

3D Data Economy Satakunta's Success Factor in Digital Green Growth
***Datan hyödyntämisen konsepti teollisuus- ja
automaatioympäristössä***

Jari Turunen
Tampereen yliopisto

Teollisuusdata

- Monessa teollisuus- ja automaatioympäristössä kerätään monenlaista dataa kunnossapidon tarpeisiin

- Miksei sitä voitaisi hyödyntää myös tuotannon kehittämiseen?



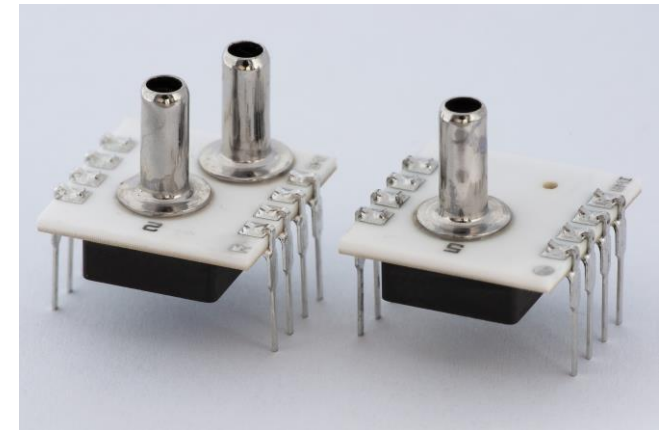
Flickr.com creative commons

Datan hyödyntäminen teollisuudessa

- Teollisuudesta saatavaa dataa voidaan käyttää:
 - Kapasiteetin lisäämiseen / pullonkaulojen löytämiseen teollisuusprosessista
 - Laadun parantamiseen (esim. kappaleen maalin paksuuden vaihteluvälin kaventaminen)
 - Lisäksi automatisoimalla tiedonkeruuta voidaan lisätä tuottavuutta (esim. manuaalisen kirjauksien poisto työvaiheesta, enemmän kirjauksia-> enemmän analysoitavaa)

Teollisuudesta saatava data

- Automaattista tai anturitietoa
 - asiakastietojärjestelmät, tilaukset yms.
 - erilaiset liikeradat (x,y,z) robotiikkakoordinaatistoissa,
 - paine, nopeus, lämpötila, aika...
 - maalikerroksen paksuus, paino...
- Manuaalisesti tallennettavaa tietoa
 - Hankalaa sensoroida (ei esimerkiksi koneluettavaa EAN yms. koodia)
 - Epäsäännöllisesti, harvoin tapahtuva mittaus (muistettava kirjata!)



[Wikimedia Commons](#)

Ulkoiset datalähteet integroitavissa

- Säätila, lämpö, ilman kosteus, etäluettavissa ilmatieteen laitokselta
 - Esim. ulkomaalaus, kosteusvaihteluille herkät työt
- Muut ulkoiset datalähteet
 - Kartat, satelliittidata
 - Avaruussää, aurinkomyrskyt
 - Geomagnetismi...
- Löytyy paljon ilmaisia koneellisen rajapinnan yli luettavia tietoja, sekä maksullisia

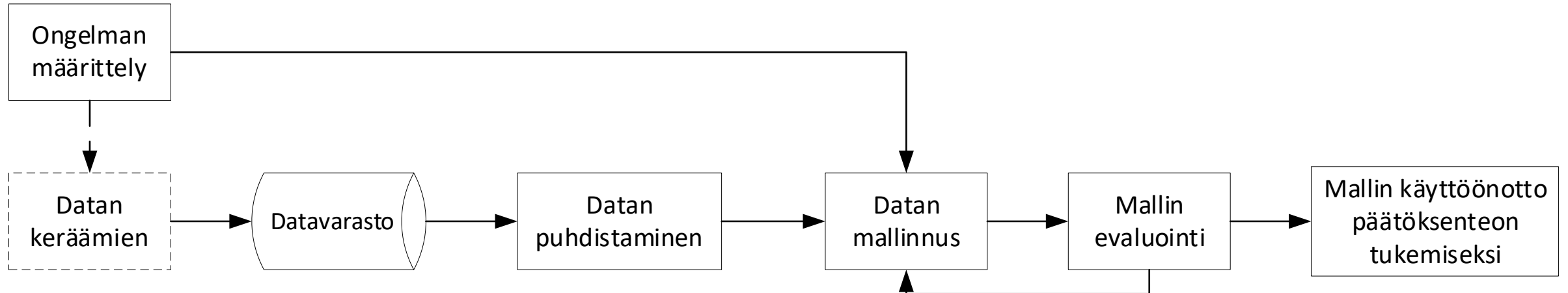
Mitä datalla voidaan tehdä?

- Datalla voidaan tehdä kaikenlaista, mutta ensiksi täytyy löytää peruskysymy(s/kset), mihin halutaan vastaus.
 - Kysymys voi olla: toimiiko tuotantolinja optimaalisesti?
 - Mikä osa linjassa vie eniten aikaa ja miksi?
- Jos linjasto on yksinkertainen ja syyt käytännössä tiedetään pitkäaikaisilla havainnoilla, dataa ei ole syytä lähteä penkomaan ja käytännön toimet linjaston parannukseen voidaan aloittaa heti.

