



Data-analytiikan ja tekoälyn hyödyntäminen ennustavaan kunnossapitoon.

Janne Hakala

OPINNÄYTETYÖ
Tammikuu 2021

Dataosaamisen ja tekoälyn ylempi tutkinto-ohjelma
(Insinööri ylempi AMK)

TIIVISTELMÄ

Tampereen ammattikorkeakoulu
Dataosaamisen ja tekoälyn ylempi tutkinto-ohjelma
(Insinööri ylempi AMK)

HAKALA, JANNE:

Data-analytiikan ja tekoälyn hyödyntäminen ennustavaan kunnossapitoon

Opinnäytetyö 63 sivua, joista liitteitä 9 sivua
tammikuu 2021

Monissa kunnossapito-organisaatioissa keskeisenä päivittäisen toiminnanohjauksen työkaluna on kunnossapidon tietojärjestelmä. Näihin järjestelmiin kerääntyy päivittäin suuria määriä ihmisten tuottamaa dataa, mutta niiden jatkohyödyntämismahdollisuudet ovat usein jääneet vähemmälle huomiolle.

Uusimpien data-analytiikan ja tekoälyn menetelmien avulla datasta on mahdollista saada esiin malleja, joita voi hyödyntää kunnossapitotoiminnassa muun muassa tehokkaampaan vikaantumisten ennustettavuuteen.

Tässä opinnäytetyössä tutkittiin vikaantumisten ennustettavuutta kunnossapitojärjestelmään kerääntynyttä dataa käyttämällä. Opinnäytetyössä on tutkittu kunnossapidon nykytilaa ja tulevaisuuden näkymiä sekä uusimpien data-analytiikkamenetelmien periaatteita ja hyödyntämismahdollisuuksia.

Tämän opinnäytetyön Case study -tutkimusosiossa käytettiin suomalaisen teollisuusyrityksen kunnossapitojärjestelmään keräämää dataa kahdeksan vuoden ajalta, yli miljoonan tapahtumarivin verran. Tämän datan perusteella pyrittiin löytämään tekoälyn ja koneoppimisen menetelmää käyttämällä vikaantumisten ennustemallia, joita verrattiin toteutuneisiin vikaantumisiin.

Lähtökohdiltaan tutkimusongelma oli haasteellinen, sillä kunnossapitojärjestelmään kerättyä dataa ja datamalleja ei lähtökohtaisesti oltu suunniteltu vikaantumisten ennustamiseen. Tämän takia myös vikaantumisennusteet olivat parhaimmillaankin vain kohtalaisia.

Tutkimustyö oli teoriaa ja käytäntöä hyödyntävä toimintatutkimus, jonka perusteella voitiin antaa kunnossapitojärjestelmän kehitykseen ja toimintatapoihin liittyviä parannusehdotuksia. Tämä osaltaan mahdollistaa paremman ennustettavuuden sekä yhdistämisen osaksi kognitiivista kunnossapitoa. Tutkimuksen tuloksista ei voida yleisesti päätellä ennustettavuuden mahdollisuuksia olemassa olevan datan perusteella, mutta ne parantavat käsityksiä datan laadun merkityksestä ennustettavuuden parantamiseksi.

ABSTRACT

Tampere University of Applied Sciences
Data Expertise and Artificial Intelligence master's program

HAKALA JANNE:

Utilization of data analytics and artificial intelligence for predictive maintenance

Bachelor's thesis 63 pages, appendices 9 pages

January 2021

In many maintenance organizations, the maintenance management system is a key tool for day-to-day operations management. These systems accumulate large amounts of human-generated data on a daily basis, but their potential for further utilization has often received less attention.

Thanks to the latest data analytics and artificial intelligence methods, it is possible to derive models from the data that can be utilized in maintenance operations of more efficient predictability of failures for instance.

In this thesis, the predictability of failures was investigated using data accumulated in the maintenance management system. The current state and future prospects of maintenance as well as the principles and utilization possibilities of the latest data analysis methods have been studied in the thesis.

In the case study section of this thesis, the data collected by the Finnish industrial company for the maintenance management system over a period of eight years, using more than a million event lines, was used. Based on this data, an attempt was made to find a method of artificial intelligence and machine learning using a failure prediction model that was compared to actual failures.

From the point of view, the research problem was challenging, as the data and data models collected in the maintenance system were not, in principle, designed to predict failures. Because of this, even the failure predictions were only moderate at best.

The research work was an action research utilizing theory and practice, on the basis of which suggestions for improvement related to the development and operating methods of the maintenance system could be given. This contributes to better predictability as well as integration into cognitive maintenance. The results of the study do not generally suggest predictability based on existing data, but they improve perceptions of the importance of data quality to improve predictability.

SISÄLLYS

1	JOHDANTO	7
1.1	Opinnäytetyön tausta	7
1.2	Opinnäytetyön tavoitteet ja rajaukset	7
1.3	Case-yritys	8
2	TUOTANTO-OMAISUUDEN HALLINTA.....	9
2.1	Kunnossapito	9
2.2	Kunnossapitolajit	10
2.3	Kunnossapidon kustannukset	10
3	KUNNOSSAPIDON TOIMINNANOHAUSJÄRJESTELMÄ.....	12
3.1	Käyttäjryhmät	12
3.2	Master data	13
3.3	Tapahtumadata	14
4	KUNNOSSAPIDON TULEVAISUUS.....	15
4.1	Ennustava kunnossapito	16
4.2	Kognitiivinen kunnossapito.....	17
4.3	Toiminnanohjausjärjestelmän rooli.....	17
5	DATA-ANALYTIikka	19
5.1	Data-analytiikan tasot.....	19
5.2	Älykäs data-analytiikka.....	20
5.3	Ennakoiva analytiikka.....	21
6	TEKOÄLY	23
6.1	Koneoppiminen	24
6.2	Neuroverkot	25
6.3	Neuroverkon rakentaminen.....	26
6.4	Neuroverkon kouluttaminen ja optimointi	27
7	CASE: YRITYS OY	28
7.1	Tietojärjestelmä ja datan valinta.....	28
7.2	Yleiskatsaus dataan	28
7.3	Kohteiden luokittelu.....	29
7.4	Työtilausten analysointi ja esikäsittely.....	30
7.5	Datan sovittaminen koneoppimismalliin	36
7.6	Datan esikäsittely	37
7.7	Neuroverkon rakentaminen opettaminen ja analysointi.....	38
7.7.1	Kirjastojen tuonti ja datan käsittely	38
7.7.2	Neuroverkon luominen	40
7.7.3	Mallin tulosten raportointi ja visualisointi.....	41

7.8 Mallin tulosten arviointi.....	43
8 EAM-JÄRJESTELMÄN KEHITYS.....	47
8.1 Tietorakenteet.....	47
8.2 Master data.....	47
8.3 Tapahtumadata.....	48
8.4 Järjestelmän käyttö.....	49
8.5 Kehitysehdotusten vienti käytäntöön.....	50
9 POHDINTA.....	51
LÄHTEET.....	53
LIITTEET.....	55
Liite 1. Datan esikäsittely 1 (2).....	55
Liite 1. Datan esikäsittely 2 (2).....	56
Liite 2. Ote käsitellystä datasta (csv).....	57
Liite 3. Ennustemallien toimivuus laitetyypeittäin 1 (6).....	58
Liite 3. Ennustemallien toimivuus laitetyypeittäin 2 (6).....	59
Liite 3. Ennustemallien toimivuus laitetyypeittäin 3 (6).....	60
Liite 3. Ennustemallien toimivuus laitetyypeittäin 4 (6).....	61
Liite 3. Ennustemallien toimivuus laitetyypeittäin 5 (6).....	62
Liite 3. Ennustemallien toimivuus laitetyypeittäin 6 (6).....	63

LYHENTEET JA TERMIT

Kohde	Laiteyksilö tai muu kohde johon työtilaus on kohdistettu
Työtilaus	Yksittäinen järjestelmään kirjattu työsuorite raportointeeseen
Preventive maintenance	Ennakoiva kunnossapito (suunniteltu)
Reactive maintenance	Reagoiva kunnossapito (häiriökorjaukset)
IoT	Esineiden ja asioiden internet Internet of Things
CMMS	Kunnossapitojärjestelmä (Computerized Maintenance Management System)
EAM	Tuotanto-omaisuuden hallintajärjestelmä (Enterprise Asset Management)
CM	Kognitiivinen kunnossapito (Cognitive Maintenance)
ANN	Keinotekoinen neuroverkko (Artificial Neural Network)

1 JOHDANTO

1.1 Opinnäytetyön tausta

Kunnossapitotoiminta on tärkeä osa organisaation tuotanto-omaisuuden hallintaa. Kunnossapidon avulla varmistetaan tuotantoprosessien häiriötön toiminta ja näin varmistetaan mahdollisimman tehokas tuotanto. Kunnossapidon rooli yrityksen tuottavuuden kehittämisessä ja kehittämisessä on merkittävä.

Kunnossapitotoiminnassa kerätään erilaista kunnossapitodataa useista eri lähteistä. Yksi merkittävä kunnossapitodatan lähteistä on kunnossapidon toiminnanohjausjärjestelmä. Toiminnanohjausjärjestelmät ovat olleet käytössä jo useiden vuosikymmenten ajan. Näihin järjestelmiin kertynyt data on lähes yksinomaan ihmisten tuottamaa ja siksi datan laatu voi vaihdella eri kunnossapito-organisaatioiden välillä huomattavastikin.

Toiminnanohjausjärjestelmän kautta hallitaan kunnossapidon päivittäisiä prosesseja, joita ovat muun muassa koneiden ja laitteiden vikakorjaukset ja niiden ennakko- ja huollot. Lisäksi kunnossapitodataa voidaan kerätä kunnossapidettäviin kohteisiin liitetyistä antureista, jotka siirtävät tietoa erilaisiin tietovarastoihin myöhemmää hyödyntämistä varten. Tätä kutsutaan usein nimellä IoT, joka muodostuu sanoista *Internet of Things* (IoT = asioiden ja esineiden internet).

1.2 Opinnäytetyön tavoitteet ja rajaukset

Kerätyllä kunnossapitodatalla ei sinällään ole itseisarvoa, ellei sitä jalosteta ja hyödynnetä kunnossapidon päätöksentekotilanteissa (Ali-Marttila, Kärri, Marttonen-Arola, Metso, Sinkkonen & Ylä-Kujala 2019, 46). Kunnossapidon toiminnanohjausjärjestelmässä oleva data on usein kehitetty palvelemaan ensisijaisesti kunnossapidon toimintaa ja sen raportointia. Toiminnanohjausjärjestelmien pää-tarkoituksena ei siis ole ollut erityisesti kunnossapidon ennustamiseen liittyvän tiedon kerääminen.

Uusien data-analytiikka- ja tekoälytyökalujen sekä niiden käyttösovellutusten kehitys on mahdollistanut erilaisten analytiikkamallien tekemisen ja tehokkaan data-analytiikan. Tekoälyn avulla on mahdollista tehdä näitä malleja suuresta datamäärästä, joita muun muassa kunnossapidon toiminnanohjausjärjestelmät sisältävät.

Tämä opinnäytetyö on teoriaa ja käytäntöä yhdistävä toimintatutkimus, jonka pyrkimyksenä on selvittää case-yrityksen kautta laitteiden vikojen ennustamisen mahdollisuudet kunnossapitojärjestelmään kerätyn datan avulla. Tutkimusten perusteella pohditaan lopuksi ratkaisuja kunnossapitojärjestelmätuotteen ja sen käytön kehittämiseksi, tukemaan ennustettavuuden parantamista osana muita kunnossapidossa käytettyjä datan keräyksen käsittelyn menetelmiä. Tässä opinnäytetyössä ei tutkita tai analysoida anturidataa tai muuta toiminnanohjausjärjestelmän ulkopuolista dataa.

1.3 Case-yritys

Tässä opinnäytetyössä on käytetty todellisen suomalaisen teollisuusyrityksen tuottamaa kunnossapitodataa, jota on tutkimukseen käytettävissä yhteensä 9 vuoden ajalta. Case-yrityksen kunnossapitotoiminta edustaa perinteistä mallia, jossa kunnossapitohenkilöstö työskentelee samassa yrityksessä, jonka tuotantomaisuutta kunnossapidetään ja käyttää yrityksen omaa kunnossapitojärjestelmää toiminnan ohjaukseen ja raportointiin. Lisäksi tämä yritys ostaa alihankkijoilta joitakin kunnossapidon palveluita, kuten kiinteistön ylläpitoa. Case-yrityksen kunnossapitotoiminta koostuu pääsääntöisesti koneille ja laitteille kohdistuvista jaksotetuista huolloista sekä niiden häiriökorjauksista. (Kunnossapitopäällikkö 2020.)

2 TUOTANTO-OMAISUUDEN HALLINTA

Tuotanto-omaisuuden hallinnalla tarkoitetaan yrityksen fyysisen omaisuuden hallintaa. Tämä omaisuuden hallinta sisältää kaikki ne toimenpiteet, joiden avulla omistaja pyrkii saamaan pitkällä aikajänteellä ja kestäväällä tavalla mahdollisimman hyvän taloudellisen tuloksen. Tuotanto-omaisuuden hallinta kattaa laitteiden kunnossapidon lisäksi tuotantojärjestelmiin liittyvät elinkaaren vaiheet investointisuunnitelmista aina käytöstä luopumiseen. Teollisuudessa tämä on osa yrityksen liiketoiminnan tavoitteiden tukemista tuotanto-omaisuuden hallinnan strategian mukaisesti (SAMP = Strategic Asset Management Plan). (Komonen 2019, 8.)

Tuotanto-omaisuuden hallintaan on käynnistynyt useampiakin standardointiprojekteja 2010-luvulla. Tämän tuloksena on julkaistu muun muassa 2014 ISO 55000, 55001 ja 55002 standardisarjat. Standardissa EN 16646 määritellään fyysisen omaisuuden hallinta seuraavasti: *”Organisaation koordinoitu toiminta, jolla hyödynnetään fyysisen omaisuuden arvo.”* (Komonen 2019, 9.)

Tuotanto-omaisuuden hallinnan tueksi on olemassa erilaisia EAM-tietojärjestelmiä (EAM = Enterprise Asset Management). EAM on tietojärjestelmä, jolla hallitaan omistajan fyysistä tuotanto-omaisuutta. Tähän omaisuuteen kuuluvat teollisuudessa kaikki koneet ja laitteet, jotka ovat joko suoraan tai epäsuoraan kytkettynä tuotantoon. Näitä ovat esimerkiksi tuotantokoneet ja -laitteet, siirto- ja kuljetuskalusto sekä rakennukset ja muu yrityksen infrastruktuuri. Lisäksi EAM-järjestelmällä hallitaan tuotanto-omaisuutteen liittyviä materiaalivirtoja, kuten varaosien varastohallintaa.

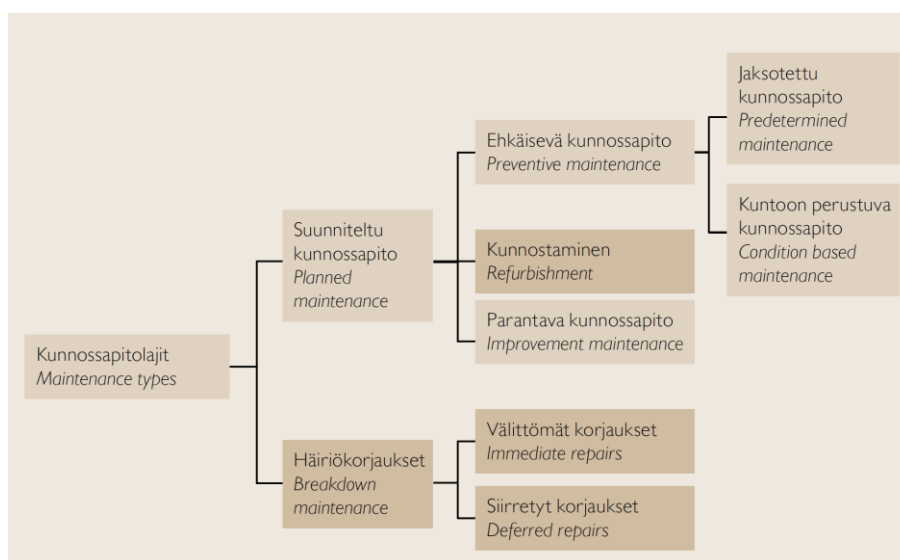
2.1 Kunnossapito

Kunnossapito on osa tuotanto-omaisuuden hallintaa. PSK 6201 -standardissa kunnossapito määritellään seuraavasti: *”Kunnossapito on kaikkien niiden teknisten, hallinnollisten ja johtamiseen liittyvien toimenpiteiden kokonaisuus, joiden tarkoituksena on säilyttää kohde tilassa tai palauttaa se tilaan, jossa se pystyy*

suorittamaan vaaditun toiminnon sen koko elinjakson aikana.” (PSK 6201 2011). Standardin ja vakiintuneiden käytänteiden mukaan kunnossapito alkaa fyysisen tuotanto-omaisuuden käyttöönotosta ja päättyy sen käytön loppumiseen.

2.2 Kunnossapitolajit

Kunnossapitolajit jakautuvat PSK 6201 standardin mukaan suunniteltuun kunnossapitoon ja häiriökorjauksiin (PSK 6201 2011). Kunnossapitolajit jakautuvat näiden päälaajien jälkeen kuvion 1 mukaisesti. Suunniteltuun kunnossapitoon lukeutuvat kaikki kunnossapidettäviin kohteisiin kohdistetut ennalta suunnitellut toimenpiteet, kuten ennakkohuollot, kunnonvalvonta ja parantava kunnossapito. Häiriökorjauksiin puolestaan lukeutuvat ne toimenpiteet, joiden tehtävä on kunnossapidon määrittämisen palauttaa kohde tilaan, jossa se pystyy suorittamaan vaaditun toiminnon.



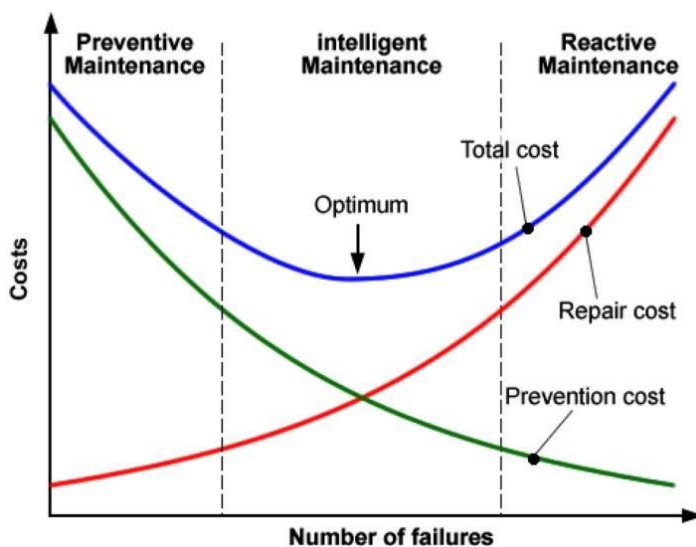
KUVIO 1. Kunnossapitolajit (PSK 6201 2011). Kunnossapitolajien PSK-standardin mukainen jaottelu

2.3 Kunnossapidon kustannukset

Kunnossapidon välittömät kustannukset muodostuvat palkoista, varaosista sekä ulkopuolisten palveluiden ostoista ja hankinnoista. Tuotannollisessa toiminnassa yritykset käyttävät kunnossapitoon noin 5 % liikevaihdosta (Kortelainen & Tennilä

2019, 27). Kunnossapidon välittömien kustannusten lisäksi on tuotannollisesta näkökulmasta kunnossapidon kustannuksia hyvin vaikea laskea, johtuen sen moniulotteisuudesta.

