

OPTIMERING AF PRODUKTION

Mere værdi i produktionen gennem dataanalyse:
Et Industri 4.0-samarbejde for at skabe mere værdi
og forudse driftsstop

Baggrunden for denne artikel er arbejdet med produktionsvirksomheder, og hvordan en lille eller mellemstor virksomhed kan opsamle data fra egen produktion og analysere dette data med det formål at optimere den fremtidige produktion. Dette har et stort potentiale for små og mellemstore virksomheder i Danmark.

Denne artikel beskriver et projekt, hvor der i samarbejde med en lokal produktionsvirksomhed er opsamlet data på en effektiv og struktureret måde samt benyttet maskinlæring til at analysere data for at forudsige driftsstop. Uventede driftsstop kan være omkostningstunge og forsinke den planlagte produktion.

Metoden til at arbejde med dette benytter aktionsforskning og elementer fra Design Thinking. Der udvikles løsninger i fællesskab, som prøves af i praksis, og der læres af disse løsninger. Den udviklede softwareplatform er i konstant forandring, krav og ønsker fra virksomheden generaliseres og benyttes som en generel løsningsmodel for andre SMV'er i fremstillingssektoren.

I samarbejde med en lokal SMV er der udviklet en prototype på et system med en fleksibel, robust arkitektur. Systemet kan lagre store datamængder (big data) og fremvise analyseresultater near real-time.

Endvidere beskrives en arbejdsproces, som er tilpasset SMV'ens kompetenceniveau og organisation. Processen skal støtte SMV'er, som vil i gang med at udnytte deres produktionsdata.

FORFATTERE

Brian Hvarregaard, adjunkt,
IT-uddannelserne, UCN

Finn E. Nordbjerg, lektor,
IT-uddannelserne, UCN

Nadeem Iftikhar, lektor,
IT-uddannelserne, UCN

INTRODUKTION

Små og mellemstore produktionsvirksomheder står foran store udfordringer i forhold til at optimere deres produktion ud fra de data, som skabes under produktionen.

Mindre end 30 % af danske SMV'er analyserer disse data, og under 10 % anvender kunstig intelligens/maskinlæring til denne analyse (Hofman-Bang, direktør Industriens

Fond, 2019). Disse data har stort potentiale, da de kan bruges til at optimere fremtidig produktion, herunder til at undgå dyre driftsstop af produktionsmaskiner (Hofman-Bang, direktør Industriens Fond, n.d.).

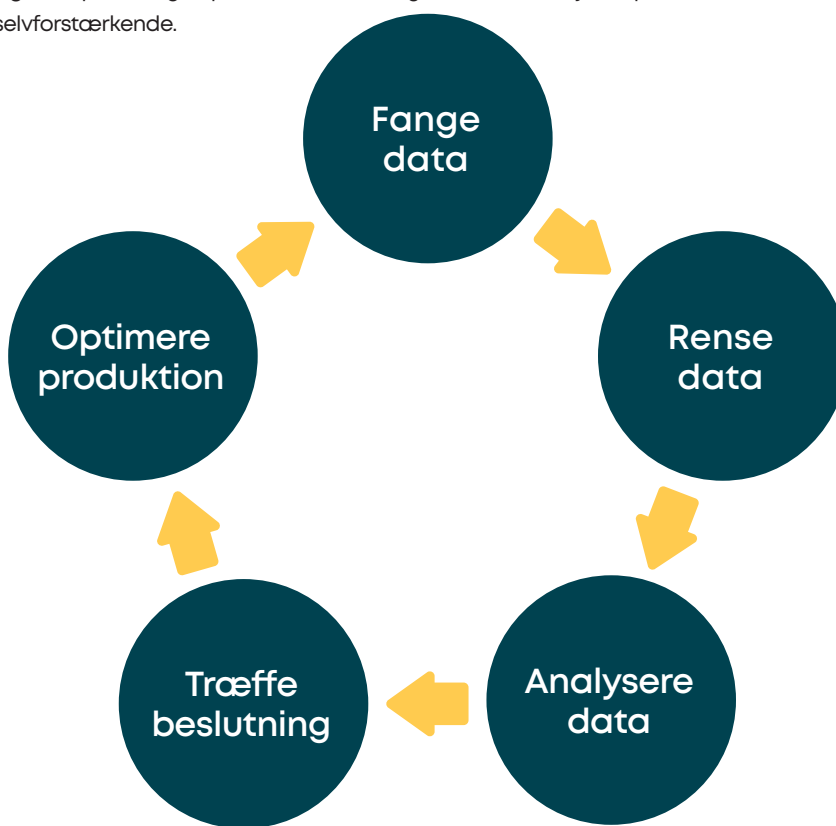
Grunden til dette manglende fokus er ofte, at man mangler en teknisk platform, som kan opsamle, behandle og præsentere data. De

fleste SMV'er har ikke ressourcer og teknisk indsigt til at starte sådan et projekt på egen hånd (Hofman-Bang, direktør Industriens Fond, 2019).

I en årrække har UCN Teknologi & Business samarbejdet med nordjyske produktionsvirksomheder om opsamling og analyse af produktionsdata, dvs. data genereret direkte i produktionsprocessen af diverse sensorer (Iftikhar et al., 2019a og Iftikhar et al., 2020). Der er udviklet en prototype på software, som er i stand til at udføre deskriptiv og prædiktiv analyse af sensordata kombineret med ERP-data. Analyseresultaterne kan vises near real-time på et dashboard, hvilket er ret unikt. Der anvendes eksisterende open-source-platforme, big data-metoder/-teknologier og maskinlæring. Resultaterne omfatter dels ovennævnte prototype på software og dels erfaringerne fra arbejdet med virksomhedens transition mod øget udnyttelse af produktionsdata. Det igangværende arbejde omfatter generalisering af erfaringerne fra samarbejdet til en model for transitionsprocessen, som kan anvendes af andre SMV'er. En del af modellen vil være anvisninger på identifikation af virksomhedens datapotentiale og dens digitale udviklingstrin. SMV'er i fremstillingssektoren befinder sig på forskellige teknologiske/digitale modenhedsniveauer, så en del af projektet vil være at tilpasse modellen til forskellige modenhedsniveauer. Den udviklede software skal generaliseres fra en konkret løsning til et software-framework, som kan implementeres af andre SMV'er. I denne forbindelse skal et software-framework forstås som en samling programkomponenter, som forholdsvis simpelt kan konfigureres, tilpasses og sammensættes til open source-løsninger for forskellige konkrete virksomheder.

