

Maskinlärningsmodeller för att förutse efterfrågan i butiker

Ordbok	3
Sammanfattning	4
1. Inledning	5
2. Maskininlärning	7
2.1 Supervised Learning	7
2.1.1 SVR	9
2.1.2 ElasticNet	10
2.2 Unsupervised Learning	11
2.3 Reinforcement Learning	11
3. Data set	12
4. Byggande av modellerna	13
5. resultat	14
6. Diskussion	16
Referenser	18

Ordbok

Partiskhet = Systematisk tendens som skapar skillnad mellan resultat och fakta i statistisk analys.

Överanpassning = det att en algoritm som har utvecklats med maskininlärning alltför precist speglar just de data som den har tränats på.

Underanpassning = det är att en algoritm som har utvecklats med maskininlärning inte alls lyckas spegla relationen mellan input och output.

Determinationskoefficient = koefficient som visar hur bra en modell passar i ett set av observationer, där 1 är att modellen passar perfekt och 0 att den passar dåligt

Noggrannhet = Hur nära en maskininlärnings algoritm har lyckats förutsäga ett resultat

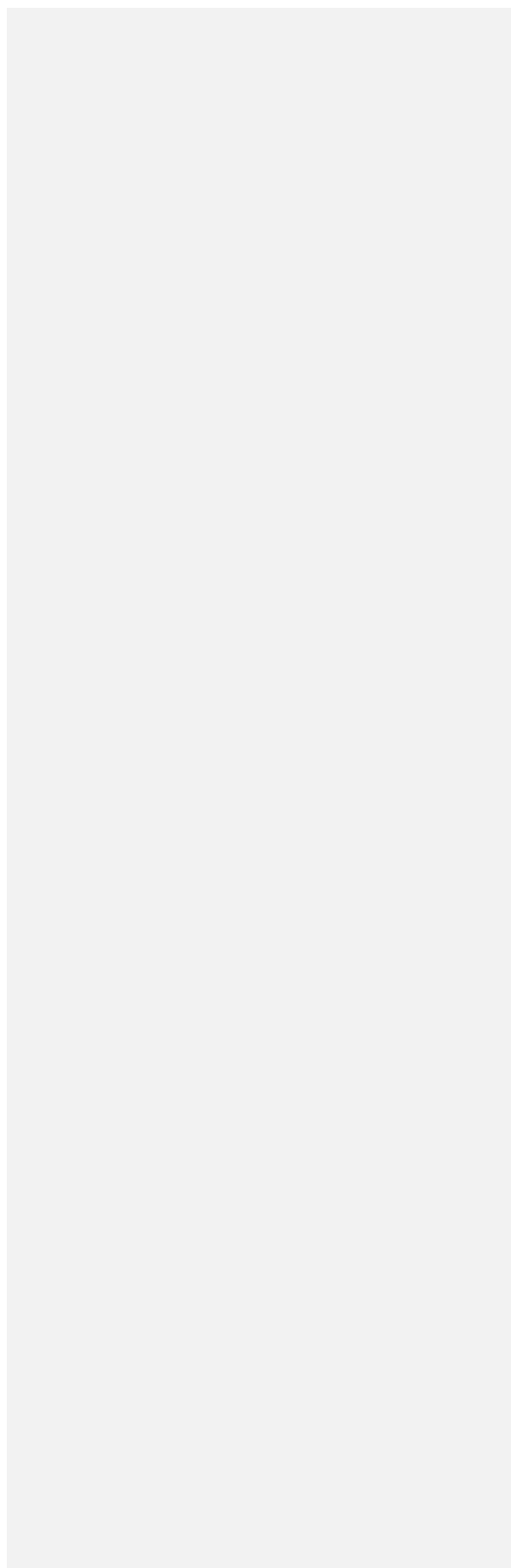
Historisk data = Data som ett företag eller människa har samlat in tidigare men kanske inte använt till något ännu

Gradient descent = Optimerings algoritm som söker ett lokalt minimum/maximum i en given funktion

Riktmärke = ett mätresultat (riktmärket) som används som standardmått vid utvärdering av teknisk utrustning

Sammanfattning

Den här avhandlingen undersöks.....



1. Inledning

Om man vill maximera sin försäljningsvinst säger det sig själv att man vill köpa in just så många produkter som det kommer att gå åt. Man vill inte köpa in för mycket eftersom man kanske måste slänga bort produkter. Men man vill heller inte köpa in för lite eftersom det betyder att man kunde ha fått mera sålt, samt det gör kunderna besvikna. Speciellt svårt kan det här vara vid högtider när någon säsong vara går mycket åt, men sedan säljs den inte alls resten av året. Då blir det speciellt viktigt att kunna förutse hur mycket av varan kommer att gå åt.