Kunnossapidon kustannusten optimointia voidaan havainnollistaa kuvion 2 mukaan. Perinteisessä jaottelussa kunnossapito on jaettu ennakoivaan kunnossapitoon (suunniteltu kunnossapito) ja reagoivaan kunnossapitoon (häiriökorjaukset). Kuvion vasen osio edustaa niin sanottua ylihuoltamista, joka minimoi vikaantumisten määrän, kasvattaen samalla suunnitellun kunnossapidon kustannuksia. Oikea puoli edustaa alihuoltamista, joka minimoi suunnitellun kunnossapidon kustannukset, mutta johtavat suureen vikaantumistiheyteen ja sitä kautta kalliisiin tuotannon menetyksiin, eli epäkäytettävyyuskustannusten nousuun. Optimitasossa kunnossapidon kustannukset ja epäkäytettävyyuskustannukset ovat yhteisvaikutukseltaan edullisimmat.



KUVIO 2. Kunnossapidon kustannusten optimointi (Tchakoua, Wamkeue, Slaoui-Hasnaoui, Tameghe & Ekemb 2013). Kunnossapidon kokonaiskustannusten ja kautuminen ennakoivassa ja reagoivassa kunnossapidossa.

3 KUNNOSSAPIDON TOIMINNAHOJAUSJÄRJESTELMÄ

Kunnossapidon toiminnanohjausjärjestelmät, joita on yleisesti kutsuttu CMMS-järjestelmiksi (CMMS = Computerized Maintenance Management Systems), tarkoitetaan järjestelmää, jolla hallitaan kunnossapidon toiminnanohjausta sekä kunnossapidon materiaalivirtoja. Kunnossapidon tietojärjestelmän pääasiallisena käyttäjänä on tuotantolaitoksen kunnossapitohenkilöstö, joka tuottaa kunnossapidon tietoa järjestelmään. Tuotanto-omaisuuden hallintajärjestelmä EAM (EAM = Enterprise Asset Management) sisältää usein kaikki CMMS järjestelmän toiminnot.

3.1 Käyttäjärhytät

Kunnossapidon tietojärjestelmien käyttäjärhytät määrättyvät roolikohtaisesti ja mahdollisesti myös toimipaikkakohtaisesti. Käyttäjärhytälle annetaan ryhmäkohtaiset käyttöoikeudet järjestelmään. Käyttöoikeuksia voidaan rajoittaa käyttäjärhytän kohtaisesti järjestelmän tietojen näkyvyyden, muokkausten, lisäysten ja poistojen osalta. Käyttöoikeudet perustuvat ryhmien rooleihin. Esimerkiksi tuotantohenkilöstöllä on usein vain rajalliset käyttöoikeudet järjestelmään. Tämän henkilöstön käyttöoikeudet rajoittuvat yleisesti vain vikojen raportointiin tuotannosta sekä niiden seuraamiseen. Kunnossapidon suunnittelijoilla ja tehokäyttäjillä on sen sijaan huomattavasti laajemmat oikeudet järjestelmän datan hallintaan. Roolikohtaisten käyttöoikeuksien lisäksi järjestelmään voidaan määritellä käyttäjille toimipaikkakohtaisia rajoituksia, eli mahdollisuudet käyttää järjestelmää vain yhden tai useamman toimipaikan osalta, oman roolinsa mukaisilla oikeuksilla.

Toiminnanohjausjärjestelmän käyttöoikeudet jakautuvat karkealla tasolla taulukon 1 mukaisesti. Taulukossa käyttöoikeusryhmän käyttöoikeudet sisältävät lähtökohtaisesti oman ryhmän oikeuksien lisäksi alempana mainitut käyttöoikeudet. Lisäksi ulkopuolisille yrityksille, jotka toimivat kunnossapidon alihankkijoina voidaan antaa rajalliset oikeudet järjestelmän käyttöön.

TAULUKKO 1. Toiminnanohjausjärjestelmän tavanomaiset roolipohjaiset käyttöoikeusryhmät VincitEAM-järjestelmässä

KÄYTTÄJÄRYHMÄT	KÄYTTÖOIKEUDET
Järjestelmänvalvojat	Käyttäjähallinta ja käyttöoikeuksien hallinta
Tehokäyttäjät	Laajat laitehallinnan ominaisuudet (poistot ja massatoiminnot)
Kunnossapidon suunnittelijat	Huoltosuunnitelmien hallinta, suppeat laitehallinnan ominaisuudet (lisäys, muokkaus)
Työnjohto	Kunnossapidon resursointi ja aikataulutus
Kunnossapitohenkilöstö	Töiden hallinta, materiaaliotot
Tuotantohenkilöstö	Vikailmoitusten laatiminen ja niiden seuranta

3.2 Master data

Master datalla tarkoitetaan dataa, joka on luonteeltaan pitkäikäistä ja hitaasti muuttuvaa (Ari Hovi 2015, blogi). Kunnossapidon toiminnanohjausjärjestelmän master datana toimii pääsääntöisesti järjestelmään perustettava tuotanto-omaisuus ja organisaation perustiedot näihin liittyvine tietoineen. Master dataa ovat muun muassa seuraavat toiminnanohjausjärjestelmän tiedot:

- Kunnossapidettävät kohteet (laitteet ja toimintopaikat)
- Kohteiden sijainti-, luokittelu- ja attribuuttitiedot
- Huoltosuunnitelmat
- Resurssiryhmät ja työntekijät
- Työlajit ja töiden luokitukset
- Varastot
- Nimikkeet
- Varaosatiedot
- Valmistajat

- Toimittajat

Master dataa usein myös integroidaan ulkopuoliseen järjestelmään joiltakin osin, jolloin sitä myös hallitaan toisaalla, esimerkiksi ERP-järjestelmässä. Näissä tapauksissa kunnossapidon toiminnanohjausjärjestelmä lukee kyseistä tietoa ulkopuolisesta integroidusta järjestelmästä.

3.3 Tapahtumadata

Kunnossapidon toiminnanohjausjärjestelmän tapahtumadatana toimii varsinaisen kunnossapidon toiminnan kautta järjestelmään tuotettu data. Tätä dataa on muun muassa:

- Vikailmoitukset
- Työt ja tehtävät
- Tuntikirjaukset
- Materiaalikirjaukset

Tapahtumadata kytkeytyy aina järjestelmän master dataan. Esimerkkinä työ, joka on tapahtumadataa, sisältää tiedon aina sen kohteesta, joka puolestaan on master dataa. Lisäksi tapahtumadata voi kytkeytyä myös toiseen tapahtumadataan. Esimerkiksi tuntikirjaustapahtuma sisältää tiedon työstä, johon se on kirjattu ja lisäksi kirjauksen tehneen työntekijän. Tässä tapauksessa ainoastaan työntekijä luokitellaan master dataksi ja tuntitapahtuma sekä työ tapahtumadataksi.

Kunnossapidon toiminnanohjausjärjestelmään raportoidaan kunnossapitohenkilöstön toimesta heidän tekemänsä vikakorjaus- ja huoltotapahtumat. Tapahtumadata sisältää paljon kunnossapitohenkilöstön järjestelmään syöttämää tekstimuotoista dataa sekä mahdollisesti valokuvia ja videoita tekstien tueksi.

4 KUNNOSSAPIDON TULEVAISUUS

Laaja-alaisen digitalisaation kehityksen myötä myös kunnossapito on voimakkaassa kehitysvaiheessa. Kunnossapitoa on ohjattu vuosikymmeniä lähinnä yksittäisillä vikailmoituksilla ja työpyynnöillä (häiriökorjaukset) sekä kalenteripohjaisilla huoltosuunnitelmissa. Ennen laajamittaisen digitalisaation hyödyntämistä, on laitteisiin kytkettyjen kunnossapitotietojen hallinta usein ollut rajoittunut laitteen käyttömäärien seurantaan tai laitteiden yksittäisten häiriöilmoitusten kytkemiseen kunnossapidon toiminnanohjausjärjestelmään. Näin tuotanto-omaisuuden hallintaan on saatu kytkettyä ensimmäiset käyttömäärään ja kuntoon perustuvat ohjaukset. Alla olevassa kuviossa 3 esitetään kunnossapidon maturiteetin kasvua, jossa tästä kalenteriaikaan pohjautuvasta kunnossapidosta ollaan siirtymässä kohti ennustavaa ja kognitiivista kunnossapitoa.



KUVIO 3. Kunnossapidon maturiteetin kasvu (Genesis Solutions 2020). Kunnossapidon maturiteettia kuvaava kuva, jossa kalenteripohjaisesta kunnossapidosta siirrytään asteittain kognitiiviseen kunnossapitoon, joka samalla kasvattaa kunnossapidon liiketoiminnan arvoa kulusta investoinniksi.

4.1 Ennustava kunnossapito

Reaktiivisesta ja kalenteripohjaisesta kunnossapidosta ollaan siirtymässä kasvavissa määrin ennustavaan kunnossapitoon, jossa laitteiden vikaantumista pyritään ennustamaan riittävällä luotettavuudella. Näin vikaantumiset voidaan ennalta välttää oikea-aikaisilla kohdistetuilla kunnossapitotoimilla.

Voimakas teknologinen kehitys ja erityisesti teollinen internet on voimistanut ennustavan kunnossapidon (predictive maintenance) mahdollisuuksia. Tässä ennustavassa kunnossapidossa hyödynnetään laitteilta kerättyä dataa, jonka avulla voidaan tehdä vikaantumisenennusteita. Näistä laitteista järjestelmiin kerättävä data voidaan jakaa karkealla tasolla kahteen, joita ovat anturi- ja laitteistodata sekä kunnossapitotoiminnassa syntynyt data. Laitteistodataa ovat muun muassa erityyppiset antureiden suurearvot raja-arvojen ylitystiedot ja käyntimäärätiedot.

Tehokkaasti toimiva kunnossapidon ennustaminen mahdollistaa kunnossapitotoimet jo ennen laitteen vikaantumista. Ennustavassa kunnossapidossa laiteyksilön vikaantumistodennäköisyys voidaan selvittää tietyn tulevaisuuteen viittavan aikajakson sisällä. Ennustavan kunnossapidon perusedellytyksiä on riittävä määrä laadukasta dataa, joka avulla ennustemalleja voidaan luoda. Vikaantumisten ennustamisessa käytetään usein montaa eri tietolähdettä, josta tiedot kerätään yhteen ja analysoidaan.

Toiminnanohjausjärjestelmään kerätty data on pääosin ihmisten sinne kirjaamaa dataa. Tämä data ei ole aina täysin reaaliaikaista ja siihen sisältyy inhimillinen näkökulma. Toiminnanohjausjärjestelmän käytettävyydestä ja saavutettavuudesta riippuen, datan laatu ja määrä voi vaihdella suurestikin. Ihmisten järjestelmään syöttämä data voi olla puutteellista ja siitä on voitu tehdä subjektiivisia päätelmiä. Toiminnanohjausjärjestelmään syötetty data on usein tarkoitettu palvelemaan vain omien töiden ja kunnossapito-organisaation toimintaa, eikä niinkään ennustavaa kunnossapitoa.

4.2 Kognitiivinen kunnossapito

Pitkään hiljaisen tiedon varassa toiminut kunnossapito on tulossa elinkaarensa päähän. Tämän henkilöriippuvaisen tiedon varassa toimivan kunnossapidon kehityspotentiaali on heikko tulevaisuuden haasteiden edessä. (Kaskikallio & Niittymaa 2017, 87.) Kunnossapidon asiantuntijoiden työpaikkavaihdosten ja eläköitymisten myötä organisaatiosta poistuu hiljaista tietoa eli tietopääomaa. Näin ollen järjestelmiin kerääntyneen tietopääoman tehokas hyödyntäminen nousee tulevaisuudessa suureen arvoon.

Kognitiivinen kunnossapito tulee yhdistämään useat kunnossapidon eri tietolähteet ja -järjestelmät. Kognitiivisessa kunnossapidossa yhdistyvät määrällinen, tilastollisilla menetelmillä analysoitu data sekä laadullisin menetelmin analysoitu data. Määrällistä dataa edustaa muun muassa laitteilta IoT-tekniikalla kerätty data. Laadullista dataa edustaa puolestaan ihmisen tuottama data, jota on kerääntynyt esimerkiksi kunnossapidon eri tietojärjestelmiin. Laadullinen kunnossapitodata sisältää pääosin ihmisen kokemuspohjaista osaamista, sisältäen tietoa esimerkiksi vikaantumisten syistä ja korjaustoimenpiteistä, sekä mahdollisista ennaltaehkäisevistä suosituksista. (Genesis Solutions 2020.)

Kognitiivisessa kunnossapidossa yhdistyvät järjestelmien automaattisesti tuottama ennustava kunnossapito sekä mahdolliset korjaus- ja huoltoehdotukset. Lisäksi järjestelmä tuottaa käyttäjälleen näihin liittyvän ohjeistuksen ja lisädokumentation, kuten käyttöohjeet. Tämä vähentää kunnossapidon henkilöriippuvuutta, kun kunnossapidon asiantuntijalle välitetään tietoa esimerkiksi tiettyä laiteyksilöä uhkaavasta vikaantumisesta sekä huoltosuosituksesta ja vaihdettavista varaosista kuvineen ja ohjeineen. (Kaskikallio & Niittymaa 2017, 93–95.)

4.3 Toiminnanohjausjärjestelmän rooli

Toiminnanohjausjärjestelmään kunnossapitäjien toimesta syöttämä data sisältää mitä suurimmassa määrin kunnossapitohenkilöstön omaa päättelyä ja tietotaitoa, jota voidaan jälkikäteen hyödyntää. Tämä järjestelmään tallennettu substanssi-osaamisen data on usein myös hyvin otsikoitua, joten sitä voidaan hyödyntää

koneoppimisen malleissa. Edellisessä kappaleessa mainitun hiljaisen tiedon tehokas hyödyntäminen aiheuttaa kasvavia muutospaineita järjestelmien toiminnalle ja niiden käytön tehokkuudelle.

Useimmat Suomen markkinoilla olevat toiminnanohjausjärjestelmät ovat perusratkaisultaan jopa vuosikymmeniä vanhoja. Aikojen saatossa niiden ominaisuuksia on muutettu vastaamaan paremmin nykyajan tarpeisiin. Erityisesti järjestelmien käyttöliittymät ovat kehittyneet vastaamaan nykyajan käyttötottumuksiin. Tästä on esimerkkinä mobiilisovellukset, joita järjestelmien käyttöön on kehitetty.

Kunnossapidon toiminnanohjausjärjestelmien taustalla olevat tietorakenteet ovat kuitenkin useimmissa tapauksissa edelleen samoja tai lähes samoja kuin ne olivat aikanaan ensimmäisessä markkinoille saatetussa versiossa. Näin ollen järjestelmiin syötetty data ei aina ole parhaalla mahdollisella tavalla optimoitu tukemaan ennustavaa ja kognitiivista kunnossapitoa.

Järjestelmän tietorakenteiden muuttaminen jälkikäteen aiheuttaa uusien ohjelmistoversioiden toimitusten myötä runsaasti konversiotyötä. Konversioissa järjestelmän olemassa oleva tietosisältö, jota voi olla jopa miljoonia rivejä, pitäisi sovittaa uuteen tietorakenteeseen. Lisäksi järjestelmät ovat usein integroitu muihin tietojärjestelmiin, joka hankaloittaa konversion tekemistä entisestään, sillä myös liitettyjen järjestelmien tulee aina olla yhteensopivat muutetun tietorakenteen kanssa.

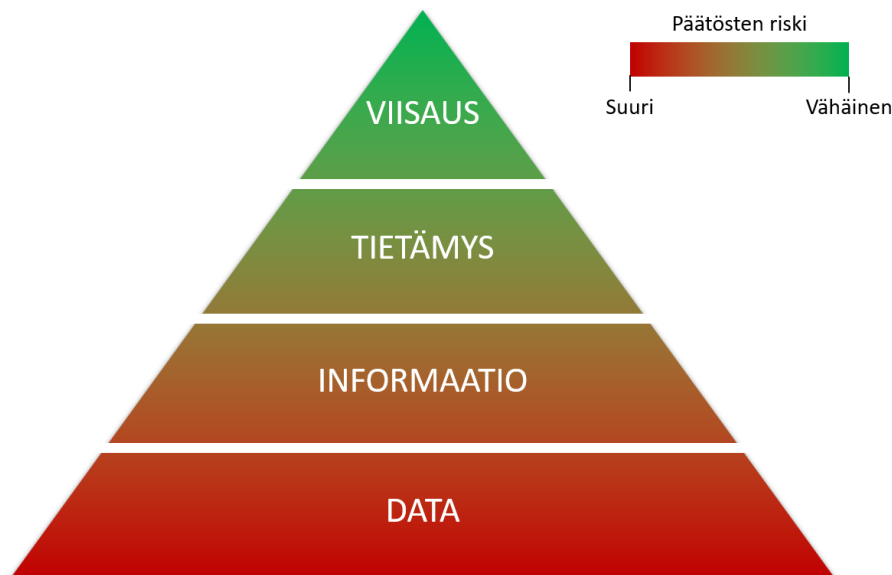
5 DATA-ANALYTIikka

Kunnossapidossa data-analytiikalla on merkittävä rooli menneen toiminnan arvioinnissa sekä toiminnan suunnittelussa. Toiminnan kannalta merkittävät päätökset tukeutuvat usein toiminnasta kerättyyn ja analysoituun dataan. Analysointi toteutetaan useimmiten erilaisten analysointityökalujen avulla, joilla tuotetaan visualisointeja järjestelmän käyttäjille ja kunnossapidon johdolle.

Toiminnanohjausjärjestelmään kerätystä datasta voidaan siis tuottaa käyttäjille informaatiota. Tämä tuotettu informaation on yleisimmin järjestelmään kerätyn datan esikäsittelyä ja visualisointeja, jotka esitetään käyttäjille raporttien muodossa. Näin ollen kunnossapidossa käytetty data-analytiikka on usein menneen toiminnan visualisointia käyttäjien nähtäville. Tulevaisuuteen viittaavat ennusteet rajoittuvat tässä perinteisessä kunnossapidon analytiikassa pääosin tulevien, ennalta aikataulutettujen huoltojen visualisointiin.

5.1 Data-analytiikan tasot

Data-analytiikan vaiheet kerätystä datasta päätöksenteon viisauteen, voidaan jakaa karkeasti neljään portaaseen. Nämä portaat voidaan visualisoida DIKW-pyramidilla (DIKW = Data, Information, Knowledge, Wisdom) kuvion 4 mukaan. Datasta jalostetaan informaatiota lisäämällä kerättyyn dataan konteksti, jolloin data voidaan esittää visuaalisessa muodossa esimerkiksi erilaisten raporttien avulla. Lisättäessä saatuun informaatioon sen merkitys, päästään tietämyksen tasolle, jolle ominaista ovat informaatiosta saadut merkitykselliset havainnot, opit ja vertailut. Tällä tasolla vaaditaan erityisesti myös kontekstisidonnaista tietotaitoa. Pyramidin huipulla on viisaus, jossa tietämys on jalostettu päätöksenteon tueksi. Ylimmällä tasolle päästyä, saadaan mukaan päätöksenteon viisaus, jossa voidaan esimerkiksi saada riittävä ymmärrys eri päätöksien vaikutuksesta lopputulokseen. Mitä ylemmäksi portailla päästään, sitä vähemmän päätöksenteot sisältävät riskejä.