Dette projekt er et aktivt forskningsprojekt, som søger at udvikle en generel softwareplatform, som

Figur 1: Optimering af produktion ud fra egne data er en cyklisk proces, som er selvforstærkende.



tilgodeser små og mellemstore produktionsvirksomheder til at forbedre deres konkurrenceevne i markedet gennem en optimering af deres produktion. Denne optimering skabes gennem opsamling og behandling af data, således at disse data giver mening for den operationelle produktion og planlægning. Der arbejdes med løbende at opsamle data fra produktionen, at rense og analysere disse data samt at træffe beslutninger ud fra dataene for at forbedre selvsamme produktion.

Denne proces med at forbedre produktionen ved at analysere egne produktionsdata er iterativ og selvforstærkende, på den måde at mere effektiv produktion giver bedre data, som igen analyseres og bruges til at forbedre produktionen. Denne proces gentages og optimeres.

Gennem projektet er der observeret flere problemstillinger, herunder blandt andet omkring det

organisatoriske og den transitionsproces, som SMV'er gennemgår for at modnes til at udnytte data til optimering af produktionen.

OPSAMLING AF DATA TIL FORBEDRING AF PRODUKTIONEN

Problemerne, som virksomhederne oplever, er todelt i form af et teknisk problem omkring opsamling og klargøring af data samt et organisatorisk problem. Det organisatoriske problem er observeret gennem forskningsprojektet og samarbejdet med lokale virksomheder. Problemet diskuteres nærmere i afsnittet om "Virksomhedsejerens vej til at udnytte egne data".

Hovedformålet med dataopsamlingen og den efterfølgende behandling af data er at kunne beskrive løbende, hvordan produktionen går sammenlignet med historiske data, samt at forudse driftsstop hos en produktionsvirksomhed for herigennem at optimere driften og produktionen.

Følgeligt består dette projekt af to delproblemer:

1. Hvordan opbygges den tekniske platform til at kunne opsamle store datamængder, rense disse data, analysere dem og gøre denne data tilgængelig for det operationelle personale inden for rimelig tid (near real-time), således der kan handles på dataanalysen?
2. Hvordan skaber man resultater for en organisation, som i forvejen ikke udnytter egne data til optimering af produktionen, og således hjælper denne organisation til at flytte sig fra passiv rapportering og information til optimering af produktionen gennem prædiktiv analyse (predictive analytics), jf. Gartner (Hofman-Bang, direktør Industriens Fond, 2019, p. 21)?

Det er ikke alle virksomheder, som har adgang til professionelt overvågnings- og automatiseringsudstyr og software. Store virksomheder har dette integreret, men for en gennemsnitlig dansk SMV er dette ikke økonomisk tilgængeligt. Det søger vi at ændre, således at SMV'er får mulighed for at kunne berige deres produktion med viden fra selvsamme og dermed kunne træffe flere og bedre beslutninger. Den udviklede softwareplatform forventes at kunne stilles frit til rådighed for interesserede SMV'er.

Når en produktionslinje stopper, kan det ofte tage væsentligt længere tid at standse produktionen helt, udbedre fejlen og få produktionen op at køre igen, end det vil tage at forebygge fejlen ved et mindre, kontrolleret stop. Der er således en sammenhæng mellem, hvor tidligt en fejl findes og udbedres, og den påvirkning, det har på omkostningen ved et utilsigtet produktionsstop.

METODE

Forskningsprojektet er primært baseret på et samarbejde med en lokal produktionsvirksomhed, hvor der gennem længere tid er udviklet prototyper af softwaresystemet til

optimering af produktionen for at kunne tilgodese og håndtere de store, varierende datamængder, som produktionssystemet skaber. Samtidig med at der læres fra problemområdet og følgende rettes til i forhold til de benyttede teorier og praksis omkring implementeringen. Dette er gjort gennem aktionsforskning (Design Science Research (DSR)), hvor der veksles mellem viden og teorier omkring problemområdet og den faktiske praksis, hvor dette resulterer i en softwareprototype, som afprøves, og der justeres efterfølgende i såvel praksis og den viden, dette genererer, således teorien og praksis følges ad og supplerer hinanden (Nordbjerg, 2021) og (Peffer et al., 2007).

For at adressere det første problemområde, som er beskrevet tidligere omkring de tekniske udfordringer, er det hensigten at udvikle og forbedre den tekniske softwareplatform, således denne softwareplatform er i stand til at opsamle, rense, analysere og præsentere data til virksomheden. Dette gøres gennem deskriptive algoritmer til at forudsige anomalier i produktionen (descriptive analysis anomaly detection) samt forudsigende algoritmer (predictive algorithms). Disse algoritmer er i konstant udvikling med formålet at blive bedre til at løse de problemer, som virksomhederne står overfor, jvf. DSR.

For at løse det andet problemområde omkring modning af virksomhederne, som ikke benytter egne data, udvikles der i samarbejde med den lokale virksomhed en softwareplatform gennem flere iterationer med det formål at få dybere forståelse af problemområdet og gennem denne forståelse forbedre softwareplatformen og modtage feedback fra den lokale virksomhed, omkring hvorvidt den nuværende iteration løser foreningsmængden af problemerne samt i hvor høj grad.

