BRNO UNIVERSITY OF TECHNOLOGY

BRNO UNIVERSITY OF TECHNOLOGY

FACULTY OF MECHANICAL ENGINEERING

FACULTY OF MECHANICAL ENGINEERING

INSTITUTE OF PROCESS ENGINEERING

INSTITUTE OF SOLID MECHANICS, MECHATRONICS AND BIOMECHANICS

COMBINED HEAT AND POWER PRODUCTION PLANNING IN A WASTE-TO-ENERGY PLANT USING MACHINE LEARNING

PLÁNOVÁNÍ VÝROBY TEPLA A ELEKTŘINY V ZAŘÍZENÍ NA ENERGETICKÉ VYUŽITÍ DPADU S VYUŽITÍM STROJOVÉHO UČENÍ

MASTER'S THESIS

DIPLOMOVÁ PRÁCE

AUTHOR Bc. Marek Kollmann

AUTOR

SUPERVISOR Ing. Michal Touš, Ph.D.

VEDOUCÍ PRÁCE

BRNO 2023

Faculty of Mechanical Engineering, Brno University of Technology / Technická 2896/2 / 616 69 / Brno

Assignment master’s Thesis

Institute:	Institute of Process Engineering
Student:	**Bc. Marek Kollmann**
Degree program:	Process Engineering
Branch:	no specialisation
Supervisor:	**Ing. Michal Touš, Ph.D.**
Academic year:	2022/23

As provided for by the Act No. 111/98 Coll. on higher education institutions and the BUT Study and Examination Regulations, the director of the Institute hereby assigns the following topic of Master's Thesis:

Combined heat and power production planning in a Waste-to-Energy plant using machine learning

Brief Description:
Short-term planning of combined heat and power production is an important part of the operation of Waste-to-Energy plants. The process is often affected by a number of unknowns, which complicates the planning. Modern machine learning algorithms are already at a very good level and have the potential to support such planning.

Master's Thesis goals:
– analysis of operational data
– model for planning of combined heat and power production
– discussion of results

Recommended bibliography:

GÉRON, Aurélien. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: concepts, tools, and techniques to build intelligent systems. Second edition. Beijing: O'Reilly, 2019. ISBN 1492032646.

Deadline for submission Master's Thesis is given by the Schedule of the Academic year 2022/23

In Brno,

L. S.

prof. Ing. Petr Stehlík, CSc., dr. h. c. Director of the Institute	doc. Ing. Jiří Hlinka, Ph.D. FME dean

ABSTRACT

This research deployed machine learning to optimize day-ahead production planning in Waste-to-Energy (WtE) plants, grappling with issues like noisy data, uncontrollable external consumption, and fluctuating steam production due to waste as a fuel source. The primary aim was to accurately predict the power transferred to the grid, which was achieved by creating a comprehensive model consisting of seven sub-models in cascade. Each sub-model was critically evaluated using standard metrics like R^2 and Mean Relative Error. Findings revealed a significant improvement in prediction accuracy, resulting in more balanced production plans and reduced operational penalties. The approach led to an estimated annual increase of power delivered by 13% and profit by 2.6 million CZK for a specific plant.

Key words:

Waste-to-Energy, WtE, Machine Learning, ML, Predictive Algorithms, Artificial Neural Networks, ANN, Light Gradient Boosting Machine, LGBM, Energy Management, Energy Management, Quantile Models, Combined heat and power, CHP, CHP planning, Steam production planning, Power planning.

ABSTRAKT

V rámci tohoto výzkumu bylo použito strojové učení k optimalizaci plánování výroby na den dopředu zařízení na energetické zpracování odpadu (Waste-to-Energy, WtE), které se potýká s problémy, jako jsou nekvalitní data, nekontrolovatelná externí spotřeba a kolísající výroba páry v důsledku použití odpadu jako zdroje paliva. Hlavním cílem bylo předpovídat s vysokou přesností výkon přenášený do sítě, čehož bylo dosaženo vytvořením komplexního modelu sestávajícího ze sedmi dílčích kaskádovitě uspořádaných modelů. Každý dílčí model byl kriticky vyhodnocen pomocí standardních ukazatelů, jako je R^2 a průměrná absolutní chyba. Zjištění odhalila významné zlepšení přesnosti předpovědí, což vedlo k vyváženějším výrobním plánům a snížení provozních penále. Tento přístup vedl k odhadovanému ročnímu zvýšení dodaného výkonu o 13 % a zisku o 2,6 milionu Kč pro konkrétní závod.

Klíčová slova:

Waste-to-Energy, WtE, Strojové učení, ML, Prediktivní algoritmy, Umělé neuronové sítě, ANN, Light Gradient Boosting Machine, LGBM, Řízení plánování energií, Řízení energie, Kvantilové modely, Kogenerační výroba tepla a elektřiny, CHP, Plánování CHP, Plánování výroby páry, Plánování energie.

ROZŠÍŘENÝ ABSTRACT

Vzhledem ke značným problémům při plánování výroby na den dopředu potřebují zařízení na energetické zpracování odpadu (dále jen WtE, z anglického Waste-to-Energy) účinné strategie k zajištění přesných předpovědí výroby páry. Ta je klíčová složka procesů výroby elektřiny a přenosu tepla. Ke složitosti těchto předpovědí přispívají různé komplexní faktory, jako jsou zkreslená data, nepředvídatelná nižší výhřevnost odpadu (LHV z anglického Lower Heating Value) a následky nekontrolované extrakce z odběrové turbíny Pro řešení těchto problémů zavádí tento výzkum strojové učení (ML), známé svou schopností učit se z historických dat a rozeznávat zákonitosti, které jsou lidským analytikům často skryté, jako nové řešení pro posílení energetického managementu v zařízeních WtE.

Hlavním cílem této studie bylo vytvoření plánovacího algoritmu založeném na black-box modelování pomocí algoritmů ML. Tato metodika reaguje na zmíněné výzvy a snižuje riziko pokut z důvodu nesplnění smlouvy, zatímco zlepšuje celkový výkon zařízení a udržuje přijatelné riziko pokuty.

V průběhu této studie byly vybrány tři různé ML algoritmy: Lineární regrese (LR), Light Gradient Boosting Machine (LGBM) a umělé neuronové sítě (ANN). Tato různorodá sada modelů umožnila zkoumat kompromisy mezi interpretovatelností, složitostí a výkonností modelu. Každý dílčí model byl vyhodnocen pomocí běžných ukazatelů, jako je R^2 a průměrná relativní chyba, aby se určila jeho účinnost při předpovídání sledovaných veličin.

Klíčovým zjištěním analýzy reziduí bylo, že chybná předpověď celkové produkce páry představovala 80 % průměrné absolutní chyby v přeneseném výkonu. Snaha o plán s vysokou mírou úspěšnosti – definovanou jako nenadhodnocení dodané energie o více než 0,5 MWh – vedla k zavedení kvantilových modelů s parametrem q = 0,05, jejichž cílem bylo 5 % nadhodnocení vyrobené páry.

Dále byla pro zvýšení efektivity predikce sekundárních toků včetně externí spotřeby, odfuku, odvzdušňovací a spotřeby předehřívací páry použita kaskádovitá predikce založená na korelační matici. Tyto pokročilé techniky posílily schopnost modelů poskytovat přesné a spolehlivé předpovědi a prokázaly potenciál ML při řešení složitých a nejistých systémů.

K dispozici byl souboru dat z let 2011 až 2017, který byl po procesech čištění a zohlednění rozložení dat v pozdějších letech zredukován na přibližně čtyři roky dat. I při ne zcela ideální kvalitě dat se ukázalo, že ML modely jsou schopny se z dostupných informací učit. Zejména model LGBM vykazoval výjimečné výsledky při předpovídání celkové produkce páry a zachycení stochastické povahy externí spotřeby.

Pro zohlednění většího přítoku do turbíny v důsledku minimalizace využití bypassu byly navíc extrapolovány údaje o výrobě energie. To modelu poskytlo data vykazující odlišné distribucí, protože v trénovacích datech bylo používání bypassu časté. Byl proveden důkladný ověřovací proces, aby se zajistilo, že modely při extrapolaci neselžou, což je u data-driven modelů známý problém. Bylo zjištěno, že model LGBM vykazuje při extrapolaci známky selhání, což posílilo rozhodnutí upřednostnit model ANN.

Použití uvedených modelů pro údaje za rok 2016 přineslo úspěšnost 95 %. Z těchto úspěšných případů se 43 % nacházelo ve smluvním tolerančním poli, což znamená podstatné zlepšení oproti stávající strategii, při níž se do tolerančního intervalu vešlo pouze 34 % případů. Výsledkem bylo celkové zvýšení zisku o 2,6 milionu Kč, což ilustruje ekonomickou životaschopnost modelu.

Závěrem lze říct, že tento výzkum vyzdvihuje slibný potenciál ML modelů, zejména kvantilového LGBM, při zlepšování plánování výroby den předem v elektrárnách WtE. Tyto

modely mohou s vysokou přesností předvídat výrobu páry a další provozní proměnných, otevírá nové možnosti flexibilnějšího plánování a tím zvyšovat efektivitu a maximalizovat zisky.

Navzdory omezením v kvalitě dat, výsledky studie podtrhují hodnotu ML pro provozní plánování v zařízeních WtE a podobných kombinovaných závodech. Další výzkum by měl vycházet z těchto poznatků, dále optimalizovat model a zahrnout do něj další proměnné, jako jsou denní tržních ceny energií a kombinace vícero zdrojů jako jsou například obnovitelné zdroje.

BIBLIOGRAFICKÁ CITACE

KOLLMANN, Marek. *Plánování výroby tepla a elektřiny v zařízení na energetické využití odpadu s využitím strojového učení.* Brno, 2023. Dostupné také z: https://www.vut.cz/studenti/zav-prace/detail/150609. Diplomová práce. Vysoké učení technické v Brně, Fakulta strojního inženýrství, Ústav procesního inženýrství. Vedoucí práce Michal Touš.

PROHLÁŠENÍ

I hereby declare that I have independently prepared my master's thesis on the topic of "Combined Heat and Power Production Planning in a Waste-to-Energy Plant Using Machine Learning" using professional literature and resources cited in the attached bibliography, which forms an appendix to this work.

26/05/2023 *Bc. Marek Kollmann*

I would like to express my deepest gratitude to my supervisor, Michal Touš, who has been instrumental in my journey through this research. He not only introduced me to the fascinating world of machine learning but also offered constant motivation and invaluable guidance throughout. His expertise and support made this journey significantly enriching and insightful.

Moreover, I owe a profound debt of gratitude to my two sisters, who have been my pillars of emotional support. Their unending encouragement and faith in me carried me through the most challenging periods of the last five years. Without them, this journey would have been far more arduous. Their unwavering support helped transform these years into a period of growth, learning, and personal development.

Table of Contents

Chapter 1: Introduction

Combined heat and power (CHP) production has become increasingly important in the energy industry to increase efficiency and reduce emissions. CHP is the simultaneous production of electricity and usable heat from a single energy source, which can be a more efficient use of energy than producing them separately. In the context of Waste-to-Energy (WtE) plants, specifically incineration, CHP production can be an effective way to recover energy from waste and generate electricity while also providing heat for district heating systems [1].

This work aims to improve energy management in the WtE plant by developing a planning approach based on black-box modelling using machine learning (ML) to algorithms. The proposed tool addresses the inherent unpredictability of heat generation from inhomogeneous waste. This, in turn, minimizes the risk of penalties resulting from contract non-fulfilment, and to enhance the plant's performance with an acceptable risk of penalties. This study is motivated by the need to address the plant's efficiency and risk of not meeting the plan at the same time, with the goal of preparing a balanced production plan that maximizes the plant's performance from an economic point of view.

Building upon the motivation of previous work [2], we explore the potential of machine learning further enhance CHP production planning in Waste-to-Energy plants. Machine learning algorithms, with their ability to learn from data and make predictions or decisions without explicit programming, are particularly well-suited for managing complex and uncertain systems like CHP production planning in WtE plants.

One of the challenges in CHP production planning is handling the uncertain lower heating value (LHV) of waste. The LHV of waste refers to the amount of heat that can be obtained from burning a unit of waste. This value can vary depending on the composition of the waste, which can be challenging to predict. Machine learning algorithms could be employed to develop models that accurately predict the LHV of waste based on historical data and other relevant factors [3].

Another challenge involves addressing the process of live steam extraction, which is also referred to as bleeding, which occurs in some CHP configurations, a process that can impact turbine performance and overall plant efficiency. Traditional methods for predicting the influence of live steam extraction require detailed turbine models. However, machine learning algorithms can circumvent this requirement by learning from data and making accurate predictions without such models, offering a more efficient, cost-effective solution suitable for rapid prototyping [4].

In summary, machine learning holds the potential to significantly improve CHP production planning in Waste-to-Energy plants by providing accurate predictions and decision support in the face of uncertainty and complexity. By emphasizing the importance of CHP production in WtE plants and its role in improving energy efficiency, this study highlights the potential benefits of employing machine learning in CHP production planning and aims to develop a balanced production plan that maximizes plant performance from an economic perspective.

Research objectives and research questions

Accurate day-ahead production planning is critical for avoiding costly fines due to discrepancies between contracted and actual energy delivery in WtE plants. This involves predicting the amount of steam that will be produced and subsequently used for electricity generation and heat delivery. However, predicting steam production can be challenging due to noisy data, uncertain LHV of waste, and the impact of live steam extraction on turbine performance and overall plant efficiency.

Machine learning has the potential to improve these predictions by learning from historical data and identifying patterns that may not be immediately apparent to human analysts. By developing accurate machine learning models for steam production prediction and addressing these challenges, it is possible to improve day-ahead production planning and avoid costly fines.

In light of these challenges and opportunities, the main research objectives of this thesis are to:

1) Develop machine learning models to optimize energy and heat contract planning and improve efficiency in a Waste-to-Energy plant.
2) Evaluate the performance of these models in terms of their accuracy, reliability, and confidence levels.
3) Analyse the potential financial benefits that could result from implementing the new planning strategy, highlighting the economic value of integrating machine learning techniques into industrial production planning.

To achieve these objectives, we will address the following research questions:

1) What are the key factors affecting the minimization of penalties and maximization of efficiency in energy and heat contract planning in a Waste-to-Energy plant?
2) How can machine learning algorithms be used to create predictions with specific confidence levels for steam production and energy and heat?
3) What machine learning algorithms are best suited for predicting steam production, and energy/heat delivery?
4) What are the potential financial benefits and risks associated with integrating machine learning techniques into the day-ahead production planning process?
5) How can the financial benefits of the new planning strategy be optimized or maximized in practice?

Scope and Limitations

The scope of this thesis is limited to day-ahead prediction of production and delivery of heat and electricity in Waste-to-Energy plant using machine learning. While the WtE serves as an important context for this research in terms of source of disturbance in data and primary need for more sophisticated approach it is not the focus of the study, and our results should be transferable to most CHP plants.

In order to develop and evaluate our models, we will be using historical data from a specific WtE plant from years 2011 to 2017. It is important to note that quality of the data is suboptimal.

Furthermore, our goal was aligned with previous work we are expanding upon [2]– creating a planning tool with inputs defined by the operators at specific WtE plant and not profit maximization. Future work could include additional models that would tackle economy of the plant more in-depth.

Overall, while our research has the potential to significantly improve day-ahead production planning in Waste-to-Energy plants, it is essential to acknowledge these limitations and their potential impact on our results. Recognizing these limitations allows for a better understanding of the study's findings and their implications for both the Waste-to-Energy industry and future research endeavours.

Outline of the thesis:

Chapter 1: Introduction

This chapter delves into the subject of the research, establishing the motivation, objectives, and delineating the study's scope.

Chapter 2: Machine Learning and Its Applications in Energy Systems

This chapter contextualizes the study, furnishing critical background knowledge. It expounds on the fundamentals of machine learning and reviews its applications in energy systems, specifically in combined heat and power production and Waste-to-Energy plants.

Chapter 3: WtE production forecasting Case study

This chapter presents a case study on forecasting in Waste-to-Energy Combined Heat and Power production, laying out the methodology employed. It details steps such as data pre-processing and feature engineering, followed by the selection and validation of various machine learning algorithms. Subsequently, it discloses the results of the model development process and contemplates their implications.

Chapter 4: Error! Reference source not found. **Conclusion**

Concludes the thesis by summarizing the main findings and suggesting potential improvements or directions for future research. This chapter provides an overview of the contributions of this study and its potential impact on the field.

Chapter 2: Machine Learning and Its Applications in Energy Systems

The energy industry is undergoing a significant transformation, driven by the need for increased efficiency, sustainability, and reliability. This transformation is facilitated by the emergence of advanced technologies, particularly machine learning, which has shown great promise in optimizing energy systems, combined heat and power systems, and WtE plants.

This chapter provides an overview of the fundamentals of machine learning and its applications in energy systems and specifically in Waste-to-Energy plants. It begins with an introduction to the basics of machine learning, including its types, key concepts, and common algorithms. It then delves into the role of machine learning in energy systems, highlighting its potential in demand forecasting, system optimization, fault detection, and renewable energy. The chapter further explores the specific applications of machine learning in combined heat and power production planning and Waste-to-Energy plants.

By providing a comprehensive overview of machine learning and its applications in energy systems, this chapter sets the stage for the subsequent chapters, which delve into the development and evaluation of machine learning models for day-ahead production planning in a Waste-to-Energy plant.

2.1 Fundamentals of Machine Learning

Machine learning is a subfield of artificial intelligence (AI) (Figure 2.1) that enables computers to learn from data and improve their performance over time [5]. This concept can be defined by a quote from Tom Mitchell, a renowned computer scientist:

"A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E." —Tom Mitchell, 1997 [6].

In other words, the essence of machine learning lies in a computer program's ability to adapt its behaviour based on accumulated experiences to enhance its performance in a specific task.

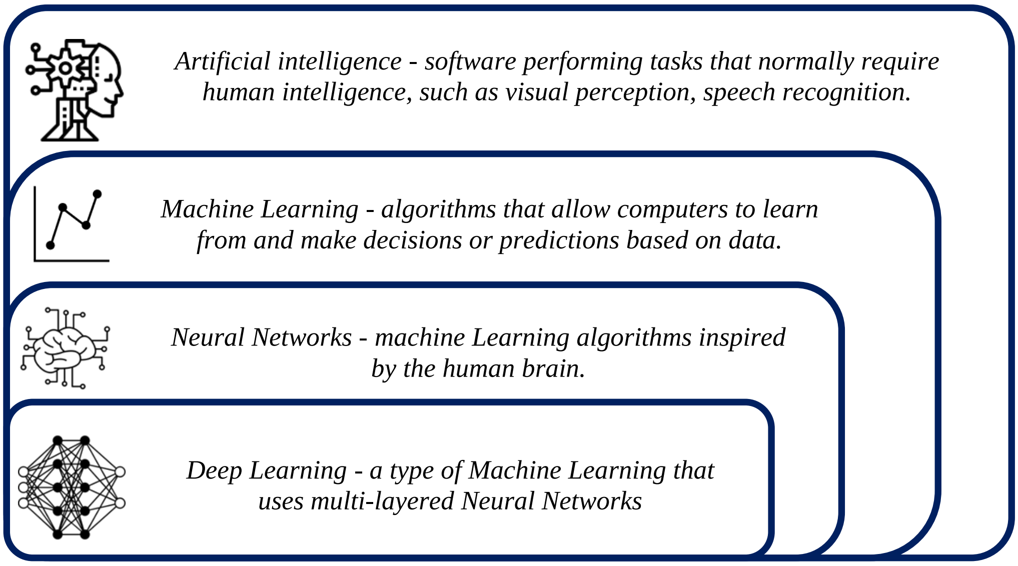

Figure 2.1: Hierarchical Subfields of Artificial Intelligence.

Machine learning algorithms can be categorized into three main types: supervised learning, unsupervised learning, and reinforcement learning. Supervised learning involves learning from labelled data, where the correct output is provided for each input. Unsupervised learning, on the other hand, does not require labelled data, and the algorithm aims to discover patterns or structures within the data as illustrated by Figure 2.2.

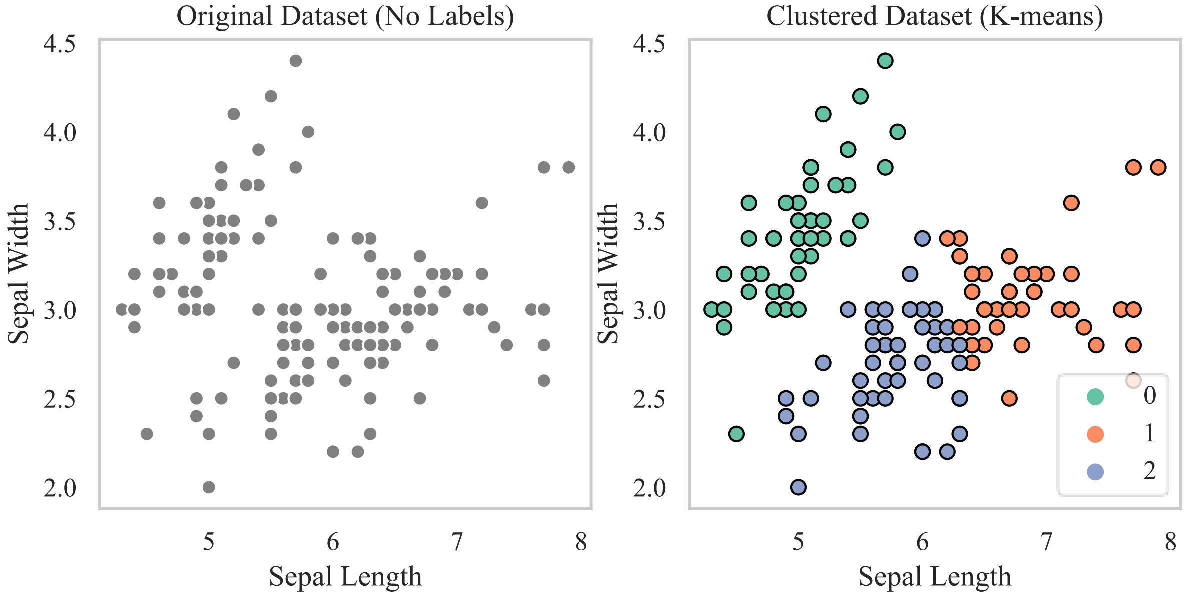

Figure 2.2: Comparison of unsupervised clustering using K-means algorithm on the Iris dataset [7, 8].

Reinforcement learning involves an agent learning to make decisions by interacting with an environment and receiving feedback in the form of rewards or penalties (penalties in the form of negative rewards, as shown in Figure 2.3) [5].

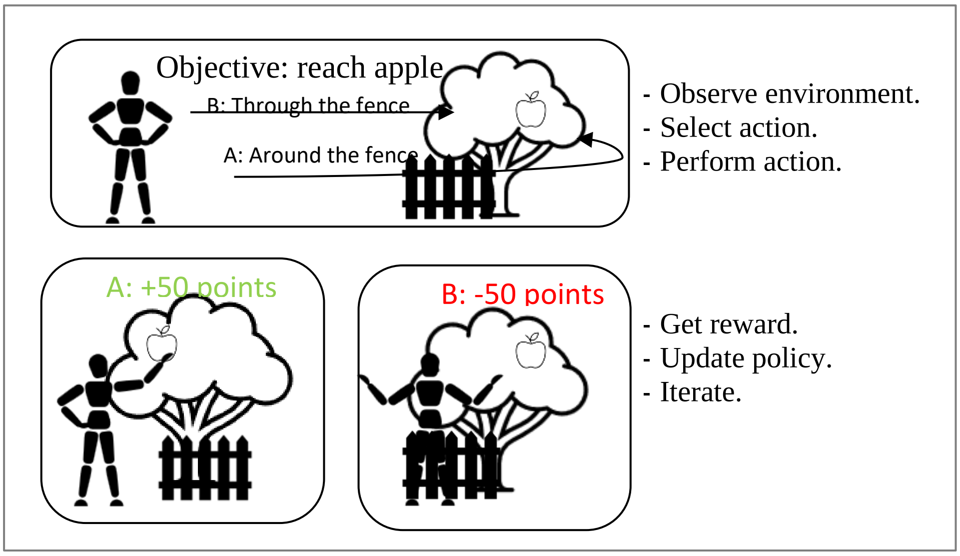

Figure 2.3: Reinforcement learning, negative feedback example [5].

1.1.1 Supervised Learning: Regression and Classification

In supervised learning, regression and classification are the two main types of problems. Regression involves predicting continuous values, such as the price of a house based on its location, size, and age or size of a tip based on total bill amount as shown in Figure 2.4.

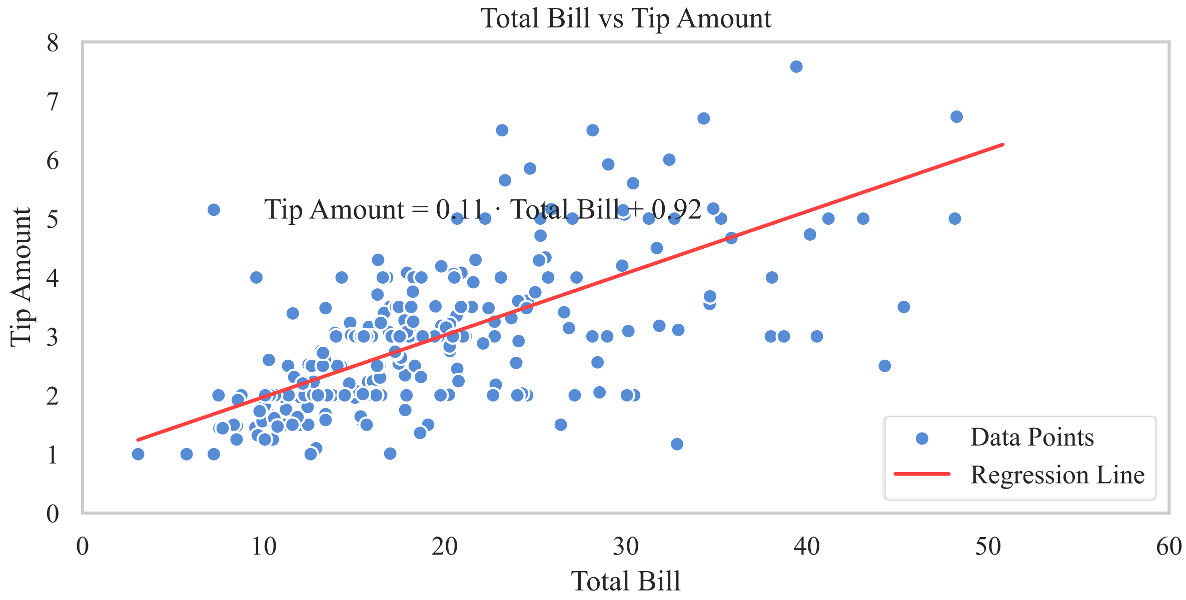

Figure 2.4: Total Bill vs Tip Amount fitted with Linear Regression[7, 9].

