



Teuvo Isomaa

# Tekoälymallin kysyntäjoustoennusteen validointi

Metropolia Ammattikorkeakoulu

Insinööri (AMK)

Sähkö- ja automaatiotekniikka

Insinöörityö

3.5.2023

# Tiivistelmä

Tekijä: Teuvo Isomaa  
Otsikko: Tekoälymallin kysyntäjoustoennusteen validointi  
Sivumäärä: 27 sivua  
Aika: 3.5.2023

Tutkinto: Insinööri (AMK)  
Tutkinto-ohjelma: Sähkö- ja automaatiotekniikka  
Ammatillinen pääaine: Automaatiotekniikka  
Ohjaajat: Jouni Häkkinen, Ryhmäpäällikkö  
Erkki Räsänen, Lehtori

---

Insinöörityön tavoitteena oli selvittää Siemensin talotekniikan ja rakennusautomaation osastolle, voiko manuaalista sähkönsäätäjöiden ennustamista korvata tekoälymallin ennusteella tekemällä numeerista vertailua ennusteen ja toteutuneen tehontarpeen välillä.

Selvitystyössä luotiin analysointityökalu sekä analysoitiin tekoälymallin ennustetarkkuutta ja tarkasteltiin tekoälymallin skaalautuvuutta opettamalla malli toiselle käyttökohteelle.

Tutkimus luo perustan tuleville kysyntäjouston kehitysprojekteille ja tarjoaa Siemens Osakeyhtiölle tietoa datan hyödyntämisen mahdollisuuksista. Luotu tekoälymalli voi mahdollistaa kysyntäjouston laajemman käyttöönoton ja tarjota lisätuloja sekä Siemensille että käyttökohteen omistavalle asiakkaalle.

Avainsanat: Koneoppiminen, tekoäly, data-analyysi, tehon kysyntäjousto

## Abstract

Author: Teuvo Isomaa  
Title: Validation of Power Demand Response Forecasting of an AI Model  
Number of Pages: 27 pages  
Date: 3 May 2023

Degree: Bachelor of Engineering  
Degree Programme: Electrical and Automation Engineering  
Professional Major: Automation Engineering  
Supervisors: Jouni Häkkinen, Group Manager  
Erkki Räsänen, Senior Lecturer

---

The bachelor's thesis work was implemented to determine whether manual power demand response forecasting can be replaced with an artificial intelligence model by performing a numerical comparison between the forecast and the actual power demand response for Siemens building technology and automation department.

The thesis analyzes the forecasting accuracy of the AI model and examines the scalability of the AI model by training the model for another use case.

The study lays the foundation for future power demand response forecasting development projects and provides Siemens Corporation with information on the potential of data utilization. The developed AI model can enable a broader adoption of power demand response forecasting and offer additional income for both Siemens and the customer owning the use case site.

Keywords: Machine learning, AI, Data-analysis, Power Demand Response Forecasting

# Sisällys

## Lyhenteet

1	Johdanto	1
2	Koneoppimismallien aikaisemmat käyttökohteet	2
2.1	Koneoppimisen menetelmät	2
2.2	Koneoppimisen kehitys viime aikoina	2
2.3	Koneoppimisen yleisimmät käyttökohteet	3
2.4	Koneoppimisen käyttökohteet rakennusautomaatiossa	4
3	Reservimarkkinat	4
4	Menetelmät ja työkalut	6
4.1	Tekoäly	6
4.1.1	Tekoälymallin kerrokset	6
4.1.2	Tekoälyalusta ja -malli	8
4.2	Data	9
4.2.1	Syötedata ja muuttujat	9
4.2.2	Datan käsittelyohjelmisto ja ohjelmointikirjastot	10
4.3	Analysointityökalu	11
4.4	Analysointimenetelmät	13
4.4.1	Keskineliövirheen neliöjuuri	13
4.4.2	Prosentuaalinen ennustevirhe	13
5	Validointiprosessi	14
5.1	Koneoppimismallin suorituskyvyn arviointi	14
5.1.1	Datan käyttäytyminen	14
5.1.2	Numeeriset vertailut	17
5.2	Koneoppimismallin validointi uuteen käyttökohteeseen	19
5.2.1	Datan käyttäytyminen	20
5.2.2	Numeeriset vertailut	20
6	Johtopäätökset	22
6.1	Datan laatu	22
6.2	Tulokset	23

6.3 Parannusehdotukset	24
7 Yhteenveto	25
Lähteet	1

## Lyhenteet

ChatGPT-4 Generative pre-trained transformer. Multimodaalinen kielimalli.

JupyterLab Interaktiivinen kehitysympäristö koodille ja datalle.

LVI Lämmitys, vesi ja ilmastointi.

NumPy Numerical python. Python-ohjelmistokirjasto taulukoille ja matriiseille.

Pandas Pythonin ohjelmistokirjasto data-analysointiin.

Plotly Pythonin ohjelmistokirjasto graafien piirtoon ja tulkintaan.

ReLU Rectified Linear Unit. Neuroverkoissa käytetty aktivointifunktio.

SPARCS Sustainable energy positive and zero carbon communities. Kansainvälinen hanke, jonka tavoitteena on luoda kestävien energiatehokkaiden ja hiilivapaiden yhteisöjen verkosto.

## 1 Johdanto

Opinnäytetyön tarkoituksena on selvittää Siemensin talotekniikan ja rakennusautomaation osastolle, voidaanko käsin tehdyn tehon kysyntäjouaston ennustamista korvata tekoälymallin ennusteella. Työssä tarkastellaan tekoälymallin ennusteen suorituskykyä ja verrataan sitä ihmisen tekemään ennusteeseen. Lisäksi selvitetään tekoälymallin skaalautuvuutta opettamalla malli uudelle käyttökohteelle.

Selvitystyö on osana Siemensin SPARCS-projektia, ja se toimii siltana kysyntäjouaston tuleville kehitysprojekteille. Selvitystyön tarkoituksena on testata tekoälymallin opetukseen tarvittavia työkaluja, datansiirtoja ja luoda pohjaa mallin opetusprosessille sekä niiden käyttöönotolle.

Selvitystyö antaa Siemens Osakeyhtiölle tietoa ja suuntaa rakennusautomaation datan hyötykäytön mahdollisuuksista. Tavoitteena on tehdä dokumentaatio uuden käyttökohteen validoinnista, joka antaa pohjan uusien mallien kehitykseen ja käyttöönottoon erilaisille rakennusautomaatiokohteille. Tekoälyn ennuste mahdollistaisi kysyntäjouaston käyttöönoton huomattavasti nykyistä laajemmin. Tämänhetkinen ennustaminen tehdään ihmisen toimesta ja on osittain hidasta sekä työlästä. Prosessin automatisointi vapauttaisi ihmisresursseja.

Kysyntäjouaston kehittäminen hyödyttää kantaverkkoa omistavaa Fingrid julkista Oyj:tä, paikallisia sähkönsiirtoyhtiöitä, Siemensiä sekä käyttökohteen omistavaa asiakasta. Tekoälymalli voi parhaimmillaan tarjota huomattavia lisätuloja sekä Siemensille että käyttökohteen omistavalle asiakkaalle. Tämän lisäksi kysyntäjouaston laajempi käyttöönotto tuo enemmän joustoja sähkönsiirtomarkkinoille, mikä auttaa sähköverkon taajuuden vaihteluihin uusiutuvien energiatuotantojen lisääntyessä.

## 2 Koneoppimismallien aikaisemmat käyttökohteet

### 2.1 Koneoppimisen menetelmät

Koneoppiminen on tekoälyn osa-alue, jossa toimintaa ei ole ohjelmoitu valmiiksi. Koneoppimisessa malli opetetaan käyttämään algoritmeja sekä tilastollisia toimintamalleja opetusdatan avulla. Opetusdatalla malli opetetaan tekemään erilaisia painotuksia tai riippuvuuksia lähdetiedon ja haluttujen tulosten mukaan. Malli osaa niiden avulla toimia tilanteissa, joihin sitä ei ole nimenomaisesti ohjelmoitu. Koneoppimisalgoritmit voidaan jakaa ohjattuun, ohjaamattomaan ja vahvistavaan oppimiseen. Jokaisella on omat algoritmit ja tekniikat.

Ohjattu oppiminen on koneoppimisen oppimismalli, jossa tiedetään haluttu lopputulos. Oppimismalli opetetaan opetusdatalla ennustamaan haluttu tietämätön lopputulos. Lopputuloksessa pyritään minimoimaan ennustetun ja tapahtuneen tuloksen välinen ero.

Ohjaamattomassa oppimisessä muodostetaan malli, joka pyrkii löytämään datasta samankaltaisuuksia tai ryppäitä ilman ennalta määriteltyä luokkaa. Ohjaamattoman oppimisen yleisiä sovellutuksia ovat klusterointi, dimensioiden vähentäminen ja poikkeamien tunnistaminen. [2.]

Vahvistusoppiminen puolestaan keskittyy tekoälymallien kouluttamiseen siten, että ne tekevät päätöksiä interaktiivisessa ympäristössä saaden palautetta palkkioiden tai rangaistusten kautta. Mallin tavoitteena on löytää haluttu lopputulos, jolla saavutetaan eniten palkkioita. Vahvistusoppimisen sovelluksia ovat muun muassa pelien pelaaminen, robotiikka ja autonomiset ajoneuvot. [3.]

