



UiT Norges arktiske universitet

Handelshøgskolen ved UiT

Har maskinlæringsteknikker bra prediksjonsevne på nordiske indekser?

En komparativ studie av maskinlæringsteknikker for prediksjon av nordiske aksjemarkeder

Henning Reitan Brynjulfsen og Runar Reitan Brynjulfsen

Masteroppgave i økonomi og administrasjon, BED-3901, juni 2024

Forord

Denne masteravhandlingen i økonomisk analyse markerer slutten på vår toårige mastergrad og totalt femårige utdanningsreise ved Handelshøgskolen, UiT Norges Arktiske Universitet. Det har vært en reise preget av mange utfordringer og tidvis lange dager, men mest av alt en følelse av mestring og stolthet.

Gleden av å kunne anvende læringsstoff vi har lært over disse årene i en stor og krevende avhandling har vært lærerikt og gøy. Muligheten til å ta i bruk teorier og fagstoff på temaer vi brenner for og anvende dette på den virkelige verden, er vi takknemlig for. Arbeidet med avhandlingen har ikke bare styrket vår faglige kompetanse, men også vår evne til kritisk tenkning og problemløsning. Denne prosessen har gitt oss en dypere forståelse av økonomiske analyser og den krevende utfordringen med å gjøre dette anvendbart i den virkelige verden.

Vi ønsker videre å rekke en stor takk til vår veileder, Thomas Leirvik, som har vært en betydelig støtte gjennom hele perioden. Dine innsikter, faglige veiledninger og nyttige tilbakemeldinger har markant bidratt til å heve kvaliteten på denne oppgaven. Vi er svært takknemlig for samarbeidet og veiledningen vi har mottatt fra deg.

Til slutt ønsker vi også å takke våre to medstudenter, Terje Fagerli Pedersen og Steffen Skeimo, som har bidratt til økt motivasjon og sørget for at arbeidet har blitt utført jevnt og trutt. Selv om det tidvis har vært uproduktivt, ville vi ikke hatt det på noen annen måte. Deres støtte og vennskap har gjort hele prosessen minneverdig og morsom, og vi kommer til å se tilbake på denne tiden med stor glede.

Henning Reitan Brynjulfsen og Runar Reitan Brynjulfsen

Tromsø, 03. juni 2024

Sammendrag

Denne masteroppgaven undersøker hvorvidt maskinlæring kan utkonkurrere kjøp-og-hold-strategier i de nordiske aksjemarkedene. Arbeidet omfatter en komparativ studie av forskjellige maskinlæringsteknikker brukt til å forutsi aksjemarkedsbevegelser.

Problemstillingen er: "Har maskinlæringsteknikker en god prediksjonsevne på nordiske aksjeindekser med lengre intervaller?"

Studien benytter seg av maskinlæringsmodeller utledet i dataprogrammet R-Studio, ved bruk av pakkene "caret", "dplyr", "e1071", "ggplot2", "glmnet", "neuralnet", "pROC", "randomForest", "TTR" og "quantmod". Random forest, kunstige nevralt nettverk, support vector machine og en kombinasjonsmodell ble anvendt for å predikere retningen på de nordiske markedene som omfatter Norge (OBX 25), Sverige (OMXS30), Danmark (OMXC25) og Finland (OMXH25). Modellenes ytelse ble evaluert ved hjelp av portefølje-backtesting målt opp mot en kjøp-og-hold-strategi.

Studien viste at prediksjonsevnen til maskinlæringsteknikkene varierte mellom de ulike markedene. Random forest oppnådde best avkastning i det norske markedet med en positiv informasjonsrate, mens nevralt nettverk presterte best i den finske indeksen. Kombinerte modeller hadde generelt jevnere ytelse enn enkeltstående modeller, på tvers av indeksene i Norge, Sverige og Finland. For den danske indeksen klarte ingen av maskinlæringsmodellene å overgå kjøp-og-hold-strategien, noe som kan skyldes kompleksiteten og volatiliteten av mindre markeder.

Resultatene antyder at modeller som tar en mer optimistisk tilnærming til prediksjoner av oppgang, har vært mest vellykkede i perioder med økonomisk vekst. Oppgaven konkluderer med at maskinlæringsteknikkene random forest, kunstige nevralt nettverk og support vector machine har en variabel prediksjonsevne over lengre tidsintervaller i nordiske aksjeindekser. Prediksjonsmodellene viste seg mest effektive når de primært forutså markedsoppgang. Modellene som kombinerte flere metoder, viste seg å være mer stabile og pålitelige i sine prediksjoner sammenlignet med enkeltstående modeller. Funnene indikerer at videre forskning bør fokusere på modelljustering og dybdeanalyse av markedsdynamikk for å forbedre prediksjonsnøyaktigheten ytterligere.

Nøkkelord: Finans, Maskinlæring, Nordiske aksjemarkeder, Prediksjon, Modellering

Innholdsfortegnelse

1	Innledning	1
1.1	Problemstilling	2
2	Teori	2
2.1	Teorier om prisutviklinger i finansmarkedet	2
2.2	Makroøkonomisk faktormodell	4
2.3	Sharpe-ratio	5
2.4	Informasjonsraten	6
3	Maskinlæring	7
3.1	Empiriske studier maskinlæring	7
3.1.1	Generelle empiriske studier maskinlæringsmetoder	7
3.1.2	Empiriske studier maskinlæringsmetoder for Norge	8
3.1.3	Empiriske studier maskinlæringsmetoder for Sverige	8
3.1.4	Empiriske studier maskinlæringsmetoder for Finland	9
3.1.5	Empiriske studier maskinlæringsmetoder for Danmark	9
3.2	Maskinlærings rolle og funksjon	10
3.2.1	Veiledet læring	11
3.3	LASSO-regresjon	11
3.4	Random Forest	12
3.5	Nevrale Nettverk	13
3.6	Support vector machine	14
3.7	Ytelsesmål i maskinlæring – Forvirringsmatrise	15
4	Metode	17
4.1	Datainnsamling	17
4.2	Dataforberedelse	18
4.2.1	Variabelseleksjon	20
4.3	Modellspekifikasjoner	21
4.3.1	Random Forest	21
4.3.2	Nevrale nettverk	22
4.3.3	Support vector machine	22
4.4	Gjennomføring av portefølje-backtesting	23
4.5	Evaluering av avkastningsdriverne	24
5	Resultater	25
5.1	Resultater random forest-modeller	28
5.1.1	Klassifiseringsresultater: Random forest-modeller	28
5.1.2	Ytelse RF modeller	29
5.1.3	Portefølje 1 – Backtesting Random forest modell	30
5.2	Resultater Nevrale nettverksmodeller	31
5.2.1	Klassifiseringsresultater: Nevrale nettverksmodeller	31
5.2.2	Ytelse Nevrale nettverksmodellene	32
5.2.3	Portefølje 2 – Backtesting nevrale nettverksmodell	33
5.3	Resultater support vector machine-modeller	34

5.3.1	Klassifiseringsresultater: Support vector machine-modeller	34
5.3.2	Modellytelse support vector machine.....	35
5.3.3	Portefølje 3 – Backtesting support vector machine-modellene.....	36
5.4	<i>Resultater kombinasjonsmodeller</i>	<i>37</i>
5.4.1	Portefølje 4 – Backtesting kombinasjonsmodeller	37
5.5	<i>Multipelregresjon</i>	<i>39</i>
6	Konklusjon.....	40
6.1	<i>Begrensninger/Svakheter med oppgaven</i>	<i>42</i>
6.2	<i>Videre forskning.....</i>	<i>42</i>
	Referanseliste.....	44
	Vedlegg.....	47
	<i>Vedlegg A: Output multipelregresjon.....</i>	<i>47</i>

Figur liste

Figur 1.	Illustrasjon av en forvirringsmatrise	16
Figur 2.	Graf av portefølje-utviklingene for Norge.....	25
Figur 3.	Graf av portefølje-utviklingene for Sverige	26
Figur 4.	Graf av portefølje-utviklingene for Finland	27
Figur 5.	Graf av portefølje-utviklingene for Danmark.....	27
Figur 6.	Forvirringsmatrise for random forest-modellene	28
Figur 7.	Forvirringsmatrise for nevrale nettverksmodellene.....	32
Figur 8.	Forvirringsmatrise for support vector machine-modellene	35

Tabelloversikt

Tabell 1.	Ytelse for random forest-modellene basert på forvirringsmatrisene	29
Tabell 2.	Prestasjon portefølje-backtesting for random forest-modellene	30
Tabell 3.	Ytelse for nevrale nettverksmodellene basert på forvirringsmatrisene.....	32
Tabell 4.	Prestasjon portefølje-backtesting for de nevrale nettverksmodellene.....	33
Tabell 5.	Ytelse for support vector machine-modellene basert på forvirringsmatrisene	35
Tabell 6.	Prestasjon portefølje-backtesting for support vector machine-modellene	36
Tabell 7.	Prestasjon portefølje-backtesting for kombinasjonsmodellene.....	38
Tabell 8.	Signifikante variabler basert på multipelregresjon	39

1 Innledning

De siste årene har den gjennomsnittlige personens interesse for aksjemarkedet økt betydelig, noe som har resultert i en eksponentiell vekst med flerfoldige milliarder i omløp (Badolia, 2016). Aksjemarkedene er dynamiske, og mange aktører forsøker kontinuerlig å utvikle og finne metoder for å oppnå konkurransefortrinn. I en verden med akselererende informasjonsflyt og stadige teknologiske fremskritt, spesielt innen kunstig intelligens, er det rimelig å anta at nye metoder og hjelpemidler vil oppstå.

Teknologiske fremskritt påvirker finansmarkedene, og en sentral del av denne utviklingen er maskinlæring, en underkategori av kunstig intelligens. Maskinlæring har unike egenskaper som har gjort det mulig å effektivisere oppgaver som tidligere var både tidskrevende og kostbare. Eksempelvis har det blitt vesentlig billigere å gjennomføre prediksjon: evnen til å bruke tilgjengelig informasjon for å generere ny informasjon som tidligere var utilgjengelig (Agrawal et al., 2017).

Maskinlæringsteknikker anvendes hyppig for prediksjon i aksjemarkeder globalt. Disse teknikkene brukes primært på store aksjemarkeder som New York Stock Exchange (NYSE), Tokyo Stock Exchange (TSE) og Shanghai Stock Exchange (SSE). Det er ikke bare markedsverdien som gjør at disse markedene er mest analyserte, men også deres betydning og innflytelse på den globale økonomien. Tilgjengeligheten av store mengder kvalitative og kvantitative data med lang historie gjør også disse markedene mer tilgjengelig for analyse.

Mindre aksjemarkeder, som de nordiske, kan ha mindre data tilgjengelig. Dette kan imidlertid gi muligheter for at gode modeller kan gi en investor en konkurransefordel, noe som igjen kan føre til større nytteverdi av prediksjonsresultatene. I motsetning til større markeder, hvor datamengden kan være overveldende, kan de nordiske markedene tilby en unik utfordring som krever mer sofistikerte og tilpassede modeller.

Når det gjelder predikering basert på kunstig intelligens og maskinlæring, finnes det en rekke metoder som kan anvendes. Gitt det brede spekteret av tilgjengelige modeller og teknikker, vil denne studien benytte flere av de mest anerkjente metodene basert på empiriske data, og evaluere deres evne til å forutsi aksjemarkedsbevegelser i nordiske aksjemarkeder.

Målet med denne oppgaven er å utvikle en kompleks modell for hver av de valgte maskinlæringsteknikkene. Modellene vil bli tilpasset de ulike landsindeksene ved hjelp av

hyperparameter-tuning, for å sikre optimal ytelse. Dette vil gjøre det mulig å vurdere hvorvidt modellene er anvendelige på tvers av forskjellige markeder, eller om de er begrenset til markeder med spesifikke markedsforhold. Dette danner grunnlaget for problemstillingen.

1.1 Problemstilling

For å undersøke i hvilken grad maskinlæringsteknikker kan lykkes med predikering av utviklingen til mindre indekser, samt skille oss ut ved bruk av månedlige data, har vi formulert følgende problemstilling:

"Har maskinlæringsteknikker en god prediksjonsevne på nordiske aksjeindekser med lengre intervaller?"

For å svare på problemstillingen, utvikles prediksjonsmodeller basert på ulike maskinlæringsteknikker. Hvordan maskinlæringsteknikkene presterer er sentralt. I den sammenhengen har vi utledet tre forskningsspørsmål som skal gjøre oss i stand til å gi en grundig vurdering av prediksjonsevnen til modellene:

1. "Hvordan varierer prediksjonsevnen til forskjellige maskinlæringsteknikker blant de ulike indeksene?"
2. "Forbedres prediksjonsevnen gjennom å kombinere flere maskinlæringsteknikker?"
3. "Slår prediksjonsmodellene en kjøp-og-hold-strategi?"

2 Teori

2.1 Teorier om prisutviklinger i finansmarkedet

Hypotesen om effisiente markeder sier et marked er effisient når prisen på aksjer og gjeldsinstrumenter reflekterer all tilgjengelig informasjon om verdipapirene (Fama, 1970, p. 383). Teorien hevder at all relevant informasjon er reflektert i prisen, som fører til at det er umulig å konsekvent oppnå en avkastning som overgår markedet ved å basere seg på historisk informasjon. Hypotesen har betydelige implikasjoner for prediksjonsmodeller som baserer seg på tidligere prisinformasjon, samt utfordrer ideen om regelmessig profitt ved å utnytte markedsfluktuering.

Effisiente markeder deles inn i tre former: svak-, halvsterk- og sterk effisiens. Den svake formen, hevder at all historisk informasjon allerede er reflektert i prisen og at tidligere kurser eller priser ikke har noen betydning for fremtidig utvikling. Halvsterk effisiens sier at all

offentlig tilgjengelig informasjon umiddelbart blir priset inn i et verdipapir. Sterk effisiens er mest rigid i sin tolkning og sier at all informasjon, inkludert privat og utilgjengelig for allmenheten, allerede er reflektert i prisen til et verdipapir. Dette impliserer at verken aksjeinvestorer eller insidere kan oppnå kontinuerlig gevinst basert på informasjon som ikke er tilgjengelig for allmenheten (Fama, 1970, p. 383).

Andre teorier som samsvarer med markedseffisiens er random walk- (RWT) og monkey theory (MT). Disse påpeker at tidligere priser ikke gir nyttig informasjon som kan brukes til å forutsi fremtidige markedsutviklinger ettersom prisene utvikler seg tilfeldig (Malkiel, 1985). Malkiel (1985) anbefaler en tilnærming basert tilfældige kjøp kombinert med langsiktig eierskap av aksjer, og argumenterer for mer effektivitet enn ved mer komplekse analytiske tilnærminger. MT poengterer markedets uforutsigbarhet, og argumenterer mot profesjonelle investorer og analyser gjennom sitatet: «A blindfolded monkey throwing darts at a newspaper's financial pages could select a portfolio that would do just as well as one carefully selected by the expert» (Malkiel, 1985).

Dersom en legger nevnte teorier til grunne, er oppgavens problemstilling nytteløs. Det har derimot kommet en rekke kritiske svar, deriblant Shostak, Lee, Prasch og Lo & MacKinlay (Lee, 2006; Lo & MacKinlay, 1988; Prasch, 2011; Shostak, 1997). Forutsetningen at alle markedsdeltakere kommer frem til rasjonelle forventinger impliserer at alle deltakere antas å ha de samme forventningene om fremtidige avkastninger. Dersom alle markedsaktører har de samme forventinger, hva motiverer gjennomføringen av transaksjoner, siden heterogene forventinger blant kjøpere og selgere danner grunnlaget for deltagelse i handel av verdipapir (Shostak, 1997, p. 29).

Forutsetningen at markedsdeltakere har homogen kunnskap og informasjon over simplifiserer kompleksiteten i virkelige markedsdynamikker. Selv om moderne teknologi gir lik tilgang på nyheter betyr det ikke at informasjonsflyten alltid er jevn, eller at investorers tolkning er lik. Teoriene impliserer at forventinger er konsentrert rundt den sanne verdien, og avvik fra verdien er tilfeldig, noe som innebærer at profitter eller tap er tilfeldige hendelser. Dette vil si at alle kjenner den sanne verdien, og derfor er læring av feil helt meningsløst. Dersom individer har ulik kunnskap, vil det påvirke prognosene deres (Shostak, 1997, p. 29).

