Forecasting Better Prices for Trip Packages based on Historical Sales Data and Related Factors

(In the context of Europe Railway Tourism)

Authors:

A.M.K.B.Athapaththu	15020071
E.H.Grero	15020258
S.M.Perera	15020551

Supervisor:

Dr. Noel Fernando

This dissertation is submitted to the University of Colombo School of Computing in partial fulfillment of the requirements for the Degree of Bachelor of Science Honours in Information Systems

University of Colombo School of Computing
35, Reid Avenue, Colombo 07
Sri Lanka
January 2020

Declaration

I, A.M.K.B.Athapaththu (15020071), hereby certify that this dissertation entitled Forecasting Better Prices for Trip Packages based on Historical Sales Data and Related Factors(In the context of Europe Railway Tourism) is entirely my own work and it has never been submitted nor is currently been submitted for any other degree.

Signature: ....

Date: February 20, 2020

I, E.H.Grero (15020258), hereby certify that this dissertation entitled Forecasting Better Prices for Trip Packages based on Historical Sales Data and Related Factors(In the context of Europe Railway Tourism) is entirely my own work and it has never been submitted nor is currently been submitted for any other degree.

Signature: ....

Date: February 20, 2020

I, S.M.Perera (15020551), hereby certify that this dissertation entitled Forecasting Better Prices for Trip Packages based on Historical Sales Data and Related Factors(In the context of Europe Railway Tourism) is entirely my own work and it has never been submitted nor is currently been submitted for any other degree.

Signature: ....

Date: February 20, 2020

I, Dr. Noel Fernando, certify that I supervised this dissertation entitled Forecasting Better Prices for Trip Packages based on Historical Sales Data and Related Factors(In the context of Europe Railway Tourism) conducted by A.M.K.B.Athapaththu, E.H.Grero, S.M.Perera in partial fulfillment of the requirements for the degree of Bachelor of Science Honours in Information Systems.

Signature: ....

Date: February 20, 2020

Acknowledgement

We would like to take this opportunity to express our gratitude to the people who have lent their support all along this thesis and thus enabled us to concentrate on our work. Foremost, we would like to express our sincere gratitude to our supervisor Dr. Noel Fernando for the continuous support to our final year research. His patience, motivation, enthusiasm, and immense knowledge helped us very much in doing our research. His guidance helped us in all times of the research and in writing of this thesis. We would like to thank Mr. Thushan Gunawardhane, Mr. Sanjeth Srikathirkamanathan and Mr. Amith Dissanayake for their insightful comments, hard questions, immense support and encouragement as advisors during our research duration. We also thank our 4th year fellow mates, for stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had in the last year. Last but not the least, our families and friends for continuous support and encouragement throughout the year.

Abstract

Traveling and tourism is a world-leading industry. Having a competitive price is crucial to survive in the current competitive environment.This study is related to forecasting better prices for a railway based tourism company located in the European region, by considering the past sales patterns with the external factors such as weather, season and holidays. Currently, they are deciding their prices by past experiences.

The dataset received contained all products of flam railway and by preprocessing and feature extraction relevant records were chosen only relating to flam railway. Extracted data was combined with external data such as weather,season and holidays which were collected from APIs. Then trip packages were identified and dataset was broken based on different packages.

In this research researchers try to evaluate and compare the performance of various models traditionally used for price prediction. Here performance of Deep Neural Network,Ordinary Least Squares Multiple Linear Regression Model (sklearn), SARIMAX Model,Ordinary Least Squares Multiple Linear Regression Model (statsmodels), Support Vector Machine and Extreme Learning Machine was evaluated. Root mean square error was used to compare the performance of the models. Model with least root mean square error was selected to predict the price of the model.

75 packaged had Deep Neural Network as the best performing model while 11 and 3 packages had Ordinary Least Squares Multiple Regression (statsmodels) Model and Linear Regression Model(sklearn) respectively.

The hybrid model was created using above models, deflated price and respective sales volume. Estimated increase of revenue when using the better price of the hybrid model had a maximum of 120.59% increase, minimum of 12.12% increase and average of 79.25%

Table Of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background of the Problem

There are various platforms including Book Visit by Visit Group, where service providers can list their packages and trips to attract customers. Currently they have to set their prices using past experience and expertise knowledge. At present there is no method to get price suggestions based on sales patterns in seasons, weather, holidays etc. But a huge volume of data is available in this domain. Furthermore these prices may change due to external factors such as weather, holidays, seasons etc. which cannot be derived by analysing only historical data.So identifying a better price based on sales patterns is an important timely problem in the current context.

1.2 Problem Statement

Europe Railway Tourism Company 'Visit Flam' is currently estimating the price for their railway tourism trips manually. Sometimes Visit Flam does not achieve a good revenue from their railway tourism trips by using the manually estimated trip prices.Europe is an area where there is rapid weather and seasonal changes. Therefore these factors also affect how people go on trips.Therefore it is a necessity for them to predict a better price for their trip packages by also analyzing weather,season and holiday data. Therefore by analysing and detecting hidden sales patterns and external information including weather data,seasonal data and holidays, we can successfully forecast a better price to the service provider. Thereby the service provider can gain a competitive advantage over other competitors. If we can address this problem Visit Flam can achieve a significant revenue, reputation and customer base increase.

1.3 Research Questions

1. Research Q1: What are the current techniques for pattern recognition, business intelligence and forecasting better prices?
2. Research Q2: What are the gaps in the current techniques described in Research Q1 with respect to forecasting better prices based on sales patterns considering related factors in Europe railway tourism?

3. Research Q3: How to propose a better price model to forecast prices based on sales patterns to the Europe railway tourism to overcome gaps in Research Q2 by considering the other factors?

1.4 Significance of the Research

Many travel providing companies and even railways can adopt this research approach of price prediction for their trips. This approach will be suitable for trips of any country as weather,season and holidays definitely affect the participation of passengers or tourists in trips in any country. Even the Sri Lankan Railway can get a better price for the train journeys based on these external factors. We also suggest that rather than just manually giving a price to a journey or a trip it is better to take other factors which influence the trip or journey into account when giving a trip or journey price.
A major significance of this research is that for any trip or journey this research approach and the models could be adopted to give a better price for the trip or journey. Not only for train trips but even for bus trips this research approach and models could be implemented.

1.5 Research Approach

This research will take a quantitative research approach to meet the main objective of the project which is to forecast a better price using the generated models. Different models will be evaluated quantitatively so that a fair comparison can be done among them. This will allow the researchers to identify the shortcomings of each model in the current context of the research and to to measure the effectiveness of each model.The machine learning models used are Deep Neural Network, Extreme Learning Machine and Support Vector Regression. The statistical models used for price prediction are time series model, linear regression and ordinary least squares model.A hybrid model was built by taking the best features from the models.Root mean square error was taken in choosing the best model.Hybrid model will be able to predict a better price for all the 89 trip packages by considering the external factors which are weather,season and holidays.

1.6 Limitations, Delimitations and Assumptions

1.6.1 Limitations

1. The prediction is based on the data set provided by the Norway based train tour company.

2. The data is based on only for the recorded daily trips based on several locations following 89 trip packages.

3. Trip requests for certain trips are not considered and only the participated trips are considered.

4. Base fee for a train ride for the certain package is considered as the Trip Price but other expenses are not considered.

5. Weather (Average Temperature ,Snow Fall, Sun Hour and UV Index), Holidays and Seasons are only considered as the external factors.

1.6.2 Delimitations

1. Only applies to destinations of Flam Railway.

1.6.3 Assumptions

1. All trips chosen only contain train travel between stations.
2. Lowest price for each trip was taken assuming it only contains train travel.

1.7 Contribution

We contribute following to the research community,

- Comparison of various models used to predict the price of a package using not only historic data but also external factors such as weather, season and holidays.

Furthermore we deliver the following to the research community for future research in this area.

- A script to generate the whole journey of a trip from trip legs.
- A script to identify least price for the day and substitute it as price for all trips in that package for the day.

1.8 Remaining Chapters

Chapter 2 is the Literature Review and Background chapter and it shows what are the current models available for price prediction and the limitations of the current models.Also Chapter 2 has a description about Flam Railway.Chapter 3 is the methodology chapter and it explains the research design and the research approach. Chapter 3 shows how data was collected and pre-processed and descriptions of the models that were used for predicting a better price. Chapter 4 is the results chapter and it shows the results we got from each model and also the relationships between variables we identified when pre-processing. Chapter 5 is the discussion chapter and it discussed the results of the results chapter and also the findings from our research. Chapter 5 contains also the conclusion of our research.

Chapter 2

Literature Review and Background

In this Chapter, a review of related work on forecasting better prices for trip packages based on historical sales data and related factors in the context of Europe Railway Tourism is discussed. Section 2.1 provides a brief introduction about Europe Railway Tourism and Flam Railway. Section 2.2 provides an introduction to price forecasting and trip price forecasting and section 2.3 discusses the existing approaches for price forecasting. Finally section 2.4 discusses the weaknesses in existing approaches for price forecasting.

Literature Review provides answers to the following Research Questions

- What are the current techniques for pattern recognition, business intelligence and forecasting better prices while also considering external factors?
- What are the gaps in the current techniques described in Research Q1 with respect to forecasting better prices based on sales patterns considering related factors in Europe railway tourism?

2.1 Introduction to Europe Railway Tourism and Flam Railway

Today travelling and tourism has become a rapidly increasing industry. It has become a major income method for some countries which have a lot of places with natural and historical importance. Because of that, the tourism industry has contributed to the local and also global economy massively. Tourism corporations have played a major role in this contribution to the tourism industry. Because of that there are a lot of tourism corporations which organise tours locally and globally with the help of information and Technology. Competition between these corporations to attract customers is that they are using a lot of offers, fascination packages etc. To gain the competitive advantage there is a trend to use machine learning and price recommending systems to offer most profitable offers to their customers.

2.1.1 Europe Railway Tourism

Europe is a major tourist attraction place in the world due to the significant, historic and attractive places. And also millions of tourist who love to experience variations of climates, weather, seasons and culture.Considering the shares of arrivals in 2017, Europe has the highest amount of shares recorded as 51% [1]. Accoring to the statistics on 2017, there were more than 672 Million tourist arrivals have been recorded and it is an increment of +8% when considering the previous year [1]. Tourist industry has contributed to the economy of Europa Region by 519 Billion USD in 2017 and it was an increment of +8%. Considering the statistics on 2017, the most popular transportation mode among the tourist is by Air [1]. But considering the Purpose of visit, 55% of tourists have arrived in 2017 for leisure, recreations and holiday [1]. Due to that most of the tourists use air for the arrival to a country but for the inbound tourism most of them are using Railways as the transportation mode.

According to the NationalGeographic.com [2], Railway tourism has been promoted due to the facilities such dinning, comfortable and sleeper compartments. And also it an economical mode of transport with huge facilities. Due to the Tourism Railways are routed among most attractive tourists destinations, tourists are able to locate different site seen locations. According to NationalGeographic.com, the Top 10 Railway trips in Europe are listed below [2].

- Sweet Switzerland: The Chocolate Train
- Tunnels Galore: The Bernina Express
- A Hotel on Wheels: Trenhotel
- The Epic Journey: Trans-Siberian Railway
- Waterworld: The Flam Railway
- Bavarian Bullet: InterCity-Express (ICE)
- The Elegance of Yesteryear: Venice Simplon-Orient-Express
- Roughing it by Rail: Balkan Flexipass
- Luxury on Wheels: The Balkan Odyssey
- A Nostalgic Journey: El Transcantábrico Gran Lujo

2.1.2 VisitFlam and Flam Railway

VisitFlam

VisitFlam is the destination company for sales and marketing of the Flåm- and Aurland region on both national and international extent. Key elements in the company marketing are the brand Visit Flåm and through Internet portal www.visitflam.com. The vision behind Flåm AS is to create the world's best fjord destination and its mission is to provide and sell excellent sustainable experiences [22]. The VisitFlam is a group of companies that provide tourist facilities in different tourist attractive places in the Norwegian countries with different regions. And also they provide variation of what facilities to their customers. there is awesome highlighting facilities and tourist destinations such as FLÅMSBANA - THE FLÅM RAILWAY, FLÅM AS, THE FJORDS DA, FRETHEIM HOTEL, FRETHEIM HOTEL and AURLANDSKOEN ÉCONOMUSÉE.

The railway is the major tourist attraction in the company that runs on daily all year round. . From Flåm by the shore of the Aurlandsfjord up to Myrdal station on the Bergen Railway, 866 meters above sea level. The purpose and vision behind Flåm Utvikling AS is to maintain and develop the Flåm Railway as one of Norway's major tourist attractions and develop Flåm and its surroundings into one of the prime tourist destinations in Scandinavia. Flåm AS and NSB – Norwegian State Railways, own the company 50 percent each [3].

Flam AS is the company that offers regular sightseeing tours, activities, accommodation and package tours by boat and bus in the Flåm and Sognefjord area. In addition, Flåm AS offers a wide range of package tours combining boat, train and coach as well as charter [3].

The company The Fjords DA offers regular sightseeing tours and charter between Flåm – Gudvangen, in other parts of the Sognefjord area, as well as between Geiranger – Hellesylt, Hjørundfjord, Oslofjord and on the Lys Fjord near Stavanger. Owners of The Fjords DA are Flåm AS and Fjord1 with 50 percent each [3].

The Flåm Railway Museum gives an insight into the building of one of the world's steepest railways, the technical development and the people behind this unique engineering work. The museum is open all year and offers free entrance.

The traditional Aurland Shoe is being produced in Aurland, at the only remaining shoe factory in Norway producing the traditional shoe.

Products and Services

Flam Group is conducting their sales of their products and services in four main regions in Norway. [4]

1. One entry in the list
2. Another entry in the list
3. Stavanger Region
4. Oslo Region

In Flåm Region there are some sub products and services conducted by the company [4].

1. The Flåm Railway
2. The Fjords Product
3. Fretheim Hotel
4. Flåm Area and Port

5. Attractions and activities in the Flåm area

In This research, it is only considering the product of the Flam railway service that was offered with the 50% combination of the Norwegian State Railway along the Aurlandsfjord.

- **The Fjords** Owned by Fjord 1 Flåm AS
  - Boats on Geirangerfjord
  - Boats on Lysefjord
  - Boats on Nærøyfjord
  - Boats on Sognefjord
  - Boat on Hjørundfjord
  - Boat on Oslofjord
- **Flåm Utvikling AS** Owned by VY Flåm AS
  - Flåmsbana/ Flam railway
- **Own products** Owned by Flåm AS
  - Fretheim Hotel
  - Heimly Pension
  - Toget Café
  - Bakkestova Café
  - Stegastein Viewpoint
  - Shuttle Bus
  - Flam railway
  - Museum
  - Aurland shoe factory
  - Sognefjorden AS
  - Activities
- **External products** Minor or no ownership
  - Myrkdalen resort
  - Boat on Fjærlandfjord
  - Boat charter
  - Bus services
  - Activities
  - Other

Flam Railway

Flam railways is one of the few private railways that has decided to offer the journey from Flam to Myrdal which is situated 876 metres from the sea level along the UNESCO heritage Fjord valley. The journey is decided to Flam to Myrdal, which is owned by the Norwegian State Railway. The distance of the journey is 20 km and its 80 percent of the track has 5.5 percent of gradient and its smallest curve is 130 metres.

According to their reports [3], the Flam Railway 1435 gauge railway tracks as the railroad is located with a narrow route. The main power train of the rail is electric powered locomotive after the Norwegian State Railway has converted their locomotives as electric locomotives in 1944 [5]. The power stations situated in the suburb area which are driven by the waterfall generate electricity of the drive trains by empowering the "Green" concept in the UNESCO heritage region. According to the report, Flam is going to implement a fully electric vessels for the cruising in other products by converting the boats into electrified propellers[3].

In regular trains there are two electric locomotive engines which are powered by the 15000V of grid connection [6] which weighs over 64 tons each (According to the reports, the secondary break engine of the locomotives provide electricity for the national grid by dynamic braking [5]) . There are six carriages with 525 seats weighing 40 tons without passengers and 45 tons with passengers. Overall train is more than 182 metres in length with 6 carriages and 2 engines. Trains are transported according to the published time table from Flam - Myrdal - Flam. Number of train turns will be varied according to the passengers and according to seasonal changes.

Facilities and Features

The Flam Railway provides group or individual bookings for the selected destination will it return tickets one-way tickets. And also the trains are equipped with cycle carriages and also special facilities for the disabled people. The train fare will be varied according to the age groups such as it can be varied for the elders and for children.

Departure and Destination

Flam Railway started from the city of Flam which is 2 meters from the sea level. Its final destination is the city of Myrdal which is 865 meters above sea level Along 20 km. This journey grew through eight Railway stations, 20 tunnels and 4 water tunnels [2]. The maximum speed of upwards is 40 kilometres per hour and downward maximum speed is what I would do to the huge descent. Along with this Railway track. So many beautiful places such as mountains, waterfalls, historical places and also site scenes. To be slower, visitors are able to observe the environment and sometimes trains will be stopped or slowed down near some specified places. For the return tickets visitors are able to travel Myrdal and return to the Flam again. One way tickets I am able to go to the specified destination and travel to be the package such as Train hiking, cycling, boat travel, hotels or for the ancient places.

Bookings and Seat reservations

The booking for the tickets can be done by online store of the flam railway or using the authorized ticket outlets from the Norwegian State Railway (NSB). Or It can be done by the authorized third-party agents such as online Citybreak Agent Network, - or on the spot in premises in Flåm Visitor Centre Train Station. There is a total capacity of 525 cheats seats divided into six carriages. Companies are recommended to reserve seats in advance because there will be some fully booked dates frequently. Users are able to book their seats by groups or as individuals. If the seats are booked they are marked as reserved and if there is a name for the group it will also be displayed. If the visitor is an individual person it is able to reserve seats as individuals and if the carriages are full, there will be an additional carriage for individual visitors. If the group of the visitors require different language for guidance, the LCD screens are displayed by their preferred language. Tickets can be booked using the age group or the type of the package that the user selects. For the cancellation of the ticket, it is able to contact the agent or directly to the company for the mutual cancellations.

Business and Information Terminals

VisitFlam's main terminal for attracting customers and place orders is their online website www.visitflam.com. In this website they have included their destinations tourism packages and their timetable and also charges for each and every package [1].
According to the report, there are some agents, who can provide tickets and packages to the customer on behalf of VisitFlam. Their agent is Norwegian State Railway (NSB) conduct rate tourism packages with Cooperatively them [3].
According to the figure they have predefined charges for the tourism packages for the railway for 2020 [3]. There are tickets called One Way Ticket and also return tickets which allow customers to select their packages and destinations.
There are so many train turns in the year. But according to the weather and their past experience they have free define the number of train tours for some seasons.

Flåm Myrdal or Myrdal–Flåm (One way) – Flåm–Myrdal–Flåm (Return)			
01 January-30 April		**One way**	**Return**
Ordinary adult ticket		NOK 360	NOK 490
Ordinary child ticket 6-17 years		NOK 180	NOK 244
01 May-30 September		**One way**	**Return**
Ordinary adult ticket		NOK 430	NOK 630
Ordinary child ticket 6-17 years		NOK 215	NOK 314
01 October-31 December		**One way**	**Return**
Ordinary adult ticket		NOK 370	NOK 500
Ordinary child ticket 6-17 years		NOK 185	NOK 250

Figure 2.1: Proposed prices for Flam Railway 2020

2.2 Introduction to Price Forecasting and Trip Price Forecasting

Price Forecasting has become very crucial as nowadays as customers seek to get a lower price while companies try to keep their revenue as high as possible and maximize their profit. Therefore it is becoming very much essential to use various computational techniques to increase the company's revenue such as demand prediction and price discrimination[7].Similarly Trip price forecasting is used to determine a best price for the trips that the company provides to their customers while considering also the external factors that affect the trip price using the computational techniques.

One of the key issues from a customer point of view is determining the minimum price or the best time to buy a ticket. The concept that tickets bought in advance are cheaper is no longer working[8]. It is indeed possible that customers who bought a ticket earlier than others pay more than those who buy a ticket later. It is also true that early purchasing implies a risk of commitment to a certain schedule[9]. Early booking will not see the weather and seasonal changes which will happen later.Therefore it is clear that trip ticket price may be affected by several factors and thus the trip price may change continuously.

Many companies of the tourism industry are now intensively using data analysis and sophisticated mathematical techniques to predict the willingness of various customer segments to pay and also to maximize their profits. This has naturally resulted in a price discrimination phenomenon which depicts that the same service which can be same category of seat for a trip being transacted at different prices. This fluctuation of prices is a worry for the customers. Therefore it becomes very difficult for the customers to decide to buy a ticket or not to a certain price in a certain time .This also naturally opens the way for supply of possibly new services which could be used as decision making tools for customers[10].

Therefore various price forecasting techniques are used to predict the price of a trip which could be good from both the customer and the trip offering companies points of views. According to Aggarwall et.al[11],it is seen that according to the nature of the model price forecasting techniques can be categorized as below.

1. Heuristic models
2. Simulation models
3. Statistical models

Below also shows the diagrammatically the approaches which could be used based on the above models[11] (Figure 2.2).

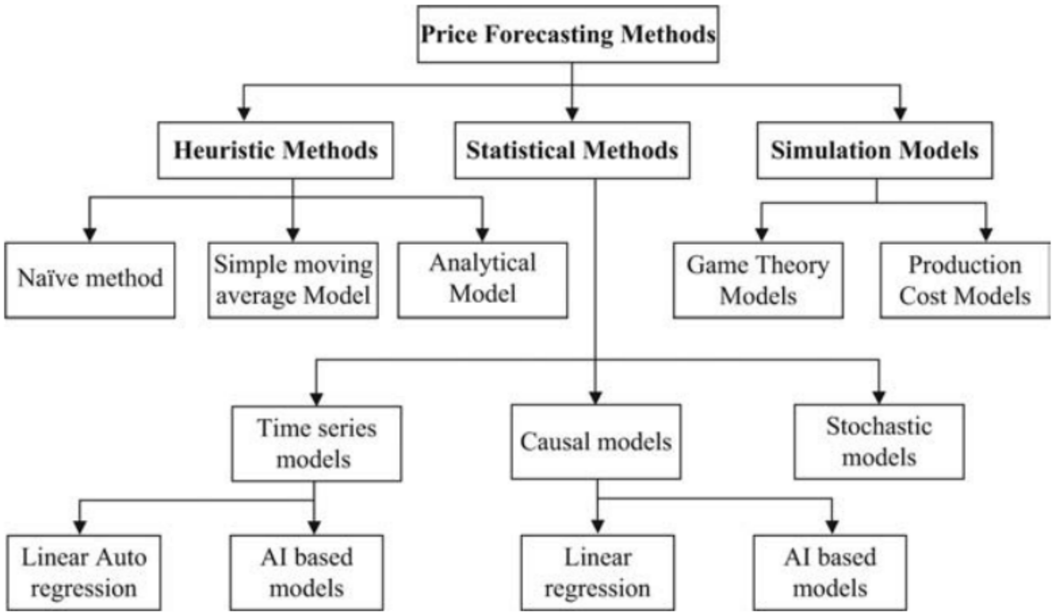

Figure 2.2: Price Prediction Model

2.3 Existing Approaches for Price Forecasting

Price Forecasting has been done using various approaches and even some have done using more than one method. Price Forecasting Approaches are very much essential in identifying the approaches to be used for the trip price forecasting.

According to Makkonen et.al [12], electronic commerce especially in the business-to-consumer context has become very popular. Therefore they have identified the most typical sales patterns of online stores in the B2C context. They have segmented the monthly sales time series of 399 online stores with time series clustering. They were able to identify four approximately equalled sized segments of which two were characterized by a clear upward or downward trend in sales and two were characterized by strong seasonal variation. They also got this result by exploring 14 business and technical parameters which includes Monthly sales, item categories, payment methods,customers,campaigns etc.Therefore from this it is very clear that analysis of sales data could be done using time series and variations of prices could be seen based on different external factors such as seasons.

As our dataset contains seasonal and time series we need a model which supports variations in seasons and which can be used to see the seasonal trend. According to Permanasari et.al [13], due to the seasonal trend in time series the Seasonal ARIMA can be selected as the model to be used. Here they have forecasted the number of malaria incidence by using the time series model SARIMA. They have used Mean Absolute Percentage Error (MAPE) in measuring the error percentage or the accuracy of the SARIMA model. Similarly SARIMA model could be used to forecast the price.

In our research we consider only 3 external factors in predicting trip prices which are weather,season and holidays.According to Arunraj et.al[14], the daily demand for fresh food products is affected by external factors such as seasonality, price reductions and holidays. They have divided holidays into 4 categories as regular holidays, Christmas, Easter and school vacations. The two types of price reductions they have used are discount price reduction and promotional price reduction. The seasonality they have taken is according to weeks and as on sundays the retail shops in Germany are closed they have taken the seasonal length as 6 considering from monday to saturday. This represents the sales pattern in a week. Then they developed a Seasonal Autoregressive Integrated Moving Average with external variables (SARIMAX) model which accounts for all the above external factors to forecast the daily sales considering also the demand. It was found that the proposed SARIMAX model improves the traditional Seasonal Autoregressive Integrated Moving Average (SARIMA) model and is very successful.Therefore from these facts we can conclude that SARIMAX model can be used in our research.

Two external factors we are considering when considering a better price for a trip are holidays and the weather. According to Chen et.al [15], temporal factors they have selected are months,weekdays and holidays and the weather forecasts which were selected as predictive factors were daily mean temperature,amount of rainfall and the amount of snowfall. Accordingly they used seasonal auto-regressive integrated moving average model with external regressors (SARIMAX) as they used the above temporal factors and predictive factors as external factors. A series of SARIMAX models of different orders was estimated and diagnosed.A general linear model (GLM) was also developed and was compared with the SARIMAX models by validity measures. They concluded from evidence of model validations that both SARIMAX and General Linear Model (GLM) can be used in forecasting daily collisions.Similarly trip prices also could be forecasted.General Linear model cannot be used in our research as it is used when there are more than one dependant variables but we have only one dependent variable which is the price.

One of the statistical methods which could be used for price prediction is linear regression. According to Tudy et.al [16], predicting daily behavior of a stock market is a serious challenge for corporate stakeholders and investors and could help them to invest more confidently by taking into consideration the risks and fluctuations. Therefore to prove that stakeholders can confidently invest prediction of SP 500 index was done by using linear regression. To do linear regression they had first divided the dataset as training and testing. It was able to obtain the coefficients, slope , error and the intercept by applying the linear regression model. Through this they were successful in proving that a good performance can be seen by applying the linear regression model.Here the linear regression method is done is Multiple Linear Regression Method. Multiple Linear Regression is also called Multiple Regression. Multiple linear regression which is considered a statistical technique uses several explanatory variables to predict the outcome of a single response variable.

According to Shinde et.al [17], for predictions of the sales price of houses various machine learning algorithms could be used. Housing sales prices are normally determined considering numerous factors such as the location of the house,area of the property, material used for construction, number of bedrooms,age of the property, and garages etc. Here they have used decision trees, Logistic Regression and Support Vector machines.

They have considered housing data of 3000 properties.Here from Logistic Regression and SVM R-squared values of 0.96 and 0.81 were obtained showing that both models show good performance. This was further proved by the comparisons of the algorithms based on parameters such as Mean Absolute Error (MAE), Mean Square Error (MSE) and Root Mean Square Error (RMSE). Logistic Regression cannot be used to our research as our dependent variable price is not binary in nature but linear regression can be used as price is continuous and regression is linear. From these facts it is clear that for our research linear regression could be used in price forecasting.According to our dataset multiple regression is the most suitable one as we have more than one independent variable.Support Vector machine is also an option to be considered.

Linear Regression can be done using several methods.One of the methods is Multiple linear regression which is discussed above and one approach of Multiple Linear Regression which can be used is Ordinary Least Squares method and it is the most commonly used multiple linear regression method. According to Lin et.al[18], Ordinary least-squares (OLS) regression is emphasized as a generalized linear modeling technique that could be used to model a single response variable, which has been recorded on at least an interval scale. The technique may be applied to single or multiple explanatory variables and also categorical explanatory variables that have been appropriately coded. As we have weather,season and holidays as explanatory variables this method is considered suitable.

According to Dubin et.al[19], the ordinary least squares (OLS) linear regression is the standard method used to build hedonic price models.Also it is shown that ANN algorithm used in the study was outperformed by ordinary least squares (OLS) regression.Therefore ordinary least squares model is certainly an option we should try and use in our research.

P. Pai et al has compared Seasonal Support Vector Regression models with Autoregressive Integrated Moving Average and Support Vector Regression models and has concluded that the SSVR is much better at predicting time series data [20].

Z. Sun et al have proposed an artificial neural network [ANN] learning algorithm called Extreme Learning Machine (ELM) for sales forecasting in the fashion retailing industry. This a single-hidden-layer feedforward neural network where the input weights and hidden biases are randomly chosen and the output weights are analytically determined by using the Moore– Penrose (MP) generalized inverse. [21] Artificial neural networks and support vector machines are used extensively forecasting time series data.

P. Chang and Y. Wang states that the most rewarding method for sales forecasting is integrating artificial neural networks with fuzzy theory. They have used the Step wise Regression Analysis and Fuzzy Delphi Method to filter the key variables to be analyzed. They have concluded that the Fuzzy Back Propagation Network is performing significantly better than Grey Forecasting, Multiple Regression Analysis and traditional Back Propagation Network[22].This shows that Artificial Neural Networks perform well.

R.Kuo and K. Xue also have used a Fuzzy Delphi method in creating a forecasting system which can handle external factors such as promotions. [23]

2.4 Weaknesses in Existing approaches for Price Forecasting

According to M.Makkonen et.al [12], prediction from a time series can only be done for the time period of the dataset. That is we can divide the dataset as training and testing and can predict for the testing dataset but we cannot predict for another dataset. If we want to predict a pattern for future data, then data needs to be there till that date to predict the pattern. Therefore the weakness in time series or SARIMAX is that continuous data is needed for prediction and also can only predict for the time period of the dataset.

According to Shinde et.al[17], they have concluded that decision trees overfit although they get a 100 percent accuracy in predicting house prices. They also say that doing house price prediction using decision trees is a problem because decision trees take into consideration the noises that are present around. Although they have tried to divide the dataset for training and testing as 50:50 they have not succeeded in solving the overfitting problem. Therefore because of overfitting in decision trees in price prediction we are not using decision trees.

According to Z. Sun et al[21], over tuning and high log computation time is inherited problems in using Extreme Learning Machine. Over tuning of parameters results in the model not being accurate. Although Extreme Learning Machine is fast it is not very accurate because of the over tuning.Also another weakness of Extreme Learning Machine model is that it has a high log computational time.This is because of the higher complexity. Therefore we did not take the Extreme Learning Machine for the final hybrid model.

According to Palmer et.al[24], one of the weaknesses of linear regression is that it provides only a linear relationship. Also linear regression has a weakness of correlation measures in providing diagnostic. Because only providing linear relationships linear regression or Normal Linear Regression can only be used in scenarios which require linear relationships. Also another weakness is that Linear regression is sensitive to outliers. Therefore it is wise not to use linear regression if there are many outliers.

According to Wang et.al[25], In Support Vector Machine(SVM) numerical optimization in a high-dimensional space may suffer from the course of dimensionality.This will result in obtaining a non optimal solution. As it is dealing with numeric data it is not wise to implement SVM if there is a higher-dimensional space. Weather,season data and price are numerical so we should be cautious when implementing SVM.

H. Cao et al propose a support vector regression trained by particle swarm optimization to predict warrant prices. "A Support vector regression (SVR) technique is a learning procedure based on statistical learning theory, which employs structured risk minimization principle." The accuracy of the SVM depends on the training parameters. H. Cao et al have used particle swarm optimization when selecting the training parameters and applied it to warrant price prediction [26].

According to the Murat et al. (2016), The price of the tourism products will depend on internal and external factors. The internal factors are marketing cost structure and cost decision mechanism. The external factors of market structure, product price of opponents, chain of distribution in legal legislations. Most of the tourist companies determine their prices by analyzing various factors suggesting market structure and conditions of competition. The suppliers choose a pricing structure which allows them to share their products on the online channels and they have three very important results with the help of Van Westendorp PSM method. These are Optimal price point, Acceptable price point, and Indifference price point. There are some factors that affect the pricing of the tourism industry. They are cost, competition, judicial factors, demand in the market structure, factors depending on distribution channels and consumer behavior [24].

According to our research we mainly consider the external factors weather, season and holidays. The pricing structure is organized according to the different trip packages. Also, in order to give a better price, we are considering an Acceptable price point. But we don't have the cost for trips. It is not possible to calculate profit using our models and a gap is that different models use different data formats in processing based on the factors used.

