Norwegian School of Economics
Bergen, Fall 2020

Predicting Housing Prices With Machine Learning

A macroeconomic analysis of the Norwegian housing market

Matej Dimoski & Markus Pettersen

Supervisor: Torfinn Harding

Master thesis in Economics and Business Administration
Major: Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

This thesis explores the applicability of machine learning in macroeconomic housing price predictions in Norway. We apply three machine learning models Elastic Net, Random Forest and Neural Network on historical time-time series data and predict quarterly and yearly growth rates between 2013 and 2019. The performance is evaluated upon predictions from Norges Bank, DNB and SSB.

Our results indicate that machine learning can produce predictions with the same accuracy as professional institutions. Among the machine learning models, Elastic Net produces the most accurate quarterly predictions. Compared to Norges Bank, Elastic Net's predictions are more accurate in 29,6% of the quarters, but less precise in the overall evaluation. Large deviations during 2018 and 2019 are decisive for the lacking performance, after new mortgage regulations were introduced from Finanstilsynet. Random Forest predicts the most accurate yearly predictions but is outperformed by Norges Bank. Still, Random Forest surpasses both DNB and SSB throughout the evaluation process.

The thesis contributes to the existing literature in several aspects. First, by outperforming housing experts, we challenge traditional macroeconomic approaches in the choice of predictive models. Second, our results indicate that linear models are more suited in shorter time spans, while nonlinear models perform better over longer horizons. Third, the machine learning models have identified household debt as the most influential variable to determine the housing prices in Norway. Overall, we believe machine learning approaches could become valuable in further academic and professional macroeconomic research.

Keywords - Machine Learning, Prediction, Forecasting, Housing Market, Macroeconomics

Preface

This thesis is written as a part of our Master of Science in Financial Economics and marks the end of our time at NHH. Writing the thesis has certainly been a worthy obstacle and a humbling exercise in persistence. Through this process we have developed a profound respect for the effort needed to produce presentable and reliable results.

Initially we chose the topic due to our common interest in macroeconomics and a curiosity towards new technologies. In addition, we were inspired by the master thesis from Bankson and Holm, written in 2019, predicting the GDP-growth in Norway using machine learning. We are thankful for using their work as inspiration. Our ambition is that the thesis could contribute to further professional and academic research, raising the applicability of machine learning.

We would like to extend our sincere gratitude to our supervisor, Torfinn Harding, for his apt comments and guidance through the writing process. Further, we would like to thank Bjørn Naug and Adnan Muneer in Norges Bank, for providing useful data and expert knowledge in the housing market. Also, thank you Genaro Sucarrat, from BI, and Oddmund Berg, from DNB, for valuable conversations and knowledge. Lastly, we want to thank friends and family for support and motivation through the process of writing this thesis.

Norwegian School of Economics

Bergen, December 2020

Matej Dimoski

Markus Pettersen

Contents

1. Introduction

The housing market has important economic implications for the wellbeing of a nation. Residential real estate provides shelter, ensures household savings and is one of the main drivers in the Norwegian economy through finance and construction. Therefore, providing accurate predictions is just as important for the central bank, property investors and house owners, as for political decision makers. The thesis explores the ability of machine learning models to predict quarterly and yearly aggregated housing growth rates in Norway. We evaluate the performance with professional institutions, respectively Norges Bank, DNB and SSB. More specifically, we employ Elastic Net, Random Forest and Neural Network to produce out-of-sample predictions between 2013 and 2019. Predictive accuracy and direction are computed and evaluated towards the professional institutions. Moreover, economic intuitions and influential events in the Norwegian Housing market are analyzed through the lens of machine learning. Lastly, we evaluate model predictability and specifications from our analysis, with national and international research.

Predicting housing growth rates have several challenging aspects. First, only short lengths of macroeconomic time series data are currently available. Housing price indices (HPI), and influential indicators, are usually computed at monthly or quarterly frequencies, with limited historical data. This limits the size of the dataset, creating difficulties with model building and testing. Second, uncertain booms and busts indicate a degree of nonlinear effects in the market (Miles, 2007). Therefore, traditional models could be less suited in capturing underlying relationships. Third, due to low frequencies in sales, transaction costs and individual attributes, the housing market has previously proven to have a high degree of heterogeneity (Vanags et al., 2017).

Despite the challenges, a large portion of institutions, banks and housing experts are frequently voicing their future projections. Therefore, it is interesting to evaluate attributes between institutions. Furthermore, professional institutes publish their estimates for different reasons. DNB Markets' projections are a part of their overall macroeconomic overview, aiming to help businesses and investors with their investment decisions. Norges Bank's objective is to ensure the correct fiscal and monetary policies, where the housing market is considered a leading indicator in the macroeconomic environment. Differences in objectives could likely influence the methodologies and perhaps the prediction accuracy.

Traditional statistical models have been fundamental in existing prediction analysis. For example, Norges Bank rely on their SAM (System of averaging models), while DNB apply regression models to produce their predictions. Combined with expert perspectives, final predictions are produced. Machine learning has previously increased its field of application within various sectors, with the sole purpose of increasing efficiency (Jung et al., 2018). Ideally, machine learning is able to analyze data quicker, cheaper, more systematic and find unobservable correlations that the human eye, and traditional statistical methods, might oversee. Still, the technology has not yet been exploited to its full potential in the macroeconomic environment. If the machine learning models are suitable for the purpose of predicting housing prices, and could yield credible results, they could potentially serve as supportive tools in further research and discussions.

Due to the macroeconomic importance and an interest in the housing market, we will analyze the following question: *Is machine learning suitable for predicting price growth in the Norwegian housing market?*

Our results indicate that machine learning can produce as accurate predictions as professional institutions. However, abnormal events decrease the overall performance. Elastic Net produces the most accurate quarterly predictions. Compared to Norges Bank's predictions, Elastic Net is more accurate in 29,6% of the quarters, but less precise overall. Especially large deviations after the introduction of new mortgage regulations, affect the Elastic Net's performance during 2018. Random Forest produces the most accurate yearly predictions among the machine learning models. Compared to Norges Bank, Random Forest is outperformed over the whole period, but produces the most accurate predictions in 2013 and 2017. Additionally, Random Forest's predictions have a higher accuracy than both DNB and SSB, outperforming two professional institutions.

Our thesis contributes to the existing literature in several ways. Firstly, by including professional institutional predictions, we raise the benchmark for what could be considered as efficient results. In the thesis, efficient results indicate predicted values are closer than, or as close as, professional institutions projections to the actual growth. Existing literature has focused on statistical models for their performance evaluation. Discovering that machine learning can surpass housing experts, we challenge traditional macroeconomic predictions. Secondly, our results support previous literature stating that linear models are more suited in shorter time spans, while nonlinear models perform better in longer horizons (Gupta & Miller, 2015; Milunovich, 2019). The consistency contributes to further model building and explores interactions and combinations between explanatory variables. Thirdly, our results are consistent with previous macroeconomic research on the predictability in the Norwegian housing market (Røed Larsen & Weum, 2008). Still, market complexity and uncertainties negatively affect the predictive ability. Lastly, our approach enables one to evaluate the importance of influential factors. Household debt stands out as the variable with the highest contribution towards determining the housing prices in Norway.

The remainder of the thesis is organized as follows: In section 2, we present background and the most relevant existing literature. Section 3 describes the dataset, as well as relevant adjustments and assumptions. Section 4 briefly presents the relevant machine learning methodology. Section 5 presents model implementation, and section 6 describes the quarterly and yearly results from the predictions. Section 7 is twofold, discussing both machine learning performance and relevant aspects in the housing market. Ultimately, the last section concludes our findings.

2. Background

Artificial intelligence and machine learning have previously been explored in the Norwegian housing market. In 2017, a robot algorithm from Veidekke won Boligtempen[1] with the most accurate housing growth forecast (Finansavisen, 2018). Their model which included publicly available historical data surpassed 20 housing experts. However, after gathering excessive publicity, the robot predicted 2% lower prices for 2019, misinterpreting the growth direction and magnitude. In 2019, Mari Mamre, Doctoral Research Fellow from NMBU, developed a Neural Network model for housing price predictions in Norway and Oslo (Stranden, 2020). The model consists of 50 explanatory variables, and aims to capture regulatory changes, macroeconomic events, and housing specific factors. We therefore recognize the topic as relevant both for professional and academic institutions.

The thesis' purpose is to explore modern technology and test its relevance for Norwegian businesses and policymakers. Thus, we imply simple machine learning models to enlighten their applicability. We have intentionally selected machine learning models that are easily implemented and understood. Therefore, the thesis is relevant both for those with limited knowledge of machine learning, as well as industry experts. Additionally, we shed light on strengths and weaknesses of the implementation, and potential improvements for further models.

2.1 Litterature Review

The thesis contributes towards a small, but growing, list of literature on machine learning in the housing market. Various literature on housing price predictions, including traditional regression and autoregressive models, are available. In this section, we only present corresponding machine learning literature. For example, housing prices are generally influenced both by macroeconomic[2] and microeconomic[3] factors (Lam et al., 2009). Since the thesis' purpose is predicting the aggregated growth in Norway, it belongs to the macroeconomic field. Thus, the literature review focuses on conducted macroeconomic research to predict housing price indices on national or state levels.

[1] Finansavisen's competition for determining the best housing price forecaster of the year
[2] Attributes that describe the social and economic situation
[3] For example, location, esthetics and neighborhood attributes

Rene de Borst (1991) implemented the first Neural Network in housing price predictions in New England. His results significantly outperformed multiple OLS-regression, arguing Neural Networks could become the next calibration technology in the housing market. The research indicated machine learning's relevance, even in its early stages.

Neural Network predicted Property Price Indices in Malaysia (Shukry et al., 2012). The indicators unemployment rate, population size, interest rate and household income were included in the model. A quarterly training set from 2000 to 2009 was extracted and tested out-of-sample in 2010 and 2011. Neural Network produced a Mean absolute percentage error (MAPE)[4] of 8%, consequently outperforming traditional multiple regression, generating a MAPE of 15%. The researchers argued Neural Network could be a good alternative to traditional multiple regression, allowing for nonlinearity and multicollinearity between indicators.

Elastic Net was included in the variable selection process in a fitted Support Vector Regression (SVR), forecasting the yearly U.S. Real Housing Price Index (Plakandaras et al., 2015). A substantially richer dataset from 1890 to 2012, was extracted. The explanatory variables GDP, interest rate, inflation, construction cost, stock price index, oil price, and budget deficit/surplus were included. A combined linear SVR achieved an out-of-sample MAPE of 2.5% outperforming a Random Walk (5.35%) and Bayesian autoregressive model (5.42%). Plakandaras et al. (2015) argued SVR was better suited as an early warning system for forecasting sudden housing price drops, compared to traditional models.

Neural Network predicted the Property Price Index more accurately than an ARIMA in Hong Kong (Abidoye et al., 2019). A quarterly dataset from 1985 to 2016 was extracted and tested between 2013 and 2016. The out of sample performance from Neural Network generated a RMSE of 7.01, which is substantially lower that an ARIMA of 23.35. Additionally, the researchers claimed interest rate, unemployment rate and household size were the most influential indicators for predicting the Property Price Index. Lastly, they argued Neural Network could help policy makers and property investors predicting booms and busts in the housing market.

George Milunovich (2019) applied 47 different algorithms forecasting the Australian Housing Price Index and growth rates. The algorithms consisted of traditional time-series models, machine learning procedures and deep learning neural networks. Quarterly data from 1972 to 2017 was

---

[4] $MAPE = \frac{1}{k}\sum_{t=1}^{k}\left|\frac{A_t - F_t}{A_t}\right|$. Generates the percentage deviation relatively to the actual growth

utilized to produce predictions one, two, four, and eight quarters ahead. In predictions one and two quarters ahead, Elastic Net ranked number five with regards to Mean Square Error (MSE). Also, most algorithms had significantly more precise estimates than a Random Walk benchmark. For predictions four and eight quarters ahead, some algorithms predicted more accurately than a Random Walk, but the overall performance was weakened. The study concluded that Support Vector Regression (SVR) generated the most precise estimates across all horizons. Additionally, Milunovich recognized a pattern that linear models performed better in shorter timespans, while nonlinear models were preferable given longer horizons.

The overall impression from existing literature indicate machine learning models have predicted more effectively out-of-sample than statistical benchmark models. Despite promising results, the existing literature have not yet focused on further implications in the housing market, through the lens of machine learning. These implications include analyzing influential events, contributing variables and the market predictability.

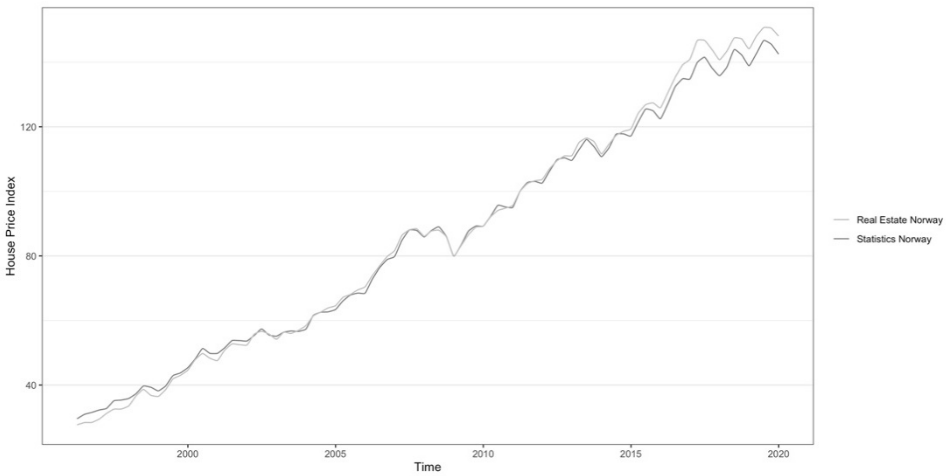
3. Data

Since the time perspective is present in our analysis, the dataset is categorized as time-series. Compared to cross-sectional data, time series are considered more complex for statistical modelling. Due to dimension and order of the observations, the assumption of independence is violated (Wooldridge, 2012). Through the data and implementation section, assumptions and model specifications are elaborated to handle these challenges. The following section provides assumptions and adjustments to the housing price index, included variables and benchmark data.

3.1 The Norwegian House Price Index

Statistics Norway (SSB) and Real Estate Norway (Eiendom Norge) publish indices for the aggregated development in the Norwegian housing market. Both indices are based on secondhand sales from the market platform Finn.no, which covers 70% of the total turnover in the housing market (Real Estate Norway, 2020). The deviations occur in the terms of classifications and weights. Real Estate Norway puts emphasis on *sales weights*, modelling the aggregated transactional values. Meanwhile, SSB prefers *inventory weights* and the value of the whole housing stock. A distinctive assumption in SSB's index imply unsold properties follow the same price trend as sold ones (Lundesgaard, 2019).

Figure 3.1: Housing Price Indices in Norway between 1990 and 2020



Note: Unadjusted housing price indices from Rel Estate Norway and Statistics Norway. Despite following each other closely, the two indices deviated up to 5% around 2016.

The choice of index depends on the problem's objective. If the objective is to measure the price development for properties traded and purchased by households, the Real Estate Norway index is favorable. However, ff the purpose is to measure the total development in the housing stock, SSB's index is preferable. Traded and purchased properties seem to be most common in national and international literature. Furthermore, DNB and Norges Bank base their projections on the Real Estate Norway Index (DNB Markets, 2020; Norges Bank, 2020) Therefore, the Real Estate Norway Index is preferable. Simultaneously, by excluding the SSB Index, their predictions become less comparable in the evaluation section.

