vAI:lla tuottavuutta?

Tuottavuuspilotti, kevät 2025

Skaala IFN: Teränvaihdon optimointi anturidatan ja koneoppimisen avulla



Euroopan unionin osarahoittama

Tampereen yliopisto

SEAMK
Seinäjoen ammattikorkeakoulu
Seinäjoki University of Applied Sciences

Vaasan yliopisto
UNIVERSITY OF VAASA

ETELÄ-POHJANMAAN LIITTO
Regional Council of South Ostrobothnia

Pilotin tiedot

- Seinäjoen Ammattikorkeakoulu ja vAI:lla tuottavuutta? -hanke toteutti pilotin yhteistyössä Skaala IFN Oy:n kanssa heidän höylälinjastosta kerätyllä sensoridatalla keväällä 2025.
- Kiitämme myös Tampereen yliopiston Open Iot Hub (MOI) –hanketta datan keräämisessä pilotin toteuttamisen mahdollistamiseksi.
  - Seinäjoki Multidisciplinary Open IoT Hub (MOI) | Tampereen korkeakouluyhteisö

Pilotin aihe 1/2

- Pilotin tavoitteena oli **pilotoida tekoälyn hyödyntämistä matalilla kustannuksilla höylälinjan viimeisen teräpakan vaihdon ajoittamiseksi** tuotantokatkosten minimoimiseksi ja laadun varmistamiseksi.
- Pilotissa hyödynnettiin MOI Hub -hankkeen keräämää sensoridataa, joka koostui kolmiakselisesta (X, Y, Z) tärinädatasta sekä metadatasta eli terienvaihtoajankohdista.
  - Dataa kerättiin jatkuvasta tuotannosta n. neljän viikon ajan kattaen useita teränvaihtosyklejä.
  - Tärinädata analysoitiin Pythonilla rakennetulla koneoppimismallilla, jossa hyödynnettiin Satunnaismetsäalgoritmia eli Random Forest –algoritmia.

Pilotin aihe 2/2

- Mallin opettamisessa käytettiin sensoridataa käsityksen muodostamisesta siitä, miten tärinäprofiilit muuttuivat terän kulumisen edetessä ja missä vaiheessa vaihtotarve tyypillisesti ilmeni.
- Mallia testattiin käyttämällä kerättyä sensoridataa, jossa algoritmin ennustamia vaihtopisteitä verrattiin todellisiin, työntekijöiden tekemiin päätöksiin teränvaihdon ajankohdasta.
- **Pilotin tulokset olivat lupaavia:** malli kykeni tunnistamaan vaihtotarpeen useissa tapauksissa lähellä todellisia vaihtokohtia. Malli antoi kuitenkin myös osaksi ylimääräisiä signaaleja. Tulokset kuitenkin osoittavat matalakustannuksisen koneoppimisen potentiaalin tuotantoprosessien kunnossapidon ja päätöksenteon tukena.

Käytetyt teknologiat

- Kolmiakselinen tärinädata (X, Y, Z, aikaleima), kerätty laitteen ulkoseinästä anturilla
  - Mittaustaajuus n. 400 Hz, 10 minuutin välein 30 sekuntia dataa -> ~12000 riviä dataa / akseli
  - Ei johdotuksia tai vastaavia kytkentöjä koneeseen, tärinädata lähetettiin langattomasti palvelimelle
  - Datankeruuseen voi perehtyä tarkemmin MOI-Hub –hankkeen aineistosta
- **Python:** Käytetty pääohjelmointikielenä kaikessa datankäsittelyssä ja analyysitehtävissä.
- **Pandas:** Hyödynnetty CSV-datan lukemiseen, sen lajitteluun ajan mukaan sekä ryhmittelyyn ajanjaksoihin.
- **NumPy**: Käytetty numeeristen laskelmien suorittamiseen ja taulukoiden käsittelyyn ominaisuuksien valmistelussa.
- **datetime:** Hyödynnetty aikaleimojen hallintaan ja aikaerojen laskemiseen (esimerkiksi ikkunoiden ja tapahtumien ajoituksen määrittelyssä).
- **scikit-learn:** Käytetty RandomForestClassifier-mallin koulutukseen sekä sen suorituskyvyn arviointiin eri mittareiden avulla.
- **Matplotlib**: Käytetty käyrien ja graafien piirtämiseen, jotta datan trendejä, ennusteita ja arviointituloksia voidaan visualisoida.
- **matplotlib.dates:** Käytetty päivämääräakselien muotoiluun ja lukujen selkeyttämiseen graafien esityksissä.

