (ingl k *random forest*, edaspidi RF). Seda on varem kasutatud Lending Clubi p2p laenukeskkonna andmete peal, kus RF lähenemine andis krediitiskooride alusel investeerimisega võrreldes oluliselt paremaid tulemusi (vt lk 20). RF lähenemine baseerub otsustuspuudel, mida on korruga kasutusel mitu (näiteks 50 tükki). RF lähenemise puhul rakendub juhuslik valik kahes etapis. Esiteks on iga puu saanud valimist juhusliku hulga liikmeid. Teiseks, iga otsustuspuu on saanud endale samast andmekomplektist juhusliku hulga tunnuseid hargnemise otsustamiseks (Breiman 2001, 10; Cubiles-De-La-Vega *et al.* 2013, 6914). Lõpuks annavad kõik otsustuspuud konkreetse laenuaotluse kohta hinnangu ehk hääletavad (Breiman 2001, 6). P2p laenuäri kontekstis tähendab see, et mudelid hääletavad, kas laen makstakse edukalt tagasi või mitte. Järgneval joonisel on kujutatud RF lähenemise tööpõhimõtte visuaalselt.



Joonis 2. *Random forest* põhimõttel töötava mudeli tööpõhimõtte skeem. Allikas: autori koostatud *random forest* tööpõhimõtete kirjelduste ja illustreerimiste alusel

RF lähenemist on peetud üheks efektiivsemaks masinõppe meetodiks ning klassikalise otsustuspuu mudeli tulemusi parandavaks meetodiks. Peamiselt sellepärast, et vaid ühel otsustuspuul võib olla suur tundlikkus andmete suhtes või liigne kallutatatus (viimase puhul üle kohaldamise probleem). (Walker 2013) Lisaks on järeldatud, et kuna iga RF meetodi abil kasvatatud puu hõlmab endas juhuslikkust, siis on nende puude omavaheline korrelatsioon madal (Florez-Lopez, Ramon-Jeronimo 2015, 5740; Malekipirbazari, Aksakalli 2015, 4627). See lisab RF puude „arvamusele“ rohkem usaldusväärsust, sest nad käituvad kui mõnevõrra erinevatel kaalutlustel otsuseid tegevad üksused.

Töö sissejuhatuses püstitas autor hüpoteesi, et otsustuspuude kasutamise abil on võimalik Bondoras investeerimisel tootlust parandada, võrreldes keskkonna poolt pakutavate krediitdiskooride lähenemisega.

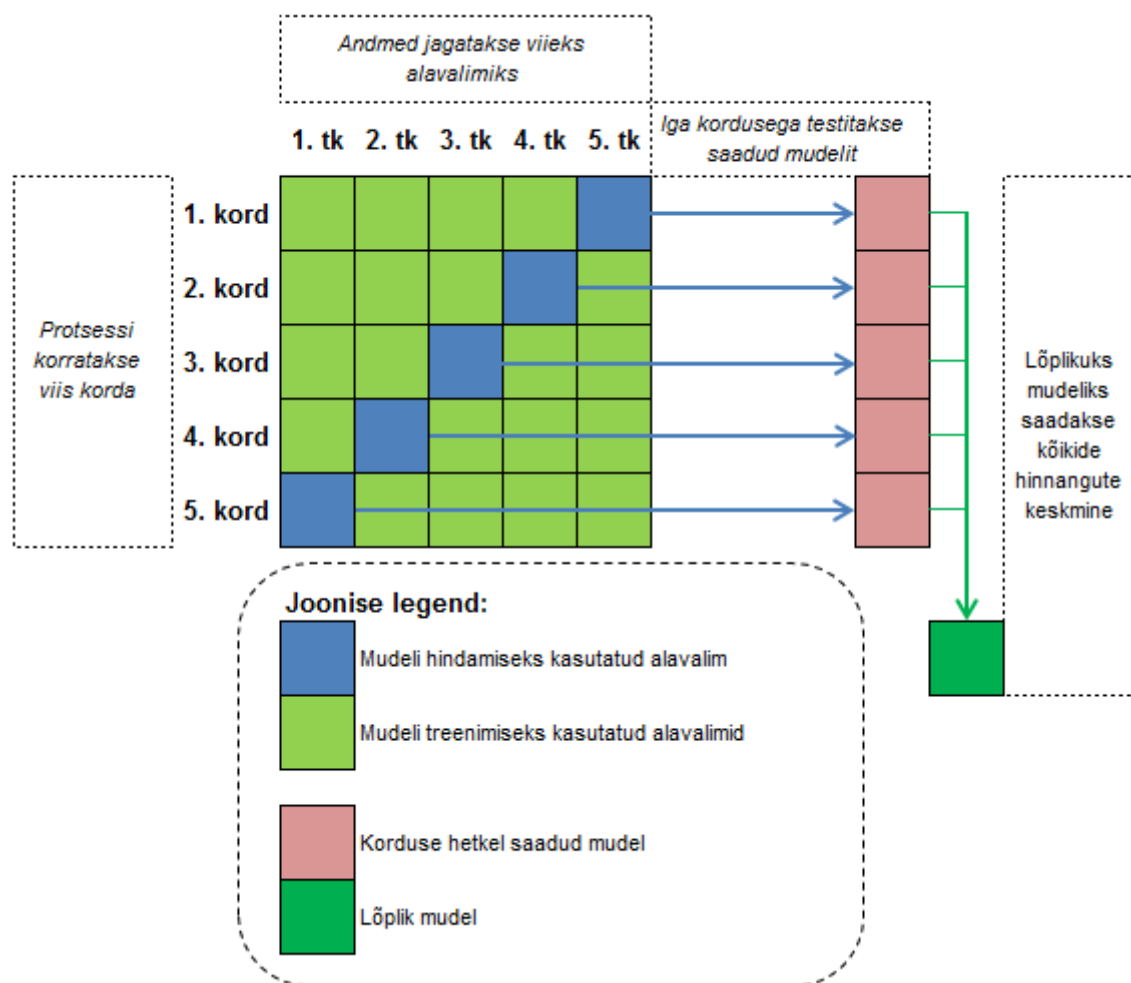
Selle hüpoteesi tõestamiseks on autor otsustuspuudele tuginevaid mudeleid rakendanud järgnevalt, selgitamaks välja täpsemat mudelit.

- Otsustuspuu ja RF mudelite kasutamine lähtudes laenaja tunnustest, mida saab valida Bondora otsingufiltrist (tunnuste arv on väiksem).
- Otsustuspuu ja RF mudelite kasutamine lähtudes kõikidest laenaja tunnustest, mida saab laenuaotlustest leida ja mille kohta Bondora on avaldanud ajaloolist statistikat (tunnuste arv on maksimaalse suurusega).

Autor otsustas mudeleid katsetada kahel erineva mahuga tunnuste grupil tulenevalt sellest, et välja selgitada, kas väiksemate tunnuste arvuga mudel osutub rohkemat infot omavast mudelist oluliselt erinevaks. Kui mitte, siis on mõistlikum praktiliselt kasutada mudelit, mis on andmete suhtes vähenõudlikum ja seeläbi efektiivsem. Ühtlasi on sellisel juhul võimalik lihtsamini mudeli jaoks vajalikku andmehulka koostada (kulub vähem aega).

Otsustuspuu mudelite treenimisel ja valideerimisel on autor kasutanud nn ristvalideerimise meetodikat (ingl k *cross-validation*), mis töötab järgneval põhimõttel. Kogu valim jagatakse osadeks arvuga k . Üks osa k -st jäetakse lõpliku mudeli kontrollimiseks ning ülejäänud $k-1$ osad kasutatakse mudeli treenimiseks (andmed, mis lähevad mudeli ehitamiseks). Protsessi korratakse k korda kuni mudel on lõpuks testitud k andmehulga peal, hindamistulemustest võetakse keskmine. Ristvalideerimise meetodikat seletab visuaalselt skeem järgneval leheküljel.

Ristvalideerimise eeliseks on peetud asjaolu, et k suurenedes mudeli tulemuse dispersioon väheneb (ehk mudel muutub stabiilsemaks, tugevamaks). (Schneider 1997) Sama meetodit on otsustuspuude juures ka varem kasutatud (Malekipirbazari, Aksakalli 2015, 4628; Florez-Lopez, Ramon-Jeronimo 2015, 5738; Cubiles-De-La-Vega *et al.* 2013, 6914). Autor on valinud k väärtuseks 5 nagu seda tehti Malekipirbazari ja Aksakalli töös otsustuspuude rakendamise kohta.



Joonis 3. Ristvalideerimise tööpõhimõte. Allikas: autori koostatud ristvalideerimise meetodi kirjelduste ja skeemide alusel (Malekipirbazari, Aksakalli 2015, 4628)

Töö teoreetilises osas (vt lk 14) toodi laenuportfelli iseärasuste osas välja, et laenude tootlused on tavapäraselt ebasümmeetrilised – maksimaalne tootlus on piiratud, aga võimalik kaotus on kuni 100% laenu investeeritud summast. See tähendab, et tegelikult võib mudeli poolt antud positiivne hinnang laenule, mis hiljem pankrotistub, olla investorile kulukam (kaotab 100% investeeritud summast) kui pankrotti prognoosiv hinnang laenule, mis tegelikult pankrotti ei lähe (investor jääb ilma vaid teenitavast intressitulust, sest ta otsustab mitte tehingusse astuda). Ebatäpsete hinnangute erinevat kulukust on käsitletud ka teised autorid, otsustuspuude juures näiteks Vadera, kes tõi näite, et meditsiinivaldkonnas on parem vältida vale-positiivset diagnoosi (terve patsient diagnoositakse haigeks) kui vale-negatiivset diagnoosi (haige patsient diagnoositakse terveks) (Vadera 2005, 2). Klassifitseerimiskulude osas tundlikku mudelit kasutasid ka Malekipirbazari ja Aksakalli otsustuspuud käsitlevas töös,

võttes nende kulude suhteks 5 ehk halva laenu ekslikult heaks hindamise kulu hinnati viis korda suuremaks hea laenu halvaks lugemisest (Malekipirbazari, Aksakalli 2015, 4628). Kim *et al.* rõhutasid enda töös klassifitseerimiskuludest, et mudelid, mis klassifitseerimiskulusid ei arvesta, võivad viia finantsinstitutsioonid valedele otsustele ning seetõttu põhjustada tõsisaid rahalisi kahjusid (Kim *et al.* 2012, 4014). Kulude osas tundlikku otsustuspuud on kasutatud ka muudes valdkondades nagu tarkvaraarendamise vigade hindamine (Siers, Islam 2015, 62), aga näiteks ka laenude mahakandmise hindamisel (Czajkowski *et al.* 2015, 1). Võttes aluseks, et Bondora pakub investori keskmiseks aastaseks tootluseks *ca* 20% (vt lisa 1 lk 65) võib teha üldise (lihtsustatud) järelduse, et keskmised laenuintressid jäävad 20% lähedale (loomulikult sõltub see eelistatud riskiklassist ja muudest teguritest ning võib leiduda keskmisest oluliselt kõrgemaid ja madalamaid intressimäärasid). Tuleneb, et investor võiks pankrotistunud laenuga kaotada viis korda rohkem⁵ (100:20) kui võita – sestap kehtestab käesoleva töö autor mudelile reegli, et pankrotistuva laenu heaks lugemine on viis korda kulukam kui edukaks osutuva laenu pankrotistuvaks hindamine.

Klassifitseerimiskulude rakendamiseks on autor kasutanud lähenemist nimega MetaCost, mis oli autori kasutatavas Rapidminer tarkvaras toetatud meetod. MetaCost lähenemise töötas algselt välja Pedro Domingos, kes kirjeldas klassifitseerimiskulude arvestamise protsessi järgnevalt (Domingos 1999, 157).

1. Mudel teab eelnevalt klassifitseerimiskulusid läbi kulude maatriksi. Treeningvalimist võetakse m alavalimit (võetakse sisse osa andmeid, koos asendamisega⁶).
2. Luuakse mudel igal alavalimil (mudel, mida parasjagu kasutatakse, see ei sõltu MetaCost meetodist). Iga mudel annab teatud tulemuse, tulemused agregeeritakse, vaadeldes saadud hinnangute täpsuse tõenäosust (kas mudel on hinnanud tulemuse õigesti) ja hinnangu kulu.
3. Iga valimi liige märgitakse uuesti ära, kirjutades tulemusse eelnevalt saadud prognoositud tulemuse vastavalt eelnevas etapis saadud mudelite hinnangutele.
4. Protsessi korratakse m korda.

⁵ Kui eeldada, et investor jääb lisaks kaotatud põhiosale ilma intressitulust, oleks koefitsiendiks kuus, autor jätab selle eelduse käesolevas töös siiski kõrvale (subjektiivne eelistus lähenemise valikul)

⁶ Võtmine koos asendamisega (ingl k *sampling with replacement*) tähendab, et valimist andmete väljavõtmise järel jääb üldkogum endiseks, mistõttu on kõikidel valimi liikmetel endiselt sama suur võimalus uuesti valitud saada (Parker 2015)

Teisiti väljendudes mõjutab Domingose meetodika mudelit järgnevalt. Kui näiteks mudeli valepositiivsete ja valenegatiivsete hinnangute kulu oleks ühesugune (ehk õiged ja valed otsused oleksid sama kaaluga), siis annaks töös kasutatav RF põhimõttel töötav otsustuspuude mudel oma positiivse hinnangu ka siis, kui laenu loeksid õigeks 50%+1 mudelitest. Kui tuua sisse valepositiivsete hinnangute lisakulud, siis 50%+1 hinnang ei ole enam piisav laenu heaks laenuks klassifitseerimisel, vaid see protsent peab olema suurem (kui palju täpselt, sõltub täpsest kulude maatriksist ja ilmselt ka kasutatud meetodikast kulude mudelisse arvamisel).

Domingose tööst lähtus, et lähenemine töötab klassifitseerimiskulude vähendamisel (Domingos 1999, 163). Sama tulemuseni jõudsid Kim *et al.* kes katsetasid lisaks MetaCost meetodile CSC nimelist lähenemist (ingl k *cost-sensitive-classifier*) ning otsustuspuud ilma klassifitseerimiskulude vähendamiseta ning järeldasid, et MetaCost vähendab klassifitseerimiskulusid kõige enam (Kim *et al.* 2012, 4019).

Kokkuvõttes iseloomustab autori meetodikat otsustuspuude rakendamisel järgnev tegevuste loetelu.

1. Laenude toorandmed laeti alla Bondora kodulehelt.
2. Valimi ettevalmistus. Eemaldati rahastamata laenud, veel jooksva maksegraafikuga laenud, ettevõtete võetud laenud, tühistatud laenud. Eemaldati tunnused, mida investor ei saa laenuotsuse tegemisel arvesse võtta.
3. Viidi läbi ristvalideerimise protseduur (eraldi kõikidele mudelitele). Algne valim jagati sellega seoses viieks alavalimiks.
4. Viidi läbi MetaCost protseduur (ristvalideerimise protseduuri sees), millega rakendati klassifitseerimiskulusid. Korduste arvuks valiti kuus.
5. Mudelid (neli tk) said valmis ning selgusid mudeleid iseloomustavad statistikud. Mudeleid kasutati edaspidises töös.

Saadud mudelite omavahelisel võrdlemisel vaatles autor mudelite täpsust ja seda, kuidas nad laene klassifitseerisid ehk millisel määral hindas mudel laenu tulemust õigesti. Tulemuse kõrvutamiseks Bondora krediidireitingutega kasutas autor reitingute alusel investeerimist paralleelselt otsustuspuude mudelitega sama andmebaasi peal, tulemusi on võrreldud alapeatükis 3.3.

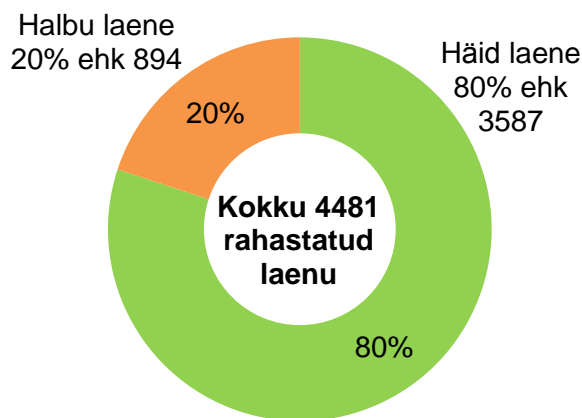
3. PRAKTILINE LAENUPORTFELLI KOOSTAMINE BONDORAS

3.1. Otsustuspuude meetodite rakendamise tulemused

Käesolevas töös kontrollis autor kahte otsustuspuudel põhinevat lähenemist kahel mõnevõrra erineval tunnuste hulgal, tekitades sedaviisi neli erinevat mudelit. Esimeseks mudeli tüübiks oli klassikaline otsustuspuu, mis koostati läbi klassifitseerimiskulusid arvestava algoritmi MetaCost. Tulenevalt MetaCost algoritmi tööpõhimõttest, genereeris autor kuus otsustuspuud, mis lõpptulemuse hindamisel andsid igaüks oma hääle, tekitades seeläbi ühe lõpliku hinnangu. Teiseks mudeli tüübiks oli *random forest* (RF) lähenemine, mille puhul koosnes iga nn mets kümnest erinevast otsustuspuust, mille erinevus seisnes asjaolus, et igal otsustushetkel oli puudel kasutada juhuslik valik laenaja tunnuseid ja valimi liikmeid. Kuna ka RF lähenemise puhul oli kasutuses MetaCost algoritm klassifitseerimiskulude arvestamiseks, tekkis RF meetodi kasutamisel kokkuvõttes $6 \times 10 = 60$ erinevat otsustuspuud (mis nõ hääletasid kümnestes gruppides). Mõlema mudeli peal katsetas autor kahte tüüpi andmebaasi: esimesel juhul oli kasutada vaid niipalju laenaja tunnuseid kui neid kajastus Bondora otsingufiltri tööriistas (turul olevate laenude otsimiseks ja filtreerimiseks), teisel juhul olid mudeli käsutuses kõik sobivad laenajate tunnused, mida Bondora laenuandmete andmebaasis leidis.