Chapter 3

Methodology and Design

3.1 Introduction

3.1.1 Research Purpose

This chapter elaborates on the design of the research where researchers strive to compare the price predicting models to circumvent the inherent weaknesses in the traditional methods in recommending and predicting a price.
This study uses Railway Tourism data provided by the Flam Railway of the Visit Group, Norway obtained via collaboration between Creative Software and the University of Colombo School of computing and selected external factors related to the tourism industry.

3.2 Research Questions

- **Research Q1**: What are the current techniques for pattern recognition, business intelligence, and forecasting better prices?
- **Research Q2**: What are the gaps in the current techniques described in Research Q1 with respect to forecasting better prices based on sales patterns considering related factors in Europe railway tourism?
- **Research Q3**: How to propose a better price model to forecast prices based on sales patterns to the Europe railway tourism to overcome gaps in Research Q2 by considering the other factors?

3.3 Research Design

This research strives to implement the most popular price forecasting models and evaluate and compare those models to find the weaknesses in each of them. Then the researchers will implement a hybrid model to mitigate the impact of these weaknesses to the forecasted price. Research design (Figure 3.1) highlights the main 11 steps researchers are following to get the expected results.

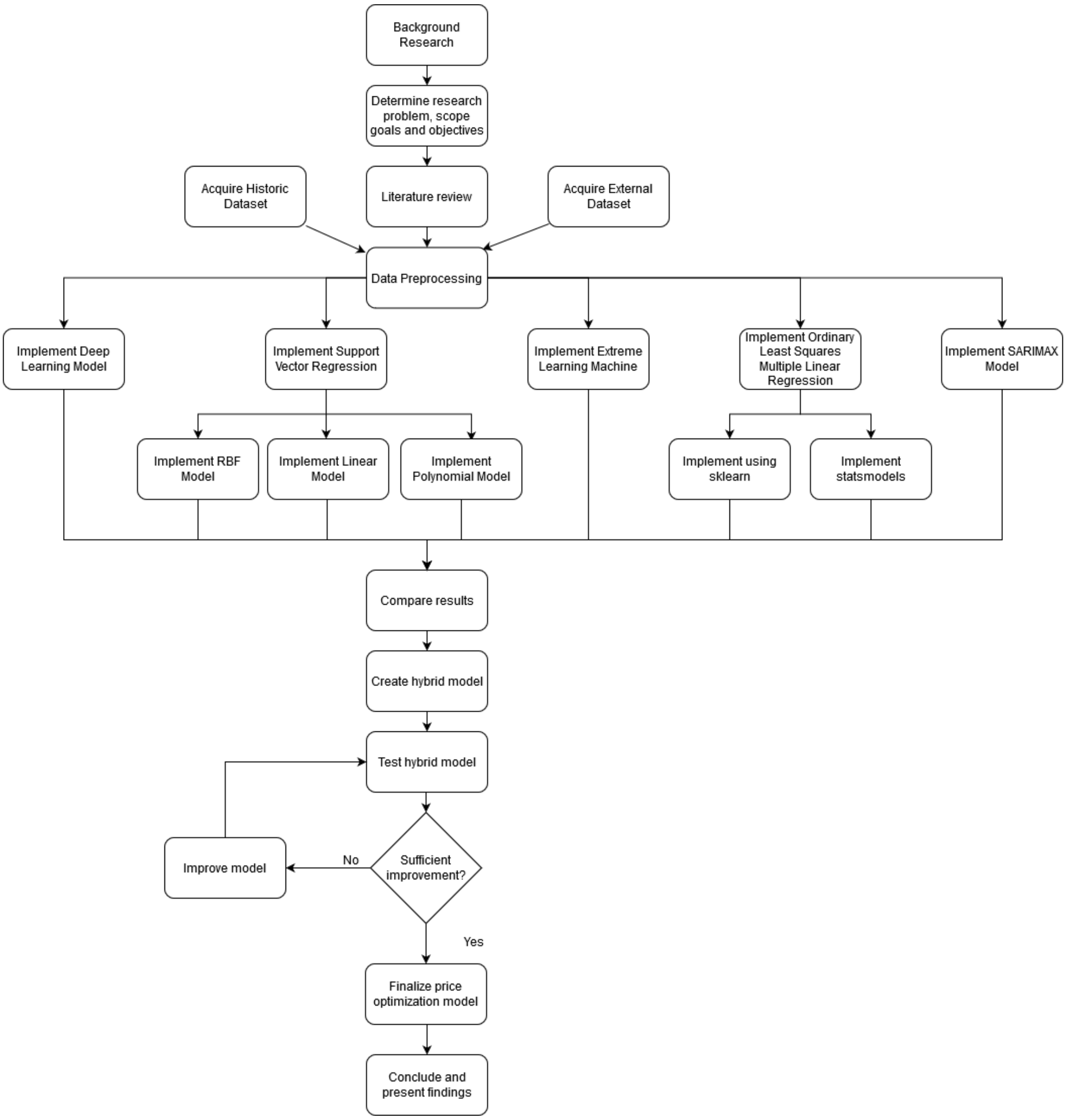

Figure 3.1: Research Design

3.4 Methodology

Initially, data was provided by the Visit Flam with related tables for all trips handled from 2008 to 2019. Due General Data Protection Regulations, the personal and identical data of the customers are not provided and set as null values in the database. Data for the train tours are filtered with relevant tables. Each and every trip has obtained its own TripId by automated Online Reservation System. According to the package that the customer chose, Trip Legs are created. Trip Leg is a single arrival and departure between two destinations (two railway stations) and a trip should be obtained at least one Trip Leg. By considering the sequence of arrivals and departures, every Trip Legs were concatenated according to Trip ID to create a unique trip based on the Trip Package. There are 89 identical trip packages and 3 of them are return trips. The base fee for the trip package for a certain day is considered as the Price for that Trip package on that day. The acquired location based weather data, holidays and Seasons are joined to the trip data set and analyzed by the methodologies mentioned below.

3.4.1 Background Research

A background study on recommender systems, business intelligence, price forecasting, and railway tourism was done by the researchers to gain a high-level understanding of the problem domain.

3.4.2 Determine the research problem, scope, goals, and objectives

After the background research problem, scope, goals, and objectives were decided as mentioned in the introduction.

3.4.3 Literature Review

A literature review was done to get a comprehensive understanding of current solutions to forecast a better price as discussed in the related work.

3.4.4 Acquire historic data

Dataset for the research was provided by client company with the data of the other packages and tourism activities. The provided data set has trip records from 2008 to 2018 with over 3.1 Million Trip records.

3.4.5 Acquire external data

External data was acquired using two online APIs. Weather and seasonal data were obtained through the World Weather Online API [27]. Holidays data was obtained through Time and Data API [28].

3.4.6 Data preprocessing

Data preprocessing was done using the Microsoft SQL server management studio 2017. Data was obtained as a BACPAC file which was then imported to SQL server management studio. Then according to the product flam railway of Visit Flam 7 tables were selected and the data set was created by joining them using certain conditions. Then the external data of weather, season and holidays which was collected was joined with the dataset created by joining the seven tables. In data preprocessing mainly rows which are null and also repeating rows were avoided. Furthermore, data that was identified as test data was removed.

The trip was considered as the selling product and trip packages were identified as the factor to which price has to be forecasted. Trip packages are built by the combination of departure legs. Trips go on different days through the same locations belong to one trip package. After preprocessing 89 trip packages were identified with 3,166,844 rows of trips.

Below shows the final query and results in the Microsoft SQL server management studio by joining all other data with the weather table.These preprocessing was done considering that Instance Id 24 is Flam Railway product of Visit Flam and PriceGroup Id 92 is Adults for which only we predict prices.

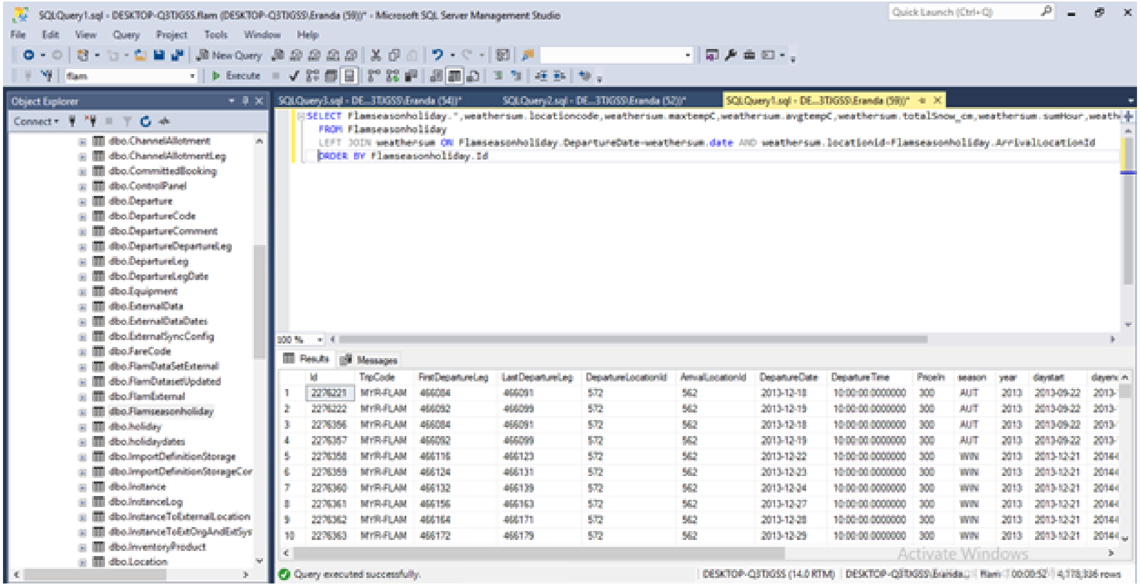

Figure 3.2: Sample Data Set

Trip Package Creating

The Trips of the Flam Railway is based on the railroads across eleven stations around the site seen areas. Tourists are able to join to the train trip at any station and end the train trip at any station and join for the other kind of sightseeing tours such as boat rides and hotels. By considering the patterns and the frequency of the usage of the arrival and departure stations of the tourists, Flam Railway has been designed pre-defined train packages that cover all the significant stations. For example, if a tourist reserved a trip from the Station A along the Station B and C to Station D, it is identified as a unique trip package. In Flam Railway, there are 89 distinct trip packages including one way and return trips. In the system, the locations/Stations are identified using its unique code

In the automated reservation system, train trips are recorded in two tables.

- Trips
- Trip Legs

Trips are the unique reserved trips that included the above unique trip packages. For example, if a tourist reserved a trip from the Station A along the Station B and C to Station D package it was identified as a trip from A to D.

The 'Trip' table, which contains the following details in the record row.

- **TripID** - Unique ID for the Trip
- **First Departure Location** - Initial departure location of the Trip
- **Last Arrival Location** - Trip's Final Destination
- **Trip Code** - Indicate the First Departure and Last Arrival Stations (For example, if the Trip is Scheduled from Station A to Station B, It was Recorded 'A - B ' as the trip Code

- **Departure Time** - Departure time and the date of the trip from the first Departure Location according to the train schedule
- **Arrival Time** - Arrival time and the date of the trip to the Last Arrival Location according to the train schedule
- **PriceIn** - Price for the Trip

Trip Legs are the departures and arrivals between the intermediate stops for the unique trip with its package. For the train ride, trains have been allocated from the main departures and destinations frequently. The number of trains and their schedules is predefined by the company according to the fast experience. The train will stop at each station along the train route.
For example, if the tourist reserves a trip from Station 'A' to 'D 'along the B and C stations, the train ride between A to B is considered as a Trip Leg. Simply the train will depart at Station A on the scheduled time and arrive at Station B. For the above-considered trip, there are three Trip Legs that can be identified as a Trip Leg. Then it departs from Station B and arrives at Station C. It can be considered as another TripLeg. Then It departs from station C and arrives at the station D. It is the last TripLeg for the package. Then the tourist will go to the destination.
Considering the above trip A-D is the trip code for the trip, the trip package plans to cover 4 stations. The trip has three Trip Legs such as A to B, B to C and C to D. and also these trips have two intermediate stops such as B and C station.

Every trip includes records for the number of TripLegs including the following details

- **ID** - Unique id for the TripLeg
- **TripId** - Unique Trip Id that trip leg belong to
- **First Departure Location** - First Departure Location of the considered trip
- **Last Departure Location** - Last Departure Location of the considered Trip
- **Departure Location** - Departure Station of the TripLeg for the considered Trip
- **Arrival Location** - Arrival station of the Trip Leg for the considered Trip
- **Departure time** - Departure date and time of the departure station according to the Train schedule
- **Arrival time** - Arrival date and time of the arrival station according to the Train Schedule

To determine the trip packages of the Flam Railway, the following steps were executed

- Determine the Unique TripIds and First Departure Location from the 'Trip' table.
- Select the TripLeg which are included in each and every TripId.
- Short the TripLegs according to the ascending order of the Departure time.

- Concatenate the First Departure Location with Sorted Arrival Locations with the '>' symbol.
- Created Trip Package was Appended to the particular TripID Record.

Example Scenario

Consider the Trip which has 1 for the TripId which Departs from Station A at 20/10/2019 09:00:00 along Station B and Station C to arrive at Station D at 20/10/2019 10:00:00 [Table 3.1]

TripID	First Dep. Locat.	Last Arrival Locat.	Trip Code	Depa. Time	Arrival Time	PriceIn (USD)
1	A	D	A-D	20/10/2019 09:00:00	20/10/2019 10:00:00	100

Table 3.1: Sample Trip Data

Then all the TripLegs are selected and sorted according to the Departure Time. [Table 3.2]

ID	TripID	First Dep. Locat.	Last Arrival Locat.	Dep. Locat.	Arrival Locat.	Dep. Time	Arrival Time
1	1	A	D	A	B	20/10/2019 09:00:00	20/10/2019 09:20:00
2	1	A	D	B	C	20/10/2019 09:20:00	20/10/2019 09:40:00
3	1	A	D	C	D	20/10/2019 09:40:00	20/10/2019 10:00:00

Table 3.2: Sample TripLeg Data

Then the First Departure Location and Arrival Location of the Sorted TripLegs were concatenated with > symbol
A > B > C > D

Created Package was appended to the considered Trip.

TripID	First Dep. Locat.	Last Arrival Locat.	Trip Code.	Dep. Time	Arrival Time	Price IN	Package
1	A	D	A-D	20/10/19 09:00:00	20/10/19 10:00:00	100	A > B > C > D

Table 3.3: Sample Trip Data with Trip Packages

By considering the above scenario 89 Unique Trip Packages were identified

Trip Price Identification Flam Railway Trip Packages provides facilities such as meals, beverages and bike rides for the tourist. If the tourists consume these facilities within the train rides, they have included the charges for those facilities for the price of the train ticket. As the facilities can be noise for the data set, it is needed to remove those additional charges from the train ticket. As the base price of the train ticket has not been provided due to the company policies, under the guidance of the Flam Railway, the Base Price for the train rides was assumed under the following assumption.

Assumptions

1. At least one trip has been recorded for the particular day and particular packages with the base price of the trip package without any additional charges.
2. The minimum price should be considered as the base price of all the trips of the considered day and considered package.

Base Price Identification
To assign the base price of the trip package for the particular date, the below steps are followed.

1. Select a trip package
2. Select a single day for the selected trip package.
3. Select all the trips for the selected day and trip package.
4. Identify the minimum price from selected trips
5. Assign the minimum price for all the selected trips.

Example Scenario

TripID	First Dep. Locat.	Last Arrival Locat.	Trip Code.	Dep. Time	Arrival Time	Price IN	Package
1	A	D	A-D	20/10/19 09:00:00	20/10/19 10:00:00	100	A > B > C > D
2	A	D	A-D	20/10/19 09:00:00	20/10/19 10:00:00	120	A > B > C > D
3	A	D	A-D	20/10/19 09:00:00	20/10/19 10:00:00	110	A > B > C > D

Table 3.4: Sample Trip Data

By following the above steps, base trip prices can be determined as the table given below.

TripID	First Dep. Locat.	Last Arrival Locat.	Trip Code.	Dep. Time	Arrival Time	Price IN	Package
1	A	D	A-D	20/10/19 09:00:00	20/10/19 10:00:00	100	A > B > C > D
2	A	D	A-D	20/10/19 09:00:00	20/10/19 10:00:00	100	A > B > C > D
3	A	D	A-D	20/10/19 09:00:00	20/10/19 10:00:00	100	A > B > C > D

Table 3.5: Sample Updated Trip Data

3.4.7 Implementing Models

Implementing machine learning models

In this step following models that are determined as most successful to determine a better price will be implemented.

- Deep Learning Model
- Support Vector Regression
  - RBF Model
  - Linear Model
  - Polynomial Model
- Extreme Learning Machine

Implementing statistical models

In these steps following statistical models are used in determining a better price with greater accuracy.

- Ordinary Least Squares Multiple Linear Regression
  - Using sklearn
  - Using statsmodels
- Seasonal ARIMA model with Exogenous Variables (SARIMAX)

3.4.8 Comparing results

Each model which was implemented was tested for the accuracy of price prediction. Results of the testing can be found in the Chapter 4 Results and Evaluation. The model that has the lowest root mean square error was considered the best in predicting the price for each package.

Deep Neural Network

Deep Neural Network was implemented using Tensorflow and Keras using Python language. In the first phase a classification deep neural network was created with 183 input nodes, 256, 128, 64, 32, 16, 8, 16, 32, 64 nodes in 1st to 9th hidden layers and 120 output nodes. This model was able to predict the price with 21.74% validation accuracy.
In the second phase a regression single layer neural network was created with the intention of assessing the effect of number of nodes in the layer on the accuracy of prediction. Neural Network was trained for number of nodes from 1 to 340 and the testing was abandoned as the researchers couldn't find any significant relationship between number of nodes and the prediction accuracy. Pearson correlation coefficient for the above relationship was -0.13738 which suggests there is no correlation between number of nodes and the prediction accuracy and validates the researchers' observation. Maximum validation accuracy gained by this model was 10.0128%. Following Accuracy vs Nodes graph shows all the observations gathered by the researchers (Figure 3.3).

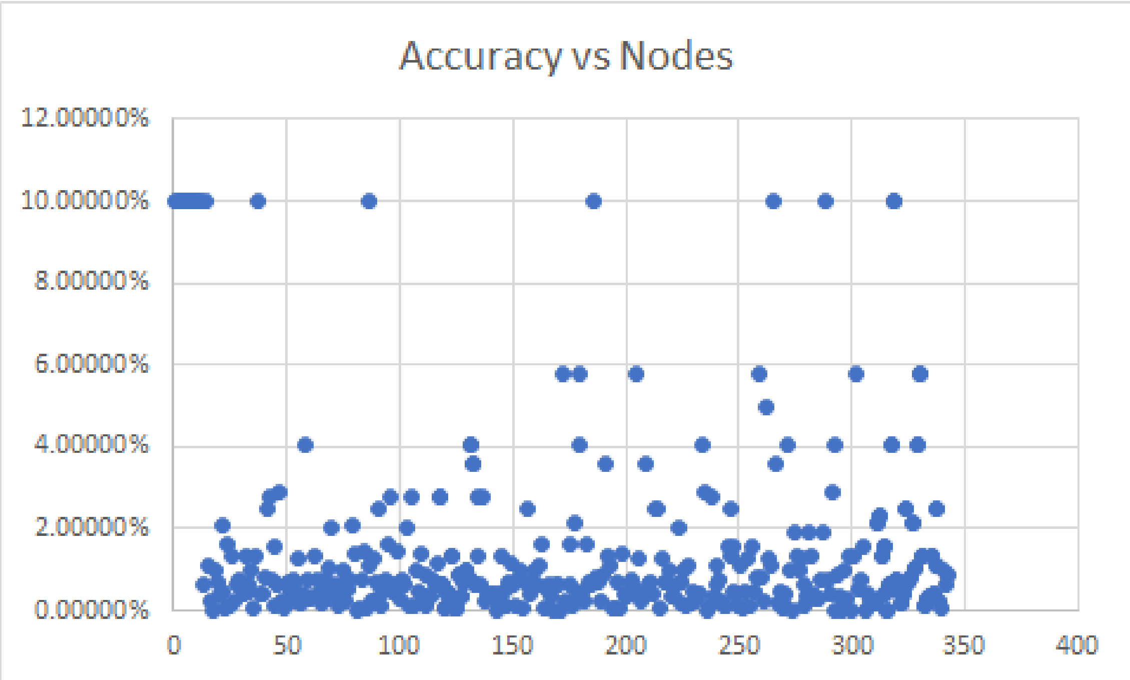

Figure 3.3: Accuracy vs Nodes for Neural Network

After that it was determined that there are too many redundant variables in the dataset. Correlation matrix of the numeric variables are as follows (Figure 3.4).

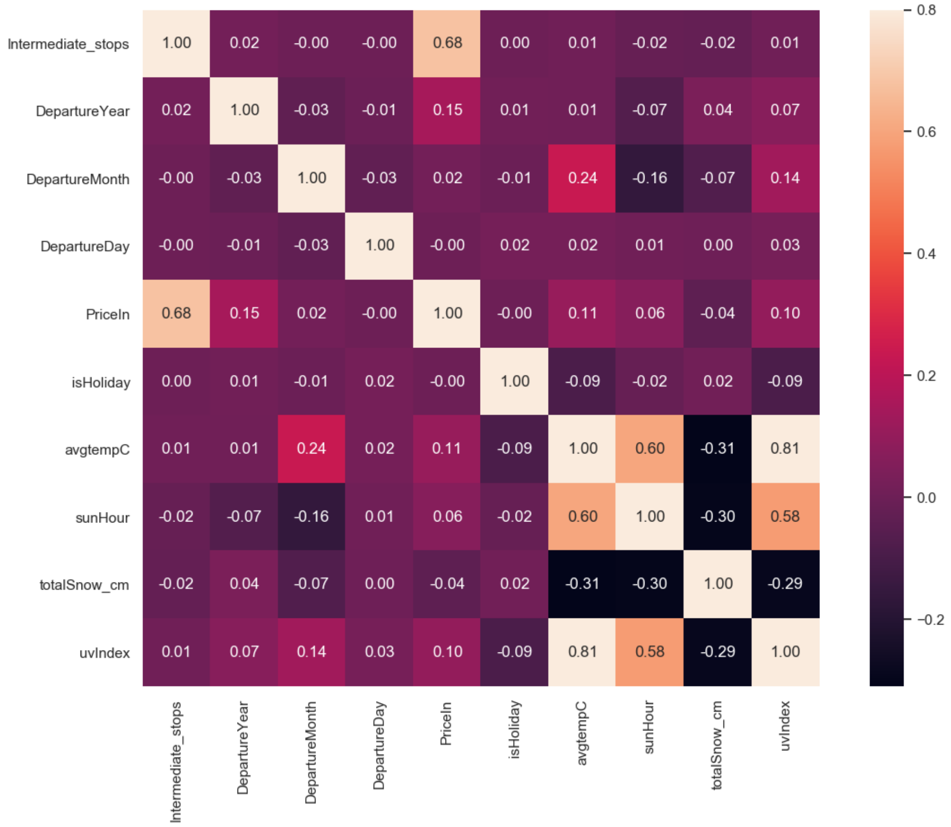

Figure 3.4: Correlation Matrix for Numeric Features

Next version of the Deep Neural Network used Standard Scalar to scale each attribute such that the mean is 0 and standard deviation is 1. It had 11 nodes in the input layers and 1024, 512, 256 nodes in hidden layers and 1 node in the output layer. A dropout layer of 0.2 was used between each hidden layer.
Unlike in the previous case instead of the whole dataset, model was trained for each package as it was determined to have much less root mean square error compared with using whole dataset.

Ordinary Least Squares Multiple Linear Regression (sklearn)

Linear regression model used here is a multiple linear regression model.Multiple linear regression was used because we have more than one explanatory variable to forecast the trip package price. In order to build the Ordinary Least Squares Multiple Linear Regression model the main library used here is sklearn.sklearn and statsmodels are the most popular libraries used in doing multiple linear regression. Other libraries used are pandas and numpy libraries were used and was written using python.In Ordinary Squares Multiple Linear Regression in sklearn, it fits a linear model with coefficients w = (w1, . . . , wp) to minimize the residual sum of squares between the observed targets in the dataset and the targets which are predicted by the linear approximation[29].
The dataset was split for training and testing as testing dataset size being 20% and the training dataset size being 80% of the whole dataset respectively.To predict the price explanatory variables can also be given sample values as shown below and then by using Ordinary Least Squares Multiple Linear Regression model a price is predicted.

```
# prediction with sklearn
New_PackageNo = 83
New_Intermediate_stops = 0
New_DepartureDay = 1
New_DepartureMonth = 1
New_DepartureYear = 2018
New_SeasonNo = 4
New_isHoliday = 0
New_avgtempC = 4
New_sunHour = 1.5
New_totalSnow_cm = 0
New_uvIndex = 1

print ('Predicted Trip Price: \n', regr.predict([[New_PackageNo ,New_Intermediate_stops,New
```

Figure 3.5: Manual Price Prediction

Ordinary Least Squares Multiple Linear Regression (statsmodels))

In statistics, ordinary least squares (OLS) is a type of linear least squares method for estimating the unknown parameters in a linear regression model. Under the additional assumption that the errors are normally distributed, OLS is the least squares estimator.Under the additional assumption that the errors are normally distributed, OLS could also be the maximum likelihood estimator.Thereby Ordinary Least Squares model can be considered as a method of Linear Regression Model.
Ordinary Least Squares model is also called OLS model.Unlike sklearn, statsmodels library doesn't automatically fit a constant. Therefore it is needed to use the method sm. add_constant(X) is used in order to add a constant.Adding a constant while it is not necessary will make the line fit much better[30].Other than sklearn statsmodels is the most popular linear regression library and it uses sm from statsmodels library for prediction as shown below. (Figure 3.6).

```
# OLS model
model = sm.OLS(Y, X).fit()
predictions = model.predict(X) # make the predictions by the model

print(Y)
print(predictions)

# Print out the statistics
print_model = model.summary()
print(print_model)
```

Figure 3.6: OLS Model

SARIMAX Model (Time series model)

The Time series model chosen was the SARIMAX model because the forecasting of trip package prices need to be done by considering season and trend details.
In order to build the SARIMAX model matplotlib, pandas, numpy and seaborn libraries were used and were written using python.
Similar to the linear regression model, training and testing sizes were split as 80% and 20% of the whole dataset. Also the data was scaled in order to get a more accurate result. SARIMA model with exogenous variables(SARIMAX) was used as we forecast the

price based on weather,season and holiday factors which need to be input as exogenous variables to get a better price[31].
Today's Trip package price is predicted by considering weather,season and holiday data of today and the manual price given by Flam railway. This is how the time series model works. The trip package 'BER->BLO' below shows how the season,trend, trip package of the next day and the correlation changes.

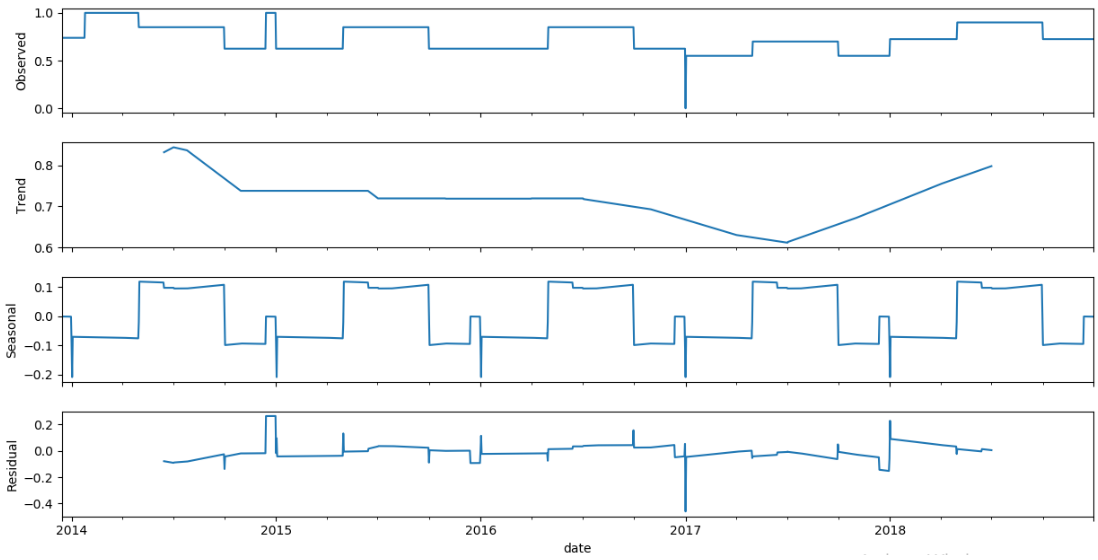

Figure 3.7: Trends of the Data Set

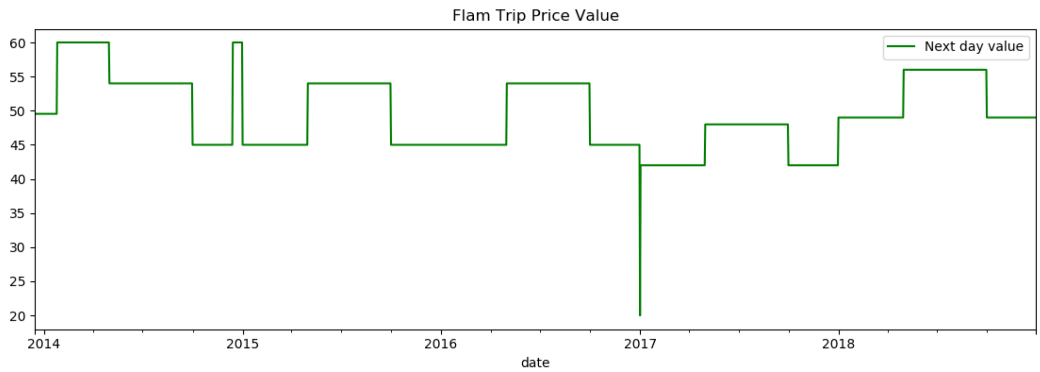

Figure 3.8: Trip Prices with Years

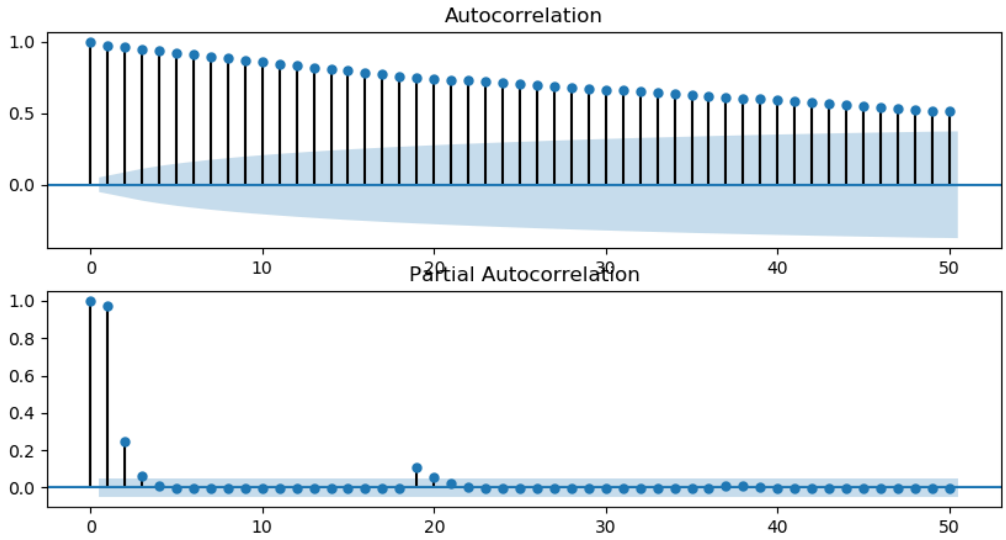

Figure 3.9: Correlation

Time series frequency chosen was daily as we have details for a trip package daily with the daily prices assigned for the trip packages. But we cannot use SARIMAX model to predict for the future dates outside the dataset.

Support Vector Regression (SVR)

Support Vector Regression can be used to predict the prices by analyzing the pattern of the impact from the external parameters.The data set is separated according to every package acquire better results. By training the data set by separately, it is planned to identify the error of the predicted price from selected parameter. According to the Correlation Matrix on figure , it can be assumed that the Temperature is the most correlated factor for the price (due to the intermediate stops are unique can constant for the Trip Package, it is not considered as a factor for the single package). Rather than temperature, Sun hour, Uv Index and snowfall is also trained to predict the prices. The prediction of prices are done by three sub models.