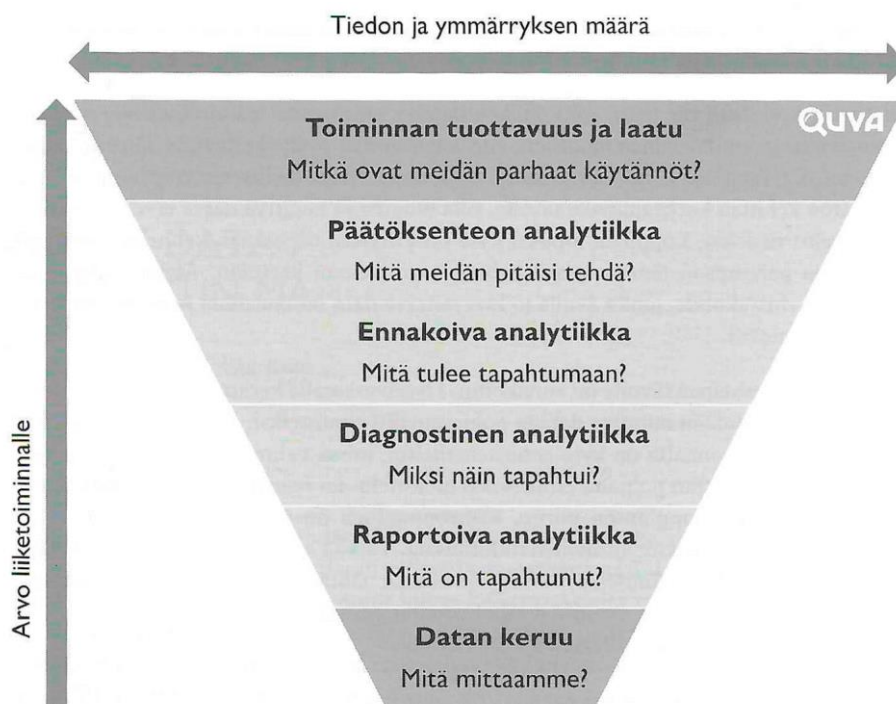
KUVIO 4. Data-analytiikan pyramidi. Pyramidi, jossa alimmasta datasta siirrytään asteittain viisauteen, joka samalla pienentää päätöksenteon riskejä.

5.2 Älykäs data-analytiikka

Keskeinen tiedon hyödyntämisen tavoite on ennakointi. Älykäs data-analytiikka hyödyntää usein erilaisia koneoppimisen malleja, kuten luokittelu- ja regressiomalleja, jotka ovat osa tekoälyn syväoppimisen tekniikoita. Älykäs data-analytiikka tulee olemaan merkittävässä asemassa, datan määrän lisääntyessä. (Ackerman & Ruusuvaori 2017, 125.)

Älykäs data-analytiikka vastaa suurten datamassojen tehokkaan käsittelyn avulla myös henkilöriippuvuuteen liittyviin haasteisiin. Henkilösidonnaisen tietopääoman hyödyntäminen, ja yhdistäminen suurten datamassojen tehokkaaseen käsittelyyn, mahdollistaa analytiikan portailla nousemisen älykkään analytiikan keinoin. Opetettaessa analytiikkaa systemaattisesti, ottamalla mukaan toiminnasta ja laitteista kerättyä dataa, saadaan päätöksenteon tueksi enemmän automaattisesti tuotettuja raportteja ja jopa ehdotuksia päätöksien tekemisestä sekä niiden lopputuloksien vaikutuksista liiketoimintaan.

Tiedon ja ymmärryksen määrää voidaan suhteuttaa liiketoiminnan arvoon. Analytiikan arvo liiketoimintaan perustuu mallin kykyyn vastata haluttuihin kysymyksiin. Kuviossa 5 on visualisoitu analytiikan portaavat, jossa kuvataan liiketoiminnan arvon kehitys, suhteessa tiedon ja ymmärryksen määrään eri portailla.



KUVIO 5. Analytiikan pyramidi (Ackerman & Ruusuvuori 2017). Pyramidissa datan keruusta siirrytään asteittain toiminnan tuottavuuden ja laadun parantamiseen. Samalla tieto ja ymmärryksen määrä sekä liiketoiminnan arvo kasvaa.

5.3 Ennakoiva analytiikka

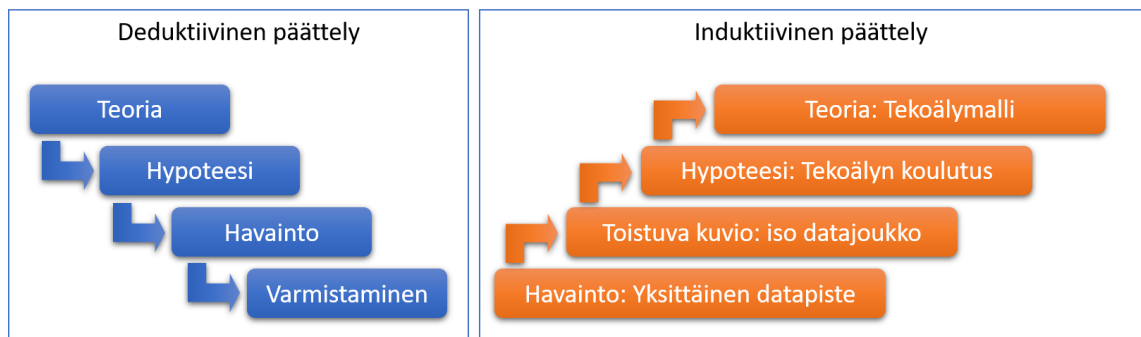
Ennakoivalla analytiikalla pyritään mahdollisimman hyvään ja tarkkaan tapahtumien ennakointiin. Datan prosessointitehon sekä varastointi- ja muistikapasiteetin voimakas kasvu on mahdollistanut ennakoivan analytiikan kehittymisen. Ennakoivan analytiikan ytimessä onkin usein tehokas opetettu koneoppimismalli. Edellytyksenä tehokkaalle ennakoivalle analytiikalle on suuri määrä laadukasta dataa sekä laaja-alaista liiketoiminnan ymmärtämistä syy-seuraussuhteiden opettamiseen tekoälymalleille. (Ackerman & Ruusuvuori 2017, 126–127.)

kunnossapito-organisaatiossa datan tehokas analysointi, sen alemmillakin kehitysportailla, toimii hyvänä esiasteena laajempaan analytiikan kehittämiseen tekoälyn avulla. Vaikka teknisten järjestelmäratkaisujen avulla tehty automaattinen analysointi päättyisikin informaatioportaalle, on siitä hyötyä jatkokehityksen kannalta. Keräämällä ja analysoimalla dataa, organisaatiossa ollaan tietoisia mitä dataa kerätään ja miten kerättyä dataa voidaan hyödyntää. Toisaalta näissä organisaatioissa on ymmärrys myös siitä, mitä dataa puuttuu ja mitä tarvittaisiin lisää ennakoivan analytiikan kehittämiseksi. Data-analytiikka ja tekoäly tukevat teknologioina toinen toistaan (Hovi 2020).

6 TEKOÄLY

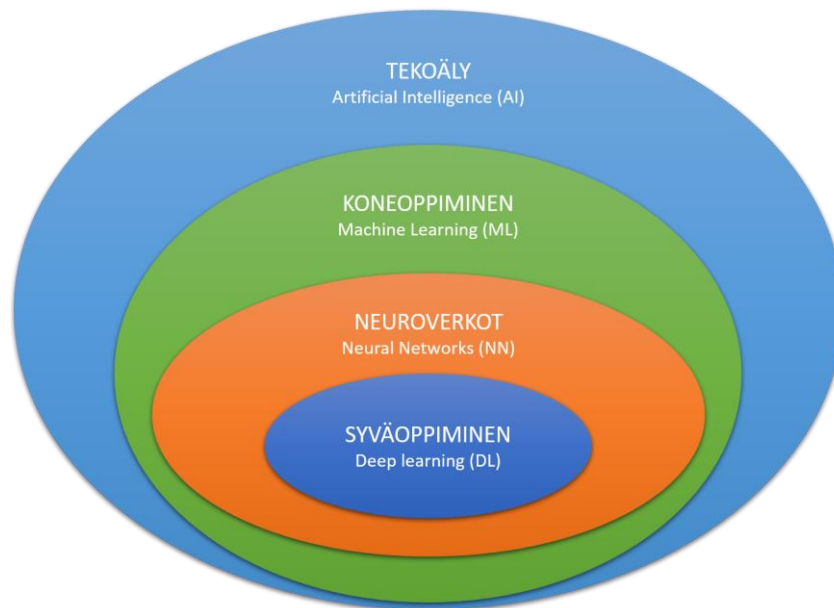
Tekoäly on pohjimmiltaan ohjelmointia, joka perustuu erilaisiin matriiseihin, derivointiin sekä tilastollisiin todennäköisyyksiin. Tekoälystä käytetään yleisesti lyhennettä AI, joka tulee englannin kielen sanoista *Artificial Intelligence*. Tekoälyn sovellutuksissa tietokoneen laskenta perustuu pääpiirteiltään samaan päättelyketjuun, kuin mitä tapahtuu ihmisen aivoissa. Tästä tulee nimitys tekoäly. (Siukonen & Neittaanmäki 2019.) Vakiintunutta määritelmää tekoälylle ei toistaiseksi ole olemassa. Tietokirjailija Antti Merilehto (2018, 18) kuvaa tekoälyä sellaisena koneen toimintana, joka ihmisen suorittamana olisi älykästä.

Perinteisessä sääntöpohjaisessa ohjelmoinnissa, ohjelmalle luodaan säännöt, jonka mukaan ohjelmoinnin tuottamaa lopputulosta verrataan havaintoihin ja varmistetaan vastaako havainnot ohjelmoitua mallia. Tätä kutsutaan *deduktiiviseksi päättelyksi*. Tekoälypohjaisessa ohjelmoinnissa algoritmit sen sijaan oppivat datasta eli yksittäisistä havainnoista. Tästä päättelyketjusta käytetään nimitystä *induktiivinen päättely*. (Kananen & Puolitaival 2019, 31.)



KUVIO 6. Deduktiivinen ja Induktiivinen päättely. Päättelyn portaat kuvattuna deduktiivisessa ja induktiivisessa päättelyssä.

Tekoäly jakautuu useisiin osa-alueisiin. Yksi tekoälyn yleisimmistä osa-alueista on ML eli koneoppiminen (ML = Machine Learning). Kuviossa 7 on esitetty tekoälyn tasot syväoppimiseen saakka.



KUVIO 7. Tekoälyn tasot, syväoppiminen. Syväoppiminen osana tekoälyä.

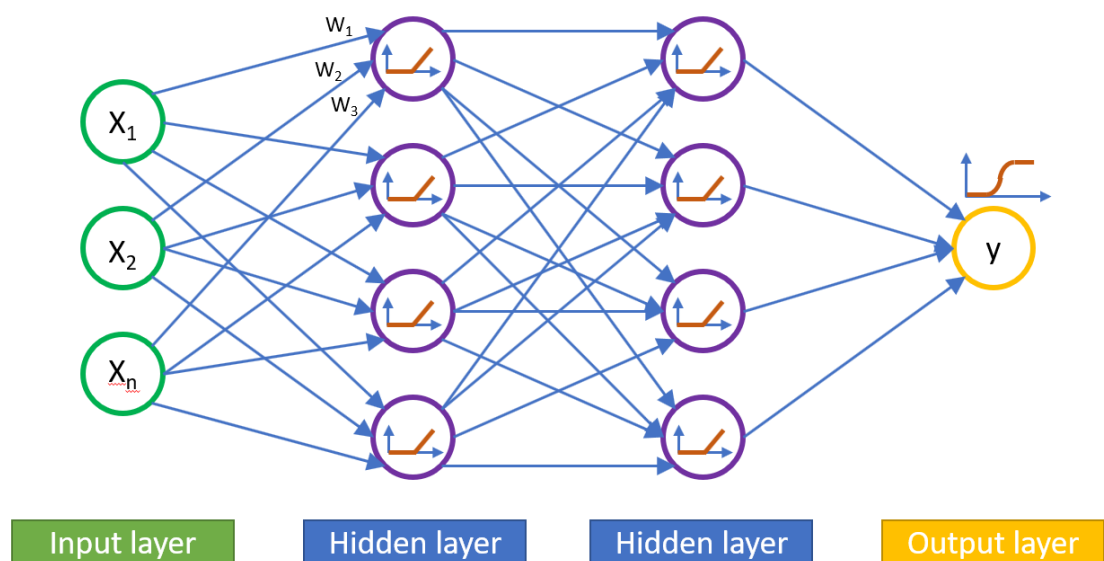
6.1 Koneoppiminen

Koneoppiminen on yksi tekoälyn sovellutuksista. Koneoppimisessa koneelle ei määritetä algoritmia jokaista tilannetta varten erikseen, vaan kone oppii datasta itsenäisesti (Merilehto 2018, 19). Kun malli on opetettu riittävällä tarkkuudella ja validoitu, voidaan se ottaa käyttöön uuden datan kanssa.

Ohjatussa koneoppimisen mallissa koneelle esitellään valmiita esimerkkejä, joihin on liitetty myös vastaus. Ohjatun ANN-koneoppimisen datasetti koostuu usein riippumattomista muuttujista X ja riippuvasta muuttujasta y . Data – vastaus -pareja tarvitaan suuri määrä, jotta malli voi oppia annetusta datasta. Tätä datasettiä kutsutaan *koulutusdataksi*. Kun malli on koulutettu, sille esitellään datasetti, jossa on pelkät riippumattomat muuttujat X . Tätä datasettiä kutsutaan *testausdataksi*, josta koulutettu malli pyrkii ennustamaan vastaukset y . Saatua ennustetta verrataan tunnettuihin vastauksiin, jolloin voidaan päätellä opetetun mallin hyvyys. (Kananen & Puolitaival 2019, 46–49).

6.2 Neuroverkot

Keinotekoiset neuroverkot ANN (ANN = Artificial Neural Network) koostuvat useista toisiinsa kytketyistä neuroneista, jotka opetetaan isolla datamäärällä. Neuroverkon toiminta perustuu lähtökohdiltaan ihmisen aivoissa tapahtuvan päättelyketjun jäljittämiseen. Neuroverkoissa neuronit toimivat itsenäisinä koneina, jotka ottavat sisään tietoa, prosessoivat sitä sekä välittävät tiedon eteenpäin seuraavalle neuronille. (Merilehto 2018, 47–52.)



KUVIO 8. Esimerkki syvästä neuroverkosta. Sisääntulokerros, kaksi piilotettua neuronikerrosta ja ulostulokerros.

Kuviossa 8 on esitetty neuroverkko, jossa vasemmalla puolella on sisääntulokerros, välissä piilokerrokset ja oikealla ulostulokerros. Piilokerrosten määrä määrittelee neuroverkon syvyyden. Neuroniyhteyksillä on jokaisella oma painokerroin, jotka ovat kuvassa esitettynä W_1 , W_2 ja W_3 . Tilan säästämiseksi kuvassa, painokertoimet ovat merkittynä vain yhden neuronin sisääntulon osalta. Kuvassa neuronien yhteydessä on valittu aktivaatiofunktio. Neuronin toiminta perustuu aktivaatiofunktion toimintaan sisään tulevien arvojen ja painokertoimien perusteella. Neuroverkon äly on painokertoimissa (Kananen & Puolitaival 2019, 129.)

Neuroverkkojen avulla on mahdollisuus löytää tehokkaasti asioiden väliset yhteydet. Ohjatussa koneoppimisessa neuroverkon oppiminen perustuu neuronien

painokertoimien säätämiseen niin sanotun vastavirta-algoritmin avulla. Tässä mallissa neuroverkon neuroneille annetaan aluksi satunnaiset painokertoimet, jolla neuroverkko laskee lopputuloksen (output). Tämän perusteella algoritmi laskee saadun tuloksen ja todellisten arvojen välisen eron. Tästä erosta saadaan laskettua mallin kustannusfunktio. Mitä pienempi ero on, niin sitä pienempi kustannusfunktion arvo on ja sitä parempi on opetettu malli. Yksi yleisimmin käytetyistä menetelmistä tämän kustannusfunktion arvon laskemiseksi on keskineliövirhe eli MSE-funktio (MSE = Mean Squared Error) kaavalla:

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2, \quad (1)$$

jossa n on datapisteiden määrä, \hat{Y} on mallin tuottama arvo ja Y on todellinen arvo.

Neuroverkossa kustannusfunktion arvo on tavoitteena saada mahdollisimman pieneksi. Tämä tapahtuu neuroverkon painokertoimia säätämällä, jossa yhtä painokerrointen säätökertaa kutsutaan epookiksi. Näitä epookkeja voi olla koulutusmallista riippuen jopa satoja.

6.3 Neuroverkon rakentaminen

Neuroverkko voidaan rakentaa esimerkiksi Python-ohjelmalla, joka perustuu avoimen lähdekoodin ohjelmaan. Pythonissa neuroverkko voidaan rakentaa käyttämällä valmiita Python kirjastoja, kuten Tensorflow-kirjastoa. Syvän neuroverkon runko rakennetaan antamalla verkolle tarvittavia parametreja. Näitä parametreja ovat:

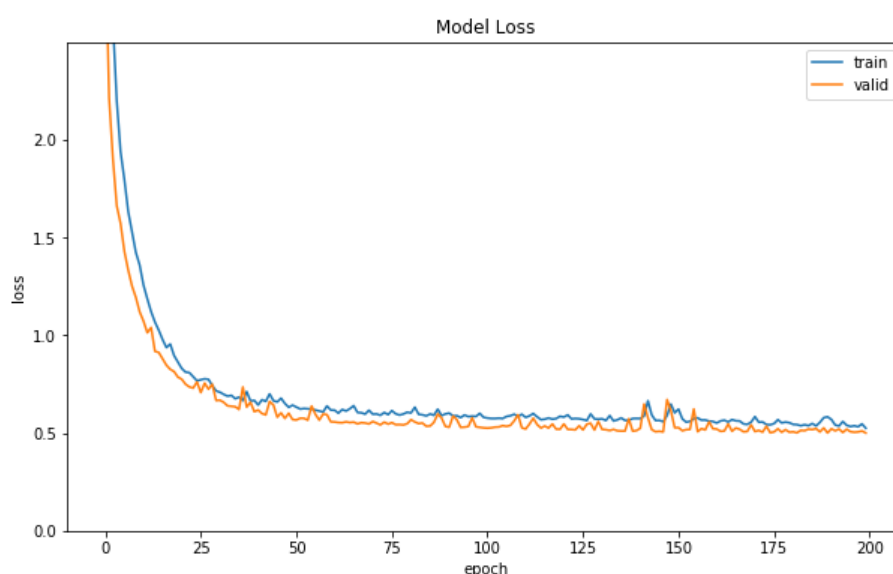
- Sisääntulokerroksen neuroneiden määrä
- 1. piilokerroksen neuronien määrä ja aktivaatiofunktio
- 2. piilokerroksen neuronien määrä ja aktivaatiofunktio
- ...
- n. piilokerroksen neuronien määrä ja aktivaatiofunktio
- Ulostulokerrosten määrä ja aktivaatiofunktio

Muita neuroverkon tarvitsevia parametreja ovat esimerkiksi optimizer ja loss-parametrit.

6.4 Neuroverkon kouluttaminen ja optimointi

Neuroverkon kouluttamisessa neuroverkolle annetaan muun muassa epookkien eli koulutuskertojen määrä sekä tietueiden määränä yhden koulutuserän koko. Neuroverkon oppiminen perustuu painokertoimien säätämiseen, jossa malli pyrkii löytämään neuroneille parhaat mahdolliset painokertoimet, joilla luvussa 6.2 mainittu kustannusfunktion arvo saataisiin mahdollisimman pieneksi.

