

Opsamling på prototype afprøvning	Ansvarlig	THA
	Oprettet	11-12-2023
Projekt: Datadrevet management i mælkeproduktionen	Side	1 af 6

Prototype afprøvning af AI algoritmer

Notat her i beskriver den gennemførte praksisafprøvning af alarmerne udviklet i Datadrevet management i mælkeproduktionen. Selve modellerne, teknikken bag prototypen og designet af prototypen er beskrevet i beskrevet i notatet [AI algoritme og prototype](#).

Sammendrag

I løbet af sommer og efterår 2023 har 6 besætningsejere og 2 driftsledere dagligt modtaget en mail med alarmer i besætningen, fra de algoritmer udviklet i Datadrevet management i mælkeproduktionen.

Besætningerne til afprøvningen er selekteret i forhold til: datakvalitet, geografi samt ønske om at indgå i afprøvningen.

Hovedkonklusion:

- Modtagerne af alarmmails har generelt fastholdt interesse for at kigge mails. Frekvensen for åbning af mail er dog faldet i løbet af afprøvningsperioden.
- Modtagerne af alarmmails føler ikke, at prototypen gav dem nævneværdig mere information end de havde tilgængeligt fra deres normale managementsystemer.
- Modtagerne af alarmmails har en lidt blandet opfattelse af handlingsanvisningerne. Personligt har de, som erfarne ko folk, den nødvendige viden til at håndtere sundheds- og produktionsalarmer. Flere fremhæver dog at handlingsanvisningerne giver en tryghed ved at bekræfte egen viden, samt en påmindelse om at tjekke det, man egentlig godt ved man bør gøre. Samtidigt nævner modtagerne, at handlingsanvisningerne kan være gode for ansatte med mindre erfaring.
- Live afprøvning af AI løsninger, giver et mere helt andet nuanceret indblik i modellerne end ved en normal data validering på et datasæt.

Formål med afprøvning

Projektet har til formål at automatisere management i malkekvægbesætningerne ved anvendelse af nye digitale teknologier, der udnytter store datamængder fra forskellige datakilder. Det formål er neddelt i 3 hypoteser, hvor afprøvningen skal afklare den sidste. De 3 hypoteser er:

- a) Vi kan levere valide forudsigelser med de data vi har på kvægdatabase i form af alarmer
- b) Vi kan knytte handlingsanvisninger på en alarm
- c) Landmændene kan udnytte alarmer og handlingsanvisninger i deres management

Målet med afprøvningen er at afklare den overordnede hypotese c. For at specificere dette blev der opstillet 7 detail hypoteser centreret omkring afprøvningen. De 7 hypoteser er:

1. Ejer/driftsleder åbner mailen daglig
2. Ejer/driftsleder fastholder interesse for mailen gennem hele afprøvningsperioden
3. Ejer/driftsleder vurderer, at alarmer giver værdi til deres driftsledelse på alle alarmtyper
4. Ejer/driftsleder undersøger udpegede dyr eller sender mail videre til medarbejder for at få dyr undersøgt
5. Antallet af dyr på en alarm liste har indflydelse på overblikket
6. Driftsleder bruger handlingsanvisninger som inspiration for sin indsats
7. Der er et kommercielt grundlag for at implementerer alarmerne i mælkeproducenternes normale managementværktøjer

Besætninger i afprøvning

Til udvælgelse af besætninger til afprøvningen, blev der trukket en liste med de besætninger fra Kvæg-databasen. Der var for at komme på listen krav om daglige vejninger og daglige ydelsesdata på koni-veau. Efterfølgende blev disse bedrifter kvalitetstjekket i forhold til geografisk placering (køreafstand til Aarhus) samt datakvalitet. Listen blev suppleret med besætninger, som havde givet input til udformning af prototypen i 2022. Der har sideløbende været efterlyst bedrifter til afprøvningen i KvægNyt (publiceret i 2022). Denne efterlysning gav intet respons.