1. RBF Model
2. Linear Model
3. Polynomial Model

Extreme Learning Machine (ELM)

When the data set was analyzed by the extreme learning Machine (ELM), It takes only significantly less amount of time rather than the other machine learning algorithms because of the slow learning speed of feedforward neural networks due to the slow gradient-based learning algorithm and all the parameters of the learning algorithm are tuned iteratively. To avoid the bottlenecks ELM uses Single Hidden Layer Feedforward Neural Networks (SLFN) to analyze a huge amount of data set. It will be suitable for the research and for the domain as the data set is around 3.1 Million rows. Rather than other

Model, ELM can use to predict the Price Class (Suitable Existing Price) by considering the External Factors.

3.4.9 Create Hybrid Model

Hybrid model was created by combining Deep Learning Model and Ordinary Least Square Model to predict prices and a Linear Interpolation function with deflated price and sales volume to find the better price. Workflow of predicting a better price is shown in the Figure 3.10.

Reason for not using the SARIMAX model even though it showed the least error will be discussed in the Chapter 5 Discussion.

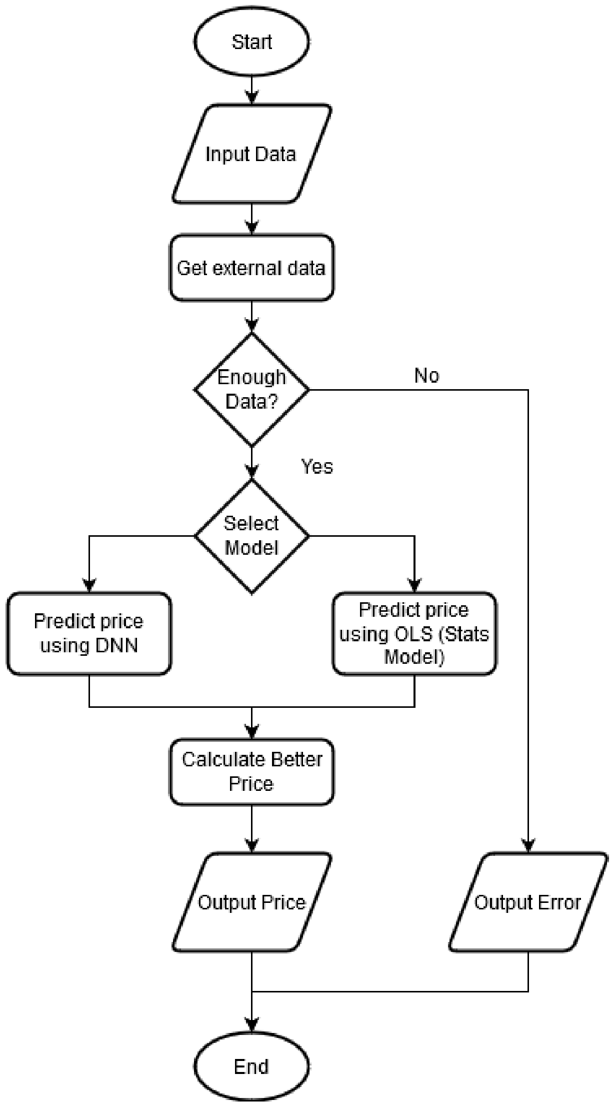

Figure 3.10: Better Price Prediction Workflow with Hybrid Model

Hybrid model consists of 2 parts. First part predict the current manual price use by Flam railway and second part calculate a better price.

Predicting Price

In this part the hybrid model will try to predict the current price of the package using the date, package name and external data. A deep leaning model is used to predict the price of 75 packages and Ordinary Least Squares Multiple Regression (statsmodels) for other 3. Price of 11 packages was not predicted due to reasons explained in the next sub section.

Finding better price

For finding the better price a linear interpolation function was used. Here unique deflated price of packages is used as independent variable and number of sales for that price as dependent variable. Prices were adjusted for inflation of year 2018 as it was the latest year with data.

Selecting the interpolation method

To select the best interpolation method a cross validation was by using all the datapoints available for each package. First a single data point was selected. Then the linear, quadratic and cubic interpolation functions were generated using the remaining datapoints. After that the volume of sales was predicted for the excluded datapoint. Finally the interpolation function with least root mean square error was selected as the best interpolation method for the package. Table 3.6 shows the first 15 packages with their cross validation errors for first 2 data points. Full table is availabe at apendix C.1.

Package	Y	Linear	Quadratic	Cubic	Best	Y	Linear	Quadratic	Cubic	Best
1	1678	2927.292458	247.3242173	-971.3438714	Linear	428	630.907586	446.9102504	294.9050239	Cubic
2	612	760.0048635	805.7245094	818.8497447	Linear	702	554.9179144	527.5024993	514.0433839	Linear
3	2832	1673.54457	1881.316074	1886.089955	Cubic	856	1020.589891	451.0405268	186.3861543	Linear
4	1071	1608.919702	1659.462885	2188.876186	Linear	3933	3779.138584	21456.92725	20282.96467	Linear
5	6327	6041.646039	6743.653407	7331.62891	Linear	7428	5514.779074	6987.336571	5665.709636	Cubic
6	9235	7050.286626	9506.516722	12186.89348	Cubic	2078	3815.914021	1195.066723	1272.034776	Cubic
7	1598	6397.613038	6399.229691	6559.21008	Linear	7428	6483.440491	10836.71049	11527.45009	Linear
8	462	503.1079136	526.1410184	519.089279	Linear	154	2.373883194	-88.57103999	-162.9289225	Linear
9	565	486.1001443	619.74017	603.129535	Cubic	462	503.1079136	526.1410184	519.089279	Linear
10	2078	3490.896354	-27848.04281	-34116.62576	Linear	1598	7343.764182	2064.417958	594.6150855	Cubic
11	3327	8416.475012	9604.258319	9499.681608	Linear	7613	6271.931792	10834.29968	10932.78687	Linear
12	2078	3815.914021	1194.312427	1271.297821	Cubic	1598	2251.001157	7482.573861	27841.16288	Linear
13	7434	1656.615844	1120.078586	1588.241597	Linear	7428	6420.759346	10760.89732	11421.66464	Linear
14	7494	6946.720566	7605.825912	7054.4543	Cubic	3882	2254.734888	2688.620822	3441.902219	Cubic
15	8874	7006.100413	7532.71676	7870.757289	Cubic	7494	5744.265868	7117.824008	7018.959041	Cubic

Table 3.6: Error of each interpolation function for each package

Total error of the first 15 packages and best interpolation method for each package is shown in the Table 3.7. Full table can be found in appendix C.2.

Package	Linear	Quadratic	Cubic	Best
1	257.6511294	487.3490563	2082.462483	Linear
2	240.2507302	328.1442973	330.0676115	Linear
3	326.0287906	1559.55462	1538.775012	Linear
4	326.0287906	1559.55462	1538.775012	Linear
5	863.705575	3585.009762	29386.15524	Linear
6	772.9680212	1089.771352	2575.497169	Linear
7	1023.235541	2002.988142	1976.859088	Linear
8	47.54122617	57.90440917	63.95030834	Linear
9	47.54122617	57.90440917	63.95030834	Linear
10	875.419008	3695.96075	30264.59594	Linear
11	863.7221972	3585.016645	29386.1566	Linear
12	773.0902434	1089.811636	2575.186374	Linear
13	1010.839392	1989.416855	1961.91566	Linear
14	669.4675446	858.8526582	854.4173713	Linear
15	473.0578001	1780.095517	1675.92193	Linear

Table 3.7: Total error of interpolation methods and best interpolation function for each package

Table 3.7 shows that the interpolation method with least error is Linear Interpolation. If the package contained one or two data points cross validation method can't be applied. Therefore no interpolation method was selected for those 11 packages.

Calculating a better price Figure 3.11 show the overview of the better price calculating process. First the input price was adjusted for inflation[32] and then checked to confirm that it is inside the bounds of the interpolation function. Then using linear interpolation, sales for the input price was calculated. Using the sales and the price current revenue was estimated.Then the price is iteratively incremented to get the maximum revenue better than the current revenue until the upper bound of the interpolation function reached. Then the price which will give a better revenue is returned as better price after readjusting for the inflation. If no better revenue can be calculated current price is assumed to be the best price.

Here only the prices which are higher than the predicted price was used because researchers were unable to gain access to data about cost of each package. Therefore if the better price is actually less than the cost of package, there will be a loss of revenue instead of projected increase of revenue. By considering higher prices this risk can be mitigated.

When there is only one or two data points it is not possible to find a better price using this method. Therefore 11 packages with one or two data points was considered to have insufficient data for price prediction.

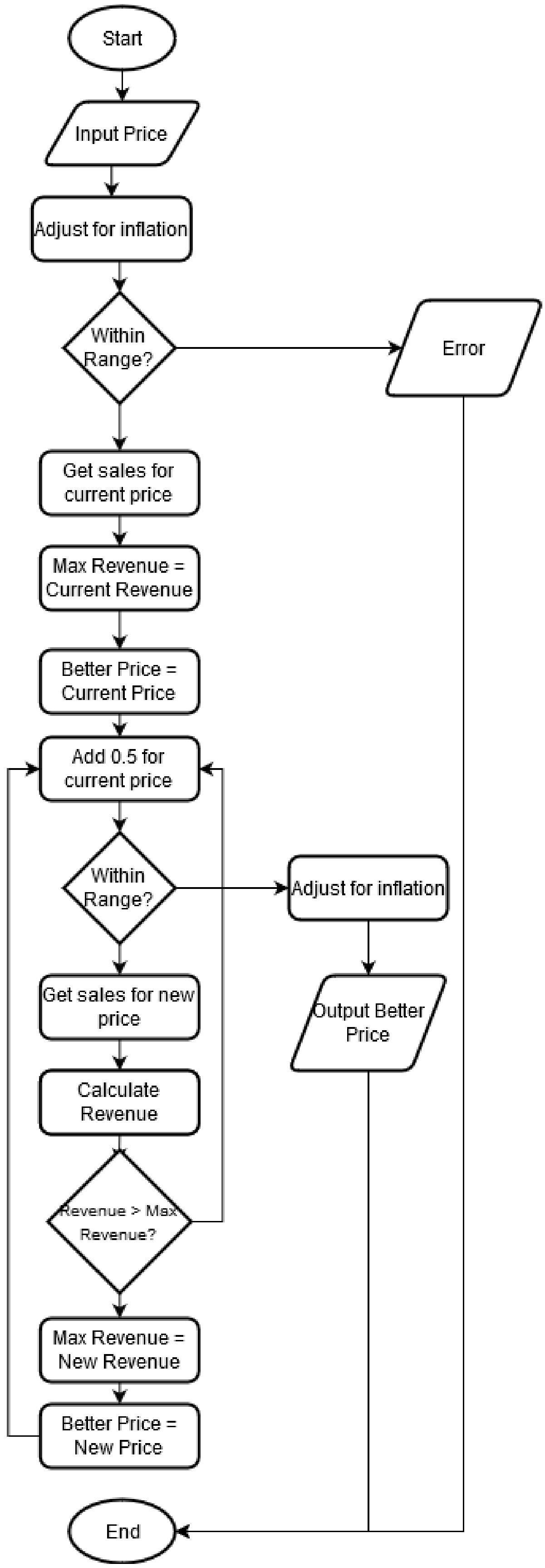

Figure 3.11: Process of calculating better price

3.4.10 Testing the Hybrid Model

The hybrid model was tested by comparing the revenue from unique prices of each package to revenue from using the correspondingly better price. Table 3.8 shows the revenue increase percentage of first 15 packages. Full table can be found in the appendix C.3.

More about the performance of hybrid model is discussed in the Chapter 4 Results and Evaluation.

Package Number	Percentage Increase of Revenue
1	47.93%
2	50.00%
3	100.27%
4	100.27%
5	84.80%
6	120.59%
7	91.12%
8	34.65%
9	34.65%
10	85.77%
11	84.81%
12	120.58%
13	90.09%
14	72.04%
15	112.63%

Table 3.8: Percentage increase of revenue from better prices

Chapter 4

Results and Evaluation

4.1 Introduction

We aimed at finding a better price for trip packages of Flam Railway by considering the external factors weather, season and holidays. In order to study how other factors affect the price we studied and plotted the relationship between other main factors to the price.

4.2 Identify Trip Packages

For the purpose of categorizing the locations where customers travel uniquely, we determined a special category called trip packages which explains the sequence of the route or the stations that customers ordered as a trip. These trips are uniquely identified by "TripID" and it was generated automatically when a customer booked a trip. We derived the Distinct TripIds from the provided database and there were over 3.1 Million TripIds recorded. After that, the First Departure location and the Last Departure location were found among all stations that customers selected as the start and the end stations of the journey. We also found the Arrivals and Departure Locations that have been recorded among the journey which was described in the sample table given below (Figure 4.1).

	Id	FirstDepartureLocation	LastArrivalLocation	DepartureLocation	ArrivalLocation
1	2276221	MYR	FLAM	REI	KJO
2	2276221	MYR	FLAM	HAR	LUN
3	2276221	MYR	FLAM	LUN	FLAM
4	2276221	MYR	FLAM	MYR	VAT
5	2276221	MYR	FLAM	VAT	REI
6	2276221	MYR	FLAM	KJO	BLO
7	2276221	MYR	FLAM	BER	HAR
8	2276221	MYR	FLAM	BLO	BER
9	2276222	MYR	FLAM	BER	HAR
10	2276222	MYR	FLAM	REI	KJO

Figure 4.1: Sample Trip Data

To determine the unique trip packages, initially, the distinct TripIds and their first

departure location were selected. Then the first departure location was taken as the first record from departure location and then the corresponding arrival location was taken. Then that arrival location was used to find the next departure location and then the corresponding arrival location was taken. Likewise according to that algorithm until the last arrival location is the Arrival location the loop would run and finally the trip package could be retrieved.
By processing all 3.1 million distinct Trip Ids from the 12.4 million Ids, 89 Trip Packages were identified.

4.3 Identify the impact on external factors

To identify more on how the external factors weather, season and holidays vary with the price and the relationships of them to the price we have generated graphs as below.
We took the mean price as the y-axis as it would show a clear value and a graph for each and every value from the external factor in the x-axis.

Firstly to get an idea of how the variables relate and vary according to the price we plotted how mean price varied with the weather factors average temperature, total snowfall, amount of sun hours and UV index as shown below.

Temperature
Below shows how mean Price varies with Average Temperature in Celsius (Figure 4.2).

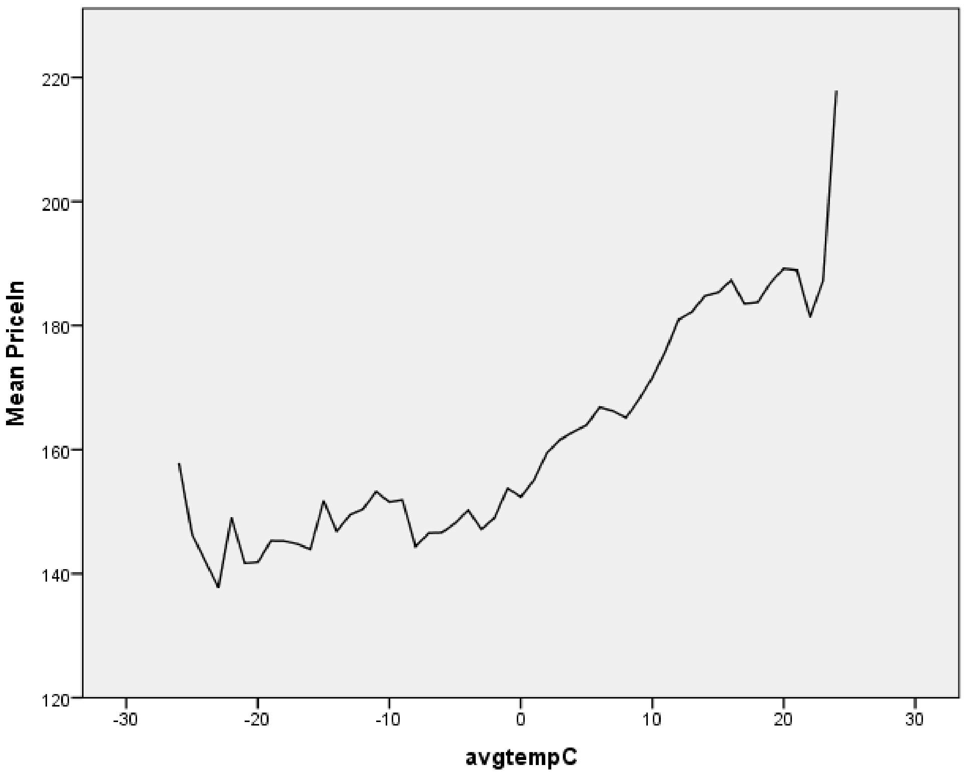

Figure 4.2: Average Temperature Vs Mean Price

This (Figure 4.2) shows that when the average temperature is less than 0 price doesn't increase or decrease drastically and is in the same range but when the average temperature increases from 0 price increases.

Snow Fall
Below shows how mean Price varies with Total SnowFall in cm (Figure 4.3).

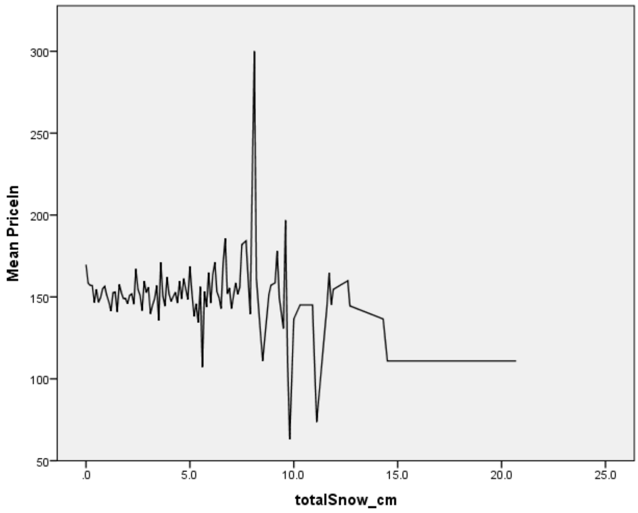

Figure 4.3: Snow Fall Vs Mean Price

.
This (Figure 4.3) shows that Price has increased rapidly when Total SnowFall in cm is about 8.0cm.

Sun Hour
Below shows how mean Price varies with Amount of sun hours (Figure 4.4).

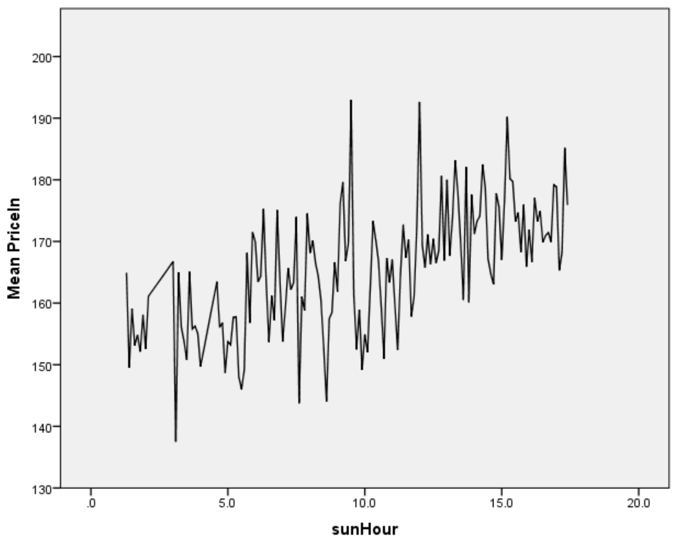

Figure 4.4: Sun Hour Vs Mean Price

According to different amounts of sun-hours mean price has varied between 130 and 200.

UV Index

Below shows how mean Price varies with the UV index.
According to different values of the UV index mean price has varied between 150 and 200 (Figure 4.5).

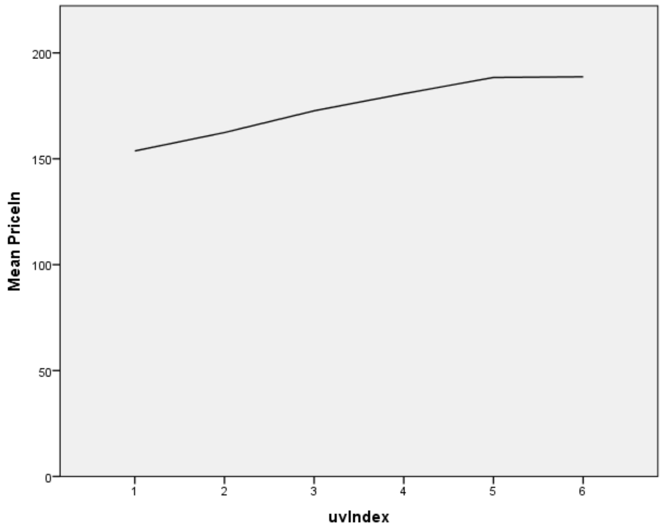

Figure 4.5: UV Index Vs Mean Price

.

Season

Next, it was identified how the mean price varies with the season. For better visualization, a bar graph is used (Figure 4.6).

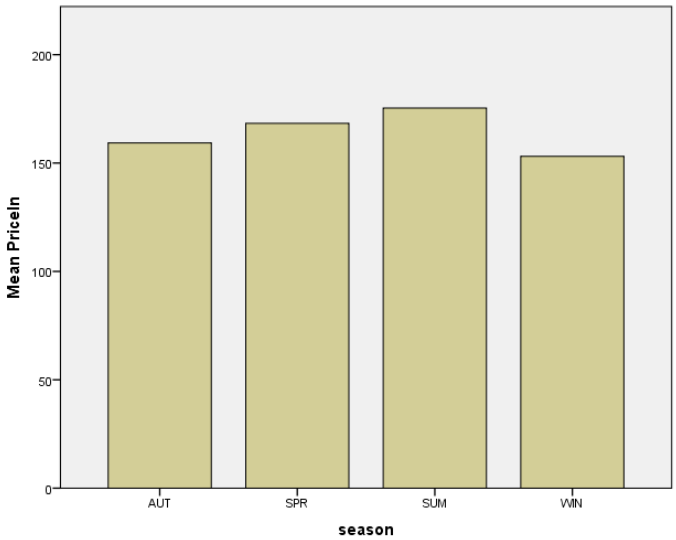

Figure 4.6: Season Vs Mean Price

.

From this (Figure 4.6), it can say that Trip Package prices are higher in summer when compared to other seasons.

Holidays

Next, it was identified how the mean price varies with the trip date being a holiday or not. For better visualization, a bar graph is used (Figure 4.7).

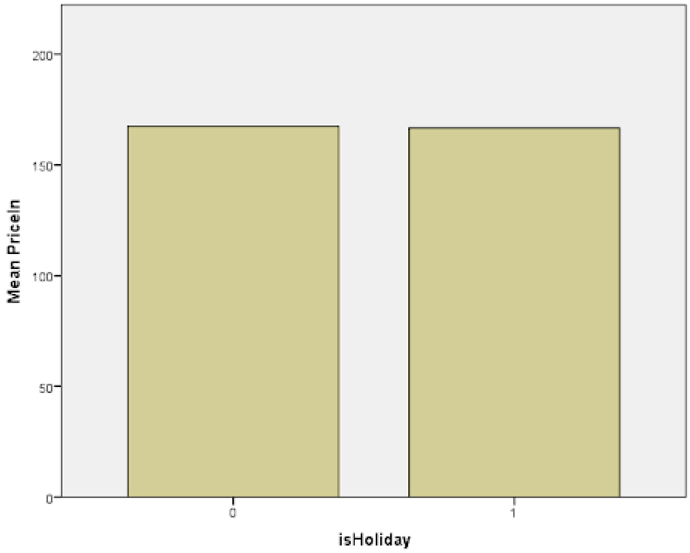

Figure 4.7: Holiday Vs Mean Price

.

From this (Figure 4.7), it can clearly say that prices have not increased because of holidays. The main reason for the results provided above, it is only consider the reserved trip data but not the requests for the reservation. According to the Flam Railway [1], as a highly tourist attractive there is a high demand for the reservations. Because of that, Trips are reserved everyday without any shortage or difference on holidays. Because of that, considering the trip reservation data, price variations for Holidays couldn't be observable. (By considering reservation requests for the each and every trips, variations for the demands for trip reservations would be observable.)

4.4 Data Analyzing

Root Mean Square error (RMSE),Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) for each package are as below

Deep Neural Network

Package	**Root Mean Square Error**	**Mean Absolute Error**	**Mean Absolute Percentage Error**
BER>BLO	1.858566	1.468043	3.152602
BER>BLO>KJO>REI	6.484673	4.624062	3.104236
BER>BLO>KJO>REI>VAT	3.873239	2.846733	2.861113
BER>BLO>KJO>REI>VAT>MYR	4.521151	2.849314	2.791244
BER>HAR	1.713967	1.399042	2.895124
BER>HAR>LUN	1.978418	1.599791	1.948576
BER>HAR>LUN>FLAM	2.374	1.813088	2.042161
BER>KJO>VAT	5.219659	3.38342	3.176681
BER>KJO>VAT>MYR	4.859553	3.03486	2.898397
BLO>BER	0.826601	0.478076	0.966286
BLO>BER>HAR	0.646029	0.470578	0.944132
BLO>BER>HAR>LUN	3.014169	2.468362	3.163834
BLO>BER>HAR>LUN>FLAM	2.292972	1.735622	1.925343
BLO>KJO>REI	4.908975	3.830123	2.427808
BLO>KJO>REI>VAT	2.04234	1.417867	1.305954
BLO>KJO>REI>VAT>FLAM	3.007701	2.502922	2.46558
BLO>KJO>REI>VAT>MYR	2.6565	1.821656	1.659015
FLAM>BER	4.486399	3.393734	3.348277
FLAM>BER>KJO>VAT	4.286934	3.172611	1.887499
FLAM>BER>KJO>VAT>MYR	5.150316	3.463431	2.431822

Table 4.1: Results of Deep Neural Network

The Full table can be found on Table B.1 on Appendix Chapter

Ordinary Least Squares Multiple Linear Regression(sklearn)

Average estimation of increase of revenue when using forecasted better prices was 79.25%

Package	Mean Absolute Error (MAE)	Root Mean Square Error (RMSE)	R Square
BER>BLO	3.458736	4.33187	0.269019
BER>BLO>KJO>REI	8.00214	11.15019	0.382836
BER>BLO>KJO>REI>VAT	25.15824	29.10985	0.527894
BER>BLO>KJO>REI>VAT>MYR	25.34503	29.35536	0.51995
BER>HAR	2.567015	3.27245	0.435294
BER>HAR>LUN	3.592226	4.49435	0.739918
BER>HAR>LUN>FLAM	13.56142	15.76643	0.31188
BER>KJO>VAT	15.2408	18.88602	0.190285
BER>KJO>VAT>MYR	15.09802	18.71001	0.205172
BLO>BER	2.52495	3.22201	0.450156
BLO>BER>HAR	2.566653	3.27242	0.435275
BLO>BER>HAR>LUN	3.592673	4.49432	0.739917
BLO>BER>HAR>LUN>FLAM	13.5749	15.78416	0.312912
BLO>KJO>REI	7.352974	9.37399	0.350237
BLO>KJO>REI>VAT	21.07545	25.08915	0.586398
BLO>KJO>REI>VAT>FLAM	0	1.03E-10	1
BLO>KJO>REI>VAT>MYR	21.19356	25.27063	0.578075
FLAM>BER	8.174771	9.65797	0.224025
FLAM>BER>KJO>VAT	12.16878	14.42549	0.199307
FLAM>BER>KJO>VAT>MYR	1.3228	4.08657	0.844064

Table 4.2: Results of Ordinary Least Squares Multiple Linear Regression(sklearn)

The Full table can be found on Table B.2 on Appendix Chapter

Ordinary Least Squares Multiple Linear Regression(statsmodels)

Package	Mean Absolute Error (MAE)	Root Mean Square Error (RMSE)	R Square
BER>BLO	3.314898	4.33243	0.229472
BER>BLO>KJO>REI	9.368491	11.15069	0.310618
BER>BLO>KJO>REI>VAT	34.5362	29.11086	0.128623
BER>BLO>KJO>REI>VAT>MYR	34.56456	29.3563	0.127731
BER>HAR	3.017193	3.27251	0.247479
BER>HAR>LUN	7.014595	3.27265	0.112332
BER>HAR>LUN>FLAM	13.37857	15.76653	0.295916
BER>KJO>VAT	13.13729	18.90905	0.072426
BER>KJO>VAT>MYR	12.94204	18.89474	0.097285
BLO>BER	3.106777	3.22213	0.25324
BLO>BER>HAR	3.017148	3.27258	0.247491
BLO>BER>HAR>LUN	7.014578	4.49448	0.112317
BLO>BER>HAR>LUN>FLAM	13.38688	15.78449	0.297101
BLO>KJO>REI	7.460728	9.37418	0.343274
BLO>KJO>REI>VAT	31.92491	25.08987	0.153661
BLO>KJO>REI>VAT>FLAM	12.28308	0	0.277542
BLO>KJO>REI>VAT>MYR	31.83661	25.27103	0.148097
FLAM>BER	8.219025	9.66347	0.222679
FLAM>BER>KJO>VAT	12.25438	14.42642	0.143115
FLAM>BER>KJO>VAT>MYR	8.553088	4.08795	0.052731

Table 4.3: Results of Ordinary Least Squares Multiple Linear Regression(statsmodels)

The Full table can be found on Table B.3 on Appendix Chapter

SARIMAX Model (Time Series)

Package	**Mean Absolute Error (MAE)**	**Root Mean Square Error (RMSE)**	**R Square**
BER>BLO	0.05265	0.12028	0.23756
BER>BLO>KJO>REI	0.10209	0.24315	0.34674
BER>BLO>KJO>REI>VAT	0.04518	0.10202	0.84451
BER>BLO>KJO>REI>VAT>MYR	0.04367	0.09955	0.85193
BER>HAR	0.03970	0.08657	0.55951
BER>HAR>LUN	0.02577	0.05824	0.96797
BER>HAR>LUN>FLAM	0.10334	0.25002	0.53191
BER>KJO>VAT	0.03800	0.14137	0.83426
BER>KJO>VAT>MYR	0.03871	0.13932	0.83905
BLO>BER	0.03811	0.08360	0.58922
BLO>BER>HAR	0.038016	0.08410	0.58429
BLO>BER>HAR>LUN	0.02577	0.05824	0.96797
BLO>BER>HAR>LUN>FLAM	0.10334	0.25002	0.53191
BLO>KJO>REI	0.06873	0.14836	0.74077
BLO>KJO>REI>VAT	0.06178	0.14925	0.71885
BLO>KJO>REI>VAT>FLAM	0.00165	0.02334	0.99550
BLO>KJO>REI>VAT>MYR	0.06027	0.14593	0.73120
FLAM>BER	0.03324	0.12098	0.78200
FLAM>BER>KJO>VAT	0.00758	0.04490	0.97012
FLAM>BER>KJO>VAT>MYR	0.01046	0.04759	0.95526

Table 4.4: Results of SARIMAX Model

The Full table can be found on Table B.4 on Appendix Chapter

Support Vector Regression (SVR)

The results of the Price prediction for separate packages using three models are given below.

Package	**RBF Model Root Mean Square Error (RMSE)**	**Liner Model Root Mean Square Error (RMSE)**	**Polynomial Model Root Mean Square Error (RMSE)**
BER>BLO	6.32198	5.71895	6.32507
BER>BLO>KJO>REI	19.09588	16.91226	17.66532
BER>BLO>KJO>REI>VAT	45.50209	44.69185	44.44279
BER>BLO>KJO>REI>VAT>MYR	50.22534	47.39201	49.77362
BER>HAR	6.14854	5.24405	5.23937
BER>HAR>LUN	10.59189	9.60826	9.4346
BER>HAR>LUN>FLAM	21.30226	21.78177	21.47419
BER>KJO>VAT	24.96192	23.87669	24.29451
BER>KJO>VAT>MYR	47.65939	26.88151	26.96798
BLO>BER	5.99866	5.38207	5.01755
BLO>BER>HAR	6.14854	5.24405	5.23937
BLO>BER>HAR>LUN	10.59189	9.60826	9.4346
BLO>BER>HAR>LUN>FLAM	21.30226	21.78177	21.47419
BLO>KJO>REI	16.85568	14.87762	15.36246
BLO>KJO>REI>VAT	16.85568	14.87762	15.36246
BLO>KJO>REI>VAT>FLAM	24.11035	24.18478	21.38035
BLO>KJO>REI>VAT>MYR	48.36985	47.83985	49.39595
FLAM>BER	48.36985	47.83985	49.39595
FLAM>BER>KJO>VAT	16.71602	16.75301	16.84424
FLAM>BER>KJO>VAT>MYR	44.8251	11.07076	11.66367

Table 4.5: Results of Support Vector Regression

The Full table can be found on Table B.5 on Appendix Chapter

Extreme Learning Machine (ELM)

Package	**Accuracy of Prediction**
BER>BLO	0.32596685
BER>BLO>KJO>REI	0.23756906
BER>BLO>KJO>REI>VAT	0.38674033
BER>BLO>KJO>REI>VAT>MYR	0.24861878
BER>HAR	0.32596685
BER>HAR>LUN	0.33149171
BER>HAR>LUN>FLAM	0.37016575
BER>KJO>VAT	0.22857143
BER>KJO>VAT>MYR	0.2
BLO>BER	0.37016575
BLO>BER>HAR	0.29281768
BLO>BER>HAR>LUN	0.32596685
BLO>BER>HAR>LUN>FLAM	0.3480663
BLO>KJO>REI	0.25414365
BLO>KJO>REI>VAT	0.40883978
BLO>KJO>REI>VAT>FLAM	0.45454545
BLO>KJO>REI>VAT>MYR	0.3038674
FLAM>BER	0.25714286
FLAM>BER>KJO>VAT	0.45714286
FLAM>BER>KJO>VAT>MYR	0.14285714

Table 4.6: Results of Extreme Learning Machine (ELM)

The Full table can be found on Table B.6 on Appendix Chapter

Chapter 5

Discussion

5.1 Introduction

Train travelling is increasing day by day as it is a cost effective and also a quicker way of travelling. People in European regions now like to go on trips by train. Therefore train trip offering companies has a good opportunity. For these companies in order to get good revenue they need to identify a better price for each and every trip package they offer. To cater this we have used machine learning and statistical models to predict a better price.

