Abstract Zhang, F. 2020. Applications of artificial neural networks for time series data analysis in energy domain. Dalarna Licentiate Theses 14. Borlänge: Dalarna University. ISBN 978-91-88679-08-6. With the development of artificial intelligence techniques and increased installation of smart meters in recent years, time series analysis using historical data in the energy domain becomes applicable. In this thesis, microdata analysis approaches are used, which consist of data acquisition, data processing, data analysis and data modelling, aiming to address two research problems in the energy domain. The first research problem is short-term electricity price forecasting of a deregulated market and the second one is anomaly detection of heat energy usage in district heating substations. As a result of electricity market deregulation, third party suppliers can enter the market and consumers are free to choose electricity suppliers, which leads to a more transparent and competitive market. Accurate short-term electricity price forecasting is crucial to the market participants in terms of maximizing profits, risk management and other short-term market operations. Literature review is performed aiming to identify the suitable methods. It is concluded that long short-term memory (LSTM) based methods are superior to other methods for time series analysis. Since the gating mechanisms of long short-term memory alleviate the problem of gradient vanishing. Another conclusion form the literature is that hybrid approach that consists of two or more artificial intelligence algorithms complimenting each other is more effective to solve complex real world problem. Based on the conclusions derived, a hybrid approach based on bidirectional LSTM (BDLSTM) and Catboost is proposed for short-term electricity price forecasting of NordPool. Performance of support vector regression (SVR), ARIMA, ensemble tree, multi-layer perception (MLP), gated recurrent unit (GRU), BDLSTM and LSTM are evaluated. Experiment results show that BDLSTM outperforms the other models in terms of Mean percentage error (MAPE), root mean square error (RMSE) and mean absolute error (MAE). Statistics show that market shares of district heating have increased steadily in the past five decades. District heating shares approximately 55% of the heat supply market in Sweden. Therefore, energy efficiency of district heating systems is of great interest to energy stakeholders. Anomalies are rare observations deviated significantly from the majority of the data, and such suspicious observations are important indicators of potential faults. To reduce the financial loss and improve energy efficiency, detecting anomalies from meter readings is essential. Another type of neural network architecture, LSTM variational autoencoder (LSTMVAE) combined with a heat signature model is proposed for anomaly detection using the dataset from an anonymous substation in Sweden. Results show that the proposed method outperforms other two baseline models LSTM, LSTM autoencoder (LSTMAE) in terms of F1 score and AUC. In this thesis, various approaches based on neural networks are explored to solve different time series data analysis in the energy domain, aiming for supporting decision makings of market participants to maximize profits, enhancing risk managements and improving energy efficiency. Although, two problems domains are covered, methods reviewed and applied in the thesis can be tailored for other energy time series analysis problems as well. Keywords: Deregulated energy market, electricity prices, district heating, energy efficiency, neural networks.