Måden, dette gøres på, er ved at mødes med produktionschefen hos en lokal virksomhed omkring dennes oplevelser og behov for herefter at

beskrive det specifikke behov hos virksomheden. Dette behov generaliseres så til funktionalitet i den generelle platform for herefter at blive verificeret hos den specifikke virksomhed. Der søges således at lave en løsning på foreningsmængden af problemer, som observeres hos SMV'er og ikke kun på fællesmængden af problemerne. I alle møderne med virksomheder er alle regler for dataindsamling og opbevaring overholdt.

Det er denne vekselvirkning af afdækning af krav hos virksomheden, der generaliseres for så at blive verificeret (som generel funktionalitet) hos den specifikke virksomhed, der verificerer, at den enkelte funktionalitet kan generaliseres til andre virksomheder også.

For at kunne implementere dette i den praktiske softwareplatform benyttes der kendte integrationsmønstre (Hohpe, 2003), dette således der sikres en fleksibilitet for at kunne tilvælge eller fravælge bestemt funktionalitet hos den specifikke virksomhed, i takt med at der læres mere og mere. Dette sikrer en omstillingsparathed i softwareplatformen, som understøtter den omstillingsparathed, som også findes i den teoretiske del af dette arbejde (DSR).

VIRKSOMHEDSEJERENS VEJ TIL AT UDNYTTE EGNE DATA

Noget af det, som er særligt interessant at kigge på ved SMV'er, er den viden og erfaring, der er til brugen af data i produktionen. Noget af det, som gør sig gældende ved mange små og mellemstore virksomheder, som gerne vil videre med at optimere deres produktion, er, at mange af disse virksomheder ved, at en af vejene til at forbedre deres produktion er gennem analyse af deres egne produktionsdata.

Der er lang vej fra at have en viden om mulighederne for brug af egne data til at forbedre produktionen til faktisk at nå til et sted, hvor man som virksomhed har kompetencerne og

den tekniske indsigt (handlemuligheder) til at påbegynde arbejdet med optimering af produktionen ud fra egne data.

Den typiske ejer af en SMV-produktionsvirksomhed har mange gange en teknisk uddannelse som fx smed eller lignende og er så med tiden blevet mere ansvarlig for produktion og drift. Dette betyder også, at man som leder måske mangler den indsigt i, hvad dataanalyse og maskinlæring oven på egne produktionsdata kan betyde for virksomheden.

Mange gange rækker ledelsen af virksomheden ud til deres erhvervsnetværk, men her kommer man også til kort. Det er svært at gå fra at vide, hvilken retning man skal bevæge sig i, til rent faktisk at tage første skridt.

Når ledelsen for virksomheden når til en erkendelse af, at der skal data i spil for at kunne få værdi ud af det, så handler det om et tidsproblem – det vil sige at jo længere tid man opsamler data, jo mere nøjagtigt kan man forudsige eventuelle driftsstop i fremtiden. Problemet er således, at man mangler data fra tidligere, og at det først er fra starten af opsamlingen, at man kan begynde at forbedre virksomhedens evne til at forudsige driftsstop.

Ledelsen af disse SMV'er står til regnskab for beslutningen om at bruge resurser på at analysere data. Det er vigtigt for projektets succes, at ledelsen forpligter sig til projektet og giver fuld støtte til, at det kan gennemføres. Det er vigtigt for projektet, at der skabes hurtige, synlige og værdiskabende resultater, for at bevare ledelsens opbakning (Iftikhar and Nordbjerg, 2021).

Det vil sige at der mange gange er langt fra at begynde at opsamle produktionsdata, til man faktisk ser en forbedring. Dette skyldes, at jo mere og bedre data der er til rådighed, jo mere præcist kan analysen og brugen af data forbedre fremtidig produktion. Tid er en væsentlig faktor i denne type af softwaresystemer, da der opsamles og akkumuleres mere (og med tiden bedre) data til brug for

de underliggende dataanalyser. Dette betyder, at værdien af egne data forstærkes over tid, og på baggrund af en optimeret produktion bliver data også bedre, hvilket igen forstærker fremtidig produktion.

Dette giver anledning til at stille nogle spændende spørgsmål omkring, hvorledes man kan kvalificere og bruge den data, man opsamler. Herunder nogle af disse spørgsmål:

- Hvor hurtigt kan man komme i gang med at give feedback til brugeren?
- Kan man bruge data opsamlet hos en anden virksomhed til at forudsige driftsstop i denne specifikke virksomhed?

Opsamlingen af data giver et billede af produktionens effektivitet. Dette billede er dog kun så nuanceret som kvaliteten (og mængden) af den data, som opsamles.

Denne betragtning kan godt lede til den konklusion, at så længe man har data, så kan man skabe værdi. Udfordringen er, om det data, man opsamler, har en forbindelse til de effektiviseringsmål, man har som virksomhed. Der er på flere møder med den lokale virksomhed diskuteret et generelt spørgsmål, som virker som en form for anti-regel, det vil sige at dette er et faresignal for både den lokale virksomhed og for forskergruppen om at holde fokus på at skabe værdi og så opsamle relevante data og ikke blot opsamle data uden nogen kobling til den værdi, man vil skabe gennem produktionsoptimeringen. Dette spørgsmål lyder således: *"Nu har jeg opsamlet data, hvad kan vi så få af værdi?"*

Både forskergruppen og den lokale virksomhed er enige om, at dette spørgsmål vender på hovedet, i forhold til hvordan man bør betragte det at skabe værdi ud fra data. Denne betragtning antyder nemlig, at alt data er lige. Dette er ikke tilfældet, for data i sig selv har minimal værdi, det er mere et spørgsmål om at få data beriget

med kontekst og viden omkring forretningsværdien og sammenhængen til de mål for produktionsoptimering, man ønsker.

