

## *Projekti MWh*

Kari Luonteri

## Sisällys

<b>Projekti MWh</b> .....	1
Projektin tarkoitus .....	3
Ongelman kuvaus .....	3
Miten lähestyä ongelmaa .....	3
Asuntojen lämmitykseen kuluva energia.....	3
Lämpimän veden kierrätyksessä tapahtuva lämpöhäviö. ....	4
Käyttöveden lämmitykseen kuluva energia .....	4
Ongelman ratkaisu.....	4
Patteriverkoston lämpöhäviöt ja niiden tarjoama vuotuinen säästömahdollisuus .....	4
Käytännön sovellus ongelman ratkaisusta.....	5
Lämpimän kiertoveden verkoston lämpöhäviöt ja niiden tarjoama vuotuinen säästömahdollisuus .....	6
Loppusanat .....	<b>Virhe. Kirjanmerkkiä ei ole määritetty.</b>

## Projektin tarkoitus

Projekti on harjoitustyö, joka tehdään Taitotalon järjestämän osatutkinnon ”Datan käsittely ja koneoppimisen hyödyntäminen” kurssia.

Harjoitustyössä opiskelija kuvaa datalla ratkaistavan ongelman ja käsittelee olemassa olevaa dataa ja ratkaisee ongelman koneoppimisen avulla.

## Ongelman kuvaus

Vuonna 2022 [Suomen energian loppukäyttö](#) oli n. 294 TWh. Rakennusten lämmityksen osuus tuosta oli n. 27% eli lähes 80 TWh. Asumiseen käytetään vuosittain energiaa n. [60 TWh](#). [Tästä asuintilojen lämmitykseen kuluu n. 43 TWh ja käyttöveden lämmitykseen n. 10 TWh](#).

Asuinrakennuksissa lämmin käyttövesi kiertää putkistossa, jotta hanasta tulevaa vettä ei jouduta juoksuttamaan turhaan. Koska tällä voidaan samalla vähentää käytettävän veden määrää, ei veden kuuman veden kierrätys ole siinä mielessä turhaa. Veden kierrätys putkistossa aiheuttaa kuitenkin lämpöhäviöitä, jotka voivat olla luokkaa 10-20% käyttöveden lämmitykseen käytettävästä energiasta. Sellaisissa rakennuksissa, joissa huoneistojen lämmittäminen tapahtuu vesikiertoisesti, osa lämmittämiseen kuluvasta energiasta kuluu samalla tavalla lämpöhäviöihin.

Lämpöhäviöitä ja niihin kuluvaa energiaa voidaan todennäköisesti pienentää, jos veden kierto pysäytetään silloin kun sille ei ole tarvetta. Niinpä ongelma voitaneen muotoilla seuraavasti.

*Voidaanko koneoppimista hyödyntää niin, että tiedossa olevan datan avulla löydetään sellaiset ajankohdat, jolloin käyttöveden kierto ja lämmitysveden kierto olisi järkevää pysäyttää tai tuottaa tieto, jonka avulla kiertovesipumppuja voidaan ohjata.*

## Miten lähestyä ongelmaa

Ongelma on kaksitahoinen, koska vettä kierrätetään kaksissa putkistoissa: 1) Lämpimän käyttöveden putkissa ja 2) patteriverkostossa. Ensimmäinen selvitettävä asia on tutkia, milloin energia näissä järjestelmissä kuluu. Käyttöveden lämmitykseen tarvittava energian voidaan arvioida olevan kohtalaisen vakio vuoden ympäri, joskin se korreloi vahvasti asunnossa käytetyn veden kulutuksen kanssa. Vedenkulutus vaihtelee vuorokauden ja viikonpäivien mukaan. Jonkin verran vaihtelua esiintyy myös kuukausien välillä. Tätä kysymystä voisi tutkia aikasarjan avulla. Lämmitykseen kuluva energia sen sijaan riippuu ulkolämpötilasta oletettavasti lineaarisesti. Tämä voidaan varmistaa vaikkapa lineaarisen regression avulla.