5.2 Findings

- SARIMAX Model was the model with the least mean square error. But it was not suitable for the prediction of future prices and involve significant amount of assumptions when used.
- There were 75 packages whose price was best predicted using deep neural network.
- There were 3 packages whose price was best predicted using Ordinary Least Squares Multiple Linear Regression(statsmodels).
- There were 11 packages with too little data to get an estimation for a better price.
- Maximum estimation of increase of revenue when using forecasted better prices was 120.59%
- Minimum estimation of increase of revenue when using forecasted better prices was 120.12%
- Average estimation of increase of revenue when using forecasted better prices was 79.25%

5.3 Discussion

According to our dataset the number of records for each trip package and the number of distinct prices in each trip package varies as below (Figure 5.1 and Figure 5.2).

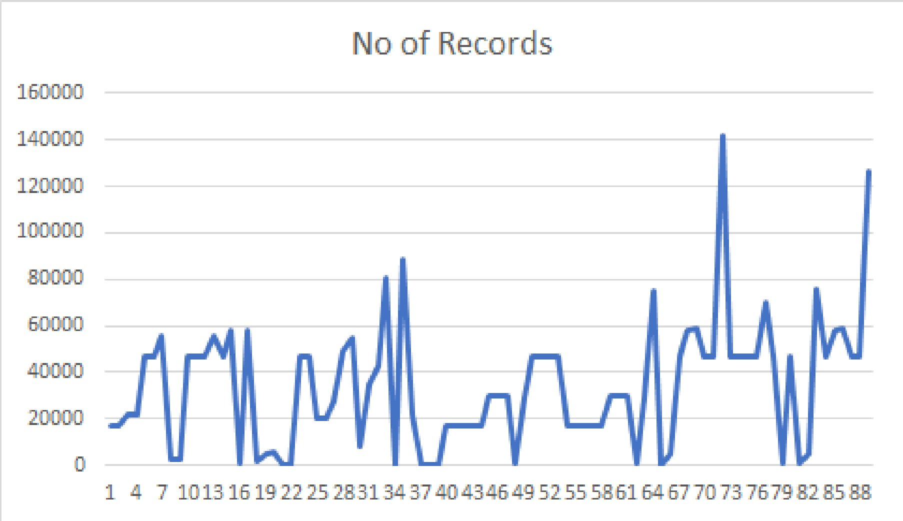

Figure 5.1: Number of Trip Records

The Full table can be found on Table A.1 on Appendix Chapter

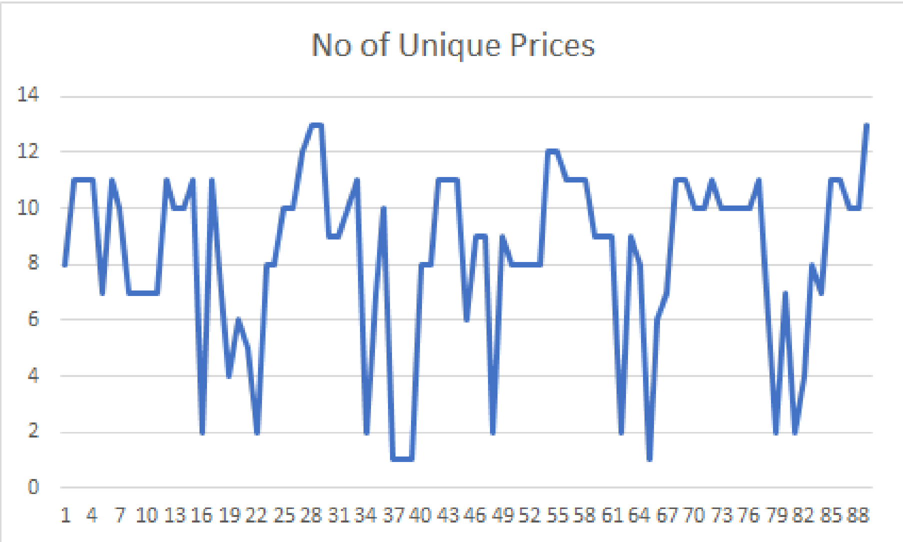

Figure 5.2: Distinct Prices

The Full table can be found on Table A.2 on Appendix Chapter

When Pearson's Correlation Coefficient calculated the relationship between the number of records for each package and number of unique prices for each package with Root Mean Square Error, a coefficient of -0.43658 and -0.3557 received. This means that there is a weak connection between Root Mean Square Error and number of records for each package and Root Mean Square Error and number of unique prices for each package.
Below shows how mean absolute error and root mean square error change for each of the models we implemented.

Deep Neural Network

When considering the results of deep neural network, line graphs of Root Mean Square Error Mean Absolute Error and Mean Absolute Percentage Error are as follows (Figure 5.3,Figure 5.4 and Figure 5.5).

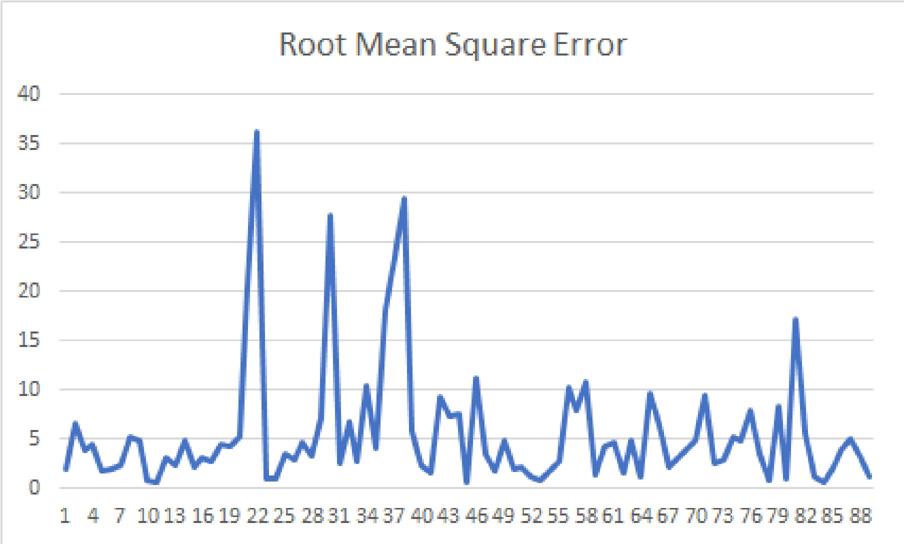

Figure 5.3: RMSE of Deep Neural Network

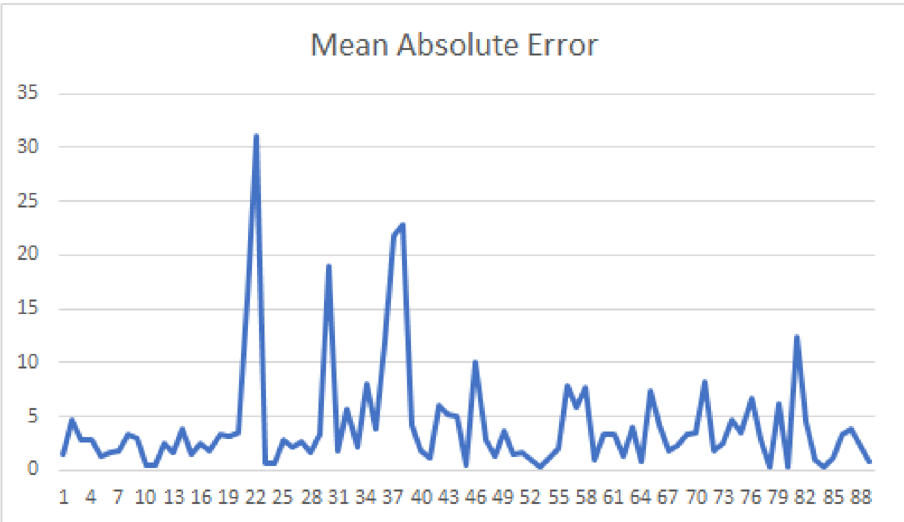

Figure 5.4: MAE of Deep Neural Network

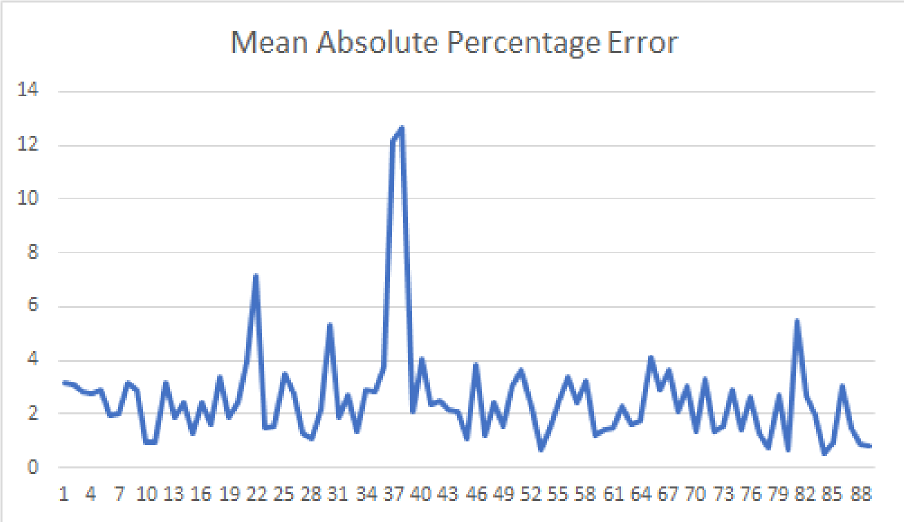

Figure 5.5: MAPE of Deep Neural Network

Ordinary Least Squares Multiple Linear Regression(sklearn)
When considering the results of Ordinary Least Squares Multiple Linear Regression(sklearn), the line graphs of Root Mean Square Error Mean Absolute Error and Mean Absolute Percentage Error are as follows (Figure 5.6 , Figure 5.7 and Figure 5.8).

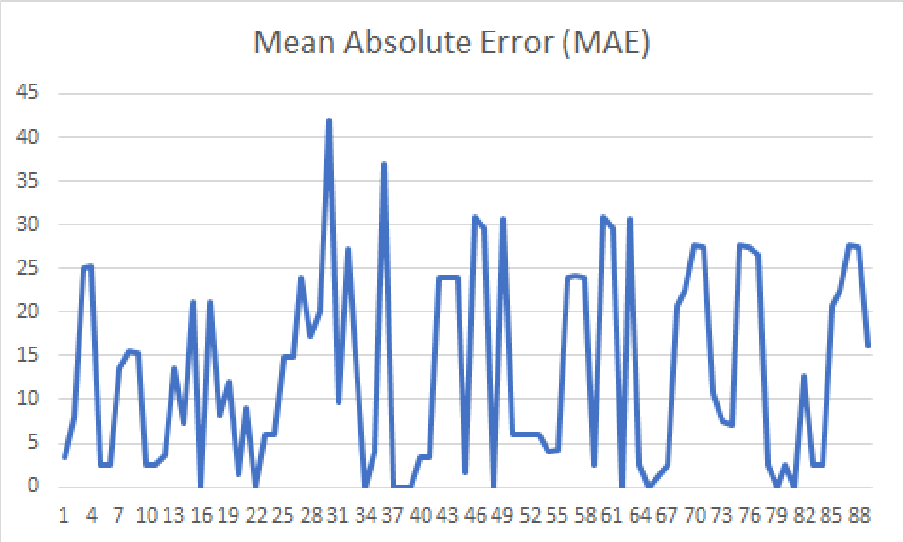

Figure 5.6: MAE of Ordinary Least Squares Multiple Linear Regression(sklearn) Model

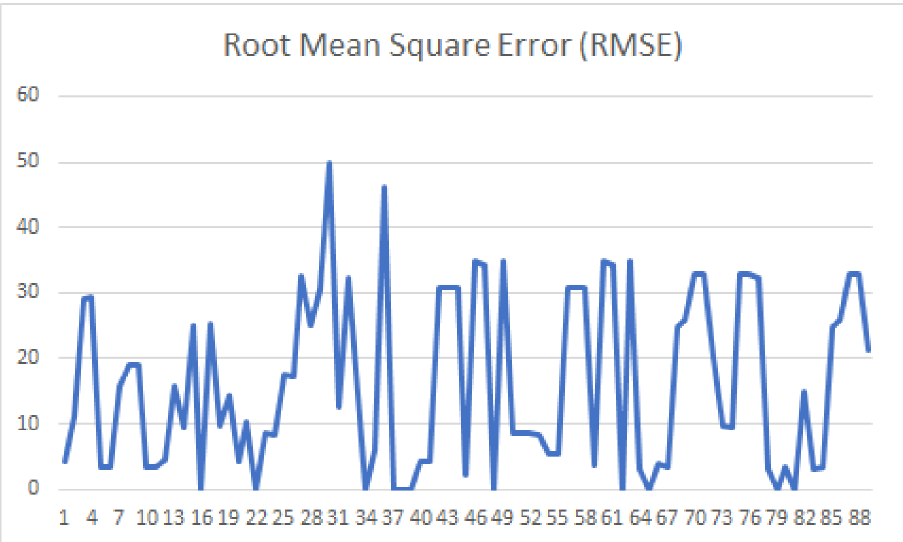

Figure 5.7: RMSE of Ordinary Least Squares Multiple Linear Regression(sklearn) Model

Ordinary Least Squares Multiple Linear Regression(statsmodels)

When considering the results of Ordinary Least Squares Multiple Linear Regression(statsmodels), the line graphs of Root Mean Square Error Mean Absolute Error and Mean Absolute Percentage Error are as follows (Figure 5.8).

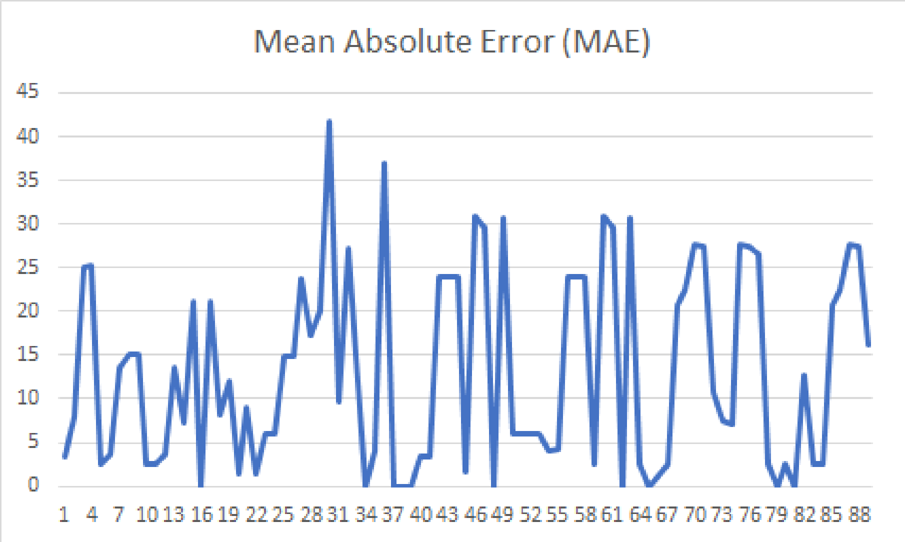

Figure 5.8: MAE of Ordinary Least Squares Multiple Linear Regression(statsmodels)

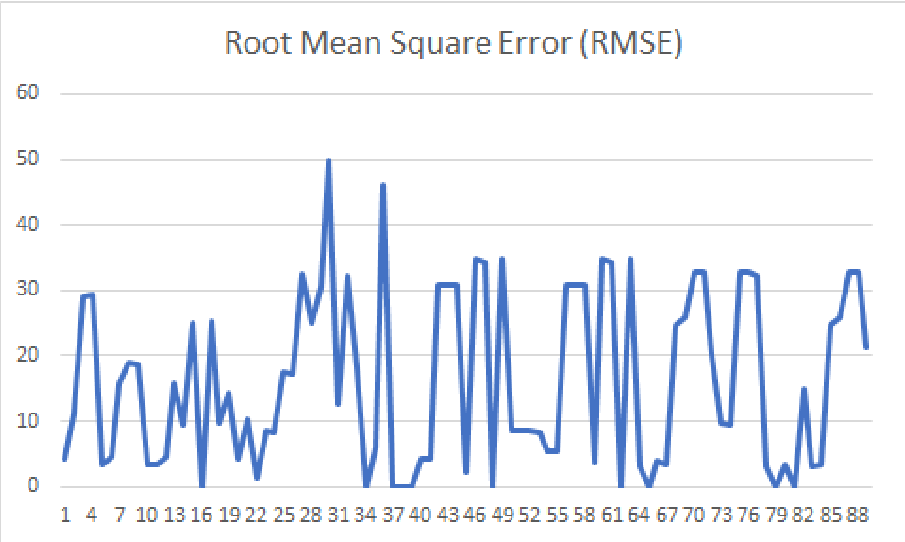

Figure 5.9: RMSE of Ordinary Least Squares Multiple Linear Regression(statsmodels)

SARIMAX (Time series)

When considering the results of SARIMAX model which is a time series model, the line graphs of Root Mean Square Error Mean Absolute Error and Mean Absolute Percentage Error are as follows (Figure 5.10 , Figure 5.11).

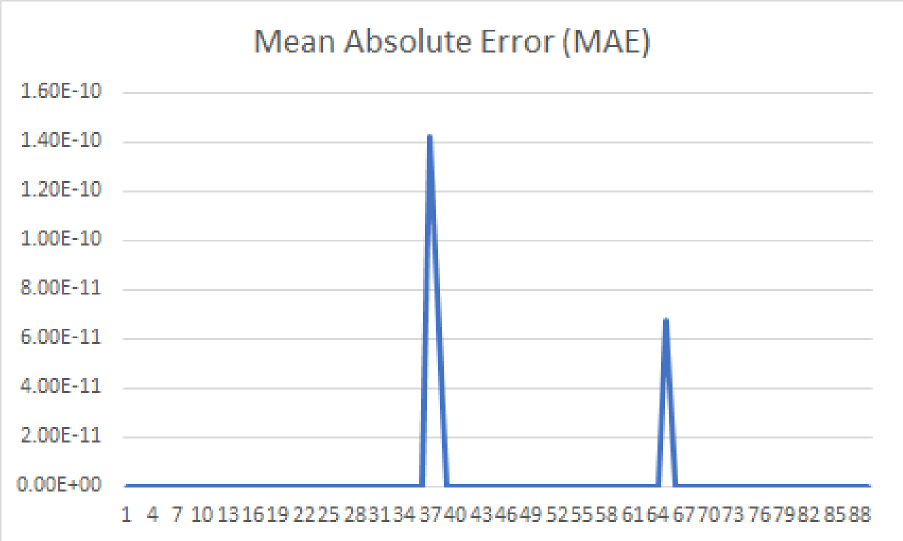

Figure 5.10: MAE of SARIMAX

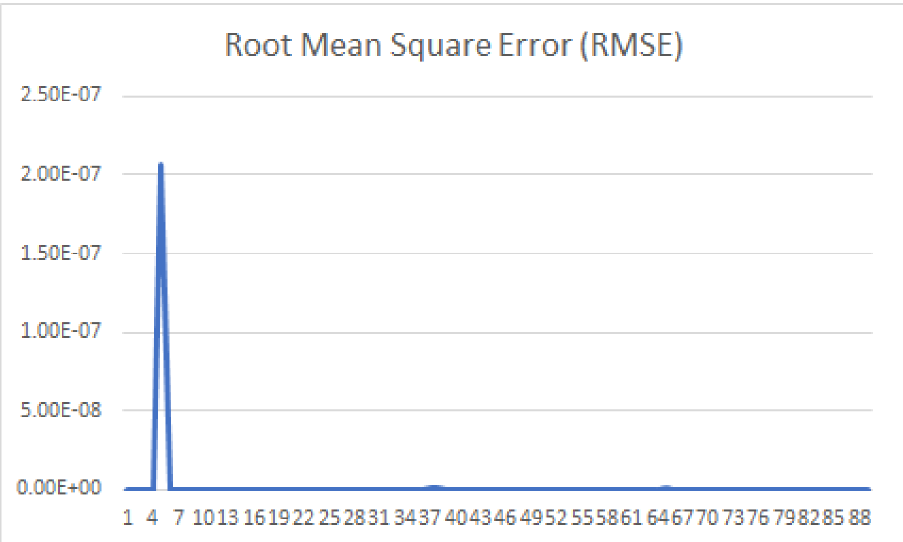

Figure 5.11: RMSE of SARIMAX

Support Vector Regression (SVR)

In Support Vector Regression there are 3 sub models and below shows how Root Mean Square error changes for the 3 sub models.

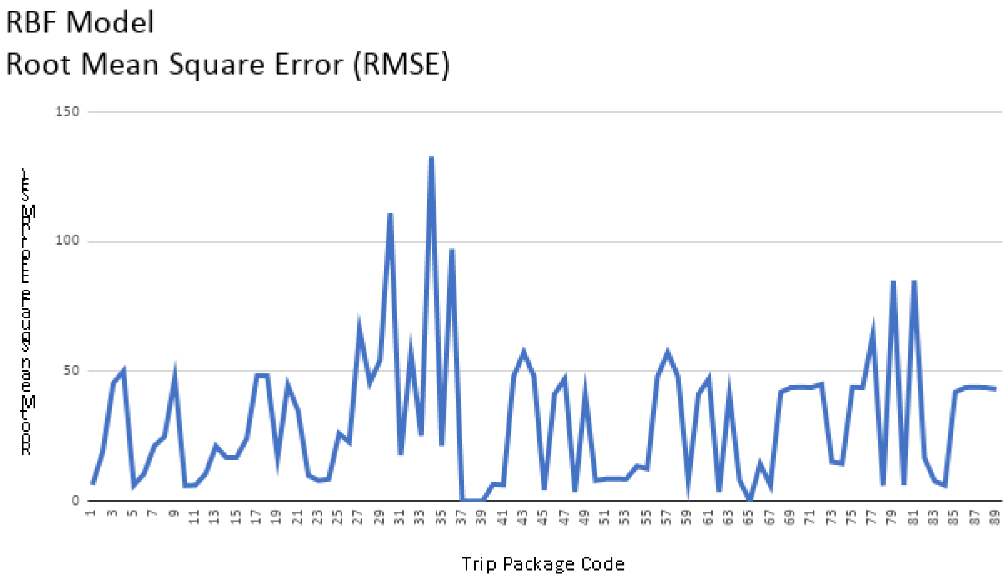

Figure 5.12: RMSE of RBF Model of SVR

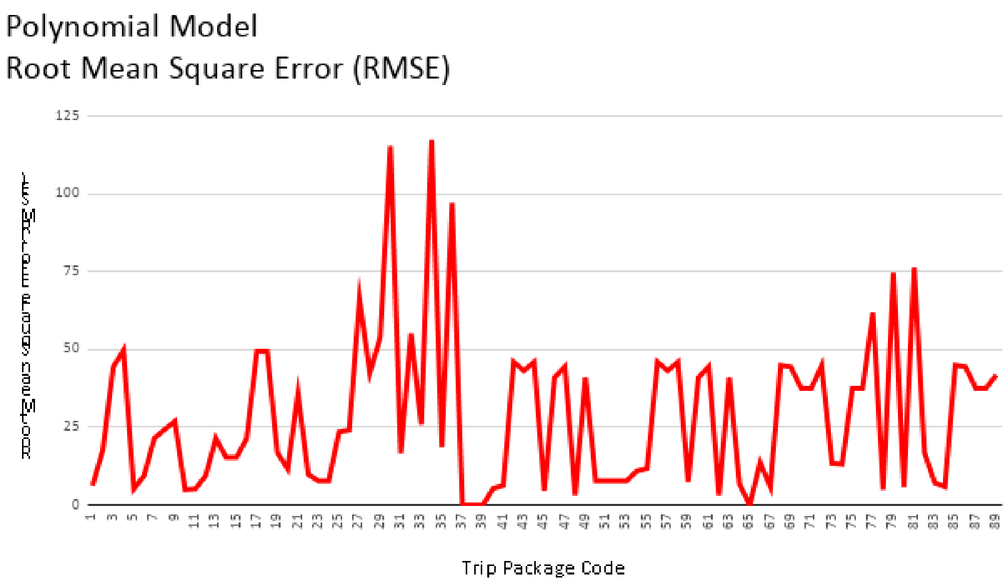

Figure 5.13: RMSE of Polynomial Model of SVR

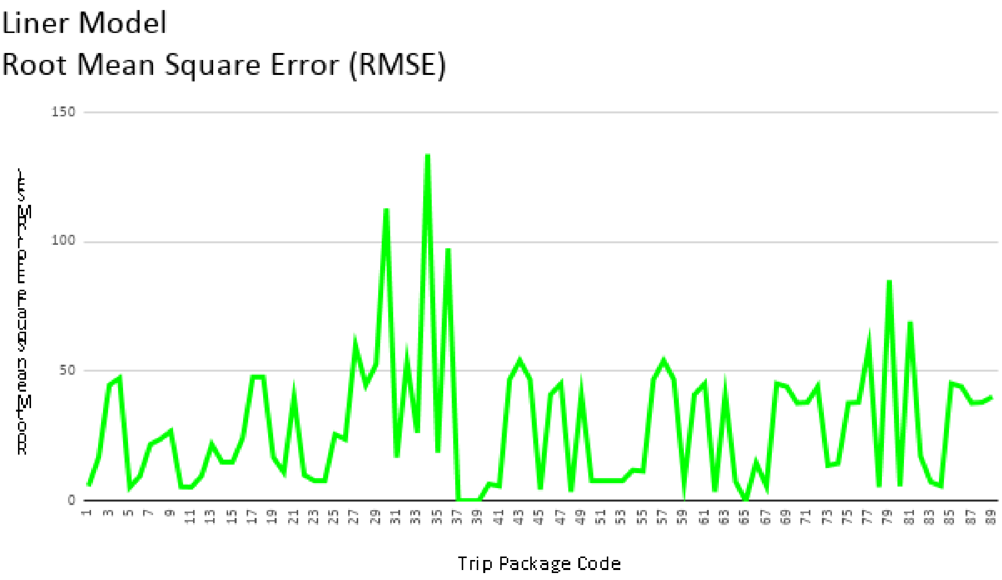

Figure 5.14: RMSE of Linear Model of SVR

Extreme Learning Machine (ELM)

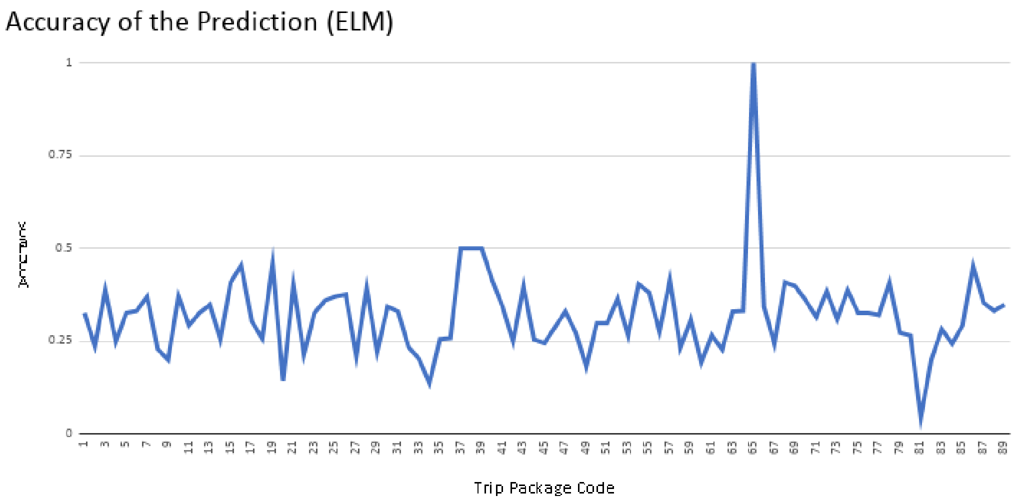

Figure 5.15: RMSE of Linear Model of SVR

Hybrid Model

Hybrid model consists of two parts. First part will decide which model to be used to price prediction and predict a price for the package for the given day. This predicted price is the equivalent of the manual price allocated by the Flam railways staff.

The two models used in this part is the deep neural network and Linear Regression model implemented with statsmodels library.

The second part of the hybrid model will calculate a better price than the predicted price using past sales data and deflated prices using a linear interpolation function.

This hybrid model is cpable to

5.4 Comparison

Based on the above models below shows the final comparison of the models using the Root Mean Square Error and the best suitable model to predict the trip price for each trip package.