Tekoäly ratkaisuna

- **Miten tämä on tekoälyä**
  - Koneoppimismalli "oppii" datan perusteella, eli se analysoi suuria määriä esikäsiteltyä tietoa ja tunnistaa monimutkaisia malleja, joita ihminen ei välttämättä huomaa. Tämä oppimisprosessi ja automaattinen päätöksenteko ovat tekoälyn ydinelementtejä. Tekoälyä käytetään siis tässä ratkaisussa erottamaan normaalit toimintatilanteet poikkeavista tilanteista (teränvaihdot), mikä mahdollistaa ennakoivan huollon.
- **Ominaisuuksien generointi ja datan valmistelu**
  - Dataa käsitellään ensin perinteisillä tilastollisilla menetelmillä, jolloin raakadata jaetaan aikaväleihin ja lasketaan erilaisia tilastollisia arvoja (kuten keskiarvo, keskihajonta, liukuvat tilastot, viiveominaisuudet, muutoksen nopeus sekä aikaominaisuudet kuten kellonaika ja viikonpäivä). Nämä ominaisuudet muodostavat "piirteet", jotka kuvaavat datan käyttäytymistä eri ajankohtina.
- **Mallin koulutus tekoälymenetelmällä**
  - Kun data on valmiiksi esikäsitelty, sitä käytetään kouluttamaan satunnaismetsämalli. Satunnaismetsä eli Random Forest on koneoppimisen malli, joka koostuu useista päätöspuista. Satunnaismetsä perustuu erilaisten päätöspuiden rakentamiseen, jotka oppivat tunnistamaan monimutkaisia kuvioita ja trendejä esikäsitellyistä ominaisuuksista. Näiden puiden tulokset yhdistetään, jotta saadaan päätös siitä, onko kyseessä normaalia vaihtelua vai todennäköisesti terän vaihdon aika (huoltotoimenpide).
- **Ennustaminen ja poikkeavien tilanteiden tunnistus**
  - Koulutettu malli käyttää oppimiaan kuvioita tehdäkseen ennusteita annetusta datasta. Se laskee todennäköisyyden, että tietty ajanjakso sisältää terän vaihdon ajan.

Python-koodi

- Pilotin avainasemassa on python-koodi, joka käsitteli sensoridataa.
- Python-koodi on saatavilla osoitteesta https://projektit.seamk.fi/alykkaat-teknologiat/vailla-tuottavuutta/
- Esimerkkitiedosto sensoridatasta löytyy myös saatavilla olevista tiedostoista, mutta sitä ei voi käyttää mallin kouluttamiseen. Tiedosto on saatavilla lähinnä formaatin näyttämiseksi.
- Seuraavissa dioissa avataan ratkaisua tarkemmin.

