Technical University of Košice

Faculty of Mining, Ecology, Process Control
and Geotechnologies

Advanced prediction models for sales forecasting

Master thesis

2023 **Bc. Aleš Jandera**

Technical University of Košice

Faculty of Mining, Ecology, Process Control and Geotechnologies

Advanced prediction models for sales forecasting

Master thesis

Study Programme:	Process control of raw materials and material extraction and processing
Field of Study:	Cybernetics
Department:	Faculty of Mining, Ecology, Process Control and Geotechnologies (FBERG)
Supervisor:	doc. Ing. Tomáš Škovránek, PhD.

Košice 2023

Bc. Aleš Jandera

Abstract in English

Sales forecasting can be divided into two main categories: short-term and long-term forecasting. Short-term forecasting is generally done on a weekly or monthly basis. Long-term forecasting is done on a quarterly or annual basis. There are many different methods that can be used for sales forecasting. The most common method is trend analysis. Trend analysis looks at past sales data to identify patterns and trends that can be used to predict future sales. Other methods include regression analysis and time series analysis. Advance prediction modeling is a type of long-term forecasting. It uses historical data and statistical techniques to predict future sales. Advance prediction modeling is often used by companies to make strategic decisions about inventory, pricing, and marketing.

Keywords in English

Mathematic modeling, forecasting, linear prediction

Abstract in Slovak

Prognózy predaja možno rozdeliť do dvoch hlavných kategórií: krátkodobá a dlhodobá prognóza. Krátkodobé prognózy sa vo všeobecnosti vykonávajú týždenne alebo mesačne. Dlhodobé prognózy sa robia štvrťročne alebo ročne. Existuje mnoho rôznych metód, ktoré možno použiť na predpovedanie predaja. Najbežnejšou metódou je analýza trendov. Analýza trendov sa zameriava na údaje o minulých predajoch, aby identifikovala vzory a trendy, ktoré možno použiť na predpovedanie budúceho predaja. Ďalšie metódy zahŕňajú regresnú analýzu a analýzu časových radov. Predbežné predikčné modelovanie je typ dlhodobého predpovedania. Na predpovedanie budúceho predaja využíva historické údaje a štatistické techniky. Pokročilé predikčné modelovanie často používajú spoločnosti na strategické rozhodnutia o zásobách, cenách a marketingu.

Keywords in Slovak

Matematicke modelovanie, predpoved, linearna predikcia

Bibliographic Citation

JANDERA, Aleš. *Advanced prediction models for sales forecasting*. Košice: Technical University of Košice, Faculty of Mining, Ecology, Process Control and Geotechnologies, 2023. 19s. Supervisor: doc. Ing. Tomáš Škovránek, PhD.

71130

TECHNICKÁ UNIVERZITA V KOŠICIACH
FAKULTA BANÍCTVA, EKOLÓGIE, RIADENIA A GEOTECHNOLÓGIÍ
Ústav riadenia a informatizácie výrobných procesov

Z A D A N I E
D I P L O M O V E J P R Á C E

Študijný odbor: **Kybernetika**
Študijný program: **Riadenie procesov získavania a spracovania surovín**

Názov práce:

Pokročilé predikčné modely pre prognózu predaja

Advanced prediction models for sales forecasting

Študent: **Bc. Aleš Jandera**
Školiteľ: **doc. Ing. Tomáš Škovránek, PhD.**
Školiace pracovisko: **Ústav riadenia a informatizácie výrobných procesov**
Konzultant práce:
Pracovisko konzultanta:

Pokyny na vypracovanie diplomovej práce:
1. Úvod
2. Analýza súčasného stavu
3. Návrhová časť
4. Zhodnotenie prínosu práce a návrh odporúčaní pre ďalší postup
5. Záver

Odporúčaný rozsah práce: 50 - 70 strán

Jazyk, v ktorom sa práca vypracuje: anglický
Termín pre odovzdanie práce: 21.04.2023
Dátum zadania diplomovej práce: 31.10.2022

....
Dr. h. c. prof. Ing. Michal Cehlár, PhD.
dekan fakulty

Declaration

I hereby declare that this thesis is my own work and effort. Where other sources of information have been used, they have been acknowledged.