Esimerkinomaisesti tutkitaan yhdestä taloyhtiöstä kerättyä dataa, joka sisältää vedenkulutuksen ja kaukolämmön kulutuksen ajan funktiona. Data on kerätty tunneittain. Ulkolämpötila saadaan poimittua esim. ilmatieteenlaitokselta.

### [Asuntojen lämmitykseen kuluva energia.](#)

Asuinrakennuksen lämmityksessä käytettävä energia kuluu pattereiden kautta säteilylämpönä ja osa siirrossa tapahtuvana lämpöhäviönä. Jos samalla energialähteellä lämmitetään sekä asuinhuoneistot, että käyttövesi, tarkoittaa tämä sitä, että pitää pystyä eristämään 1) asumiseen käytettävä energia käyttöveden veden lämmittämiseen kuluvasta energiasta 2) lämpöhäviöiden osuus. Käyttöveden lämmitykseen kuluva energia on mahdollista poistaa yhtälöstä niin, että käytetään vain sitä dataa, jolloin vedenkulutus on nolla tai voidaan perustellusti olettaa, että käytetään vain kylmää vettä, koska tällöin energiaa ei kulu käyttöveden lämmitykseen. Selvitetään asuntojen lämmityksen lineaarinen yhtälö esim. lineaarisen regression avulla. ( $y = ax + b$ ). Jos oletetaan, että lämpöhäviöiden osuus on kaikissa olosuhteissa jokseenkin vakio, niin niiden

osuus voidaan poistaa, kun tarkastellaan 1) sellaista dataa, jolloin vedenkulutus on nolla (tai käytetään vain kylmää vettä) ja 2) rakennusten lämmitykseen ei kulu energiaa.

### Lämpimän veden kierrätyksessä tapahtuva lämpöhäviö.

Kun vettä kierrätetään käyttövesiputkistoissa ja patteriverkostossa on seurauksena lämpöhäviö. Tämän suuruus voidaan arvioida (tai laskea) esim. tarkastelemalla minimienergiankulutusta milloin kaksi ehtoa täyttyy: 1) lämmintä vettä ei käytetä ja 2) Asuinrakennusten lämmitykseen ei kulu energiaa. On huomattava, että lämpöhäviöitä tapahtuu sekä patteriverkostossa, että käyttövesiputkistossa. Voitaneen arvioida, että lämpöhäviöt jakaantuvat tasan kahden putkiston välillä.

### Käyttöveden lämmitykseen kuluva energia

Käyttöveden lämmittäminen kuluttaa energiaa kahdella tavalla: 1) Käytetty veden lämmittäminen sekä 2) lämpimän käyttöveden kierrättämisen aiheuttama lämpöhäviö. Lämpöhäviöt saadaan arvioitua kohtalaisen luotettavasti, kuten aiemmin todettiin ja toisaalta kokonaisenergian, lämpöhäviöiden ja asuntojen lämmittämiseen kuluvan energian avulla saadaan laskettua kuinka paljon käyttöveden lämmitykseen kuluu energiaa.

Käyttöveden lämmittämisessä kuluva energia saadaan yksinkertaisella vähennyslaskulla: Kulunut kokonaisenergia - Asuinrakennusten lämmitykseen kulunut energia - lämpöhäviöiden energia.

Energiankulutus on toki mahdollista laskea, mutta, jotta käyttöveden kierrossa tapahtuvia lämpöhäviöitä voitaisiin eliminoida, edellyttäisi se, että lämpimän käyttöveden kierto katkaistaan. Tämä puolestaan edellyttäisi sitä, että pystytään luotettavasti ennustamaan ne tunnit, jolloin taloyhtiössä käytetään vain kylmää vettä tai ei käytetä vettä lainkaan. Ongelmana on, että jos ennuste menee pieleen, joudutaan vettä juoksuttamaan, jolloin vedenkulutus kasvaa.