ID	Package	DNN	SVR RBF	SVR Linear	SVR Polynomial	Linear Regression (statsmode	Linear Regression (sklearn)	SARIMAX	Min Error	Model	Min Error Without SARIMAX	Best Model Except SARIMAX
1	BER>BLO	1.8585656	6.32198	5.71895	6.32507	4.33243	4.33187	0.12028	0.12028	SARIMAX	1.858565598	DNN
2	BER>BLO>KJO>REI	6.48467258	19.0959	16.9123	17.66532	11.15069	11.15019	0.24315	0.24315	SARIMAX	6.484672578	DNN
3	BER>BLO>KJO>REI>VAT	3.87323933	45.5021	44.6919	44.44279	29.11086	29.10985	0.10202	0.10202	SARIMAX	3.873239332	DNN
4	BER>BLO>KJO>REI>VAT>MYR	4.52115085	50.2253	47.392	49.77362	29.3563	29.35536	0.09955	0.09955	SARIMAX	4.52115085	DNN
5	BER>HAR	1.713967	6.14854	5.24405	5.23937	3.27251	3.27245	0.08657	0.08657	SARIMAX	1.713967003	DNN
6	BER>HAR>LUN	1.97841815	10.5919	9.60826	9.4346	3.27265	4.49435	0.05824	0.05824	SARIMAX	1.978418145	DNN
7	BER>HAR>LUN>FLAM	2.37399954	21.3023	21.7818	21.47419	15.76653	15.76643	0.25002	0.25002	SARIMAX	2.373999544	DNN
8	BER>KJO>VAT	5.21965878	24.9619	23.8767	24.29451	18.90905	18.88602	0.14137	0.14137	SARIMAX	5.219658782	DNN
9	BER>KJO>VAT>MYR	4.85955276	47.6594	26.8815	26.96798	18.89474	18.71001	0.13932	0.13932	SARIMAX	4.85955276	DNN
10	BLO>BER	0.82660103	5.99866	5.38207	5.01755	3.22213	3.22201	0.0836	0.0836	SARIMAX	0.826601027	DNN
11	BLO>BER>HAR	0.64602889	6.14854	5.24405	5.23937	3.27258	3.27242	0.0841	0.0841	SARIMAX	0.646028887	DNN
12	BLO>BER>HAR>LUN	3.01416935	10.5919	9.60826	9.4346	4.49448	4.49432	0.05824	0.05824	SARIMAX	3.014169348	DNN
13	BLO>BER>HAR>LUN>FLAM	2.29297223	21.3023	21.7818	21.47419	15.78449	15.78416	0.25002	0.25002	SARIMAX	2.292972226	DNN
14	BLO>KJO>REI	4.90897542	16.8557	14.8776	15.36246	9.37418	9.37399	0.14836	0.14836	SARIMAX	4.908975422	DNN
15	BLO>KJO>REI>VAT	2.04234032	16.8557	14.8776	15.36246	25.08987	25.08915	0.14925	0.14925	SARIMAX	2.042340322	DNN
16	BLO>KJO>REI>VAT>FLAM	3.00770121	24.1104	24.1848	21.38035	0	1.03E-10	0.02334	0	Linear Regression (statsmodel)	0	Linear Regression (statsmodel)
17	BLO>KJO>REI>VAT>MYR	2.6564999	48.3699	47.8399	49.39595	25.27103	25.27063	0.14593	0.14593	SARIMAX	2.656499901	DNN
18	FLAM>BER	4.48639859	48.3699	47.8399	49.39595	9.66347	9.65797	0.12098	0.12098	SARIMAX	4.486398595	DNN
19	FLAM>BER>KJO>VAT	4.28693436	16.716	16.753	16.84424	14.42642	14.42549	0.0449	0.0449	SARIMAX	4.286934361	DNN
20	FLAM>BER>KJO>VAT>MYR	5.15031614	44.8251	11.0708	11.66367	4.08795	4.08657	0.04759	0.04759	SARIMAX	4.08657	Linear Regression (sklearn)
21	FLAM>BER>KJO>VAT>MYR>VAT>KJO>FLAM	20.3424932	34.8903	40.0512	35.95806	10.39904	10.37733	0.03501	0.03501	SARIMAX	10.37733	Linear Regression (sklearn)
22	FLAM>BER>KJO>VAT>MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	36.2223107	9.951	9.95038	9.89981	0	1.3678	0.0234	0	Linear Regression (statsmodel)	0	Linear Regression (statsmodel)
23	FLAM>LUN	1.04953939	7.97389	7.75363	7.79226	8.46827	8.46809	0.07053	0.07053	SARIMAX	1.049539386	DNN
24	FLAM>LUN>HAR	1.0573686	8.57759	7.732	7.78108	8.393	8.39286	0.07087	0.07087	SARIMAX	1.057368597	DNN
25	FLAM>LUN>HAR>BER	3.55641568	26.0964	25.6568	23.54452	17.47837	17.47779	0.32512	0.32512	SARIMAX	3.556415676	DNN

Figure 5.16: Comparison Chart 1

26	FLAM>LUN>HAR>BER>BLO	2.81293173	22.7801	23.7857	24.10695	17.15539	17.15474	0.32768	0.32768	SARIMAX	2.812931733	DNN
27	FLAM>LUN>HAR>BER>BLO>KJO>REI	4.67348333	66.0205	59.9078	66.37645	32.57737	32.57474	0.02951	0.02951	SARIMAX	4.67348333	DNN
28	FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT	3.28081128	45.7666	44.8364	42.57949	25.14763	25.14734	0.04209	0.04209	SARIMAX	3.280811278	DNN
29	FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT>MYR	7.20204752	54.4851	52.7897	53.93302	30.5659	30.56546	0.03969	0.03969	SARIMAX	7.202047518	DNN
30	FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT>MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	27.680385	111.073	112.965	115.4014	49.98332	49.97189	0.1071	0.1071	SARIMAX	27.680385	DNN
31	FLAM>LUN>HAR>BLO	2.55946591	17.7519	16.7382	16.64998	12.54078	12.53978	0.13469	0.13469	SARIMAX	2.559465915	DNN
32	FLAM>LUN>HAR>BLO>KJO>REI	6.76635556	57.0818	53.8537	55.18569	32.12023	32.11961	0.04575	0.04575	SARIMAX	6.766355563	DNN
33	T	2.63775533	25.3106	26.319	25.92051	18.2222	18.22201	0.08097	0.08097	SARIMAX	2.637755328	DNN
34	FLAM>LUN>HAR>BLO>KJO>REI>VAT>FLAM	10.4958917	132.764	133.885	117.3839	0	2.04E-10	0.02334	0	Linear Regression (statsmodel)	0	Linear Regression (statsmodel)
35	FLAM>LUN>HAR>BLO>KJO>REI>VAT>MYR	4.03569768	21.3747	18.6711	18.67107	5.91971	5.91963	0.04272	0.04272	SARIMAX	4.035697681	DNN
36	FLAM>LUN>HAR>BLO>KJO>REI>VAT>MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	18.0900278	97.2033	97.582	97.17584	46.05223	46.05014	0.10885	0.10885	SARIMAX	18.09002784	DNN
37	FLAM>VAT	23.5693401	0	0	0	0	3.63E-14	3.31E-10	0	SVR RBF/ SVR Liner / SVR Polynomial /	0	Linear Regression (statsmodel)
38	FLAM>VAT>MYR	29.4633199	0	0	0	0	2.92E-14	6.67E-12	0	SVR RBF/ SVR Liner / SVR Polynomial /	0	Linear Regression (statsmodel)
39	FLAM>VAT>MYR>VAT>FLAM	5.88221103	0	0	0	0	5.38E-14	0.03294	0	SVR RBF/ SVR Liner / SVR Polynomial /	0	Linear Regression (statsmodel)
40	HAR>BER	2.29865368	6.56816	6.4953	5.38852	4.26252	4.26212	0.11937	0.11937	SARIMAX	2.29865368	DNN
41	HAR>BER>BLO	1.54285651	6.32198	5.71895	6.32508	4.33264	4.3323	0.12028	0.12028	SARIMAX	1.542856514	DNN
42	HAR>BER>BLO>KJO>REI	9.34062935	48.0325	46.8007	46.02116	30.69226	30.69029	0.02832	0.02832	SARIMAX	9.340629347	DNN
43	HAR>BER>BLO>KJO>REI>VAT	7.37792352	57.4225	54.2004	43.24679	30.68308	30.6822	0.02995	0.02995	SARIMAX	7.377923519	DNN
44	HAR>BER>BLO>KJO>REI>VAT>MYR	7.48821177	48.0325	46.8007	46.02116	30.68517	30.68365	0.02832	0.02832	SARIMAX	7.488211769	DNN

Figure 5.17: Comparison Chart 2

45	HAR>BLO	0.62385604	4.32108	4.31541	4.62342	2.14537	2.14491	0.04928	0.04928	SARIMAX	0.623856038	DNN
46	HAR>BLO>KJO>REI	11.2409129	41.3033	41.0659	41.08314	34.98612	34.98533	0.0362	0.0362	SARIMAX	11.24091289	DNN
47	HAR>BLO>KJO>REI>VAT	3.56431097	47.2614	45.2754	44.54644	34.2084	34.20656	0.03504	0.03504	SARIMAX	3.564310975	DNN
48	HAR>BLO>KJO>REI>VAT>FLAM	1.73217266	3.57209	3.4716	3.24752	0	9.77E-12	0.02334	0	Linear Regression (statsmodel)	0	Linear Regression (statsmodel)
49	HAR>BLO>KJO>REI>VAT>MYR	4.87625239	41.3033	41.0659	41.08314	34.8454	34.84482	0.0362	0.0362	SARIMAX	4.876252387	DNN
50	HAR>LUN	1.96801535	7.99863	7.78515	7.82409	8.51926	8.51911	0.07051	0.07051	SARIMAX	1.968015348	DNN
51	HAR>LUN>FLAM	2.16775915	8.60274	7.76582	7.81742	8.44698	8.44687	0.07085	0.07085	SARIMAX	2.167759149	DNN
52	LUN>FLAM	1.20357194	8.60274	7.76582	7.81742	8.44697	8.44687	0.07085	0.07085	SARIMAX	1.203571943	DNN
53	LUN>HAR	0.74647162	8.57759	7.732	7.78108	8.39301	8.39286	0.07087	0.07087	SARIMAX	0.746471618	DNN
54	LUN>HAR>BER	1.81793962	13.4922	11.852	10.96682	5.38053	5.37999	0.05917	0.05917	SARIMAX	1.817939619	DNN
55	LUN>HAR>BER>BLO	2.62286528	12.558	11.455	11.84734	5.48725	5.48698	0.05888	0.05888	SARIMAX	2.622865282	DNN
56	LUN>HAR>BER>BLO>KJO>REI	10.1822847	48.0325	46.8007	46.02116	30.69259	30.69146	0.02832	0.02832	SARIMAX	10.18228468	DNN
57	LUN>HAR>BER>BLO>KJO>REI>VAT	7.93478426	57.4225	54.2004	43.24679	30.68835	30.68753	0.02995	0.02995	SARIMAX	7.934784259	DNN
58	LUN>HAR>BER>BLO>KJO>REI>VAT>MYR	10.8722828	48.0325	46.8007	46.02116	30.69218	30.69029	0.02832	0.02832	SARIMAX	10.87228277	DNN
59	LUN>HAR>BLO	1.27242847	7.93038	7.41939	7.51247	3.51002	3.50978	0.02533	0.02533	SARIMAX	1.272428473	DNN
60	LUN>HAR>BLO>KJO>REI	4.2429816	41.3033	41.0659	41.08314	34.98697	34.98533	0.0362	0.0362	SARIMAX	4.242981598	DNN
61	LUN>HAR>BLO>KJO>REI>VAT	4.5715837	47.2614	45.2754	44.54644	34.20834	34.20722	0.03504	0.03504	SARIMAX	4.571583702	DNN
62	LUN>HAR>BLO>KJO>REI>VAT>FLAM	1.64062646	3.57209	3.4716	3.24752	0	4.70E-12	0.02334	0	Linear Regression (statsmodel)	0	Linear Regression (statsmodel)
63	LUN>HAR>BLO>KJO>REI>VAT>MYR	4.75171103	41.3033	41.0659	41.08314	34.84665	34.84565	0.0362	0.0362	SARIMAX	4.751711025	DNN
64	MYR>VAT	1.10324232	8.03924	7.36756	6.7296	3.19414	3.19406	0.04085	0.04085	SARIMAX	1.103242318	DNN
65	MYR>VAT>FLAM	9.66528372	0	0	0	0	2.35E-13	1.78E-09	0	SVR RBF/ SVR Liner / SVR Polynomial /	0	Linear Regression (statsmodel)
66	MYR>VAT>KJO>FLAM	6.43383606	14.4113	14.7696	13.71924	4.06011	4.05993	0.04676	0.04676	SARIMAX	4.05993	Linear Regression (sklearn)
67	MYR>VAT>REI	2.13636211	6.15244	5.75581	5.92813	3.23087	3.23085	0.08803	0.08803	SARIMAX	2.136362114	DNN
68	MYR>VAT>REI>KJO>BLO	3.09819869	41.9864	45.3167	44.93226	24.66787	24.66754	0.14498	0.14498	SARIMAX	3.098198686	DNN
69	MYR>VAT>REI>KJO>BLO>BER	3.91107698	43.8582	44.0742	44.42567	25.97117	25.9706	0.14799	0.14799	SARIMAX	3.911076975	DNN

Figure 5.18: Comparison Chart 3

70	MYR>VAT>REI>KJO>BLO>BER>HAR	4.73892896	43.9627	37.7619	37.59945	32.77882	32.77831	0.02408	0.02408	SARIMAX	4.738928957	DNN
71	MYR>VAT>REI>KJO>BLO>BER>HAR >LUN	9.49517569	43.8425	38.0032	37.60524	32.80756	32.80666	0.02415	9.4951757	DNN	9.49517569	DNN
72	MYR>VAT>REI>KJO>BLO>BER>HAR >LUN>FLAM	2.55931246	45.1157	44.0879	44.8914	20.13379	20.13327	0.06618	0.06618	SARIMAX	2.559312459	DNN
73	REI>KJO>BLO	2.8568719	15.1788	13.685	13.51189	9.61727	9.61705	0.14466	2.8568719	DNN	2.856871904	DNN
74	REI>KJO>BLO>BER	5.2596708	14.533	14.5042	13.22763	9.34145	9.34108	0.14476	5.2596708	DNN	5.259670795	DNN
75	REI>KJO>BLO>BER>HAR	4.80192525	43.9627	37.7619	37.59945	32.77826	32.77749	0.02408	0.02408	SARIMAX	4.801925246	DNN
76	REI>KJO>BLO>BER>HAR>LUN	7.90194648	43.8425	38.0032	37.60524	32.807	32.80638	0.02415	7.9019465	DNN	7.901946478	DNN
77	REI>KJO>BLO>BER>HAR>LUN>FLAM	3.56546795	64.3484	59.821	61.91822	32.26719	32.26658	0.02539	0.02539	SARIMAX	3.565467948	DNN
78	REI>VAT	0.70337584	5.99393	5.34335	4.97517	3.17165	3.17151	0.08337	0.08337	SARIMAX	0.70337584	DNN
79	REI>VAT>FLAM	8.35706839	84.8811	85.1288	74.71557	0	2.19E-11	0.02334	0	Linear Regression (statsmodel)	0	Linear Regression (statsmodel)
80	REI>VAT>MYR	0.89942177	6.14973	5.73058	5.90008	3.19906	3.19893	0.08776	0.08776	SARIMAX	0.899421775	DNN
81	VAT>FLAM	17.1471454	85.0696	69.3978	76.37011	0	2.42E-10	0.02329	0	Linear Regression (statsmodel)	0	Linear Regression (statsmodel)
82	VAT>KJO>FLAM	5.68194328	16.7688	17.3771	16.93783	14.90718	14.90529	0.04475	0.04475	SARIMAX	5.681943279	DNN
83	VAT>MYR	1.10480961	7.72999	7.2708	7.22816	3.15673	3.15665	0.05496	0.05496	SARIMAX	1.104809606	DNN
84	VAT>REI	0.58829711	6.15288	5.76	5.93294	3.23603	3.23591	0.08657	0.08657	SARIMAX	0.588297113	DNN
85	VAT>REI>KJO>BLO	2.10824443	42.0205	45.3453	44.96285	24.68203	24.68168	0.14406	0.14406	SARIMAX	2.108244426	DNN
86	VAT>REI>KJO>BLO>BER	3.88601963	43.8841	44.0972	44.46143	25.98228	25.98159	0.14914	0.14914	SARIMAX	3.886019626	DNN
87	VAT>REI>KJO>BLO>BER>HAR	4.96589968	43.9627	37.7619	37.59945	32.7778	32.77741	0.02408	0.02408	SARIMAX	4.965899682	DNN
88	VAT>REI>KJO>BLO>BER>HAR>LUN	3.30127273	43.8425	38.0032	37.60524	32.80719	32.80612	0.02415	0.02415	SARIMAX	3.301272733	DNN
89	VAT>REI>KJO>BLO>BER>HAR>LUN> FLAM	1.15912514	43.2844	40.1876	41.8495	21.22076	21.22023	0.03428	0.03428	SARIMAX	1.159125135	DNN

Figure 5.19: Comparison Chart 4

5.5 Conclusion

Least Root Mean Square Error for most of the packages was recorded by Time Series Model. But it is not suitable for the prediction of non-continuous data.

Therefore next best model was the Deep Neural Network. DNN had the least root mean square error in 75 pakages. Next was Ordinary Least Squares Multiple Linear Regression(statsmodels) model with 3 packages. In the packages which were discarded due to the lack of data points, linear regression model resulted the least root mean square error. This is due to one or two datapoints present in the dataset.

As there is a weak relationship between Root Mean Square Error, number of records for each package and number of unique prices for each package, we can hypothesis that with a bigger dataset, error can be reduced in the Deep Neural Network.

Hybrid model resulted in average of 79.25% increase of approximated revenue while minimum and maximum increase of improvement is 12.12% and 120.59% respectively. Therefore hybrid model was able to achieve significant increase of revenue to the company.

5.6 Future Work

- Considering more external factors that can affect the price of a trip package will increase the accuracy of data prediction.
- Including the cost for each package will improve the calculating better price.

Bibliography

[1] The Organization for Economic Co-operation and Development (OECD), “OECD Tourism Trends and Policies 2018 - Highlights”, https://www.oecd.org/cfe/tourism/, 2019 [Online]. Available: https://www.oecd.org/cfe/tourism/2018-Tourism-Trends-Policies-Highlights-ENG.pdf, [Accessed:20-Apr-2019].

[2] 10 of the Best European Train Trips, NationalGeographic.com, https://www.nationalgeographic.com/travel/destinations/europe/european-train-trips/,[Accessed:16-Aug-2019]

[3] Flam-Product Manual 2020, https://www.visitflam.com/information/travel-trade/, 2019 [Online],
Available: https://adobeindd.com/view/publications/f1d916d1-c7e5-4c3d-9fad-ce11b339662d/cg8q/publication-web-resources/pdf/2019-04-08-Fl_m-Produktmanual-2020.pdf [Accessed:25-Apr-2019].

[4] 22] The World Travel Tourism Council (WTTC), “SRI LANKA 2019 ANNUAL RESEARCH: KEY HIGHLIGHTS”, https://www.wttc.org/ , 2019 [Online]. Available: https://www.wttc.org/economic-impact/country-analysis/country-data/ [Accessed:22-Apr-2019].

[5] visitflam.com/press,
https://www.visitflam.com/information/corporate-information/press/, 2019 [Online]. Available: https://www.visitflam.com/globalassets/dokumenter/press-information-2014.pdf [Accessed:25-Apr-2019].

[6] Flam Group, “Experience Flåmsbana”, https://www.visitflam.com/flamsbana/, 2019 [Online]. Available: https://www.visitflam.com/flamsbana/ [Accessed:20-Apr-2019].

[7] J. Ahmed, N. Zaki, K. Shuaib, and F. Khan, “Airline ticket price and demand prediction : A survey,” J. King Saud Univ. - Comput. Inf. Sci., no. xxxx, 2019.

[8] W. Groves and M. Gini, “An agent for optimizing airline ticket purchasing An Agent for Optimizing Airline Ticket Purchasing (Extended Abstract),” no. May 2013, pp. 71–73, 2014.

[9] J. Ahmed, N. Zaki, K. Shuaib, and F. Khan, “Airline ticket price and demand prediction : A survey,” J. King Saud Univ. - Comput. Inf. Sci., no. xxxx, 2019.

[10] T. Wohlfarth et al., “A Data-Mining Approach to Travel Price Forecasting To cite this version : HAL Id : hal-00665041 A Data-Mining Approach to Travel Price Forecasting,” 2012.

[11] S. K. Aggarwal and A. Kumar, "Short term price forecasting in deregulated electricity markets A review of statistical models and key issues," no. July 2014, 2009.

[12] M. Makkonen and L. Frank, "Identifying the Sales Patterns of Online Stores with Time Series Clustering," pp. 491–505, 2018.

[13] A. E. Permanasari, I. Hidayah, and I. A. Bustoni, "SARIMA (Seasonal ARIMA) Implementation on Time Series to Forecast The Number of Malaria Incidence," no. October, 2013.

[14] N. S. Arunraj and D. Ahrens, "Application of SARIMAX Model to Forecast Daily Sales in Food Retail Industry Application of SARIMAX Model to Forecast Daily Sales in Food Retail Industry," no. April, 2016.

[15] Y. Chen and S. Tjandra, "Daily Collision Prediction with SARIMAX and Generalized Linear Models on the Basis of Temporal and Weather Variables," no. 2432, pp. 26–36, 2014.

[16] S. Tudy, "A LINEAR REGRESSION APPROACH TO PREDICTION OF STOCK MARKET TRADING VOLUME : A C ASE," vol. 4, no. 3, pp. 25–31, 2013.

[17] N. Shinde and K. Gawande, "VALUATION OF HOUSE PRICES USING PREDICTIVE," no. 6, pp. 34–40, 2018.

[18] Lin, Y., 2015. Auto Car Sales Prediction: A Statistical Study Using Functional Data Analysis and Time Series (Doctoral dissertation).

[19] R. Dubin, "Predicting House Prices Using Multiple Listings Data," no. February 1998, 2016.

[20] P. F. Pai, K. P. Lin, C. S. Lin, and P. T. Chang, "Time series forecasting by a seasonal support vector regression model," Expert Syst. Appl., vol. 37, no. 6, pp. 4261–4265, 2010.

[21] Z. L. Sun, T. M. Choi, K. F. Au, and Y. Yu, "Sales forecasting using extreme learning machine with applications in fashion retailing," Decis. Support Syst., vol. 46, no. 1, pp. 411–419, 2008.

[22] P. C. Chang and Y. W. Wang, "Fuzzy Delphi and back-propagation model for sales forecasting in PCB industry," Expert Syst. Appl., vol. 30, no. 4, pp. 715–726, 2006.

[23] R. J. Kuo and K. C. Xue, "Fuzzy neural networks with application to sales forecasting," Fuzzy Sets Syst., vol. 108, no. 2, pp. 123–143, 1999.

[24] P. Palmer and D. O. Connell, "Regression Analysis for Prediction :," no. September, 2009.

[25] W. Wang, C. Men, and W. Lu, "Online prediction model based on support vector machine," no. January 2008, 2019.

[26] H. Cao and M. Ahmed, "Application of support vector regression trained by particle swarm optimization in warrant price prediction," ICIMA 2010 - 2010 2nd Int. Conf. Ind. Mechatronics Autom., vol. 1, pp. 358–361, 2010.

[27] World Weather Online API,

https://www.worldweatheronline.com/developer/api/,[Accessed:16-Aug-2019]

[28] Time and Date Free Holiday API, https://www.timeanddate.com/services/api/holiday-api.html, [Accessed:16-Aug-2019]

[29] Scikit-learn.org.(2020).sklearn.linear_model.LinearRegression-scikit-learn 0.22.1 documentation. [online] Available at: https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html [Accessed 19 Feb. 2020].

[30] Statsmodels.org. (2020). Ordinary Least Squares-statsmodels. [online] Available at: https://www.statsmodels.org/dev/examples/notebooks/generated/ols.html [Accessed 19 Feb. 2020].

[31] Brownlee, J. (2020). A Gentle Introduction to SARIMA for Time Series Forecasting in Python. [online] Machine Learning Mastery. Available at: https://machinelearningmastery.com/sarima-for-time-series-forecasting-in-python/ [Accessed 19 Feb. 2020].

[32] Imf.org. (2020). [online] Available at: https://www.imf.org/external/pubs/ft/weo/2019/02/wedata/v
[Accessed 7 Jan. 2020].

Appendices

Appendix A

Trip Package Details

A.1 No of Trips for each Package

Table A.1: No of Trips for each Package (Full Table).

Package	No of Trips
BER>BLO	17279
BER>BLO>KJO>REI	17279
BER>BLO>KJO>REI>VAT	21669
BER>BLO>KJO>REI>VAT>MYR	21669
BER>HAR	46620
BER>HAR>LUN	46619
BER>HAR>LUN>FLAM	55515
BER>KJO>VAT	2135
BER>KJO>VAT>MYR	2135
BLO>BER	46775
BLO>BER>HAR	46621
BLO>BER>HAR>LUN	46621
BLO>BER>HAR>LUN>FLAM	55286
BLO>KJO>REI	46849
BLO>KJO>REI>VAT	58366
BLO>KJO>REI>VAT>FLAM	473
BLO>KJO>REI>VAT>MYR	57893
FLAM>BER	1956
FLAM>BER>KJO>VAT	5005
FLAM>BER>KJO>VAT>MYR	5407
FLAM>BER>KJO>VAT>MYR>VAT>KJO>FLAM	397
FLAM>BER>KJO>VAT>MYR>VAT>REI>KJO>BLO> BER>HAR>LUN>FLAM	150
FLAM>LUN	46855
FLAM>LUN>HAR	46854
FLAM>LUN>HAR>BER	20508
FLAM>LUN>HAR>BER>BLO	20508
FLAM>LUN>HAR>BER>BLO>KJO>REI	27737
FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT	49089

Continued on next page

Table A.1 – continued from previous page

Package	No of Trips
FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT>MYR	55168
FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT>MYR >VAT> REI>KJO>BLO>BER>HAR>LUN>FLAM	8194
FLAM>LUN>HAR>BLO	35032
FLAM>LUN>HAR>BLO>KJO>REI	42975
FLAM>LUN>HAR>BLO>KJO>REI>VAT	80304
FLAM>LUN>HAR>BLO>KJO>REI>VAT>FLAM	373
FLAM>LUN>HAR>BLO>KJO>REI>VAT>MYR	89021
FLAM>LUN>HAR>BLO>KJO>REI>VAT>MYR>VAT >REI>KJO>BLO>BER>HAR>LUN>FLAM	21963
FLAM>VAT	54
FLAM>VAT>MYR	54
FLAM>VAT>MYR>VAT>FLAM	85
HAR>BER	17280
HAR>BER>BLO	17280
HAR>BER>BLO>KJO>REI	17280
HAR>BER>BLO>KJO>REI>VAT	17279
HAR>BER>BLO>KJO>REI>VAT>MYR	17279
HAR>BLO	29572
HAR>BLO>KJO>REI	29572
HAR>BLO>KJO>REI>VAT	29571
HAR>BLO>KJO>REI>VAT>FLAM	473
HAR>BLO>KJO>REI>VAT>MYR	29098
HAR>LUN	46619
HAR>LUN>FLAM	46619
LUN>FLAM	46619
LUN>HAR	46854
LUN>HAR>BER	17281
LUN>HAR>BER>BLO	17281
LUN>HAR>BER>BLO>KJO>REI	17281
LUN>HAR>BER>BLO>KJO>REI>VAT	17280
LUN>HAR>BER>BLO>KJO>REI>VAT>MYR	17280
LUN>HAR>BLO	29572
LUN>HAR>BLO>KJO>REI	29572
LUN>HAR>BLO>KJO>REI>VAT	29572
LUN>HAR>BLO>KJO>REI>VAT>FLAM	473
LUN>HAR>BLO>KJO>REI>VAT>MYR	29099
MYR>VAT	74614
MYR>VAT>FLAM	54
MYR>VAT>KJO>FLAM	5211
MYR>VAT>REI	46628
MYR>VAT>REI>KJO>BLO	58125
MYR>VAT>REI>KJO>BLO>BER	59120
MYR>VAT>REI>KJO>BLO>BER>HAR	46627
MYR>VAT>REI>KJO>BLO>BER>HAR>LUN	46627

Continued on next page

Table A.1 – continued from previous page

Package	No of Trips
MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	141716
REI>KJO>BLO	46625
REI>KJO>BLO>BER	46623
REI>KJO>BLO>BER>HAR	46623
REI>KJO>BLO>BER>HAR>LUN	46623
REI>KJO>BLO>BER>HAR>LUN>FLAM	70470
REI>VAT	46848
REI>VAT>FLAM	473
REI>VAT>MYR	46374
VAT>FLAM	527
VAT>KJO>FLAM	4809
VAT>MYR	75750
VAT>REI	46625
VAT>REI>KJO>BLO	58122
VAT>REI>KJO>BLO>BER	59117
VAT>REI>KJO>BLO>BER>HAR	46625
VAT>REI>KJO>BLO>BER>HAR>LUN	46625
VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	126240

A.2 No of Trips for each Package

Table A.2: No of Trips for each Package (Full Table).

Package	No of Unique Prices
BER>BLO	8
BER>BLO>KJO>REI	11
BER>BLO>KJO>REI>VAT	11
BER>BLO>KJO>REI>VAT>MYR	11
BER>HAR	7
BER>HAR>LUN	11
BER>HAR>LUN>FLAM	10
BER>KJO>VAT	7
BER>KJO>VAT>MYR	7
BLO>BER	7
BLO>BER>HAR	7
BLO>BER>HAR>LUN	11
BLO>BER>HAR>LUN>FLAM	10
BLO>KJO>REI	10
BLO>KJO>REI>VAT	11
BLO>KJO>REI>VAT>FLAM	2
BLO>KJO>REI>VAT>MYR	11
FLAM>BER	7
FLAM>BER>KJO>VAT	4

Continued on next page

Table A.2 – continued from previous page

Package	No of Unique Prices
FLAM>BER>KJO>VAT>MYR	6
FLAM>BER>KJO>VAT>MYR>VAT>KJO>FLAM	5
FLAM>BER>KJO>VAT>MYR>VAT>REI>KJO>BLO >BER>HAR>LUN>FLAM	2
FLAM>LUN	8
FLAM>LUN>HAR	8
FLAM>LUN>HAR>BER	10
FLAM>LUN>HAR>BER>BLO	10
FLAM>LUN>HAR>BER>BLO>KJO>REI	12
FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT	13
FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT>MYR	13
FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT>MYR >VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	9
FLAM>LUN>HAR>BLO	9
FLAM>LUN>HAR>BLO>KJO>REI	10
FLAM>LUN>HAR>BLO>KJO>REI>VAT	11
FLAM>LUN>HAR>BLO>KJO>REI>VAT>FLAM	2
FLAM>LUN>HAR>BLO>KJO>REI>VAT>MYR	7
FLAM>LUN>HAR>BLO>KJO>REI>VAT>MYR>VAT >REI>KJO>BLO>BER>HAR>LUN>FLAM	10
FLAM>VAT	1
FLAM>VAT>MYR	1
FLAM>VAT>MYR>VAT>FLAM	1
HAR>BER	8
HAR>BER>BLO	8
HAR>BER>BLO>KJO>REI	11
HAR>BER>BLO>KJO>REI>VAT	11
HAR>BER>BLO>KJO>REI>VAT>MYR	11
HAR>BLO	6
HAR>BLO>KJO>REI	9
HAR>BLO>KJO>REI>VAT	9
HAR>BLO>KJO>REI>VAT>FLAM	2
HAR>BLO>KJO>REI>VAT>MYR	9
HAR>LUN	8
HAR>LUN>FLAM	8
LUN>FLAM	8
LUN>HAR	8
LUN>HAR>BER	12
LUN>HAR>BER>BLO	12
LUN>HAR>BER>BLO>KJO>REI	11
LUN>HAR>BER>BLO>KJO>REI>VAT	11
LUN>HAR>BER>BLO>KJO>REI>VAT>MYR	11
LUN>HAR>BLO	9
LUN>HAR>BLO>KJO>REI	9
LUN>HAR>BLO>KJO>REI>VAT	9

Continued on next page

Table A.2 – continued from previous page

Package	No of Unique Prices
LUN>HAR>BLO>KJO>REI>VAT>FLAM	2
LUN>HAR>BLO>KJO>REI>VAT>MYR	9
MYR>VAT	8
MYR>VAT>FLAM	1
MYR>VAT>KJO>FLAM	6
MYR>VAT>REI	7
MYR>VAT>REI>KJO>BLO	11
MYR>VAT>REI>KJO>BLO>BER	11
MYR>VAT>REI>KJO>BLO>BER>HAR	10
MYR>VAT>REI>KJO>BLO>BER>HAR>LUN	10
MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	11
REI>KJO>BLO	10
REI>KJO>BLO>BER	10
REI>KJO>BLO>BER>HAR	10
REI>KJO>BLO>BER>HAR>LUN	10
REI>KJO>BLO>BER>HAR>LUN>FLAM	11
REI>VAT	7
REI>VAT>FLAM	2
REI>VAT>MYR	7
VAT>FLAM	2
VAT>KJO>FLAM	4
VAT>MYR	8
VAT>REI	7
VAT>REI>KJO>BLO	11
VAT>REI>KJO>BLO>BER	11
VAT>REI>KJO>BLO>BER>HAR	10
VAT>REI>KJO>BLO>BER>HAR>LUN	10
VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	13

Appendix B

Data Analyzing

B.1 Deep Neural Network

Table B.1: Results of Deep Neural Network (Full Table).