Om man ser på t.ex. matbutiker vid julen. Några veckor innan julafton kommer det att säljas stora mängder julskinka, men genast efter att julen är över kommer det inte att säljas speciellt många skinkor resten av året. Då behöver matbutiken på något sätt räkna ut hur många julskinkor som kommer det att säljas under de här helgerna, och köpa in rätt mängd. Det här sköts av en inköpare, en person i butiken som har erfarenhet av arbetet som köper varor till butiken.

Om man tar julskinkan igen som exempel. Traditionellt sett kommer den här inköparen att se hur mycket julskinka det gick åt de tidigare åren, och sedan med hjälp av sin tidigare erfarenhet och intuition göra en gissning på hur mycket julskinka kommer att gå åt det här året. Det här fungerar åtminstone för mindre butiker där man har en inköpare för en avdelning. Men för större butiker blir det problem där en inköpare måste hålla reda på mycket historiska data för många produkter och gör flera beslut om dessa produkter. Förstås har hen olika datorprogram som håller redan på den här data, men besluten och estimeringen sköter hen själv. Enligt Arvan, M, [1] så är det just i försäljningskedjan som den mänskliga faktorns fel oftast kommer fram. Desto fler och större beslut som måste göras desto alvarligare blir felen, och flera automatiserade verktyg behövs för att stöda människan i fråga. Därför skulle det vara ett alternativ att använda sig av maskininlärningsmetoder för att kunna eliminera en del av den mänskliga faktorn i beslutsprocessen.

Med automatiserad datainsamling och databehandling kommer det faktum att man kan förutse efterfrågan automatiskt. Exakta och kvalitativa data gör att påfyllning av lagrens kan möta framtidens behov. Det som alla matbutiker har gemensamt och är en svaghet hos dem är att de måste förnya sina varor innan de blir gamla. Det här tillsammans med att folk är mera klimatmedvetna och vill ha varor som är hanterade på ett sådant sätt att de inte skadar miljön, gör

att det blir viktigt att förutse konsumtions trenderna hos kunderna. Det här kan man göra tack vare att man på senare tid har börjat automatisera processer inom butikerna och samla in stora mängder data för att underlätta att kunna göra rätt beslut. Data driven prognoser gör att alla delar försäljnings kedjan kan förutse vilka produkter som skall tillföras till butiken så det matchar efterfrågan. Felix weber och Reinhard Schütte [2] undersökte i sin artikel hur butiker inom detaljhandeln använder sig av maskininläring och för vilka typer av problem. Där märkte dom att det som butiker mest använder maskininläring för är för att göra förutsägingar (*se bild 1*). Målet med förutsägingarna är att bättre kunna förbereda sig på vad som kommer att hända och därmed vad kommer att behövas. Dessa analyser baserar sig på historiska data som företaget ifråga har samlat in.

Problem Type	Retail Use-Case	Share of Use among top Retailers	
Classification	Freshness and markdown detection	20%	80%
Prediction	predicting the sales data, predict shelf-out scenarios, personalized medicine	80%	20%
Clustering	Customer segmentation, advertisement, personalized communication	10%	90%
Optimization	optimize in the supply chain, logistics, automatization, pricing	30%	70%
Anomaly Detection	medical abnormalities, detect fraudulent offers or scam and fake reviews	30%	70%
Ranking	Product listings, advertising	10%	90%
Recommendation	Customer personalization, cross- and up-selling	20%	80%

In Use
Not Used

Bild 1: Överblick av problemtyper, maskininlärnings tekniker och hur stor andel av butiker använder maskininläring för dessa typer av problem

Vi kan alltså se att butiker använder sig av maskininläring redan till olika grader redan för att underlätta vissa delar av försäljningen och marknadsföringen (*se bild 1*). Det verkar finnas belägg för att maskininläring kan användas i butiker.

2. Maskininlärning

Maskininlärning är gren av AI (artificiellintelligens), vilket är definierat som att en maskin kan imitera mänskligt tankesätt. Den försöka lösa ett komplext problem på samma sätt som en människa skulle göra. Det är datorn som lär sig programmera sig själv med erfarenhet och data. Programmeraren behöver inte specifikt programmera programmet att lära sig, utan den gör det själv efterhand. [3] Som vi kan se i bild 2 så kan man vidare dela in maskininlärning i tre stycken grenar, övervakad inlärning ,övervakad inlärning och Förstärkningsinlärning. Dessa under kategorier av maskininlärning lämpar sig olika bra för olika typer av problem. För den typen av problem som jag ser på i den här artikeln, kommer fokus att vara på övervakad inlärning eftersom jag tror det kommer lämpa sig bäst.