A taxonomy of different types of ANNs that have been used for electricity price forecasting is shown in Figure 3. The taxonomy of ANNs may vary. For example, a shallow multi-layer perception neural network (MLPNN) can become deep by adding more hidden layers. However, the structure of a conventional MLPNN is shallow. Figure 3. Taxonomy of ANN based models for electricity price forecasting Past studies with respect to electricity prices forecasting using a single ANN based models from year 2010 afterwards are summarized briefly below: In [4], a MLPNN based model was proposed for day ahead locational marginal prices (LMPs) in the pool based energy markets. Main contribution of this study was the presented Bayesian learning algorithm that avoids the over fitting problem by controlling model complexity. A MLPNN based method was proposed in [5] to forecast the day ahead electricity of European market. Main contributions of this study were the use of a deeper MLPNN architecture with two hidden layers and a novel feature selection algorithm based on Bayesian optimization. RBFNN was applied in [6] for short-term electricity price forecasting. The main contribution of this study is optimization of learning rate parameters using Orthogonal Experimental Design (OED) algorithm. A ELM based approach was proposed in [7] for electricity price forecasting. Major improvement of this work was incorporating bootstrapping algorithm for uncertainty estimation. A Local Linear Wavelet Neural Network (LLWNN) approach was proposed in [8] for day ahead electricity price forecasting. Comparing with conventional WNN, weights between hidden and output layer of WNN was replaced by a local linear model in LLWNN. Main advantage of the proposed method is its local capacity of Wavelet basis functions with less hidden units. The methods used in aforementioned studies were feedforward neural networks, while a recurrent neural networks (RNN) based approach was used in [9] for one hour ahead electricity spot price forecasting. Expectation Maximization (EM) algorithm with Kalman filtering and smoothing were used to estimate both noise in the data and model uncertainty. In addition, as type of recurrent neural network, Elman neural network was used in studies [10] and [11]. A Gated Recurrent Unit (GRU) based recurrent neural network was used in [12]. A comparison between shallow and deep ANNs were performed in this study, concluding that overall deep ANNs outperform shallow ANNs. 2.2 literature review of electricity prices forecasting using ANN based hybrid models ANN based hybrid models are models that consist of ANNs and other machine learning algorithms [13]. In general, by combing various models scientifically, hybrid models become more robust and capable of dealing with problems that are complex in nature than a single model can. There are eight ANN based hybrid approaches covered in the past studies from year 2015 afterwards. The number of studies for each hybrid approach from year 2015 to 2018 is shown in Figure 4. Figure 4. Number of studies for each hybrid model per year In study [14], a combination of K-Nearest Neighbor (KNN) with ANN was proposed. KNN algorithm is utilized for selecting relevant input and reducing the size of training data, aiming to improve the performance of a single ANN. Capability of modeling relationship between input and output of each subset as well as the dependencies between input variables of a subset is the major Advantage of k-NN algorithm. Similarly, in [15], ANN was combined with another clustering algorithm based on Self-Organizing Maps (SOM). SOM was used to partition electricity price data into clusters according to the similarity of dynamic properties, and each cluster is an approximately stationary series. As a result, ANN performs better using the pre-processed data by SOM. An ensemble ANN approach was proposed in [16], to form such an ensemble ANN model, different learning algorithms, such as Levenberg-Marquardt (LM), Levenberg-Marquardt (BFGS) and Bayesian Regularization (BR), or different types of ANN were selected to be used for each ANN involved. A combination of Fuzzy logic and ANN was proposed in [17], fuzzy logic was applied to transform linguistic information in to numeric data and soften high frequency changes of series data, which is commonly observed in the electricity price series. Hybrid structured deep ANNs models were used in [5] and [18]. Different types of deep ANNs were combined to form a hybrid model in these two studies and experiments showed that such combinations outperform a single deep ANN model. Another type of commonly used hybrid models were combining Evolutionary Algorithms (EA) with ANN [19] and its variations [20], [21], [22], [23]. In general, EA were used to search for the optimal parameters of ANN in such combinations. 