### 2.2 Koneoppimisen kehitys viime aikoina

Tekoälystä on odotettu tulevan huomattava osa ihmisten elämää jo vuosikaudet. Se ei kuitenkaan ole vielä tehnyt varsinaista läpimurtoaan, mutta sen odotetaan muuttuvan lähivuosien aikana varsinkin Googlen ja Microsoftin tekoälyllä varustettujen virtuaalisten asiakaspalvelijoiden myötä. Erityisesti OpenAI:n



kehittämä ChatGPT-4 on jo nyt herättänyt suurta kohua viime kuukausina ja sen ansiosta tekoälyn on ennustettu mullistavan ihmisten elämää erityisesti lähivuosina. Tekoälyä on yleisesti rinnastettu olevan yhtä merkittävää kuin tulen ja sähkön tuleminen osaksi ihmisen elämää. Toisaalta on myös ennustettu, että se tulee aikanaan rikkomaan koko internetin. [4.]

GPT-4 on lyhenne sanoista ”Generative pre-trained transformer”, joka viittaa syväoppimistekniikkaan ja 4 viittaa tekoälymallin versioon. Syväoppimistekniikka käyttää keinotekoisia hermoverkkoja kirjoittaakseen kuten ihminen. ChatGPT-4 on multimodaalinen, ja se voi siis käsitellä videokuvaa, kuvia, ääntä ja tekstiä. Sen ennustetaan mullistavan varsinkin työelämää esimerkiksi Microsoftin Office 365-liitännäisellä, joka pystyy tekemään data-analyysiä, muistiinpanoja kokouksista ja sähköpostihistoriasta kertomalla tietyille käyttäjälle räätälöidyt avainkohdat viimeaikaisista tapahtumista tai kokouksessa käsitellyistä asioista. Tekoäly on kehittynyt varsinkin matemaattisissa kyvyissä, kuvan tunnistuksessa ja asiayhteyksien tunnistuksessa. Lisää on ilmeisesti vielä tulossa, sillä ChatGPT-4 ei ole tekoälymallin viimeinen versio. ChatGPT-4 on edistynein tekoälymalli ja sen ennustetaan tekevän suuria harppauksia myös muilla aloilla kuten lääketieteessä. [5; 6; 7.]

### 2.3 Koneoppimisen yleisimmät käyttökohteet

Koneoppimista on käytetty hyvin monipuolisesti eri aloilla. Tähän mennessä sitä on tyypillisesti käytetty esimerkiksi terveydenhuollossa, logistiikassa, virtuaalisissa asiakaspalvelijoissa sekä sähköisessä kaupankäynnissä. [8.]

Lääketieteessä käytetään erityisesti kuvantunnistusta. Koneoppimisella pyritään tunnistamaan diagnostiikka kuvista erilaisia vikoja, kasvaimia tai vaurioita. Sähköisessä kaupankäynnissä koneoppimista voidaan käyttää luottokorttipetosten havaitsemiseen. [9; 10.]

Logistiikassa koneoppimista käytetään tunnistamaan muodostuvia liikenneuhkia sekä itseohjautuvia autoja. Virtuaalisia asiakaspalvelijoita ovat esimerkiksi

Alexa, Google Maps ja asiakaspalvelu botit. Virtuaaliset asiakaspalvelijat perustuvat yleensä puheen tai tekstin tunnistukseen.

Suomessa koneoppimista on ruvettu hyödyntämään talousrikoksien havaitsemiseen kuten erilaisten kartellien löytämiseen. [11.]

## 2.4 Koneoppimisen käyttökohteet rakennusautomaatiossa

Koneoppimista käytetään yleisesti rakennusautomaatiossa energiatehokkuuden ja rakennuksen suorituskyvyn parantamiseen. Varsinkin energiatehokkuuden parantaminen ja erilaiset sähköhallintapalvelut ovat lisääntyneet viime vuosina energiakriisin vauhdittamana. Suomessa erityisesti vihreän energian, vety- ja akkuteollisuuden lisääntyessä energiatehokkuuden ja energian käytön optimoinnin ennakoitaan lisääntyvän huomattavasti seuraavien vuosien aikana ja sen myötä koneoppimisen merkitys tulee korostumaan.

Koneoppimista hyödynnetään ennakoivaan huoltoon esimerkiksi rakennustietojen analysointiin ja huollon tai korjausten tarpeen ennustamiseen. Koneoppimista voidaan käyttää myös vikojen havaitsemiseen kuten rakennusjärjestelmien, kuten LVI- ja valaistusjärjestelmien, vikojen havaitsemiseen ja diagno-sointiin. Koneoppimista hyödynnetään LVI-järjestelmien, valaistuksen ja muiden energiaa kuluttavien järjestelmien ohjaamiseen ja energiahävikin vähentämiseksi. Koneoppimista voidaan hyödyntää myös energian kysynnän arvioimiseen. Energiankulutuksen optimointiin huippukysynnän aikoina energiakustannusten vähentämiseksi ja lisäsähkötuotannon tarpeen minimoimiseksi. Läsäolotunnistuksen avulla voidaan määrittää huoneen käyttöaste ja säätää lämpötilaa, valaistusta ja ilmanvaihtoa sen mukaan. [12.]

## 3 Reservimarkkinat

Sähköä tulee käyttää niin paljon, kuin sitä tuotetaan. Mikäli sähköä tuotetaan enemmän kuin sitä käytetään, voi sähköverkon taajuus nousta liian korkealle. Jos sähkön kulutus taas on tuotantoa suurempaa, voi sähköverkon taajuus

laskea liian alas. Taajuus pyritään pitämään 49,9-50,1 hertsin välissä. Liian alhainen taajuus johtaa valojen vilkkumiseen ja lopulta sähköverkon lamaan tuloon. Liian korkea taajuus voi taas johtaa laiterikkoihin. Sähköverkon taajuuden vaihtelut ovat lisääntyneet varsinkin 2000-luvun alussa vesivoiman ja tuulienergian tuotannon kasvaessa. [13.]

Fingrid Oy ylläpitää ja kehittää sähkön reservimarkkinoita, jolla pyritään vastaamaan sähköverkon taajuuden vaihtelun haasteisiin. Fingrid ostaa sähkön tuotannon ja kulutuksen joustoa reservimarkkinatoimijoilta, jotka suunnittelevat oman sähkönkulutuksensa etukäteen. Mahdollisesta joustomahdollisuudesta tehdään tarjous Fingridille, joka päättää myöhemmin hyväksyykö tarjouksen. Mikäli tarjous hyväksytään, on reservimarkkinatoimija velvollinen joustamaan sopimuksen mukaan. [14.]

Sähköverkon taajuuden vaihteluun pyritään vastaamaan tuomalla joustavuutta käyttökohteisiin. Rakennusautomaation osalta se voi tarkoittaa rakennuskohdasta joustavuutta. Rakennuksien datankeruun ja älyn lisääntyessä, voidaan sähköntuotannon ylijäämää tai alijäämää myydä Fingridille. Tämä luo uusia mahdollisuuksia kaupungeille ja laitoksille, jotka pystyvät saamaan joustavuudesta huomattavaa kompensatiota.

Sähköpaneeli- ja akustoratkaisuista tulee näin myös huomattavasti houkuttavampia vaihtoehtoja. Kauppakeskuksissa tai tuotantolaitoksissa se tarkoittaa halvempaa sähköntuotantoa aurinkoenergiasta tai akustosta silloin, kun sähköverkon energia on kallista. Tarvittaessa sitä pystytään myös myymään, mikäli sähköä ei tarvita.

Datan avulla pystytään ennustamaan laitoskohtaista joustoa. Siemensillä ennustusta on tehty pääosin ihmisen toimesta. Ihmisen tekemä ennustus on työllästä ja sitoo ihmisresursseja. Siemens tekee kysyntäjoustotarjouksen Fingridille joka päivä klo 17:00 mennessä. Tätä varten tehonjoustoennuste pyritään tekemään samana päivänä puolen päivän aikoihin. Siemensin kysyntäjoustoennuste tehdään aina seuraavalle 36 tunnille, josta viimeiset 24 tuntia ovat

merkitseviä ennustuksia. Ensimmäiset 12 tuntia eivät ole olennaisia tarjouksen kannalta, sillä tarjoukset alkavat vasta puolen yön jälkeen. Tässä tutkimuksessa päivittäiset ennustukset on kerätty yhteen yhdeksi suureksi datasetiksi ja niitä tarkastellaan suurempana kokonaisuutena.

## 4 Menetelmät ja työkalut

### 4.1 Tekoäly

#### 4.1.1 Tekoälymallin kerrokset

Monikerroksisia hermoverkkoja käytetään yleisesti tekoälymalleina. Niitä kutsutaan myös monikerroksisiksi perceptroni-verkoiksi. Keinotekoinen hermoverkko koostuu hermosoluista, jotka muodostavat erilaisia kerroksia. Jokaisella neuronilla on tulo, joka sisältää ja tulostaa tietoja seuraavan neuronin tuloa varten.