Košice, 21.4.2023 . . . .

Signature

Acknowledgement

Very strong thanks to whole teaching staff at the Institute of Control and Informatization of Production Processes for their patience, leadership and knowledge which helped me to write this thesis. I would like to express my deepest appreciation to my supervisor Tomáš Škovránek for his time and advices.

Contents

Introduction

Linear prediction [1] is a method used in signal processing to predict future values of a time series based on past observations. The technique is based on the assumption that the signal can be modeled as a linear combination of past values and a noise term. Long-term linear prediction refers to the application of this method to predict values over a longer period of time, such as months or years. It requires a greater amount of data and is more complex than short-term prediction, but can be useful in areas such as stock market forecasting and weather prediction. The goal of this master thesis is to developing new algorithms and mathematical models to improve the accuracy of long-term predictions in sales forecasting. Matlab [1] livescript [2] will be used as development enviroment.

Task formulation

Proposed a mathematical model and alghoritms for sales forcasting based on long-term prediction with improved Levinson - Durbin scheme which should have better performance and acuraccy than known linear prediction mechanism.

[1]MATLAB is a fourth-generation programminglanguage and numerical analysis environment. Uses for MATLAB include matrix calculations, developing and running algorithms, creating user interfaces (UI) and data visualization.

1 Analytical part

1.1 Introduction

Linear prediction is a statistical method used to predict future values based on historical data. The Durbin-Levinson algorithm is a method for solving the linear prediction problem for autoregressive (AR) models[1], which are models where the current output depends on previous outputs. The algorithm solves the linear prediction problem by finding the coefficients of the AR model that minimize the prediction error. The resulting AR coefficients can be used to make predictions about future values based on past observations. This method should be used with using the pattern of linear relationship between the independent and dependent variables. Here's a basic outline of the steps involved in using linear prediction to forecast sales data:

1. Collect sales data: Gather the historical sales data for the product or service that you want to forecast.
2. Plot the data: Plot the sales data over time to visually inspect the trend and identify any patterns.
3. Choose a model: Select an appropriate linear model to represent the relationship between the independent and dependent variables in the data. For example, you might choose a simple linear regression model.
4. Train the model: Train the selected model on the historical sales data using a method such as least squares regression.

[1]Autoregressive (AR) models are time series models that describe the relationship between the current value of a variable and its past values. In an autoregressive model, each observation is modeled as a linear combination of past observations, with weights called AR coefficients. AR models are widely used in various fields such as economics, engineering, and finance for modeling and forecasting time series data. The order of the AR model, denoted as "p", refers to the number of past values used to predict the current value. For example, an AR(1) model uses only the previous observation to predict the current value, while an AR(2) model uses the previous two observations.

5. Make predictions: Use the trained model to make predictions on future sales data. You may want to generate predictions for several months or years in advance.

6. Evaluate the model: Assess the accuracy of the predictions by comparing them to the actual sales data. Use metrics such as mean absolute error or root mean squared error to quantify the model's performance.

7. Refine the model: If necessary, refine the model by adding additional independent variables or transforming the existing variables. Repeat the training and evaluation steps until you have a model that provides accurate forecasts.

1.2 Levinson-Durbin sheme

The Levinson-Durbin algorithm [3] is an iterative numerical method used to solve the autoregressive (AR) model of a time series. AR models are used to model and forecast time series data [4], such as sales data, by assuming that each future value of the series depends on a linear combination of previous values.
The Levinson-Durbin algorithm solves the AR model by iteratively updating the coefficients of the model to minimize the prediction error between the model and the actual data. The algorithm is fast and efficient, and it is widely used in digital signal processing, speech processing, and control systems, among other applications.
The Levinson-Durbin algorithm is often used as an alternative to the Yule-Walker equations, which are another commonly used method for solving AR models. Unlike the Yule-Walker equations, the Levinson-Durbin algorithm can be easily modified to handle non-stationary time series data, and it is also more robust to numerical issues such as rounding errors.