Neuroverkossa on lukuisia niin sanottuja hyperparametreja, joilla mallin oppimista ja lopputulosta voidaan optimoida. Optimointi on usein yritys ja erehdys -menetelmään perustuvaa parantelua, jossa parametrien säätämisen jälkeen tehdään uusi koulutuskierrös ja verrataan lopputulosta ennen säätöä saatuun lopputulokseen. Mallin optimoimisessa auttaa hyvin valittujen visualisointien käyttö, joilla voidaan havainnollistaa esimerkiksi mallin tarkkuuden ja kustannusfunktion arvon käyttäytyminen epookkien mukaan, kuten kuviossa 9 on esitetty. Hyväkään neuroverkko ei pysty ennustamaan absoluuttista totuutta, vaan kyseessä on aina todennäköisyyttä kuvaava malli (Kananen & Puolitaival 2019, 140).



KUVIO 9. Kustannusfunktion arvo voidaan laskea kouluttamisen yhteydessä. Kustannusfunktion arvon muutos opetus- ja testausaineistolla epookkien kasvaessa.

7 CASE: YRITYS OY

Case-study tutkimuksessa käytettiin suomalaisen elintarviketeollisuudessa toimivan yrityksen kunnossapitodataa. Kunnossapitodata on saatu suoraan käytössä olleen kunnossapitojärjestelmän tietokannasta, yrityksen kunnossapitopäällikön toimittamana.

7.1 Tietojärjestelmä ja datan valinta

Yrityksen kunnossapidon toiminnanohjausjärjestelmänä on vuoteen 2019 asti toiminut Powermaint. Tämä kunnossapidon järjestelmä on ollut tietokoneissa työasemaversiona, eikä siihen ole ollut saatavissa mobiilikäyttöliittymää. Powermaint järjestelmään on kerätty kunnossapitodataa vuodesta 2000 (Kunnossapitopäällikkö 2020). Tarkasteltavaksi dataksi valittiin vuodesta 2010 kerätty data, jolloin tutkimusta voitiin tehdä järjestelmään viimeisen kahdeksan vuoden aikana kerätystä datasta. Analysoitavaksi dataksi valittiin yksittäiset työtilaukset, suunnitellut huollot sekä laitteet, joihin työtilaukset ja huollot kohdistuivat. Datan tutkimuksessa ja esikäsittelyssä analysointia varten käytettiin Tableau Prep Builder -ohjelmaa (ver. 2020.3.3). Datan analysoinnissa käytettiin saman ohjelmistotuotteen Tableau Desktop -ohjelmaa (ver. 2020.3.3).

7.2 Yleiskatsaus dataan

Kunnossapitojärjestelmän tietokannasta saadun csv-data-aineiston lähtötutkimukset tehtiin varmistamalla aluksi relaatiot työtilausten, tehtävien ja kohteiden eli laitteiden välillä. Aineisto todettiin näiden osalta moitteettomaksi ja rakenteellisesti hyväksi. Epäselvät tapaukset käsiteltiin erikseen yrityksen kunnossapitopäällikön toimesta. Kokonaisuudessaan aineisto todettiin kelpolliseksi tarkempia analysointeja varten.

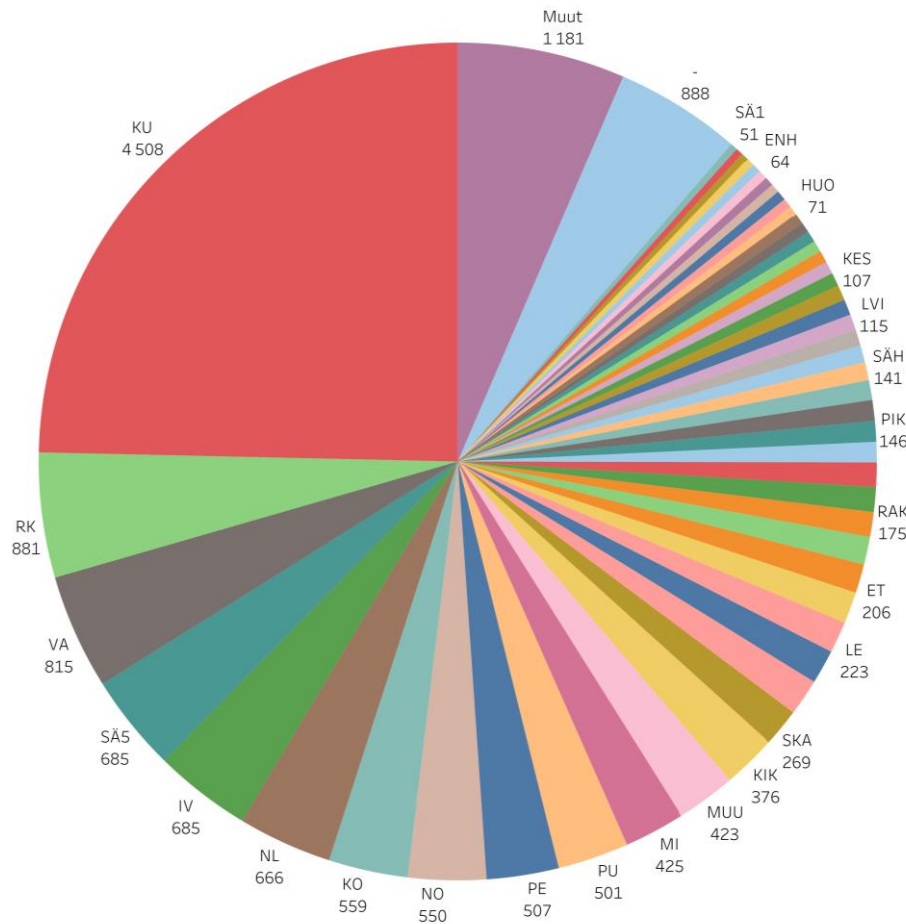
Työtilauksia oli aineistossa yhteensä 468 916 kappaletta ja laitteita 18 299 kappaletta. Työtilaukset ovat kohdistettuna laitteille joko suoraan tai tehtävien kautta.

Osa ennakkohuollon työtilauksista on jaettu tehtäviin, joihin on erikseen määritetty laite kullekin tehtävälle. Laitteille kohdistuvia työtilauksia tai tehtäviä oli kaikkiaan 1 461 318 kappaletta. Jäljempänä tässä raportissa kohteista ja laitteista tarkoitetaan samaa asiaa.

7.3 Kohteiden luokittelu

Järjestelmässä kohteiden luokittelevana tietona on ainoastaan laitetyyppi. Muuta laitteen luokitteludataa ei järjestelmästä ole saatavilla (Kunnossapitopäällikkö 2020). Kohteiden luokittelu on yksi koneoppimismallien tarkkuutta parantava parametri, joita neuroverkko hyödyntää painokertoimia asettaessa. Suurempi määrä luokittelevaa dataa, parantaa usein ennustemallien tarkkuutta.

Kohteiden luokittelussa ei olla käytetty standardeja tai muutakaan dokumentoitua ja vakiintunutta mallia. Luokittelujen peruslähtökohtana on ollut lähes yksinomaan hakutoimintojen helpottaminen (Kunnossapitopäällikkö 2021). Kaikista kohteista noin neljännes on kuljettimia (KU = Kuljettimet). Kuvion 10 Muut-luokka sisältää kaikki ne laiteluokat, joissa laiteyksilöiden määrä on alle 50 kpl. Näitä luokkia on yhteensä 76 kpl.



KUVIO 10. Kohteiden jakautuminen luokkiin. Kohteiden jakautuminen esitettyä ympyrädiagrammina.

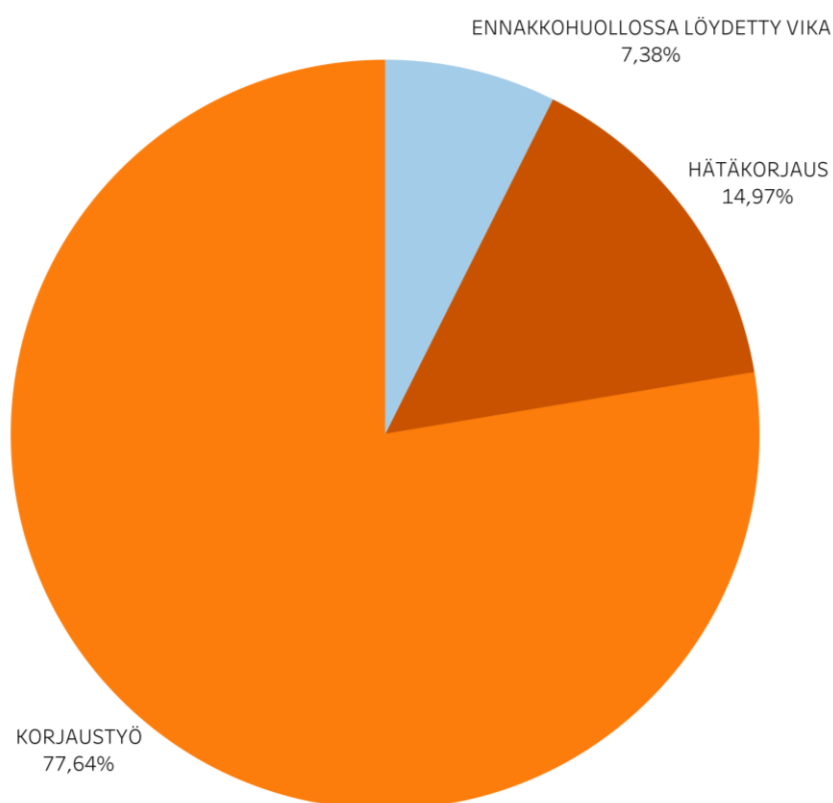
7.4 Työtilausten analysointi ja esikäsittely

Työtilaukset on luokiteltu työlajien mukaan. Luokittelu ei noudata standardeja. Sen sijaan luokat ovat aikanaan määritelty vapaasti liiketoiminnan ja kunnossapidon omista lähtökohdista (Kunnossapitopäällikkö 2021). Työt voidaan ole-massa olevan datan perusteella jaotella karkeasti standardin mukaisiin luokkiin. PSK-standardin mukaiset pääluokat ovat suunniteltu kunnossapito ja häiriökor-jaukset, kuten kappaleessa 2.2 on esitetty.

Suunniteltuun kunnossapitoon määriteltiin vain Ennakkohuolto ja Vuosihuolto -työlajin työt. Sen sijaan häiriökorjauksille saman lajin sisälle ryhmiteltiin Hätäkor-

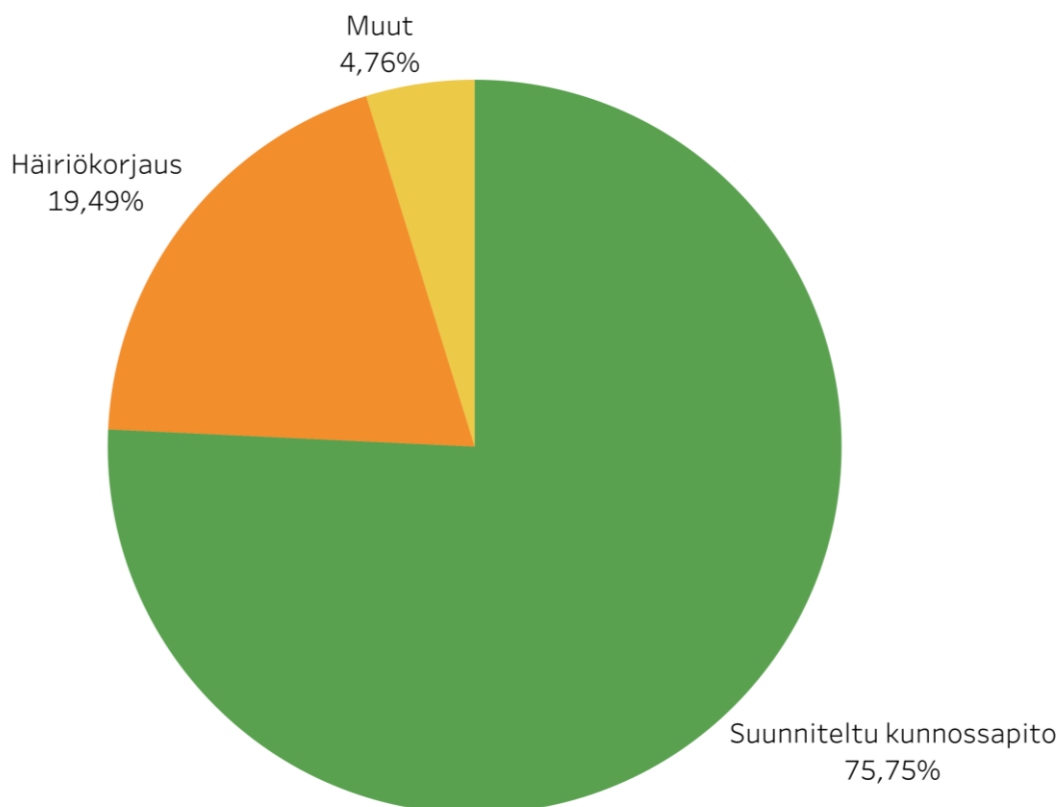
jaus- ja Korjaustyö -työlajit, joita ei käytännössä voida yksiselitteisesti erottaa toisistaan, ja joiden luokittelussa on käytetty kunnossapitoasentajien subjektiivista näkemystä (Kunnossapitopäällikkö 2020). Lisäksi ennakkohuollon löytämä vika lisättiin Häiriökorjausluokkaan, koska tässä tapauksessa kyseessä on varsinaisesti vikakorjaus.

Häiriökorjausluokan työt jakaantuivat Korjaustyöhön (78 %), Hätkäkorjauksiin (15 %) ja Ennakkohuollossa havaittuihin vikoihin (7 %), kuten kuviossa 11 on esitetty.



KUVIO 11. Häiriökorjausten jakautuminen. Häiriökorjausten jakautuminen esitettyinä ympyrädiagrammina.

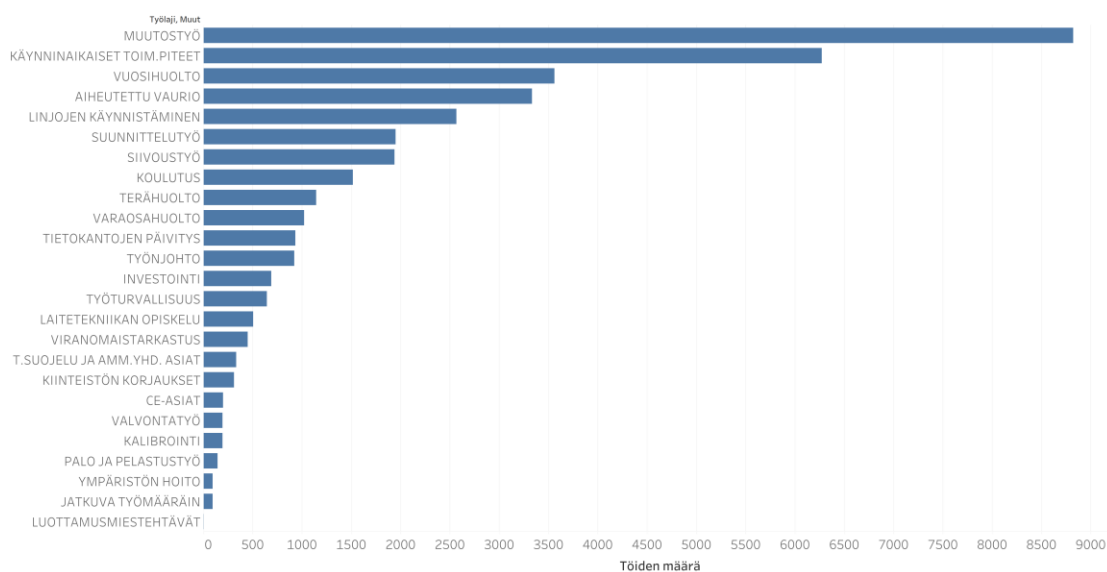
Uuden ryhmittelyn jälkeen työt jakautuvat työlajeittain kuvion 12 mukaisesti Suunniteltuun kunnossapitoon (76 %), Häiriökorjauksiin (19 %) ja Muihin (5 %).



KUVIO 12. Työtilausten luokittelu standardin mukaisille päätasoille. Työtilaukset luokiteltu standardin mukaisiin päätasoihin ympyrädiagrammissa.

Tavoitteena oli määritellä datasta häiriöiden ennustettavuus. Häiriöiden määrää käytetään ennustemallissa riippuvana muuttujana. Ennustemallin riippumattomana muuttujana töiden osalta olivat suunnitellun kunnossapidon määrät valittuna ajanjaksona. Tämän lisäksi muista työlajeista oli syytä ottaa mukaan ne työlajit, jotka voivat vaikuttaa ennustettavuuteen. Näiden työlajien yhteispiirteinä olivat kohteelle vähintäänkin välillisesti PSK 6201 -standardin mukaiset kunnossapidon toimenpiteet, kuten kunnonvalvonta (PSK 6201 2011). Muut-lajin työt jakautuvat määrällisesti kuvion 13 mukaisesti. Näistä Muut-työlajin töistä luetaan pois ne työt, joilla ei oletettavasti ole työn luonteen tai niiden vähäisyyden vuoksi vaikutusta häiriökorjausten ennustettavuuteen. Nämä työt ovat vain välillisesti kunnossapitoon liittyviä töitä, ja joiden yhteys kunnossapitoon tulee lähinnä työn

suorittajien kautta, jotka työskentelevät kunnossapito-organisaatiossa (Kunnossapitopäällikkö 2021).

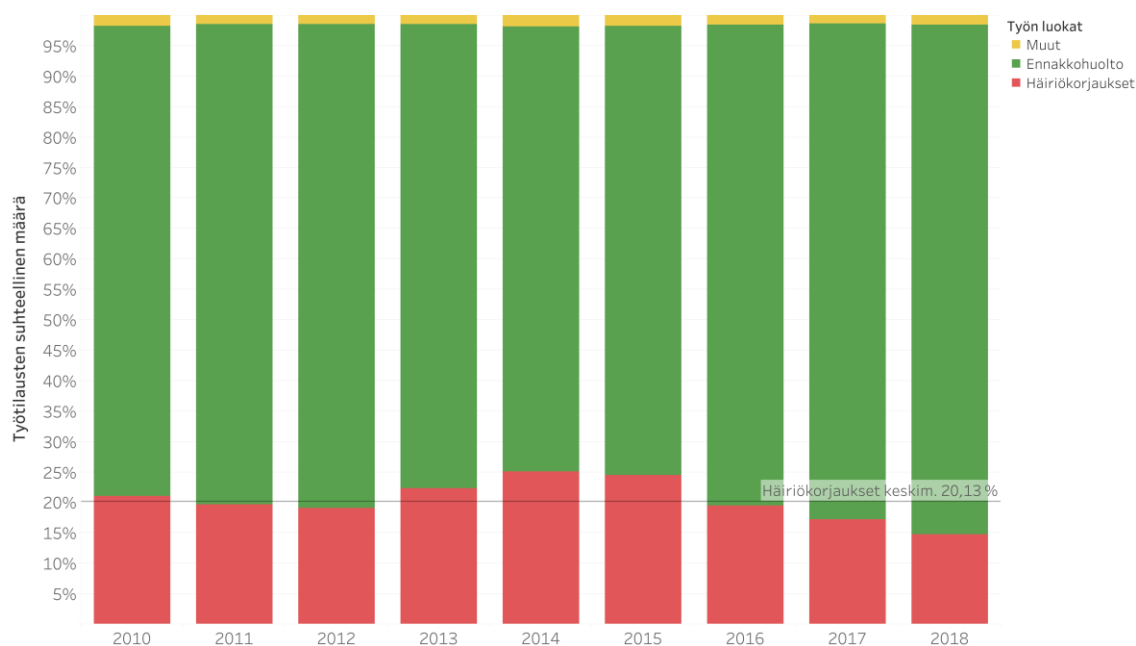


KUVIO 13. Muut-ryhmän työtilausten määrät luokittain. Muut-ryhmän työt esitettyinä palkkidiagrammissa.