Classification involves predicting categorical values, such as whether an email is spam or not spam. Both regression and classification involve using labelled data to train a model that can make predictions on unseen data as illustrated by Figure 2.5.

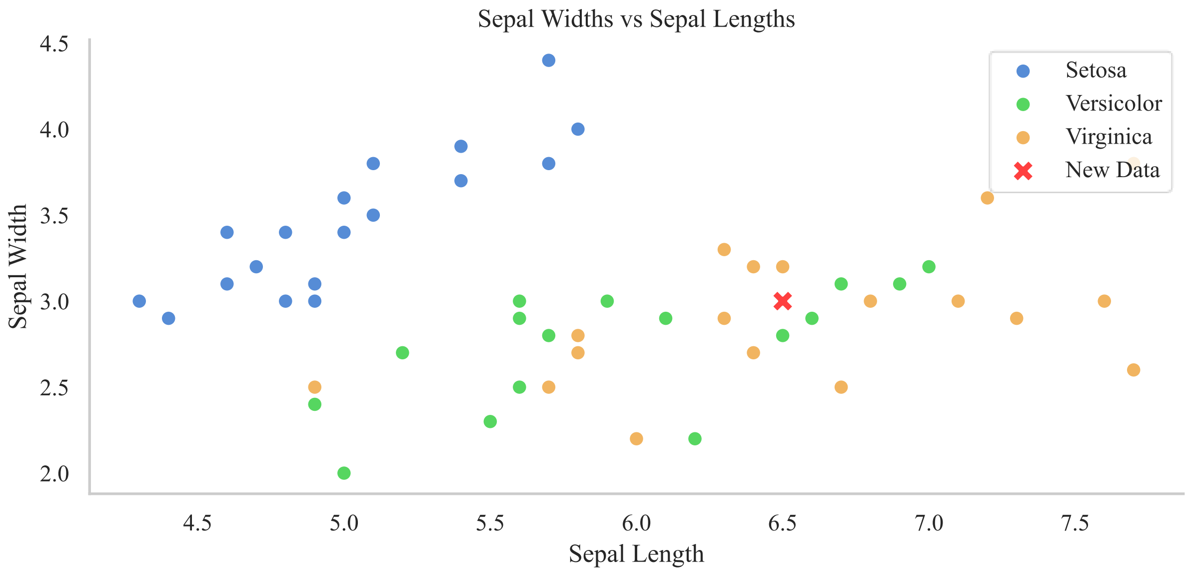

Figure 2.5: Scatter plot of sepal widths vs sepal lengths for different iris flowers [7, 8].

1.1.2 Data Splitting: Train, Test, and Validation Sets

To train and evaluate a machine learning model, the dataset is typically split into three subsets: train, test, and validation sets. The train set (commonly 60% of dataset) is used to fit the model, while the test set (commonly 20% of dataset) is used to assess its performance on unseen data. The validation set (commonly 20% of dataset), often a subset of the train set, is used to fine-tune model hyperparameters and select the best model before evaluating on the test set. This splitting process helps prevent overfitting and ensures the model generalizes well to new data. Example of this split is depicted in Figure 2.6.

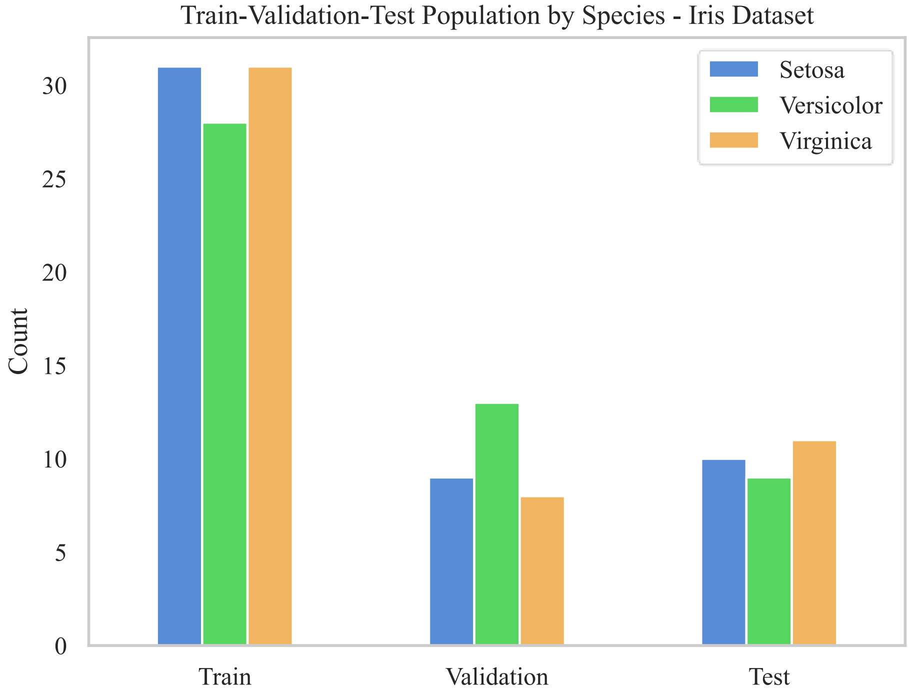

Figure 2.6: Train-Validation-Test Population by Species – Iris Dataset [7, 8].

1.1.3 Overfitting

Overfitting occurs when a machine learning model captures noise (as illustrated in Figure 2.7) in the training data, resulting in poor performance on unseen data. This is often caused by an overly complex model that fits the training data too closely. To address overfitting, it is essential to use techniques such as cross-validation, regularization, and early stopping.

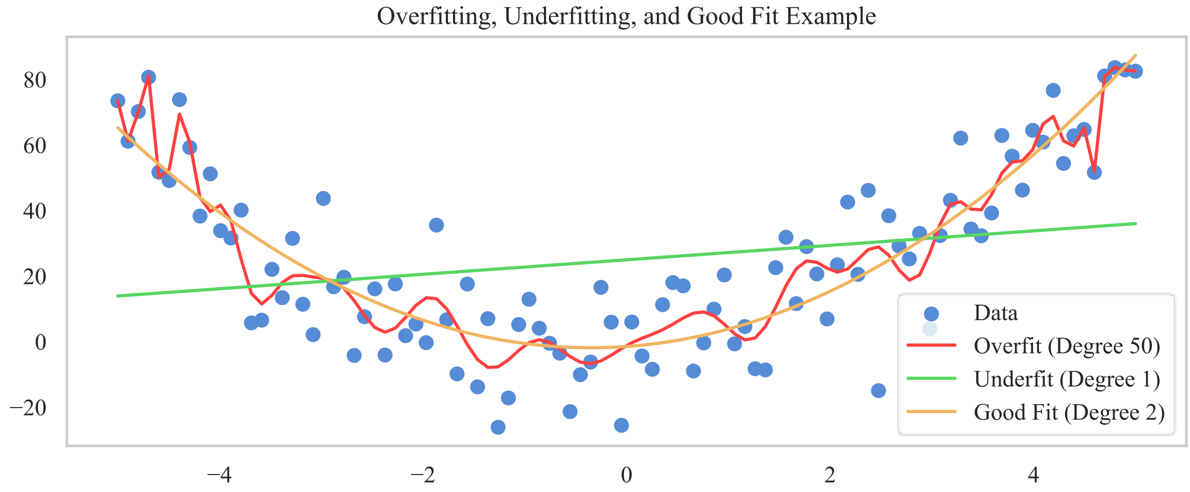

Figure 2.7: Overfitting, Underfitting, and Good Fit Example.

1.1.4 Loss Function

A loss function, or cost function, measures the difference between the predicted values and the actual values for a given dataset. The objective of a machine learning model is to minimize this loss function during the training process. Different loss functions are used for different types of problems, such as Mean Squared Error for regression tasks and cross-entropy for classification tasks.

1.1.5 Regularization

Regularization is a technique used to prevent overfitting by adding a penalty term to the model's objective function. Two common regularization methods are Lasso (L1) and Ridge (L2) regression. These methods differ in the type of penalty they apply. In Ridge Regression, the penalty is the squared magnitude of the coefficients. Mathematically, the loss function in Ridge Regression is:

$$\mathrm{L1} = \sum \left(y_i - \beta_0 - \Sigma\beta_j \cdot x_{ij}\right)^2 + \lambda \cdot \sum \beta_j^2 \tag{2.1}$$

Lasso Regression, on the other hand, uses the absolute value of the magnitude of the coefficients as the penalty. Its loss function is:

$$\mathrm{L2} = \sum \left(y_i - \beta_0 - \Sigma\beta_j \cdot x_{ij}\right)^2 + \lambda \cdot \sum |\beta_j| \tag{2.2}$$

In these equations:
y_i is the observed outcome, β_0 is the intercept of the model, β_j is the coefficient for the j^{th} predictor, x_{ij} is the value of the j^{th} predictor for the i^{th} observation, and λ is a tuning parameter that controls the strength of the penalty.

In both cases, larger values of λ lead to greater penalty and thus simpler models. These methods can greatly help to improve the model's ability to generalize to unseen data.

1.1.6 Cross-Validation

Cross-validation is a technique used to evaluate the performance of machine learning models, particularly during hyperparameter tuning. The most common method is k-fold cross-validation, where the train set is divided into k subsets or folds. The model is trained on k-1 folds and validated on the remaining fold, with this process repeated k times. The average performance across all folds is used as the model's performance metric (as depicted in Figure 2.8). This helps prevent overfitting and ensures a more reliable evaluation.

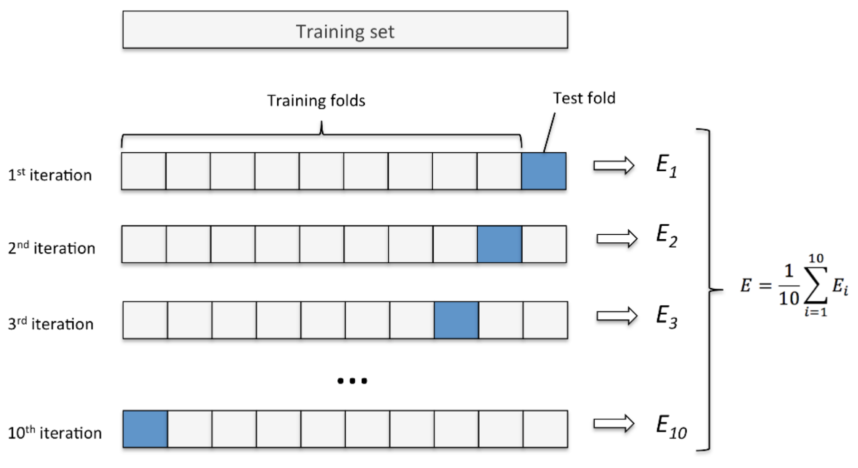

Figure 2.8: k-fold Cross-validation with k = 10 [10].

1.1.7 Hyperparameters and Model Tuning

Hyperparameters are parameters in machine learning models that control their complexity and are not learned during training. They can significantly impact a model's performance, and finding the optimal set of hyperparameters is an important part of the model development process. Techniques such as grid search and random search can be used to explore the hyperparameter space, with cross-validation used to evaluate performance.

1.1.8 Model Evaluation Metrics

Several evaluation metrics are used to assess the performance of machine learning models, such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R^2) for regression tasks, and accuracy, precision, recall, and F1 score for classification tasks. Each metric provides insights into different aspects of the model's performance, such as how well it generalizes to new data and its ability to predict true positives and negatives (see Table 2.1 and Table 2.2). These metrics provide various ways to evaluate the performance of a machine learning model, and the choice of metric depends on the specific problem and objectives of the analysis.

For regression:

a. Mean Squared Error (MSE): MSE is the average of the squared differences between the predicted and actual values. It is calculated as follows:

$$\mathrm{MSE} = \frac{1}{n}\sum\left(\mathrm{Y_i} - \hat{\mathrm{Y}}_\mathrm{i}\right)^2 \tag{2.3}$$

b. Mean Absolute Error (MAE): MAE is the average of the absolute differences between the predicted and actual values. It is calculated as follows:

$$MAE = \frac{1}{n}\sum |Yi - \hat{Y}i| \tag{2.4}$$

c. Mean Relative Error (MRE): MRE is the average of the relative differences between the predicted and actual values. It is calculated as follows:

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \widehat{Y_i}}{Y_i}\right| \tag{2.5}$$

d. Signed Mean Error (SME): SME is the average of the signed differences between the predicted and actual values taking into account the direction of the differences (positive or negative). It provides information about the overall bias in the predictions. It is calculated as follows:

$$MSME = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \widehat{Y_i}\right) \tag{2.6}$$

In these equations:
n is the total number of observations or instances, in these equations, Y_i is the actual value of an observation and $\widehat{Y}_i$ is the predicted value of an observation.

e. R-squared: It is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable(s) in a regression model. Typical values for R-squared fall within the interval of (0, 1), where a negative value indicates a particularly poor fit. It is calculated as follows:

$$R^2 = 1 - \frac{SSR}{SST} \tag{2.7}$$

where,
- SSR is the sum of the squared residuals (predicted - actual values).
- SST is the total sum of squares (actual - mean of actual values).

Table 2.1: Pros and cons of standard metrics in regression tasks.

Metric	Pros	Cons
MSE	- Penalizes large errors more due to squaring. - Provides a smooth, differentiable function, useful for gradient-based optimization.	- May be too sensitive to outliers due to squaring. - Not easily interpretable as it doesn't have the same units as the input.
MAE	- Easier to interpret as it's in the same units as the input. - Less sensitive to outliers compared to MSE.	- Provides less emphasis on large errors compared to MSE. - Not as mathematically convenient for gradient-based optimization.
MRE	- Good for comparing errors in datasets with wide value ranges. - Normalizes the absolute error by the actual values, hence providing a scale-free measure.	- Undefined or sensitive to zero values in the actual data. - Not always appropriate when actual values are close to zero.
SME	- Provides information about the overall bias in the predictions. - Can help identify if the model is consistently over or under predicting.	- Doesn't take into account the magnitude of the errors. - May mislead accuracy if positive and negative errors cancel each other out.
R^2	- Measures how much of the variance in data is explained by the model, giving a scale-free score. - Allows for comparing different regression models.	- Value can be artificially inflated by adding unnecessary variables to the model. - Not suitable for comparing models across different datasets.

For classification:

a. 4. Accuracy: This is the proportion of true results among the total number of cases examined. It is calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.8}$$

b. 5. Precision: This is the proportion of true positive predictions among the total positive predictions. It is calculated as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2.9}$$

c. 6. Recall (Sensitivity): This is the proportion of true positive predictions among the total actual positives. It is calculated as follows:

$$Recall = \frac{TP}{TP + FN} \tag{2.10}$$

d. 7. F1 Score: This is the harmonic mean of precision and recall. It is calculated as follows:

$$\text{F1 Score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{2.11}$$

In these equations:
- TP = True Positives
- FP = False Positives
- TN = True Negatives
- FN = False Negatives

Table 2.2: Pros and cons of standard metrics in classification tasks.

Metric	Pros	Cons
Accuracy	- Easy to understand and interpret. - Gives a good measure when the classes are balanced.	- Can be misleading when the classes are imbalanced. - Doesn't consider the type of error (FP, FN).
Precision	- Useful when the cost of False Positives is high. - Good for imbalanced datasets.	- Can be misleading if the cost of False Negatives is high but not considered. - Doesn't provide information about the True Negatives.
Recall (Sensitivity)	- Useful when the cost of False Negatives is high. - Good for imbalanced datasets.	- Can be misleading if the cost of False Positives is high but not considered. - Doesn't provide information about the True Negatives.
F1 Score	- Balances the trade-off between Precision and Recall.	- Might not be the best metric when you care more about precision or recall over the other. - Still not suitable when the cost of False Positives and False Negatives are very different.

1.1.9 Feature Engineering and Selection

Feature engineering involves transforming raw data into meaningful features that can be used as input for machine learning models. This may include scaling, normalization, encoding categorical variables, and creating new features based on domain knowledge. For example, extracting text length feature for sentiment analysis in tweets.

1.1.10 Quantile Regression

Quantile regression is an extension of linear regression that predicts specific quantiles of the target variable, rather than its mean. This approach provides a more comprehensive view of the target variable's distribution, allowing for better estimation of conditional quantiles and understanding of the uncertainty associated with predictions. Quantile regression can be especially useful in situations where the target variable's distribution is not symmetric or has extreme values as illustrate in Figure 2.9.

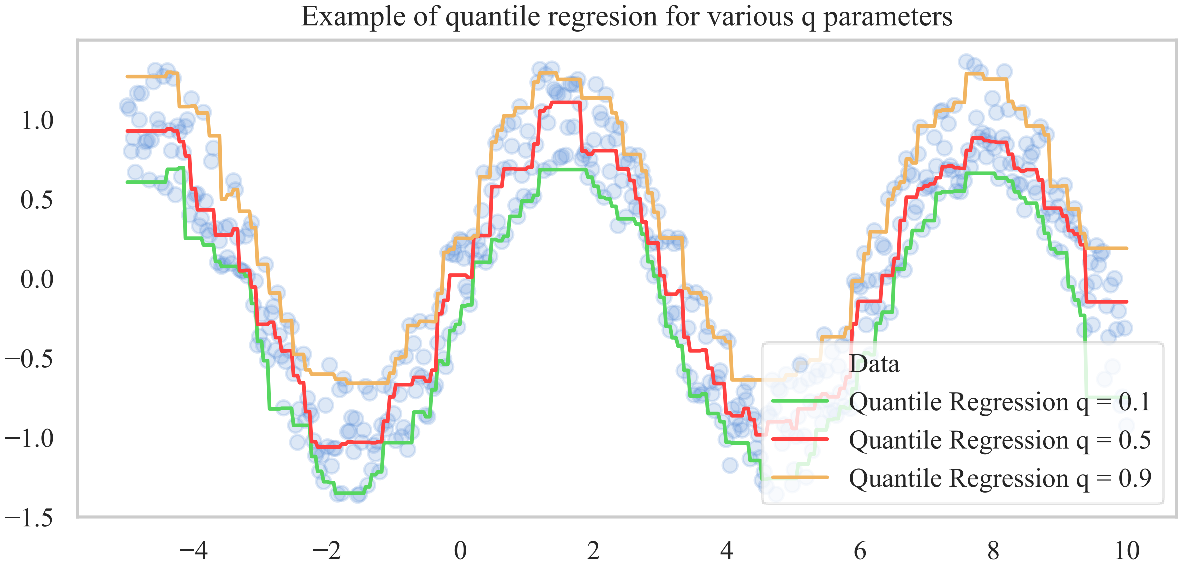

Figure 2.9: Quantile Regression Example.

In most models this can be achieved by using equation 4.3 as a loss function

$$q_{loss} = max\left(q \cdot (Y_i - Y_i), (1 - q) \cdot \left(\widehat{Y_i} - Y_i\right)\right) \tag{2.12}$$

Where:

q represents the desired quantile, indicating the specific percentile of the target variable that is being predicted. It lies between 0 and 1, where 0.5 represents the median, Y_i is the actual value of an observation and $\widehat{Y_i}$ is the predicted value of an observation.

1.1.11 Common ML Algorithms

In our exploration of machine learning algorithms, we recognize that there are a multitude of options, each with their own strengths and weaknesses (see Table 2.3).

Table 2.3: Pros and cons for classic regression algorithms with varying complexity.

Algorithm	Pros	Cons
Linear Regression	- Simple and interpretable. - Fast to train. - Good for well-defined linear relationships.	- Assumes linear relationship between features and target. - Can be outperformed by complex models on non-linear data.
Decision Trees (for Regression)	- Can model non-linear relationships. - Interpretable (if the tree is not too deep). - Doesn't require feature scaling.	- Can easily overfit or underfit. - Not as accurate as other algorithms for regression.
Random Forest	- Reduces overfitting compared to decision trees. - Can model complex, non-linear relationships.	- Less interpretable than decision trees. - Training can be computationally intensive with large datasets.
Support Vector Regression (SVR)	- Can model non-linear relationships (with suitable kernel). - Robust to outliers.	- Choice of kernel and parameters can have a big impact on performance - Can be slow to train with large datasets.
Gradient Boosting (LGBM)	- Often provides very good predictive accuracy. - Can model complex, non-linear relationships.	- Less interpretable. - Can be slow to train and requires careful tuning.
Neural Networks (for Regression)	- Can model complex, non-linear relationships. - Can handle large datasets and high dimensional inputs.	- Requires a lot of data to train. - Can be difficult to interpret. - Needs careful pre-processing and parameter tuning.

The choice of the appropriate algorithm often hinges on the specific problem at hand, and the nature of the dataset available. Python libraries were employed for their implementation [11, 12].

For the purposes of our thesis, we have chosen one representative algorithm from each of the complexity classes:

Low Complexity Algorithms:

These are relatively simpler algorithms, which can be easily interpreted and have fewer hyperparameters to tune. These are often the first choice for preliminary data analysis.

- MA & ES (Moving Average and Exponential Smoothing)
- **MLR (Multiple Linear Regression)**
- DT (Decision Trees)
- GLR (Generalized Linear Regression)

Mid Complexity Algorithms:

These algorithms are more sophisticated and often involve tuning more hyperparameters. They offer more flexibility and can model more complex relationships, but they might be harder to interpret.

- SVM (Support Vector Machines)

- RF (Random Forest)
- **GB (Gradient Boosting)**
- HM (Harmony Search)

High Complexity Algorithms:

These algorithms are typically the most complex, capable of modelling highly complex and nonlinear relationships. They often involve substantial computational resources and expertise to implement and tune effectively. ANN (Artificial Neural Networks) and NN (Neural Networks) are included in this category due to their ability to model complex, non-linear relationships.

- **ANN (Artificial Neural Networks)**
- NN (Neural Networks)
- DL (Deep Learning)
- RL (Reinforcement Learning)

Each of these algorithms was selected for its unique attributes, and its ability to address the specific demands of the supervised regression problems in our study. In choosing one from each class, we aim to provide a comprehensive and comparative exploration of these different algorithms, and their applicability to the tasks at hand.

2.2 Energy Systems overview

Energy systems encompass the processes involved in the production, distribution, and consumption of energy (see Figure 2.10**Error! Reference source not found.**). These systems are complex and multifaceted, comprising several components that work together to ensure a consistent and reliable supply of energy [13].

At the heart of any energy system are the energy sources. These can be classified into renewable and non-renewable sources. Renewable energy systems utilize resources that are naturally replenished, such as solar, wind, and hydro power. Non-renewable energy systems, in contrast, rely on finite resources such as fossil fuels and nuclear energy. Fossil fuel-based systems burn coal, oil, or natural gas to generate electricity. These systems have been the backbone of global energy supply for many years, but they produce greenhouse gases and other pollutants, contributing to climate change.

The produced energy is then transmitted and distributed through power grids. These grids are complex networks of power plants, transformers, transmission lines, and distribution lines that deliver electricity from the point of generation to consumers.

In the energy sector, different actors play crucial roles. These include energy producers, regulators, utility companies, and consumers. Energy producers generate electricity, regulators oversee the industry to ensure fair practices and safety standards, utility companies manage the distribution of electricity, and consumers use the electricity.

However, the energy industry faces numerous challenges. These include managing the unpredictable nature of renewable energy sources, the environmental impact of non-renewable energy sources, aging infrastructure, and the increasing demand for energy. Addressing these challenges requires innovative solutions and advancements in technology, including the use of machine learning and artificial intelligence [14].

Role and Potential of Machine Learning in Energy Systems

This overview of energy systems draws extensively from a highly comprehensive and up-to-date review in the field [15]. This review serves as a cornerstone for understanding the current landscape of energy systems, providing a rich and expansive overview that has been distilled and summarized in this section, Table 2.4 further summarizes the algorithms used in reviewed works.

Table 2.4: Count of machine learning models applications, sorted by complexity levels, along with their respective three most common algorithms as presented by Forootan et al. [15]

Complexity Level	Count	Most Common Algorithms
Low	25	MA & ES (8), GLR (7), MLR (5)
Mid	45	SVM (15), HM (10), RF (9)
High	55	ANN (25), NN (20), DL (10)

Machine Learning presents a significant opportunity to revolutionize energy systems, providing tools to analyse large amounts of data, make predictions, and optimize operations.

Optimalisation

One of the main applications of ML in energy systems is energy consumption optimization. By analysing consumption patterns and other influencing factors, ML algorithms can provide valuable insights for consumers and utility companies to manage and reduce energy usage effectively.