1.3 Linear prediction

Linear prediction is a statistical technique used to forecast future values based on past observations. It is a method for modeling the relationship between a dependent variable and one or more independent variables in a linear form. The goal of linear prediction is to find the best linear approximation of the relationship between the variables, which can then be used to make predictions about future values of the dependent variable.

Linear prediction can be performed using simple linear regression or multiple linear regression, depending on the number of independent variables involved [5]. In simple linear regression, a single independent variable is used to predict the value of the dependent variable, while in multiple linear regression, multiple independent variables are used to make the prediction.
Linear prediction models are commonly used in finance, economics, and engineering, among other fields, to forecast future values of time series data, such as stock prices, sales, or demand. The accuracy of linear prediction models depends on several factors, including the quality of the data, the choice of independent variables, and the degree of linearity in the relationship between the variables.

1.3.1 Short-term linear prediction

Short-term linear prediction [6] refers to the use of linear prediction techniques to make predictions about the near-term future values of a time series. It is used to forecast future values of a dependent variable based on its past values and any relevant independent variables.
In short-term linear prediction, the focus is on accurately predicting the next few values of the dependent variable, typically in the range of several weeks to a few months. The linear prediction models used for short-term forecasting are typically simple and straightforward, often using a small number of independent variables. The goal is to provide a quick and easily interpretable forecast that can be used to make operational decisions in the short-term.
Common techniques for short-term linear prediction include moving average models, exponential smoothing, and autoregressive models. These methods use the historical data of the time series to model the relationship between the dependent and independent variables, and to make predictions about future values. The accuracy of short-term linear predictions can be evaluated using metrics such as mean absolute error, mean squared error, or the correlation coefficient between the actual and predicted values.

1.3.2 Long-term linear prediction

Long-term linear prediction refers to the use of linear prediction techniques to make predictions about the future values of a time series over an extended period of time, typically several months to several years. Unlike short-term linear prediction, which focuses on forecasting the near-term future, long-term linear prediction aims to provide a more comprehensive and accurate forecast of future

values.
In long-term linear prediction, more sophisticated models [7] are typically used, such as multiple linear regression or time series models, and a larger number of independent variables may be considered. The models are also trained on a larger historical dataset to ensure that they capture any long-term trends or patterns in the data.
Long-term linear prediction is commonly used in fields such as finance, economics, and marketing, to make long-term projections about variables such as sales, demand, or stock prices. The goal is to provide a comprehensive and accurate forecast that can be used to make strategic decisions in the long-term. The accuracy of long-term linear predictions can be evaluated using the same metrics as for short-term linear predictions [8], as well as additional metrics such as mean absolute percentage error or mean absolute scaled error.
In some long-term prediction use cases it needs to solve the suppression of Late Reverberation Effect[2] this can be done by with minimal performance degradation by framework developed by Keisuke Kinoshita [9] for both single-channel and multichannel scenarios.

1.4 Models used for sales data

There are several mathematical models used for sales prediction, including:

1. Time series models: These models are used to analyze and forecast sales data over time, such as seasonal patterns, trends, and fluctuations. Examples include ARIMA (AutoRegressive Integrated Moving Average), SARIMA (Seasonal ARIMA), and exponential smoothing.
2. Regression models: These models use historical data to determine the relationship between sales and one or more independent variables, such as price, promotion, and advertising. Examples include linear regression, logistic regression, and multiple regression.
3. Decision tree models: These models use a tree-like structure to make decisions based on the relationship between sales and multiple independent

[2]Late Reverberation Effect refers to the decay of sound in an environment after the initial sound source has stopped. This effect results in the persistence of sound in a space for a short period of time and helps create the characteristic ambiance of a room or space. It is an important aspect of room acoustics and is used in sound design and music production to enhance the perceived sound quality and spatial experience of audio.

variables. Examples include CART (Classification and Regression Tree) and Random Forest.

