



TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

EERIK PAKARINEN
PAPERIKONEKÄYTTÖJEN ANALYSOINTI TEKOÄLYLLÄ

Diplomityö

Tarkastaja: professori Risto Ritala
Tarkastaja ja aihe hyväksytty
1. marraskuuta 2018

TIIVISTELMÄ

EERIK PAKARINEN: Paperikonekäyttöjen analysointi tekoälyllä

Tampereen teknillinen yliopisto

Diplomityö, 52 sivua

Elokuu 2018

Automaatiotekniikan diplomi-insinöörin tutkinto-ohjelma

Pääaine: Systemien hallinta

Tarkastaja: professori Risto Ritala

Avainsanat: paperikone, ohjausjärjestelmät, tekoäly

Tässä työssä selvitetään tekoälyn mahdollisuuksia paperikoneen sähkökäyttöjen analysointiin. Työ rajataan leikkurin mekaanisten värähtelyiden analysointiin. ABB:n sähkökäytöt mittaavat antureilla valtavan määrän dataa talteen. Laitteiden vikaantuessa toimitaan tällä hetkellä reaktiivisesti, vaikka prediktiivisesti toimiminen olisi kaikkien osapuolten kannalta parempi vaihtoehto. Tekoälyn ja tiedonlouhinnan avulla voisi olla mahdollista löytää mekaanisista värähtelyistä joitain tärkeitä suorituskykyilmaisimia eli KPI:tä.

Työ suoritetaan analysoimalla leikkurin sähkökäyttöjen mekaanisten värähtelyiden mittausdataa. Mittausdata on kerätty dataloggereille. RTDB-tietokanta (Real-Time Database) tulee ensin muuttaa sopivaan muotoon, kuten xls-tiedostoksi, jotta sen käsittely onnistuu kolmannen osapuolen pilvipalveluissa. Datan perusteella määritellään valmis koneoppimisalgoritmi, jonka avulla pyritään löytämään datasta mahdollisia toistuvuuksia.

Teoriaosuudessa käsitellään työn toimintaympäristö sekä työkalut, joita työssä käytetään. Toimintaympäristössä esitellään paperitehdas ja paperikone yleisesti ja leikkuri tarkemmin. Sähkökäyttöjen osalta kerrotaan erikseen moottoreista, taajuusmuuttajista ja ohjausjärjestelmistä. Viimeisenä toimintaympäristön osana esitellään teollisuus 4.0, joka on uusin teollisuuden vallankumous ja johon pilvipalvelut ja IoT liittyvät merkittävästi.

Työkalut, jotka työssä esitellään, ovat tekoäly, IoT ja tiedonlouhinta. Tekoäly on tietokoneen ajattelukykyä. Sitä voidaan kehittää loogisesti tai ihmismäisesti. Tekoälyn suurimmat odotukset kohdistuvat tällä hetkellä tietokoneen itseoppimiseen ja varsinkin luonnollisen kielen ymmärtämiseen. Esimerkkinä odotuksia täynnä olevasta tekoälystä on IBM Watson, joka pystyy nykyään puhumaan ja ymmärtämään yhdeksää kieltä natiivisti. Tekoälyn kognitiivisiin mahdollisuuksiin lukeutuvat myös ennustavat huoltotoimenpiteet. Uudet tekniikat, kuten esimerkiksi 5G, ovat välttämättömiä, jotta IoT voi yleistyä tehtaissa. Tiedonlouhinta on myös tärkeää datan määrien kasvaessa.

ABSTRACT

EERIK PAKARINEN: Analyzing paper machine drive data with AI
Tampere University of Technology
Master of Science Thesis, 52 pages
August 2018
Master's Degree Programme in Automation Engineering
Major: System Operation
Examiner: Professor Risto Ritala

Keywords: paper machine, control systems, AI

This study explores the possibilities of artificial intelligence in analyzing the electrical drives of a paper machine. The study is limited to the mechanical vibrations of the winder. At present, ABB's electrical drives produce huge amounts of data. Currently all equipment failures are handled reactively, even though it would be a better option to act predictively. With AI and datamining, it might be possible to find some important key performance indicators (KPIs) for mechanical vibrations.

The study is conducted by analyzing the mechanical vibration measurement data of the electric drives of a winder. Measurement data has been collected on data loggers. The RTDB database must first be converted to i.e. XLS format, so that it is possible to analyze the data with third-party cloud computing services. Based on the data, a machine learning algorithm will be defined in order to find possible patterns in the data.

The theoretical part of the thesis consists of the related operating environment and the tools used in this work. In the section for the operating environment, paper mill and paper machine are presented in general and winder in more detail. As for electric drives, there will be separate introductions for motors, frequency converters and control systems. The last part of this section will cover Industry 4.0, the latest industrial revolution that involves the phenomena of cloud services and IoT.

The tools that will be presented in the study are artificial intelligence, IoT and data mining. Artificial intelligence is a computer's ability to think. It can be developed logically or humanly. The greatest expectations are currently focused on self-learning and particularly understanding natural language. An example of this is IBM Watson, which now speaks and understands nine languages natively and has been able to win the Jeopardy! game show. The cognitive potential of artificial intelligence includes predictive maintenance. New techniques in wireless networks and quick connections are important in order to make IoT more commonly utilized in factories. Data mining also plays an important role in analyzing big data.

ALKUSANAT

Kiitos tästä työstä kuuluu työni tarkastajalle professori Risto Ritalalle, ohjaajalle DI Markku Haikolalle sekä muille ABB:n asiantuntijoille, joilta sain paljon apua ja opetusta aiheesta. Kiitos kuuluu myös perheelleni, jolta sain paljon tukea työni kanssa.

Helsingissä, 09.08.2018

Eerik Pakarinen

SISÄLLYSLUETTELO

1.	JOHDANTO	1
1.1	Työn tausta	1
1.2	Työn tarkoitus.....	1
1.3	Työn rakenne	2
2.	TOIMINTAYMPÄRISTÖ	4
2.1	Paperiteollisuus.....	4
2.2	Pituusleikkuri	5
2.3	Sähkökäytöt.....	8
2.3.1	Moottorit.....	8
2.3.2	Kytkimet ja vaihteistot	10
2.3.3	Taajuusmuuttajat.....	11
2.3.4	Ohjausjärjestelmät.....	12
2.3.5	Mekaaniset värähtelyt	14
3.	TYÖKALUT.....	16
3.1	Tekoäly	16
3.1.1	Havaitseminen	17
3.1.2	Ymmärtäminen	17
3.1.3	Oppiminen	18
3.1.4	Tulevaisuus.....	19
3.2	IoT.....	19
3.2.1	Teollisuus 4.0.....	21
3.2.2	IoT:n tulevaisuus.....	22
3.3	Tiedonlouhinta.....	23
3.3.1	Esikäsittely	24
3.3.2	Riippuvuusanalyysi.....	25
3.3.3	Luokittelu	26
3.3.4	Tulosten tarkastus	27
3.4	Ennakoiva kunnossapito	27
4.	ALUSTAT	29
4.1	Pilvipalvelut	29
4.2	IBM Watson	30
4.2.1	Käyttökohteet.....	31
4.2.2	Käyttökohteet paperiteollisuudessa	34
4.3	Microsoft Azure.....	35
4.3.1	Käyttökohteet.....	35
4.3.2	Käyttökohteet paperiteollisuudessa	35
5.	DATA JA SEN ANALYYSI.....	37
5.1	Raakadatan esittely	37

5.2	Datan muokkaus	37
5.3	Tiedonlouhinta IBM Watson Analyticsilla ja Azure Machine Learning Studiolla	38
5.4	Tiedon esittäminen	41
6.	YHTEENVETO	44
6.1	Yleistä	44
6.2	Tekoälyn hyödyntäminen paperikonekäytössä	44
6.2.1	Ennakoiva kunnossapito.....	44
6.2.2	Vianmääritys.....	45
6.2.3	Parempi suunnittelu.....	45
6.3	Mahdollisuudet	45
6.4	Tulevaisuuden töitä.....	46
	LÄHTEET	47

LIITE A: IBM WATSON ANALYTICSIN ENNUSTEMALLIN SÄÄNTÖLUETTELO
KIREYDEN REFERENSSIN KESKIARVOLLE

KUVALUETTELO

<i>Kuva 1. Paperikoneen piirroskuva. Perustuu lähteeseen [3].</i>	5
<i>Kuva 2. Patenttikuva leikkurin toiminnasta. Perustuu lähteeseen [6]</i>	6
<i>Kuva 3. Paperin levitystelat [5, Luku 21.1]</i>	6
<i>Kuva 4. Kantotelaleikkurin toimintaperiaate [7].</i>	7
<i>Kuva 5. Keskiörullaimen toimintaperiaate [5, Luku 21.1].</i>	7
<i>Kuva 6. Sähkökäytön tyypillinen järjestelmä. Muokattu lähteestä [8].</i>	8
<i>Kuva 7. Kaavioesitys kolmivaihekkäämityksen synnyttämästä pyörivästä magneettikentästä [9, s. 120].</i>	9
<i>Kuva 8. Tasasähkömoottorin periaatekuva. Perustuu lähteeseen [11].</i>	10
<i>Kuva 9. Taajuusmuuttajan periaatekuva [13], termit suomennettu.</i>	12
<i>Kuva 10. Taajuusmuuttajan säätörakenne.</i>	12
<i>Kuva 11. Teollisuus 4.0. Muokattu lähteestä [31].</i>	22
<i>Kuva 12. Gartnerin hypekäyrä IoT:n kehitykselle. Muokattu lähteestä [32].</i>	23
<i>Kuva 13. Neuroverkkoregression arkkitehtuuri yksinkertaisilla sisään- ja ulostulosolmuilla. Muokattu lähteestä [39, Luku 4].</i>	26
<i>Kuva 14. Watson Analyticsin ehdottamat kysymykset datasta.</i>	31
<i>Kuva 15. Watson Analyticsin visualisointi parametrin ajureista.</i>	32
<i>Kuva 16. IBM Watson Analyticsin visualisointimahdollisuudet.</i>	32
<i>Kuva 17. Watson Analyticsin ehdottamat analyysilöydöt.</i>	33
<i>Kuva 18. AIEQ pörssinoteeratun rahaston ja NASDAQ-indeksin (COMP) vertailu vuoden aikajaksolta [50].</i>	34
<i>Kuva 19. Microsoft Azuren ennustavan kunnossapidon esimerkkiarkkitehtuuri [54].</i>	36
<i>Kuva 20. IBM Watson Analyticsin ennustemallin sääntöluettelo.</i>	38
<i>Kuva 21. Luettelo Azure Machine Learning Studion palikkaryhmistä.</i>	39
<i>Kuva 22. Azure Machine Learning Studion alusta, johon on rakennettu regressioiden vertailun ohjelma.</i>	40
<i>Kuva 23. Azure Machine Learning Studion taulukko, jossa on kuvattu eri regressioiden ominaisuuksia testidatan pohjalta.</i>	41
<i>Kuva 24. IBM Watson Analyticsin spiraalikaavio.</i>	42
<i>Kuva 25 IBM Watson Analyticsin ennustavan mallin päätöspuu.</i>	42
<i>Kuva 26. IBM Watson Analyticsin pylväskaavio.</i>	43

LYHENTEET JA MERKINNÄT

AGI	engl. Artificial General Intelligence, yleinen tekoäly
CPS	engl. Cyber Physical System, verkko-fyysinen järjestelmä
IaaS	engl. Infrastructure-as-a-Service, infrastruktuuri palveluna
IoE	engl. Internet of Everything, kaiken internet
IoP	engl. Internet of People, ihmisten internet
IoS	engl. Internet of Services, palveluiden internet
IoT	engl. Internet of Things, esineiden internet
NFC	engl. Near Field Communication, lähikenttäviestintä
Paas	engl. Platform-as-a-Service, sovellusalusta palveluna
RFID	engl. Radio Frequency Identification, radiotaajuusidentifiointi
RTDB	engl. Real-Time Database, reaaliaikainen tietokanta
SaaS	engl. Software-as-a-Service, sovellus palveluna
SOA	engl. Service Oriented Architecture, palveluorientoitunut arkkitehtuuri
TCP	engl. Transmission Control Protocol, siirronohjausprotokolla
WSAN	engl. Wireless Sensor and Actuator Networks, langattomien anturien ja toimilaitteiden verkko

1. JOHDANTO

1.1 Työn tausta

Suomen talouden tukipilari on pitkään ollut metsäteollisuus. Suomen viennistä noin viidesosa on metsäteollisuudesta. Suuret havumetsäalueet ovat antaneet loistavat edellytykset suomalaisille paperinvalmistukseen, mutta nykyään digitalisaatio on lisännyt paperittomuutta esimerkiksi liiketoiminnassa. Paperia kuitenkin tarvitaan yhä, vaikka sen käyttö on vähentynyt huomattavasti vuosikymmenien saatossa. Lisäksi globalisaation takia lisääntynyt verkkokauppa on kasvattanut kartonkien kysyntää. Kilpailun koventuessa paperinvalmistajien on varmistettava asemansa markkinoilla optimoimalla koneensa tuottamaan laadukasta paperia nopeasti.

Paperikoneen sähkökäytöt ovat keskeisessä osassa tätä kehitystä, sillä ne ohjaavat paperinvalmistuksen nopeutta ja laatuun vaikuttavia parametreja. Paperikoneiden sähkökäyttöjen mekaanisesta värähtelystä saa tärkeää informaatiota esimerkiksi ennakoivaan huoltoon.

Paperikoneen leikkuri leikkaa paperia pituussuunnassa sopivan kokosiin rulliin, jotta kuljetus jatkovalmistukseen olisi helpompaa. Tässä työssä tutkitaan mahdollisuuksia kehittää paperinvalmistuksen tehokkuutta analysoimalla tekoälyalustojen avulla leikkureiden tehdasdataa. Tekoälyyn perustuvia analyysityökaluja on monia, joista mainittakoon IBM Watson, Googlen Identity-Aware Proxy ja Amazonin Web Services. Tässä työssä on tutkittu dataa IBM Watson- sekä Microsoft Azure -sovelluksilla.

1.2 Työn tarkoitus

Paperinvalmistuslinja menettää potentiaalista tuottoa aina paperikoneen ollessa pysähtyneenä, joten sähkökäyttöjen tulee toimia moitteettomasti. Sähkökäyttöjen vikaantuminen voi myös johtaa ratakatkoihin, sillä sähkökäytöt ohjaavat telojen nopeuksia ja niiden myötä paperirainan kireyttä. Väärin ohjatut telat voivat hajottaa paperirainan. Paperikone on nykyään yleensä noin 10 metriä leveä ja tuottaa paperia jopa kymmeniä metrejä sekunnissa. Ohjaus on vaativaa, kun nopeutta ja laatua tahdotaan lisätä samanaikaisesti. Paperinvalmistuksessa tekoäly pystyy tekemään päätöksiä ihmistä nopeammin ja yritysten tulee olla muutosten mukana, jotta pitävät kilpailuetunsa.

Työn hypoteesina on, että tekoäly osaa löytää analysoitavasta leikkuridatasta relaatioita ja trendejä, joita ihmissilmä ei osaa havaita. IBM Watsonin Analytics-työkalun avulla voi

löytää syitä esimerkiksi ratakatkojen todennäköisille aiheuttajille. Watson osaa myös kertoa useiden asioiden yhteisvaikutuksen ja laskea niiden mahdollisen syy-seuraussuhteen. IBM Watson Analytics -työkalulla ei pysty kuitenkaan – ainakaan vielä – kovin hyvin ennustamaan, mutta ennustamisominaisuutta kehitetään aktiivisesti.

On kiinnostavaa tietää, miten muuten tekoälyä voisi hyödyntää sähkökäyttöjen analysoinnissa ja huoltotoimenpiteissä. Sähkökäytöistä luetaan tällä hetkellä kymmeniä parametreja jopa 10 millisekunnin välein. Jos näitä tietoja voitaisiin käyttää tehokkaasti, huolto ja tuotekehitys voisivat parantua huomattavasti ja myös paperinvalmistajat olisivat varmasti tyytyväisiä. Tekoälysovellusten, kuten tässä työssä käytettyjen IBM Watsonin ja Microsoft Azuren, oletetaan kehittyvän seuraavan vuosikymmenen aikana jopa pelottavan paljon. Tekoäly on vielä kuitenkin melko kapeaa, eli sitä ei voi soveltaa kovinkaan moneen asiaan. Nykyiset tekoälyohjelmistot ovat tehokkaita datan analysoinnissa, mutta maailmaa niillä ei valloiteta.

1.3 Työn rakenne

Työssä käsitellään aluksi luvussa 2 kirjallisuuden avulla sovellusympäristöä. Työn ympäristönä esitellään paperiteollisuus sekä paperikoneen pituusleikkuri ja sähkökäytöt. Sähkökäytöistä käsitellään tarkemmin sähkömoottorit, kytkimet ja vaihteistot, taajuusmuuttajat, ohjausjärjestelmät sekä toimintaa häiritsevät mekaaniset värähtelyt.

Luvussa 3 esitellään tutkimuksessa käytettävät työkalut, jotka ovat tekoäly yleisesti, Internet of Things, tiedonlouhinta ja ennakoiva kunnossapito. Tekoällyn näkökulmista työssä käsitellään havaitseminen, ymmärtäminen, oppiminen ja tekoällyn tulevaisuudennäkymiä. Internet of Thingsin käsittelyyn kuuluvat Saksassa käsitteellistetty teollisuuden neljäs vallankumous sekä sen mahdolliset tulevaisuudennäkymät. Tiedonlouhinnasta käsitellään tiedonlouhinnan esiprosessointia, tehtäviä sekä tuloksien arviointia.

Luvussa 4 kuvataan alustoja, joiden avulla paperikoneen leikkurin dataa voidaan analysoida. Työssä on tutkittu ja sovellettu IBM:n Watsonia sekä Microsoftin Azurea, jotka lukeutuvat markkinoiden edistyneimpiin tekoälysovelluksiin.

Luvussa 5 esitellään tutkimusdatan keruussa ja analysoinnissa käytetyt menetelmät. Data on rakennettu Pythonin avulla taulukoksi, jossa paperileikkurin muuttojen keskeiset piirteet on identifioitu.

Luvussa 6 esitellään tutkimuksesta saadut tulokset. Tulosten analysointiin käytetyt työkalut ovat IBM Watsonin Analytics ja Microsoft Azuren Machine Learning Studio. Tutkielma päättyy yhteenvetoon ja jatkotutkimusehdotuksiin. Tutkimuksen tuloksena on se, että tekoälyä voidaan hyödyntää paperikoneiden sähkökäytöissä ennakoivaan

kunnossapitoon, vian määrittämiseen sekä parempaan suunnitteluun. Jatkotutkimuksen tarvetta työlle on, koska käytetystä datassa ei esiintynyt riittävästi vikaantumistapauksia.

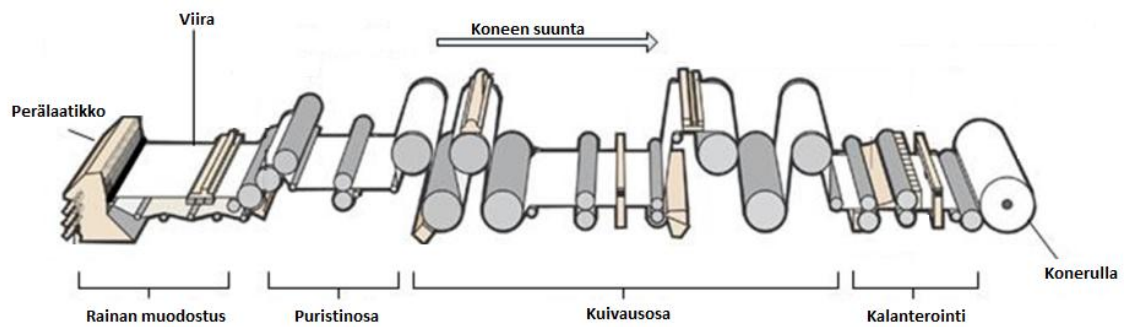
2. TOIMINTAYMPÄRISTÖ

2.1 Paperiteollisuus

Suomessa paperiteollisuudella on pitkät juuret. Suomi on tunnettu vihreästä kullastaan eli metsistään. Näiden resurssien avulla metsä- ja paperiteollisuus ovat olleet pitkän aikaa maamme merkittävimpiä vientiteollisuuden aloja. Suomessa on useita paperitehtaita, vaikkakin paperin kysyntä on viimeisen kymmenen vuoden aikana pudonnut huippuvuosista. Suomessa paperikonerakennusosaaminen on maailmanluokan tasoa.

Paperin tärkeimpiä ominaisuuksia ovat neliöpaino ja kuitumäärä [1]. Neliöpaino voi olla 10–500 g/m² paperilaadusta riippuen. Kevyimmät paperilaadut ovat pehmopapereita ja painavimmat raskaita paperilevyjä. Paperin kuivapaino on tyypillisesti 91–95 % kokonaispainosta. Korkeammalla kuiva-ainepitoisuudella paperista tulee liian haurasta ja paperin lujuus kärsii. Yksi paperitehdas voi tehdä eri päivinä erilaatuisia papereita, mutta usein on kuitenkin kannattavampaa minimoida valmistettavien lajien määrää. Kartongeille ja pehmopapereille on omat koneensa, vaikka niiden toimintaperiaate onkin melko samanlainen kuin painopaperikoneen.