Adjustments were necessary to fit the index to our purpose. To better analyze the short-term development, the seasonally adjusted index is extracted, as seasonal effects might conceal the true underlying development in the market (Statistics Norway, 2008). Furthermore, our dataset consists of quarterly observations, while Real Estate Norway publishes monthly indices. We use the same approach as Norges Bank, by computing a quarterly averaged index from the relevant months (Personal communication, 2020). Additionally, Real Estate Norway started publishing their index in 2003, while our dataset dates to 1996. However, Norges Bank constructed a housing price index in the period between 1996 and 2003. These indices are chained as a discontinued index, which Econ Pöyry produced[5].

3.2 Covariates

Our dataset consists of 14 explanatory variables, deducted quarterly from 1996Q1 to 2019Q4. Variable selection is a source of model bias. To decrease including noninfluential indicators, the chosen variables are collected based on their importance in previous housing literature. Including explanatory variables are limited to avoid overfitting. Overfitting might arise due to the model being too fitted to the limited in-sample data points, thereupon the out-of-sample estimates become less accurate (Kenton, 2019). For example, Gupta, et al. (2011) predicted more accurate housing prices with a dynamic model including ten variables, rather than the 120-variable model in the US. Assuming similar patterns in Norway, extracting 14 variables limits seems sufficient to our purpose. Furthermore, international macroeconomic variables have not been included. As housing expert Erling Røed Larsen claimed, the market has been, and will be, local, with local drivers and boundaries (NRK Debatten, 2020). Moreover, The International Monetary Fund (IMF, 2020) proposed two extinctive categories for explaining housing prices: *Business Cycle* and *Housing Specific Factors*. Thus, the chosen variables must fit into one of the sections.

[5] In collaboration with Norges Eiendomsmeglerforbund (NEF), Eiendomsmeglerforetakenes Forening (EEF) and Finn.no

The dataset is grounded in the housing prediction model from Jacobsen and Naug (2004). It consists of Business cycle factors such as *interest rate after tax, unemployment* and *inflation*, and the Housing specific factors *household stock and household income*. Furthermore, the variables *oil price* (Plakandaras, et al., 2015), *national budget surplus/deficit* (Abidoye et al., 2019) and *stock indices* (Milunovich, 2019) are included. A full description of the included variables is presented in **Table A.2.1**.

The extracted variables are either on index, growth or absolute form. Variables collected as indices are computed to quarterly year on year growth. Variables extracted on absolute or growth-form, are implied directly without any computations. Seasonally adjusted explanatory variables are extracted when available. SSB, Bloomberg, Real Estate Norway and Norges Bank have been the main sources and were extracted during the 3rd quarter of 2020. With monthly and daily publications, arithmetic averages are computed to transform the frequency to quarterly observations.

Most traditional time series regression models rely on assumptions regarding stationarity in the data (Palachy, 2019). It implies constant average, variance and covariance between the observations (Wooldridge, 2012). Problems such as spurious correlations could occur if not managed correctly. This is commonly solved by transforming the variables into logarithmic or growth form (Hyndman, 2016). However, machine learning does not require stationary variables. To illustrate differentiating strengths, we have not put emphasis on adjusting the data to ensure stationarity. Consequently, the traditional benchmark models ARIMA and Random Walk are not applicable to our dataset.

Certain of the included explanatory variables have been revised. When revised variables are available, we consistently extract the last publications. Consequently, our dataset is considered more accurate, compared to the available data the institutions possessed. Unfortunately, few actions are available to reduce the informational advantage since unrevised publications have not been found. Revision issues could serve as a potential weakness for the machine learning credibility.

Our approach is only suitable with full datasets. Therefore, the historical starting point depends on the most recent variable's publishing date. At first, the intentions were to extract data back to 1980, with the purpose of increasing the total number of observations. However, it became clear that influential variables were not published during the 1980's. Therefore, a trade-off between prioritizing the number of observations or including variables was weighted. Generally, machine learning prefers to discover frequent complex correlations and interactions given large portions of data. However, in macroeconomic time series forecasting, shorter training periods have been effective. For example, OECD predicted the GDP-growth more accurately out-of-sample, only including 5 years of historical data in their training sets. Due to rapidly changing economies, they claim recent time series are more informative to near future than remote past (Woloszko, 2017). After testing with starting points, the preferable models prioritized variables instead of observations, confirming OECD's perceptions. Our dataset therefore spans from 1996Q1 to 2019Q4.

3.3 Data for Evaluation

In Norway, 5 institutions have produced frequent yearly predictions: DNB, Norges Bank, Real Estate Norway, Statistics Norway (SSB) and The Confederation of Norwegian Enterprise (NHO). Additionally, private institutions such as Nordea, Prognosesenteret, Sparebank 1 and Swedbank have reported housing predictions in the media. Still, these predictions have not been published publicly, limiting their applicability. To include predictions from all the institutions has not been achievable with our format. As mentioned in section 3.1, SSB predicts their inventory weighted index, differentiating the prediction target. SSB`s predictions are therefore only evaluated upon their overall performance and are not further discussed. NHO started producing yearly predictions in 2018, limiting the available number of predictions. Real Estate Norway publishes 12-months growth rates, violating the growth computation, further explained in 3.4. Therefore, projections from DNB and Norges Bank are fundamental in the evaluation section.

Regarding quarterly predictions, Norges Bank published year on year growth rates from 2013Q2 to 2019Q4 in their Monetary Policy Reports (MPR). They predicted the current quarter (nowcasting) and the following quarter (forecasting) simultaneously. We have consistently extracted the predictions for the latter quarter. Additional adjustments were needed in two instances. In MPR 3/17 and 3/19, Norges Bank published monthly predictions. In these instances, we calculate the quarterly arithmetic averages. The 27 extracted predictions are shown in **Table A.1.1**.

Yearly predictions from DNB, Norges Bank and SSB are extracted from 2013 to 2019. To minimize informational advantages, we have consistently extracted publications from the 4th quarter. The institutions are still publishing with one month in between. SSB publishes around the 1st of December, Norges Bank around the 18th of December, while DNB reports on the 17th of January the following year. DNB has therefore an additional month of informational advantage. However, the additional month is not further emphasized in the analysis. All extracted predictions are shown in **Table A.1.2**.

3.4 Problems with growth rate computations

Through the process, we discovered a misconception between journalists and experts, with regards to the interpretability of yearly growth rates. Specifically, distinguishing between average yearly growth- and the 12-month growth rates. The computations can generate opposite conclusions, confusing prediction estimates and historical growth rates. Therefore, we briefly explain the differences.

The Average Yearly Growth Rate computes an arithmetic average of monthly housing price indices, divided by the previous year's average. It represents the growth between average property in year t, and the average property the previous year. Throughout the thesis, the average yearly growth rate is computed.

$$Average\ Yearly\ Growth\ Rate_t = \left(\frac{\overline{HPI_t}}{\overline{HPI_{t-1}}} - 1\right) * 100 \tag{3.1}$$

The 12-month Growth Rate is computed by dividing the index value in December, by the index in January. It represents the development in prices during the last 12 months.

$$12\ month\ Growth\ Rate_t = \left(\frac{HPI\ December_t}{HPI\ January_t} - 1\right) * 100 \tag{3.2}$$

During 2017, the computations generated contrasting overall conclusions. The 12 months growth rate indicated a 4.1% decline in prices for 2017. Computing the average yearly growth rate generated a positive 5.7% compared to 2016 prices. Clearly, two unfortunate effects occur from the computational differences. First, different conclusions are drawn towards the growth direction. Second, predictions become less comparable, since most predictions do not specify computation rate (Senneset, 2018). After personal communication with Norges Bank (2020), and

DNB (2020), both institutions produce average yearly growth rates. Therefore, we follow this computation during yearly predictions.

Predicting the average yearly growth rate yields one advantage. Standing in December of 2017, the index value is 249.0, while the yearly average for 2017 is 240. In 2018 predictions, the expert already knows today's value is above the yearly average. To generate negative growth rates, the price indices need to fall below 240 during 2018. By possessing this knowledge, the next year's growth rates have a higher probability of being positive, compared to the 12th month growth.

4. Machine Learning Theory

Machine learning is an application of artificial intelligence, which provides systems and models to automatically learn from historical experience without being explicitly programmed (Expert System, 2020). Within the field of statistics, machine learning belongs to the class of algorithmic modelling. Compared to traditional regression, less focus is put on the relationship between the dependent and independent variables, since the overall objective is to compute the most efficient predictions.

The machine learning process can be split into two forms: unsupervised and supervised learning (Soni, 2018). Supervised learning refers to cases where there exists prior knowledge of the relationship between variables, and a specific output is requested. Unsupervised learning methods are applied to discover relationships between variables without prior knowledge. This enables the latter to discover hidden structures and combinations within the data. Since the thesis` output variable is specified, supervised learning is preferred. Furthermore, the machine learning process consists of two main elements: One learning process that best fits the independent variables to the dependent variable, and an algorithm that, based on the learning, models the relationship between the two categories of variables (Jung et al., 2018).

Our models have previously been applied in similar studies, thus ensuring applicability and relevant comparisons. The chosen models also cover sub-groups of machine learning: Elastic Net combines two linear models. Random Forest is an ensemble nonlinear model computed from multiple decision trees, and Neural Network is a nonlinear Black-Box[6] structure. The relative performance between the machine learning models could illustrate strengths and weaknesses towards further model specifications. In our analysis, we have not put emphasis on whether the chosen models provide possibilities to analyze its interpretability.

4.1 Elastic Net

Elastic Net was developed at Stanford University in 2005 (Zou & Hastie, 2005). The model builds on the Ordinary Least Squares model (OLS), while including additional penalty terms from Lasso

[6] Black-Box refers to models where less knowledge of the model's internal workings is provided, and where its interpretability is less available.

and Ridge regressions. The following paragraphs explains linear models and the penalty terms, before the model is presented.

Elastic Net represents linear models in our thesis. In a linear model, parameters are either a constant (β_0), or a parameter (β_i) multiplied by an independent variable (x_j). Therefore, Elastic Net does not capture interactions and combinations of the parameters. However, Elastic Net implies simplicity and interpretability compared to nonlinear models (Frost, 2017).

$$Y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_j x_j \tag{4.1}$$

4.1.1 Ridge Regression

Ridge uses regularization by shrinking included coefficients from OLS. The regularization term is equal to the squared magnitude of the coefficients. By doing this, Ridge reduces coefficients of highly correlated variables. Ridge regression accomplishes to decrease parameter variance without omitting variables.

As with OLS, the objective is to minimize the sum of squared residuals. Additionally, including the penalty term decreases the coefficients that are close to zero (Hoerl & Kennard, 1970). When both the sum of squared residuals and the penalty term are subject to the minimization problem, Ridge achieves the optimal result by shrinking the regressors that are highly correlated. Regressors that explain the same variance will have a lower coefficient than an OLS-estimator.

$$\hat{\beta} = argmin_{\hat{\beta}_j} \left[\left(\sum_{i=1}^{n} \left(Y - X\hat{\beta} \right)^2 + \lambda \sum_{j=1}^{p} \left(\hat{\beta}_j \right)^2 \right) \right] \tag{4.2}$$

Y represents the actual housing price growth, while X is the true value of the explanatory variables. n is the number of observations and p is the number of explanatory variables. The first term represents the traditional OLS-regression, while the penalty term is represented in the second term. The extent of the penalty term is determined by the parameter lambda. The optimal value of lambda is achieved by the cross-validation process.

4.1.2 Lasso Regression

$$\hat{\beta} = argmin_{\hat{\beta}_j} \left[\left(\sum_{i=1}^{n} \left(Y - X\hat{\beta} \right)^2 + \lambda \sum_{j=1}^{p} \left| \hat{\beta}_j \right| \right) \right] \tag{4.3}$$

Lasso operates with a variable selection penalty term, where highly correlated variables are omitted. A higher lambda indicates the threshold for omitting variables is lower. By eliminating highly correlated variables, multicollinearity issues are reduced. A previous critique of Lasso regression is not specifying which correlated variable should be omitted. Therefore, influential variables might be excluded from the model. For example, population can be omitted due to the correlation with the housing stock. Still, both explanatory variables have an individual effect on the housing prices, which will not be captured in a Lasso regression.

4.1.3 Elastic Net model

Elastic Net combines OLS-regression with the penalty terms from Ridge and Lasso. Lambda determines the penalty weighting, and alpha weights the relative penalty between the regressors. A low alpha would prefer the penalty term from Ridge regression. By combining the penalty terms, Elastic Net reduces model variance and eliminates strongly correlated variables.

$$\hat{\beta} = argmin_{\hat{\beta}_j} \left[\left(\sum_{i=1}^{n} \left(Y - X\hat{\beta} \right)^2 + \lambda \sum_{j=1}^{p} [(1-\alpha) \left(\hat{\beta}_j \right)^2 + \alpha \left| \hat{\beta}_j \right|] \right) \right] \quad (4.4)$$

The penalty terms in Elastic Net are ideal in situations with more explanatory variables than observations. A Lasso model bounds to have more observations than variables, which is permitted in Elastic Net (Zou & Hastie, 2005). Influential variables could therefore be omitted, only due to limited observations available in Lasso.

When lambda is zero, the estimator is equal to the OLS-estimator. When lambda is greater than zero, the minimizing coefficient constraints are added. A higher lambda would lower the threshold for minimizing coefficients. Overall, Elastic Net includes both variable selection and regularization in a linear model. Variables with high correlations can be included without increasing the parameter variance. This makes Elastic Net resistant to problems such as multicollinearity.

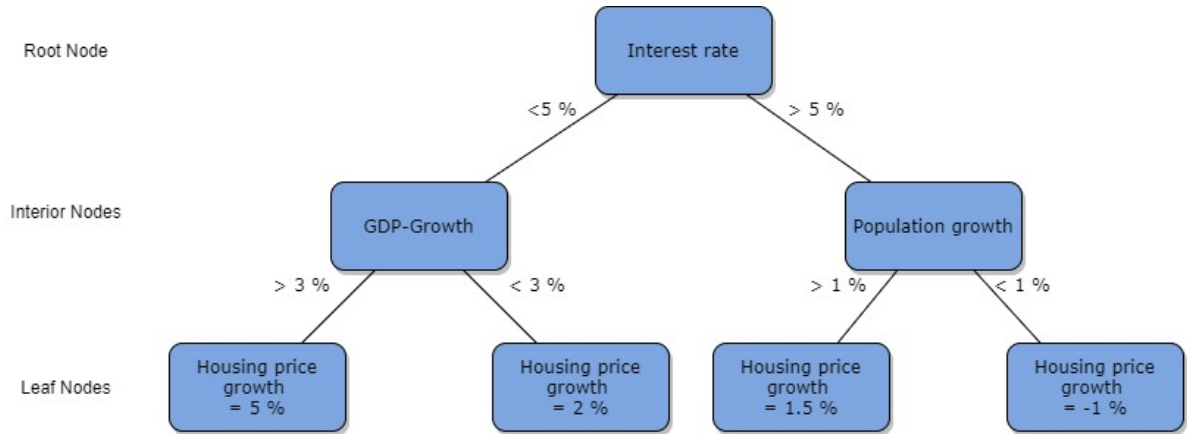
4.2 Random Forest

Random Forest was developed by Professor Leo Breiman (2001) at University of California Berkeley. The model consists of a combination between multiple decision trees, instead of an individual regression tree (Donges, 2019). By ensembling multiple decision trees, individual model bias is reduced. The following section presents decision trees, bootstrap aggregating, and the combined Random Forest Model.