4. Machine learning models: These models use algorithms such as neural networks and support vector machines to make predictions based on patterns in the data.

The choice of mathematical model depends on the characteristics of the data, the desired level of accuracy, and the computational resources available.

1.4.1 Time-series models

Time-series models are mathematical models used to analyze and forecast data that are collected over time [10]. These models are used to study and make predictions about the trends, patterns, and behavior of the data over time, taking into account historical values and their relationship with the present. Time-series models are widely used in areas such as economics, finance, and weather forecasting, among others. The models are based on various statistical techniques, including ARIMA (AutoRegressive Integrated Moving Average), SARIMA (Seasonal ARIMA), and exponential smoothing, among others. The goal of time-series modeling is to build a mathematical representation of the underlying process that generates the time-series data, allowing for accurate prediction of future values.

1.4.2 Regression models

Regression models are a type of statistical models used to examine the relationship between a dependent variable and one or more independent variables [11]. The goal of regression analysis is to model the relationship between these variables and make predictions about the dependent variable based on the values of the independent variables. Regression models are widely used in many fields, including economics, finance, marketing, and social sciences, to make predictions and understand the relationship between variables.
There are several types of regression models, including:

1. Linear regression: a simple regression model where the relationship between the dependent and independent variables is modeled using a linear equation.

2. Logistic regression: used for binary classification problems where the dependent variable is binary and the goal is to model the relationship between the independent variables and the probability of the dependent variable being either 0 or 1.

3. Multiple regression: used when there are multiple independent variables and the goal is to model the relationship between all of these variables and the dependent variable.

4. Polynomial regression: used when the relationship between the dependent and independent variables is non-linear and can be modeled using a polynomial equation.

The choice of regression model depends on the nature of the data and the research question being asked.

1.4.3 Decision tree models

Decision tree models are a type of machine learning models used for both regression and classification tasks [12]. They are tree-like structures that make predictions by breaking down a dataset into smaller and smaller subsets, based on the values of the input variables. At each internal node of the tree, a decision rule is used to split the data based on the value of a feature, and the process continues until the data are separated into homogeneous groups, or leaves. The predictions are then made based on the average or majority class in each leaf node.
Decision trees have several advantages, including ease of interpretability, handling of non-linear relationships, and ability to handle both categorical and numerical data. Some examples of decision tree algorithms are CART (Classification and Regression Tree) and Random Forest.
The decision tree model is trained using a dataset, and the tree structure is built using a greedy algorithm that seeks to maximize the reduction in impurity of the target variable at each split. The model can then be used to make predictions on new data by following the decision rules in the tree.

1.4.4 Machine learning models

Machine learning models are a subset of artificial intelligence that allows computers to learn and make predictions or decisions without being explicitly programmed. Machine learning models are based on algorithms that use statistical methods to find patterns in data and make predictions about new, unseen data. There are several types of machine learning models, including:

1. Supervised learning: where the model is trained on labeled data, with the goal of learning the relationship between the input features and the target variable, and making predictions about the target variable for new, unseen data.

2. Unsupervised learning: where the model is trained on unlabeled data, with the goal of finding patterns or structure in the data, such as clustering or dimensionality reduction.

3. Reinforcement learning: where the model learns by receiving rewards or penalties for its actions in an environment, with the goal of maximizing the reward over time.

4. Deep learning: a subset of machine learning that uses artificial neural networks with multiple hidden layers to model complex relationships in the data.

The choice of machine learning model depends on the problem being solved and the type of data being used. Machine learning models have been applied to a wide range of tasks, including image and speech recognition, natural language processing, and predictive modeling. Advances in machine learning (ML), faster processors and the availability of digitized healthcare data have contributed to a growing number of papers describing ML applications in healthcare. [13]

2 Goal of the thesis

In this thesis we are focus on improving linear prediction in sales and financial forecasting and combine it with modern machine learning approaches. Machine learning techniques are used to improve the accuracy of linear prediction models. Specifically, there are two ways in which machine learning are applied to linear prediction:

1. Feature engineering: In this approach, machine learning is used to extract relevant features from the input signal that can be used as input to a linear prediction model. The extracted features can capture complex patterns in the data that are not captured by the raw input. Feature engineering can be done using techniques such as principal component analysis, wavelet transform, and Fourier transform.

