Kan maskinlæring forutsi hastigheten på norske veier?



Forfattere:
Vilde Hedemark Valen-Sendstad, masterstudent, NMBU
Øyvind Lervik Nilsen, Førsteamanuensis II, NMBU/ Sjefsingeniør Statens vegvesen
Kristian Hovde Liland, professor, NMBU

Denne artikkelen er basert på hovedforfatterens masteroppgave i Industriell Økonomi ved NMBU, våren 2026. Masteroppgaven omfatter 30 studiepoeng og er utarbeidet ved Fakultetet for Realfag og Teknologi.

I Nasjonal Transportplan 2025–2036 er kunstig intelligens fremhevet som et sentralt verktøy for å optimalisere trafikkflyt på veinettet. En type kunstig intelligens er maskinlæringsmodeller. Men hvor godt fungerer maskinlæringsmodeller egentlig på norske veier? I denne studien har vi testet fem ulike maskinlæringsmodeller på data fra 30 veistrekninger i Oslo og Akershus. To av modellene viste seg svært treffsikre på veier de allerede hadde data fra, med en gjennomsnittlig feil på rundt 4 prosentpoeng, tilsvarende omtrent ±3 km/t i en 80-sone. Da modellene ble testet på veier de ikke hadde sett tidligere, falt nøyaktigheten markant. Dette tyder på at modellene i stor grad lærer seg særtrekk ved den enkelte veien, fremfor generelle mønstre som kan overføres til nye strekninger. Funnene viser at maskinlæring kan bli et nyttig verktøy for trafikkstyring på veier vi allerede har god kunnskap om, men at det trengs mer informasjon om veienes faktiske egenskaper for at modellene skal kunne brukes på strekninger de ikke er trent på.

INTRODUKSJON

Effektiv trafikkstyring er en av de største utfordringene i moderne byer. Når trafikken øker, blir det stadig viktigere å kunne forutsi når og hvor utfordringer vil kunne oppstå for å kunne gi god informasjon til brukerne om fremkommeligheten på vegnettet. I Nasjonal Transportplan 2025–2036 trekkes kunstig intelligens frem som et sentralt virkemiddel for å optimalisere bruken av eksisterende infrastruktur og transportnett (Samferdselsdepartementet & Nærings- og fiskeridepartementet, 2024). Likevel finnes det lite forskning på hvor godt disse modellene fungerer i en norsk setting.

Maskinlæring har de siste årene blitt et stadig viktigere verktøy i arbeidet med smartere transportsystemer. Slike modeller kan finne mønstre i store datamengder og ta hensyn til mange ulike faktorer samtidig, som trafikkmengde, tidspunkt og værforhold (Hastie et al., 2009). Dette gjør maskinlæring særlig egnet for oppgaver som prediksjon av hastighet og estimering av kødannelse.

Samtidig skiller Norge seg fra andre land der forskning på trafikkprediksjon er gjennomført. Norge har store sesongvariasjoner, markante forskjeller mellom kyst og innland, og vesentlig variasjon i nedbør, frost og vind på tvers av regioner. Derfor kan resultater fra internasjonale studier, som ofte er gjennomført i Asia, ikke nødvendigvis overføres direkte til norske forhold (Razali et al., 2021; Zhou et al., 2022). Dette tydeliggjør behovet for forskning basert på norske data.

Forskning på trafikkhastighetsprediksjon er mindre utbredt enn forskning på trafikkmengdeprediksjon. I tillegg er eksisterende studier i stor grad dominert av dyplæringsmetoder som krever flere ressurser og mer data for utvikling og drift (Razali et al., 2021). Mange studier tester heller ikke om modellene fungerer på veistrekninger de ikke er trent på (Pavlyuk, 2021), og værvariabler utelates ofte eller behandles ulikt fra studie til studie (Razali et al., 2021; Shaygan et al., 2022). Tradisjonelle maskinlæringsmetoder kan være mer tolkbare og gir ofte høy prediksjonsytelse, selv med begrensede datamengder (Hastie et al., 2009). Disse metodene representerer derfor et attraktivt alternativ når datatilgangen er begrenset og forståelse av modellens beslutninger er viktig.

