

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond
Informaatikainstituut

IDK40LT

Kaspar Triebstok 134678 IAPB

KINNISVARA HINDAMISE
AUTOMATISEERIMISE VAJADUSED,
VÕIMALUSED JA TAKISTUSED

Bakalaureusetöö

Juhendaja: Tarvo Treier
Lektor

Tallinn 2016

Kokkuvõte

Töö eesmärgiks oli kaardistada kinnisvara hinnastamise automatiseerimise vajadused, võimalused ja takistavad probleemid. Lisaks sellele pakkuda välja automaatse hindamise protsess.

Kaardistati järgmised punktid:

1. hinnastamise meetod, mida automatiseerida,
2. andmed, mis on vajalikud selle tegemiseks,
3. kanalid, mille kaudu on võimalik vajalikud andmed pärida.

Koostati veebiteenuse arhitektuuriline plaan ning tehniline kirjeldus selle teostamiseks. Samuti töötati välja kontrollprotsess tulemuste kontrolliks ning valiti välja kõige potentsiaalikum algoritm masinõppe rakendamiseks.

Olulisemad järeldused:

Suurimad väljakutsed esinesid just kõige mõjukamate tegurite leidmisel ja nende parameetrite mõjutegurite määramisel.

Kuna piirkondi ja tegureid on palju, siis on õigete tegurite ja nende suuruse määramisel kõige optimaalsem kasutada masinõpet.

X-tee kasutamine on seoses versioonide muutusega keeruline. Aastal 2016 on X-tee läinud üle versioonilt 5 versioonile 6. Sellest tulenevalt on vajalik turvaserverite uuendamine ja ülemineku versiooni paigaldamine.

Võimalikud edasiarendused:

- Korterite planeeringu arvesse võtmine hoone keskmise turuväärtuse arvutamisel.
- Masinõppe realiseerimine.
- Hinnastamise kvaliteedi hinnete rangemaks muutmine (pt 4.2)
 - Vähendada vea protsenti, mille puhul toimuks kontrollprotsessi tsükli läbimine ning seeläbi viga tekitava teguri leidmine ja parandamine.

- Hinnastamise kontrollprotsessis teada saadud vea põhjuste leidmise algoritm ja selle ühendamine kontrollprotsessi.

1. Kas eesmärk saavutati

(vt jaotis 1.1, lk 11)

Töö eesmärgiks oli kaardistada ja välja tuua probleeme, mis takistavad täna kinnisvara hinnastamist automatiseerida. Lisaks sellele välja tuua võimalikud lahendused nendele probleemidele.

Antud lõputöö saavutas mõningate mõõndustega eesmärgid. Analüüsi tulemusena sai välja toodud pudelikaelad automatiseerimise juures ja samal ajal pandi kirja ka võimalikud lahendused. Selle käigus püstitati küsimused, millele tänaseks veel lõplikku vastust ei ole:

- Kuidas leida peatükis 4.3 pakutud kontrollprotsessi käigus selgunud vea protsendi põhjused?
- Sai välja toodud põhjused, miks Ehitisregistri andmete kvaliteet varieerub, kuid ei vastatud küsimusele, kuidas seda parandada.
- Kui täpne on pakutud lahenduse kvaliteet?

Kuna see lõputöö on analüütiline pool alles käimas olevale projektile, siis võib väita, et töö täitis üldiselt oma eesmärgi. Sai selgemalt kirjeldatud probleemid, millele on vaja tehniline lahendus leida. Sai kaardistatud lahenduse realiseerimisel esineda võivad probleemid.

2. Põhitulemuste loetelu

(vt jaotis 1.4, lk 13)

Milline näeb välja kinnisvara hindamine täna ja mis oleks majanduslik motivatsioon selle automatiseerimiseks.

Analüüsides varade hindamise standardiseeriat [11] ja intervjuuerides elukutselist kinnisvara hindajat [12] sai välja valitud hindamise meetod, mille automatiseerimine oleks kõige loogilisem. Selleks oli võrdlusmeetod. Selle tulemusel sai kaardistatud, kuidas seda meetodit hindajad kasutavad ning püstitatud probleem, kuidas saaks seda automatiseerida.

Analüüsi olemasolevaid kinnisvarade automatiseeritud hinnastamise lahendusi ettevõtte Zillow näitel ja kaardistatud Eestis populaarsed kinnisvaraportaalid.

Võttes aluseks võrdlusmeetodi pandi kirja, mis andmeid on vaja, et kinnisvara hindamist antud meetodiga teostada.

Puudused vajalikes kinnisvara andmetes.

Toodi välja puudused kinnisvara olemasolevates andmetes ning intervjuerides Ehitisregistri spetsialisti [26] kaardistati nende puuduste põhjused. Suurimaks põhjuseks on ehituseadustiku ja seal tulenevate nõuete pidev muutumine.