Paperitehtaan tehtävä on valmistaa puumassasta paperia. Paperitehtaassa voi olla useita paperikoneita. Paperikone on suuri, satoja metrejä pitkä kokonaisuus. Paperin valmistusta voidaan pitää vedenpoistoprosessina. Paperikone (ks. kuva 1) koostuu märkämästä ja kuivapästä. Märkämässä kuiva-aineen osuus on noin 1 % ja kuivapäessä noin 95 %. Märkämää koostuu rainan muodostuksesta ja puristinosasta. Rainan muodostuksessa perälaatikolta syötetään viiralle paperimassaa, jolloin osa vedestä imetään viiran läpi. Puristinosassa puristaa rainaa ja saa selluloosakuidut yhdistymään. Kuivapää koostuu kuivausosasta, kalanteroinnista ja rullaimesta. Kuivausosassa paperi ajetaan useiden kuumien kuivatussylintereiden kautta, mikä kuivattaa paperin lähes lopulliseen kosteuteensa. Kalanteroinnissa käsitellään paperin pintaa. Esimerkiksi kopiopaperi hiotaan kalanterissa sileämmäksi, jotta painomuste tarttuu paremmin paperiin. Rullain rullaa paperirainan isoiksi konerulliksi. Asiakkaiden vaatimuksesta konerullien paperi yleensä käsitellään jollain tavalla. Käsittelyyn voi kuulua esimerkiksi päällystys ja kiillotus. Kun paperi on käsitelty ja rullattu, se täytyy aukirullata, leikata ja rullata pienemmiksi asiakasrulliksi. [2], [3]

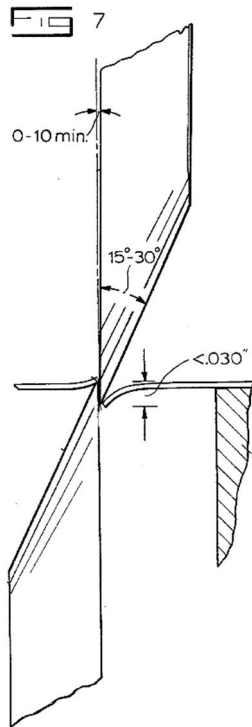


Kuva 1. Paperikoneen piirroskuva. Perustuu lähteeseen [3].

2.2 Pituusleikkuri

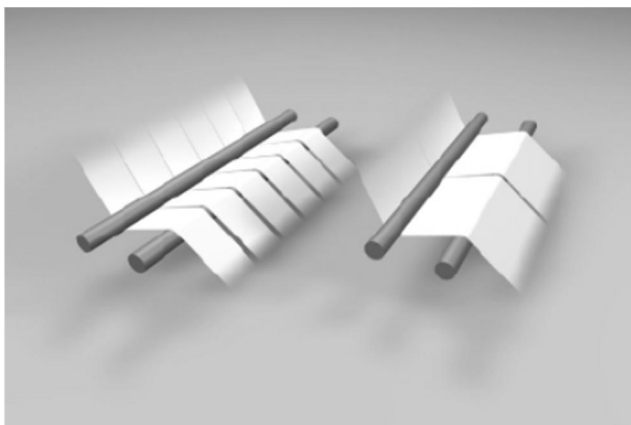
Paperikoneen pituusleikkuri leikkaa leveän paperirullan useaksi pienemmäksi rullaksi, jotta asiakas pystyy tekemään jälkikäsittelyn paperille. Paperin kuljetus on myös helpompaa paperirullien ollessa pienempiä. Yksi 1–3 metrin levyinen asiakasrulla voi painaa 200–2000 kilogrammaa riippuen rullan halkaisijasta ja paperin tiheydestä [4]. Paperitehtaassa on yleensä vain yksi pituusleikkuri paperikonetta kohden, minkä johdosta pituusleikkurin tuotantonopeuden tulee olla hieman korkeampi kuin paperikoneen. Varastotilaa ei yleensä ole niin paljon, että konerullia voisi säilyttää ennen niiden leikkaamista. Yleensä riittää, että pituusleikkuri on 30–40 % jatkuvakäyttöistä paperikonetta nopeampi. Pituusleikkurissa katkon tapahtuessa rainaa voidaan jatkaa kuumasaumautuvalla nauhalla. Asiakasrulliin ei näitä liitoksia kuitenkaan haluta, joten rulla yleensä hylätään saumauksen takia ja sekoitetaan hylkyyn käytettäväksi uudelleen raaka-aineena. [2, Luku 7.2] Pituusleikkurin nopeutta, kuten paperikoneenkin nopeuksia, ohjataan sähkökäyttöillä.

Pituusleikkuri koostuu aukirullaimesta, leikkausteristä ja kiinnirullauksesta. Pituusleikkurissa paperiraina ajetaan leikkausterien läpi, jotka koostuvat pyörivistä teloista. Leikkausterä koostuu ylä- ja alateristä [2, Luku 7.2]. Toisinaan käytetään myös laserleikkausta, vesisuihkuleikkausta tai muita erikoistekniikoita, mutta niiden käyttö on kuitenkin suhteellisen vähäistä [5, Luku 21.1]. Pituusleikkurissa myös konerullan reunat leikataan pois hylkyyn, koska paperiradan sivulinja vaeltaa eikä reuna ole laadukasta. Leikkausterien nopeuden tulee olla hieman suurempi kuin sille tulevan paperin, jotta leikkauksesta tulee siisti. Kuvassa 2 on esitelty leikkausterän toimintaperiaate pölyn määrän vähentämiseksi (patentti vuodelta 1979 [6]). Jos ylä- ja alaterä leikkaavat paperia limittäin kuten sakset, leikkauskohdasta ei tule täydellinen.



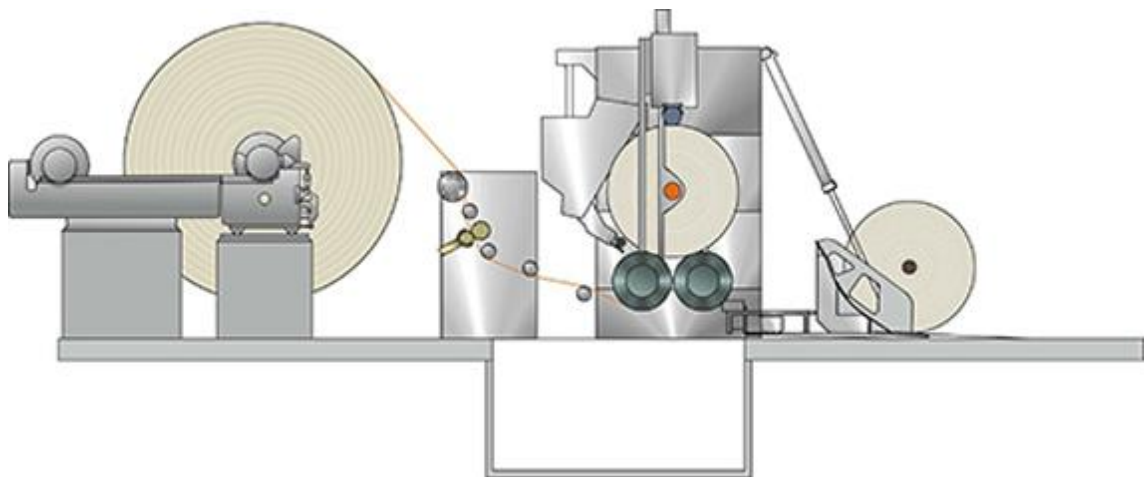
Kuva 2. Patenttikuva leikkurin toiminnasta. Perustuu lähteeseen [6]

Leikkausterien jälkeen leikatut rainat täytyy erottaa niin, että rainat eivät mene limittäin. Riittävä ratakreisy voi myös riittää erottamaan rainat. Yleisesti tähän käytetään kahta rinnakkaista levitystelaa kuten kuvassa 3 [5, Luku 21.1].



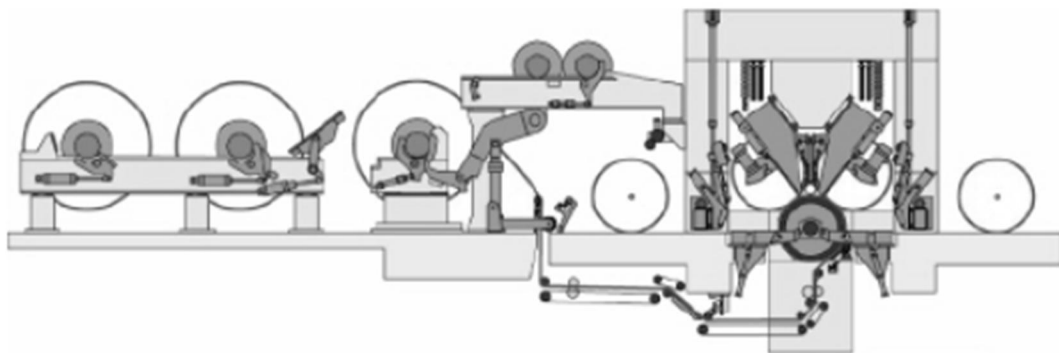
Kuva 3. Paperin levitystelat [5, Luku 21.1]

Pituusleikkuri on usein joko kantotelaleikkuri tai keskiörullain. Kantotelaleikkurissa asiakasrullat rakentuvat etu- ja takakantotelan päällä. Kantotelaleikkurin toiminta perustuu vetoon, momenttiin ja nippikuormaan [5, Luku 21.1]. Takakantotelan momenttia kasvatettaessa suhteessa etukantotelaan rullasta tulee tiukempi. Takakantotela on nopeussäädetty ja etukantotela on momenttisäädetty. Momenttien suhteella saadaan säädettyä rullan tiheyttä. Nippikuorman avulla mahdollistetaan momenttisuhteen luominen, kun rulla ei ole vielä tarpeeksi painava. Kuvassa 4 esitetään kantotelaleikkurin toimintaperiaate. Vasemmalla kuvassa näkyy konerulla, joka kulkee keltaisen leikkausterien läpi, minkä jälkeen rainat siirtyvät levitystelojen kautta kantotelojen päälle rullattavaksi asiakasrulliksi. Oikealla kuvassa on valmis asiakasrulla, jotka siirretään sivummalle kohti pakkausta.



Kuva 4. Kantotelaleikkurin toimintaperiaate [7].

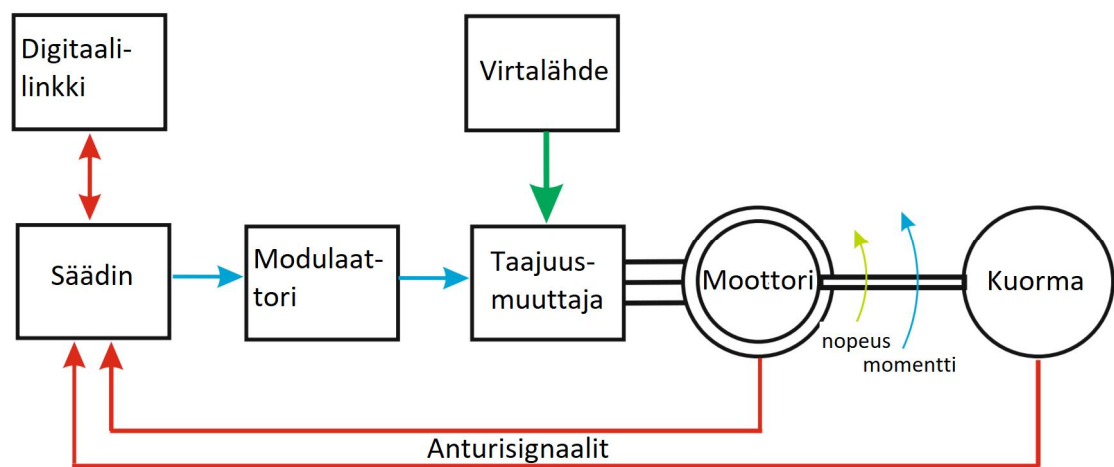
Keskiörullain on kalliimpi investointi kuin kantotelaleikkuri, mutta sillä saa tehtyä isompia asiakasrullia. Keskiörullaimessa asiakasrullat muodostuvat yhden telan päälle sen molemmille puolille. Joka toinen asiakasrulla muodostuu etupuolelle ja joka toinen takapuolelle. Kuvassa 5 on esitetty keskiörullaimen toimintaperiaate, jossa raina tulee telalle alhaalta. Valmiit asiakasrullat kerätään molemmille puolille telaa. [5, Luku 21.1]



Kuva 5. Keskiörullaimen toimintaperiaate [5, Luku 21.1].

2.3 Sähkökäytöt

Sähkökäyttö on sähköverkon ja koneen välinen energiamuunnin. Sähkökäyttöjä voidaan käyttää useissa eri sovelluksissa. Muita energiamuuntimia ovat esimerkiksi hydraulikkakäytöt, pneumatiikkakäytöt ja polttomoottorit. Sähkökäyttöjen etuja muihin muuntimiin verrattuna ovat esimerkiksi puhtaus, suuri käyttöalue, hiljaisuus, kuluvien osien vähäinen määrä ja nopea käynnistyminen [8]. Eräs iso etu on myös mahdollisuus pyörittää konetta eri suuntiin ja jarruttaa. Jarrutuksen energian saa kehittyneissä sähkökäytöissä takaisin verkkoon. Tyypillisesti sähkökäyttö konfiguroidaan taajuusmuuttajan parametreilla. Paperikoneissa sähkökäytöt on useimmiten toteutettu linjakäyttösovelluksina, jolloin prosessi sitoo käyttöryhmät toisiinsa. Tämän seurauksena käyttöryhmillä on yhteinen ohjauslogiikka. Sähkökäytön ohjaus voidaan jakaa prosessin, käytön, moottorin ja suuntaajan ohjaukseen. Kuvassa 6 on esitelty tyypillinen sähkökäytön järjestelmä. Moottori pyörittää kuormaa, joista molemmista anturit lähettävät signaalit säätimelle. Säädin kommunikoi ulkoisen tietokoneen kanssa ja laskee säätöparametreille uudet arvot. Modulaattori toimii säätimen ja taajuusmuuttajan kommunikointilinkkinä. Taajuusmuuttaja on sähköverkon ja moottorin välinen tehoelektrinen laite, joka manipuloi energiansiirtoa tällä välillä.



Kuva 6. Sähkökäytön tyypillinen järjestelmä. Muokattu lähteestä [8].

Sähkökäytön mekaaninen värähtely voi olla iso ongelma, koska se rasittaa ja häiritsee prosessia. Ääritapauksessa akselit voivat katketa, jolloin tuotanto pysähtyy ja syntyy tuotantotappioita. Mekaanisia värähtelyjä voidaan havaita moottorista, vaihteistosta, antureista ja taajuusmuuttajan mittauksista.

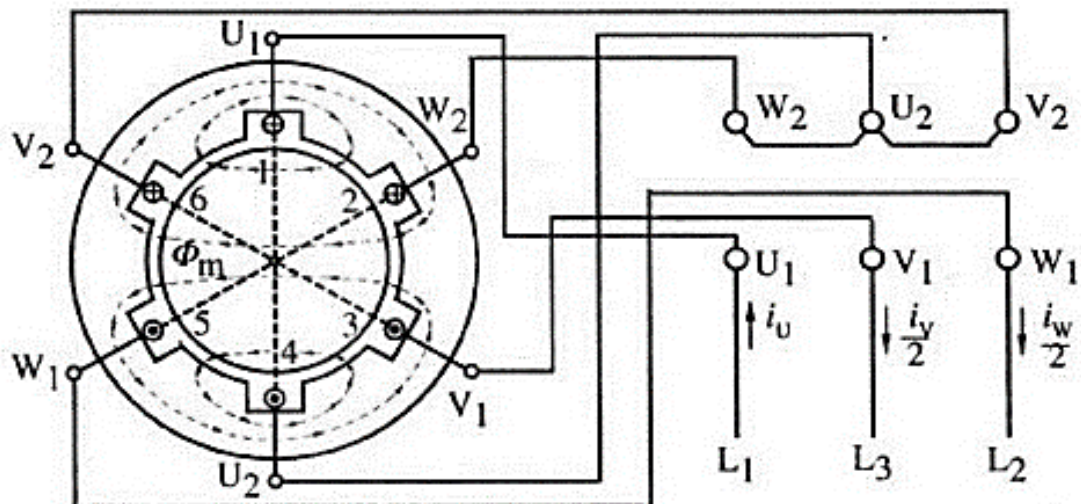
2.3.1 Moottorit

Sähkömoottorit ovat sähkökäytön osa, joka muuntaa sähköenergian pyörimisenergiaksi ja pyörittää paperikoneiden teloja. Tarvittaessa sähkömoottori voi myös ottaa telan

inertiasta muodostuvan jarrutusenergian talteen ja generaattorina syöttää energian takaisin sähköverkkoon. Sähkömoottorit ovat suosittuja niiden laajan ja portaattoman pyörimisnopeusalueen sekä säädettävyyden takia. Paperikoneiden sähkökäytöissä käytetään nykyään yleisimmin vaihtosähkökoneita ja vähenevässä määrin tasasähkömoottoreita.

Epätahtimoottori eli asynkronimoottori on vaihtosähkömoottori, ja sen toiminta perustuu pyörivään magneettikenttään. Nimensä mukaisesti epätahtimoottorin pyörimisnopeus poikkeaa kuormitettuna tahtinopeudesta: epätahtimoottoreissa moottorien pyörivä osa eli roottori ei pyöri magneettikentän kanssa samassa tahdissa toisin kuin tahtimoottorissa. Epätahtimoottorin oikea nopeus saadaan laskettua jättämällä avulla, joka kertoo, kuinka monta prosenttia roottorin nopeus on tahtinopeutta pienempi. Yhteistä kaikille epätahtimoottoreille on se, että vaihtosähköllä luodaan staattoriin pyörivä magneettikenttä. Kolmivaiheisen vaihtosähkömoottorin pyörivä magneettikenttä syntyy ilman lisälaitteita symmetrisen kolmivaihekäämityksen ja siinä kulkevan symmetrisen kolmivaihevirran avulla. [9, Luku 3.1]

Kuvassa 6 on kolmivaihemoottorin periaatekuva. Vaihekäämin muodostaa yksi vyyhti, jonka sivu on yhdessä urassa. Käämitys on kytketty tähtikytkentään yhdistämällä vaihekäämien loppupäät U_2 , V_2 ja W_2 yhteen. Kun vaihekäämien alkupäät U_1 , V_1 ja W_1 liitetään kolmivaihejohtoon L_1 , L_3 ja L_2 , käämityksen läpi alkaa kulkea kolmivaihevirta, joka symmetrisessä tapauksessa on kuvan 7 mukainen.



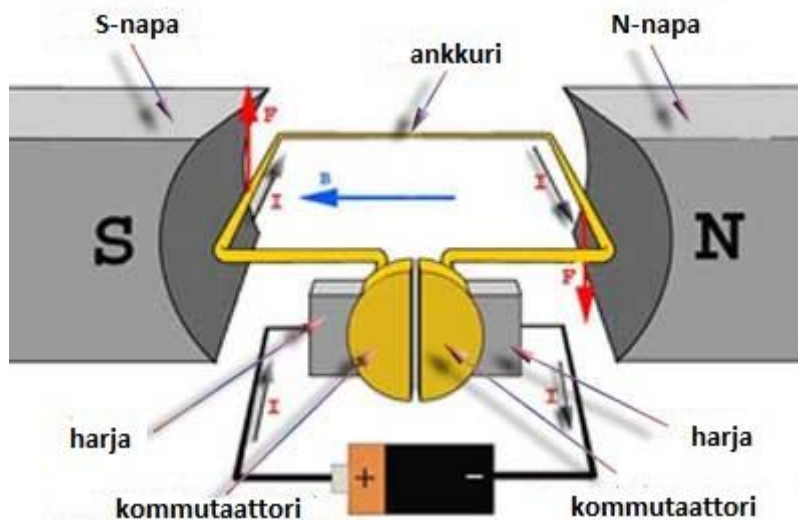
Kuva 7. Kaavioesitys kolmivaihekäämityksen synnyttämästä pyörivästä magneettikentästä [9, s. 120].

Kestomagneettimoottori on vaihtosähkömoottori, jonka staattorissa on käämitykset ja roottorissa kestopagneetit ja sen avulla saadaan aikaan suuri momentti pienilläkin

kierroksilla. Kestomagneettimoottori ei tarvitse erillistä vaihdetta, koska se voidaan suunnitella luonnostaan pienille nimellispyörimisnopeuksille. Vaihteiston puuttuminen on iso etu huollon kannalta, koska pyörivien osien määrä vähenee. Muita kestomagneettimoottorin hyötyjä epätahtimoottorin sovelluksiin verrattuna ovat pienempi varaosien määrä, helpompi asennus ja vaihteiston vähäisemmät energiahäviöt. [10]

Tasasähkökone voi olla joko moottori tai generaattori, koska näiden rakenteet eivät eroa toisistaan. Tasasähkömoottorin staattori on täysrautaa ja se on magnetoitu sähköisesti. Siten moottori muodostaa tasamagneettikenttä kehänsä sisälle. Tasasähkömoottorin pyörijä on kytketty tasasähkövirtalähteeseen. Pyörijän kommutaattori muuntaa ankkuriin tulevan tasasähkön vaihtosähköksi. Pyörijässä on ankkurikämmityksiä, joihin tasamagneettikenttä luo induktion. Indusoituva ankkuri pyörii tasasähkökentässä. Tasasähkömoottorissa on yleensä useita napapareja, jotta pyöriminen on mahdollisimman tasaista. [9, Luku 5]

Kuvassa 8 on esitelty tasasähkömoottorien toiminnan yksinkertaistettu periaate. Staattorin muodostama tasamagneettikenttä muodostuu N- ja S-napojen välille. Ankkuriin syntyy kommutaattorin synnyttämä vaihtojännite, joka saa ankkurin pyörimään liikkeeseen.