I denne studien undersøker vi hvor godt fem ulike maskinlæringsmodeller kan forutsi hastighetsavvik på 30 veistrekninger i Oslo og Akershus. Vi undersøker også hvor godt modellene fungerer på veier de ikke har sett tidligere, hvor godt de klarer å forutsi trafikken i fremtidige tidsperioder, og hvilke faktorer som har størst betydning for hastigheten.

METODISK RAMMEVERK

Data og målvariabel

Datagrunnlaget består av fire kilder: Trafikkdata og reisetidsdata fra Norsk vegdatabank, fartsgrensedata fra Statens vegvesens vegkart, samt værdata fra Meteorologisk institutts Frost API. Dataene dekker kalenderåret 2025, og datasettene fra Statens vegvesen kommer fra 30 ulike strekninger i Oslo og Akershus. Det endelige datasettet inneholder 262 800 observasjoner der hver observasjon er et aggregert gjennomsnitt på time-nivå. Variablene i datasettet inkluderer blant annet trafikkmengde, fartsgrense, trafikktetthet, andel tungtrafikk, værforhold, tidspunkt på døgnet og hvilken veistrekning målingene kom fra.

Studien undersøker hvor mye den faktiske gjennomsnittlige hastigheten avviker fra fartsgrensen og beregnes som vist i Formel 1. Det betyr at dersom bilistene kjører i 76 km/t på en strekning med fartsgrense 80 km/t, er hastighetsavviket -5%.

$$Hastighetsavvik\ (\%) = \frac{Observert\ hastighet - Fartsgrense}{Fartsgrense} \quad (1)$$

Denne måten å måle hastigheten på gjør det mulig å sammenlikne strekninger med ulike fartsgrenser direkte og synliggjør også strekninger der avvikene er større. Det gir verdifull informasjon om hvor og når trafikken flyter dårligere enn forventet, noe som er viktig for både fremkommelighet og forutsigbarhet i transportsystemet.

Modeller

Fem modeller ble valgt som kandidatmodeller: Random Forest, Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), K-Nearest Neighbors (KNN) og Support Vector Machine (SVM). Trebaserte modeller som Random Forest, XGBoost og LightGBM presterer ofte godt på ikke-lineære og flerdimensjonale datasett, mens KNN og SVM ble valgt for å representere algoritmer med fundamentalt ulike modelleringstilnærminger.

Evaluering

For å undersøke hvor godt modellene egnet seg til oppgaven, ble de testet på tre ulike måter. Dette gjorde det mulig å undersøke hvor godt modellene fungerer på veier de ikke har sett tidligere, i tidsperioder de ikke er trent på, og på data som ellers ligner treningsdataene. De tre modellene som presterte best, ble videreutviklet og evaluert grundigere.

I den endelige evalueringen ble de tre beste modellene vurdert ut fra hvor presise prediksjonene var (RMSE), hvor store feil de gjorde i gjennomsnitt (MAE) og hvor godt de klarte å

forklare variasjonene i hastigheten (R^2). Suksesskriteriene ble satt på forhånd til $R^2 \geq 0{,}80$, MAE ≤ 4,3 % og RMSE ≤ 7,5 %, motivert av grenser satt i andre studier knyttet til trafikkprediksjon. For disse tre modellene undersøkte vi også hvilke faktorer som påvirket prediksjonene mest og hvordan frost og nedbør påvirker hastigheten.

RESULTATER

Resultatene fra den innledende evalueringen er presentert i Tabell 1. Analysen viser at modellene presterer godt både når dataene er delt opp tilfeldig og når de deles opp med hensyn til tid, som representert ved kolonnene «Tilfeldig» og «Tid» i Tabell 1. Når modellene testes på ukjente strekninger, faller derimot ytelsen vesentlig, som vist i kolonnen «Rom» i Tabell 1. De tre trebaserte modellene (Random Forest, LightGBM, XGBoost) oppnår generelt best resultater og ble derfor utviklet videre. SVM og KNN ekskluderes basert på både ytelse og beregningsmessige kostnader ved kjøring av modellene.