2. Model selection and training: In this approach, machine learning is used to select the best linear model for the prediction task and to estimate its parameters from the data. This can involve selecting the best set of input variables for the linear model, choosing the best regularization parameter to avoid overfitting, and optimizing the model hyperparameters. Common machine learning algorithms used for linear prediction include linear regression, support vector regression, and artificial neural networks.

The goal of long-term linear prediction is to estimate future values of a signal or time series based on its past values using a linear model. The linear prediction model uses a set of coefficients to weight past values of the signal and produce a prediction for future values. The accuracy of the prediction depends on the quality of the model and the complexity of the underlying signal. Combine this principles get to us the best results from linear prediction over sales and financial datasets.

3 Methodology

3.1 Characteristics of the research object

Sales data and financial datasets are two broad categories of data that have different characteristics and are used for different purposes. Here are some general characteristics of each type of dataset:
Sales data
Typically contains transactional information, such as the date, time, location, and amount of a purchase Can include additional information about the customer, such as their demographic profile, purchase history, and preferences Often analyzed to understand customer behavior, such as buying patterns, trends, and preferences May have seasonal or cyclical patterns, depending on the nature of the product or service being sold Can be used to optimize marketing and sales strategies, such as targeting specific customer segments, promoting certain products, or adjusting prices and discounts
Financial datasets
Typically contains financial information, such as the revenue, expenses, assets, liabilities, and cash flow of a company or organization Can include additional information about the market, such as interest rates, exchange rates, and stock prices Often analyzed to evaluate the financial performance and health of a company, such as profitability, solvency, and liquidity May have regulatory or compliance requirements, such as financial reporting standards or tax laws Can be used to make strategic decisions, such as investment, merger and acquisition, or divestiture Both sales data and financial datasets can be used for forecasting and modeling, but they have different analytical techniques and tools. Sales data often requires customer segmentation, predictive analytics, and machine learning algorithms, while financial datasets require financial ratio analysis, time series forecasting, and risk assessment.

3.2 Methods

3.2.1 Linear regression

Linear regression [**linear**] attempts to model the relationship between two variables by fitting a linear equation to observed data. One variable is considered to be an explanatory variable, and the other is considered to be a dependent variable. For example, we want to relate the weights of individuals to their heights using a linear regression model. Before attempting to fit a linear model to observed data, we should first determine if there exists a relationship between the variables of interest. This does not necessarily mean that one variable causes the other, but that there is some significant association between them. To determine the strength of relationship a scatterplot can be a helpful tool. If there appears to be no association between the proposed explanatory and dependent variables, then fitting a linear regression model to the data probably will not provide a useful model. A valuable numerical measure of association between two variables is the correlation coefficient. That is a value from $[-1, 1]$ range indicating the strength of the association of the observed data for the two variables.
A linear regression line has an equation of the form $Y = a + bX$, where X is the explanatory variable and Y is the dependent variable. The slope of the line is b, and a is the intercept. Example of some basic linear models:

$$\begin{aligned}
Y &= ax + b \\
Y &= a + bx + c \\
Y &= a \sin x + b
\end{aligned} \tag{3.1}$$

3.2.2 Linear prediction

Linear prediction is a statistical technique used to forecast future values based on past observations. It is a method for modeling the relationship between a dependent variable and one or more independent variables in a linear form. The goal of linear prediction is to find the best linear approximation of the relationship between the variables, which can then be used to make predictions about future values of the dependent variable.