Denne tilgang til at få værdi til beslutningsstøtte uanset dataenes beskaffenhed er at vende beslutningsstøtten på hovedet.

Hele tanken omkring, hvordan man kan forbedre sin virksomhed og derefter produktionen, bør bunde i et forretningsmæssigt synspunkt, og således bør spørgsmålet hellere være: *"Jeg har brug for at træffe en bestemt type forretningsmæssige beslutninger, hvilket data skal der opsamles fra vores produktion, og hvilke analyser af disse data kan understøtte beslutningsprocessen?"*

I dette tilfælde opsamler man relevant data for beslutningerne og ikke blot data for datas skyld.

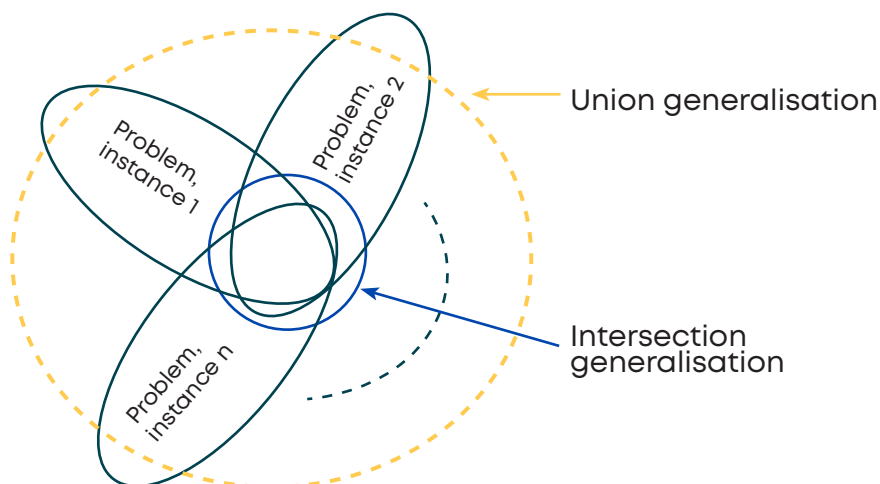
Denne holdning til data er ikke ny og kommer ej heller bag på forskningsgruppen. Dette er en meget normal tilgang hos SMV'er, som er på et vist modningsniveau i forhold til dataanalyse. Disse modningsniveauer understøttes af Industriens Fond (Hofman-Bang, direktør Industriens Fond, 2019).

Det er ikke, fordi den enkelte SMV ikke har behov for svar og beslutningsstøtte. Disse virksomheder har spørgsmål, som alle andre produktionsvirksomheder har. Disse spørgsmål omhandler mange gange de samme styrings- og beslutningsstøtteredskaber. Det er velkendt fra arbejdet med Business Intelligens (Howson, 2014), at problemet for disse virksomheder er, at de ofte ikke har en platform, hvor de kan stille disse spørgsmål. Dette projekt søger at skabe denne platform, således man som virksomhed har en entydig kilde til beslutningsstøtte omkring optimering af produktionen.

ANALYSE

Der er ikke to virksomheder, som er ens. Selv om man vil kunne finde mange fællestræk mellem en stor

Figur 2: Generalisering af virksomhedsspecifikke krav til fælles funktionalitet.
(Udarbejdet af Finn Nordbjerg.)



del af produktionsvirksomhederne, er det langtfra muligt at lave noget, som er specifikt for en enkelt virksomhed, og som også passer til andre virksomheder.

Måden, vi tilgår dette på, er ved at starte med at generalisere, det vil sige at vi tager udgangspunkt i den enkelte virksomhed, som er en del af dette projekt, og ser på, hvad de har af problemer, hvordan deres produktion er sammensat, og hvordan der skabes data. Herefter skal dette generaliseres til en generel struktur/arkitektur for herefter at specialiseres ned til den enkelte virksomhed igen. Så har vi en model, som virker.

For at understøtte den tekniske del af problemet skal der bygges en robust softwarearkitektur til at håndtere virksomhedernes forskelligheder (antal maskiner, typer af maskiner osv.). Vores fremgangsmåden er at starte med en specifik virksomhed og dens behov og problemstillinger. Den udviklede løsning generaliseres til en generisk løsningsmodel, som valideres gennem en verificering hos den specifikke virksomhed. Denne fremgangsmåde er i tråd med den overordnede metode, som er beskrevet i DSR. Se ovenstående figur.

Betragter man et system ud fra dette udgangspunkt (generalisering), kan hvilket som helst system reduceres til en mængde sensorer med forskelligt input, og det er så konteksten og informationen omkring den enkelte sensor, som bliver afgørende for selve dataanalysen og den efterfølgende beslutningsstøtte.

Selve arkitekturen af et system, som skal tilgodese mange forskellige datakilder i mange forskellige formater, skal ligeledes gennemarbejdes og valideres. Dette betyder, at der skal kigges ind i at have mange generelle dele af arkitekturen, som genbruges på tværs af alle virksomheder, og få specifikke dele, som kun benyttes af den ene virksomhed.

Risikoen ved dette er, at lykkes dette ikke, er applikationen ikke rentabel i udvikling og vedligehold, da der vil være for mange tilpasninger i denne applikation, hvilket vil være direkte i strid med filosofien om at kunne etablere et generelt beslutningsstøttesystem hos SMV.

Løsningen til denne problemstilling er at anlægge et agnostisk syn på data. For softwareudviklingen og analysen af data er det ikke relevant, hvilket produktionsapparat der er tale om. Det er ej heller

relevant, hvordan data opsamles, eller de tekniske detaljer i dette. Det vigtige er, at man har abstraheret sensor og data væk fra produktionsprocessen.

