International Journal of Mathematics, Statistics, and Computer Science
Vol. 2, 2024
ISSN: 2704-1077 eISSN 2704-1069, DOI: 10.59543/ijmscs.v2i.7851

# A COMPARATIVE ANALYSIS OF MACHINE LEARNING ALGORITHMS FOR THE PURPOSE OF PREDICTING NORWEGIAN AIR PASSENGER TRAFFIC

**Aleksander Stanulov[1], Sahar Yassine[1,*]**
[1]Department of Applied Data Science, Noroff University College, Norway; sahar.yassine@noroff.no

*Received 07.06.2023, Revised 02.07.2023, Accepted 05.07.2023, Published 08.07.2023*

**ABSTRACT** This analysis aims to provide an overview of potential machine learning algorithms that may aid the aviation industry in predicting future air passenger traffic flow, which can help increase stakeholder value as well as improve customer experiences. A review and discussion of the aviation industry's past, current, and future challenges is provided, as well as an overview of machine learning algorithms, neural networks, and learning methods. Further, an overview and discussion of the architecture of the Long Short-Term Memory (LSTM) network, Support Vector Regression Machine (SVRM), and Random Forest (RF) algorithms is provided. The comparative analysis provides an overview and comparison of the performance of the LSTM, SVRM, and RF models based on Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). The dataset used includes the hourly number of passengers from scheduled flights at Oslo Airport Gardermoen for the period of January 1, 2009, to December 31, 2019, including the datetime features such as Time (hour), day, month, and year, as well as the weather features of air temperature and mean wind speed, with a total of 96185 samples. The Long Short-Term Memory model exhibited the highest generalization ability, with a performance evaluation on the testing dataset of 0.00445/0.06667 MSE/RMSE. Additionally, the performance of the SVRM and RF models on the testing dataset is 0.00511/0.07147 and 0.00543/0.07368 MSE/RMSE respectively. In addition to the performance, each of the models' complexity, stability, and ability to predict the hourly and daily fluctuations of passengers are discussed.

***Keywords:*** Machine Learning, LSTM, SVRM, RF, Aviation.

## 1. INTRODUCTION

The training and application of supervised Machine Learning (ML) algorithms, for the purpose of predicting future air passenger traffic in Norway, can greatly benefit the aviation industry to make better decisions based on empirical data. The aviation industry has experienced significant growth in the amount of air traffic during the past two decades, with two of the main reasons for said growth being a country's increase in Gross Domestic Product (GDP) and population growth [1], [2]. Numerous examples where the growth of the aviation industry can be observed include a steady increase in air passenger traffic, the expansion or building of new airports, as well as airlines expanding their fleet of aircraft by purchasing additional ones. This unparalleled historical growth has in recent years been challenged and is presented with numerous challenges that it must solve to continue growing in the future.
Events such as the COVID-19 pandemic sent the industry to a standstill following a significant reduction in demand during the height of the pandemic [3]. One of the main drivers behind the reduction in demand came because of the implementation of air travel restrictions by the Norwegian and other governments.

Although a steady increase in population worldwide and in Norway has historically been observed [4], which has benefitted the growth of the aviation industry, an emerging long-term challenge to the industry is the stagnation of birth rates in Norway [5] and abroad [6]. To meet these short and long-term challenges, as well as improve customer experiences, the aviation industry's ability to predict future air passenger traffic flow needs to be enhanced.

There is a need to enhance the prediction of future air passenger traffic flow within the aviation industry in order to increase stakeholder value, as well as to improve customer experiences. Incorrect or insufficient forecasting of air traffic can lead to significant losses for the aviation industry, which can be observed if an airliner purchases a large amount of aircraft to meet a demand which may later change, which may result in a portion of the fleet not being used, not generating revenue. Other large investments such as building long-term aviation infrastructure such as airports would also greatly benefit from air traffic forecasting. The accurate forecasting of air traffic demand, both short-term and long-term, would greatly benefit the industry.

The aim of this research effort is to assist in that by providing a comparative analysis of a Long Short-Term Memory (LSTM), Support Vector Regression Machine (SVRM), and Random Forest model (RF) model. The models will be trained and evaluated on a dataset including the Passenger amounts at the Oslo Airport Terminal Gardemoen, for each hour of the day, ranging from Jan 1, 2009, to Dec 31, 2019. In addition to the datetime features of hour, day, month, and year, the dataset also includes the mean wind speed and air temperature for each respective hour. The passenger data was collected with the help of [7], and the weather data was collected from [8], and the yearly CSVs of both were cleaned and merged at each corresponding hour. The comparative analysis will compare the performance of the models based on Mean Squared Error (MSE) and Root Mean Squared Error (RMSE).