4.2.1 Decision Trees

Decision trees are series of sequential decisions to separate the data to reach pre-specified goals (James et al., 2013). **Figure 4.1** illustrates the mechanisms in a decision tree.

Figure 4.1: Illustrative decision tree with regressions



Note: With background in the chosen dataset, observations are split based on whether the interest rate in the same period is higher than or lower than 5%. The process is replicated in each node, splitting the observations into new nodes. The final Leaf Nodes generate distinctive predictions and are homogeneous with regards to the housing price growth.

Additional Interior Nodes could increase precision. Since the model catches more signal from the data, the complexity increases. However, noise would also be captured. At the very extent, additional interior nodes could repeat the whole dataset, which would imply an overfitted model. Decision trees therefore consider a trade-off between signal and noise to find the optimal nodes.

4.2.2 Bootstrap Aggregating

To reduce overfitting issues with decision trees, bootstrap aggregating is included in Random Forest. The concept generates various training subsamples, repeatedly selecting random samples with replacements from the full dataset. Individual decision trees are computed from randomly selected subsets. A final prediction is computed as a weighted average from the individual decision trees (James et al., 2013).

No precaution rules are present in the splitting variables process. Consequently, highly correlated trees could arise since influential variables are frequently chosen as root and interior nodes. For example, household debt and interest rate have significant influence on the overall housing prices. When debt and interest rate are included in the subsets, the trees could frequently choose these parameters, resulting in correlated trees. Correlated trees would to some degree be influenced by overfitting.

4.2.3 Random Forest Model

Random Forest applies a bootstrap aggregating method by repeatedly selecting random samples with replacements. Additionally, random selection of root and interior nodes are introduced. Splitting variables are not chosen from their separability, but randomly selected. Therefore, individual decision trees become more distinctive, minimizing the overfitting concern with bootstrap aggregating. This results in every variable contributes to the overall growth, regardless of influence degree.

Random Forest represents nonlinear ensemble models in our thesis. The nonlinearity is captured from the bootstrap aggregating process since different subsets identify interactions and combinations between influential variables. The ensembling methodology is introduced by averaging individual trees. Overall, the model reduces variance and increases precision compared to individual decision trees.

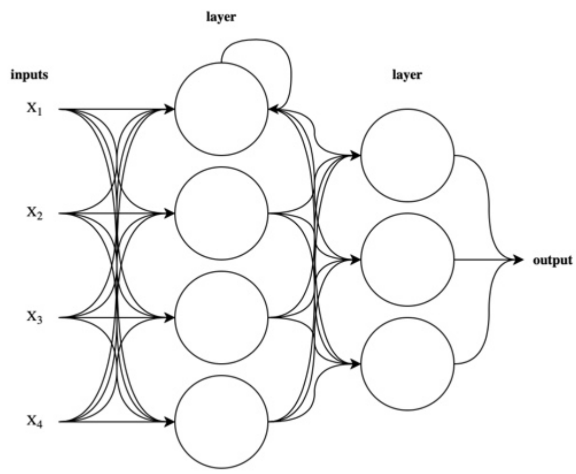
Despite being intuitive and applicable, Random Forest is less emphasized in existing literature. Moreover, Random Forest proves less implications and arguments on model performance. Arguments and economic intuitions are important for ensuring credibility in macroeconomic predictions. Still, the model is interpretable, and captures nonlinearities and highly correlated variables. By reducing overfitting and model variance, we believe Random Forest fits the purpose of this thesis.

4.3 Neural Network

Neural Network is among the first machine learning models developed and roots back to 1943. The model encompasses a large degree of network architectures. The initial intention was mimicking the human brain, running input through learning nodes to produce the desired output (Lantz, 2013).

To provide outputs, each input-variable is given an overall weight (Rosenblatt, 1958). Each weight represents the importance of input and allows for individual contribution to a greater or lesser amount to the sum of signals. The weights are determined through activation functions where the output signal is decided by the sum of input signals. Furthermore, each signal must reach a certain threshold value to either be included or excluded. The determination process is complex, and different layers of perceptions are linked to each other in a system of nodes. A typical representation is shown in **Figure 4.2**, where the x_1, x_2, x_3 and x_3 are the input variables.

Figure 4.2: Illustrative Neural Network composition



Note: Each input sends information to all the layers. The importance of each sent information is determined by weights. Information is also allowed to travel outside an into the same layer. The information is gathered to produce the desired output.

4.3.1 Network Topology

The capacity of a neural network is rooted in its topology. The network architecture can be explained by two key characteristics. The first characteristic is that the number of layers defines how many groups of neurons are included in the model (Lantz, 2013). Each layer has a defined set of connection weights for each input variable. Adding layers will increase the network complexity. The second characteristic is the number of nodes in each layer, which also affects model complexity. Currently, no universal applications on the determination on number of nodes exist. Existing literature suggests best practice is to use the fewest number of nodes that still results in adequate performance in the predictions. This can be achieved through hyperparameter tuning (Lantz, 2013), further explained in section 5.1.3.

Another aspect of the neural network architecture is the direction of information flow. Recurrent Neural Networks (RNNs) are networks well suited for sequential data processing, such as time series (Elman, 1990). In contrast to traditional "feed-forward networks", where information is passed in one direction through the neurons, RNNs are extended to include feedback connections allowing information to travel in multiple directions. The cycles of information enable network memory, where the neurons have different states. Allowing information to travel in both directions makes RNNs suitable for analyzing time-series data, by creating complex networks. In our thesis, the Neural Network model contains the parameter *delay* in Caret, which incorporates the recurrent aspect.

Neural Network models are known to be one of the more accurate algorithms within machine learning (Lantz, 2013), which makes it applicable to numerous problems. Also, Neural Networks incorporate complex patterns which are difficult to analyze in detail. Understanding mathematical intuitions and model operations becomes less visible for the user. This is commonly referred to as a Black Box problem (Maroto, 2017).

Neural Network allows for non-linear relationships between the dependent and independent variables, but these relationships are favorable in processing large amounts of data. The model also incorporates mechanisms to prevent over- and underfitting, which is important in cases where problems with multicollinearity might occur (Lawrence, 1997). For the rest of this thesis, we refer to the model as the Neural Network model.

5. Implementation

This section outlines the implemented model approach. The section presents the Caret-package, available in the statistical programming language R. In the following we will explain the technicalities in model training, and how the data is split into training and testing sets. Ultimately, our evaluation method is presented, ensuring equal comparisons with the professional institutions. The implementation is inspired by Bankson & Holm's (2019) thesis, predicting the GDP growth in Norway utilizing machine learning.

5.1 Caret

For the technical part of the implementation, the R-package *Caret*[7] is applied. The package contains functions that attempt to streamline the process of creating predictive models (Kuhn, 2008). Tools for data splitting, pre-processing, model tuning, and other functionalities are provided for panel data, time series and cross-sectional data. In total, Caret provides 230 different machine learning models. Caret's strength is due to its simplicity for less skilled programmers. However, due to the streamline procedures, less functionalities for parameter tuning are available. Therefore, a trade-off between simplicity and functionality is considered, when choosing Caret.

5.1.1 Data-splitting

Necessary preparations are needed to ensure credible predictions. First, a clear distinction between training and testing subsets is set. If models are tested on already trained data, predictions will be invalid. Furthermore, due to time series data, the observations' order must be maintained. Common data-splitting methods are therefore not applicable since they require independent observations. A train-test split that respects the temporal order of observations solves this concern (Brownlee, 2016).

The training set (in-sample) consists of data with the aim of learning and tuning the models. A training set normally includes 70% to 90% of the observations. An early assumption was to ensure sufficient quarterly evaluation with Norges Bank. Our testing set therefore follows their available housing price projections from Q2 2013 to Q4 2019. This returns a final data-split between training and test of 75% and 25% respectively. Given an already small dataset, less data is delegated in the training process, which is considered a weakness. The remaining dataset (out of

[7] Short for **C**lassification **A**nd **Re**gression **T**raining

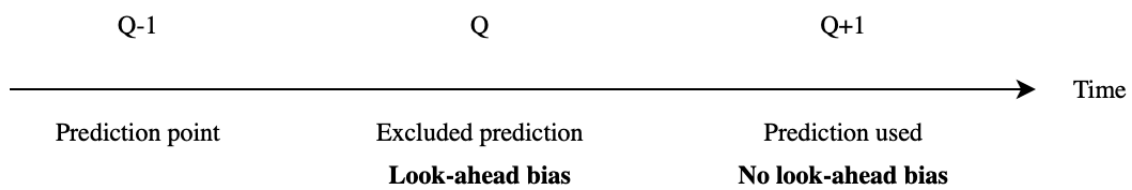
sample) is used for prediction purposes, to test the models on unseen data. The training-test mechanisms are further presented through The Rolling Forecast Origin in section 5.1.3.

5.1.2 Look-ahead Bias

The thesis aims to simulate realistic prediction circumstances. Therefore, unpublished data at the prediction point need to be excluded. In case we did not consider this problem, our predictions would have been influenced by look-ahead bias. Look-ahead bias occurs when an analysis uses information that would not have been available during the analyzed period (Kenton, 2020). The problem can unintentionally sway prediction results closer to the desired outcome, creating results that are too optimistic. The following example further explains the concept. SSB publishes the housing stock and household income for Q2 on the 2nd of September. Therefore, the variables are published two months afterwards. If our models predict the Q3 growth with data from Q2, look-ahead bias would occur. In reality, these publications would not have been available at the historical prediction point. Hence, the models would return optimistic predictions by possessing unpublished information.

In our approach, look-ahead bias is solved through computing quarterly growth for period Q and Q+1 simultaneously. The prediction for period Q is exposed to look-ahead bias, since information for period Q-1 is published during period Q. Using Q+1 solves this problem, by not using information from the previous quarter Q. Therefore, we consistently extract predicted values for time Q+1. The approach itself is illustrated in **Figure 5.1.**

Figure 5.1: Approach for avoiding look-ahead bias in quarterly predictions

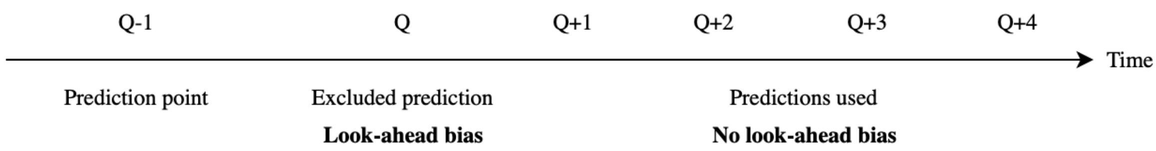


Note: Q-1 is the prediction point, while Q and Q+1 represent quarterly predictions produced simultaneously. Only the predicted values for Q+1 are extracted in further analysis.

A weakness with our approach is excluding information that would have been available at the prediction point, i.e., information that is published on monthly and daily basis. Therefore, the professional institutions can potentially include more information when producing their predictions, possessing an informational advantage. Still, we believe that few other actions are applicable without violating the look-ahead bias.

Yearly predictions apply the same methodology. When predicting the quarterly values for each year, we exclude information from the last quarter. For example, when predicting the four quarterly values for 2015, we only apply information available up and until Q3 of 2014. Hence, we exclude information made available in the fourth quarter when predicting for 2015. This approach replicates real prediction processes, ensuring comparability with DNB, Norges Bank and SSB. The approach, and intuition, is illustrated in **Figure 5.2**.

Figure 5.2: Approach for avoiding look-ahead bias on yearly predictions



Note: Q-1 is the quarter where the prediction is made, while Q, Q+1, Q+2, Q+3 and Q+4 represent the quarterly predictions produced simultaneously. The average of Q+1, Q+2, Q+3 and Q+4 is used as yearly growth rates in the analysis.

5.1.3 Cross-validation

Machine learning uses comprehensive tuning in the training processes. The process is called cross-validation and customizes the model to optimize accuracy (Lu et al., 2019). Throughout the training process, model parameters are tuned to minimize in-sample error and overfitting. Through tuning, the machine learning model searches for the best way to optimize the model. For the model training process, cross validation is crucial to discover relationships between the input and the output, enabling the production of accurate out-of-sample predictions.

Each model tunes unique hyperparameters that help fit the input to the output. While hyperparameters are used to control the learning process, other model parameters are derived through the learning process. Furthermore, hyperparameters are available for tuning and customization in machine learning models and can be set by the user in advance. Examples of relevant hyperparameters are the numbers of hidden nodes in Neural Network, the trees' depth in Random Forest, or the lambda and alpha in Elastic Net.

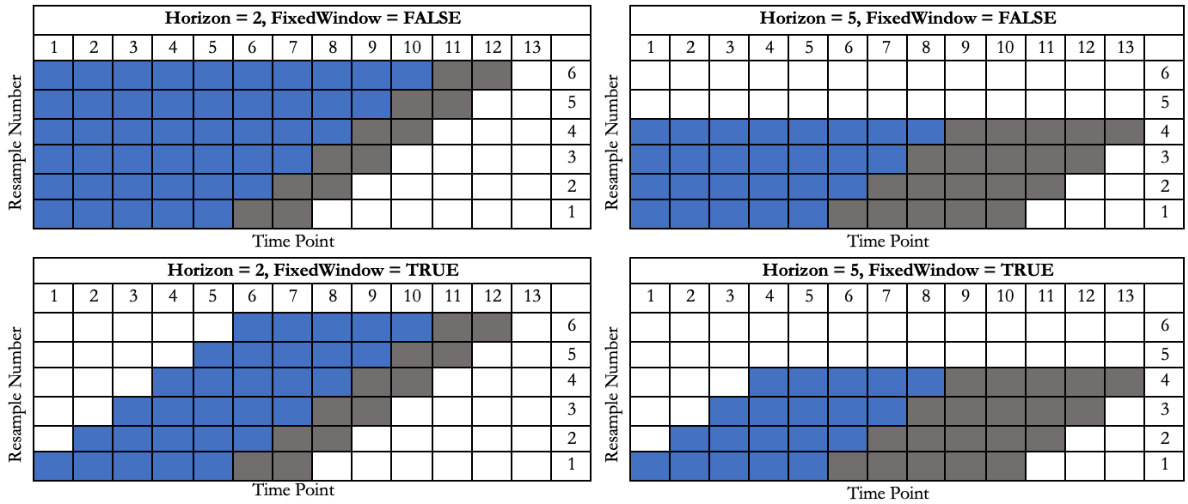
In a cross-validation process, the training set is divided into multiple sub-groups. One group is set aside, and later used for validation. Through multiple iterations of training and validating, the model tests different combinations of hyperparameter-values. The best combinations are chosen in the final model, ensuring that the model has the best prerequisites to produce accurate out-of-sample predictions.

Rolling Forecast Resampling

The order of the observations highly affects the underlying data-splitting and cross-validation. Hyndman and Koehler (2006) presented a process for cross-validation and data-splitting in time series. The method is called ***The Rolling Forecast Origin*** and splits the data into multiple individual training and test-sets. Each training set contains information that is available until the corresponding prediction point. For example, when predicting the yearly growth rate of 2015, the model uses all available information, including 2014. When predicting in 2016, the model includes information from 2015, adding new observations to the training set. A general explanation is that both the training and the test sets shifts over time. By introducing this method, the importance of the time-dimension becomes visible. Macroeconomic variables are highly dependent on recent information from previous periods (Woloszko, 2017). Thus, the time dimension makes observations in period T-1, T and T+1 highly correlated. In contrast, when using cross-sectional data, each observation should be independent. The rolling window replicates real circumstances, where new information is made available from one period to the next.