Dette vil gøre, at den generelle del af systemet kun skal kende til begrebet "sensor" og vide, at en sensor kan producere data over tid. Dette betyder, at vi får en meget generel sammenhæng og kobling i systemet og mulighed for sammenholdelse af data mellem forskellige maskiner (en maskine vil således med denne abstraktion kunne betragtes som en logisk gruppering af sensorer), og dermed vil vi kunne skabe en generel platform til analyse af data.

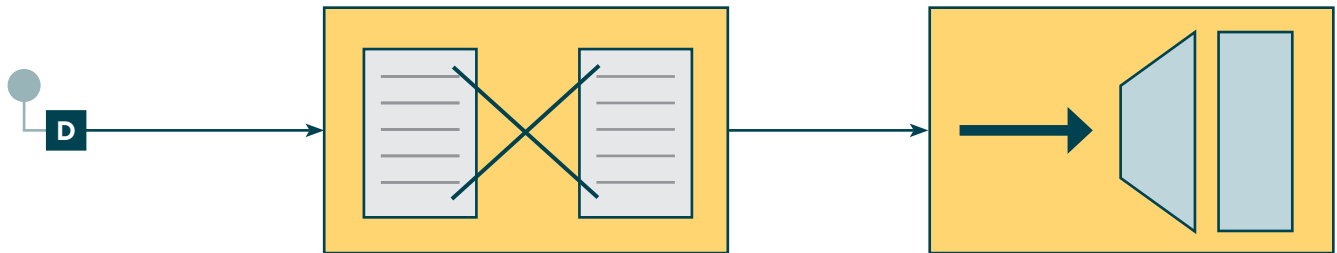
Ved denne generalisering er det vigtigt, at man beriger data med en logik for at skabe kontekst og sammenhæng. Logikken kan fx være navn, placering, led i produktionen eller et bestemt navn på sensoren, som giver mening for virksomheden. Data i sig selv siger ikke ret meget: Det er først, når man får domæneviden kombineret med disse data, at der skabes et fundament for at træffe beslutninger. Denne berigelse må nødvendigvis være en del af "konfigurationen" af den specifikke virksomhedsløsning, domæneviden omkring maskiner, sensorer, værdier og deres betydning, fx alarm og takt.

I et testmiljø er det relativt let at teste med automatisk genereret data, men noget andet er, når det kommer ud i produktion og skal testes i drift. Da virksomheder har mange forskellige typer af sensorer, kan den data, som sendes ind til systemet, være meget forskellig. Dette kan eksempelvis være følgende:

- Heltal
- Kommatal
- Binær sekvens
- Billede (fx termisk billede eller QR-kode)

Disse forskellige typer af data skal man kunne opsamle og benytte i systemet også. Dette stiller høje

Figur 3: Transformation af virksomhedsspecifikt data til generelt input.
(Udarbejdet af Brian Hvarregaard med udgangspunkt i Hohpe, 2003.)



krav til arkitekturen, da den skal kunne håndtere ethvert format.

Løsningen til dette problem er ikke at placere det som en del af den generelle arkitektur, men som en del, der eksisterer uden for arkitekturen – en oversætter. Opgaven for denne oversætter er at oversætte fra det lokale virksomhedsformat (eller sensorformat) til et fast defineret, ensartet format [se figur med sensor og værdien]. Sammenholdt med domæneviden i konfigurationen vil dette stadig kunne bruge den generelle arkitektur til at skabe værdi med.

AT BEARBEJDE STORE MÆNGDER DATA I REALTID

Den tredje del af problemstillingen, der adresseres, er den varians, der er i forskellige produktionsapparaters data, den hastighed, de skabes med, samt den mængde data, som skal analyseres.

En udfordring ved dette system er, at det vokser over tid. Der akkumuleres data over tid, da denne data er med til at forudsige fremtidige driftsstop. Disse historiske data indeholder blandt andet data omkring tidligere driftsstop, og dette betyder, at nye systemer kan trænes ud fra dette data, og med et trænet system vil man hurtigere og mere pålideligt kunne forudsige driftsstop. Man kan populært sige, at ”jo mere man træner, jo bedre bliver man”.

Udfordringen ved at analysere data af så stor varians, volumen og hastighed er tofold. Der skal sikres en umiddelbar deskriptiv analyse i

form af et dashboard eller en rapport, således at det kan vurderes, om den nuværende tilstand er tilfredsstillende. Samtidig skal der bruges avancerede algoritmer og maskinlæring til at lave den prædiktive og præskriptive analyse, således man vil kunne forudsige driftsstop og foreslå afværgeforanstaltninger. Principielt er det ikke vanskeligt. Det, som vanskeliggør det i dette tilfælde, er, at databehandlingen skal gøres med en sådan hastighed, at operatøren advares i tilstrækkelig god tid, inden fejl opstår.

For at dette skal kunne lade sig gøre, er der implementeret en såkaldt lambdaarkitektur, hvor de indkomne data splittes i to hovedspor – et til deskriptiv analyse og et til prædiktiv analyse. Det spor, som beregner den fremtidige risiko for driftsstop, er implementeret via et almindeligt messaging system (Hohpe, 2003), således dette kan skaleres, og de store mængder data kan behandles parallelt på flere servere, om man vil (evt. gennem en cloudløsning). Dette valg til arkitekturen vil naturligt skalere, således at enhver virksomhed, uanset antallet af produktionsmaskiner, kan benytte denne platform. Dette er en forbedring i forhold til mere klassiske arkitekturen, hvor skalering har været akilleshælen.