Kuva 8. Tasasähkömoottorin periaatekuva. Perustuu lähteeseen [11].

2.3.2 Kytkimet ja vaihteistot

Kytinten tarkoituksena on liittää kaksi pyörivää akselia toisiinsa. Usein kyseessä on moottorin ja kuorman akselit. Kytkimiä on kolmen tyyppisiä: kiinteitä, jäykkiä ja joustavia. Kiinteät kytkimet on useimmiten kiinnitetty pulteilla laippaliitoksiksi. Kiinteiden kytkimien akselit tulee kohdistaa tarkasti, joten niiden käyttö on

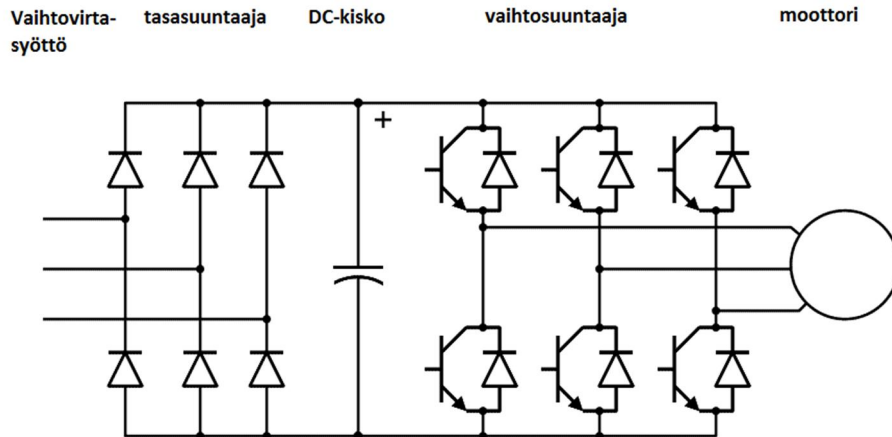
sähkökäytöissä vähäistä. Jäykät kytkimet ovat usein kaarihammaskytкимиä, joissa hammastetut laipat akseleiden päässä yhdistetään holkeilla, joiden sisäpinnoissa on samanlaiset hammastukset. Tällainen rakenne on jäykkä väännön suhteen. Joustavat kytkimet vaimentavat vääntömomentti-iskuja ja -värähtelyitä. Joustavat kytkimet voidaan valmistaa esimerkiksi metallista tai kumista. Materiaalien ominaisuuksien erot ovat vaimennuksessa ja kestävyudessa. Sekä jäykällä että joustavilla kytkimillä voi ilmetä välystä, eli momentin suunnan vaihtuessa kontakti akselien välillä saattaa kadota hetkeksi. [12, Luku 1.5].

Vaihteita käytetään moottorien ja kuormien pyörimisnopeuksien sovittamiseen toisiinsa ja voimansiirtoon akselista toiselle. Vaihteen sijaan voidaan käyttää myös hihnavälitystä. Tyypillisesti moottorin optimaalinen pyörimisnopeus on nopeampi kuin kuorman optimaalinen pyörimisnopeus, jolloin vaihteistoa tai hihnavälitystä käytetään hidastamaan pyörimisnopeutta. Hihnavälityksen etuna on hyvä asennusepätarckkuuden sieto, äänetön käynti, halpuus ja nykäyksien vaimennus. Hihnavälitys ei tarvitse myöskään erikseen kytkintä. Hammasvaihteen etuna on luistoton toiminta, suuri tehonvälityskyky, pieni tilantarve ja reagoimattomuus kemikaaleille. [12, Luku 1.5].

2.3.3 Taajuusmuuttajat

Taajuusmuuttajan tehtävä on välittää moottorille sen vaatima energia. Taajuusmuuttaja toimii ikään kuin sähkömoottorin vaihdelaatikkona. Ilman taajuusmuuttajaa sähkömoottori pyörisi täydellä nopeudella ja säätö tehtäisiin mekaanisilla säätimillä prosesseissa. Näin ollen sähkömoottori ei kuluta yhtä paljon sähköenergiaa taajuusmuuttajan kanssa. Taajuusmuuttaja siis pienentää energiankulutusta ja pidentää laitteiden käyttöikää moottorin pyöriessä taloudellisemmin. [9, ss. 458–562]

Taajuusmuuttaja ottaa vaihtosähköverkosta virtaa ja muuttaa sen tasasuuntaajilla tasasähköksi. Kuvassa 9 on esitetty taajuusmuuttajan periaate. Vaihtovirtasyötöstä tuleva sähkö muunnetaan tasasähköksi tasasuuntaimilla. Tasasähkö kulkee DC-kiskon yli vaihtosuuntaajille, jotka muuntavat tasasähkön taas vaihtosähköksi. Vaihtosähkö saadaan näin moottorille sopivaksi.

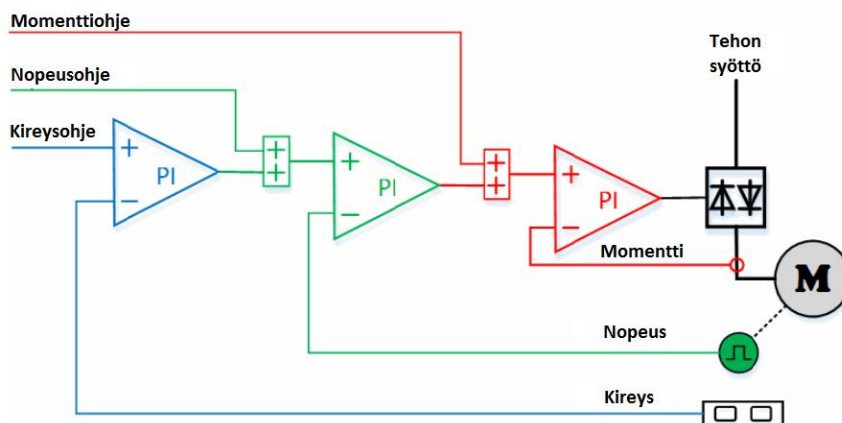


Kuva 9. Taajuusmuuttajan periaatekuva [13], termit suomennettu.

2.3.4 Ohjausjärjestelmät

Paperin valmistuksessa sähkökäytöt ovat suuria linjakäyttöjä ja niiden on toimittava sulavasti yhteisymmärryksessä, jotta paperirata ei katkea eikä tuotanto pysähdy. Yleensä ohjausjärjestelmät toimivat ohjelmoitavalla logiikalla.

Sähkökäyttöjen säätäminen on tehokas tapa hallita parametrien ja kuorman muutosten häiriöitä [12, Luku 7]. Yleisin tapa säätää on PI-säätö sen yksinkertaisuuden ja hyvän sopivuuden takia. Nopeussäätö on keskeisin kaikista sähkökäytön säädöistä. Nopeussäätöä vaikeuttavat kuorman muutosvaihtelut sekä akselijärjestelmien mekaaniset resonanssit. Sähkökoneiden vuota yleisesti ohjataan eikä säädetä, koska sähkökoneen voi mallintaa melko tarkasti ja vuota on vaikea mitata suoraan. Kuvassa 10 on esitelty taajuusmuuttajan säätörakenne. Moottorille tulevaa tehoa ohjataan momentin, nopeuden ja kireyden olo- ja asetusarvoilla. Ainoastaan kireyssäätö tapahtuu kuvan 6 esittämässä säätäjässä.



Kuva 10. Taajuusmuuttajan säätörakenne.

Sähkökäyttöjen tärkeimmät mekaaniset anturit ovat nopeutta, asentoa ja voimaa mittaavia. Moottorien kulmanopeuksia mitataan kulmanopeusantureilla tai pulssiantureilla. Pulssiantureilta saatavilla mittauksilla saadaan laskettua roottorin asento ja nopeus. Ratakireyttä voi mitata voima-antureilla. Sähköisiä suureita eli virtaa ja jännitettä mitataan useista kohdista sähkökäytöissä. Taajuusmuuttajalta mitataan välipiirin jännite ja moottorin vaiheiden virrat. Verkkovirta ja -jännite, moottorin pääjännite ja verkon vaihevirtojen summa ovat myös tärkeitä mitattavia sähköisiä suureita. Sähkövirta voidaan mitata shunttivastuksella tai virtamuuntimella. [12, Luku 8]

Moottorin lämpötila on hyvä mitata, koska lämpötila vaikuttaa paljon sähkökäytön hyötysuhteeseen ja tarkkuuteen. Lämpötilan mittauksella pyritään myös estämään laitteiden ylikuumentuminen. Lämpötila on hyvä mitata ainakin moottorista sekä sähkökäytön tehoasteista. Vastuslämpötila-anturilla, jonka resistanssi vaihtelee lämpötilan vaihteluiden mukaan, saadaan mitattua sähkökäytön lämpötiloja helposti. Yleisin vastuslämpötila-anturi on PT100-anturi, joka on valmistettu platinasta ja jonka resistanssi on 100 Ω , kun lämpötila on nolla celsiusastetta. Sen etuna on lähes lineaarinen suhde lämpötilan ja vastuksen välillä. [12, Luku 8]

Tiedonsiirto on digitaalisessa tehtaassa yhä suuremmissa asemassa. Dataa kertyy paljon, ja sitä kerätään talteen analyysia varten. Tiedonsiirtoa tarvitaan käytön sisällä varsinkin anturiviestien ja puolijohdetehokytkimien ohjausviestien välittämiseksi. Tiedonsiirron siirtotie koostuu yleensä useista vastaanottimista ja lähettimistä, jolloin sitä kutsutaan väyläksi. Tieto voi myös liikkua pisteestä pisteeseen muodostaen suoran yhteyden verkkolaitteiden välille. Digitaaliset viestit kulkevat väylissä protokollien eli yhteyskäytäntöjen määrittelemällä tavalla. Protokollat määrittelevät, kuinka väylän sisältö muodostetaan ja tulkitaan. Tiedonsiirrolta vaaditaan usein virhevapaita ja nopeita välityksiä, sillä heikko viestinvälitys heikentää prosessin turvallisuutta ja laatua, mikä voi johtaa pahimmillaan henkilö- ja laitevahinkoihin. Tiedonsiirtoon liittyy usein myös sähkömagneettisia häiriöitä. Häiriöitä voi vähentää käyttämällä esimerkiksi optisia kaapeleita, suuria signaalitasoja tai parikerrettyä tai parisuojattua kaapelia. [12, Luku 9]

Tiedonsiirto toteutetaan yleensä jonkinlaisella verkkotopologialla. Yksinkertaisin verkkotopologia on kahden laitteen välinen yhteys. Jos yhteen keskusyksikköön yhdistetään useita laitteita, kyseessä on tähtitopologia. Laitteet voidaan myös kytkeä renkaaksi, jolloin yhteys laitteisiin säilyy, vaikka yksi yhteys katkeaisi. Yhtä kaapelia, johon kaikki laitteet kytketään, kutsutaan väyläksi. Verkossa voi olla myös keskuslaitteita (master tai server) ja muita laitteita (slave, client), jotka keskustelevat hierarkian mukaisessa järjestyksessä. Slave-laitteet eivät tee keskustelun aloitteita vaan vastaavat keskuslaitteiden kyselyihin. [12, Luku 9]

Tehdasautomaation alimman tason tiedonsiirrossa käytetään usein kenttäväyliä [12, Luku 9.5]. Kenttäväylä välittää antureiden, toimilaitteiden, tulo- ja lähtöyksiköiden,

ohjelmoitavien logiikoiden ja ohjausjärjestelmien viestejä. Kenttäväylästandardeja on useita, joista esimerkkejä ovat PROFIBUS, FIELDBUS Foundation ja CAN. Teollisuusautomaatioissa yli puolet yhteyksistä muodostetaan kenttäväylillä. Noin 40 % tiedonsiirrosta tapahtuu Ethernetillä. Langattomia yhteyksiä ei ole vielä tehdasympäristössä yleisesti käytössä. Tehdasympäristöissä on suuret määrät sähkölaitteita ja rautaisia elementtejä, jotka häiritsevät langattomia signaaleja. HMS Industrial Networks'n mukaan langattomia yhteyksiä on tällä hetkellä noin 5 % tehdasympäristöiden kenttäväylistä [14].

Sähkökäyttöä ohjataan usein erilaisten sovellusten avulla. Sovellusohjelmien ja laitteiston välillä toimii käyttöjärjestelmä. Käyttöjärjestelmän tehtäviä ovat sovellusohjelmien suoritusten tulkinta, ohjelmien välisen tiedonsiirron hallinta, aikapalvelut, muistinhallinta, poikkeustilanteiden hoito ja oheislaitteiden ohjaus. Sähkökäytön käyttöjärjestelmältä vaaditaan tarkkaa reaaliaikaisuutta, sillä lyhyet suoritusvälit ovat niille tyypillisiä. Sähkökäytöt ja niiden ohjelmistot tulee testata. Testauksessa on tarpeellista seurata alan standardeja, jotta sähkökäyttö todetaan turvalliseksi, eikä vahinkoja synny. [12, Luku 10]

2.3.5 Mekaaniset värähtelyt

Mekaaniset värähtelyt ovat mekaanisen järjestelmän osien keskinäisiä jaksollisia liikkeitä. Mekaaninen järjestelmä pyrkii värähtelemään järjestelmän resonanssitaajuuksilla. Jos mekaaniseen järjestelmään kohdistuu pakotettu värähtely, se voi aiheuttaa järjestelmälle suuria rasituksia, mikäli värähtelytaajuus on lähellä resonanssitaajuutta. Ainoastaan mekaniikan vääntövaimennus voi rajoittaa amplitudin kasvua. Rasitukset aiheuttavat järjestelmissä materiaalin väsymistä, mikä johtaa rikkoihin. [9, Luku 8]

Paperikoneen sähkömoottorikäyttö on monimutkainen järjestelmä ja siihen liittyy useita värähtelytyyppejä. Kieppuminen on moottorin akselin värähtelyä. Pitkät ja ohuet kappaleet altistuvat kieppumiselle enemmän kuin jäykät lyhyet kappaleet. Moottorin pyörimisnopeus vaikuttaa kieppumiseen. Jos moottorin akselia ei ole tasapainotettu kunnolla ja kieppumisvärähtely on resonanssitaajuutta lähellä, värähtely voimistuu keskipakoisvoiman avulla vakavinkin seurauksin. Resonanssitaajuudet voidaan ohittaa nopeasti, jolloin ne eivät ehdi aiheuttamaan häiriöitä. [9, Luku 8]

Akselien osien kiertymistä toistensa suhteen kutsutaan vääntövärähtelyksi. Vääntövärähtelyistä seuraa momentin muutos. Vääntövärähtelyitä voidaan vaimentaa nopeussäädöllä tai akseleiden kytkimillä. Vääntövärähtely on sähkökäytöissä tärkeä värähtelytyyppi, koska se vaikuttaa sekä momenttiin että nopeuteen, jotka ovat merkittäviä säätösuureita. [9, Luku 8]

Perustuksen värähtelyt ovat sähkökäytön alustan värähtelyitä. Mikään alusta ei voi olla äärettömän jäykkä, ja tästä syystä sähkökäyttö aiheuttaa värähtelyitä myös alustaan. Koska alusta kokee värähtelyitä eri suuntiin, sillä voi olla erisuuruisia resonanssitaajuuksia. [9, Luku 8]

3. TYÖKALUT

3.1 Tekoäly

Tekoäly on kehittynyt sitä mukaa kuin tietokoneetkin. Tekoälyn määrittely ei ole yksiselitteinen, mutta määrittelyn apuna voi käyttää kahta kysymystä:

1. Onko tekoäly rationaalinen vai inhimillinen?
2. Ajatteleeko vai käyttäytyykö tekoäly?

Näiden kysymysten perusteella tekoälyn määrittelyä voidaan lähestyä neljästä eri näkökulmasta, jotka on esitelty taulukossa 1 [15]:

Taulukko 1. Tekoälyn määritelmien lähestymistavat.

Systeemi ajattelee kuin ihminen	Systeemi ajattelee rationaalisesti
Systeemi käyttäytyy kuin ihminen	Systeemi käyttäytyy rationaalisesti

Ensimmäinen lähestymistapa on tutkia koneen ajattelutavan inhimillisyyttä, eli koneen taipumusta ratkoa samanlaisia ongelmia samoilla menetelmillä kuin ihminen. Toinen ja kenties käytetyin lähestymistapa on koneen käyttäytymisen inhimillisuus. Alan Turingin kehittämä Turingin testi määrittelee älykkyyden sen mukaan, kuinka hyvin systeemi pystyy imitoimaan ihmisen käyttäytymistä. Läpäistäkseen Turingin testin tietokoneella tulee olla kyky luonnolliseen kielen prosessointiin, tiedon esittämiseen, automaattiseen päättelyyn ja koneoppimiseen [15].

Kolmas lähestymistapa on ajattelumallien rationaalisuus. Kone siis toimii täysin loogisesti logiikan sääntöjen mukaan. Tämä lähestymistapa on kuitenkin ongelmallinen, sillä ongelmien ratkaisu on yleensä huomattavasti monimutkaisempaa käytännössä kuin teoriassa.

Neljäs lähestymistapa on käyttäytymisen rationaalisuus. Kone siis tekee sen, mitä sen halutaan tekevän. Yleensä se on juuri inhimillistä tekemistä, mutta siihen eivät vaikuta

inhimilliset virheet, ellei niitä siltä haluta ja vaadita. Esimerkiksi itseohjautuvien autojen tahdotaan toimivan rationaalisesti, jotta ne vähentäisivät ihmisten ajovirheiden määrää.

Tekoäly on edistynyt merkittävästi lähiaikoina. Nykyään tietokone pystyy voittamaan ihmisen lähes kaikissa loogista päättelyä vaativissa peleissä, kuten shakissa ja gossa [16]. Tietokone pystyy oppimaan ja tekemään päätöksiä oppimiensa asioiden perusteella. Tietokone toimii usein tehokkaammin ja virheettömämmin kuin ihminen, ja siksi tietokoneohjatuilla roboteilla halutaan korvata ihmisille vaaralliset työt.

Jotta ihmiset hyötyisivät mahdollisimman paljon tekoälyn kehityksestä, tekoälyn on opittava ymmärtämään ja tuottamaan ihmiselle ymmärrettävää materiaalia kuten kuvia, puhetta tai tekstiä. Tämä voi olla kuitenkin haastavaa. Jos esimerkiksi kielen oppimisen perustana käytetään sosiaalista mediaa, tekoäly voi oppia tuottamaan rasistista tai muuten asiattonta sisältöä [17]. Kehittyneet tekoälyt pystyvät kuitenkin työskentelemään ihmisten kanssa sujuvasti, koska ne pystyvät käyttämään samoja tietolähteitä kuin ihmiset. Tekoälyn ja koneoppimisen avulla ihmisten ostoskäyttäytymistä ja nettihistoriaa analysoimalla markkinoijat voivat kohdistaa mainokset oikeille henkilöille [18].

3.1.1 Havaitseminen

Tekoälyn yksi tärkeä ominaisuus on havaitseminen. Ilman havaintoja tuloksien ja päätelmien tuottaminen on pelkkää arvailua. Havainnointi voidaan tehdä esimerkiksi konenäöllä, puheenkäsittelyllä, neuroverkoilla ja hahmontunnistuksella [19]. Data pitää pystyä lokeroimaan oikein, minkä jälkeen sen perusteella ryhdytään jatkotoimenpiteisiin. Hahmontunnistuksella tarkoitetaan luonnollisista kohteista tehtyjä mittauksia ja havaintoja, kohteiden mittauksen automaattista analyysiä sekä kohteiden tunnistusta analyysin perusteella [20].

Konenäkö pyrkii vastaamaan ihmisen näköaistia ja kameroiden avulla havaitsemaan ja tunnistamaan erilaisia kohteita. Nykyään konenäkö on laajalti käytössä teollisuudessa viantunnistuksessa ja esimerkiksi itseajavissa autoissa. Nykyaikaiset reaaliaikaiset konenäkösovellukset [21] pystyvät tunnistamaan nopeasti esimerkiksi CAPTCHA-tunnistuksia ja etsintäkuulutettuja rikollisia valvontakameroiden avulla.

