

Prosjektplan

Navn på sandkasseprosjekt: Prediksjon av sykefravær for mer effektiv oppfølging

Navn på virksomhet(er): NAV

Deltakere fra virksomhet:

Navn	Rolle	E-post
Mari Håvardsholm Dramdal	Jurist	
Elisabeth Njøs Hatle	Jurist	
Anders Holt	Personvernombud	
Helene Lund	Kommunikasjonsrådgiver	
Cathrine Pihl Lyngstad	Seksjonssjef, dataseksjonen	
Robindra Prabhu	Data scientist	
Tor Halle Rise	Produkteier	
Sigurd Skogvoll	Fag, sykefraværsoppfølging	
Lars Sutterud	Data scientist	
Andrea Venge Tollefsen	Jurist, ekstern konsulent	

Deltakere fra Datatilsynet:

Navn	Rolle	E-post	Telefon
Dag Grytli	Prosjektleder		
Hallstein Husand	Teknolog		
Berit Bye Rinnan	Jurist		

1. Kort beskrivelse av prosjektet

Om sykefraværsoppfølgingen

Sykefraværsoppfølgingen i Norge baserer seg på definerte stoppunkter. Ved fire ukers sykefravær er arbeidsgiver pålagt å utarbeide en oppfølgingsplan. Innen syv uker skal sykmeldt og arbeidsgiver gjennomføre dialogmøte 1. Ved åtte uker er NAV pålagt å kontrollere om den sykmeldte er i aktivitet, eller om den kan unntas fra aktivitetsplikten. Ved 17 ukers sykefravær er NAV pålagt å bestemme om et dialogmøte med den sykmeldte, arbeidsgiver og sykmelder er nødvendig. På alle disse stoppunktene må de fleste sykmeldte vurderes.

Dette prosjektet tar utgangspunkt i stoppunktet ved 17 ukers sykefravær - beslutningen om å innkalle til dialogmøte.

En av våre hypoteser er at det avholdes for mange unødvendige dialogmøter, som stjeler tid fra de involverte partene.

Ved å bruke maskinlæring for å predikere sykefraværlengde utover 17 uker ønsker vi å understøtte veileders beslutning om behov for dialogmøte 2.

Samfunns- og brukernytte

Gjennom sandkasseprosjektet ønsker NAV å teste ut og drøfte fremgangsmåter, metoder og vurderinger for å klargjøre hvilke føringer og terskler som gjelder ved bruk av kunstig intelligens (KI). Dette for å kunne danne førende praksis for hvordan NAV skal sikre kontroll og ansvarlighet gjennom et KI-utviklingsløp. Prosjektet vil være et viktig bidrag på veien for å danne retningsgivende praksis for hvordan NAV skal arbeide når vi ønsker å benytte KI, og for å klargjøre hvilket handlingsrom NAV har til å bruke KI. Dette arbeidet mener vi vil kunne bidra til å:

- Gi større muligheter for å benytte KI der dette er lovlig og ansvarlig, herunder klargjøre om NAV som offentlig forvaltning må gis klare føringer for når vi kan og hvordan vi skal benytte KI.
- Skape større trygghet for at våre vurderinger og praksis er i tråd med både regelverk og god praksis.
- Korte ned vei fra idé til implementert (ansvarlig) KI på andre områder i NAV, samt for andre virksomheter som ser potensialet i lignende KI-anvendelser.

Prosjektet vil med andre ord ha generell nytte og overføringsverdi for NAV generelt, men også for andre virksomheter. I tillegg vil selve løsningen vi diskuterer i prosjektet ha konkret samfunns- og brukernytte:

- Bedre oppfølging for sykmeldte som har behov for dialogmøte, ved at kun de som har behov blir innkalt
- Partene involvert i sykefraværet spares for tid ved å i større grad unngå å kalle inn til unødvendige dialogmøter
- Redusert tidsbruk på vurdering av behov for dialogmøte for NAVs sykefraværsveiledere.

Ansvarlighet

For NAV er det essensielt at vi utforsker og tar i bruk KI på en *ansvarlig* måte. Dette gjelder alle, men som en stor offentlig aktør er vi vårt særlige ansvar bevisst. Anvendelsen skal ikke bare være lovlig, men også innrettes i tråd med samfunnsetiske normer, på en måte som møter individet med respekt og verdighet, og ikke minst bygger tillit. Vi er kjent med hvilken fallhøyde som følger teknologien, ikke minst innenfor kompliserte tema som rettferdighet, forklarbarhet og tillit. Samtidig gir teknologien oss muligheter til å "diagnostisere" og blottlegge slike problemstillinger. Vi tror at en ansvarlig tilnærming i praksis må gripe denne muligheten gjennom hele utviklingsløpet og gjøre slike problemer synlige og til gjenstand for tverrfaglig diskusjon. Gjennom deltakelse i sandkassen ønsker vi å gjøre nettopp dette, bl.a. ved å bringe konkrete vurderinger av lovlighet, rettferdighet og forklarbarhet til torgs. Vi håper med det å kunne ta i bruk KI-løsningen raskere, og med større trygghet for at våre vurderinger og praksis er i tråd med både regelverk og god praksis.

2. Mål og forventede leveranser i prosjektet

Mål 1: Avklare behandlingsgrunnlag knyttet til planlagt behandling.