Bruttolisten indeholdt samlet 20 bedrifter, hvor der fra 8. maj dagligt blev beregnet alarmer og genereret alarmmails til interne test af løsningen. Med baggrund i denne interne test kørsel, blev der foretaget mindre tilpasninger af prototypen samtidigt med at alarmernes driftsstabilitet blev overvåget. I forbindelse med vurdering af driftsstabilitet blev 9 bedrifter fravalgt, da der var ustabilitet i synkronisering af data fra malkerobotten eller der var flere mælkeleverende bedrifter på samme ejendom. Blandt de resterende blev ejer kontaktet. 4 ønskede ikke at deltage i afprøvningen. Efterfølgende blev afprøvningen udvidet med 2 besætninger, da en af værterne ønskede alarmmail på 2 mælkeleverende besætninger med samme ejer.

Bedrifterne i afprøvningen blev efterfølgende introduceret til prototypen enten telefonisk eller ved et møde på bedriften. Når besætningsejer havde underskrevet samarbejdsaftalen og fået en introduktion, blev mailen dagligt sendt til ejer og/eller driftsleder. Tabel 1 indeholder en anonymiseret liste med de besætninger, der indgik i afprøvningen, deres datagrundlag, afprøvningsperiode samt hvilken rolle modtageren af alarmmailen havde på ejendommen.

Tabel 1. Anonymiseret liste over bedrifter i afprøvning, med datagrundlag, afprøvningsperiode og modtagers rolle på bedriften. Bemærk at besætning E, G og H indgår i samme bedrift og der derfor er samme ejer som modtager alle 3 mails.

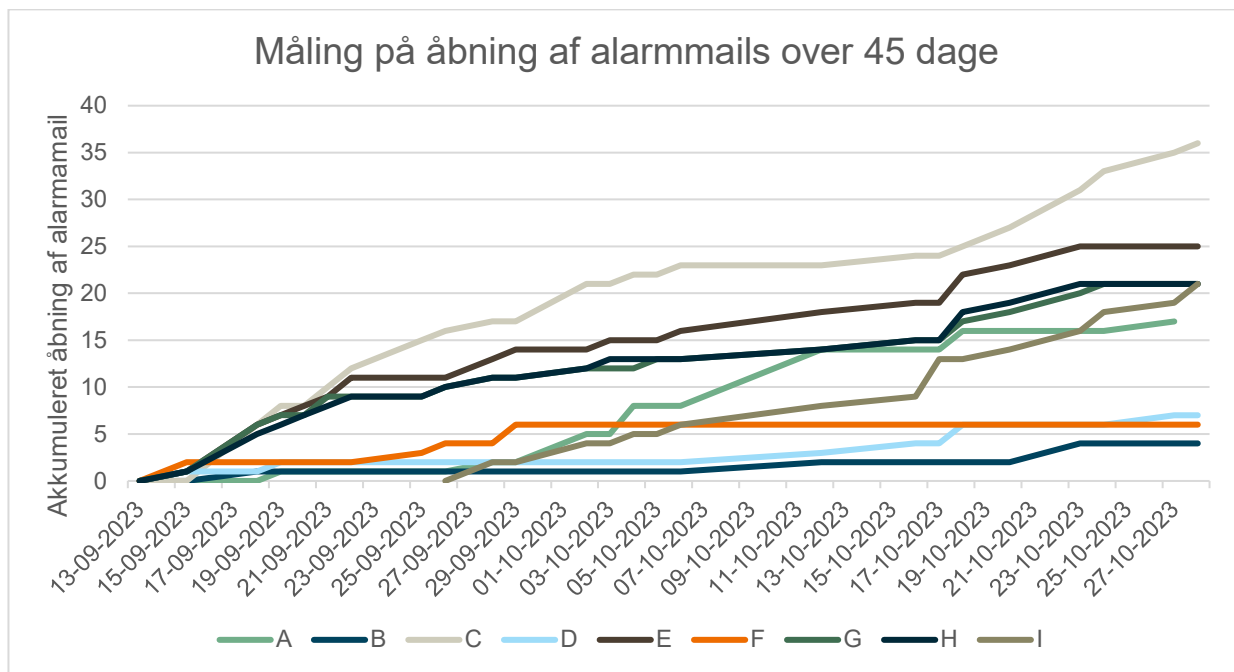
Besætning	Datagrundlag	Afprøvningsperiode	Modtager af alarmmail
A	Komplet	6. juli – 17. november	Ejer og driftsleder
B	Komplet	6. juli – 17. november	Ejer og driftsleder
C	Komplet	6. juli – 17. november	Ejer
D	Tankmælk og DMI	6. juli – 17. november	Ejer
E*	Tankmælk og DMI	14. juli – 17. november	Ejer
F	Tankmælk og DMI	14. juli – 17. november	Ejer
G*	Tankmælk og DMI	7. september – 17. november	Ejer
H*	Tankmælk og DMI	7. september – 17. november	Ejer
I	Tankmælk og DMI	23. september – 17. november	Driftsleder