Enne mudelite väljundite lähemat selgitamist on autor pööranud tähelepanu andmetele, mis lõppkokkuvõttes mudelitesse jõudsid ehk vaadeldud nõ suurt pilti taustsüsteemist. Mõlema andmebaasi korral jäi algsest laenude tabelist lõplikuks kasutamiseks alles 4481 laenu. Valimi liikmete vähenemise põhjustas rahastamata või tühistatud laenuaotluste elimineerimine ning kuupäeva reegel, et kõik vaatlusalused laenuaotlused peavad olema maksegraafiku lõpuni jõudnud (et oleks selge laenu edukas tagasimaksmine või pankrotistumine).

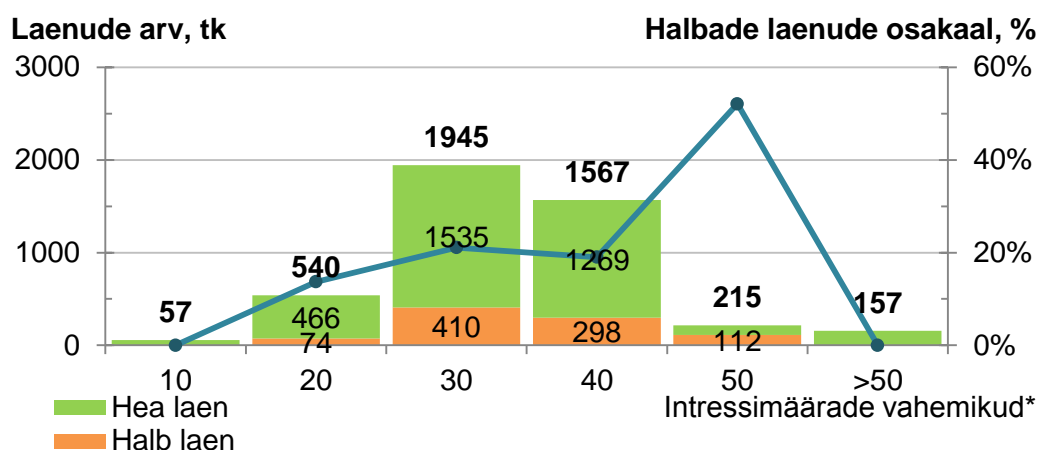
Kui vaadata kõikide laenude jaotust riski aspektist, siis järgnevalt jooniselt selgub, et pankrotistunud laenude hulk moodustab laenude koguhulgast suhteliselt väikese osa.



Joonis 4. Bondora laenude jaotus pankrotistunud (halbade) ja edukalt tagasi makstud (heade) laenude järgi. Allikas: autori koostatud Bondora laenude andmebaasi alusel

Seega võib öelda, et kogu valim on väga tugevalt kallutatud heade laenude suunas ja isegi täiesti juhusliku valiku puhul peaks iga 100 laenu kohta pankrotistuma 20 laenu. Ühtlasi tähendab see, et vaatamata suhteliselt suurele valimile (4481 liiget) saavad mudelid tegelikult õppida tuvastama pankrotistuvaid laene vaid 894 laenu andmete tuginedes.

Pöörates tähelepanu riskile ja tootlusele üheaegselt, on autor koostanud järgneva joonise laenude jaotusest intressimäärade alusel, tuues sealjuures esile pankrotistunud laenude hulga ja nende osakaalu kogu laenude seas teatud intressimäära vahemikus.



* nt väärtus 40 horisontaaltejel tähendab, et intressimäärade vahemikus (30;40] on 1567 laenu

Joonis 5. Bondora laenude jaotus intressimäärade alusel heade ja halbade laenude lõikes. Allikas: autori koostatud Bondora laenude andmebaasi alusel

Eelnevalt jooniselt lähtub, et reeglina on Bondora laenude intressimäärad jäänud vahemikku 30-40% aastas (3512 laenu 4481-st), intressivahemikes 20-30% ja 30-40% on laene vastavalt 1945 ja 1567 tk. Halbade laenude osakaal jääb nendes intressimäärade vahemikes ligi 20% tasemele. Halbade laenude osakaal rohkem kui kahekordistub kui intressimäärad jõuavad 40-50% vahemikku. Oluliselt vähem laene on välja antud kuni 20% intressitasemeni (597 laenu), samas on nende intressimäärade juures esinenud osakaaluliselt vähem pankrotistumisi.

Kasutatud valimit lähemalt uurides selgus, et 157 laenu, mis on välja antud intressimääraga üle 50% (vt joonis 5 lk 36), kuid mille halbade laenude osakaal on 0%, puhul on tegu Bondora poolt kõrge riskiga laenudeks kvalifitseeritud valdavalt Hispaania taotlejate laenudega. Laenud on Bondora laenude toorandmete andmebaasist lähtudes tasutud ennetähtaegselt ehk nende algne lepingujärgne viimase makse tasumise kuupäev on reeglina hilisem kui tegelik kuupäev, mis on määratud antud töös maksimaalselt 10. aprillile 2015. aastal (mil on allalaetud laenuandmete andmebaas).

Kokkuvõttes võib valimi kohta öelda, et ühelt poolt on Bondora senine poliitika laenuaotluste turulepääskamisel olnud pigem edukas, sest halbade laenude arv on kogu välja antud laenude hulgas suhteliselt väike (20% kõikidest rahastatud laenudest). Teiselt poolt on see prognoose tegevate mudelite koostamise vaatevinklist pigem negatiivne asjaolu, sest pankrotistuvate laenude kohta on proportsionaalselt vähem infot (nad on vähem valimis esindatud). Otsustuspuu vaatenurgast võib see tähendada, et kõik puu poolt tehtud otsused pigem hindavad laenud heaks, sest häid laene võib igasse otsustuspuu harusse proportsionaalselt rohkem jagada sõltumata valikutest. Seda ebavõrdset olukorda peaks autori hinnangul tasakaalustama mudelitesse toodud MetaCost klassifitseerimiskulude algoritm, mille tõttu suhtuvad mudelid tegelikult halva laenu kvaliteetseks hindamisele (vale-positiivne hinnang) viis korda tõsisemalt kui tegelikult hea laenu halvaks hindamisele (vale-negatiivne hinnang).

Katsetatud mudelite hindamiseks on olulisemad karakteristikud esitatud tabelis järgneval leheküljel.

Tabel 2. Autori genereeritud otsustuspuu (DT) ja *random forest* (RF) tüüpi mudelite karakteristikud*

Mudel ja andmed	Hinnangute õigesti klassifitseerimise määr (<i>accuracy</i>), %	Heaks loetud hinnangute õigesti klassifitseerimise määr (<i>precision</i>), %	AUC, %	Valepositiivsete** määr, %	Valenegatiivsete määr, %
DT, vähendatud tunnused	69,6	91,7	74,2	24,7	31,8
DT, kõik tunnused	68,2	92,2	73,8	22,4	34,2
RF, vähendatud tunnused	82,0	88,5	82,8	46,4	10,9
RF, kõik tunnused	82,5	87,2	81,3	53,7	8,5

* tabelis on rasvaselt rõhutatud antud kategooria parim tulemus

** positiivne tulemus tähendab head laenu

Allikas: autori koostatud

Ülaltoodud tabelist tuleneb esiteks, et toodud karakteristikute alusel võiks mudelite paremusjärjestuse koostada erinevalt, sõltuvalt sellest, millist parameetrit pidada olulisemaks. Kui vaadata kõikide hinnangute õigesti klassifitseerimist, siis eristuvad RF mudelid, mis paigutavad ligikaudu 12% või 14% rohkem hinnangutest õigesse kategooriasse, võrreldes lihtsamate DT mudelitega (tabelis näidik *accuracy*). Siinjuures ei eristu tulemused oluliselt sellest, millist tunnuste hulka kasutatud on. Kui jälgida vaid mudeli positiivseks klassifitseeritud näite (ehk heaks hinnatud laene, tabelis näidik *precision*), siis kõikide mudelite puhul on positiivseks loetud hinnangutest ka tegelikult õiged 87% või 89% hinnangutest – eristuvad DT mudelid, kelle puhul on määr RF mudelitest pisut suurem (mõlemal 92% tasemel).

Teistes valdkonna sarnastes töodes on mudelite juures välja toodud AUC karakteristik (Cubiles-De-La-Vega et al. 2013, 6914; Fawcett 2006, 868; Florez-Lopez ja Ramon-Jeronimo 2015, 5738; Malekipirbazari ja Aksakalli 2015, 4629). AUC tähendab ingl k *area under ROC curve* ehk binaarsete väärtustega klassifikatsiooni puhul tõenäosust, et suvaliselt valitud negatiivse väärtusega valimi liige omab väiksemat võimalust kuuluda positiivsesse klassi kui suvaliselt valitud positiivse väärtusega valimi liige (H. Zhang ja Su 2006, 892). Ehk teiste sõnadega tähendab AUC antud kontekstis tõenäosust, et mudel hindab suvalist head laenu hea klassi kuuluvaks rohkem kui suvalist halba laenu hea klassi kuuluvaks. Näiteks juhul kui AUC oleks 0,5 ehk 50%, tähendaks see, et mudel viskab iga hinnangu puhul kulli ja kirja –

pooled hinnangud annab õigesti, pooled valesti. AUC karakteristikku poolt osutusid paremaks RF mudelid.

Väga oluline on käesoleva töö autori hinnangul valepositiivsete määra karakteristik, mis antud juhul tähendab seda, kui suures osas hindab mudel tegelikult pankrotistuvaid laene headeks laenudeks. Siin eristuvad selgelt lihtsamad otsustuspuu (DT) mudelid RF mudelitest, hinnates halbadest laenudest (894 tk valimis) 22% või 25% headeks laenudeks. RF mudelid hindavad valesti 46% või 54%, sõltuvalt kättesaadavate tunnuste hulgast. Vale-negatiivseid hinnanguid, ehk hinnates hea laenu pankrotistuvaks, annavad RF mudelid jällegi DT mudelitest oluliselt vähem (11% või 9% võrreldes DT mudelite 32% või 34% määraga). Siinjuures tuleb tähelepanu juhtida, et juhul kui valim on üldiselt täidetud pigem positiivsete hinnangutega (heade laenudega), nagu ta Bondora puhul paistab olevat, võib endiselt paremaks osutada mudel, mis omab kõikide hinnangute puhul suuremat täpsust kui vaid valede hinnangute äratundmisel. Seega on RF mudelid vaid eeltoodud karakteristikuid vaadates endiselt atraktiivsed investeerimisotsuste tegijad.

Kui vaadelda otsustuspuid eraldi, saamaks aru, mille alusel tegelikult hinnanguid anti, saab põhjalikuma visuaalse ülevaate anda vaid lihtsamatest DT mudelitest, mida tulenevalt MetaCost algoritmi kasutamisest konstrueeriti 6 tükki (sest rohkemate korduste rakendamine ei parandanud mudeli täpsust oluliselt). Lihtsamate DT mudelite korral kasutas autor puude harvendamist, mis tähendab nende harude eemaldamist ja asendamist lehega, mis oluliselt ei paranda mudeli täpsust (nimetatud funktsiooni sai autori kasutatud tarkvaras soovi korral eraldi sisse lülitada).

Järgneval leheküljel paiknevas tabelis on autor esile toonud, milliseid tunnuseid mudelid otsustamisel oluliseks hindasid ehk mis mõjutas kõige enam mudeli hinnangut laenule. Tunnused on esitatud kõige olulisemast harust alustades.

Tabelist lähtub, et mudelite valikud eelistatud laenaja tunnuste osas erinevad kui anda mudelitele rohkem infot võrreldes Bondora otsingufiltris sisalduva informatsiooniga. Siit lähtub, et Bondora võiks kaaluda otsingufiltri täiustamist juhul kui rohkemat informatsiooni kasutav mudel soovitab investeerimisel paremaid otsuseid (rohkem infot kasutav mudel annab parema tulemuse).

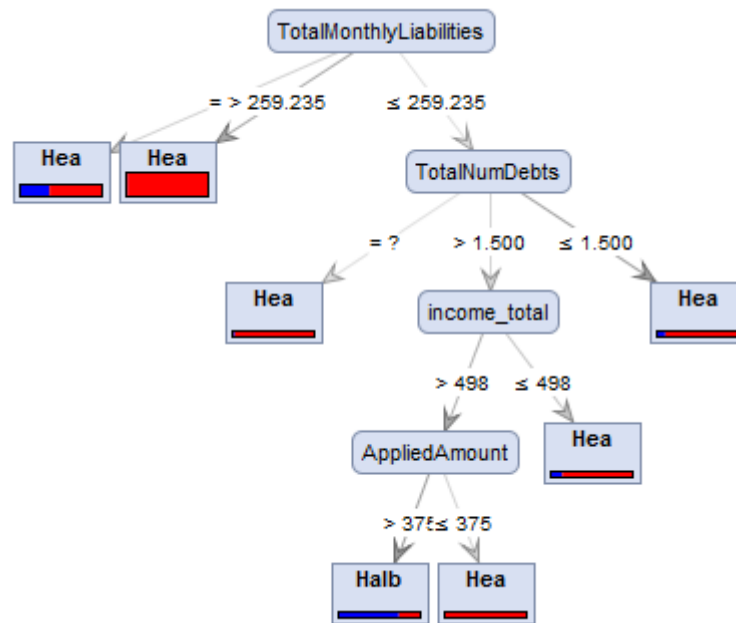
Tabel 3. Otsustuspuude (DT) mudelite poolt arvestatud laenaja tunnused hinnangute langetamisel

Otsustuspuu nr	1	2	3	4	5	6
DT mudel vähendatud tunnustega	TotalLiabilitiesBeforeLoan		DebtToIncome		TotalLiabilitiesBefore Loan	
					AppliedAmount	
					DebtToIncome	LoanDuration
DT mudel kõikide tunnustega	TotalMonthlyLiabilities		OtherLiabilitiesBeforeLoan		TotalMonthlyLiabilities	
	BondoraCreditHistory	Employment_Duration_Current_Employer			NumDebtOther	TotalNumDebts
	AppliedAmountToIncome/AmountofPreviousLoans	AppliedAmountToIncome			FreeCash	income_total
						Applied Amount

Allikas: autori koostatud

Üldiselt selgub, et mõlemad mudelid vaatavad laenaja finantskohustustega seonduvat informatsiooni nagu eelnevate finantskohustuste arvu (TotalLiabilitiesBeforeLoan), laenukoormust (DebtToIncome), aga ka laenu kestust ja taotletud summat. Rohkemaid tunnuseid kasutav mudel vaatleb viimast suhtena sissetulekusse (AppliedAmountToIncome). Võib öelda, et mudelite leitud tunnuseid saab pidada sisuliselt loogiliseks laenaja krediidi-võimekuse hindamisel.

Nagu eelnevalt esitatud tabelist näha võib, olid otsustuspuude mudelid maksimaalselt vaid kolme haru sügavused, mis tähendab, et nad on visuaalselt lihtsamini hoomatavad. Seda kinnitab järgnev joonis, mis on esitatud laiemat tunnustevalikut kasutava otsustuspuu nr 6 kohta.



Joonis 6. Otsustuspuu valikute langetamise protsess sõltuvalt laenaja tunnustest, ühe otsustuspuu näide. Allikas: autori koostatud tarkvara Rapidminer kaudu

Eelnevalt jooniselt on otsustusprotsessi kujunemist küll lihtne mõista, aga siiski on autori hinnangul otsustuspuude paljususe tõttu igapäevasel rakendamisel mõistlikum kasutada tehniliselt automatiseeritud lähenemist, mis ei vaja erinevate puude visuaalset hindamist.

Random forest tüüpi otsustuspuude korral autor kõiki tunnuseid töös eraldi välja ei too, kuna neid on otsustuspuude paljususe tõttu (10 otsustuspuud x 6 MetaCost kordust = 60 otsustuspuud kokku) ja puude mitteharendamise tõttu (tulenevalt *random forest* lähenemise eripärast ei ole see vajalik) ebamõistlik ühekaupa läbi vaadata. Hoolimata sellest on oluline märkida, et kasutaja saab Rapidminer tarkvara kaudu soovi korral kõigi mudelite kõigi otsuste harudega tutvuda. Läbi pistelise kontrolli on näha, et mudelid kasutavad tunnustena samuti valdavalt laenaja finantskohustusi kirjeldavaid tunnuseid (näiteks TotalLiabilitiesBeforeLoan).

Kokkuvõttes selgus eelnenud alapeatükist, et Bondora laenuandmete baas võimaldab käesolevas töös kasutada kokku ca 4500 laenu andmeid, millest 20% ehk ca 900 laenu on pankrotistunud. See tähendab, et andmed on kallutatud pigem heade laenude poole ning mudelitel võib selle võrra raskem olla pankrotistuvaid laene vastavalt klassifitseerida. Saadud mudelite võrdlemisel selgus, et *random forest* (RF) mudelid annavad küll kõikide hinnangute kohta täpsemad tulemused (hindavad ca 80% laenudest õigesti), samas jäävad RF mudelid

vaid halbu laene silmas pidades lihtsamatele DT mudelitele alla, sest viimaste valepositiivsete määr jääb 22-25% piirimaile kui RF mudelite oma vahemikku 46-54%. Individuaalsete mudelite visuaalse vaatluse pealt võib öelda, et mudelid tunduvad sisuliselt loogilised, st võtavad arvesse laenajat iseloomustavad finantsinfot.

Autori koostatud mudelite tulemusi investeerimisotsuste langetamisel on Bondora krediidireitingute süsteemiga võrreldud alapeatükis 3.3. Järgnevas alapeatükis on käsitletud intervjuu tulemusi.

3.2. Bondora esindajaga läbiviidud intervjuu tulemused

Nagu selgus käesoleva töö teoreetilisest osast mitmete lähenemiste olemasolu p2p platvormidel investeerimise kohta (vt lk 16), ilmnas Bondora esindajaga läbiviidud intervjuust, et p2p platvormides igapäevaselt kasutatud võimalused investorile otsuste langetamiseks on samuti praktikas mõnevõrra erinevad. Lühidalt võib p2p portaalide soovitude maailmapraktika kokku võtta järgmiselt (Pertman 2015).