- Leveranse 1.1: Avklare om det er behov for et tydelig rettslig grunnlag for selve metoden. Trenger NAV en uttrykkelig hjemmel for å predikere og/eller bruke maskinlæring? Er det forskjell på om vi implementerer en ekspertbasert regelmotor som gir anbefalinger, eller om den er basert på data og maskinlæring?

- Leveranse 1.2a: Avklare om folketrygdloven § 8-7a alene gir et dekkende rettslig grunnlag? Er det flere bestemmelser i folketrygdloven som må drøftes i vurderingen av om NAV har supplerende rettslig grunnlag? Vi ønsker å se nærmere på hva som er lovgiverintensjonen bak § 8-7a, og om prediksjonsmodellen bidrar til å oppnå dette formålet?
- Leveranse 1.2b: Behandlingens lovlighet avhenger av at den også er forholdsmessig. Gir § 8-7a alene hjemmel til å kunne behandle personopplysninger i det antallet og i det omfanget det er behov for?

Mål 2: Rettferdighet - gjennomgå NAVs vurdering av rettferdighet for maskinlæringsmodellen

- Leveranse 2.1: Vurdere om vi har valgt gode/riktige rettferdighetsmål for utfallet og i overenstemmelse med lovkrav. Hvis en vurdering av utfallsrettferdighet ikke er tilstrekkelig (og det er det sannsynligvis ikke), hvordan kan/bør en rettferdighetsvurdering/evaluering av datasett og innretning se ut?
- Leveranse 2.2: Avklare hvem som har særlig krav på beskyttelse og hvordan diskriminering av disse kan vurderes i praksis. Er det eksempelvis mulig/tilrådelig å benytte særskilte kategorier for å evaluere diskriminering mot disse?

Mål 3: Forklarbarhet - gjennomgå NAVs tilnærming til å forklare en maskinlæringsmodell

- Leveranse 3.1: Vurdere hvorvidt vårt valg av forklarbarhetsmetode og innretning for enkeltprediksjoner harmonerer med lovkrav og hvor/hvorfor den evt kommer til kort.
- Leveranse 3.2: Hvordan innrette forklaringer mot ulike brukere for å imøtekomme forsvarlig forvaltning både på enkeltsaksnivå og systemnivå? Hvordan ser en god forklaring på systemnivå ut?
- Leveranse 3.3: A29WP og ICO etterspør begge det sistnevnte kaller "impact explanation", dvs. "explain significance and envisage consequences of the processing". Hvordan kan og bør en slik forklaring se ut i dette tilfellet?

3. Aktiviteter, metode, arbeidsfordeling og tidsplan

Tema	Leveranse	Aktivitet	Metode	Arbeidsfordeling	Periode
Behandlingsgrunnlag	1.1 Lovanvendelse	Diskusjoner	Workshop(s) + asynkront ved kommentarer i dokumenter	NAV leder prosessen. Datatilsynet stiller med	10.5-18.6

				folk + evt. aktuelle verktøy	
Behandlingsgrunnlag	1.2a og 1.2b	Diskusjoner	Workshop(s) + asynkront ved kommentarer i dokumenter	NAV og Datatilsynet	10.5-18.6
Rettferdighet	2.1 Lovkrav rettferdighet	Kartlegging av praksis (inkl. dokumentasjon av vurderinger)	Uformelt tilsyn (tentativt)	Datatilsynet leder prosessen	25.5-18.6
Rettferdighet	2.2 Særlige krav til beskyttelse	Diskusjoner	Workshop(s) + asynkront ved kommentarer i dokumenter	NAV og Datatilsynet	25.5-18.6
Forklarbarhet	3.1 Lovkrav forklarbarhet	Kartlegging av praksis (inkl. dokumentasjon av vurderinger)	Workshop(s)	Datatilsynet leder prosessen	8.6-18.6
Forklarbarhet	3.2 Utforme forklaringer på enkeltsaksnivå og systemnivå	Diskusjoner	Workshop(s) + asynkront ved kommentarer i dokumenter	NAV og Datatilsynet	8.6-18.6
Forklarbarhet	3.3 Impact explanation	Diskusjoner	Workshop(s) + asynkront ved kommentarer i dokumenter	NAV og Datatilsynet	8.6-18.6
Sluttrapport	Sluttrapport	Rapport, kronikk, podcast	Utarbeide en sluttrapport og gjennomføre utadrettede aktiviteter	Datatilsynet og NAV	18.6 – 15.8

4. Kommunikasjon

Prosjektet vil dele resultater og erfaringer både eksternt og internt. Som et minimum vil prosjektplanen og sluttrapporten gjøres tilgjengelig utad, i tillegg til at milepæler og resultater underveis i prosjektet kan publiseres på prosjektets nettside eller i sandkassens podcast.

Det utarbeides en egen kommunikasjonsplan for prosjektet med detaljer rundt mål, målgrupper og planlagte tiltak for eksternt og intern kommunikasjon. Ansvar for framdrift og oppfølging av kommunikasjonsplanen ligger hos Datatilsynets kontaktperson i sandkasseprosjektet.

Den utadrettede virksomheten bør fokusere på den potensielle samfunnsnyttene både i form av effektiv forvaltning og gevinster for enkeltindivider. I tillegg bør det fremheves at et vellykket prosjekt vil få betydning for andre aktører som ønsker å utforske lignende metoder.