2.2 Makroøkonomisk faktormodell

I en makroøkonomisk faktormodell er faktorene uventede endringer (overraskelser) i makroøkonomiske variabler, som har en betydelig innvirkning på avkastningene. For eksempel, i aksjemarkedet kan disse faktorene påvirke selskapenes forventede fremtidige kontantstrømmer eller rentesatsen som anvendes for å neddiskontere kontantstrømmene til dagens verdi. Typiske makroøkonomiske faktorer som benyttes i slike modeller inkluderer rentesatser, inflasjonsrisiko, forretningsrisiko, og kredittspredninger (Pinto & Podkaminer, 2019)

I makroøkonomiske faktormodeller antas det at avkastningen på aksjeindekser er korrelert kun med overraskelser i makroøkonomiske faktorer, som eksempelvis inflasjon eller reell produksjon. Overraskelse i denne sammenhengen defineres som differansen mellom den faktiske verdien av en økonomisk faktor og den tidligere forventede verdien av samme faktor. En faktoroverraskelse representerer dermed den delen av faktorens avkastning som var uventet (Pinto & Podkaminer, 2019). Disse faktoroverraskelsene fungerer som uavhengige variabler i modellen. Denne tilnærmingen kontrasteres med en annen modellpresentasjon der de uavhengige variablene er selve avkastningene, som illustrert i en alternativ modellformulering. Dette understreker variasjonen i hvordan uavhengige variabler kan representeres i forskjellige økonomiske modeller. Gitt at flere makrofaktorer, si K faktorer, forklarer avkastningene på aksjeindekser, kan en makroøkonomisk faktormodell uttrykkes som:

$$R_I = a_I + b_{I1}F_1 + b_{I2}F_2 + \dots + b_{IK}F_K + \varepsilon_I \quad (1)$$

hvor:

R_I = avkastning til indeks I

a_I = den forventede avkastningen til indeks I

b_{IK} = følsomheten til avkastningen for indeks I for en overraskelse i faktor k, $k=1, 2, \dots, K$

F_K = overraskelsen i faktor k, $k=1, 2, \dots, K$

ε_I = et feilledd med null gjennomsnitt som representerer den delen av avkastningen til indeks I som ikke forklares av faktormodellen.

Hvis man analyserer månedlige avkastninger for aksjer, kan en makroøkonomisk overraskelse illustreres ved å undersøke ulike sektorer og markedsindekser. For eksempel, ved månedens start kan analytikere forvente at oljepriser holder seg stabile på 70 dollar per fat, basert på økonometriske modeller, geopolitiske hendelser og produksjonsdata. Ved månedens slutt kan

uventede geopolitiske spenninger ha drevet prisen opp til 75 dollar per fat. Denne økningen på 5 dollar kan føre til høyere avkastning for indekser med mange energiselskaper, på grunn av bedre enn forventede inntekter.

Aksjeavkastninger kan brytes ned i tre komponenter: Forventet avkastning, som reflekterer effekten av de forutsette makroøkonomiske verdiene, uventet avkastning fra nye opplysninger avdekket gjennom månedens overraskelser, og et feilledd som håndterer uforutsette faktorer (Pinto & Podkaminer, 2019). Vår modell er justert for risikofri avkastning for å gi en mer nøyaktig forståelse av hvordan meravkastningen til investeringen skapes. Den spesifikke formelen for vår multippelregresjon:

$$r_{p,t} - r_f = \alpha + \beta_1 \cdot (r_{energi,t} - r_f) + \beta_2 \cdot (r_{gull,t} - r_f) + \beta_3 \cdot (r_{biotek,t} - r_f) + \beta_4 \cdot (r_{industri,t} - r_f) + \beta_5 \cdot (r_{finans,t} - r_f) + \beta_6 \cdot (r_{s\&p500,t} - r_f) + \beta_7 \cdot (r_{vix,t} - r_f) + \varepsilon_I \quad (2)$$

Hvor:

- $r_{p,t}$ er avkastningen på porteføljen i periode t
- r_f er risikofri rente
- α er den gjennomsnittlige avkastningen utover det som forklares av modellen
- β_i er sensitiviteten til porteføljen mot faktor i
- $r_{faktor,t}$ er avkastningen for hver faktor i periode t
- ε_I er en feilledd som representerer delen av avkastningen til eiendel I som ikke kan forklares av faktormodellen

2.3 Sharpe-ratio

Risikjustert avkastning kan beregnes ved bruk av sharpe-ratio (SR). SR beregner forskjellen mellom avkastningen på en investering og den risikofrie renten, før dette deles med standardavviket for investeringen (Muralidhar, 2015). SR anvendes for å måle hvor effektivt avkastningen fra en investering kompenserer investoren for den pådratte risikoen, sett i forhold til en risikofri investering. Formelen for SR er:

$$S_p = \frac{r_p - r_f}{\sigma_p} \quad (3)$$

Hvor:

- S_p er sharpe-raten til porteføljen p
- r_p er avkastningen til porteføljen p
- r_f er den risikofrie renten

σ_p er standardavviket til porteføljen p

Generelt sett foretrekkes en høyere SR, men som Muralidhar (2015) påpeker, er det nødvendig med et sammenligningsgrunnlag for å forstå dens verdi fullt ut (Muralidhar, 2015). Dette gjør det mulig å avgjøre om den risikojusterte avkastningen skyldes høy volatilitet og risiko, eller om andre faktorer spiller inn. Ved å bruke SR kan vi måle hvor mye ekstra avkastning vi får for hver enhet av risiko vi tar, noe som gir oss en dypere forståelse av effektiviteten til våre prediksjonsmodeller. Dette gjør det også mulig å sammenligne deres risikojusterte avkastning både med tradisjonelle investeringsmetoder og referanse-indekser, noe som er viktig for å bedømme modellenes prestasjon under ulike markedsforhold.

2.4 Informasjonsraten

Informasjonsraten (IR) er et nøkkelmål i porteføljeforvaltning som sammenfatter en aktiv porteføljes gjennomsnittlige meravkastning over en referanseindeks justert for risiko. Dette prinsippet er basert på Harry Markowitz forskning innen porteføljeteori, hvor spesielt avkastning og varians ble fremhevet som essensielle elementer for å evaluere prestasjonen til en investering (Goodwin, 1998). Informasjonsraten kan uttrykkes som:

$$IR = \frac{\alpha_p}{\sigma(e_p)} \quad (4)$$

Hvor $\sigma(e_p)$ er:

$$\sigma(e_p) = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (e_t - \bar{e})^2} \quad (5)$$

Hvor

- α_p er porteføljens meravkastning i forhold til markedsporteføljen
- $\sigma(e_p)$ er den usystematiske risikoen til en portefølje
- e_t er meravkastning for en periode
- \bar{e} er gjennomsnittlig meravkastning
- T er antall perioder

En høyere informasjonsrate indikerer at den aktive porteføljen har oppnådd en risikojustert meravkastning gjennom sine investeringsstrategier. Dette antyder at potensialet for å skape α , eller meravkastningen utover referanseindeksen med risikojustering, er betydelig (Goodwin, 1998).

3 Maskinlæring

3.1 Empiriske studier maskinlæring

Maskinlæring for predikering av aksjemarkedet er mye brukt i eksisterende litteratur. Denne studien bruker flere av de tidligere for å forbedre prediksjonene. Videre introduseres analyse av tidligere studier for å identifisere anbefalinger, samt styrker og svakheter ved ulike metoder og fremgangsmåter. De empiriske studiene danner fundamentet for valg av maskinlæringsmetoder både ved bruk av generelle anbefalinger og landsspesifikke.

3.1.1 Generelle empiriske studier maskinlæringsmetoder

Sheth og Shah (2023) gjennomførte en meta-analyse der de undersøkte hvilke maskinlæringsmetoder blant kunstig nevralt nettverk, support vector machine (SVM) og long short-term memory (LSTM) som var de beste- og mest nøyaktige metodene for å predikere fremtidige aksjepriser. Meta-analysen fant ut at de nevralt nettverkene ga beste resultater, hovedsakelig som følger av evnen til å gjenkjenne mønstre og håndtere ikke-lineære relasjoner (Sheth & Shah, 2023). SVM og LSTM presterte bra ved prediksjon av store datasett, men hadde fallende ytelse ved mindre datasett.

Singh (2022) brukte åtte ulike maskinlæringsteknikker til prediksjoner, deriblant lineær regresjon, random forest, adaboost, support vector machine og kunstige nevralt nettverk. Resultatene viste at lineær regresjon og nevralt nettverk presterte nesten identisk, men at nevralt nettverk krevde mer tid til trening og validering. Support vector machine hadde høy effektivitet på større datasett, i likhet med random forest og adaboost som forbedret ytelsen med økende datamengde, i motsetning til nevralt nettverk som underpresterte ved større datasett (Singh, 2022).

Krauss et al. (2017) brukte maskinlæringsteknikkene dype nevralt nettverk (DNN), gradient-boostede trær (GBT) og random forest (RF) for å se på effektiviteten i statistisk arbitrasje. I denne studien viste ensembletilnærmingen som kombinerte forskjellige modeller høyere ytelse enn enkeltmodeller (Krauss et al., 2017). Selv om resultatene avtok etter hvert i analysert periode viste kombinasjonsmodellen potensiale under visse markedsforhold. Dette indikerer en diversifisert kombinasjon av metoder kan være en effektiv strategi ved usikkerhet om modellvalg og tilgjengelig databehandlingskapasitet.

3.1.2 Empiriske studier maskinlæringsmetoder for Norge

Aronsen & Markussen (2023) publiserte en avhandling der random forest, support vector machine og logistisk regresjon forsøkte å predikere det norske aksjemarkedet. Resultatene viste at random forest-modeller hadde spesielt god prestasjon og konsekvent utkonkurrerte andre modeller frem til 2020. Den mest effektive porteføljen basert på random forest oppnådde betydelig avkastning før transaksjonskostnader, og viste seg økonomisk fordelaktig før markedsendringene inntrådte i 2020 som konsekvens av Covid-19. Studiet til Aronsen og Markussen avdekket at selv om maskinlæringsporteføljene opprinnelig hadde en fordel, mistet de sin prediktive kraft etter pandemien, noe som indikerte en nedgang i modellenes evne til å tilpasse seg nye markedsforhold (Aronsen & Markussen, 2023).

I 2016 undersøkte Olden potensialet for å oppnå lønnsom aksjehandel ved bruk av maskinlæringsteknikker gjennom sammenligning av binære klassifiseringsalgoritmer og deres prestasjon. Særlig ble en 'Stacked Ensemble Learning-algoritme' utviklet, som integrerte prediksjoner fra andre algoritmer som ekstra funksjoner. Studiet viste at de beste modellene overgikk Oslo Børs' Benchmark Index (OBX), men utfordringer med testperioden og generell aksjeprediksjon begrenset konklusjoner om et langsiktig lønnsomt handelssystem. Det var ingen bevis for at stacked ensemble learning-algoritmer presterte bedre enn andre teknikker. Selv om enkelte algoritmer var lønnsomme i testperioden, var det utfordringer med å forutsi aksjemarkedet og vurdere representativiteten av testperioden. Resultatene antydte at visse maskinlæringsalgoritmer kan være lønnsomme under gitte forhold, men ytterligere forskning er nødvendig (Olden, 2016).

3.1.3 Empiriske studier maskinlæringsmetoder for Sverige

I 2023 gjennomførte Wik en studie for å forutsi én-dags avkastning på den svenske aksjeindeksen og to store svenske børsnoterte aksjer. Mens prediksjonene for indeksen var vellykkede, var resultatene for enkeltaksjene mindre imponerende. Studien utforsket også prediksjon av svenske aksje- og indekspriser ved hjelp av ulike maskinlæringsalgoritmer, inkludert lineær regresjon, k-nearest neighbors, random forest, support vector machine og nevrale nettverk, basert på historiske priser fra ledende aksjer og indekser på tvers av ulike sektorer. Resultatene viste at lineær regresjon, support vector machine og nevrale nettverk presterte best, sannsynligvis på grunn av deres robusthet mot store prisbevegelser. Lineær regresjon og support vector machine ble fremhevet for deres tidsøkonomiske trening kombinert med ytelse på nivå med komplekse nevrale nettverk. Random forest presterte

svakere over tid, og oppnådde verken meravkastning eller viste treningseffektivitet. Studien påpekte også at alle modeller led av overtilpasning, sannsynligvis på grunn av modellkompleksitet eller valg av attributter (Wik, 2023).

3.1.4 Empiriske studier maskinlæringsmetoder for Finland

Määttä (2020) undersøkte hvordan nevrale nettverk av typen lang korttidsminne (LSTM) gjorde det ved predikering av aksjeavkastningsbevegelser. Studien sammenlignet effektiviteten av LSTM-nettverk med konvensjonelle rekurrente nevrale nettverk og logistisk regresjon, ved bruk av to ulike sett med inngangsfunksjoner. Resultatene viste at LSTM-modellene konsekvent overgikk de andre metodene, med den mest effektive multivariate LSTM-modellen som den mest økonomisk fordelaktige blant de testede porteføljene (Määttä, 2020). Studien avdekket at LSTM-nettverkene beholdt sin fordel gjennom hele perioden, men lønnsomheten var hovedsakelig fra første halvdel av testperioden. Videre analyse viste at LSTM-modellene kunne identifisere mønstre som likner på kjente kapitalmarkedsanomalier, som kortvarig gjennomsnittsreversering og momentum. Forfatteren antyder at aksjeavkastningene delvis er drevet av langsiktige signaler, som LSTM-nettverkene effektivt kan trekke ut fra støyende markedsdata, og dermed utfordrer den tradisjonelle markedseffisiens-teorien i det finske markedet. Til sammenligning viste studien at en 'kjøp-og-hold'-strategi var mindre effektiv i den samme perioden.

3.1.5 Empiriske studier maskinlæringsmetoder for Danmark

En studie av Pedersen og Lyngø (2023) utforsket bruken av LSTM- og LSTM-X-modeller for prognose av aksjepriser, med fokus på enkeltaksjer i det danske markedet representert ved OMXC25-indeksen. Studien analyserte hvordan tilpasninger i modellarkitektur og integrering av eksterne makroøkonomiske variabler påvirker prognosenøyaktigheten under markedsureligheter. Spesielt undersøkte de hvordan ulike datalengder påvirket modellens evne til å forutsi volatilitet, med Novo Nordisk som et nøkkelfelle hvor LSTM-modellen utmerket seg sammenlignet med andre aksjer (Pedersen & Lyngø, 2023). Dette antyder at LSTM-modeller kan være spesielt nyttige for å forutsi prissvingninger under spesifikke økonomiske forhold. Selv om ingen av modellene hadde overlegen prediksjonsevne generelt, viste LSTM-modellen potensial for forbedret nøyaktighet for Novo Nordisk, noe som antyder behovet for tilpassede modeller for individuelle aksjer. Pedersen og Lyngøs forskning gir viktige innsikter i bruk av avanserte prognosemodeller på det danske aksjemarkedet, og

understreker betydningen av kontinuerlig forskning og utvikling for å forbedre prediktiv nøyaktighet og pålitelighet, spesielt for selskaper som Novo Nordisk.

Den siste empiriske studien vi undersøkte var av Jørgensen et al. (2021) som predikerte aksjepriser i Danmark ved hjelp av kunstige nevralt nettverk (KNN), inkludert modeller som feed-forward nevralt nettverk (FNN), rekurrente nevralt nettverk (RNN) og long short-term memory (LSTM). Også denne studien undersøkte seks enkeltaksjer, men for perioden 2010 til 2019, med de første åtte årene for trening, og det siste året for testing av prognosene. Resultatene viste at selv om enkelte FNN-strukturer fremviste noe nøyaktighet, ble de slått av ARIMA(X)-modeller for de fleste av de seks analyserte aksjene (Jørgensen et al., 2021). Tillegg av ekstra variabler førte til reduksjon av KNN-modellenes ytelse. Funnene antyder at det danske aksjemarkedet oppfører seg som en tilfeldig vandring og dermed underbygger random walk teorien, som reflekteres gjennom effektiviteten til de beste KNN-modellene. Også dette studiet påpeker behov for videre forskning før nevralt nettverks-modeller kan anses som en pålitelig metode i tilfeldige og komplekse markeder som det danske markedet virker å være.