To serve this objective, this paper is structured as the following. The literature review section provides an overview of the aviation industry, its historic growth, current and future challenges, as well as an overview of the LSTM, SVRM, and RF algorithms. Furthermore, it also discusses previous implementations of ML and NNs within the aviation industry and for forecasting. After that, the Methodology section provides an overview of the data collection and processing, network, and algorithm modeling, and hyperparameter tuning for each specific model. Then, the data collection & processing section goes into detail about how the data was collected and processed for the creation of the dataset. The Comparative Analysis section discusses the difference in performance between the three models that were trained and evaluated, focusing on model performance, stability, and complexity differences, as well as comparing predicted vs. actual testing values. Finally, we demonstrate the Conclusion and the future work which summarize the research contribution and discuss some of the different avenues possible for future work.

## 2. LITERATURE REVIEW

The literature review aims to provide an overview of the historical growth and future challenges within the aviation industry, as well as discuss the architectures of the LSTM, SVRM, and RF algorithms. Furthermore, it will cover existing research within the field of forecasting, as well as previous research that has used ML and NNs within the aviation industry and for the purpose of forecasting.

### 2.1. Historical Growth and Future Challenges Within the Aviation Industry

Historically the aviation industry has had unprecedented growth (Figure 1), which can be attributed to various factors including Gross Domestic Product (GDP) growth and world population increase. As observed by the World Bank [9], which shows a steady growth within the industry, a significant increase is observed between the years 2009 and 2019. In 2009, there were 2.25 billion passengers carried, and in 2019 peaks at 4.56 billion, which is a 102.6% increase over a relatively short period of 10 years. The remarkable growth within the industry, for the pre-COVID 19 years is also reviewed in the International Air Transport Association's (ICAO) annual 2019 report [10], which observes an average of 4.9% increase in scheduled passenger traffic measured in revenue passenger-kilometres (RPK), in the year 2019.

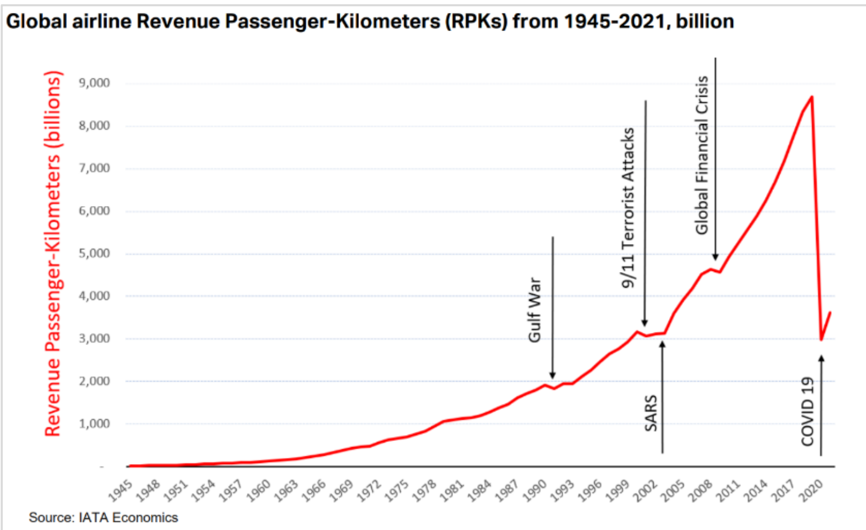

Figure 1. Global airline Revenue Passenger-Kilometers (RPKs) from 1945-2021, billion. Source: [11]

To a certain extent, it is possible to correlate the historical growth of the aviation industry to varying factors such as GDP increase, social factors such as increased tourism, as well as economic evolution in developing regions, which all serve as an indication for market demand increase [11]. Airport infrastructure, such as the number of runways available at an airport and the size of waiting rooms, to name a few, is greatly influenced by the demand for air traffic [12]. While impressive growth has been observed for decades, the number of passengers carried for the year 2020 dropped to 1.81 billion. This drop in passengers carried is a consequence of the air travel restrictions implemented by various governments because of the COVID-19 pandemic (Figure 2).