3.1.2 Ymmärtäminen

Pelkkä datan havainnointi ei riitä, vaan dataa pitää pystyä myös tulkitsemaan. Yksi suuri tekoälyn haaste on ymmärtää ja tuottaa ihmisen kieltä. Kieliä on maailmassa tuhansia, ja kaikissa on oma kielioppinsa ja sanastonsa. Tietokone pystyy käsittelemään sekunnissa tuhansia tiedostoja, joiden käsittelyyn ihmisellä menisi jopa vuosia. Esimerkiksi potilastietokantoja pystytään lukemaan ja analysoimaan erittäin nopeasti. Tietokoneen laskentanopeus ihmiseen verrattuna on suuri kannustin tekoälyn kehittämiseen. [19]

Kieli voidaan erotella eri tasoille, jotta ymmärtämisen opettelu helpottuisi. Kielen eri tasot voidaan luokitella viiteen tasoon: fonologia, morfologia, syntaksi, semantiikka ja pragmatiikka. Jokaisen tason kohdalla on tietokoneella erilaisia haasteita ymmärtää kieltä. Teknologia on kehittynyt hurjasti viimeisen kahdenkymmenen vuoden ajan luonnollisen kielen ymmärtämisessä. Jotkin sovellukset osaavat nykyään jo natiivisti puhua ja ymmärtää eri kieliä. Googlen Translaten käännökset olivat erittäin alkeellisella tasolla vielä 10 vuotta sitten, mutta nykyään käännöspalvelut kehittyvät nopeasti. Tietokoneen ymmärrys on lähiaikoina kehittynyt paljon ja varsinkin syväneuroverkkojen avulla ollaan saatu tuloksia aikaan. Syväneuroverkkojen avulla tietokone voi käsitellä lukuisia dokumentteja ja ymmärtää eri asioiden korrelaatiot. Dan Ciraganin ym. [22] artikkelissa on testattu syväneuroverkkojen käyttöä erilaisissa hahmontunnistuksissa ja tulokset olivat parempi kuin aiemmat. Syväneuroverkkojen tehokkaan ymmärtämisen ansiosta sitä käytetään tunnistamaan esimerkiksi käsinkirjoitettuja tekstejä ja liikennemerkkejä. Näistä on hyötyä, kun kehitetään käännössovelluksia ja itseohjautuvia autoja.

Tiedon voi myös ymmärtää loogisesti. Klassinen looginen ajattelu on usein mustavalkoista, eikä se kuvaa reaali prosesseja tarkasti. Klassisen logiikan mukaan tieto on joko totta tai sitten ei. Loogisen tiedon esittämiseen liittyy vahvasti päättely.

3.1.3 Oppiminen

Oppiminen on älykkyyden keskeisimpiä osa-alueita. Tekoälyn voidaan sanoa oppivan, kun se pystyy laajentamaan tietämystään tai kykyään suorittaa omia tehtäviään. Tekoälyn oppiminen voi olla ihmislähtöistä tai konelähtöistä. Koneoppiminen on tehokkaampi keino kuin ihmisopetus, jos opetettava asia on matemaattisesti riittävän yksinkertainen. Tietokoneen on helpompi löytää eri asioiden korrelaatiota suuresta määrästä dataa, kun taas ihmisellä siinä voi kestää kauankin. Ihminen voi esimerkiksi opettaa käsin ohjaamalla tietyt liikeradat robotille, joiden mukaan roboti jatkossa aina liikkuu. Robotin liikeradat voidaan myös ohjelmoida robotille. Koneoppimisen tapauksessa tietokone voi oppia esimerkiksi tunnistamaan annettujen kuvien perusteella, onko kuvassa koira vai susi.

Oppiminen voi olla ohjattua, vahvistettua tai ohjaamatonta. Ohjattu oppiminen tapahtuu ihmisen opettamana, jolloin tietokone ei pääse oppimaan omatoimisesti. Ohjatussa oppimisessa on tietokoneelle annettu sekä data, josta kone oppii halutut tulokset, että ratkaisu, jota oppimiselta odotetaan. Vahvistettu oppiminen ei enää anna tarkkaa haluttua tulosta oppimiselle. Oppimisprosessin aikana tietokoneelle annetaan kuitenkin positiivista palautetta, jos data on ollut hyvä. Ohjaamatonta oppimista voi kutsua myös itsenäiseksi oppimiseksi. Tietokone määrittelee itse mallin datasta ilman palautetta mallin

hyvyydestä. Tietokone voi myös päätyä erilaisiin malliratkaisuihin samalla datalla. [15], [19]

3.1.4 Tulevaisuus

Tekoäly on puhuttanut paljon populaarikulttuurissa sekä mediassa. Viihdekulttuurissa on lukemattomia dystopisia esimerkkejä robottien valtaannoususta. Esimerkiksi elokuvissa WALL•E ja Terminator teknologian kehitys on johtanut tilanteeseen, jossa robotit uhkaavat ihmiskuntaa tai ovat vallanneet yhteiskunnan täysin. Monet merkittävät tieteen ja tekniikan henkilöt kuten Elon Musk [23], Bill Gates [24] ja Ray Kurzweil [25] ovat todenneet, että tekoäly on todennäköinen uhka ihmiskunnalle. Myös Vladimir Putin on todennut, että tekoälykilpailun voittaja tulee hallitsemaan maailmaa [26].

On pelätty, että kun tietokone alkaa oppia itsenäisesti ja koodata omia ohjelmiaan, se pystyy nopeasti kehittymään entisestään, koska tietämys kasvaa maailmassa eksponentiaalisesti. Ray Kurzweil, joka vuonna 1988 ennusti internetin kovan nousun 2000-luvun taitteessa, on ennustanut, että ihmiskunta kykenee kehittämään AGI:n (Artificial General Intelligence). Tekoälyn kehittyessä työ muuttuu muotoaan valtavasti. Korkeakoulutetut insinööritkin saattavat joutua miettimään työuraansa uusiksi.

Tekoäly tuo kuitenkin paljon hyviäkin asioita. Työt, jotka vaativat paljon toistuvia toimenpiteitä, voidaan poistaa robottien ja tekoälyn ansiosta. Ihmiset pystyvät kommunikoimaan toistensa kanssa helpommin, kun kielen kääntäjäohjelmistot kehittyvät vauhdilla. Prosesseja voidaan optimoida tehokkaammiksi ja vähemmän resursseja vieviksi datan ja koneoppimisen ansiosta. Tekoäly mahdollistaa paljon hyviä asioita, kunhan sen tavoitteet on määritelty hyviksi.

3.2 IoT

IoT (Internet of Things, suomeksi esineiden internet) viime vuosina yhä enenevässä määrin yleistynyt termi, jonka Kevin Ashton loi vuonna 1999 [27]. IoT:n perusidea on se, että suuret määrät elektronia laitteita, kuten mobiililaitteita ja antureita kytketään keskenään langattomasti, ja yhdessä laitteet toimivat yhteisen tavoitteen saavuttamiseksi [28].

Älyä on tuotu kaikkialle: jääkaappeihin, pesukoneisiin ja paperikoneen sähkökäyttöihin. Kaikista laitteista saadaan kerättyä dataa lukuisin anturein. Tämä vaatii valtavat määrät tiedonsiirto- ja palvelinkapasiteettia, mikä kasvattaa sähkönkulutusta. Hyödyt voivat kuitenkin olla viime kädessä energiaa säästäviä, sillä datan ansiosta laitteita voi käyttää optimiparametreilla, jotka säästävät energiaa. IoT:n lisäksi puhutaan myös käsitteistä IoP (Internet of People), IoS (Internet of Services) ja IoE (Internet of Everything). Älykkäiden verkkojen yhteistyö voidaan jakaa kolmeen osa-alueeseen: ihmisten väliseen, ihmisen ja

koneen väliseen sekä koneiden väliseen, joista viimeistä käytetään yleensä IoT:n yhteydessä [29].

IoT-verkoissa käytetään yhä enemmän väliohjelmistoja, jotka jakavat teknologian sovelluskohtaisiin komponentteihin helpottaen siten yksittäisten sovellusten kehittämistä. Tämä on niin sanottu palveluorientoitunut malli (SOA, service oriented architecture). Väliohjelmisto toimii käyttäjäliittymän ja itse laitteiden välissä ja on yleensä piilossa sekä kehittäjiltä että käyttäjiltä. Väliohjelmisto koostuu yleensä palvelun kokoamisesta, palvelun hallinnasta ja esineiden erottelusta. Palvelun kokoaminen tarkoittaa keskenään verkottuneiden kohteiden kokoamista yhtenäisiksi sovelluksiksi. Palvelujen hallinnalla varmistetaan, että jokaisella kohteella on sille määritellyt ominaisuudet, jotka mahdollistavat sen hallitsemisen osana IoT-kokonaisuutta. Esineiden erottelussa kohteille määritellään yhdenmukaiset toiminnot, joiden avulla niitä voidaan hallita tehokkaammin. [28]

Reaalimaailmaa digitalisoitaessa tärkeimpiä nykYTEknologioita ovat radiotaajuinen etätunnistus (RFID, radio frequency identification), lähikenttäviestintä (NFC, near field communication) ja langattomat sensori- ja toimilaitteverkot (WSAN, wireless sensor and actuator networks). Muita tärkeitä IoT:ta kannattelevia teknologioita ovat robotiikka, data-analyysi, koneoppiminen, M2M-kommunikaatio, pilvilaskenta ja älyrakennustekniikat [30].

RFID sisältää antennin, jotta se voi lähettää ja vastaanottaa radiosignaaleja. Tähän teknologiaan perustuen laitteet voidaan tunnistaa ja paikantaa. RFID voi olla joko passiivinen, semipassiivinen tai aktiivinen. Useimmiten RFID on passiivinen, mikä tarkoittaa sitä, että siinä ei ole virtalähdettä kuten paristoa. Passiivinen RFID saa tarvittavan lähetysvirtansa induktiolla. Semipassiivinen RFID käyttää paristoa vain silloin, kun se saa signaalin lukijalta. Aktiivinen RFID on jatkuvasti pariston avulla toiminnassa, ja sillä on laajin radiokantama. RFID-tekniikkaa käytetään nykyään esimerkiksi kauppojen varkaudenestojärjestelmissä sekä kirjastojen lainausjärjestelmissä. [28]

NFC hyödyntää RFID-tekniikkaa, mutta se pystyy lisäksi myös itse tunnistamaan laitteita. NFC:n avulla laitteet tunnistavat toisensa ja pystyvät lähietäisyydellä välittämään paljon tietoa. NFC on nykyään käytössä esimerkiksi maksukorttien ja kännyköiden lähimaksuissa. [28]

Sähkökäytöissä dataa harvemmin siirretään RFID-tekniikalla. Data, jota hyödynnetään, saadaan antureilta ja ohjausjärjestelmältä. RFID voisi kuitenkin lähettää tietonsa nopeasti verkkoon, jolloin äkkinäiset rikkoutumiset voitaisiin havaita nopeammin. Sensoriverkot ovat usean solmun muodostamia langattomia verkkoja. Usein solmut raportoivat yhteen erikoissolmuun. Sensoriverkon solmujen saumaton liittäminen internetiin nykyisillä

menetelmillä on kuitenkin ongelmallista muun muassa siksi, ettei IP-osoiteavaruudessa ole lähimainkaan riittävästi osoitteita. [28]

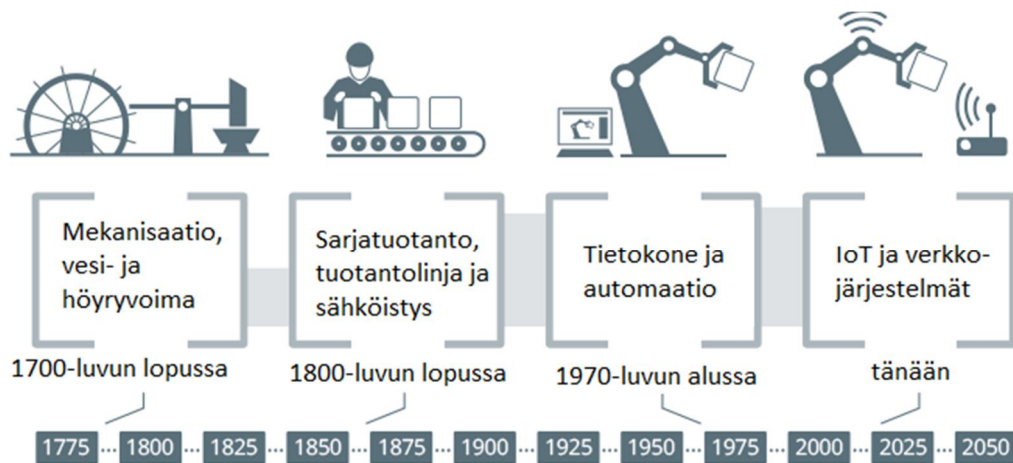
Suuri ongelma IoT:n kasvamisessa on internetin valmius valtavaan määrään dataa. Nykyiset IP-osoitteet IPv4-protokollassa tarjoavat vain noin neljä miljardia solmua. Sen seuraajaksi kehitelty IPv6-protokolla tarjoaisi jo 340 sekstiljoonaa (10^{38}) solmua. IoT tarvitsee huomattavasti tehokkaamman kuljetusprotokollan kuin tämän hetken yleisin TCP (transmission control protocol). TCP yhdistää laitteet verkossa kolmen kädenpuristuksen avulla, mikä on hidasta ja IoT:n kannalta tarpeetonta. TCP myös ehkäisee ruuhkan syntymistä, mikä on IoT:n tapauksessa tehokkuutta syövä ominaisuus. TCP-protokolla vaatii datan jäävän puskuriksi sekä lähetys- että vastaanottovaiheessa, mikä voi olla ongelma passiivisissa RFID-laitteissa.

Toinen IoT:n yleistymisestä seuraava ongelma on tietoturva. IoT:n komponentit ovat usein vailla valvontaa, langattomia laitteita on helppo kuunnella ja pienen energia- ja laskentatehon takia ne eivät pysty tarjoamaan monimutkaisia suojauksia. IoT-komponentit voivat altistua myös kolmannen osapuolen hyökkäysten kohteeksi. Esimerkiksi itseohjautuvan auton ei haluta olevan yhteydessä kaikkialle, jottei sitä voi kaapata. Varkaudet ja identiteettivarkaudet ovat lisääntyneet, kun NFC on yleistynyt maksukorteissa. NFC-kortin voi nykyään suojata vain peittämällä kortin esimerkiksi foliolla. IoT ei ole vielä saanut omaa standardiaan, mikä vaikeuttaa yleistä kehitystyötä. [28]

3.2.1 Teollisuus 4.0

Yritysmaailmassa puhutaan nykyään paljon uudesta teollisuusvallankumouksesta, joka on neljäs laatuaan. Kuvassa 10 on esitelty neljä teollista vallankumousta, jotka ovat:

1. Mekanisaatio, höyry- ja vesivoima,
2. Massatuotanto ja sähkö,
3. Tietokone ja automaatio sekä
4. IoT ja verkkojärjestelmät.

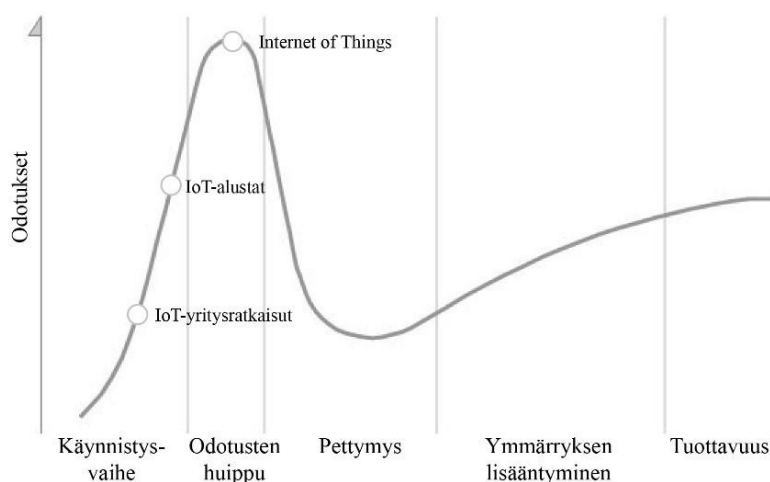


Kuva 11. Teollisuus 4.0. Muokattu lähteestä [31].

Neljäs vallankumous sai nimensä 2011, kun Saksassa pohdittiin autonvalmistajien kilpailukyvyn vahvistamista [29]. Neljäs vallankumous koostuu IoT:stä, pilvipalveluista, kognitiivisesta oppimisesta ja muista kyber-fyysisistä järjestelmistä (CPS, cyber physical system). Teollisuus 4.0:n tarkoituksena on lisätä koneiden välistä kommunikointia ja siten kasvattaa tuottavuutta. Ihmisten määrää tehtaissa voidaan vähentää ja lopulta työvoima voidaan poistaa jopa kokonaan, jos tehdas pystyy täysin autonomisesti valmistamaan tuotteitaan ja kehittämään prosessiaan. Teollisuus 4.0 tuottaa valtavan määrän uutta dataa. Kaikki laitteet on kytketty toisiinsa langattomasti tai langallisesti.

3.2.2 IoT:n tulevaisuus

IoT ja teollisuus 4.0 tulevat kasvamaan lähivuosina. Pian 5G-verkkoliikenne yleistyy sekä teollisuudessa että operaattoreilla. Tästä seuraa se, että laitteita saadaan yhdistettyä verkoilla nopeammin, tarkemmin ja laajemmin. 5G lupaa jopa 1000-kertaisia datansiirron nopeuksia 4G:hen verrattuna. Reaaliaikainen data on joissain prosesseissa erittäin kriittistä ja 5G:n viive on 1 millisekunnin luokkaa, mikä on yksi viideskymmenesosa 4G-verkon viiveestä. Yhdessä pilvipalveluiden kanssa IoT:n mahdollisuudet ovat valtavat. Kuvassa 11 on Gartnerin hypekäyrä [32] esineiden internetille vuodelta 2016. Sen mukaan IoT-alustat saavuttavat pian hypepiikkinsä ja teknologian arkipäiväistyminen kestää 2–5 vuotta. Gartnerin mukaan IoT on alle kymmenen vuoden päästä jo tuottavuusasteella.



Kuva 12. Gartnerin hypekäyrä IoT:n kehitykselle. Muokattu lähteestä [32].

Tulevaisuudessa IoT:llä on kasvun takia myös suuria ongelmia. Datan määrän kasvaessa huolimattoman käsittelyn riski kasvaa, joten tietoturva on iso haaste, kun älykkäitä laitteita on joka puolella. Muita ongelmia ovat esimerkiksi poliittisen ohjauksen puute ja siihen liittyvä epävarmuus, kehittyvien maiden verkkojen riittämätön taso ja standardien puuttuminen. [30]

3.3 Tiedonlouhinta

Tiedonlouhinta tarkoittaa menetelmiä, joilla poimitaan suuresta määrästä dataa tärkeä informaatio hyötykäyttöön. Digitalisaatio on lisännyt maailmassa olevan datan määrää valtavasti. Datan määrä kasvaa eksponentiaalisesti ja sanotaan, että 90 % kaikesta datasta on kertynyt parin viime vuoden aikana [33]. Datan varastoimisesta on tullut halpaa ja helppoa nykyisten pilvipalveluiden ansiosta. Ongelmaksi on muodostunut kuitenkin datan hyödyntäminen ja sen analysointi.

Datan on tärkeää olla laadukasta, sillä väärää tai virheellistä dataa ei saada analytiikan keinoin hyödylliseksi. Oikeiden analytiikkatyökalujen avulla saadaan luotettavia tuloksia. Datan oikeellisuudella saadaan useita hyötyjä: johdon toimet perustuvat varmemmin faktoihin, raportointi voi olla nopeampaa tarkistusajkojen lyhetyssä ja resursseja säästyy organisaatiolla. [34, Luku 4]

Tiedonlouhintaprosessi koostuu seuraavista osaprosesseista: [35]

1. Datan puhdistus (häiriöiden poisto)
2. Datan yhdistäminen (useiden datalähteiden käyttö)

3. Datan valinta (validin datan valinta omaan käyttöön)
4. Datan muuntaminen (data käsitellään oikeaan muotoon tiedonlouhintaa varten)
5. Tiedonlouhinta (älykkäiden menetelmien käyttö datamallien etsimiseen)
6. Mallien arviointi (kiinnostavien mallien valinta)
7. Tiedon esittäminen (datan esittely käyttäjille).