Koodin vaiheet 1/2

- **Datan lukeminen ja järjestely:**
  - CSV-tiedostosta luetaan tärinädata, jossa aikaleimat jäsennetään automaattisesti.
  - Data lajitellaan aikajärjestykseen, jotta se etenee kronologisesti.
- **Aikaväleihin ryhmittely ja tilastojen laskenta:**
  - Data jaetaan 1 minuutin väleihin.
  - Kussakin ikkunassa lasketaan tilastollisia arvoja: keskiarvo, keskihajonta, minimi, maksimi ja havaintojen lukumäärä.
  - Ikkunat, joissa ei ole dataa, suodatetaan pois.
- **Tunnettujen teränvaihtoaikojen määrittely:**
  - Tunnetut teränvaihtoajat on määritelty ja muunnettu datetime-olioiksi oikealla aikavyöhykkeellä.
  - Näitä käytetään vertailukohtana, jotta voidaan myöhemmin arvioida ennusteiden tarkkuutta.
- **Datan jakaminen ja ominaisuuksien valmistelu:**
  - Data jaetaan aikajärjestyksen perusteella koulutus- (70 %) ja testiaineistoon (30 %).
- **Luodaan useita ominaisuuksia:**
  - **Viiveominaisuudet:** Edellisten ikkunoiden keskihajonta ja keskiarvo.
  - **Liukuvat tilastot:** Liukuva keskiarvo, keskihajonta ja vaihteluväli eri ikkunoilla.
  - **Muutoksen nopeus:** Standardipoikkeaman muutossuhteet.
  - **Aikaominaisuudet:** Kellonaika ja viikonpäivä.

Koodin vaiheet 2/2

- **Mallin koulutus:**
  - Luodaan tavoitemuuttuja "blade_change", jossa ennustusikkuna (esim. 2 tuntia ennen tunnettuja muutoksia) merkitään arvolla 1.
  - Koulutusaineistosta muodostetaan ominaisuusjoukko ja tavoitemuuttuja.
  - Koulutetaan RandomForestClassifier, joka oppii erottamaan normaalit aikavälit poikkeavista (terämuutokset).
- **Ennusteet ja arviointi:**
  - Ennustetaan testidatasta mallin avulla teränvaihtotodennäköisyyksiä.
  - Ennustetut todennäköisyydet muunnetaan kynnysarvoa (0.07) käyttäen binaarisiksi ennusteiksi (kyllä/ei).
  - Ennusteiden osuvuutta arvioidaan vertaamalla niitä todellisiin (tunnettuihin) teränvaihtoaikoihin:
    - Lasketaan esimerkiksi herkkyys ja tarkkuus.
- **Visualisointi:**
  - Piirretään graafi, joissa:
    - Näytetään tärinädatan keskihajonnat (X-, Y- ja Z-akselit) aikajanalla.
    - Merkitään todelliset ja ennustetut terämuutokset pystysuorilla viivoilla (eriväriset).
    - Esitetään esimerkkiaikajakso oikein ennustetun tapahtuman ympäriltä, jotta havaitaan ennusteen osuvuus visuaalisesti.

Koneoppimisen tulokset (sokkotestijakso)