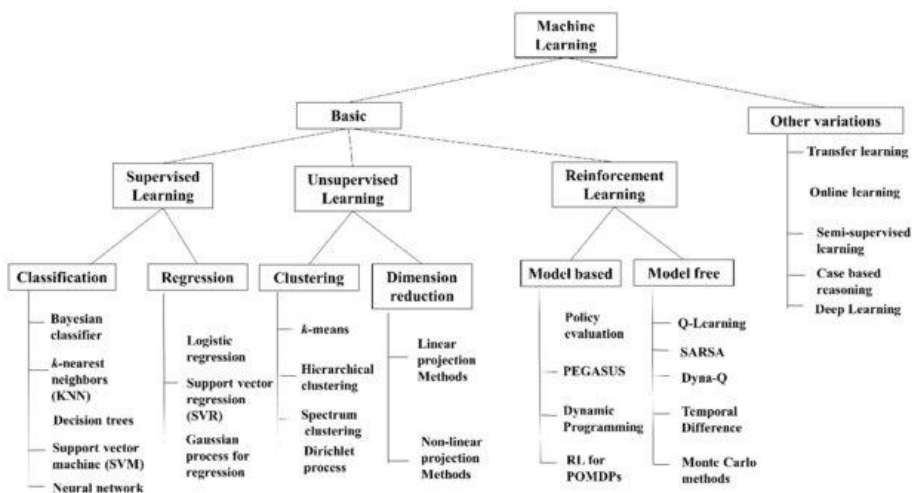


Bild 2: Maskininlärningens taxonomi (Comprehensive Survey of Machine Learning Approaches in Cognitive Radio-Based Vehicular Ad Hoc Networks, Mohammad Asif Hossain)

2.1 övervakad inlärning

övervakad inlärning består av en funktion som lär sig att kartlägga en relation mellan attribut (input) och dess output. [5] Man gör ett tränings set som innehåller alla inputtar och de korrekta outputten. Som sedan modellen tränar sig på för att bli mera precis tills man testar den på ett test

set. Det man försöker göra är att få modellen att lära sig tillräckligt bra på tränings setet att när den får ny input data skall den kunna förutse outputen tillräckligt bra. övervakad inläring behöver övervakade eller märkta data för att kunna fungera, den kan inte som oövervakad inläring hitta egna mönster i data.[6] Man kan vidare dela in övervakad inläring i klassifikation och regression (bild 2). Klassifikation handlar om att sätta data in i specifika kategorier den hittar specifika entiteter i data och försöker dra en slutsats på hur dessa entiteter skall definieras, (se bild 4) några vanliga algoritmer är decision trees, k-nearest neighbor(KNN), support vector Machines (SVM) (se bild 2). Regression fungerar bättre när man vill hitta hur en variabel påverkar en annan. Outputn ger ett nummer inom en numerisk räckvidd [7] Några vanliga algoritmer för regressions problem är Support vector regression (SVR) och logistic regression och Random Forest.

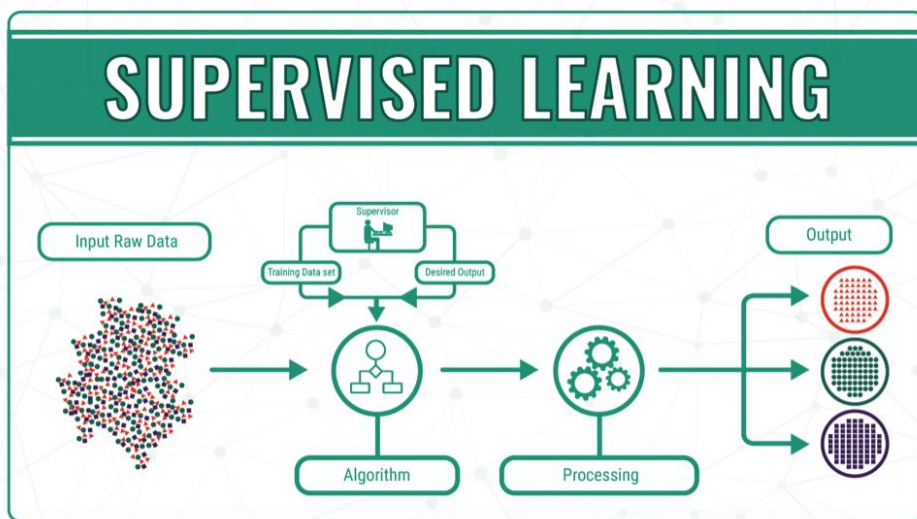


Bild 3: övervakad inläring (<https://medium.com/@michaelgraw/putting-supervised-and-unsupervised-learning-to-work-for-your-business-c7bb68f50efa>)