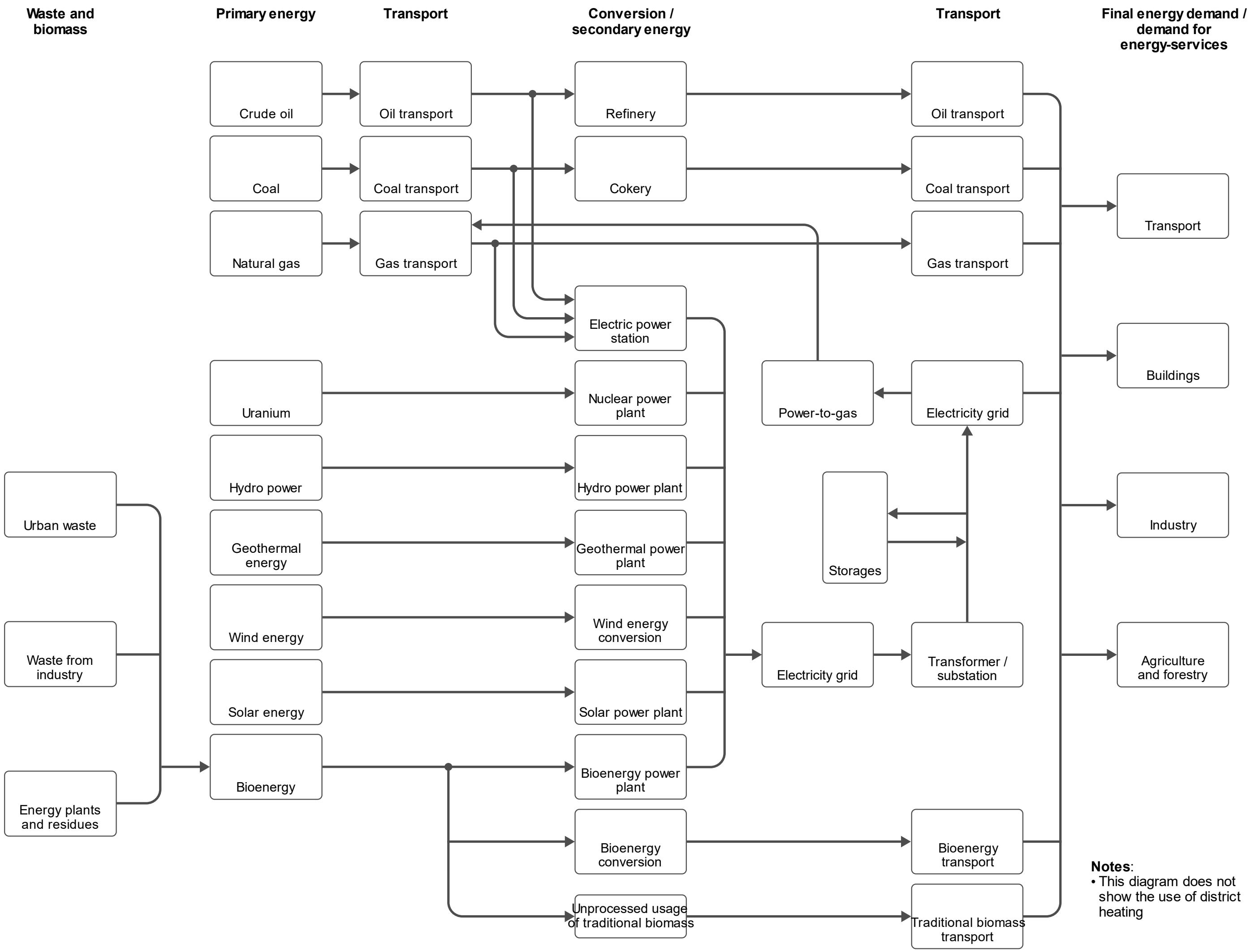

Figure 2.10: Physical components of a generic energy system supplying fuels and electricity (but not district heat) to end-users [13].

Machine learning aids in managing energy consumption efficiently, as illustrated by Ilbeigi et al.'s study, which demonstrated a 35% reduction in energy use [16]. ML also enhances renewable energy equipment performance and longevity, like in the case of Wen et al.'s wind turbine airfoil optimization [18]. ML-based optimization even indirectly improves energy efficiency in electric vehicles and fuel cells [20]. Zhou et al.'s study showcased ML's potential in improving overall energy production efficiency [22].

Overall, ML's role in optimizing energy systems spans from managing consumption to improving production efficiency and equipment performance. It showcases the versatility and potential of ML algorithms in addressing complex problems in the energy sector.

Demand forecasting

Another critical role of ML is in energy demand prediction. Accurate forecasts of energy demand are essential for efficient energy management, allowing energy providers to balance supply and demand and avoid energy wastage. ML algorithms can analyse historical consumption data and other relevant factors to predict future energy demand with high accuracy.

Several studies highlight the advantages of ML and DL in energy consumption forecasting. For instance, Amasyali et al. emphasized the need for ML-based models in commercial and educational buildings [24]. A study by Walker et al. demonstrated that ML algorithms like Random Forests are efficient in predicting electricity demand [26]. Hybrid Models (HM) have also emerged as powerful tools for predicting energy consumption. Similarly, Kazemzadeh et al. suggested a hybrid model for long-term prediction of peak electrical load and total electrical energy demand [28].

Despite some gaps in the literature, especially in areas like long-term and energy consumption forecasts, ML and DL algorithms have shown promising results in the energy sector. They not only assist in reducing energy consumption and mitigating the impacts of climate change [30] but also contribute to improving the efficiency and cost-effectiveness of energy systems [32]. Future research should continue to focus on optimizing these algorithms and exploring under-represented areas to further advance energy demand forecasting and management.

Fault and Defect Detection

Fault and Defect Detection (FDD) in industrial processes is crucial, especially with human errors causing 70% of accidents [34]. Monitoring tools are essential for the safety and reliability of equipment in energy systems[36, 38]. Advanced detection technology is required for complex systems like wind turbines [40].

AI and ML methods are increasingly used in FDD, improving speed and efficiency [42]. Various models from ensemble learning [44, 46] have been explored. Future studies should continue refining these methodologies for improved system performance.

Renewable energy

Renewable energy sources, particularly solar and wind, are integral to the future of sustainable energy systems. Predicting the output power of these systems accurately remains a critical challenge, but significant strides have been made through Machine Learning (ML) models.

Solar energy systems' output power is affected by various factors like weather conditions and cell positioning. Traditional methods for estimating solar radiation have been replaced by ML models due to their ability to manage complex relationships [48]. Voyant et al. evaluated different ML methods and suggested that methods like ANN, ARIMA, SVM, and SVR are effective for predicting solar radiation [50]. Alizamir et al. found that the Gradient Boosting Tree model outperformed others in predicting solar radiation in the U.S. and Turkey [52].

Wind energy prediction is difficult due to the inherent randomness and nonlinearity of wind behaviour. ML and Deep Learning (DL) models have been developed to predict wind energy based on wind speed and direction data. Zendehboud et al. suggested the SVM model for predicting wind power due to its speed, reliability, and accuracy [53]. Demolli et al. found that XGBoost, SVR, and RF were effective in predicting long-term wind power, with RF performing best [54].

Overall, ML models outperform traditional methods in predicting the output power of renewable energy systems. Continued research and development of these models can lead to improved prediction accuracy, ultimately enhancing the reliability of energy systems.

Specific ML algorithms relevant to these applications include decision trees, support vector machines, neural networks, and ensemble methods. Each of these algorithms has its strengths and weaknesses and is suited to different types of problems.

However, the use of ML in energy systems is not without challenges. One major issue is the quality of data. ML algorithms rely on large amounts of accurate data for training. If the data is noisy, incomplete, or biased, this can affect the accuracy of the ML models [15].

Another challenge is interpretability. While ML models can make accurate predictions, they are often seen as "black boxes" because their decision-making process can be hard to understand. This can be a problem in situations where it's necessary to understand why a particular decision was made [5, 15].

In conclusion, while there are challenges to overcome, the potential of ML in transforming energy systems is immense. With continued research and development, we can expect to see more advanced and efficient energy systems in the future.

2.3 Combined Heat and Power

Combined Heat and Power systems, also known as cogeneration systems, are a type of energy system that simultaneously generates both electricity and useful heat from the same energy source, common configurations be seen in Figure 2.11. This method of energy production is highly efficient, as it reduces the waste that typically occurs in traditional energy production systems, where heat and electricity are produced separately [17].

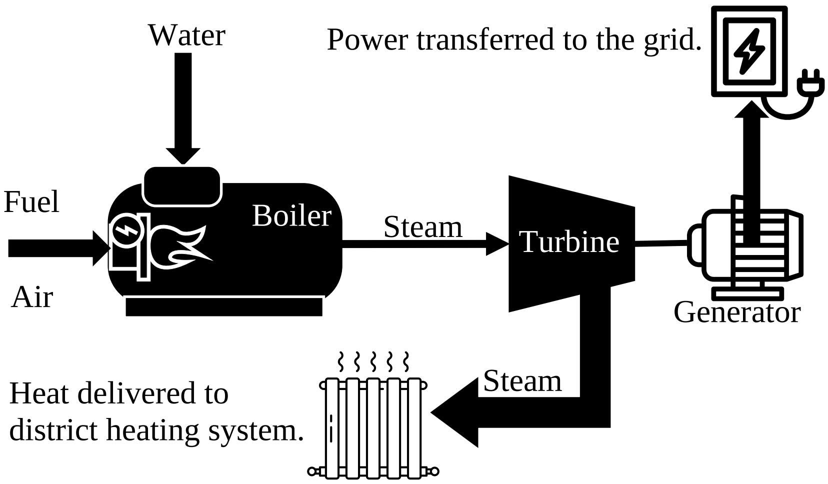

Figure 2.11: Common CHP configuration: Boiler-turbine-generator & heat exchanger.

At the core of a CHP system are three primary components: an electricity generator, a heat recovery system, and an exhaust treatment system. The electricity generator is typically driven by a turbine or an engine, which is powered by a variety of fuel sources such as natural gas, biomass, or coal. The heat recovery system captures waste heat from the electricity generator and repurposes it for useful applications, such as space heating or water heating. Lastly, the exhaust treatment system ensures that emissions from the CHP system are within acceptable limits [17].

CHP systems play a significant role in the energy landscape due to their high efficiency and their potential to reduce greenhouse gas emissions. By using waste heat productively, CHP systems can achieve energy efficiencies of up to 65-80%, compared to the approximately 50% efficiency of traditional separate heat and power (SHP) systems. Furthermore, by generating heat and power close to the point of use, CHP systems can reduce energy transmission losses and improve the reliability of energy supply [17].

Despite these advantages, CHP systems face several challenges. These include the need for a consistent and relatively high heat demand, the significant upfront capital costs, and the complexity of integrating CHP systems into existing energy infrastructure. Furthermore, the operation and management of CHP systems require careful planning and optimization to maximize their benefits.

Role of Machine Learning in CHP Production Planning

Machine Learning (ML) has the potential to substantially improve the performance and efficiency of Combined Heat and Power (CHP) systems. Through complex data analysis, pattern recognition, and predictive capabilities, ML can play a crucial role in proactive CHP production forecasting, operation, and servicing. Figure 2.12 is showing the number of papers considered in survey [19] arranged according to the ML techniques discussed in this section and the main applications.

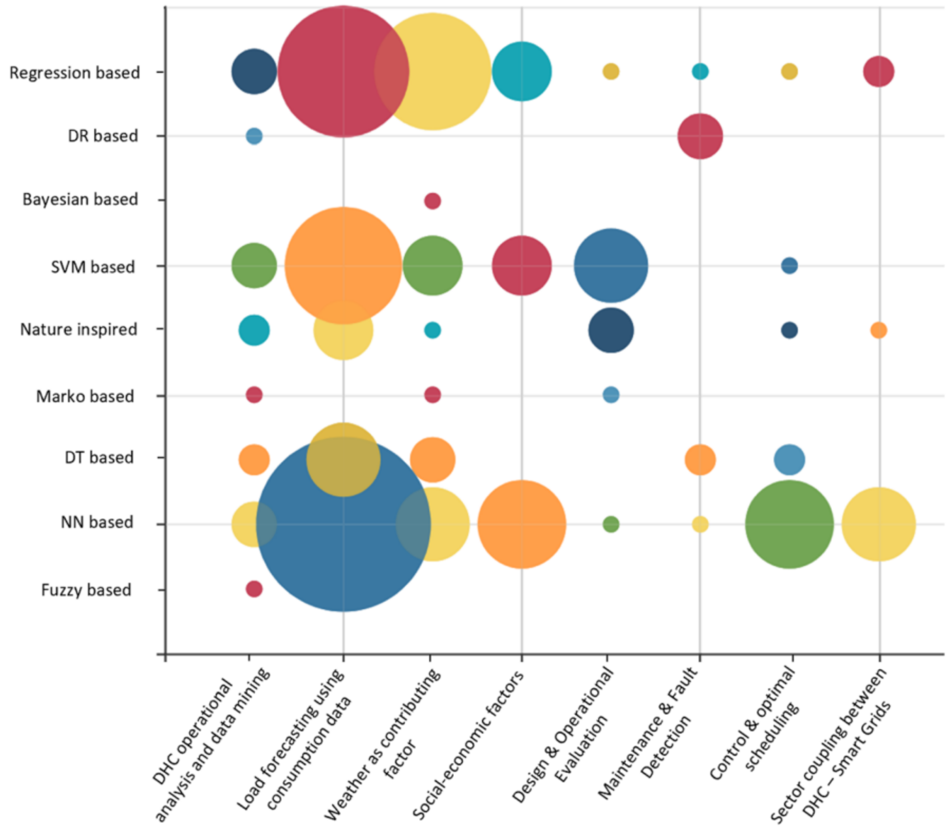

Figure 2.12: Bubble chart displaying paper count based on ML techniques and applications, allowing for multiple counts per paper [19].

A primary application of ML in CHP systems is demand forecasting. Accurate predictions of both electricity and heat demand are essential for efficient CHP operation. ML can utilize historical consumption data, weather predictions, and other relevant factors to forecast future energy needs. Consequently, CHP operators can pre-emptively adapt their production schedules based on ML forecasts with varying time frames (Figure 2.13), minimizing losses and enhancing efficiency [19, 21, 23].

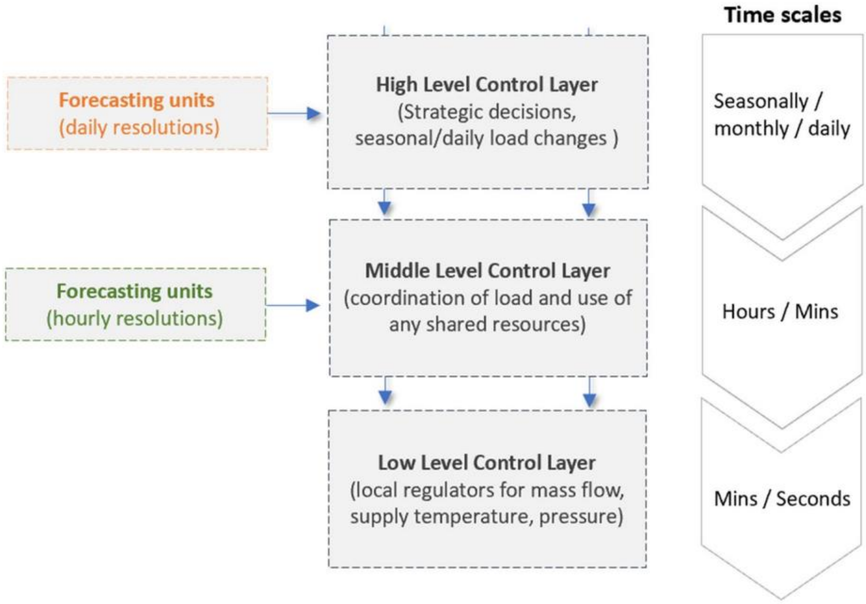

Figure 2.13: ML-empowered decision making and control hierarchy for DHC networks [19].

2.4 Waste-to-Energy Plants

Waste-to-Energy plants are facilities that convert municipal solid waste (MSW) into energy, usually in the form of electricity and/or heat (as illustrated in Figure 2.14). This conversion process not only helps to manage waste but also contributes to sustainable energy production, making WtE plants an integral part of the modern energy landscape [25].

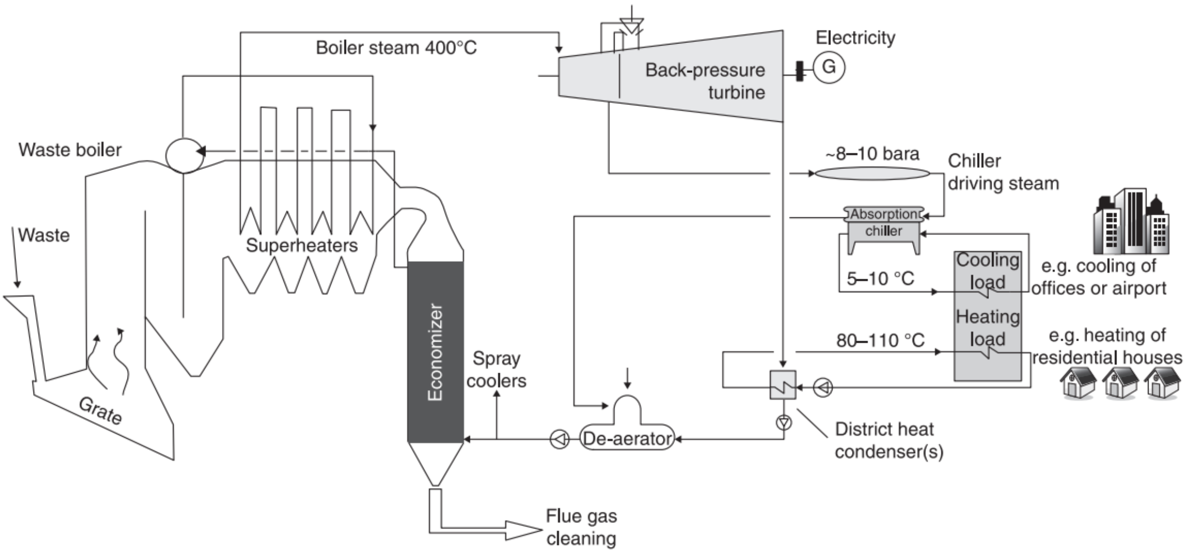

Figure 2.14: A schematic diagram of waste to energy (WtE) plant [27].

The primary process in a WtE plant involves thermal treatment of waste. The most common form of this is incineration, where waste is burned at high temperatures. The heat generated from this process is then used to produce steam, which drives a turbine to generate electricity. Other methods of Waste-to-Energy conversion include gasification, in which waste is converted into a gas that can be burned for power or heat, and anaerobic digestion, which uses bacteria to decompose organic waste and generate biogas [23, 29].

WtE plants offer numerous benefits. They provide a practical solution for waste management, particularly in urban areas where landfill space is limited. They generate energy from a resource—waste—that would otherwise be discarded, thus contributing to resource efficiency. Furthermore, by reducing the volume of waste that ends up in landfills, WtE plants can help to reduce greenhouse gas emissions [29].

However, operating WtE plants is complex and challenging. They need to handle a wide variety of waste types, each with different energy content and combustion characteristics. Fluctuations in waste input can affect the efficiency and stability of energy production. Emissions from WtE plants, including pollutants such as dioxins and heavy metals, need to be carefully managed to minimize environmental impact. These challenges necessitate advanced technologies and strategies for efficient operation and management of WtE plants.

Potential of Machine Learning in Waste-to-Energy Plants

Machine Learning presents a significant opportunity to enhance the operation of WtE plants. Through its data-driven algorithms, ML can analyse complex datasets, identify patterns, and make accurate predictions, all of which can be applied to various aspects of WtE operations [23, 31].

One of the key applications of ML in WtE plants is in predicting the calorific value of waste. The calorific value, which indicates the amount of energy that can be extracted from waste, varies widely depending on the type and composition of the waste. ML algorithms can analyse waste composition data to predict the calorific value accurately, enabling WtE operators to adjust their operations accordingly [33, 35].

Furthermore, ML can be applied to optimize the combustion process in WtE plants. ML algorithms can analyse operational data to identify the optimal operating conditions for efficient combustion, such as the optimal air-to-fuel ratio or the ideal temperature profile. This can help to maximize energy production, reduce fuel consumption, and minimize emissions [23, 37].

In the realm of fault detection and diagnosis in Waste-to-Energy plants, Machine Learning presents untapped potential. Through the analysis of operational data, ML can spot anomalies, signalling potential equipment faults or failures. Such early fault detection can mitigate expensive downtime and prolong equipment lifespan. However, this field remains relatively under-researched, with existing models often being complex and demanding [39, 41]. Thus, it represents a substantial opportunity for research and advancement.

Despite these promising applications, the use of ML in WtE plants also poses challenges. These include the need for high-quality data, the complexity of ML algorithms, and the potential for overfitting or underfitting. Nevertheless, with continued research and development, ML holds great promise for improving.

Chapter 3: Case study: WtE production forecasting

In this study, the application of machine learning in WtE plant will be showcased through the planning of heat and electricity production. To illustrate this Figure 3.1 presents the configuration of the WtE facility consisting of 4 boilers and a turbine with unregulated extraction. Furthermore, the relevant data for this case study can be found in Table 3.1.

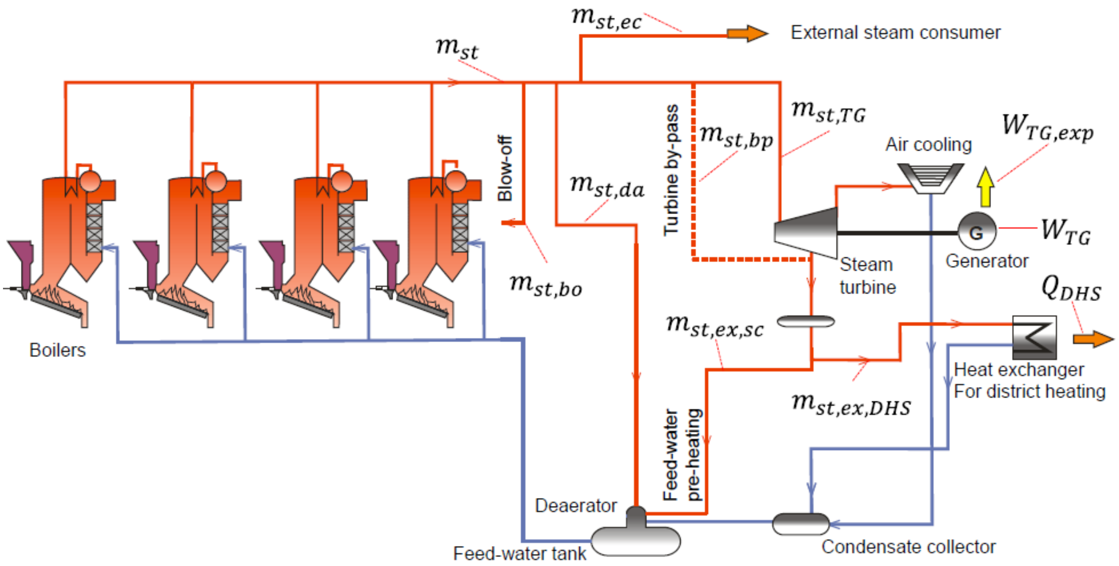

Figure 3.1: A simplified flowsheet of the steam condensate cycle used in WtE technology (red lines represent steam, blue lines represent water, flue gas treatment system excluded) [2] .

Table 3.1: Given parameters by operator, their notation, description, and unit.

Variable notation	Description	Unit
$\dot{m}_{st,boiler1}$	Boiler 1 output	[t/h]
$\dot{m}_{st,boiler2}$	Boiler 2 output	[t/h]
$\dot{m}_{st,boiler3}$	Boiler 3 output	[t/h]
$\dot{m}_{st,boiler4}$	Boiler 4 output	[t/h]
$\dot{m}_{st,turb.inflow}$	Flow rate of steam to the steam turbine	[t/h]
$\dot{m}_{st,turb.inflow,calc}$	Flow rate of steam to the steam turbine - balance equations	[t/h]
$p_{st,ex}$	Pressure of extraction steam	[MPa]
$\dot{m}_{st,ex}$	Flow rate of extraction steam	[t/h]
$\dot{m}_{st,external}$	Flow rate of steam to externa	[t/h]
$\dot{m}_{st,demi}$	Feed water tank pre-heating	[t/h]
$\dot{Q}_{DHS}$	Heat delivery to a district heating system	[MWh]
$\dot{m}_{st,blow-off}$	Flow rate of steam for boilers' blow of	[t/h]
$\dot{m}_{st,deair}$	Flow rate of steam to deaerator	[t/h]
$W_{generated}^{\overline{bp}>0}$	Generated power	[MW]
$W_{transferred}^{\overline{bp}>0}$	Transmitted power	[MW]
$\mathrm{bp}_{\mathrm{valve}}$	Bypass valve opening	[%]

CHP production is governed by contracts specifying heat and electricity delivery, prices, penalties, and more. Figure 3.2 and Figure 3.3 illustrate the influence of different months of the year on these contract values.

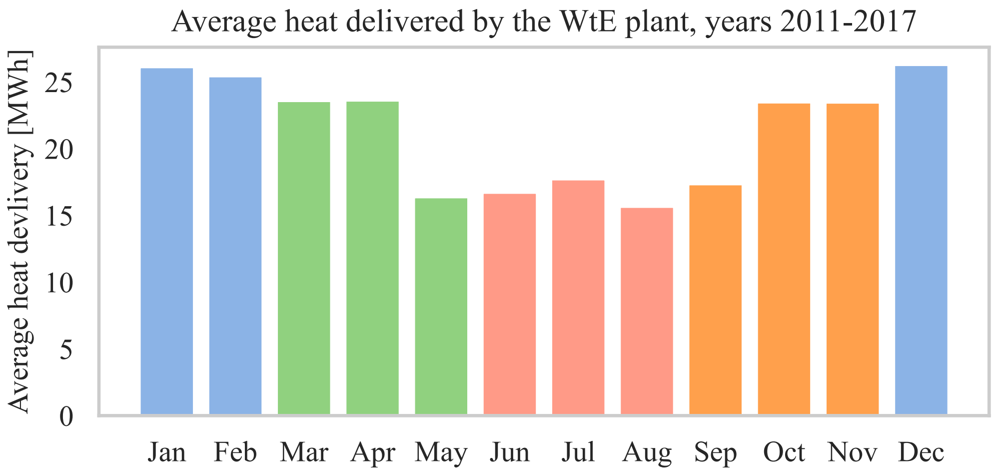

Figure 3.2: Average heat delivery of the WtE plant throughout the years 2011 to 2017.