Package	RMSE	MAE	MAPE
BER>BLO	1.858566	1.468043	3.152602
BER>BLO>KJO>REI	6.484673	4.624062	3.104236
BER>BLO>KJO>REI>VAT	3.873239	2.846733	2.861113
BER>BLO>KJO>REI>VAT>MYR	4.521151	2.849314	2.791244
BER>HAR	1.713967	1.399042	2.895124
BER>HAR>LUN	1.978418	1.599791	1.948576
BER>HAR>LUN>FLAM	2.374	1.813088	2.042161
BER>KJO>VAT	5.219659	3.38342	3.176681
BER>KJO>VAT>MYR	4.859553	3.03486	2.898397
BLO>BER	0.826601	0.478076	0.966286
BLO>BER>HAR	0.646029	0.470578	0.944132
BLO>BER>HAR>LUN	3.014169	2.468362	3.163834
BLO>BER>HAR>LUN>FLAM	2.292972	1.735622	1.925343
BLO>KJO>REI	4.908975	3.830123	2.427808
BLO>KJO>REI>VAT	2.04234	1.417867	1.305954
BLO>KJO>REI>VAT>FLAM	3.007701	2.502922	2.46558
BLO>KJO>REI>VAT>MYR	2.6565	1.821656	1.659015
FLAM>BER	4.486399	3.393734	3.348277
FLAM>BER>KJO>VAT	4.286934	3.172611	1.887499
FLAM>BER>KJO>VAT>MYR	5.150316	3.463431	2.431822
FLAM>BER>KJO>VAT>MYR>VAT>KJO>FLAM	20.34249	15.95083	3.972168
FLAM>BER>KJO>VAT>MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	36.22231	31.01369	7.108944
FLAM>LUN	1.049539	0.665165	1.499233
FLAM>LUN>HAR	1.057369	0.662838	1.538762
FLAM>LUN>HAR>BER	3.556416	2.849285	3.520059
FLAM>LUN>HAR>BER>BLO	2.812932	2.215114	2.778917

Continued on next page

Table B.1 – continued from previous page

Package	RMSE	MAE	MAPE
FLAM>LUN>HAR>BER>BLO >KJO>REI	4.673483	2.634299	1.316589
FLAM>LUN>HAR>BER>BLO >KJO>REI>VAT	3.280811	1.589091	1.082199
FLAM>LUN>HAR>BER>BLO >KJO>REI>VAT>MYR	7.202048	3.353522	2.150453
FLAM>LUN>HAR>BER>BLO >KJO>REI>VAT>MYR>VAT> REI>KJO>BLO>BER>HAR >LUN>FLAM	27.68039	19.00272	5.288353
FLAM>LUN>HAR>BLO	2.559466	1.858472	1.886723
FLAM>LUN>HAR>BLO>KJO >REI	6.766356	5.684449	2.678014
FLAM>LUN>HAR>BLO>KJO >REI>VAT	2.637755	2.189242	1.384624
FLAM>LUN>HAR>BLO>KJO >REI>VAT>FLAM	10.49589	8.022769	2.902952
FLAM>LUN>HAR>BLO>KJO >REI>VAT>MYR	4.035698	3.824775	2.817942
FLAM>LUN>HAR>BLO>KJO >REI>VAT>MYR>VAT>REI >KJO>BLO>BER>HAR>LUN >FLAM	18.09003	11.94107	3.786264
FLAM>VAT	23.56934	21.87919	12.1551
FLAM>VAT>MYR	29.46332	22.7781	12.6545
FLAM>VAT>MYR>VAT>FLAM	5.882211	4.223087	2.111544
HAR>BER	2.298654	1.89993	4.012732
HAR>BER>BLO	1.542857	1.119171	2.390324
HAR>BER>BLO>KJO>REI	9.340629	6.114926	2.52653
HAR>BER>BLO>KJO>REI >VAT	7.377924	5.231702	2.158643
HAR>BER>BLO>KJO>REI >VAT>MYR	7.488212	4.978387	2.089367
HAR>BLO	0.623856	0.56772	1.121232
HAR>BLO>KJO>REI	11.24091	9.987472	3.809465
HAR>BLO>KJO>REI>VAT	3.564311	2.913637	1.208618
HAR>BLO>KJO>REI>VAT >FLAM	1.732173	1.378804	2.449808
HAR>BLO>KJO>REI>VAT >MYR	4.876252	3.640034	1.54463
HAR>LUN	1.968015	1.419105	3.050428
HAR>LUN>FLAM	2.167759	1.636852	3.609924
LUN>FLAM	1.203572	1.058201	2.236696
LUN>HAR	0.746472	0.312553	0.669944
LUN>HAR>BER	1.81794	1.156857	1.504521

Continued on next page

Table B.1 – continued from previous page

Package	RMSE	MAE	MAPE
LUN>HAR>BER>BLO	2.622865	1.95878	2.502383
LUN>HAR>BER>BLO>KJO>REI	10.18228	7.910404	3.357073
LUN>HAR>BER>BLO>KJO>REI>VAT	7.934784	5.931449	2.431512
LUN>HAR>BER>BLO>KJO>REI>VAT>MYR	10.87228	7.715033	3.225315
LUN>HAR>BLO	1.272428	0.995268	1.21836
LUN>HAR>BLO>KJO>REI	4.242982	3.27355	1.414428
LUN>HAR>BLO>KJO>REI>VAT	4.571584	3.426566	1.460505
LUN>HAR>BLO>KJO>REI>VAT>FLAM	1.640626	1.286168	2.282473
LUN>HAR>BLO>KJO>REI>VAT>MYR	4.751711	4.089745	1.643411
MYR>VAT	1.103242	0.82301	1.754781
MYR>VAT>FLAM	9.665284	7.443003	4.135002
MYR>VAT>KJO>FLAM	6.433836	4.230544	2.881315
MYR>VAT>REI	2.136362	1.814792	3.656908
MYR>VAT>REI>KJO>BLO	3.098199	2.325438	2.076524
MYR>VAT>REI>KJO>BLO>BER	3.911077	3.32042	3.052266
MYR>VAT>REI>KJO>BLO>BER>HAR	4.738929	3.503313	1.336157
MYR>VAT>REI>KJO>BLO>BER>HAR>LUN	9.495176	8.294823	3.335257
MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	2.559312	1.902278	1.362386
REI>KJO>BLO	2.856872	2.440422	1.535398
REI>KJO>BLO>BER	5.259671	4.655553	2.919897
REI>KJO>BLO>BER>HAR	4.801925	3.556793	1.423036
REI>KJO>BLO>BER>HAR>LUN	7.901946	6.745264	2.637164
REI>KJO>BLO>BER>HAR>LUN>FLAM	3.565468	2.798954	1.299343
REI>VAT	0.703376	0.362151	0.728513
REI>VAT>FLAM	8.357068	6.167291	2.670213
REI>VAT>MYR	0.899422	0.334246	0.690492
VAT>FLAM	17.14715	12.47981	5.424099
VAT>KJO>FLAM	5.681943	4.555935	2.682559
VAT>MYR	1.10481	0.925405	1.931521
VAT>REI	0.588297	0.283793	0.57171
VAT>REI>KJO>BLO	2.108244	1.205626	0.972983
VAT>REI>KJO>BLO>BER	3.88602	3.376073	3.034362
VAT>REI>KJO>BLO>BER>HAR	4.9659	3.765417	1.522494

Continued on next page

Table B.1 – continued from previous page

Package	RMSE	MAE	MAPE
VAT>REI>KJO>BLO>BER>HAR >LUN	3.301273	2.309422	0.907479
VAT>REI>KJO>BLO>BER>HAR >LUN>FLAM	1.159125	0.841962	0.841962

B.2 Ordinary Least Squares Multiple Linear Regression (sklearn)

Table B.2: Results of Ordinary Least Squares Multiple Linear Regression(sklearn) (Full Table).

Package	MAE	RMAE	S Square
BER>BLO	3.458736	4.33187	0.269019
BER>BLO>KJO>REI	8.00214	11.15019	0.382836
BER>BLO>KJO>REI>VAT	25.15824	29.10985	0.527894
BER>BLO>KJO>REI>VAT>MYR	25.34503	29.35536	0.51995
BER>HAR	2.567015	3.27245	0.435294
BER>HAR>LUN	3.592226	4.49435	0.739918
BER>HAR>LUN>FLAM	13.56142	15.76643	0.31188
BER>KJO>VAT	15.2408	18.88602	0.190285
BER>KJO>VAT>MYR	15.09802	18.71001	0.205172
BLO>BER	2.52495	3.22201	0.450156
BLO>BER>HAR	2.566653	3.27242	0.435275
BLO>BER>HAR>LUN	3.592673	4.49432	0.739917
BLO>BER>HAR>LUN>FLAM	13.5749	15.78416	0.312912
BLO>KJO>REI	7.352974	9.37399	0.350237
BLO>KJO>REI>VAT	21.07545	25.08915	0.586398
BLO>KJO>REI>VAT>FLAM	0	1.03E-10	1
BLO>KJO>REI>VAT>MYR	21.19356	25.27063	0.578075
FLAM>BER	8.174771	9.65797	0.224025
FLAM>BER>KJO>VAT	12.16878	14.42549	0.199307
FLAM>BER>KJO>VAT>MYR	1.3228	4.08657	0.844064
FLAM>BER>KJO>VAT>MYR >VAT>KJO>FLAM	8.92412	10.37733	0.827748
FLAM>BER>KJO>VAT>MYR >VAT>REI>KJO>BLO>BER>HAR >LUN>FLAM	0	1.3678	1
FLAM>LUN	5.939065	8.46809	0.023409
FLAM>LUN>HAR	5.982399	8.39286	0.040256
FLAM>LUN>HAR>BER	14.87469	17.47779	0.300008
FLAM>LUN>HAR>BER>BLO	14.86165	17.15474	0.325515
FLAM>LUN>HAR>BER>BLO >KJO>REI	23.7517	32.57474	0.57058

Continued on next page

Table B.2 – continued from previous page

Package	MSE	R MAE	R Square
FLAM>LUN>HAR>BER>BLO >KJO>REI>VAT	17.25859	25.14734	0.329558
FLAM>LUN>HAR>BER>BLO >KJO>REI>VAT>MYR	20.12627	30.56546	0.202169
FLAM>LUN>HAR>BER>BLO >KJO>REI>VAT>MYR>VAT >REI>KJO>BLO>BER>HAR >LUN>FLAM	41.71987	49.97189	0.854676
FLAM>LUN>HAR>BLO	9.730596	12.53978	0.164577
FLAM>LUN>HAR>BLO>KJO >REI	27.33079	32.11961	0.572159
FLAM>LUN>HAR>BLO>KJO >REI>VAT	15.01603	18.22201	0.245608
FLAM>LUN>HAR>BLO>KJO >REI>VAT>FLAM	5.82E-11	2.04E-10	1
FLAM>LUN>HAR>BLO>KJO >REI>VAT>MYR	3.963822	5.91963	0.757222
FLAM>VAT>MYR	0	46.05014	1
FLAM>VAT>MYR>VAT>FLAM	0	3.63E-14	1
HAR>BER	3.346099	2.92E-14	0.29217
HAR>BER>BLO	3.457021	5.38E-14	0.269108
HAR>BER>BLO>KJO>REI	23.9379	4.26212	0.556302
HAR>BER>BLO>KJO>REI>VAT	24.02054	4.3323	0.556455
HAR>BER>BLO>KJO>REI>VAT >MYR	23.88176	30.69029	0.556474
HAR>BLO	1.617802	30.6822	0.579821
HAR>BLO>KJO>REI	30.90655	30.68365	0.393981
HAR>BLO>KJO>REI>VAT	29.61568	2.14491	0.420634
HAR>BLO>KJO>REI>VAT >FLAM	0	34.98533	1
HAR>BLO>KJO>REI>VAT >MYR	30.74584	34.20656	0.384989
HAR>LUN	6.002632	9.77E-12	0.0212
HAR>LUN>FLAM	6.035311	34.84482	0.037312
LUN>FLAM	6.039005	8.51911	0.037313
LUN>HAR	5.969588	8.44687	0.040242
LUN>HAR>BER	4.090605	8.44687	0.755865
LUN>HAR>BER>BLO	4.270951	8.39286	0.746225
LUN>HAR>BER>BLO>KJO>REI	23.93593	5.37999	0.55636
LUN>HAR>BER>BLO>KJO >REI >VAT	24.04601	5.48698	0.556289
LUN>HAR>BER>BLO>KJO>REI >VAT>MYR	23.8977	30.69146	0.556284
LUN>HAR>BLO	2.593102	30.68753	0.781044
LUN>HAR>BLO>KJO>REI>VAT	29.62398	30.69029	0.420643

Continued on next page

Table B.2 – continued from previous page

Package	MSE	R MAE	R Square
LUN>HAR>BLO>KJO>REI>VAT>FLAM	0	3.50978	1
LUN>HAR>BLO>KJO>REI>VAT>MYR	30.74842	34.98533	0.384998
MYR>VAT	2.455062	34.20722	0.699289
MYR>VAT>FLAM	0	4.70E-12	1
MYR>VAT>KJO>FLAM	1.267803	34.84565	0.846842
MYR>VAT>REI	2.529295	3.19406	0.448484
MYR>VAT>REI>KJO>BLO	20.78968	2.35E-13	0.602224
MYR>VAT>REI>KJO>BLO>BER	22.46956	4.05993	0.577384
MYR>VAT>REI>KJO>BLO>BER>HAR	27.66727	3.23085	0.400531
MYR>VAT>REI>KJO>BLO>BER>HAR>LUN	27.42854	24.66754	0.399645
MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	10.74465	25.9706	0.122201
REI>KJO>BLO	7.5034	32.77831	0.304184
REI>KJO>BLO>BER	7.166963	32.80666	0.342632
REI>KJO>BLO>BER>HAR	27.64706	20.13327	0.40052
REI>KJO>BLO>BER>HAR>LUN	27.41628	9.61705	0.39964
REI>KJO>BLO>BER>HAR>LUN>FLAM	26.52986	9.34108	0.552691
REI>VAT	2.502754	32.77749	0.463126
REI>VAT>FLAM	0	32.80638	1
REI>VAT>MYR	2.524949	32.26658	0.454278
VAT>FLAM	5.82E-11	3.17151	1
VAT>KJO>FLAM	12.7093	2.19E-11	0.113015
VAT>MYR	2.408913	3.19893	0.700571
VAT>REI	2.533192	2.42E-10	0.447464
VAT>REI>KJO>BLO	20.79601	14.90529	0.602061
VAT>REI>KJO>BLO>BER	22.48452	3.15665	0.577337
VAT>REI>KJO>BLO>BER>HAR	27.65406	3.23591	0.4005
VAT>REI>KJO>BLO>BER>HAR>LUN	27.422	24.68168	0.399666
VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	16.20492	25.98159	0.415398

B.3 Ordinary Least Squares Multiple Linear Regression (statsmodels)

Table B.3: Results of Ordinary Least Squares Multiple Linear Regression(statsmodels) (Full Table).

Package	MAE	RMAE	S Square
BER>BLO	3.458736	4.33243	0.269019
BER>BLO>KJO>REI	9.368491	11.15069	0.310618
BER>BLO>KJO>REI>VAT	34.5362	29.11086	0.128623
BER>BLO>KJO>REI>VAT>MYR	34.56456	29.3563	0.127731
BER>HAR	3.017193	3.27251	0.247479
BER>HAR>LUN	7.014595	3.27265	0.112332
BER>HAR>LUN>FLAM	13.37857	15.76653	0.295916
BER>KJO>VAT	13.13729	18.90905	0.072426
BER>KJO>VAT>MYR	12.94204	18.89474	0.097285
BLO>BER	3.106777	3.22213	0.25324
BLO>BER>HAR	3.017148	3.27258	0.247491
BLO>BER>HAR>LUN	7.014578	4.49448	0.112317
BLO>BER>HAR>LUN>FLAM	13.38688	15.78449	0.297101
BLO>KJO>REI	7.460728	9.37418	0.343274
BLO>KJO>REI>VAT	31.92491	25.08987	0.153661
BLO>KJO>REI>VAT>FLAM	12.28308	0	0.277542
BLO>KJO>REI>VAT>MYR	31.83661	25.27103	0.148097
FLAM>BER	8.219025	9.66347	0.222679
FLAM>BER>KJO>VAT	12.25438	14.42642	0.143115
FLAM>BER>KJO>VAT>MYR	8.553088	4.08795	0.052731
FLAM>BER>KJO>VAT>MYR>VAT>KJO>FLAM	13.31782	10.39904	0.489475
FLAM>BER>KJO>VAT>MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	8.148933	0	0.093024
FLAM>LUN	5.982837	8.46827	0.022141
FLAM>LUN>HAR	5.94707	8.393	0.039545
FLAM>LUN>HAR>BER	14.85727	17.47837	0.282961
FLAM>LUN>HAR>BER>BLO	14.99674	17.15539	0.291922
FLAM>LUN>HAR>BER>BLO>KJO>REI	43.58817	32.57737	0.029905
FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT	17.08108	25.14763	0.316439
FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT>MYR	19.50396	30.5659	0.182051
FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT>MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	106.6376	49.98332	0.161713

Continued on next page

Table B.3 – continued from previous page

Package	MSE	R MAE	R Square
FLAM>LUN>HAR>BLO	9.951744	12.54078	0.155892
FLAM>LUN>HAR>BLO>KJO >REI	43.68381	32.12023	0.020163
FLAM>LUN>HAR>BLO>KJO >REI>VAT	15.11403	18.2222	0.24393
FLAM>LUN>HAR>BLO>KJO >REI>VAT>FLAM	73.66834	0	0.288836
FLAM>LUN>HAR>BLO>KJO >REI>VAT>MYR	9.639293	5.91971	0.114537
FLAM>LUN>HAR>BLO>KJO >REI>VAT>MYR>VAT>REI >KJO>BLO>BER>HAR>LUN >FLAM	86.8283	46.05223	0.107412
FLAM>VAT	3.89E-14	0	0
FLAM>VAT>MYR	2.95E-13	0	0
FLAM>VAT>MYR>VAT>FLAM	4.85E-14	0	0
HAR>BER	3.167343	4.26252	0.264574
HAR>BER>BLO	3.315324	4.33264	0.229485
HAR>BER>BLO>KJO>REI	33.81067	30.69226	0.15555
HAR>BER>BLO>KJO>REI>VAT	33.56005	30.68308	0.161505
HAR>BER>BLO>KJO>REI>VAT >MYR	33.80991	30.68517	0.155522
HAR>BLO	2.430873	2.14537	0.192832
HAR>BLO>KJO>REI	36.28417	34.98612	0.041666
HAR>BLO>KJO>REI>VAT	35.50844	34.2084	0.064417
HAR>BLO>KJO>REI>VAT >FLAM	1.836281	0	0.282382
HAR>BLO>KJO>REI>VAT>MYR	35.46547	34.8454	0.040573
HAR>LUN	6.044727	8.51926	0.02033
HAR>LUN>FLAM	5.997736	8.44698	0.036389
LUN>FLAM	5.997736	8.44697	0.036389
LUN>HAR	5.94707	8.39301	0.039545
LUN>HAR>BER	8.189063	5.38053	0.155867
LUN>HAR>BER>BLO	7.975336	5.48725	0.195554
LUN>HAR>BER>BLO>KJO>REI	33.81591	30.69259	0.155615
LUN>HAR>BER>BLO>KJO>REI >VAT	33.56052	30.68835	0.161543
LUN>HAR>BER>BLO>KJO>REI >VAT>MYR	33.81067	30.69218	0.15555
LUN>HAR>BLO	5.512178	3.51002	0.208917
LUN>HAR>BLO>KJO>REI	36.28417	34.98697	0.041666
LUN>HAR>BLO>KJO>REI>VAT	35.51158	34.20834	0.064401
LUN>HAR>BLO>KJO>REI>VAT >FLAM	1.836281	0	0.282382

Continued on next page

Table B.3 – continued from previous page

Package	MSE	R MAE	R Square
LUN>HAR>BLO>KJO>REI>VAT>MYR	35.46889	34.84665	0.040551
MYR>VAT	3.420865	3.19414	0.414155
MYR>VAT>FLAM	1.94E-13	0	0
MYR>VAT>KJO>FLAM	5.74602	4.06011	0.39006
MYR>VAT>REI	3.100595	3.23087	0.255917
MYR>VAT>REI>KJO>BLO	31.08371	24.66787	0.152829
MYR>VAT>REI>KJO>BLO>BER	33.26097	25.97117	0.117477
MYR>VAT>REI>KJO>BLO>BER>HAR	32.74802	32.77882	0.047466
MYR>VAT>REI>KJO>BLO>BER>HAR>LUN	33.17767	32.80756	0.034197
MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	10.89411	20.13379	0.12212
REI>KJO>BLO	7.567031	9.61727	0.30156
REI>KJO>BLO>BER	7.212857	9.34145	0.338657
REI>KJO>BLO>BER>HAR	32.74674	32.77826	0.047444
REI>KJO>BLO>BER>HAR>LUN	33.1764	32.807	0.034188
REI>KJO>BLO>BER>HAR>LUN>FLAM	37.38946	32.26719	0.156684
REI>VAT	3.047735	3.17165	0.27981
REI>VAT>FLAM	43.0114	0	0.276849
REI>VAT>MYR	3.064963	3.19906	0.272791
VAT>FLAM	42.18889	0	0.254432
VAT>KJO>FLAM	12.43643	14.90718	0.078077
VAT>MYR	3.407812	3.15673	0.418911
VAT>REI	3.106937	3.23603	0.253326
VAT>REI>KJO>BLO	31.10044	24.68203	0.152357
VAT>REI>KJO>BLO>BER	33.27923	25.98228	0.117015
VAT>REI>KJO>BLO>BER>HAR	32.74768	32.7778	0.047455
VAT>REI>KJO>BLO>BER>HAR>LUN	33.17734	32.80719	0.034194
VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	16.76845	21.22076	0.372263

B.4 SARIMAX Model (Time Series)

Table B.4: SARIMAX Model (Time Series) (Full Table).

Package	MAE	RMAE	S Square
BER>BLO	0.05265	0.12028	0.23756
BER>BLO>KJO>REI	0.10209	0.24315	0.34674
BER>BLO>KJO>REI>VAT	0.04518	0.10202	0.84451
BER>BLO>KJO>REI>VAT>MYR	0.04367	0.09955	0.85193

Continued on next page

Table B.4 – continued from previous page

Package	MSE	R MAE	R Square
BER>HAR	0.0397	0.08657	0.55951
BER>HAR>LUN	0.02577	0.05824	0.96797
BER>HAR>LUN>FLAM	0.10334	0.25002	0.53191
BER>KJO>VAT	0.038	0.14137	0.83426
BER>KJO>VAT>MYR	0.03871	0.13932	0.83905
BLO>BER	0.03811	0.0836	0.58922
BLO>BER>HAR	0.038016	0.0841	0.58429
BLO>BER>HAR>LUN	0.02577	0.05824	0.96797
BLO>BER>HAR>LUN>FLAM	0.10334	0.25002	0.53191
BLO>KJO>REI	0.06873	0.14836	0.74077
BLO>KJO>REI>VAT	0.06178	0.14925	0.71885
BLO>KJO>REI>VAT>FLAM	0.00165	0.02334	0.9955
BLO>KJO>REI>VAT>MYR	0.06027	0.14593	0.7312
FLAM>BER	0.03324	0.12098	0.782
FLAM>BER>KJO>VAT	0.00758	0.0449	0.97012
FLAM>BER>KJO>VAT>MYR	0.01046	0.04759	0.95526
FLAM>BER>KJO>VAT>MYR> VAT>KJO>FLAM	0.0039	0.03501	0.99066
FLAM>BER>KJO>VAT>MYR> VAT>REI>KJO>BLO>BER> HAR>LUN>FLAM	0.00268	0.0234	0.99754
FLAM>LUN	0.03081	0.07053	0.79646
FLAM>LUN>HAR	0.0311	0.07087	0.79449
FLAM>LUN>HAR>BER	0.14673	0.32512	0.32323
FLAM>LUN>HAR>BER>BLO	0.14761	0.32768	0.31255
FLAM>LUN>HAR>BER>BLO> KJO>REI	0.00961	0.02951	0.81862
FLAM>LUN>HAR>BER>BLO> KJO >REI>VAT	0.01484	0.04209	0.43003
FLAM>LUN>HAR>BER>BLO> KJO>REI>VAT>MYR	0.01686	0.03969	0.50595
FLAM>LUN>HAR>BER>BLO> KJO>REI>VAT>MYR>VAT>REI> KJO>BLO>BER>HAR>LUN> FLAM	0.05238	0.1071	0.90256
FLAM>LUN>HAR>BLO	0.04814	0.13469	0.85493
FLAM>LUN>HAR>BLO>KJO> REI	0.01395	0.04575	0.98327
FLAM>LUN>HAR>BLO>KJO> REI>VAT	0.03001	0.08097	0.86968
FLAM>LUN>HAR>BLO>KJO> REI>VAT>FLAM	0.00165	0.02334	0.9955
FLAM>LUN>HAR>BLO>KJO> REI>VAT>MYR	0.01425	0.04272	0.95328

Continued on next page

Table B.4 – continued from previous page

Package	MSE	R MAE	R Square
FLAM>LUN>HAR>BLO>KJO> REI>VAT>MYR>VAT>REI>KJO> BLO>BER>HAR>LUN>FLAM	0.03971	0.10885	0.87063
FLAM>VAT	1.95E-10	3.31E-10	0
FLAM>VAT>MYR	3.51E-12	6.67E-12	0
FLAM>VAT>MYR>VAT>FLAM	0.00193	0.03294	0.99112
HAR>BER	0.05255	0.11937	0.24907
HAR>BER>BLO	0.05265	0.12028	0.23756
HAR>BER>BLO>KJO>REI	0.00964	0.02832	0.72334
HAR>BER>BLO>KJO>REI>VAT	0.01018	0.02995	0.69062
HAR>BER>BLO>KJO>REI>VAT >MYR	0.00964	0.02832	0.72334
HAR>BLO	0.01348	0.04928	0.97323
HAR>BLO>KJO>REI	0.01002	0.0362	0.98284
HAR>BLO>KJO>REI>VAT	0.00775	0.03504	0.98392
HAR>BLO>KJO>REI>VAT> FLAM	0.00165	0.02334	0.9955
HAR>BLO>KJO>REI>VAT>MYR	0.01002	0.0362	0.98284
HAR>LUN	0.03069	0.07051	0.79865
HAR>LUN>FLAM	0.03093	0.07085	0.7967
LUN>FLAM	0.03093	0.07085	0.79670
LUN>HAR	0.0311	0.07087	0.79449
LUN>HAR>BER	0.02626	0.05917	0.96734
LUN>HAR>BER>BLO	0.02605	0.05888	0.96766
LUN>HAR>BER>BLO>KJO>REI	0.00964	0.02832	0.72334
LUN>HAR>BER>BLO>KJO>REI >VAT	0.01018	0.02995	0.69062
LUN>HAR>BER>BLO>KJO>REI >VAT>MYR	0.00964	0.02832	0.72334
LUN>HAR>BLO	0.00518	0.02533	0.99461
LUN>HAR>BLO>KJO>REI	0.01002	0.0362	0.98284
LUN>HAR>BLO>KJO>REI>VAT	0.00775	0.03504	0.98392
LUN>HAR>BLO>KJO>REI>VAT >FLAM	0.00165	0.02334	0.99550
LUN>HAR>BLO>KJO>REI>VAT >MYR	0.01002	0.0362	0.98284
MYR>VAT	0.00754	0.04085	0.98239
MYR>VAT>FLAM	1.08E-09	1.78E-09	0.0
MYR>VAT>KJO>FLAM	0.00714	0.04676	0.95683
MYR>VAT>REI	0.04007	0.08803	0.54362
MYR>VAT>REI>KJO>BLO	0.06026	0.14498	0.73599
MYR>VAT>REI>KJO>BLO>BER	0.06156	0.14799	0.72494
MYR>VAT>REI>KJO>BLO>BER >HAR	0.00598	0.02408	0.74901

Continued on next page

Table B.4 – continued from previous page

Package	MSE	R MAE	R Square
MYR>VAT>REI>KJO>BLO>BER >HAR>LUN	0.0067	0.02415	0.74762
MYR>VAT>REI>KJO>BLO>BER >HAR>LUN>FLAM	0.02659	0.06618	0.91131
REI>KJO>BLO	0.06641	0.14466	0.75244
REI>KJO>BLO>BER	0.06626	0.14476	0.75209
REI>KJO>BLO>BER>HAR	0.00598	0.02408	0.74901
REI>KJO>BLO>BER>HAR>LUN	0.0067	0.02415	0.74762
REI>KJO>BLO>BER>HAR>LUN >FLAM	0.00538	0.02539	0.84632
REI>VAT	0.03805	0.08337	0.5857
REI>VAT>FLAM	0.00165	0.02334	0.9955
REI>VAT>MYR	0.04069	0.08776	0.54098
VAT>FLAM	0.00139	0.02329	0.99556
VAT>KJO>FLAM	0.00635	0.04475	0.97032
VAT>MYR	0.00934	0.05496	0.97705
VAT>REI	0.0397	0.08657	0.55951
VAT>REI>KJO>BLO	0.05941	0.14406	0.73956
VAT>REI>KJO>BLO>BER	0.06168	0.14914	0.72086
VAT>REI>KJO>BLO>BER>HAR	0.00598	0.02408	0.74901
VAT>REI>KJO>BLO>BER>HAR >LUN	0.0067	0.02415	0.74762
VAT>REI>KJO>BLO>BER>HAR >LUN>FLAM	0.0117	0.03428	0.58942

B.5 Support Vector Regression (SVR)

Table B.5: Support Vector Regression (SVR) (Full Table).

Package	RBF RMAE	Liner RMAE	Polynomial RMAE
BER>BLO	6.32198	5.71895	6.32507
BER>BLO>KJO>REI	19.09588	16.91226	17.66532
BER>BLO>KJO>REI>VAT	45.50209	44.69185	44.44279
BER>BLO>KJO>REI>VAT>MYR	50.22534	47.39201	49.77362
BER>HAR	6.14854	5.24405	5.23937
BER>HAR>LUN	10.59189	9.60826	9.4346
BER>HAR>LUN>FLAM	21.30226	21.78177	21.47419
BER>KJO>VAT	24.96192	23.87669	24.29451
BER>KJO>VAT>MYR	47.65939	26.88151	26.96798
BLO>BER	5.99866	5.38207	5.01755
BLO>BER>HAR	6.14854	5.24405	5.23937
BLO>BER>HAR>LUN	10.59189	9.60826	9.4346
BLO>BER>HAR>LUN>FLAM	21.30226	21.78177	21.47419
BLO>KJO>REI	16.85568	14.87762	15.36246

Continued on next page

Table B.5 – continued from previous page

Package	RBF RMAE	Liner RMAE	Polynomial RMAE
BLO>KJO>REI>VAT	16.85568	14.87762	15.36246
BLO>KJO>REI>VAT>FLAM	24.11035	24.18478	21.38035
BLO>KJO>REI>VAT>MYR	48.36985	47.83985	49.39595
FLAM>BER	48.36985	47.83985	49.39595
FLAM>BER>KJO>VAT	16.71602	16.75301	16.84424
FLAM>BER>KJO>VAT>MYR	44.8251	11.07076	11.66367
FLAM>BER>KJO>VAT>MYR >VAT>KJO>FLAM	34.89031	40.0512	35.95806
FLAM>BER>KJO>VAT>MYR >VAT>REI>KJO>BLO> BER >HAR>LUN>FLAM	9.951	9.95038	9.89981
FLAM>LUN	7.97389	7.75363	7.79226
FLAM>LUN>HAR	8.57759	7.732	7.78108
FLAM>LUN>HAR>BER	26.09637	25.65682	23.54452
FLAM>LUN>HAR>BER>BLO	22.7801	23.78572	24.10695
FLAM>LUN>HAR>BER>BLO >KJO>REI	66.02045	59.90782	66.37645
FLAM>LUN>HAR>BER>BLO >KJO>REI>VAT	45.76663	44.83644	42.57949
FLAM>LUN>HAR>BER>BLO >KJO>REI>VAT>MYR	54.48506	52.78967	53.93302
FLAM>LUN>HAR>BER>BLO >KJO>REI>VAT>MYR>VAT >REI>KJO>BLO>BER>HAR >LUN>FLAM	111.07268	112.96514	115.4014
FLAM>LUN>HAR>BLO	17.75194	16.73817	16.64998
FLAM>LUN>HAR>BLO>KJO >REI	57.08184	53.85372	55.18569
FLAM>LUN>HAR>BLO>KJO >REI>VAT	25.31063	26.31903	25.92051
FLAM>LUN>HAR>BLO>KJO >REI>VAT>FLAM	132.76394	133.88456	117.38385
FLAM>LUN>HAR>BLO>KJO >REI>VAT>MYR	21.37466	18.67107	18.67107
FLAM>LUN>HAR>BLO>KJO >REI>VAT>MYR>VAT>REI >KJO>BLO>BER>HAR>LUN >FLAM	97.2033	97.58202	97.17584
FLAM>VAT	0	0	0
FLAM>VAT>MYR	0	0	0
FLAM>VAT>MYR>VAT>FLAM	0	0	0
HAR>BER	6.56816	6.4953	5.38852
HAR>BER>BLO	6.32198	5.71895	6.32508
HAR>BER>BLO>KJO>REI	48.03246	46.80069	46.02116

Continued on next page

Table B.5 – continued from previous page

Package	RBF RMAE	Liner RMAE	Polynomial RMAE
HAR>BER>BLO>KJO>REI>VAT	57.42248	54.20039	43.24679
HAR>BER>BLO>KJO>REI>VAT>MYR	48.03246	46.80069	46.02116
HAR>BLO	4.32108	4.31541	4.62342
HAR>BLO>KJO>REI	41.30332	41.06588	41.08314
HAR>BLO>KJO>REI>VAT	47.26136	45.27535	44.54644
HAR>BLO>KJO>REI>VAT>FLAM	3.57209	3.4716	3.24752
HAR>BLO>KJO>REI>VAT>MYR	41.30332	41.06588	41.08314
HAR>LUN	7.99863	7.78515	7.82409
HAR>LUN>FLAM	8.60274	7.76582	7.81742
LUN>FLAM	8.60274	7.76582	7.81742
LUN>HAR	8.57759	7.732	7.78108
LUN>HAR>BER	13.49223	11.85197	10.96682
LUN>HAR>BER>BLO	12.55803	11.45499	11.84734
LUN>HAR>BER>BLO>KJO>REI	48.03246	46.80069	46.02116
LUN>HAR>BER>BLO>KJO>REI>VAT	57.42248	54.20039	43.24679
LUN>HAR>BER>BLO>KJO>REI>VAT>MYR	48.03246	46.80069	46.02116
LUN>HAR>BLO	7.93038	7.41939	7.51247
LUN>HAR>BLO>KJO>REI	41.30332	41.06588	41.08314
LUN>HAR>BLO>KJO>REI>VAT	47.26136	45.27535	44.54644
LUN>HAR>BLO>KJO>REI>VAT>FLAM	3.57209	3.4716	3.24752
LUN>HAR>BLO>KJO>REI>VAT>MYR	41.30332	41.06588	41.08314
MYR>VAT	8.03924	7.36756	6.7296
MYR>VAT>FLAM	0	0	0
MYR>VAT>KJO>FLAM	14.41129	14.76957	13.71924
MYR>VAT>REI	6.15244	5.75581	5.92813
MYR>VAT>REI>KJO>BLO	41.98637	45.31673	44.93226
MYR>VAT>REI>KJO>BLO>BER	43.85821	44.07417	44.42567
MYR>VAT>REI>KJO>BLO>BER>HAR	43.96271	37.76192	37.59945
MYR>VAT>REI>KJO>BLO>BER>HAR>LUN	43.8425	38.00317	37.60524
MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	45.11573	44.08786	44.8914
REI>KJO>BLO	15.17875	13.68504	13.51189

Continued on next page

Table B.5 – continued from previous page

Package	RBF RMAE	Liner RMAE	Polynomial RMAE
REI>KJO>BLO>BER	14.53298	14.50421	13.22763
REI>KJO>BLO>BER>HAR	43.96271	37.76192	37.59945
REI>KJO>BLO>BER>HAR >LUN	43.8425	38.00317	37.60524
REI>KJO>BLO>BER>HAR >LUN>FLAM	64.34844	59.82099	61.91822
REI>VAT	5.99393	5.34335	4.97517
REI>VAT>FLAM	84.88108	85.12875	74.71557
REI>VAT>MYR	6.14973	5.73058	5.90008
VAT>FLAM	85.06959	69.39775	76.37011
VAT>KJO>FLAM	16.76882	17.37712	16.93783
VAT>MYR	7.72999	7.2708	7.22816
VAT>REI	6.15288	5.76	5.93294
VAT>REI>KJO>BLO	42.02052	45.34533	44.96285
VAT>REI>KJO>BLO>BER	43.88407	44.09716	44.46143
VAT>REI>KJO>BLO>BER >HAR	43.96271	37.76192	37.59945
VAT>REI>KJO>BLO>BER >HAR>LUN	43.8425	38.00317	37.60524
VAT>REI>KJO>BLO>BER >HAR>LUN>FLAM	43.28443	40.18755	41.8495

B.6 Extreme Learning Machine (ELM)

Table B.6: Extreme Learning Machine (ELM) (Full Table).