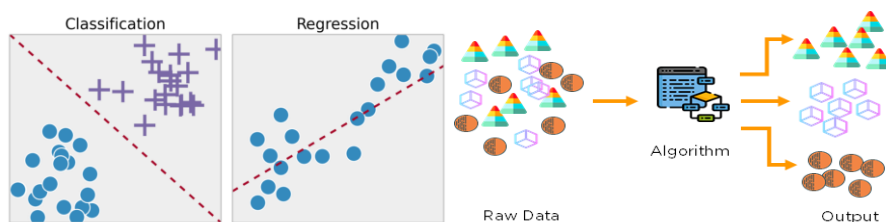


Bild 4: Regression och Klassifikation

Bild 5: oövervakad inläring

$$\text{MIN} \sum_{i=1}^n (y_i - w_i x_i)^2 \quad |y_i - w_i x_i| \leq \epsilon$$

Bild 6: OLS ekvation

Bild 7: SVR val av epsilon

2.1.1 SVR

I många regressionsmodeller är målet att minimera summan av kvadratfelen, d.v.s. skillnaden mellan det observerade värdet och värdet man försöker förutsäga.[8] Om man ser på vanliga minsta kvadratekvationen (OLS, se bild 6) där y är det man försöker förutsäga, w är koefficienten och x är egenskaperna (engelska features). Med den här funktionen försöker man minimera kvadrat felen och den kommer att ge dig en linje som i bild 4 genom ditt data.[9] Med SVR är det möjligt att välja noggrannheten ϵ (se bild 7 och 8) för att hitta en mera lämplig linje genom data.[10] SVR är en av metoderna som kommer att användas för att undersöka om man kan använda maskinlärning för att förutsäga efterfrågan. Detta eftersom den är bra på att hantera extremvärden vilka kommer att existera, eftersom vissa veckor kommer det att säljas mera av en produkt p.g.a. helgveckor som t.ex. påsk. Den är även mycket bra på att generalisera, med hög precision. En nackdel är att den behöver mycket tränings data för att fungera optimalt, men förhoppningsvis kommer den data som ges att räcka till för att ge ett bra resultat [11] Vidare visade Nazish A, Muhammad I [12] I sin studie över vilka maskinlärnings regressorer som fungerar bäst för att förutsäga priser på aktiemarknaden att just SVR hade högre noggrannhet än sina motparter som besluts träd och klassifierings algoritmer. Vilket också stöder teorin att regression är rätt

riktning att ta med den här studien.

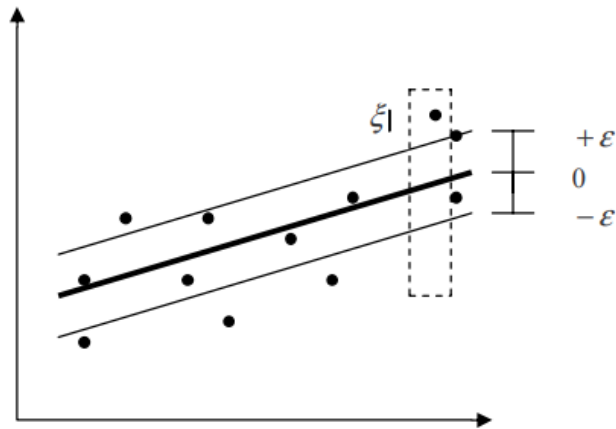


Bild 8: SVR

2.1.2 ElasticNet

ElasticNet är en utveckling av Ridge och lassoregression, som i sin tur är mera avancerade versioner av OLS (se bild 6). OLS fungerar som vanlig statistisk analys där man försöker hålla partiskheten så låg som möjligt. Men det här kan leda till att variansen i en modell blir väldigt stor. Både Ridge och lassoregression löser det här problemet med regularization. Vilket betyder att man gör variansen mindre genom att introducera lite partiskhet. Ridge regression gör det här genom att minska på modellens komplexitet. Den straffar värden som är för långt från 0 vilket gör den mera kompakt utan att förlora värden i modellen. Den gör det här genom att se på summan av kvadrat felen (se bild 9). Lasso gör samma sak men på dess absolut belopp (se bild 10) Dessa två modeller har båda sina för- och nackdelar och fungerar olika bra till olika problem. ElasticNet kombinerar dessa två metoder för att få det bästa av två världar.[13] En stor del av regressionsanalys handlar om att avväga hur mycket partiskhet och varians man vill ha i sin modell. Ols som tidigare sagts ger en hög varians men låg partiskhet, det här vill man undvika. Därför borde ElasticNet som tar de goda sidorna av ridge och lasso att ge ett bättre resultat. Den är även bra på att undvika under- och överanpassning, vilket är en egenskap som den här studien kommer att dra nytta av. [14] I samma studie av Nazish A, Muhammad I som redan nämndes, visade också att just Ridge, Lasso

och ElasticNet gav resultat som var bättre än besluts träd och klassificerings algoritmer. ElasticNet är också det säkrare alternativet av Ridge och Lasso eftersom det undviker nakdelarna av båda två. Det här ställs in med en alpha variabel i ElasticNet.