- Valdavalt (paljudes suurtes p2p platvormides) on kasutusel krediidiskoorid, samas erineb nende arvutamise meetoodika.
- P2p laenuäri valdkonna tuntud esindajad nagu Zopa ja Ratesetter (UK) ei kasuta krediidiskoore (tähelepanuväärne erand), ega paku võimalust taotluseid valida. Enne investeerimist tutvustatakse investorile indikatiivset tootlust, mida võib teenida eelnevalt valitud ajaperioodil.
- P2p platvormide praktikas esineb (vähesel määral) üksikuid eristuvaid näiteid nagu sotsiaalmeedias esineva info kasutamine p2p portaali poolt krediidivõimekuse hindamiseks.

Valdkonna tuntumad ettevõtted nagu USA p2p platvormid Lending Club ja Prosper kasutavad krediidiskoore, mis on väga sarnased Bondora meetoodikale. UK suured p2p platvormid eelistavad passiivse lahenduse loomist investori jaoks. Sotsiaalmeedia kasutamist krediidivõimekuse hindamiseks on kasutatud, aga vähe, ja selle edukuse kohta on vara hinnangut anda. (Pertman 2015) Intervjuu vastustest on näha, et p2p platvormide vahel ei ole tekkinud üksmeelt parima lahenduse osas, mille järgi investor võiks portaalis käituda, et investeringuid valida. Samas on käesoleva töö autori hinnangul võimalik, et näiteks USA ja

UK portaaliid ja nende klientuur on piisavalt erinevate hoiakutega, et tegemist ongi kõige eelistatute valikutega mõlema maailmajao klientidele.

Krediidireitingu kohta selgus intervjuust, et krediidireitingute lähenemise (nagu Bondora kasutab) teeb eriliselt väärtuslikuks asjaolu, et investoritele antakse juurde riski element, mida otsustel arvesse võtta. Bondora esindaja tõi võrdleva näite nende varasemalt propageeritud laenajat iseloomustavast tunnusest nagu krediidiskoor (mis on portaalis jätkuvalt avalikustatud). Krediidiskoor, mis tugineb infole laenaja varasemast kehvast maksekäitumisest või selle puudumisest võib ühesuguse hinnangu anda laenajale, kes on varasemalt oma kohustused tasunud kui ka laenajale, kes ei ole lihtsalt varem laenukohustusi omanud. Lisaks toodi välja, et krediidireitingute süsteem on maailmas laialt levinud traditsioonilises pangandussektoris (näiteks võlakirjade puhul). (Pertman 2015) Järeldub, et krediidireitingute alusel investeerimine on vähemalt varasemalt esitatud variantide (nagu ainult indikatiivse intressi pakkumine) hulgast ainuke, mille alusel saab lisaks soovitud tulumäärale ka eelistatud riskitaset silmas pidada. Siiski tuli intervjuust välja, et Bondora krediidireitingute süsteem on veel arengujärgus ja mudeleid pidevalt täiendatakse (Pertman 2015).

Osaliselt seoses eelnevaga, et Bondora krediidireitingute süsteem on pidevas arendusfaasis, selgus intervjuust asjaolu, et alternatiivsed lähenemised investeerimisotsuste tegemiseks (näiteks kasutades otsustuspuu meetodit) võivad tõenäoliselt anda parema tulemuse. Siinjuures on oluline silmas pidada järgmiseid aspekte (Pertman 2015).

- Tõenäoliselt ei esine üksikuid kindlaid tunnuseid, mida silmas pidada näiteks sama krediidireitinguga laenajate vahel valimisel – pigem on parema valiku tegemine kompleksne lahendus ja hõlmab mitmete tunnuste koos vaatlemist.
- Juhul kui investoril õnnestub selgitada välja tunnused, mis laenaja tagasimakse tõenäosust mõjutavad, võib selle teadmise rakendamine osutuda keerukaks, sest sobivaid laene peab otsima (tekib ajakulu) ning neid ei pruugigi teatud ajahetkel piisavalt leiduda (piisavalt, et kõik vabad vahendid sobivatesse laenudesse paigutada).
- Juhul kui investori investeeritav rahasumma ületab tavapäraselt sellises keskkonnas paigutatava mahu, on sobivate laenude leidmine eriti raskendatud ning see võib viia tagasi olukorda, kus investor kas lõdvendab investeerimis-

kriteeriume või otsustab ühte laenu paigutada tema portfelli mahtu arvestades ebaproportsionaalselt suure summa. Viimane asjaolu tähendab omakorda ebapiisavat diversifitseerimist ehk liigsete riskide võtmist.

Siiski tuleb eelnevast välja, et juhul kui investori paigutatav rahasumma on suhteliselt väike, võrreldes näiteks institutsionaalsete investoritega (väikeinvestori puhul võib sageli nii olla), võib alternatiivsete meetmete, näiteks otsustuspuude, rakendamine osutada edukaks siis kui leitakse piisavalt suurt lisatulu pakkuv investeerimispõhimõte, mida on igapäevaselt võimalik järjepidevalt rakendada. Samas tõi intervjuueritav Bondora esindaja välja, et näiteks USA p2p portaalis Lending Club on arendatud krediidireitingu mudelit kaua ning on märgata, et seal alternatiivselt investorite poolt kasutatust leidnud laenamirobotitel (mis kasutavad enda väljatöötatud mudelit) on krediidireitingute ees küll edumaa, aga saavutatud lisatulu on olnud ajas vähenev ehk Lending Clubi reitingusüsteem on ajas paranenud (Pertman 2015).

Otsustuspuu igapäevast rakendamist investeerimisel (või mõne teise investeerimisreeglitiku rakendamist) võib pärssida protsessi võimalik keerukas ülesehitus. Siiski tõi Bondora esindaja välja, et sarnaselt teistele suurtele p2p platvormidele nagu Lending Club ja Prosper on Bondoral plaanis tulevikus välja arendada API lahendus (Pertman 2015). API ehk ingl k *application programming interface* on standardiseeritud reeglite kogum, mille põhjal arvutiprogrammid omavahel suhtlevad (Encyclopedia Britannica 2013). Seega oleks võimalik investoril esiteks sobivate investeeringute otsinguparameetrid ja teiseks sobivate tulemuste pihta suunatud investeerimissoovid programmiks (või pigem käskluste jadaks) muuta – see lubaks väga kiiresti toimida (ka väikeinvestoril) ning sobivate keerukamate investeerimiskriteeriumite portaalis rakendamine muutuks vaid ühekordseks tehniliseks, mitte igapäevast ajaressurssi nõudvaks probleemiks. Siiski nentis Bondora esindaja, et nendepoolne API lahenduse arendamine on küll plaanis, aga täpsem visioon lõpptulemusest on veel väljatöötamata ning valmimise tähtaeg ei ole teada (Pertman 2015).

Käesoleva töö turuanalüüsi osas (vt lk 10) selgus, et hiljem turule tulnud väiksemad p2p ettevõtted on pakkunud investoritele pigem kõrgemaid intresse kui need, kes on turul tegutsenud kaua ja saavutanud teistest suuremad laenumahud – seega võib eksisteerida teatud tendents, kus p2p platvormi kasvades ja kasutajabaasi suurenedes investorite teenitud tootlused pigem vähenevad (koos riskitasemega). Bondora sõnul ei ole nende eesmärgiks eelistada portaalis laenutaotluse teinud isikute vahel kõrgema tootlusega või väiksem riskiga

laenajaid – pigem on eesmärk olla vahendaja ja tuua turule laenuid, mis Bondora poolse kontrolli läbivad (ehk on Bondora hinnangul investoritele tulusad). Sealjuures tõi Bondora esindaja välja, et kui näiteks USAga võrrelda, siis seal on kõik laenuintressid madalamad ehk raha on põhimõtteliselt odavam. Eestis on aga intressitasemed kõrgemad, kõik tarbimislaenuid on klientidele kallimad. Lisati, et mõnes riigis piiravad intresse riiklikult seadistatud maksimaalsed määrad, näiteks Saksamaal. (Pertman 2015) Seega tuleb arvestada, et p2p laenuurul võib keskmiste intressitasemete kujunemisel suurt rolli mängida ka vastava riigi eripära ja seda mitte laenaja, vaid riigi riskitaset ja sealseid intressimäärasid üldiselt silmas pidades.

Selleks, et p2p platvorm saaks investorile pakkuda veelgi paremat tootlust ja laenajale õiglasemat intressimäära, tõi Bondora välja mõned probleemid, mis käesoleva töö kirjutamise hetkel aktuaalsed on ning mille lahendamise vajab aega, aga ka suhtumise muutumist regulatiivsel tasemel. Ühe aspektina toodi välja, et laenaja kohta informatsiooni saamine võib p2p platvormina tegutsevale vahendajale olla problemaatiline. Näitena toodi, et näiteks Eestis ei ole veel tekkinud andmebaasi, kus kajastuks kodanike kohta nende positiivne finantsajalugu ehk näiteks see, et inimene on 20 aastat õigeaegselt enda laenu tagasi maksnud. See annaks laenu andjatele lisainformatsiooni laenaja krediidiriski kohta ning taoline mudel on edukalt kasutusel USA-s FICO skooride näol. (Pertman 2015) Täpselt samasuguse arengusoovi kohta annab aimu 2015. aastal moodustatud Eesti Vabariigi valitsuse koalitsioonileppe punkt 11.12, kus valitsus lubab kaaluda koostöös erasektoriga positiivse krediidiregistri loomist (Rõivas *et al.* 2015).

Teiseks p2p laenuportaalide arengut pärssivaks asjaoluks võib pidada ebasoosivat suhtumist p2p laenuärises üldiselt, seda ka Eestis. Toodi välja, et näiteks UK-s soovib riik p2p laenuäri edendada ja selles suunas regulatsioone luua – üheks positiivseks näiteks toodi asjaolu, et sealsed p2p platvormide investorid saaksid intressituludelt maksusoodustusi (Bondora esindaja ei olnud kindel, kas esitatud näide on veel planeerimise tasemel või juba jõustunud). Vastupidiselt on Eestis hakatud regulatiivsel tasandil kõiki laenajaid mõõtma nõu puuga, mitte eristades tarbijatele kalleid SMS-laenu ärisid, pankasid ega p2p laenuplatvorme. (Pertman 2015) Viimasest tähelepanekust koorub välja asjaolu, et tegelikult võib väikeinvestori tootluse parandamiseks leiduda ruumi regulatiivsel tasandil, mida kahjuks üksikindiviidid ei saa kontrollida. Siiski on selge, et näiteks maksusoodustuse korral suureneks tõenäoliselt väikeinvestorite netotootlus investeringutelt ainuüksi seepärast, et maksumäär

alaneks. Sama sündmus võib kaasa tuua soodsamad tingimused laenuvõtjale, sest pakkumise pool muutuks aktiivsemaks.

Selleks, et aidata kaasa p2p laenuäri uurivale teadustööle, küsis autor alakaetud valdkondade kohta p2p platvormide uurimisel. Bondora esindaja nentis, et p2p valdkond on akadeemilisel maastikul veel vähe uuritud ning pigem keskendub olemasolev materjal investeerimisstrateegia lahtimõtestamisele ning turu suuruse ja kasvu analüüsile. Samas tuleks tegeleda ka p2p valdkonna regulatsiooni olemasolu ja mõju uurimisega. (Pertman 2015)

Kokkuvõttes selgus Bondora esindajaga läbiviidud intervjuust, et igapäevases p2p laenuäri praktikas esineb erinevaid lähenemisi, mille alusel platvormid ise investoritele investeerimisotsuste langetamiseks informatsiooni pakuvad või investeerida soovivad. Krediidireitingu lähenemist kasutavad paljud suuremad p2p platvormid ning selle eeliseks peetakse juurdeantavat infot laenu riski kohta. On tõenäoline, et investor suudab iseseisvalt laenustatistikat uurides ja mudeleid välja töötades leida meetodi, mis annab talle kõrgema tootluse kui krediidireitingu abil saavutatu, aga sealjuures võivad tekkida takistused meetodi rakendamisel seoses suurenenud ajakuluga investeringute otsimisel, saadaolevate laenude ebapiisava mahuga võrreldes investeeritava rahasummaga, isiklikust mudelist saadud teadmise rakendamise keerukusega ning viimaks ja kokkuvõtteks sellega, kas saavutatud tootluste vahe on piisav, et lisavaeva näha. Samas on võimalik, et keerukamad tehnilised lahendused ei ole takistuseks kui neid on võimalik automatiseerida näiteks API lahendust kasutades (mida hetkel Bondoras teha ei saa). Viimaks selgus, et p2p laenuäri juures on arenguruumi regulatiivses vallas, mis täna pigem takistab investoril maksimaalset tootluse saavutamist.

3.3. Otsustuspuude rakendamise võrdlus Bondora krediidireitingu meetodiga ning tulemuste igapäevane kasutamine

Mudelite tulemuste võrdlemiseks oli autori käsutuses Bondora laenuandmete andmebaas, millele töö käigus koostatud mudelid lisasid iga laenu tegeliku tulemuse (hea või halb laen) kõrvale oma hinnangu laenu edukuse kohta.

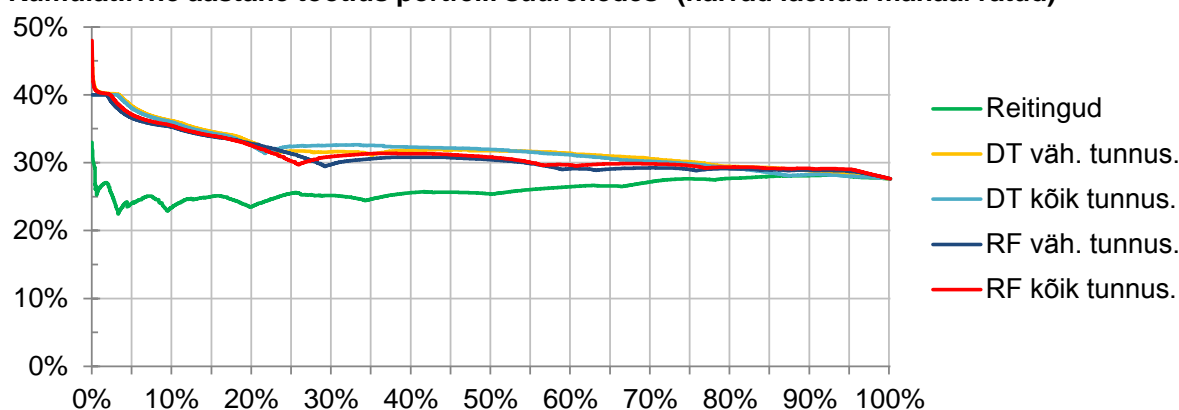
Tulemuste võrdlemisel võttis autor aluseks sarnase metoodika nagu tehti seda Malekipirbazari ja Aksakalli töös Lending Clubi p2p platvormi laenude analüüsimisel otsustuspuu meetodiga. Nimelt võtsid autorid eelduseks investori (näitlikult) soovitava minimaalse riskitaseme ja sellest tuleneva saadolevate laenude mahu (% kogu saadaolevatest laenudest). Järgnevalt vaadati samale laenumahule investeerimissoovitust andvate otsustuspuude hinnangut (kasutati RF tüüpi mudelit). Viimaks võrreldi soovitusel saanud laenude pankrotistumise määra ning jõuti järeldusele, kumb meetod tõi samas mahus investeringute tegemisel kaasa vähem pankrotistunud laene. (Malekipirbazari, Aksakalli 2015, 4629). Tulenevalt sellest, et käesoleva töö eesmärgiks on lisaks riskile hõlmata teenitava tulu poolt, on autor lähenemiste võrdlemisel sisse toonud intressimäärade aspekti. Töö alguses selgitati, et Markowitzi teooria järgi peaks investor korraga pidama silmas nii riski ja tulu, sealjuures püüdes hoida soovitud tulutaseme juures riski minimaalsena või soovitud riskitaseme juures tulu maksimaalsena (vt lk 13). Krediidireitingute ja autori mudelite võrdlemise ning üheaegselt riski ja tulu komponentide fookuses hoidmise tarvis tegi autor järgnevad eeldused.

- Konstrueeriti mitu strateegiat, mille juhul on kõikidel ühesugune stardipositsioon (turg samade laenudega), strateegiad erinesid vaid investeringute valimise loogikas.
- Kehtis eeldus, et riskikartlik investor soovib hoida riski minimaalsena, sealjuures maksimeerides saadavat tulu (antud juhul laenude intressimäära).
- Ühe strateegia järgi investeeris investor reitingute alusel Bondora krediidireitingute süsteemi kõige madalama riskiga laenudesse (reiting AA), liikudes paigutatava rahasumma suurenedes järk-järgult kõrgema riskiga laenudesse (Bondoras on 7 astet kuni tasemeni HR).
- Paralleelselt investeeris investor otsustuspuude soovitusel järgi (4 erinevat mudelit), järjestades investeringud otsustuspuude hääletustulemuse kindluse alusel (*confidence*).
- Juhul kui korraga oli saadaval mitu sama reitinguga või otsustuspuu kindlustasemega investeringut, investeeriti nendest kõrgema intressimääraga laenu (tulu maksimeerimine sama riski juures).
- Soovides hõlmata käesolevas töös kasutatud andmeid, võeti aluseks andmebaasis krediidireitingut omavad laenud (1985 laenu kogusummas 2 571 750 eurot). Sellest mahust eemaldati kõrge intressiga (>50%) laenud, mis olid

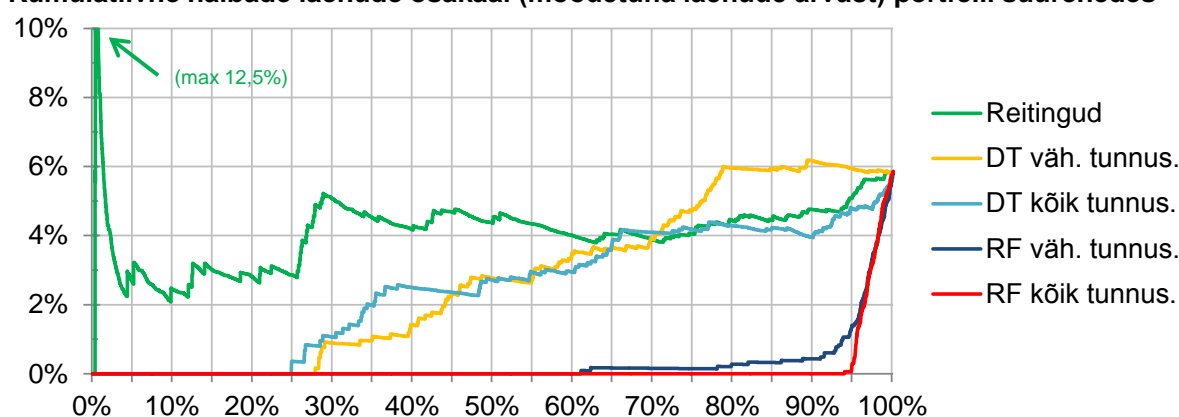
ennetähtaegselt tagasi makstud (vt joonis 5 lk 36), jättes valimisse lõpuks 1827 krediidireitingutega laenu mahus 2 445 750 eurot). Eeldati, et investor investeerib kogu summa ulatuses hõlmates lõpuks kogu portfelli. Igale strateegiale on üks samasugune turg samasuguste investeeringutega. Teiste sõnadega tähendab see, et kõik 1827 laenu on üheaegselt turul ning paigutatava rahasumma suurenedes saab kõiki takistusteta osta.