Jotta lämpimän käyttöveden kierto kannattaa katkaista olisi pystyttävä ennustamaan riittävän luotettavasti ne tunnit, jolloin veden kulutus on nolla tai alle 10 ehkä 20 litraa tunnissa.

### Ongelman ratkaisu

Jaetaan ongelma ratkaisu kahteen osaan, joihin haetaan erilliset ratkaisut. Tämä on järkevää siinäkin mielessä, että verkostot ovat erilliset ja niiden toimilaitteet (i.e. pumput) ovat erilliset.

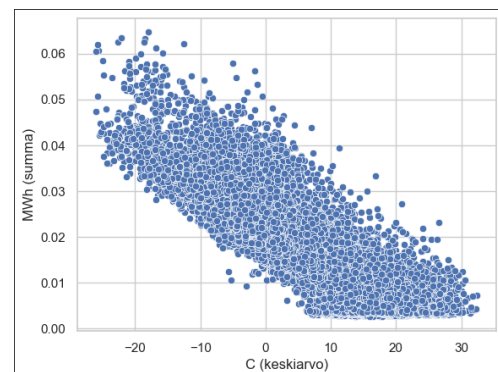
### Patteriverkoston lämpöhäviöt ja niiden tarjoama vuotuinen säästömahdollisuus

Asuntojen lämmityksen vaatima energiankulutus on jokseenkin suoraan verrannollinen ulkolämpötilaan nähden, joten on perusteltua käyttää lineaarista regressiota energiankulutuksen yhtälön selvittämiseksi. Lineaarisuus voidaan todeta helposti esim. hajontakuvion avulla:

On mahdollista, että energiankulutukseen vaikuttaa myös muita ulkoisia muuttujia kuin vain lämpötila. Tällaisia voivat olla esimerkiksi tuuli. Koska ilmatieteenlaitoksen datassa on mukana myös tuulenvoimakkuus, niin lisätään se mukaan laskentaan.

Oikealla oleva hajontakuviio näyttää

kokonaisenergiankulutuksen verrattuna ulkolämpötilaan, jossa mukana ovat vielä sekä lämpöhäviöt, että käyttöveden lämmityksen vaatima energiankulutus. Voidaksemme laskea asuntojen lämmityksen vaatiman energiankulutuksen poistetaan nuo kaksi muuta komponenttia.



Käyttöveden lämmityksen vaatima energia saadaan poistettua helposti poistamalla datasta sellaiset havainnot, jolloin käytetään lämmintä vettä. Tämä onnistuu esim. niin, että poistetaan ne havainnot, jolloin tuntikulutus on suurempi kuin 0.0 m<sup>3</sup>/h.

Edellisen lisäksi poistetaan vielä lämpöhäviöt, joka tapahtuu vähentämällä jäljelle jääneiden havaintojen energiankulutuksesta minimienergiankulutus, joka on n. 0.0025 MWh/h.

Tällöin oikealla oleva hajontakuviokuva onkin huomattavan paljon siistimpi.

Tällä datalla laskettuna lineaarinen regressio antaa tuloksena suunnilleen (eri ajokerroissa esiintyy pientä vaihtelua) yhtälön:

$$y = -0.001009x_1 + 0.000596x_2 + 0.014410,$$

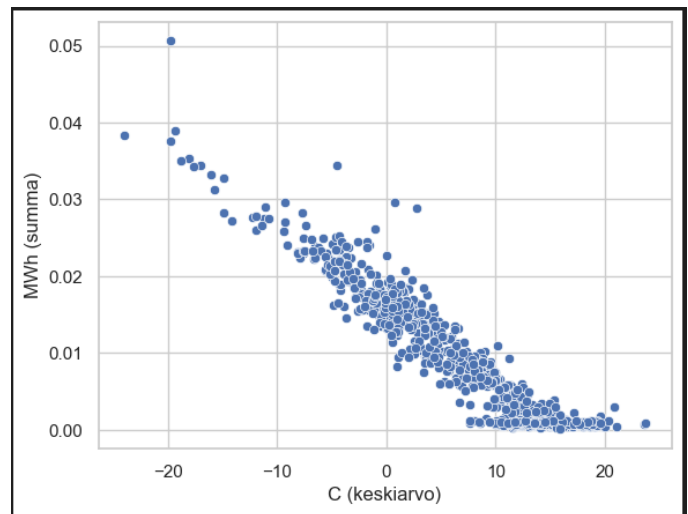
jossa y on energiankulutus, x<sub>1</sub> ulkolämpötila, x<sub>2</sub> tuulen nopeus ja vakiotermi saa arvon 0.014729. Rajalämpötila, jossa asuinrakennusten lämmitykseen käytettävä energiankulutus putoaa nolnaan, saadaan laskettua tuosta yhtälöstä helposti merkitsemällä tuulen nopeus ja energiankulutus nolliksi. Tämä rajalämpötila on n. 14.3 °C. Rajalämpötila oletettavasti vaihtelee eri rakennuksissa, koska rakennusten eristyksissä ja rakennustavoissa voi olla eroja. Tulkinta tästä rajalämpötilasta on siis se, että kun ulkolämpötila ylittää rajalämpötilan, ei asuntojen lämmitykseen pitäisi kulua enää energiaa. Todellisuudessa joitain poikkeamia tuohon on, joka on nähtävissä myös hajontakuviossa, mutta kyse lienee niissä tapauksissa lämpöhäviöiden vaihtelusta, koska energiamäärä niissä on hyvin vähäinen.

Tämän ratkaisun [Python-koodi löytyy GitHub:sta](#).

### Käytännön sovellus ongelman ratkaisusta

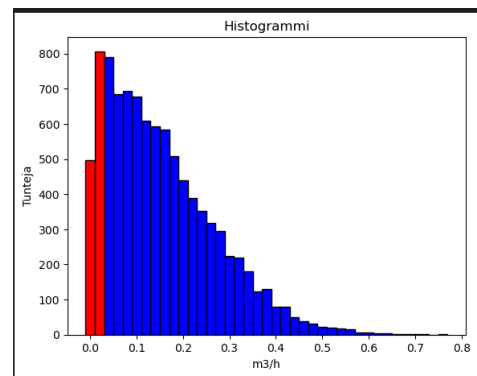
Koska löydettiin selkeä rajalämpötila, jonka yläpuolella rakennusten lämmittäminen ei enää vaadi lämpimän veden kierrätystä patteriverkostossa, niin potentiaalinen säästö vuotuisesta lämpöhäviöiden vähenemisestä saadaan arvioitua kohtuullisella tarkkuudella. Kyseisen esimerkin taloyhtiössä päädyttiin arvioon, että vuotuinen säästöpotentiaali on n. 240 EUR/v. Toimilaitteita, jolla veden kiertoa voidaan ohjata sääennusteeseen perustuen, näyttää löytyvän jo markkinoilta. Esimerkki laitteesta: [Weather compensation | Danfoss](#)

Laitteen myyntihinta näyttää olevan alimmillaan n. 450 EUR, joten ilman asennuskustannuksia laitteen takaisimaksuaika on hieman alle 2 vuotta.



## Lämpimän kiertoveden verkoston lämpöhäviöt ja niiden tarjoama vuotuinen säästömahdollisuus

Jotta lämpimän kiertoveden kierto voidaan katkaista, on pystyttävä ennustamaan riittävällä tarkkuudella (esim. 95% varmuudella) ne tunnit, jolloin kiertoa ei tarvita. Tämä on hankalaa erityisesti tilanteessa, jossa vedenkulutusta taloyhtiössä mitataan vain yhdellä mittarilla. Asiaa kuitenkin auttaa, jos pyritään ennustamaan ainoastaan sellaisia hetkiä, jolloin tunnin veden kulutus on nolla tai hyvin alhainen, Esim 10-20 litraa tunnissa. WC-altaan huuhtelu kuluttaa vettä n. 10 litraa, joten yöaikaan tuollainen kulutus lienee pelkän kylmän veden käyttöä. Kuva vasemmalla



Ennuste tehtiin käyttäen FB Prophet algoritmia ja sillä päästiin ennustetarkkuudessa tasolle 53%, joka ei ole lähellekään riittävä.