$$\lambda \sum_{j=1}^p \beta_j^2$$

Bild 9: Ridge regression

$$\lambda \sum_{j=1}^p |\beta_j|$$

Bild 10: Lasso regression

2.2 oövervakad inlärning

oövervakad inlärning använder algoritmer i omärkta data set för att hitta mönster och grupperingar, utan att en människa är involverad (se bild 5).[15] oövervakad inlärning kan vidare delas in i Clustering och Dimension Reduktion. Clustering grupperar omärkt data tillsammans baserat på dess likheter och olikheter, den kan hitta mönster i datat och gruppera den därefter, som man ser i bild 5. Några vanliga algoritmer är A-means och Hierarcical clustering. Mycket data leder ofta till mera precisa resultat men det kan också leda till ett vanligt fel som kallas till overfitting. För att undvika det här kan du använda dig av Dimension reduction algoritmer. Dessa reducerar mängden data inputs tillika medan den håller integriteten av data setet helt. Dessa använd mest när man bearbetar datat i ett problem.[16]

Commented [JN1]: nytt ord?

2.3 Förstärkningsinlärning

Förstärkningsinlärning handlar om att en agent t.ex. ett datorprogram eller en robot som lever i någon slags miljö, där den måste hitta den optimala beslutsprocessen för att lösa ett problem inom miljön. T.ex. försöka hitta ut ur en labyrint på snabbaste sätt. Men agenten behöver nödvändigtvis inte få all information om miljön utan kan agera på ett sådant sätt att miljön påverkas på ett positivt eller negativt sätt. Agenten blir därmed belönad om den gör något positivt, och blir straffad om den gör något negativt. Målet är att agenten lär sig maximera den mängd belöning den får för att uppnå ett optimalt resultat.[17] Förstärkningsinlärning kan vidare delas upp i modell baserad och modell fri Förstärkningsinlärning. Modell baserad RL försöker förstå hela miljön och göra bäst möjliga agerande. Modell fri RL försöker inte förstå hela miljön, utan föröker förut säga hur en

handling kommer vara bra eller dålig utifrån själva agerande mot hurdant tillstånd agenten är i just nu.[18]

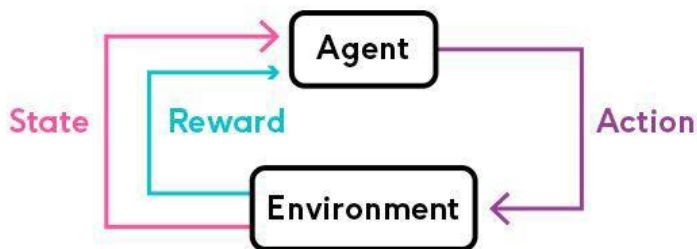


Bild 6: Förstärkningsinlärning

3. Data set

Data setet som kommer att användas för undersökningen är Sales_Transactions_Dataset_Weekly[19] Data Set (se bild 7). Det innehåller 819 produkter, som är benämnda som P1 – P818, samt hur mycket av dem som har sålts vecka för vecka. Veckorna är representerade som “W0 – W51”, alltså det finns försäljnings data för 1 år. Vidare finns de normaliserade värdet för varje vecka representerat i kolumnerna “Normalized 0 – Normalized 51”. Det normaliserade värdet 0 kommer att representera det minsta värdet för en produkt som sålts under veckorna medan 1 kommer att representera det högsta. Detta kan behövas eftersom många regressions algoritmer använder sig av gradient descent vilket vill använda sig av data som har en skala när den optimeras, vilket betyder att dessa kolumner kommer att hjälpa dessa typer av algoritmer att arbeta snabbare.

Product_Code	W0	W1	W2	W3	W4	W5	W6	W7	W8	...	Normalized 42	Normalized 43	Normalized 44	Normalized 45	Normalized 46	Normalized 47	Normalized 48	
0	P1	11	12	10	8	13	12	14	21	6	...	0.06	0.22	0.28	0.39	0.50	0.00	0.22
1	P2	7	6	3	2	7	1	6	3	3	...	0.20	0.40	0.50	0.10	0.10	0.40	0.50
2	P3	7	11	8	9	10	8	7	13	12	...	0.27	1.00	0.18	0.18	0.36	0.45	1.00
3	P4	12	8	13	5	9	6	9	13	13	...	0.41	0.47	0.06	0.12	0.24	0.35	0.71
4	P5	8	5	13	11	6	7	9	14	9	...	0.27	0.53	0.27	0.60	0.20	0.20	0.13
...
806	P815	0	0	1	0	0	2	1	0	0	...	0.00	0.33	0.33	0.00	0.00	0.33	0.00
807	P816	0	1	0	0	1	2	2	6	0	...	0.43	0.43	0.57	0.29	0.57	0.71	0.71
808	P817	1	0	0	0	1	1	2	1	1	...	0.50	0.00	0.00	0.50	0.50	0.00	0.00
809	P818	0	0	0	1	0	0	0	0	1	...	0.00	0.00	0.00	0.50	0.50	0.00	0.00