3.2 Maskinlærings rolle og funksjon

Maskinlæring utgjør et bredt spekter av verktøy som gir datamaskiner evnen til å handle og tenke på intelligente måter. Som undergruppe av kunstig intelligens, konsentrerer maskinlæring seg om å forstå data gjennom å identifisere årsakssammenhenger og mønstre, og brukes til å forutsi fremtidige hendelser. Dette innebærer at algoritmer utvikler evnen til å ta informerte beslutninger ved å lære fra eksempler. Algoritmer analyserer informasjon for å gjenkjenne gjentakende mønstre og anvender denne innsikten for å gjøre informerte prediksjoner om fremtidige utfall (DeRose et al., 2019).

Maskinlæring kan påvirke investeringsstyring ved å bidra til risikostyring og prediksjon av aksjereturer. Maskinlæring muliggjør presise prediksjoner av aksjeavkastninger ved å utføre tekstuell analyse av finansielle rapporter (DeRose et al., 2019). Ved å identifisere signifikante endringer i ledelsesdiskusjon og risikoseksjoner, kan maskinlæring utvinne verdifull innsikt som kan påvirke aksjekjøp og -salg, og dermed gi en strategisk fordel i markedet. Disse teknikkene kan styrke både presisjonen og effektiviteten i porteføljestyling og handelsstrategier (DeRose et al., 2019). Maskinlæring kategoriseres hovedsakelig i tre klasser: Veiledet læring, ikke-veiledet læring, og forsterket læring. Vår analyse vil konsentrere seg om veiledet læring.

3.2.1 Veiledet læring

Veiledet læring innebærer å lære algoritmer hvordan man fra inndata (X) kan forutsi en ønsket utdata (Y), ved å finne mønstre som binder disse sammen. Dette krever et datasett med både inndata og korresponderende utdata, kjent som et merket datasett. Gjennom treningsprosessen tilpasser algoritmen seg for å modellere forholdet mellom X og Y slik at den kan forutsi Y for nye, uobserverte X (DeRose et al., 2019). I denne sammenhengen kalles de avhengige variablene for mål, mens de uavhengige variablene fungerer som egenskaper. Evalueringen av en veiledet maskinlæringsmodell skjer ved å sammenligne modellens prediksjoner med faktiske data, slik at modellens nøyaktighet og evne til generalisering kan vurderes.

Hovedformene for veiledet læring er regresjon og klassifisering. Mens regresjon behandler kontinuerlige målvariabler, anvender klassifisering kategoriske eller ordinale målvariabler. Klassifisering bruker maskinlæringsalgoritmer til å forutsi om aksjekurser vil stige eller falle. Prosessen omfatter analyse av historiske data og markedsindikatorer for å utvikle en modell som identifiserer to mulige utfall: 'opp' eller 'ned'. Dette defineres som et binært klassifiseringsproblem. Nøyaktige forutsigelser av prisbevegelser er kritiske. En pålitelig modell støtter utviklingen av investeringsstrategier ved å tilby grunnlag for informerte beslutninger (DeRose et al., 2019).

3.3 LASSO-regresjon

LASSO-regresjon, introdusert av Tibshirani (1996), er en statistisk metode som anvendes for å forbedre prediksjonene av en modell og utfører variabelutvalg (Tibshirani, 1996). "Least absolute shrinkage and selection operator", forkortet LASSO, uttrykkes som:

$$\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j \chi_{ij})^2 + \lambda \sum_{j=1}^p |\beta_j| = RSS + \lambda \sum_{j=1}^p |\beta_j| \quad (6)$$

Hvor:

- λ er en konstant som kontrollerer styrken av straffen som pålegges koeffisientenes størrelse.
- $\sum_{j=1}^p \beta_j$ er summen av de absolutte verdiene av regresjonskoeffisientene, kjent som L1-straffen.
- β_0 er skjæringspunktet
- β_j er stigningstallet for j -prediktoren

- χ_{ij} er verdien av j-prediktoren og i observasjonen.
- y_i er den faktiske verdien av responsvariabelen for i observasjonen.
- RSS er summen av kvadrerte ulikheter mellom de observerte responsverdiene og de estimerte responsverdiene

LASSO fungerer ved å legge til en straffeledd i regresjonsmodellen som er proporsjonal med summen av de absolutte verdiene av regresjonskoeffisientene, og denne straffen styres av en parameter λ . Når λ er satt til en verdi som ikke er for lav, vil metoden tvinge noen av koeffisientene til å bli nøyaktig null. Dette betyr at LASSO effektivt velger bort mindre viktige variabler, siden de ikke bidrar til modellen etter straffe-prosessen, og dermed oppnås variabelseleksjon. Denne metoden er spesielt verdifull i situasjoner hvor det finnes et stort antall mulige forklaringsvariabler, hvor antallet observasjoner er begrenset.

LASSO bidrar til å øke nøyaktigheten av prediksjoner og gjør modeller enklere å forstå ved å innlemme de beste egenskapene fra både ridge-regresjon og valg av delmengder av variabler. Når prediktorene er sterkt korrelerte, velger LASSO ut en forklaringsvariabel og reduserer de øvrige til ingenting. Dette fører til at variasjonen i de estimerte koeffisientene minker, siden noen koeffisienter faktisk blir satt til null, hvilket resulterer i modeller som er enklere å tolke (Muthukrishnan & Rohini, 2016).

3.4 Random Forest

Random forest er en ensemble-teknikk som forbedrer nøyaktighet og robusthet ved å kombinere mange beslutningstrær. Metoden bruker bagging, en type bootstrap hvor forskjellige treningsdatasett skapes ved tilfeldig utvalg med erstatning fra det opprinnelige datasettet. Hver modell trenes på sitt datasett, og deres prediksjoner samles gjennom flertallsstemme for klassifisering for en endelig prediksjon (DeRose et al., 2019).

Random forest-algoritmen øker variasjonen i beslutningstrær ved å introdusere tilfeldig utvalg av egenskaper ved hvert splittpunkt. Denne metoden begrenser trærnes avhengighet av dominerende forklarende variabler ved kun å vurdere en tilfeldig undergruppe av egenskaper for hvert skille. Dette fremmer mangfold blant trærne og sikrer at beslutninger baseres på ulike egenskapskombinasjoner, noe som reduserer risikoen for at dominerende egenskaper forvrenger hele modellens resultat (James et al., 2021).

Gjennom en tilfeldig seleksjon av datapunkter og variabler fra datasettet, danner algoritmen en diversifisert 'skog' av beslutningstrær. Hvert tre trenes på sin egen bootstrappede prøve og deres prediksjoner kombineres for økt nøyaktighet. Modellens presisjon kan ytterligere forbedres ved å justere hyperparametere som antall trær, antall funksjoner ved hver splitting, minimum antall prøver for en splitting, og maksimal dybde for hvert tre (DeRose et al., 2019).

I random forest samler hver trestamme sine uavhengige prediksjoner for å danne en felles endelig beslutning. I klassifiseringsaker velges den kategori som får flest stemmer fra alle trærne. Denne prosessen drar nytte av innsikt fra flere modeller, noe som minsker risikoen for overtilpasning og øker den samlede prediksjonsnøyaktigheten (Speiser et al., 2019).

Random forest er også en effektiv metode for variabelutvalg, som utføres i to trinn. Først rangeres variabler etter deres viktighet i random forest, og deretter velges de mest informative variablene ut gjennom en stegvis prosess. Denne metoden fremmer et nøye utvalg av variabler, som forbedrer både modellens enkelhet og prediksjonsevne, spesielt i datasett med mange dimensjoner (Genuer et al., 2010).

3.5 Nevrale Nettverk

Kunstige nevrable nettverk (KNN) er en overordnet kategori innen maskinlæring som tar utgangspunkt i datamodeller som er inspirert av den menneskelige hjernens nettverk av nevroner. De er designet for å etterligne måten hjernen behandler og tolker kompleks informasjon på. Et KNN består av flere neuroner eller noder arrangert i lag: et inputlag, et eller flere skjulte lag, og et outputlag. Neuronene i hvert lag er forbundet med vektene som representerer styrken på forbindelsene mellom dem (Abiodun et al., 2018).

Videre har KNN to underkategorier av nettverk-arkitektur og oppbygging basert på samhandlingen mellom neuronene: Rekurrente nevrable nettverk (RNN) og feed-forward nevrable nettverk (FNN). Hvis neuronene kan behandle informasjon i ulike retninger og er i stand til å sende tilbakemeldingssignaler til hverandre, klassifiseres de som en del av rekurrente nevrable nettverk (RNN) (Sazlı, 2006). Hvis neuronene kun behandler informasjon i én retning, havner de derimot inn under kategorien feed-forward nevrable nettverk (FNN). FNN har videre to undergrupper som baserer seg på antall skjulte lag, enkeltlag og multilag. Enkeltlags-FNN har kun ett skjult lag mellom input- og output-lagene, mens multilags-FNN består av flere skjulte lag, noe som muliggjør mer komplekse beregninger og læringsprosesser.

Læring i et KNN skjer gjennom en prosess kjent som backpropagation, hvor nettverket justerer sine vekter basert på feilen mellom den ønskede utgangen og nettverkets faktiske utgang. Denne prosessen innebærer å optimalisere nettverkets vekter for å minimere en tapfunksjon, som ofte er kryssentropi i klassifiseringssammenhenger. Backpropagation sikrer at hver vekt bidrar til å redusere den totale feilen i nettverkets prediksjoner over tid (James et al., 2021). En læringsheuristikk som baserer seg på dette er 'Resilient backpropagation (RPROP)' og er en algoritme som forbedrer læringen i nevralt nettverk ved kun å justere vekter basert på retningen av feilen, noe som gjør tilpasningen klarere og mer effektiv (Riedmiller & Braun, 1993).

Nevrale nettverk er effektive for å forutsi bevegelser i aksjemarkedet, takket være deres evne til å håndtere store og komplekse, ikke-lineære data. Disse nettverkene kan brukes til å identifisere mønstre, trekke ut data, klyngedata, og lære statistiske parametere. En hybrid tilnærming som kombinerer fundamental og teknisk analyse brukes ofte for å forutsi aksjepriser mer nøyaktig (Sheth & Shah, 2023).

Nevrale nettverk medbringer både fordeler og utfordringer. De er effektive til å håndtere ikke-lineære problemstillinger og læring fra data, noe som gjør dem svært fleksible og kraftige (Abiodun et al., 2018). Imidlertid krever de store datamengder for trening, og deres "svarte boks"-natur kan gjøre det vanskelig å forstå hvordan de kommer fram til spesifikke konklusjoner (James et al., 2021). Overtilpasning er en vanlig utfordring, og det kreves spesifikke teknikker for å sikre at nettverket generaliserer godt til ny data.

3.6 Support vector machine

Support vector machine (SVM) er en klassifiseringsteknikk som ble populær i datavitenskap på 1990-tallet. Den er anerkjent som en av de mest effektive klassifikatorene, kjent for sin gode ytelse over et bredt spekter av applikasjoner (James et al., 2021). SVM skiller seg ut ved sin bruk av hyperplan, en avgjørende grenseflate i klassifiseringsprosessen. Dette hyperplanet fungerer som skillelinje mellom to klasser av datapunkter i et flerdimensjonalt rom. I enklere rom som to eller tre dimensjoner representerer et hyperplan henholdsvis en linje og et plan (James et al., 2021). Høyere dimensjoner tar dette konseptet form av flater som effektivt kan separere datapunkter med komplekse relasjoner. Målet med SVM er å optimalisere plasseringen av hyperplanet for å maksimere avstanden mellom de nærmeste medlemmene av hver klasse, og dermed forbedre klassifiseringens nøyaktighet og robusthet. Hyperplan kan uttrykkes som:

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p = 0 \quad (7)$$

Kjernefunksjoner, eller 'kernels', spiller en sentral rolle i SVM-modeller ved at de muliggjør håndtering av ikke-lineære forhold mellom klassene i et datasett. Ved å implementere en kernel, utvides modellens kapasitet til å konstruere en ikke-lineær separasjon mellom klassene. Dette er viktig for effektiv klassifisering av komplekse datasett (James et al., 2021). Blant de ulike typer kjernefunksjoner er den radiale basisfunksjonen (RBF), også kjent som den radiale kernelen, spesielt populær. Denne typen kernel er svært effektiv for datasett som er komplekse og ikke-lineære, da den transformerer datapunktene inn i et høyere dimensjonalt rom. I dette rommet kan et lineært hyperplan anvendes for å adskille klassene effektivt. Den radiale kernelen kan matematisk uttrykkes som:

$$K(x_i, x_j) = \exp(-\gamma \sum_{j=1}^p (x_{ij} - x_{ij})^2) \quad (8)$$

3.7 Ytelsesmål i maskinlæring – Forvirringsmatrise

Forvirringsmatriser måler prestasjon av en klassifiseringsalgoritme. Matrisen sammenligner prediksjonene satt opp mot reelt resultat (Singh et al., 2021). Dette hjelper å fremstille hvor ofte en prediksjonsmodell predikerer både rett og feil. Forvirringsmatrisen består av fire karakteristikk, og er som følgende:

- I. TP. (True Positive): TP representerer antall riktig predikerte utfall, eksempelvis når en avkastning predikeres opp og den faktisk beveger seg oppover.
- II. TN. (True Negative): TN representerer antall feilpredikerte utfall, eksempelvis når en avkastning predikeres ned og den faktisk beveger seg ned.
- III. FP. (False Positive): FP representerer antall feilklassifiserte positive utfall, eksempelvis når en avkastning predikeres opp og den faktisk går ned. Kjent som en type I feil og er det mest kritiske å unngå i aksjemarkedet.
- IV. FN. (False Negative): FN representerer antall feilpredikerte negative utfall, eksempelvis når en avkastning predikeres ned og den faktisk går opp. Kjent som en type II feil.

Disse karakteristikkene plasseres deretter i en forvirringsmatrise, som videre brukes til å regne ut ytelsesmål fra.

		Faktisk klasse	
		Negativ	Positiv
Predikert klasse	Negativ	TN	FN
	Positiv	FP	TP

Figur 1. Illustrasjon av en forvirringsmatrise

Prestasjonen til algoritmen regnes videre ut basert på scorene fra forvirringsmatrisen, og består av følgende nøkkeltall:

I. Nøyaktighet: Forholdet mellom korrekte prediksjoner og totalt antallet prediksjoner, uttrykt ved: $Accuracy = \frac{(TP+TN)}{(TP+FP+FN+TN)}$ (9)

II. Presisjon: Forholdet mellom korrekte prediksjoner og totalt antall positive prediksjoner, også kjent som "positiv prediktiv verdi", uttrykt ved: $Presisjon = \frac{TP}{TP+FP}$ (10)

III. Fall-out: Andelen faktisk negative tilfeller som feilaktig klassifiseres som positive av prediksjonen, uttrykt ved: $Fall-out = \frac{FP}{TN+FP}$ (11)

IV. Recall: Forholdet mellom korrekte prediksjoner og det totale antallet prediksjoner som faktisk er positiv, uttrykt ved: $Recall = \frac{TP}{TP+FN}$ (12)

V. F1-score: Representerer likevekts punktet mellom nøyaktige positive prediksjoner og evnen til å fange opp alle faktiske positive tilfellene, uttrykt ved: $F1Score = \frac{2 \times precision \times recall}{precision + recall}$ (13)

VI. Spesifisitet: Modellens evne til å identifisere negative tilfeller, uttrykt ved: $Spesifisitet = \frac{TN}{TN+FP}$ (14)

VII. Negativ prediktiv verdi: Sannsynligheten av at negative prediksjoner faktisk er korrekt, uttrykt ved: $NPV = \frac{TN}{TN+FN}$ (15)

I tillegg utregnes AUC, området under kurven i en ROC- (receiver operating characteristic) kurve, som er nært knyttet til forvirringsmatrisen og fungerer som et viktig prestasjonsmål. ROC-kurven, som illustrerer forholdet mellom recall (sann positiv rate) og fall-out (falsk positiv rate), avhenger av forskjellige terskelinnstillinger. Når terskelen heves, økes klassifiseringen av tilfeller som negative, noe som reduserer både antallet falske positive (FP)

og sanne positive (TP), og omvendt (EL Khal, 2021). Fordi ROC-kurven gir en grafisk fremstilling snarere enn et spesifikt tall, er AUC-verdien avgjørende da den indikerer modellens evne til å skille mellom recall og fall-out, uavhengig av valgt terskel. AUC-verdien, som er et tall mellom 0 og 1, indikerer effektiviteten i modellens identifikasjon og differensiering av positive fra negative tilfeller, en verdi $> 0,5$ betyr at modellen diskriminerer bedre enn tilfeldig gjetning (Jiménez-Valverde, 2012). AUC blir beregnet ved hjelp av pakken 'pROC' i dataprogrammet Rstudio (R) (Robin et al., 2023).