Mitä tahansa tarkoituksellista dataa voi louhia. Esimerkiksi säädataa voi käyttää erittäin moneen tarkoitukseen, kuten jalkapallopelien tulosten ennustamiseen ja kauppohen kampanjoiden asiakasmääriin, mutta säädatalla ei kannata ennustaa yliopistojen hakijamääriä. Tiedonlouhinnalla luotuja malleja voi olla tuhansia, ellei jopa miljoonia. Malli on kiinnostava, jos se on helposti ymmärrettävissä, validi, mahdollisesti hyödyllinen sekä uusi. Tiedonlouhinta-algoritmit eivät löydä kuitenkaan kaikkia kiinnostavia malleja, koska se olisi liian monimutkaista. Algoritmit eivät myöskään usein pysty tuottamaan pelkästään kiinnostavia malleja, vaikka se onkin kehittäjien tavoite. Mallit, joita tiedonlouhinnalla saadaan, ovat joko esittäviä tai ennustavia. Esittävä tieto kuvaa valitun datan joitain ominaisuuksia, ja ennustavaa tietoa voidaan käyttää hyväksi tulevaisuuden ennustamiseen. [35, Luku 1]

Tiedonlouhinta on monen eri tieteenalan yhdistämistä yhteiseen tarkoitukseen. Tärkeimpiä näistä tieteenaloista ovat tilastotieteet, koneoppiminen, tietokantajärjestelmät ja tiedonhakutekniikka. Tiedonlouhinnan suurimmat haasteet ovat louhinnan metodologia, käyttäjävuorovaikutus, tehokkuus, skaalautuvuus, datatyyppien erilaisuus ja tiedonlouhinnan eettiset ongelmat yhteiskunnassa. [35, Luku 1]

3.3.1 Esikäsittely

Tiedonlouhinnan esikäsittely sisältää aiemmin esitellyn tiedonlouhintaprosessin neljä ensimmäistä vaihetta: datan puhdistuksen, yhdistämisen, valinnan ja muunnon. [35, Luku 3]

Tiedonlouhintaa varten täytyy ensin määritellä ongelma, joka tiedonlouhinnalla pyritään ratkaisemaan. On selvitettävä, mitkä ovat eniten tarkastelun kohdetta kuvaavia muuttujia. Datan määrä on myös tärkeä huomioida, koska mitä enemmän dataa on saatavilla, sitä parempaa informaatiota datasta voidaan louhia. Toisaalta iso datamäärä on monimutkaisille algoritmeille hidastava tekijä. Aluksi on myös hyvä pohtia, millaisia hahmoja olisi datasta mielekästä etsiä. Asiantuntijoiden tulee osata kertoa, mitä datasta voi pystyä näkemään ja mistä taustateorioista olisi hyötyä. [35, Luku 3]

Datan esikäsittelyyn kuuluu mahdollisten datajoukkojen yhdistäminen ja normalisointi. Data voi olla useissa tiedostoissa. Tiedonlouhinnan kannalta on hyvä, jos kaikki data on samassa normalisoidussa tiedostossa. Dataa voi alustavasti analysoida ennen varsinaista louhintaa esimerkiksi laskemalla muuttujien keskiarvoja, keskihajontoja tai

korrelaatioita. Datan siivous kuuluu myös esikäsittelyyn. Puuttuvia arvoja voidaan merkitä nolliksi tai keskiarvoiksi ja virheet tulee siistiä datasta pois. Taitava datan käsittelijä osaa myös johtaa uusia hyviä muuttujia esimerkiksi yhdistelemällä muuttujia tai diskretoimalla. Koska suuremmalla datamäärällä saa paremmat tulokset, on parempi pitää aluksi kaikki datan muuttujat mukana tiedonlouhinnassa. Laskennallisesti on toki parempi karsia tarpeettomia muuttujia. [35, Luku 3]

3.3.2 Riippuvuusanalyysi

Riippuvuusanalyysit tarkoittavat määrällisiä analyysimenetelmiä, joilla pyritään selittämään syy-seuraussuhteita muuttujien välillä. Sen todentaminen edellyttää kokeellista tutkimusasetelmaa, jossa koe- ja kontrolliryhmiä verrataan toisiinsa. Regressioanalyysin avulla syy-seuraussuhdetta pystytään mittaamaan heikommin, mutta sen etuna on mahdollisuus tutkia samanaikaisesti monen muuttujan vaikutusta kohdemuuttujaan. [36]

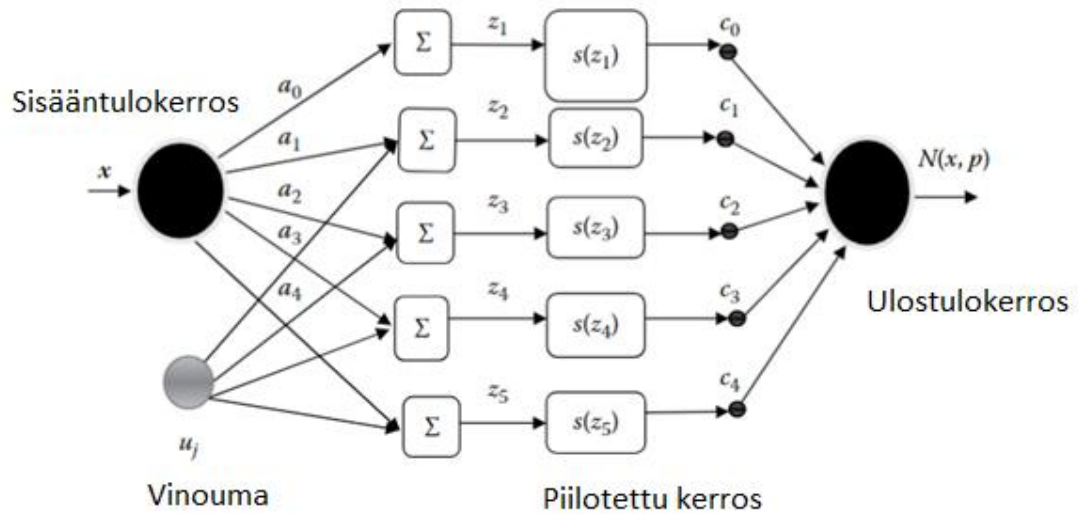
Erilaisia regressioita ovat esimerkiksi lineaarinen bayesilainen regressio, neuroverkkoregressio ja päätöspuuregressio. Lineaarisen bayesilaisen regression mallissa [38] vektori \mathbf{x} kerrotaan kerroinmatriisilla \mathbf{A} ja siihen lisätään bayesilainen, normaalijakautunut virhe e , jotta saadaan vektori \mathbf{y} :

$$\mathbf{y} = \mathbf{Ax} + e \quad (1)$$

$$e \sim N(0, \mathbf{V}) \quad (2)$$

$$p(\mathbf{y}|\mathbf{x}, \mathbf{A}, \mathbf{V}) \sim N(\mathbf{Ax}, \mathbf{V}). \quad (3)$$

Neuroverkkoregression tarkoituksena on luoda neuroverkoilla regressiomalli, joka parhaiten kuvaa dataa [39, Luku 4]. Kuvassa 13 on esitelty neuroverkkoregression arkkitehtuuri. Käyttäjälle näkyvät vain sisääntulo- ja ulostulokerrokset ja neuroverkko luo piilotettuja kerroksia, joiden avulla neuroverkko yrittää etsiä parhaan mallin esimerkiksi pienimmän neliösumman menetelmällä. Vinoumien ja painokertoimien muuntelun avulla neuroverkko pyrkii löytämään parhaan mahdollisen ratkaisun mallille.



Kuva 13. Neuroverkkoregression arkkitehtuuri yksinkertaisilla sisään- ja ulostulosolmuilla. Muokattu lähteestä [39, Luku 4].

Päätöspuuregressio käyttää päätöspuuluokittelumenetelmää löytääkseen parhaan ratkaisun datamallin kuvantamiseen. Päätöspuun haarat jakautuvat niin, että se minimoi yläoksien variaation. Alimpana juurioksana on siis ominaisuus, joka parhaiten vähentää variaatioiden määrää. Päätöspuuregressio pystyy käsittelemään valtavan määrän mahdollisia lopputuloksia tehokkaasti ja sillä saadaan aikaiseksi pienempiä päätöspuita kuin perinteisillä menetelmillä.

3.3.3 Luokittelu

Luokittelu on data-analyysin muoto, joka erottelee datamalleja tärkeisiin kuvaaviin luokkiin. Luokittelun tarkoituksena on jakaa esimerkiksi paperikatkon syyt eri luokkiin virhetilanteen mukaan. Onnistuneella luokittelualgoritmeilla voidaan mahdollisesti ennakoita vikaantumisen tai vähintäänkin tunnistaa nopeasti vikaantumisen syy. Datan luokittelu on kaksivaiheinen prosessi. Ensimmäinen vaihe on oppiminen ja toinen itse luokittelu. Luokittelu vaatii aina opetusdatan, jonka avulla algoritmi oppii datassa esiintyvät mallit mahdollisimman hyvin. Opetetun algoritmin avulla malli osaa luokitella lopun datasta mahdollisimman tarkasti. Luokittelua voidaan tehdä esimerkiksi päätöspuilla tai neuroverkoilla. [35, Luku 8.1]

Klusterointi on analysointimenetelmä, jolla pyritään muodostamaan haluttu määrä datapisteryhmiä. Klusterointi on ohjaamatonta luokittelua, jonka avulla datajoukot saadaan jaettua toisistaan erilleen. Suuresta datasta voidaan saada irti erilaiset ryhmittymät ja mahdollisesti havaita joitain kaavamaisuuksia. [35, Luku 10]

3.3.4 Tulosten tarkastus

Tiedonlouhinnan tuloksista voi olla paljonkin hyötyä, ja ne saattavat näyttää erinomaisilta. Tuloksista saa kuitenkin helposti sen näköisiä kuin niistä haluaa, jos valikoi alkuperäiseen dataan vain halutut parametrit tai siistii kaikki virheet pois.

Tuloksien tarkastelussa on tärkeää tietää prosessista paljon. Asiantuntijoiden tietotaidon avulla pystytään etsimään tieto, jota tiedonlouhinnalla on alun perinkin haluttu etsiä. Pelkkä datan visualisointi ei riitä, jos sen informaatioarvo on nolla. Tiedon taso saadaan nostettua datan ja informaation kautta oikeaksi hyödylliseksi tiedoksi, jota voi käyttää prosessien kehittämiseen. [34, Luku 4]

3.4 Ennakoiva kunnossapito

Laitteiden kunnossapito on yrityksille iso menoerä. Jos olisi mahdollista minimoida kalliiden laitteiden huoltokustannukset, yritysten kannattavuus paranisi. Isot toimijat tarjoavat jo ennakoivaa kunnossapitoa tuoteportfoliossaan. Kunnossapidon palveluliiketoiminta on noussut laitteiden valmistuksen rinnalle merkittäväksi liiketoiminnaksi. Kunnossapitopalvelua kutsutaan termillä Maintenance-as-a-Service (MaaS). Laitteiden kunnossapitoa voi tehdä monella eri tavalla. Laite voidaan huoltaa vasta kun se hajoaa, jolloin tulee arvaamattomia tuotannon pysäytyksiä. Laitetta voidaan huoltaa tietyn ajan välein, esimerkiksi vuosittain. Vuosittainen huolto saattaa olla turhan usein, mutta laitevalmistaja säästyy asiakkaiden vihaiselta palautteelta, kun laitteet eivät hajoa huollon puutteen takia. Tiedonlouhinnan ja tekoälyn avulla ennakoivaa kunnossapitoa voidaan soveltaa ja kustannukset saadaan optimoitua. Ennakoiva kunnossapito hyödyntää mitatun datan ja laitteet voidaan korjata, kun sille on oikeasti tarvetta. Ennakoiva kunnossapito parantaa laitteen laatua, vähentää laitteen hukka-aikaa sekä välttää hallitsemattomat hajoamiset [40].

Laitteista ja tehtaan toiminnasta kerätään usein jo paljon erilaista dataa. Taloudellisia ja teknisiä tietokantoja kerätään, mutta niitä harvemmin hyödynnetään yhtenä suurena datamassana. Jos tietokantojen tiedot yhdistettäisiin, saataisiin paljon ennakoivaa kunnossapitoa hyödyttävää dataa. Pilvilaskenta tarjoaa ennakoivalle kunnossapidolle uusia mahdollisuuksia. Sen avulla voidaan käyttää laajempaa määrää dataa tehokkaammin ja luovemmin. Internet of Things ja ennakoiva kunnossapito kulkevat käsi kädessä. Jotta kaikki laitteet voisivat olla huollettuina ja koko prosessi toimisi moitteettomasti, on kaikkien laitteiden datan oltava tiedossa. [40]

Ennakoivan kunnossapidon haasteina Schmidtin ym. [40] mukaan ovat oikean datan hyödyntäminen, systemaattinen lähestymistapa, tietojohtaminen ja epävarmuuden hallinta. Oikean datan hyödyntämisessä pitää ottaa huomioon eri laitteiden ominaisuudet,

jos vertailuja tehdään. Myös ympäristö ja laitehuollot saattavat muokata joistakin laitteista saatavaa dataa ja siten vääristää tuloksia.

4. ALUSTAT

4.1 Pilvipalvelut

Datan valtavan määrän takia palvelimet ja suuret konehuoneet ovat monelle yritykselle välttämättömiä, mutta nykyään niiden korvaaminen pilvipalveluilla on yhä useammalle yritykselle tehokkaampi vaihtoehto. Pilvipalvelut ovat ulkoistettuja tietotekniikkapalveluja. Pilvipalvelun ominaispiirteitä ovat National Institute of Standards and Technologyn mukaan:

- itsepalvelullisuus,
- pääsy palveluihin eri päätelaitteilla,
- resurssien yhteiskäyttö,
- nopea joustavuus ja
- käytön tarkka mittaaminen. [41]

Monet yritykset tarjoavat pilvipalveluita, jotka voidaan luokitella kolmeen eri luokkaan: IaaS (Infrastructure-as-a-Service), PaaS (Platform-as-a-Service) ja SaaS (Software-as-a-Service). IaaS tarjoaa palvelimia ja palvelinsaleja käyttäjilleen nopeasti. IaaS-palveluun sisältyy usein itse tallennustila, mutta myös niiden ylläpito. PaaS tarjoaa käyttäjilleen pilvipalvelualustan, jonka ominaisuuksiin kuuluvat operaatiojärjestelmä sekä ajoaika. SaaS tarjoaa käyttäjilleen sovellusten käyttölisenssejä sovellusten ostamisen sijaan.

Pilvipalvelut ovat käyttäjille yleensä edullisempi ja joustavampi ratkaisu kuin kiinteät palvelinratkaisut. Tämä perustuu siihen, että maksetaan vain siitä, mitä käytetään. Pilvipalvelutarjoajat pitävän myös huolen siitä, että tarjolla on parasta tekniikkaa käytössä ja ettei käyttäjien tarvitse maksaa huoltokatkoja itse. Lisäksi pilvipalvelut ovat nopeampia kuin perinteiset IT-palvelut.

Pilvipalveluihin liittyy myös tietysti riskejä. Palvelimet ovat usein ulkomailla, joten käyttäjä voi kokea oman datansa haavoittuvaksi. Harvoin käyttäjät tietävät tarkalleen, kuinka heidän datastaan loppujen lopuksi pidetään huolta. Verkossa liikkuva data on myös alttiimpi tietomurroille kuin omissa palvelimissa oleva data. Nykyään ja varsinkin tulevaisuudessa pilvipalvelimien turvallisuuden takaamiseksi voidaan käyttää kolmannen osapuolen auditointipalveluita [42]. Kolmas osapuoli testaa pääsyä pilvipalvelimille ja raportoi, jos onnistuu purkamaan salauksen. Turvallisuuden lisäksi käyttäjiä huolettavat pilvipalveluiden kasvava hinta, heikentyvä suorituskyky, standardien puuttuminen, riittämätön räätälöitävyys sekä yhteensopivuusongelmat nykyisten järjestelmien kanssa. [34, ss. 28–31], [43]

Suurimmalle osalle yrityksistä ohjelmistot ja ICT-palvelut ovat vain työkaluja ydinliiketoiminnan tukemiseksi. Tästä syystä pilvipalvelut ovat toimiva ratkaisu monelle yritykselle, jotta yrityksen resurssit voidaan käyttää ydinliiketoiminnan palvelemiseen. Tyypillisiä käyttötarkoituksia pilvipalveluille ovat esimerkiksi kehitystyö, testaus, toiminnallinen ulkoistaminen ja resurssireservi. Yrityksen pohtiessa pilvipalveluun siirtymistä sen on otettava huomioon nykyisen toimintamallin hyödyt ja kustannukset, pilvipalveluun siirtymisen hyödyt ja kustannukset, pilvipalvelun riskit ja yrityksen valmius siirtyä pilvipalveluiden käyttäjäksi.

Pilvilaskenta on tehokas analytiikkatyökalu. Moni yritys kerää tällä hetkellä valtavia määriä dataa. Datan tehokkaaseen analysointiin vaadittava laskentateho on riittävä usein vain pilvilaskennan avulla. Suuret teknologiayritykset, kuten Google, Amazon, IBM ja Microsoft, kilpailevat markkinaosuuksista. Tässä työssä vertaillaan IBM:n Watson Analyticsiä sekä Microsoftin Azure Machine Learning Studiota.

4.2 IBM Watson

IBM Watson on monipuolinen sovellusjärjestelmä. Se on oppiva järjestelmä, joka tekoälyn avulla pystyy tekemään päätöksiä ja ennustamaan tulevia ongelmia. Watson on ennen kaikkea kognitiivinen järjestelmä, mutta se hyödyntää luonnollisen kielen, analytiikan ja koneoppimisen tekniikkoja [44, Luku 9]. Tällä hetkellä IBM Watson osaa esimerkiksi puhua kieliä natiivisti. Suomen kieltä Watson on oppinut Turun yliopiston kehittelemällä tekniikalla [45]. Vuonna 2011 Watson voitti Jeopardy!-ohjelman kaksi aiempaa mestaria [44, Luku 9]. Watson IoT for Manufacturing -palvelu pystyy parantamaan laitteiden luotettavuutta ja tuottavuutta paremman näkyvyyden, luotettavuuden ja operoinnin ansioista. Watson pystyy pienentämään kustannuksia, lyhentämään läpimenoaikoja ja ennustamaan mahdolliset huoltotarpeet.

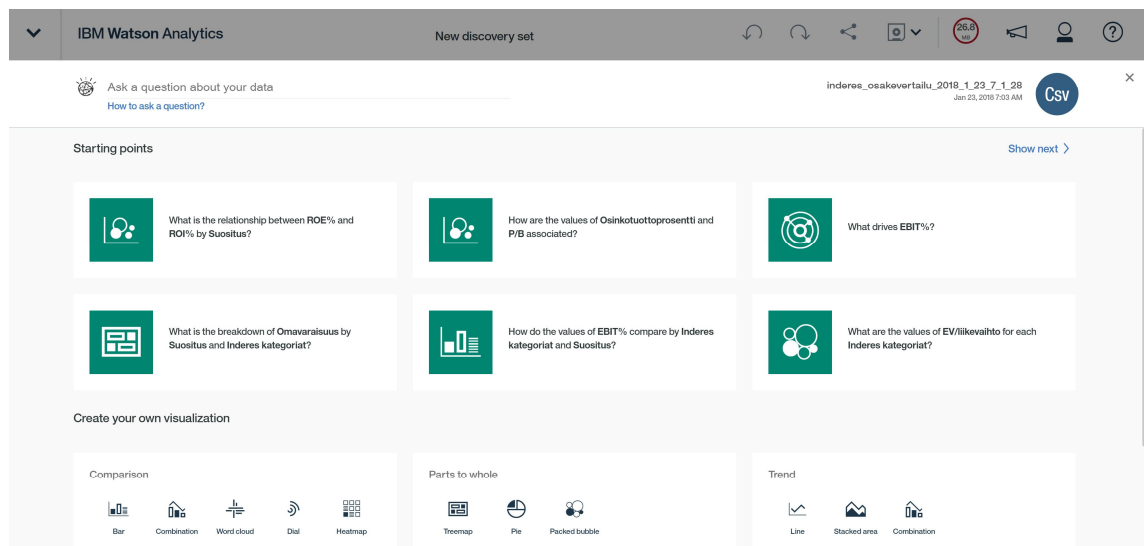
Watsonin kognitiivinen tiedonkäsittely pyrkii ajattelemaan kuten ihminen. Data-analyysissä perinteinen numeronmurskaus on Watsonia tehokkaampi vaihtoehto. Watsonin kognitiiviset kyvyt tulevat paremmin käyttöön, jos data sisältää kuvaa tai ääntä. Tapahtumalokeja tutkittaessa Watson on erinomainen, koska se pystyy löytämään avainasiat hakusanojen avulla tai monimutkaisten löytöjen perusteella. Tapahtumalokit ovat haasteellisia: vapaata tekstiä on hankala lukea, tekstissä voi olla paljon kirjoitusvirheitä eikä teksti ole strukturoitua. Tapahtumalokeista saatava tieto on kuitenkin tärkeää konerikkojen arviointia tulkittaessa ja siksi sitä kannattaa tehdä. Watson pystyy löytämään tietokannasta eri dokumenttien väliset korrelaatiot nopeasti. Siihen ihmisellä meni useita vuosia, kun taas koneelta tapahtumalokien läpikäymiseen kuluu muutama sekunti.

Chenin ym. mukaan [46] IBM Watsonin kognitiiviset kyvyt ovat jo lääketieteessä edenneet hyödylliselle tasolle. Watsonin avulla saadaan potilaille oireiden ja tutkimustulosten mukaan määrättyä todennäköisimmin tehokkain lääke. Lääkäreiden työ myös helpottuu, kun kaikki mahdolliset lääkevaihtoehdot ovat tietokannassa.

4.2.1 Käyttökohteet

Watsonilla on työkaluja ennakoivaan kunnossapitoon. Watson yhdistää aikaisempaa ja nykyistä dataa luodakseen ennusteen mahdollisista tulevaisuuden tapahtumista kuten laiterikoista. Jokaiseen tapahtumaan voidaan liittää todennäköisyyslaskelma. Yritykset voivat myös optimoida huoltovälejä sekä useiden tehtaiden huoltoreittejä Watsonin analytiikkatyökalujen avulla.