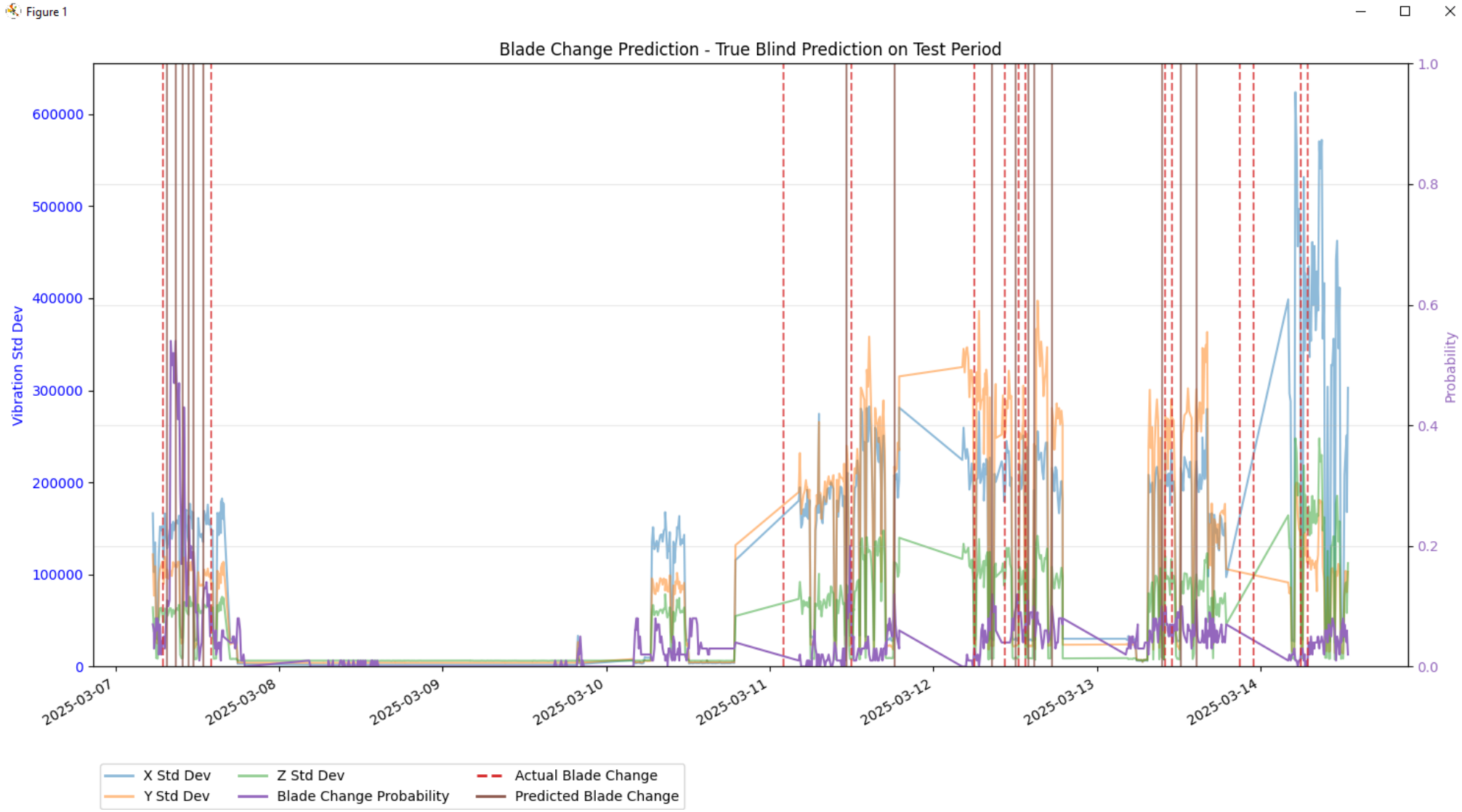
14 todellista teränvaihtoa testijaksolla

- 3 teränvaihtoa työvuorojen ulkopuolella
- 1. päivänä ensimmäinen vaihto heti testidatan alussa
  - Tekoäly antoi toistuvia signaaleja ennen 2. vaihtoa
- 3. päivänä signaali ennen oikeaa vaihtoa + päivän lopussa → oikea vaihto heti seuraavan päivän alussa
- 4. päivänä 4 vaihtoa:
  - Ensimmäinen signaali ajoissa ennen vaihtoa
  - Seuraavat vaihdot → tekoälyn signaalit osuivat vaihtojen lähelle
  - Vaihtojen jälkeen vielä 2 ylimääräistä signaalia
- 5. päivänä:
  - Signaali ennen kahta lähekkäistä vaihtoa
  - Myös vaihtojen jälkeen 2 signaalia, vaikka terät oli jo vaihdettu
  - Terät vaihdettiin työpäivien välissä
- 6. päivänä:
  - Vaihto päivän alussa, ei tekoälyn signaalia

Koneoppimisen tulokset (sokkotestijakso)