Ülaltoodud võrdlusprotseduuri eesmärgiks oli võtta maksimaalselt arvesse krediidireitingute poolt pakutavat infot ehk hõlmata neid võimalikult palju, loomaks võimalikult head võrdlusbaasi. Eeltoodud strateegiate tulemusel on autor konstrueerinud graafiku, millel on näha nii halbade laenude osakaal portfelli suurenedes kui investori kumulatiivne aastatootlus, kusjuures on tootluse puhul arvestatud halbade laenude mõjuga (investeeringud hakkavad portfelli suurenedes hõlmama ka halbade laenude koguseid).

Kumulatiivne aastane tootlus portfelli suurenedes* (halbad laenud mahaarvatud)



Kumulatiivne halbade laenude osakaal (möödetuna laenude arvust) portfelli suurenedes*



* 100% tase horisontaalteljel tähendab, et investor omandab kõik laenud turul

Joonis 7. Bondora krediidireitingute meetodi, otsustuspuude (DT) ja *random forest* (RF) mudelite võrdlus investeerimisotsuste langetamisel. Allikas: autori koostatud

Eelnevalt esitatud jooniselt selgub, et krediidireitingute alusel investeerimine osutus kõige kehvemaks strateegiaks, valides koheselt portfelli mõned pankrotistuvad laenu väiksema riskiga krediidigrupist AA (3 laenu summas 1850 eurot oli pankrotistunud). See vähendas kumulatiivset intressitulu nõ investeerimise alguses. Kuna krediidireitingute järgi käituv investor valis kõigepealt väikseima riskitaseme ja siis sellele riskitasemele vastavast laenude hulgast investeringud alates kõrgemast intressist, hakkas reitingutepõhisel investoril intressikõver kumulatiivselt kasvama tulenevalt sellest, et üha kõrgema riskiga laenude juurde liikudes hakkasid nende laenude keskmised intressid suurenema. Ühtlasi on näha, et halbade laenude osakaal püsis portfellis edaspidi ligikaudu 4% juures.

Otsustuspuud (DT või RF meetoditel) andsid krediidireitingutega võrreldes erineva tulemuse. Tulenevalt sellest, et investor valis investeringuid riski minimeerimiseks esiteks otsustuspuude kindlustaseme järgi (mitte vaid hea/halb määratluse järgi), alustades kõrgemast, hakkasid otsustuspuude (DT) mudelid halbu laene ekslikult portfelli võtma alles 25% ja 28% paigutamise järgselt (vastavalt DT kõikide ja vähendatud tunnustega mudelid). Tulenevalt sellest, et sama kindlustaseme juures eelistas investor võimalusel kõrgema intressiga laene (et maksimeerida tulu sama riskitaseme juures), sattusid esimeses järjekorras eelistatud laenude hulka kõrge intressiga laenud (näiteks 49% intressimääraga laenud krediidireitinguga HR). Seepärast olid kumulatiivsed aastatootlused DT või RF mudelitel tähelepanuväärselt kõrgemad siis kui portfelli osakaal kogu turust oli veel väike (näiteks olles investeerinud 10% laenedesse kogu turust, olid nii DT kui RF mudelid saavutanud *ca* 36% suurused aastased tootlused, olenemata tunnuste arvust mudelites).

RF mudelite tulemus ületas halbade laenude vältimisel DT mudeleid, sõltumata kättesaadava informatsioonihulga (laenuvõtjat iseloomustavate tunnuste hulga) maksimaalsest või vähendatud versioonist. Eelnevalt jooniselt (vt joonis 7 lk 48) selgus, et vähendatud informatsioonihulgaga RF mudel korjas esimesed halvad laenud portfelli siis, kui oli investeeritud 61% kogu turu mahust. Seevastu kogu saadaolevat informatsiooni kasutav RF mudel eksis esimest korda alles siis, kui ta oli eelnevalt andnud soovitusel investeerimaks 94% portfelli.

Autor juhib tähelepanu, et kõik mudelid investeerisid lõpuks kõikidesse laenedesse. See tingimus lubati, et näidata, kuidas kokkuvõttes jõutakse ühe ja sama intressimäära ja halbade laenude osakaaluni olenemata eelnevalt tehtud valikutest (suurendamiseks kindlust, et

kõikide strateegiate puhul on kasutatud ühesugust võrdlusbaasi). Siinjuures on oluline mainida, et otsustuspuude mudelite alusel investeeriv investor korjas seega kõige viimases järjekorras portfelli laenud, mida mudel juba ette pankrotistuvaks prognoosis. Näiteks RF mudelite puhul on seda hästi näha eelneva joonise (vt joonis 7 lk 48) parempoolses servas, kui pankrotistunud laenude osakaal tõusis järsult turu keskmisele tasemele.

Võrreldes varasemate töödega sai autor võrreldava tulemuse Malekipiribazari ja Aksakalli kasutatud RF lähenemisega Lending Clubi andmete peal, mida käsitleti käesoleva töö teoreetilises osas (vt lk 18). Mainitud töös leidsid autorid, et RF meetod annab võrreldes p2p portaali enda krediidiskooridega investeerides oluliselt väiksema pankrotistunud laenude osakaalu (Malekipiribazari, Aksakalli 2015, 4630). Sarnaselt mainitud autorite tööga tuli käesolevas töös otsustuspuude kasutamisel välja, et kuni teatud piirini on võimalik täielikult vältida pankrotistuvate laenude valimist portfelli. Teoreetilises osas mainitud Yap *et al.* töö tulemused, milles selgus, et otsustuspuu ei paranda mh krediidiskooridega võrreldes investeerimistulemust (vt lk 18) on käesoleva töö resultaadiga vastuolulised (autor sai otsustuspuude kaudu oluliselt väiksema pankrottide määra ehk mudel suutis klassifitseerida häid ja halbu laenusid krediidireitingutest paremini). Samas, nagu käesoleva töö teoreetilises osas mainitud (*Ibid.*), ei olnud viidatud töö autorid kasutanud finantsandmeid, vaid tulemusi üldistati terviseklubi andmebaasi alusel – sestap võib olla, et tegelikult ei osutunud mainitud autorite kasutatud andmed piisavalt hästi finantssektoris krediidikäitumist peegeldavaks.

Eelnevalt esitatud investeerimisstrateegiate võrdluse tulemusel teeb autor järelduse, et otsustuspuudel põhinevad mudelid suudavad Bondoras investeerimisel investori tootlust parandada, võrreldes krediidireitingute alusel investeerimisega. Otsustuspuude mudelid parandasid tootlust keskmiselt 12% võrra siis kui investor oli valinud kõikidest turuolevatest laenudest eelistatuid 10%, järjestades neid esiteks mudelite kindlustaseme ja teiseks laenude intressimäära alusel. 10% tase on autori hinnangul mõistlik üldistus näiteks olukorras, kus regulaarselt investeerides (näiteks iga kuu või kvartali lõpus) valib investor kõikidest parajasti turul olevatest laenudest välja kümnendiku, mis tema hinnangul on parimad (eeltoodud tingimustel). Autori eelnevalt läbiviidud uurimise käigus selgus, et korraga võib Bondora keskkonnas üleval olla ca 400 laenuaotlust (vt lk 22). Tõenäoliselt kujuneb tegelik ühe korraga investeerimiseks valitav laenude maht 10% tasemest väiksemaks arvestades Bondora laenuturu suurust ja p2p valdkonna üleüldist kasvutempot (vt lk 11).

Autor juhib tulemuste tõlgendamisel tähelepanu, et otsustuspuude mudelitel olid kasutada laenude andmed, mis ei hõlmanud ajaliselt sellist perioodi, milles sisalduks näiteks majanduslanguse faas (tulenevalt sellest, et esimesed Bondora laenud väljastati 2009. aasta alguses, vt lk 24). Seevastu Bondora krediidireitingud võtavad väidetavalt arvesse riigi riski aspekti (vt lk 22). Seega ei saa tulemuste alusel kindlalt väita, et otsustuspuude mudelite alusel investeerides oleksid krediidireitingutest paremad tulemused pikas perspektiivis investeerides (nii, et investeerimisperioodi jääb mõni uus majanduskriis või muu turušokk).

Kui vaadata eelneval joonisel toodud (vt joonis 7 lk 48) intressikõverate asetust, siis koorub töö tulemustest üks klassikalise finantsteooriaga mõnevõrra vastuoluline tähelepanek. Nimelt kui arvestada, et eelneval joonisel on autori koostatud mudelid järjestanud investeringud esiteks mudeli kindluse alusel, siis jaotub investeringute riskitase investori jaoks vasakult paremale, alustades kõige madalamast. Samuti järjestuvad kõrgema intressiga laenud vasakusse serva, sest kõikide otsustuspuude mudelite alusel koostatud investeerimisportfellide kumulatiivsed intressikõverad on langevad. Siit lähtub, et otsustuspuude abil investeringuid analüüsides ei pruugi kehtida üldiselt aktsepteeritud teenitava tulu ja võetava riski vaheline pöördvõrdeline seos. Antud töö tulemuste pealt võib järeldada, et minimaalse riskitaseme juures on tegelikult leitavad investeringud, mis annavad ka kõrgeima tootluse. Muidugi võib see nii olla ka seepärast, et turg ise hindab vastavad laenud nõ ebaefektiivselt, nähes suuremat riski seal kus seda tegelikult võibolla ei ole.

Negatiivse asjaoluna võib välja tuua, et hoolimata sellest, et otsustuspuud (DT või RF) on suutnud halbu ja häid laene klassifitseerides tuua empiirilisel parema tulemuse, on nende igapäevane rakendamine tehniliselt keerulisem kui krediidireitingute järgi investeerides. Nagu autor Bondora esindajaga läbiviidud intervjuu abil välja selgitas, ei esine tõenäoliselt laenudes kindlaid üksikuid tunnuseid, mida näiteks otsingufiltri abil välja filtreerides saab kohe eristada head ja halvad laenud – pigem tuleb vaadata mitmeid tunnuseid korraga (vt lk 43). Otsustuspuude lähenemise empiiriline katsetus kinnitas seda arvamust. Ühe otsustuspuu visuaalne jälgimine ja hindamine võib investorile olla lihtsasti hoomatav, aga soovides arvestada klassifitseerimiskuludega tekib ka kõige lihtsama otsustuspuu meetodi rakendamisel korraga mitmeid puid, rääkimata RF lähenemisest, mille korral tuleb korraga silmas pidada kümnete otsustuspuude hääletustulemust. Sestap tuleb tõdeda, et igapäevane otsustuspuude rakendamine investeerimisel võib osutada edukaks vaid juhul kui protsessi on võimalik

automatiseerida, näiteks saades ligipääsu kõikidele turuolevatele investeeringutele Exceli faili kujul (nii nagu laenu on kajastatud Bondora toorandmete failis), ning fail siis mudelitele hindamiseks esitada. Tulevikus võib automatiseerimise võimaluse puudujääki kompenseerida API lahendus, mis Bondora esindaja sõnul on plaanis välja arendada.

Kuna käesoleva töö kirjutamise hetkel kirjeldatud API lahendus puudub, on üks võimalus otsustuspuude rakendamiseks investeerimisotsuse tegemise ajahetkel kättesaadavate investeeringute info kirjutamine käsitsi Bondora portaalist Exceli tabelisse, et väljatöötatud mudel seda analüüsida saaks. Siinkohal lihtsustaks oluliselt vaeva mudeli kasutamine, mis võtab arvesse võimalikult vähe laenaja tunnuseid (autori mudelites võtsid vähendatud tunnuste arvuga otsustuspuude mudelid arvesse 21 tunnust). Samas tuleb tunnistada, et taoline käsitööd nõudev lahendus on endiselt ajamahukas ning vajadus suurema automatiseerimise järele ei kao.

Autori hinnangul oleks vajalik töö tulemusi edasi uurida läbi lõpliku mudeli testimise meetodika täiustamise. Oluline oleks vaadelda, mil määral käituvad erinevad mudelid siis kui olukord on igapäevase investeerimistegevusega veelgi kattuvam. Nagu Bondora esindaja välja tõi, on isiklike mudelite abil portaali krediidireitingutest parema tulemuse saavutamisel oluline, et investor suudab enda võimalikku eelist järjepidevalt rakendada (vt lk 44), ehk antud juhul peaks regulaarselt leiduma investeeringuid, millesse mudelid soovivad suure kindlusega investeerida. Juhul kui sobivaid investeeringuid ei leidu piisavalt, tuleb mudeli kriteeriume lõdvendada või olemasolevatesse laenudesse veelgi suuremaid summasid paigutada (diversifitseerides vähem).

Eelnevalt selgus, et mudelite üks rakendusvõimalus olukorras, kus Bondora ei ole veel API lahendust välja arendanud, eeldaks käsitsi kättesaadavate andmete Bondora andmebaasist Excelisse organiseerimist, mis aga võib osutada väga ajamahukaks. Investori poolseks olukorra leevenduseks tuleks kasuks töö edasiarendamine suunas, mis uurib laenaja tunnuste arvu veel enamat vähendamist nii, et saadud mudel säilitaks piisava hinnangute täpsuse. Sellisel juhul oleks mudeli igapäevane rakendamine lihtsam. Autori hinnangul ei saa kindlalt väita, et mudelitesse sisestatud kõikidest laenaja andmetest võttis mudel igat tunnust arvesse – samas nõuab ebavajalike tunnuste eemaldamine tööd, et selgitada välja mudeli jaoks kõige ebaolulisemad tunnused või nende kombinatsioonid.

Nagu Bondora esindaja intervjuus kinnitas, et Bondora krediireitingute süsteemi mudelid vaatavad iga riigi klienti erinevate parameetrite järgi, oleks võimalik töö edasiuurimise valdkond mudelite täiustamine, püüdes arvestada riigi erinevuse aspekti (käesolevas töös autor seda ei teinud).

Kokkuvõttes selgus eelnevast alapeatükist, et krediireitingute strateegia alusel investeerimine, alustades kõige madalamast oodatud kahjumäärast (ehk riskist), annab investorile kõige väiksema tootluse (võrreldes otsustuspuudega) ning pidevalt umbes 4% juurde jääva hapuks läinud laenude osakaalu portfelist. Otsustuspuude mudelid annavad investorile krediireitingutest parema tootluse, sest eelistavad esimestena kõrgema intressimääraga laene. Esimesed halvad laenud korjavad lihtsamad DT mudelid portfelli, olles eelnevalt edukalt selekteerinud 25%-28% kogu turul olevatest laenudest. Keerukamatest RF mudelitest eksis laenaja kohta vähem infot omanud mudel esimest korda, olles eelnevalt vigadeta selekteerinud 61% turu mahust, laenuvõtja tunnuste kohta rohkemat infot omanud mudel andis esimese valepositiivse hinnangu alles 94% tasemel. Otsustuspuude lähenemist saab järelikult pidada krediireitingutest paremaks, aga igapäevaselt on mudeleid keeruline kasutada seoses tehniliste piirangutega laenuandmete ligipääsemisel. Probleemi aitaks ühelt poolt lahendada vajaliku tehnilise arenduse tegemine Bondoras, teiselt poolt võib abiks olla mudelite lihtsustamine nii, et nad vajaksid vähem informatsiooni, andes endiselt krediireitingutest parema tulemuse. Lisaks on oluline teadmiseks võtta, et otsustuspuud ei saanud laenude puhul arvestada võimalikku majanduskriisi olukorda, sest valimi andmetes seda ei sisaldunud – selle ettevaatusega tuleb suhtuda tulemuste tõlgendamisse pikas perspektiivis.

KOKKUVÕTE

Käesolevas töös uuriti investeerimist Bondora p2p laenukeskkonnas eraisikust investori vaatenurgast. Töö eesmärgiks oli selgitada välja optimaalse laenuportfelli koostamise põhimõtted p2p laenamise keskkonnas, rakendades otsustuspuu kontseptsiooni Bondora laenuplatvormi empiirilistele andmetele. Pidades silmas p2p laenukeskkondade turgu, püstitati hüpotees, et Bondora puhul on tegemist konkurentidega võrreldes investoritele kõrgemat tootlust pakkuva keskkonnaga. Teine hüpotees väitis, et Bondoras investeerimisel on otsustuspuude abil võimalik tootlust parandada, võrreldes Bondora krediidiskooride lähenemisega.

Töö eesmärk täideti. Töö tulemus oli, et otsustuspuude abil õnnestus investeerimistootlust parandada, võrreldes krediidireitingute alusel investeerimisega. Autori hüpotees Bondora kõrgemast intressitootlusest võrreldes kogu p2p turuga leidis kinnitust, Bondora poolt pakutud *ca* 20% intress on konkureerivatest platvormidest kõrgem (keskmine intressitulu vaadeldud ettevõtete hulgas oli 9%). Siiski sõltub Bondora intressimäär valitud laenudest ning intressimäärade võrdlus kaasatud ettevõtetest võrdlusbaasis.