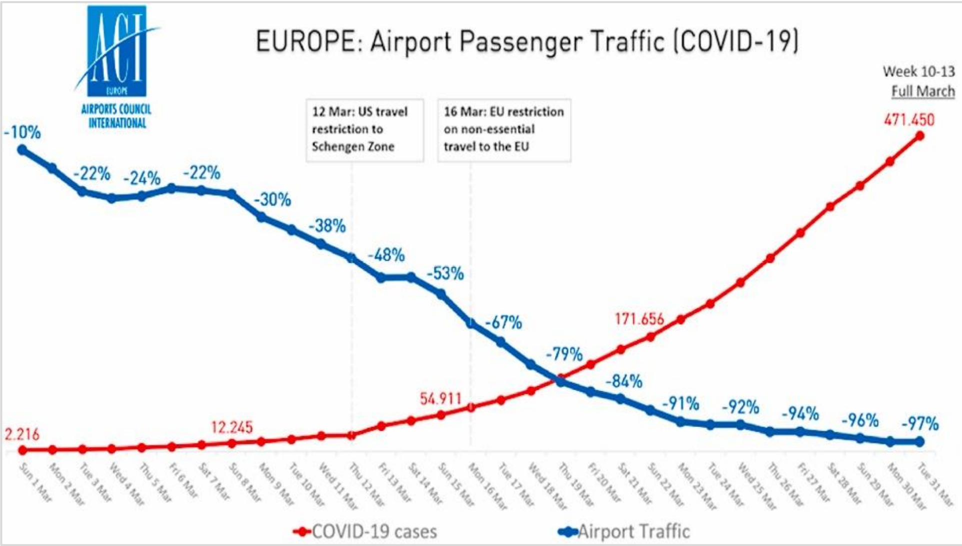


Figure 2. Air passenger traffic & COVID-19 cases for March of 2020 in Europe. Source: ACI Europe [14]

The major reduction in airport traffic correlates to major revenue loss for airlines and airports alike, with the extent of the loss being dependent on, among others, the effectiveness of economic stimulus [13]. More recent prognosis shows that although the industry was ill-equipped to deal with the major consequences of COVID-19, the industry has been resilient in its recovery. This resilience is outlined by the Airbus Global Market Forecast of 2022 by [14], which shows that the aviation market is expected to recover to 2019 levels already between 2023 and 2025. A thorough investigation into Artificial Intelligence's ability to assist the aviation industry, specifically in times of crisis such as COVID-19 has been done by [15]. The researchers discuss the opportunity for the industry to leverage AI tools to enhance its business model.

Economic and social factors such as gross domestic product (GDP) and population growth are, among others, contributing factors to the historical growth of air traffic [2], [16]. The impact of these factors will vary from region to region. Regions that have a larger GDP and/or population growth will on average also experience an increase in domestic and international air traffic. Considering correlations between population and air traffic growth have been observed, it is also one of the challenges that exist for the long-term economic stability of the aviation industry. The Norway 2022 National Population Projections [5] report estimates a net positive population growth in Norway, from 2022's population of 5.4 million to 6.1 million in 2060 and 6.2 million in 2100. The report also describes an increase in aging, with over 25% of the population being aged 70 or over by 2100, with there being more older people than children and teenagers already by 2031.

The United Nations report [6] provides a comprehensive overview of the population projections for different regions in the world. The UN report observes a noticeable decrease in population growth from 1950 and onwards, with the rate of population growth falling below 1% for the first time in 2020. The report elaborates that reduced levels of fertility is the main reason behind the steady decrease in growth.

### 2.2. Machine Learning Algorithms

There are many different machine learning algorithms, each with its benefits and drawbacks. These algorithms are often split into separate categories, where one of the main deciding factors when choosing an algorithm is the type of data that will be used during training. Within the context of this paper, which focuses on the supervised learning algorithms, LSTM, SVRM, and RF. Supervised Learning algorithms are algorithms that use labeled training data to improve their ability to make correct predictions or infer a function based on the labeled data [17]–[19]. The following three chapters will provide a short theoretical overview of the SVRM, LSTM, and RF algorithms, as those are the algorithms that are trained and evaluated on the passenger data within the context of this paper.

### 2.3. Support Vector Machines & Support Vector Regression Machines

Support Vector Machine (SVM) is a learning machine first introduced by Cortes and Vapnik [20] for the purpose of two-class classification problems with high generalization capability. For a two-class linear classification task, SVM is a supervised learning model that attempts to find the best fit for a hyperplane in an n-dimensional space to accurately classify the data points [21], [22].