Kuvan 1 mukaan hermoverkko koostuu syöttökerroksesta, yhdestä tai useammasta piilokerroksesta ja tulostekerroksesta. Syöttökerros vastaanottaa tietoja suoraan lähteestä, kuten kuvia tai säätietoja. Piilotettu kerros saa syöttönsä syöttökerrokselta sekä muuntaa ja yhdistää ärsykkeet. Tämä muunnos voi olla painotettu summa, ja tulos on aktivointifunktion tuottama tulos.

Aktivointifunktioita käytetään lisäämään hermoverkkojen kykyä oppia ja mallintaa monimutkaisia toimintoja. Ne lisäävät epälineaarisuutta malliin, joka mahdollistaa sen sopeutumisen paremmin erilaisiin tietojen muotoihin. Aktivointifunktioita ovat esimerkiksi hyperbolinen tangentti, ReLu-funktio ja sigmoid-funktio.

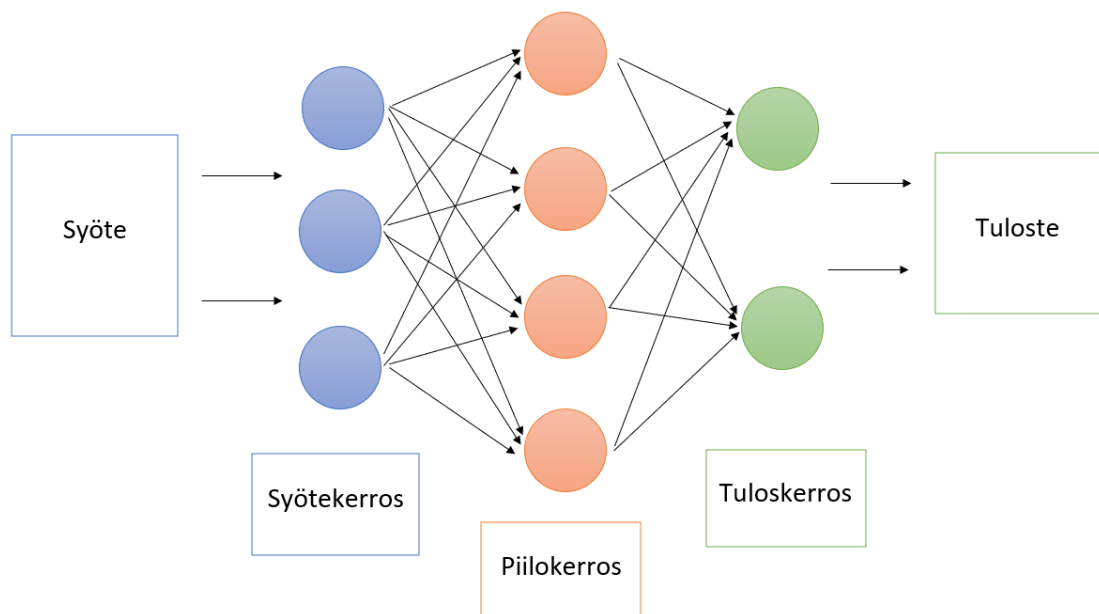
[15.]

Takaisinkytkentäalgoritmeja, kuten vastavirta-algoritmia, käytetään hermoverkkojen harjoittelussa. Vastavirta-algoritmi välittää virhesignaalin takaisin verkon läpi ja päivittää neuronien painoarvoja virheen minimointiin tähtäävän gradienttimenetelmän avulla.

Kerrosten ja neuronien lisääminen voi parantaa mallin ennustevoimaa, mutta liian monta kerrosta voi johtaa liian monimutkaiseen malliin, kun taas liian harvat kerrokset voivat yksinkertaistaa ennustetta liikaa. Hermoverkkojen yleisiä haasteita ovat ylisovitus ja alisovitus. Ylisovitus tapahtuu, kun malli oppii liian hyvin koulutusdatan erityispiirteet, mutta ei pysty yleistämään uusiin tietoihin. Alisovitus taas tarkoittaa sitä, että malli ei ole tarpeeksi monimutkainen kuvataksien koulutusdatan piirteitä. [16.]

Tekoälymallin tulosten hienosäätämiseksi hyperparametrien viritys on yleinen lähestymistapa. Hyperparametrit asetetaan ennen harjoittelua, ja ne voivat sisältää oppimismopeuden, piilotettujen kerrosten sekä kunkin piilokerroksen hermosolujen määrän. Näiden hyperparametrien säätäminen voi parantaa mallin suorituskykyä.

Muita lähestymistapoja tekoälymallin hienosäätämiseen ovat erilaisten harjoitusalgoritmien käyttö, säännöstelytekniikat ylisovituksen estämiseksi ja erilaiset aktivointitoiminnot piilossa olevissa kerroksissa. Lisäksi monipuolisemman tietojoukon käyttö voi myös parantaa mallin suorituskykyä. [17.]



Kuva 1. Monikerroksinen neuroniverkko.

#### 4.1.2 Tekoälyalusta ja -malli

Työ aloitettiin ottamalla Siemensin pääkonttorin tekemä tekoälyalusta käyttöön, mihin pystytään liittämään erilaisia tekoälymalleja. Alusta toimii pilvessä ja tekoälymalleja ajetaan alustalla omassa kontti ympäristössä. Tekoälyalustaan on tarkoitus istuttaa tekoälymalleja, jotka lukevat ja tallentavat dataa pilvessä.

Seuraavaksi tutkittiin tekoälymallia, jonka Siemensin pääkonttori oli toimittanut. Malli on luotu käyttämällä halutun käyttökohteen teho- sekä paikallista sääennustusdataa. Tekoälymallia varten asennettiin paikallinen virtuaaliympäristö, jossa malli koulutettiin käyttökohteen datan avulla. Työkalu sisälsi datan esikäsittelytyökaluja, mutta niihin ei perehdytty työssä kovin syvästi. Datan esikäsittelyasetukset oli valmiiksi asetettu ja mallin luominen suoritettiin esivalmistetuilla asetuksilla. Esikäsittely sisälsi ainakin datan suodatusta ja normalisointityökaluja.

Tämän jälkeen etsittiin toinen samantapainen käyttökohde, johon tuloksia verrattiin. Mallin opetus ajettiin samoilla esikäsittelyasetuksilla ja tuloksia verrattiin keskenään. Näin pyrittiin selvittämään tekoälymallin skaalautuvuus erilaisille käyttökohteille. Tavoitteena oli, että mallin opetus onnistuu sujuvasti, ja mallia voidaan käyttää useille käyttökohteille, joille löytyy samankaltaista tehodataa. Näin opetusprosessi kysyntäjouston ennustamiseen tarkoitetulle tekoälymallille saataisiin sujuvaksi ja käyttökohteita saataisiin sujuvasti lisää.

Tekoälymallin opetusprosessissa käytettiin yksinkertaista monikerroksista neuroniverkkoa. Koska kyseessä on yksinkertainen monikerroksinen neuroniverkko, on se myös haavoittuva syötekerroksen datakatkoille. Luomisvaiheessa todettiin, että on järkevämpää lähestyä aihetta yksinkertaisella mallilla, jotta validointi saadaan tehtyä sujuvasti ja mahdollisia hienosäätöjä päästään tekemään helposti projektin edetessä. Datakatkoksien haavoittuvuuksien korjaamiseen todettiin olevan muita tapoja, kuten jälkikäteen täyttämällä datakatkoksia olettamalla ennustetulos vanhan datan perusteella tai muuttamalla neuroniverkkomallia. Neuroniverkkomallin muuttaminen olisi vaatinut suurempaa perehtymistä

mallin ominaisuuksiin ja olisi vaatinut syvällisempää perehtymistä tekoälyyn. Tässä työssä perehdyttiin alkuperäisen mallin validointiin. Mallin ennustus keskittyi pelkästään rakennusautomaation osuuteen tehonjoustossa.

Lähtökohtana oli, että ennustuksen pitää olla tarpeeksi hyvä verrattuna ihmisen tekemään ennustukseen, jotta kehitystyötä pystytään jatkamaan. Epätarkempi tulos voidaan hyväksyä, mikäli prosessin automatisoinnin todetaan säästävän työntekijöiden työaika.

## 4.2 Data

### 4.2.1 Syötedata ja muuttujat

Mallin opetukseen käytettyä dataa on kerätty useammalta vuodelta. Osa datan aikaleimauksista on erittäin epäsäännöllisiä. Sama toistuu opetusdatassa sekä toteutuneessa tehonjouston datassa. Poikkeuksena on sähkömittarin mittaama teho- sekä säädata. Osa aikaleimoista on myös eri standardia ja ilmoitettu eri aikavyöhykkeen mukaan.

Ensimmäisenä muutettiin aikaleimat samaan aikastandardiin, jotta luvut saatiin piirrettyä graafille. Aikavyöhykkeet tasattiin ja otettiin huomioon kesä- ja talviajan muutokset mittausajalta.