- Testijaksolla ennustettiin 16 mahdollista terän vaihtoa.
- Kun arvioitiin tunnistuksen tarkkuutta 1 tunnin aikavälin sisällä, todettiin:
  - **Todelliset onnistumiset:** Malli ennusti oikein 5 tapahtumaa.
  - **Väärät positiiviset:** Malli ennusti 11 kertaa tapahtuman, vaikka sellaista ei tapahtunut.
  - Havaittuja tapahtumia, jotka jäävät ennustamatta: 9 terävaihtoa, joita malli ei tunnistanut.
- Näistä lasketut arviointimittarit ovat:
  - **Herkkyys (Recall):** 0.36, mikä tarkoittaa, että mallin ennustuksista 36 % kaikista oikeista terämuutoksista havaittiin.
  - **Tarkkuus (Precision):** 0.31, eli mallin ennusteista 31 % osui oikeaan.
  - **F1 Score:** 0.33, joka yhdistää tarkkuuden ja herkkyyden tasapuolisesti.
- Tulokset osoittavat, että signaalit eivät ole satunnaisia, vaan ne seuraavat mitattavissa olevia piirteitä tärinädatassa. Tarkkuus ei kuitenkaan riitä luotettavaan tuotantokäyttöön tällaisenaan.

Koneoppimisen tulokset (sokkotestijakso)



Yhteenveto: Hyödyt

- **Automaattinen ja järjestelmällinen datankäsittely**
  - Koodi lukee ja järjestää tärinädataa automaattisesti, mikä varmistaa, että kaikki tiedot ovat ajantasaisia ja kronologisesti oikeassa järjestyksessä.
  - Data jaetaan 1 minuutin väleihin, mikä auttaa jäsentämään raakadataa systemaattisesti ja vähentää käsittelyvirheiden mahdollisuutta.
- **Ominaisuuksien generointi**
  - Ominaisuuksien tärkeysanalyysi antaa arvokasta tietoa siitä, mitkä seikat vaikuttavat eniten mallin päätöksiin.
  - Erilaiset ominaisuudet (kuten viiveominaisuudet, liukuvat tilastot, muutoksen nopeus) muodostavat tietoperustan mallin oppimiselle.
- **Selkeä arviointi ja jatkuva kehitys**
  - Vaikka malli tunnistaa testijaksolla 16 potentiaalista tapahtumaa ja antaa 5 tarkkaa osumaa, se toimii lähtökohtana, jota voidaan jatkokehittää ja säätää entistä tarkemmaksi.
- **Skaalautuvuus ja tuotantokelpoinen toteutus**
  - Ratkaisu soveltuu käytettäväksi tuotantoympäristössä, missä dataa kertyy jatkuvasti.
  - Automatisoitu tietojen käsittely ja ennustaminen mahdollistavat reaaliaikaisen seurannan, mikä on erittäin tärkeää koneiden ennakoivassa huollossa.
- **Ennakoiva huolto ja poikkeavuuksien tunnistus**
  - Malli (RandomForestClassifier) on koulutettu tunnistamaan poikkeavia tapahtumia, mikä mahdollistaa ennakoivan huollon.
  - Ennakointi auttaa tunnistamaan, milloin koneessa voi olla huoltotarvetta, jolloin voidaan aikatauluttaa tarvittavat toimenpiteet ennen vikoja.