LAMBDAARKITEKTUREN

Formålet med at implementere en specifik arkitektur til at håndtere disse data med er at gå væk fra en

traditionel, sekventiel processing af data for herefter at give et resultat. De datamængder, der bearbejdes, er alt for store, og processeringen tager alt for lang tid. Ligeledes giver denne arkitektur en mulighed for at være bedre forberedt og mere agil i forhold til fremtidige ændringer til produktionen eller til de algoritmer, som ligger til grund for dette.

Der er to store typer af resultater, der er interessante. Som tidligere nævnt er dette en deskriptiv analyse, som bedst kan sammenlignes med et traditionelt data warehouse, hvor man bruger Business Intelligence eller en form for rapportering til at beskrive statistiske og historiske data (fx gennemsnitlig produktion over tid eller på produkter).

Den anden type resultater, der er interessante, er data, som er bearbejdet af maskinlæring for at skabe en forudsigtelse af driftsstop. Det vil sige en prædiktiv analyse. Denne type af data skal lære af alt den eksisterende data og bruge sine maskinlæringsalgoritmer til at genkende mønstre for, hvornår noget er tæt på at gå galt.

Lambdaarkitekturen (Marz and Warren, 2015) er unik, fordi den tager højde for den individuelle procesring af de forskellige typer af data. Den vigtigste del af denne arkitektur er dens evne til at behandle data parallelt. For hvert datapunkt, som bliver modtaget fra en produktionsmaskine, kan dette datapunkt behandles af både den deskriptive analyse og den præskriptive analyse (og en hvilken som helst

anden type, skulle man ønske at udvide systemet) parallelt (evt. cloud-baseret databehandling).

Man kan sige, at denne arkitektur kopierer hvert datapunkt og fordele disse kopier af data ud til hver sin "swim lane", som fungerer uafhængigt af de andre "swim lanes". En "swim lane" er et isoleret spor i en applikation, som kan fungere uafhængigt af andre "swim lanes". Dette benyttes blandt andet til fejlisolering, således at skulle en "swim lane" bukke under for presset eller fejle på anden vis, så vil det ikke påvirke de andre "swim lanes", og disse "swim lanes" vil så fortsætte med at behandle data.

Det er essentielt for denne type af arkitektur, at hver "swim lane" isoleres, således eventuelle fejl i

databehandlingen ikke påvirker de andre "swim lanes".

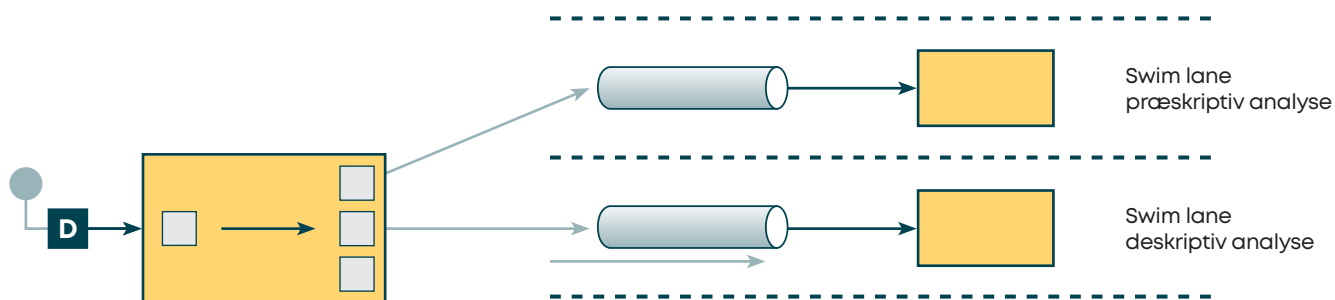
Denne fleksibilitet i arkitekturen er vigtig i dette tilfælde, hvor man kan udvide arkitekturen, i takt med at man lærer mere omkring behandlingen af data. Denne fleksibilitet gør det muligt at udvide arkitekturen enkelt ved blot at tilføje et antal nye "swim lanes" til behandling af data. Denne udvidelse er trivial på baggrund af denne arkitektur.

Noget andet, som er essentielt for denne type af arkitektur, er, at hver enkelt "swim lane" kan skalere uafhængigt af de andre "swim lanes". Datapunkterne bevæger sig hurtigst gennem den "swim lane", hvor data processeres mindst. Dette betyder, at i de tilfælde, hvor store algoritmer skal behandle data, vil

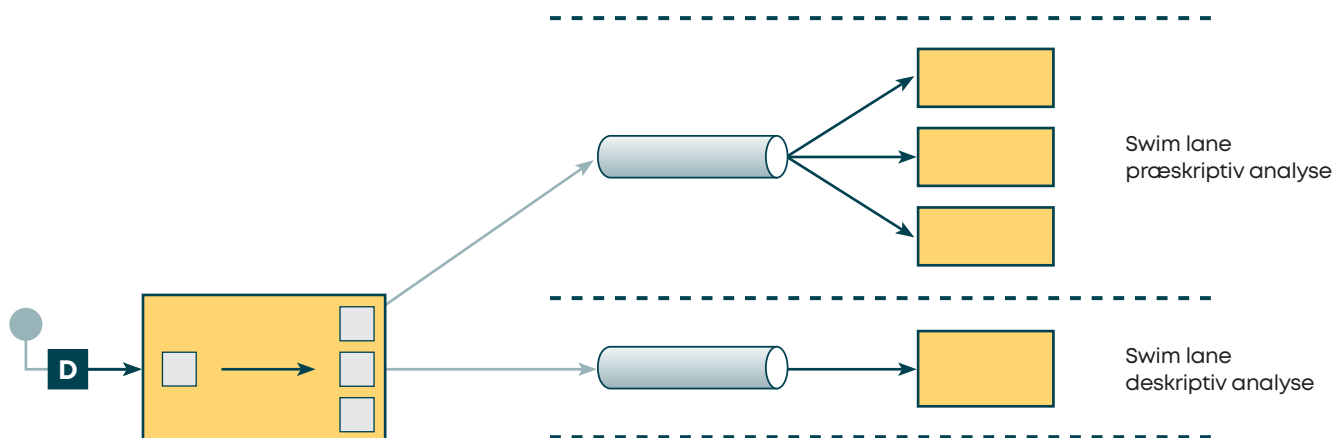
denne behandling tage længere tid. Data fra produktionsapparaterne kan ikke kontrolleres, og hastigheden kan variere såvel op som ned, samtidig med at antallet af produktionsapparater også kan variere. Lambdaarkitekturen skal kunne håndtere såvel en stigning i antallet af produktionsapparater (volumen) samt hastigheden som data leveres med (Velocity).