Tekoälymallin syötedata koostuu neljästä eri muuttujasta, jotka ovat power adjust, Max active capacity negative, max active capacity positive sekä active power. Muuttujat perustuvat seuraavaan laskukaavaan:

$$\text{Reference power} + \text{Power adjust} \times \text{Max active capacity} = \text{Active power} \quad (1)$$

Reference power on kiinteistön teho, johon ei sisälly kysyntäjoustoa. Power adjust kertoo rakennusautomaation joustopyynnöstä. Power adjust saa lukuja -100 ja 100 prosentin väliltä. Max active capacity kertoo, kuinka paljon

tehonkulutuksesta pystytään joustamaan. Tämä luku on aina nolaa suurempi luku ja sitä käsitellään kilowateissa. Jousto tapahtuu sähköverkon suunnasta, ja se sisältää kaksi komponenttia. Positiivisen ja negatiivisen komponentin. Positiivinen tehonjousto tarkoittaa kantaverkon taajuuden nostoa, jolloin halutaan pienentää rakennuksen käyttämää tehoa. Negatiivinen tehonjousto puolestaan tarkoittaa kantaverkon taajuuden laskemista eli rakennuksen käyttämää tehoa halutaan nostaa.

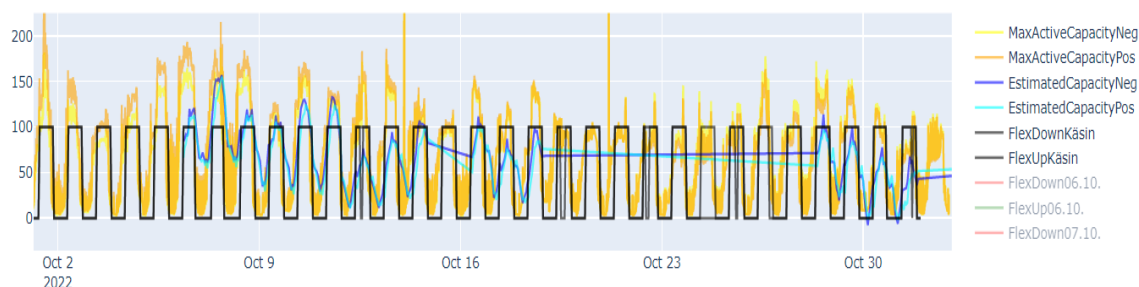
#### 4.2.2 Datan käsittelyohjelmisto ja ohjelmointikirjastot

Datan käsittelyyn käytettiin Jupyter Lab -ohjelmistoa, joka tarjoaa interaktiivisen ympäristön eri ohjelmointikielten, kuten esimerkiksi Pythonin käyttöön. Jupyter Lab mahdollistaa tehokkaan datan tutkimisen, analysoinnin ja visualisoinnin erilaisten kirjastojen ja työkalujen avulla. Ohjelmointikielenä käytettiin python 3.10.8-versiota.

Analysointityökalua luodessa käytettiin avuksi Python-kirjastoja. Näitä oli mm. Pandas, Numpy sekä Plotly. Pandas on tehokas data-analyysi- ja käsittelykirjasto, joka tarjoaa joustavia ja tehokkaita datarakenteita, kuten datakehysä. NumPy puolestaan on kirjasto numeeriseen laskentaan, ja se tukee suurten moniulotteisten taulukoiden ja matriisien käsittelyä. Yhdessä Pandas ja NumPy helpottavat erilaisia datan esikäsittelytehtäviä, kuten datan puhdistamista, muuntamista ja yhdistämistä.

Käsiteltävän datan visualisoimiseksi käytettiin Plotly-kirjastoa. Plotly on monipuolinen ja helppokäyttöinen kirjasto interaktiivisten ja visuaalisten graafien sekä kaavioiden luomiseen. Plotlyn avulla on mahdollista tuottaa havainnollistavia visualisointeja, jotka auttavat ymmärtämään kuvioita ja datan käyttäytymistä.





Kuva 2. Lokakuun ennuste graafilla.

Kuvassa 2 näkyy projektin alkupuolella tehtyjä graafeja. Graafissa näkyy mustalla värillä käsin tehtyjä ennusteita positiiviselle ja negatiiviselle tehonjoustolle. Tässä graafissa keltainen kuvaa toteutunutta negatiivista tehonjoustoja ja oranssi väri positiivista tehonjoustoja. Siniset värit ovat tehonjouston ennusteita, kun taas punainen ja vihreä ovat yksittäisten päivien ennusteita, jotka on piilotettu graafista.

### 4.3 Analysointityökalu

Validointiprosessin tarkoituksena on tarkastella, onko tekoälymallin ennustus riittävän tarkka ja tehdä ennusteen sekä toteutuneen tehon välillä numeerisia vertailuja. Koneoppimismallin validointiin kuuluu sen suorituskyvyn arvioiminen numeerisesti ja graafisesti. Tekoälymallin validointidataa ei ole käytetty mallin opetukseen.

Validointia varten tehtiin analysointityökalu, jolla pystyttiin tekemään varsinainen data-analyysi. Työkalu sisältää noin 1000 riviä Python-koodia, ja se pyrittiin pitämään selkeänä, jotta sitä voidaan käyttää myös toisiin käyttökohteisiin mahdollisimman helposti ja sujuvasti. Työkalu toteutettiin niin, että työkalulle voidaan määrittää syöttö- ja tulostedata tiedostopolkuina. Syöttödata sisältää esimerkiksi tehonarvoja tietyllä aikaleimalla. Tulostedata sisältää esimerkiksi keskineliövirheen neliöjuuren, datakatkoksien aikaleimoja ja datahäviöitä, joita muodostuu, kun datasettejä verrataan toisiinsa. Datahäviöt muodostuvat, kun esimerkiksi todellista tehonjoustoja vertailtaessa tietyltä tunnilta ei löydy vastaparia ennusteen puolelta, ja toisen datakehiksen arvo on tyhjä. Tyhjät arvot poistetaan

datakehysistä, jotta numeerista vertailua tehtävässä laskutoimituksessa ei jaeta tyhjällä arvolla. Tämä johtaisi ohjelman kaatumiseen tai tyhjään lopputulokseen.

Analysointityökalu tehtiin niin, että se pitää kirjaa datakatkoksista ja datahäviöistä. Jos todelliselle tehonjoustolle ei löydy arvoa tietylle tasatunnille, niin analysointityökalu poistaa tyhjät arvot datakehyksestä ja lisää ajanjakson erilliselle tiedostolle. Tämä tehdään sekä positiiviselle että negatiiviselle tehonjoustolle ja niistä paljastuvat mahdolliset datakatkot. Sama tehdään myös tehonjouston ennusteelle. Jos vertailtavaa paria ennusteen ja reaaliarvon välillä ei ole, tulostetaan se erilliseen tiedostoon. Tästä paljastuvat mahdolliset datahäviöt.

Analysointityökalu muuttaa datan aikaleimat samaan aikavyöhykkeeseen ja tallentaa ne sopivaan datakehykseen. Aikaleimojen avulla tehonjousta tehtiin tuntikohtaisia arvoja. Tehonjoustolle muodostettiin tuntikohtainen arvo laske-  
malla ensin positiiviselle ja negatiiviselle tehonjoustolle sekuntikohtainen keskiarvo ja edelleen minuuttikohtainen keskiarvo. Tehonjousten minuuttikohtaiset arvot yhdistettiin ottamalla negatiivisesta ja positiivisesta tehonjousta pienempi arvo yhteiseen datakehykseen. Näistä minuuttikohtaisista arvoista muodostettiin vielä tuntikohtainen keskiarvo. Tuntikohtainen keskiarvo on todellinen toteutunut tuntikohtainen joustoarvio seuraavalle alkavalle tunnille ja sitä verrattiin tekoälymallin ennusteeseen. Tämän lisäksi positiivisesta ja negatiivisesta tehonjousta tehtiin omat tuntikohtaiset keskiarvot, jotta niitä voitiin vertailla positiiviseen ja negatiiviseen tehonjouston ennusteeseen. Myös ennusteesta tehtiin tuntikohtainen keskiarvo valitsemalla pienempi arvo negatiivisesta ja positiivisesta tehonjousta yhteiseen datakehykseen. Tästä muodostui todellinen tehonjoustoennuste. Todellinen tehonjoustoennuste ja todellinen tehonjousto pyöristettiin vielä lähimpää 100 kilowattituntiin. Tämä kuvastaa sitä, toteutuiko kyseiselle tunnille tarjoutua tehonjousta. Fingridille tehtävä joustotarjous voi olla minimissään 100 kilowattia, ja pyöristys kertoo, toteutuuko tarjouspyyntö.

Analysointityökalu tekee numeerista vertailua toteutuneelle tehonjouston positiiviselle ja negatiiviselle arvolle ja vertaa sitä ennusteen tekemään positiiviseen ja

negatiiviseen arvoon. Näistä laskettiin keskineliövirheen neliöjuuri sekä prosentuaalisen eroavaisuus. Lisäksi todellista tehonjoustoennustetta sekä todellista tehonjousto verrattiin graafisesti.

Analysointityökalusta tehtiin joustava. Ohjelmaa kirjoittaessa huomattiin, että datasetti voi olla suuri ja joskus voi olla kätevämpää tarkastella tiettyä aikajanaa suurelta datasetiltä. Työkaluun siis ohjelmoitiin samat toiminnot erilliselle aikajanelle. Keskineliövirheen neliöjuuri sekä prosentuaalinen eroavaisuus ennusteen ja reaaliarvon välillä saadaan laskettua koko datasetille sekä määritetylle aikajanelle. Tämä helpottaa huomattavasti datasetin tarkastelua esimerkiksi viikon ajalta koko datasetin ollessa kuukausissa.