Töö teoreetiline osa alustas p2p laenukeskkondade mõiste selgitamisest. P2p laenukeskkonna võib kokku võtta kui era- või juriidilistest isikutest laenuandjaid ja –võtjaid ühendava turu. Kehtib informatsiooni asümmeetria, mis tähendab, et laenuandja ei oma laenuvõtja kohta nii palju infot kui viimane ise. Võrreldes pankade rahastatud laenudega, on p2p laenud tihti kõrgemate intressidega. Üheks põhjuseks on, et valdavalt on p2p laenud tagatiseta. Lisaks esineb p2p keskkondades nn valikunihe (*selection bias*), mis tähendab, et p2p keskkondades tegutsevad riskantsemad taotlejad. P2p laenukeskkonnas investeerimisele lisandub keskkonna risk, mis tuleneb olukorrast, kus p2p platvorm lõpetab tegevuse.

P2p laenukeskkondade turg on maailmas traditsioonilise pangandussektori kõrval suhteliselt uus (vanusega 10 aastat), aga kiirelt kasvav. USAs ja Euroopas väljastati 12. kuu jooksul kuni 2015. aasta märtsini ligikaudu 9 mld euro väärtuses laene, aastane kasv oli 145%. Euroopa kontekstis rahastati 2014. aastal alternatiivrahastuse turul 3 mld euro ulatuses

tehinguid, 2015. aastaks prognoositakse mahu rohkem kui kahekordistumist. Bondora pakub investoritele rahvusvahelist investeerimisvõimalust.

Võrreldes väärtpaberitega rakenduvad laenuportfelli investeerimisel iseärasused nagu laenutulude ebasümmeetrilisus. Laenuportfelli koostades tuleks üheaegselt silmas pidada nii teenitavat tulu kui võetavat riski ning nende tasakaalu. Varasemalt on p2p laenude teemal tehtud mitmeid töid, milles on keskendutud laenukvaliteedi hindamiseks vajalikele tunnustele. Masinõppe ja otsustuspuude uuringuid on tehtud vähem, aga paljulubavate tulemustega.

Bondora laenukeskkond pakub investoritele valikute tegemiseks kahte lähenemist – portfellihalduri teenuse kasutamist ilma konkreetseid laenuaotluseid analüüsivõimata ning aktiivse investori rolli, kes valib investeeringud individuaalsete laenuaotluste seast.

Uurimismeetoditena kasutas autor masinõppe meetodeid loomaks otsustuspuud Bondora laenuandmetele tuginedes. Otsustuspuu valikut põhjendasid meetodi lihtsus ja võimekus. Ühelt poolt on otsustuspuu visuaalselt mõistetav, teiselt poolt annab täpse tulemuse ja saab hakkama suurte andmemahudega. Otsustuspuu täiendus *random foresti* näol, mis kätkeb endas mitme otsustuspuu korruga kasutamist, parandab mudeli täpsust. Väheneb risk, et otsustuspuu on üle-kohaldatud ja seetõttu uute andmetega eksiv. Autor tekitas neli otsustuspuude mudelit, nendest kaks *random forest* kontseptsioonil, kahe puhul kasutati üksikute otsustuspuude lähenemist. Mudelid arvestasid klassifitseerimiskulusid. Autor eeldas, et iga heaks hinnatud laen, mis tegelikult pankrotistub, põhjustab investorile viis korda suurema kulu kui edukas laen, mille mudel halvaks hindab. Lõplik valim koosnes 4481 laenust.

Lisaks viidi Bondora esindajaga läbi intervjuu eesmärgiga kontrollida autori eelduseid ning saada rohkem taustainfot p2p keskkondadest. Bondora pidas otsustuspuude rakendamise ideed perspektiivikaks. Seisukoht oli, et otsustuspuu kasutamine tõenäoliselt parandab tootlust Bondora krediidireitingute süsteemiga võrreldes. Siiski võivad meetodi igapäevasel rakendamisel tekkida probleemid seoses ajakuluga sobivate investeeringute otsimisel. Kokkuvõttes taandub küsimus sellele, kas mudelite investeerimissoovitused pakuvad piisavalt suuremat tootlust võrreldes Bondora krediidireitingute süsteemiga.

Mudelid paistsid kirjeldavate karakteristikute poolest erinevalt silma. Kõikide hinnangute klassifitseerimise alusel omasid suurimat täpsust *random forest* mudelid, samas

andsid lihtsate otsustuspuude mudelid vähem valepositiivseid hinnanguid. Mudelite võrdlemisel omavahel ja Bondora krediidireitingutega koostas autor viis ühesuguselt stardipositsioonilt alustavat investeerimisstrateegiat. Krediidireitingute alusel investeerimine kätkes alustamist madalama riskiga laenudest, lõpetades kõrgema riskiga lepingutega. Otsustuspuud järjestasid investeringud kõigepealt otsuse kindluse koefitsiendi järgi. Teisena järjestasid kõik strateegiad investeringud intressimäära alusel. Autor võrdles omavahel strateegiate kumulatiivset tootlust ja halbade laenude osakaalu.

Otsustuspuude tulemus ületas krediidireitingute alusel investeerimist tähelepanuväärselt nii halbade laenude osakaalu minimeerimise kui teenitava intressitulu maksimeerimise osas. Kõige kehvem klassikalise otsustuspuu mudel korjas portfelli esimesed halvad laenud, olles eelnevalt edukalt investeerinud 25% kogu vahenditest. Halbade laenude vältimises parimaks osutunud *random forest* mudel eksis 94% investeringute tegemise järgselt. Kõik mudelid parandasid investori tootlust keskmiselt 12% võrra, olles sel hetkel paigutanud 10% portfelist. Tulemuste tõlgendamisel peab arvestama, et pikas perspektiivis võivad need praegustest erineda, sest otsustuspuud ei saanud valimi tõttu arvestada majanduskriisi võimaliku mõjuga.

Käesoleva töö kirjutamise ajal on keeruline igapäevaste investeerimisotsuste langetamisel mudeleid rakendada, sest ei esine eristuvaid laenuvõtja tunnuseid, mille abil Bondoras investeringuid otsida või filtreerida. On tarvis ligipääsu kõikidele turuolevatele investeringutele Exceli tabeli kujul, et infot mudelisse sisestada. Selleks automaatselt vahendit Bondora ei paku, aga tulevikus on ettevõttel plaanis vastav lahendus välja arendada.

Töö edasiarenduse valdkonnad puudutavad mudeli testimise meetoodika täiustamist, muutes seda igapäevase investeerimisega võrreldavamaks. Olukorras, kus Bondora ei toeta automaatselt turuolevate investeringute informatsiooni allalaadimist Exceli tabelisse, tuleks uurida, kas otsustuspuude mudelitesse antava informatsiooni mahtu saab kärpida mudeli prognoosimisvõimet vähendamata. Mudelite täiustamise kohapealt tuleb kõne alla riikide kaupa mudelite katsetamine, eesmärgiga mõista, kas õnnestub luua erinevaid, vastava riigi laenaja analüüsimiseks täpsemaid otsustuspuuid. Bondora esindaja väitel loob väärtust p2p laenuäri valdkonda puudutavate regulatsioonide uurimine, mis on akadeemilises valdkonnas ebapiisavalt kaetud.

VIIDATUD ALLIKAD

3.3.3 Kodumajapidamistele antud laenude jääk ja arv laenuliigi, valuuta ja tagatise lõikes.

Eesti Pank.

<http://statistika.eestipank.ee/?lng=et#listMenu/1172/treeMenu/FINANTSSEKTOR/147/650> (30.03.2015)

Altfi.com. Altfi portaal. <http://www.altfi.com/data/?level1=data&level2=&level3=&level4=> (03.04.2015)

Altman, E., Haldeman, R., Narayanan, P. (1977). ZETA ANALYSIS A New model to identify bankruptcy risk of corporations. - *Journal of Banking and Finance*. No 1, pp. 29–54.

Andmete eksportimine. Bondora. <https://www.bondora.ee/et/et/datasets> (11.04.2015)

API. (2013). Encyclopedia Britannica.

<http://www.britannica.com/EBchecked/topic/1472947/API> (14.04.2015)

Arya, S., Eckel, C., Wichman, C. (2013). Anatomy of the credit score. - *Journal of Economic Behavior and Organization*. Vol 95, pp. 175–185.

Automaatne investeerimine. Bondora. <https://www.bondora.ee/et/invest/portfolio-manager> (30.03.2015)

Bachmann, A., Becker, A., Buerckner, D., Hilker, M., Kock, F., Lehmann, M., Tiburtius, P., Funk, B. (2011). Online peer-to-peer lending - A literature review. - *Journal of Internet Banking and Commerce*. Vol 16, pp 1-18.

Banking without banks. The Economist. <http://www.economist.com/news/finance-and-economics/21597932-offering-both-borrowers-and-lenders-better-deal-websites-put-two> (31.03.2015)

Basti, E., Kuzey, C., Delen, D. (2015). Analyzing Initial Public Offerings' Short-Term Performance Using Decision Trees and SVMs. - *Decision Support Systems*. Vol 73, pp. 15–27.

Bondora laenutaotlus - remont. Bondora. <https://www.bondora.ee/et/Auction/Show/1c7ebd5c-c5f1-4ab4-bc4e-a46b00b95917?title=Remont> (30.03.2015)

Bondora rating. Bondora. <https://www.bondora.ee/et/invest/rating> (04.04.2015)

- Breiman, L. (2001). Random forests. - *Machine learning*. Vol 45, pp. 5–32.
- Czajkowski, M., Czerwonka, M., Kretowski, M. (2015). Cost-sensitive global model trees applied to loan charge-off forecasting. - *Decision Support Systems*. Pp 1-23.
- Cubiles-De-La-Vega, M-D., Blanco-Oliver, A., Pino-Mejías, R., Lara-Rubio, J. (2013). Improving the management of microfinance institutions by using credit scoring models based on Statistical Learning techniques. - *Expert Systems with Applications*. Vol 40, iss 17, pp. 6910–6917.
- Decision Trees – Using The Available Data to Identify Lending Opportunities on Bondora. P2P-Banking. <http://www.p2p-banking.com/countries/baltic-decision-trees-using-the-available-data-to-identify-lending-opportunities-on-bondora-part-1/> (30.03.2015)
- Domingos, P. (1999). MetaCost: A General Method for Making Classifiers Cost-Sensitive. - *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*. Vol 55, pp. 155–164.
- Eeldatav tootlus. Bondora. <https://www.bondora.ee/et/investeeri/statistika/intress> (04.04.2015)
- Emekter, R., Tu, Y., Jirasakuldech, B., Lu, M. (2014). Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending. *Applied Economics*, Vol. 47, iss 1, pp. 54–70.
- Fawcett, T. (2006). An introduction to ROC analysis. - *Pattern Recognition Letters*. Vol 27, iss. 8, pp. 861–874.
- Florez-Lopez, R., Ramon-Jeronimo, J. M. (2015). Enhancing accuracy and interpretability of ensemble strategies in credit risk assessment. A correlated-adjusted decision forest proposal. - *Expert Systems with Applications*. Vol. 42, iss. 13, pp. 5737–5753.
- Funded accumulative monthly Lending Club. Nickelsteamroller. http://www.nickelsteamroller.com/#!/charts?vendor=lendingclub&chart=total_funded&resolution=1&accumulated=true (03.04.2015)
- Funded accumulative monthly Prosper. Nickelsteamroller. http://www.nickelsteamroller.com/#!/charts?vendor=prosper&chart=total_funded&resolution=1&accumulated=true (03.04.2015)
- Gollinger, T. L., Morgan, J. B. (1993). Calculation of an Efficient Frontier for a Commercial Loan Portfolio. - *The Journal of Portfolio Management*. Vol 19, pp. 39–46.
- Herzenstein, M., Dholakia, U. M., Andrews, R. L. (2011). Strategic Herding Behavior in Peer-to-Peer Loan Auctions. - *Journal of Interactive Marketing*, Vol. 25, iss. 1, pp. 27–36.

- Hulme, M., Wright, C. (2006). Internet based social lending: Past, present and future. - *Social Futures Observatory*. pp. 1-115
- Iyer, R., Khwaja, A.I., Luttmer, E.F.P., Shue, K. (2009). Screening in New Credit Markets Can Individual Lenders Infer Borrower Creditworthiness in Peer-to-Peer Lending? - *Harvard Kennedy School Faculty Research Working Papers Series*. pp. 1-42.
- Jacobson, T., Roszbach, K. (2003). Bank lending policy, credit scoring and value-at-risk. *Journal of Banking and Finance*. Vol 27, iss. 4, pp. 615–633.
- Kim, J., Choi, K., Kim, G., Suh, Y. (2012). Classification cost: An empirical comparison among traditional classifier, Cost-Sensitive Classifier, and MetaCost. - *Expert Systems with Applications*. Vol 39, iss. 4, pp. 4013–4019.
- Klaft, M. (2008). Online Peer-to-Peer Lending: A Lenders' Perspective. - *SSRN Electronic Journal*. pp. 1-5.
- Korduma kippuvad küsimused. Bondora.
<https://www.bondora.ee/et/investeeri/investeerimisjuhis> (04.04.2015)
- Laenuaotlused. Bondora. <https://www.bondora.ee/et/invest/primary-market#> (30.03.2015)
- Lee, E., Lee, B. (2012). Herding behavior in online P2P lending: An empirical investigation. *Electronic Commerce Research and Applications*. Vol 11, iss. 5, pp. 495–503.
- Liberum AltFi Volume Index Europe. AltFi. <http://www.altfi.com/data/indices/EURvolume> (03.04.2015)
- Liberum AltFi Volume Index UK. AltFi. <http://www.altfi.com/data/indices/UKvolume> (03.04.2015)
- Malekipirbazari, M., Aksakalli, V. (2015). Risk assessment in social lending via random forests. - *Expert Systems with Applications*. Vol 42, iss. 10, pp. 4621–4631.
- Markowitz, H. (1952). Portfolio Selection. - *Journal of Finance*. Vol 7, pp. 77–91.
- Mild, A., Waitz, M., Wöckl, J. (2015). How low can you go? — Overcoming the inability of lenders to set proper interest rates on unsecured peer-to-peer lending markets. - *Journal of Business Research*. pp. 1-15.
- P2P Lending Services Open to International, Non-Resident Investors. P2P-Banking.
<http://www.p2p-banking.com/p2p-lending-services-open-to-international-non-resident-investors/> (30.03.2015)
- Parker, M. (2015). Sampling With Replacement and Sampling Without Replacement.
<http://www.ma.utexas.edu/users/parker/sampling/repl.htm> (10.05.2015)

- Patil, D., Bichkar, R. (2012). Issues in Optimization of Decision Tree Learning: A Survey. - *International Journal of Applied Information Systems*. Vol. 3, iss. 5, pp. 13–30.
- Pertman, Taavi. Bondora esindaja. Autori intervjuu. Helisalvestis. (14.04.2015)
- RahaFoorum. Rahafoorum. <http://rahafoorum.ee/> (13.04.2015)
- Rõivas, T., Mikser, S., Reinsalu, U. (2015). EESTI REFORMIERAKONNA, SOTSIAALDEMOKRAATLIKU ERAKONNA ning ERAKONNA ISAMAA JA RES PUBLICA LIIT kokkulepe valitsuse moodustamise ja valitsusliidu tegevusprogrammi põhialuste kohta. <http://www.sotsdem.ee/wp-content/uploads/2015/04/RE-SDE-ja-IRLi-valitsusliidu-lepe.pdf> (15.04.2015)
- Samm-sammult. Omaraha. <https://www.omaraha.ee/et/info/investing-step-by-step/> (03.03.2015)
- Schneider, J. (1997). Cross validation. <http://www.cs.cmu.edu/~schneide/tut5/node42.html> (16.04.2015)
- Seasoned Returns as of September 30, 2014. Prosper. <https://www.prosper.com/invest/marketplace-performance/> (30.03.2015)
- Siers, M. J., Zahidul Islam, M. (2015). Software defect prediction using a cost sensitive decision forest and voting and a potential solution to the class imbalance problem. - *Information Systems*. Vol. 51, pp. 62–71.
- Stevenson, B., Fadil, M. (1995). Modern Portfolio Theory : Can It Work for Commercial Loans ? - *Lending rev*. Vol 4, pp. 1-10.
- Zhang, D., Zhou, L. (2004). Discovering golden nuggets: Data mining in financial application. - *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*. Vo. 34, pp. 513–522
- Zhang, H., Su, J. (2006). Learning probabilistic decision trees for AUC. - *Pattern Recognition Letters*. Vol. 27, iss. 8, pp. 892–899.
- Teenitud tootlus. Bondora. <https://www.bondora.ee/et/investeeri/statistika/tootlus> (03.04.2015)
- Vadera, S. (2005). Inducing Cost-Sensitive Non-Linear Decision Trees. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.108.240&rep=rep1&type=pdf> (14.05.2015)
- Walker, M. (2013). Random Forests Algorithm. - <http://www.datasciencecentral.com/profiles/blogs/random-forests-algorithm> (16.04.2015)

Wardrop, R., Zhang, B., Rau, R., Gray, M. (2015). - The European Alternative Finance Benchmarking Report. <http://www.jbs.cam.ac.uk/index.php?id=6481#.VT0tICGqpBd> (17.04.2015)

Viires, K. (2015). Intervjuu tüübid ja meetodid. <http://intervjuu.weebly.com/intervjuu-tuumluumlbid-ja-meetodid.html> (06.04.2015)

Üldine statistika. Bondora. <https://www.bondora.ee/et/investeeri/statistika> (04.04.2015)

Yap, B. W., Ong, S. H., Husain, N. H. M. (2011). Using data mining to improve assessment of credit worthiness via credit scoring models. - *Expert Systems with Applications*. Vol. 38, iss. 10, pp. 13274–13283.

Yum, H., Lee, B., Chae, M. (2012). From the wisdom of crowds to my own judgment in microfinance through online peer-to-peer lending platforms. - *Electronic Commerce Research and Applications*. Vol. 11, iss. 5, pp. 469–483.