3.2.3 Backpropagation

Backpropagation is a widely used algorithm for computing the gradients of the loss function with respect to the weights of a neural network. It is the backbone of training deep neural networks using stochastic gradient descent (SGD) or its variants. The backpropagation algorithm computes the gradient of the loss function with respect to each weight in the network by recursively applying the chain rule of differentiation. The algorithm is typically implemented in two phases:

1. Forward pass: The forward pass involves computing the output of each layer in the network, starting from the input layer and propagating through the hidden layers to the output layer. The output of each layer is computed as a function of the input to the layer and the weights of the layer. The forward pass computes the predictions of the network on a given input.

2. Backward pass: The backward pass involves computing the gradients of the loss function with respect to each weight in the network, starting from the output layer and propagating backwards through the hidden layers to the input layer. The gradients are computed by applying the chain rule of differentiation to the output of each layer. The gradients are then used to update the weights of the network using an optimization algorithm such as SGD.

The backpropagation algorithm can be optimized using various techniques, such as parallel computing, weight sharing, and regularization. It is a powerful tool for training deep neural networks with many layers and millions of parameters, and has enabled significant advances in many areas of machine learning, including computer vision, natural language processing, and speech recognition.

3.3 Datasets

3.4 Comparison criteria

Finally we define the method to compare our models results. Absolute number of income value prediction should not be important for the store owners because of that we calculated the aberration for each month prediction and then we easily calculate quarterly and yearly results. The sum of squared errors (SSE), defined by:

$$SSE = \sum_{i=1}^{n} w_i (y_i - \overline{y_i})^2,$$

between the fitting models and the used data serves as the fitting criterion, with values closer to 0 indicating a smaller random error component of the model.

Also some other quality measures were evaluated, *i.e.* the R-square from interval $[0,\ 1]$, that indicates the proportion of variance satisfactory explained by the fitting-model (*e.g.* R-square $= 0.7325$ means that the fit explains 73.25% of the total variation in the data about the average); R-square is defined as the ratio of the sum of squares of the regression (SSR) and the total sum of squares (SST). SSR is defined as

$$SSR = \sum_{i=1}^{n} w_i(\overline{y_i} - \overline{y_i})^2.$$

SST is also called the sum of squares about the mean, and is defined as

$$SST = \sum_{i=1}^{n} w_i(y_i - \overline{y})^2,$$

where SST = SSR + SSE. Givenm these definition, R-square is expressed as

$$\frac{SSR}{SST} = 1 - \frac{SSE}{SST}.$$

The adjusted R-square statistic, with values smaller or equal to 1, where values closer to 1 indicate a better fit; the root mean squared error (RMSE):

$$RMSE = s = \sqrt{\frac{SSE}{v}}$$

with values closer to 0 indicating a fit more useful for prediction.

3.5 Statistics methods

Median

The median is a statistical measure that represents the middle value of a dataset. It is a value that separates the dataset into two halves: half of the values are greater than the median, and half of the values are less than the median. In other words, the median is the value that is exactly in the middle of the dataset when the values are arranged in order of magnitude.

To compute the median of a dataset, we first sort the values in ascending or descending order. If the dataset has an odd number of values, the median is the middle value. For example, in the dataset $1, 2, 3, 4, 5$, the median is 3. If the dataset has an even number of values, the median is the average of the two middle values. For example, in the dataset $1, 2, 3, 4$, the median is $(2 + 3)/2 = 2.5$.

The median is a robust statistic, meaning that it is not affected by outliers or extreme values in the dataset, unlike the mean. This makes it a useful measure of central tendency in datasets with a large number of outliers or skewed distributions. The median is commonly used in various applications, such as finance, economics, and social sciences, to summarize and compare datasets.

Standard deviation

Standard deviation is a statistical measure that quantifies the amount of variation or dispersion in a dataset. It measures how far the values in a dataset deviate from the mean, or average, of the dataset. The standard deviation is a non-negative number and has the same units as the data being measured.

To compute the standard deviation of a dataset, we first calculate the mean of the dataset. Then, we calculate the difference between each value in the dataset and the mean, square each difference, and sum up the squared differences. Finally, we divide the sum of squared differences by the number of values in the dataset, and take the square root of the result. This gives us the standard deviation of the dataset.