## 4.4 Analysointimenetelmät

### 4.4.1 Keskineliövirheen neliöjuuri

Numeerisina vertailuina käytettiin prosentuaalista ennustevirhettä todellisten tehonjouston arvojen sekä tehonjouston ennusteen välillä. Tämän lisäksi laskettiin keskineliövirheen neliöjuuri reaali- sekä ennustearvojen välille jokaiselle kuukaudelle sekä koko datasetille. Keskineliövirhe laskettiin kaavan 2 mukaan.

$$\text{Keskineliövirheen neliöjuuri} = \sqrt{\frac{1}{n} \sum_{j=1}^n (p - x)^2} \quad (2)$$

Keskineliövirheen neliöjuuri kuvaa ennusteen laatua. Se kertoo, miten hyvin tekoälymallin ennuste vastaa reaaliarvoja. Keskineliövirheen neliöjuuren haluttu arvo riippuu ongelman kontekstista sekä datan laajuudesta. Keskineliövirheen neliöjuuren tulosta voi pitää hyväksyttävänä, mikäli se olisi noin 10, mutta mielellään jopa alle.

### 4.4.2 Prosentuaalinen ennustevirhe

Prosentuaalinen ennustevirhe laskettiin kaavan 3 mukaan. Prosentuaalinen ennustevirhe kertoo mallin ennustetarkkuuden poikkeaman todellisesta arvosta.

Se ei anna kattavaa käsitystä mallin suorituskyvystä, koska se ei ota huomioon virheiden suuntaa. Kaavassa on myös otettu huomioon mahdolliset virheelliset negatiiviset arvot ja siksi jakajassa on käytetty itseisarvoa.

$$\text{Prosentuaalinen ennustevirhe} = \frac{\text{Reaaliarvo} - \text{Ennustearvo}}{|\text{Reaaliarvo}|} * 100 \quad (3)$$

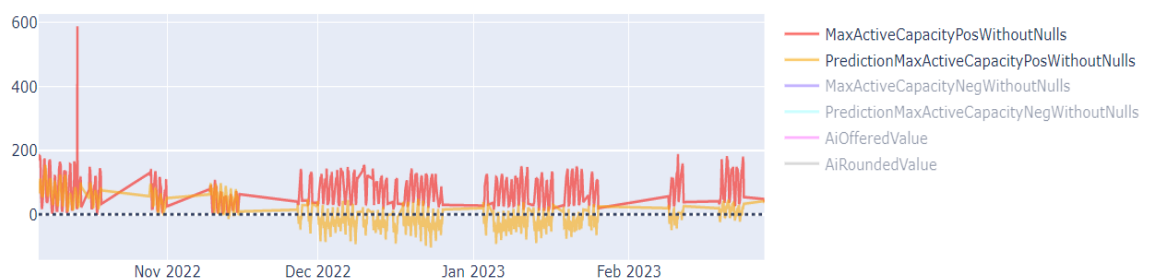
Reaaliarvon datakehystä on myös poistettu tyhjät arvot, jotta laskukaava toimi.

## 5 Validointiprosessi

### 5.1 Koneoppimismallin suorituskyvyn arviointi

#### 5.1.1 Datan käyttäytyminen

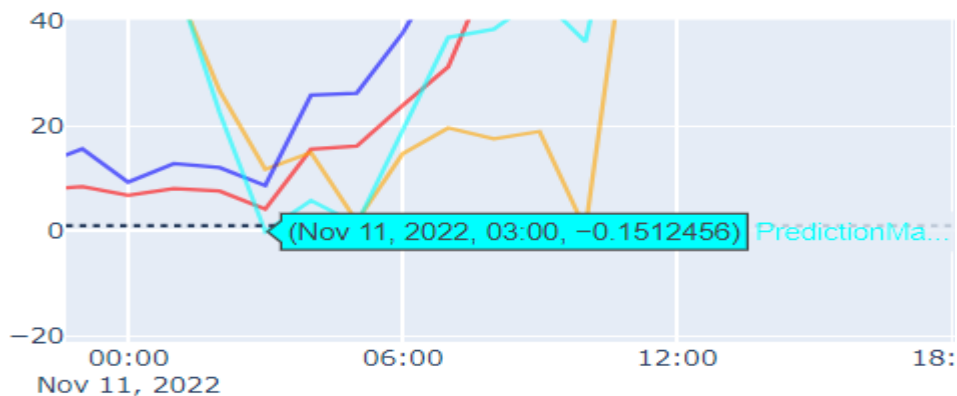
Graafeja katsomalla havaittiin, että graafissa oli datakatkoksia ja vuoden vaihteessa ennustus muuttui negatiiviseksi. Kuvasta 3 nähdään, kuinka ennuste karkaa nollan alapuolelle toteutuneen positiivisen tehonjouston ollessa nollan yläpuolella. Lisäksi huomattiin, että todellinen tehonjousto karkasi yli 2000 kilowattiin. Kuvaajaan luvut on muutettu 600 kilowattiin paremman skaalautuvuuden takia. Luvut eivät kuitenkaan vastaa todellisia lukuja ja ovat selkeästi jonkinlaisen häiriön merkki.



Kuva 3. Ennusteen vaihtelu kylmillä säillä.

Datakatkoksia havaittiin reilusti, ja ne näkyvät graafissa suorana viivana. Kuvan 3 kuvaajassa punainen väri on toteutunut positiivinen tehonjousto ja keltainen väri ennuste positiiviselle tehonjoustolle. Graafista on poistettu tyhjät arvot. Graafista näkyy, kun toteutuneelle tehonjoustolle ei löydy arvoja, jolloin tekoälymalli ei kykene ennustamaan tehonjoustoja. Graafia tutkittaessa kävi ilmi, että toteutuneelta tehonjoustosta puuttuu tunnin tai kahden arvot ja sen läheisyydestä puuttui usein tehonjoustoennuste. Malli ei välttämättä osaa ennustaa tulevien tuntien tehonjoustoennustetta, mikäli syötetiedoista puuttuu dataa. Tehonjoustoennusteen datahäviöt näkyvät graafissa suorana viivana. Ennusteen sekä todellisen tehonjousto arvojen puuttuminen viittaa jonkinlaiseen datakatkoon. Koska datakatkoksia oli runsaasti ja marraskuusta eteenpäin pitkälle helmikuuhun ennuste oli selkeästi virheellinen, keskityttiin erityisesti lokakuun ja maaliskuun ennusteisiin.

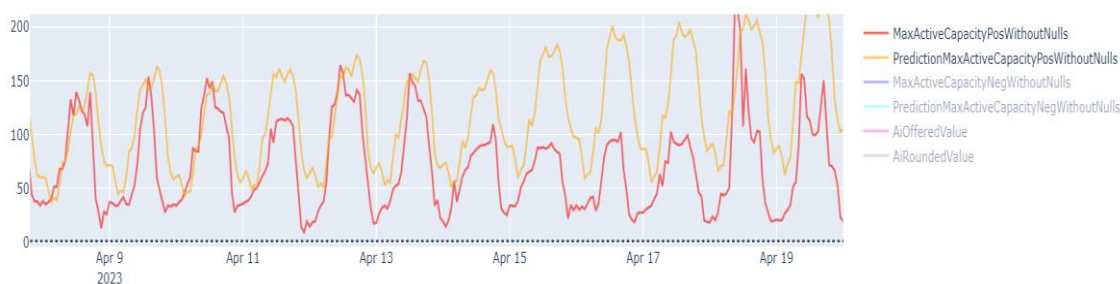
Lokakuun aikana havaittiin useampia datakatkoja todellisen tehonjousto dassa. Datakatkot aiheuttivat datahäviöitä myös ennusteen puolella. Positiivisen todellisen tehonjousto datakatkokset tapahtuivat yleensä aamusta tai aikaisin aamuyöstä. Ennuste positiiviselle tehonjoustolle puuttui yleensä yöaikaan ja kesti kahdesta päivästä viikkoon. Suoraa yhteyttä datakatkoille ei kuitenkaan löytynyt. Esimerkiksi negatiivisen tehonjousto ennuste puuttui 14.10.2022 klo 22:00 ja 16.10.2022 klo 9:00 ajanjaksolta, kun taas todellisen negatiivisen tehonjousto arvot puuttuivat vain 17.10.2022 klo 12:00. Kokonaisuudessaan ennuste puuttui lähes puolelta kuukaudelta ja ehjiä viikkoja lokakuulta löytyi vain viikon ajalta ja kahdelta ajanjaksolta.



Kuva 4. Tehonjouston ennuste menee negatiiviseksi.

Marraskuun tehonjousto oli hyvin rikkonainen. Datakatkoksia on lähes koko kuukaudelta. Marraskuusta löytyi vain yksi ehjä viikko, jolta löytyy jonkinlainen ennuste. Ennuste muuttuu negatiiviseksi marraskuun loppupuolella kuvan 4 mukaan. Ennuste muuttuu negatiiviseksi aamu 3:00 ja 5:00 välillä, mutta nousee takaisin positiiviseksi. Ennuste muuttuu joulukuun puolella lähes kokonaan negatiiviseksi.

Kuvan 4 vaalean sininen on ennuste negatiiviselle tehonjoustolle. Kuvassa näkyy, kun ennuste menee negatiiviseksi. Kun ennuste menee negatiiviseksi, on kyseessä jonkinlainen häiriö. Huomattavaa oli, että ennuste meni lähes poikkeuksetta negatiiviseksi heti aamuyöstä.