Denne håndtering af volumen af data og hastigheden af indkomne data gøres individuelt på hver "swim lane", og hver "swim lane" kan automatisk skalere sin behandling op og ned, i takt med at der kommer mere eller mindre data (man antager her, at der er variation over tid, fx mindre data i nogle perioder, fx weekender og ferier).

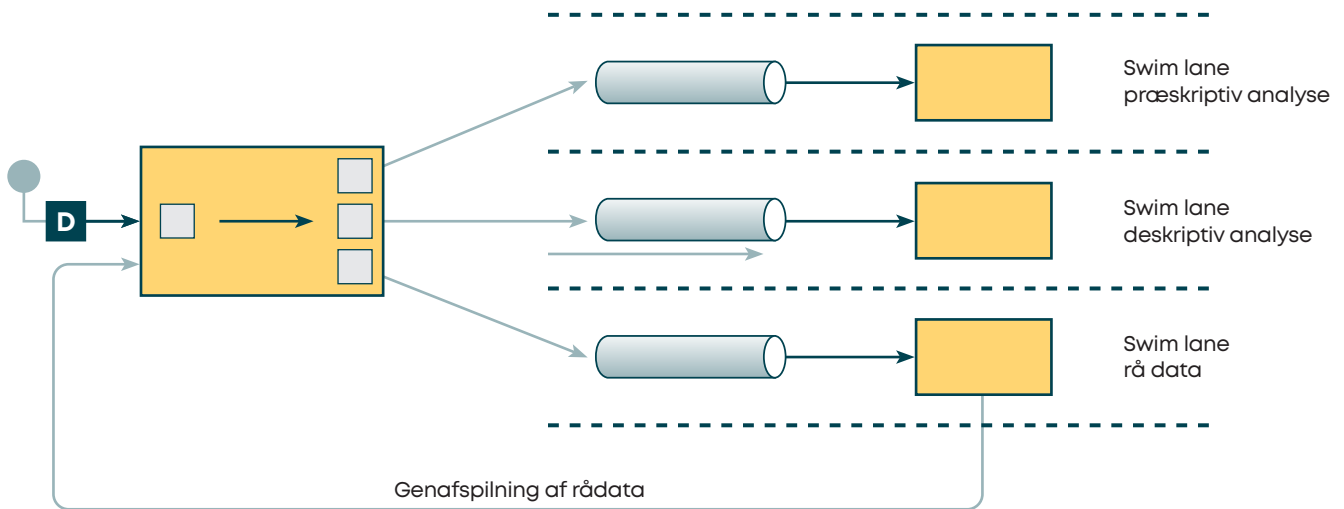
Figur 4: Swim lanes til præsriptiv og deskriptiv analyse. Kopi af datapunkter og isolation af transaktioner inden for Swim lanes. (Udarbejdet af Brian Hvarregaard med udgangspunkt i Hohpe, 2003.)



Figur 5: Uafhængig skalering af den ene swim lane uden at påvirke skaleringen af den anden swim lane. (Udarbejdet af Brian Hvarregaard med udgangspunkt i Hohpe, 2003.)



Figur 5: Uafhængig skalering af den ene swim lane uden at påvirke skaleringen af den anden swim lane. (Udarbejdet af Brian Hvarregaard med udgangspunkt i Hohpe, 2003.)



Med denne fleksibilitet af arkitekturen, hvor det er muligt at tilføje nye "swim lanes" til databehandling, er netop denne arkitektur blevet udviklet med et selvforbedrende element. Denne udvidelse har til formål at sikre data og sikre fremtidig forbedring af såvel algoritmer som arkitektur. I denne specifikke arkitektur er der, for at sikre denne fremtidige fleksibilitet, tilføjet en ny "swim lane" til at gemme rå og ubehandlede data i en database.

Det er essentielt for denne type af systemer, at man behandler data for netop at få deskriptive og præskriptive analyser. Dette har dog den ulempe, at data nu er modificeret af algoritmerne, og det er ikke muligt at gå tilbage til originalformatet, og algoritmerne til de prædiktive og præskriptive analyser fungerer kun på rådata.

Netop den sidste "swim lane" er kendetegnet ved kun at gemme rådata med det eneste formål, at man, i tilfælde af at algoritmerne for de præskriptive analyser opdateres, vil kunne "genafspille" alle data, som er i databasen, i rå format for at træne algoritmerne igen med nyt data. Denne metode sikrer, at lambdaarkitekturen

forbliver relevant for produktionen, og at den kontinuerligt forbedres med nye og forbedrede algoritmer.

DISKUSSION

Der er gennem samarbejde med en lokal produktionsvirksomhed opbygget en robust lambdaarkitektur, som kan skalere og håndtere de store datamængder, således der er et grundlag for at kunne verificere dette hos andre virksomheder, som oplever en lignende problemstilling.

Der er indledningsvist implementeret en sektion af lambdaarkitekturen, som omhandler den deskriptive analyse og en visuel præsentation af denne analyse i form af et dashboard, som kan bruges til beslutningsstøtte. Dette giver maskinoperatøren på gulvet en mulighed for at se og agere på informationer, som beskriver den nuværende tilstand på maskinen (fx alarmer, takt og generel information).