2.3 Conclusions from past studies of electricity prices forecasting using ANN based methods From Section 2.1, although all reviewed single ANN based models have been successfully applied to electricity price forecasting, characteristics of each individual ANN is different. Shallow feed forward neural networks such as MLPNN, RBFNN and ELM were more adopted before 2013, popularity of these neural networks decreases in recent years due to certain drawbacks. The main reason is that MLPNN is vulnerable to outliers present in electricity prices. Performance of RBFNN suffers when the dimension of input data is high. In terms of ELM, the introduced randomness leads to uncertainty in modeling dependencies presented in electricity prices. On the other hand, WNN was presented with respect to electricity price forecasting in one study in 2015. Although WNN is proved to be efficient in dealing with chaotic electricity prices and capable of modelling non-stationary high frequency price series, WNN suffers from the curse of high dimensionality. Besides, the choice of an appropriate mother wavelet function is difficult. Therefore, applications of WNN for electricity price forecasting are not widely found in recent literature. Comparing with the feed forward neural networks, RNNs are more suitable for capturing dependencies present in time series data. Therefore, RNNs are the most widely adopted neural networks for electricity prices forecasting in recent studies. However, the length of decencies captured by shallow RNNs proposed in early years is limited, which is a major drawback. Deep neural networks, which in general are more capable of solving complex problems, attract researchers’ attentions in this field. There are two relevant studies [5][18] in 2018 in which deep neural networks are adopted. Results of these two studies show that deep neural networks outperform shallow neural networks. Besides, among all deep neural networks LSTM is regarded as the most suitable model for time series forecasting due to the capability of capturing long-term dependencies of electricity prices. Tackling the problem of vanishing gradient is another major advantage of LSTM. However, LSTM is not comprehensively studied for electricity price forecasting in existing literature. Another conclusion is that a single neural network model sometimes is not efficient enough to solve complex real-world problems and its performance can be improved by combing ANN models with other AI models. From Section 2.2, two types of hybrid approaches [15], [14] were proposed in several studies from 2015 and 2016: combination of clustering algorithms with ANNs and combing a single EA with ANN. These two types are relatively easy to implement. However, there are a few drawbacks. To be more specific, in terms of the most commonly applied K-NN algorithm, the hyper parameter K needs to be carefully pre-specified. On the other hand, the single EA with ANN approach was improved by adding Wavelet transformation to the hybrid model. In fact, such improved approaches with respect to electricity price forecasting were popular among researchers in 2017 and 2018. Major enhancement made by this approach is the utilization of Wavelet transformation that decomposes original electricity prices data into a set of well-behaved subseries which is easier for prediction. Another variation of Wavelet EA ANN hybrid model is Wavelet EA ARMA ANN. Motivation of adding ARMA to the hybrid model is that ARMA is proved to be efficient in modeling the linear part of electricity prices. However, results of hybrid models with EA algorithms are not necessarily reproducible or globally optimal. Therefore, justification of the result is difficult. In 2017, combination of fuzzy and ANN was proposed. Though fuzzy c-means algorithm was adopted to reduce the number of fuzzy rules, the task of deriving fuzzy rules itself is non-trivial. With the development of parallel-processing techniques and GPU, efficiency of deep neural networks is improved. In addition, deep neural networks generally outperform shallow neural networks when solving complex problems. Due to the aforementioned factors, hybrid deep neural network models became the most widely used approaches by the researchers in recent studies. Another conclusion is that there is a lack of recognized benchmarking procedure and a standard dataset for benchmarking of different models, which make the direct comparison of different results difficult. Most researchers choose the dataset of a particular electricity market that is aligned with their own research interest, use different error measures, length of training/testing time steps and so forth and thus results from different papers are not directly comparable.