Muut-ryhmän töistä analysointiin otettiin mukaan harkinnan seuraavat työläjit, jotka ryhmiteltiin uudeksi Muut-päätyölajiksi:

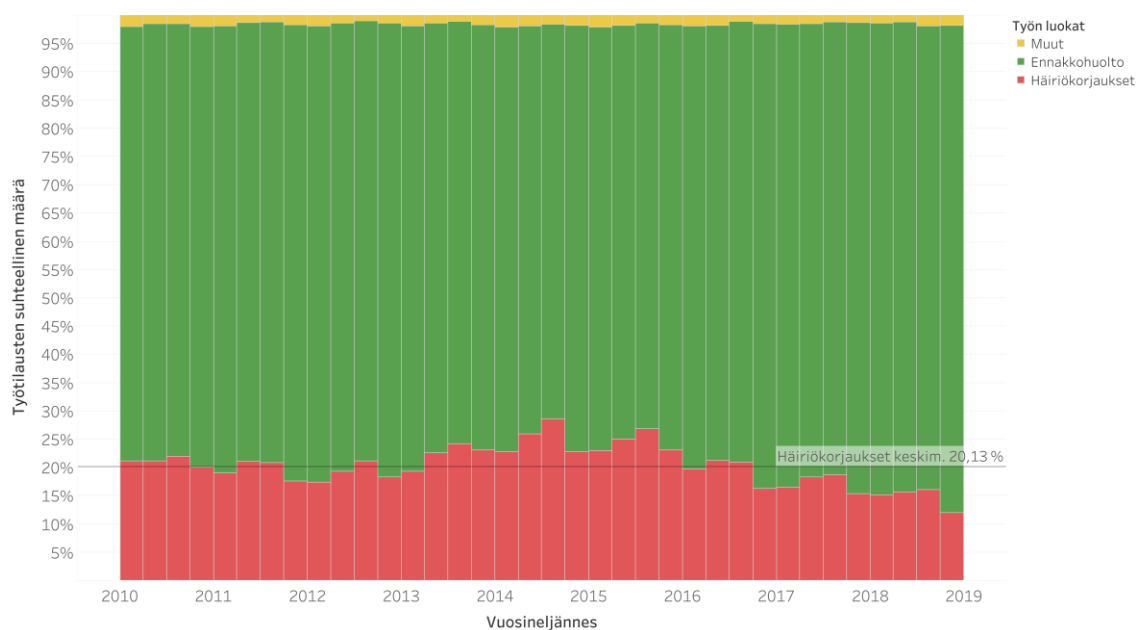
- Käynnin aikaiset toimenpiteet
- Suunnittelutyö
- Muutostyö
- Terähuolto

Uuden luokittelujen mukaiset päätyölajit sijoitettiin vuosittaiseen aikasarjaan kuvion 14 mukaisesti. Analysoinnista voidaan havaita, että työtilausten päätyölajien mukaista vuosittaista hajontaa ei ole havaittavissa merkittävässä määrin, vaan töiden jakautuminen eri vuosilla pysyy pääosin samalla tasolla.

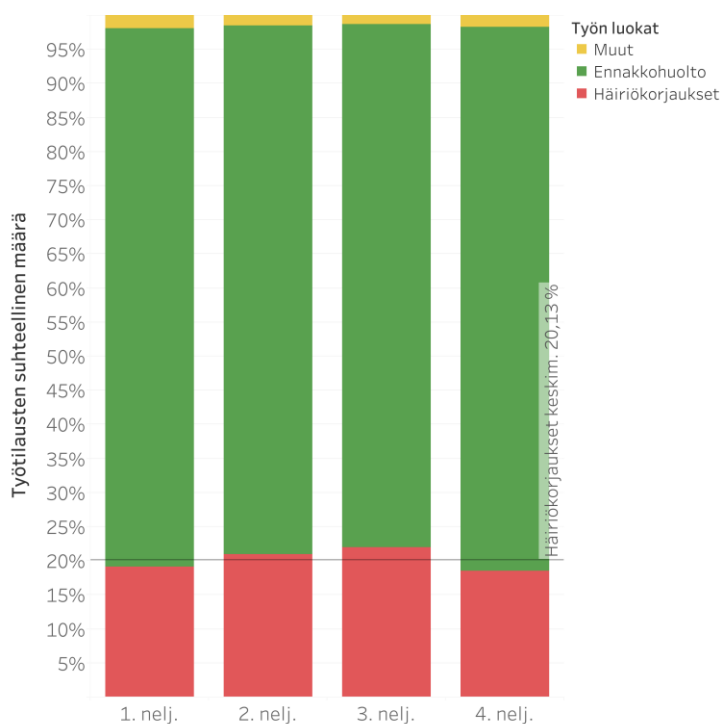


KUVIO 14. Työtilausten jakautuminen eri päätyölajeihin vuosittain. Pylväsdiagrammi työtilausten jakautumisesta vuosittain häiriöihin, suunniteltuihin töihin ja muihin töihin vuodesta 2010 vuoteen 2018.

Analysoitaessa työtilauksia vuosineljänneksittäin kuvioiden 14 – 16 mukaisesti, voidaan todeta, että vaihtelu ei ole merkittävää myöskään vuosineljänneksittäin. Ennustettavuuden kannalta vuoden merkitys on vähäistä, mutta vuosineljänneksillä voi toimialasta riippuen olla suurikin merkitys ennustettavuuteen. Vuosineljänneksien ennustettavuus perustuu valmistavan tuotannon syklisyyteen ja sesonkikausien mukaisiin tuotantomääriin, jolloin korkeammat tuotantomäärät vaikuttavat kasvavasti häiriökorjausten määriin. Case yrityksessä myös vakihenkilökunnan loma-ajat saattavat osaltaan näkyä toisen vuosineljänneksen korkeammassa vikaantumismäärissä (Kunnossapitopäällikkö 2021).



KUVIO 15. Työtilausten jakautuminen vuosien neljänneksittäin. Pylväsdiagrammi työtilausten jakautumisesta vuosien neljänneksittäin häiriöihin, suunniteltuihin töihin ja muihin töihin vuodesta 2010 vuoteen 2018.

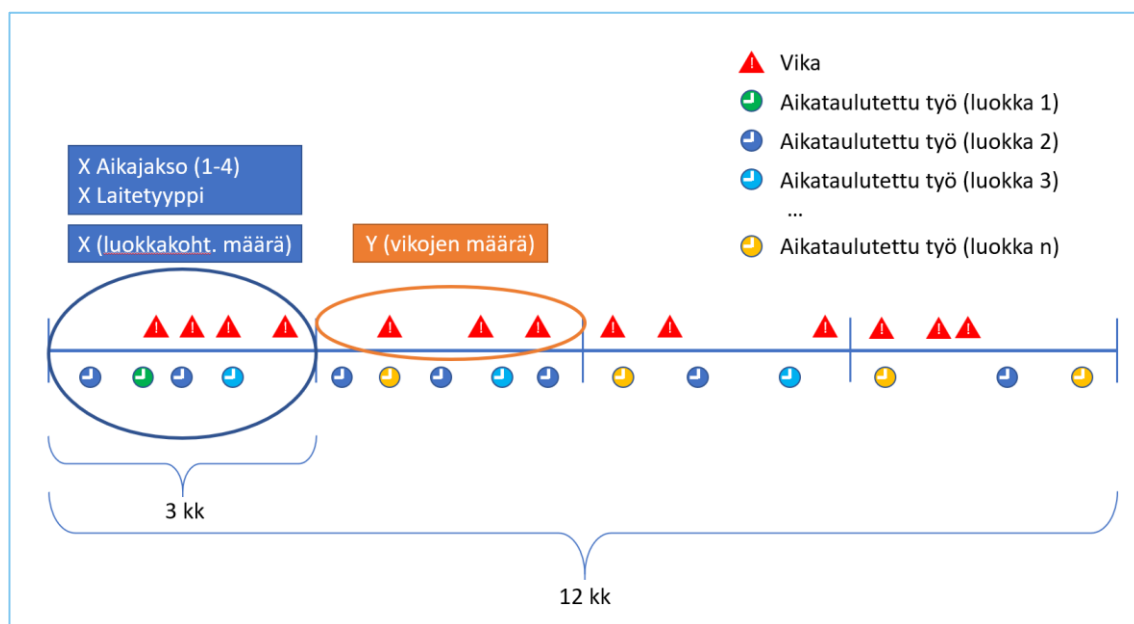


KUVIO 16. Työtilausten jakautuminen vuosineljänneksittäin. Pylväsdiagrammi työtilausten jakautumisesta vuosineljänneksittäin häiriöihin.

7.5 Datan sovittaminen koneoppimismalliin

Kyseiseen dataan voidaan soveltaa useita erilaisia koneoppimismalleja, kuten aikasarjaan perustuvaa RNN (RNN = Recurrent Neural Network). Tässä tutkimuksessa päädyttiin malliin, jossa regressioanalyysillä ennustetaan 3 kuukauden aikajakson vikaantumisten määrää edellisen vastaavan aikajakson datalla. Aikajaksojen valinnalla saadaan mukaan myös vuodenaikoihin perustuvaa luokittelevuutta. Koneoppimismallina käytettiin syväoppimiseen ja ohjattuun oppimiseen perustuvaa ANN-neuroverkkoa (ANN = Artificial Neural Network).

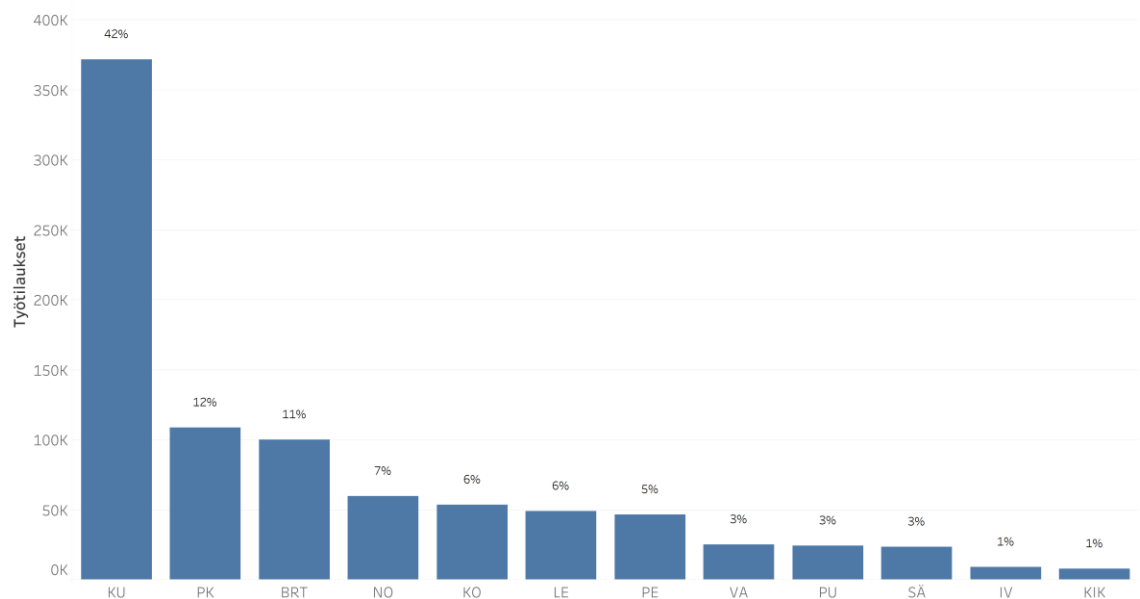
Alla olevalla kuvalla havainnollistetaan datan valinta koneoppimismallin rakentamiseksi. Neuroverkon riippumattomiksi ominaisuuksiksi (X) valitaan aikajakson toteutunut vikojen ja aikataulutettujen töiden määrä sekä laitetyyppi ja vuoden aikajakso (1–4). Riippuviksi muuttujiksi (y) valitaan seuraavan jakson vikojen määrä.



KUVIO 17. Datan valinta koneoppimismallille. Visualisointi, jossa on esitetty datan valinta koneoppimismallille, käyttäen yhden vuoden aikajanaa ja sen jaksoista vuosineljänneksiin.

7.6 Datan esikäsittely

Data valmisteltiin koneoppimismallille valitun tekniikan mukaisesti. Tässä datan esikäsittelyssä käytettiin Tableau Prep Builder -ohjelmaa (ver. 2020.3.3), jonka käsittelyn vaiheet ja mallit kuvataan liitteessä 1. Kunkin havaintopisteen eli työsuorituksen luokittelevaksi dataksi määriteltiin työluokka, laitetyyppi ja havaintopäivämäärän mukainen aikajakson luokka (1–4). Kvantitatiivisen analytiikan vuoksi, datasta suodatettiin pois kaikki ne laiteryhmitt, jotka sisälsivät alle 2000 työsuoritetta aikajakson aikana vuodesta 2010 vuoteen 2019. Lisäksi datasta suodatettiin pois niiden laitteiden sisältämät työsuoritteet, joissa kohteena ei ollut laiteyksilö vaan esimerkiksi ylätasen kohde, kuten kokonainen osasto tai solu. Laiteluokiksi valikoitui 12 luokkaa, joiden työsuoritteiden jakautuminen on esitetty kuviossa 18. Datan esikäsittelyn jälkeen rivien määräksi saatiin yhteensä 126 798 riviä. Ote esikäsitellystä datasta on esitetty liitteessä 2.



KUVIO 18. Työsuoritteiden jakautuminen valittujen laitteiden osalta. Pylväsdiagrammiesitys, jossa valittujen laiteluokkien työsuoritteiden jakautuminen on visualisoitu.

7.7 Neuroverkon rakentaminen opettaminen ja analysointi

Esikäsittelyn ja datan analysoinnin jälkeen datan käsittelyä jatkettiin avoimen lähdekoodin Spyder-sovelluksessa (ver. 4.0.1) Python 3 -ohjelmointialustaa käyttäen (ver. 3.7.6). Koneoppimismallissa on useita eri parametreja, joita säätämällä voidaan vaikuttaa oppimisen tuloksiin. Jäljempänä kuvatussa ohjelmakoodissa on useiden kokeellisten tulosten kautta saavutetut parhaimmat parametrit.

7.7.1 Kirjastojen tuonti ja datan käsittely

Ohjelmakoodin kirjastoina käytetään Pythonin valmiita avoimen lähdekoodin ohjelmakirjastoja:

```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler
from tensorflow.python.keras.models import Sequential
from tensorflow.python.keras.layers import Dense
from sklearn.metrics import r2_score
from sklearn.metrics import mean_absolute_error
```

Valitaan tutkittavat laitetyypit ja alustetaan mallin opettamisen raportoinnin lista:

```
types = ['PK', 'NO', 'KIK', 'PE', 'KIK', 'PU',
         'BRT', 'IV', 'PK', 'LE', 'VA', 'SÄ']
report = []
```

Seuraavat koodirivit suoritetaan silmukan

```
for t in types:
```

sisällä, jossa jokaisen laitetyypin havainnot opetetaan ja arvioidaan erikseen

Silmukan sisällä ensimmäisenä luetaan tietojoukko sisään ja määritellään se muuttujaan df. Tämän jälkeen täytetään Null-arvot 0:ksi ja järjestetään sekä indeksoidaan luotu data frame uudelleen:

```
df = pd.read_csv('failure_predictions_3month.csv', sep=';')
df.fillna(0, inplace=True)
df.sort_values(by=['Period'], inplace=True)
df.reset_index(drop=True, inplace=True)
```

Poistetaan luodusta data framesta kaikki muut objektityypit, lukuun ottamatta sil-
mukassa vuorossa olevaa tarkasteltavaa objektityyppiä:

```
df.drop(df[df['O_Type'] != t].index, inplace = True)
del df['O_Type']
```

Luodaan luokittelevasta tiedosta omat sarakkeet ja poistetaan yksi luoduista sa-
rakkeista välttämällä niin sanottu Dummy variable -ansa (Algosome 2020). Li-
säksi poistetaan tarpeeton luokittelun lähdesarake Qx:

```
dummies_qx = pd.get_dummies(df['Qx'], drop_first=True)
df = df.join(dummies_qx)
del df['Qx']
```

Järjestetään syntyneet sarakkeet uudelleen siten, että viimeiseksi sarakkeeksi
tulee riippuva muuttuja y:

```
cols = df.columns.tolist()
cols = cols[2:10] + cols[11:] + cols [10:11]
df = df[cols]
```

Määritellään riippumattomien muuttujien X, ja riippuvan muuttujan y sarakkeet:

```
X = df.iloc[:, :-1].values
y = df.iloc[:, -1].values
```

Jaetaan luodut tietojoukot X ja y mallin opetus- ja testiaineistoksi suhteessa 70/30
sekä muotoillaan y-tietojoukko kaksiulotteiseksi:

```
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size = 0.3, random_state=21)
y_train = y_train.reshape(-1, 1)
y_test = y_test.reshape(-1, 1)
```

Jossa *random_state* -arvolla määritellään kiinteästi opetus ja testiaineiston rivit.
Tämä helpottaa mallin optimointia, jolloin parametreja voidaan säätää käyttä-
mällä samoja testi- ja opetusaineistorivejä eri opetuskertojen välillä.

Skaalataan X- ja y-tietojoukot, käyttäen MinMaxScaleria:

```
scaler_x = MinMaxScaler()
X_train = scaler_x.fit_transform(X_train)
X_test = scaler_x.transform(X_test)
```

```

scaler_y = MinMaxScaler()
y_train = scaler_y.fit_transform(y_train)
y_test = scaler_y.transform(y_test)

```

jossa kaikki arvot skaalataan arvoihin nolasta yhteen.

7.7.2 Neuroverkon luominen

Kirjastojen tuonnin ja datan esikäsittelyn jälkeen voidaan ohjelmoida varsinainen neuroverkko:

```

model = Sequential()
model.add(Dense(units=25, input_dim=(X_train.shape[1]),
kernel_initializer='normal', activation='relu'))
model.add(Dense(units=30, kernel_initializer='normal',
activation='relu'))
model.add(Dense(units=30, kernel_initializer='normal',
activation='relu'))
model.add(Dense(units=1, kernel_initializer='normal',
activation='relu'))
model.compile(optimizer='adam', loss='mse',
metrics=['mse', 'mae'])

```

jossa neuronien määrä ensimmäisellä tasolla on 25 ja seuraavilla kahdella tasolla 30. Aktivaatiofunktiona käytetään ReLu-funktiota. Optimizeriksi valittiin *adam* ja loss-funktioksi *mse* (*mse* = Mean Squared Error). Metriikaksi valittiin *mse* ja *mae* (*mae* = Mean Absolute Error).

Mallin opettaminen voidaan aloittaa seuraavalla ohjelmarivillä:

```

history = model.fit(X_train, y_train, epochs=50,
batch_size=16, verbose=1, validation_data=(X_test,y_test))

```

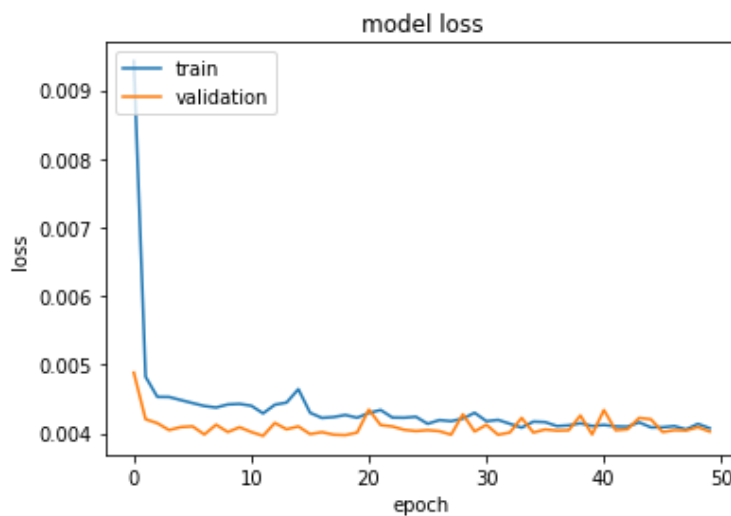
jossa opettavien erien suuruus on 16 tietuetta ja epookkeja 50 kappaletta.