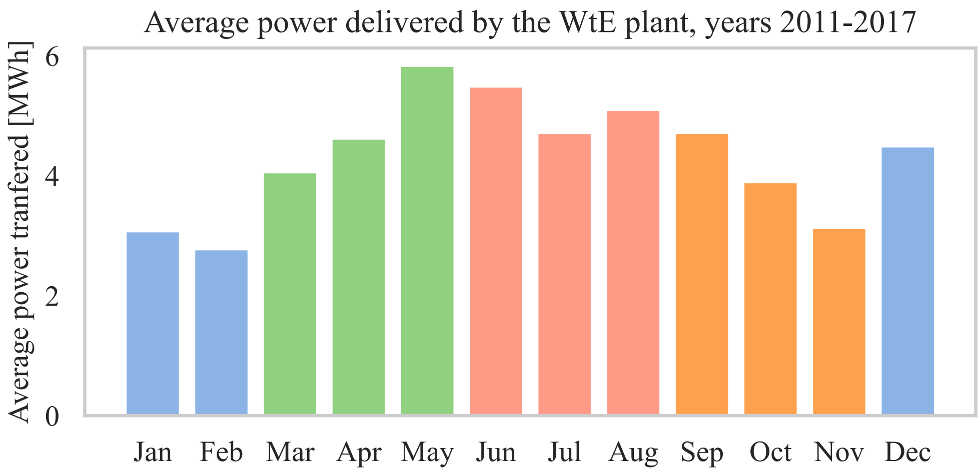

Figure 3.3 :Average electricity delivery of the WtE plant throughout the years 2011 to 2017.

These contracts conditions, along with the plant's actual performance, dictate the planning of heat and electricity delivery on an hourly basis for the next day. The goal is to create a balanced production plan that maximizes the plant's economic performance while considering efficiency and the risk of not meeting the plan. Waste-to-Energy heat and electricity planning is a challenging task due to the inhomogeneous nature of waste, fluctuating external steam demand, and operational uncertainties such as the lower heating value of waste and live steam extraction.

Contracts impose a high penalty for electricity delivery deviations beyond ±0.5 MWh. In contrast, there is a low penalty for short-term heat delivery deviations. As a result, the focus in planning and subsequent operation is on electricity delivery, and the plant opts for a conservative approach to ensure it can meet the delivery requirements. If steam production is higher, the steam turbine by-pass is used to decrease it, resulting in higher heat delivery but with lower penalties compared to electricity deviations. However, this conservative approach is not beneficial for maximizing CHP production and financial revenue. The use of the steam turbine by-pass along with the situations that may occur in relation to a proposed production plan as illustrated in Figure 3.4, underscore the challenges faced in the current planning strategy.

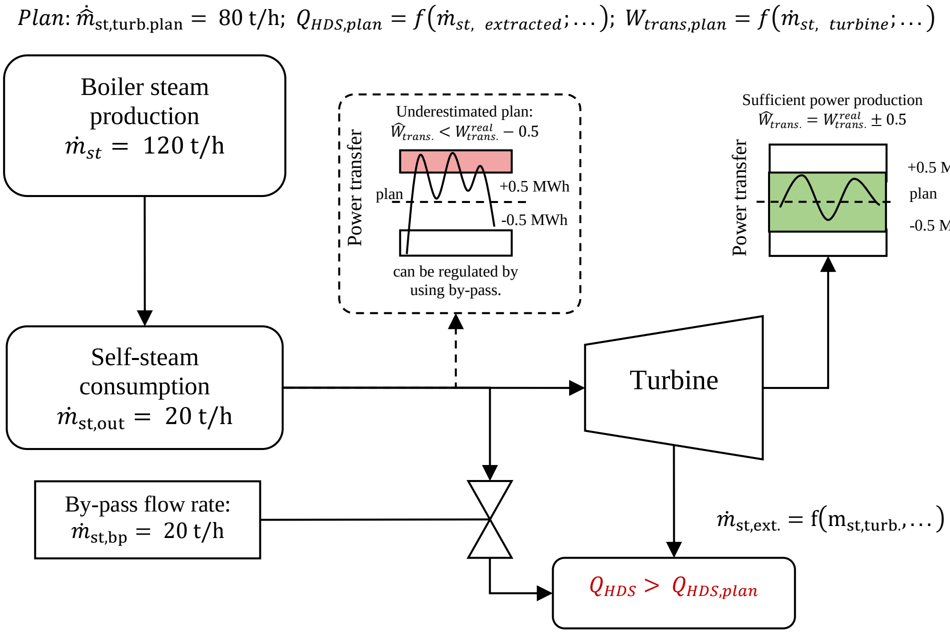

Figure 3.4: Scheme depicting current planning strategy's approach when faced with underestimated plan.

Rationale for Using Machine Learning for CHP Production Planning

To address the challenges faced in CHP production planning, the study employs data-driven models, including linear regression and artificial neural networks models. Machine learning techniques like ANN can successfully identify nonlinear relationships between variables and are generally suitable for regression-type problems. Additionally, linear regression models offer a lower level of complexity, which can be advantageous in further applications.

The adoption of machine learning in CHP production planning provides several benefits:

Improved accuracy: Machine learning models can better estimate the plant's actual performance, leading to more balanced production plans and reduced penalties.

Enhanced efficiency: Accurate planning reduces the need for energy-wasting actions during operation, such as turbine bypassing or heat releasing into the environment.

Increased revenue: Better net thermal efficiency and CHP production result in higher financial returns for the plant.

This case study demonstrates the potential of machine learning techniques in enhancing the efficiency, accuracy, and economic viability of CHP production planning in WtE plants. Building upon the work of Touš and others, our study aims to further explore the potential of machine learning in this context. Most notably, Teng [44] proposed a Waste-to-Energy management tool that offers forecasting and real-time optimization of power generation, considering anomalies. Their framework, based on Hierarchical Temporal Memory (HTM), a type of biological neural network, demonstrated promising results in an industrial case study.

However, their approach, while innovative, is complex and focuses on real-time optimization and doesn't address the bypass usage minimalization. In contrast, our study extends Touš et al.'s work by incorporating more accessible and interpretable machine learning models, such as ANN and LR, to predict steam and power production while maintaining the original objectives. This focus on simplicity, combined with a risk management approach, distinguishes our study from previous research and contributes novel insights to the field.

3.1 Methodology

This chapter describes the methodology used in this study to develop and evaluate machine learning models for predicting various aspects of steam and energy production in Waste-to-Energy plants. The methodology includes data pre-processing, feature engineering, and the selection and evaluation of machine learning algorithms within each of the three distinct model group as illustrated in Figure 3.5.

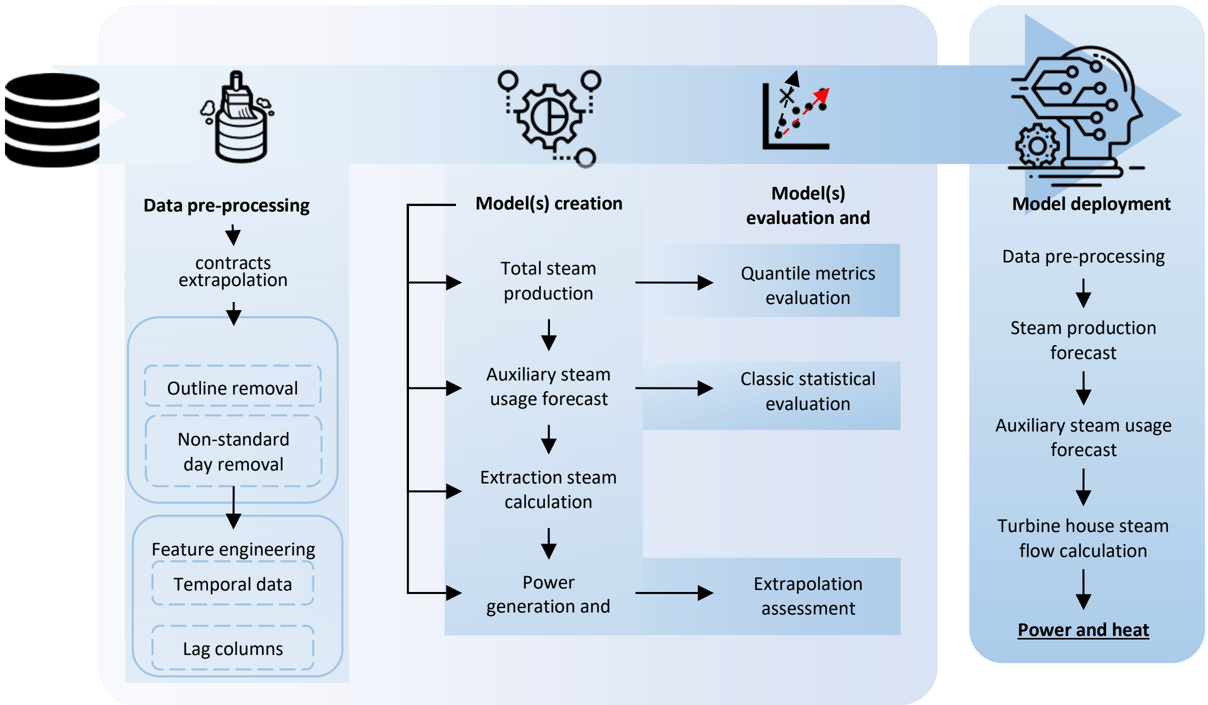

Figure 3.5: Structure of Deployed Methodology - Workflow for Building a Robust ML Model.

3.1.1 Data Pre-processing

The data used in this study were provided by the plant operators of a Waste-to-Energy facility. The dataset covers a time period from 2011 to 2017. Out of the extensive dataset, only variables that directly or indirectly affect energy and heat/power production were chosen, while variables related to flue gas processing operations were omitted. A comprehensive scheme and list of variables provided are illustrated in Figure 3.1 and Table 3.1, respectively.

Planned Heat and Power Delivery Data extrapolation

The extraction steam flow rate calculation is not straightforward due to the uncontrolled extraction. The approach employed in previous work by Touš et al. utilized the heat plan contracts value $\dot{Q}_{DHS,plan}$ as a variable in the calculation of the extraction steam flow rate (refer to chapter 3.1.3) Regrettably, the contracts provided to us only cover the period from 2012 to a portion of 2013. It is worth noting that the actual heat delivered data is available to us; however, these values are frequently distorted due to the plant's inclination to deviate from the heat plan in order to adhere to its power plan. The utilization of the steam bypass prior to the turbine, as indicated in the preceding chapter, serves as an indicator of this practice.

Planned heat extrapolation

Data extrapolation method was utilized, aiming to extend the provided data across the entire dataset. Essentially, an average annual plan was constructed using approximately a year and a half's worth of data, which was then scaled down by a factor of 0.75. This reduction was implemented to ensure the derived plan mirrored the adjusted heat delivered, which served as a reference point as per equation (3.1).

$$\dot{Q}_{DHS,adjusted} = \dot{Q}_{DHS} - \frac{\dot{m}_{st,bypass} \cdot h_{st}}{3.6} \tag{3.1}$$

In this equation:
$\dot{Q}_{DHS,adjusted}$ refers to the adjusted heat delivered, $\dot{Q}_{DHS}$ represents the original heat delivery plan, $\dot{m}_{st,bypass}$ is the bypass steam mass flow, and h_{st} is the enthalpy of steam.

$\dot{m}_{st,bypass}$ was calculated assuming a linear valve characteristic.

$$\dot{m}_{st,bypass} = \sum_{i=1}^{4} \dot{m}_{st,boiler,i} - \left(\dot{m}_{st,external} + \dot{m}_{st,blow-off} + \dot{m}_{st,deair}\right) \cdot (1 - bp_{valve}). \tag{3.2}$$

where:
$\sum \dot{m}_{st,boiler,i}$ is the total steam production, $\dot{m}_{st,external}$ is external steam consumption, $\dot{m}_{st,blow-off}$ is steam used for blow-off, $\dot{m}_{st,deair}$ is steam used in deaerator and bp_{valve} is the ratio representing the degree to which the bypass valve is opened, ranging from 0 (completely closed) to 1 (completely open).

By employing this method, estimations of the heat delivery plan were generated, demonstrating a distribution comparable to that of the original data (Figure 3.7).

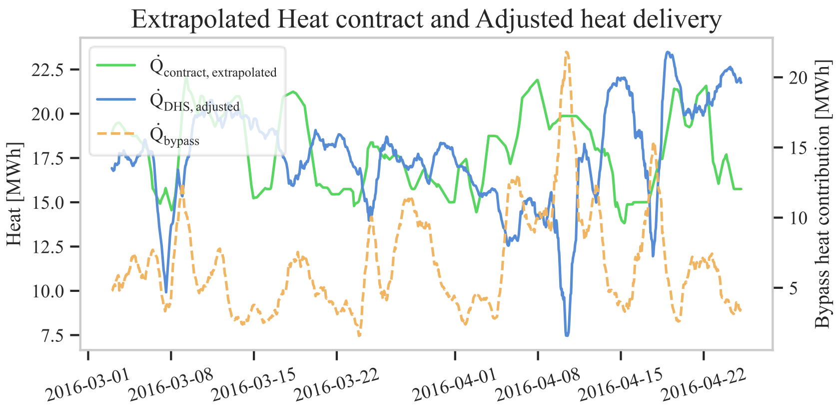

Figure 3.6: Comparison of Adjusted Heat Delivery, Extrapolated Heat Plan, and Bypass Heat.

Power extrapolation

In line with the methodology adopted by Touš et al., the electricity delivery corresponding to a zero bypassed steam flow rate, denoted as $W_{generated}^{\overline{bp}=0}$ for power generation and $W_{transferred}^{\overline{bp}=0}$ for power transferred, was utilized. It is important to note that the bypass is predominantly employed during colder months (Figure 3.7), indicating that the occurrence of $W_{transferred}^{\overline{bp}=0}$ is infrequent in the operational data throughout the year.

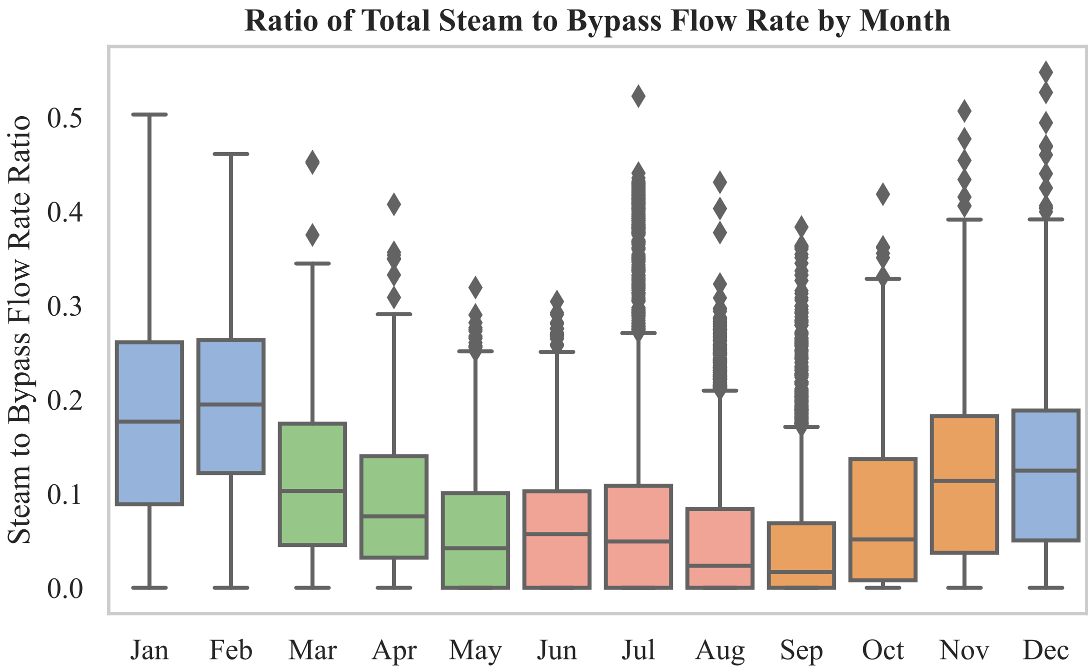

Figure 3.7:Average monthly usage of bypass throughout the years 2011-2017.

To compute the $W_{generated}^{\overline{bp}=0}$, a turbine model must first be established based on the available data (impacted by bypass usage). Following this, the turbine inflow and live extraction need to be recalculated, and these recalculated values should be applied as inputs to the aforementioned model. A similar process is followed for $\mathrm{W}_{\mathrm{transferred}}^{\overline{\mathrm{bp}}=0}$ calculation, as it is primarily based on $W_{generated}^{\overline{bp}>0}$. There is no need for further recalculation, as only $W_{generated}^{\overline{bp}=0}$ and temporal data are needed. The procedure for data extrapolation is represented in Figure 3.8.

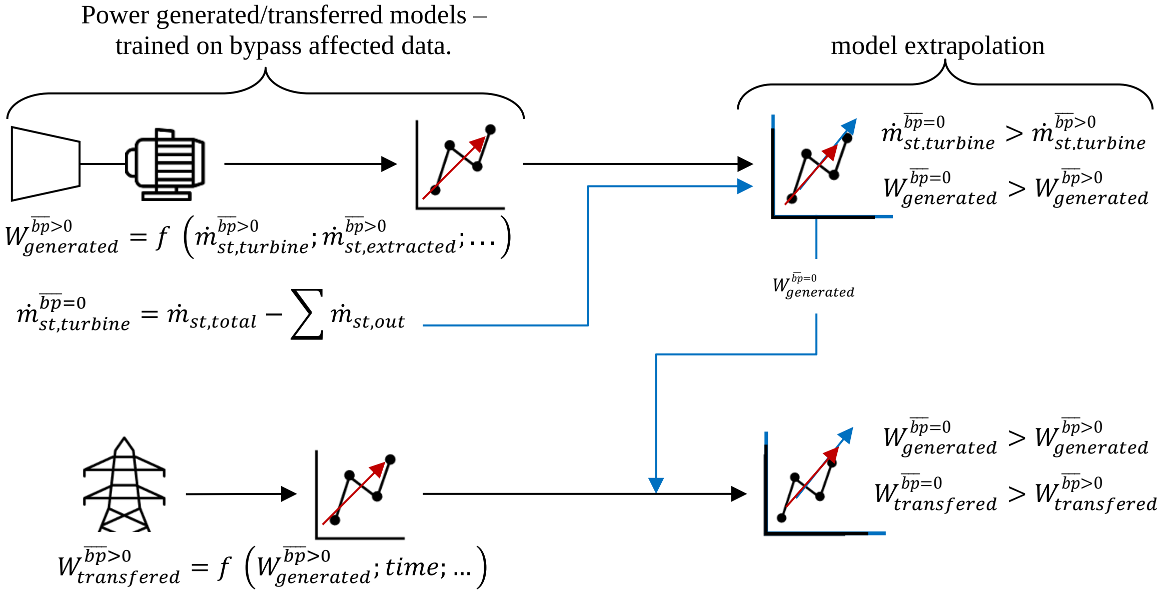

Figure 3.8. Process of extrapolating power generated and transferred from provided data.

Handling of outliners and missing values

The dataset contained several missing values and inconsistencies, such as months without some features recorded and cases with improper material balances. To handle missing values, rows with missing key values (eg: power generated, boiler production) were dropped since the missing data were often due to external causes or issues in the plant. Additionally, the dataset contained enough data points to allow for the removal of rows with missing values without significantly impacting the analysis. Methods used were as follows:

a) Z-score outlier removal: The Z-score (given by equation 3.3) measures the number of standard deviations a data point is from the mean of the dataset. Data points with a Z-score greater than a specified threshold (e.g., 2 or 3, specific for each variable) were considered outliers and were removed from the dataset.

$$Z = \frac{(x - \mu)}{\sigma} \tag{3.3}$$

Where:
x is the value of a data point, μ is the mean of the dataset and σ is the standard deviation of the dataset

b) Non-standard operation days were filtered out in the dataset, encompassing periods of repairs, maintenance, or plant shutdowns. These data points, which may not accurately represent the Waste-to-Energy plant's typical operation, have the potential to negatively impact the performance of the machine learning models. To mitigate this concern, data from non-standard operation days were excluded through the utilization of information provided by identifying anomalous patterns in the data, such as no steam or power production.

Furthermore, it should be noted that there are instances where the measured steam entering the turbine does not match the calculated steam using balance equations. This discrepancy has been acknowledged by the data provider. Using Kernel Density Estimation (KDE) plots, that smooth out the data distribution by estimating the density at various points [7]. Instances where this discrepancy was severe (as illustrated in the Figure 3.9) were removed.

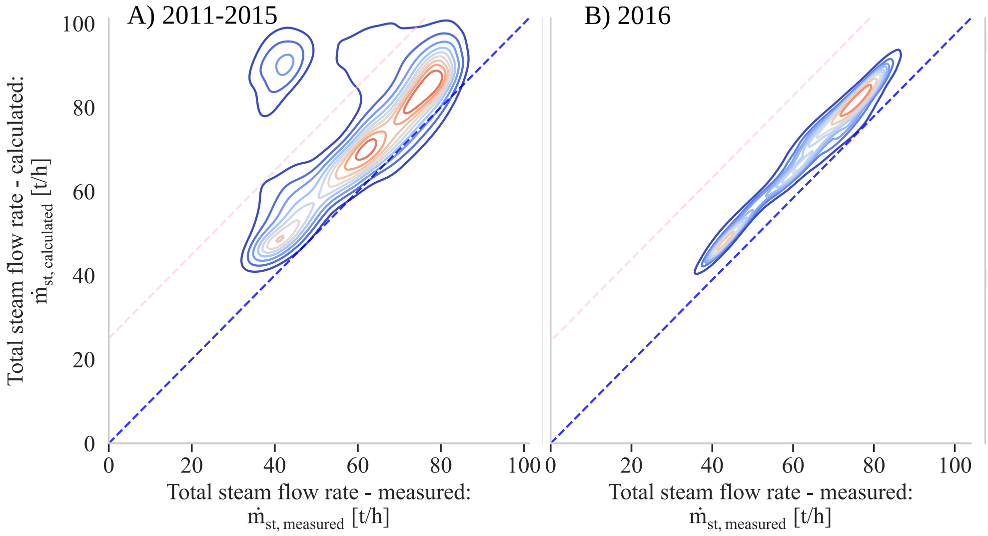

Figure 3.9: Showcase of misalignment between calculated and measured flowrates. A) for the years 2011 2015 representing out training dataset, B) for the year 2016 representing our test dataset.

Normalization:

Categorical variables, such as seasons, were encoded using one-hot encoding. For example, if the season was summer, the summer column would have a value of 1 while the other season columns would have a value of 0.

Data normalization was performed using Python's *sklear StandardScaler*. The method scales the data to have a mean of 0 and a standard deviation of 1. The StandardScaler calculates the mean and standard deviation from the training data and applies the normalization formula (equation (3.3) to both the training and testing datasets. It is vital to apply the same scaling parameters (mean and standard deviation) to both training and testing data to maintain consistency and prevent bias in the model evaluation process.

Feature Engineering

New features were derived from the existing data, including, temporal data: seasons (winter, summer, fall, spring), day of the week, month and lag columns (24h to 48h) for all important steam flows. This practice is a standard approach in time series regression [43]. The addition of these lag columns is further justified by pronounced seasonality evident in our dataset, as demonstrated in Figure 3.10.

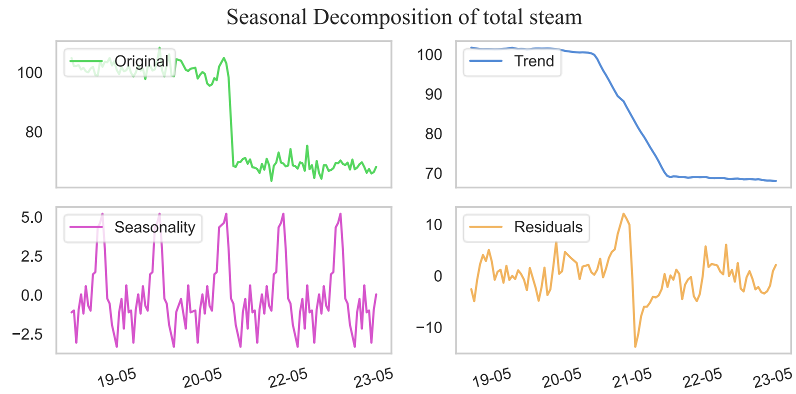

Figure 3.10: Seasonal Decomposition Using the statsmodels Python Library's [45].

The total boiler production, represented as $\dot{m}_{st,total}$ was chosen for forecasting as an aggregate due to the absence of distinct characteristics observed in individual boilers. This observation is supported by the data presented in Table 3.2 and Figure 3.11.

Table 3.2: Boiler 1 to 4 output characteristics.

Boiler name	usage	mean [t/h]	max [t/h]	Std [t/h]
Boiler 1	67%	35.46	48.55	17.06
Boiler 2	72%	34.45	47.38	16.19
Boiler 3	71%	35.04	44.93	16.54
Boiler 4	87%	34.26	46.85	15.62

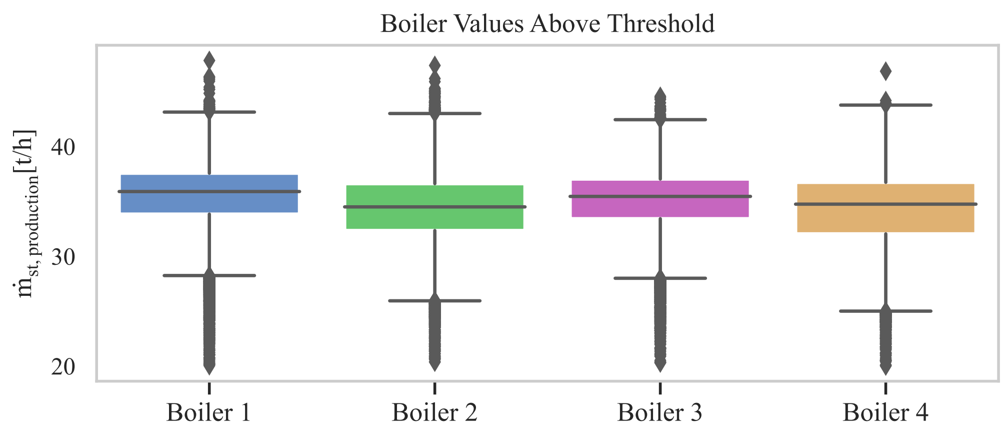

Figure 3.11: Boxplot illustrating the distribution of production values for Boilers 1-4.

It is worth noting that the occurrence of four boilers running simultaneously is an infrequent event (see Figure 3.12), Consequently, it is expected that these periods, being data-driven in nature, will experience diminished performance in our models.