Tabell 1: R^2, MAE og standardavvik for alle modeller og valideringsstrategier. De beste resultatene er markert i fet skrift.

	MAE (%)			**R^2**		
Modell	Tilfeldig	Rom	Tid	Tilfeldig	Rom	Tid
XGBoost	**6.74**	8.32	**6.79**	**0.66**	**0.29**	**0.65**
Random Forest	6.78	8.54	6.88	**0.66**	**0.29**	**0.65**
LightGBM	8.10	8.26	8.08	0.54	0.28	0.54
KNN	8.19	10.01	8.26	0.51	0.01	0.49
SVM	9.07	**6.14**	9.16	0.34	0.27	0.31

Det viktigste funnet er at modellene blir betydelig mindre treffsikre når de brukes på veier de ikke kjenner fra før. Her utgjør SVM et interessant unntak. Ved testing på ukjente strekninger har SVM den laveste gjennomsnittlige feilen med en MAE på rundt 6%. Likevel er det verdt å merke seg at modellen har en relativt lav forklaringsgrad (R^2) på 0.27, hvilket betyr at modellen kun forklarer 27% av variasjonen i hastighetsavviket. Dette tyder på at det er andre årsaker til at SVM oppnår lavere feil enn at modellen faktisk har lært mønstre som kan overføres til nye veier.

Optimaliserte modeller

Tabell 2 sammenligner R^2, MAE og RMSE for de optimaliserte modellene.

Tabell 2: R^2, MAE og RMSE for XGBoost, LightGBM og Random Forest. Beste resultater er markert i fet skrift.

	Resultater for optimaliserte modeller			
Modell	MAE (%)	RMSE (%)	R^2	Optimaliserings-tid
LightGBM	4.34	**7.37**	**0.809**	701.25
Random Forest	**4.22**	7.45	0.804	2272.85
XGBoost	4.96	8.32	0.756	**223.18**

Random Forest og LightGBM møter alle forhåndsdefinerte suksesskriterier med tilnærmet lik ytelse ($R^2 \approx 0{,}80$ og MAE ≈ 4,3 %, RMSE ≈ 7,4 %). Med en gjennomsnittlig feil på om lag 4,3 prosentpoeng fra fartsgrensen, betyr det i praksis at modellen i en 80-sone typisk treffer innenfor ±3,4 km/t. XGBoost presterer noe svakere med $R^2 = 0{,}756$ og overskrider RMSE-terskelen med 0,32 prosentpoeng. LightGBM foretrekkes fremfor Random Forest fordi den krever vesentlig kortere beregningstid, noe som er relevant dersom modellen skal oppdateres jevnlig.

Å fjerne strekningsidentifikatoren (Strekningsnavn) fra modellen reduserer forklaringsgraden med 9–14 prosentpoeng for alle modeller. Dette tyder på at variabelen fanger opp reell, stedsspesifikk informasjon som ikke finnes eksplisitt i datasettet.

Feature importance og betydningen av værvariabler

Analyser av variablenes viktighet viser at fem faktorer hadde størst betydning for modellene: hvilken veistrekning det er snakk om, trafikktetthet, antall kjøretøy, tidspunkt på døgnet og andelen tungtrafikk. Høy trafikktetthet og høyt volum var konsekvent forbundet med lavere hastighet enn fartsgrensen. Det samme gjaldt rushtiden, der mye kø naturlig nok ga lavere hastighet. Om natten var derimot mønsteret motsatt: da kjørte folk oftere fortere enn fartsgrensen, sannsynligvis fordi det er mindre trafikk.

Værforholdene hadde mindre betydning enn trafikkforholdene, men dypere analyser avslører at været likevel påvirker hastigheten. Lavere temperaturer gir mer negative hastighetsprediksjoner, med en spesielt tydelig nedgang rundt 0°C. Nedbør er assosiert med redusert hastighet både med og uten frost, men den initiale nedgangen er brattere under frostforhold. Analysen viser også at observert hastighetsavvik synker med 2–10 prosentpoeng under nedbør, både med og uten frost, hvilket bekrefter at sjåfører kjører saktere i dårlig vær. Dette viser at modellene fanger opp reelle effekter av værforholdene, selv om værvariablene samlet sett betyr mindre enn trafikkvariablene. Den mest sannsynlige forklaringen er at værvariablene overskygges av den dominerende innflytelsen fra trafikkvariablene under trening.