In Caret, the functions *TrainControl* and *CreateTimeSlices* cross-validate and split the data into training- and test-sets. *CreateTimeSlices* contains three parameters that are set in advance. First, *initial window* sets the initial length of the first training-set. This indicates how many consecutive observations are included in the first training-set iteration. Second, *horizon* defines the number of predictions in each iteration. The horizon is set to two quarters for quarterly predictions, and five for yearly predictions. Lastly, *fixed window* defines whether the size of the training set should be held constant or be expanding. **Figure 5.3** illustrates all possible specifications for the *CreateTimeSlice*-method.

Figure 5.3: Illustrative figure showing all different combinations in the CreateTimeSlice-function in Caret.



Note: Overview of settings in CreateTimeSlice. Horizon defines the number of predictions made each iteration. FixedWindow determines whether the training set should expand or be held constant (just moved one period forward after each iteration). Time points show at what period the training set starts and ends, as well as the desired predictions. Resampling number is the number of iterations. Blue fields = training sets, Grey fields = test sets.

We set the fixed window parameter to FALSE, to include all training sets from previous observations until the prediction point. Consequently, more noise in the training set is captured, but the number of observations in the training process is increased. Given our small dataset, prioritizing additional observations could be favorable. Furthermore, we set the initial window parameter so that the first predicted value is 2013Q2 for the quarterly predictions, and 2013Q1 for the yearly predictions. In **Figure 5.3**, the two upper boxes illustrate the parameter composition in our thesis.

5.2 Evaluation

To evaluate model performance, accuracy and direction measurements are needed. We need to quantify to which extent the predicted values are following the actual observed values (Sucarrat, 2019). We apply the measurements Mean Directional Accuracy (MDA), Rooted Mean Squared Error (RMSE) and Mean Absolute Error (MAE).

Throughout this section, the following notations are applied: Y_t represents the actual growth in period t, and $\widehat{Y}_t$ represents the predicted value. K represents the total number of relevant prediction periods. The measurements represent percentages, since all predictions are computed as percentage growth rates.

MDA represents the percentage of predictions that matches the correct growth direction. MDA is directly applicable to compare the performance across models. For example, MDA of 80% indicates directional correctness in 80% of the predicted values. The MDA should be higher than 50% in large test sets, to provide predictive performance.

$$MDA = \frac{1}{K}\sum_{t=1}^{K} 1_{sign(Y_t - Y_{t-1}) == sign(\hat{Y}_t - Y_{t-1})} \tag{5.1}$$

RMSE represent the standard deviation of the residuals. It measures the spread of all residuals and illustrates how concentrated the predictions are around the actual growth (Holmes et al., 2000). We apply a rolling resampling in our implementation, so that a new residual is computed after each iteration. This creates a more representative RMSE when predicting time series, and when comparing predicted values to the benchmark (Clark & McCracken, 2001). RMSE is expressed as the root of the averaged squared residuals.

$$RMSE = \sqrt{\frac{1}{K}\sum_{t=1}^{K}\left(Y_t - \hat{Y}_t\right)^2} \tag{5.2}$$

Due to the squared term, large deviations are punished harder than smaller deviations. Therefore, a direct intuition of RMSE is more complicated, even though the punishment of larger deviations is crucial when evaluating predictions of macroeconomic variables.

MAE represents the average absolute difference between the predictions and actual growth (James et al., 2013). A MAE of 1.2, represents an average error of 1.2 percentage points compared to the actual growth (Vandeput, 2019). Compared to the RMSE, MAE does not put emphasis on whether the deviations are large or small.

$$MAE = \frac{1}{K}\sum_{t=1}^{K}\left|Y_t - \hat{Y}_t\right| \tag{5.3}$$

In addition to the presented error measurements, absolute individual errors are computed. The absolute errors are used to index the performance of models and predictors, while creating a favorable comparison of individual predicted values.

6. Results

We use described data and implementation to evaluate the performance of the machine learning models. Firstly, the quarterly predictions are presented and compared to predictions from Norges Bank. Lastly, yearly predictions are presented and evaluated upon predictions from DNB, Norges Bank and SSB.

6.1 Model Performance - Quarterly Predictions

The first part of the analysis compares quarterly predictions from the machine learning models to Norges Bank's predictions as benchmark. The projections are extracted from the Monetary Policy Report (MPR) in the period from 2013Q2 to 2019Q4. In the report, Norges Bank compute the quarterly (y/y) growth for the current and next quarter. We consistently gather the latter projection.

Table 6.1: Overall results from quarterly predictions for 2013Q2 to 2019Q4

Measurements	Elastic Net	Random Forest	Neural Network	Norges Bank
RMSE[1]	3.08	3.27	3.30	1.32
MDA[2]	62.96%	59.26%	66.67%	85.19%

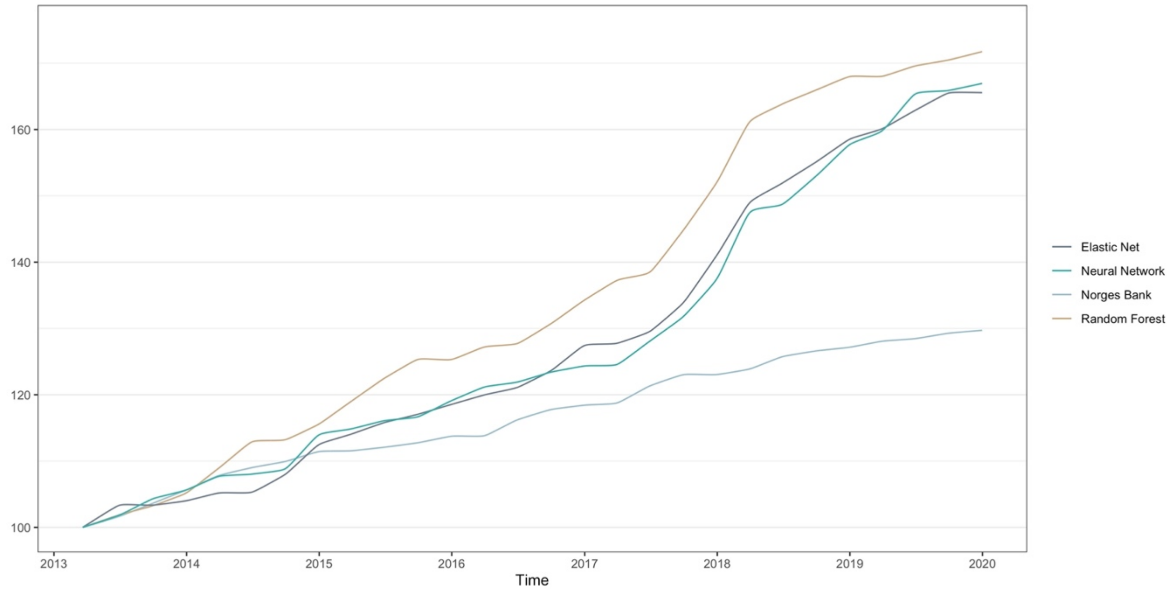
[1] Absolute %-point error
[2] Mean Directional Accuracy – number of predictions (%) in the correct growth direction

The overall results, shown in **Table 6.1**, create insights into several aspects. Neither of the machine learning models are able to predict more accurately than Norges Bank. Norges Bank's RMSE of 1.32, is considerably lower than all machine learning models. Among the machine learning models, Elastic Net produces the most accurate predictions. The RMSE of 3.08 indicates that the model, on average, deviates 3.08 percentage points from the actual growth over the same period. All machine learning models predict the correct direction in more than 50% of the quarters. Most notably, however; Norges Bank predicts the correct direction in 85.19% of the quarters. This is considerably higher than all machine learning models. Overall, Norges Bank's predictions have been the closest to the actual growth on average between 2013 and 2019.

6.1.1 Indexed Absolute Errors

Based on the results, it is interesting to identify which periods the machine learning model have predicted accurately or not, i.e., when the models are least precise. **Figure 6.1** visualizes a representation of the absolute errors across the models, where the base-quarter is 2013Q1.

Figure 6.1: Indexed absolute errors from quarterly predictions in the period of 2013Q2 to 2019Q4



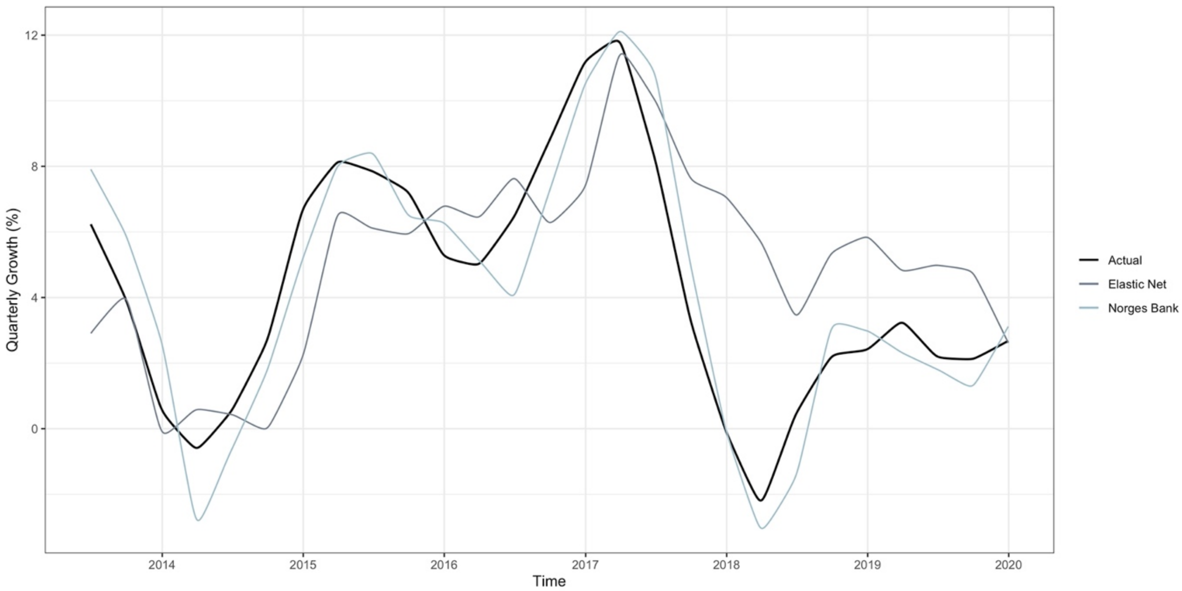
Note: Absolute errors for quarterly housing price predictions, where the first quarter of 2013 is set as base-quarter, and the second quarter as the first growth value. Each quarter, the absolute errors are accumulated development in absolute errors. Due to only accumulating absolute (positive) values, the curve cannot decline.

The distance between the lines and the horizontal axis are accumulated absolute errors. A flat curve indicates a prediction closer to the actual growth. In contrast, a steeper curve represents an inaccurate prediction. Analyzing the curves, all models performed similarly during the first period until 2015. Interestingly, both Elastic Net and Neural Network predicted more accurate than Norges Bank during the period. The results could indicate that the two machine learning models capture the underlying macroeconomic mechanisms better than Norges Bank. However, steeper curves during the late quarters 2017 and early 2018 indicate large deviations across all machine learning models. These deviations increase the overall RMSE, reducing overall performance. Hence, analyzing the periods of 2014 and 2017/2018 more closely are preferable, and is discussed in section 7.3.

6.1.2 Head to head – Elastic Net and Norges Bank

In further analysis, the best performing machine learning model Elastic Net is compared to predictions from Norges Bank. From **Figure 6.2**, both models seem to follow the development in the housing market, despite deviations in 2014 and 2018.

Figure 6.2: Plotted quarterly predictions from Elastic Net and Norges Bank



Note: Housing price growth in Norway. The actual housing price growth (quarterly (y/y)) is computed from Real Estate Norway's seasonally adjusted HPI.

A descriptive analysis on the performance from Norges Bank and Elastic Nets is presented in **Table 6.2**.

Table 6.2: Descriptive comparison of quarterly predictions

	Elastic Net	Norges Bank
Quarters Won[1]	8	19
Percentage Won[1]	29.63 %	70.37 %
Most Precise prediction[2]	0.03 (2013Q3)	0.03 (2017Q4)
Least Precise Prediction[2]	7.87 (2018Q1)	2.57 (2017Q2)
MAE[2]	2.43	1.10

[1] Number of predictions where the predicted value is closer to the actual value
[2] Absolute %-points error

A direct comparison show that Norges Bank predicts with the most precision in 70.37% of the quarters. Despite being generally less accurate than Norges Bank, Elastic Net seems to produce more accurate predictions during the quarters of 2014. From **Figure 6.2** we see that Norges Bank overestimates the fall in housing prices during 2014. In this period, the housing market was affected by the decline in oil prices. Hence, Elastic Net seems to capture this decline more accurately than Norges Bank. One should also mention that Elastic Net's MAE of 2.43 is highly affected by large deviations through 2017 and 2018, indicating the model proved a higher predictive accuracy during the first period.

6.2 Model Performance - Yearly Predictions

In market analysis and news articles, yearly growth rates are considered the most relevant. Additionally, from section 3.4, the included institutions report the average annual growth rate. To ensure equal comparisons, we consistently compute the average of our quarterly growth rates each year. We first evaluate the machine learning performance, before the best performing model is compared upon the professional institutions.

Table 6.3: Results from yearly predictions for 2013 to 2019

Measurements	Elastic Net	Random Forest	Neural Network
MDA[1]	42.9%	57.1%	42.9%
RMSE[2]	5.86	4.35	5.73
MAE[2]	4.46	3.69	4.62

[1] Mean Directional Accuracy - number of predictions (%) predicting the right direction
[2] Absolute %-point error

From **Table 6.3**, we find Random Forest to be superior among the machine learning models. A RMSE of 4.35 shows that Random Forest predicts 24.95%[8] more accurately compared to Elastic Net and Neural Network, with RMSEs of 5.86 and 5.73 respectively. In terms of MDA, only Random Forest fulfills the 50% requirement, predicting the correct direction in 57.1% of the seven years. Therefore, Random Forest represents machine learning in the further analysis.

[8] Average of the relative comparison to Elastic Net and Neural Network

6.2.1 Random Forest Compared to Norges Bank, DNB and SSB

To evaluate the machine learning performance, Random Forest is compared to predictions from Norges Bank, DNB and SSB. The results are summarized in table **Table 6.4**.