Kuva 5. Huhtikuun ennusteita.

Maaliskuun aikana ennuste muuttuu taas positiiviseksi ja muistuttaa enemmän lokakuun ennustetta. Sään viilentyessä ennuste ei seuraa reaaliarvoa kuten kuvassa 5 näkyy. 14.4. ja 18.4. välillä sää kylmenee, ja lämmitystarpeen kasvessa tehonjousto laskee. Ennusteen mukaan joustoa olisi jopa alkukuukautta enemmän. Kuitenkin ennusteen kannalta maaliskuun ja huhtikuun alku näyttää huomattavasti tarkemmalta verrattuna alkuvuoteen.

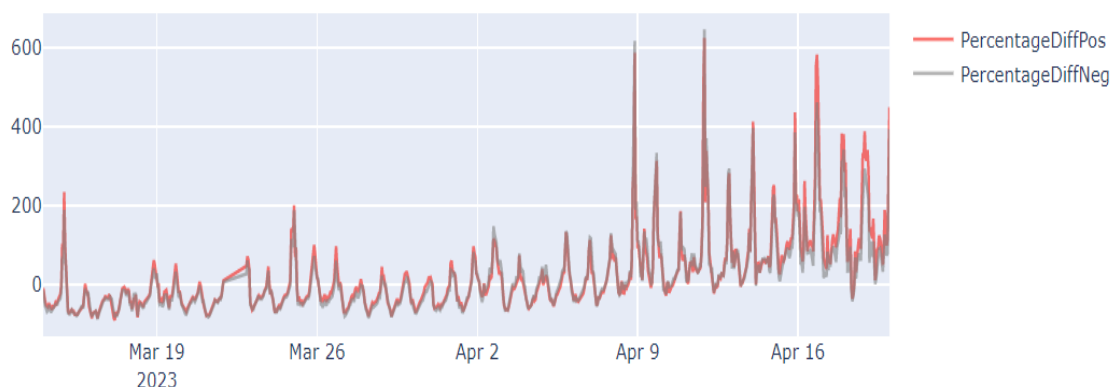
### 5.1.2 Numeeriset vertailut

Keskineliövirheen neliöjuuren tulokset olivat melko kehoja. Koko datasetin ajalle keskineliövirheen neliöjuuri oli 94,9 negatiiviselle ja 96,1 positiiviselle tehonjoustolle. Tuloksessa näkyvät negatiiviset ennusteet, jotka nostavat lukemaa selkeästi. Lokakuulle saatiin 38,1 negatiiviselle ja 53,4 positiiviselle tehonjoustolle. Marraskuulle saatiin 47,1 negatiiviselle ja 56,7 positiiviselle. Joulukuun lukemat olivat yli 100, kun ennuste oli joulukuun ajalta negatiivinen. Maaliskuun alusta huhtikuun alkuun olevalle ajanjaksolle saatiin lukemiksi 39,1 positiiviselle ja 44,4 negatiiviselle tehonjoustolle. Tulokset olivat siis kehoja.

Koko datasetin prosentuaalinen ennustevirhe vaihtelee -500 ja 2300 prosentin väliltä. Negatiiviset arvot kertovat, että ennuste on ollut reaaliarvoa suurempi. Suurimmat ennustevirheet ovat taas ajalta, jolloin ennuste on negatiivinen ja reaaliarvo on positiivinen. Lokakuun ajalta virheprosentti on pääosin 0 ja 500 prosentin välissä muutamaa isoa virhepiikkiä lukuun ottamatta.

Marraskuun osalta prosentuaalinen ennustevirhe on suurimmillaan ja vaihtelee -180 ja 2300 prosentin välillä. Marraskuussa vaihtelu on hyvin suurta ja epätaisaista. Joulukuussa ennustevirhe tasaantuu ja on -400 ja 160 prosentin välissä. Ennuste on tältä ajalta negatiivinen ja vaihtelee -100 ja 50 prosentin välillä. Sama toistuu myös tammi- ja helmikuussa. Prosentuaalinen ennustevirhe vaihtelee melko vähän, mutta ennuste on negatiivinen. Näitä lukuja ei voida

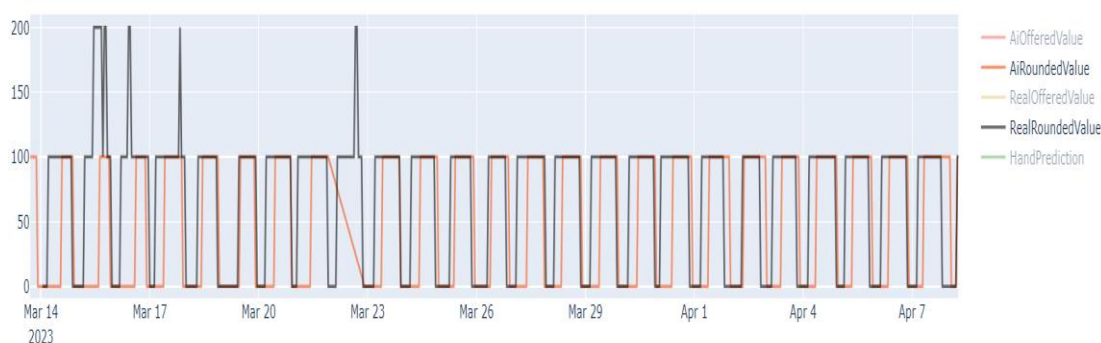
huomioida, sillä ennusteen ei pitäisi mennä negatiiviseksi.



Kuva 6. Prosentuaalisia ennustevirheitä.

Maaliskuussa prosentuaalinen ennustevirhe näyttää tasaiselta kuvan 6 tavoin. Punainen väri kuvaa positiivisen tehonjouston prosentuaalista ennustevirhettä ennusteen ja toteutuneen tehonjouston välillä. Musta väri kuvaa prosentuaalista ennustevirhettä toteutuneen negatiivisen tehonjouston ja tehonjouston ennusteen välillä. Maaliskuussa ennustevirhe vaihtelee 5 ja 235 prosentin välillä. Ennuste on kuitenkin tältä ajanjaksolta kuitenkin positiivinen. Maaliskuun ennusteet ovat siis huomattavasti parempia verrattuna muihin kuukausiin. Prosentuaalinen ennustevirhe kuitenkin huononee huhtikuulle ja vaihtelee -40 ja 650 prosentin välillä. Ennustevirhe vaihtelee voimakkaasti eikä ole tasainen.





Kuva 7. Pyöristetyt tulokset

Kun positiivisesta ja negatiivisesta tehonjousta muodostetaan todellinen joustotarjous, pyöristetään tehonjousto lähimpään sataseen. Mikäli tehonjousto on suurempi kuin 50 kilowattia, tehonjousto pyöristetään 100 kilowattiin ja tästä tehdään tarjous Fingridille. Vastaavasti tehonjouston ollessa alle 50 kilowattia, luku pyöristetään nolnaan eikä joustotarjousta voida tehdä. Kuvassa 7 näkyy pyöristettyjen lukujen vertailua. Kuvaajasta nähdään, kuinka tekoälyn tekemä tarjous tulee hieman toteutunutta tarjousta jäljessä. Graafi näyttää suhteellisen säännölliseltä, ja maaliskuun ennuste näyttää jopa hyvältä.

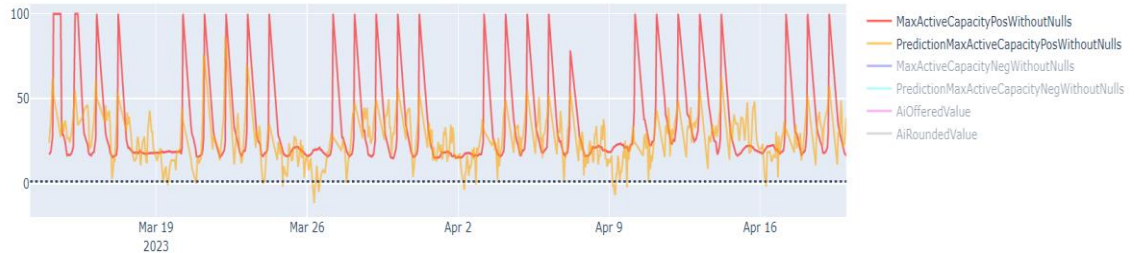
Käsin tehty ennuste oli huomattavasti tarkempi ja vakaampi tekoälyn ennusteeseen verrattuna. Käsin tehdystä ennusteesta ei löytynyt dataa maaliskuulta, joten sitä ei verrattu tekoälyn ennusteeseen. Projektin alkupuolella käsin tehty ennuste oli huomattavasti tarkempi tehonjoustoennustukseen nähden. Kuitenkin datan rikkonaisuuden takia suurempia johtopäätöksiä on vaikea tehdä pidemmältä aikaväliltä.

## 5.2 Koneoppimismallin validointi uuteen käyttökohteeseen

Tekoälymallia yritettiin toistaa toiselle käyttökohteelle. Toisella käyttökohteella oli samankaltaista syötedataa, joten sama tekoälymalli opetettiin toiselle käyttökohteelle. Kyseinen käyttökohte otettiin käyttöön alkuvuodesta ja ennustedataa löytyy vain maaliskuusta lähtien.