MCC (Matthews Correlation Coefficient) anvendes for å evaluere binære klasser og måler korrelasjonen mellom de faktiske og predikerte klassifiseringene. MCC introduseres for å gi et målbart tall som potensielt er mer pålitelig enn nøyaktighet og f1-score da disse målene kan vise overoptimistiske resultater. MCC gir bare høy score når prediksjonen viser gode resultater i samtlige kategorier i forvirringsmatrisen, vurdert opp mot antall positive og negative elementer i datasettet (Chicco & Jurman, 2020). Det argumenteres for at MCC alltid bør foretrekkes fremfor nøyaktighet og f1-score, spesielt når det er skjev klassefordeling i resultatene (Chicco & Jurman, 2020; Chicco et al., 2021). Scoren for MCC er en skala som strekker seg fra -1 til +1. MCC utregnes også basert på forvirringsmatrisen, der følgende formel legges til grunne:

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \quad (16)$$

4 Metode

4.1 Datainnsamling

Børsdataene for oppgaven omfatter priser og variabler fra det norske markedet (OBX 25), det svenske markedet (OMXS30), det danske markedet (OMXC25) og det finske markedet (OMXH25), og er hentet fra pålitelige Investing.com (Investing.com, 2024). Indeksene representerer de største og mest omsatte selskapene på respektive børser i Norden.

Studien baserer seg primært på månedlig statistikk og indikatorer for perioden 1. april 2001 til 1. desember 2023. Datainnsamlingen omfatter økonomiske og finansielle variabler samt råvarepriser, som antas å påvirke børsindeksene. Før modellutledningen ble datamaterialet imidlertid avgrenset til perioden 1. januar 2002 til 1. desember 2023, grunnet inkludering av ytterligere variabler som resulterte i manglende verdier, som kunne ha påvirket analysens kvalitet. Til tross for at studien i hovedsak bygger på månedlig statistikk og indikatorer, er det

et unntak: intramånedlig volatilitet. Denne ble manuelt beregnet basert på daglige prisdata (\pm 20 dager) i Excel, basert på data hentet fra Investing.com.

For å kartlegge mulige prisfaktorer som kan påvirke markedsutviklingen, ble relevante råvarepriser inkludert i analysen. Disse råvarene omfatter olje, gass, gull, sølv, aluminium, kobber, jernmalm, nikkel og hvete. Råvareprisene er omregnet til avkastning for å reflektere investeringsverdien over tid. Dataene for olje, gass, gull og sølv ble hentet fra Investing.com (Investing.com, 2024), mens data for nikkel, kobber, jernmalm, aluminium og hvete ble innhentet fra Federal Reserve Economic Data (FRED, 2024), en anerkjent kilde for internasjonale økonomiske indikatorer.

Datasettet inkluderer videre økonomiske og finansielle variabler spesifikke for hvert land. Disse dataene er inkludert for å gi en helhetlig forståelse av de ulike faktorene som kan påvirke markedsutviklingen, da økonomiske indikatorer ofte har direkte eller indirekte innvirkning på markedstrender. Makrovariablene omfatter 10-års statsobligasjoner, inflasjon, valutakurser og arbeidsledighetstall. Dataene for statsobligasjoner og inflasjon er hentet fra Federal Reserve Economic Data, mens arbeidsledighetstallene er innhentet fra Eurostat (Eurostat, 2024), den offisielle statistiske avdelingen for EU.

4.2 Dataforberedelse

Etter datainnsamlingen ble videre dataforberedelser hovedsakelig gjennomført i R. Først ble prisendringer beregnet ved å finne differansen mellom etterfølgende prisnivåer. Deretter ble en binær variabel introdusert for å indikere om prisendringen var positiv (> 0) eller negativ (< 0). Ved en positiv prisendring ble den binære variabelen satt til 1, som indikerer en økning i indeksverdien, og ved en negativ prisendring ble den binære variabelen satt til 0, som indikerer en nedgang i indeksverdien. Den binære variabelen, i likhet med andre nye variabler, ble lagt til ved bruk av R-pakken "dplyr" (Wickham et al., 2023).

Det ble så inkludert forsinkede verdier for de numeriske variablene (lags). Dette innebar å forskyve dataene nedover for et bestemt antall perioder (i), noe som resulterte i "lags" for opptil tre perioder. Formålet var å analysere hvordan tidligere verdier påvirker fremtidige verdier. For hver numeriske variabel og den binære variabelen ble nye kolonner skapt for å representere disse forsinkede verdiene. Disse kolonnene ble fylt med manglete verdier (NA) for de første i radene, hvor " i " refererer til antall perioder dataene var forskjøvet.

For å styrke prediksjonsevnen, særlig i perioder med nedadgående markeder, inkluderte vi noen tekniske indikatorer i analysen basert på Macchiarulos (2018) studie, som viste at maskinlæring presterte bedre i oppgangsperioder, mens teknisk analyse var mer effektiv i nedgangsperioder (Macchiarulo, 2018). Indikatorene som ble inkludert er "Relative Strength Index (RSI) " og "Moving Average Convergence Divergence (MACD) " .

De tekniske indikatorene RSI og MACD er beregnet ved bruk av R-pakken 'TTR' (Ulrich, 2023). RSI, basert på J. Welles Wilders (1978) studie, måler momentum i en indeks eller aksje (Wilder, 1978). Siden våre modeller benytter månedlige observasjoner og vi ønsket tilstrekkelig antall observasjoner for å beregne langsiktig momentum, ble RSI-perioden justert til 12 måneder. MACD, som er forskjellen mellom to eksponentielle glidende gjennomsnitt (EMA), bruker vanligvis en kortere og en lengre periode EMA. Standardinnstillingene er 12 dager for den korte perioden og 26 dager for den lange perioden. For å tilpasse disse innstillingene til månedlige data, ble den korte perioden justert til $n_{Fast}=2$, den lange perioden til $n_{Slow}=5$, og signallinjen til $n=3$. Justeringene ble gjort for å opprettholde sammenhengen mellom de opprinnelige daglige parameterne og den månedlige frekvensen, slik at indikatorene fortsatt reflekterer markedets bevegelser på en konsistent måte (Nison, 2001). Det ble, i likhet med alle variabler, opprettet forsinkende verdier av de tekniske indikatorene.

Før videre modellering ble igangsatt, gjennomførte vi det viktige steget med å håndtere manglende data (NA-verdier) som oppsto ved opprettelsen av forsinkede verdier for variablene. Ved å fjerne disse verdiene sikret vi at analysen ikke ble kompromittert av manglende data. NA-verdiene oppstår fordi de første radene ikke har tilstrekkelig tidligere data til å opprette forsinkede verdier, og å fjerne disse verdiene er en avgjørende prosess for å håndtere manglende data korrekt slik at man i tillegg unngår 'look-ahead bias'. Dette sikrer at prediksjonene kun er basert på informasjon som var tilgjengelig frem til det aktuelle tidspunktet, og eliminerer risikoen for at modellen tar i bruk fremtidig informasjon som den ikke burde ha tilgang til. Dette førte til at det endelige datasettet ble begrenset til perioden 01. september 2002 til 01. desember 2023

Sist i dataforberedelsesfasen gjennomførte vi en inndeling av datasettet. Datasettet ble delt inn i et treningssett og et testsett basert på en 70/30-splitt. 70 prosent av dataene ble anvendt for å trene modellen, mens de resterende 30 prosentene ble brukt til testing. Oppdelingen fulgte dataenes kronologiske rekkefølge for å sikre at modellen lærte av historiske data for å kunne

forutsi fremtidige verdier. Dette oppsettet muliggjorde også testing av porteføljene for å evaluere prediksjonsmodellens prestasjon mot den faktiske markedsutviklingen.

4.2.1 Variabelseleksjon

Variabelseleksjonen og håndteringen av en binær klassifiseringsoppgave ble blant annet utført ved hjelp av en Lasso-regresjon og R-pakken "glmnet" (Friedman et al., 2023). For å utføre Lasso-regresjon forberedte vi først en matrise av forklaringsvariabler (X), hvor vi utelot ikke-forsinkende variabler for å unngå å introdusere ikke-tilgjengelig informasjon.

Responsvariabelen (Y) ble formatert til numeriske verdier (0 og 1) for å møte kravene til binær klassifisering. Ved å bruke Lasso-regresjon, som inkluderer en regulariseringsledd som minsker betydningen av mindre viktige variabler, sikret vi et best mulig sett av forklaringsvariabler. Målet var å identifisere de variablene som gir mest prediktiv verdi for å forutsi prisretningen, som ville forbedre nøyaktigheten og stabiliteten til våre modeller.

Kritisk for denne tilnærmingen var bruk av kryssvalidering for å finne en gunstig verdi av regulariseringsparameteren, λ , som balanserer mellom modellens tilpasning og kompleksitet. Denne optimaliseringen sikret at vi valgte en modell som verken var over- eller underjustert. Variabler assosiert med ikke-null koeffisienter i den resulterende Lasso-modellen ble betraktet som viktige, da de overlever prosessen med regularisering og bidrar signifikant til modellens prediksjonsevne.

Parallelt med Lasso-regresjonen anvendte vi random forest for å evaluere variabelens viktighet fra et ikke-lineært perspektiv. Ved å konstruere et stort antall beslutningstrær basert på ulike utvalg av treningsdataen og variabler, kunne vi måle hver variabels gjennomsnittlige bidrag til forbedringen av modellens prediktive nøyaktighet, ofte uttrykt gjennom en reduksjon i Gini-urenhetsgrad, som måler hvor blandet en node er i en beslutningstrestruktur. En lav Gini-urenhetsgrad indikerer at noden i hovedsak består av eksempler fra én klasse, noe som fører til mer presise prediksjoner.

En nøkkelkomponent i vår tilnærming var etableringen av en terskelverdi for variabelviktighet, satt til medianen av de observerte viktighetsverdiene. Dette tillot oss å fokusere på de øvre 50% av variablene målt ved deres bidrag til modellens prestasjon, og effektivt filtrere bort variabler som bidro marginalt. Denne metoden sikret at kun de mest informative variablene ble valgt for videre analyse.

Etter å ha identifisert viktige variabler fra både lasso-regresjon og random forest for alle fire indeksene, kombinerte vi disse og utarbeidet et datasett med variabler ansett som betydningsfulle fra begge perspektiver. Denne prosessen sikrer at modellen trenes på de mest relevante variablene, med formål at fører til bedre prestasjon og generaliseringsevne. Datasettene ble på nytt delt opp i 70/30 fordeling med, og endte på 178 observasjoner i trening og 76 observasjoner i testdatasettet.

4.3 Modellspeifikasjoner

For å kunne sammenligne prestasjon mellom modellene for de ulike landene, samt skape representative modeller som kan være anvendbar på en større mengde data har vi lagt til grunne relativt like modeller. Ulikhetene er modellspeifikasjonene, som introduseres videre.

4.3.1 Random Forest

Den første metoden vi brukte for prediksjon var ensemble-metoden random forest (RF), som kombinerer beslutningstrær og bagging for klassifisering og prediksjon. RF-prediksjonene baserer seg på Breiman og Cutlers RF for klassifisering og regresjon ettersom modellene ble utviklet ved hjelp av R-pakken "randomForest" (Liaw & Wiener, 2022).

Random forest-modellens hyperparameterne ble optimalisert ved hjelp av kryssvalidering. Vi benyttet 10- eller 15-folds kryssvalidering for de ulike landene, hvor datasettet ble delt inn i henholdsvis 10 eller 15 deler. Av disse ble 9- eller 14 deler brukt til trening, mens den gjenværende delen ble brukt til validering. Denne metoden bidrar til å redusere risikoen for overtilpasning og forbedrer modellens evne til å generalisere til ny og ukjent data.

For å sikre stabilitet og redusere varians i modellene ble det brukt et konsistent antall trær, $n_{tree} = 1000$, for alle landene. Hyperparameteren 'mtry', som styrer antall variabler som vurderes ved hver splitting, ble satt til å variere mellom 2 og 5. Denne variasjonen ble undersøkt ved hjelp av et gridsøk, ved bruk av R-pakken "caret" (Kuhn et al., 2023). Gridsøket identifiserte den optimale mtry-verdien for hver iterasjon. Kombinert med kryssvalidering utgjør disse justeringen de nødvendige modell-tuningene for RF-modellene.

Til slutt spesifiserte vi at prediktorers viktighet skulle vurderes under modellkjøringen. Øvrige parametere ble ikke spesifisert og forble dermed på standardinnstillingene for R-pakken. For å sikre reproduserbarhet ble "set. seed(123)" angitt før hver modellkjøring, som ble konsekvent gjennomført for alle maskinlæringsteknikkene.

4.3.2 Nevrale nettverk

De nevrale nettverksmodellene ble utledet ved hjelp av Fritschs m.fl. R-pakke "neuralnet" (Fritsch et al., 2019). Aller først ble de numeriske variablene standardisert for å oppnå en enhetlig skala. Denne prosessen innebærer sentralisering ved å trekke gjennomsnittet fra hver verdi, etterfulgt av skalering til et standardisert område. Standardisering bidrar til å fremme raskere og mer stabil konvergens under læringsprosessen i det nevrale nettverket. Deretter ble de standardiserte variablene kombinert med binær prisinformasjon (opp/ned), som resulterte i et komplett og relevant datasett for våre nevrale nettverksmodeller.

Preliminære eksperimenter, som indikerte en balanse mellom modellens læringskapasitet og generaliseringsevne, dannet grunnlaget for valg av antall noder i de nevrale nettverkene. Flere av de andre maskinlæringsmodellene bruker innebygde funksjoner for hyperparameter tuning, men dette gjelder ikke nevrale nettverksmodellen. Derfor gjennomførte vi manuell tuning for å finne det optimale antallet noder for de nevrale nettverksmodellene for de ulike indeksene. De nevrale nettverkene ble til slutt trent med skjulte lag som inneholdt mellom 4 og 9 noder, avhengig av indeksen. Vi spesifiserte også at modellene skulle tolkes som klassifikasjonsoppgaver. Øvrige parametere beholdt vi som standard, som betyr at modellene er et enkeltlags feed-forward nevralt nettverk. Mer spesifikt benyttet vi RPROP med vekt-tilbakestilling fra Riedmiller og Braun (1993) og Riedmiller (1994), som er den spesifikke algoritmen referert til som 'nevrale nettverk' gjennomgående i denne oppgaven.

4.3.3 Support vector machine

Support vector machine (SVM) modellene ble utviklet ved hjelp Meyers m.fl. R-pakke "e1071" (Meyer et al., 2023). Hyperparameteroptimalisering ble gjennomført ved hjelp av kryssvalidering, spesifikt 10-fold kryssvalidering med tre repetisjoner. Denne metodikken ble valgt for å teste modellens stabilitet over ulike deler av dataene, samtidig som den sikrer at modellen har god evne til å generalisere til nye, ukjente data.