DISKUSJON

Det viktigste funnet i studien er at modellene fungerer godt på veier de allerede kjenner, men betydelig dårligere på nye veier. Dette tyder på at modellene i stor grad lærer seg særtrekk ved den enkelte veistrekning, fremfor generelle mønstre som kan brukes overalt. Funnet samsvarer med tidligere forskning, som viser at modeller ofte virker mer treffsikre enn de faktisk er dersom de ikke testes på nye geografiske områder (Sweet et al., 2023, Pavlyuk, 2021).

Et annet viktig funn er hvor stor betydning veistrekningen i seg selv har. Strekningsnavnet fungerer i praksis som en indikator på forhold som ikke finnes direkte i datasettet, som veibredde, svinger, siktforhold og lokale trafikkmønstre. Når strekningsnavnet fjernes, blir modellene merkbart dårligere. Det skyldes ikke svakheter ved modellene, men at viktig informasjon om veiene mangler i datagrunnlaget.

Hvor problematisk dette er, avhenger av hva modellen skal brukes til. Dersom målet er å følge trafikkutviklingen på veier man allerede kjenner godt, fungerer modellene svært bra. Da er det også naturlig å bruke strekningsnavn, siden det fanger opp forskjeller mellom veiene som ikke er registrert andre steder i dataene.

Dersom målet derimot er å forutsi trafikken på helt nye veier, er situasjonen en annen. Resultatene viser at modellene foreløpig ikke er like treffsikre når de møter veier de ikke har sett før. For å lykkes med dette trengs mer detaljert informasjon om veienes egenskaper, som bredde, kurvatur og stigningsforhold.

Tidligere forskning på trafikkhastighetsprediksjon har i stor grad benyttet mer komplekse dyplæringsmodeller som CNN, LSTM og GCN (Zhou et al., 2022; Razali et al., 2021). Det er vanskelig å sammenligne resultatene direkte med tidligere studier, både fordi forskerne bruker ulike datasett og fordi trafikk måles på forskjellige måter. Resultatene i denne studien er likevel sammenlignbare med publiserte resultater fra mer komplekse modeller, samtidig som modellene er lettere å forstå og krever mindre beregningskraft. LightGBM oppnådde omtrent samme nøyaktighet som Random Forest, men brukte langt kortere tid på å optimaliseres. Det kan være en viktig fordel dersom modellen skal oppdateres ofte eller brukes i sanntid. Derfor fremstår LightGBM som den mest egnede modellen i dette datasettet, med både høy ytelse og relativt lav optimaliseringstid.

Modellene er ikke klare for full utrulling ennå, men resultatene viser at maskinlæring kan være et nyttig verktøy i trafikkforvaltningen. Slike modeller kan gi bedre grunnlag for å forutse endringer i fremkommelighet og trafikkflyt under ulike trafikk- og værforhold.

Forbedret trafikkflyt og redusert kødannelse gir direkte samfunnsøkonomiske gevinster (Kim, 2019) og er i tråd med FNs bærekraftsmål 8 og 9.

Begrensninger og videre arbeid

Denne studien har tre sentrale begrensninger. For det første påvirkes trafikken sterkt av det som skjedde kort tid i forveien. Modellen tar ikke fullt ut hensyn til slike sammenhenger, noe som kan gjøre resultatene noe mer optimistiske enn de ville vært i praksis. En mulig løsning er å inkludere informasjon om trafikkforholdene den foregående timen. For det andre viser alle modeller økt prediksjonsfeil ved ekstreme verdier av målvariabelen, det vil si i situasjoner der hastighetsavviket er enten veldig lavt eller

veldig høyt. Modellene er dermed minst pålitelige nettopp i situasjoner som ulykker, veiarbeid, ekstremvær, der presis prediksjon er mest kritisk, og dette er konsistent med annen litteratur der modeller er optimalisert for gjennomsnittsfeil (Ellis, 2024). Derfor bør beslutningstakere som bruker modellene til å identifisere bekymringsfulle strekninger ta høyde for at alvorlighetsgraden av ekstreme tilfeller systematisk underestimeres. For det tredje er dataene begrenset til kun ett kalenderår og kun 30 strekninger i Oslo og Akershus. Dette gjør at overførbarheten til andre regioner, særlig i Nord-Norge med mer ekstreme klimaforhold, ikke er testet.