Opfølgning på anvendelse

Adspurgte på åbningsfrekvens svarede afprøvningsværterne, at den er faldet i løbet af afprøvningsperioden. De angiver både bortfald af nyhedsværdien og perioder med sæsonarbejde i marken, som årsager til den mindre interesse. Hovedparten af afprøvningsværter har dog oplyst, at de har set mails mindst et par gange om ugen i den sidste del af forløbet. En enkelt vært oplyste, at han via emnefeltet fik en god fornemmelse for status i besætningen. Hvis der ikke var obs-punkter, åbnede han ikke mailen.

Fra medio september har vi målt på åbning af alarmmails. Resultatet fra målingen fremgår af figur 1, og viser en akkumuleret optælling af antal mails, der er åbnet pr. besætning. Systemet giver en indikation på åbningen. Systemet kræver dog, at mailklienten accepterer link til en webside. Derfor kan der være mailklienter, hvor en åbning ikke bliver registreret af systemet. Det betyder for eksempel, at besætning B godt kan have åbnet mere end 4 mails i perioden, men at det er 4 mails, som er åbnet fra en telefon/PC, som ikke er den primære kilde til åbning af alarmmails.

Det fremgår af figur 1, at hovedparten af de deltagende landmænd åbner alarmmails flere gang om ugen også sidst i afprøvningsperioden. Vi ser også at alle besætninger har perioder i efteråret, hvor der ikke er åbnet mails. Dette bekræfter interview oplysninger om at travlhed i forbindelse med høst af grovfoder og andet markarbejde har reduceret interessen for prototypen. Der er også besætninger, hvor der er indikationer på at interessen for at åbne mails faldet meget. Dette gælder fx besætning D, som ved interview har oplyst, at de tidligt i afprøvningsforløbet havde en indsats baseret på viden opnået fra prototypen, men at interessen efterfølgende var faldet kraftigt.



Figur 1. Akkumuleret åbning af alarmmails, målt pr. besætning målt fra 13. september til 28. oktober. Bemærk at besætning I blev igangsat i måleperioden. Besætning G og H igangsættes også i perioden, men her har besætningsejer modtaget alarmmails fra besætning E.

Det lidt lave, men stabile niveau sidst i afprøvningsperioden, kan hænge sammen med at besætnings-ejerne vidste, at alarmmailen ville stoppe. De fortsætter derfor med at bruge de informationskilder, de anvendte før afprøvningen.

Input til alarmområderne

Det kendetegner de landmænd, som har indgået i afprøvningen, at de var meget tilfredse med deres eksisterende løsning til at følge produktion og sundhed i besætningen.

For alarmerne, som kræver daglige oplysninger om vægt og ydelse på ko niveau, gælder, at besætningsejerne er yderst fortrolige med det managementsystem, som leverer grunddata til modellen. De oplyser samlet, at prototypen finder de samme dyr, men de oplever måske, at prototypen er lidt forsinket. Samtidigt har deres managementsystem også data fra aktivitetsmålere, som giver en ekstra sikkerhed på udpegning af køer med behov for ekstra tilsyn. Deres brug af prototypen er dermed kun lidt ekstra i forhold til de lister, de dagligt anvender til produktionsstyring.