Suuret, monimutkaiset järjestelmät

- Mikäli dataa on 10...1000 mitattua kanavaa, joihin tulee koko ajan uutta tietoa kanavien antureista, antureiden välisen relaation tulkitseminen voi olla vaikeaa
 - Peruskysymys voi silti olla sama eli esimerkiksi voidaanko linjaa nopeuttaa tai poistaa pullonkauloja?
 - Tässä vaiheessa analytiikkakonsepti voidaan ottaa käyttöön

Data-analytiikkakonseptin prosessikuvaus

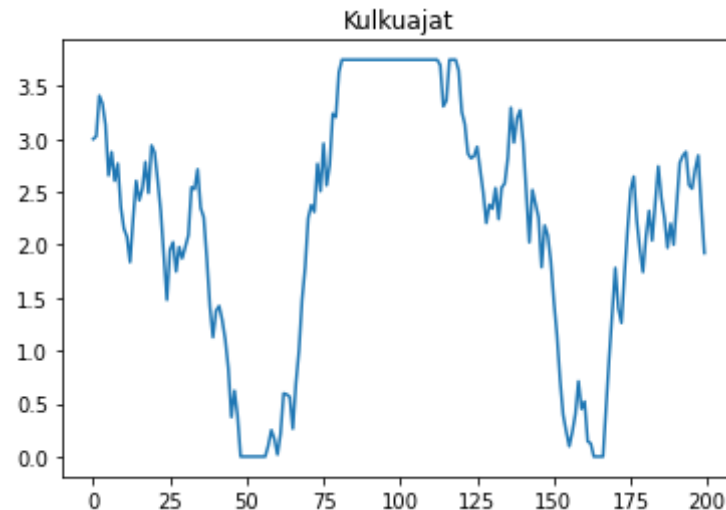


Ongelman määrittely

- Konkreettinen kysymys / sarja kysymyksiä, mihin datasta toivotaan löytyvän vastaus
 - Realistinen oltava suhteessa datan määrään ja laatuun

Datan puhdistaminen

- Dataa on jo olemassa, se on mahdollisesti puhdistettava ja tarkastettava visuaalisesti
 - Anturien vikaantuminen ja uudelleen säätäminen
 - Rajojen yli menevät arvot, NaN = not a number jne.



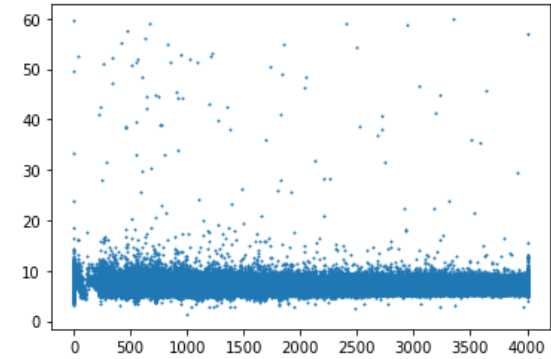
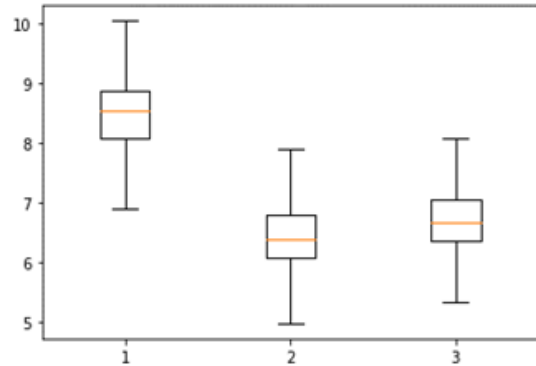
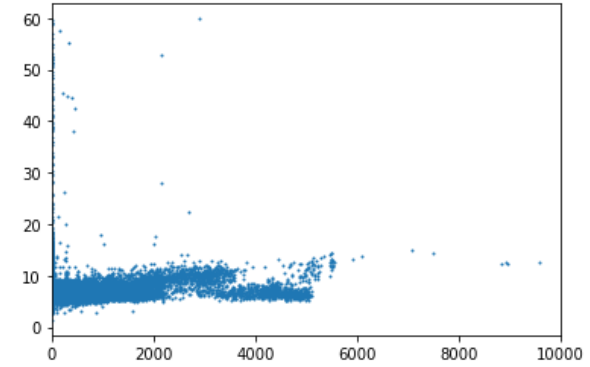
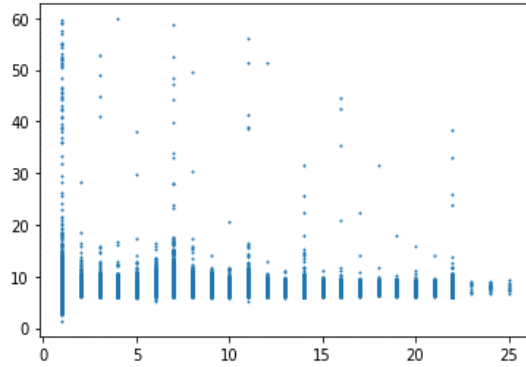
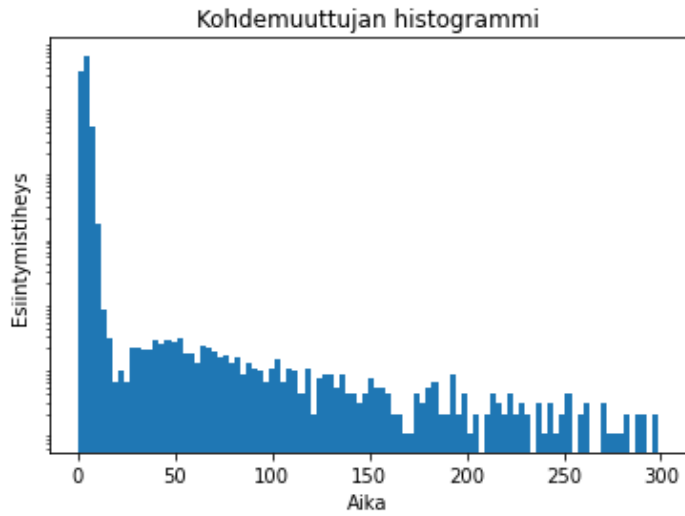
Esimerkki saturoitumisesta