SUMMARY

BUILDING A LOAN PORTFOLIO WITH OPTIMAL RETURN FOR A PRIVATE INVESTOR – APPLYING DECISION TREE METHODOLOGIES IN PEER-TO-PEER LOAN ENVIRONMENT BONDORA

Martin-Leo Kisand

Financing small loans via internet-based environment acting as an intermediary is a relatively new field. The first enterprise allowing investment into loans through the internet by initially setting a preferred risk level is Zopa, a UK based peer-to-peer (p2p) loan business founded in 2005. P2p lending is an alternative to financing via banks by providing an opportunity to finance small loans in a more transparent way and with smaller transaction costs. The central research problem in this paper is that different approaches exist in making everyday investment decisions in p2p environments. This suggests that there is room for improvement moving towards a best standard in investment decision-making.

The aim of this paper is to determine the principles of building a loan portfolio with optimal return by applying decision tree approach on the empirical data of Bondora, a p2p loan environment located in Estonia but allowing access to international investors. Author constructs a research hypothesis that Bondora provides the highest average return among other p2p environments, considering the risk as well. The second research hypothesis is that investment returns can be improved using decision trees when compared to investment decisions based on credit rating system in Bondora.

Investing in p2p loan environments entails different risks compared to investing *e.g.*, in stocks. A selection bias in p2p platforms implies that usually p2p borrowers are riskier than the ones receiving funding from banks. Furthermore, p2p loans usually come with no collateral. Finally there is a risk that the specific p2p platform decides to end its business. Nevertheless, p2p loan market is growing rapidly having increased on average by 145% in 12

months until March 2015 when it financed loans in the amount of 9 billion euros. It is estimated that financed amounts will double during 2015.

The data machine learning methods used in this paper consisted of decision tree and its enhanced version random forest. Author built four models using total or decreased amount of data about borrowers. A database of Bondora historical data in the volume of 4481 loans was used to train the models. In addition an interview was made with representative of Bondora to validate the author's assumptions and obtain background information on p2p loan market.

The results of the interview suggested that decision tree approach could enable investors to achieve higher returns (compared to investing based on credit ratings) but it may become difficult to apply the models in practice because of need to search for suitable loans.

The decision tree and random forest models were matched against Bondora's credit rating system. Comparison was carried out by constructing different investment strategies with access to same loans and amount of monetary resources. Strategies differed only in sorting and choosing among all the loans on the market. The author measured the cumulative net return investor received (taking into account the losses generated when investing into a bad loan) and the cumulative ratio of bad loans in total number of loans invested in.

Results showed that decision tree models (incl. random forest) substantially succeeded investing based on credit ratings. The models reduced the amount of bad loans in portfolio and provided on average a 12% higher annual return compared to credit ratings when investor had selected the first 10% of the total loans available. The most accurate model was based on random forest and made its first bad decision upon exceeding the 94% threshold of all loans. Simpler decision tree model invested 25% of the market when it first suggested a bad loan.

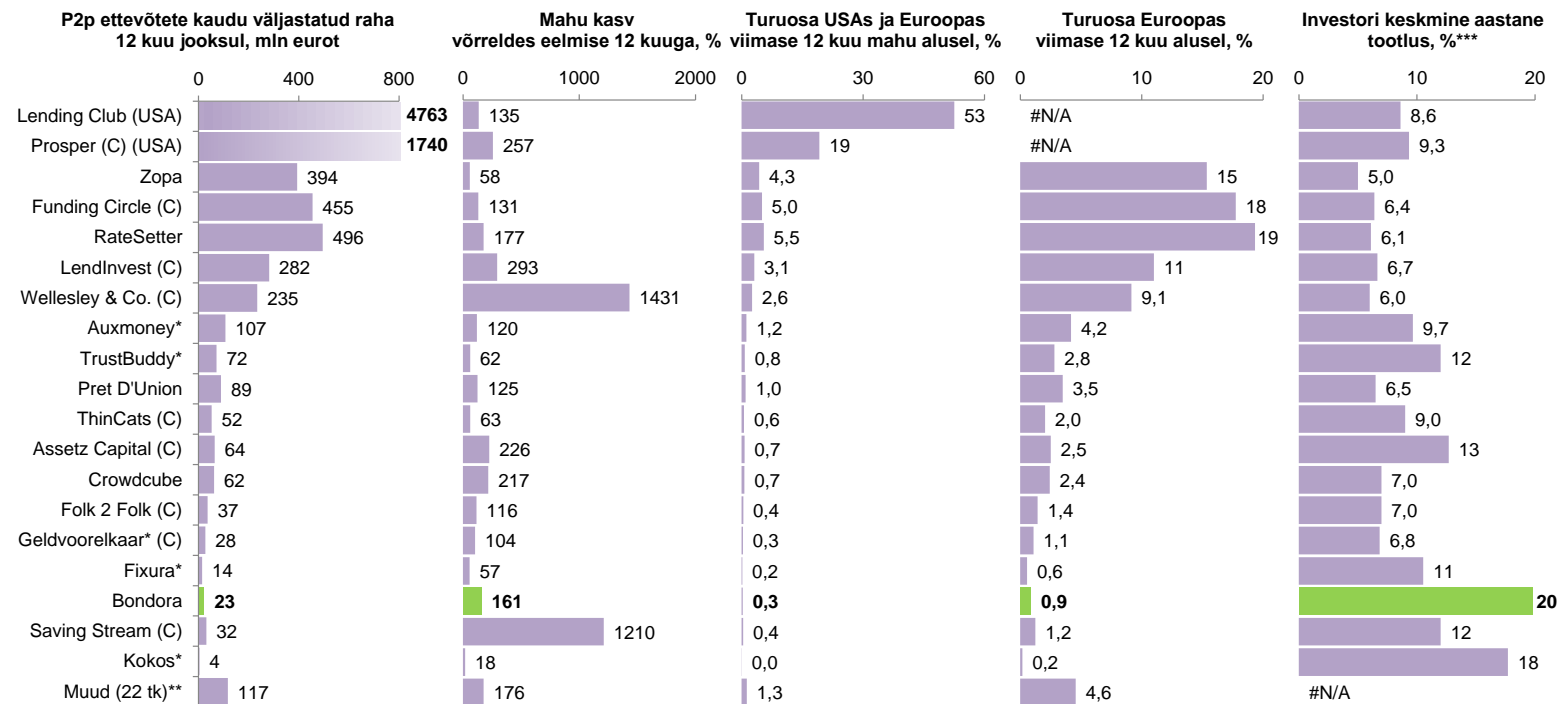
In conclusion, the author proved that decision tree models (incl. random forest) improve investor return compared to investing based on credit ratings. Still the results should be treated with some caution because the sample used when training the models did not include any loans experiencing economic recession. By analysing the p2p loan market, author proved that Bondora's average interest rate is higher than elsewhere in US or Europe. The latter conclusion is sensitive to specific loans and p2p loan providers taken into consideration.

When it comes to everyday appliance of the decision tree models (incl. random forest), technical difficulties exist. At the moment of writing this paper, there is no automatic way to access available loans in the market in a form which is similar to database used when training the models. As the interview suggested, a solution may become available in the future.

Future research concerning p2p loan environment should include improving the testing of decision tree models by making it more comparable to real-life investing. Additionally, it should be studied if the required amount of information by decision tree models could be reduced without hampering their predictive power. Also it should be studied if lenders national origin impacts his or her credit risk. Finally, as p2p loan market is growing in a fast pace, further studies concerning the impact of market regulations is necessary.

LISAD

Lisa 1. USA ja Euroopa p2p laenuurgu iseloomustavad näitajad



(USA) - asukoht USA-s (ülejäänud Euroopas); (C) - rahastatakse osaliselt või täielikult ettevõtteid (ülejäänud ainult eraisikuid); * - AltFi.com hinnang, ** - ettevõtteid kumulatiivse rahastatud sumмага < 20 mln eurot on summeeritud; *** ettevõtete kodulehtedel esitatud hinnanguline keskmine tootlus seisuga 03.04.2015 (autori koostatud); #N/A - andmeid ei ole

Joonis 8. Suurimad P2p platvormid USAs ja Euroopas ja nende olulisemad turuosa ja tootluse näitajad

Allikas: autori koostatud p2p portaalide ja ettevõtete kodulehtede andmete alusel (Liberum AltFi Volume Index Europe 2015; Liberum AltFi Volume Index UK 2015; Funded accumulative monthly Lending Club 2015; Funded accumulative monthly Prosper 2015)

Lisa 2. Bondora portfelli halduri seadistuste illustatsioon

Portfelli üksikasjad

Portfelli investeeritav summa € ?

VABA KONTOJÄÄK: 18,34€

Portfelli jaotus

LÄHTESTA

BONDORA RATING	€	AA	A	B	C	D	E	F	HR
UUS PORTFELLI JAOTUSE SIHT	0 %	1 %	2 %	5 %	9 %	20 %	16 %	10 %	37 %
INVESTEERITAV SUMMA ÜHTE LAENU		15 €	15 €	20 €	20 €	40 €	20 €	10 €	40 €
PORTFELLI JAOTUS HETKEL	1,57%	0%	3,05%	5,08%	14,23%	15,63%	12,6%	13,25%	34,59%
Ajalooline osakaal		0,30 %	2,12 %	6,17 %	12,13 %	18,91 %	20,45 %	11,53 %	28,39 %
Prognoositav keskmine intressimäär		14,86 %	16,44 %	18,19 %	21,83 %	26,45 %	32,03 %	40,53 %	65,91 %
Oodatav kahjumäär		1,67 %	2,57 %	4,42 %	7,28 %	10,74 %	15,13 %	21,48 %	45,52 %
Eeldatav tootlus		13,19 %	13,87 %	13,77 %	14,55 %	15,71 %	16,90 %	19,05 %	20,40 %

Joonis 9. Bondora portfelli halduri seadistusfunktsioonid

Allikas: (Automaatne investeerimine 2015)

Lisa 3. Bondora investeeringute otsingufiltri illustatsioon

Summa	<input type="text"/> - <input type="text"/>	Sugu	<input type="text" value="Kõik"/>
Soovitud intress	<input type="text"/> - <input type="text"/>	Vanus	<input type="text"/> - <input type="text"/>
Pikkus	<input type="checkbox"/> 3 <input type="checkbox"/> 9 <input type="checkbox"/> 12 <input type="checkbox"/> 18 <input type="checkbox"/> 24 <input type="checkbox"/> 36 <input type="checkbox"/> 48 <input type="checkbox"/> 60	Omandisuhe elukohaga	<input type="text" value="-"/>
Laenuaotluse vanus	<input type="text" value="Kõik"/>	Haridus	<input type="text" value="-"/>
Riik	<input type="checkbox"/> Eesti <input type="checkbox"/> España <input type="checkbox"/> Slovensko <input type="checkbox"/> Suomi	Perekonnaseis	<input type="text" value="-"/>
Krediidigrupp	<input type="checkbox"/> A <input type="checkbox"/> B <input type="checkbox"/> C	Ülalpeetavate arv	<input type="text" value="-"/>
Maksekäitumine	<input type="checkbox"/> 600 <input type="checkbox"/> 700 <input type="checkbox"/> 800 <input type="checkbox"/> 900 <input type="checkbox"/> 1000	Töösuhte liik	<input type="text" value="-"/>
Laenukoormus	<input type="text"/> - <input type="text"/>	Tegevusala	<input type="text" value="-"/>
Sissetulek	<input type="text"/> - <input type="text"/>	Ametinimetus	<input type="text" value="-"/>
% Täidetud	<input type="text"/> - <input type="text"/>	Kehtiva töösuhte aeg	<input type="text" value="-"/>
Eeldatav tootlus %	<input type="text"/> - <input type="text"/>	Tööstaaž kokku (aastates)	<input type="text" value="-"/>
Oodatava kahju määr %	<input type="text"/> - <input type="text"/>	Refin	<input type="text" value="Kõik"/>
Bondora Rating	<input type="checkbox"/> AA <input type="checkbox"/> A <input type="checkbox"/> B <input type="checkbox"/> C <input type="checkbox"/> D <input type="checkbox"/> E <input type="checkbox"/> F <input type="checkbox"/> HR	Palun kirjeldage oma igakuiseid püsikohustusi	<input type="text" value="Kõik"/>
Andmed on kinnitatud	<input type="checkbox"/> Sissetulek kinnitamata <input type="checkbox"/> Sissetulek kinnitatud <input type="checkbox"/> Sissetulek ja kohustused kinnitatud	Kasutaja	<input type="text"/>
		Otstarve	<input type="text" value="Kõik"/>
		Minu investeering	<input type="text" value="Kõik"/>
			<input type="button" value="Otsi"/> <input type="button" value="Kõik"/>

Joonis 10. Bondora investeeringute otsingufiltrite seadistusfunktsioonid

Allikas: (Laenuaotlused 2015)

Lisa 4. Bondora laenuaotluse näide

3800€ - 5000€
Laenusumma

18,25%
Intress

60k
Laenu pikkus

LAENUTAOTLUSE FINANTSEERIMINE

Intress: **18,25%**

Investeeringuid: **3765€ (85)**
(75,30% täidetud) 35€ puudu

Taotluse aeg: **30.märts.15 11:58**

Taotlus sulgub: **Kiirtaotlus**
06.apr.15 20:00

Aega investeerimiseks: **6p 22h 8m** ?

LAENUTAOTLEJA

Maksepäev: **17**

Maksekäitumine: **A 1000**

Vanus: **48 (Mees)**

Elukoht: **Hollola**

Kasutaja: **BOK91K63A**

Klient alates: **11.08.2014**

TEE INVESTEERING

Summa* € ?

Intress: **18,25%** ?

Investeeri

Vaba raha: 139€

Investeeringud: 0€

Laenujääk: 0€

Lisa jälgimisnimekirja

ÜLDINE **SEADED**

LAENAJA AJALUGU

Võlgasid:	0€	Tehtud tagasimakseid:	850€
Minu teenitud intressid:	0	Tagastatud ennetähtaegselt:	0 / 0€
Minu teenitud viivised:	0	Investeeringute jääk:	0€

Laenu sihtotstarve:

Muu info: ●

Bondora Rating: **B**

Lisa 4 järg

VIIMASED VIIS TAOTLUST					
LAENUTAOTLUS	LAENUSUMMA	INTRESS	LAENU PIKKUS	INVESTEERINGUID	TAOTLUS SULGUB
Muu tarbimine	3800€ - 5000€	31%	48 k	100,00% 5000€ (223 pakkumist)	20.okt.14
Remont	4700€ - 6200€	31%	60 k	27,10% 1680€ (63 pakkumist)	19.aug.14

ANDMED LAENUTAOTLUSELT			
Töösuhe:	Täiskohaga palgatööline	Haridus:	Kesk/keskeri
Tegevusala:	Tervishoid ja sotsiaalhoolekanne	Perekonnaseis:	Abielus
Ametikoht:	Keskastme juht	Ülalpeetavaid:	0
Töösuhte algus:	Rohkem kui 5 aastat	Eluaseme liik:	Omanik pangalaenuga
Töökogemus, a:	15 kuni 25 aastat		


LAENUTAOTLEJA IGAKUINE SISSETULEK JA KOHUSTUSED		
Palgatulu:	3400€	Laenukohustused ilma kodulaenu: 706,67€
Pension :	0€	Kodulaen: 1000€
Peretoetus :	0€	Uue laenu tagasimakse: 144€
Sotsiaaltoetused:	0€	Muud kohustused: 1020€
Vanemahüvitis :	0€	Andmed on kinnitatud: Sissetulek ja kohustused kinnitatud
Elatis :	0€	Kohustused peale refinantseerimist: 2870,67€
Muud tulud:	0€	
Igakuine sissetulek:	3400€	
	Vaba raha:	529,02€
	Laenukoormus:	25,03%

Lisa 4 järg

PALUN KIRJELDAGE OMA IGAKUISEID PÜSIKOHUSTUSI					
REFIN	KOHUSTUSE LIIK	LAENUANDJA	KUUMAKSE(€)	JÄÄK(€)	TAGATISE LIIK
PÜSIKOHUSTUSED					
	Elamiskulud		1020		
LAENUKOHUSTUSED					
Laen	Säästöpankki		500	18000	Kinnisvara
Laen	Santander		329	19000	Auto
Laen	Nordea		500	18000	Kinnisvara
Laen	Danske bank		180	8000	
Laen	isePankur		197,67	4777,48	
		KOKKU:	2726,67	67777,48	

KÜSIMUSED JA VASTUSED ×

MAKSEHÄIRED ×

 Laenajal puuduvad ametlikud maksehäired

Lisa 4 järg

INVESTEERINGUD				
LAENUANDJA	SUMMA	VÕITEV SUMMA	AEG	STAATUS
BOK7A694A	5€	5€	30.märts.15 13:17	Võitev
BO442946	5€	5€	30.märts.15 13:23	Võitev
BO9K3765A	10€	10€	30.märts.15 14:53	Võitev
BOA33974	200€	200€	30.märts.15 14:55	Võitev
metss	5€	5€	30.märts.15 15:04	Võitev
BO4979A5	5€	5€	30.märts.15 15:44	Võitev
BO22A577	10€	10€	30.märts.15 15:46	Võitev
BO495361	5€	5€	30.märts.15 16:10	Võitev
BO3K5769A	20€	20€	30.märts.15 16:50	Võitev
BO96A1KK3	5€	5€	30.märts.15 18:12	Võitev
BOA5KK93A	5€	5€	30.märts.15 18:49	Võitev
BOA7K757	10€	10€	30.märts.15 18:49	Võitev
BOA34K/4A	5€	5€	30.märts.15 19:41	Võitev
BO563231A	5€	5€	30.märts.15 19:48	Võitev
spaecker	5€	5€	30.märts.15 20:14	Võitev
BO491271A	5€	5€	30.märts.15 20:28	Võitev
BO62AA34A	25€	25€	30.märts.15 20:46	Võitev
lincolnman	10€	10€	30.märts.15 20:50	Võitev
BO291A42A	10€	10€	30.märts.15 20:57	Võitev
BO24A119	5€	5€	30.märts.15 21:45	Võitev
BO212331	5€	5€	30.märts.15 21:45	Võitev
Madisp	5€	5€	30.märts.15 21:47	Võitev
KOKKU: 3765€ (85)		3765€ (85)		