Specifikt alarmerne på klove gav i den ene bedrift uforholdsmæssigt mange alarmkøer. Det betød, at ejer mistede tiltroen til alarmen. For de øvrige 2 bedrifter var der kun få klovalarmer.

Input til handlingsanvisninger

Direkte adspurgt udtrykte langt hovedparten at handlingsanvisningerne ikke gav dem ny information. Modsat mente de også at handlingsanvisningerne indeholdt de handlinger, som de ville udføre.

Anvisningerne var derfor rigtige, men lidt overflødige. Enkelte sagde dog, at det gav dem en sikkerhed i forhold deres egne valg. Flere påpegede, at anvisningerne kunne hjælpe deres mindre erfarne medarbejdere. Det blev også nævnt, at anvisningerne gav den en god påmindelse om at få igangsat en indsats. Det virker lidt bedre end bare en liste med dyrenumre.

Opfølgning på hypoteser

Der var forud for afprøvningen defineret 7 konkrete områder som skulle afdækkes i afprøvningen. I tabel 2 er der for hver hypotese angivet udfald og en uddybning. Hypoteserne er valideret på basis af interview med afprøvningsværterne.

Tabel 2. Hypoteser og test af hypoteser i afprøvningen. Udfaldet af hypoteserne er primært valideret på basis af interview med afprøvningsværterne.

Hypotese til test	Udfald	Bemærkning
Ejer/driftsleder åbner mailen daglig.	Falsk	I starten var der en stor interesse. Den bedrift, som havde 3 besætninger med i afprøvningen, kiggede dagligt emnetekst i mailklient. Derved fik vedkommende et hurtigt overblik uden at åbne mails. Hovedparten åbnede dog ikke mailen dagligt.
Ejer/driftslederen fastholder interesse for mailen gennem hele afprøvningsperioden.	Sandt	Frekvensen for åbning af mail falder sidst i afprøvningsperioden, men for hovedparten af landmændene åbnes mails stadig flere gange om ugen.
Ejer/driftsleder vurderer at alarmer giver værdi til deres driftsledelse på alle alarmtyper.	Falsk	Ved interview er der kun oplyst én konkret situation, hvor alarmer gav anledning til en konkret indsats. Alle værter oplyste dog, at alarmerne var plausible og i høj grad bekræftede den de informationer, de kunne finde i deres eksisterende managementværktøjer.
Ejer/driftslederen undersøger udpegede dyr eller sender mail videre til medarbejder for at få dyr undersøgt.	Falsk	3 af de besætninger, som indgik i afprøvningen, modtog alarmer på enkelt dyr. De anvendte alle en styringsliste fra deres malkeudstyr som basis for deres daglige arbejde i besætningen. Alle oplyste, at alarmmailen viste det samme, men de ændrede ikke deres management rutiner.
Antallet af dyr på en alarm liste har indflydelse på overblikket	Sandt	På alarmerne vedrørende klovproblemer oplevede en bedrift mange dyr på listerne. Det fik vedkommende til tidligt at opgive brugen af listen.
Driftslederen bruger handlingsanvisninger som inspiration for sin indsats.	Sandt	Hovedparten af værterne kunne se en værdi ved handlingsanvisningerne. De var ikke overrasket over indholdet, men det gav den en bekræftelse på egen viden. Enkelte fremhævede at handlingsanvisningerne ville være gode for deres ansatte med mindre erfaring.
Der er et kommercielt grundlag for at implementerer alarmerne i mælkeproducenternes normale managementværktøjer.	Falsk	Der er intet i afprøvningen, som indikerer, at der er en betalingsvillighed i forhold til at købe alarmmails som til produkt. Der vil dog være en kommerciel mulighed for at implementere dele af algoritmerne i eksisterende produkter. Algoritmerne kan også indgå som et supplement til salg af hardware (fx vejeudstyr).