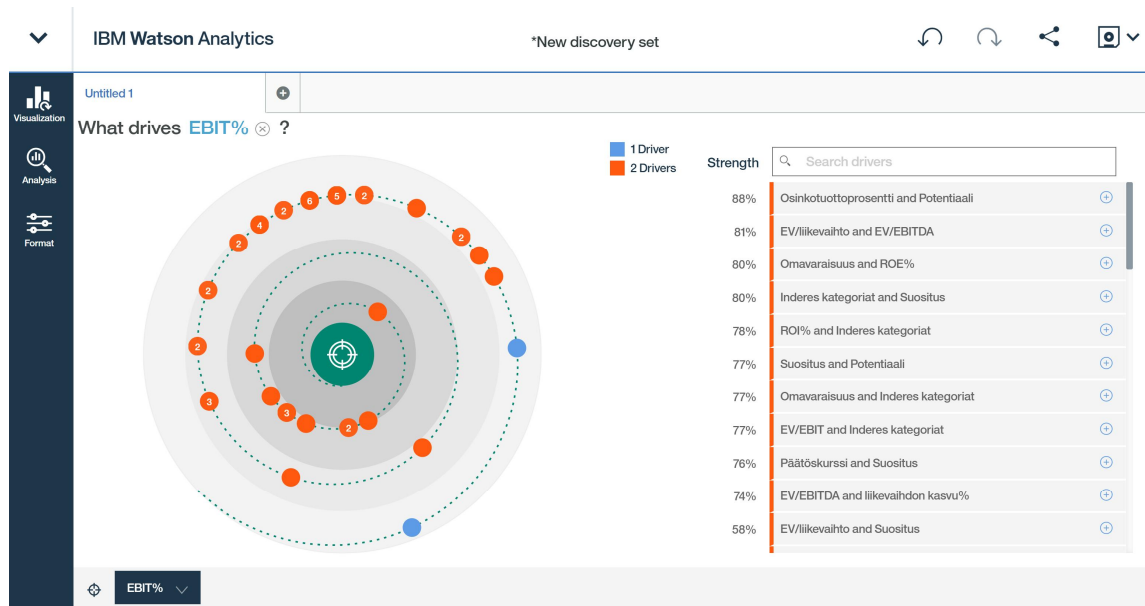
Watson Analytics on tehokas työkalu, kun halutaan etsiä tekstitiedostojen riippuvuuksia ja tiettyjen parametrien välisiä riippuvuussuhteita. Esimerkkinä on esitelty suomalaisten osakkeiden vertailu. Data on saatu Inderes-analyysitalon osakevertailusta [47]. IBM Watson Analytics ehdottaa datasta mahdollisia kysymyksiä, joiden se kuvittelee asiakasta kiinnostavan. Kuvassa 14 on Watson Analyticsin näkymä, jossa se ehdottaa kysymyksiä, joita annetusta datasta voisi mahdollisesti esittää ja joista voisi saada kiinnostavaa informaatiota.



Kuva 14. Watson Analyticsin ehdottamat kysymykset datasta.

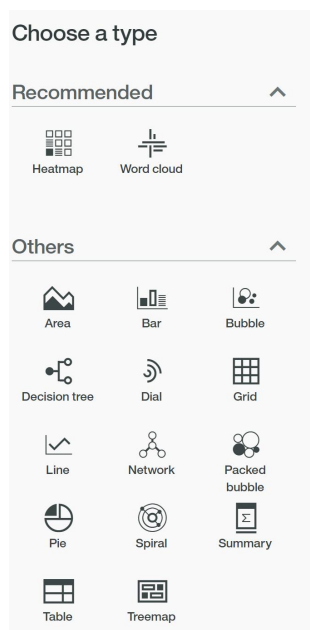
Kysymysehdotuksista voi ottaa mallia, mutta käyttäjän on tiedettävä itse, mitä haluaa tietää datasta. Esimerkiksi jos tutkitaan, mitkä ovat EBIT%:n ajurit, saadaan kuvassa 15 oleva näkymä. Kuvasta huomataan, että datan perusteella tärkeimpänä ajurina EBIT%:lle

ovat osinkotuottoprosentin ja Inderes-palvelun laskema tuottopotentiaalin yhteisvaikutus. Watson Analyticsin visuaaliset ominaisuudet ovat erittäin kuvaavia.



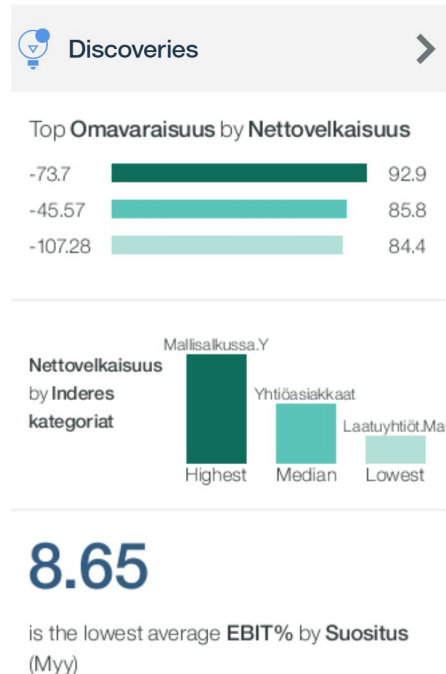
Kuva 15. Watson Analyticsin visualisointi parametrin ajureista.

Watson Analyticsin visualisointimahdollisuudet on esitelty kuvassa 16. Niitä on lämpökartta, sanapilvi, aluekaavio, pylväskaavio, kuplakaavio, päätöspuu, mittarikaavio, verkko, linjakaavio, verkosto, pakattu kuplakaavio, piirakkakaavio, spiraalikaavio, yhteenveto, taulukko ja puukartta.



Kuva 16. IBM Watson Analyticsin visualisointimahdollisuudet.

IBM Watson Analytics antaa myös ehdotuksia, kun ensimmäinen analyysi on tehty. Kuvassa 17 on Watsonin ehdottamat kiinnostavat löydöt, joita käyttäjä on saattanut etsiäkin.



Kuva 17. Watson Analyticsin ehdottamat analyysilöydöt.

IBM Watsonia käytetään jo nykyään osakkeiden poiminnassa. AIEQ-niminen pörssinoteerattu rahasto käyttää hyväkseen Watsonin kognitiivisia kykyjä. Watson käy läpi osakeanalyysijä, talouslukuja ja esimerkiksi sosiaalisen median julkaisuja, joiden perusteella AIEQ valitsee matalasti arvostetut osakkeet ja ostaa ne rahastonsa. [48], [49]Kuvan 18 perusteella AIEQ ei kuitenkaan ole pystynyt tuottamaan parempaa tuottoa kuin NASDAQ-indeksi.



Kuva 18. AIEQ pörssinoteeratun rahaston ja NASDAQ-indeksin (COMP) vertailu vuoden aikajaksolta [50].

IBM Watson Analytics ei varsinaisesti tarjoa yksinään mitään kognitiivisia analyysityökaluja. Yhdistettynä muihin maksullisiin Watsonin työkaluihin se on kuitenkin hyvä lisä, sillä sen avulla voi helposti visualisoida datasta haluamansa tiedon mukavasti pureskeltavaan muotoon.

4.2.2 Käyttökohteet paperiteollisuudessa

Watsonin sivuilla [51], [52] kerrotaan Watson IoT for Manufacturing -palvelusta. Prosessiautomaatioon se soveltuu mainiosti. Palvelu sisältää toimintoja, joiden avulla laitoksen data saadaan hyödynnettyä, kuten suorituskyvyn arvioinnin ja ennakoivan huollon.

Suorituskyvyn arviointi mahdollistaa laitoksen laitteiden tunnistamisen ja kykenee ennustamaan laitteet, jotka todennäköisesti hajoavat pian. Paperiteollisuudessa prosessit ovat jatkuvia, joten laitteen hajoamisesta johtuvat rikot ovat kalliita. Datasta palvelu pyrkii etsimään epäsäännöllisyyksiä, jotka ovat aiemminkin johtaneet rikoihin. Palvelu käyttää ISA95-standardia laite- ja tehdastason väliseen kommunikointiin, jotta koko laitoksen data voidaan käyttää kokonaisvaltaisiin ennusteisiin. ISA95 on amerikkalaisen ISA:n (International Society of Automation) kehittämä kansainvälinen standardi, joka määrittelee valmistusohjausjärjestelmän rakenteet ja tehtävät.

4.3 Microsoft Azure

Microsoftin Azure on pilvipalvelin, kuten IBM:n Watsonkin. Azure on monipuolinen järjestelmäkokonaisuus, jota voi hyödyntää joko vain IaaS-verkkopalvelimena tai PaaS- tai SaaS-palveluna. Azure on tehokas apuväline pilvilaskentaan, joka tarjoaa paljon erilaisia työkaluja datan murskaamiseen. Watsonin Analyticsiin verrattuna Azure on ketterämpi ja sitä käyttäessä voi laajasti käyttää eri ohjelmointikieliä. Microsoft markkinoi Azurea erittäin turvallisena palveluna ja sitä käyttää moni merkittävä virasto, kuten esimerkiksi FBI [53]. Azure on sitoutunut siihen, ettei kehitä omaa palveluaan asiakkaidensa datalla. Microsoft Azuren *Machine Learning Studio* on pilvilaskenta-alusta, jonka avulla on helppo suunnitella graafisesti tehokkaita ja ennustavia analyysejä. Alustalla on lukuisia määriä algoritmeja ja omia algoritmeja on helppo luoda itse R:llä tai Pythonilla.

4.3.1 Käyttökohteet

Azuren käyttäjäkuntaan kuuluvat kaikki yritykset, jotka haluavat digitalisoida toimintaansa. Tekoälytuotteita, joita Azure markkinoi omilla sivuillaan, on useita, kuten bottipalvelu, luetun- ja puheenymmärrys sekä perinteinen koneoppiminen.

Azuren tuotteistukset ovat moduuleita, joista tarvittava sovellusarkkitehtuuri rakennetaan. Joku asiakas saattaa tarvita pelkät tietokannat, mutta toinen tarvitsee kaiken virtuaalikoneista analyysipalveluihin ja visualisointiin. Tämä tuo joustoa Azuren ja asiakkaan välille, ja kaikki saavat sopivan tuotteen. Azuren käytöstä maksetaan käytön eikä lisenssien mukaan, mikä on varsinkin pienille yrityksille sopivaa, sillä silloin ei tarvitse tehdä suuria investointeja sovelluksiin tai palvelimiin.

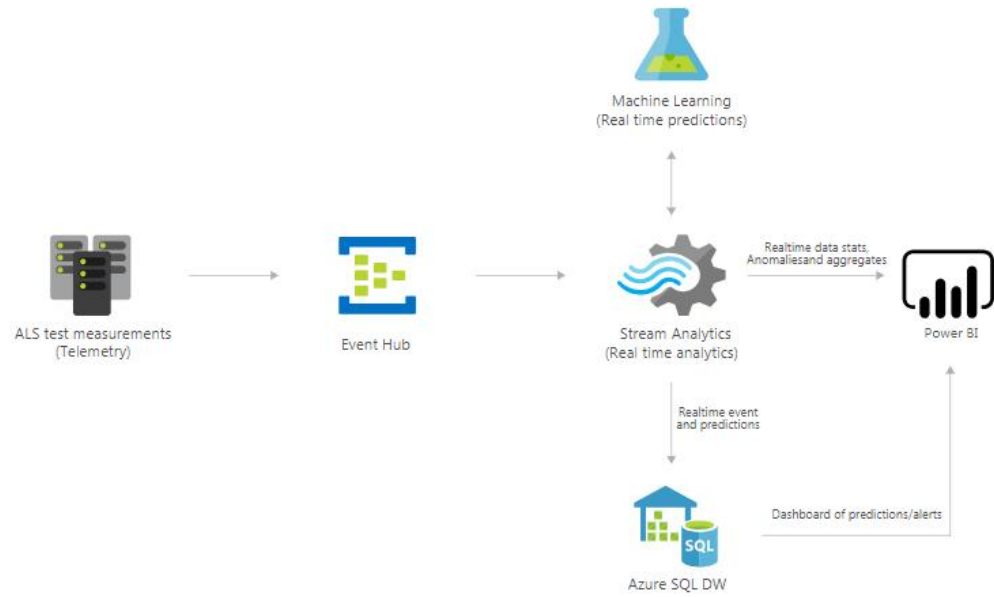
4.3.2 Käyttökohteet paperiteollisuudessa

Microsoft Azuren käyttökohteet paperiteollisuudessa ovat samanlaisia kuin Watsonillakin. Ennustavaan kunnossapitoon löytyy sovellusratkaisuja valmiina. Azuren koneoppimisalgoritmit voivat tunnistaa laitteiden vikaantumisen ennakoivasti. Azuren turvalliset pilvipalvelimet siirtävät nopeasti dataa, jolloin tehtaan laitteistoa voidaan tutkia etäältä ja auttaa mahdollisissa korjaustoimenpiteissä.

Kuvassa 19 on kuvattu Azuren tarjoama sovellusarkkitehtuuri ennakoivalle huollolle. Tässä arkkitehtuurissa on viisi tuotetta:

- Stream Analytics, välittää reaaliajassa sensoridataa muille tuotteille
- Event Hub, kerää sensoridatan ja lähettää sen Stream Analyticsille
- Machine Learning Studio, ennustaa mahdolliset vikaantumiset

- SQL Data Warehouse, tallentaa prosessidatan ja vikatieidot sekä
- Power BI, visualisoi datan operaattorille.



Kuva 19. Microsoft Azuren ennustavan kunnossapidon esimerkkiarkkitehtuuri [54].

5. DATA JA SEN ANALYYSI

5.1 Raakadatan esittely

Analysoitava data on dataloggerilta saatu 10 viikon mittainen paperikoneen leikkurin mittaussaineisto. Data on joukko aikasarjoja, joissa sensoreiden arvot on tallennettu tietokantaan tiheimmillään 50 millisekunnin välein. Analysointia varten valittiin seuraavat muuttujat, jotka kuvaavat parhaiten paperikoneen leikkurin sähkökäytön toimintaa:

- rullaimen moottorin todellinen momentti
- rullaimen moottorin todellinen nopeus
- todellinen kireys
- kireyden referenssiarvo
- kireyssäädön korjaus
- aukirullaimen todellinen halkaisija ja
- kiihtyvyys/hidastuvuus.

Datassa ei ole havaittavissa huomattavia virheitä, joten dataa ei tarvitse puhdistaa häiriöistä. Datan mittaustaajuus ei kuitenkaan ole riittävä mekaanisten värähtelyiden tarkasteluun. Mekaanisten värähtelyiden havaitsemiseksi sensoreiden arvot tulisi tallentaa vähintään 10 millisekunnin välein.

5.2 Datan muokkaus

Aikasarjadata on muutettu taulukkomuotoon, jossa jokaisella rivillä on yksi muutto ja jokaisessa sarakkeessa on muutolle lasketut tunnusluvut. Tunnuslukuja on muuton aloitus- ja lopetusaika, muuton kesto ja jokaisesta valitun muuttujan keskiarvo, varianssi, minimi sekä maksimi. Dataloggerin datan tiedostomuoto oli data-analyysiä varten muunnettu Parquet-tietokannaksi viikon mittaisiin jaksoihin. Parquet on Twitterin ja Clouderan kehittämä avoimen lähdekoodin Apache-ohjelmistojärjestön tietokanta [55]. Parquet on sarakepohjainen tietokanta ja pakkaa tiedostot sarakkeisiin perustuen, mikä tekee pakkaamisesta tehokkaampaa.

Analysoitava taulukkorakenne on tehty Python-koodilla. Työkaluna tähän on käytetty Spyder-ohjelmistoa, jonka rakenne ja käyttö vastaa hieman Matlabia. Python on nousemassa data-analytiikan käytetyimmäksi kieleksi sen suuren yhteisön ja avoimen lähdekoodin kirjastojen takia [56]. Avoimen lähdekoodin kirjastoista työssä oli käytössä NumPy, PyArrow sekä pandas. Python-ohjelma lukee Parquet-tiedoston ja etsii siitä

rullaimen moottorin todellisen nopeuden, jonka mukaan muuton alku- ja loppukohtat määritetään. Yhden muuton alkukohta määritetään siitä kohdasta kun nopeuden arvo kasvaa positiiviseksi. Muuton lopussa nopeus palaa takaisin nolnaan. Muutto on rajattu niin, että muuton keston tulee olla 5 ja 30 minuutin välillä ja muuton nopeuden maksimiarvon on ylitettävä 300 metriä minuutissa. Ohjelma määrittelee jokaiselle muutolle valittujen muuttujien tunnusluvut ja lisää ne taulukkoon. Ohjelma käsittelee yhden viikon jakson kerrallaan. Kun ohjelma on ajettu, se tulostetaan csv(comma-separated values)-tiedostoksi, jota analysointityökalut pystyvät tulkitsemaan.

5.3 Tiedonlouhinta IBM Watson Analyticsilla ja Azure Machine Learning Studiolla

IBM Watson Analytics ei tarjoa juuri muuta kuin datan visualisointia. Watson Analytics -alusta ei sisällä kognitiivista laskentaa, mutta oli järkevin vaihtoehto Watsonin portfolioista, jota pystyi käyttämään ilmaiseksi.

IBM Watson Analytics käyttää arvojen ennustamiseen CHAID (chi-squared automatic interaction detection)-menetelmää, joka yhdistelee ohjelmalle syötettyjä arvoja. Kuvassa 20 on kireyden referenssin keskiarvon ennustemallin sääntöluettelo. Listan sarakkeet ovat ennustettu arvo, säännöt sekä tuloksien määrä. Esimerkiksi ensimmäinen rivi kertoo, mitä muut muuttujat todennäköisesti ovat, jos kireyden referenssin keskiarvo on 926,9. Tämän ennusteen todennukseksi datassa on siitä 54 tulosta. Koko taulukko on listattu liitteeseen A. Saadut tulokset ovat hyvin triviaaleja. Maksimi ja varianssi kulkevat keskiarvon kanssa käsi kädessä ja kireyden referenssi määräytyy aukirullauksen halkaisijan mukaan.

What is a predictive model for **Mean_TensionRef** ? (Predictive strength: 97%)

Decision Rules Tree















Decision rules show that Max_TensionRef and 21 other inputs predict Mean_TensionRef.

▲▼ Predicted value	Rules	Records
962.90	Max_TensionRef > 1,400 S_Var_TensionRef = 153613.58313 to < 201249.67877 S_Mean_UnwindDiam = 1321.75181 to < 1694.72564; 1694.72564 to < 2067.69946 S_Var_Kiihtyvyyshidastuvuus = 0.20572 and above	54
956.18	Max_TensionRef > 1,400 S_Var_TensionRef = 153613.58313 to < 201249.67877 S_Mean_UnwindDiam = 1321.75181 to < 1694.72564; 1694.72564 to < 2067.69946 S_Var_Kiihtyvyyshidastuvuus = 0.07142 to < 0.11618; 0.11618 to < 0.16095; 0.16095 to < 0.20572	96
955.07	Max_TensionRef > 1,400 S_Var_TensionRef = 153613.58313 to < 201249.67877 S_Mean_UnwindDiam = 2067.69946 to < 2440.67328 S_Min_UnwindDiam = 1643.66255 to < 2025.19683	60
948.47	Max_TensionRef > 1,400 S_Var_TensionRef = 153613.58313 to < 201249.67877 S_Mean_UnwindDiam = 2067.69946 to < 2440.67328 S_Min_UnwindDiam = 2025.19683 to < 2406.7311	55

Kuva 20. IBM Watson Analyticsin ennustemallin sääntöluettelo.

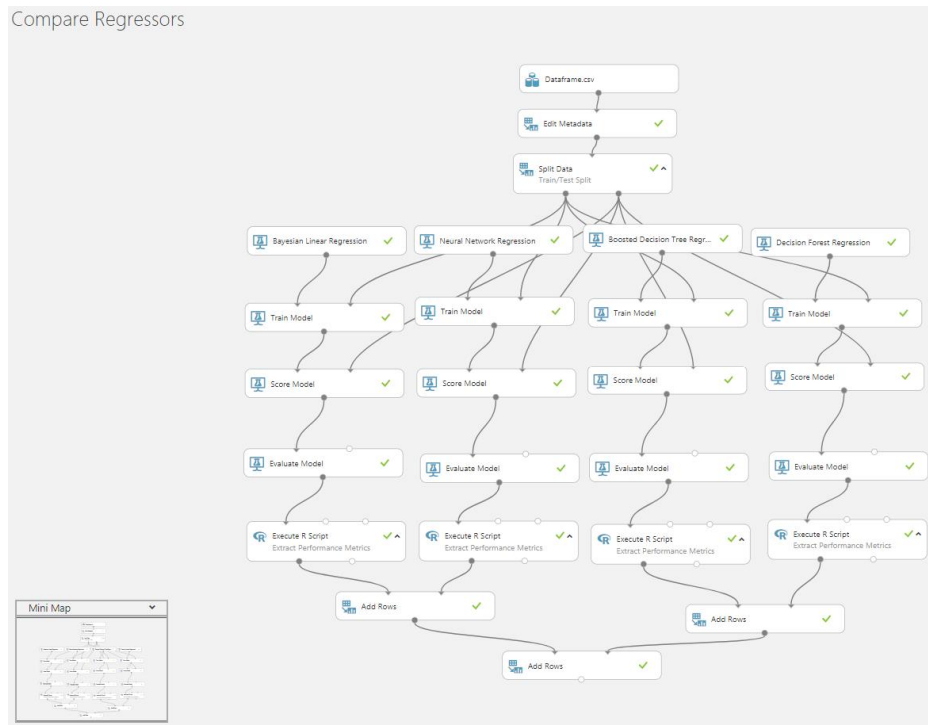
Microsoft Azuren Machine Learning Studio (MLS) on ilmainen työkalu datan käsittelyyn ja analysointiin. IBM Watson Analyticsiin verrattuna se sisältää huomattavasti enemmän koneoppimista ja tekoälyä. Lisäksi MLS mahdollistaa datan käsittelyn sekä useilla valmiilla työkaluilla että omien koodien kirjoittamisen esimerkiksi R:llä tai Pythonilla. Kuvassa 21 on luettelo MLS:n valmiista vaihtoehdoista. Datan esikäsittelyyn on paljon valmiita palikoita. Machine Learning -osion alla on ennustemallien rakennuspalikoita

datan opettamista ja pisteytystä varten. Valmiita malleja on luokittelulle, regressiolle, poikkeamien tunnistukselle sekä klusteroinnille.