Yhteenveto: Haasteet ja huomioitavaa

- **Datankeruun oikeellisuus ja laatu**
  - **Ihmisten tekemä data:** Ratkaisun toimivuus perustuu kerättyyn tärinädataan ja tapahtumamerkintöihin. Jos dataa merkitään manuaalisesti, ihmisten tekemät virheet voivat johtaa epätarkkoihin tai puutteellisiin tietomerkintöihin. Tämä vaikuttaa suoraan mallin kykyyn oppia.
  - **Tarkka ajoitus:** Terien vaihdon ajankohta on kriittinen, koska mallin oppiminen ja ennustusten tarkkuus riippuvat siitä, että tapahtumat kirjataan tarkasti. Pienetkin viiveet tai virheet aikaleimoissa voivat johtaa virheellisiin tuloksiin.
- **Mallin oppimisen varmistaminen**
  - **Ominaisuuksien valinta:** Mallin suorituskykyyn vaikuttavat käyttäytymisen kuvaamiseen luodut ominaisuudet (lag-ominaisuudet, liukuvat tilastot, muutoksen nopeus). On tärkeää arvioida, että ominaisuudet todella kuvaavat relevantteja signaaleja, eikä malli opi virheellisiä riippuvuuksia.
- **Live-datan hyödyntäminen tulevaisuudessa**
  - **Nykyinen tilanne:** Tässä vaiheessa koodi toimii ennalta kerätyn datan (CSV-tiedosto) kanssa, joten se ei ole reaaliaikainen ratkaisu.
  - **Mahdollinen jatkokehitys:** Ratkaisua voidaan jatkokehittää siten, että se hyödyntää live-dataa esimerkiksi reaaliaikaisessa tilannekuvassa. Tämä edellyttää jatkokehitystä, joka varmistavat jatkuvan ja oikea-aikaisen datavirran sekä sen yhteensopivuuden olemassa olevan analytiikkaputken kanssa.
- **Yleiset haasteet tuotantoympäristössä**
  - **Mallin suorituskyky:** Nykyiset arviointitulokset (Recall 0.36, Precision 0.31, F1 Score 0.33) osoittavat, että mallilla on parannettavaa. Jatkossa on tärkeää kehittää malli niin, että se tunnistaa mahdolliset muutostilanteet tehokkaammin ja vähentää virhepositiivisia ennusteita.
  - **Jatkuva monitorointi ja validointi:** Mallin koulutuksen jälkeen on tärkeää monitoroida sen suorituskykyä jatkuvasti, jotta varmistetaan, että se oppii ajantasaisesti oikeanlaisia kuvioita ja reagoi muutoksiin tuotantoympäristössä.

Ajatukset

- Tämä ratkaisu voi mahdollistaa ennakoivan huollon tuottavassa teollisuudessa hyödyntämällä koneoppimista ja älykästä datankäsittelyä. Ratkaisu analysoi tärinädataa, jakaa sen aikaväleihin ja laskee tilastollisia mittareita, joiden perusteella ennustetaan potentiaalisia teränvaihtotilanteita. Tärkeää on, että kerätty data on tarpeeksi laadukasta ja tapahtumien ajoitus kirjataan tarkasti, sillä pienet virheet tai pyöristykset voivat vaikuttaa mallin oppimiseen.
- Toteutus käyttää ennalta kerättyä dataa, mutta se on helposti laajennettavissa reaaliaikaiseen (live) dataan, mikä mahdollistaisi nopeamman ja tarkemman reagoinnin tuotantolinjan muutoksiin. Näin ennakoiva huolto voisi merkittävästi vähentää odottamattomia seisokkeja ja huoltokustannuksia, mikä parantaa teollisuuden tuottavuutta ja kilpailukykyä.
- Matalakustannuksinen IoT-sensorointi on toteutettavissa teollisessa ympäristössä avoimen teknologian avulla. Pilotin aineisto toimii hyvänä alkuna jatkokehittämiselle.

Jatkokehitettäväksi

- Enemmän dataa mallin koulutukseen
  - Metadata mahd. tarkkaan kirjattua
  - Muu data, kuten tuotantonopeus, ym.
- Reaaliaikainen datansyöttö mahdollistaa reaaliaikaisen hälytyksen
- Eri mittauspisteiden testaus, lisäantureiden käyttö

Lisätietoa

- Linkit Python -koodiin ja esimerkkidataan löytyvät vAI:lla tuottavuutta? –hankkeen verkkosivuilta
  - https://projektit.seamk.fi/alykkaat-teknologiat/vailla-tuottavuutta/
- Lisätietoa datankeruuseen liittyen on saatavilla MOI-Hub –hankkeen kautta



Vaasan yliopisto
UNIVERSITY OF VAASA

SEAMK
Seinäjoen ammattikorkeakoulu
Seinäjoki University of Applied Sciences

Tampereen yliopisto

Euroopan unionin
osarahoittama

ETELÄ-POHJANMAAN LIITTO
Regional Council of South Ostrobothnia