In Paper III “, lagged series were created by shifting the original series to create features for time series analysis. Then, autocorrelation analysis was performed to understand which lagged series are important to determine the current series values. According to the autocorrelation analysis results, lagged series that indicate potential autocorrelations at a 95% confidence interval were selected as the initial candidate input features. To reduce the complexity of the model and avoid potential overfitting, a second phase of feature ranking and feature selection was performed using a boosting-based algorithm. As there were categorical variables involved and conventional algorithms are not efficient to deal with them. For instance, one-hot encoding algorithm introduce extra memory cost when processing categorical variables. Therefore, CatBoost algorithm that was chosen to tackle this problem. The importance score of each feature was calculated and ranked by CatBoost, less important features were filtered out according to a pre-defined threshold score. The dataset was standardized into zero mean and unit standard deviation before model building. In Paper IV, the distribution of original heat energy usage series was plotted against outdoor temperature for examining the relationship between outdoor temperature and heat energy usage visually. According to the literature, heat load consists of space heating usage, hot water usage and other processes. In addition, there exists a strong relationship between the distribution of outdoor temperatures and space heat power demand. Theoretically, the energy signature [34] is the proportionality constant between outdoor temperature and space heat power. Therefore, space heating of the heat load is estimated given the outdoor temperature using the energy signature. After that, the original heat energy series is decomposed into two parts, the regular space heating usage part and random or semi-random residuals that consist of hot water preparation and other processes. The processed series for data modeling consists of the residuals component only by subtracting the space heating usage from the original heat load series. 3.3 Data modelling In Paper III, the processed dataset is divided into 80% training and 20% testing. MLP, SVR, ensemble tree, ARIMA and BDLSTM models are trained. Mean percentage error (MAPE), root mean square error (RMSE) and mean absolute error (MAE) are used as performance measures in this study. In Paper IV, the processed dataset is divided into 75% for training, 10% for validation and 15% for testing. Two types of models are built in this study. First, a physical model, heat signature is built to model the relationship between outdoor temperature and heat energy usage. Then, three neural network-based models, LSTM, LSTM autoencoder (LSTMAE) and LSTM variational autoencoder (LSTMVAE) are trained to reconstruct the original series. Then, the distribution of reconstruction errors is visualized and analyzed. Different threshold values are applied to filter the anomalies. F1 score and AUC are used as performance measures of each model in this study. 3.4 Paper III Proposed Method In Paper III, a hybrid approach based on BDLSTM and Catboost algorithms is proposed for short-term electricity prices forecasting of a deregulated market is proposed. The overall process of the proposed approach is illustrated in Figure 5. Figure 5. The overall process of the proposed approach From Figure 5, after data collection and visualization, a two-phase feature selection process is performed. In phase one, autocorrelation tests of four series data are performed to filter statistically significant legs of the original series, which indicate potential autocorrelations at a 95% confidence interval. In phase two, the filtered legs along with the three categorical variables are used as initial candidate features and fed into the Catboost algorithm for a second round feature selection. The motivations of using Catboost are it process categorical variables more efficiently than other conventional algorithms. In addition, it eliminates less important features in terms of a lower ranking score so that the model is light weighted and less likely to be overfitted. After feature selection and data preprocessing, the dataset is split into 80% for training and 20% for testing. The training data is used to build BDLSTM, SVR, MLP, ARIMA, ensemble tree, GRU and LSTM models for forecasting. MAPE, RMSE and MAE of these models are measured as well as training and testing time of each model. 3.5 Paper IV Proposed Method The neural network-based methods reviewed in the literature with respect to electricity prices forecasting can be modified and used for performing other time series data analysis in the energy domain as well. In this study, autoencoder architectures are designed and developed to address another research problem in energy domain: Anomaly Detection of Heat Energy Usage in District Heating Substations. The overall process of the proposed approach is illustrated in Figure 6. Figure 6. The overall process of the proposed approach According to the literature, there exists a strong relationship between the distribution of outdoor temperatures and space heat power demand. Theoretically, the energy signature is the proportionality constant between outdoor temperature and space heat power. Therefore, a physical model ‘heat signature’ is used to model the relationship between outdoor temperatures and heat energy usage. Then, the original series is decomposed into space heating component and hot water, other processes that are random or semi random residuals. 75% of the normalized residuals are fed to the neural network models for training, while 10% of the dataset is used for validation and the rest 15% of data is kept for testing. Anomalies are filtered by applying different threshold values to the reconstruction errors of the trained model. F1 score and AUC of three models, LSTM, LSTM AE and LSTM VAE are measured. 4. Results and discussions Paper I & II included in this thesis Identify potential methods for electricity prices forecasting, while Paper III addresses the research problem of short-term electricity prices forecasting using an approach based on the conclusions derived from Paper I & II. Paper IV addresses the research problem of identifying anomalies of a district heating system. Results of Paper III and IV are presented in Section 4.1 and 4.2 along with the corresponding discussions.