### 5.2.1 Datan käyttäytyminen

Graafeja katsomalla havaittiin, että datan arvoja on selkeästi pyöristetty. Ennusteen ja todellisen tehonjouston noustessa lähelle 100 kilowattia. Tehonjouston arvo on pyöristetty tasalukuun. Syytä tähän ei löytynyt. Tehonjousto ei ylitä 100 kilowattia missään vaiheessa kuten kävi ensimmäisellä kohteella.

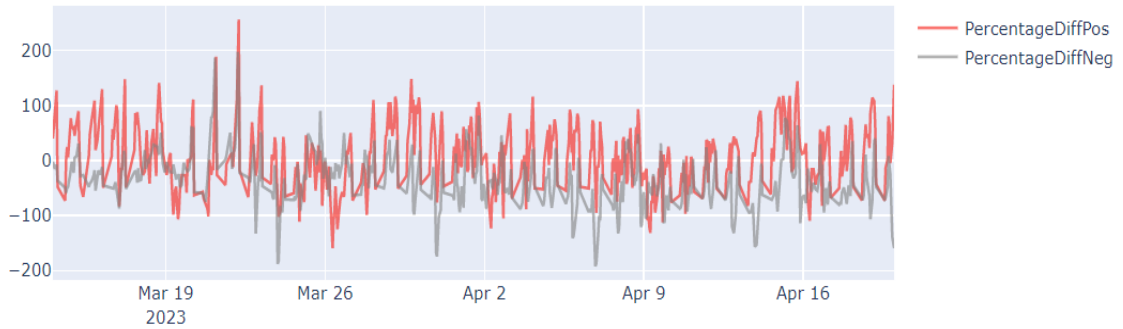


Kuva 8. Toisen käyttökohteen ennusteita.

Toisen käyttökohteen malli käyttäytyy jokseenkin samoin tavoin kuin ensimmäinen käyttökohde. Kuvassa 8 näkyy, kuinka ennuste karkaa ajoittain negatiiviseksi. Kuitenkin ennuste jää huomattavasti alemmalle tasolle kuin toteutunut tehonjousto. Huomattavaa oli, että jokainen toteutunut tehonjouston piikki ajoittuu jokaiselle päivälle klo 5:00. Poikkeuksena on 7.4., jolloin toteutunut tehonjousto ylittää vain 78 kilowattiin. Tehonjoustoennuste puolestaan lähentelee 50 kilowattia, mutta ei juurikaan ylitä sitä. Toteutunut tehonjousto on kuitenkin liian säännöllinen. Vaikuttaisi siltä, että dataa on muutettu tai malli on oppinut jonkinlaisen käyttäytymismallin.

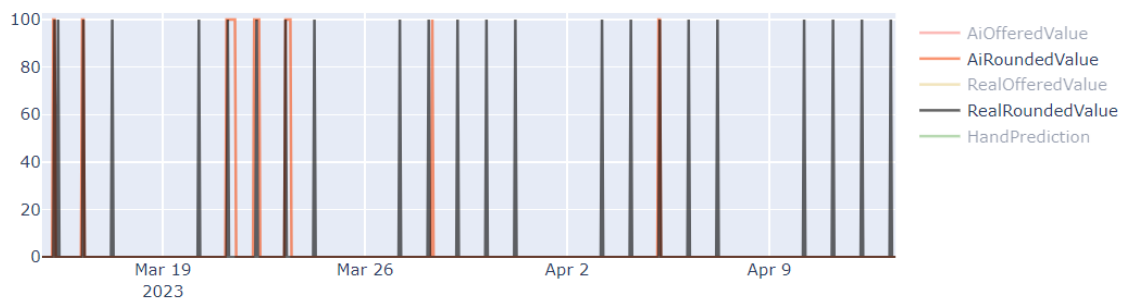
### 5.2.2 Numeeriset vertailut

Toisen käyttökohteen keskineliövirheen neliöjuuren tulokset olivat huomattavasti paremmat. Koko datasetin virhe oli 17,6 positiiviselle sekä 21,1 negatiiviselle tehonjoustolle. Tulos lähentelee jo hyväksyttävää tasoa. Dataa löytyi vain kuukauden ajalta, joten vertailukelpoista kuukautta ei löydy. Tulos on kuitenkin hyvinkin lupaava heti ensimmäiselle ajanjaksolle.



Kuva 8. Toisen käyttökohteen prosentuaalinen ennustevirhe.

Prosentuaalinen ennustevirhe oli myös huomattavasti maltillisempi. Ennustevirhe vaihtelee -180 ja 260 prosentin välillä. Kuvassa 8 näkyy, kuinka suurimmat virhepiikit osuvat klo 5:00 aikoihin. Ennustevirhe pysyy kuitenkin kohtuullisena suurimman osan ajasta.



Kuva 9. Toisen käyttökohteen pyöristetyt tulokset.

Pyöristetyissä joustoarvoissa nähdään taas huomattava ero ensimmäiseen käyttökohteeseen. Nopeat tehonjoustopiikit näkyvät entistä selvemmin pyöristetyissä arvoissa. Toteutunut tehonjoustotarjous syntyy jokaiselle päivälle klo 5:00. Tehonjoustoennuste toisaalta ei juurikaan yllä 100 kilowattiin muutamaa päivää lukuun ottamatta. Ennusteen mukaan tehonjoustotarjous muodostuu vain neljälle päivälle klo 5:00 ja klo 10:00 välillä. Kohteesta ei ollut käsin tehtyä ennustetta saatavilla, joten sitä ei vertailtu.

## 6 Johtopäätökset

### 6.1 Datan laatu

Datan esikäsittely on erityisen tärkeää. Tekoälyn suorituskyky perustuu täysin datan käyttäytymiseen. Tästä syystä tekoälyn suorituskyky on niin hyvä kuin data, mihin tekoälyn päätökset perustuvat. Varsinkin pakkassäällä malli näyttäisi käyttäytyvän arvaamattomasti. Juurisyihin tulisi perehtyä syvällisemmin. Kohteiden opetusdatan esikäsittelyyn tulisi perehtyä syvällisemmin, kuten esimerkiksi datan normalisointiin ja puhdistusprosessiin. Mikäli esikäsittely tehdään huolellisesti, voisi se mahdollisesti parantaa mallin suorituskykyä. Dataa oli runsaasti opetusvaiheessa, joten mallin suorituskyvyn voisi olettaa olevan hyväksyttävällä tasolla. Se tulisi kuitenkin varmistaa.

Datassa oli huomattavan paljon datakatkoja. Datakatkojen syyhyn tulisi perehtyä tarkemmin. Johtopäätöksiä on vaikea tehdä, kun ennusteesta puuttuu kokonaisuuksia. Datakatkojen syytä on osittain selvitetty ja ainakin osan oletetaan johtuvan dataympäristön rakennusprosessista. Katkokset voivat liittyä tiedonsiirtoon tai puutteelliseen datankeruuseen. Jos datakatkoja ei saada korjattua, tulisi miettiä erilaista neuroniverkkomallia, jossa datakatkokkien sietokyky on parempi. Toinen vaihtoehto olisi esikäsitellä data ja täyttää data-aukot edellisen tunnin arvoilla tai jonkinlaisella arviolla tunnetun datan perusteella. Dataaukkoja pyrittiin täyttämään jälkikäteen, mutta graafeja tutkittaessa huomattiin, että datakatkokkia oli vielä jäljellä. Data-aukkojen täyttö ei vielä poistanut datakatkokkia ja datakatkojen pitkittyessä, aukot voivat pahentaa ennustetta. Lyhyet datakatkokset voivat aiheuttaa pieniä virheitä ennustuksissa, ja se myös osittain selittää kehua ennustetta.

Kohteen soveltuvuutta olisi hyvä myös tarkastella. Vaikka dataa on riittävästi, niin pitää myös tarkastella, kuinka soveltuvia kohteet ovat käytetyille tekoälymallille. Mikäli kohteen suuruusluokka vaihtelee, esimerkiksi kaupungista ostoskeskukseen, tulee se myös huomioida tekoälymallia luodessa. Joihinkin kohteisiin

tulee mahdollisesti soveltaa erilaista hermoverkostoa, mikäli käyttökohteen luonne muuttuu.

## 6.2 Tulokset

Ensimmäisen käyttökohteen tulokset olivat osittain lupaavia ja osittain erittäin huonoja. Projekti on edennyt hitaasti ja dataympäristön pystytyksessä on mennyt odotettua kauemmin. Ensimmäisen kohteen dataa ruvettiin seuraamaan lokakuun alusta, mutta datan perusteella malli on vasta viime viikkoina saatu toimimaan kohtuullisella varmuudella. Seuraavien kuukausien datasetit tulevat määrittämään miten projektissa kannattaa jatkaa. Ensimmäisen käyttökohteen ennustus on ainakin näyttänyt osittain lupaavalta, vaikka ennustus tuleekin toteutuneen tehonjoustoja jäljessä, mutta pienellä optimoinnilla ennuste osuisi melko hyvin toteutuneeseen tehonjoustoan. Mikäli todetaan, että ennuste kulkee muutaman tunnin jäljessä, tulisi kysyntäjousto laskentaa tarkastaa. On mahdollista, että lukujen pyöristyksessä on tullut virhe ja joko ennusteen tai todellisen arvon lukemat ovat datasetissä kirjattu väärälle tunnille. Aikaleimassa voi silloin olla muutaman tunnin heitto. Muutamat viikot antavat kuitenkin toivoa, että mallista saadaan riittävän hyvä.