Justeringen av hyperparameterne ble utført for kostnadsparameteren (C) og σ for RBF-kernel. C -parameteren, som regulerer avveiningen mellom å oppnå en bred margin og minimere klassifiseringsfeil, ble testet i et intervall fra 0,1 til 2 med en økning på 0,2 per sekvenstesting. σ , som påvirker formen på beslutningsgrensen som skapes av RBF-kernelen, ble testet i et intervall fra 0,01 til 0,05 med en økning på 0,01 per sekvenstesting. I likhet med hyperparameteroptimaliseringen for RF-modellene, ble R-pakken "caret" benyttet for å gjennomføre et omfattende gridsøk for å finne de optimale verdiene for σ og C . Dette søket

evaluerer kombinasjoner av C - og σ -verdier for å finne den beste konfigurasjonen basert på kryssvalideringsresultatene.

4.4 Gjennomføring av portefølje-backtesting

Porteføljene er utformet med det spesifikke målet å speile den faktiske markedsavkastningen når modellen predikerer oppgang, men forbli utenfor markedet når modellen predikerer nedgang. Dette betyr at porteføljene, gjennom bruken av prediktive modeller, forsøker å oppnå avkastningen som faktisk ble realisert i markedet under oppgangsperioder, mens de unngår tap ved å være ute av markedet under nedgangsperioder. Ved å basere handelsbeslutninger på modellens prediksjoner om markedets bevegelser, justerer porteføljene investeringene for å reflektere de reelle markedsforholdene. Dette gjør det mulig for investorer å måle potensialet i sine handelsstrategier mot faktiske historiske data.

Porteføljene består av fire hovedtyper: en basert på random forest, en på nevrale nettverk, en support vector machine, og til slutt en kombinasjonsportefølje. Den kombinerte maskinlæringsporteføljen benytter alle tre modellene i samme prediksjon. Hvis minst to av de tre metodene predikerer samme markedsbevegelse (kjøp, hold eller selg), vil handlinger utføres basert på flertallets prediksjon.

Porteføljene starter med en startkapital på kr 100.000, som representerer det tilgjengelige beløpet for handel. Modellene foretar handelsbeslutninger (kjøp/salg/hold) basert på prediksjoner om markedets opp- eller nedgang, hvor all handel skjer med hele beløpet. Dette betyr at posisjoner alltid tas med det fullstendige beløpet.

Porteføljene inkluderer transaksjonskostnader på 0,049% ved kjøp og salg for å gjenspeile et realistisk handelsscenario. Denne kostnaden er basert på Nordnets gebyr for aksjehandler over 52 667 NOK per transaksjon i Norden (Nordnet, n.d.). Kostnadene trekkes fra balansen ved hver handelstransaksjon, enten det er kjøp eller salg, og simulerer de handelsgebyrene en investor ville møtt ved bruk av en megler som Nordnet. Disse kostnadene påvirker nettoavkastningen ved å redusere gevinsten fra handelsaktiviteter.

Handelsmekanismer for porteføljene gjennomkjøringene:

- **Kjøp:** Utføres når modellen predikerer en markedsoppgang og ingen posisjoner er holdt. Kjøpet reduserer balansen med de påløpte transaksjonskostnadene.

- **Salg:** Aktiveres når modellen predikerer en markedsnedgang og en posisjon er holdt. Posisjonen selges, og balansen økes med den oppnådde avkastningen, fratrukket transaksjonskostnadene.
- **Hold:** Velges når modellen ikke foreslår en ny handel. Dette kan være når den forventer fortsatt oppgang og en posisjon allerede er holdt, eller ved forventet videre nedgang uten en eksisterende posisjon. Balansen oppdateres med den løpende avkastningen hvis en posisjon er holdt, mens den forblir uendret hvis ingen posisjon er holdt. Det påløper ingen ytterligere kostnader.

Til slutt evalueres de økonomiske resultatene av porteføljene gjennom både deres endelige saldo og avkastning, uttrykt i absolutte tall og prosent. For å bedømme effektiviteten og risikoen til modellene som styrer disse porteføljene, beregnes nøkkeltallene sharpe-ratio og informasjonsrate.

4.5 Evaluering av avkastningsdriverne

For å undersøke hvilke faktorer som påvirker avkastningen til de utvalgte indeksene, har vi gjennomført en multippel regresjonsanalyse hvor porteføljeavkastningen er analysert mot globale børsnoterte fond (ETF-er). ETF-variablene benyttes som indikatorer på generelle økonomiske sykluser og deres innvirkning på verdensøkonomien. Dataene er hentet direkte i R ved bruk av pakken "quantmod" (Ryan & Ulrich, 2023), og består av følgende børsnoterte fond:

1. S&P 500 (SPDR S&P 500 ETF Trust). Denne indeksen representerer 500 av de største selskapene i USA og fungerer som en viktig global referanseindeks.
2. VIX (ProShares Short VIX Short-Term Futures ETF). Denne ETF-en måler volatiliteten og risikoen i det brede amerikanske markedet og brukes til å forstå markedsusikkerhet på globalt nivå.
3. Gull (SPDR Gold Shares). Denne ETF-en følger gullpriser og tilbyr eksponering mot gullmarkedet, en investering som tradisjonelt anses som en beskyttelse mot inflasjon og valutarisiko.
4. Energi (iShares Global Energy ETF). Følger priser på olje- og andre energiresurser, som spiller en avgjørende rolle for energiselskaper og industriproduksjon.
5. Finans (iShares MSCI Europe Financials ETF). Gir innsikt i den økonomiske helsen og stabiliteten i Europa.

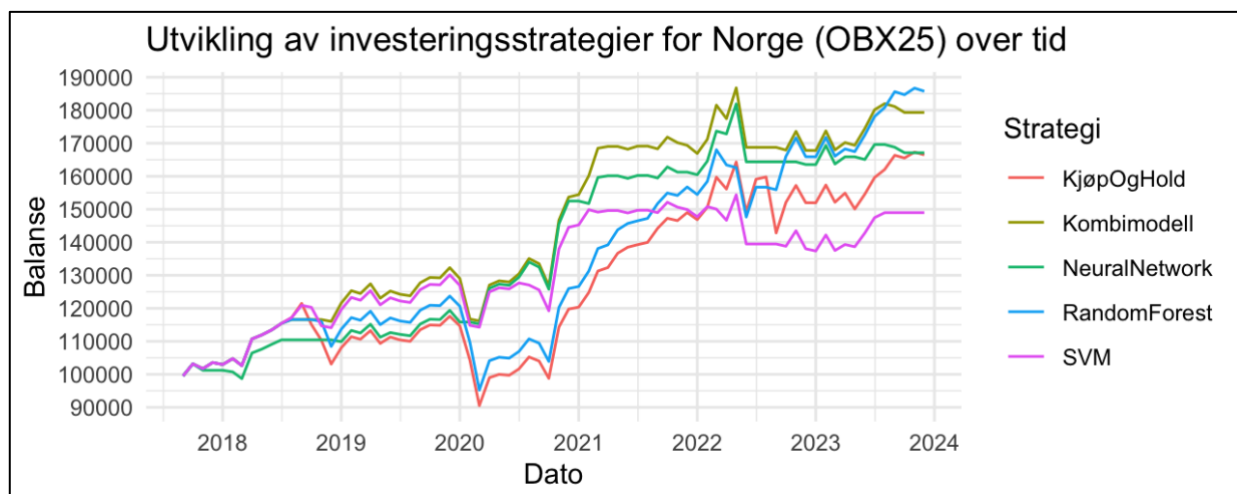
6. Helseindustri (iShares Global Healthcare ETF). Gir innsikt i både utviklingen og de økonomiske forholdene i den globale helseindustrien.
7. Industri (iShares Global Industrials ETF). ETF-en speiler den globale industrielle sektoren, og måler industriell aktivitet og produksjonsnivåer.

Kun syv variabler ble valgt i analysen. Alle representerer globale faktorer som potensielt kan forklare deler av avkastningen. Selv om ulike ETF-er kan påvirke indeksene, er målet at minst én av disse variablene vil bidra til å forklare deler av markedsavkastningen og utviklingen.

5 Resultater

Videre presenteres de samlede resultatene som illustrerer variasjonene i prediksjonene. Styrker og svakheter ved prediksjonsmodellene diskuteres mer inngående i de påfølgende underkapitlene, hvor ytelsesmål benyttes for å forklare de underliggende fenomenene som påvirker modellene. Før underkapitlene presenteres, fremstilles prisutviklingen og korrelasjon mellom prediksjonsporteføljene og den faktiske markedsutviklingen for hvert land. Dette fremstilles ved hjelp av R-pakken "ggplot2" (Wickham et al., 2024).

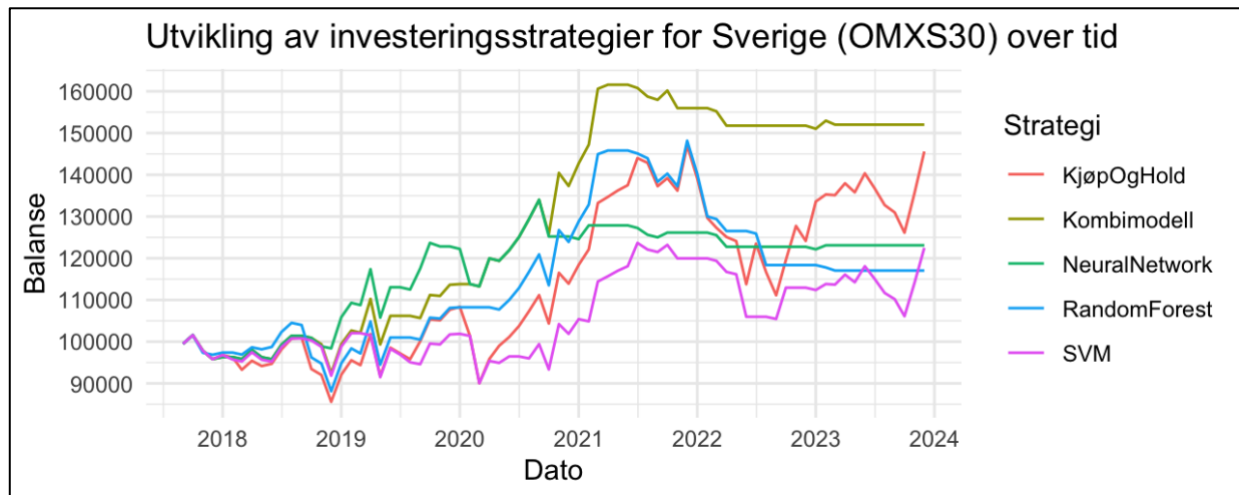
Porteføljene bygger på binære prediksjoner fra maskinlæringsmodellene. Modellens evne til å nøyaktig forutsi markedsbevegelser er avgjørende for å skape meravkastning sammenlignet med en kjøp-og-hold-strategi, noe som demonstrerer dens verdi for investorer. Den overordnede modellprestasjonen illustreres gjennom porteføljeutviklingen, som viser hvilke modeller som gir meravkastning. Disse er fremstilt i figur 2, 3, 4 og 5.



Figur 2. Graf av portefølje-utviklingene for Norge

Figur 2 illustrerer utviklingen av de ulike porteføljene for Norge. Både random forest og kombinasjonsmodellen oppnår en betydelig del av sin meravkastning ved nøyaktig å forutsi

nedgang i perioden mellom 2022 og 2023. Dette beskyttet investeringene mot større markedskorreksjoner, som påvirket kjøp-og-hold-porteføljen. Selv om alle maskinlæringsporteføljene nøyaktig forutså den respektive nedgangen, hadde den nevrале nettverksmodellen og SVM-modellen en tendens til å forutsi nedgang også under oppgangsperioder. Dette førte til at de ofte holdt seg utenfor markedet, noe som resulterte i at de oppnådde henholdsvis tilsvarende eller lavere avkastning sammenlignet med kjøp-og-hold-strategien.



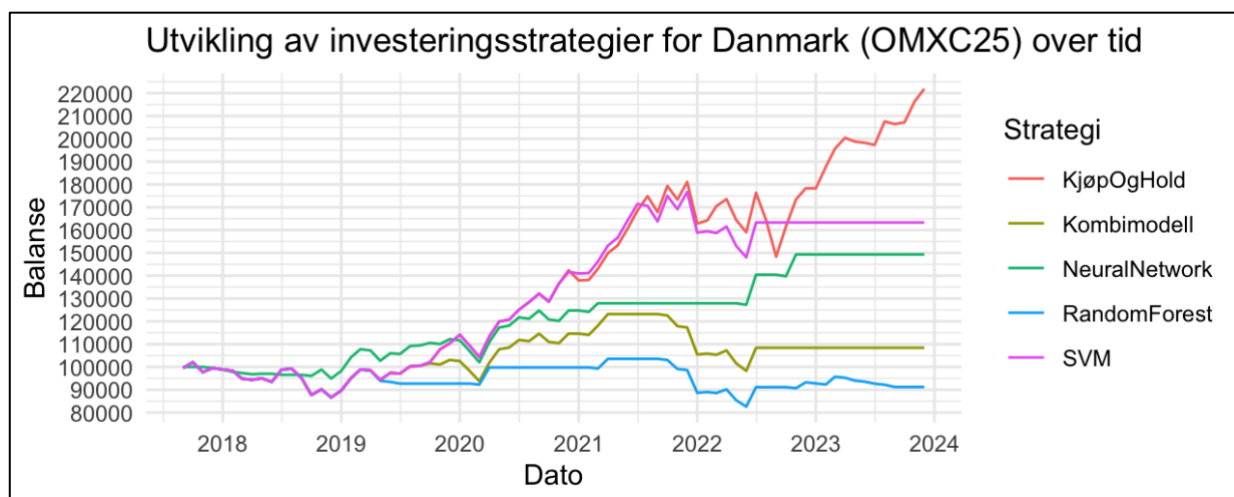
Figur 3. Graf av portefølje-utviklingene for Sverige

Kombinasjonsmodellen dominerer den svenske indeksen fra midten av 2020 til starten av 2021. Etter denne perioden forblir modellen stort sett utenfor markedet på grunn av prediksjoner om nedgang, som vist i figur 3. Den nevrале nettverksmodellen begynner å slite fra 2021, som kan indikere problemer med å tilpasse seg nye markedsdynamikker eller sensitivitet for overtilpasning i perioder med høy volatilitet og raske markedsendringer. Fra 2022 møter også random forest-modellen utfordringer, noe som reflekterer dens begrensninger i å håndtere markedsendringer som avviker fra de historiske dataene den er trent på. Samtidig opplever support vector machine-modellen en betydelig forbedring i sine prediksjoner fra midten av 2022.



Figur 4. Graf av portefølje-utviklingene for Finland

For Finland viser alle modellene en relativt lik trend frem til midten av 2022, som illustrert i figur 4. Dette antyder at de reagerer på tilsvarende markedsdynamikker. Fra slutten av 2022 observeres en betydelig divergens mellom strategiene, noe som kan indikere at modellene responderer forskjellig på økt markedsusikkerhet. SVM-modellen trekker seg ut av markedet i 2023 og returnerer mot slutten av året, mens den nevrale nettverksmodellen viser sterk prestasjon i hele 2022. Random forest-modellen følger markedet tett fra 2023 og utover. Det er verdt å påpeke at den nevrale nettverksmodellen og kombinasjonsmodellen avslutter helt lik og oppnår samme avkastning i perioden.



Figur 5. Graf av portefølje-utviklingene for Danmark

Den danske indeksen, illustrert i figur 5, viser tydelig kjøp-og-hold-strategiens dominans. Strategien utmerker seg ved å utkonkurrere alle maskinlæringsbaserte strategier mot slutten av perioden, spesielt etter midten av 2022. Både den nevrale nettverksmodellen og random forest-modellen viser betydelige avvik fra faktisk markedsutvikling. SVM-modellen er den

eneste som matcher kjøp-og-hold-strategien i effektivitet frem til midten av 2022. Imidlertid følger SVM-modellen de andre maskinlæringsmodellene etter dette ved å predikere nedgang, noe som fører til at modellen går glipp av den påfølgende oppgangen og avkastningen som kjøp-og-hold-strategien oppnår.