Bild 7: Datas set weekly sales

Det här data setet kommer att representera hur mycket en butik säljer varje vecka av en produkt. Sedan med hjälp av algoritmerna SVR och ElasticNet undersöks om det går att förutse hur mycket kommer att säljas under en utvald vecka, i det här fallet är vecka 51 valt att vara målet att försöka förutsäga.

4. Byggande av modellerna

Maskininlärnings modellen byggs genom att dela upp ditt data set i tränings data set och ett data set som validerar resultatet. Modellen tränas på träningsdata och sedan testas den mot data setet som validerar det, om man tycker resultatet inte är bra så ändrar man på hyperparametrarna. Sedan ser man hur bra modellen klarar sig på ett på ett testdata set, och ser att modellen fungerar. Hur man delar upp sitt data beror i huvudsak på 2 faktorer, hur många sampel man har i sitt data set och vad frösorts modell man tränar. Men i regel är de ungefär 80% tränings data 20% validations data (se bild 2).[4]

Commented [JN2]: kan behöva förklaring

Commented [JN3]: hitta bättre ord

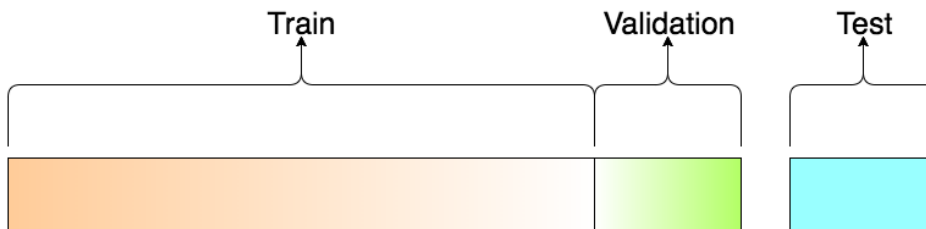


Bild 8: Hur man delar upp data set

När dessa modeller för den här studien byggs kommer de att göras i programmeringsspråket python, och använda sig av biblioteket scikit-learn. Scikit-learn innehåller en rad olika funktioner som låter en både träna och testa sina modeller med olika maskinlärnings algoritmer. Först delas dat på data setet upp i den tidigare nämnda 80% tränings data och 20% testdata.

```
x_train,x_test,y_train,y_test = train_test_split(features,target,test_size=0.2,random_state=7)
```

Där features är alla veckor utom vecka 51 och normaliserade vecka 51 från data setet, och target är vecka 51 alltså det som skall förutses av algoritmerna. X_train och x_test är de variabler som representerar tränings data, medan y_train och y_test representerar testdata. Sedan ställs hyperparametrarna in och modellen tränas.

```
svr = SVR(epsilon=0.2)
svr.fit(x_train,y_train)
e_net = ElasticNet(alpha=1.5,l1_ratio=1)
e_net.fit(x_train,y_train)
```

Efter att modellerna har tränats skall de testas hur bra förutsägnings förmåga de har samt vad dess summa av kvadratfel är.

```
yfit = svr.predict(x_test)
svrscore = round(svr.score(x_test,y_test)*100,2)
mean_squared_error_svr(y_test,yfit)
enypred = e_net.predict(x_test)
mean_squared_error_e_net(y_test,enypred)
e_netScore=round(e_net.score(x_test,y_test)*100,2)
```

I nästa kapitel tas upp vad testerna visade och hur bra modellerna blev.

5. resultat

Båda modellerna visar upp ett nöjaktigt resultat, där ElasticNet var noggrannare och lyckades därmed bättre. Svr hade en determinationskoefficient av 0,8094 och ElasticNet hade 0,8657 vilket tyder på att ElasticNet lyckades skapa en bättre regressionslinje som passade bättre mot data. Båda dessa resultat är acceptabla och vi kan se att ingendera algoritmen har över- eller underanpassats.

Man kan se i bilderna 9 och 10 att ElasticNet har en mera kompakt samling av data runt regressionslinjen än Svr, och därmed lyckats hålla variationerna lägre. Det här syns i bilderna 11

och 12 där man ser att Svr har vissa variationer som är väldigt höga medan den nog har många som är nära 0. ElasticNet har flera små variationer men har nästa aldrig en variation där den har gissat 10 ifrån den riktiga mängden. Man kan vidare se att i bild 13 så har svr problem med att förutsäga mängder som är över 23 utan kommer att göra sin högsta gissning runt just 23. ElasticNet som vi ser i bild 12 har inte det här problemet utan lyckas ganska nära följa kurvan över solda produkter.