Der er udviklet et simpelt betjeningspanel til en maskinoperatør, således denne kan orientere sig omkring evt. alarmer og standardinformation. Som maskinoperatør er det vigtigt, at der ikke vises irrelevant data, men i tilfælde af behovet

for at maskinoperatøren indblendes, så skal der vises relevant data, i det omfang det er nødvendigt.

Der er lavet flere algoritmer omkring forudsigelsen af driftsstop, og disse vil i høj grad forudsige driftsstop. Dette er for nuværende implementeret som en prototype i den anden sektion af lambdaarkitekturen. Algoritmerne er afprøvet på historiske data og kan ud fra disse forudsige driftsstop med mere end 80 %'s sikkerhed (Iftikhar et al., 2019b).

Ved et samarbejde som dette vil der være en indlejret konflikt mellem virksomheden, som ønsker værdiskabende og handlingsorienterede resultater, og projektgruppen, som indledningsvist i projektet har som fokus at skabe en robust arkitektur, som kan udbredes generelt.

I forbindelse med verificeringen af systemet er dette implementeret, så det er muligt at se data omkring alarmer i systemet samt den aktuelle takt pr. sensor near real-time. Indledningsvist er der fokuseret på alarmer og den umiddelbare gevinst i at kunne gengive en alarm. Dette skyldes, at alarmer i systemet er grundlaget for den maskinlæring,

som skal kunne forudsige fremtidige driftsstop (som udtrykkes som en alarm).

Dette projekt er den første implementering af såvel den tekniske platform som de rapporter og analyser, som en SMV skal bruge for at arbejde med dataanalyse. Der er i dette projekt kun opsamlet empiri fra en enkelt virksomhed, fremover vil der blive opsamlet empiri fra flere SMV'er med henblik på at skabe et endnu bedre fundament. Dette betyder, at der stadig er en usikkerhed omkring, at dette er generelle tendenser, og at såvel den tekniske platform som de organisatoriske udfordringer kan generaliseres på tværs af flere SMV'er.

Gennem hele forløbet er der opsamlet erfaringer omkring arbejdsprocessen og samarbejdet med virksomheden. De vigtigste konklusioner på dette samarbejde beskrives nedenfor.

KONKLUSION

Gennem dette projekt er der designet og implementeret en stærk softwarearkitektur (lambdaarkitekturen), som gør det muligt at opbygge en softwareplatform til at modtage store mængder af data og behandle disse for at kunne tilgodese den umiddelbare beslutningsstøtte i "near real-time". Der er udviklet og testet maskinlæringsalgoritmer, som med god sikkerhed (>80 %) kan forudsige driftsstop. Samtidig konkluderes der, at der stadig er udfordringer med at lave en enkelt løsning, som passer på tværs af alle virksomheder – dette skal konfigureres til hver enkelt virksomhed.

På den organisatoriske del er der overordnet to delkonklusioner, som gør sig gældende:

1. Det er vigtigt at finde en business-case fra virksomheden, som tilgodeser produktionen inden for en kort tidsperiode, for at kunne

bekræfte ledelsen i, at arbejdet med dataanalyse er værdiskabende. Det er projektgruppens erfaring, at dette ikke nødvendigvis er i tråd med den tekniske implementering. Her må den tekniske implementering vige for at sikre, at der er den nødvendige gevinstrealisering for virksomheden med hensyn til en fortsættelse af projektet.

2. En anden central del ved dette projekt er, at det er blevet synligt, at det er vigtigt at have en forankring af projektet hos ledelsen fra starten af. Denne organisatoriske forankring er central for at kunne få gennemført ændringer og være samskabende (fx skrive videnskabelige artikler) omkring projektet. Der er i dette projekt udgivet flere videnskabelige artikler med virksomhedens IT- og udviklingschef som medforfatter.

Litteraturliste

- Hofman-Bang Direktør Industriens Fond, T., 2019. *Netværks-og Informationsmøde*.
- Hohpe, G., 2003. *Enterprise integration patterns : designing, building and deploying messaging solutions*. Pearson Professional Education, Harlow.
- Howson, Cindi., 2014. *Successful business intelligence : unlock the value of BI & big data*.
- Iftikhar, N., Baatrup-Andersen, T., Nordbjerg, F.E., Bobolea, E., Radu, P.-B., 2019a. Data Analytics for Smart Manufacturing: A Case Study, in: *8th International Conference on Data Science, Technology and Applications (DATA 2019)*.
- Iftikhar, N., Baatrup-Andersen, T., Nordbjerg, F.E., Bobolea, E., Radu, P.B., 2019b. Data analytics for smart manufacturing: A case study. *DATA 2019 - Proceedings of the 8th International Conference on Data Science, Technology and Applications 392–399*. <https://doi.org/10.5220/0008116203920399>
- Iftikhar, N., Lachowicz, B.P., Madatasz, A., Nordbjerg, F.E., Baatrup-Andersen, T., Jeppesen, K., 2020. Real-time Visualization of Sensor Data in Smart Manufacturing using Lambda Architecture, in: *Data2020*.
- Iftikhar, N., Nordbjerg, F., 2021. *Implementing Machine Learning in Small and Medium-sized Manufacturing Enterprises* — UC Knowledge - University Colleges Knowledge database.
- Marz, N., Warren, J., 2015. *Big Data - Principles and best practices of scalable realtime data systems*. Manning.
- Nordbjerg, F., 2021. *Scientific Methods 2: Investigations in Software Development*.
- Peffers, K., Tuunanen, T., Rothenberger, M., Chatterjee, S., 2007. A design science research methodology for information systems research. *Journal of Management Information Systems 24*, 45–77.