Joonis 11. Näide Bondora keskkonnas üleval olevast laenuaotlusest

Allikas: (Bondora laenuaotlus - remont 2015)

Lisa 5. Bondora laenuandmetes sisalduvad laenaja tunnused ja nende selgitused

Tabel 4. Bondora laenude toorandmetes laenuaotluseid iseloomustavad tunnused ja nende selgitused (kogu andmebaas)

Jrk nr.	Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnuse väärtused (ühikud, vahemikud)
1	14D FromFirstPayment	Kas esimese graafikujärgse makse kuupäevast on möödunud 14 päeva	0 - Ei; 1 - Jah
2	1D FromFirstPayment	Kas esimese graafikujärgse makse kuupäevast on möödunud 1 päev	0 - Ei; 1 - Jah
3	30D FromFirstPayment	Kas esimese graafikujärgse makse kuupäevast on möödunud 30 päeva	0 - Ei; 1 - Jah
4	60D FromFirstPayment	Kas esimese graafikujärgse makse kuupäevast on möödunud 60 päeva	0 - Ei; 1 - Jah
5	AD	Laenu tegelik pankrotistumine; laen loetakse pankrotistunuks, kui 3 järjestikust makset on tasumata	0 - Ei; 1 - Jah
6	Age	Laenaja vanus	Numbrilised väärtused (aastad)
7	AmountOfBids	Pakkumiste kogusumma	Numbrilised väärtused (eurot)
8	AmountOfInvestments	Investeeringute väärtus	Numbrilised väärtused (eurot)
9	AmountOfPreviousApplications	Varasemate laenuaotluste kogusumma	Numbrilised väärtused (eurot)
10	AmountOfPreviousLoans	Varasemate laenude kogusumma	Numbrilised väärtused (eurot)
11	ApplicationSignedHour	Laenuaotluse esitamise kellaeg tundides	Numbrilised väärtused (tund)

Lisa 5 järg

Jrk nr.	Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnuse väärtused (ühikud, vahemikud)
12	ApplicationSignedWeekday	Laenuaotluse esitamise nädalapäev	Numbrilised väärtused (nädalapäev)
13	ApplicationType	Millal makstakse laenusumma välja	1 Viitrahastus, 2 Kiirrahastus
14	AppliedAmount	Laenaja soovitatav laenumaht	Numbrilised väärtused (eurot)
15	AppliedAmountToIncome	Taotletud summa ja sissetuleku suhe, %	Numbrilised väärtused (%)
16	AssociatedDuplicateLoanId	Selgitus puudub Bondora andmebaasis	Tekstijada (kood)
17	BidsInvestmentPlan	isePakkujate tehtud investeerimispakkumiste arv	Numbrilised väärtused (tk)
18	BidsManual	Käsitsi tehtud investeerimispakkumiste arv	Numbrilised väärtused (tk)
19	BondoraCreditHistory	Kategooriamuutuja kolme võimaliku väärtusega:	Krediidialugu puudub, Hea krediidialugu, Halb krediidialugu
20	CancelledWithin14Days	Laenud, mis on tühistatud 14 päeva jooksul	0 - Ei; 1 - Jah
21	City	Laenuvõtja elukohalinn	Linna nimetus
22	ContractEndDate	Laenulepingu lõppemise kuupäev	Kuupäev
23	CountOfBankCredits	Pangalaenude arv	Numbrilised väärtused (tk)
24	CountOfOtherCredits	Muude laenukohustuste arv	Numbrilised väärtused (tk)
25	CountOfPaydayLoans	Kiirraenude arv	Numbrilised väärtused (tk)
26	Country	Laenaja asukohariik	Variandid: Eesti, Hispaania, Soome
27	County	Laenuvõtja elukohamaakond	Maakonna nimetus
28	credit_score	Laenaja varasemat maksekäitumist iseloomustav skoor	Variandid: 600, 700, 800, 900, 1000
29	CreditDecision	Bondora tehtud laenuotsus	0 tagasi lükatud, 1 vastu võetud, tühi Vastuvõtmisprotsessi pole

Lisa 5 järg

Jrk nr.	Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnuse väärtused (ühikud, vahemikud)
30	CreditGroup	Bondora liigitus laenaja krediidigruppi	Variandid: A, B, C
31	CurrentDebtDays	Laenu võlas oldud päevade koguarv toorandmete ajakohastamise ajal	Numbrilised väärtused (päevade arv)
32	CurrentLoanHasBeenExtended	Laenuvõtja on kehtiva laenu laenugraafikut muutnud	0 - Ei; 1 - Jah
33	DebtLiabilitiesBeforeLoan	Finantskohustuste arv enne laenu võtmist	Numbrilised väärtused (tk)
34	DebtRestructuringDate	Võlas oleva laenu uue tagasimaksegraafiku koostamise kuupäev	Kuupäev
35	DebtToIncome	Laenaja olemasolev laenukoormus	Numbriline vahemik (protsent)
36	Default_StartDate	Laenu pankrotistumise kuupäev	Kuupäev
37	DefaultedOnDay	Laenu pankrotistumiseni kulunud päevade arv	Numbrilised väärtused (päevade arv)
38	EAD1	Riskipositsioon pankrotistumisel, tagastamata põhiosa pankrotistumisel	Numbrilised väärtused (eurot)
39	EAD2	Riskipositsioon pankrotistumisel, laenusumma miinus kõik pankrotistumise ajaks tehtud maksed	Numbrilised väärtused (eurot)
40	education_id	Haridus	Variandid: algharidus või puudub, põhiharidus, kutseharidus, kesk/keskeri, kõrgharidus
41	EL	Oodatav kahjumäär (agregeeritud väärtused 0 ja 1 mudelitest)	Numbrilised väärtused (protsent)
42	EL_V0	Oodatav kahjumäär, mis põhineb versiooni 0 mudelil	Numbrilised väärtused (protsent)
43	EL_V1	Oodatav kahjumäär, mis põhineb versiooni 1 mudelil	Numbrilised väärtused (protsent)

Lisa 5 järg

Jrk nr.	Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnuse väärtused (ühikud, vahemikud)
44	Employment_Duration_Current_Employer	Kehtiva töölepingu sõlmimisest möödunud aeg	Variandid: katseaeg, vahemikud eraldi kuni 5 aastani, rohkem kui 5 aastat
45	employment_status_id	Millises töösuhtes laenaja hetkel on	Variandid: töötu, poole kohaga palgatööline, täiskohaga palgatööline, FIE, ettevõtja, pensionär
46	Exposure_at_DebtRestructuring	Tagastamata põhiosa ning kogunenud tasumata intress ja viivised kuupäeval DebtRestructuringDate	Numbrilised väärtused (eurot)
47	Exposure_at_Default	Võla kogusumma kuupäeval Default_StartDate	Numbrilised väärtused (eurot)
48	FirstPaymentDate	Esimese makse kuupäev algse laenugraafiku kohaselt	Kuupäev
49	FreeCash	Vabalt kasutatav sissetulek pärast igakuiste kohustuste täitmist	Numbrilised väärtused (eurot)
50	FundedAmount	Laenuvõtja saadud summa	Numbrilised väärtused (eurot)
51	Gender	Laenaja sugu	Mees, naine
52	GracePeriodEnd	Maksepuhkuse lõpu kuupäev	Kuupäev
53	GracePeriodStart	Maksepuhkuse algus kuupäev	Kuupäev
54	HasExtendedSchedule	Laenuvõtja on varasema laenu laenugraafikut muutnud	0 - Ei; 1 - Jah
55	home_ownership_type_id	Selgitus, kas laenaja on kinnisvaraomanik või näiteks üürnik	Variandid: kodutu, omanik, elan vanematega, üürnik eelmöbleeritud pinnal, üürnik eelmöbleerimata pinnal, munitsipaalomand, kaasüürnik, kaasomand, omanik pangalaenuga, omanik - omandil hüpoteek
56	income_from_principal_employer	Töötasu	Numbrilised väärtused (eurot)

Lisa 5 järg

Jrk nr.	Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnuse väärtused (ühikud, vahemikud)
57	income_other	Muu sissetulek	Numbrilised väärtused (eurot)
58	income_total	Laenaja igakuine sissetulek	Numbriline vahemik (eurot)
59	IncomeFromChildSupport	Lapsetoetus	Numbrilised väärtused (eurot)
60	IncomeFromFamilyAllowance	Peretoetus	Numbrilised väärtused (eurot)
61	IncomeFromLeavePay	Puhkusetasu	Numbrilised väärtused (eurot)
62	IncomeFromPension	Pension	Numbrilised väärtused (eurot)
63	IncomeFromSocialWelfare	Sotsiaalhoolekanne	Numbrilised väärtused (eurot)
64	InDebt14Day	See laen on mingil hetkel 14 päeva võlas olnud	0 - Ei; 1 - Jah
65	InDebt14Day_Exposure	Võla kogusumma kuupäeval InDebt14Day_StartDate	Numbrilised väärtused (eurot)
66	InDebt14Day_Interest	Kuupäevaks InDebt14Day_StartDate kogunenud tasumata intressisumma	Numbrilised väärtused (eurot)
67	InDebt14Day_Principal	Laenu tagastamata põhiosa summa kuupäeval InDebt14Day_StartDate	Numbrilised väärtused (eurot)
68	InDebt14Day_PrincipalProportion	Kuupäeval InDebt14Day_StartDate kehtinud laenu tagastamata põhiosa summa ja laenusumma suhe	Numbrilised väärtused (protsent)
69	InDebt14Day_StartDate	Esimene kuupäev, mil lepinguliste maksete tegemine oli 14 päeva hilinenud	Kuupäev
70	InDebt14Day_TotalRepayments	Kuupäevaks InDebt14Day_StartDate tasutud maksete (põhiosa + intress + viivised) kogusumma	Numbrilised väärtused (eurot)

Lisa 5 järg

Jrk nr.	Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnuse väärtused (ühikud, vahemikud)
71	InDebt1Day	See laen on mingil hetkel 1 päev võlas olnud	0 - Ei; 1 - Jah
72	InDebt1Day_Exposure	Võla kogusumma kuupäeval InDebt1Day_StartDate	Numbrilised väärtused (eurot)
73	InDebt1Day_Interest	Kuupäevaks InDebt1Day_StartDate kogunenud tasumata intressisumma	Numbrilised väärtused (eurot)
74	InDebt1Day_Principal	Laenu tagastamata põhiosa summa kuupäeval InDebt1Day_StartDate	Numbrilised väärtused (eurot)
75	InDebt1Day_PrincipalProportion	Kuupäeval InDebt1Day_StartDate kehtinud laenu tagastamata põhiosa summa ja laenusumma suhe	Numbrilised väärtused (protsent)
76	InDebt1Day_StartDate	Esimene kuupäev, mil lepinguliste maksete tegemine oli 1 päev hilinenud	Kuupäev
77	InDebt1Day_TotalRepayments	Kuupäevaks InDebt1Day_StartDate tasutud maksete (põhiosa + intress + viivised) kogusumma	Numbrilised väärtused (eurot)
78	InDebt21Day_Exposure	Võla kogusumma kuupäeval InDebt21Day_StartDate	Numbrilised väärtused (eurot)
79	InDebt21Day_Interest	Kuupäevaks InDebt21Day_StartDate kogunenud tasumata intressisumma	Numbrilised väärtused (eurot)
80	InDebt21Day_Principal	Laenu tagastamata põhiosa summa kuupäeval InDebt21Day_StartDate	Numbrilised väärtused (eurot)
81	InDebt21Day_PrincipalProportion	Kuupäeval InDebt21Day_StartDate kehtinud laenu tagastamata põhiosa summa ja laenusumma suhe	Numbrilised väärtused (protsent)
82	InDebt21Day_StartDate	Esimene kuupäev, mil lepinguliste maksete tegemine oli 21 päeva hilinenud	Kuupäev

Lisa 5 järg

Jrk nr.	Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnuse väärtused (ühikud, vahemikud)
83	InDebt21Day_TotalRepayments	Kuupäevaks InDebt21Day_StartDate tasutud maksete (põhiosa + intress + viivised) kogusumma	Numbrilised väärtused (eurot)
84	InDebt30Day	See laen on mingil hetkel 30 päeva võlas olnud	0 - Ei; 1 - Jah
85	InDebt30Day_Exposure	Võla kogusumma kuupäeval InDebt30Day_StartDate	Numbrilised väärtused (eurot)
86	InDebt30Day_Interest	Kuupäevaks InDebt30Day_StartDate kogunenud tasumata intressisumma	Numbrilised väärtused (eurot)
87	InDebt30Day_Principal	Laenu tagastamata põhiosa summa kuupäeval InDebt30Day_StartDate	Numbrilised väärtused (eurot)
88	InDebt30Day_PrincipalProportion	Kuupäeval InDebt30Day_StartDate kehtinud laenu tagastamata põhiosa summa ja laenusumma suhe	Numbrilised väärtused (protsent)
89	InDebt30Day_StartDate	Esimene kuupäev, mil lepinguliste maksete tegemine oli 30 päeva hilinenud	Kuupäev
90	InDebt30Day_TotalRepayments	Kuupäevaks InDebt30Day_StartDate tasutud maksete (põhiosa + intress + viivised) kogusumma	Numbrilised väärtused (eurot)
91	InDebt60Day	See laen on mingil hetkel 60 päeva võlas olnud	0 - Ei; 1 - Jah
92	InDebt60Day_Exposure	Võla kogusumma kuupäeval InDebt60Day_StartDate	Numbrilised väärtused (eurot)
93	InDebt60Day_Interest	Kuupäevaks InDebt60Day_StartDate kogunenud tasumata intressisumma	Numbrilised väärtused (eurot)
94	InDebt60Day_Principal	Laenu tagastamata põhiosa summa kuupäeval InDebt60Day_StartDate	Numbrilised väärtused (eurot)
95	InDebt60Day_PrincipalProportion	Kuupäeval InDebt60Day_StartDate kehtinud laenu tagastamata põhiosa summa ja	Numbrilised väärtused (protsent)

Lisa 5 järg

Jrk nr.	Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnuse väärtused (ühikud, vahemikud)
		laenusumma suhe	
96	InDebt60Day_StartDate	Esimene kuupäev, mil lepinguliste maksete tegemine oli 60 päeva hilinenud	Kuupäev
97	InDebt60Day_TotalRepayments	Kuupäevaks InDebt60Day_StartDate tasutud maksete (põhiosa + intress + viivised) kogusumma	Numbrilised väärtused (eurot)
98	InDebt7Day_Exposure	Võla kogusumma kuupäeval InDebt7Day_StartDate	Numbrilised väärtused (eurot)
99	InDebt7Day_Interest	Kuupäevaks InDebt7Day_StartDate kogunenud tasumata intressisumma	Numbrilised väärtused (eurot)
100	InDebt7Day_Principal	Laenu tagastamata põhiosa summa kuupäeval InDebt7Day_StartDate	Numbrilised väärtused (eurot)
101	InDebt7Day_PrincipalProportion	Kuupäeval InDebt7Day_StartDate kehtinud laenu tagastamata põhiosa summa ja laenusumma suhe	Numbrilised väärtused (protsent)
102	InDebt7Day_StartDate	Esimene kuupäev, mil lepinguliste maksete tegemine oli 7 päeva hilinenud	Kuupäev
103	InDebt7Day_TotalRepayments	Kuupäevaks InDebt7Day_StartDate tasutud maksete (põhiosa + intress + viivised) kogusumma	Numbrilised väärtused (eurot)
104	Interest	Laenaja soovitatav intressimäär (aastas)	Numbrilised väärtused (protsent)
105	Interest_at_DebtRestructuring	Kogunenud tasumata intressisumma uue tagasimaksegraafiku koostamise ajal	Numbrilised väärtused (eurot)
106	InterestAndPenaltiesPaid	Laenulepingu raames makstud intress ja viivised	Numbrilised väärtused (eurot)

Lisa 5 järg

Jrk nr.	Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnuse väärtused (ühikud, vahemikud)
107	InterestDebt_at_Default	Kuupäevaks Default_StartDate kogunenud tasumata intressisumma	Numbrilised väärtused (eurot)
108	IsBusinessLoan	Kas laenu taotles juriidiline isik	0 - Ei; 1 - Jah
109	IsFirstPaymentDefault	Laen läks ilma esimest kuumakset tegemata 14 päeva viivisesse	0 - Ei; 1 - Jah
110	IsInactiveDuplicate	Selgitus puudub Bondora andmebaasis	0 - Ei; 1 - Jah
111	IssuedInterest	Lõplik intressimäär	Numbrilised väärtused (protsent)
112	language_code	1 Eesti	1 Eesti, 2 Inglise, 3 Vene, 4 Soome, 5 Saksa, 6 Hispaania, 7 Slovakkia
113	LastPaymentOn	Kuupäev, millal laenusaja viimati tagasimakse tegi	Kuupäev
114	LeftMoneyForFirstPayment	Kas laenaja jättis oma kontole raha esimese tagasimakse tegemiseks	0 - Ei; 1 - Jah
115	LiabilitiesToIncome	Kohustuste ja sissetuleku suhe, %	Numbrilised väärtused (protsent)
116	loan_id	Laenu taotlusele antud unikaalne ID	Kood
117	LoanApplicationStartedDate	Laenu taotluse alustamise kuupäev	Kuupäev
118	LoanDate	Laenu väljastamise kuupäev	Kuupäev
119	LoanDuration	Laenu perioodi pikkus	Variandid: 3, 9, 12, 18, 24, 36, 48 ja 60 (kuud)
120	LoanNumber	Bondora süsteemis laenu tähistav unikaalne number	Kood
121	marital_status_id	Perekonnaseis	Variandid: abielus, vabaabielus, vallaline, lahutatud, lesk
122	MaturityDate_Last	Laenu täieliku tagastamise kuupäev raporti koostamise kuupäeva seisuga	Kuupäev