Package	Accuracy of the Prediction
BER>BLO	0.32596685
BER>BLO>KJO>REI	0.23756906
BER>BLO>KJO>REI>VAT	0.38674033
BER>BLO>KJO>REI>VAT>MYR	0.24861878
BER>HAR	0.32596685
BER>HAR>LUN	0.33149171
BER>HAR>LUN>FLAM	0.37016575
BER>KJO>VAT	0.22857143
BER>KJO>VAT>MYR	0.2
BLO>BER	0.37016575
BLO>BER>HAR	0.29281768
BLO>BER>HAR>LUN	0.32596685
BLO>BER>HAR>LUN>FLAM	0.3480663
BLO>KJO>REI	0.25414365
BLO>KJO>REI>VAT	0.40883978
BLO>KJO>REI>VAT>FLAM	0.45454545
BLO>KJO>REI>VAT>MYR	0.3038674

Continued on next page

Table B.6 – continued from previous page

Package	Accuracy of the Prediction
FLAM>BER	0.25714286
FLAM>BER>KJO>VAT	0.45714286
FLAM>BER>KJO>VAT>MYR	0.14285714
FLAM>BER>KJO>VAT>MYR>VAT>KJO>FLAM	0.4
FLAM>BER>KJO>VAT>MYR>VAT>REI>KJO>BLO >BER>HAR>LUN>FLAM	0.21428571
FLAM>LUN	0.32596685
FLAM>LUN>HAR	0.35911602
FLAM>LUN>HAR>BER	0.37016575
FLAM>LUN>HAR>BER>BLO	0.37569061
FLAM>LUN>HAR>BER>BLO>KJO>REI	0.20994475
FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT	0.39226519
FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT>MYR	0.22162162
FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT> MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	0.34285714
FLAM>LUN>HAR>BLO	0.32978723
FLAM>LUN>HAR>BLO>KJO>REI	0.23404255
FLAM>LUN>HAR>BLO>KJO>REI>VAT	0.20212766
FLAM>LUN>HAR>BLO>KJO>REI>VAT>FLAM	0.13636364
FLAM>LUN>HAR>BLO>KJO>REI>VAT>MYR	0.25531915
FLAM>LUN>HAR>BLO>KJO>REI>VAT> MYR>VAT>REI>KJO>BLO>BER>HAR> LUN>FLAM	0.25806452
FLAM>VAT	0.5
FLAM>VAT>MYR	0.5
FLAM>VAT>MYR>VAT>FLAM	0.5
HAR>BER	0.41436464
HAR>BER>BLO	0.34254144
HAR>BER>BLO>KJO>REI	0.24861878
HAR>BER>BLO>KJO>REI>VAT	0.39779006
HAR>BER>BLO>KJO>REI>VAT>MYR	0.25414365
HAR>BLO	0.24468085
HAR>BLO>KJO>REI	0.28723404
HAR>BLO>KJO>REI>VAT	0.32978723
HAR>BLO>KJO>REI>VAT>FLAM	0.27272727
HAR>BLO>KJO>REI>VAT>MYR	0.18085106
HAR>LUN	0.29834254
HAR>LUN>FLAM	0.29834254
LUN>FLAM	0.36464088
LUN>HAR	0.26519337
LUN>HAR>BER	0.40331492
LUN>HAR>BER>BLO	0.38121547
LUN>HAR>BER>BLO>KJO>REI	0.27624309
LUN>HAR>BER>BLO>KJO>REI>VAT	0.41436464
LUN>HAR>BER>BLO>KJO>REI>VAT>MYR	0.2320442
LUN>HAR>BLO	0.30851064

Continued on next page

Table B.6 – continued from previous page

Package	Accuracy of the Prediction
LUN>HAR>BLO>KJO>REI	0.19148936
LUN>HAR>BLO>KJO>REI>VAT	0.26595745
LUN>HAR>BLO>KJO>REI>VAT>FLAM	0.22727273
LUN>HAR>BLO>KJO>REI>VAT>MYR	0.32978723
MYR>VAT	0.33149171
MYR>VAT>FLAM	1
MYR>VAT>KJO>FLAM	0.34285714
MYR>VAT>REI	0.24309392
MYR>VAT>REI>KJO>BLO	0.40883978
MYR>VAT>REI>KJO>BLO>BER	0.39779006
MYR>VAT>REI>KJO>BLO>BER>HAR	0.35911602
MYR>VAT>REI>KJO>BLO>BER>HAR>LUN	0.31491713
MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	0.38378378
REI>KJO>BLO	0.30939227
REI>KJO>BLO>BER	0.38674033
REI>KJO>BLO>BER>HAR	0.32596685
REI>KJO>BLO>BER>HAR>LUN	0.32596685
REI>KJO>BLO>BER>HAR>LUN>FLAM	0.32044199
REI>VAT	0.40883978
REI>VAT>FLAM	0.27272727
REI>VAT>MYR	0.26519337
VAT>FLAM	0.04545455
VAT>KJO>FLAM	0.2
VAT>MYR	0.28176796
VAT>REI	0.24309392
VAT>REI>KJO>BLO	0.29281768
VAT>REI>KJO>BLO>BER	0.45303867
VAT>REI>KJO>BLO>BER>HAR	0.35359116
VAT>REI>KJO>BLO>BER>HAR>LUN	0.33149171
VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM	0.3480663

Appendix C

Hybrid Model

C.1 Error of each interpolation function for each package

Table C.1: Error of each interpolation function for each package

Package	Y	Linear	Quadratic	Cubic	Best
1	1678	2927.292458	247.3242173	-971.3438714	Linear
2	612	760.0048635	805.7245094	818.8497447	Linear
3	2832	1673.54457	1881.316074	1886.089955	Cubic
4	1071	1608.919702	1659.462885	2188.876186	Linear
5	6327	6041.646039	6743.653407	7331.62891	Linear
6	9235	7050.286626	9506.516722	12186.89348	Cubic
7	1598	6397.613038	6399.229691	6559.21008	Linear
8	462	503.1079136	526.1410184	519.089279	Linear
9	565	486.1001443	619.74017	603.129535	Cubic
10	2078	3490.896354	-27848.04281	-34116.62576	Linear
11	3327	8416.475012	9604.258319	9499.681608	Linear
12	2078	3815.914021	1194.312427	1271.297821	Cubic
13	7434	1656.615844	1120.078586	1588.241597	Linear
14	7494	6946.720566	7605.825912	7054.4543	Cubic
15	8874	7006.100413	7532.71676	7870.757289	Cubic
17	6210	5360.230609	6319.506182	6481.115287	Cubic
18	565	250.3886106	-36.59904367	-135.0216194	Linear
19	1891	234.2744522	83.14924464	63.47432063	Linear
20	2021	866.1304862	1001.992458	936.2248265	Cubic
21	105	83.01125027	82.00687084	82.85204327	Linear
23	6332	2249.815001	-7684.765475	-346161.0166	Linear
24	3882	3019.906339	23570.47199	38560.20379	Linear

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
25	1071	1006.961457	1274.325126	1403.8153	Linear
26	1005	2365.225184	7002.328741	6731.219329	Linear
27	612	808.0258154	492.3337588	257.21659	Cubic
28	952	2217.094915	4003.907675	4105.613715	Linear
29	1111	1891.686029	2759.511784	2840.561834	Linear
30	7	1127.956869	1254.305195	1757.902353	Linear
31	4648	1929.92532	1975.805277	2073.123143	Cubic
32	127	4382.964888	1838.176494	2151.279412	Cubic
33	1015	2771.753116	2452.47216	2457.493628	Cubic
35	18107	3083.558872	2693.607749	2548.060697	Linear
36	3547	3023.018949	3048.039261	3047.543553	Cubic
40	428	630.907586	446.9108439	294.9054291	Cubic
41	428	630.907586	446.9108439	294.9054291	Cubic
42	702	519.5653179	399.8039113	245.5076008	Linear
43	1678	3076.170616	3244.890982	3278.416339	Linear
44	2833	3164.546653	3661.862996	3896.076371	Linear
45	400	571.1166337	-28705.45214	-126500.4256	Linear
46	5475	414.6536255	332.9348415	312.5908124	Linear
47	503	1550.961333	3359.810145	4307.531289	Linear
49	5508	4527.818573	6926.829862	9744.481771	Linear
50	2078	3490.896354	-26677.20012	-28023.11655	Linear
51	2078	3490.896354	-26677.20012	-28023.11655	Linear
52	1598	7343.764182	2373.760467	913.3927969	Cubic
53	1837	7378.115492	2473.886713	1018.768637	Cubic
54	428	2519.271523	13020.0965	37062.78453	Linear
55	702	1000.22406	1322.080164	1331.609649	Linear
56	612	811.1963329	1272.783449	1593.339456	Linear
57	952	975.25	-1432.364963	-7337.738528	Linear
58	918	843.352794	-1237.458791	-2344.111934	Linear
59	400	706.1965713	-4037.875709	-5289.658637	Linear
60	5629	4527.818573	6926.749059	9744.039552	Linear
61	5629	4527.818573	6926.749059	9744.039552	Linear
63	732	479.5654621	-2179.692057	-3127.504844	Linear
64	9486	7344.006723	8178.554652	8508.573266	Cubic
66	9	78.12567845	-360.3526955	-951.6138501	Linear
67	3852	5242.396736	1438.787948	-16.35363881	Linear

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
68	7434	6712.386323	5970.313916	5784.693934	Linear
69	3327	2687.905549	2371.026615	2511.264898	Linear
70	3852	3932.147931	3461.031196	-2720.859377	Linear
71	3852	3932.147931	3461.031196	-2720.859377	Linear
72	27834	11170.42522	10797.20127	10513.60892	Linear
73	7614	8094.827346	10989.53985	12205.24488	Linear
74	3327	2684.670687	1990.329846	2171.616842	Linear
75	7429	1306.4	-1392.009102	-12413.31964	Linear
76	3895	3274.005315	-188.8147083	-1436.45071	Linear
77	12573	10383.70555	13604.18432	14003.83228	Cubic
78	3882	1771.596128	29101.92579	324472.6651	Linear
80	7650	6363.885068	10896.41803	10995.31807	Linear
82	1891	234.9584259	101.4599377	85.25492527	Linear
83	9486	7825.40519	8731.561667	9091.117621	Cubic
84	3852	5236.550979	1420.951172	-46.18247639	Linear
85	4844	4604.601175	25495.30708	23790.21448	Linear
86	4844	4604.601175	24157.72369	21719.77936	Linear
87	6328	5107.309997	8111.921664	11975.53061	Linear
88	6328	5107.309997	8111.921664	11975.53061	Linear
89	20543	3178.230219	-51952.43773	-47054.01021	Linear
1	1678	2927.292458	247.3242173	-971.3438714	Linear
2	612	760.0048635	805.7245094	818.8497447	Linear
3	2832	1673.54457	1881.316074	1886.089955	Cubic
4	1071	1608.919702	1659.462885	2188.876186	Linear
5	6327	6041.646039	6743.653407	7331.62891	Linear
6	9235	7050.286626	9506.516722	12186.89348	Cubic
7	1598	6397.613038	6399.229691	6559.21008	Linear
8	462	503.1079136	526.1410184	519.089279	Linear
9	565	486.1001443	619.74017	603.129535	Cubic
10	2078	3490.896354	-27848.04281	-34116.62576	Linear
11	3327	8416.475012	9604.258319	9499.681608	Linear
12	2078	3815.914021	1194.312427	1271.297821	Cubic
13	7434	1656.615844	1120.078586	1588.241597	Linear
14	7494	6946.720566	7605.825912	7054.4543	Cubic
15	8874	7006.100413	7532.71676	7870.757289	Cubic
17	6210	5360.230609	6319.506182	6481.115287	Cubic

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
18	565	250.3886106	-36.59904367	-135.0216194	Linear
19	1891	234.2744522	83.14924464	63.47432063	Linear
20	2021	866.1304862	1001.992458	936.2248265	Cubic
21	105	83.01125027	82.00687084	82.85204327	Linear
23	6332	2249.815001	-7684.765475	-346161.0166	Linear
24	3882	3019.906339	23570.47199	38560.20379	Linear
25	1071	1006.961457	1274.325126	1403.8153	Linear
26	1005	2365.225184	7002.328741	6731.219329	Linear
27	612	808.0258154	492.3337588	257.21659	Cubic
28	952	2217.094915	4003.907675	4105.613715	Linear
29	1111	1891.686029	2759.511784	2840.561834	Linear
30	7	1127.956869	1254.305195	1757.902353	Linear
31	4648	1929.92532	1975.805277	2073.123143	Cubic
32	127	4382.964888	1838.176494	2151.279412	Cubic
33	1015	2771.753116	2452.47216	2457.493628	Cubic
35	18107	3083.558872	2693.607749	2548.060697	Linear
36	3547	3023.018949	3048.039261	3047.543553	Cubic
40	428	630.907586	446.9108439	294.9054291	Cubic
41	428	630.907586	446.9108439	294.9054291	Cubic
42	702	519.5653179	399.8039113	245.5076008	Linear
43	1678	3076.170616	3244.890982	3278.416339	Linear
44	2833	3164.546653	3661.862996	3896.076371	Linear
45	400	571.1166337	-28705.45214	-126500.4256	Linear
46	5475	414.6536255	332.9348415	312.5908124	Linear
47	503	1550.961333	3359.810145	4307.531289	Linear
49	5508	4527.818573	6926.829862	9744.481771	Linear
50	2078	3490.896354	-26677.20012	-28023.11655	Linear
51	2078	3490.896354	-26677.20012	-28023.11655	Linear
52	1598	7343.764182	2373.760467	913.3927969	Cubic
53	1837	7378.115492	2473.886713	1018.768637	Cubic
54	428	2519.271523	13020.0965	37062.78453	Linear
55	702	1000.22406	1322.080164	1331.609649	Linear
56	612	811.1963329	1272.783449	1593.339456	Linear
57	952	975.25	-1432.364963	-7337.738528	Linear
58	918	843.352794	-1237.458791	-2344.111934	Linear
59	400	706.1965713	-4037.875709	-5289.658637	Linear
60	5629	4527.818573	6926.749059	9744.039552	Linear

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
61	5629	4527.818573	6926.749059	9744.039552	Linear
63	732	479.5654621	-2179.692057	-3127.504844	Linear
64	9486	7344.006723	8178.554652	8508.573266	Cubic
66	9	78.12567845	-360.3526955	-951.6138501	Linear
67	3852	5242.396736	1438.787948	-16.35363881	Linear
68	7434	6712.386323	5970.313916	5784.693934	Linear
69	3327	2687.905549	2371.026615	2511.264898	Linear
70	3852	3932.147931	3461.031196	-2720.859377	Linear
71	3852	3932.147931	3461.031196	-2720.859377	Linear
72	27834	11170.42522	10797.20127	10513.60892	Linear
73	7614	8094.827346	10989.53985	12205.24488	Linear
74	3327	2684.670687	1990.329846	2171.616842	Linear
75	7429	1306.4	-1392.009102	-12413.31964	Linear
76	3895	3274.005315	-188.8147083	-1436.45071	Linear
77	12573	10383.70555	13604.18432	14003.83228	Cubic
78	3882	1771.596128	29101.92579	324472.6651	Linear
80	7650	6363.885068	10896.41803	10995.31807	Linear
82	1891	234.9584259	101.4599377	85.25492527	Linear
83	9486	7825.40519	8731.561667	9091.117621	Cubic
84	3852	5236.550979	1420.951172	-46.18247639	Linear
85	4844	4604.601175	25495.30708	23790.21448	Linear
86	4844	4604.601175	24157.72369	21719.77936	Linear
87	6328	5107.309997	8111.921664	11975.53061	Linear
88	6328	5107.309997	8111.921664	11975.53061	Linear
89	20543	3178.230219	-51952.43773	-47054.01021	Linear
1	702	748.8541735	616.1392101	551.4497971	Linear
2	918	1760.842577	2597.374593	3059.799503	Linear
3	3700	388.7801731	523.5162089	553.0217066	Cubic
4	952	1119.137417	1140.516019	1423.131851	Linear
5	3852	5235.781729	1421.878875	-44.50440768	Linear
6	7428	4052.784585	5740.945261	5946.334392	Cubic
7	5229	4295.844493	5069.814961	5268.341207	Cubic
8	4	290.5491642	323.7947348	320.6345323	Linear
9	4	290.5491642	323.7947348	320.6345323	Linear
10	3327	8551.976853	9833.803757	9732.227704	Linear

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
11	3852	5235.781729	1420.941846	-45.28034886	Linear
12	9235	7050.286626	9506.623973	12187.28745	Cubic
13	1267	6551.263898	6129.131242	5318.22648	Cubic
14	1836	3822.489175	3824.197778	2615.838192	Cubic
15	1836	3858.5962	2906.608296	2915.242067	Cubic
17	4850	6843.972014	2613.481396	68.26598147	Linear
18	4	177.2234645	178.862468	191.7115	Linear
19	119	1796.021603	3625.97311	5612.991903	Linear
20	9	85.18526688	-371.2787927	-954.8869754	Linear
21	106	38.95510963	243.2310536	685.9166457	Linear
23	9235	5051.799076	5538.075021	5716.095237	Cubic
24	9235	5051.799076	5538.075088	5716.095429	Cubic
25	702	838.0617457	20.45611804	-6249.80877	Linear
26	1837	3463.4113	3784.216244	3930.170627	Linear
27	1161	3522.51109	5486.532409	6316.500707	Linear
28	3935	8437.225728	8558.565426	8542.410571	Linear
29	1792	1376.757417	1356.372164	1285.411783	Linear
30	231	2098.473201	3167.41294	21297.14049	Linear
31	6543	7898.138654	8589.758732	8810.221089	Linear
32	6732	3003.389875	15025.4851	28793.44086	Linear
33	2106	2612.79467	12584.89992	42976.37924	Linear
35	2106	2786.961617	-152172.9358	-1157558.808	Linear
36	2610	3396.139722	7441.305634	7356.606734	Linear
40	612	540.3286319	646.5515868	933.1731554	Cubic
41	3150	1431.092025	1365.3752	27364.96297	Linear
42	612	811.1963329	1272.783449	1593.339456	Linear
43	3148	1842.319781	1795.295906	1782.637305	Linear
44	612	811.1963329	1272.768151	1593.323062	Linear
45	503	4894.587214	4689.324922	4676.518503	Cubic
46	503	1550.961333	3359.81131	4307.532557	Linear
47	5629	4527.818573	6926.749059	9744.039547	Linear
49	5475	412.946407	328.3476024	306.3681843	Linear
50	9235	5031.662689	5530.627053	5711.820705	Cubic
51	7428	2998.164808	9533.130428	24158.31281	Cubic
52	6326	2247.082225	-7740.527044	-347128.748	Linear
53	9235	5051.799076	5538.075088	5716.095429	Cubic
54	918	2639.608015	2856.110981	2745.183465	Linear
55	3150	1318.022292	1319.915115	1324.262831	Cubic
56	918	843.352794	-1237.458791	-2344.111933	Linear
57	3168	2762.350891	2301.774868	2066.175225	Linear

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
58	428	588.0187119	739.497669	1320.781821	Linear
59	5629	711.5282447	-1805.419698	-1378.181526	Linear
60	732	491.1884993	-2165.654771	-3113.253562	Linear
61	732	491.1884993	-2165.654771	-3113.253562	Linear
63	400	5515.488581	6374.307541	6481.093322	Linear
64	7485	7784.207089	6757.258255	24.25753398	Linear
66	2021	825.0845238	931.9419221	856.0618343	Cubic
67	1620	7344.692596	2380.747102	956.6889104	Cubic
68	4480	473.3090706	675.9173584	832.5804889	Cubic
69	1620	3851.618101	2912.243056	2942.303058	Cubic
70	1641	6115.987933	11088.13333	14792.35158	Linear
71	3327	4817.043753	7335.168199	8604.338343	Linear
72	17073	12963.44204	8545.313023	3134.23524	Linear
73	7430	6928.315882	7586.329621	7034.671269	Cubic
74	2078	5317.874461	4507.659307	4513.535179	Cubic
75	7613	4801.342411	-18579.79548	-28799.30715	Linear
76	7613	4801.342411	-18579.79548	-28799.30715	Linear
77	1640	6115.359648	2248.813175	1429.817079	Cubic
78	7494	5497.430184	7000.822014	5792.73213	Cubic
80	3323	8416.475012	9592.458567	9491.20967	Linear
82	1205	233.3731809	-1754.165968	-5234.907484	Linear
83	7208	7658.256931	6455.435793	-1007.397159	Linear
84	9235	5032.060417	5531.275496	5716.033058	Cubic
85	3327	2669.764784	2349.519619	2501.441478	Linear
86	3327	2669.764784	2351.161254	2491.742972	Linear
87	3327	4817.043753	7335.174287	8604.344987	Linear
88	7430	1306.4	-1392.009138	-12413.32043	Linear
89	7430	11213.60364	18607.28122	22976.66031	Linear
1	918	2638.959082	3002.163719	3002.567571	Linear
2	952	735.0257357	368.2200841	241.6450442	Linear
3	1836	2424.0508	2710.633011	2673.151239	Linear
4	702	967.5046556	620.3430342	202.3061413	Cubic
5	3878	1771.447388	29172.08443	325094.2249	Linear
6	3327	1886.941855	-3061.716803	-3097.718024	Linear
7	4850	1911.963842	-1247.462711	-1419.311789	Linear

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
8	558	508.600297	515.6112842	522.0254399	Cubic
9	154	2.373883194	-88.57103999	-162.9289225	Linear
10	3878	1771.447388	30011.30567	334696.1374	Linear
11	6327	6042.258356	6744.295707	7332.19898	Linear
12	3878	4731.224126	4598.282249	4538.407128	Cubic
13	1598	6326.165628	6327.581923	6482.179681	Linear
14	3335	2686.018844	1934.217448	2104.683713	Linear
15	7704	8222.500703	8599.262259	8551.143169	Linear
17	3323	2813.305405	2489.77947	2620.254253	Linear
18	154	3.831543284	-23.1228386	-42.67712351	Linear
20	1399	1056.631739	773.4389901	748.3338039	Linear
23	7650	6364.661045	10890.70047	11000.10276	Linear
24	3852	5634.595763	-2228.386516	-6816.86139	Linear
25	1837	3463.4113	3784.216244	3930.170627	Linear
26	2833	689.1699643	627.368736	641.8649115	Linear
27	8357	4686.096382	4736.632214	4480.111199	Cubic
28	1837	2464.398426	6683.743744	8811.693045	Linear
29	1406	1693.653716	897.5854883	697.9197903	Linear
30	72	1020.247158	1293.202468	1298.376137	Linear
31	6608	557.1874858	828.8865304	2797.856011	Cubic
32	5475	1872.621382	1920.697099	2132.023022	Cubic
33	14928	9525.57074	5294.969704	5291.344441	Linear
35	20198	11227.54813	-6793.387724	-12682.73592	Linear
36	704	1395.212227	1952.721238	2273.578658	Linear
40	952	750.6369448	947.4088729	1270.406383	Cubic
41	612	540.3286319	646.5515868	933.1731554	Cubic
42	3148	1842.319781	1795.295895	1782.63727	Linear
43	428	587.6130845	739.2789896	1320.636371	Linear
44	3168	2762.350891	2301.774869	2066.175228	Linear
45	6543	4746.953757	6477.462029	7002.769129	Cubic
46	732	491.1884993	-2165.654771	-3113.253562	Linear
47	5475	414.6536255	332.9348633	312.5908768	Linear
49	6732	3673.137185	-12564.39556	-19965.37018	Linear
50	7613	6271.931792	10835.77121	10945.09832	Linear
51	7613	6271.931792	10835.77121	10945.09832	Linear
52	3878	3018.577839	23630.20111	38650.81899	Linear
53	6332	2249.815001	-7684.765456	-346161.0143	Linear
54	3148	844.6871701	838.7056924	832.8959422	Linear

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
55	2833	1741.288011	-5215.001215	-7257.636986	Linear
56	3148	1842.319781	1795.295896	1782.637273	Linear
57	428	588.0187119	739.497669	1320.781821	Linear
58	584	1775.720153	3629.905939	5379.260233	Linear
59	6732	4207.934395	3546.714026	2351.657079	Linear
60	400	5515.488581	6369.601103	6474.647438	Linear
61	5475	414.6536255	332.9348415	312.5908124	Linear
63	3424	3276.643894	3031.897309	-1507.488201	Linear
64	12167	6434.082425	10672.62664	10758.72219	Cubic
66	1311	974.2190963	636.1239942	615.6905853	Linear
67	7431	5513.247679	6990.738065	5682.281373	Cubic
68	7704	8202.788185	8585.534268	8544.749433	Linear
69	4480	473.3090706	742.6463204	1073.267168	Cubic
70	6329	5107.309997	8111.921662	11975.53058	Linear
71	3895	3274.036728	-188.8085897	-1436.444164	Linear
72	11521	27115.10669	30838.27895	31303.94241	Linear
73	3878	2080.983737	2527.22732	3300.095631	Cubic
74	3878	2080.983737	2527.227685	3300.097698	Cubic
75	2078	9182.754591	10326.37214	10474.80679	Linear
76	2078	9182.754591	10326.37214	10474.80679	Linear
77	10458	11575.15817	11612.06556	10980.4501	Cubic
78	3852	5293.546554	1518.577113	157.6686014	Linear
80	3852	5199.698093	1618.506313	228.8060879	Linear
83	13179	6517.499953	10647.98882	10731.36035	Cubic
84	6328	6042.870673	6745.042271	7333.229348	Linear
85	4480	473.3090706	675.9176653	832.5867804	Cubic
86	4480	473.3090706	742.6465728	1073.27242	Cubic
87	3895	3274.005315	-188.8172383	-1436.454763	Linear
88	7614	4801.342411	-18579.79549	-28799.3072	Linear
89	4807	10944.45277	10850.92568	10915.6195	Cubic
1	612	540.3286319	646.5520498	933.1734667	Cubic
2	428	768.1879753	893.9785719	1294.432751	Linear
3	826	1365.440986	194.5106761	276.8900963	Linear
4	856	1020.589891	451.0405268	186.3861543	Linear
5	3327	8416.475012	9604.258342	9499.681724	Linear
6	3878	4730.728288	4597.809628	4537.944865	Cubic
7	4844	3861.109058	22133.42123	21963.03623	Linear
10	7429	5514.939592	6990.868171	5724.262735	Cubic
11	1598	7343.764182	2378.161321	954.0403597	Cubic
12	1267	6718.669841	5953.446585	5740.09854	Cubic

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
13	4850	1911.963842	-1165.823126	-1335.722203	Linear
14	9235	3492.458159	1445.861416	1719.233659	Linear
15	7434	6747.591542	5983.475961	5823.7329	Linear
17	1836	3848.129051	2973.915001	2994.67091	Cubic
23	7496	2955.885928	9231.362878	23143.34989	Cubic
24	1837	7378.115492	2473.886713	1018.768637	Cubic
25	1005	2365.225184	7002.328741	6731.219329	Linear
26	581	2388.621436	2690.660194	2590.634182	Linear
27	952	1063.929223	2305.159391	3145.116568	Linear
28	2492	1317.416449	1013.030136	873.1781577	Linear
29	2417	1736.215402	1311.765003	1189.205376	Linear
30	1310	660.057082	-298.3473053	-460.8372141	Linear
31	503	3496.398069	2214.331028	1796.932363	Cubic
32	732	4629.32184	4375.672941	3945.213154	Cubic
33	19701	12450.53879	10826.15909	11668.1892	Linear
35	16152	18571.04403	6614.941216	-96738.9908	Linear
36	3067	3474.822012	3591.712696	3589.052827	Linear
40	918	2638.959082	3002.163719	3002.56757	Linear
41	918	2638.959082	3002.163719	3002.56757	Linear
42	3168	2762.350891	2301.774868	2066.175225	Linear
43	2833	3164.546653	3661.862996	3896.076371	Linear
44	428	587.6130845	739.2789896	1320.636371	Linear
45	5475	2980.448411	2755.895092	2815.33367	Linear
46	6732	3750.672345	-12373.6744	-19793.3814	Linear
47	127	4382.336603	8039.21388	459625.9302	Linear
49	127	4168.719659	8050.085599	461328.9264	Linear
50	3852	5511.040526	-2719.432424	-7554.531016	Linear
51	9235	5031.662689	5530.627053	5711.820705	Cubic
52	3852	5511.040526	-2719.432424	-7554.531016	Linear
53	3882	3019.906339	23570.47199	38560.20379	Linear
54	2833	1741.288011	-5215.001215	-7257.636986	Linear
55	1678	3114.280843	4809.554075	4784.781403	Linear
56	2833	3164.546653	3661.862945	3896.076317	Linear
57	3148	1842.319781	1795.295895	1782.63727	Linear
58	3168	2762.350891	2301.774868	2066.175225	Linear
59	134	5964.934552	6043.995263	5999.299732	Linear
60	127	4382.964888	8039.612065	459626.1861	Linear
61	6732	3750.672345	-12373.6744	-19793.3814	Linear
63	6732	3673.137185	-12564.39556	-19965.37021	Linear

