**GROUP -10**

**TIME SERIES FORECASTS FOR TOURISM FROM UK to THAILAND**

Chaitra Putumbaka(50442616), Sreelekha Komma(50442289), Rohith Chandra Balagiri(50442139)

Department of Engineering Science – Data Science

University at Buffalo, Buffalo, NY 14260

**ABSTRACT**

Travelling has become everyone’s go to in order to get relief from work stress. The number of people visiting Thailand has increased day to day. Thailand has many tourist attractions where number of people visiting these places has increased and the services which are provided by the travel industry has become inadequate. So, in order to keep track of the demand in this project we have done time series analysis on the number of people visiting Thailand from other countries over a period of time from 2010-2016.

**INTRODUCTION**

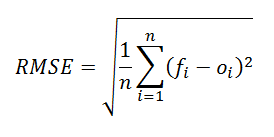
Due to the characteristics of the travel industry and the peculiar nature of human behavior, demand for tourist forecasting in the industry is generally problematic. At the same time, it is crucial for the firm to plan and organize regarding the proper asset categorization. Time series examination is a noticeable procedure for expecting the travel industry interest. It takes advantage of time series information for the point of deciding the ways of behaving, insights, and other significant components of said information. Along these lines, the objective of this part is to tell the best way to perform time series examination for the travel industry interest. Regional execution contrasts remarkably, assortments are outrageous and negative certified and per capita, traveler improvement is entirely expected. To forecast the future, the examination of the seasonality of current and historical data is required. In this project, we’ve implemented time series forecasting models like ARIMA, ETS, Holt-winters to observe the trend and variety.

**Evaluating metrics:**

The RMSE measures the deviation of predicted values with the actual values.

The RMSE root-mean square error is used to calculate the accuracy of the forecast which is considered for finding the best model.

RMSE is calculated by



**Dataset Description:**

The dataset consists of information on number of tourist visitors to Thailand from other countries over a period. The dataset has been taken from Date.world website.

**Preprocessing:**

Removed Null values from the data.

Replaced the missing values with mean values.

Text

Description automatically generated with medium confidence

**EDA:**

**Exploratory Data Analysis:**

Data visualization of all countries over the period is as follows

Chart, bar chart

Description automatically generated

Seasonal Plot on number of tourists visiting Thailand annually is as follows:

Chart, line chart

Description automatically generated

The Decomposition plot of tourists visiting Thailand from UK.

Diagram

Description automatically generated with low confidence

**Methods Implemented:**

**ARIMA Model:**

ARIMA model is used to statistically analyze and forecast the timeseries of data. ARIMA is acronym for AutoRegressive Integrated Moving Average model.

In AutoRegression, a model that takes use of the dependable connection between a perception and a certain number of delayed perceptions.

Chart, histogram

Description automatically generated

Blue curve is the one fitted by model on training data and red curve is the one forecasted by model. And the forecasted curve follows the seasonality of the past years and also follows the increasing trend.

**ETS:**

In Time Series Decomposition ETS model, which stands for Error-Trend-Seasonality.  Error, trend, and seasonality are the three categories used to divide the series. It is a univariate forecasting model When dealing with time-series data. It focuses on seasonal and trend characteristics. Some of the ideas used in this model are the ETS decomposition, the exponential smoothing algorithm, and the trend methods model. It is a model which captures the level, trend and seasonality.

Chart

Description automatically generated

**Holt-Winters Model:**

Holt-Winters forecasting is a technique to model and predict the behavior of a sequence of values across time-to-time series. Holt-Winters is one of the most used forecasting approaches for time series. Three characteristics of the time series can be modeled using the Holt-Winters method: an average value, a trend across time, and a cyclical repeating pattern (seasonality).

Chart, histogram

Description automatically generated

**RESULTS:**

|  |  |  |
| --- | --- | --- |
| MODEL | RMSE | MAE |
| ARIMA | 4589.23 | 3352.966 |
| ETS | 4501.834 | 3341.63 |
| HOLT WINTERS | 4317.513 | 3244.545 |

**CONCLUSION:**

As RMSE and MAE values are low for Holt Winters model compared to ARIMA and ETS, we have decided to use Holt Winters for further forecasting as it performed well on this data. We would also like to do forecasting using a Neural network model like Long Short-term Memory (LSTM) model as it is one of the best models which forecast well. As the data has monthly predictions, if it records daily instead of monthly, then we will be able to generalize it more.

**GitHub Link:**

[**https://github.com/sreelekha2407/Group-10--EAS-509-Time-series-Forecasts-for-tourism-in-the-UK**](https://github.com/sreelekha2407/Group-10--EAS-509-Time-series-Forecasts-for-tourism-in-the-UK)

**REFERENCES:**

**<https://data.world/payapdatasci/thai-tourism>**

[**https://openforecast.org/adam/ARIMAandETS.html**](https://openforecast.org/adam/ARIMAandETS.html)

[**https://www.analyticsvidhya.com/blog/2021/08/holt-winters-method-for-time-series-analysis/**](https://www.analyticsvidhya.com/blog/2021/08/holt-winters-method-for-time-series-analysis/)