Table 6.4: Results from yearly predictions of 2013 to 2019

	Random Forest	Norges Bank	DNB	SSB[1]
MDA[2]	57.1%	85.7%	42.9%	42.9%
RMSE[3]	4.35	2.23	4.49	4.53
MAE[3]	3.69	2.1	4.24	4.13

[1]Computed based on Statistics Norway's Inventory weighted HPI
[2] Mean Directional Accuracy - number of predictions (%) being in the right direction
[3] Absolute %-point error

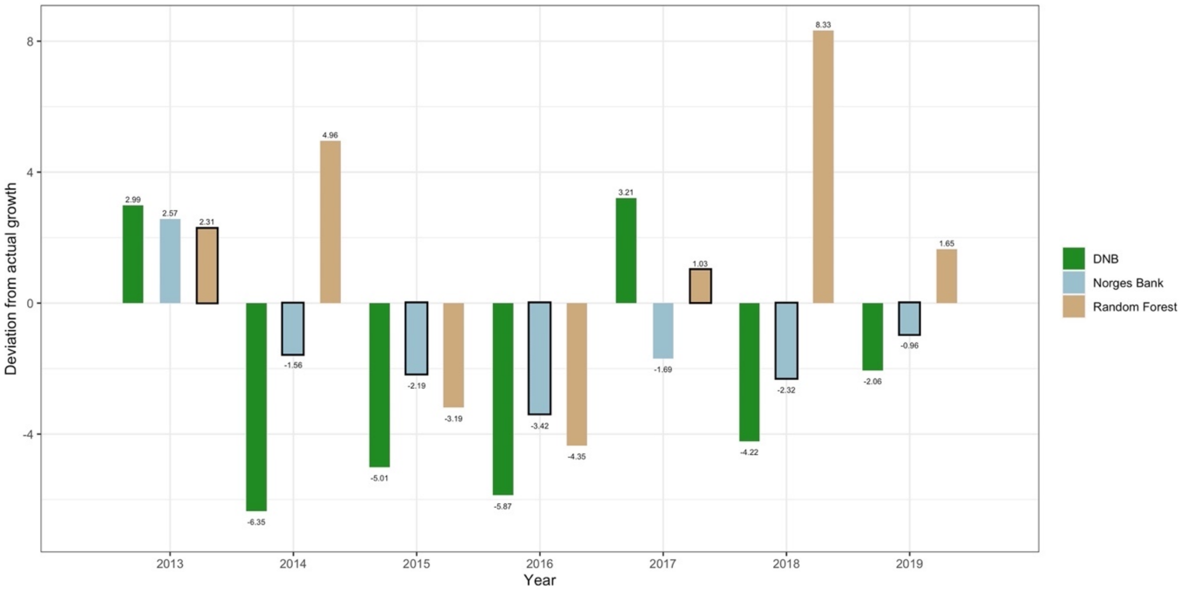
The yearly results show Norges Bank is overall superior to Random Forest, DNB and SSB. Norges Bank predicts 48.73% more accurate than Random Forest in terms of RMSE. However, neither DNB nor SSB outperform Random Forest, with RMSEs of 4.49 and 4.53 respectively. A relative comparison shows that Random Forest on average predicts 3.6%[9] more accurate compared to DNB and SSB. Hence, the results strengthen the arguments for applying machine learning, since it outperformed two of the three professional institutions. Additionally, neither DNB nor SSB satisfies the MDA-requirement of 50% accuracy.

The overall performance is considerably weakened with regards to overall RMSE, compared to the quarterly results. Questions regarding the degree of predictability across horizons in the housing market are further discussed in **section 7.4**. Still, computing the MAE, all models deviate on average between two and four percentage points from the actual growth. The actual average growth rate between 2013 and 2019 is 4.4%. Therefore, all institutions deviate on average between 50% and 100% from the actual growth.

[9] Average of relative comparisons to DNB and SSB

6.2.2 Periodic Absolute Errors

Figure 6.3: Deviations from actual growth – yearly predictions from 2013 to 2019



Note: Deviations computed from the difference between predicted values and actual growth. The lowest deviation is highlighted with a black frame in each period.

Figure 6.3 shows individual yearly errors. The lowest annual deviation is highlighted with a black frame. Similar to the quarterly results, two events affect overall performance. First, the expected fear in the housing market in 2014. Norges Bank and DNB are pessimistic in their predictions. However, Random Forest is not affected by the market sentiment, leading to a somewhat optimistic view for 2014. Second, the effects after the implemented mortgage regulations for 2018. Random Forest is not able to capture the effect, while DNB and Norges Bank overestimate the fall. These events highly affect the overall performance and are further discussed in section 7.3.1 and 7.3.2.

6.2.3 Descriptive Performance

Table 6.5: Descriptive analysis of absolute errors from yearly predictions between 2013 and 2019

	Random Forest	Norges Bank	DNB
Lowest Deviation[1]	1.03	0.96	2.06
Highest Deviation[1]	8.33	3.42	6.35
Median[1]	3.19	2.19	4.22
Closest to true growth[2]	2	5	0
Percentage Won[3]	29%	71%	0%
MAE[1]	3.69	2.1	4.24

[1] Absolute %-point error
[2] Number of predictions where the predicted value is closes to the actual value
[3] Number of quarters closes to the true value, as percentage of total predicted values

Descriptive results are shown in **Table 6.5**. Norges Bank had the lowest deviation of 0.96 percentage points during the period. Interestingly, only one prediction was deviating less than one percentage point from the actual growth. Furthermore, Random Forest predicts the most accurate rates in 2013 and 2017. Random Forest also produced the highest deviation of 8.33 percentage points in 2018. However, the model was not able to capture the effect of the mortgage regulation introduced in 2017. Lastly, DNB produces the least accurate predictions in six out of seven years, indicating a 0% winning rate. Also, DNB did not predict within a 2-percentage points interval between 2013 and 2019.

The overall results indicate a predictive performance to machine learning. However, high deviation through 2018 still indicate clear weaknesses with the models. The inaccurate predictions for 2018 also affects the overall RMSE, reducing overall performance. Still, the testing period is short, and conclusions are not drawn to future model selection preferences.

7. Discussion

This section discusses further implications for future macroeconomic predictions, respectively machine learning modelling and the Norwegian housing market. We will first evaluate underlying aspects between the machine learning models. Our results and methodologies are compared to previous literature, identifying consistencies and improvements. Further, we find strengths and weaknesses of applying machine learning during influential events in the Norwegian housing market. Moreover, the predictability in the housing market is elaborated through the lens of our results. Additionally, variable contribution is discussed compared to existing literature. Lastly, model limitations and further research questions are examined.

7.1 Machine learning Performance

We observe relative differences in performance among the machine learning models. The linear model Elastic Net is superior to Random Forest and Neural Network in quarterly predictions. In yearly predictions, the nonlinear model Random Forest produces the most accurate predictions. Similar patterns have occurred in American and Australian housing literature (Gupta & Miller, 2015; Milunovich, 2019). It implied that linear models are more suited in making one-quarter ahead housing predictions, while nonlinear models dominate over longer time horizons. The increased complexity is present in nonlinear models, where interactions and combinations between explanatory variables are captured. We believe further model builders could benefit examining these consistencies.

A deeper analysis of the Elastic Net model can be found in its cross-validation, see **Figure A.3.2**. Elastic Net is an ensemble model combining OLS regression and penalty-terms from Lasso and Ridge. Ridge allows variable reduction, while Lasso introduces variable selection. Through cross-validation, the model seeks to minimize in-sample deviation, by testing different values for alpha and lambda. Elastic Net implies an alpha of 0.10 and a lambda of 0.525 in the optimal model. The alpha states a model prioritization to Ridge regression, reducing coefficients of highly correlating variables. The lambda creates a lower weighting of coefficients compared to regular OLS regression. Thus, the overall model is a modified linear OLS-regression which prioritizes minimizing coefficients of the covariates.

Despite having an intuitive approach, the Random Forest has fewer possibilities to provide reasoning for the underlying performance (Breiman, 2001). Still, Caret provides insights to the influential hyperparameters: the number of generated trees (*ntrees*) and the number of random variables selected for each tree (*mrty*). As shown in **Figure A.3.1**, the final model consists of 500 trees and five chosen variables in each tree. Every decision tree would have five splitting points, generating a total of 32 (2^5) root nodes. The average from 500 decision trees is computed, returning the predictions for Random Forest. The final model allows for individual contribution and interactions between variables. Capturing nonlinear relationships between variables are probable for the promising yearly results.

Developing macroeconomic machine learning models is challenging due to limited availability of historical data. This reduces captured relationships in the training process. Model ensembling could serve as a technique to prevent this challenge. Previous studies have found that ensembled models are more robust in smaller datasets, since individual model errors could be minimized (Tiffin, 2016; Valland, 2019). Ensemble techniques are prior in both Elastic Net and Random Forest. However, these techniques are not present in Neural Network. Possibly, Neural Network could increase model performance, given larger datasets. Abidoye et al. (2019) extracted data back to 1985, utilizing an additional eleven years of observations. Further applications could possibly prioritize observations instead of variables to increase Neural Network's performance, following the intuitions from Abidoye et al. (2019).

7.2 Insights from Existing Litterature

Applying machine learning in housing predictions are still at an early stage in academic research. Previous research in Australia, Hong Kong, Malaysia and USA have yielded promising results, by outperforming OLS-regressions, Random Walks and ARIMA-models. However, as Sucarrat (2019) argued, prediction models should at least outperform a Random Walk, in order to have any predictive accuracy. Hence, our thesis is, to the authors' knowledge, the first to evaluate machine learning performance compared to professional institutions. Still, insights can be extracted from the previous literature results and methodologies. Therefore, we compare our methodologies with Abidoye et al. (2019) in Hong Kong, and Milunovich (2019) in Australia.

7.2.1 Hong Kong Prediction Study

Abidoye et al. (2019) predicted the Property Price Index in Hong Kong with Neural Network. Their objective was predicting booms and bubbles in the housing market. While we have predicted quarterly growth rates, Abidoye et al. predicted the quarterly property index directly.

Hence, to ensure comparable results, we compute the quarterly (y/y) growth from the predictions in the paper. An overall performance analysis could provide additional insights and is presented in **Table 7.1**.

Table 7.1: Comparison of results in Norway and Hong Kong

	Norway (NN)	Hong Kong (ANN)
MAE[1]	2.48	3.31
Starting Quarter	2013Q2	2013Q1
Ending Quarter	2019Q4	2016Q3
Predicted Quarters	27	15

[1]Mean absolute error computed from the quarterly (y/y) growth. Predictions from Hong Kong are available in the paper

The studies investigate distinctive markets, time spans and training sets. Still, the included variables, model specifications and intuitions share similarities. If we assume direct comparisons between the Neural Network models, the predicted values are 33%[10] more accurate in Norway. Market specifications probably intervene with the comparison assumption. Hong Kong is the number one city in the world with the most skyscrapers, leading to a high density in the population (Burton, 2018). In addition, property prices have previously been considered extremely volatile, resulting being one of the most expensive cities in the world (Abidoye et al., 2019). Therefore, the computed MAE would naturally be higher in Hong Kong, given higher fluctuations in prices. Still, Abidoye et al. (2019) argued Neural Network could be used as a decision tool for predicting booms and bubbles in Hong Kong, helping government policy makers and real estate investors. Neural Network predicted more accurate than Norges Bank during the macroeconomic events between 2013 and 2015, substantiating the hypothesis from Abidoye et al. (2019).

7.2.2 Australian Forecasting Study

The machine learning study from Milunovich (2019) found similar results in the Australian Housing market. Interestingly, his study deviates in three aspects in the model implementations. First, his approach implies a K-Fold[11] cross-validation method. He also preferred a fixed window in the testing set, i.e., that the size of the training set is held constant throughout the training process. Lastly, his method required data to be adjusted to satisfy statistical assumptions. This includes non-correlating error terms and stationarity in the dataset.

[10] Relative comparison of the MAEs

[11] K-fold cross-validation randomly selects K training and testing subsets when tuning the model hyperparameters. For each iteration, the model is trained on K-1 subsets and tested on the remaining subsets. After all iterations are finished, the average of the measurements is calculated and is representing the training accuracy of the model (Brownlee, 2018)

Our methodology implies a rolling forecasting origin in the cross-validation. Hyndman (2016) considers the rolling forecasting origin to be preferable in time-series predictions, since K-fold has problems handling inherent serial correlation and non-stationarity. Furthermore, we imply a non-fixed window, by expanding the number of observations in every prediction. Consequently, our approach could capture more noise in the dataset. Lastly, we have not put emphasis on stationarity in the explanatory variables in our analysis since this is not required in the machine learning models.

Elastic Net ranked 5th among the 47 algorithms in one quarter ahead predictions. The studies have examined different markets, limiting performance evaluation. Still, if we allow for an overall comparison, model selection strengths and weaknesses could occur. Milunovich computed quarterly logarithmic growth rates (q/q). We therefore change the growth rate formula to follow logarithmic quarter on quarter growth rates. The performance is shown in **Table 7.2**.

Table 7.2*:* Comparison of results in Norway and Australia

	Norway (Elastic Net)	Australia (Elastic Net)
RMSE[1]	1.97	1.41[2]
Starting Quarter	2013Q2	2005Q1
Ending Quarter	2019Q4	2017Q3
Predicted Quarters	27	47

[1]Absolute %-point errors computed from quarterly (q/q) predictions
[2] In the paper, Milunovich reports the MSE. Hence the RMSE is computed by rooting the MSE

If we allow for relative comparisons, Elastic Net is 34.1%[12] more accurate in the Australian market. This could either indicate that the approaches are influential in model building, or that unfortunate events intervene with the comparison assumption. Assuming model influence, it is interesting that the K-fold cross validation still has produced accurate predictions, even though experts prefer a rolling forecast origin. Also, the choice of using a non-fixed window could have weakened our performance. Simultaneously, abnormal events probably intervene with the comparison assumption. Elastic Net was highly inaccurate during 2017 to 2018, decreasing overall performance.

[12] Relative comparison of RMSEs from the Elastic Net predictions in Norway and Australia

Lastly, one should mention that Milunovich's study and our results are not absolute comparable. Differences in datasets and market attributes weakens the relative comparison. Hence, the RMSE values does not state that his results are superior. However, this comparison aims to highlight that variations in approach might influence the machine learning performance.

7.3 Machine learning in abnormal events

Expert predictions fluctuate in emphasis and arguments, possibly leading to disagreeing conclusions. Zarnowitz and Braun (1992) found the group mean of forecasts (consensus) outperformed individual macroeconomic predictions, since individual prediction bias is reduced. Later, Bennett et al. (1997) argued that when companies prioritize publicity instead of accuracy in macroeconomic predictions, the predictions deviate more from consensus. The studies indicate potential biases expert forecasters experience in their macroeconomic analysis. Machine learning could ideally reduce this bias, by solely focusing on precision. Simultaneously, machine learning has difficulties identifying non-numeric market consensus. We therefore explore strengths and weaknesses to machine learning compared the professional projections in the two most influential events, 2014 and 2018.

7.3.1 2014 – Machine learning in an uncertain housing market

In the autumn of 2013, market consensus had a pessimistic view towards growth rates for 2014. Still, the institutions differed in arguments and intuitions. Sudden price drops in October 2013 were prominent, justified through psychological fear (E24, 2013). Further, Nordea Markets characterized the housing market as troubling, due to high degrees of debt collections and forced sales, predicting -7.8% growth (Juel, 2013). DNB (2014) focused on higher unemployment levels and interest rates during 2014, predicting a negative growth of 4%. The machine learning models, however, have not found underlying relationships supporting market consensus. All machine learning models predicted positive growth rates for 2014.

The yearly growth was 2.35% positive in 2014, diverging from market consensus. Low interest rates and limited supply of new properties were highlighted as influential for the 2014 market (Halvorsen, 2014). Potentially, negative consensus could influence a large degree of predictions, indicating a group thinking mentality. After reading through several market reports for 2014, only a few highlighted positive projections for 2014. This could question the individuality in analysis between forecasting institutions. In this scenario, machine learning could potentially decrease the degree of group thinking, by limiting expert sentiment in the analysis. We consider this a potential strength in further use of machine learning.

7.3.2 2018 – New Mortgage Regulations

Norwegian authorities implemented new mortgage regulations from January 2017, with the objective to create a more sustainable housing growth (Finansdepartementet, 2016). This was, among other factors, due to rapid growth around Oslo. The regulations created a mortgage roof of five times gross income. In addition, the regulations increased equity requirements to 40% on secondary properties in Oslo. The main objective for the regulations was decreasing the household debt levels.

DNB and Norges Bank predicted falling housing prices for 2018, stating drops of 3.5% and 1.6% respectively. The mortgage regulations were prominent for the negative sentiment. First, the restrictions could exclude potential house buyers in bidding rounds, leading to a decrease in demand. Second, threat of increased future interest rates, due to the stable economic situation, would reduce the risk willingness for debt holders. In addition, DNB and Norges Bank argued large stocks of unsold new properties would flood the market supply. Therefore, consensus was negative towards the 2018 market.