SVMs can also be used in non-classification applications such as regression tasks, in which case they are referred to as Support Vector Regression Machine (SVRM), or Support Vector Regression (SVR). Introduced by [23], and similar to an SVM, SVRM is a supervised learning algorithm, which for a regression task, outputs a continuous value. SVRMs attempt to define a hyperplane that would fit the largest possible amount of data points [21]. SVRM has also been widely used for time series forecasting, showing that the algorithm can be adapted to a wide range of time series tasks, in part due to its flexibility of using different kernels which allows it to understand non-linear relationships by mapping the input data to higher dimensions [24], [25].

### 2.4. Recurrent and Long Short-Term Memory Neural Networks

A Recurrent Neural Network (RNN) can recall past events during training through information loops within the hidden layer, which allows the network to recall previous information that may aid it in learning about the current sequence. An example use case of an RNN can often be found within natural language processing, where it may be useful for the network to know what word came before the word or sequence it is attempting to predict [26]. While there are different types of RNN, a standard RNN refers to, in part, the network having one or more hidden layers which contain recurrent neurons. Recurrent neurons, when compared to the more basic ANN neurons such as the Perceptron, may receive or output information from and to other neurons, which may reside in the current hidden layer or in other hidden layers from within the same network, all of which depends on the specific RNN setup [26], [27].

The LSTM architecture is an extension of RNNs and addresses the issue of decaying and vanishing error backflow by adding additional units to the network such as memory cells and gate units. These memory cells and gate units, like the recurrent neurons, reside within the hidden layer of the network, which may also include other types of units depending on the type of network [28]. The LSTM architecture includes numerous activation layers each with its own activation function. The activation layers, as well as the corresponding activation functions, to a large extent, are there to control the flow of information to the cell state. The cell state is effectively the long-term part of the LSTM model, which includes the information being retained within the memory cell. That information can be, in addition to the information added through the gated units, passed on to the next cell [26], [27].

### 2.5. Decision Trees and Random Forests

A Random Forest, first introduced by Breiman [29] is an ensemble machine learning algorithm that consists of numerous decision trees that train on a random subset of training data and features. An ensemble method refers to the combination of the predictions of numerous algorithms which, together, can provide a more accurate final prediction. Bootstrap Aggregation (Bagging) is an example of an ensemble method, also introduced by Breiman [30], which helps by reducing variance within the algorithm by voting, for classification, or averaging, for regression predictors. A decision tree algorithm can be used both for classification or regression problems and is a supervised learning method that uses a tree-like structure.

### 2.6. Applications of Machine Learning & Forecasting Algorithms

Ojo and Ogunnusi [31] propose a Support Vector Regression Machine (SVRM) based model for the purpose of predicting air passenger traffic for the Muritala Mohammed International Airport in Nigeria. The basis for the study includes 132 months of data for the period of 2007-2018 and uses Root Mean Squared Error (RMSE) as the basis for model evaluation. The research is in part focused on finding out which SVRM Kernels are more effective given the particular task and data.

An application of Artificial Neural Network (ANN) was used by Putra and Safrilah [32] for the prediction of the capacity of a runway at Juanda International Airport, Indonesia from January 2016 to March 2016. The researchers found that the ANN was generally able to imitate the fluctuations of runway capacity and suggest that similar models can be used by airports to improve the management of runway capacity. A Neural Network (NN) developed by Blinova [33] is a multi-layered “focused time-lagged feedforward network”, which the researcher describes as a network with a short-term memory concentrated in the input layer, using the standard backpropagation algorithm for learning. The researchers report the forecasting error for the NNs during the adaption stage to be below 5%. The data that was used was between the intra-regional and inter-regional passenger traffic in Russia between the years 2006 and 2010, with the models having the ability to adequately predict the air passenger traffic demand for the future 2-3 years.

An autoregressive neural network called DeepAR developed by Salinas et al. [34] has the potential to deal with a wide range of probabilistic time-series based forecasting problems. The model is based on the recurrent neural network architecture and can produce accurate probabilistic forecasts. The model can be applied to a wide range of tasks, including traffic prediction.

### 2.7. Time Series Forecasting Algorithms

Statistical forecasting algorithms, such as the Auto Regressive Integrated Moving Average (ARIMA) and Seasonal Auto-Regressive Integrated Moving Average (SARIMA), first introduced in 1976 by Box and Jenkins [35], can be used for the purpose of making time series forecasting. Some of these algorithms have also been used for the purpose of making predictions within the aviation industry [36]. ARIMA models consist of 3 main parameters, the autoregressive, the integrated, which is the differencing factor, and the moving average parameter, which leverages previous forecasting errors instead of values to forecast future values [36], [37]. The SARIMA algorithm expands upon ARIMA by adding seasonality to the forecast. SARIMA expands upon ARIMA, keeping the previous 3 parameters of ARIMA, autoregressive, integrated, and moving average, expanding the algorithm by adding the seasonal autoregressive, seasonal integrated, and seasonal moving average parameters [37].