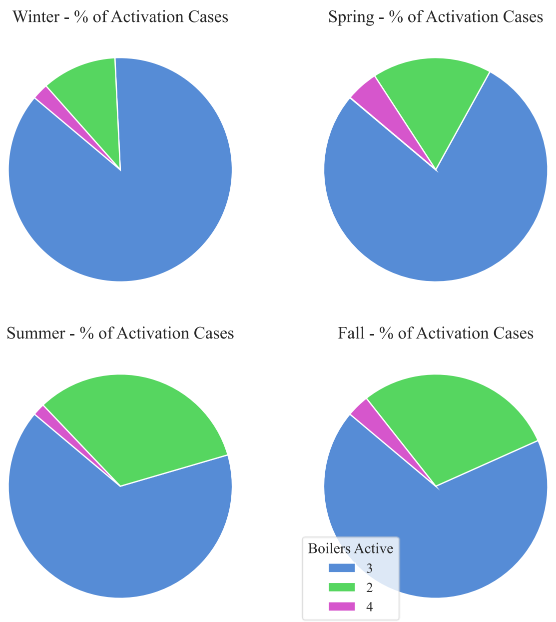

Figure 3.12: Pie charts representing the percentage of boiler activation for each season.

No feature transformations, such as log transformation, were applied as they did not produce any improvements when tried. Dimensionality reduction techniques were not applied due to the limited number of columns in the dataset.

3.1.2 Model Grouping Strategy

During the initial stages of the development process, the need for three distinct model groups to tackle the specific challenges and requirements of the prediction process was identified.

The first group of models is centred around the prediction of steam production from all boilers, a critical variable that significantly impacts other variables within the system. It was revealed that approximately 80%[1] of the mean absolute error in power transfer predictions resulted from inaccuracies in the total steam prediction. To achieve a high and well-defined level of confidence in the predictions, quantile regression models were employed.

This approach to $\dot{m}_{st,total}$ prediction differs from that of Touš et al. Where the boiler output was modelled using random walk algorithm and confidence interval was determined using Monte Carlo simulation [2].

The second group of models deals with auxiliary factors that use total steam produced but their individual magnitude is a small fraction of $\dot{m}_{st,total}$. These factors include $\dot{m}_{st,external}, \dot{m}_{st,blow-off}, \dot{m}_{st,deair}, \dot{m}_{st,demi}$. To increase the amount of potentially useful factors for their prediction, a unique approach was deployed as depicted.

Unlike the first group of models, which operated at a predetermined confidence level, the prediction of these auxiliary variables was carried out using a cascading framework. The

[1] This figure was obtained by substituting the real total steam for the predicted one throughout the prediction pipeline.

utilization of predetermined confidence levels has been avoided in order to address the issue of accumulating underestimations. Instead, the framework employed a cascading prediction system, where variables were arranged based on their correlations and interrelationships, thereby mitigating potential underestimations and enhancing the accuracy of individual predictions.

The ordering of the variables was achieved through the computation of sum of their correlations. The correlation between each pair of variables was calculated using Pearson correlation coefficient. This coefficient *r* was computed using the following formula:

$$r = \frac{\sum_{i=1}^{n}\left((x_i - \bar{x}) \cdot (y_i - \bar{y})\right)}{(n-1) \cdot s_x \cdot s_y} \tag{3.4}$$

Where:
r is the Pearson correlation coefficient, x_i and y_i are individual sample points indexed with i, $\bar{x}$ and $\bar{y}$ are the means of x and y variables respectively, n is the number of data points, s_x and s_y are the standard deviations of *x* and *y* variables respectively.

After obtaining the correlation matrix from these coefficients, a sum of the correlations for each variable was calculated by the following equation:

$$\Sigma Corr(k) = \sum Corr(k,l) \tag{3.5}$$

where:
$\Sigma Corr(k)$ is a sum of correlations between variable $\underline{k}$ and the rest of the variables in cascade, *Corr(k,l)* is the Pearson correlation coefficient between the main variable *k* and factor *l*

The resulting sums are ordered in ascending order, as the variable with largest $\Sigma Corr$ benefits the most from being at the end of the cascade.

the third group of models is centred on the generation and transfer of power. In an effort to maintain simplicity, the factors contributing to energy generation include steam input, live extraction estimates, and temperature forecasts. As this model is dedicated to the estimation of $W_{generated}^{\overline{bp}=0}$, it requires low level of complexity. Second model within this group is dedicated to the energy transfer to the power grid, primarily based on the energy generated and temporal variables to account for the plant's self-consumption.

In summary, our approach involves three distinct groups of models, each addressing specific challenges and requirements of the prediction process with final model order depicted in Figure 3.13.

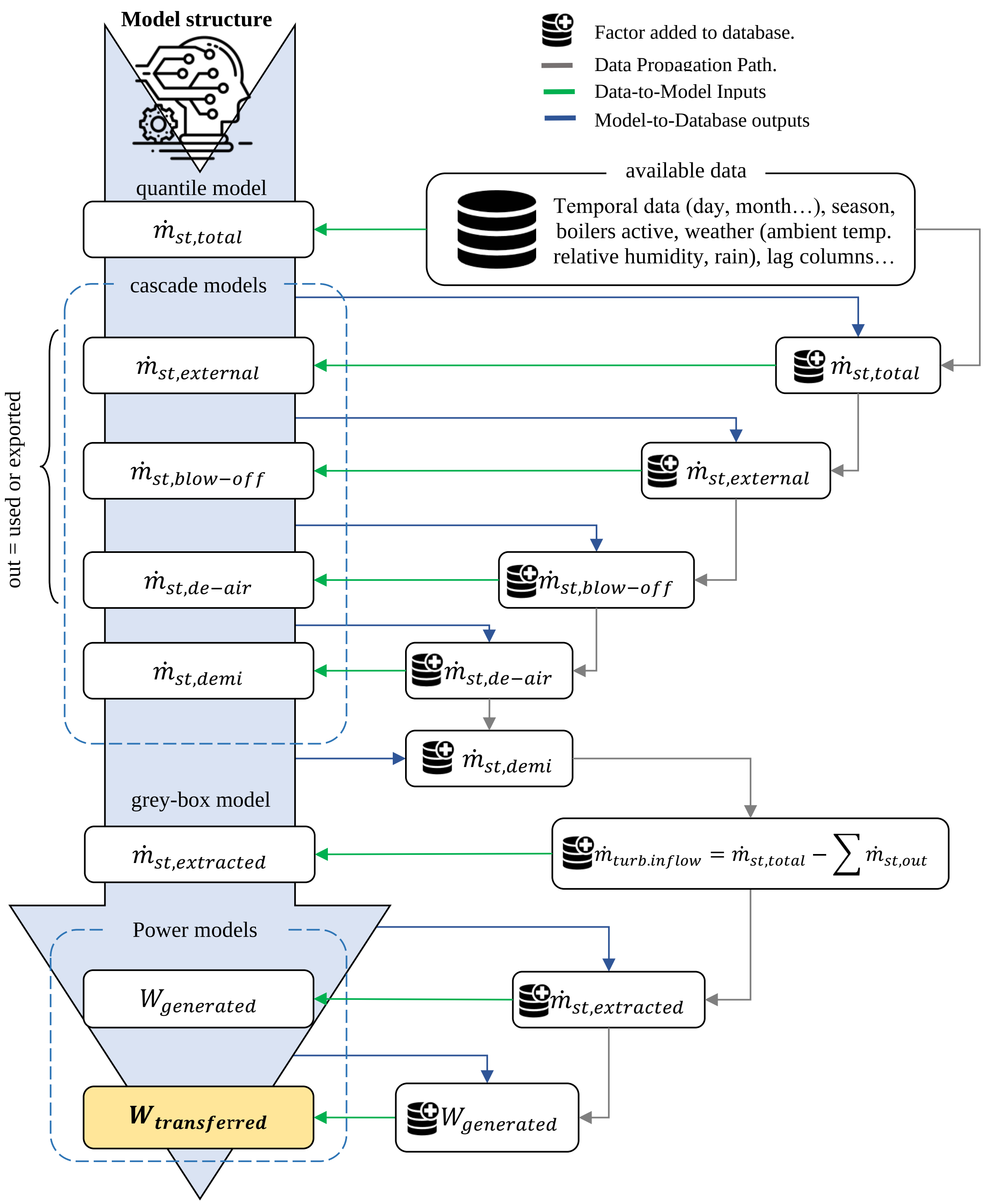

Figure 3.13.Compelte model structure. Going from database through the cascade of models.

3.1.3 Live steam extraction

As mentioned earlier, the plant's power is generated using a turbine with uncontrolled extraction referred to as bleeding turbine. The extraction steam flow rate significantly influences power and electricity production, making it crucial to develop accurate models for units within the turbine house.

The extraction steam flow rate was calculated based on Touš's work. The model consists of several components, including the self-consumption steam flow rate (grouped under the axillaries umbrella), mass and energy balance equations, and an algorithm for calculating extraction steam flow rates. Additionally, the model addresses the use of a turbine bypass to maintain higher steam temperatures in the DHS exchanger. An algorithm is presented that encompasses the entire live steam extraction process, incorporating the calculation of extraction steam flow rates, bypassed steam flow rates, as well as their corresponding temperatures and enthalpies (Figure 3.14).

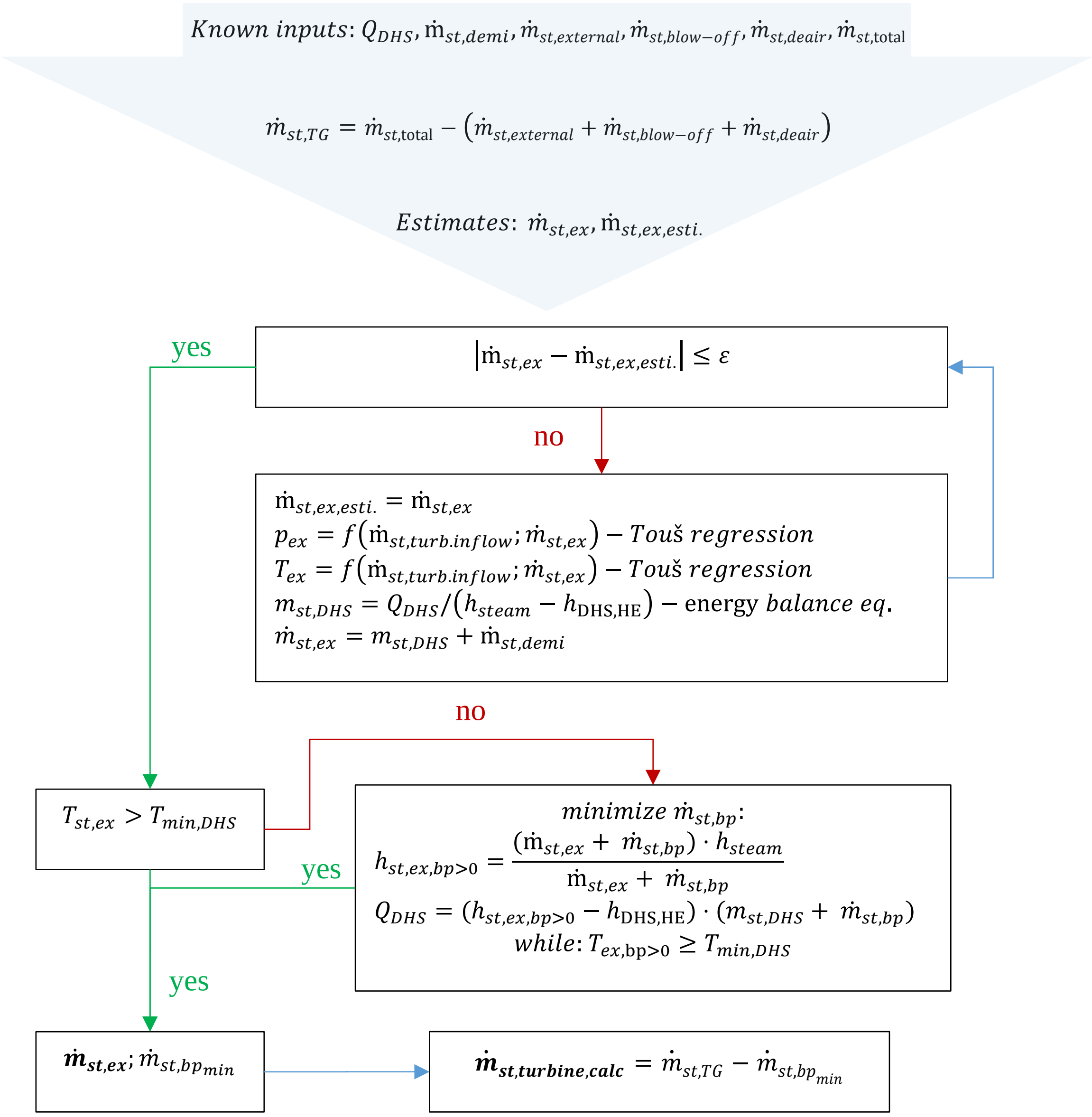

Figure 3.14: Extraction steam flow rate calculation algorithm.

Selection and Evaluation of Machine Learning Algorithms

Regression algorithms were the focus of our study, as the objective was to predict continuous values. Three algorithms were selected, representing varying levels of complexity: Linear Regression (low complexity), Light Gradient Boosting Machine (medium complexity), and Artificial Neural Networks (high complexity). This selection encompassed a range of complexity, enabling an examination of the trade-offs between model interpretability, complexity, and performance (see Table 3.3).

Table 3.3: Pros ans Cons of selected ML models.

	Pros	*Cons*
LR	- Simple and interpretable	- Limited model complexity
	- Fast training and prediction	- Assumes a linear relationship between variables
	- Minimal parameter tuning	- May underperform on complex data
LGBM	- Handles large datasets efficiently	- More complex than LR
	- Good performance on various problems	- Requires extensive parameter tuning
	- Supports categorical features without one-hot encoding	- Less interpretable than LR
	- Handles missing data and outliers well	
ANN	- Can model complex, non-linear relationships	- Requires extensive parameter tuning
	- Good performance on a wide range of problems	- Less interpretable than LR and LGBM
	- Can approximate any continuous function	- Slower training and prediction times
	- Can learn hierarchical representations	- Prone to overfitting

It should be noted that there are many other machine learning algorithms that could be applied to this problem (see Table 2.3), such as Long Short-Term Memory (LSTM), Support Vector Regression (SVR), and Random Forest Regression. However, the selected algorithms provide a diverse representation of the available techniques and offer a comprehensive understanding of their strengths and weaknesses when applied to the prediction of steam production in Waste-to-Energy plants.

Linear Regression

Linear Regression is a fundamental and widely used machine learning algorithm for regression tasks. It models the relationship between a dependent variable and one or more independent variables by fitting a linear equation to the observed data. In this study, we used a multiple linear regression model to predict steam production based on the selected features. Parameter tuning LR has minimal parameters to tune, making it a simple and interpretable model. The primary parameter is the regularization term, which helps prevent overfitting [5].

Light Gradient Boosting Machine (LGBM)

LGBM is a gradient boosting framework that employs tree-based learning algorithms. It is specifically designed for efficiency and scalability, enabling it to handle large datasets with a smaller memory footprint compared to other gradient boosting algorithms. LGBM has gained popularity owing to its ability to manage large datasets while delivering high performance. Gradient boosting algorithms work by combining several weak learners (Simple or low-complexity models), usually decision trees, to create a strong learner (as illustrated in Figure

3.15). LGBM improves upon traditional gradient boosting methods by using a leaf-wise growth strategy rather than a level-wise one. This leaf-wise approach allows the algorithm to focus on the most significant splits, leading to faster convergence and improved accuracy [5, 12].

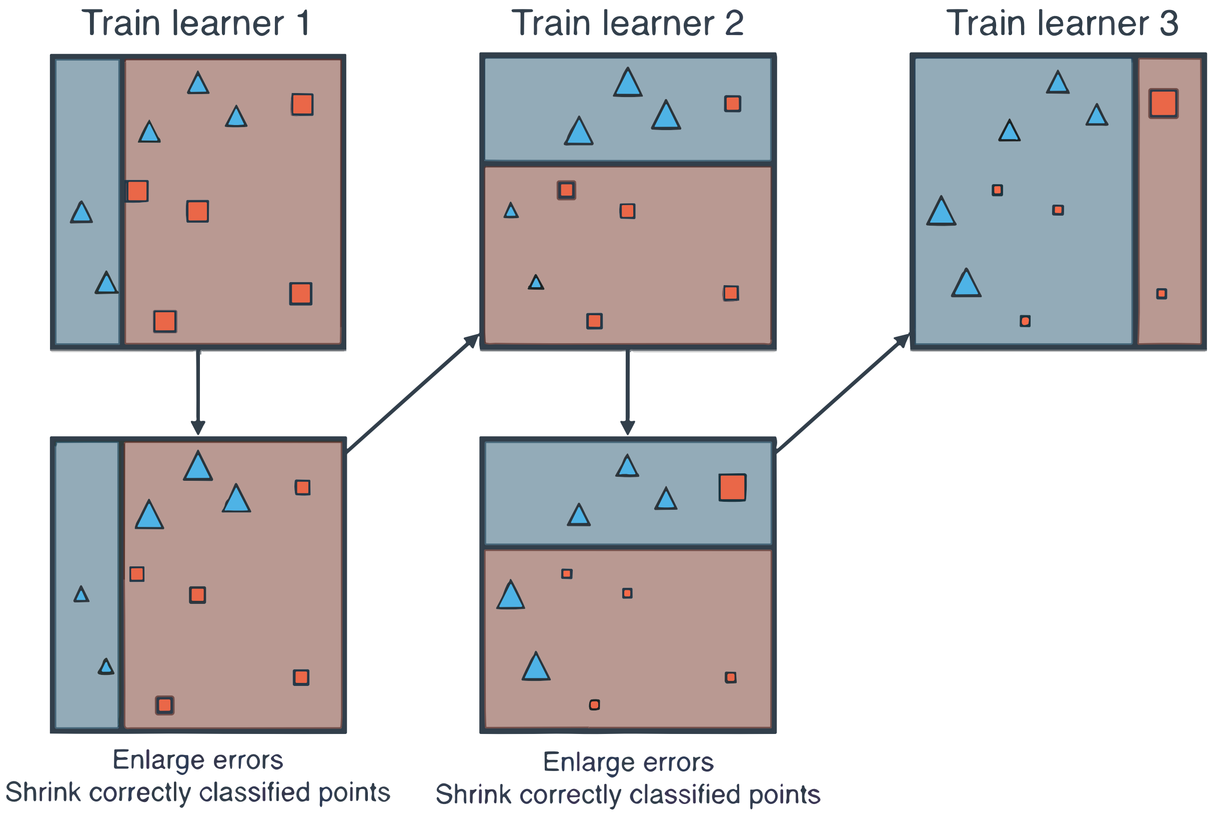

Figure 3.15 An example of weak learners being used to train a strong learner [47].

Artificial Neural Networks (ANN)

Artificial Neural Networks are a class of machine learning models inspired by the structure and function of the human brain. They consist of interconnected layers of nodes, also known as neurons (Figure 3.16). Each neuron receives input from previous neurons, processes it, and sends the output to the next layer. ANNs are versatile and can be used for various tasks, including regression problems. In a feedforward neural network, the most common type of ANN, information moves in one direction, from the input layer through hidden layers (if any) and to the output layer. The network learns by adjusting the weights and biases of connections between neurons during training. ANN models have numerous hyperparameters, including the number of hidden layers, neurons per layer, activation functions, learning rate, and regularization techniques. We used a systematic approach, such as grid search or random search, in conjunction with cross-validation to find the optimal set of hyperparameters for our specific problem [5].

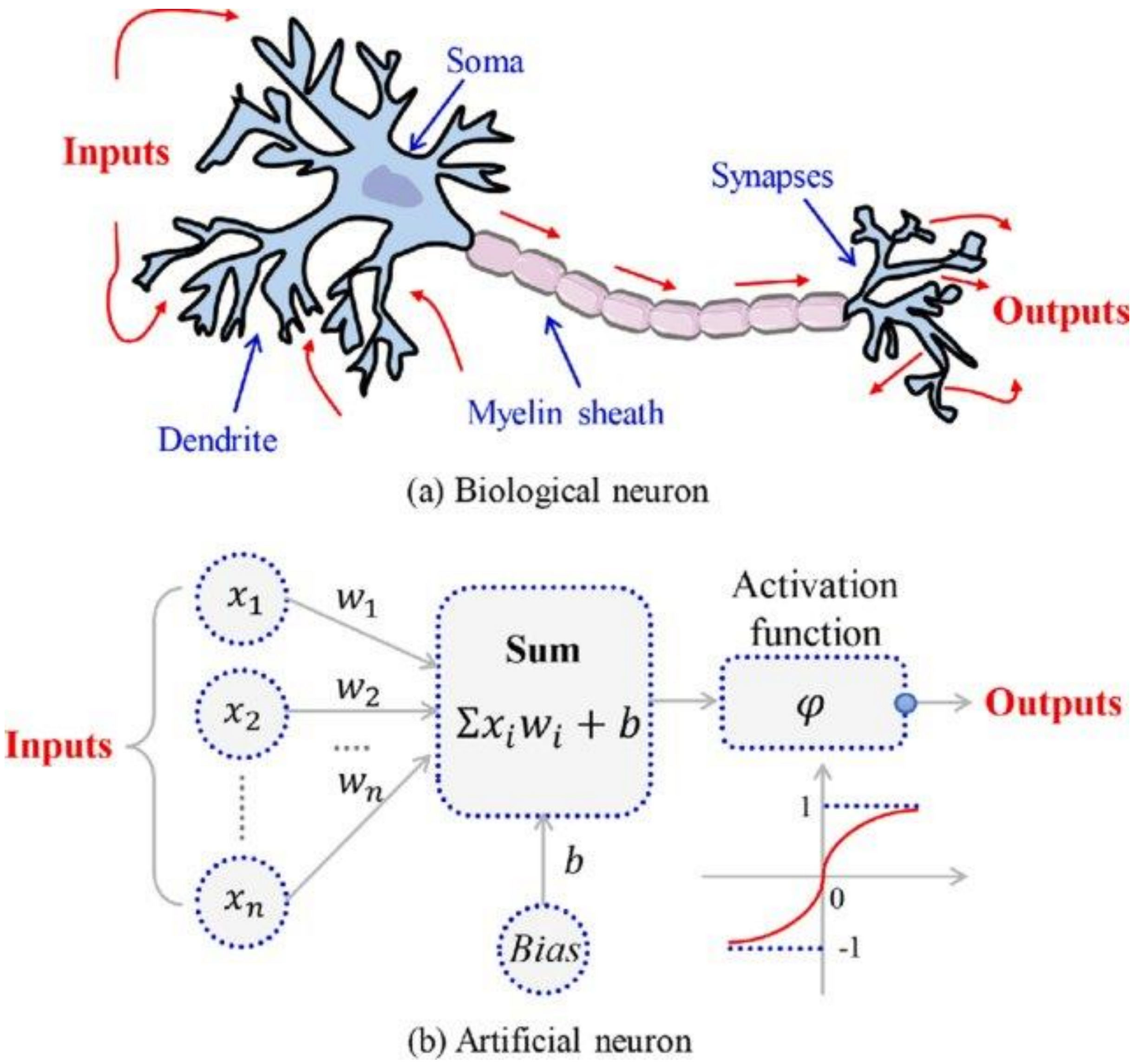

Figure 3.16: Comparison between biological neuron and artificial neuron [49]

3.1.4 Group-wise Model Comparison and Selection

The main driving factor for categorizing the models into distinct groups was the differing requirements of various variables in our predictive framework. Notably high degree of confidence in total steam production prediction and the demands for managing extrapolation in power models and a necessitated such a grouping.

Quantile-Based Performance

In the context of total steam production prediction, quantile-based performance metrics play a crucial role. We evaluate the models based on the following quantile-based criteria:

1) Coverage Probability: The proportion of observed values that fall within the 95% prediction intervals. The desired value for this metric is 0.95 or higher.
2) Performance metrics: Various performance metrics, including Mean Absolute Error, Mean Relative Error, and R-squared, were used to evaluate the models, providing insights into their accuracy and precision.

Auxiliary Performance

Models predicting auxiliary variables where selection was based common metrics, this is also why the group was originally named auxiliary as this grouping is made out of less essential streams, these metrics therefore are:

1. Performance metrics: Various performance metrics, including MAE, MRE and R^2 were used to evaluate the models, providing insights into their accuracy and precision.
2. Model complexity: The complexity of the models was considered, with a preference for simpler models that offered similar performance to more complex ones to minimize the risk of overfitting and improve model interpretability.

Power generation/transfer Performance

As previously mentioned, the creation of a turbine model capable of calculating $W_{generated}^{\overline{bp}=0}$ using data that may have a different distribution than the data on which it was trained. Hence both, the extrapolation capability and regression metrics of the models need to be taken into account.

1. Performance metrics: The models were evaluated using a range of performance metrics, such as MAE, MRE, and R^2. These metrics provided valuable insights into the accuracy and precision of the models.
2. Extrapolation capability: The models' ability to extrapolate beyond the training data was examined using the Standardized Mean Error (SME), with Linear Regression (LR) serving as the benchmark. This analysis allowed for a comparison of the performance of more complex models in terms of their ability to handle extrapolation.

The model that best balances the criteria across all three groups and demonstrates the optimal combination of predictive performance, complexity, and generalization ability is selected as the final model for predicting it’s assigned variable.

3.2 Model Development and Performance Evaluation

In this chapter, the selected machine learning models are developed and evaluated. Building upon the outlined methodology. The specifics of model development, validation, performance evaluation, and final selection are explored. Emphasis is placed on the implementation details and customization of the models to address the problem at hand, while validating their performance and evaluating their effectiveness in predicting steam production.

Data is split into three sets: train (80% of data from 2011 to 2015), validation (remaining 20% of data from 2011 to 2015), and test (year 2016), as visualized in Figure 3.17. Year 2017 is excluded due to the large gap in the dataset, aiming to calculate the annual benefit of this approach.