- ▶  Saved Datasets
- ▶  Data Format Conversions
- ▶  Data Input and Output
- ▶  Data Transformation
- ▶  Feature Selection
- ▶  Machine Learning
- ▶  OpenCV Library Modules
- ▶  Python Language Modules
- ▶  R Language Modules
- ▶  Statistical Functions
- ▶  Text Analytics
- ▶  Time Series
- ▶  Web Service
- ▶  Deprecated

Kuva 21. Luettelo Azure Machine Learning Studio palikkaryhmistä.









Kuvassa 22 on MLS:n alusta, jolle palikoista rakennetaan ohjelma omaa dataa varten. Esimerkkinä on vertailtu neljää eri regressiota: bayesilaista lineaariregressiota, neuroverkkoregressiota, tehostettua päätöspuuregressiota ja päätöspuuregressiota.



Kuva 22. Azure Machine Learning Studion alusta, johon on rakennettu regressioiden vertailun ohjelma.

Data on jaettu opetus- ja testausdataksi, minkä jälkeen data on opetettu ja pisteytetty erikseen jokaista neljää regressiomentelmää käyttäen. Tämän jälkeen R-koodi laskee, kuinka hyvin eri regressiot suoriutuvat. Lopuksi vielä yhdistetään suorituskyyvyt yhdeksi taulukoksi, joka on esitetty kuvassa 23. Taulukossa on vertailtu eri regressioiden hyvyiksiä, ja tämän analyysin perusteella tehostettu päätöspuuregressio olisi paras vaihtoehto neljästä valitusta. Sen virheet ovat pienimpiä ja determinaatikerroin (R^2) on suurin. Päätöspuuregressio sai myös positiivisen determinaatikertoimen, mutta bayesilainen ja neuroverkkoregressio saivat negatiivisen. Determinaatikerroin kuvaa regression suhdetta lineaariregressioon. Suhteen ollessa 1 regressio on täysin lineaarinen. Negatiivisen arvon saadessaan regression suhde lineaariregressioon on huonompi kuin vaakasuoran viivan.

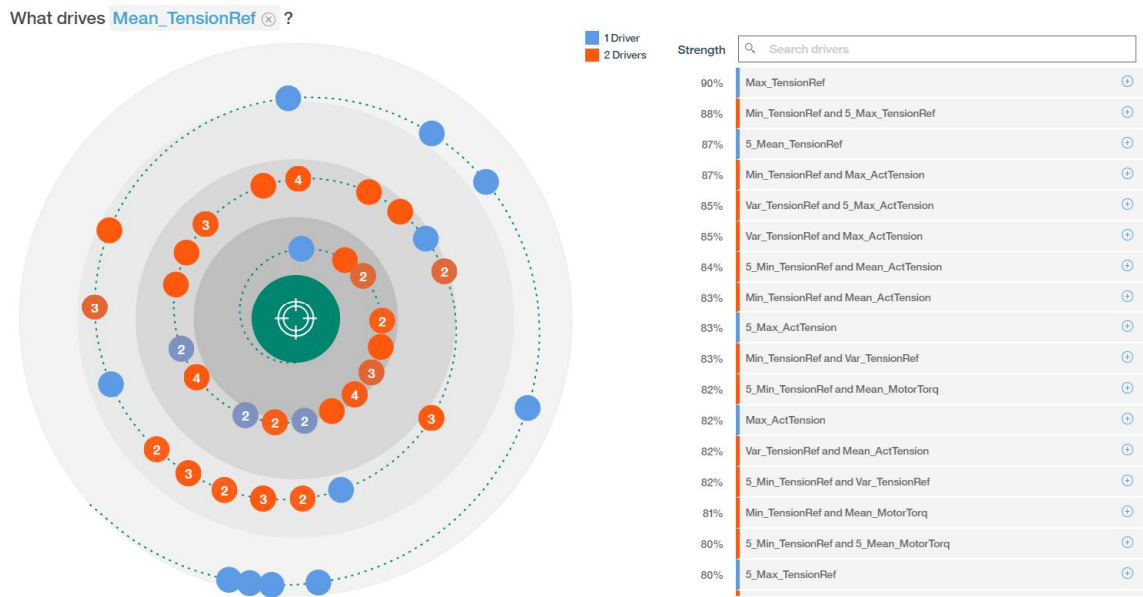
Compare Regressors - Copy > Add Rows > Results dataset

rows	columns						
4	6	Algorithm	Mean Absolute Error	Root Mean Squared Error	Relative Absolute Error	Relative Squared Error	Coefficient of Determination
							
		Bayesian Linear Regression	0.437028	2.561442	0.452992	4.040613	-3.040613
		Neural Network Regression	0.382127	1.412916	0.396086	1.229451	-0.229451
		Boosted Decision Tree Regression	0.267127	0.505667	0.276885	0.157474	0.842526
		Decision Forest Regression	0.359774	0.609429	0.372916	0.22873	0.77127

Kuva 23. Azure Machine Learning Studion taulukko, jossa on kuvattu eri regressioiden ominaisuuksia testidatan pohjalta.

5.4 Tiedon esittäminen

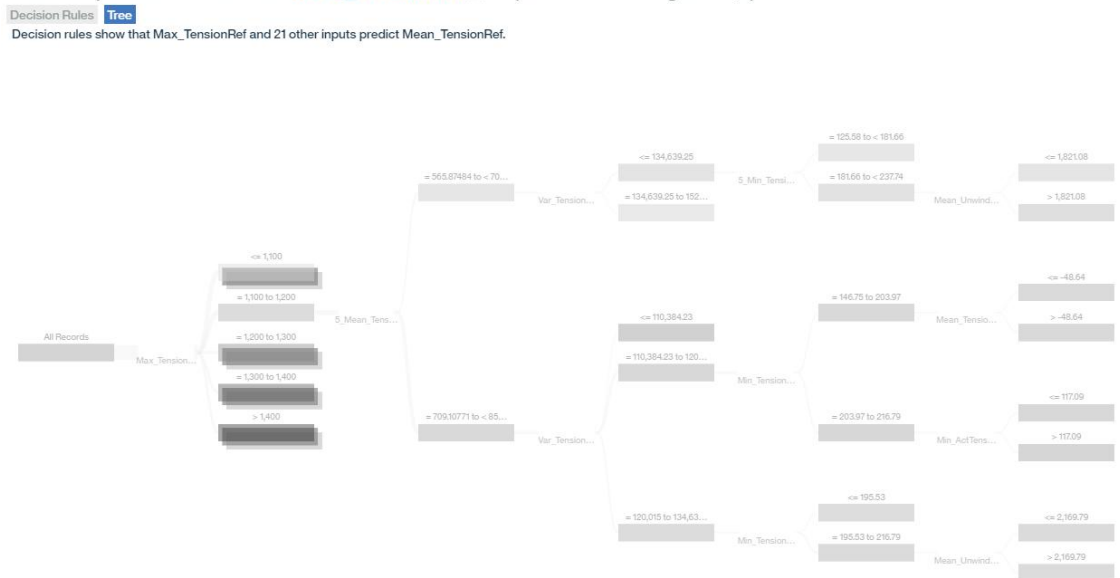
IBM Watson Analytics on datan esittämiseen hyvin soveltuva työkalu, koska se sisältää runsaasti erilaisia selkeitä kaavioita. Päättöspuu sekä spiraalikaavio ovat IBM Watson Analyticsin ennusteita tarjoavia kaavioita. Kuvassa 24 on spiraalikaavio, joka kuvaa valitun muuttujan kanssa eniten korreloivat muut muuttujat. Mitä lähempänä kuvaajan piste on keskipistettä, sitä tarkemmin muuttuja tai kahden muuttujan yhdistelmä korreloi valitun muuttujan kanssa. Kuvan esimerkissä on valittu muuttujaksi referenssikireyden keskiarvo. Kuvassa vasemmalla on spiraali, josta näkee sinisellä yhden muuttujan korrelaatiot ja oranssilla kahden muuttujan yhdistelmän korrelaatiot. Kuvassa oikealla on lista muuttujista korrelaatiojärjestyksessä. Listasta näkee, kuinka vahvasti muuttujat korreloivat keskenään. Klikkaamalla muuttujaa aukeaa uusi välilehti, joka kertoo enemmän muuttujien välisestä korrelaatiosta pylväsdiagrammin avulla.



Kuva 24. IBM Watson Analyticsin spiraalikaavio.

Kuvassa 25 on IBM Watson Analyticsin ennustavan mallin päätöspuu kireyden referenssin keskiarvolle. Tämä perustuu samaan malliin kuin päätössäännötkin, mutta antaa visuaalisemman kuvan muuttujien vuorovaikutuksista. Päätöspuun vasemmassa reunassa on kireyden referenssin keskiarvoon eniten vaikuttava muuttuja, ja siirryttäessä kohti oikeaa reunaa muuttujien vaikutus kireyden referenssin keskiarvoon pienenee.

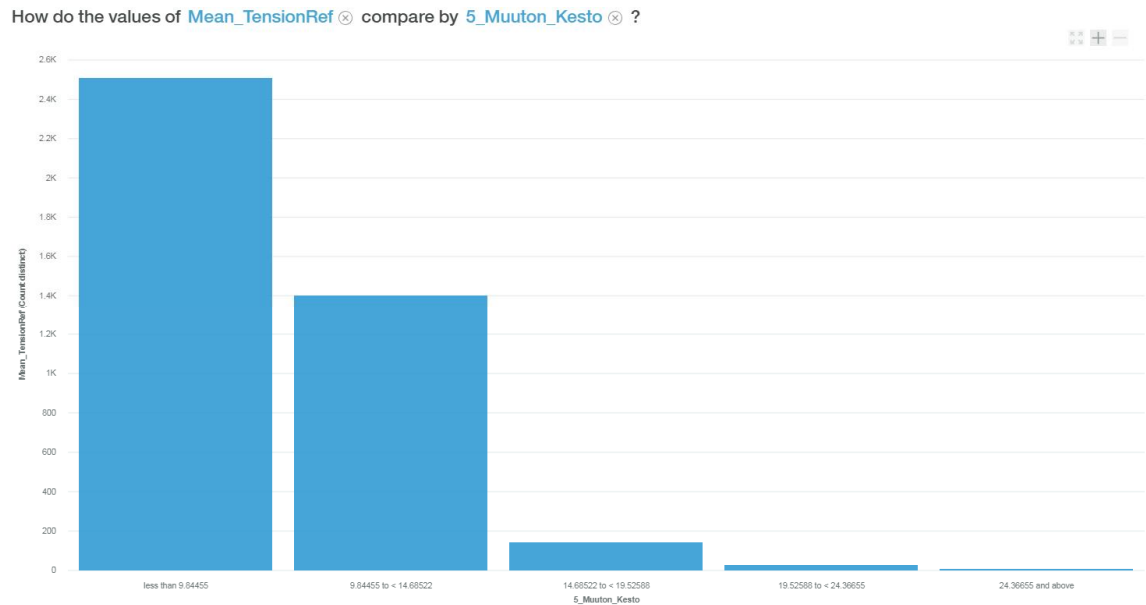
What is a predictive model for **Mean_TensionRef** ? (Predictive strength: 97%)



Kuva 25 IBM Watson Analyticsin ennustavan mallin päätöspuu.

Kuvassa 16 esitellyistä muista työkaluista ei ole muuta kuin visuaalista hyötyä, mutta visualisointikin voi olla joissain tapauksissa kiinnostavaa ja hyödyllistä. Dataa voi muokata useilla eri tavoilla näyttämään esimerkiksi havaintojen lukumäärää tai kertymää.

Kuvan 26 pylväskaavio yksi esimerkki IBM Watson Analyticsin visuaalisista näkymistä. Kuvasta nähdään kireyden referenssin keskiarvon jakauma muuton keston mukaan.



Kuva 26. IBM Watson Analyticsin pylväskaavio.

Microsoft Azure Machine Learning Studiolla ei ole yhtä kattavia visualisointimahdollisuuksia kuin IBM Watson Analyticsilla. Pylvästaulukko on palvelun ainut oma visualisointiväline, eikä sekään ole kovin kuvaava. Azurella on kuitenkin muita työkaluja (esimerkiksi Power BI), joiden avulla analysoidusta datasta saa paljonkin erilaisia visuaalisia kaavioita.

6. YHTEENVETO

6.1 Yleistä

Työkalut, joita tässä työssä käytettiin, olivat kokeiluversiona oikeista analyysityökaluista. Sekä IBM Watson Analytics että Microsoft Azure Machine Learning Studio tarjosivat 30 päivän kokeiluversion, joten työn tekemiseen tarvittiin monta eri sähköpostiosoitetta, koska työkalujen tutkimiseen ja datan tarkasteluun kului enemmän kuin 30 päivää. Watsonilla ja Azurella on kattava tuoteportfolio. Portfolioissa on erilaisia pilvilaskenta- ja -varastopalveluita.

Tässä työssä käytetyllä datalla on vaikea saada mielenkiintoisia tuloksia. Paperitehtaissa tehdään useita eri paperilaatuja. Paperimassan koostumus voi vaihdella ja haluttu paperin neliömassakin voi vaikuttaa eri tavalla paperileikkurin datan käyttäytymiseen. Paperikoneiden operaattoreilla voi olla erilaisia tapoja käyttää koneita. Tässä datassa oli vähän ratakatoja, joiden avulla olisi saanut luokiteltua datan onnistuneisiin ja epäonnistuneisiin muuttoihin. Paperikoneiden sähkökäytöt ovat monimutkaisia järjestelmiä, ja onnistuneet tulokset vaativat onnistunutta datan keruuta ja käsittelyä.

6.2 Tekoälyn hyödyntäminen paperikonekäytöissä

Tekoälyä voidaan paperikonekäytöissä hyödyntää monella tavalla. Kolme tärkeää käyttökohdetta ovat ennakoiva kunnossapito, vikatapausten selvitys ja parempi suunnittelu.

6.2.1 Ennakoiva kunnossapito

Laitteiden huolto on tehtaille yleensä aikaperusteista. Vuosihuollon aikana vaihdetaan osia uusiin, ja laitteita huolletaan ja rasvataan. Tekoälyn avulla on mahdollista pystyä tunnistamaan mahdolliset vikaantumiset esimerkiksi laitteiden lämpötilojen noususta. Koneoppimisen algoritmeilla on mahdollista oppia, mitkä ovat olleet syitä mahdollisiin laiterikkoihin. Tällöin tietokone voi kertoa huollon järjestäjälle, milloin laite todennäköisesti vikaantuu.

IBM Watson tarjoaa ennakoivaan kunnossapitoon IBM Prescriptive Maintenance for Manufacturing -tuotetta. Tuote seuraa laitteen sekä ympäristön dataa ja osaa ennustaa tulevasta rikkoutumisesta koneoppimisen ja analytiikan avulla. Tuotteen vähimmäistietovaatimuksena on laitteen aikaisemmat vikatiedot, normaali operointidata sekä riittävän korkea näytteenottotaajuus.

Azure tarjoaa arkkitehtuuria, jolla ennakoiva kunnossapito onnistuu. Arkkitehtuuriin kuuluu portfolion eri tuotteita kuten SQL-tietokantaa, koneoppimista ja analytiikkatyökaluja. Azure opastaa sivuillaan, kuinka ennakoivan kunnossapidon koneoppimisen voi ohjelmoida Machine Learning Studion avulla. Azuren sivuilla on Azure IoT -ratkaisusta kattava esittely, jossa kerrotaan tuotteesta ja opastetaan asiakasta sen käyttöön.

6.2.2 Vianmääritys

Vikatilanteissa tai ratakatkojen sattuessa ilman ennakointia voi tekoälyn avulla etsiä nopeasti todennäköisimmät syyt rikkoon. Tekoälyn avulla yllättävän rikon tiedot ja analyysi saadaan nopeasti operaattoreiden tietoon, ja vika päästään korjaamaan ilman suurempaa viivettä.

Sekä IBM Watson että Azure tarjoavat IoT-ratkaisujensa mukana myös vianetsintätyökaluja, joiden avulla voidaan selvittää, minkä takia laite on hajonnut. Ratakatkotiedon ollessa järjestelmässä myös ratakatkon syyt saadaan selville.

6.2.3 Parempi suunnittelu

Tekoälyn käytöstä saatu hyöty voi myös olla laitteiden parempi suunnittelu. Kun ennakoivan kunnossapidon ja vianmääritysten datasta on opittu laitteiden ja prosessien ongelmakohtia, niihin voi tulevia projekteja ja niiden mekaanisia tai ohjelmallisia ominaisuuksia suunniteltaessa puuttua jo etukäteen. Säätoratkaisut saadaan optimoituja, kun tiedetään, millä moottorin kiihtyvyyksillä on saatu parhaat ominaisuudet.

Paperitehtaiden tehtävänä on valmistaa riittävän hyvää paperia niin, että asiakkaat eivät tee reklamaatioita. Näin ollen laitteet on optimoitava niin, että prosessi toimii, mutta mahdollisimman kustannustehokkaasti. Tekoälyn mahdollisuudet ovat rajattomat ja jos vain järjestelmät sen sallivat, voi paperitehdas optimoida koko prosessin hyödyntäen esimerkiksi raaka-aineiden hintaa tai ulkolämpötilaa. Ihmisen on vaikeaa tulkita monen muuttujan korrelaatioita, mutta tietokone pystyy siihen vaivatta.

6.3 Mahdollisuudet

Paperikoneiden sähkökäytöt ovat monimutkaisia prosesseja ja paperinvalmistuksessa on useita tuloksiin vaikuttavia parametrejä. Asiakkaita kiinnostaa lähinnä prosessin jatkuvuus, eli ratakatkoja tulisi olla mahdollisimman vähän eivätkä koneet saisi vikaantua. Paperin laadun tulee olla riittävää, jotta reklamaatioita ei tule. Tässä työssä tehty taulukoitu data ei ole riittävä, jotta siitä saisi asiakkaille hyödyllisiä tuloksia.

Kuitenkin datarakenne ja sen prosessointi ovat hyödyllisiä mahdollisia tulevaisuuden tutkimuksia varten.

Nykyiset pilvipalvelut, kuten tässä työssä tutkitut IBM Watson ja Microsoft Azure, tarjoavat yrityksille kattavia portfolioita eri tarpeisiin. Molemmissa palveluissa on työkalut ennustavaan kunnossapitoon ja prosessien optimointiin. Pilvipalveluiden etuna on niiden hinta käytön mukaan, mutta paljon tarkkaa dataa keräävillä prosesseilla se voi käydä liian kalliiksi. Ennakoiva huolto on monelle tehtaalle hyödyllinen, sillä sen avulla saadaan optimoitua huoltovälit mahdollisimman pitkiksi. Paperitehtailla huoltoajat on usein vakioitu, mutta jos laitteet eivät vaikuta vikaantuneilta, niitä ei kannata vaihtaa. Paperikoneen sähkökäyttäjien valmistajana ABB voi halutessaan myydä MaaS-palvelua ja optimoida asiakkaiden huoltokustannukset. ABB pystyy kerätyn datan perusteella tekemään parempaa tuotesuunnittelua tulevaisuudessa.

Tekoälyn mahdollisuudet paperikoneiden sähkökäytössä ovat rajattomat, kunhan vain käytettävissä on riittävästi resursseja. Prosessit voidaan antaa täysin tietokoneiden hallintaan, jotta ne tyydyttäisivät kustannuksien ja laadun tarpeet optimaalisesti. Ihmismielen on vaikea luopua päätäntävällästä, joten siihen ei olla varmastikaan vielä menossa, että tietokoneet ohjaisivat koko tehdasta ilman ihmisen lupaa tehdä tärkeitä muutoksia.