7.7.3 Mallin tulosten raportointi ja visualisointi

Opetus ja tulokset visualisoitiin käyttämällä *matplotlib.pyplot* -kirjastoa:

```
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'validation'], loc='upper left')
plt.show()
```

josta tuloksena saadaan alla olevan kuvan mukainen visualisointi



KUVIO 19. Mallin oppiminen. Kustannusfunktion arvon visualisointi opetus- ja testidatalla.

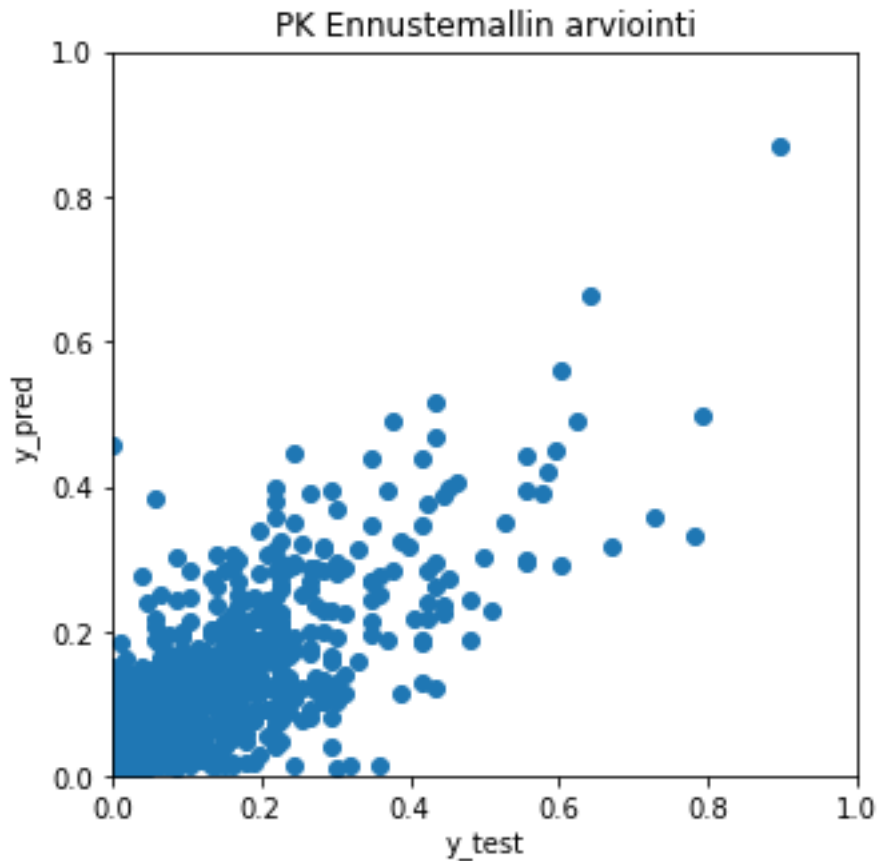
Testidatalla luotiin ennusteet:

```
y_pred = model.predict(X_test)
```

ja lisättiin ennusteet seuraavaan visualisointiin

```
plt.figure(figsize=(5,5))
plt.scatter(y_test,y_pred)
plt.title(t + ' Ennustemallin arviointi')
plt.xlabel('y_test')
plt.ylabel('y_pred')
plt.ylim(0,1)
plt.xlim(0,1)
plt.show()
```

jossa t on silmukassa vuorossa oleva laitetyyppi. Visualisoinnin tuloksena tulostuu pistekaavio, jossa vaaka-akselina on toteutuneet vikojen määrät ja pystyakselilla koneoppimismallin ennustamat vikojen määrät. Kuviossa 20 esitetään PK-laitetyypin ennustemallin toimivuus. Muiden analysoitujen laitetyyppien osalta tulokset kuvataan pistekaavioina liitteessä 3.



KUVIO 20. Ennustemallin arviointi PK laitetyyppi. Pistekaavio, jossa esitettyinä PK-laitetyypin häiriöiden toteutumat ja mallin tuottamat ennusteet.

Lopuksi report-data frameen lisättiin mallin oppimisen tulokset:

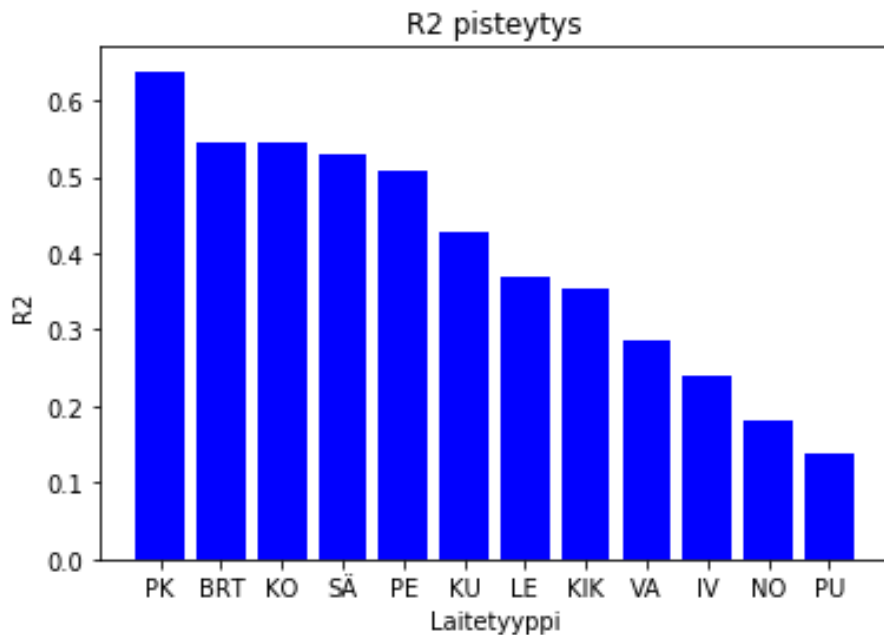
```
rep_r.append([t,r2_score(y_test, y_pred),
             mean_absolute_error(y_test, y_pred)])
```

jossa taulukkoon lisättiin silmukassa oleva laitetyyppi, r^2 -pisteytys ja keskimääräinen absoluuttinen virhe.

7.8 Mallin tulosten arviointi

Tuloksien arvioinnissa tarkoitus on aluksi löytää ne laitetyypit, joilla ennustemallin toimivuus on parhain. Tämän jälkeen voidaan pohtia mitkä tekijät mahdollisesti vaikuttavat mallin toimivuuteen paremmin kuin muilla laitteilla. Mallin toimivuutta voidaan analysoida ennustemallin lineaarisuuden korrelaatioasteella R^2 ja keskimääräisellä absoluuttisella virheellä.

Korkea korrelaatioaste R^2 tarkoittaa parempaa korrelaatiota todellisten arvojen ja ennustettujen arvojen välillä. Heikko pisteytys sen sijaan tarkoittaa heikkoa korrelaatiota eli huonoa ennustettavuutta. Kuvioista 21 selviää, että paras R^2 -pisteytys on PK-laitetyypillä (0,63) ja huonoin PU-laitetyypillä (0,13).

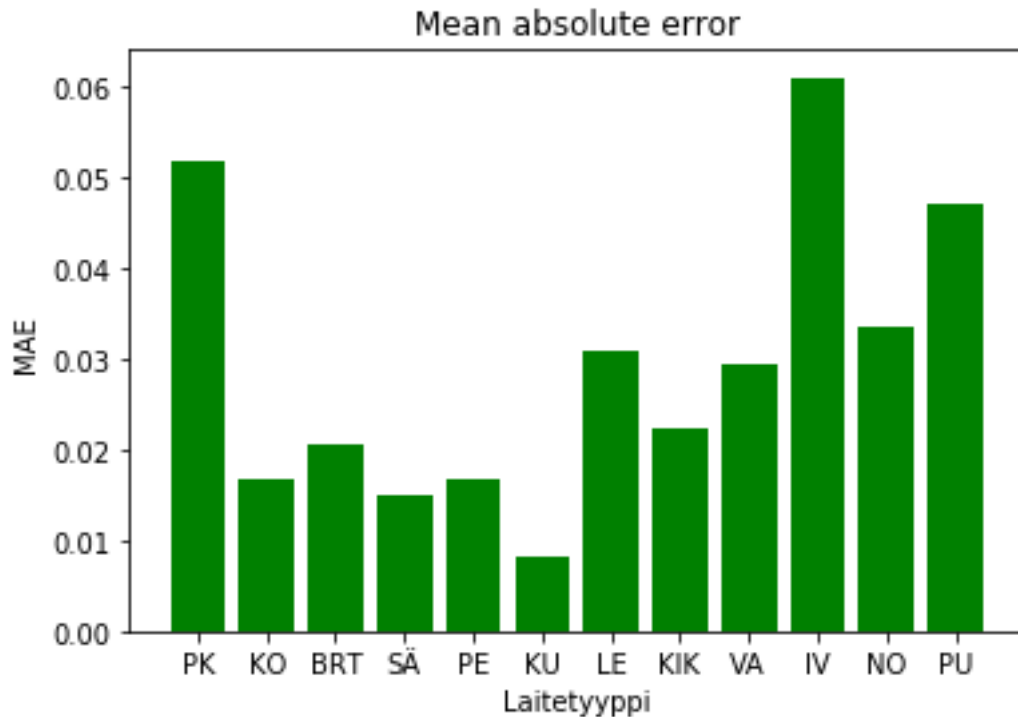


KUVIO 21. Ennustemallin R^2 -pisteytys. Pylväsdiagrammina visualisoitu R^2 -pisteytys laiteluokittain.

Alla olevassa kuviossa 22 on esitetty keskimääräinen absoluuttinen virhe, joka esittää laitekohtaisesti kaikkien havaintojen keskimääräisen virheen. Virhe lasketaan kaavalla:

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - x|, \quad (2)$$

jossa n on datapisteiden määrä, x_i on mallin tuottama ennuste ja x on todellinen arvo.



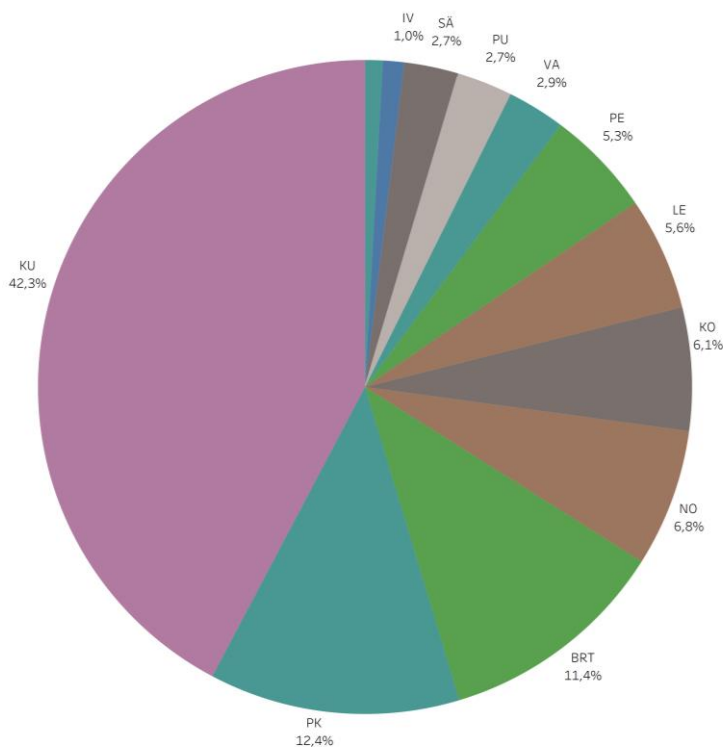
KUVIO 22. Ennustemallin keskimääräinen absoluuttinen virhe

Keskimääräistä virhettä arvioitaessa on otettava huomioon vikojen määrän vaihtelu eri laitetyyppien välillä. Parhaan R^2 -pisteytyksen omaavalla PK-laitetyypillä keskimääräinen virhe on huomattavan suuri, suhteessa pienemmän R^2 -pisteytyksen laitetyypeillä. Tämä selittyy pääosin suuremmalla vikojen määrällä eli vikaantumistaajuudella, jolloin ennusteiden virhe on alttiimpi isommille poikkeamille.

Korkeimman korrelaatioasteen omaavan PK-laitetyypin virhearvo on lähes yhtä suuri kuin matalimman korrelaatioasteen PU-laitetyypillä. Näiden kahden laitetypin pistekaavioista havaitaan, että ennustettavuus on kuitenkin huomattavasti parempi PK-laitetyypillä (vrt. liite 3). Saatujen tulosten perusteella R^2 arvo kuvaa ennustettavuutta paremmin kuin keskimääräinen absoluuttinen virhe, kun laitetyyppien välinen vikaantumistaajuus vaihtelee voimakkaasti.

Absoluuttista vikaantumisen ennustettavuutta arvioitaessa ennustemallin tulokset ovat parhaimmillaankin vain kohtalaisia. Samalla malli kuitenkin antaa arvokasta tietoa ennustettavuuden eroavaisuuksista laitetyyppien välillä, joka mahdollistaa tarkempien jatkopäätelmien ja ehdotusten tekemisen ennustettavuuden parantamiseksi.

Tarkasteltaessa datan määrää voidaan alla olevasta kuviosta todeta, että kuljetimiin (KU) on kohdistunut enemmän kuin neljännes kaikista työsuoritteista valittuna ajanjaksona vuodesta 2010 vuoteen 2020. Kun tätä laiteluokkakohdaista työsuoritteiden määrää verrataan R^2 -pisteytykseen, voidaan todeta, että ennustettavuus ei suoraan korreloi datan määrään.



KUVIO 23 Työtilausten määrä laitetyypeittäin. Ympyrädiagrammina esitetty työsuoritteiden määrä laitetyypeittäin.

Ennustettavuuden tulkintaa käsiteltiin osin myös kvalitatiivisin menetelmin. Yrityksen kunnossapitopäällikköä haastateltiin ennustettavuuden eroavaisuuksista ja niiden mahdollisista syistä. Kunnossapitopäällikön mukaan parhaimman R^2 -pisteytyksen PK-laitteet, ovat erityisen tärkeitä ja tuotantokriittisiä laitteita. Näin ollen esimerkiksi PK-laitteiden suunnitellut huollot ovat tehty noudattaen erityistä suunnitelmallisuutta ja huolellisuudella. Laiteluokan tärkeys on tiedostettu läpi

koko kunnossapito-organisaation. Kunnossapitopäällikön mukaan tämä on saat-
tanut johtaa myös täsmällisempiin työkirjauksiin ja parempaan datan laatuun.
(Kunnossapitopäällikkö 2021.)

8 EAM-JÄRJESTELMÄN KEHITYS

Saatujen ennustemallien tulosten perusteella on mahdollista antaa parannusehdotuksia kunnossapidon tietojärjestelmien kehittämiseen, master dataan ja käyttöön. Ennusteisiin käytetyn datan tulisi olla riittävällä tasolla yksikäsitteistä ja tarkkarajaista, jotta koneoppimisen algoritmit voisivat hyödyntää niitä ennustemallien tekemisessä.

8.1 Tietorakenteet

Järjestelmän tietorakenteiden kehitys on järjestelmätoimittajan vastuulla. Tietorakenteiden kehityksellä voidaan vaikuttaa ennustettavuuteen ja samalla se tukee järjestelmän master datan hallintaa. Tietorakenteissa tulisi ennustettavuuden parantamiseksi suosia erityisesti luokittelevaa ja numeraalista dataa, jota pystytään parhaiten hyödyntämään koneoppimismallien kehityksessä. Niin sanottuja vapaatekstikenttiä ei voida hyödyntää samoissa määrin. Luokittelevassa datassa olisi hyvä olla lisäksi lisäparametreja, joilla saadaan luokittelevuuteen mahdollisesti syvyyttä enemmän.

8.2 Master data

Mahdollisesti jopa suurin merkitys ennustettavuuteen on järjestelmän master datalla, jonka hallintaa tukevat laadukkaasti määritetyt ja kehitetyt tietorakenteet. Tuotanto-omaisuuden tarkka ja syvä luokittelu koetaan usein työlääksi toteuttaa ja ylläpitää, mutta osaltaan se auttaa koneoppimismallien kouluttamista. Täsmällisellä luokittelulla neuroverkon painokertoimet saadaan asetettua tarkemmalle tasolle.

Case-yrityksessä laitedatan luokittelun granulariteetti on hyvin suurta. Esimerkkinä kuljetin luokan laitteita on lähes neljännes kaikista laitteista, kappalemäärän

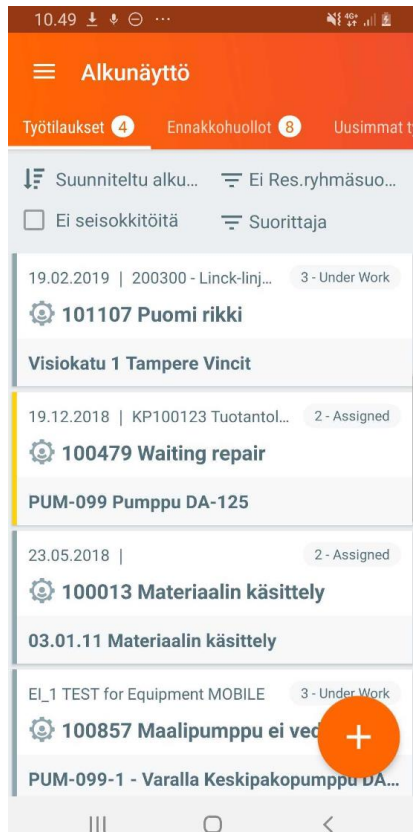
ollessa 4508 (ks. kuvio 10). Kaikkiaan laiteluokkia on 124 kpl, joista jopa 38 luokkaa sisältää alle 10 laiteyksilöä. Yhdellä tasolla tapahtuva luokittelu ei ole antanut mahdollisuutta syventää luokittelua ja täten parantaa master datan laatua.

Koneen teknisten attribuuttien merkitys perinteisessä reaktiivisessa ja kalenteri-perustaisessa kunnossapidossa on jäänyt vähäiseksi. Nämä attribuutit ovat usein palvelleet vain erilaisia listauksia, jolla voidaan listata esimerkiksi kaikki koneet, joiden moottoritehot ovat välillä 1 kW – 2 kW. Attribuutit ovat tämän vuoksi jääneet vähemmälle huomiolle, niitä hallitaan erillisillä laitedokumenteilla tai niiden hallinnasta ja ylläpidosta on luovuttu jopa kokonaan (Saunajoki 2020). Koneoppimisen hyödyntämisen kautta laitteiden tekniset attribuutit saavat kokonaan uuden merkityksen ja ne toimivat osana laitteiden luokittelevaa dataa. Järjestelmien teknisten attribuuttien tukea tulisi kehittää, jotta ne antaisivat mahdollisuuden koneoppimisen hyödyntämiseen vikojen ennustamisessa.