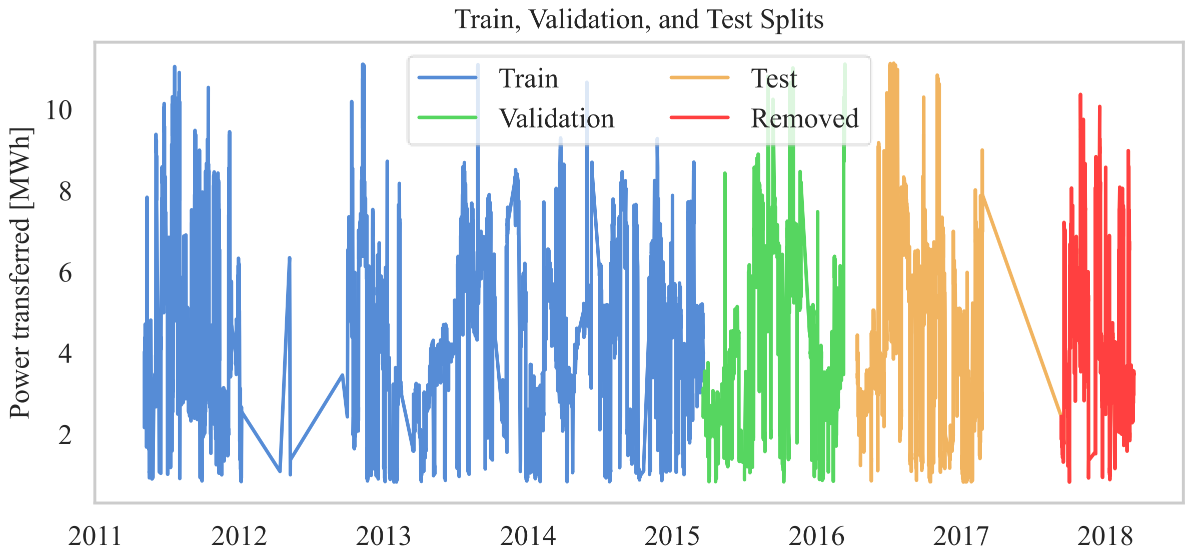

Figure 3.17: Visualization of Train/Validation/Test split along with unused data.

3.2.1 Data Pre-processing

Following the general process outlined in the methodology sections, data pre-processing is tailored to the specific dataset and problem. The results and specifics of the data pre-processing approach applied to this dataset are presented.

Handling Missing Values

Analysis of the dataset revealed 8.1% of the data points with missing key values. As discussed in the methodology section, the decision was made to remove these rows from the dataset.

Outlier Detection and Removal

After applying the Z-score outlier removal method and filtering out non-standard operation days, a total of 35877 datapoints remained, approximately 4 years' worth of data. Half of the year 2011 and most of the year 2012 were largely removed as they fell outside the distribution of the test dataset (see Figure 3.18 and Figure 3.19).

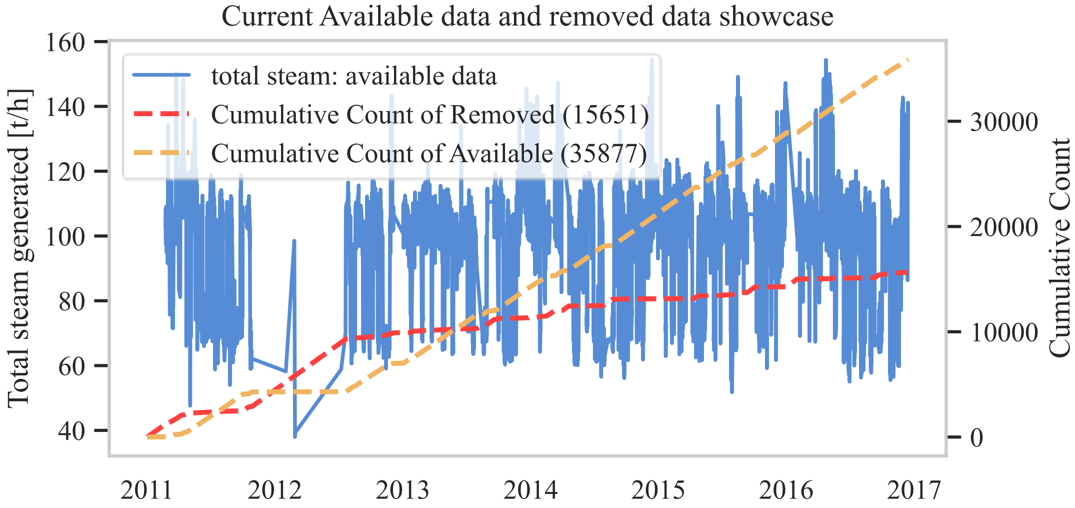

Figure 3.18: Showcase of the total available data of total steam generated along with the cumulative counts of removed and available values.

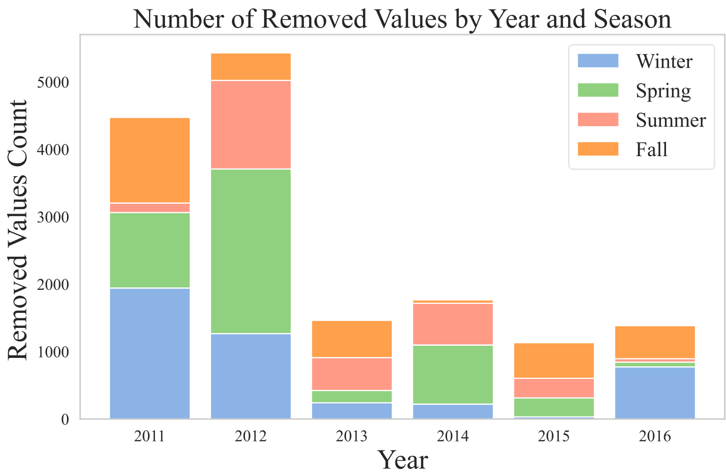

Figure 3.19: Removed values by year and season.

In summary, the data pre-processing approach resulted in a cleaned, encoded, and normalized dataset containing approximately 4 years' worth of data points and 139 features, ready for model development and evaluation.

3.2.2 Model Development: General Process

A general model development process was adhered to for all groups of machine learning models, comprising of the following steps:

1. Model Development: Depending on the chosen model, the appropriate libraries and settings were used (e.g., LGBM library for LGBM Quantile models).
2. Parameter Tuning: A range of hyperparameters was explored, and random search in combination with cross-validation was used to find the optimal set of hyperparameters for each model.
3. Model Training: Models were trained using the optimal set of hyperparameters and the pre-processed dataset. Performance was evaluated using cross-validation and the performance metrics

Group 1 - ANN, LR, and LGBM Quantile Models

For the first group of machine learning models, the general model development process was followed. Specific steps unique to each model in this group include:

- LGBM: The objective function was set to 'quantile' and an appropriate quantile value was specified.
- Linear Regression and Artificial Neural Networks: Custom loss function was created.

Group 2 - AUXILIARY Models

The general model development process was adhered to, with one notable exception: the top 7 features, as determined by the feature importance function of LGBM (refer to Figure 3.20), were utilized. This strategic selection of key factors was beneficial for both LR and ANN, as these models encountered difficulties handling an excessive number of unhelpful columns. In certain instances, the inclusion of these columns resulted in significant underperformance of the models. However, it is noteworthy to mention that this limitation on the number of factors did not compromise the efficacy of the models, as evidenced by their satisfactory performance.

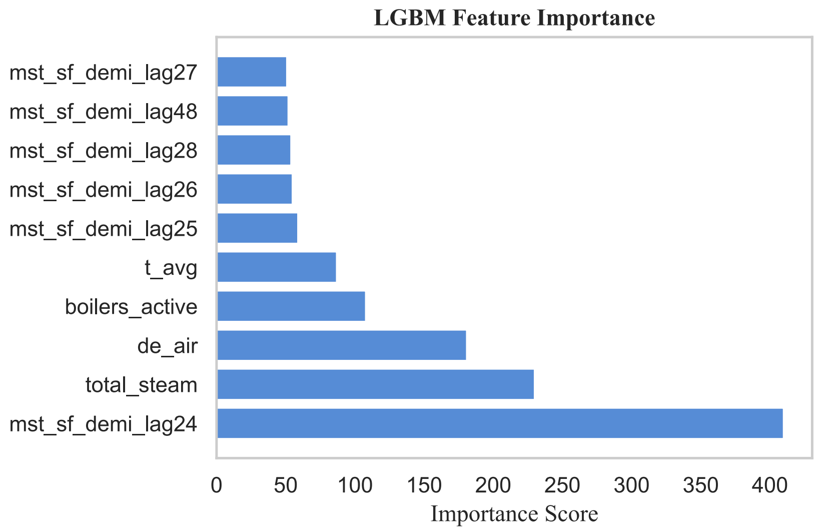

Figure 3.20: LGBM importance score for pre-heating steam LGBM auxiliary model.

Second difference for this group was the incorporation cascade. The resulting correlation matrix and order the of prediction (from top to bottom) depicted in Figure 3.21.

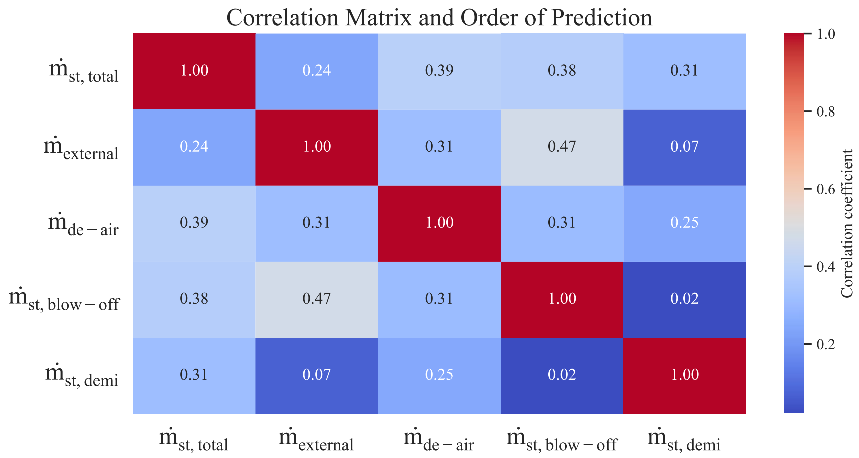

Figure 3.21: Correlation matrix for auxiliary steam usage ordering.

Group 3 - Turbine Power Generation and Transfer Models

In the development of these models, we placed a significant focus on simplicity and interpretability, adhering to the standard model development process. In the context of power generation, we chose to forgo the use of lag columns and temporal variables. This decision ensures that the generated power is fundamentally tied to the steam in the turbine and the prevailing weather conditions, aligning with the specifications detailed in the previous chapter. For the model of power transfer, our approach was to exclusively use data on power generation and temporal aspects. This strategy helps capture the dynamics of the plant's internal electricity consumption more accurately.

3.2.3 Model Comparison and Selection

For each model group, the performance was evaluated based on the criteria outlined in the methodology chapter. The model that offered the optimal blend of performance, complexity, and generalizability within each group was selected. The models selected for each group are as follows:

Group 1 - Total steam production:

When the performance of the models was compared at the 5% quantile, the results were clear. The LGBM model was found to be the most viable candidate based on the results shown in Figure 3.22.

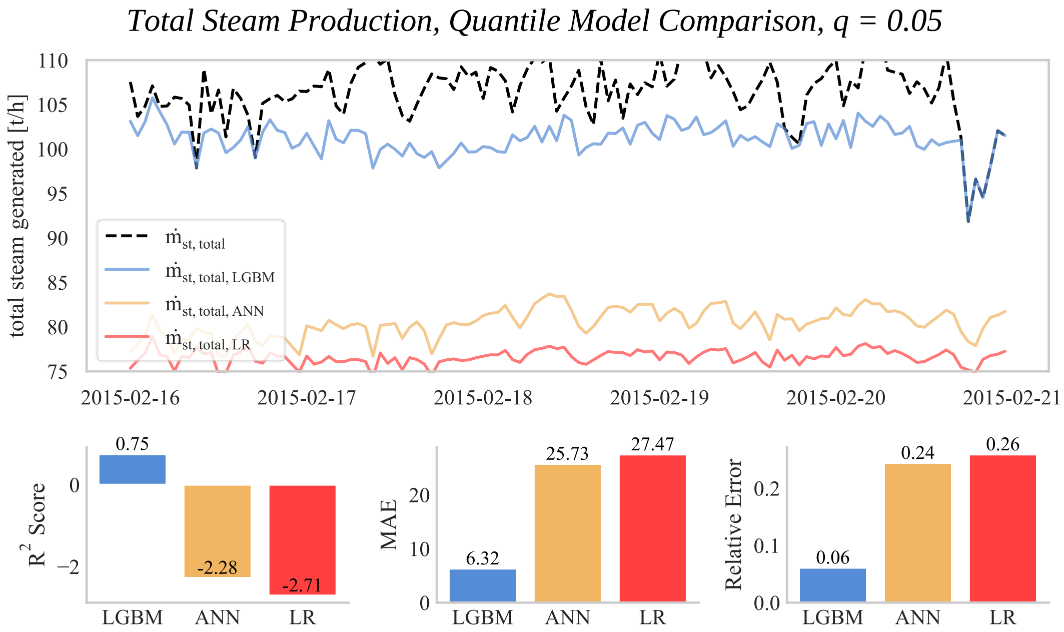

Figure 3.22: Comparison of algorithms' capability to predict total steam production flowrate using quantile loss parameter q = 0.05.

To ensure the results were not caused by the poor model quality, the same models were compared with the quantile parameter set to 50%. Despite increased competitiveness among the models in this setting, the LGBM model was found to perform the best Figure 3.23.

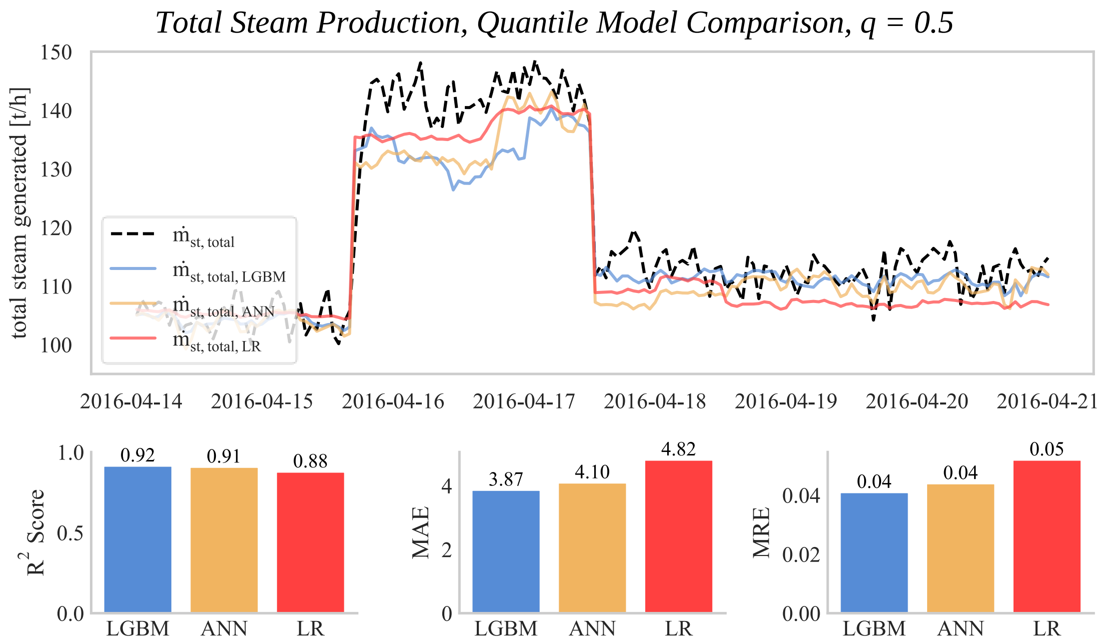

Figure 3.23: Comparison of algorithms' ability to predict total steam production flowrate defined quantile loss parameter q = 0.5 (mean).

Group 2 - blow-off, deaerator, external consumption, pre-heating:

Blow off (Figure 3.24) and deaerator air steam (Figure 3.25) models exhibited relatively simple behaviour, efficiently captured by the linear regression model.

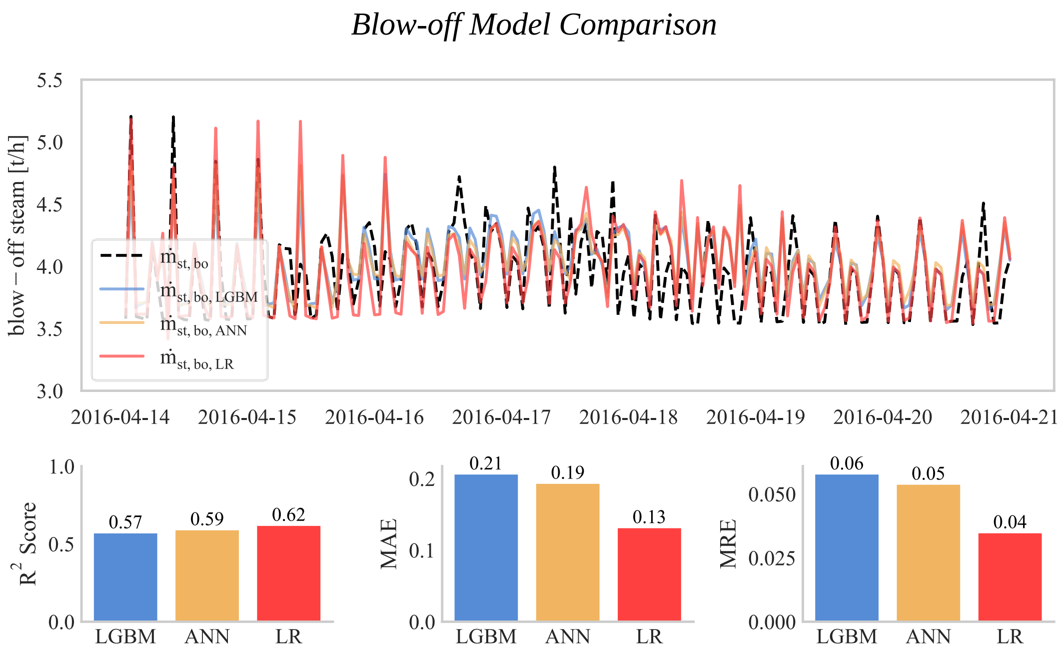

Figure 3.24: Comparison of algorithms' ability to predict blow-off flowrate.

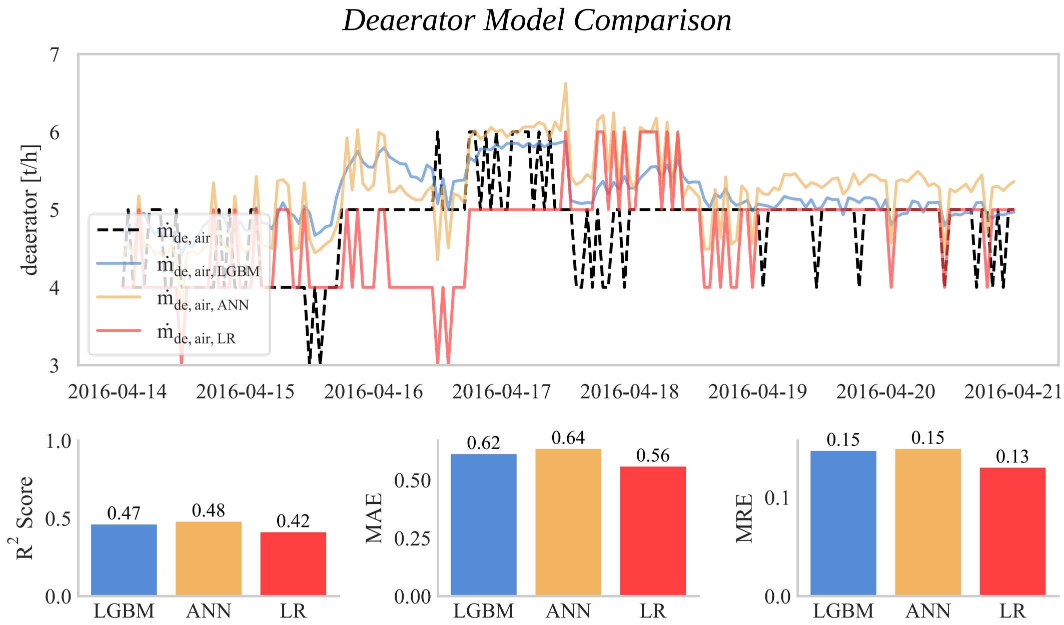

Figure 3.25: Comparison of algorithms' ability to predict deaerator flowrate.

The external consumption (Figure 3.26) presented a more complex scenario Since it is controlled by an external company, there are no internal variables that could provide meaningful insight. In this case, the LGBM was found to be the most suitable model by a thin margin.

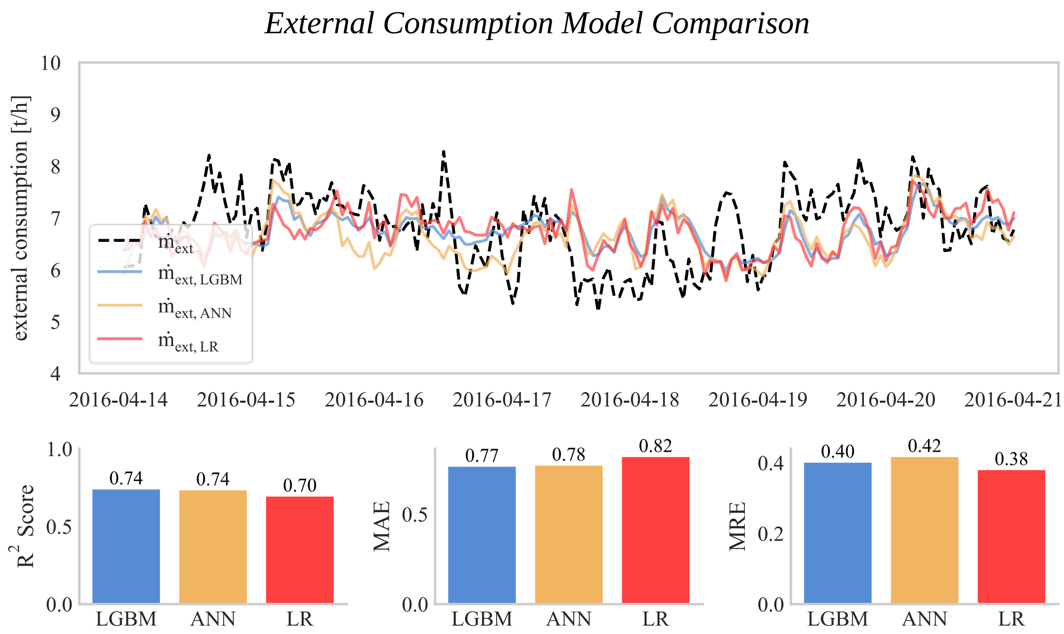

Figure 3.26: Comparison of algorithms' ability to predict deaerator flowrate.

For the pre-heating self-consumption in Group 2, the LGBM was found to perform the best across all metrics (Figure 3.27). However, it is important to note that the R^2 score across the models suggests that reassuring results were not yielded by any of the algorithms. Thankfully, the impact of this stream is minimal, as it is solely used for live extraction. Furthermore, considering the low overall value of the pre-heating flow rate and the small MRE, the results are deemed satisfactory.

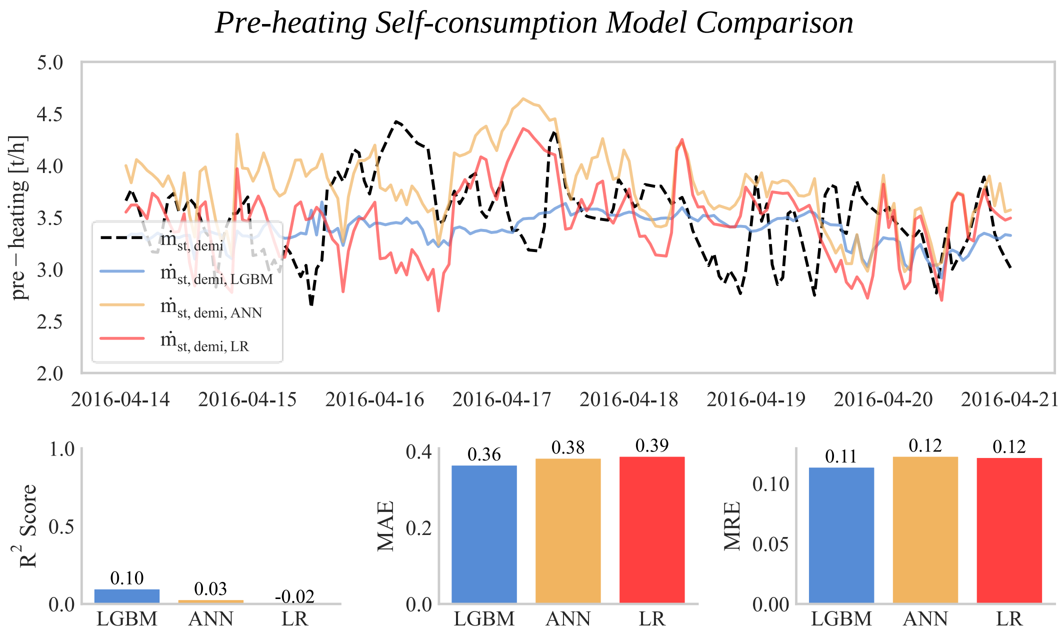

Figure 3.27: Comparison of algorithms' ability to predict pre-heating self-conception flowrate.

Group 3 - Power generation and transfer:

As stated in chapters 3.1.1, challenge of the third group is not the performance of the models' themselves (Figure 3.28 and Figure 3.29), but trustworthiness of theirs outputs when given data unaffected by bypass.