Grafene illustrert i figur 2, 3, 4 og 5 fremhever korrelasjonen mellom prediksjonsmodellene og de langsiktige kjøp-og-hold-strategiene. For at prediksjonsmodellene skal anses som vellykkede, må de skape meravkastning utover kjøp-og-hold-strategien. Dette innebærer at minimumsavkastningen for at prediksjonsmodellene skal kunne vurderes som vellykkede, er henholdsvis 66,79 % for Norge, 45,81 % for Sverige, 13,08 % for Finland og 122,48 % for Danmark. For å forstå de underliggende årsakene til avkastningen av prediksjonsmodellene, er det nødvendig å analysere de spesifikke prestasjonsindikatorene. Hver maskinlæringsmetode utforskes separat for å kunne trekke konklusjoner om prediksjonsevnen til de forskjellige teknikkene og identifisere eventuelle teknikker som utmerker seg positivt eller negativt.

5.1 Resultater random forest-modeller

5.1.1 Klassifiseringsresultater: Random forest-modeller

Figur 6 viser forvirringsmatrisen for random forest-modellene. Denne matrisen gir innsikt i modellens nøyaktighet og feilklassifiseringer, ved å sammenligne faktiske mot predikerte klasser for de nordiske landene. Det er spesielt interessant å merke seg de hyppige forekomstene av både falske positive og falske negative klassifiseringer.

		Faktisk klasse	
		Negativ (0)	Positiv (1)
Predikert klasse	Negativ (0)	Norge ($n = 2$)	Norge ($n = 5$)
		Sverige ($n = 19$)	Sverige ($n = 16$)
		Finland ($n = 6$)	Finland ($n = 4$)
		Danmark ($n = 13$)	Danmark ($n = 27$)
	Positiv (1)	Norge ($n = 26$)	Norge ($n = 43$)
		Sverige ($n = 16$)	Sverige ($n = 25$)
		Finland ($n = 29$)	Finland ($n = 37$)
		Danmark ($n = 15$)	Danmark ($n = 21$)

Figur 6. Forvirringsmatrise for random forest-modellene

5.1.2 Ytelse RF modeller

I Tabell 1 sammenlignes ytelsesmålinger for random forest-modellene på tvers av de nordiske landene. Dette gir en detaljert forståelse av hvordan disse modellene presterer når det gjelder nøyaktighet, presisjon, og andre viktige indikatorer, noe som belyser deres effektivitet i å klassifisere markedsutviklinger.

Tabell 1. Ytelse for random forest-modellene basert på forvirringsmatrisene

	Norge	Sverige	Finland	Danmark
Nøyaktighet	0,5921	0,5789	0,5658	0,4474
Presisjon	0,6232	0,6098	0,5606	0,5833
Fall-out	0,9286	0,4571	0,8286	0,5357
Recall	0,8958	0,6098	0,9024	0,4375
F1-Score	0,7350	0,6098	0,6916	0,5000
Spesifisitet	0,0714	0,5429	0,1714	0,4643
NPV	0,2857	0,5429	0,6000	0,3250
AUC	0,4866	0,6028	0,6408	0,5982
MCC	-0,0546	0,1526	0,1089	-0,0949

Generelt viser modellene høy presisjon for alle indeksene, noe som indikerer en god evne til å korrekt identifisere positive tilfeller. Imidlertid viser den norske random forest-modellen en lav AUC-verdi (0,4866), noe som indikerer begrensninger i å skille mellom positive og negative tilfeller. Dette står i kontrast til den høye recall-verdien som observeres både for den norske (0,8958) og svenske (0,6098) indeksen, noe som indikerer at modellene er effektive i å identifisere faktiske positive tilfeller. Når det gjelder nøyaktighet, utmerker den danske indeksen seg ved å ha vesentlig lavere nøyaktighet (0,4474) og F1-score (0,5000) sammenlignet med de andre indeksene. Dette understøttes av observasjoner fra forvirringsmatrisen, som indikerer problemer med balansen mellom presisjon og recall. Disse funnene antyder at modellen har utfordringer med å klassifisere både positive og negative tilfeller korrekt.

For spesifisitet demonstrerer den svenske modellen høy effektivitet i å identifisere negative tilfeller (0,5429). Derimot viser modellene for Norge (0,0714) og Finland (0,1714) svakheter i dette området. For NPV, som måler sannsynligheten for at negative prediksjoner faktisk er korrekte, utmerker modellene for Sverige (0,5429) og Finland (0,6000) seg med sterk ytelse. I

kontrast viser modellene for Norge (0,2857) og Danmark (0,3250) dårlig evne til å korrekt predikere negative markedsutviklinger.

Som en indikasjon på effektiviteten til klassifiseringssystemets evne til å skille mellom klasser i datasettene, ser vi på ytelsesmålet MCC. Norge (-0,0546) og Danmark (-0,0949) viser dårlig evne til å klassifisere korrekt, med score som indikerer at effektiviteten er dårligere enn tilfeldig gjetting. Disse prediksjonene tyder på en ubalanse i forvirringsmatrisen, der antall feilprediksjoner påvirker resultatene negativt. MCC-skalaen går fra -1 til +1, og en score nær 0 indikerer at modellenes prestasjoner er tilnærmet lik tilfeldig gjetting. Sverige (0,1526) og Finland (0,1089) viser bedre evne til å klassifisere korrekt, men er fortsatt nærme 0.

5.1.3 Portefølje 1 – Backtesting Random forest modell

Tabell 2 illustrerer avkastningen og risikjusterte ytelsesmål for random forest-porteføljene i Norge, Sverige, Finland og Danmark. Ved å sammenligne med andre maskinlæringsmetoder og en kjøp-og-hold-strategi, kan vi identifisere hvilke modeller som tilbyr høyest meravkastning under ulike markedsforhold.

Tabell 2. Prestasjon portefølje-backtesting for random forest-modellene

	Norge	Sverige	Finland	Danmark
Avkastning random forest modellene	85,74%	17,05%	6,35%	-8,78%
vs. Kjøp-og-hold	+18,95%	-28,76%	-6,73%	-131,26%
vs. Nevrale nettverk	+18,62%	-6,05%	-20,81%	-58,1%
vs. Support vector machine	+36,77%	-5,43%	-10,62%	-72,09%
vs. Kombinasjon	+7,43%	-34,95%	-20,92%	-17,18%
Sharpe ratio	0,6804	0,2038	0,0923	-0,1474
Informasjonsraten	0,0803	-0,1046	-0,0638	-0,3993

Random forest-modellen for Norge utmerker seg som den beste presterende modellen blant de nordiske landene, og er den eneste random forest-modellen som oppnår meravkastning målt opp mot de andre maskinlæringssteikkene. Dette reflekteres i en positiv informasjonsrate, som indikerer modellens evne til å skape meravkastning sammenlignet med markedet.

Disse funnene står i kontrast til tidligere studier gjort av Aronsen og Markussen (2023) for Norge. De poengterer at random forest-modeller hadde spesielt god prestasjon og konsekvent utkonkurrerte andre modeller frem til 2020. Den mest effektive porteføljen basert på random

forest oppnådde betydelige avkastninger før transaksjonskostnader, og viste seg økonomisk fordelaktig før markedsendringene som følge av Covid-19 i 2020. Imidlertid viser våre funn at modellene fortsatte å fungere godt i perioden 2020 til midten av 2022. Det er viktig å merke seg at Aronsen og Markussen fokuserte på OSEBX-indeksen med enkeltaksjer i deres porteføljer, mens vår studie ser på OBX-indeksen med indeks i porteføljen, noe som kan forklare noen av de observerte forskjellene i modellens ytelse under ulike markedsforhold.

For Sverige og Finland har random forest-modellene imidlertid utfordringer. De klarte ikke å skape meravkastning mot verken noen av de andre maskinlæringsteknikkene, ei heller mot kjøp-og-hold-strategien. For Sverige samsvarer dette med funnene til Wik (2023), som viste at random forest hadde den svakeste prestasjon blant de evaluerte modellene i det svenske markedet (Wik, 2023).

Mest utfordrende er det for random forest-modellen som predikerer opp-eller nedgang i Danmark. Her er sharpe-ratioen negativ (-0,1474), noe som signaliserer en ugunstig risikojustert avkastning. Det er spesielt bekymringsverdig at random forest er den desidert dårligste modellen for Danmark, og oppnår negativ avkastning over testperioden (-8,78%). Grunnen til den negative avkastningen er at modellen hovedsakelig har predikert nedgang siden midten av 2019 og forble ute av markedet store deler av tida. Når den var inne i markedet, tok den stort sett feil, noe som ytterligere understreker dens dårlige prestasjon.

5.2 Resultater Nevrale nettverksmodeller

5.2.1 Klassifiseringsresultater: Nevrale nettverksmodeller

I likhet med random forest-modellene, viser de nevralt nettverksmodellene en betydelig evne til å identifisere et stort antall positive tilfeller som vist i figur 7. Spesielt for den finske indeksen observeres en sterk prestasjon, men det er også merkbare feilklassifiseringer. Disse feilklassifiseringene inkluderer både falske positive og falske negative resultater. Det er spesielt bemerkelsesverdig at det forekommer mange falske negative tilfeller i Danmark og mange falske positive tilfeller i Finland.

		Faktisk klasse	
		Negativ (0)	Positiv (1)
Predikert klasse	Negativ (0)	Norge ($n = 13$)	Norge ($n = 19$)
		Sverige ($n = 24$)	Sverige ($n = 23$)
		Finland ($n = 10$)	Finland ($n = 5$)
		Danmark ($n = 17$)	Danmark ($n = 30$)
	Positiv (1)	Norge ($n = 15$)	Norge ($n = 29$)
		Sverige ($n = 11$)	Sverige ($n = 18$)
		Finland ($n = 25$)	Finland ($n = 36$)
		Danmark ($n = 11$)	Danmark ($n = 18$)

Figur 7. Forvirringsmatrise for nevrale nettverksmodellene

5.2.2 Ytelse Nevrale nettverksmodellene

I tabell 3 fremstilles resultatene til de nevrale nettverkspredikeringene av den binære variabelen basert på forvirringsmatrisen.

Tabell 3. Ytelse for nevrale nettverksmodellene basert på forvirringsmatrisene

	Norge	Sverige	Finland	Danmark
Nøyaktighet	0,5526	0,5526	0,6053	0,4605
Presisjon	0,6591	0,6207	0,5902	0,6207
Fall-out	0,5357	0,3143	0,7143	0,3929
Recall	0,6042	0,4390	0,8780	0,375
F1-Score	0,6304	0,5142	0,7059	0,4675
Spesifisitet	0,4643	0,6857	0,2857	0,6071
NPV	0,4063	0,5106	0,6667	0,3617
AUC	0,4301	0,6056	0,6049	0,5141
MCC	0,0669	0,1280	0,2051	-0,0177

Tabell 3 viser at nevrale nettverksmodellene varierer i sin evne til å klassifisere korrekt avhengig av det spesifikke landet. Generelt sett varierer modellens ytelse, med Finland som viser den høyeste nøyaktigheten (0,6053) og recall (0,8780), noe som indikerer en sterk evne til å korrekt klassifisere og identifisere faktiske positive tilfeller. På den andre siden viser Danmark lavest nøyaktighet (0,4605) og F1-score (0,4675), som reflekterer utfordringer med

balansen mellom presisjon og recall, og antyder en inkonsistens i modellens evne til å levere jevn ytelse.

Spesifisitet er relativt høy i Sverige (0,6857) og Danmark (0,6071), noe som tyder på en sterk evne til å identifisere negative tilfeller korrekt. Når det gjelder NPV, har Finland (0,6667) og Sverige (0,5106) høyere verdier, noe som viser at modellene i disse landene er bedre til å forutsi negative tilfeller korrekt. Imidlertid har disse modellene generelt lav AUC-verdi, spesielt i Norge (0,4301), noe som antyder begrensninger i modellens diskriminerende evne.

MCC-verdiene gir en samlet vurdering av modellens prediksjonsytelse ved å ta hensyn til både sanne, falske positive og negative tilfeller. Finland har den høyeste MCC (0,2051), som indikerer en relativt god balanse mellom de forskjellige ytelsesmålene, mens Danmark har den laveste MCC (-0,0177), noe som antyder at modellen har en prestasjon nær tilfeldig gjetning. Dette understøtter funnene fra de andre ytelsesmålene, og viser at det er betydelige variasjoner i modellens ytelse mellom de ulike landene.

5.2.3 Portefølje 2 – Backtesting nevralt nettverksmodell

Tabell 4 presenterer resultatene av backtesting for de nevralt nettverksmodellene i de nordiske markedene, og viser totalavkastningen kjøp-, salg- eller hold- porteføljen har oppnådd i løpet av testperioden.

Tabell 4. Prestasjon portefølje-backtesting for de nevralt nettverksmodellene

	Norge	Sverige	Finland	Danmark
Avkastning nevralt nettverk modellene	67,12%	23,10%	27,16%	49,31%
vs. Kjøp-og-hold	+0,33%	-22,71%	+14,08%	-73,17%
vs. Random forest	-18,62%	+6,05%	+20,81%	+58,09%
vs. Support vector machine	+18,15%	+0,62%	+10,19%	-14%
vs. Kombinasjon	-11,19%	-28,9%	-0,11%	+40,91%
Sharpe ratio	0,6555	0,3240	0,2902	0,7284
Informasjonsraten	-0,0138	-0,0756	0,0766	-0,1693

Generelt viser de nevralt nettverksmodellene forbedringer i risikjustert avkastning i de nordiske markedene. I Norge har modellen fått en ytelse som nesten speiler markedet, noe som resulterer i en nedgang i den absolutte avkastning fra random forest-porteføljen, og er på linje med avkastningen til en langsiktig kjøp-og-hold-strategi. Til tross for dette,

oppretholder modellen en solid sharpe-ratio (0,6555), som indikerer at risikohåndteringen og effektiviteten i den norske modellen fortsatt er god.

I kontrast til dette, øker Finland sin avkastning betydelig med en totalavkastning på 27,16%, kombinert med en positiv informasjonsrate (0,0766). Kombinasjonsmodellen har tilnærmet nøyaktig lik avkastning. Dette reflekterer en vellykket tilpasning til markedsdynamikken og en forbedret prediktiv-nøyaktighet i modellen. Dette samsvarer med Määttä (2020) som påpekte at LSTM-nevralt nettverk var effektive i å forutsi aksjeavkastninger på det finske markedet, med god risikojustert avkastning indikert ved sharpe-ratioen (Määttä, 2020).

Sverige viser moderate forbedringer i sharpe-ratioen (0,3240), og den svenske nevralt nettverksmodellen presterer noe bedre enn random forest-modellen, med en betydelig økning i sharpe-ratioen. Fra Figur 3 kan man se at utviklingen i Sverige viser en god trend frem til starten av 2021, hvor modellen stort sett er ute av markedet de neste to årene.

Danmark skiller seg ut med en betydelig forbedring i både sharpe-ratio (0,7284) og avkastning (49,31%), til tross for utfordringer med konsistens i meravkastningen. Pedersen og Lyngø (2023) hevder at avanserte prognosemodeller er viktige for å forbedre nøyaktigheten og påliteligheten av prediktive modeller på det danske aksjemarkedet, noe som er spesielt relevant for selskaper som Novo Nordisk (Pedersen & Lyngø, 2023). Jørgensen et al. (2021) understøtter dette ved å påpeke at det danske aksjemarkedet oppfører seg som en tilfeldig vandring, og understreker behovet for mer forskning for å vurdere effektiviteten av nevralt nettverksmodeller i tilfeldige og komplekse markeder (Jørgensen et al., 2021).

Samlet sett viser nevralt nettverksmodeller en variert, men positiv utvikling i de nordiske markedene. Finland og Danmark fremstår som de mest forbedrede markedene sammenlignet med random forest-porteføljene, mens Norge og Sverige viser mer moderate endringer, der Norge ender med samme avkastning som markedet.

5.3 Resultater support vector machine-modeller

5.3.1 Klassifiseringsresultater: Support vector machine-modeller

Resultatene fra forvirringsmatrisen i figur 8 for viser at Support vector machine-modellen har en variabel ytelse på tvers av de nordiske landene. I Norge viser modellen en relativt høy evne til å identifisere positive tilfeller med 35 sanne positive, men samtidig en høy andel falske positive (20 tilfeller). Sverige viser en lignende trend med 29 sanne positive, og 23 falske

positive. Finland har den laveste andelen falske negative (5 tilfeller), men høy andel falske positive (26 tilfeller). Danmark har en balanse mellom sanne positive (34 tilfeller) og falske positive (21 tilfeller), men også en betydelig mengde falske negative (14 tilfeller). Disse resultatene indikerer at SVM-modellen har utfordringer med nøyaktigheten, spesielt i å redusere antall falske negative, men viser samtidig styrker i å identifisere sanne positive tilfeller, spesielt i Danmark.