6.4 Tulevaisuuden töitä

Prosessi datan keräämiseen on tällä hetkellä käsityötä ja suuret tietokannat siirtyvät todella hitaasti etäyhteyttä pitkin. Työssä käytetty data siirrettiin ulkoisella kiintolevyllä. Pilviverkon kautta datan siirtäminen olisi nopeampaa, mutta tietoturvariski ja hinta nousevat. Datan analysoinnin tulisi olla mahdollisimman reaaliaikaista, jotta vikaantumiset saataisiin nopeasti selvitettyä. Paperitehtaiden suurin motiivi on pitää koneet pyörimässä mahdollisimman jatkuvasti. Tästä syystä on hankala kerätä dataa, jossa vikaantumisia on paljon. Vikojen yleisyys helpottaisi vikatilanteiden syiden tutkimista.

Mielenkiintoisena tutkimuskohteena voisi olla uusi paperikone, jota tekoäly ohjaa alusta asti ilman ohjeita. Näin ollen tietokone oppisi vikojen myötä optimaalisimman tavan käyttää kyseistä konetta. Parametrien tulisi määritellä tarkasti, jotta tulokset olisivat hyödyllisiä.

LÄHTEET

- [1] M. O. Santos ja B. Almada-Lobo, ”Integrated pulp and paper mill planning and scheduling”, *Comput. Ind. Eng.*, vsk. 63, nro 1, ss. 1–12, elo 2012.
- [2] Metlas, ”Paperikoneet: yleistä ”, vsk. M-502.;M-5. Metlas , Hki? , 1990.
- [3] D. Chu, M. Forbes, J. Backstrom, C. Gheorghe, ja S. Chu, ”Model Predictive Control and Optimization for Papermaking Processes”, teoksessa *Advanced Model Predictive Control*, InTech, 2011.
- [4] P. Rautiainen, *Papermaking science and technology: Book 10 Papermaking : part 3, finishing*, 2nd, total p. Helsinki: Finnish Paper Engineers’ Association, 2010.
- [5] H. Holik, *Handbook of paper and board*. Wiley-VCH Verlag GmbH, 2013.
- [6] K. G. Frye ja D. C. Fitzpatrick, ”Slitter for high bulk traveling paper web material”, US4274319 A, 1979.
- [7] Valmet, ”Two-drum winders”. [Verkossa]. Saatavissa: <http://www.valmet.com/board-and-paper/board-and-paper-machines/winders-and-roll-handling/winding/two-drum-winders/>. [Viitattu: 10-marras-2017].
- [8] A. Veltman, D. W. J. Pulle, ja R. W. De Doncker, *Fundamentals of Electrical Drives*, Second; 2; Cham: Springer Verlag, 2016.
- [9] L. Aura ja A. J. Tonteri, *Sähkökoneet ja tehoelektroniikan perusteet*. Helsinki; Juva; Porvoo: WSOY, 1996.
- [10] ”Kestomagneettimoottorit - SynRM (IEC-pienjännitemoottorit) | ABB”. [Verkossa]. Saatavissa: <http://new.abb.com/motors-generators/fi/iec-pienjannitemoottorit/SynRM/kestomagneettimoottorit>. [Viitattu: 22-marras-2017].
- [11] ”Working or Operating Principle of DC Motor | Electrical4u”. [Verkossa]. Saatavissa: <https://www.electrical4u.com/working-or-operating-principle-of-dc-motor/>. [Viitattu: 19-marras-2017].
- [12] J. Niiranen, *Sähkomoottorikäytön digitaalinen ohjaus*, 2. korj. p., vsk. 590.; 590. Helsinki: Otatieto, 2000.
- [13] ”Electric Drives - Control Systems - Description and Applications”. [Verkossa]. Saatavissa: <http://www.mpoweruk.com/motorcontrols.htm>. [Viitattu: 03-maaliskuu-2018].
- [14] ”EtherNet/IP leads Industrial network market shares | Automation.com”. [Verkossa]. Saatavissa: <https://www.automation.com/automation-news/industry/ethernetip-leads-industrial-network-market-shares>. [Viitattu: 04-

joulu-2017].

- [15] S. J. Russell, P. Norvig, J. F. Canny, J. M. Malik, ja D. D. Edwards, *Artificial intelligence: a modern approach*. London: Prentice-Hall International, 1995.
- [16] E. Gibney, ”Google masters Go”, *Nature*, vsk. 529, nro 7587, s. 445, 2016.
- [17] M. Day, ”Microsoft takes AI chatbot offline after it repeats offensive remarks”, *The Seattle Times*, marras 2016.
- [18] B. Wierenga, ”Marketing and Artificial Intelligence: Great Opportunities, Reluctant Partners”, vsk. 258, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, ss. 1–8.
- [19] E. Hyvönen, I. Karanta, ja M. Syrjänen, *Tekoälyn ensyklopedia*. Gaudeamus, 1993.
- [20] P. Koistinen, ”Tilastollinen hahmontunnistus”, 2002.
- [21] ”Learning system in real-time machine vision”, *Neurocomputing*, vsk. 288, ss. 1–2, touko 2018.
- [22] D. Ciresan, U. Meier, ja J. Schmidhuber, ”Multi-column deep neural networks for image classification”, teoksessa *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, ss. 3642–3649.
- [23] ”ValueWalk: AI Weapons Should Be Restricted, According to Elon Musk And Other Robotics Experts - ProQuest”. [Verkossa]. Saatavissa: <https://search-proquest-com.libproxy.tut.fi/docview/1930509710?pq-origsite=summon&http://search.proquest.com/abicomplete/advanced>. [Viitattu: 08-marras-2017].
- [24] D. Helbing *ym.*, ”Will Democracy Survive Big Data and Artificial Intelligence?”, 2017.
- [25] R. Kurzweil, ”The Singularity is Near”, teoksessa *Ethics and Emerging Technologies*, R. L. Sandler, Toim. London: Palgrave Macmillan UK, 2014, ss. 393–406.
- [26] ”Putin: Leader in artificial intelligence will rule world”, 2017. [Verkossa]. Saatavissa: <https://www.cnbc.com/2017/09/04/putin-leader-in-artificial-intelligence-will-rule-world.html>. [Viitattu: 08-marras-2017].
- [27] K. Ashton, ”That ’Internet of Things’ Thing - 2009-06-22 - Page 1 - RFID Journal”. [Verkossa]. Saatavissa: <http://www.rfidjournal.com/articles/view?4986>. [Viitattu: 05-elo-2018].
- [28] L. Atzori, A. Iera, ja G. Morabito, ”The Internet of Things: A survey”, *Comput. Networks*, vsk. 54, nro 15, ss. 2787–2805, 2010.
- [29] M. Hermann, T. Pentek, ja B. Otto, ”Design Principles for Industrie 4.0 Scenarios”, teoksessa *2016 49th Hawaii International Conference on System Sciences*

- (HICSS), 2016, ss. 3928–3937.
- [30] H. Sultan, ”Internet of Things”, *Int. J. Adv. Res. Comput. Sci.*, vsk. 8, nro 2, s. 4, 2017.
- [31] ”Lösungen – Bausteine der Industrie 4.0 - Yokogawa Deutschland”. [Verkossa]. Saatavissa: <https://www.yokogawa.com/de/loesungen/industrie-4-0.htm>. [Viitattu: 03-maalis-2018].
- [32] A. V. R. S. J. Lheureux, ”Hype Cycle for the Internet of Things, 2016”, *Gartner*, ss. 1–69, 2016.
- [33] L. Bradshaw, ”Big Data And What It Means”.
- [34] T. Markkula ja A. Syväniemi, *The Analytics Journey*. Hansaprint Oy, 2016.
- [35] J. Han ja M. Kamber, *Data mining: concepts and techniques*. San Francisco (CA): Morgan Kaufmann, 2001.
- [36] M. Hazewinkel, *Encyclopaedia of mathematics*. Springer-Verlag, 2002.
- [37] J. Daintith ja E. Wright, *A Dictionary of Computing*. Oxford University Press, 2008.
- [38] ”Bayesian linear regression Thomas P. Minka”, 1998.
- [39] S. Chakraverty ja S. Mall, *Artificial Neural Networks for Engineers and Scientists*. Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742: CRC Press, 2017.
- [40] B. Schmidt, L. Wang, I. för ingenjörsvetenskap, H. i Skövde, ja F. för Virtuella system, ”Cloud-enhanced predictive maintenance”, *Int. J. Adv. Manuf. Technol.*, 2016.
- [41] P. Mell ja T. Grance, ”The NIST Final Version of NIST Cloud Computing Definition Published”, *Nist Spec. Publ.*, vsk. 145, s. 7, 2011.
- [42] ”A SURVEY: CLOUD COMPUTING SECURITY ISSUES AND TECHNIQUES”, *Int. J. Adv. Res. Comput. Sci.*, vsk. 8, nro 7.
- [43] P. Heino, *Pilvipalvelut*. Helsinki: Talentum Media, 2010.
- [44] J. Hurwitz, M. Kaufman, ja A. Bowles, *Cognitive computing and big data analytics*. .
- [45] J. Valta, ”Tekoäly Watson oppii suomea turkulaisilta”, 2017.
- [46] Y. Chen, E. Argentinis, ja G. Weber, ”IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life Sciences Research”, *Clinical Therapeutics*, vsk. 38, nro 4. ss. 688–701, 2016.

- [47] ”Inderes osakevertailu | Inderes: Osakeanalyysit, mallisalkku, osakevertailu & aamukatsaus”. [Verkossa]. Saatavissa: <https://www.inderes.fi/osakevertailu>. [Viitattu: 05-heinä-2018].
- [48] ”ABOUT AIEQ – The AI Powered Equity ETF”. [Verkossa]. Saatavissa: <http://www.equbotetf.com/about-aieq/>. [Viitattu: 05-heinä-2018].
- [49] ”A new ETF uses IBM’s AI technology to help it pick winning stocks - MarketWatch”. [Verkossa]. Saatavissa: <https://www.marketwatch.com/story/a-new-etf-uses-ibms-ai-technology-to-help-it-pick-winning-stocks-2017-10-18>. [Viitattu: 05-heinä-2018].
- [50] ”Stock Comparison: Compare AI Powered Equity ETF (AIEQ) to Other Stocks.” [Verkossa]. Saatavissa: <https://www.nasdaq.com/symbol/aieq/stock-comparison>. [Viitattu: 05-heinä-2018].
- [51] ”IBM Plant Performance Analytics - Details - United States”. [Verkossa]. Saatavissa: <https://www.ibm.com/us-en/marketplace/plant-performance-analytics/details#product-header-top>. [Viitattu: 14-heinä-2018].
- [52] ”Industry 4.0 and Connected Manufacturing - IBM Watson IoT”. [Verkossa]. Saatavissa: <https://www.ibm.com/internet-of-things/industries/iot-manufacturing>. [Viitattu: 04-maalis-2018].
- [53] ”Microsoft Azure: Planning, Deploying, and Managing Your Data Center in the Cloud - Books24x7”. [Verkossa]. Saatavissa: <http://library.books24x7.com/libproxy.tut.fi/assetviewer.aspx?bookid=101449&chunkid=962306412¬eMenuToggle=0&leftMenuState=1>. [Viitattu: 22-helmi-2018].
- [54] ”Defect prevention with predictive maintenance | Microsoft Azure”. [Verkossa]. Saatavissa: <https://azure.microsoft.com/en-us/solutions/architecture/defect-prevention-with-predictive-maintenance/>. [Viitattu: 30-touko-2018].
- [55] ”Cloudera and Twitter bring Parquet to the Hadoop floor - JAXenter”. [Verkossa]. Saatavissa: <https://jaxenter.com/cloudera-and-twitter-bring-parquet-to-the-hadoop-floor-105802.html>. [Viitattu: 29-heinä-2018].
- [56] ”Python eats away at R: Top Software for Analytics, Data Science, Machine Learning in 2018: Trends and Analysis”. [Verkossa]. Saatavissa: <https://www.kdnuggets.com/2018/05/poll-tools-analytics-data-science-machine-learning-results.html>. [Viitattu: 30-touko-2018].

LIITE A: IBM WATSON ANALYTICSIN ENNUSTEMALLIN SÄÄNTÖLUETTELO KIREYDEN REFERENSSIN KESKIARVOLLE

What is a predictive model for **Mean_TensionRef** ? (Predictive strength: 97%)

Decision Rules Tree
Decision rules show that Max_TensionRef and 21 other inputs predict Mean_TensionRef.

▲▼ Predicted value	Rules	Records
962.90	Max_TensionRef > 1400 S_Var_TensionRef = 153613.58313 to < 201249.67877 S_Mean_UnwindDiam = 132170181 to < 1694.72564; 1694.72564 to < 2067.69946 S_Var_KiikkyysHidastuvuus = 0.20572 and above	54
956.18	Max_TensionRef > 1400 S_Var_TensionRef = 153613.58313 to < 201249.67877 S_Mean_UnwindDiam = 132170181 to < 1694.72564; 1694.72564 to < 2067.69946 S_Var_KiikkyysHidastuvuus = 0.07142 to < 0.11618; 0.11618 to < 0.16095; 0.16095 to < 0.20572	96
955.07	Max_TensionRef > 1400 S_Var_TensionRef = 153613.58313 to < 201249.67877 S_Mean_UnwindDiam = 2067.69946 to < 2440.67328 S_Min_UnwindDiam = 1643.66255 to < 2025.19683	60
948.47	Max_TensionRef > 1400 S_Var_TensionRef = 153613.58313 to < 201249.67877 S_Mean_UnwindDiam = 2067.69946 to < 2440.67328 S_Min_UnwindDiam = 2025.19683 to < 2406.7311	55
939.55	Max_TensionRef > 1400 S_Var_TensionRef = 153613.58313 to < 201249.67877 S_Mean_UnwindDiam = 2440.67328 and above; less than 132170181	79
905.29	Max_TensionRef > 1400 S_Var_TensionRef = 105977.48749 to < 153613.58313; 201249.67877 and above; 58341.39185 to < 105977.48749 Max_UnwindDiam > 1722.39	119
876.72	Max_TensionRef = 1300 to 1400 S_Mean_TensionRef = 852.34058 to < 995.67346 Max_TorqueRef > 19.42	68
865.54	Max_TensionRef = 1300 to 1400 S_Mean_TensionRef = 852.34058 to < 995.67346 Max_TorqueRef = 19.42 S_Min_TensionRef = 237.74 to < 293.82; 293.82 and above	108
862.66	Max_TensionRef = 1300 to 1400 S_Mean_TensionRef = 852.34058 to < 995.67346 Max_TorqueRef = 19.42 S_Min_TensionRef = 181.66 to < 237.74	97
841.61	Max_TensionRef > 1400 S_Var_TensionRef = 105977.48749 to < 153613.58313; 201249.67877 and above; 58341.39185 to < 105977.48749 Max_UnwindDiam <= 1722.39	50
840.11	Max_TensionRef = 1300 to 1400 S_Mean_TensionRef = 709.10771 to < 852.34058; less than 565.87484 Min_UnwindDiam > 1060.54 S_Max_TensionRef = 1398.02 and above	111
815.13	Max_TensionRef = 1200 to 1300 Var_TensionRef = 110.38423 to 134.63925 S_Min_ActTension = 125.77657 to < 195.11254; 195.11254 to < 266.44971; 266.44971 and above; less than 55.4385	70
804.46	Max_TensionRef = 1300 to 1400 S_Mean_TensionRef = 709.10771 to < 852.34058; less than 565.87484 Min_UnwindDiam > 1060.54 S_Max_TensionRef = 1246.03999 to < 1398.02	65
802.74	Max_TensionRef = 1200 to 1300 Var_TensionRef = 110.38423 to 134.63925 S_Min_ActTension = 55.4385 to < 125.77657 S_Min_TensionRef = 237.74 to < 293.82	89
797.94	Max_TensionRef = 1200 to 1300 Var_TensionRef = 110.38423 to 134.63925 S_Min_ActTension = 55.4385 to < 125.77657 S_Min_TensionRef = 181.66 to < 237.74	65
795.70	Max_TensionRef = 1200 to 1300 Var_TensionRef = 134.63925 to 152.15888 Min_TensionRef = 195.53 to 203.97 Max_UnwindDiam <= 2,052.49	56
792.49	Max_TensionRef = 1200 to 1300 Var_TensionRef = 134.63925 to 152.15888 Min_TensionRef = 203.97 to 216.79	97
788.48	Max_TensionRef = 1200 to 1300 Var_TensionRef = 134.63925 to 152.15888 Min_TensionRef = 195.53 to 203.97 Max_UnwindDiam = 2,052.49 to 2,638.26	113
785.51	Max_TensionRef = 1200 to 1300 Var_TensionRef = 134.63925 to 152.15888 Min_TensionRef = 195.53 to 203.97 Max_UnwindDiam > 2,638.26	55
776.41	Max_TensionRef = 1200 to 1300 Var_TensionRef = 134.63925 to 152.15888 Min_TensionRef = 146.75 to 195.53	75
774.45	Max_TensionRef = 1300 to 1400 S_Mean_TensionRef = 709.10771 to < 852.34058; less than 565.87484 Min_UnwindDiam <= 1060.54	61
763.03	Max_TensionRef = 1,100 to 1,200 S_Mean_TensionRef = 709.10771 to < 852.34058 Var_TensionRef <= 110.38423	51
757.36	Max_TensionRef = 1200 to 1300 Var_TensionRef = 134.63925 to 152.15888 Min_TensionRef > 216.79	84
744.06	Max_TensionRef = 1,100 to 1,200 S_Mean_TensionRef = 709.10771 to < 852.34058 Var_TensionRef = 110.38423 to 120.015 more...	60

738.29	Max_TensionRef = 1,100 to 1,200 S_Mean_TensionRef = 709,10771 to < 852,34058 Var_TensionRef = 110,384,23 to 120,015 more...	227
732.07	Max_TensionRef = 1,100 to 1,200 S_Mean_TensionRef = 709,10771 to < 852,34058 Var_TensionRef = 110,384,23 to 120,015 more...	147
727.32	Max_TensionRef = 1,100 to 1,200 S_Mean_TensionRef = 709,10771 to < 852,34058 Var_TensionRef = 110,384,23 to 120,015 more...	59
723.74	Max_TensionRef = 1,100 to 1,200 S_Mean_TensionRef = 709,10771 to < 852,34058 Var_TensionRef = 120,015 to 134,639,25 more...	88
720.62	Max_TensionRef = 1,100 to 1,200 S_Mean_TensionRef = 709,10771 to < 852,34058 Var_TensionRef = 120,015 to 134,639,25 more...	106
718.97	Max_TensionRef = 1,200 to 1,300 Var_TensionRef > 152,158,88	59
717.87	Max_TensionRef = 1,100 to 1,200 S_Mean_TensionRef = 709,10771 to < 852,34058 Var_TensionRef = 120,015 to 134,639,25 Min_TensionRef <= 195,53	94
705.01	Max_TensionRef = 1,100 to 1,200 S_Mean_TensionRef = 565,87484 to < 709,10771 Var_TensionRef <= 134,639,25 more...	58
685.05	Max_TensionRef = 1,100 to 1,200 S_Mean_TensionRef = 565,87484 to < 709,10771 Var_TensionRef <= 134,639,25 more...	77
673.33	Max_TensionRef = 1,100 to 1,200 S_Mean_TensionRef = 565,87484 to < 709,10771 Var_TensionRef <= 134,639,25 S_Min_TensionRef = 125,58 to < 131,66	68
671.75	Max_TensionRef <= 1,100 Var_TensionRef <= 110,384,23 Min_TensionRef > 146,75 more...	58
666.83	Max_TensionRef <= 1,100 Var_TensionRef <= 110,384,23 Min_TensionRef > 146,75 more...	132
665.99	Max_TensionRef = 1,100 to 1,200 S_Mean_TensionRef = 565,87484 to < 709,10771 Var_TensionRef = 134,639,25 to 152,158,88	76
657.16	Max_TensionRef <= 1,100 Var_TensionRef <= 110,384,23 Min_TensionRef > 146,75 more...	192
656.69	Max_TensionRef <= 1,100 Var_TensionRef <= 110,384,23 Min_TensionRef > 146,75 more...	85
652.59	Max_TensionRef <= 1,100 Var_TensionRef <= 110,384,23 Min_TensionRef > 146,75 more...	58
651.47	Max_TensionRef <= 1,100 Var_TensionRef <= 110,384,23 Min_TensionRef <= 146,75 more...	125
648.30	Max_TensionRef <= 1,100 Var_TensionRef <= 110,384,23 Min_TensionRef <= 146,75 more...	53
643.86	Max_TensionRef <= 1,100 Var_TensionRef = 110,384,23 to 120,015 Min_MotorSpeed <= -0,46	54
639.60	Max_TensionRef <= 1,100 Var_TensionRef <= 110,384,23 Min_TensionRef <= 146,75 more...	55
637.74	Max_TensionRef <= 1,100 Var_TensionRef = 110,384,23 to 120,015 Min_MotorSpeed = -0,46 to -0,34	80
634.30	Max_TensionRef <= 1,100 Var_TensionRef = 110,384,23 to 120,015 Min_MotorSpeed > -0,34 Max_ActTension = 1,153,89 to 1,233,5	75
629.74	Max_TensionRef <= 1,100 Var_TensionRef = 110,384,23 to 120,015 Min_MotorSpeed > -0,34 Max_ActTension <= 1,153,89	65
609.68	Max_TensionRef <= 1,100 Var_TensionRef = 120,015 to 134,639,25 Min_TensionRef = 146,75 to 195,53	99