Tidsskriftet Tidsskrift for Den norske legeforening

Maskinlæring i medisinsk forskning

MEDISIN OG TALL

GURO F. GISKEØDEGÅRD

guro.giskeodegard@ntnu.no
Guro F. Giskeødegård er førsteamanuensis i biostatistikk ved K.G. Jebsen-senter for genetisk epidemiologi ved NTNU.
Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

STIAN LYDERSEN

Stian Lydersen er dr.ing. og professor i medisinsk statistikk ved Regionalt kunnskapssenter for barn og unge – psykisk helse og barnevern (RKBU Midt-Norge) ved Institutt for psykisk helse ved NTNU.
Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

Maskinlæring brukes til å finne underliggende mønster i data. Dette kan være nyttig i medisinsk forskning.

Maskinlæring er en form for kunstig intelligens og brukes til finne underliggende mønster i data. Maskinlæring kan bygge på statistiske metoder eller andre metoder fra matematikk eller informatikk som ikke legger en sannsynlighetsmodell til grunn. Maskinlæring er spesielt nyttig når man har store datasett med mange variabler, og læringen innebærer å trene opp en modell for å finne sammenhenger mellom variablene. Ofte er hensikten å bygge en modell som kan predikere et utfall. Et typisk trekk ved mange maskinlæringsmetoder er at treningsprosessen er iterativ, det vil si at man gjør én endring om gangen slik at tilpasningen til dataene blir bedre og bedre. Innen maskinlæring brukes ofte den engelske termen *features* synonymt med *variabel* innen statistikk, og termen *karakteristikk* synonymt med *utfallsvariabel*.

Ikke-veiledet maskinlæring

Man skiller ofte mellom ikke-veiledet og veiledet maskinlæring (engelsk: *unsupervised* og *supervised machine learning*). Med ikke-veiledede metoder bruker man kun de målte variablene til å finne sammenhenger og gruppestrukturer i dataene, uten å skulle predikere en utfallsvariabel. Dersom man i en studie har målt genuttrykk i blodprøver fra kreftpasienter og friske kontrollpersoner, vil man i ikke-veiledet maskinlæring kun bruke genutrykksdataene i analysen, uten informasjon om hvilke prøver som tilhører hhv. pasienter og kontrollpersoner. Modellen vil da fortelle oss om vi har naturlige gruppestrukturer i dataene, for eksempel at personenes alder har stor påvirkning på genutrykket, eller at prøver fra pasientene skiller seg fra kontrollene. Ikke-veiledede metoder fungerer også utmerket for å detektere ekstreme verdier eller ekstreme kombinasjoner av verdier. Hierarkisk klyngeanalyse, prinsipal komponentanalyse og den nyere metoden UMAP (*uniform manifold approximation and projection*) er eksempler på metoder for ikke-veiledet maskinlæring (1). Felles for disse er at de kan visualisere store datasett med mange variabler på en enkel måte i få dimensjoner.

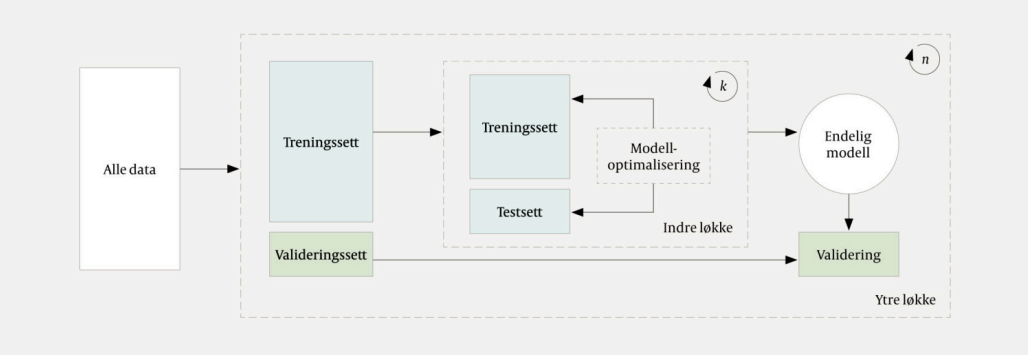
Veiledet maskinlæring

I veiledet maskinlæring brukes en eller flere karakteristikker (utfallsvariabler) mens modellen trener, og man ender opp med en modell som er optimalisert for å finne sammenhengen mellom forklaringsvariablene og karakteristikken man er interessert i. Karakteristikken kan være en kontinuerlig eller en kategorisk variabel. I eksemplet med genutrykkdata vil en veiledet maskinlæringsmetode informeres om hvilke prøver som tilhører pasienter, og hvilke som tilhører kontroller. Vi kan på denne måten bygge en modell som kan predikere om et gitt genutrykk tilhører en kreftpasient eller en kontroll.

Noen veiledede maskinlæringsmodeller fungerer som såkalte svarte bokser, det vil si at modellen kan predikere status for en ny prøve, men uten å gi informasjon om *hvorfor* prøven gis denne statusen. Nevrale nettverk og XGBoost (*extreme gradient boosting*) er eksempler på populære svart boks-modeller. Slike modeller er mindre nyttige dersom vi er interessert i å forstå den underliggende biologien som skiller pasienter fra kontrollpersoner. Det er imidlertid et stort fokus innen forskningen på å «åpne opp» slike svarte bokser slik at vi kan forstå hvorfor de modellerer som de gjør, og det er økende interesse for fagfeltet forklarbar kunstig intelligens (XAI, *explainable artificial intelligence*). Det finnes også veiledede maskinlæringsmetoder som naturlig gir informasjon om hvilke variabler som bidrar til prediksjonen, som for eksempel visse regresjonsmodeller.

Validering er svært viktig

En kompleks veiledet maskinlæringsmodell er svært god til å finne mønster i data. Faktisk kan den gjøre jobben så godt at den blir overtilpasset dataene den har lært ifra. En overtilpasset modell vil beskrive dataene den har lært fra svært godt, men fungerer dårlig for nye data. Dermed vil biologiske tolkninger av en slik modell gi oss unøyaktig eller feil informasjon. Det er derfor viktig at veiledede maskinlæringsmodeller er godt validerte (figur 1). Dette innebærer at man trener og optimaliserer modellen på en del av dataene, ofte kalt treningssettet, for eksempel tilfeldig utvalg av 80 % av dataene. Deretter valideres den endelige modellen på data som ikke ble brukt i læringsprosessen (ofte kalt valideringssettet).



Figur 1 Under validering av en veiledet maskinlæringsmodell deles dataene opp i et treningssett og et valideringssett. Modellen optimaliseres gjennom en læringsprosess, og den endelige modellen valideres med valideringssettet. Ofte deles treningssettet opp i ytterligere trenings- og testsett i en indre løkke, hvor modellen optimaliseres. Den indre løkken gjentas gjerne flere ganger (her: *k* ganger) med tilfeldig oppdeling i trenings- og testsett. Den ytre løkken kan også gjentas *n* ganger for å estimere variasjonen i modellens yteevne.

REFERENCES

1. Røislien J, Langaas M. Klynger. Tidsskr Nor Legeforen 2022; 142: 1586. [PubMed][CrossRef]

Publisert: 29. mai 2023. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.23.0196