Toisen käyttökohteen osalta tulokset ovat hieman ristiriitaisia. Toisaalta prosentuaalinen ennustusvirhe on pienempi, mutta toisaalta toteutunut tehonjousto, tehonjoustoan ennuste ja pyöristetty todellinen ennuste on lähtökohtaisesti hyvin erilaista dataa verrattuna ensimmäiseen käyttökohteeseen. Kysyntäjousto tapahtuisi pääosin yöstä, mikä ei kuulosta kovinkaan järkevältä. Graafit olivat liian säännöllisiä säästä riippumatta. Toisen käyttökohteen dataa tulisi tarkastella lähemmin, onko data validia. Jos datassa ei ole vikaa niin silloin ennustus on myös käyttäytynyt suurin piirtein toivotulla tavalla, vaikka joustotarjoukseen ei ennustuksessa aina päästy. Kuitenkin ennustus myötäilee todellisia arvoja, vaikka jääkin hieman vajaaksi.

### 6.3 Parannusehdotukset

Datan luotettavuutta tulisi parantaa. Tämä tarkoittaa luotettavampaa mittaristoa tai laitteistoa kohteelle, jotta datakatkoja ei ilmene. Mikäli datakatkoihin ei voida vaikuttaa, tulisi syöttödatan esikäsittelyä parantaa, jotta ennusteen toimintavarmuutta voitaisiin parantaa. Tämän lisäksi olisi hyvä myös tarkastaa opetusdatan laatu ja eroavaisuudet käyttökohteiden välillä. Mikäli kohteet eroavat huomattavasti toisistaan, voidaan ehkä mallin soveltavuudesta käyttökohteille tehdä jonkinlaisia johtopäätöksiä.

Tekoälymallin ennustuksen varmuutta voitaisiin parantaa ja kokeilla erilaisia muutoksia opetusprosessissa. Tekoälymallia luotaessa voitaisiin lisätä epookkien määrää. Tämä tarkoittaisi oppimisdatan läpikäynnin lisäämistä. Epookkeja lisäämällä malli pystyy oppimaan koulutusdatan ominaisuuksia syvällisemmin. Kuitenkin liiallinen määrä voi lisätä ylisovituksen riskiä. Koulutusvaihetta olisi kuitenkin hyvä tarkastella enemmän. Oppimisaikaa voitaisiin myös lisätä samalla tai monimutkaisemmalla hermoverkolla. Hermoverkkorakennetta voitaisiin muuttaa niin, että malli sietää datakatkoja paremmin. Hermoverkkoon voitaisiin tehdä myös muita muutoksia, kuten lisäämällä piilotettujen kerroksien määrää tai hermosolujen määrää hermoverkon kerroksissa. Piilotettujen kerroksien lisääminen toisi lisäkapasiteettia mallille. Tämä mahdollistaisi mallin oppimaan monimutkaisempia suhteita syötteiden ja ennusteiden välillä. Lisättyjen piilokerrosten avulla malli pystyisi oppimaan abstraktimpia piirteitä datasta ja se voi parantaa ennustuskykyä monimutkaisemmissa asioissa. Toisaalta kerroksia lisäämällä lisätään myös ylisovitusriskiä, ja malli voi oppia koulutusdatan kohinaa.

Huomattavaa on myös, että mallin käyttäytyminen vaikuttaa reagoivan lämpötilaan. Graafeja tutkiessa huomattiin, että ennustus on todennäköisesti negatiivinen pakkassäällä. Lämpötilan vaikuttavuutta mallin käyttäytymiseen tulisi myös tarkastella tarkemmin. Tässä on varmasti sellainen asia, johon voisi kiinnittää huomiota.

## 7 Yhteenveto

Ilmastonmuutoksen pysäyttämiseksi uusiutuvan energian määrä kasvaa energiantuotannossa. Näin ollen sähkötuotannosta tulee yhä enemmän sääriippuvaista, ja se luo uusia haasteita myös sähköverkostolle.

Koneoppiminen on onneksi ottanut suuria harppauksia viime vuosina, ja se voi mahdollistaa rakennusautomaation laajamittaisen kysynnänjouston käyttöönoton lähivuosina. Tässä voi olla yksi lääke stabiilimpaan sähkönjakeluun. Koneoppimisella on suuri potentiaali tehostaa rakennusautomaation energiatehokkuutta ja se toisi meidät lähemmäksi päästöttömämpää huomista. Tämän lisäksi sillä pystyisi vielä tekemään hyvää bisnestä.

## Lähteet

- 1 Kokko, K. (2020) Tekoälyn sovelluksia konetekniikassa. Kandidaatintyö. Oulun yliopisto. s.12 <<http://jultika.oulu.fi/files/nbnfioulu-202011253201.pdf>> Luettu 9.4.2023.
- 2 Elements of Ai. (2023). Verkkoaineisto. Helsingin yliopisto. <<https://course.elementsofai.com/fi/4/1>> Luettu 10.4.2023.
- 3 GPT-4 is OpenAI's most advanced system, producing safer and more useful responses. (2023) Verkkoaineisto. OpenAi. <<https://openai.com/product/gpt-4>> Luettu 15.4.2023.
- 4 Truly, Alan, 6.4.2023. GPT-4: how to use, new features, availability, and more. Digitaltrends. <<https://www.digitaltrends.com/computing/chatgpt-4-everything-we-know-so-far/>> Luettu 15.4.2023.
- 5 Marr, Bernard 24.2.2023. GPT-4 Is Coming – What We Know So Far. Forbes. <<https://www.forbes.com/sites/bernardmarr/2023/02/24/gpt-4-is-coming--what-we-know-so-far/>> Luettu 16.4.2023.
- 6 Linus tech tips, I Am Hard... [REDACTED] - WAN Show March 17, 2023, Linus media group. <<https://youtu.be/m4RolgXsoxY?t=6517>> Katsottu 17.3.2023.
- 7 Haponik, Artur. 2023. AI, big data, and machine learning in transportation. Addepto. <<https://addepto.com/blog/ai-big-data-and-machine-learning-in-transportation/>> Luettu 20.4.2023.
- 8 Deb, Rupam. 19.12.2022. What are Machine Learning Applications? Top 10 Industry and Real-World Use Cases. Emeritus. <<https://emeritus.org/blog/machine-learning-what-are-machine-learning-applications/>> Luettu 25.4.2023.
- 9 McCrea, Ollie. 4.7.2021. The use of Ai image recognition in medicine. Sentsight.ai. <<https://www.sentsight.ai/the-use-of-ai-image-recognition-in-medicine/>> Luettu 26.4.2023.
- 10 Elhusseny, Ouf, Mohamed. 2022. Future computing and informatics journal. Volume 7 Issue 1. <<https://digitalcommons.aaru.edu.jo/cgi/viewcontent.cgi?article=1152&context=fcij>> Luettu 27.4.2023.



- 11 Lindhold, Petteri, 8.3.2023. Koneoppiminen käy miltei 40 miljardin euron markkinoiden kimppuun – kartelleja etsii menetelmä, joka olisi paljastanut aiemmat kartellit. Yle uutiset. <<https://yle.fi/a/74-20018862>> Luettu 27.4.2023.
- 12 Energiatehokkuutta datan avulla Hämeenlinnan korkeakoulukeskuksessa kaukolämmön kulutuksessa. 10.8.2022. HAMK Beat. <<https://blog.hamk.fi/hamk-beat/energiatehokkuutta-datan-avulla-hameenlinnan-korkeakoulukeskuksessa-kaukolammon-kulutuksessa/>> Luettu 27.4.2023.
- 13 Kyytsönen, Jouko. 19.2.2021. Sähkökatkojen määrä kasvaa, sillä taajuuden muutokset lisääntyvät tuuli- ja vesivoiman vaihtelun vuoksi. <<https://www.maaseuduntulevaisuus.fi/uutiset/4d8f7f63-cb86-5db0-b669-835ad97a2df1>> Luettu 28.4.2023.
- 14 Kysyntäjousto. 2022. Verkkoaineisto. Fingrid. <<https://www.fingrid.fi/sahkomarkkinat/markkinoiden-yhtenaisyyys/pilottihankkeita/kysyntajousto/>> Luettu 28.4.2023.
- 15 Tekoälyn perusteita ja sovelluksia. 2023. Verkkoaineisto. Jyväskylän yliopisto. <<https://tim.jyu.fi/view/kurssit/tie/tiep1000/tekoalyn-sovellukset/kirja#aktivointifunktiot>> Luettu 28.4.2023.
- 16 Oscar Alsing Ferreira, 15.6.2017. But How Does The MultiLayer Perceptron Actually Work? <<https://www.youtube.com/watch?v=AZEfmoWBXwg>> Katsottu 29.4.2023.
- 17 Tekoälyn perusteita ja sovelluksia. 2023. Verkkoaineisto. Jyväskylän yliopisto. <<https://tim.jyu.fi/view/kurssit/tie/tiep1000/tekoalyn-sovellukset/kirja#yli-ja-alisovittaminen>> Luettu 29.4.2023.