8.3 Tapahtumadata

Järjestelmään kerätään tapahtumadataa päivittäisessä kunnossapitotoiminnassa. Järjestelmän käytettävyydellä ja saavutettavuudella erityinen merkitys tapahtumadatan laatuun ja sen määrään. Selkeät ja hyvät käyttöliittymät varmistavat osaltaan datan laadun ja määrän. Toisaalta hankalakäyttöiset ja vaikeaselkoiset käyttöliittymät johtavat usein matalampaan käyttöasteeseen ja sen takia huonompilaatuiseen tapahtumadataan.

Tapahtumadatan ensisijainen lähdekäyttöliittymä on usein mobiilisovellus, jonka kautta kunnossapitoasentajat kirjaavat tapahtumadataa. Mobiilisovellus varmistaa, että asentajilla on yksinkertainen ja selkeä käyttöliittymä aina välittömässä läheisyydessä, muun muassa töiden raportointia varten. Tämän tutkimuksen case-yrityksen Powermaint-järjestelmään ei ole ollut saatavissa mobiilisovellusta vaan kaikki tapahtumadata on kerätty tietokoneen työasemasovelluksella.



KUVIO 24. Mobiilikäyttöliittymä (VincitEAM). Esimerkki EAM-järjestelmän mobiilikäyttöliittymästä.

Kunnossapitojärjestelmän tapahtumadata voidaan karkeasti jakaa rakenteelliseen ja ei-rakenteelliseen dataan. Rakenteellista dataa edustaa muun muassa luokittelu- ja tekstimuotoiset työraportit. Ei-rakenteellista dataa puolestaan esimerkiksi raporteille liitetyt kuvat ja videot. Kognitiiviseen kunnossapitoon siirryttäessä tullaan tarvitsemaan kaikkia datatyyppejä ja -muotoja, sillä näiden hyödyntämispotentiaali tulee kasvamaan kehittyneiden tekoälypohjaisten menetelmien myötä. Järjestelmän kehityksessä onkin syytä ottaa huomioon laaja-alaisesti näiden kaikkien datatyyppeiden tehokas kerääminen, eikä keskittyä pelkästään luokittelevaan dataan, jota tässä opinnäytetyössä tutkittiin.

8.4 Järjestelmän käyttö

Järjestelmän käytöllä on suuri merkitys datan laatuun. Käyttöä tukee erityisesti hyvin toimivat käyttöliittymät ja huolellisesti suunniteltu kunnossapitotoiminta. Hyvin toimivassa kunnossapito-organisaatiossa on laadittu selkeät toimintaoh-

jeistukset, jotka mahdollisuuksien mukaan pohjautuvat standardeihin. Tämä osaltaan vähentää henkilöriippuvuutta datan laadun suhteen. Huonosti ohjeistettuna eri henkilöt saattavat luokitella saman työn täysin eri tavalla, jos ohjeistuksia ei ole käytettävissä. Näin ollen on suuri mahdollisuus siihen, että kunnossapitodataa ei pystytä hyödyntämään tehokkaasti analytiikan ja tekoälyn menetelmin. Järjestelmän käytössä tulisi pyrkiä erityisesti kerättävän datan tasalaatuisuuteen käyttämällä toimintaohjeita ja perehdyttämällä uudet henkilöt huolellisesti tehtävänsä ja järjestelmän käyttöön. Ohjeistuksien merkitys kasvaa, kun vaihtuvuus on suurta tai kunnossapidossa käytetään paljon alihankintaa. Järjestelmän käytöstä ja ohjeistuksesta on usein vastuussa organisaation operatiivisesta kunnossapidosta vastaava henkilö.

8.5 Kehitysehdotusten vienti käytäntöön

Tässä opinnäytetyössä saadut parannus- ja kehitysehdotukset viedään käsiteltäväksi kunnossapitojärjestelmän tuotehallinnan ohjausryhmälle, jonka kautta ne liitetään soveltuvilta osin käyttäjätarinoina (*User Story*) tuotekehityksen kehitysjonoon eli niin sanottuun *backlogiin*. Kehitysjonosta käyttäjätarinat poimitaan *Sprint Planning* -palaverin yhteydessä kehitykseen. Uudet ominaisuudet kehitetään tuotteen kehitysympäristöön, josta ne testausten jälkeen lisätään osaksi tuotetta. Myös tulevien uusien käyttäjätarinoiden osalta tullaan kehityksessä huomioimaan tässä opinnäytetyössä havaitut seikat vikaantumisten ennustettavuuden parantamiseksi.

Master datan hallintaan liittyvät havainnot tullaan huomioimaan järjestelmäkäyttönottoprojektien osalta projektipäälliköiden ohjeistuksissa. Näin master datan laatu voidaan varmistaa heti käytön alusta alkaen. Järjestelmän käyttöön liittyvät havainnot tullaan ottamaan huomioon käyttöohjeiden kehityksessä ja asiakkaan pääkäyttäjän ohjeistuksissa, jotta kertyvä tapahtumadata tukisi mahdollisimman hyvin vikaantumisten ennustettavuutta.

9 POHDINTA

Kunnossapitojärjestelmän data sisältää paljon arvokasta tietoa kunnossapidon toiminnasta. Tässä datassa on usein mukana myös kontekstin mukaiset päätelmät ja asiantuntijoiden raportoimat ratkaisut ongelmiin. Tässä opinnäytetyössä saatujen tutkimustulosten perusteella voitiin antaa joitakin järjestelmään ja sen käyttöön liittyviä kehitysehdotuksia ennustettavuuden parantamiseksi. Lisäksi eräs merkittävä havainto oli, että suuri datan määrä ei aina takaa parempaa ennustettavuutta. Sen sijaan hyvä datan laatu yhdistettynä laadukkaan datan määrään tukee vikaantumisten ennustettavuutta hyvin. Datan suuri määrä ei siis takaa hyvää lopputulosta, jos data ei ole laadukasta. Toisaalta laadukkaan data ei tuota toivottua tulosta, ellei sitä ole riittävästi.

Tässä opinnäytetyössä käytettyjen menetelmien ja datan perusteella voidaan todeta, että sellaisenaan kunnossapitojärjestelmän käyttö vikojen ennustamiseen ei ole riittävän luotettava. Sen sijaan osana laajempaa kunnossapidon järjestelmien ekosysteemiä, toiminnanohjausjärjestelmän data yhdessä anturidatan kanssa voi parhaimmillaan tukea tehokkaasti vikojen ennustettavuutta. Tällainen tehokkaasti toimiva kokonaisuus antaa parhaan mahdollisen tuen päivittäiseen päätöksentekoon kunnossapitotoiminnassa. Suurin haaste datan tehokkaalle hyödyntämiselle usein on, ettei organisaatio ole alun perin suunnitellut dataa hyödynnettäväksi missään muualla kuin siinä prosessissa, jossa sitä luodaan (Pankakoski 2020).

Tekoälyyn pohjautuvien mallien oppimista voidaan usein kuvata mustana laatikona, jonka päättelymallia ei pystytä tarkemmin kuvaamaan (Merilehto 2018, 165). Tämä vaikeuttaa tarkkojen kohdennettujen ohjeiden tai toimintamallien antamista ennustemallin luotettavuuden parantamiseksi.

Teoreettisiin lähtökohtiin nähden koin onnistuneeni tutkimustyössäni hyvin. Vaikka odotin saavani datasta parempaa vikaantumisten ennustettavuutta tekoälyn menetelmää käyttämällä, niin tutkimustulos antoi arvokasta tietoa siitä, millaista dataa tulisi saada lisää, jotta ennustettavuus olisi parempi. Tutkimukseen

valitsemani koneoppimisen menetelmä oli vain yksi lukuisten tekoälymenetelmien joukosta.

Eri tekniikoita tutkimalla olisi varmasti mahdollista löytyä parempikin menetelmä tässä tutkimuksessa käytetyn datan analysointiin. Eri kunnossapito-organisaatioissa on myös erilaiset toimintatavat järjestelmän käytön suhteen ja näin ollen datan määrä ja laatu voi vaihdella eri organisaatioiden välillä huomattavastikin. Useiden eri tekoälyn menetelmien ja datan määrän ja laadun vaihtelun vuoksi, nyt saaduista tutkimustuloksista ei sellaisenaan voida tehdä yleisiä päätelmiä vikaantumisten ennustettavuudesta kunnossapitojärjestelmän sisältämän datan avulla.

Luvussa 8 annettujen järjestelmän ja sen käytön kehitysehdotusten lisäksi, hyvänä jatkotutkimuksena toimisi IoT-datan ja kunnossapidon toiminnanohjausjärjestelmän datan yhdistäminen vikaantumisennusteiden tekemiseksi. Lisäksi osana tulevaa kognitiivista kunnossapitoa saattaisi tukea tutkimus, jossa järjestelmädataan pohjautuen tuotettaisiin automaattisesti kunnossapitoasentajalle avustavaa tietoa ongelman ratkaisuun. Laitteen vikaantumistapauksessa asentajalle tulisi järjestelmästä automaattisesti avustavaa tietoa aikaisempien vastaavien vikakorjausten ratkaisusta ja käytetyistä varaosista. Näin vikaantumisen aiheuttama epäkäytettävyysaika jäisi mahdollisimman lyhyeksi, kun asentajalla olisi jatkuvasti mukanaan tekoälyyn pohjautuva kunnossapitoavustaja.

Kun tekoäly otetaan mukaan yritysten arkeen osaksi, vaikka pieniäkin ongelmanratkaisuja, lisää se mahdollisuuksia tekoälyn hyödyntämispotentiaaliin, kun haetaan ratkaisuja isompiin ongelmiin. Tekoälyn hyödyntämispotentiaali kasvaa, kun ymmärrys sen mahdollisuuksista ja rajallisuudesta ymmärretään laajalti organisaatioissa ja nämä osataan ottaa huomioon kaikissa ongelmanratkaisun vaiheissa aina datan keräämisestä lähtien.

LÄHTEET

Ackerman, E. & Ruusuvuori, P. 2017. Älykäs data-analytiikka teollisen internetin mahdollistajana. Julkaisussa Teollinen internet uudistaa palveluliiketoimintaa ja kunnossapitoa. Helsinki: Promaint ry.

Algosome. Dummy Variable Trap in Regression Models. Luettu 12.12.2020.
<https://www.algosome.com/articles/dummy-variable-trap-regression.html>

Ali-Marttila, M. Kärri, T. Marttonen-Arola, S. Metso, L. Sinkkonen, T. & Ylä-Kujala, A. 2019. Verkostoituneen kunnossapitoympäristön johtaminen. Julkaisussa Kunnossapidon vuosikirja 2019. Helsinki: Promaint ry.

Genesis Solutions. Think Ahead: Where Cognitive Predictive Maintenance is Leading EAM. Luettu 12.12.2020.
<https://www.genessolutions.com/Knowledge-Center/Insights/Think-Ahead-Where-Cognitive-Predictive-Maintenance-is-Leading-EAM/>

Hovi, J. Blogi. Älä tee tekoälyä vain tekoälyn vuoksi – näin vältät epäonnistumisen. Luettu 9.12.2020.
<https://www.arihovi.com/ala-tee-tekoalya-vain-tekoalyn-vuoksi-nain-valtat-epaonnistumisen/>

Kananen, H. & Puolitaival, H. 2019. Tekoäly bisneksen uudet työkalut. Helsinki: Alma Talent ry.

Kaskikallio, K & Niittymaa, H. 2017. Kognitiivinen tietojenkäsittely teollisen internetin kunnossapitoratkaisuissa. Julkaisussa Teollinen Internet uudistaa palveluliiketoimintaa ja kunnossapitoa. Helsinki: Promaint ry.

Komonen, K. 2019. Tuotanto-omaisuuden hallinnan kenttä. Julkaisussa Kunnossapidon vuosikirja 2019. Helsinki: Promaint ry.

Kortelainen, H. & Tennilä, J. 2019. Kunnossapidon taloudellinen merkitys ja tunnusluvut. Julkaisussa Kunnossapidon vuosikirja 2019. Helsinki: Promaint ry.

Kunnossapitopäällikkö. 2020. Haastattelut ajalla 1.7.2020 – 31.12.2020. Haastattelija Hakala, J. Tampere.

Kunnossapitopäällikkö. 2021. Haastattelu 4.1.2021. Haastattelija Hakala, J. Tampere.

Merilehto, A. 2018. Tekoäly matkaopas johtajalle. Helsinki: Alma Talent ry.

Pankakoski, J. 2020. Kolumni. Datan käyttö liiketoiminnassa – Mahdollisuus vai välttämättömyys. Luettu 28.11.2020.
<https://www.alihankinta.fi/fi/kolumni/datan-kaytto-liiketoiminnassa-mahdollisuus-vai-valttamattomyys/>

PSK 6201 standardi. 2011. Helsinki: PSK Standardisointiyhdistys ry.

Saunajoki, S. konsulttipäällikkö. 2020. Haastattelu 4.1.2021. Haastattelija Hakala, J. Tampere.

Siukonen, T. & Neittaanmäki, P. 2019. Mitä tulisi tietää tekoälystä. Jyväskylä: Docendo Oy.

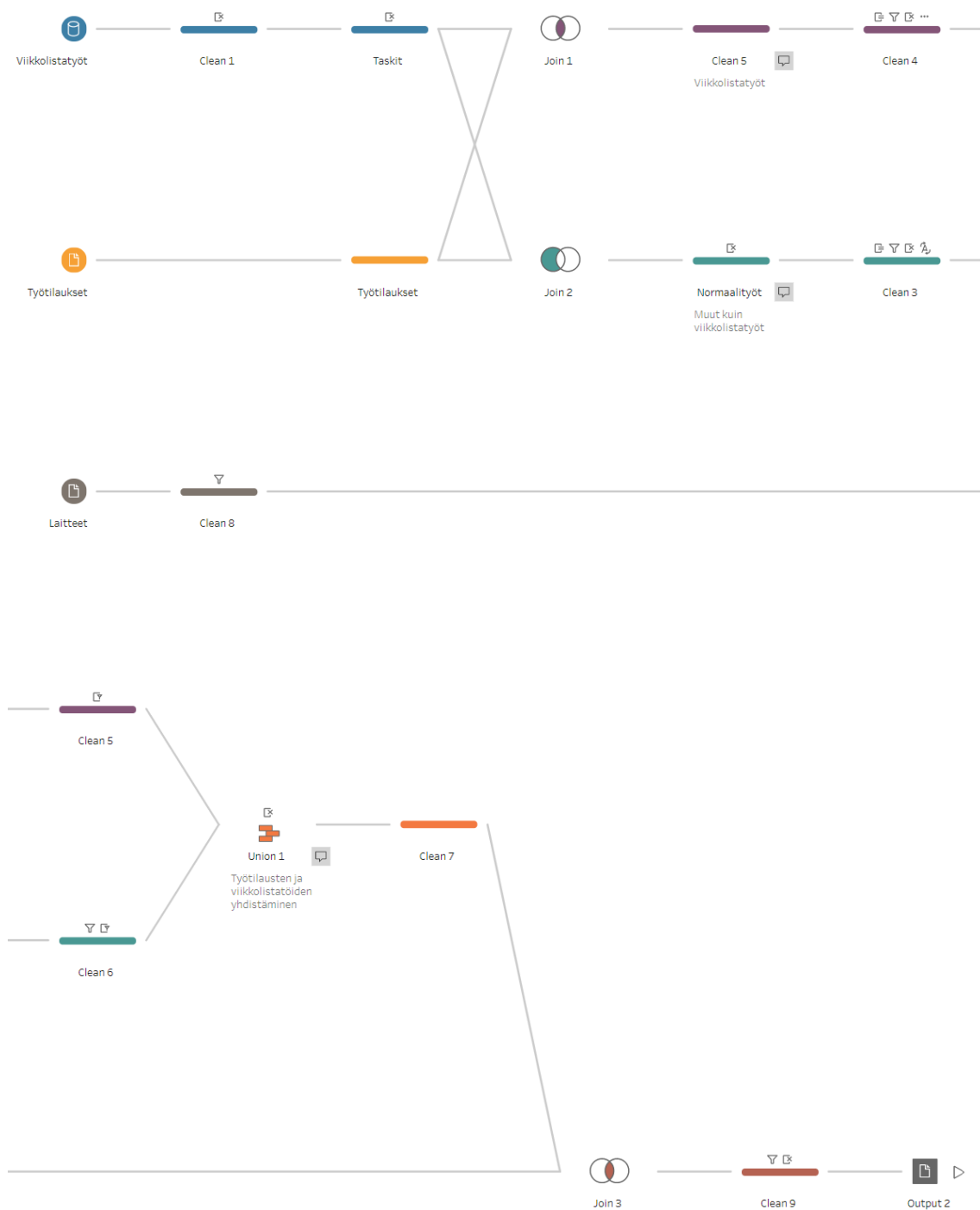
P. Tchakoua, R. Wamkeue, M. Slaoui-Hasnaoui, F. Tameghe, T. Ekemb, G. 2013. New Trends and Future Challenges for Wind Turbines Condition Monitoring. Luettu 8.12.2020.

https://www.researchgate.net/publication/271456155_New_trends_and_future_challenges_for_wind_turbines_condition_monitoring

LIITTEET

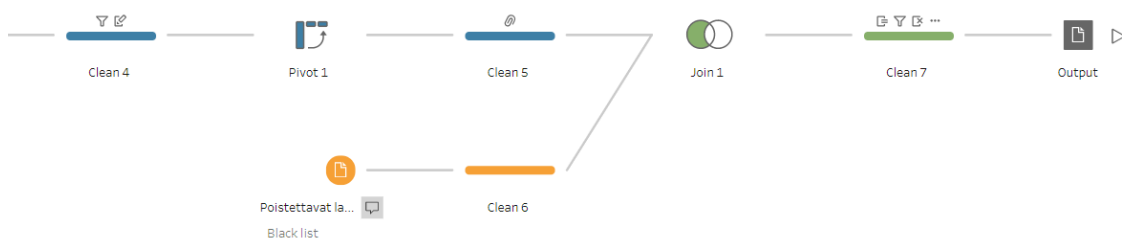
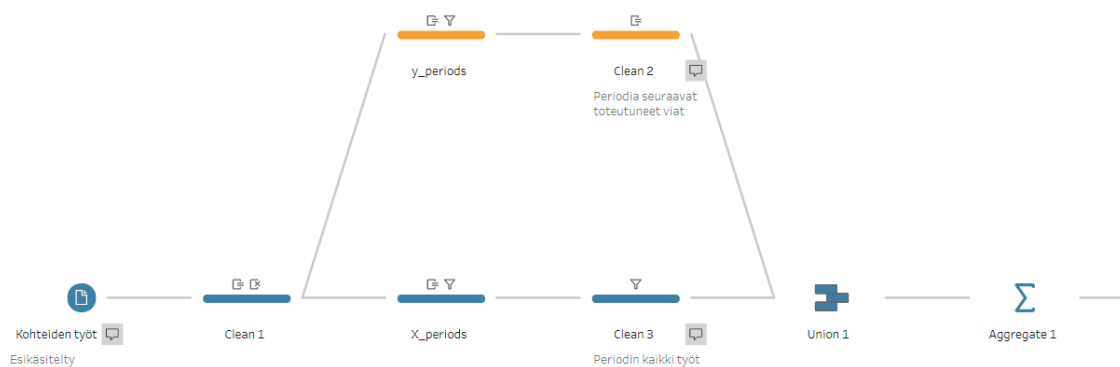
Liite 1. Datan esikäsittely

1 (2)



Liite 1. Datan esikäsittely

2 (2)