Päätettävä yhdessä yrityksen kanssa mitä virhearvoille tehdään

Datan avulla tapahtuva mallinnus

- Malli on valittava niin, että se sopii tarkoitukseen ja palvelee peruskysymystä
 - 'Digital Twin', mallinnetaan koko järjestelmä, osa järjestelmästä
- Mallinnetaan kulkuajat, pieni osa järjestelmästä, kanavien välistä riippuvuutta jne.

Datatyyppejä



Malli

- Mallivalikoima on laaja, ja niitä on valittavissa erilaisille datatyypeille:
 - Lineaariset mallit
 - regressiomallit, jne
 - Epälineaariset mallit
 - neuroverkot / tekoäly
 - Päättöspuut, jne
- Lisää esimerkkejä Prizztechin raportissa

Mallin evaluointi

- Mallin sopivuus ja evaluointi yrityksen tarpeisiin on aina tehtävä yrityksen kanssa
 - Mikäli malli ei toimi kunnolla sitä on muutettava tai mallinnuksen perusajatusta muutettava

Case Cimcorp



Kuva 1. Tuoretavaravarasto © First Production Marius Tikkanen, Cimcorp Media Hub

- Cimcorp Oy valmistaa varastonhallinta-automatiikkaa asiakkailleen. Varastonhallinta koostuu kolmiulotteisessa koordinaatistossa toimivista roboteista, kuljettimista sekä ohjausjärjestelmistä.

Case Cimcorp tavoitteet ja tulokset

- Pilotin tavoitteena oli selvittää, miten teollisuusympäristöstä kerätyn data pohjalta voidaan parhaiten selvittää kriittisiin kohtemuuttujiin (esim. suoritusaikaa tai laatua kuvaavat muuttujat) eniten vaikuttavat tekijät.
- Pilotin alkuvaiheessa oli tarkoitus käydä yhdessä pilottikohteen edustajan kanssa läpi kerätty data ja määritellä siitä ennakkotiedon pohjalta avainmuuttujat.
- Tarkoitus oli myös määritellä poikkeamat tai kohtemuuttujien ei-toivotut tilat, joita datan pohjalta lähdetään selvittämään.
- Valittuja muuttujia oli tarkoitus tarkastella visualisointien avulla ja valita data-analytiikkaan tarkoituksenmukaiset työkalut.

Tulokset

- Yritykselle esiteltiin erilaisia mallinnusmenetelmiä (Python koodit) sekä tulosten analysointi- ja visualisointimenetelmiä (esimerkiksi Shapley indeksit), joita voidaan hyödyntää datakanavien analysoinneissa ja vertailuissa

Pilotin opetukset

- Pilottiprojektissa oli huomionarvoista ongelman **selkeä määrittely**. Selkeä määrittely nopeuttaa liikkeellelähtöä merkittävästi, ja vaikka datasettiä ja muuttujakanavia vaihdetaan, lopputavoite on silti sama koko projektin ajan.
- Toinen tärkeä seikka pilotin onnistumisen näkökulmasta on **tiivis kommunikointi** projektin osapuolien kesken. Pilottikohteessa yhteydenpito tapahtui viikkopalaverien avulla, johon kaikki osapuolet olivat kutsuttuna mukaan.
- Kehitystyössä käytettiin **kohdeyritykselle tuttua työkalua**, joka oli Python-koodausympäristö. Tämä lisäsi osaltaan kehitystyön nopeutta ja joustavuutta, sillä kohdeyrityksestä pystyttiin nopeasti kommentoimaan koodissa olevia puutteita ja antamaan kehitysehdotuksia.

Tampereen yliopisto, Pori

- Data-Analytiikan ja Optimoinnin (DAO) ryhmä



- Professori, TkT Tarmo Lipping,
tarmo.lipping@tuni.fi



- Yliopistonlehtori, TkT Jari Turunen