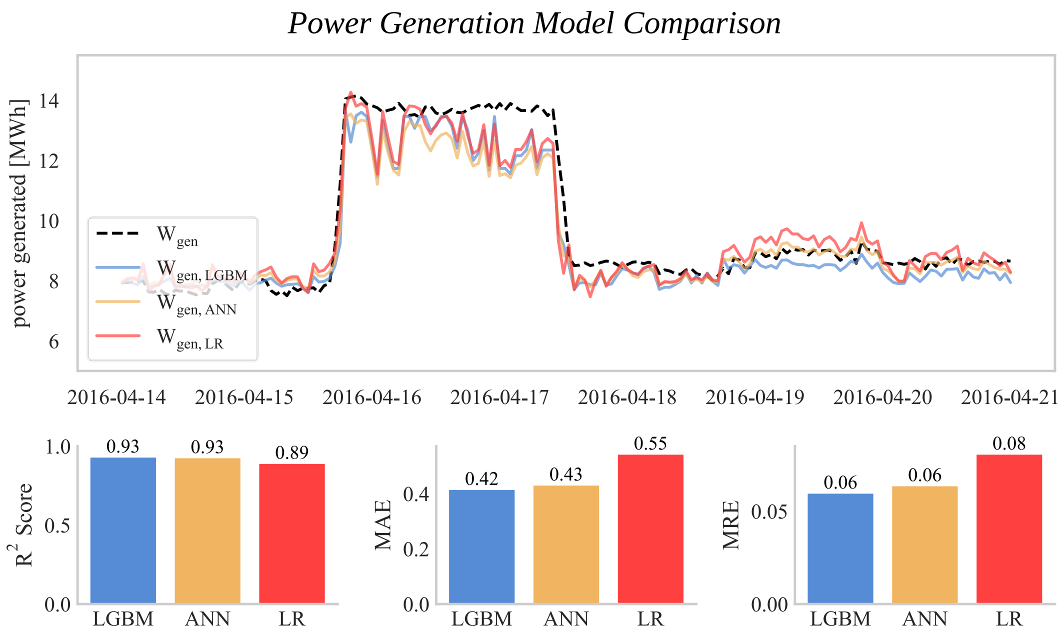

Figure 3.28: Comparison of algorithms' ability to predict generated power.

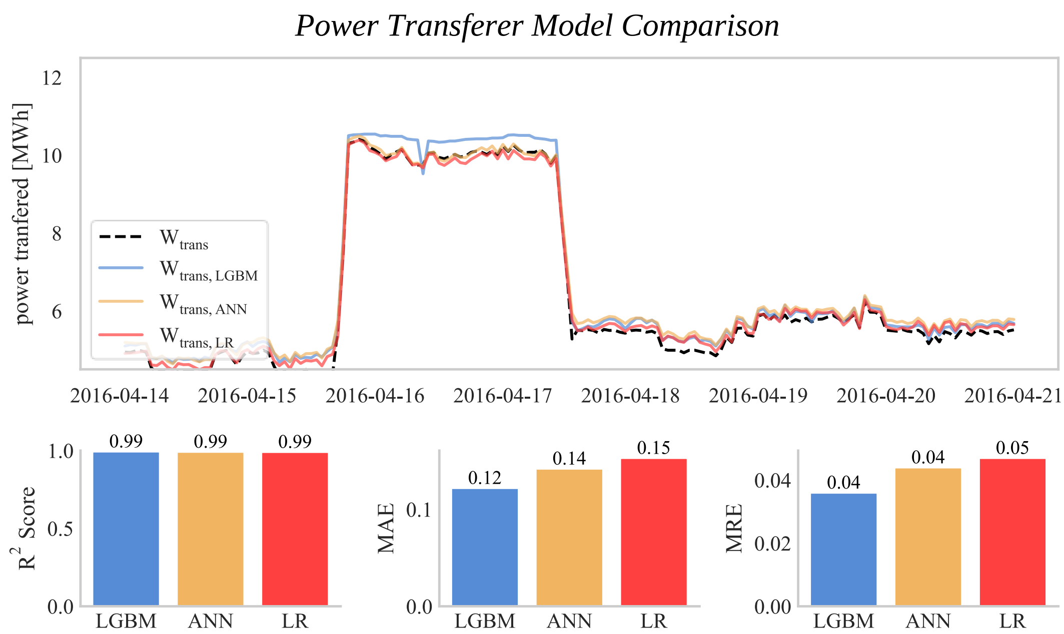

Figure 3.29: Comparison of algorithms' ability to predict transferred power to gird.

Great performance in the modelling of power generation was shown by the LR model but potential limitations was recognized, particularly concerning the correlation between power generation and weather. This significant correlation was acknowledged by both plant operators and Touš et al., and further emphasized by the feature importance function of the LGBM model (Figure 3.30) and the LR approximation equations (3.6) and (3.7)[2].

It was presumed that, with more detailed data about steam quality for power generation - such as pressure and temperature, or further information about the plant's power self-consumption - models like ANN or LGBM might exhibit improved potential for future refinement. Nevertheless, their extrapolative capabilities could be limited.

$$\widehat{W}_{generated,LR} = 2.548 \cdot \dot{m}_{st,turb.inflow} + 0.146 \cdot t_{avg} - 0.693 \cdot \dot{m}_{st,extracted} - 0.029 \cdot h_{avg} + 7.23 \quad (3.6)$$

$$\widehat{W}_{transferred,LR} = 2.477 \cdot day - 0.005 \cdot t_{avg} + 0.037 \cdot h_{avg} - 0.047 \cdot month - 0.374 \cdot boilers\ count + 0.003 \cdot hour + 4.360 \quad (3.7)$$

Where:
$\dot{m}_{st,turb.inflow}$ is total turbine steam inflow, t_{avg} is average temperature, $\dot{m}_{st,extracted}$ is turbine uncontrolled extraction flowrate, h_{avg} is average relative humidity, day is a day of a week, month is a month of a year and boilers count is count of boilers active.

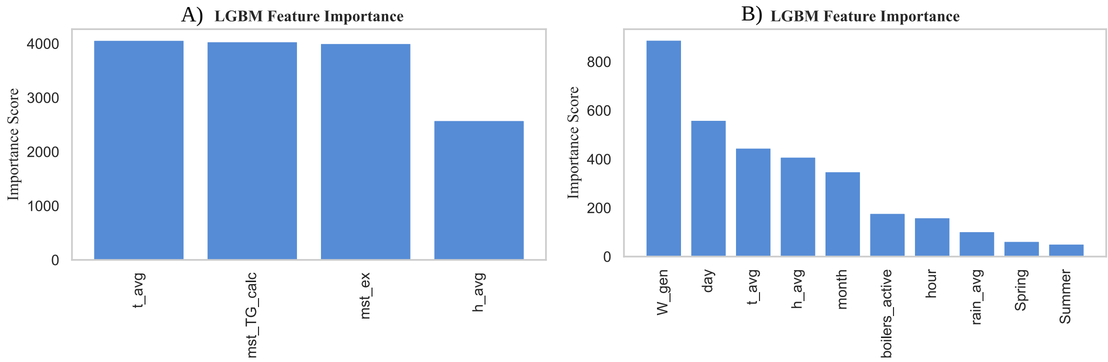

Figure 3.30: Feature importance – LGBM model. A) Power generated model, B) Power transferred model.

A test was conducted, wherein models were trained on original data and extrapolated by supplying them with steam flow rates unaffected by bypass. The LR model, due to its clear correlation between inputs and outputs, offered a robust benchmark against which the LGBM and ANN models were compared. Signed mean error and sum ratios at various quantiles $(\dot{m}_{st,total} < Q33; Q33 < \dot{m}_{st,total} < Q66;\ Q66 < \dot{m}_{st,total})$ were utilized to detect possible biases. It was anticipated that any model exhibiting a larger bias or significantly lower total sum would indicate ineffective extrapolation.

From the results (Figure 3.31 and Table 3.4), it was observed that the LGBM model was unable to extrapolate effectively - an expected outcome, given that tree-based methods like LGBM frequently struggle with extrapolation due to their inherent nature of capturing patterns within the range of the training data, but lacking the capacity to predict beyond that range [5].

[2] The equation inputs are to be normalized using their standard deviation and mean from the training set.

In contrast, the ANN model exhibited a high level of extrapolative competence, leading to its selection as the algorithm of choice for future implementation.

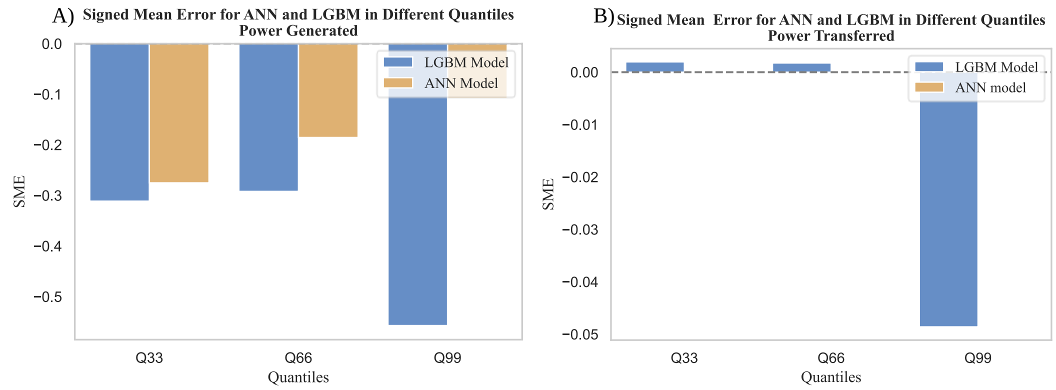

Figure 3.31: Signed Mean Error of LGBM vs LR and ANN vs LR comparisome. A) Power generated model, B) Power transferred model.

Table 3.4: Sum ration of LGBM and ANN for power generation and transfer comparison.

	∑*model ratio*	*Q33*	*Q66*	***Q99***
Power Generated	∑LGBM /∑LR	0.97	0.95	0.85
	∑ANN /∑LR	0.99	0.98	1.00
	∑model ratio	Q33	Q66	**Q99**
Power Transferred	∑LGBM /∑LR	0.99	0.99	0.86
	∑ANN /∑LR	1.00	1.00	1.00

Upon analysing the results in the context of the three groups, the optimal models for each variable were identified.

Table 3.5: Selected models with their corresponding metrics.

Variable predicted	***Best Model***	***R^2***	***MAE***	***MRE***
Total Steam Production	LGBM	0.75	6.32	0.06
Blow-off	LR	0.62	0.13	0.04
Deaerator	LR	0.42	0.56	0.13
External Consumption	LGBM	0.74	0.77	0.4
Pre-Heating	LGBM	0.1	0.36	0.11
Power Generation	ANN	0.99	0.43	0.06
Power Transferred	ANN	0.99	0.14	0.04

3.2.4 Predictive Accuracy and Economic Impact

In the pursuit of establishing a model for power transfer with defined risk of overestimation, it was observed that the system's predictability was associated with the accuracy of the estimated of total steam generated. The visualization of this correlation in KDE plots, where residuals of total steam generated were compared against residuals of power produced (Figure 3.32 – A) revealed that the data points aligned closely to the diagonal, thereby signifying a robust correlation between these variables. On the other hand, the concentration of residuals from total steam generation, when compared with those of external consumption, tended to cluster towards the centre (Figure 3.32 – B) suggested a more stochastic behaviour. The observation of such stochastic behaviour was not entirely unexpected as the control over external factors is uncontrolled and the current dataset does not provide any internal variables that could feasibly enhance the predictive capacity in this regard.

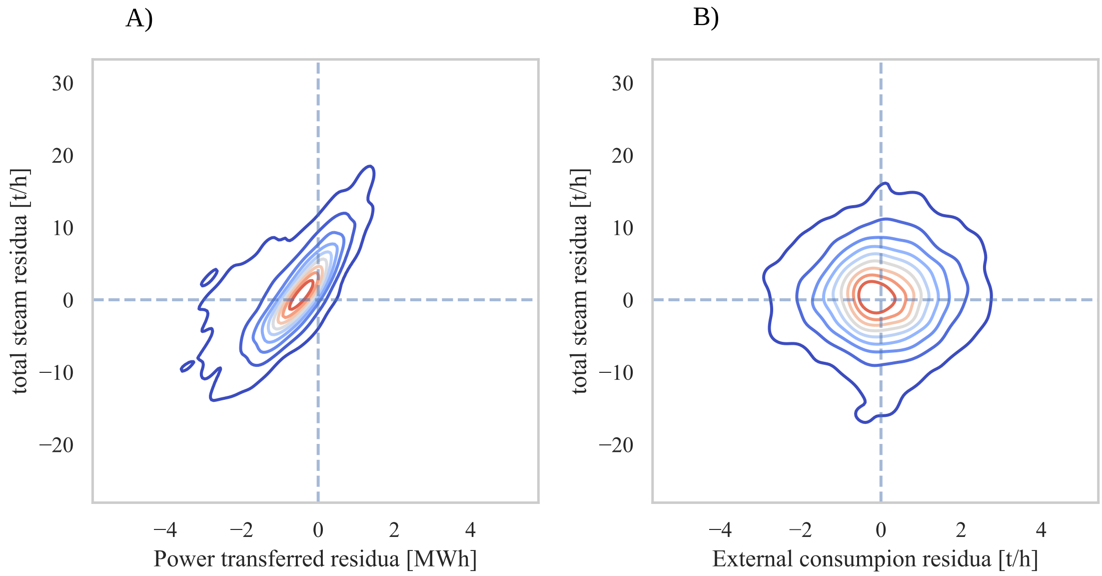

Figure 3.32: Residaul KDE plot comparing total steam prediction residua distributioin with A) Power transferred residas, B) External consumprion residuas.

The selection of q = 0.05 was made with the expectation that it would result in an overestimation rate of approximately 5%. Without constructing a tolerance interval similar in width to that of power transfer (1/16 of the mean), an overestimate rate of 20% would be observed instead (Figure 3.33). Taking into account the inherent variability of waste and the absence of factors enabling more robust boiler modelling, this outcome is considered satisfactory.

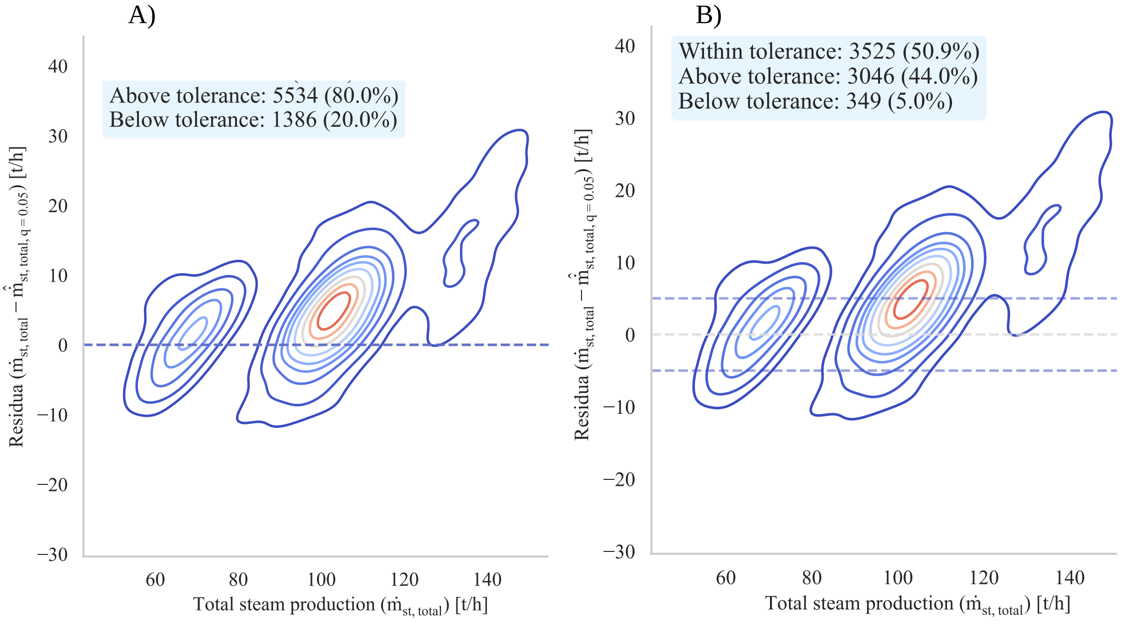

Figure 3.33: Residual KDE plot with overall success rate of total stem forecast A) without tolerance interval, B) with tolerance interval.

Electricity Delivery Predictive Accuracy

The methodology implemented was demonstrated to have attained an average predictive accuracy of 95% within the ±0.5 MWh interval. The novel approach employed was effective in pushing power transfer closer to its full potential, a significant improvement when compared with the current approach (as illustrated in Figure 3.34).

Consequently, a more optimal plan was established relative to the existing strategy. The average difference between the two plans was noted to be 13%, with an increase to 89% of maximal potential from the current 79%.

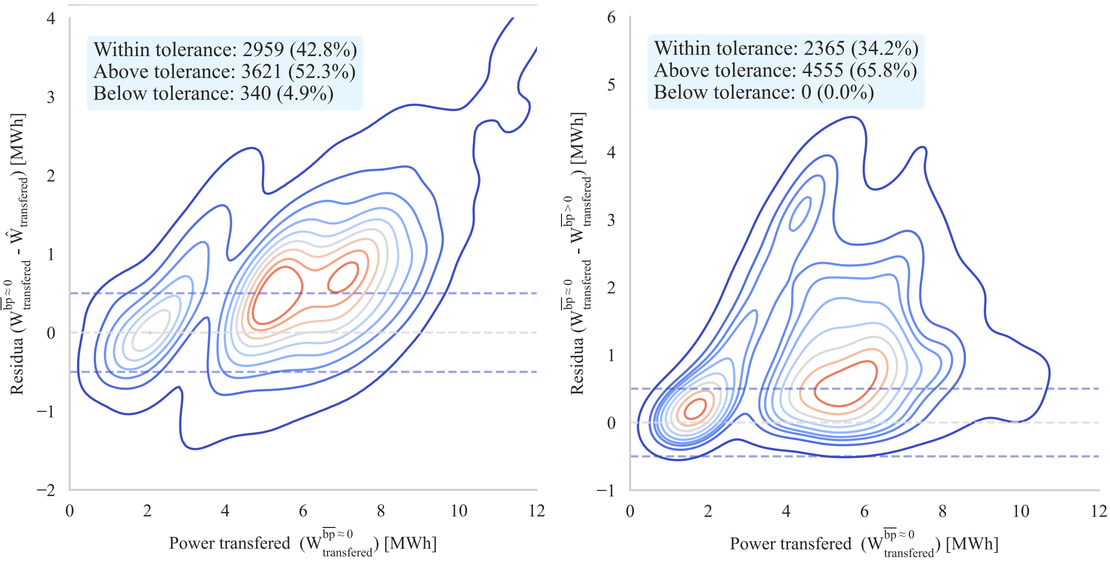

Figure 3.34: Residual KDE plot with overall successrate of power transferred. A) Novel aproach, B) current aproach.

Throughout the year, the model maintained a consistent success rate, except for a notable drop in accuracy during December. This anomaly might be attributed to increased steam production, which, in relative terms, reduced the protective effect of the ±0.5 MWh buffer zone, potentially impacting the model's prediction accuracy.

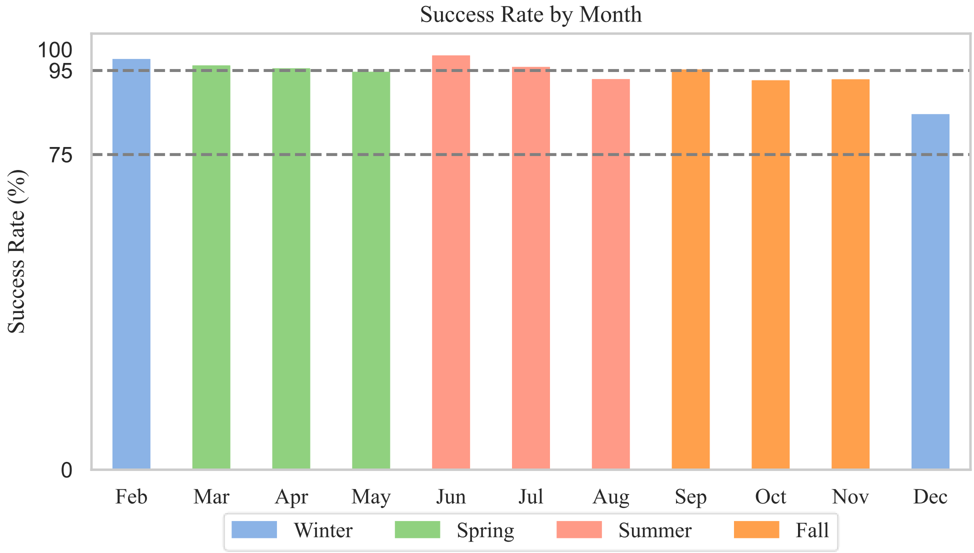

Figure 3.35: Success rate of power transferred predictions througjout the year.

Figure 3.36 serves as visual representations of characteristic periods across all four seasons, highlighting the subtleties of the prediction failures and the differences between the two approaches. It is worth noting that as the bypass usage was minimal during

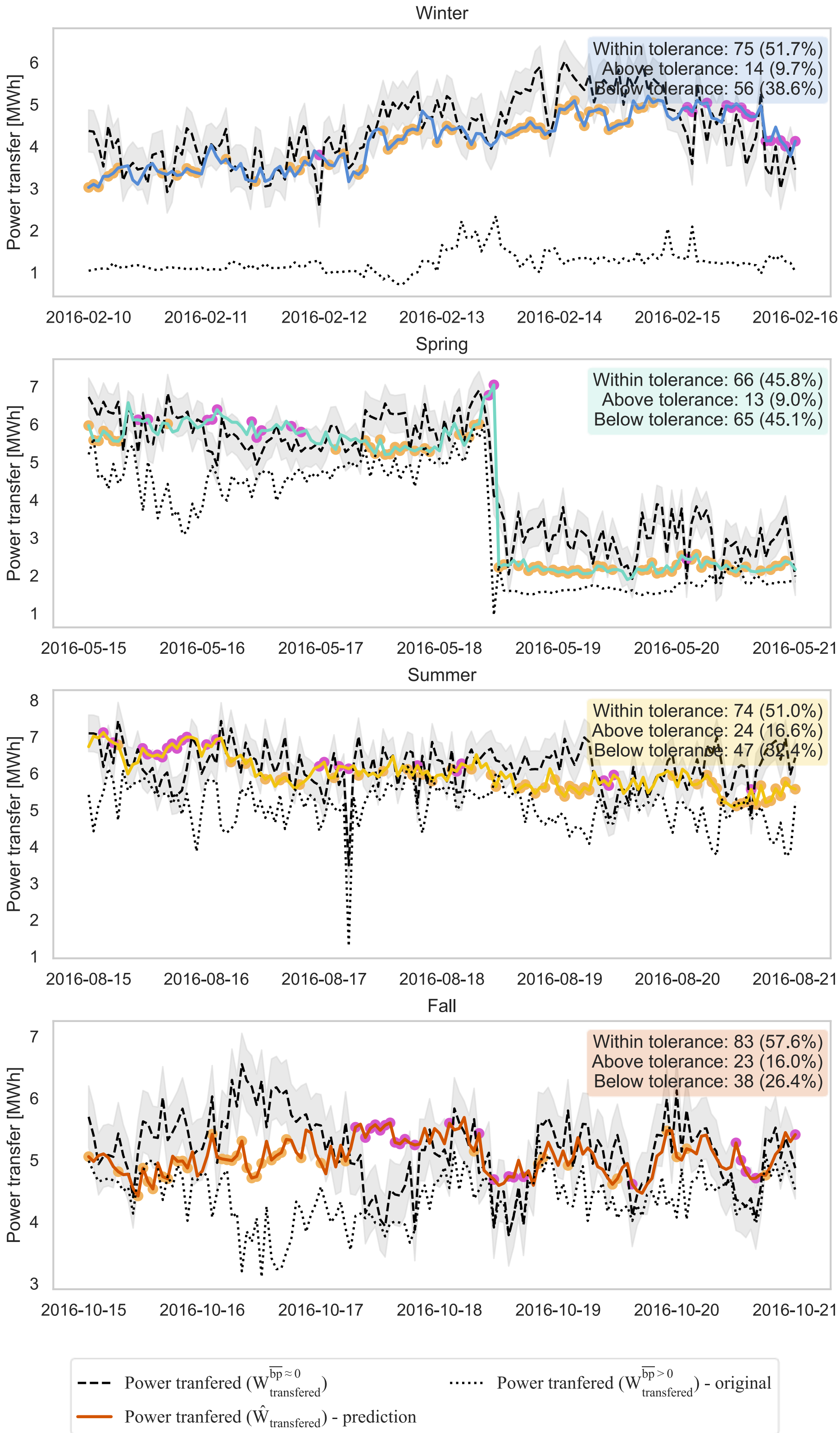
Winter
Within tolerance: 75 (51.7%)
Above tolerance: 14 (9.7%)
Below tolerance: 56 (38.6%)
Power transfer [MWh]
2016-02-10
2016-02-11
2016-02-12
2016-02-13
2016-02-14
2016-02-15
2016-02-16
Spring
Within tolerance: 66 (45.8%)
Above tolerance: 13 (9.0%)
Below tolerance: 65 (45.1%)
2016-05-15
2016-05-16
2016-05-17
2016-05-18
2016-05-19
2016-05-20
2016-05-21
Summer
Within tolerance: 74 (51.0%)
Above tolerance: 24 (16.6%)
Below tolerance: 47 (32.4%)
2016-08-15
2016-08-16
2016-08-17
2016-08-18
2016-08-19
2016-08-20
2016-08-21
Fall
Within tolerance: 83 (57.6%)
Above tolerance: 23 (16.0%)
Below tolerance: 38 (26.4%)
2016-10-15
2016-10-16
2016-10-17
2016-10-18
2016-10-19
2016-10-20
2016-10-21
Power tranfered ($W^{\overline{bp}\approx 0}_{transfered}$)
Power tranfered ($W^{\overline{bp}>0}_{transfered}$) - original
Power tranfered ($\hat{W}_{transfered}$) - prediction

Figure 3.36: Cross-seasonal comparison of novel and current method for representative time frames with local success rate of a novel approach.

Financial assessment

In the financial assessment, it was assumed that the contractual price for electricity transferred to the grid would not significantly deviate from day-ahead market prices. Therefore, electricity prices provided by OTE, a.s., a key operator in the electricity and natural gas markets in the Czech Republic [51], were utilized.

In the case of underestimation, the option to utilize the bypass was available, resulting in a minor penalty due to violation of the heat delivery contract. However, in cases of overestimation, the contractual obligations could not be met, leading to a more substantial penalty as illustrated in Figure 3.37.The penalties for the two scenarios were determined using different coefficients: $C_u = 0.15$ for underestimation and $C_o = 3$ for overestimation. The penalty for contract violation in this study was computed using equation (3.8).

$$pentaly = \left|W_{transfered} - W_{contract}\right| \cdot P_{\mathrm{OTE}} \cdot C \tag{3.8}$$

Where:

$W_{transfered}$ is power transferred to the grid, $W_{contract}$ is power to be delivered in day-ahead contract, P_{OTE} is price of electricity on day-ahead-market, C is a coefficient that varies depending on whether the production was underestimated or overestimated, with values of 0.15 and 3, respectively.