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
64	4159	9013.778444	7474.538157	7193.06877	Cubic
67	7614	6280.46754	10839.85487	10938.38214	Linear
68	7431	5741.090833	7133.641745	7075.173473	Cubic
69	4850	8554.098502	7809.240513	6154.668927	Cubic
70	7614	4801.983197	-18578.2267	-28797.90739	Linear
71	9235	2255.692993	2127.999417	2095.599461	Linear
72	1993	215.5	-1480.468827	-2847.667433	Linear
73	6328	5069.652641	5201.489198	5546.499414	Cubic
74	1598	3818.535379	3823.468501	2622.023082	Cubic
75	3852	3931.553558	3460.30232	-2721.571642	Linear
76	3327	4817.043753	7335.172489	8604.342278	Linear
77	7429	6598.212955	4504.493953	5001.805503	Linear
78	7650	6363.885068	10889.30741	10988.00763	Linear
80	9235	5043.969542	5530.607786	5712.403746	Cubic
83	1835	13061.38382	16943.61898	16166.17996	Linear
84	7430	5515.779074	6988.749768	5666.656365	Cubic
85	7430	5740.286673	7132.915254	7079.261344	Cubic
86	7434	6708.800606	5401.52916	4812.532323	Linear
87	7430	1306.4	-1392.009138	-12413.32043	Linear
88	3852	3931.553558	3460.071951	-2722.023458	Linear
89	22089	9066.612781	9249.308738	9346.68054	Cubic
1	952	750.4764268	947.3631917	1270.387843	Cubic
2	3148	1485.073187	222.9582819	279.5737797	Linear
3	3933	3779.138584	21456.92725	20282.96467	Linear
4	2832	1673.54457	1881.316074	1886.089955	Cubic
5	2078	3490.896354	-27004.99955	-33066.08592	Linear
6	1598	2251.001157	7482.92154	27845.03331	Linear
7	1267	6551.263898	6129.059466	5317.581649	Cubic
10	7613	6365.569551	11040.07705	11141.19566	Linear
11	7429	5514.779074	6987.336571	5665.709636	Cubic
12	3327	1886.941855	-3061.159734	-3096.770625	Linear
13	5229	4295.844493	5073.908069	5276.69179	Cubic
14	2078	5323.178626	4538.176193	4551.169634	Cubic
15	4850	6843.972014	2613.486898	68.21028684	Linear
17	8874	7006.100413	7532.742798	7870.602493	Cubic
23	3882	3019.906339	23570.47201	38560.20401	Linear
24	7495	2955.885928	9231.362878	23143.34989	Cubic
25	2833	689.1699643	627.368736	641.8649115	Linear

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
26	952	987.2673318	665.6386569	480.2610621	Linear
27	1346	1109.793467	-1816.146081	-2343.383929	Linear
28	584	1797.354165	2376.696067	3584.26721	Linear
29	2133	2301.808152	3194.562843	3296.023	Linear
31	134	5295.824713	4384.488947	2773.885016	Cubic
32	1943	3137.20168	4727.715124	5632.552018	Linear
33	6543	10208.02643	17310.35261	20169.7389	Linear
36	1834	1521.811129	97.56080391	-337.5566794	Linear
40	1678	2927.292458	247.3242174	-971.3438713	Linear
41	1836	894.3423427	-2767.638843	-2982.791166	Linear
42	952	975.25	-1432.364963	-7337.738528	Linear
43	3168	2762.350891	2301.774869	2066.175228	Linear
44	702	519.5653179	399.8979187	245.5819715	Linear
46	127	4382.964888	8039.612065	459626.1861	Linear
47	400	5515.488581	6369.601083	6474.647379	Linear
49	491	1550.961333	3359.424526	4307.147039	Linear
50	7428	2998.164808	9533.130428	24158.31281	Cubic
51	3878	3018.577839	23630.20111	38650.81899	Linear
52	2078	3490.896354	-26677.20012	-28023.11655	Linear
53	2078	3498.409092	-26614.1529	-27957.35529	Linear
54	1837	1917.91337	8146.632831	43060.48786	Linear
55	1005	767.7859141	134.0056498	-70.90571851	Linear
56	952	975.5625	-1432.211197	-7337.643422	Linear
57	1678	3076.170616	3244.890996	3278.416391	Linear
58	702	519.5653179	399.8039113	245.5076008	Linear
60	5475	414.6536255	332.9348415	312.5908124	Linear
61	400	5515.488581	6369.601103	6474.647438	Linear
63	491	1550.961333	3359.425691	4307.148307	Linear
64	1620	12114.45571	14243.3371	13351.86615	Linear
67	3327	8416.475012	9604.567004	9500.40439	Linear
68	8838	7006.100413	7528.355105	7860.001218	Cubic
69	9833	5580.75713	6139.086342	6962.64038	Cubic
70	3895	3274.036728	-188.8085897	-1436.444164	Linear
71	2078	9182.786802	10326.37859	10474.81244	Linear

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
72	1403	793.49023	-4338.16166	-5425.192766	Linear
73	2078	5317.874461	4507.643799	4513.514836	Cubic
74	3852	5219.022665	5468.943931	5082.259586	Cubic
75	3327	4817.043753	7335.172489	8604.342278	Linear
76	6328	5106.976316	8111.614357	11975.14247	Linear
77	9235	4601.805049	4888.052382	5054.504059	Cubic
78	2078	3497.706544	-26947.11871	-32997.64378	Linear
80	7154	5395.013376	6825.285418	5579.985991	Cubic
83	4290	9013.778444	6899.854328	6583.632566	Cubic
84	3327	8416.475012	9603.682269	9499.063429	Linear
85	7434	6708.800606	5965.609907	5777.443562	Linear
86	4850	8554.098502	7809.304961	6154.880477	Cubic
87	2078	9182.786802	10326.38321	10474.82129	Linear
88	1640	6115.987933	11088.13336	14792.35188	Linear
89	1640	6261.666667	-13645.43814	-102768.5785	Linear
1	3150	1431.092025	1365.3752	27364.96298	Linear
2	2833	2174.121806	2186.865304	2278.115528	Cubic
3	702	967.5046556	620.3430342	202.3061413	Cubic
4	1836	2424.0508	2710.633011	2673.151239	Linear
5	7613	6271.931792	10834.30004	10932.78778	Linear
6	6326	1822.180452	-140.3057677	280.4962166	Linear
7	3359	473.3090706	413.7598088	-217.9435055	Linear
10	6327	6042.258356	6738.180171	7298.819613	Linear
11	3878	1771.447388	29172.08441	325094.2233	Linear
12	7429	4052.855984	5741.005914	5946.439493	Cubic
13	4844	3861.109058	22123.85649	21941.73745	Linear
14	6331	5091.4329	5231.502503	5579.917989	Cubic
15	4848	4617.865976	25597.73804	23882.85602	Linear
17	7704	8222.500703	8599.192112	8551.473542	Linear
23	3336	8416.475012	9592.271895	9475.475257	Linear
24	3336	8416.475012	9592.271855	9475.475119	Linear
25	3933	3189.518163	16965.68278	16005.70931	Linear
26	3935	2137.665342	1815.241477	1476.071985	Linear
27	584	1775.720153	836.3488813	188.7765631	Cubic
28	1436	2587.630135	2859.648679	3057.938309	Linear
29	743	5543.667988	1826.39111	1833.288624	Cubic
33	2094	1060.73518	675.6744241	673.8943091	Linear
40	2833	2979.880434	4325.305069	4389.707241	Linear
41	1678	2927.292458	247.3242174	-971.3438713	Linear

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
42	918	843.352794	-1237.458791	-2344.111934	Linear
43	702	519.5653179	399.8979187	245.5819715	Linear
44	1678	3076.170616	3244.890982	3278.416339	Linear
50	6326	2247.082225	-7740.527044	-347128.748	Linear
51	6326	2247.082225	-7740.527044	-347128.748	Linear
52	3327	8416.475012	9603.984769	9483.777061	Linear
53	3336	8416.475012	9592.271855	9475.475119	Linear
54	1005	767.7859141	134.0056498	-70.90571851	Linear
55	952	2975.215415	4390.310162	4328.36584	Linear
56	584	1775.720153	3629.815223	5379.205218	Linear
57	2833	3164.546653	3661.862997	3896.076373	Linear
58	612	811.1963329	1272.783449	1593.339456	Linear
64	9034	9015.57199	10437.87782	13287.14562	Linear
67	6329	6043.482991	6743.493374	7325.892069	Linear
68	6329	5587.138645	6571.755742	6706.263962	Cubic
69	7434	6712.386323	5408.003816	4822.03073	Linear
70	3327	4817.043753	7335.168199	8604.338343	Linear
71	7614	4801.983197	-18578.2267	-28797.90739	Linear
72	1053	197.0311799	986.9551077	7966.483929	Cubic
73	1598	3818.535379	3823.468825	2622.026746	Cubic
74	6328	5069.312324	5201.166365	5546.124315	Cubic
75	1640	6115.359648	11087.34588	14791.68855	Linear
76	3852	3931.553558	3460.30232	-2721.571642	Linear
77	3895	8034.004183	8285.437806	7888.682948	Cubic
78	6331	6082.05897	6764.969848	7300.841815	Linear
80	6209	5873.871147	6546.487501	7106.298204	Linear
83	8905	8885.891483	10408.2764	13588.35163	Linear
84	3878	1771.447388	29171.6006	325085.7957	Linear
85	4850	6824.107358	2597.146296	63.52911668	Linear
86	9833	5580.75713	6138.962144	6962.223953	Cubic
87	1640	6115.987933	11088.13336	14792.35188	Linear
88	9235	2255.692993	2127.995998	2095.593508	Linear
89	12934	19682.03498	24523.92861	24587.52929	Linear
1	3150	1431.092025	1365.3752	27364.96298	Linear
2	2833	2174.121806	2186.865304	2278.115528	Cubic
3	702	967.5046556	620.3430342	202.3061413	Cubic
4	1836	2424.0508	2710.633011	2673.151239	Linear
5	7613	6271.931792	10834.30004	10932.78778	Linear

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
6	6326	1822.180452	-140.3057677	280.4962166	Linear
7	3359	473.3090706	413.7598088	-217.9435055	Linear
10	6327	6042.258356	6738.180171	7298.819613	Linear
11	3878	1771.447388	29172.08441	325094.2233	Linear
12	7429	4052.855984	5741.005914	5946.439493	Cubic
13	4844	3861.109058	22123.85649	21941.73745	Linear
14	6331	5091.4329	5231.502503	5579.917989	Cubic
15	4848	4617.865976	25597.73804	23882.85602	Linear
17	7704	8222.500703	8599.192112	8551.473542	Linear
23	3336	8416.475012	9592.271895	9475.475257	Linear
24	3336	8416.475012	9592.271855	9475.475119	Linear
25	3933	3189.518163	16965.68278	16005.70931	Linear
26	3935	2137.665342	1815.241477	1476.071985	Linear
27	584	1775.720153	836.3488813	188.7765631	Cubic
28	1436	2587.630135	2859.648679	3057.938309	Linear
29	743	5543.667988	1826.39111	1833.288624	Cubic
33	2094	1060.73518	675.6744241	673.8943091	Linear
40	2833	2979.880434	4325.305069	4389.707241	Linear
41	1678	2927.292458	247.3242174	-971.3438713	Linear
42	918	843.352794	-1237.458791	-2344.111934	Linear
43	702	519.5653179	399.8979187	245.5819715	Linear
44	1678	3076.170616	3244.890982	3278.416339	Linear
50	6326	2247.082225	-7740.527044	-347128.748	Linear
51	6326	2247.082225	-7740.527044	-347128.748	Linear
52	3327	8416.475012	9603.984769	9483.777061	Linear
53	3336	8416.475012	9592.271855	9475.475119	Linear
54	1005	767.7859141	134.0056498	-70.90571851	Linear
55	952	2975.215415	4390.310162	4328.36584	Linear
56	584	1775.720153	3629.815223	5379.205218	Linear
57	2833	3164.546653	3661.862997	3896.076373	Linear
58	612	811.1963329	1272.783449	1593.339456	Linear
64	9034	9015.57199	10437.87782	13287.14562	Linear
67	6329	6043.482991	6743.493374	7325.892069	Linear
68	6329	5587.138645	6571.755742	6706.263962	Cubic
69	7434	6712.386323	5408.003816	4822.03073	Linear
70	3327	4817.043753	7335.168199	8604.338343	Linear
71	7614	4801.983197	-18578.2267	-28797.90739	Linear

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
72	1053	197.0311799	986.9551077	7966.483929	Cubic
73	1598	3818.535379	3823.468825	2622.026746	Cubic
74	6328	5069.312324	5201.166365	5546.124315	Cubic
75	1640	6115.359648	11087.34588	14791.68855	Linear
76	3852	3931.553558	3460.30232	-2721.571642	Linear
77	3895	8034.004183	8285.437806	7888.682948	Cubic
78	6331	6082.05897	6764.969848	7300.841815	Linear
80	6209	5873.871147	6546.487501	7106.298204	Linear
83	8905	8885.891483	10408.2764	13588.35163	Linear
84	3878	1771.447388	29171.6006	325085.7957	Linear
85	4850	6824.107358	2597.146296	63.52911668	Linear
86	9833	5580.75713	6138.962144	6962.223953	Cubic
87	1640	6115.987933	11088.13336	14792.35188	Linear
88	9235	2255.692993	2127.995998	2095.593508	Linear
89	12934	19682.03498	24523.92861	24587.52929	Linear
1	1836	894.3423427	-2767.638846	-2982.791171	Linear
2	1836	3101.762965	3065.466508	2237.718766	Cubic
3	3935	2514.104842	-1622.832307	-3818.645889	Linear
4	3700	388.7801731	523.5162089	553.0217066	Cubic
7	3327	4218.223116	2709.95124	2223.775491	Cubic
13	3327	4218.223116	2708.042596	2219.945342	Cubic
15	6331	5598.854015	6634.383757	6834.66069	Cubic
17	7154	5644.612337	6988.590842	6898.290416	Cubic
25	2811	388.7801731	457.40582	315.3799752	Cubic
26	1071	1006.961457	1274.325126	1403.8153	Linear
27	5233	7927.549602	9242.79126	9932.052976	Linear
28	2833	1574.04797	-571.7795587	-1535.334196	Linear
29	8283	8766.697786	8911.518458	8828.008041	Linear
40	702	748.8541735	616.1186771	551.4410304	Linear
41	2833	2979.880434	4325.305069	4389.707241	Linear
42	584	1775.720153	3629.905939	5379.260233	Linear
43	952	975.25	-1432.364963	-7337.738525	Linear
44	584	1775.720153	3629.905939	5379.260231	Linear
54	952	2975.215415	4390.310162	4328.36584	Linear
55	428	2519.271523	13020.0965	37062.78453	Linear
56	1678	3076.170616	3244.890996	3278.416388	Linear
57	702	519.5653179	399.8039113	245.5076008	Linear
58	2833	3164.546653	3661.862997	3896.076373	Linear
64	5357	4937.43332	-20276.54314	-19999.92336	Linear

Continued on next page

Table C.1 – continued from previous page

Package	Y	Linear	Quadratic	Cubic	Best
68	4844	4604.601175	25495.31156	23790.24716	Linear
69	4844	4604.601175	24157.72709	21719.80538	Linear
72	10685	14608.4466	-146901.3927	-1271770.451	Linear
83	5357	5085.983021	-19011.45437	-18410.10339	Linear
85	6328	5592.698307	6572.067114	6699.869562	Cubic
86	7430	5740.286673	7132.870126	7080.020047	Cubic
89	2832	10055.67193	11738.09901	12068.23682	Linear
28	8770	11144.29981	54556.77026	150998.0354	Linear
29	2019	928.5982762	815.1022609	735.9261979	Linear
72	9153	25204.80984	50860.47554	53370.06134	Linear
89	10685	13646.10191	66185.4253	187901.396	Linear
29	973	114.1830766	-1288.808302	-16720.37806	Linear
72	930	6351.340401	6242.07359	7865.102612	Cubic
29	8770	11932.47755	21994.49335	36186.58982	Linear

C.2 Total error of interpolation methods and best interpolation function for each package

Table C.2: Total error of interpolation methods and best interpolation function for each package

Package	Linear	Quadratic	Cubic	Best
1	257.6511294	487.3490563	2082.462483	Linear
2	240.2507302	328.1442973	330.0676115	Linear
3	326.0287906	1559.55462	1538.775012	Linear
4	326.0287906	1559.55462	1538.775012	Linear
5	863.705575	3585.009762	29386.15524	Linear
6	772.9680212	1089.771352	2575.497169	Linear
7	1023.235541	2002.988142	1976.859088	Linear
8	47.54122617	57.90440917	63.95030834	Linear
9	47.54122617	57.90440917	63.95030834	Linear
10	875.419008	3695.96075	30264.59594	Linear
11	863.7221972	3585.016645	29386.1566	Linear
12	773.0902434	1089.811636	2575.186374	Linear
13	1010.839392	1989.416855	1961.91566	Linear
14	669.4675446	858.8526582	854.4173713	Linear
15	473.0578001	1780.095517	1675.92193	Linear
17	468.7122925	1780.259488	1675.87907	Linear
18	58.69023687	102.2875705	142.0968243	Linear
19	484.7375228	889.1482631	1490.117968	Linear
20	184.4973782	210.9923031	261.318607	Linear

Continued on next page

Table C.2 – continued from previous page

Package	Linear	Quadratic	Cubic	Best
21	13.81299308	27.17701693	115.147495	Linear
23	990.218184	3532.820129	32367.82559	Linear
24	990.1703554	3532.83561	32367.83124	Linear
25	390.1620177	1261.566078	1323.702165	Linear
26	390.1620177	1261.566078	1323.702165	Linear
27	441.8454953	643.6129076	755.8869007	Linear
28	573.5847628	3630.162709	10997.73489	Linear
29	416.3442765	921.6530541	2187.637885	Linear
30	279.8635114	554.6977392	2392.029296	Linear
31	1005.341133	1744.763594	2746.151019	Linear
32	953.1863376	1397.750519	2508.855054	Linear
33	1657.675463	3324.322446	9492.845422	Linear
35	2884.213659	18131.90939	129652.888	Linear
36	475.4115798	664.4036639	16891.06303	Linear
40	257.6502612	487.3490781	2082.462506	Linear
41	257.6502612	487.3490781	2082.462506	Linear
42	193.7047901	423.5589668	877.3421445	Linear
43	193.7024779	423.5569721	877.3404788	Linear
44	193.7024779	423.5569721	877.3404788	Linear
45	671.5864072	3346.892594	14124.43571	Linear
46	899.8406351	2252.954963	46039.82851	Linear
47	899.8109539	2252.940947	46039.80296	Linear
49	891.4934745	2270.957682	46211.0533	Linear
50	995.40419	3550.98052	32462.25165	Linear
51	995.40419	3550.98052	32462.25165	Linear
52	995.40419	3550.98052	32462.25165	Linear
53	990.1703554	3532.83561	32367.83124	Linear
54	402.9616016	1438.407545	4696.769776	Linear
55	402.9616016	1438.407545	4696.769776	Linear
56	193.7050509	423.5484485	877.3338239	Linear
57	193.7047901	423.5589668	877.3421445	Linear
58	193.7047901	423.5589668	877.3421445	Linear
59	1123.467378	1424.281051	1581.874692	Linear
60	899.8406351	2252.954963	46039.82851	Linear
61	899.8406351	2252.954963	46039.82851	Linear
63	891.5219286	2270.971606	46211.07885	Linear
64	1148.991292	2935.208812	14513.69416	Linear
66	190.6646535	223.5210652	272.8678276	Linear
67	862.6086513	3584.957621	29387.02669	Linear
68	477.1915313	1773.428611	1670.205833	Linear
69	630.3488677	1708.504082	1497.041637	Linear
70	1179.830294	2884.858201	4209.691239	Linear
71	1179.830294	2884.858201	4209.691239	Linear
72	2198.596347	12105.34138	91743.88362	Linear
73	679.8254444	862.218024	852.13132	Linear

Continued on next page

Table C.2 – continued from previous page

Package	Linear	Quadratic	Cubic	Best
74	679.8472723	862.2317498	852.143206	Linear
75	1179.810233	2884.913075	4209.764239	Linear
76	1179.810233	2884.913075	4209.764239	Linear
77	827.1237266	1287.465289	2592.78639	Linear
78	853.2625681	3575.912716	29328.72451	Linear
80	848.6794962	3585.464206	29398.02931	Linear
82	474.9722348	874.7511527	1468.081313	Linear
83	1224.822635	2897.371899	13508.60589	Linear
84	863.7812594	3584.939959	29385.38421	Linear
85	478.0405353	1773.6485	1670.427189	Linear
86	630.9921571	1708.785824	1497.411201	Linear
87	1179.8912	2885.036962	4209.902385	Linear
88	1179.8912	2885.036962	4209.902385	Linear
89	2092.871813	7412.844297	16805.36368	Linear

C.3 Percentage increase of revenue from better prices

Table C.3: Percentage increase of revenue from better prices

Package No	Percentage Increase of Revenue
1	47.93
2	50.00
3	100.27
4	100.27
5	84.80
6	120.59
7	91.12
8	34.65
9	34.65
10	85.77
11	84.81
12	120.58
13	90.09
14	72.04
15	112.63
17	106.14
18	44.29
19	72.75
20	29.03
21	14.16
23	100.09
24	100.09
25	76.36
26	76.36

Continued on next page

Table C.3 – continued from previous page

Package No	Percentage Increase of Revenue
27	74.91
28	12.12
29	19.47
30	78.66
31	112.03
32	95.99
33	105.16
35	77.74
36	20.14
40	47.96
41	47.96
42	47.09
43	47.09
44	47.09
45	92.16
46	108.31
47	108.31
49	110.68
50	100.68
51	100.68
52	100.68
53	100.09
54	111.76
55	111.76
56	47.09
57	47.09
58	47.09
59	93.67
60	108.31
61	108.31
63	110.67
64	80.14
66	30.27
67	84.77
68	112.20
69	112.91
70	81.23
71	81.23
72	46.19
73	72.72
74	72.72
75	81.24
76	81.24
77	65.52
78	84.47

Continued on next page

Table C.3 – continued from previous page

Package No	Percentage Increase of Revenue
80	82.50
82	75.53
83	78.40
84	84.81
85	112.21
86	112.92
87	81.24
88	81.24
89	95.46

Appendix D

Source Code SnapShot

D.1 Hybrid Model

```python
import pandas as pd
import dill
from keras.models import load_model
import math


def predict_better_price(Package_Number, DepartureDate, Season, AvgTemp, SunHour, TotalSnow, UVIndex):
    # Package Names
    packages = ['BER>BLO', 'BER>BLO>KJO>REI', 'BER>BLO>KJO>REI>VAT', 'BER>BLO>KJO>REI>VAT>MYR', 'BER>HAR',
                'BER>HAR>LUN', 'BER>HAR>LUN>FLAM', 'BER>KJO>VAT', 'BER>KJO>VAT>MYR', 'BLO>BER', 'BLO>BER>HAR',
                'BLO>BER>HAR>LUN', 'BLO>BER>HAR>LUN>FLAM', 'BLO>KJO>REI', 'BLO>KJO>REI>VAT', 'BLO>KJO>REI>VAT>FLAM',
                'BLO>KJO>REI>VAT>MYR', 'FLAM>BER', 'FLAM>BER>KJO>VAT', 'FLAM>BER>KJO>VAT>MYR',
                'FLAM>BER>KJO>VAT>MYR>VAT>KJO>FLAM', 'FLAM>BER>KJO>VAT>MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM',
                'FLAM>LUN', 'FLAM>LUN>HAR', 'FLAM>LUN>HAR>BER', 'FLAM>LUN>HAR>BER>BLO', 'FLAM>LUN>HAR>BER>BLO>KJO>REI',
                'FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT', 'FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT>MYR',
                'FLAM>LUN>HAR>BER>BLO>KJO>REI>VAT>MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM', 'FLAM>LUN>HAR>BLO',
                'FLAM>LUN>HAR>BLO>KJO>REI', 'FLAM>LUN>HAR>BLO>KJO>REI>VAT', 'FLAM>LUN>HAR>BLO>KJO>REI>VAT>FLAM',
                'FLAM>LUN>HAR>BLO>KJO>REI>VAT>MYR', 'FLAM>LUN>HAR>BLO>KJO>REI>VAT>MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM',
                'FLAM>VAT', 'FLAM>VAT>MYR', 'FLAM>VAT>MYR>VAT>FLAM', 'HAR>BER', 'HAR>BER>BLO', 'HAR>BER>BLO>KJO>REI',
                'HAR>BER>BLO>KJO>REI>VAT', 'HAR>BER>BLO>KJO>REI>VAT>MYR', 'HAR>BLO', 'HAR>BLO>KJO>REI',
                'HAR>BLO>KJO>REI>VAT', 'HAR>BLO>KJO>REI>VAT>FLAM', 'HAR>BLO>KJO>REI>VAT>MYR', 'HAR>LUN', 'HAR>LUN>FLAM',
                'LUN>FLAM', 'LUN>HAR', 'LUN>HAR>BER', 'LUN>HAR>BER>BLO', 'LUN>HAR>BER>BLO>KJO>REI',
                'LUN>HAR>BER>BLO>KJO>REI>VAT', 'LUN>HAR>BER>BLO>KJO>REI>VAT>MYR', 'LUN>HAR>BLO', 'LUN>HAR>BLO>KJO>REI',
                'LUN>HAR>BLO>KJO>REI>VAT', 'LUN>HAR>BLO>KJO>REI>VAT>FLAM', 'LUN>HAR>BLO>KJO>REI>VAT>MYR', 'MYR>VAT',
                'MYR>VAT>FLAM', 'MYR>VAT>KJO>FLAM', 'MYR>VAT>REI', 'MYR>VAT>REI>KJO>BLO', 'MYR>VAT>REI>KJO>BLO>BER',
                'MYR>VAT>REI>KJO>BLO>BER>HAR', 'MYR>VAT>REI>KJO>BLO>BER>HAR>LUN',
                'MYR>VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM', 'REI>KJO>BLO', 'REI>KJO>BLO>BER', 'REI>KJO>BLO>BER>HAR',
                'REI>KJO>BLO>BER>HAR>LUN', 'REI>KJO>BLO>BER>HAR>LUN>FLAM', 'REI>VAT', 'REI>VAT>FLAM', 'REI>VAT>MYR',
                'VAT>FLAM', 'VAT>KJO>FLAM', 'VAT>MYR', 'VAT>REI', 'VAT>REI>KJO>BLO', 'VAT>REI>KJO>BLO>BER',
                'VAT>REI>KJO>BLO>BER>HAR', 'VAT>REI>KJO>BLO>BER>HAR>LUN', 'VAT>REI>KJO>BLO>BER>HAR>LUN>FLAM']

    Package_Number = int(Package_Number)

    # Model to use
    Model_DNN = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18, 19, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32,
                 33, 35, 36, 40, 41, 42, 43, 44, 45, 46, 47, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 63, 64,
                 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 80, 82, 83, 84, 85, 86, 87, 88, 89]
    Model_OLS = [20, 21, 66]
    Model_Regression = [16, 22, 34, 37, 38, 39, 48, 62, 65, 79, 81]
```

Figure D.1: Sample Code 1

```python
    # For packages using DNN
    if Package_Number in Model_DNN:
        DepartureYear, DepartureMonth, DepartureDay = DepartureDate.split("-")
        # File Name
        package_name = packages[Package_Number - 1].replace(">", "_")
        file_name = str(Package_Number) + "_" + package_name

        if Season == "AUT":
            SeasonAUT, SeasonSPR, SeasonSUM, SeasonWIN = 1, 0, 0, 0
        elif Season == "SPR":
            SeasonAUT, SeasonSPR, SeasonSUM, SeasonWIN = 0, 1, 0, 0
        elif Season == "SUM":
            SeasonAUT, SeasonSPR, SeasonSUM, SeasonWIN = 0, 0, 1, 0
        else:
            SeasonAUT, SeasonSPR, SeasonSUM, SeasonWIN = 0, 0, 0, 1

        # Create vector to predict
        x_input = pd.DataFrame(
            {'DepartureYear': DepartureYear, 'DepartureMonth': DepartureMonth, 'DepartureDay': DepartureDay,
             'avgtempC': AvgTemp, 'sunHour': SunHour, 'totalSnow_cm': TotalSnow, 'uvIndex': UVIndex,
             'season_AUT': SeasonAUT, 'season_SPR': SeasonSPR, 'season_SUM': SeasonSUM, 'season_WIN': SeasonWIN},
            index=[0])
        print(x_input)

        with open("dnn_files/scales/" + str(Package_Number) + "_" + package_name + "_Scale.pkl", 'rb') as scale_file:
            scale = dill.load(scale_file)

        x_input = scale.fit_transform(x_input)
        print(x_input)

        # Load DNN model

        model = load_model("dnn_files/models/" + str(Package_Number) + "_" + package_name + ".h5")

        price_predicted = model.predict(x_input)

        print(price_predicted)

        with open('common_files/max_min_dict.pkl', 'rb') as min_max_dict_file:
            min_max_dict = dill.load(min_max_dict_file)
```

Figure D.2: Sample Code 2

```
        if min_max_dict[int(Package_Number)]['len'] == 1:
            print("One data point can't interpolate")
        elif min_max_dict[int(Package_Number)]['len'] == 2:
            print("Two data points. No local maxima")
        else:
            if price_predicted >= min_max_dict[int(Package_Number)]['min'] and price_predicted <= \
                    min_max_dict[int(Package_Number)]['max']:
                with open("common_files/interpolation_functions/" + file_name + "_Linear.pkl",
                          'rb') as interpolationfile:
                    f1 = dill.load(interpolationfile)
                    predicted_sales = f1(price_predicted)
                    print(predicted_sales)
                    # better_price=fminbound(lambda x:-f1(x),price_predicted,max(x))
                    better_price = find_max_revenue(f1, price_predicted, min_max_dict[int(Package_Number)]['max'])
                    print(better_price)
                    better_sales = f1(better_price)
                    print(better_sales)
            else:
                print("Value out of interpolation range")

    # For packages using OLS

    elif Package_Number in Model_OLS:
        DepartureYear, DepartureMonth, DepartureDay = DepartureDate.split("-")
        # File Name
        DepartureYear, DepartureMonth, DepartureDay = DepartureDate.split("-")
        # File Name
        package_name = packages[Package_Number - 1].replace(">", "_")
        file_name = str(Package_Number) + "_" + package_name

        if Season == "AUT":
            Season = 1
        elif Season == "SPR":
            Season = 2
        elif Season == "SUM":
            Season = 3
        else:
            Season = 4
```

Figure D.3: Sample Code 3

```
        data_vector = [
            [int(DepartureDay), int(DepartureMonth), int(DepartureYear), int(Season), float(AvgTemp), float(SunHour),
             float(TotalSnow), int(UVIndex)]]

        # print("data/Final/OLS Models/"+file_name+"_ols.pkl")

        with open("ols_files/models/" + file_name + "_ols.pkl", 'rb') as ols_model_file:
            model = dill.load(ols_model_file)

        price_predicted = model.predict(data_vector)
        # # Get Range of the package
        # df_train = pd.read_csv('data/Final/InterpolationFiles/' + file_name + "_sales.csv")
        #
        # x = df_train['DeflatePrice']

        with open('common_files/max_min_dict.pkl', 'rb') as min_max_dict_file:
            min_max_dict = dill.load(min_max_dict_file)

        if min_max_dict[int(Package_Number)]['len'] == 1:
            print("One data point can't interpolate")
        elif min_max_dict[int(Package_Number)]['len'] == 2:
            print("Two data points. No local maxima")
        else:
            if price_predicted >= min_max_dict[int(Package_Number)]['min'] and price_predicted <= \
                    min_max_dict[int(Package_Number)]['max']:
                with open("common_files/interpolation_functions/" + file_name + "_Linear.pkl",
                          'rb') as interpolationfile:
                    f1 = dill.load(interpolationfile)
                    print(price_predicted)
                    predicted_sales = f1(price_predicted)
                    print(predicted_sales)
                    # better_price=fminbound(lambda x:-f1(x),price_predicted,max(x))
                    better_price = find_max_revenue(f1, price_predicted, min_max_dict[int(Package_Number)]['max'])
                    print(better_price)
                    better_sales = f1(better_price)
                    print(better_sales)

            else:
                print("Value out of interpolation range")
    else:
        print("Too small data point to predict")
```

Figure D.4: Sample Code 4

```
def find_max_revenue(f, start, end):
    max_price = init_price = start
    max_rev = init_rev = start * f(start)

    int_start = int(math.ceil(start))
    int_end = int(math.floor(end))

    i = int_start
    while i <= int_end:
        revenue = i * f(i)

        if revenue > max_rev:
            max_rev = revenue
            max_price = i
        i += 0.5

    return max_price


# Package_Number, DepartureDate, Season,  AvgTemp, SunHour, TotalSnow, UVIndex

predict_better_price('1','2014-3-4','WIN','0','5.4','1','1')
predict_better_price('20', '2018-5-14', 'SPR', '14', '17.1', '0', '3')
```

Figure D.5: Sample Code 5