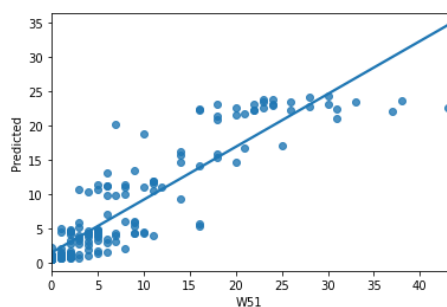


Bild 9: Svr regression

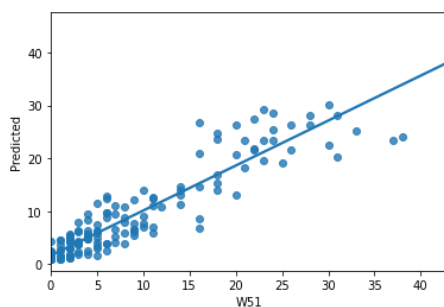


Bild 10: ElasticNet regression

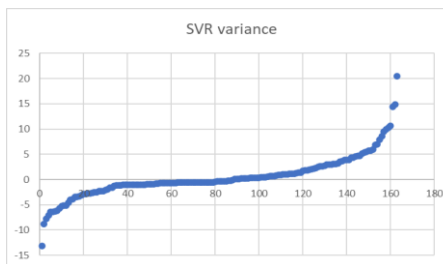


Bild 11: Svr variance

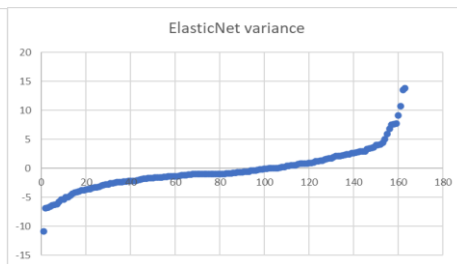


Bild 12: ElasticNet variance

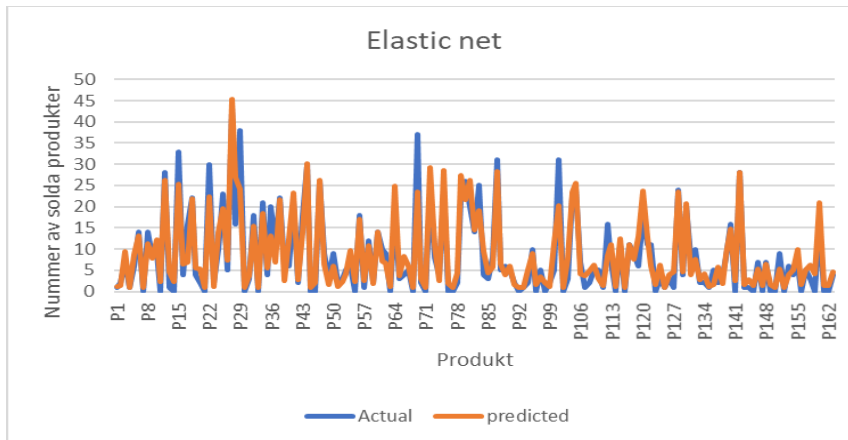


Bild 12

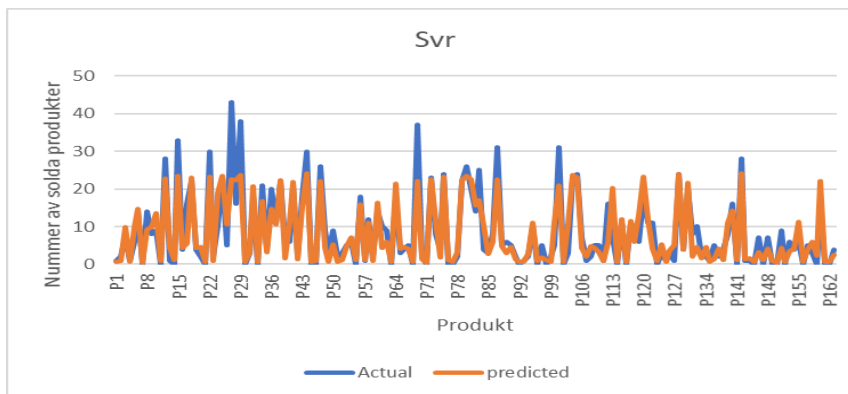


Bild 13

6. Diskussion och avslutning

Frågan blir om man kan använda sig av maskininlärning när man förutser efterfrågan, och om det är ett bättre alternativ än mera traditionella metoder. Eftersom det finns många olika maskininlärnings metoder att använda sig av, t.ex. SVR, som lämpar sig för olika slags analyser, blir frågan vidare vilka kommer att ge det bästa resultatet. Är det värt att använda sig av maskininlärning när man förutser efterfrågan eller blir felmarginalen för stor.