Lisa 5 järg

Jrk nr.	Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnuse väärtused (ühikud, vahemikud)
123	MaturityDate_Original	Laenu täieliku tagastamise kuupäev algse laenugraafiku kohaselt	Kuupäev
124	MaxDebtMonthsFinance	Kõige pikem aeg, mille vältel finantsleping on võlas olnud (kõik laenud, krediitkaardid, arved jne) (ainult Eesti kohta)	Numbrilised väärtused (kuude arv)
125	MaxDebtMonthsOther	Kõige pikem aeg, mille vältel mõni muu leping on võlas olnud (ainult Eesti kohta)	Numbrilised väärtused (kuude arv)
126	MaxDebtMonthsTelco	Kõige pikem aeg, mille vältel telekommunikatsioonileping on võlas olnud (ainult Eesti kohta)	Numbrilised väärtused (kuude arv)
127	ModelVersion	Bondora Ratingu arvutamise mudeli versioon	Versiooni number
128	MonthlyPaymentDay	Laenumaksete graafikujärgne kuupäev	Kuupäev
129	NewCreditCustomer	Kas kliendil oli varasem Bondora krediidi ajalugu	0 - Ei; 1 - Jah
130	NewLoanMonthlyPayment	Uue laenu kuumakse	Numbrilised väärtused (eurot)
131	NewOfferMade	Algse laenu taotluse hindamise käigus struktureeriti see ümber ja pakuti pikemat laenu, suuremat/väiksemat laenusummat või kõrgemat intressimäära	0 - Ei; 1 - Jah
132	NewPaymentToIncome	Uue laenu kuumakse ja sissetuleku suhe, %	Numbrilised väärtused (protsent)
133	NoOfBids	Pakkumiste arv	Numbrilised väärtused (tk)
134	NoOfInvestments	Investeeringute arv	Numbrilised väärtused (tk)
135	NoOfPreviousApplications	Varasemate laenu taotluste arv	Numbrilised väärtused (tk)

Lisa 5 järg

Jrk nr.	Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnuse väärtused (ühikud, vahemikud)
136	NoOfPreviousLoans	Varasemate laenude arv	Numbrilised väärtused (tk)
137	nr_of_dependants	Ülalpeetavate arv	Variandid: 0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10; > 10
138	NumDebtsFinance	Finantslepingutega seotud võlgade koguarv (kõik laenud, krediitkaardid, arved jne) (ainult Eesti kohta)	Numbrilised väärtused (tk)
139	NumDebtsOther	Muude lepingutega seotud võlgade arv (ainult Eesti kohta)	Numbrilised väärtused (tk)
140	NumDebtsTelco	Telekommunikatsioonilepingutega seotud võlgade koguarv (ainult Eesti kohta)	Numbrilised väärtused (tk)
141	occupation_area	Millisel tegevusalal laenaja tegutseb	Variandid sarnaselt EMTAK tegevusalade registrile (nt töötlev tööstus, ehitus, veondus ja laondus jne)
142	OtherLiabilitiesBeforeLoan	Muude kohustuste (v.a. laenukohustused) arv enne laenu võtmist	Numbrilised väärtused (tk)
143	OutstandingPrincipal	Laenu tagastamata põhiosa toorandmete ajakohastamise ajal	Numbrilised väärtused (eurot)
144	PreviousEarlyRepayments	Varasemad ennetähtaegsed tagasimaksed	Numbrilised väärtused (eurot)
145	PreviousLateFeesPaid	Varasemalt makstud viivised	Numbrilised väärtused (eurot)
146	PreviousRepayments	Varasemad tagasimaksed	Numbrilised väärtused (eurot)
147	Principal_at_DebtRestructuring	Tagastamata põhiosa uue tagasimaksegraafiku koostamise ajal	Numbrilised väärtused (eurot)
148	Principal_at_Default	Laenu tagastamata põhiosa summa kuupäeval Default_StartDate	Numbrilised väärtused (eurot)

Lisa 5 järg

Jrk nr.	Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnuse väärtused (ühikud, vahemikud)
149	PrincipalDebtAmount	Võlgnetav põhiosa summa	Numbrilised väärtused (eurot)
150	PrincipalProportion_at_DebtRestructuring	Tagastamata põhiosa ja laenusumma suhe uue tagasimaksegraafiku koostamise ajal	Numbrilised väärtused (protsent)
151	PrincipalProportion_at_Default	Kuupäeval Default_StartDate kehtinud laenu tagastamata põhiosa summa ja laenusumma suhe	Numbrilised väärtused (protsent)
152	Rating	Bondora Rating (agregeeritud väärtused 0 ja 1 mudelitest)	Variandid: AA, A, B, C, D, E, F, HR
153	Rating_V0	Bondora Rating, mis põhineb versiooni 0 mudelil	Variandid: AA, A, B, C, D, E, F, HR
154	Rating_V1	Bondora Rating, mis põhineb versiooni 1 mudelil	Variandid: AA, A, B, C, D, E, F, HR
155	Recovery	Pärast laenu pankrotistumist sissenõutud võlg, muu hulgas intressisumma ja viivised	Numbrilised väärtused (eurot)
156	ReportAsOfEOD	Raporti koostamise kuupäev	Kuupäev
157	ScoringDate	Bondora Ratingu arvutamise kuupäev (agregeeritud väärtused 0 ja 1 mudelitest)	Kuupäev
158	ScoringDate_V0	Laenu Ratingu arvutamise kuupäev	Kuupäev
159	ScoringDate_V1	Laenu Ratingu arvutamise kuupäev	Kuupäev
160	SumOfBankCredits	Pangalaenude kogusumma	Numbrilised väärtused (eurot)
161	SumOfOtherCredits	Muude laenukohustuste kogusumma	Numbrilised väärtused (eurot)
162	SumOfPaydayLoans	Kiirlaenude kogusumma	Numbrilised väärtused (eurot)
163	TotalLiabilitiesBeforeLoan	Mitu kehtivat rahalist kohustust laenuaotlejal on	Variandid: 0;1-4;5-9;10+
164	TotalMaxDebtMonths	Kõige pikem aeg, mille vältel laen on võlas olnud (ainult Eesti kohta)	Numbrilised väärtused (kuude arv)

Lisa 5 järg

Jrk nr.	Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnuse väärtused (ühikud, vahemikud)
165	TotalMonthlyLiabilities	Igakuiste kohustuste kogusumma	Numbrilised väärtused (eurot)
166	TotalNumDebts	Võlgade koguarv (ainult Eesti kohta)	Numbrilised väärtused (tk)
167	TotalRepayments_at_DebtRestructuring	Kõik enne uue tagasimaksegraafiku koostamist tasutud maksed	Numbrilised väärtused (eurot)
168	TotalRepayments_at_Default	Kuupäevaks Default_StartDate tasutud maksete (põhiosa + intress + viivised) kogusumma	Numbrilised väärtused (eurot)
169	UnpaidInterestOutstanding	Võlas olev intressi- ja viivisesumma, mis hõlmab võlas olevalt põhiosa summalt ajal, kui see põhiosa summa võlas oli, arvestatud intressi	Numbrilised väärtused (eurot)
170	UseOfLoan	Laenu kasutusotstarve	Variandid: laenude konsolideerimine, kinnisvara, remont, äri, haridus, reisimine, liiklusvahend, muu tarbimine, tervis
171	UserName	Kliendi kasutajanimi Bondoras	Kasutajanimi
172	WasFunded	Näitab, kas investorid rahastasid laenuaotluse	0 - Ei; 1 - Jah
173	VerificationType	Millisel moel on Bondora laenuaotluse andmeid kontrollinud	Variandid: sissetulek kinnitamata, sissetulek kinnitatud, sissetulek ja kohustused kinnitatud
174	work_experience	Kaua on laenuaotleja töötanud	Variandid: < 2 a; 2-5 a; 5-10 a; 15-25 a; > 25 a

Allikas: Bondora kodulehe andmete alusel (Andmete eksportimine 2015) ning autori tõlgendusel olemasolevatest arvandmetest kodulehelt puudunud selgituste korral

Lisa 6. Otsustuspuude mudelitesse valitud laenajat iseloomustavad tunnused

Tabel 5. Autori kasutatud otsustuspuude mudelitesse valitud parameetrid (märgitud tähega x).

Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnus leitav taotluses	Tunnus määratav Bondora otsingufiltris
AD	Laenu tegelik pankrotistumine; laen loetakse pankrotistunuks, kui 3 järjestikust makset on tasumata	x	
Age	Laenaja vanus	x	x
AmountOfInvestments	Investeeringute väärtus	x	
AmountOfPreviousApplications	Varasemate laenutaotluste kogusumma	x	
AmountOfPreviousLoans	Varasemate laenude kogusumma	x	
ApplicationSignedHour	Laenutaotluse esitamise kellaeg tundides	x	
ApplicationSignedWeekday	Laenutaotluse esitamise nädalapäev	x	
ApplicationType	Millal makstakse laenusumma välja	x	
AppliedAmount	Laenaja soovitatav laenumaht	x	x
AppliedAmountToIncome	Taotletud summa ja sissetuleku suhe, %	x	
BondoraCreditHistory	Kategooriamuutuja kolme võimaliku väärtusega:	x	
City	Laenuvõtja elukohalinn	x	
CountOfBankCredits	Pangalaenude arv	x	
CountOfOtherCredits	Muude laenukohustuste arv	x	
CountOfPaydayLoans	Kiirlaenude arv	x	
Country	Laenaja asukohariik	x	x
County	Laenuvõtja elukohamaakond	x	
credit_score	Laenaja varasemat maksekäitumist iseloomustav skoor	x	x
CreditGroup	Bondora liigitus laenaja krediidi gruppi	x	x
DebtLiabilitiesBeforeLoan	Finantskohustuste arv enne laenu võtmist	x	
DebtToIncome	Laenaja olemasolev laenukoormus	x	x
education_id	Haridus	x	x

Lisa 6 järg

Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnus leitav taotluses	Tunnus määratav Bondora otsingufiltris
Employment_Duration_Current_Employer	Kehtiva töölepingu sõlmimisest möödunud aeg	x	
employment_status_id	Millises töösuhtes laenaja hetkel on	x	x
FreeCash	Vabalt kasutatav sissetulek pärast igakuiste kohustuste täitmist	x	
Gender	Laenaja sugu	x	x
home_ownership_type_id	Selgitus, kas laenaja on kinnisvaraomanik või näiteks üürnik	x	x
income_from_principal_employer	Töötasu	x	
income_other	Muu sissetulek	x	
income_total	Laenaja igakuine sissetulek	x	x
IncomeFromChildSupport	Lapsetoetus	x	
IncomeFromFamilyAllowance	Peretoetus	x	
IncomeFromLeavePay	Puhkusetasu	x	
IncomeFromPension	Pension	x	
IncomeFromSocialWelfare	Sotsiaalhoolekanne	x	
Interest	Laenaja soovitatav intressimäär (aastas)	x	x
language_code	1 Eesti	x	
LiabilitiesToIncome	Kohustuste ja sissetuleku suhe, %	x	
LoanDuration	Laenuperioodi pikkus	x	x
marital_status_id	Perekonnaseis	x	x
MonthlyPaymentDay	Laenumaksete graafikujärgne kuupäev	x	
NewCreditCustomer	Kas kliendil oli varasem Bondora krediidialalugu	x	
NewLoanMonthlyPayment	Uue laenu kuumakse	x	
NewPaymentToIncome	Uue laenu kuumakse ja sissetuleku suhe, %	x	
NoOfPreviousApplications	Varasemate laenutaotluste arv	x	
NoOfPreviousLoans	Varasemate laenude arv	x	
nr_of_dependants	Ülalpeetavate arv	x	x

Lisa 6 järg

Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Tunnus leitav taotluses	Tunnus määratav Bondora otsingufiltris
NumDebtsFinance	Finantslepingutega seotud võlgade koguarv (kõik laenud, krediitkaardid, arved jne) (ainult Eesti kohta)	x	
NumDebtsOther	Muude lepingutega seotud võlgade arv (ainult Eesti kohta)	x	
NumDebtsTelco	Telekommunikatsioonilepingutega seotud võlgade koguarv (ainult Eesti kohta)	x	
occupation_area	Millisel tegevusalal laenaja tegutseb	x	x
OtherLiabilitiesBeforeLoan	Muude kohustuste (v.a. laenukohustused) arv enne laenu võtmist	x	
PreviousEarlyRepayments	Varasemad ennetähtaegsed tagasimaksed	x	
PreviousRepayments	Varasemad tagasimaksed	x	
Rating	Bondora Rating (agregeeritud väärtused 0 ja 1 mudelitest)	x	x
SumOfBankCredits	Pangalaenude kogusumma	x	
SumOfOtherCredits	Muude laenukohustuste kogusumma	x	
SumOfPaydayLoans	Kiiralaenude kogusumma	x	
TotalLiabilitiesBeforeLoan	Mitu kehtivat rahalist kohustust laenutaotlejal on	x	x
TotalMonthlyLiabilities	Igakuiste kohustuste kogusumma	x	
TotalNumDebts	Võlgade koguarv (ainult Eesti kohta)	x	
UseOfLoan	Laenu kasutusotstarve	x	x
VerificationType	Millisel moel on Bondora laenutaotluse andmeid kontrollinud	x	x
work_experience	Kaua on laenutaotleja töötanud	x	x
	Kokku	64	21

Allikas: autori koostatud

Lisa 7. Töös kasutatud valimi puuduolevad väärtused tunnuste lõikes

Tabel 6. Puuduvaid väärtuseid sisaldanud tunnused (kokku 64-st tunnusest) autori lõplikus valimis (4481 laenuaotlust)

Tunnuse nimetus Bondora toorandmete andmebaasis	Sisuline kirjeldus	Puudu- olevate väärtuste arv	Puudu- olevate väärtuste osakaal (kokku 4481 laenu)
Rating	Bondora Rating	2497	56%
TotalLiabilitiesBeforeLoan	Kehtivate rahaliste kohustuste arv	1788	40%
DebtLiabilitiesBeforeLoan	Finantskohustuste arv enne laenu võtmist	1788	40%
OtherLiabilitiesBeforeLoan	Muude kohustuste (v.a. laenukohustused) arv enne laenu võtmist	1788	40%
TotalMonthlyLiabilities	Igakuiste kohustuste kogusumma	1788	40%
DebtToIncome	Laenaja olemasolev laenukoormus	1788	40%
FreeCash	Vabalt kasutatav sissetulek pärast igakuiste kohustuste täitmist	1788	40%
LiabilitiesToIncome	Kohustuste ja sissetuleku suhe, %	1788	40%
home_ownership_type_id	Selgitus, kas laenaja on kinnisvaraomanik või näiteks üürnik	1600	36%
TotalNumDebts	Võlgade koguarv (ainult Eesti kohta)	1060	24%
NumDebtsFinance	Finantslepingutega seotud võlgade koguarv (kõik laenud, krediitkaardid, arved jne)	1060	24%
NumDebtsTelco	Telekommunikatsioonilepingutega seotud võlgade koguarv (ainult Eesti kohta)	1060	24%
NumDebtsOther	Muude lepingutega seotud võlgade arv (ainult Eesti kohta)	1060	24%
Employment_Duration_Current_Employer	Kehtiva töölepingu sõlmimisest möödunud aeg	583	13%
County	Laenuvõtja elukohamaakond	346	8%
City	Laenuvõtja elukohalinn	325	7%
employment_status_id	Millises töösuhtes laenaja hetkel on	127	3%
occupation_area	Millisel tegevusalal laenaja tegutseb	42	1%
nr_of_dependants	Ülalpeetavate arv	13	0%
CreditGroup	Bondora liigitus laenaja krediidigrupi	10	0%
credit_score	Laenaja varasemat maksekäitumist iseloomustav skoor	9	0,2%
work_experience	Kaua on laenuaotleja töötanud	2	0,04%
	Tunnuseid kokku (tk)	22	

Allikas: autori koostatud

Lisa 8. Bondora esindajaga läbiviidud intervjuu küsimused

1. Miks on teie hinnangul pakutud p2p laenuplatvormidel investeerimisel välja erinevaid meetodeid, mitte ühte parimat lähenemist (nt rusikareeglid, krediidskooride järgi investeerimine, regressioonimudelite tulemuste alusel teatud parameetrite järgi investeerimine, otsustuspuud jt andmekaevet kasutavad lähenemised)? Kas oskate välja tuua nende kitsaskohad või positiivsed küljed?
2. Miks on Bondora otsustanud oma investoritele ligipääsetavaks muuta krediidskooridel põhineva lähenemise, portfelli halduri teenuse passiivsele investorile ja laenaja parameetrite järgi investeerimise aktiivsele investorile (kas ja miks peetakse neid alternatiive parimateks)?
3. Kas ja mil määral on võimalik väikeinvestoril Bondora krediidskooride tulemusi ise kontrollida?
4. Juhul kui jätta kõrvale krediidskoorid, siis millised on Bondora hinnangul olulisemad laenu taotlusest väljaloetavat parameetrid, mis võivad laenu hilisemat tagasimaksetõenäosust oluliselt mõjutada, mida peaks investor jälgima?
5. Kas Bondora pikaajaliseks eesmärgiks on pakkuda investoritele kõrget tootlust või pigem minimaalset riski (aga ka madalamat tootlust)?
6. Milline on Bondora hinnang otsustuspuu meetodika kasutamisele investeerimisotsuste langetamisel? (Kas see võib anda pikaajaliselt krediidskooridest lähtumisega võrreldes parema tulemuse? Miks?)
7. Kas Bondora plaanib lähiajal olulisi arendusi, mis võiksid mõjutada investorite senist käitumist Bondora portaalis? Kui jah, siis millised need on?
8. Millised võimalused on Bondoras investeerimiseks mitte portaali enda graafilise tarkvaraliidese kaudu vaid näiteks kasutades kasutaja isetehtud skripte ja päringuid andmebaasidest?
9. Milliseid probleeme või kitsaskohti näete täna p2p laenuäris, mis takistavad investoritel saamaks kõrgemat või väiksema riskiga intressitulu või laenajatel saamaks oma taotlusele õiglasemat intressimäära? Milliseid nendest saab lahendada p2p platvorm, milliseid riiklik regulatsioon või muu kolmas osapool?
10. Milliseid küsimusi tuleks p2p laenuäri käsitledes veelgi enam uurida, mis on täna alakaetud teemad?