		Faktisk klasse	
		Negativ (0)	Positiv (1)
Predikert klasse	Negativ (0)	Norge ($n = 8$)	Norge ($n = 13$)
		Sverige ($n = 12$)	Sverige ($n = 12$)
		Finland ($n = 9$)	Finland ($n = 5$)
		Danmark ($n = 7$)	Danmark ($n = 14$)
	Positiv (1)	Norge ($n = 20$)	Norge ($n = 35$)
		Sverige ($n = 23$)	Sverige ($n = 29$)
		Finland ($n = 26$)	Finland ($n = 36$)
		Danmark ($n = 21$)	Danmark ($n = 34$)

Figur 8. Forvirringsmatrise for support vector machine-modellene

5.3.2 Modellytelse support vector machine

I tabell 5 fremstilles resultatene til support vector machine-modellenes prediksjoner av den binære variabelen basert på forvirringsmatrisen.

Tabell 5. Ytelse for support vector machine-modellene basert på forvirringsmatrisene

	Norge	Sverige	Finland	Danmark
Nøyaktighet	0,5658	0,5395	0,5921	0,5395
Presisjon	0,6364	0,5577	0,5806	0,6182
Fall-out	0,7143	0,6571	0,7429	0,75
Recall	0,7292	0,7073	0,8780	0,7083
FI-Score	0,6796	0,6237	0,6990	0,6602
Spesifisitet	0,2857	0,3429	0,2571	0,25
NPV	0,3810	0,5000	0,6429	0,3334
AUC	0,5298	0,5721	0,5868	0,5565
MCC	0,0161	0,0538	0,1738	-0,0449

Generelt viser modellene en jevn nøyaktighet som presentert i tabell 5, Finland har en marginal forbedring (0,5921) sammenlignet med de andre, noe som peker på en moderat konsistens i modellenes evne til korrekt å klassifisere både positive og negative tilfeller. Når det gjelder presisjon, er resultatene relativt jevne, med Danmark (0,6182) som den beste. Dette indikerer en balansert tilnærming til å unngå falske positive tilfeller over de forskjellige indeksene. Recall er generelt høy, spesielt i Finland (0,8780), noe som antyder en sterk evne til å identifisere faktiske positive tilfeller.

Til tross for styrkene ved recall-verdiene, har modellene utfordringer med spesifisitet, noe som indikerer en generell tendens til redusert evne til å korrekt identifisere negative tilfeller. Dette er en svakhet som gjenspeiles over alle indeksene. Den negative prediktive verdien (NPV) viser også rom for forbedring, spesielt i Norge (0,3810) og Danmark (0,3334), når det gjelder nøyaktigheten av negative prediksjoner.

5.3.3 Portefølje 3 – Backtesting support vector machine-modellene

Tabell 6 presenterer resultatene av backtesting for SVM-modellene i de nordiske markedene, og viser totalavkastningen SVM-porteføljen har oppnådd i løpet av testperioden.

Tabell 6. Prestasjon portefølje-backtesting for support vector machine-modellene

	Norge	Sverige	Finland	Danmark
Avkastning support vector machine modellene	48,97%	22,48%	16,97%	63,31%
vs. Kjøp-og-hold	-17,82%	-23,33%	+3,89%	-59,17%
vs. Random forest	-36,77%	+5,43%	+10,62%	+72,09%
vs. Nevrale nettverk	-18,15%	-0,62%	-10,19%	+13,99%
vs. Kombinasjon	-29,34%	-29,52%	-10,3%	+54,91%
Sharpe ratio	0,4463	0,2553	0,1958	0,6688
Informasjonsraten	-0,0673	-0,0933	0,0157	-0,1842

Fra figur 2 så vi at SVM-modellen for Norge fulgte de beste modellene frem til slutten av 2020. Deretter ble det tydelig at SVM er den svakeste modellen for Norge når det gjelder avkastning (-17,82% sammenlignet med kjøp-og-hold), sharpe-ratio (0,4463) og informasjonsrate (-0,0673).

For den svenske indeksen viser SVM seg å være den svakeste helt frem til midten av 2023 som illustrert i figur 3. Interessant nok, i den siste perioden predikerer modellen korrekt og er

inne i markedet, og ender dermed med en totalavkastning på 22,48% i løpet av perioden. Denne modellen viser også bedre resultater for både sharpe-ratio (0,2553) og informasjonsrate (-0,0933) sammenlignet med random forest-porteføljen, og er på nivå med den nevrale nettverksmodellen, som illustrert i Figur 3.

SVM-modellen for Finland viser en evne til å følge resten av markedet frem til 2023. I den siste perioden er modellen ute av markedet da den predikerer en nedgang, hvor den i store perioder har rett. Dette er også den eneste indeksen der SVM oppnår en marginal positiv informasjonsrate (0,0157).

For den danske indeksen skiller SVM-modellen seg ut som den maskinlæringsmetoden som presterer best og er helt på nivå med kjøp-og-hold-strategien frem til midten av 2022. Deretter forlater modellen markedet fordi den predikerer en nedgang, og går dermed glipp av mye av oppgangen som fører til at den oppnår 59 prosentpoeng mindre avkastning enn kjøp-og-hold-strategien.

Resultatene for SVM porteføljene indikerer en generell underprestasjon på vårt datasett, som er begrenset i størrelse. Dette samsvarer med funn i den eksisterende litteraturen som antyder at SVM kan ha utfordringer med mindre datasett. Spesifikt peker Sheth og Shah (2023) på at selv om SVM viser potensial generelt, kan metodens effektivitet være innskrenket når den brukes på mindre mengder data. Videre støtter Singh (2022) denne betraktningen ved å fremheve at SVM vanligvis yter bedre med større datasett.

5.4 Resultater kombinasjonsmodeller

5.4.1 Portefølje 4 – Backtesting kombinasjonsmodeller

Tabell 7 presenterer resultatene av backtesting for kombinasjonsmodellene, og viser totalavkastningen kombinasjonsporteføljen har oppnådd i løpet av testperioden.

Tabell 7. Prestasjon portefølje-backtesting for kombinasjonsmodellene

	Norge	Sverige	Finland	Danmark
Avkastning kombinasjonsmodellene	78,31%	52,00%	27,27%	8,40%
vs. Kjøp-og-hold	+12,52%	+6,19%	+14,19%	-114,08%
vs. Random forest	-6,43%	+34,95%	+20,92%	+17,18%
vs. Nevrale nettverk	+12,19%	+28,9%	+0,11%	-40,92%
vs. Support vector machine	+29,34%	+29,52%	+10,3%	-54,91%
Sharpe ratio	0,7042	0,6083	0,2898	0,5921
Informasjonsraten	0,0244	-0,0023	0,0843	-0,2106

Som vist i tabell 7 har kombinasjonsmodellen høy ytelse på generell basis. Dette samsvarer med Krauss et al. (2017) som hever at en kombinasjonsmodell i noen tilfeller kunne oppnå bedre resultater enn enkeltmodeller (Krauss et al., 2017).

For Sverige er modellen den beste når det gjelder avkastning (52,00%), til tross for en informasjonsrate som marginalt er i minus (-0,0023). Modellen oppnår også en høy sharpe-ratio (0,6083). Kombinasjonsmodellen hadde en overlegen avkastning frem til starten av 2020, som vist i figur 3. Etter dette var modellen periodisk inne i markedet under nedganger, men tilbrakte store deler av den gjenværende perioden utenfor markedet.

For Norge er kombinasjonsporteføljen tett på random forest-porteføljen i ytelse, men går ut av markedet i de siste månedene av perioden, og blir dermed slått av random forest-porteføljen. Likevel er det kombinasjonsporteføljen som oppnår den høyeste sharpe-ratioen (0,7042) blant porteføljene for Norge.

For Finland viser den kombinerte modellen tilnærmet lik avkastning (27,27%) som den nevrale nettverksmodellen, og er marginalt den beste modellen for det finske markedet. Den oppnår også marginalt bedre sharpe-ratio (0,2898) og informasjonsrate (0,0843) enn den nevrale nettverksmodellen. Utviklingen er nesten identisk med den nevrale nettverksmodellen, og i likhet med den nevrale nettverksmodellen er kombinasjonsmodellen også ute av markedet fra 2023 og ut, som illustrert i figur 4.

I Danmark sliter kombinasjonsporteføljen i likhet med de andre maskinlæringsporteføljene. Fra midten av 2022 er modellen ute av markedet og oppnår knapt positiv avkastning for hele

perioden (8,40%), noe som gjør den til den nest dårligste metoden for Danmark. Fra figur 5 kan man se at modellen henger greit med de øvrige modellene frem til 2022, hvor modellen tar en kjøpposisjon i et nedgående marked, før den går ut av markedet fra midten av 2022 og kun predikerer videre nedgang og dermed forblir ute av markedet resten av perioden som følger av at to eller flere prediksjonsmodeller predikerer nedgang.

5.5 Multippelregresjon

Resultatene for prediksjonsmodellene gir ikke innsikt i de underliggende driverne for avkastningen. For å identifisere hvilke makroøkonomiske faktorer som påvirker markedsavkastningen, har vi gjennomført en multippel regresjon. Kun signifikante variabler er presentert i Tabell 8, mens fullstendige detaljer og output er tilgjengelig i vedlegg A.

Tabell 8. Signifikante variabler basert på multippelregresjon

Koeffisient	Estimat	Standardfeil	T-verdi	Pr(> t)	Signifikant for
Energi	0,3018	0,0517	5,843	0,000000159***	Norge
VIX	-0,0211	0,0105	-2,000	0,0495*	Norge
Industri	0,5562	0,1787	3,113	0,00271**	Sverige
Gull	-0,2225	0,0837	-2,660	0,00974**	Sverige
Energi	-0,1123	0,0568	-1,977	0,05206.	Sverige
VIX	-0,0250	0,0122	-2,058	0,0434*	Finland
Industri	0,4897	0,2016	2,429	0,0178*	Danmark
Helseindustri	0,3333	0,1472	2,264	0,0267*	Danmark
Skjæringspunkt	0,0060	0,0035	1,721	0,0899.	Danmark

*Signifikans koder: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Indeksene for Norge, Sverige, Finland og Danmark har alle minst én makroøkonomisk variabel som viser signifikans og påvirker avkastningen betydelig. For Norge er energi-variabelen spesielt viktig, noe som understreker landets rolle som en olje- og gassnasjon med svært høy signifikans. Fryktindeksen VIX viser en negativ sammenheng for både Norge og Finland, noe som indikerer at en økning i VIX, et mål på markedsfrykt, typisk resulterer i lavere avkastninger i disse landene.

En av årsakene til at helseindustri-variabelen er inkludert er at danske Novo Nordisk ble Europas mest verdifulle selskap i 2023 (Aaser, 2023). Denne utviklingen antas å ha hatt en betydelig innvirkning på Danmarks markedsavkastning, en antagelse som styrkes av

variabelens statistisk signifikante resultater. Videre viser industri-variabelen en mer fremtredende påvirkning på Danmarks økonomi, med et koeffisientestimat på 0,4897 og en tilhørende lav p-verdi på 0,0178, som indikerer en positiv effekt på avkastningen.

For Sverige viser resultatene at avkastningen positivt korrelerer med industrisektoren. I tillegg observeres en negativ korrelasjon mellom markedsavkastningen og prisøkning på gull. Energi-variabelen er nesten signifikant med en p-verdi like over 0,05, markert ved signifikanskoden punktum. Dette reflekterer en invers relasjon hvor økte priser og avkastninger innen disse sektorene vanligvis resulterer i en reduksjon i den generelle markedsavkastningen i Sverige

6 Konklusjon

Dette studiet har utforsket ytelsen til forskjellige maskinlæringsmodeller—inkludert random forest, kunstige nevralt nettverk (KNN), og support vector machine (SVM)—i prediksjon og avkastning for det norske markedet (OBX 25), det svenske markedet (OMXS30), det danske markedet (OMXC25) og det finske markedet (OMXH25). For å vurdere effektiviteten av disse metodene, ble det laget porteføljer basert på prediksjonene fra hver modell. Disse porteføljene ble deretter sammenlignet med en kjøp-og-hold-strategi for å se om de kunne oppnå høyere avkastning.

Våre resultater fra de nordiske aksjeindeksene demonstrerer en variert effektivitet av maskinlæringsteknikkene random forest, kunstig nevralt nettverk og support vector machine i prediksjon over lengre intervaller. Spesielt bemerket vi at random forest oppnådde den beste avkastningen i det norske markedet med en positiv informasjonsrate, mens nevralt nettverk presterte særlig bra i den finske indeksen, også de med positiv informasjonsrate. I tillegg observerer vi at de kombinerte modellene generelt yter bedre enn enkeltstående modeller, spesielt i det norske, svenske og finske markedet hvor modellene også viste en positiv informasjonsrate. Dette støtter funnene i Krauss et al. (2017), som også rapporterte forbedret prediksjonsevne for kombinerte modeller mot individuelle maskinlæringsmetoder.

Igjennom analysen av de ulike markedene ser vi at de modellene som oftest predikerte oppgang, tenderte til å oppnå de beste resultatene. Dette sammenfaller med generell markedsoppførsel i perioden, hvor de fleste indekser har opplevd en oppadgående trend. Dermed ser det ut til at modeller som tar en mer optimistisk tilnærming til prediksjoner av oppgang, har vært mest vellykkede. Derimot har mer konservative modeller, som oftere

predikerer nedgang, generelt ikke klart å overgå en enkel kjøp-og-hold-strategi, spesielt gitt at børsene har hatt en overordnet positiv utvikling i den aktuelle perioden.

For den danske indeksen er prediksjonsutfordringene spesielt fremtredende. Ingen av maskinlæringsmodellene klarte å overgå kjøp-og-hold-strategien. Dette kan delvis forklares med at den danske indeksen er et mindre marked hvor prediksjoner kan være spesielt utfordrende. Særlig kan store selskaper som Novo Nordisk, som opplever markante prisendringer, skape store utslag som ikke nødvendigvis er fanget opp i modellenes treningsperioder. Disse observasjonene stemmer overens med Pedersen og Lyngbø (2023), som påpekte behovet for en dypere forståelse og justering av modellparametere. Videre støtter Jørgensen et al. (2021) sitt arbeid med den danske indeksen random walk-teorien, noe som bekrefter markedets uforutsigbarhet og understreker behovet for videre forskning på nevralt nettverk og andre maskinlæringsmetoder for mer effektiv prediksjon.

For å identifisere driverne bak avkastningene til indeksene i Norden ble det utført en multipl regressjonsanalyse. Energi er svært signifikant for Norge, og VIX-indeksen viser en negativ sammenheng med avkastningen i både Norge og Finland. I Danmark har helseindustrien en positiv effekt, sannsynligvis på grunn av Novo Nordisks fremtredende rolle, mens industri har en positiv effekt i både Danmark og Sverige. I Sverige fant vi også en negativ korrelasjon mellom markedsavkastning og gullprisen.

Samlet sett viser våre resultater en variabel prediksjonsevne av maskinlæringsstenikkene random forest, kunstig nevralt nettverk og support vector machine over lengre tidsintervaller i de nordiske aksjeindeksene. Prediksjonsmodellene viste seg å være mest effektive når de primært forutså markedets oppgang, spesielt i perioder med økonomisk vekst, og de kombinerte modellene viste seg å være mer stabile i sine prediksjoner over de ulike indeksene sammenlignet med enkeltstående modeller. Imidlertid reflekterer utfordringene i det danske markedet kompleksiteten og volatiliteten som kan oppstå i mindre indekser, spesielt påvirket av dominerende selskaper. Disse funnene antyder at fremtidig forskning bør fokusere på både modelljustering og dybdeanalyse av markedsdynamikk for å forbedre prediksjonsnøyaktigheten ytterligere.