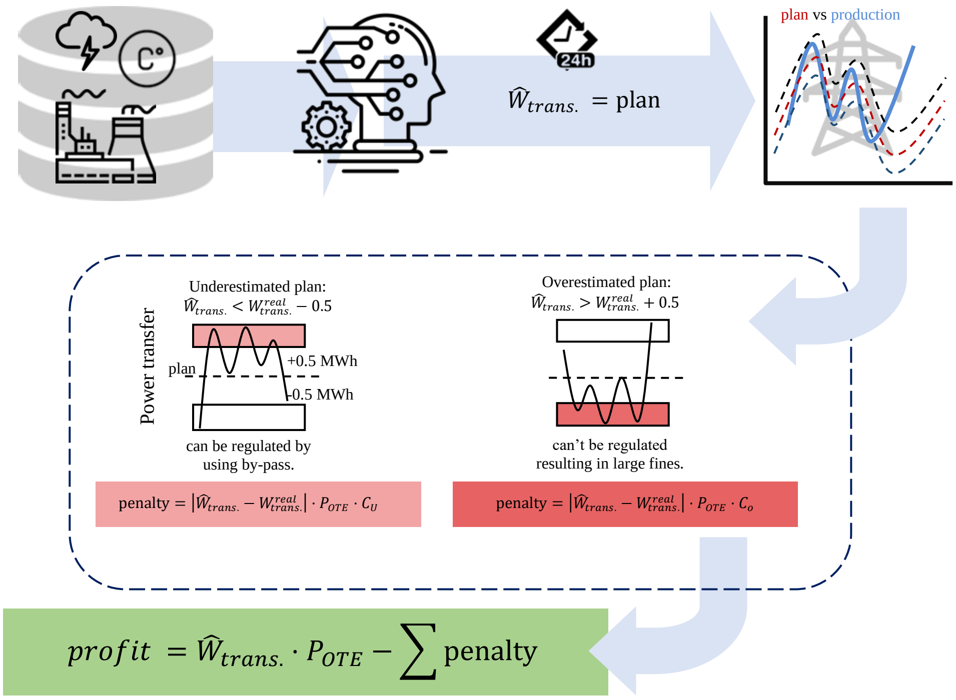

Figure 3.37. Power delivery profits calculation.

Comparing the novel approach with the current strategy, the estimated yearly profit increase was approximately 2.6 million CZK (Figure 3.38).

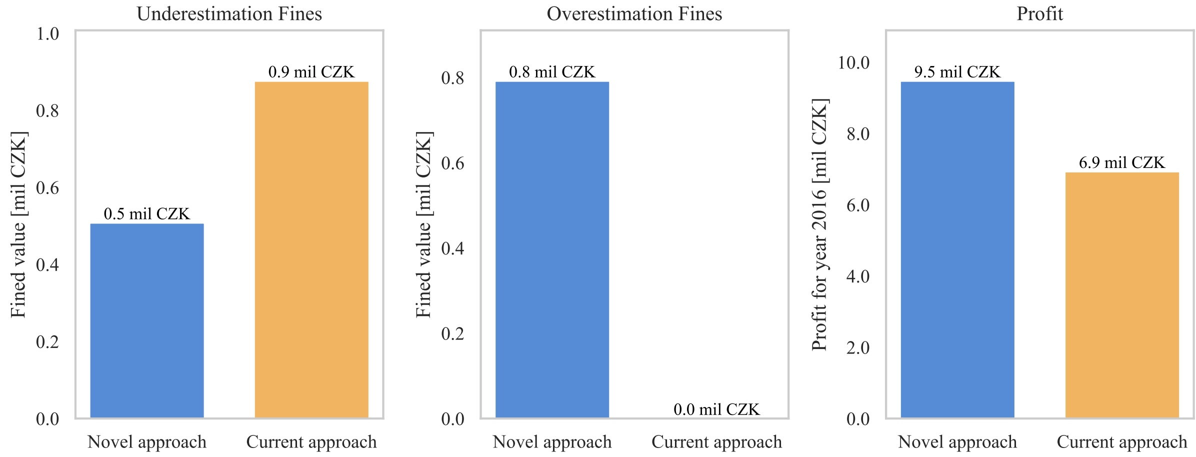

Figure 3.38: Fines and profits for both novel and current approach.

Despite the lower success rate in December, the novel model still outperformed the current strategy. Since bypass usage was most prominent in colder months (Figure 3.7), the largest profit increase was observed outside Q3 (Figure 3.39).

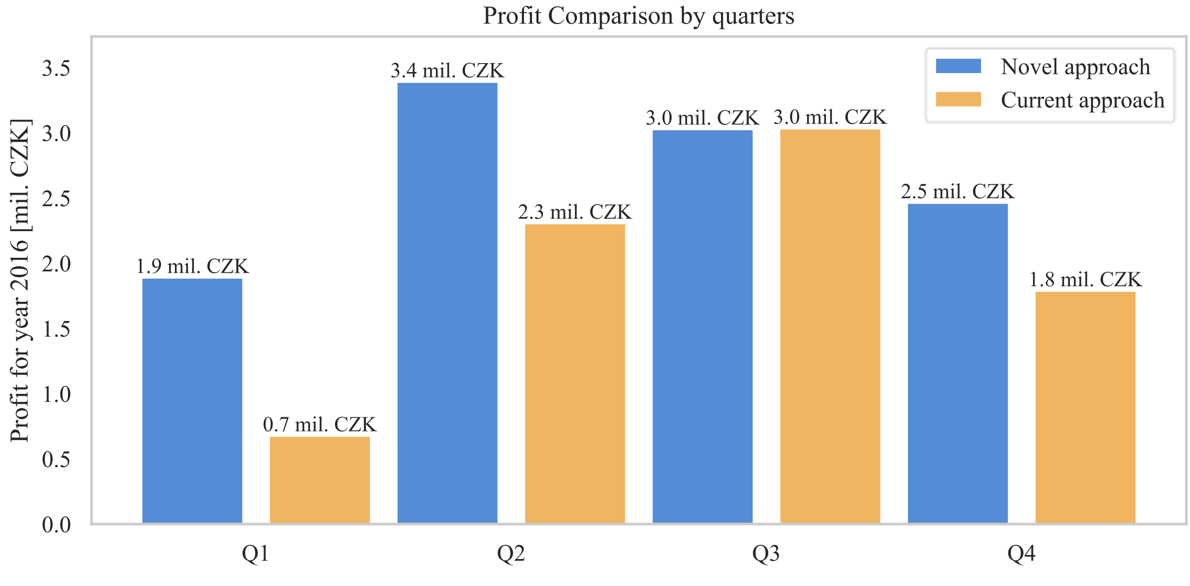

Figure 3.39: Quarterly profits for both novel and current approach.

Chapter 4: Conclusion

The objective of this thesis was to construct a machine learning model that forecasts the combined heat and power production of Waste-to-Energy plants, employing weather, temporal data, and production history as inputs. The Main task for this model is total power production forecast with a high degree of certainty, as if the production value does not fall within ±0.5 MWh of plan, plant faces severe fines. The plan should strive to be less risk averse while still economically viable.

For this purpose, we have chosen three popular machine learning algorithms of varying complexity – Linear regression (LR), light gradient boosting machine (LGBM) and artificial neural network (ANN) representing low, mid, and high levels of complexity, respectively. The model itself is made of 7 sub-models, going in sequence from steam generated in boilers to power transferred to grid. Excluding turbine live steam extraction (which is grey box algorithm), all the other variables were predicted using the three aforementioned algorithms. We used various metrics (R2 score, MAE, interpretability...) to select the best model for each variable.

The residuals analysis revealed that the prediction errors of total steam production attributes to 80% of the mean absolute error in power transferred. The goal of constructing a plan with a high success rate (where success is defined as not overestimating power production by more than 0.5 MWh) led to the implementation of quantile models with a parameter of q = 0.05, which aimed for a 5% overestimation of steam produced. Additionally, to enhance the predictive utility of auxiliary streams, including external consumption, blowoff, deaerator, and preheating steam consumption, we employed a correlation matrix-based cascading of predictions.

To account for larger turbine inflow resulting from the minimization of bypass usage, the power generation data was extrapolated. This tested the model on data exhibiting a different distribution, as frequent bypass usage was observed in the training data. Subsequently, a verification process was conducted to ensure that the models did not fail when extrapolating, a known concern for data-driven models. During this process, the LGBM model was found to exhibit signs of failure during extrapolation, an outcome anticipated for tree-based algorithms.

Applying the model to data from the year 2016, we achieved a success rate of 95%. Of these successful cases, 43% fell within the tolerance field, with the remainder slightly underestimating it. In contrast, the current strategy yielded a 100% success rate, but only 34% of these cases were within the tolerance interval. The results showed an overall increase in profit for our model by 2.6 million CZK. Notably our model outperformed the current strategy in every fiscal quarter, generating 13% more power for the grid while operating at 89% of its potential maximum.

Future work should concentrate on further optimization. Our current approach primarily targets bypass minimization, not taking into account factors such as heat prices or contract availability. To advance in this direction, we require higher-quality data and more comprehensive information about contracts. Additionally, if renewable energy sources like solar power were to be included, or if a longer timeframe were to be forecasted, the implementation of market price forecasting algorithms could prove beneficial.

Bibliography

[1] KAZDA, Kody a Xiang LI. *A critical review of the modeling and optimization of combined heat and power dispatch* [online]. 2020. ISSN 22279717. Dostupné z: doi:10.3390/PR8040441

[2] TOUŠ, Michal, Martin PAVLAS, Ondřej PUTNA, Petr STEHLÍK a Lukáš CRHA. Combined heat and power production planning in a waste-to-energy plant on a short-term basis. *Energy* [online]. 2015, **90**. ISSN 03605442. Dostupné z: doi:10.1016/j.energy.2015.05.077

[3] BIRGEN, Cansu, Elisa MAGNANELLI, Per CARLSSON, Øyvind SKREIBERG, Jostein MOSBY a Michaël BECIDAN. Machine learning based modelling for lower heating value prediction of municipal solid waste. *Fuel* [online]. 2021, **283**. ISSN 00162361. Dostupné z: doi:10.1016/j.fuel.2020.118906

[4] HACKNEY, Rick, Theoklis NIKOLAIDIS a Alvise PELLEGRINI. A method for modelling compressor bleed in gas turbine analysis software. *Applied Thermal Engineering* [online]. 2020, **172**. ISSN 13594311. Dostupné z: doi:10.1016/j.applthermaleng.2020.115087

[5] GÉRON, Aurélien. *Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems*. 2017.

[6] MITCHELL, Tom. *Machine Learning. New York: McGraw Hill*. 1997. ISSN 0885-6125.

[7] PEDREGOSA, F, G VAROQUAUX, A GRAMFORT, B MICHEL V. AND THIRION, O GRISEL, M BLONDEL, R PRETTENHOFER P. AND WEISS, V DUBOURG, J VANDERPLAS, A PASSOS, D COURNAPEAU, M BRUCHER, M PERROT a E DUCHESNAY. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011, **12**, 2825–2830.

[8] FISHER, R. A. THE USE OF MULTIPLE MEASUREMENTS IN TAXONOMIC PROBLEMS. *Annals of Eugenics* [online]. 1936, **7**(2). Dostupné z: doi:10.1111/j.1469-1809.1936.tb02137.x

[9] BRYANT, Peter a Marlene A SMITH. Practical Data Analysis: Case Studies in Business Statistics. In: . 1994.

[10] RASCHKA, Sebastian a Vahid MIRJALILI. *Python Machine Learning: Machine Learning & Deep Learning with Python, Scikit-Learn and TensorFlow 2, Third Edition*. 2019.

[11] PEDREGOSA, F, G VAROQUAUX, A GRAMFORT, B MICHEL V. AND THIRION, O GRISEL, M BLONDEL, R PRETTENHOFER P. AND WEISS, V DUBOURG, J VANDERPLAS, A PASSOS, D COURNAPEAU, M BRUCHER, M PERROT a E DUCHESNAY. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011, **12**, 2825–2830.

[12] KE, Guolin, Qi MENG, Thomas FINLEY, Taifeng WANG, Wei CHEN, Weidong MA, Qiwei YE a Tie-Yan LIU. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*. 2017, **30**, 3146–3154.

[13] IPCC. IPCC - The fifth Assessment Report. Working Group 3 - Mitigation of Climate Change. *Chapter 7 - Energy Systems*. 2014.

[14] HIGHTOWER, Mike a Suzanne A. PIERCE. *The energy challenge* [online]. 2008. ISSN 14764687. Dostupné z: doi:10.1038/452285a

[15] FOROOTAN, Mohammad Mahdi, Iman LARKI, Rahim ZAHEDI a Abolfazl AHMADI. *Machine Learning and Deep Learning in Energy Systems: A Review* [online]. 2022. ISSN 20711050. Dostupné z: doi:10.3390/su14084832

[16] ILBEIGI, Marjan, Mohammad GHOMEISHI a Ali DEHGHANBANADAKI. Prediction and optimization of energy consumption in an office building using artificial neural network and a genetic algorithm. *Sustainable Cities and Society* [online]. 2020, **61**. ISSN 22106707. Dostupné z: doi:10.1016/j.scs.2020.102325

[17] EPA. U . S . Environmental Protection Agency Combined Heat and Power Partnership. *Environmental Protection*. 2017, (September).

[18] WEN, Hao, Song SANG, Chenhui QIU, Xiangrui DU, Xiao ZHU a Qian SHI. A new optimization method of wind turbine airfoil performance based on Bessel equation and GABP artificial neural network. *Energy* [online]. 2019, **187**. ISSN 03605442. Dostupné z: doi:10.1016/j.energy.2019.116106

[19] NTAKOLIA, Charis, Athanasios ANAGNOSTIS, Serafeim MOUSTAKIDIS a Nikolaos KARCANIAS. *Machine learning applied on the district heating and cooling sector: a review* [online]. 2022. ISSN 18683975. Dostupné z: doi:10.1007/s12667-020-00405-9

[20] TENG, Teng, Xin ZHANG, Han DONG a Qicheng XUE. A comprehensive review of energy management optimization strategies for fuel cell passenger vehicle. *International Journal of Hydrogen Energy* [online]. 2020, **45**(39). ISSN 03603199. Dostupné z: doi:10.1016/j.ijhydene.2019.12.202

[21] OMIRGALIYEV, Ruslan, Nurkhat ZHAKIYEV, Ideyat BAPIYEV a Aldiyar SALKENOV. Industrial Application of Machine Learning Clustering for a Combined Heat and Power Plant: A Pavlodar Case Study. In: *UkrMiCo 2021 - 2021 IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics, Proceedings* [online]. 2021. Dostupné z: doi:10.1109/UkrMiCo52950.2021.9716605

[22] ZHOU, Yuekuan, Siqian ZHENG a Guoqiang ZHANG. Artificial neural network based multivariable optimization of a hybrid system integrated with phase change materials, active cooling and hybrid ventilations. *Energy Conversion and Management* [online]. 2019, **197**. ISSN 01968904. Dostupné z: doi:10.1016/j.enconman.2019.111859

[23] KABUGO, James Clovis, Sirkka Liisa JÄMSÄ-JOUNELA, Robert SCHIEMANN a Christian BINDER. Industry 4.0 based process data analytics platform: A waste-to-energy plant case study. *International Journal of Electrical Power & Energy Systems* [online]. 2020, **115**, 105508 [vid. 2023-05-13]. ISSN 0142-0615. Dostupné z: doi:10.1016/J.IJEPES.2019.105508

[24] AMASYALI, Kadir a Nora M. EL-GOHARY. *A review of data-driven building energy consumption prediction studies* [online]. 2018. ISSN 18790690. Dostupné z: doi:10.1016/j.rser.2017.04.095

[25] KLINGHOFFER, N B, N J THEMELIS a M J CASTALDI. 1 - Waste to energy (WTE): an introduction. In: Naomi B KLINGHOFFER a Marco J CASTALDI, ed. *Waste to Energy Conversion Technology* [online]. B.m.: Woodhead Publishing, 2013, Woodhead Publishing Series in Energy, s. 3–14. ISBN 978-0-85709-011-9. Dostupné z: doi:https://doi.org/10.1533/9780857096364.1.3

[26] WALKER, Shalika, Waqas KHAN, Katarina KATIC, Wim MAASSEN a Wim ZEILER. Accuracy of different machine learning algorithms and added-value of predicting aggregated-level energy performance of commercial buildings. *Energy and*

Buildings [online]. 2020, **209**. ISSN 03787788. Dostupné z: doi:10.1016/j.enbuild.2019.109705

[27] TOBIASEN, L. a B. KAMUK. *Waste to Energy Conversion Technology*. 2013.

[28] KAZEMZADEH, Mohammad Rasool, Ali AMJADIAN a Turaj AMRAEE. A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting. *Energy* [online]. 2020, **204**. ISSN 03605442. Dostupné z: doi:10.1016/j.energy.2020.117948

[29] SHARMA, Surbhi, Soumen BASU, Nagaraj P. SHETTI, Mohammadreza KAMALI, Pavan WALVEKAR a Tejraj M. AMINABHAVI. *Waste-to-energy nexus: A sustainable development* [online]. 2020. ISSN 18736424. Dostupné z: doi:10.1016/j.envpol.2020.115501

[30] GHODRATI, Armin, Rahim ZAHEDI a Abolfazl AHMADI. Analysis of cold thermal energy storage using phase change materials in freezers. *Journal of Energy Storage* [online]. 2022, **51**. ISSN 2352152X. Dostupné z: doi:10.1016/j.est.2022.104433

[31] LIANG, Rui, Chao CHEN, Akash KUMAR, Junyu TAO, Yan KANG, Dong HAN, Xianjia JIANG, Pei TANG, Beibei YAN a Guanyi CHEN. *State-of-the-art applications of machine learning in the life cycle of solid waste management* [online]. 2023. ISSN 2095221X. Dostupné z: doi:10.1007/s11783-023-1644-x

[32] BOT, K., A. RUANO a M. G. RUANO. Forecasting electricity demand in households using Moga-designed artificial neural networks. In: *IFAC-PapersOnLine* [online]. 2020. ISSN 24058963. Dostupné z: doi:10.1016/j.ifacol.2020.12.1985

[33] TAKI, Morteza a Abbas ROHANI. Machine learning models for prediction the Higher Heating Value (HHV) of Municipal Solid Waste (MSW) for waste-to-energy evaluation. *Case Studies in Thermal Engineering* [online]. 2022, **31**, 101823 [vid. 2023-05-13]. ISSN 2214-157X. Dostupné z: doi:10.1016/J.CSITE.2022.101823

[34] KOUJOK, Mohamed El, Ahmed RAGAB a Mouloud AMAZOUZ. A Multi-Agent Approach Based on Machine-Learning for Fault Diagnosis. In: *IFAC-PapersOnLine* [online]. 2019. ISSN 24058963. Dostupné z: doi:10.1016/j.ifacol.2019.10.007

[35] BIRGEN, Cansu, Elisa MAGNANELLI, Per CARLSSON, Øyvind SKREIBERG, Jostein MOSBY a Michaël BECIDAN. Machine learning based modelling for lower heating value prediction of municipal solid waste. *Fuel* [online]. 2021, **283**, 118906 [vid. 2023-05-13]. ISSN 0016-2361. Dostupné z: doi:10.1016/J.FUEL.2020.118906

[36] DELEPLACE, A., V. ATAMURADOV, A. ALLALI, J. PELLE, R. PLANA a G. ALLEAUME. Ensemble learning-based fault detection in nuclear power plant screen cleaners. In: *IFAC-PapersOnLine* [online]. 2020. ISSN 24058963. Dostupné z: doi:10.1016/j.ifacol.2020.12.2773

[37] PEHLKEN, Alexandra, Henriette GARMATTER, Lisa DAWEL, Fabian CYRIS, Hendrik BECK, Fenja SCHWARK, Roland SCHARF a Astrid NIEBE. How can machine learning improve waste-to-energy plant operation. In: [online]. 2023. Dostupné z: doi:10.1109/ice/itmc-iamot55089.2022.10033210

[38] ZHANG, Wei, Xiang LI, Hui MA, Zhong LUO a Xu LI. Federated learning for machinery fault diagnosis with dynamic validation and self-supervision. *Knowledge-Based Systems* [online]. 2021, **213**. ISSN 09507051. Dostupné z: doi:10.1016/j.knosys.2020.106679

[39] GHOLIZADEH, Hadi, Hamed FAZLOLLAHTABAR, Amir M. FATHOLLAHI-FARD a Maxim A. DULEBENETS. Preventive maintenance for the flexible flowshop scheduling under uncertainty: a waste-to-energy system. *Environmental Science and*

Pollution Research [online]. 2021. ISSN 16147499. Dostupné z: doi:10.1007/s11356-021-16234-x

[40] PANG, Yu, Limin JIA, Xuejia ZHANG, Zhan LIU a Dazi LI. Design and implementation of automatic fault diagnosis system for wind turbine. *Computers and Electrical Engineering* [online]. 2020, **87**. ISSN 00457906. Dostupné z: doi:10.1016/j.compeleceng.2020.106754

[41] HU, Chenlian, Xiao LIU, Jie LU a Chi Hwa WANG. Operations scheduling of waste-to-energy plants under uncertainty. *Journal of Cleaner Production* [online]. 2020, **253**. ISSN 09596526. Dostupné z: doi:10.1016/j.jclepro.2019.119953

[42] LABRADOR RIVAS, Angel Esteban a Taufik ABRÃO. Faults in smart grid systems: Monitoring, detection and classification. *Electric Power Systems Research* [online]. 2020, **189**. ISSN 03787796. Dostupné z: doi:10.1016/j.epsr.2020.106602

[43] TENG, Sin Yong, Vítězslav MÁŠA, Michal TOUŠ, Marek VONDRA, Hon Loong LAM a Petr STEHLÍK. Waste-to-energy forecasting and real-time optimization: An anomaly-aware approach. *Renewable Energy* [online]. 2022, **181**. ISSN 18790682. Dostupné z: doi:10.1016/j.renene.2021.09.026

[44] HAN, Hua, Zhan ZHANG, Xiaoyu CUI a Qinghong MENG. Ensemble learning with member optimization for fault diagnosis of a building energy system. *Energy and Buildings* [online]. 2020, **226**. ISSN 03787788. Dostupné z: doi:10.1016/j.enbuild.2020.110351

[45] SEABOLD, Skipper a Josef PERKTOLD. statsmodels: Econometric and statistical modeling with python. In: *9th Python in Science Conference*. 2010.

[46] ESKANDARI, Aref, Jafar MILIMONFARED a Mohammadreza AGHAEI. Line-line fault detection and classification for photovoltaic systems using ensemble learning model based on I-V characteristics. *Solar Energy* [online]. 2020, **211**. ISSN 0038092X. Dostupné z: doi:10.1016/j.solener.2020.09.071

[47] W.TRASK, Andrew. *grokking Deep Learning*. 2016.

[48] FAN, Junliang, Lifeng WU, Fucang ZHANG, Huanjie CAI, Wenzhi ZENG, Xiukang WANG a Haiyang ZOU. Empirical and machine learning models for predicting daily global solar radiation from sunshine duration: A review and case study in China. *Renewable and Sustainable Energy Reviews* [online]. 2019, **100**. ISSN 18790690. Dostupné z: doi:10.1016/j.rser.2018.10.018

[49] KARPATHY, Andrej. Stanford university cs231n: Convolutional neural networks for visual recognition. *URL: http://cs231n. stanford. edu/syllabus. html*. 2018.

[50] VOYANT, Cyril, Gilles NOTTON, Soteris KALOGIROU, Marie Laure NIVET, Christophe PAOLI, Fabrice MOTTE a Alexis FOUILLOY. *Machine learning methods for solar radiation forecasting: A review* [online]. 2017. ISSN 18790682. Dostupné z: doi:10.1016/j.renene.2016.12.095

[51] *OTE, a.s.* [online]. Praha: OTE, a.s. 2023. Dostupné z: https://www.ote-cr.cz/cs/kratkodobe-trhy/elektrina/

[52] ALIZAMIR, Meysam, Sungwon KIM, Ozgur KISI a Mohammad ZOUNEMAT-KERMANI. A comparative study of several machine learning based non-linear regression methods in estimating solar radiation: Case studies of the USA and Turkey regions. *Energy* [online]. 2020, **197**. ISSN 03605442. Dostupné z: doi:10.1016/j.energy.2020.117239

[53] ZENDEHBOUDI, Alireza, M. A. BASEER a R. SAIDUR. *Application of support vector machine models for forecasting solar and wind energy resources: A review* [online]. 2018. ISSN 09596526. Dostupné z: doi:10.1016/j.jclepro.2018.07.164

[54] DEMOLLI, H., Ahmet Sakir DOKUZ, A. ECEMIS a M. GOKCEK. Wind power forecasting based on daily wind speed data using machine learning algorithms. *Energy Conversion and Management* [online]. 2019, **198**. ISSN 01968904. Dostupné z: doi:10.1016/j.enconman.2019.111823

List of Abbreviations

Abbreviation	Meaning
CHP	Combined Heat and Power
WtE	Waste to Energy
ML	Machine Learning
LHV	Lower Heating Value
AI	Artificial Intelligence
MSE	Mean Squared Error
MAE	Mean Absolute Error
R^2	R-squared
SME	Signed Mean Error
SVR	Support Vector Regression
MA & ES	Moving Average and Exponential Smoothing
MLR	Multiple Linear Regression
DT	Decision Trees
GLR	Generalized Linear Regression
RF	Random Forest
GB	Gradient Boosting
HM	Harmony Search
ANN	Artificial Neural Networks
NN	Neural Networks
DL	Deep Learning
RL	Reinforcement Learning
FDD	Fault and Defect Detection
SHP	Separate Heat and Power
MSW	Municipal Solid Waste
HTM	Hierarchical Temporal Memory
KDE	Kernel Density Estimation

List of Figures

List of Tables

List of Attachments