All machine learning models predicted abnormal high growth rates for 2018. The positive correlation between household debt and housing prices could explain optimistic sentiment for machine learning. This is further explained in **section 7.5.1**. Finanstilsynets' objective was initiating action to decrease debt levels. Without prior information of the regulatory changes, the models could expect further growth in debt levels, increasing housing prices during 2018.

The yearly growth for 2018 was 0.72%, deviating from market consensus and machine learning. Abnormally high turnover volumes highlighted the 2018 market, which few experts had foreseen (Strømnes, 2019). Still, the implications from the mortgage regulations were present, slowing down the housing price growth.

Unfortunately, the new mortgage regulations were not successfully included in the machine learning models. Therefore, the models did not have prior knowledge towards the changing environment. Additionally, uncertain events are naturally hard for machine learning to capture, since these patterns have not occurred in the training sets. Such regulatory changes also affect the underlying dynamics. Furthermore, since the housing market is influenced by uncertain events and abnormalities, this serves as a weakness for the use of machine learning. Still, Mari Mamre implemented regulatory changes into her Neural Network model, indicating that future models can capture these abnormalities to some degree (Stranden, 2019).

7.4 Market Predictability and Uncertainty

Although our findings indicate machine learning is suitable for housing price predictions, market complexity and uncertainties have decreased overall performance. In macroeconomic predictions, the accuracy across time horizons has two deviating arguments. On the one hand, shorter time horizons are influenced by frequent short abnormal situations. By averaging growth rates over longer horizons, these abnormal situations become less decisive in the predictions. In this case, longer time horizons are more predictable. On the other hand, longer horizons take more future uncertainty into account. The development of underlying mechanisms can deviate throughout these periods. The prediction accuracy decreased by 68% for Norges Bank from quarterly to yearly predictions, indicating higher degrees of predictability with shorter horizons. We want to examine our results with previous national and international research on the housing market efficiency. In an efficient market, predictive models should not be able to produce consistent results over time. If markets are inefficient, uncertainties and abnormal situations might still reduce predictive performance.

In 1989, Case and Shiller argued the market for single-family homes were inefficient in the US. Through their study, they found that year-to-year prices tend to be followed by changes in the same direction as the subsequent years. However, individual housing prices were not forecastable in a micro perspective. Further, Case and Shiller (1990) investigated the predictability in a macro perspective across changing horizons. In their study, they suggest the housing market is predictable in shorter horizons, as far as one-quarter ahead predictions. Later, Kuo (1996) improved their methodology by introducing serial correlation, finding some predictability up to four quarters ahead.

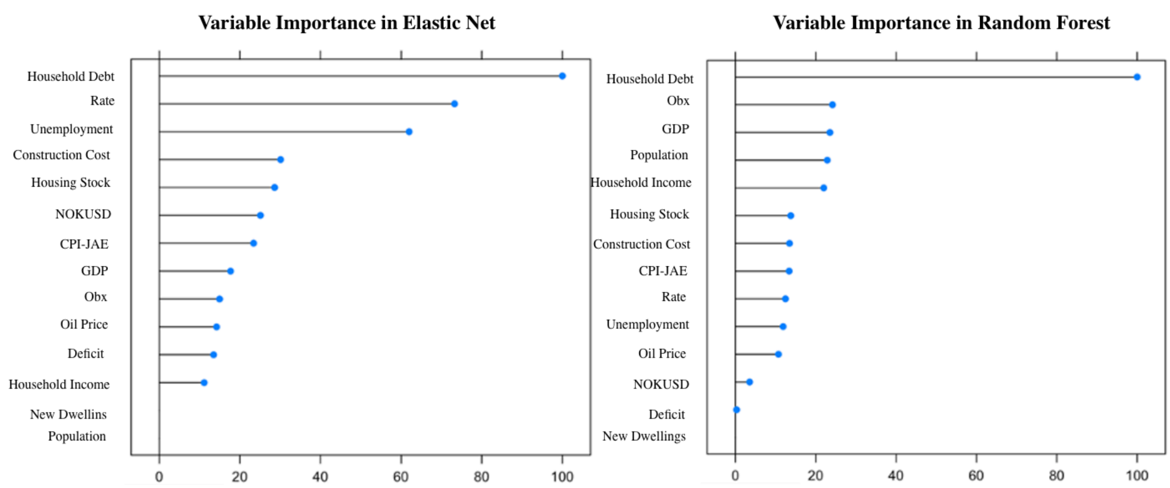
In the Norwegian housing market, two distinctive papers test the macro and micro efficiency. Røed Larsen and Weum (2008) replicated the Case & Shiller time-structure test on the price index to test the efficiency hypothesis in a Norwegian setting. Based on quarterly sales data in Oslo from 1991 to 2002, they find consistent time structures in the market, characterizing the market as inefficient on a macroeconomic level. However, Anundsen and Røed Larsen (2018) indicated housing predictability was absent at a micro level. Their main argument was the number of bidders were stochastic. Hence, they rejected the hypothesis that individual properties were predictable. The papers indicate a degree of predictability on macro level, but not at a micro level.

Our study contributes to the macroeconomic aspect. The microeconomic perspective still influences the overall market, by increasing the complexity. Our quarterly results indicate a degree of predictability from 2013 to 2019, confirming the findings from Case and Shiller (1990) and Røed Larsen and Weum (2008). Even though deviations occur, the models capture underlying movements, at least guessing the correct direction in more than 50% of the periods.

Our yearly results indicate predictability to some degree, but overall performance is weakened. Consistently with Kuo (1996), we find the housing market is forecastable to some degree up to four quarters. Still, yearly predictions deviate on average between 50% and 100 % from the average growth, indicating that predicting housing prices is considered difficult. Hence, we believe future machine learning is expected to have inaccurate predictions, as with 2018.

7.5 Variable Importance

Figure 7.1: Variable Importance for quarterly predictions from Elastic Net and Random Forest



Note: Variable importance from the *VarImp*-function in Caret. Each variable is given a percentage value based on its importance in the overall model. The importance is based on the increase in model error if the variable is omitted from the model. All scores are shown as a percentage of the most important variable *Household Debt*.

A further analysis of the contributing variables could highlight the economic intuitions from machine learning. Having intuitive and well-established arguments for the predictions strengthens the overall trustworthiness. In the Caret package, the function *VarImp* generates a ranking based on the variables' contribution. Most ideally, the models would generate weightings in every prediction, which is not provided in the package. Still, every variable is given a score from 0 to 100, based on the increase in error if the variable is omitted. Then, each variable is divided by the most important variable, returning a percentage score (Chauhan, 2017). The variable importance is gathered from the quarterly predictions from Elastic Net and Random Forest.

Common for Elastic Net and Random Forest are the importance of household debt, being the main contributor to the overall housing price growth. Household debt was also identified as most influential in the Swedish housing market (Borg, 2019). Debt level implications are discussed in detail in **section 7.5.1**. Further, the models differ in emphasized variables. Elastic Net puts emphasis on the interest rate and unemployment, which is consistent with previous Norwegian housing research (Jacobsen & Naug, 2004). The population has intuitively contributed to the housing price growth. An increasing population would increase potential buyers and influence new construction. Hence, the importance of this variable is certain in Random Forest. However, population is omitted in Elastic Net. A correlation matrix of the selected variables, see **Figure A.2.1**, shows that population is positively correlated with the overall housing stock, household income and GDP. Thus, it is plausible that high correlations between explanatory variables resulted in the lasso regression omission.

The oil price is ranked the 10th and 11th from 14 variables in Elastic Net and Random Forest. Beltratti and Morana (2010) argued oil prices have a significant effect on housing prices in the G7-area, being an important variable in the business cycle development. Similar patterns could occur in our analysis, given the oil industry's importance to the Norwegian economy. The variable importance indicates that oil prices influence the housing market but is among the weakest contributing variables nationally. A probable explanation refers to differences among regional markets. For example, in June 2020, Oslo experienced a 12-months growth of 6.1%, while Stavanger witnessed a 0.9% growth during the same period (Myrvang, 2020). Real Estate Brokers argued the aftermath from the oil crisis still contributed to the moderate growth. However, long term consequences from the oil crises were less evident in the Oslo area. We therefore believe, based on the variable importance, that the oil price influence is more present in local markets rather than nationally.

One should be careful drawing conclusions directly from the *VarImp*-function. The function provides a relative comparison of the individual contributions during the cross-validation process. As mentioned, Elastic Net performs variable reduction, which might result in reduced coefficients of important variables that are expected to influence housing price growth. Correlating variables having an impact on the housing price growth might still have their coefficients reduced. Additionally, causal interpretations cannot be drawn from the variable importance. Still, the variable importance indicates which attributes are influencing the models during the whole period. Newer programming packages LIME and SHAP are working on increasing the interpretability for machine learning models (Sharma, 2018). We believe further analysis could indicate more macroeconomic intuitions and arguments, increasing the predictions' trustworthiness.

7.5.1 Household Debt

Analyzing the variable importance from Elastic Net and Random Forest, household debt is considered the most important variable in the predictions. Questions regarding whether the household debt is considered an indicator, or a reflection of the housing price development, have been raised in existing literature. If housing prices are rising while wages are constant, the price change can be covered through higher debt levels. In this case, housing prices are affecting the household debt, indicating a reverse causality problem. Reverse causality occurs when the direction of the causal effect relationship is inverted, since the onset of the cause is not detectable (Schölkopf, 2008). However, if the supply of debt is increased through lower interest rates, more buyers can bid on the same properties, increasing prices. In this case, debt levels contribute towards higher prices, and reverse causality is not present. Debt levels in Norwegian households has been mentioned as a threat for the Norwegian economy (Finanstilsynet, 2020). They recently reported all-time high debt levels with an average 338% debt to income ratios. High debt levels increase the vulnerability to loss in future income. Based on economic intuition and previous influence, we chose to include household debt in the models.

Existing literature have different outlooks regarding the relationship between household debt and housing prices. Borio et al. (1994) indicated a positive correlation between the growth in debt and housing prices internationally. Debt appears to follow the housing price growth, with a time lag. Additionally, Hungnes (2002) found supporting results for household debt influencing the aggregated housing prices in his model for the Norwegian economy. However, Jacobsen et al. (2006) suggested the supply of debt has an individual effect on housing prices. Therefore, housing prices and household debt should be predicted simultaneously, indicating a reverse causality problem.

To illustrate the importance of household debt in our results, we exclude the variable in our models. The new performance is captured in **Table 7.3**.

Table 7.3: Results from quarterly predictions w, and w/o, household debt

Machine Learning Models	RMSE[1]	RMSE w/o *debt*	Increase in RMSE[2]
Elastic Net	3.08	3.82	24%
Random Forest	3.37	3.86	14.5%

[1] Absolute %-point error
[2] Relative comparison between the two RMSEs

Elastic Net and Random Forest are considerably weakened, increasing their RMSEs with 24% and 14.5% respectively. These results raise two interpretations. If the reverse causality from Jacobsen et al. (2006) is certain, the original models suffer from this bias, weakening the overall credibility. If the intuitions from Hungnes (2002) are certain, including household debt increases the overall performance substantially.

7.6 Limitations and Further Research

The limitations in the thesis mainly stem from the included data and methodologic approach. Our analysis seeks to apply basic and easily implemented machine learning models for predicting housing price growth in the Norwegian housing market. Hence, the limitations of this thesis provide possible suggestions and interesting ideas for future research.

An influential assumption in this thesis was to only include previously proven macroeconomic indicators. Most relevant literature included seven to ten variables, highlighting simplicity in their models. However, as Mullainathan (2017) argued, machine learning algorithms can manage large portions of data without overfitting the models. Therefore, we believe exceeding the included variables could capture more of the underlying aspects in the market, without creating overfitted models. Oddmund Berg (Personal Communication, 2020) from DNB Markets proposed to include additional data. However, this came at a later stage in the writing process. Further research could explore the inclusion of additional housing specific factors, such as property turnover volume, length of stay and average number of bidders. Furthermore, the supply of data indicators has increased exponentially, and a reasonable assumption is that macroeconomic indicators follow this trend in frequency and detail. We therefore believe machine learning models could become more applicable in the future, by utilizing the increased supply of data.

The Norwegian housing market consists of several regional submarkets (Anundsen & Mæhlum, 2017). Hence, one can either focus on the aggregated macro perspective, or the micro perspective analysis. As mentioned in section 7.5, Oslo and Stavanger experienced different reactions in the aftermath of the oil crisis. This thesis is focused on an aggregated perspective on national level. Individual or regional differences are therefore not captured in our model. Also, Alessi et al. (2011) indicated that the inclusion of microeconomic factors in a predictive analysis might capture more of the price dynamics. For further research it would therefore be interesting to include several submarkets to capture individual differences in the market.

Moreover, the main purpose was to investigate an area of machine learning yet unexplored, applying simple and intuitive models. Based on existing literature, we chose to focus on Elastic Net, Random Forest and Neural Network in our predictions. However, the inclusion of more complex models could potentially increase overall accuracy. For example, Support Vector Regression has been highlighted as the most accurate model both in the US (Plakandaras, 2015) and Australia (Milunovich, 2019). Additionally, due to unavailability of data, ensemble models tend to outperform individual models. For example, combinational models, such as Super Learner, was highlighted as effective in GDP-predictions (Jung et al., 2018; Bankson & Holm, 2019).

Lastly, this thesis implements one of many model approaches. Implementing different variations of cross-validation and training, such as K-fold, could yield other results and intuitions. Additionally, analyzing the mechanism further than model-performance to increase interpretability is preferable in macroeconomic research. The Caret-package helps us implement machine learning models in an easy and intuitive way but limits the availability of extensive model tuning and specifications. Thus, applying other packages could provide additional insights in further research.

8. Concluding Remarks

Computing accurate housing price predictions are considered important for central banks, property investors and political decision makers. In this thesis, we therefore analyze the following question: *Is machine learning suitable for predicting price growth in the Norwegian housing market?*

Overall, machine learning can capture underlying movements in the market and yield credible predictions. Still, our overall results have not outperformed predictions from Norges Bank on quarterly and yearly basis. Elastic Net predicts more accurately than Norges Bank in 29% of the testing quarters from 2013 to 2019. Additionally, Random Forest produced the most accurate yearly predictions in 2013 and 2017, as well as overall outperforming both DNB and SSB. The uncertain event after The New Mortgage Regulations influences the 2018 predictions, decreasing overall performance. Further implications are consistent with existing literature that shows how linear models are more suited in shorter horizons, while nonlinear models capture the increased complexity in longer horizons. Moreover, ensemble models tend to outperform individual models, mostly due to the small datasets available. The machine learning models have also identified to be the household debt level in Norway to be the most influential variable. Furthermore, our results indicate that predictability is certain on quarterly and yearly basis, but the overall performance is weakened with yearly predictions.

We believe machine learning is suitable in housing price predictions in normal situations but deviates during abnormal events. Further improvements and model selections could help future performance and become a valuable tool for decision makers. Still, there is a large degree of uncertainty associated with housing prices, which decreases the overall performance. The existing models have yet not captured these uncertainties to a large extent. We therefore believe this uncertainty would be present in future predictions as well as being a limitation to the overall use of machine learning.

Ultimately, predicting housing prices is considered difficult. Compared to traditional statistical models, machine learning could still extract hidden structures and combinations between variables, increasing the accuracy. Machine learning could therefore be a potential tool in future predictions. Moreover, our thesis investigated the applicability in a historical perspective. However, the thesis has not focused on forecasting future growth rates in Norway. However, after including all available data until Q4 2019, all our models have predicted yearly growth rates for 2020 and 2021 in **Table A.1.3.** Hence, it will be interesting to see whether Random Forest's 2.80%[13] prediction for 2021 coincides with the actual housing price growth in Norway.