Koska tulos on itse asiassa vain hieman napanheittoa parempi, niin katsoin aiheelliseksi ajaa datan myös logistisen regression algoritmilla. Sillä näytti olevan mahdollista päästä ennustetarkkuteen 59.9%, joka on jo jonkinlainen parannus. Parempaan lopputulokseen päästäkseni katsoin aiheelliseksi testata vielä kolmatta algoritmia. Tähän valitsin Gradient Boosting algoritmin siksi, että se on useissa Kaggle kilpailuissa osoittautunut parhaaksi classifier algoritmiksi. Sille sain koulutusdatan tarkkuudeksi 67%. Samalla kuitenkin testidatan tarkkuus jää arvoon 61.3%. Näiden kahden eron suuruus kertoo siitä, että malli alkaa olla ylisovitettu. Ilman ylisovitusta päästään koulutusdatan tarkkuudessa tasolle 63.5%.

Koska mallien tarkkuus jää useammalla algoritmilla huomattavasti ali tavoitelukeman, niin on mietittävä muita keinoja, jolla tarkkuutta voisi parantaa. Feature engineering voisi olla yksi vaihtoehtoinen tapa. Ilman lisämuuttujia, joko orgaanisesti saatuja tai Feature engineeringillä luotuja ei haluttuun tulokseen oikein taida olla mahdollisuuksia päästä.

Ratkaisua tähän ongelmaan ei siis löytynyt millään algoritmilla, mutta testattujen algoritmien [Python-koodit löytyvät GitHubista](#).

## Mitä projekti opetti

Projekti oli opettavainen monessa mielessä. Sen kestäessä oli mahdollisuus hakea käytännön tuntumaa useisiin eri algoritmeihin, joka käsittääkseni vastaa normaalia todellisuutta siksi, että etsittäessä (riittävän) hyvää algoritmia, jolla asetettu tavoite halutaan saavuttaa, ei siihen päästä välttämättä ensimmäiseksi valitulla algoritmilla. Algoritmista toiseen vaihtaminen ei loppujen lopuksi ole kovin hankalaa tai työlästä. Esimerkkejä eri algoritmien käytöstä on löydettävissä esim. Kaggle:sta. Myös Sklearn tarjoaa hyödyllisiä esimerkkejä.

Toinen asia, jota pidin projektissa hyödyllisenä, oli hyperparametrien käytön käytännön opettelu. Niillä pystyy vielä hienosäätämään algoritmin toimintaa, jotta tarkkuutta saadaan vielä hieman nostettua

Kolmas asia, joka jäi ehkä kurssin aikana vähemmälle, oli se, miten algoritmin yli- tai alisovitus selvitetään. Tämäkin oli vastaisuutta silmällä pitäen erittäin hyödyllinen asia.

Neljäs asia, jonka tämä projekti opetti, oli se että datan käsittelyyn, ennen kuin sen saa vietyä algoritmin koulutukseen, menee yllättävän paljon aikaa.

Viides opittu asia oli se, että joskus olemassa olevat mittarit eivät kuvaa oikein haluttua lopputulosta, jolloin kannattaa miettiä parempi mittari, jolla algoritmin toimivuutta mitataan.

Projektissa oli lisäksi tarkoitus testata ajaa data algoritmien läpi Azure ympäristössä. Tähän ei kuitenkaan aika aivan riittänyt koska datan käsittely sekä algoritmien opiskeluun meni hieman enemmän aikaa kuin alun perin arvioin. Tämä on ehkä sellainen asia, johon palaa vielä omalla ajalla.