KONKLUSJON

Studien viser at Random Forest og LightGBM møter alle forhåndsdefinerte suksesskriterier med en forklaringsgrad (R^2) på rundt 0,80 og en gjennomsnittlig feil på omtrent 4 % når de testes på norske riksveier i Oslo og Akershus. Resultatene viser samtidig at tradisjonelle maskinlæringsmetoder kan oppnå resultater på nivå med langt mer komplekse dyplæringsmodeller. Av modellene som er testet i denne studien vurderes LightGBM som den foretrukne modellen fremfor Random Forest.

Det viktigste funnet er at modellene fungerer langt bedre på veier de kjenner fra før enn på nye veier. Skal slike modeller brukes i større skala, trengs det mer detaljert informasjon om veienes egenskaper.

Tradisjonelle maskinlæringsmetoder er allerede i stand til å støtte de datadrevne målene i NTP 2025–2036 for forvaltning av kjente veistrekninger. For å ta teknologien i bruk i større skala trengs det imidlertid mer detaljerte data om veiene, mer data fra andre deler av Norge og videre utvikling av modellene.

KILDER

Ellis, H. (2024). To mean or not to mean: An investigation of regression to the mean [Honors research project, Williams Honors College]. IdeaExchange, Univer sity of Akron. Retrieved March 12, 2026, from https://ideaexchange.uakron. edu/honors_research_projects/1904

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer

Kim, J. (2019). Estimating the social cost of congestion using the bottleneck model. Economics of Transportation, 19, 100119. https://doi.org/10.1016/j.ecotra. 2019.100119

Pavlyuk, D. (2021). Spatiotemporal cross-validation of urban traffic forecasting mod els. Transportation Research Procedia, 52, 179–186. https://doi.org/10.1016/ j.trpro.2021.01.020

Razali, N. A. M., Shamsaimon, N., Ishak, K. K., Ramli, S., Mohamad Amran, M. F., &Sukardi, S. (2021). Gap, techniques and evaluation: Traffic flow prediction using machine learning and deep learning. Journal of Big Data, 8, 152. https: //doi.org/10.1186/s40537-021-00542-7

Samferdselsdepartementet & Nærings- og fiskeridepartementet. (2024). Nasjonal transportplan 2025-2036 (Meld. St. 14 (2023-2024)). Regjeringen Norge. Re trieved November 30, 2025, from https://www.regjeringen.no/globalassets/ departementene/sd/ntp/ntp-2025-2036/utredningsoppdraget-leveranse januar-2023/utredningsrapport-endelig-l2254365.pdf

Shaygan, M., Meese, C., Li, W., Zhao, X., & Nejad, M. (2022). Traffic prediction using artificial intelligence: Review of recent advances and emerging opportu nities. Transportation Research Part C: Emerging Technologies, 145, 103921. https://doi.org/10.1016/j.trc.2022.103921

Sweet, L.-B., Müller, C., Anand, M., & Zscheischler, J. (2023). Cross-validation strategy impacts the performance and interpretation of machine learning models. Artificial Intelligence for the Earth Systems, 2(4), e230026. https: //doi.org/10.1175/AIES-D-23-0026.1

World Health Organization. (2018). Global status report on road safety 2018. World Health Organization. https://www.who.int/publications/i/item/9789241565684

Zhou, Z., Yang, Z., Zhang, Y., Huang, Y., Chen, H., & Yu, Z. (2022). A compre hensive study of speed prediction in transportation system: From vehicle to traffic. iScience, 25(3), 103909. https://doi.org/10.1016/j.isci.2022.103909