Tuukka S. [20] undersökte i sitt magisters arbete om efterfrågan på produkter kunde förutsägas på basis av hur de hade gått åt i andra butiker. Som riktmärke användes vanlig tidsserieanalys för att testas mot mera komplicerade modeller. Där visade det sig att maskininlärnings algoritmerna klarade sig bättre än tidsserieanalysen. Det här tyder på att det inte bara går att använda sig av maskininlärning för den här typens problem, utan också att de är att föredra.

Som man kan se av resultaten av testen som gjordes, lyckades båda maskininlärnings modeller att göra förutsägingar som ganska nära reflekterade den verkliga mängden av sålda produkter. Även om en människa har hand om att försöka förutsäga efterfrågan på en produkt, är det nästan omöjligt att kunna veta exakt hur mycket en vara kommer att gå åt. Man måste ta i beaktande att modellerna inte ger exakt vad som kommer att gå åt utan gör en vägd gissning på hur mycket som kommer att gå åt. Det man såg i Tuukka S. arbete var att även vanliga statistiska analysmetoder har svårt att ge ett exakt tal på hur mycket som kommer att gå åt, och om en människa skulle göra uppskattningen skulle den också ha det svårt att få förutsägningen exakt rätt på alla produkter. Med det här i baktanken tycker jag att resultaten man ser av den här undersökningen är acceptabla.

Referenser

Bild 9,10: <https://towardsdatascience.com/l1-and-l2-regularization-methods-ce25e7fc831c>

Bild 8: Debasish Basak, Srimanta Pal, Dipak Chandra Patranabis. Support Vector Regression Article in Statistics and Computing, november 2007

Bild 6,7: <https://towardsdatascience.com/an-introduction-to-support-vector-regression-svr-a3ebc1672c2>

Bild 6 : <https://medium.com/@vishnuvijayanpv/what-is-reinforcement-learning-e5dc827c8564>

Bild 5 : <https://medium.com/analytics-vidhya/beginners-guide-to-unsupervised-learning-76a575c4e942>

Bild 4: (<https://towardsdatascience.com/supervised-vs-unsupervised-learning-14f68e32ea8d>)

[1] Arvan, M., Fahimnia, B., Reisi, M., Siemsen, E. (2018). Integrating Human Judgement into Quantitative Forecasting Methods: A Review. Omega.

[2] Felix Weber, Reinhard Schütte (2019) A Domain-Oriented Analysis of the Impact of Machine Learning—The Case of Retailing.

[3] <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained> (Mars 2022)

[4] <https://towardsdatascience.com/how-to-build-a-data-set-for-your-machine-learning-project-5b3b871881ac> (Mars 2022)

[5] Agnieszka, Ł., Volker T. (2014) Introducing Machine Learning Chapter ,

[6] Qiong L, Ying W. (2012) Supervised Learning .January

[7] <https://www.edureka.co/blog/supervised-learning/> (Mars 2022)

[8] https://hlab.stanford.edu/brian/error_sum_of_squares.html (Mars 2022)

- [9] <https://towardsdatascience.com/an-introduction-to-support-vector-regression-svr-a3ebc1672c2> (Mars 2022)
- [10] Debasish B, Srimanta P, Dipak P. Support (2007) Vector Regression Article in Statistics and Computing,
- [11] Mariette A, Rahul K (2015) Support Vector Regression In book: Efficient Learning Machines
- [12] Nazish A, Muhammad I, Dr. Zubair N (2021). A comparative study of Different Machine Learning Regressors For Stock Market Prediction
- [13] <https://www.datacamp.com/community/tutorials/tutorial-ridge-lasso-elastic-net> (Mars 2022)
- [14] <https://towardsdatascience.com/bias-variance-and-regularization-in-linear-regression-lasso-ridge-and-elastic-net-8bf81991d0c5> (Mars 2022)
- [15] Saman Siadati. (2018) What is Unsupervised Learning,
- [16] <https://www.ibm.com/cloud/learn/unsupervised-learning>
- [17] V. Heidrich-Meisner, M. Lauer, C. Igel and M. Riedmiller (2007). Reinforcement learning in a Nutshell
- [18] <https://blog.auberginesolutions.com/model-based-vs-model-free-reinforcement-learning/> (Mars 2022)
- [19] https://archive.ics.uci.edu/ml/datasets/Sales_Transactions_Dataset_Weekly (Mars 2022)
- [20] Tuukka S (2020) Predicting product sales in retail store chain
Predicting product sales in retail store chain (aalto.fi)