6.1 Begrensninger/Svakheter med oppgaven

En innledende utfordring var identifikasjonen av forklaringsvariabler som kunne anvendes universelt og være generaliserbare for de fire nordiske børsindeksene, i betraktning av deres unike markedsspesifikke karakteristikk. For å håndtere denne utfordringen innlemmet vi globale råvarepriser som innehar betydelig relevans for den globale økonomien. Videre utvidet vi analysen med landsspesifikke økonomiske og finansielle variabler tilpasset hver enkelt marked. Til slutt brukte vi variabelseleksjon ved hjelp av Lasso og Random Forest for å sikre at kun relevante variabler ble inkludert for hver indeks.

Videre støtte vi på utfordringer i forsøket på å sikre en omfattende historisk dataserie for alle de fire nordiske indeksene, særlig med hensyn til data som går tilbake før år 2000. En tilleggsutfordring har vært mangelen på volum- og sprednings-data for samtlige indekser. Disse variablene er essensielle for analysen av markedslikviditet og prissvingninger. Dessverre var de tilgjengelige historiske dataene kun fragmenterte med hensyn til disse faktorene, noe som nødvendiggjorde deres eksklusjon fra analysen.

Vi overveide bruken av kapitalverdimodellen (CAPM) for å analysere signifikansen av avkastningene til porteføljene. Tidligere forskning av Banz (1981), Fama og French (1992) har imidlertid hevdet at β -verdien ikke tilbyr tilstrekkelig forklaringskraft for variasjoner i gjennomsnittlige avkastninger (Banz, 1981; Fama & French, 1992). En etablert metode for å vurdere signifikansen av avkastninger er bruk av Fama-French-faktorer, men disse er ikke tilgjengelige for de indeksene vi har valgt. Som et alternativ brukte vi en makroøkonomisk multifaktormodell for å identifisere de underliggende driverne bak avkastningene til porteføljene.

6.2 Videre forskning

For videre forskning anbefales det å gjennomføre en grundig analyse av det danske aksjemarkedet for å identifisere årsakene til den lave prediktive nøyaktigheten som er observert både i våre egne modeller og i tidligere studier (Pedersen & Lyngbø, 2023; Jørgensen et al., 2021). Det primære målet med en slik studie ville være å avdekke de underliggende dynamikkene i markedet, blant annet hvorfor modellene predikerer så mye ned i oppgangsperioder, noe som kan bidra til en mer finjustert tilpasning av modell-parameterne for bedre å håndtere markedets kompleksitet og volatilitet. Videre, dybdeanalyse av selskaper som har stor innvirkning, som Novo Nordisk, kan gi innsikt i hvordan modellene kan tilpasses for å bedre reflektere markedets dynamikk.

En annen retning for videre forskning er å evaluere hvordan eksterne makroøkonomiske faktorer påvirker modellens prediksjonsevne. I dette studiet ble det observert at enkelte makroøkonomiske variabler hadde signifikant påvirkning på avkastningen til ulike indekser. Fremtidige studier bør derfor inkludere en bredere analyse av slike variabler, inkludert deres interaksjoner og samlede effekt på modellens ytelse. Dette kan bidra til å identifisere nye indikatorer som kan forbedre prediksjonsnøyaktigheten ytterligere.

Til slutt kan videre forskning fokusere på alternativer til våre modeller, eksempelvis mer avanserte nevralt nettverksarkitekturer, som Long Short-Term Memory (LSTM), rekurrente nevralt nettverk og multilag-feed-forward nevralt nettverk. Spesielt kan LSTM-nettverk være nyttige for å få fange opp langsiktige avhengigheter i tidsseriedata, noe som kan forbedre prediksjonene i volatile markeder.

Referanseliste

- Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, 4(11), e00938. <https://doi.org/10.1016/j.heliyon.2018.e00938>
- Agrawal, A., Gans, J., & Goldfarb, A. (2017). What to expect from artificial intelligence. In: MIT Sloan Management Review Cambridge, MA, USA.
- Aronsen, M., & Markussen, C. (2023). *Can machine learning beat the Norwegian stock market? A comparison of popular machine learning models* [UiT - The arctic university of Norway].
- Badolia, L. (2016). *How can I get started Investing in the Stock Market*. <https://books.google.no/books?id=zA9PDQAAQBAJ>
- Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1), 3-18. [https://doi.org/10.1016/0304-405X\(81\)90018-0](https://doi.org/10.1016/0304-405X(81)90018-0)
- Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, 21(1), 6. <https://doi.org/10.1186/s12864-019-6413-7>
- Chicco, D., Warrens, M. J., & Jurman, G. (2021). The Matthews Correlation Coefficient (MCC) is More Informative Than Cohen's Kappa and Brier Score in Binary Classification Assessment. *IEEE Access*, 9, 78368-78381. <https://doi.org/10.1109/ACCESS.2021.3084050>
- DeRose, K., Dixon, M., & LeLannou, C. (2019). Machine Learning. In: CFA Institute.
- EL Khal, Y. (2021). *Confusion Matrix, AUC and ROC curve and Gini clearly explained*. Retrieved 11.04.24 from <https://yassineelkhal.medium.com/confusion-matrix-auc-and-roc-curve-and-gini-clearly-explained-221788618eb2>
- Eurostat. (2024). *Unemployment by sex and age – monthly data*. https://doi.org/10.2908/une_rt_m
- Fama, E., & French, K. (1992). The Cross-Section of Expected Stock Returns. *The Journal of THE AMERICAN FINANCE ASSOCIATION*, 47(2), 427-465. <https://doi.org/10.1111/j.1540-6261.1992.tb04398.x>
- Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383-417. <https://doi.org/10.2307/2325486>
- FRED. (2024). *Federal Reserve Economic Data: Prices*. <https://fred.stlouisfed.org>
- Friedman, J., Hastie, T., Tibshirani, R., Narasimhan, B., Tay, K., Simon, N., Qian, J., & Yang, J. (2023). *glmnet: Lasso and Elastic-Net Regularized Generalized Linear Models*. In <https://CRAN.R-project.org/package=glmnet>
- Fritsch, S., Frauke, G., & Wright, M. (2019). *neuralnet: Training of Neural Networks*. In <https://CRAN.R-project.org/package=neuralnet>
- Genuer, R., Poggi, J.-M., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, 31(14), 2225-2236. <https://doi.org/10.1016/j.patrec.2010.03.014>
- Goodwin, T. H. (1998). The Information Ratio. *Financial Analysts Journal*, 54(4), 34-43. <http://www.jstor.org/stable/4480091>
- Investing.com. (2024). *Indices*. <https://www.investing.com/indices/>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning: with applications in R* (2 ed.). Springer.
- Jiménez-Valverde, A. (2012). Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*, 21(4), 498-507. <https://doi.org/10.1111/j.1466-8238.2011.00683.x>

- Jørgensen, A. B., Møller, M. K., & Høstrup, R. (2021). *Forecasting of Danish stocks prices using artificial neural networks*. [Aalborg University Business School].
- Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*, 259(2), 689-702. <https://doi.org/10.1016/j.ejor.2016.10.031>
- Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., Team, R. C., Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., & Hunt, T. (2023). *caret: Classification and Regression Training*. In <https://CRAN.R-project.org/package=caret>
- Lee, C. H. G. (2006). *A critique on efficient market hypothesis (EMH): Empirical evidence of return anomalies in 12 U.S. industry portfolios* [B.A., Simon Fraser University]. <https://summit.sfu.ca/item/4296>
- Liaw, A., & Wiener, M. (2022). *randomForest: Breiman and Cutler's Random Forests for Classification and Regression*. In <https://CRAN.R-project.org/package=randomForest>
- Lo, W. A., & MacKinlay, C. A. (1988). Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test. *The Review of Financial Studies*, 1, 41-66. <https://www.jstor.org/stable/2962126>
- Macchiarulo, A. (2018). Predicting and beating the stock market with machine learning and technical analysis. *The Journal of Internet Banking and Commerce*, 23, 1-22. <https://api.semanticscholar.org/CorpusID:158446761>
- Malkiel, B. G. (1985). *A Random Walk Down Wall Street: The Time-Tested Strategy for Successful Investing*. New York: W.W.Norton.
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2023). *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien*. In <https://CRAN.R-project.org/package=e1071>
- Muralidhar, A. (2015). The Sharpe Ratio Revisited: What It Really Tells Us. *Journal of Performance Measurement*, 19. <https://ssrn.com/abstract=2692859>
- Muthukrishnan, R., & Rohini, R. (2016). LASSO: A feature selection technique in predictive modeling for machine learning. <https://doi.org/10.1109/ICACA.2016.7887916>
- Määttä, J. (2020). Stock market prediction with long short-term memory neural networks: Empirical study on Finnish stock market 1999–2020. . https://www.utupub.fi/bitstream/handle/10024/150978/Maatta_Jarkko_ProGradu.pdf?sequence=1
- Nison, S. (2001). *Japanese Candlestick Charting Techniques: A Contemporary Guide to the Ancient Investment Techniques of the Far East* (Second ed.). Penguin Publishing Group. <https://books.google.com.pe/books?id=lo16E0A7eecC>
- Nordnet. (n.d.). *Prisliste*. Retrieved 11.03.24 from <https://www.nordnet.no/no/kundeservice/prisliste>
- Olden, M. (2016). *Predicting Stocks with Machine Learning: Stacked Classifiers and other Learners Applied to the Oslo Stock Exchange* [University of Oslo].
- Pedersen, M. L., & Lyngø, M. B. (2023). *Forecasting Danish stock prices in 2022 under macroeconomic distress using macro indicators - A deep learning approach* [Aalborg Universitet].
- Pinto, J. E., & Podkaminer, E. L. (2019). Using Multifactor Models. In C. Institute (Ed.).
- Prasch, R. E. (2011). The Instability of Financial Markets: A Critique of Efficient Markets Theory. *Edward Elgar Publishing*. https://ideas.repec.org/h/elg/eechap/13978_5.html
- Riedmiller, M., & Braun, H. (1993). A direct adaptive method for faster backpropagation learning: the RPROP algorithm. IEEE International Conference on Neural Networks,

- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2023). *pROC: Display and Analyze ROC Curves*. In <https://CRAN.R-project.org/package=pROC>
- Ryan, J., & Ulrich, J. (2023). *quantmod: Quantitative Financial Modelling Framework*. In <https://CRAN.R-project.org/package=quantmod>
- Sazlı, M. H. (2006). A brief review of feed-forward neural networks. *Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering*, 50(01), 0-0. https://doi.org/10.1501/commua1-2_0000000026
- Sheth, D., & Shah, M. (2023). Predicting stock market using machine learning: best and accurate way to know future stock prices. *International Journal of System Assurance Engineering and Management*, 14(1), 1-18. <https://doi.org/10.1007/s13198-022-01811-1>
- Shostak, F. (1997). In defense of fundamental analysis: A critique of the efficient market hypothesis. *The Review of Austrian Economics*, 10(2), 27-45. <https://doi.org/10.1007/BF02538483>
- Singh, D. G. (2022). Machine Learning Models in Stock Market Prediction. *International Journal of Innovative Technology and Exploring Engineering*, 11, 18–28. <https://doi.org/10.35940/ijtee.c9733.0111322>
- Singh, P., Singh, N., Singh, K. K., & Singh, A. (2021). Chapter 5 - Diagnosing of disease using machine learning. In K. K. Singh, M. Elhoseny, A. Singh, & A. A. Elngar (Eds.), *Machine Learning and the Internet of Medical Things in Healthcare* (pp. 89-111). Academic Press. <https://doi.org/10.1016/B978-0-12-821229-5.00003-3>
- Speiser, J. L., Miller, M. E., Tooze, J., & Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. *Expert Systems with Applications*, 134, 93-101. <https://doi.org/10.1016/j.eswa.2019.05.028>
- Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Ulrich, J. (2023). *TTR: Technical Trading Rules*. In <https://CRAN.R-project.org/package=TTR>
- Wickham, H., Chang, W., Henry, L., Pedersen, M. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., & van den Brand, T. (2024). *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. In <https://CRAN.R-project.org/package=ggplot2>
- Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A Grammar of Data Manipulation*. In <https://CRAN.R-project.org/package=dplyr>
- Wik, H. (2023). *Swedish Stock and Index Price Prediction Using Machine Learning* [Mälardalen university].
- Wilder, J. W. (1978). *New Concepts in technical trading systems*. <https://books.google.no/books?id=WesJAQAAMAAJ>
- Aaser, K. (2023). *Nå er danske Novo Nordisk Europas mest verdifulle selskap*. E24. <https://e24.no/naeringsliv/i/APMREx/naa-er-danske-novo-nordisk-europas-mest-verdifulle-selskap>

Vedlegg

Vedlegg A: Output multipelregresjon.

Tabell A.1: Output multipelregresjonsmodell for Norge

```
Residuals:
  Min       1Q   Median       3Q      Max
-0.05167 -0.01704  0.00213  0.01333  0.05017

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.003218   0.002777   1.159   0.2506
sp500rf      0.141225   0.166493   0.848   0.3993
goldrf      -0.109704   0.076054  -1.442   0.1538
vixrf       -0.021064   0.010531  -2.000   0.0495 *
energyrf     0.301833   0.051655   5.843 0.000000159 ***
healthcarerf 0.030291   0.118623   0.255   0.7992
industryrf   0.207556   0.162437   1.278   0.2057
financialrf  -0.050228   0.077352  -0.649   0.5183
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02322 on 68 degrees of freedom
Multiple R-squared:  0.7464,    Adjusted R-squared:  0.7203
F-statistic: 28.6 on 7 and 68 DF,  p-value: < 0.00000000000000022
```

Tabell A.2: Output multipelregresjonsmodell for Sverige

```
Residuals:
  Min       1Q   Median       3Q      Max
-0.052411 -0.017590  0.001866  0.015740  0.050265

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.002579   0.003071   0.840  0.40410
sp500rf      0.108761   0.183160   0.594  0.55461
goldrf      -0.222532   0.083657  -2.660  0.00974 **
vixrf       -0.008156   0.011585  -0.704  0.48384
energyrf    -0.112329   0.056809  -1.977  0.05206 .
healthcarerf 0.123108   0.130441   0.944  0.34862
industryrf   0.556233   0.178687   3.113  0.00271 **
financialrf   0.114380   0.085094   1.344  0.18336
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02555 on 68 degrees of freedom
Multiple R-squared:  0.7497,    Adjusted R-squared:  0.7239
F-statistic: 29.09 on 7 and 68 DF,  p-value: < 2.2e-16
```

Tabell A.3: Output multipelregresjonsmodell for Finland

```

Residuals:
      Min       1Q   Median       3Q      Max
-0.056218 -0.019696  0.000796  0.018741  0.062312

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.002599   0.003222  -0.806   0.4228
sp500rf      0.252382   0.192176   1.313   0.1935
goldrf      -0.095234   0.087820  -1.084   0.2820
vixrf       -0.025013   0.012155  -2.058   0.0434 *
energyrf     0.021646   0.059591   0.363   0.7176
healthcarerf 0.142555   0.136851   1.042   0.3012
industryrf   0.248128   0.187471   1.324   0.1901
financialrf  0.115958   0.089284   1.299   0.1984
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0268 on 68 degrees of freedom
Multiple R-squared:  0.6961,    Adjusted R-squared:  0.6648
F-statistic: 22.25 on 7 and 68 DF,  p-value: 2.477e-15

```

Tabell A.4: Output multipelregresjonsmodell for Danmark

```

Residuals:
      Min       1Q   Median       3Q      Max
-0.064109 -0.019330  0.001833  0.016556  0.066285

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.005969   0.003469   1.721   0.0899 .
sp500rf      0.060583   0.206672   0.293   0.7703
goldrf      0.007981   0.094431   0.085   0.9329
vixrf       -0.003450   0.013072  -0.264   0.7926
energyrf    -0.093857   0.064096  -1.464   0.1477
healthcarerf 0.333283   0.147186   2.264   0.0267 *
industryrf   0.489730   0.201614   2.429   0.0178 *
financialrf  -0.135928   0.096019  -1.416   0.1614
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02883 on 68 degrees of freedom
Multiple R-squared:  0.5797,    Adjusted R-squared:  0.5365
F-statistic: 13.4 on 7 and 68 DF,  p-value: 9.824e-11

```