Opfølgning på alarmområder

I tabel 3 er der angivet en anbefaling for hver af de 6 områder, som indgår i prototypen. Anbefalingerne bygger på interviews med de landmænd, der har deltaget i afprøvningen.

Tabel 3. Anbefaling og bemærkninger til de enkelte alarmområder i prototypen.

Område	Anbefaling	Bemærkning
Vægt	Klar til implementering	Model giver gode og anvendelige informationer til brug i det daglige management i en mælkeproduktion. Ved implementering skal det vurderes, om besætningsejer har eksisterende systemer, som præsenterer data for besætningsejer
Ydelse	Klar til implementering	Model giver gode og anvendelige informationer til brug i det daglige management i en mælkeproduktion. Ved implementering skal det vurderes, om besætningsejer har eksisterende systemer, som præsenterer data for besætningsejer
Ketose	Klar til implementering	Ketose modellen giver gode informationer til management. Datagrundlaget til ketose modellen er det samme som vægt og ydelse. Derfor er outputtet også korreleret. Derfor bør det overvejes at se de 3 modeller i en sammenhæng.
Klove	Ikke moden til implementering	Klovmodellen har i praksis vist sig fungere med meget varierende succes. I nogle bedrifter har outputtet virket plausibelt. I andre tilfælde har modellen vist uforklarlige resultater, hvor der pludselige er kommet mange dyr til eller fra listen. Den praktiske afprøvning viser derfor, at der skal arbejdes mere med klovmodellen forud for implementering i et management-program.
Tankmælk	Kan eventuelt indarbejdes i de værktøjer, hvor mælkeproducenterne i øvrigt følger udviklingen i mælketanken.	Tankmælks alarmerne ligger meget tæt op ad outputtet fra DMS ¹ modulet "KMP". Derfor oplever brugerne begrænset ny information, selvom alarmer i prototypen bygger på en mere databaseret og dynamiske opfølgning. Det bør overvejes, om der skal ske en tilpasning af alarmteksterne. Som bruger kan det være svært at adskille " <i>forholdsvis stort fald</i> " fra " <i>faldende over længere tid</i> ". Det bør også vurderes, om teksten " <i>uændret</i> " er sigende. I prototypen blev den anvendt, når der ikke var alarmer. Det omfattede også situationen, hvor der var en stigning i en positiv retning.
Varme-stress	Klar til implementering. Dog skal det afklares, hvorfor der er observeret forskelle i de data, DMI leverer fra deres API og deres lokalprognoser.	Vi har i prototypen observeret, at der er forskelle i DMI-prognoserne afhængigt af, om data leveres fra deres API eller i deres lokaludsigt. Det skaber tvivl om datagrundlaget og dermed usikkerhed på værdien af omsætning af vejrudsigten til THI.

¹ DMS er et management værktøj, som anvendes af hovedparten af danske mælkeproducenter

Brugerafprøvning som supplement til normal validering af AI modeller

I projektet, har vi helt naturligt lagt stor vægt på udvikling af AI algoritmer. Det er opgaver med datatræk, datamodellering samt flere iterationer med træning og validering. Forud for afprøvningen stod vi derfor med modeller, som var valideret efter helt normale standarder. Vi må dog erkende at det at afprøve modellerne i et live setup, med daglige ny kørsel af data, giver et helt andet indblik i hvordan en model fungerer i praksis sammenlignet med for eksempel en validering ved hjælp af en confusion matrix.

Ved live at følge alarmer i en besætning, også selvom det sker uden involvering af besætningsejer eller som en simulering på baggrund af et datatræk, får man en række andre input. Vi opdagede for eksempel at vores klovmodel gav meget varierende resultater i testbesætningerne. Så varierende at resultaterne ikke virkede tilforladelige, selvom den på dyreniveau validerede fint.

Det er derfor en klar anbefaling fra projektet, at nye indsatser omkring kunstig intelligens, skal valideres ud fra en bruger synsvinkel. Enten ved et live setup som anvendt her, eller ved en simulering.