A small standard deviation indicates that the values in the dataset are tightly clustered around the mean, while a large standard deviation indicates that the values are widely spread out from the mean. Standard deviation is commonly used in various applications, such as finance, engineering, and natural sciences, to analyze and compare datasets. It is also an important parameter in many statistical tests and models, such as the normal distribution and the t-test.

4 Syntactic part

Based on the Analytical part 1 let us create new mathematical models and approaches to made a fast and accuracy sales forcasting consist of long-therm linear prediction with individual weights calculated for each period all based on Levinson-Durbin scheme caled Extended linear prediction (ELP) We expect to get better results than by using prediction based on short-term or long-term standard linear prediction (see section 1.3). Finally, our approach will return future values for sales companies based on previous data with better aberration than linear prediction has.

4.1 Extended long-term prediction

4.2 Weights for each period

4.3 AI principles to detect best order of linear prediction

4.4 Combining all principles to forecast process

5 Evaluation

5.1 Experiment

5.1.1 Preprocessing of input data

5.1.2 Models for sales forecasting

5.1.3 Results

5.2 Matlab Live script application

6 Summary

Bibliography

1. VAIDYANATHAN, P. p. *The Theory of Linear Prediction*. 2007.
2. MATHWORKS. *What Is a Live Script or Function?* 2021. `https://www.mathworks.com/help/matlab/matlab_prog/what-is-a-live-script-or-function.html`.
3. PHAM, Dinh Tuan; LE BRETON, Alain. Levinson-Durbin-type algorithms for continuous-time autoregressive models and applications. *Mathematics of Control, Signals and Systems*. 1991.
4. DURBIN, J. *The Fitting of Time-Series Models*. 1960.
5. PARKS, P. Further comments on "An symmetric matrix formulation of the Hurwitz-Routh stability criterion. *IEEE Transactions on Automatic Control*. 1963, vol. 8, no. 3, pp. 270–271.
6. RIAHY, G.H.; ABEDI, M. Short term wind speed forecasting for wind turbine applications using linear prediction method. *Renewable Energy*. 2008, vol. 33, no. 1, pp. 35–41.
7. NAVE, G.; COHEN, A. ECG compression using long-term prediction. *IEEE Transactions on Biomedical Engineering*. 1993, vol. 40, no. 9, pp. 877–885. Available from DOI: `10.1109/10.245608`.
8. D. N. BAKER R. L. McPherron, T. E. Cayton et all. Linear prediction filter analysis of relativistic electron properties at 6.6 RE. *Space Physics*. 1990, vol. 95, no. A9, pp. 15133–15140.
9. KINOSHITA, Keisuke; DELCROIX, Marc; NAKATANI, Tomohiro; MIYOSHI, Masato. Suppression of Late Reverberation Effect on Speech Signal Using Long-Term Multiple-step Linear Prediction. *IEEE Transactions on Audio, Speech, and Language Processing*. 2009, vol. 17, no. 4, pp. 534–545. Available from DOI: `10.1109/TASL.2008.2009015`.
10. CRYER, Jonathan D. *Time series analysis*. Duxbury Press Boston, 1986.

11. FAHRMEIR, Ludwig; KNEIB, Thomas; LANG, Stefan; MARX, Brian D. Regression models. In: *Regression: Models, methods and applications*. Springer, 2022, pp. 23–84.

12. KOTSIANTIS, Sotiris B. Decision trees: a recent overview. *Artificial Intelligence Review*. 2013, vol. 39, pp. 261–283.

13. CHEN, Po-Hsuan Cameron; LIU, Yun; PENG, Lily. How to develop machine learning models for healthcare. *Nature materials*. 2019, vol. 18, no. 5, pp. 410–414.

List of Appendixes

A Flowcharts

A.1 Short-term linear prediction

A.2 Long-term linear prediction

A.3 Extended long-term linear prediction

B Modeling different situations