[13] Yearly predictions for 2020 and 2021, presented in **Table A.1.3**, for Elastic Net, Random Forest and Neural Network. The predictions are based on the underlying dataset until 2019Q4. This is solely for supplementary purposes. The aim of the thesis has been to back test quarterly and yearly predictions, so that the performance could be evaluated.

References

Abidoye, R. B. (2019). Predicting property price index using artificial intelligence techniques: Evidence from Hong Kong | Emerald Insight. *International Journal of Housing Markets and Analysis*, *12*(6). https://doi.org/10.1108\/IJHMA

Alessi, L., Matteo Barigozzi, & Capasso, M. (2011, April). Nonfundamentalness in structural econometric models: a review. Retrieved December 18, 2020, from ResearchGate website: https://www.researchgate.net/publication/48910465_Nonfundamentalness_in_structural_econometric_models_a_review

Anundsen, A. K., & Røed Larsen, E. (2018). TESTING FOR MICRO-EFFICIENCY IN THE HOUSING MARKET. *International Economic Review*, *59*(4), 2133–2162. https://doi.org/10.1111/iere.12332

Anundsen, A., & Mæhlum, S. (2017). Regionale forskjeller i boligpriser og gjeld. Retrieved December 18, 2020, from Norges Bank website: https://www.norges-bank.no/aktuelt/nyheter-og-hendelser/Signerte-publikasjoner/Aktuell-kommentar/2017/aktuell-kommentar-42017/?fbclid=IwAR0rLr4neBiAneV0nExT8H-UvYpdKlkqdcJ264aElECX0g7LoAZ_j0WOE2A

Bankson, C. A., & Holm, A. M. (2019, December). Kunstig intelligens i makroøkonomisk prognosearbeid. Retrieved from NHH website: https://openaccess.nhh.no/nhh-xmlui/bitstream/handle/11250/2644325/masterthesis.pdf?sequence=1&isAllowed=y

Beltratti, A., & Morana, C. (2010). International house prices and macroeconomic fluctuations. *Journal of Banking & Finance*, *34*(3), 533–545. https://doi.org/10.1016/j.jbankfin.2009.08.020

Borg, T. (2019, January 23). Swedish Housing Prices: Leading indicators and scenario analysis | Indicio Technologies. Retrieved November 11, 2020, from Indicio Technologies website: https://www.indicio.com/swedish-housing-prices/

Borio, C. E. V., Kennedy, N., & Prowse, S. D. (1994). Exploring aggregate asset fluctuations across countries: measurement, determinants and monetary policy implications. *BIS Economic Paper*, *40*(Bank of International Settlements(Basel)). Retrieved from Monetary and Economic Dept.

Borst, R. D. (1991). The Next Modelling/Calibration Technology for the Assessment Community? *Property Tax Journal*, *10*(1), 69–94.

Breiman, L. (2001). *Random Forests*. Retrieved from University of California, Berkeley website: https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf

Brownlee, J. (2016, December 18). How To Backtest Machine Learning Models for Time Series Forecasting. Retrieved November 4, 2020, from Machine Learning Mastery website: https://machinelearningmastery.com/backtest-machine-learning-models-time-series-forecasting/

Burton, J. (2018, February 16). Cities with the Most Skyscrapers. Retrieved November 24, 2020, from WorldAtlas website: https://www.worldatlas.com/articles/cities-with-the-most-skyscrapers-in-the-world.html

Case, K. E., & Shiller, R. J. (1989). The efficiency of the market for single-family homes. *The American Economic Review*, *125–137*.

Case, K. E., & Shiller, R. J. (1990). Forecasting prices and excess returns in the housing market. *Real Estate Economics*, *18*(3), 253–273.

Chauhan, A. (2017, June 15). What Is Variable Importance and How Is It Calculated? Retrieved November 11, 2020, from dzone.com website: https://dzone.com/articles/variable-importance-and-how-it-is-calculated

Clark, T. E., & McCracken, M. (2001, February). Tests of Equal Forecast Accuracy and Encompassing for Nested Models. Retrieved December 4, 2020, from ResearchGate website: https://www.researchgate.net/publication/222691075_Tests_of_Equal_Forecast_Accuracy_and_Encompassing_for_Nested_Models

DNB Markets. (2020). Økonomiske utsikter. In *DnB*. Retrieved from https://www.dnb.no/portalfront/nedlast/no/markets/analyser-rapporter/norske/okonomiske-utsikter/HR200827.pdf

DN.no, & Riisnæs, I. G. (2020, April 23). – Mye mer sannsynlig at boligprisene faller 25 prosent. Retrieved October 14, 2020, from www.dn.no website: https://www.dn.no/eiendom/-mye-mer-sannsynlig-at-boligprisene-faller-25-prosent/2-1-794991

Donges, N. (2019). A complete guide to the random forest algorithm. Retrieved November 3, 2020, from Built In website: https://builtin.com/data-science/random-forest-algorithm

E24. (2013, November 1). Kraftig fall i boligprisene. Retrieved November 11, 2020, from E24.no website: https://e24.no/privatoekonomi/i/LA9dQR/kraftig-fall-i-boligprisene-i-oktober

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179–221.

Expert System. (2020, May 5). What is Machine Learning? A definition - Expert System. Retrieved November 12, 2020, from Expert System website: https://expertsystem.com/machine-learning-definition/

Finansavisen. (2018, December 25). Robotråskinn spår boligprisfall. Retrieved October 15, 2020, from Finansavisen website: https://finansavisen.no/nyheter/eiendom/2018/12/robotraaskinn-spaar-boligprisfall

Finansdepartementet. (2016, December 14). Fastsetter ny boliglånsforskrift. Retrieved November 11, 2020, from Regjeringen.no website: https://www.regjeringen.no/no/aktuelt/fastsetter-ny-boliglansforskrift/id2523967/

Finanstilsynet. (2020, November 19). Boliglånsundersøkelsen 2020. Retrieved November 24, 2020, from Finanstilsynet website: https://www.finanstilsynet.no/nyhetsarkiv/pressemeldinger/2020/boliglansundersokelsen-2020/

Frost, J. (2017, August 6). How to Choose Between Linear and Nonlinear Regression - Statistics By Jim. Retrieved December 4, 2020, from Statistics By Jim website: https://statisticsbyjim.com/regression/choose-linear-nonlinear-regression/

Gupta, R., Kabundi, A., & Miller, S. M. (2011). Forecasting the US real house price index: Structural and non-structural models with and without fundamentals. *Economic Modelling*, *28*(4), 2013–2021. https://doi.org/10.1016/j.econmod.2011.04.005

Gupta, R., & Miller, S. (2015). The out-of-sample forecasting performance of nonlinear models of regional housing prices in the US. Retrieved November 25, 2020, from Applied Economics website: https://www.tandfonline.com/doi/full/10.1080/00036846.2015.1005814

Halvorsen, M. T. (2014, December 28). - Mange dommedagsprofeter for boligmarkedet i 2014. De tok feil. Retrieved December 8, 2020, from www.dn.no website: https://www.dn.no/boligpriser/makropanelet/-mange-dommedagsprofeter-for-boligmarkedet-i-2014-de-tok-feil/1-1-5274951

Hoerl, A., & Kennard, R. (1970). American Society for Quality Ridge Regression: Biased Estimation for. *American Statistical Association*, *12*(1), 55–67. Retrieved from https://www.math.arizona.edu/~hzhang/math574m/Read/RidgeRegressionBiasedEstimationForNonorthogonalProblems.pdf

Holmes, K. W., Chadwick, O. A., & Phaedon Kyriakidis. (2000, June). Error in a USGS 30-Meter Digital Elevation Model and Its Impact on Terrain Modeling. Retrieved December 4, 2020, from ResearchGate website: https://www.researchgate.net/publication/222565264_Error_in_a_USGS_30-Meter_Digital_Elevation_Model_and_Its_Impact_on_Terrain_Modeling

Hungnes, H. (2002). Boligpriser, boligkapital og boligkonsum: En Makroøkonomisk modell for norsk økonomi. *Sosiale Og Økonomiske Studier*, *5.5*(108). Statistisk Sentralbyrå.

Hyndman, & Koehler. (2006). 3.4 Evaluating forecast accuracy | Forecasting: Principles and Practice. Retrieved December 4, 2020, from Otexts.com website: https://otexts.com/fpp2/accuracy.html

Hyndman, R. (2016). Cross-validation for time series | Rob J Hyndman. Retrieved November 27, 2020, from Robjhyndman.com website: https://robjhyndman.com/hyndsight/tscv/

IMF. (2020). Residential Property Price Index (RPPI) Practical Compilation Guide. Retrieved October 20, 2020, from imf.org website: https://www.imf.org/en/Data/Statistics/RPPI-guide

Jacobsen, D. H., Solberg-Johansen, K., & Haugland, K. (2006). Boliginvesteringer og boligpriser. *Penger Og Kreditt*, *34*(4/2006). Retrieved from 229-241. Norges Bank.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *Springer Texts in Statistics An Introduction to Statistical Learning*. Retrieved from http://faculty.marshall.usc.edu/gareth-james/ISL/ISLR%20Seventh%20Printing.pdf

Juel, S. (2013, December 30). Sjeføkonomene: Slik blir boligprisene i 2014. Retrieved November 10, 2020, from Aftenbladet.no website: https://www.aftenbladet.no/forbruker/i/OnG7b3/sjefoekonomene-slik-blir-boligprisene-i-2014

Jung, J.-K. (2018, November 1). An Algorithmic Crystal Ball: Forecasts-based on Machine Learning. Retrieved November 12, 2020, from IMF website: https://www.imf.org/en/Publications/WP/Issues/2018/11/01/An-Algorithmic-Crystal-Ball-Forecasts-based-on-Machine-Learning-46288

Ka, C., Leung, Y., Chi, E., & Tang, H. (2015). Availability, Affordability and Volatility: The Case of the Hong Kong Housing Market. *REAL ESTATE REVIEW*, *383*(3), 383–428. Retrieved from https://www.um.edu.mo/fba/irer/papers/current/vol18n3_pdf/05.pdf

Kenton, W. (2019, July 2). How Overfitting Works. Retrieved November 25, 2020, from Investopedia website: https://www.investopedia.com/terms/o/overfitting.asp

Kenton, W. (2020, June 6). Look-Ahead Bias. Retrieved October 26, 2020, from Investopedia website: https://www.investopedia.com/terms/l/lookaheadbias.asp

Kuo, C. L. (1996). Autocorrelation and Seasonality in the Real Estate Market. *Journal of Real Estate Finance and Economics*, *12*. 136-62.

Lam, K. C., Lam, K. Y., & Yu, C. y. (2009, April 22). An Artificial Neural Network and Entropy Model for Residential Property Price Forecasting in Hong Kong. Retrieved December 4, 2020, from Journal of Property Research website: https://www.tandfonline.com/doi/full/10.1080/09599910902837051

Landberg, Ø. (2020, May 25). Norge er mindre oljeavhengig enn ved forrige krise. Retrieved November 24, 2020, from E24.no website: https://e24.no/norsk-oekonomi/i/dOXzd1/norge-er-mindre-oljeavhengig-enn-ved-forrige-krise

Lantz, B. (2013). Machine Learning with R: Lantz, Brett: 9781782162148: Amazon.com: Books. Retrieved December 4, 2020, from Amazon.com website: https://www.amazon.com/Machine-Learning-R-Brett-Lantz/dp/1782162143

Laster, D., Bennett, P. B., & Geoum, I. S. (1997). Rational Bias in Macroeconomic Forecasts. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.993826

Lawrence, S. (1997). *Lessons in Neural Network Training: Overfitting May be Harder than Expected*. Retrieved from National Conference of Artificial Intelligence website: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.38.6468&rep=rep1&type=pdf

Lu, H., Zou, N., Jacobs, R., & Morgan, D. (2019, July). Error assessment and optimal cross-validation approaches in machine learning applied to impurity diffusion. Retrieved November 25, 2020, from ResearchGate website: https://www.researchgate.net/publication/334513458_Error_assessment_and_optimal_cross-validation_approaches_in_machine_learning_applied_to_impurity_diffusion

Lundesgaard, E. (2019, September 30). Why do the price indices from Real Estate Norway and Statistics Norway vary? Retrieved October 20, 2020, from Eiendom Norge website: https://eiendomnorge.no/blog/why-do-the-price-indices-from-real-estate-norway-and-statistics-norway-vary-article1459-944.html

Maroto, M. (2017, April 20). Neural Networks = Black Box? - Towards Data Science. Retrieved December 4, 2020, from Medium website: https://towardsdatascience.com/neural-networks-black-box-b20723f9a417

Max, K. (2008, November). Building Predictive Models in R Using the caret Package. Retrieved December 4, 2020, from ResearchGate website: https://www.researchgate.net/publication/26568624_Building_Predictive_Models_in_R_Using_the_caret_Package

Miles, W. (2007). Boom–Bust Cycles and the Forecasting Performance of Linear and Non-Linear Models of House Prices. *The Journal of Real Estate Finance and Economics*, *36*(3), 249–264. https://doi.org/10.1007/s11146-007-9067-1

Milunovich, G. (2019, July 11). Forecasting Australian Real House Price Index: A Comparison Study of Machine Learning and Time Series Methods. Retrieved October 13, 2020, from ResearchGate website: https://www.researchgate.net/publication/334388950_Forecasting_Australian_Real_House_Price_Index_A_Comparison_Study_of_Machine_Learning_and_Time_Series_Methods

Mullainathan, S. (2017). Machine Learning and Prediction in Economics and Finance [YouTube Video]. Retrieved from https://www.youtube.com/watch?v=xl3yQBhI6vY&t=4617s

Myrvang, S. (2020, September 20). Oljekrisen svir fortsatt: Må selge med tap seks år etter boligkjøpet. Retrieved December 4, 2020, from Www.vg.no website: https://www.vg.no/nyheter/i/JJ1oLJ/oljekrisen-svir-fortsatt-maa-selge-med-tap-seks-aar-etter-boligkjoepet

Naug, B., & Jacobsen, D. H. (2004). Hva driver boligprisene? Retrieved from Norges Bank website: https://www.norges-bank.no/globalassets/upload/publikasjoner/penger_og_kreditt/2004-04/jacobsen.pdf

Norges Bank. (2020, September 24). Pengepolitisk rapport med vurdering av finansiell stabilitet 3/2020. Retrieved December 4, 2020, from Norges-bank.no website: https://www.norges-bank.no/aktuelt/nyheter-og-hendelser/Publikasjoner/Pengepolitisk-rapport-med-vurdering-av-finansiell-stabilitet/2020/ppr-32020/

NRK Debatten. (2020, October 13). Det vanskelige boligmarkedet. Retrieved October 22, 2020, from NRK TV website: https://tv.nrk.no/serie/debatten/202010/NNFA51101320/avspiller

Palachy, S. (2019, August). Detecting stationarity in time series data. Retrieved October 22, 2020, from KDnuggets website: https://www.kdnuggets.com/2019/08/stationarity-time-series-data.html

Plakandaras, V., Gupta, R., Gogas, P., & Papadimitriou, T. (2015). Forecasting the U.S. Real House Price Index. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2431627

Real Estate Norway. (2020). *Hovedrapport September 2020*. Eiendomsstatistikk: Eiendom Norge. https://eiendomnorge.no/boligprisstatistikk/statistikkbank/rapporter/manedsrapporter/.