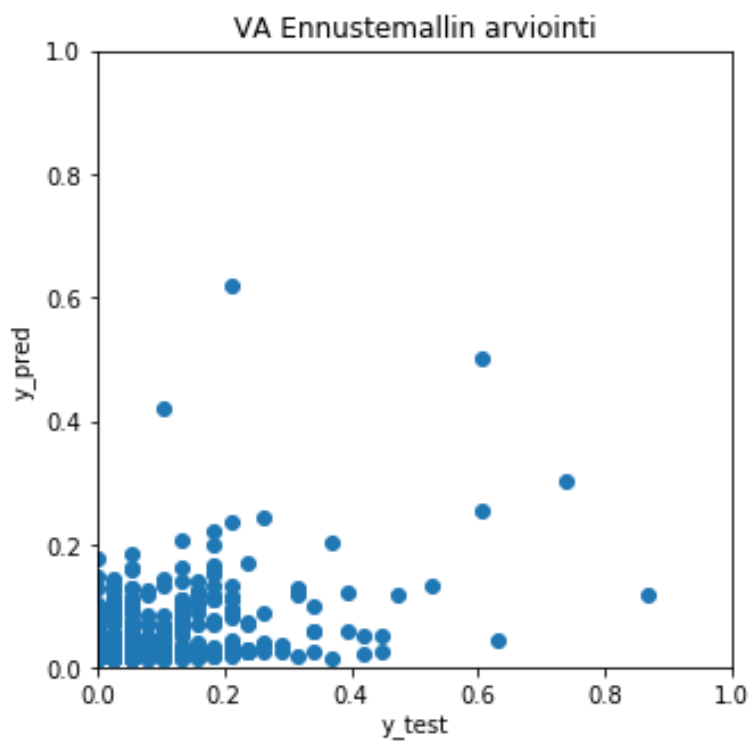
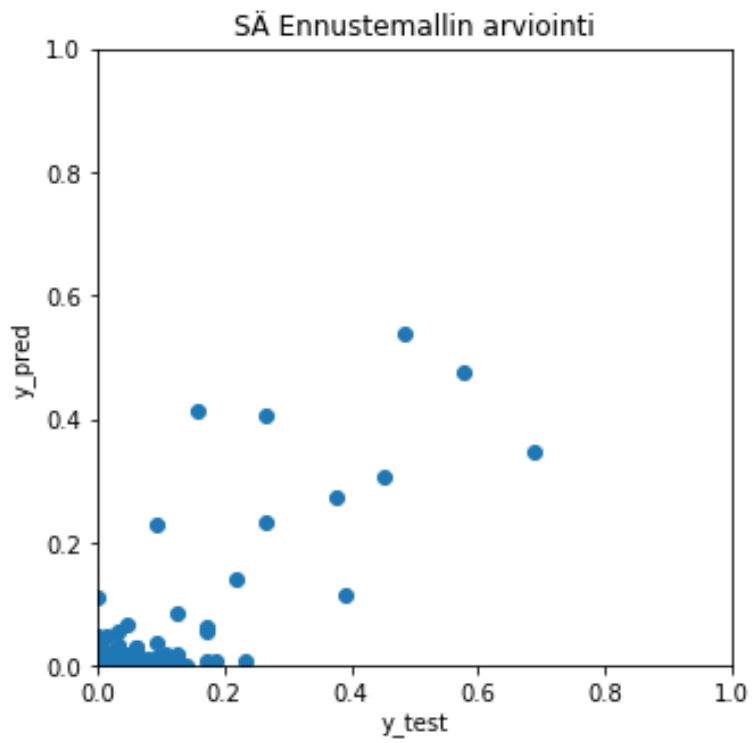


Liite 2. Ote käsitellystä datasta (csv)

OBJ_ID_N	Period	Qx	O_Type	X HK_KO	X SU	X EH	X EL	X MT	X VH	X TH	X KT	y HK_KO
14282	201001	1	PE			13						
2025	201404	4	KU	1		8						
161	201102	2	KU	1		5		1				8
25506	201802	2	NL		1							
54192	201603	3	KU									1
7764	201402	2	KU			2						
62060	201604	4	PE			3						
29626	201104	4	VA	3								
64547	201701	1	KU			4						
25329	201402	2	BRT	6		25					1	10
33779	201103	3	KU			5						
1221	201103	3	KU			6						
14713	201202	2	KIK									1
2655	201502	2	VA	3								
8981	201203	3	KU			4						
44406	201401	1	KU	1		2						1
8563	201001	1	ET			2						
29904	201104	4	IV			1						
35828	201704	4	KU			16						
54364	201702	2	KO	4		23				1		5
24346	201801	1	KU			2						
43364	201102	2	PK	2		2						2
4461	201003	3	KIK	1								1
25064	201101	1	SÄ			1						
1244	201102	2	KU			3						1
26884	201302	2	VA			29						
28506	201204	4	VA									4
36266	201602	2	KU			4						
35909	201301	1	KU			5						2
46714	201602	2	IV	1								
3937	201602	2	PK	12		6			1			14
3449	201703	3	PU			4						
3130	201101	1	KO			1						
44585	201503	3	KO			3						
14749	201601	1	IV			4						
33841	201003	3	KU			13						
55407	201703	3	VA			7						
11362	201302	2	NL	1								
13542	201003	3	NL	2								
47570	201303	3	NO									1
20584	201204	4	KU			7						
34924	201302	2	LE			1						1
46695	201302	2	KIK			1						

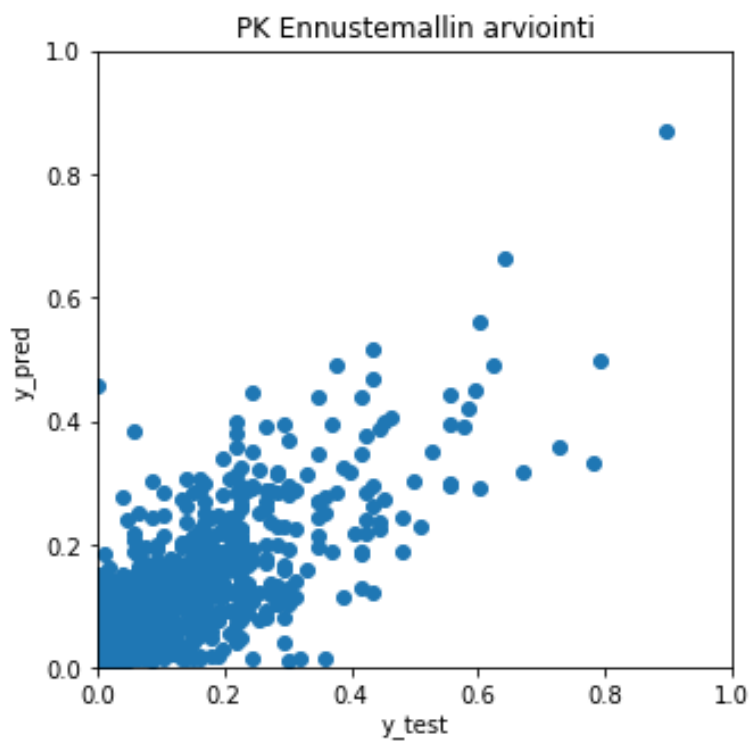
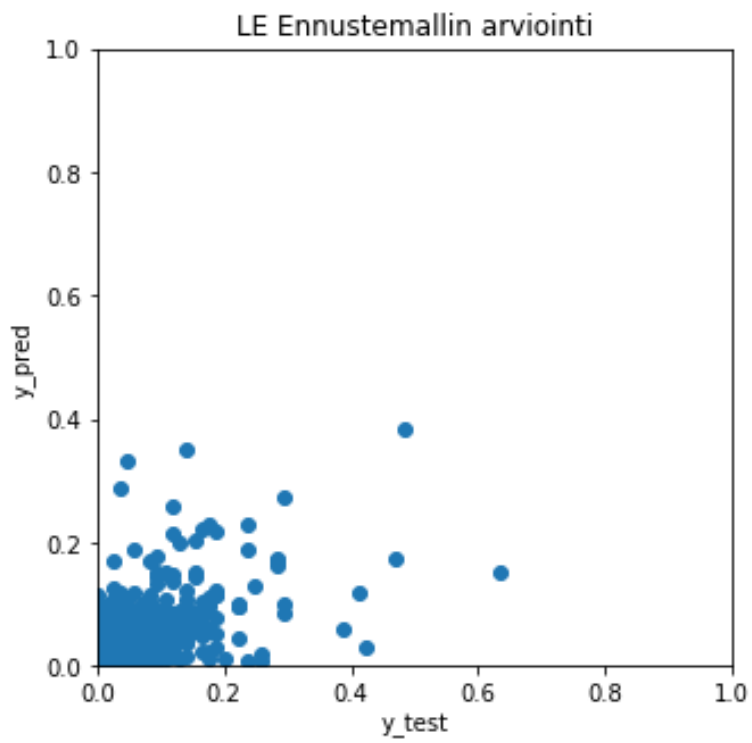
Liite 3. Ennustemallien toimivuus laitetyypeittäin

1 (6)



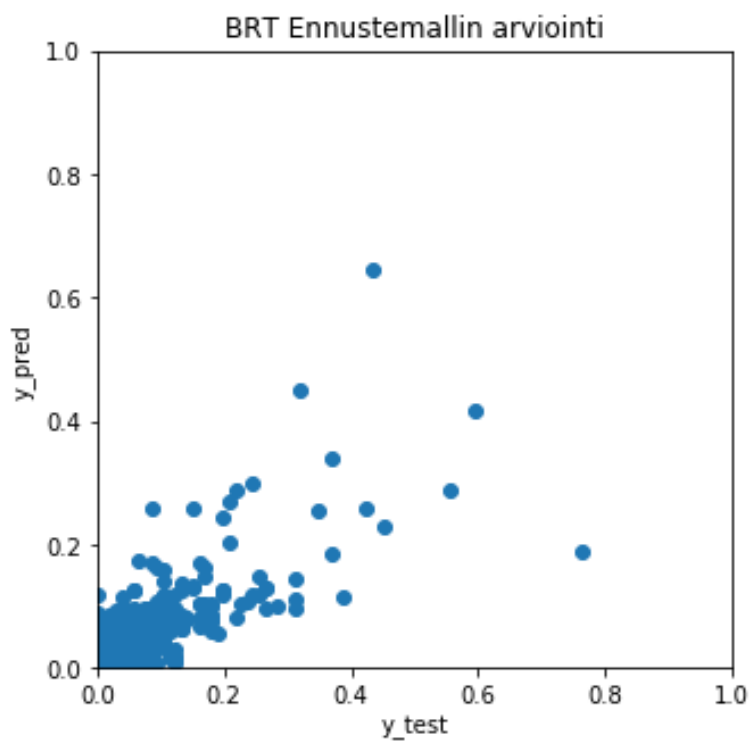
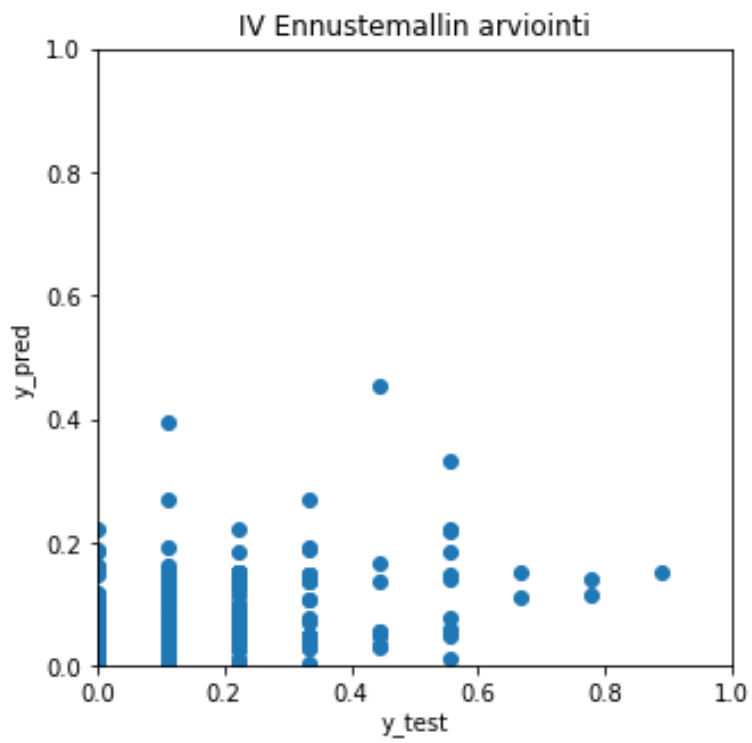
Liite 3. Ennustemallien toimivuus laitetyypeittäin

2 (6)



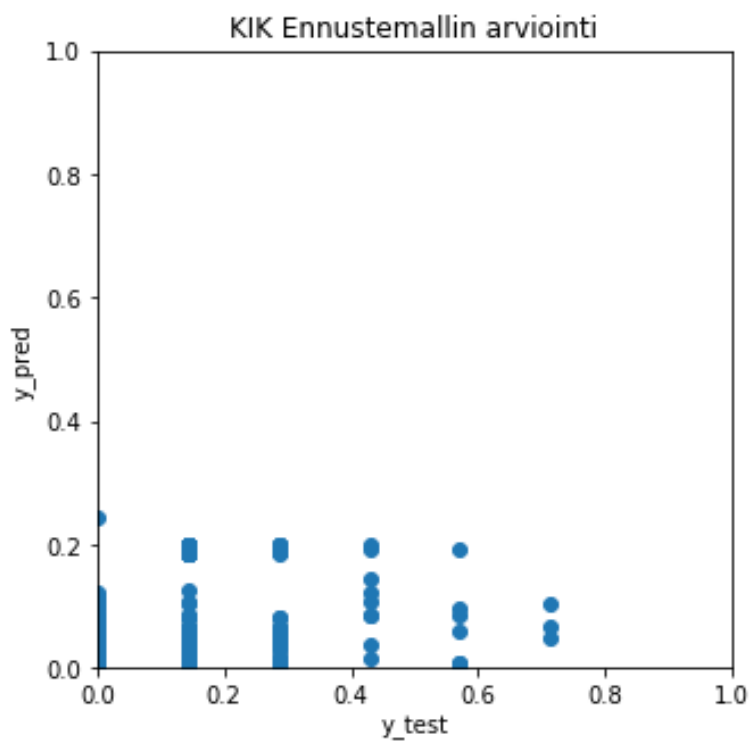
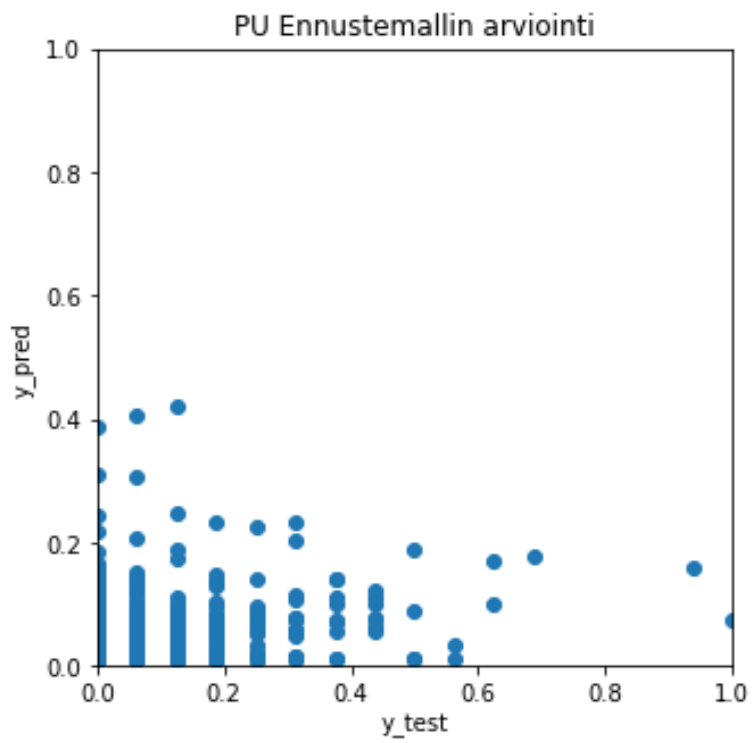
Liite 3. Ennustemallien toimivuus laitetyypeittäin

3 (6)



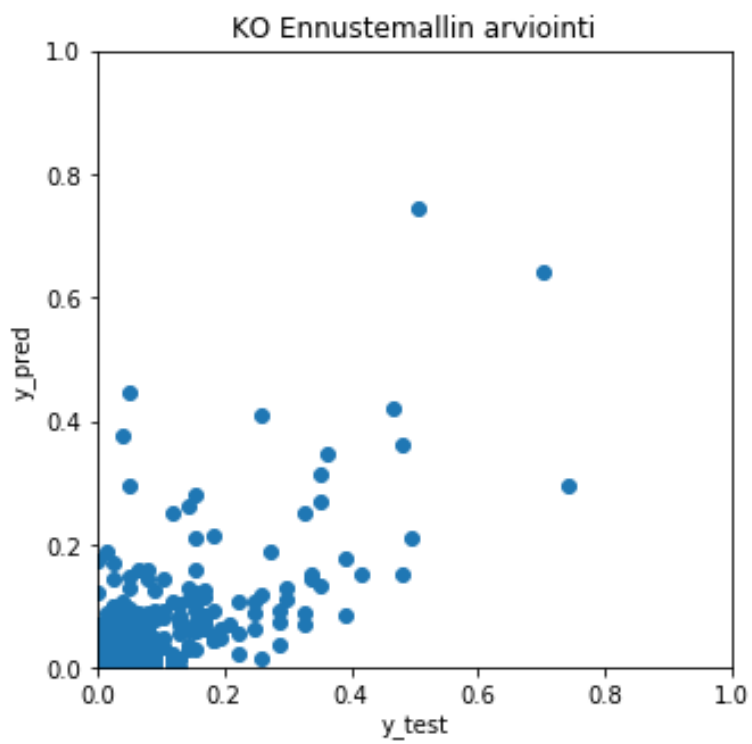
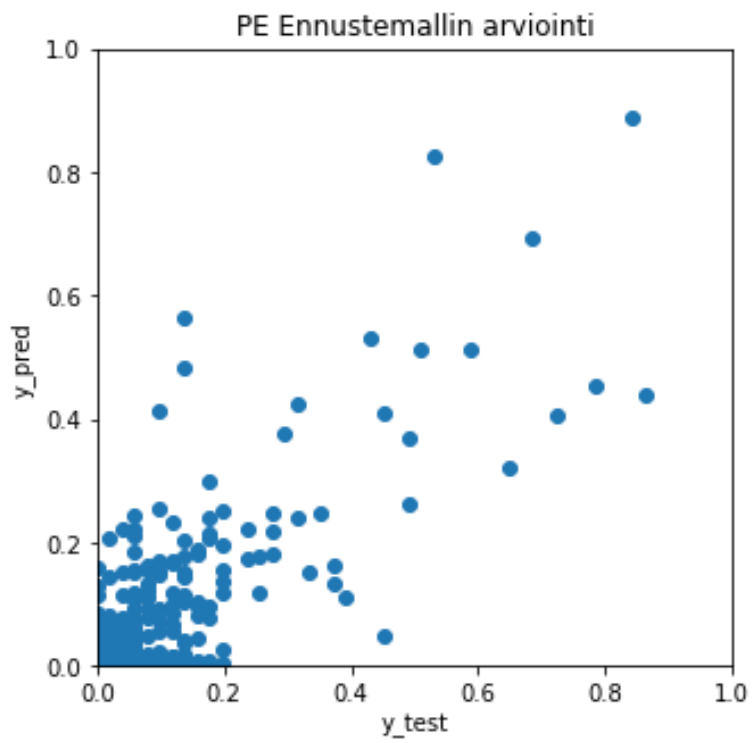
Liite 3. Ennustemallien toimivuus laitetyypeittäin

4 (6)



Liite 3. Ennustemallien toimivuus laitetyypeittäin

5 (6)



Liite 3. Ennustemallien toimivuus laitetyypeittäin

6 (6)