Samuti toodi välja Maa-ameti andmete puudused, milleks oli eelkõige registri ühenduse puudus X-teest. Kuna Maa-ameti statistilisi andmeid on keeruline kontrollida ja ei ole tingimata selle töö eesmärkide täitmiseks vajalik, siis ei antud ka hinnangut Maa-ameti andmete kvaliteedile.

Pakuti välja automaatse hinnastamise protsess järgnevate punktide näol:

- **Andmete pärimise ja töötluse tehniline lahendus.**
Leiti tehnilised võimalused, kuidas kinnisvara automaatset hinnastamist saaks realiseerida. Pandi kirja tehnilise spetsifikatsiooni osad, kuidas andmeid pärida ning protsess, mida saadud andmetega edasi teha. Lisaks selle töötati välja kontrollprotsess, millega oleks võimalik pakutud lahenduse tulemust kontrollida. Sellest tulenevalt analüüsi, kuidas ka seda protsessi võiks automatiseerida ehk sellele rakendada masinõpet.
- **Masinõppe rakendamine.**
Töötati välja protsess, millega oleks võimalik pakutud lahenduse tulemust kontrollida. Konkreetse lahenduse olemasolu tähendas seda, et sellele saaks rakendada masinõpet. Analüüsi olemasolevaid masinõppe algoritme. Lahendust ei realiseeritud, kuid toetudes teiste kogemustele [22], pakuti välja kõige potentsiaalikam algoritm võrdlusmeetodi tulemuste kontrollimiseks.

Summary

The goal of this thesis was to map out the needs, opportunities and obstacles in automating real-estate valuation. Propose a process to automate real-estate valuation.

The following topics were analyzed:

1. What valuation process to automate?
2. What is the needed data for automation?
3. What channels could be used to request this data?

An architectural plan and a technical description for implementing the plan was created. Also a control process to test the result was developed and an algorithm that could be used for implementing machine learning on the control process was chosen.

Major conclusions:

The value of a real-estate object is shaped by many factors. This thesis focused on the valuation of apartments. One of the biggest challenges was to map out the most important factors and see how much those factors alter the real-estate value in using the sales comparison approach.

Because of the large number of variables and their regression it is most optimal to find the best values via machine learning algorithms.

Using X-road in times when the entire protocol was updated and moved to from version 5 to version 6 was painfully slow and problematic

Proposed future developments:

- Implement the design plan of an apartment into the valuation algorithm.
- Realize proposed machine learning algorithm.
- Increase the quality of the machine learning algorithm by lowering the C grade error percentage in the control process proposed in chapter 4.2.

- In valuation checking process develop and algorithm to find the root cause of the error produced by the valuation algorithm.

Were the goals of this paper reached?

(See chapter 1.1 on page 11)

The goals were mainly reached with the exception of finding answers to some extra questions raised during the process. The following questions were not answered:

- How to find the root cause of the error produced by the control process proposed in chapter 4.3.
- The reasons why the quality of EHR data varies was brought out, but due to the involvement of legislative issues a solution was not developed nor proposed.
- The quality of the result produced by the proposed automation process was not assessed.

Due to the fact that this thesis is part of an ongoing project to automate the valuation of a real-estate object it can be said that it met its objective. As stated the goal was to map out the bottle necks and state more clearly the problems faced in the process of automating real-estate valuation. In regards to that this thesis achieved its goals. What is more this paper stated the next logical steps to be taken to improve the automation process.

The main results

- **How does the valuation of real-estate work today and what would be the economic value in automating that.**

Analyzing the property valuation standard [11] and interviewing a professional real-estate broker [12] the best valuation method to automate was chosen. As a result the sales comparison approach was chosen. After that a problem was stated how this method could be automated.

Excising automation methods were analyzed with the example of the Zillow Company.

What data is needed to automate the real-estate valuation process?

Information needed for the sales comparison approach was mapped. Channels that would provide the needed data were then researched. After that an architectural plan was composed on how the proposed process of automation would look like. The result of this was a concrete architectural plan of the proposed web service API and what technologies should be used for implementation.

Shortcomings in the data needed for sales comparison approach.

By analyzing the data from EHR and interviewing a developer of the EHR the shortcomings in the quality and quantity of the data needed for sales comparison approach were brought out. The main shortcoming in the EHR data was its varying quality caused by the constantly changing legislation and its requirements on the gathering of building data.

What is more deficiency's in Land Agency's data was brought out with the main being its missing connection with X-road. Since the quality of the Land Agency's data is somewhat difficult and also not necessarily needed for the purpose of this paper then it was left out of this thesis.

An automated real-estate valuation process was proposed in the form of the following points:

- **Technical description on how to acquire and process the data.**

Solutions were found on how to realize the proposed architectural plane. Specification was composed on where and how to query the needed data. After that a control process was developed on how to check the results of the initial automated valuation process.

- **How to implement machine learning?**

Different machine learning algorithms were analyzed and a potential algorithm was chosen to be used in implementing machine learning on the control process. The proposed algorithm was not realized.