Røed Larsen, E., & Weum, S. (2008). Testing the efficiency of the Norwegian housing market. *Journal of Urban Economics*, *64*(2), 510–517. https://doi.org/10.1016/j.jue.2008.05.004

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. Retrieved November 13, 2020, from Apa.org website: https://psycnet.apa.org/record/1959-09865-001

Schölkopf, G. J. (2008). *Causality: Objectives and Assessment Challenges in Machine Learning, Volume 4*. Retrieved from http://www.causality.inf.ethz.ch/ciml/CiML-v4-book.pdf

Senneset, K. (2018, May 2). Å lage boligprisprognoser – det er fali' det - Prognosesenteret. Retrieved November 12, 2020, from Prognosesenteret website: https://prognosesenteret.no/a-boligprisprognoser-fali/

Sharma, A. (2018, November 4). Decrypting your Machine Learning model using LIME - Towards Data Science. Retrieved November 28, 2020, from Medium website: https://towardsdatascience.com/decrypting-your-machine-learning-model-using-lime-5adc035109b5

Shukry, M., Radzi, M., Muthuveerappan, C., Kamarudin, N., Sarrazin, I., & Mohammad. (2012). FORECASTING HOUSE PRICE INDEX USING ARTIFICIAL NEURAL NETWORK. *International Journal of Real Estate Studies*, *7*(1), 2012. Retrieved from https://www.utm.my/intrest/files/2013/08/FORECASTING-HOUSE-PRICE-INDEX-USING-ARTIFICIAL-NEURAL-NETWORK.pdf

Soni, D. (2018, March 22). Supervised vs. Unsupervised Learning - Towards Data Science. Retrieved November 20, 2020, from Towards Science website: https://towardsdatascience.com/supervised-vs-unsupervised-learning-14f68e32ea8d

Stark, T., & Croushore, D. (2002). Forecasting with a real-time data set for macroeconomists. *Journal of Macroeconomics*, *24*(4), 507–531. Retrieved from https://econpapers.repec.org/article/eeejmacro/v_3a24_3ay_3a2002_3ai_3a4_3ap_3a507-531.htm

Statistics Norway. (2008). 1. What is seasonal adjustment? Retrieved December 4, 2020, from Www.ssb.no website: https://www.ssb.no/a/english/kortnavn/byggeareal_en/sesongjustering_en.html

Stranden, A. L. (2020, January 14). Nytt dataprogram spår fortsatt stigende boligpriser i 2020. Retrieved November 20, 2020, from Forskning.no website: https://forskning.no/data-penger-og-privatokonomi-okonomi/nytt-dataprogram-spar-fortsatt-stigende-boligpriser-i-2020/1620010

Strømnes, S. (2019, January 4). Høyt omsetningsvolum i boligmarkedet 2018 - NEF.no. Retrieved November 11, 2020, from NEF.no website: https://www.nef.no/nyheter/hoyt-omsetningsvolum-i-boligmarkedet-2018/

Sucarrat, G. (2019). *How precise are the forecasts of the Norwegian national budget?* Retrieved from https://mpra.ub.uni-muenchen.de/96850/1/MPRA_paper_96850.pdf

Tiffin, A. (2016). Seeing in the Dark: A Machine-Learning Approach to Nowcasting in Lebanon. Retrieved from IMF Working Paper website: https://www.imf.org/external/pubs/ft/wp/2016/wp1656.pdf

Valland, R. A. (2019). *Nowcasting av norsk bnp med maskinlæringsalgoritmer*. Universitetet i Bergen. Retrived: from https://bora.uib.no/bora-xmlui/bitstream/handle/1956/20904/Masteroppgave.pdf?sequence=1&isAllowed=y

Vanags, J., Ineta Geipele, Auseklis Sarkans, & Didzis Usenieks. (2017, January 26). Housing Heterogeneity Dimensions and Their Elements: A Systemic Approach. Retrieved December 4, 2020, from ResearchGate website: https://www.researchgate.net/publication/318427471_Housing_Heterogeneity_Dimensions_and_Their_Elements_A_Systemic_Approach

Vandeput, N. (2019, July 5). Forecast KPI: RMSE, MAE, MAPE & Bias - Towards Data Science. Retrieved November 5, 2020, from Medium website: https://towardsdatascience.com/forecast-kpi-rmse-mae-mape-bias-cdc5703d242d

Woloszko, N. (2017). Machine Learning for Economic Forecasting | Nicolas Woloszko [YouTube Video]. Retrieved from https://www.youtube.com/watch?v=VnoRay3YO48&t=596s

Wooldridge, J. M. (2012). Introductory Econometrics, A Modern Approach. Retrieved from https://economics.ut.ac.ir/documents/3030266/14100645/Jeffrey_M._Wooldridge_Introductory_Econometrics_A_Modern_Approach__2012.pdf

Zarnowitz, V., & Braun, P. (1992). *Twenty-two Years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisons of Forecasting Performance*. https://doi.org/10.3386/w3965

Zou, H., & Hastie, T. (2005). Addendum: Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(5), 768–768. https://doi.org/10.1111/j.1467-9868.2005.00527.x

Appendices

A.1 Data

Table A.1.1: Quarterly predictions from Norges Bank

Prediction Quarter	Predicted Housing Growth (y/y)[1]	Time	Source (MPR)[2]
2019Q4	3.12	18/09/19	MPR 3/19
2019Q3	1.32	20/06/19	MPR 2/19
2019Q2	1.81	20/03/19	MPR 1/19
2019Q1	2.32	12/12/18	MPR 4/18
2018Q4	2.98	19/09/18	MPR 3/18
2018Q3	3.05	20/06/18	MPR 2/18
2018Q2	-1.4	21/03/18	MPR 1/18
2018Q1	-3.03	13/12/17	MPR 1/18
2017Q4	-0.06	20/09/17	MPR 3/17
2017Q3	4.98	21/06/17	MPR 2/17
2017Q2	10.75	15/03/17	MPR 1/17
2017Q1	12.11	14/12/16	MPR 4/16
2016Q4	10.51	21/09/16	MPR 3/16
2016Q3	7.27	23/06/16	MPR 2/16
2016Q2	4.09	17/03/16	MPR 1/16
2016Q1	5.13	17/12/15	MPR 4/15
2015Q4	6.27	24/09/15	MPR 3/15
2015Q3	6.52	18/06/15	MPR 2/15
2015Q2	8.38	19/03/15	MPR 1/15
2015Q1	7.99	11/12/14	MPR 4/14
2014Q4	5.19	18/09/14	MPR 3/14
2014Q3	1.80	19/06/14	MPR 2/14
2014Q2	-0.63	27/03/14	MPR 1/14
2014Q1	-2.76	05/12/13	MPR 4/13
2013Q4	2.59	19/09/13	MPR 3/13
2013Q3	5.86	20/06/13	MPR 2/13
2013Q2	7.91	14/03/13	MPR 1/13

[1] Percentage Growth
[2] Monetary Policy Report from Norges Bank

Table A.1.2: Yearly predictions from professional institutions

Institution	Year	Predicted Growth[1]	Time	Source
Norges Bank	2019	1.60	12/12/18	MPR[2] 4/18
Norges Bank	2018	-1.60	13/12/17	MPR[2] 4/17
Norges Bank	2017	8.71	14/12/16	MPR[2] 4/16
Norges Bank	2016	4.77	17/12/15	MPR[2] 4/15
Norges Bank	2015	6.63	11/12/14	MPR[2] 4/14
Norges Bank	2014	-1.57	05/12/13	MPR[2] 4/13
Norges Bank	2013	7.66	31/10/12	MPR[2] 3/12
DNB	2019	0.5	24/01/19	Økonomiske utsikter 2019 nr. 1 p.52
DNB	2018	-3.5	24/01/18	Økonomiske utsikter 2018 nr. 1 p.51
DNB	2017	9.0	17/01/17	Økonomiske utsikter 2017 nr. 1 p.89
DNB	2016	2.0	17/02/16	Økonomiske utsikter 2016 nr. 1 p.105
DNB	2015	2.1	15/01/15	Økonomiske utsikter 2015 nr. 1 p.104
DNB	2014	-4.0	27/01/14	Økonomiske utsikter 2014 nr. 1 p.85
DNB	2013	7.5	16/01/13	Økonomiske utsikter 2013 nr. 1 p.93
SSB	2019	1.1	07/03/18	Konjunkturtendensene 1/2018 p.20
SSB	2018	-5.0	30/11/17	Økonomiske analyser 4/2017 p.30
SSB	2017	7.2	01/12/16	Økonomiske analyser 5/2016 p.46
SSB	2016	1.5	03/12/15	Økonomiske analyser 4/2015 p.46
SSB	2015	0.2	04/12/14	Økonomiske analyser 6/2014 p.46
SSB	2014	-2.2	06/12/13	Økonomiske analyser 5/2013 p.46
SSB	2013	6.8	01/06/12	Økonomiske analyser 6/2012 p.46

[1] Percentage average growth
[2] Monetary Policy Report

Table A.1.3: Yearly predictions for 2020 and 2021[1]

Year	Elastic Net	Random Forest	Neural Network
2020	3.17	1.99	0.97
2021	4.58	2.80	2.68

[1]Based on the dataset used in the rest of the analysis, 1996Q2 to 2019Q4

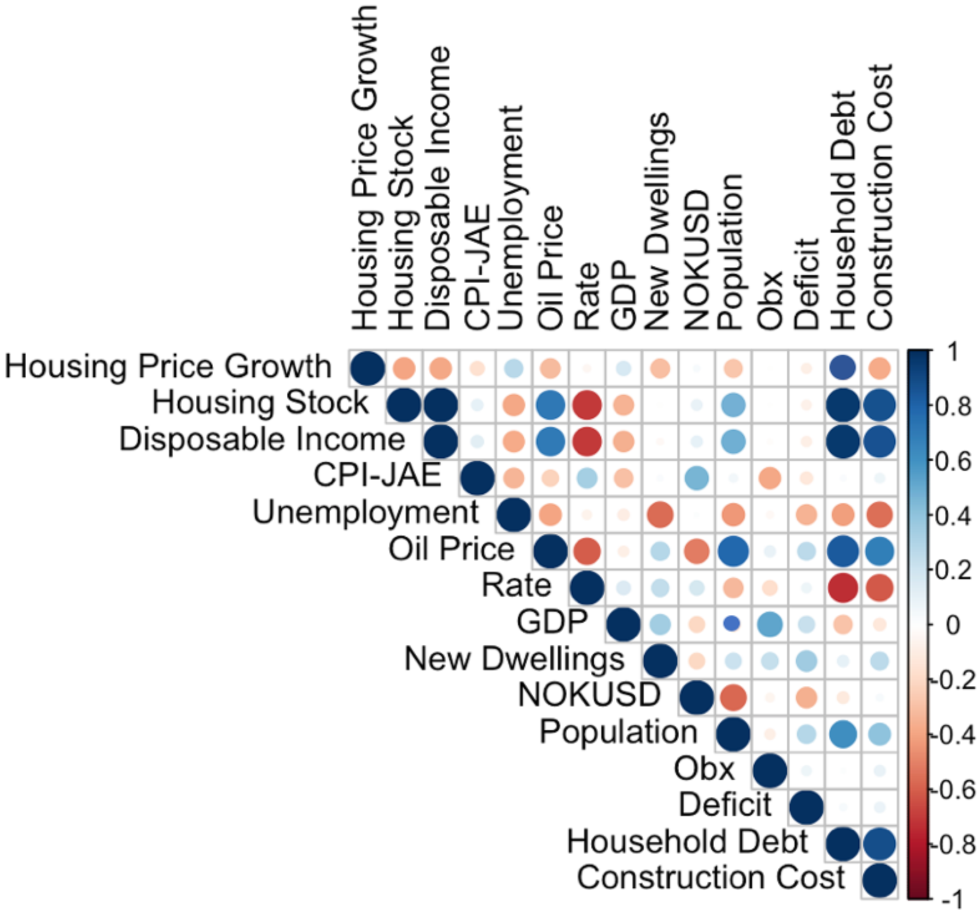
A.2 Covariates

Table A.2.1: Covariates in the dataset

Variable	Group	Type	Source	Frequency[1]
HPI Norway	House Specific	Index	Real Estate Norway	Monthly
Construction Cost Index	House Specific	Index	SSB	Monthly
Household Debt	House Specific	Percentage	SSB	Quarterly
Housing Stock	House Specific	Percentage Growth (y/y)	SSB	Quarterly
New Dwellings	House Specific	Index	SSB	Quarterly
CPI-JAE	Business Cycle	Index	SSB	Monthly
Exchange Rate NOK/USD	Business Cycle	Nominal	Bloomberg	Daily
GDP Mainland Norway	Business Cycle	Nominal	SSB	Monthly
Household Income	Business Cycle	Absolute	SSB	Quarterly
Interest rate after tax	Business Cycle	Percentage	SSB	Quarterly
OBX Stock Index	Business Cycle	Index	Bloomberg	Daily
Brent Oil	Business Cycle	Absolute	Bloomberg	Daily
Budget Surplus/Deficit	Business Cycle	Nominal	SSB	Quarterly
Population	Business Cycle	Absolute	SSB	Quarterly
Unemployment	Business Cycle	Percentage	NAV	Monthly

[1]Frequency divided into daily, monthly and quarterly. The dataset used in the analysis is, however, transformed to quarterly data

Figure A.2.1: Correlation Matrix



Note: Correlation matrix of the 14 variables used in the analysis. The depended variable is also included in the matrix. The colors range from dark blue to red, where dark blue indicates a strong positive correlation, in contrast to red which indicates a strong negative correlation.

A.3 Other Figures

Figure A.3.1: Cross-Validation from Random Forest

```
Random Forest

96 samples
14 predictors

Pre-processing: centered (14), scaled (14)
Resampling: Rolling Forecasting Origin Resampling (2 held-out with no fixed window)
Summary of sample sizes: 68, 69, 70, 71, 72, 73, ...
Resampling results across tuning parameters:

  mtry  RMSE      Rsquared  MAE
   1    3.455453  1         3.270346
   2    3.351169  1         3.157258
   3    3.309990  1         3.119608
   4    3.306348  1         3.105947
   5    3.285241  1         3.077244
   6    3.325552  1         3.119918
   7    3.306447  1         3.093930
   8    3.296405  1         3.079322
   9    3.327645  1         3.134018
  10    3.371107  1         3.163697
  11    3.347889  1         3.119470
  12    3.369966  1         3.149049
  13    3.391046  1         3.175228
  14    3.401293  1         3.190177

RMSE was used to select the optimal model using the smallest value.
The final value used for the model was mtry = 5.
```

Figure A.3.2: Cross-Validation from Elastic Net

```
glmnet

96 samples
14 predictors

Pre-processing: centered (14), scaled (14)
Resampling: Rolling Forecasting Origin Resampling (2 held-out with no fixed window)
Summary of sample sizes: 68, 69, 70, 71, 72, 73, ...
Resampling results across tuning parameters:

  alpha  lambda       RMSE      Rsquared  MAE
  0.10   0.005246505  4.857795  1         4.512515
  0.10   0.052465052  4.466824  1         4.162206
  0.10   0.524650520  3.639156  1         3.419370
  0.55   0.005246505  4.847948  1         4.503167
  0.55   0.052465052  4.348599  1         4.065839
  0.55   0.524650520  3.755119  1         3.531705
  1.00   0.005246505  4.836498  1         4.492032
  1.00   0.052465052  4.226589  1         3.970660
  1.00   0.524650520  3.774863  1         3.629379

RMSE was used to select the optimal model using the smallest value.
The final values used for the model were alpha = 0.1 and lambda = 0.5246505.
```