An application of a multivariate ARIMA model for the purpose of predicting air transport demand has been done by [12]. A SARIMA-SVR approach has been proposed by [38] for the purpose of aviation planning and capacity management.

## 3. METHODOLOGY

In this section, we will explore and discuss the methodologies, techniques, and tools used in each step of the data collection, and processing as well as hyperparameter tuning and model evaluation. Figure 3. Stages of the studydemonstrates the stages and the steps of our research efforts.

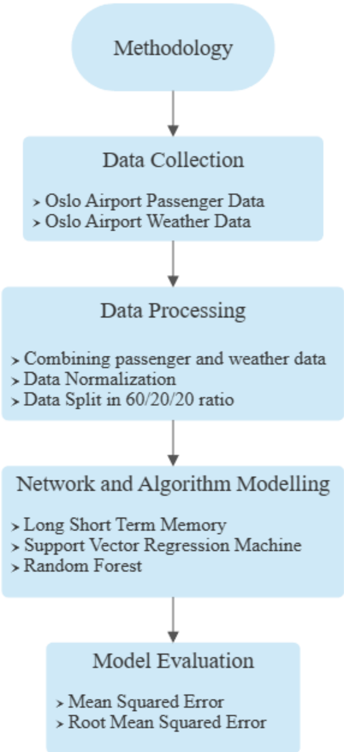


Figure 3. Stages of the study

### 3.1. Data Collection & Processing

Two components make up the dataset used in the research, both being hourly. The first is the passenger data and the second is the weather data. The passenger data describes the number of passengers present at the terminal of Oslo Gardermoen Airport, Norway, at each hour of the day. The data range between 01.01.2009 - 31.12.2019 and is provided by Avinor [7]. The passenger data includes the date, time, and amount of passengers present at the terminal at each hour. The amount of passengers at the terminal is defined by passengers that either depart or arrive on an airplane at the airport, this includes Terminating and Transferring passengers.

Terminating passengers are defined as passengers that either begin or end their trip at the airport, and transfer passengers are defined as passengers that arrive and depart the airport on different aircraft or the same aircraft bearing a different flight number, and are counted twice. In addition, the passenger data only includes scheduled passenger flights. The second part of the dataset is the weather data, which consists of the Air temperature and Mean wind speed at the Oslo Airport Weather Station Gardermoen SN4780, provided by the Norwegian Centre for Climate Services [8].

Data processing is necessary to clean and prepare the data before the model training. This includes data cleaning, removal of null data points, normalization in the range of 0-1, splitting it into training, testing, and validation datasets as well as input data reshaping. The data split is done in a 60/20/20 ratio where 60% is used for training, 20% is used for validation and the final 20% is used for testing. The allocation of 40% of the data for validation and testing datasets was deemed necessary in order to reduce the risk of the model overfitting as well as improve the process of hyperparameter tuning.

### 3.2. Network and Algorithm Modelling

We are going to apply three machine learning models which are: Long Short-Term Memory Model (LSTM), Support Vector Regression Machine Model (SVRM), and Random Forest Model (RF). The three models will be evaluated based on Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). For each of these values, the lower the value the better the performance. Mean Squared Error measures the average squared difference between the predicted and the actual value, in this case, it's the value of Passengers. The closer the value for MSE is to 0, the fewer errors that the model makes when making predictions, and their squared difference against the actual values is computed. An additional observation that is used in the model evaluation and the comparative analysis is the analysis of predicted versus actual values. Observing the predicted vs actual values is vital for the scope of this paper as it allows the research to observe if the model can emulate the seasonality and cyclic fluctuations of Passengers in an hourly format.

The reason MSE and RMSE are valuable when discussing the evaluation of the models, within the scope of this paper, is the models are being trained to emulate the seasonality and fluctuations observed in the Passenger's feature in the real data. Since the data exhibits seasonality distributed fluctuations that are repeated sequentially, 24 hours at a time, metrics that penalize larger errors, like MSE and RMSE, are more appropriate. Furthermore, to the extent of this paper, smaller errors are not as impactful as larger ones because even though the predicted values might not be identical to the actual ones, the important factor for the models' predictions is that they emulate the daily fluctuations or seasonality that the actual data exhibits. On that basis, the models will be evaluated on their performance in terms of the MSE and RMSE values, as well as discussing their ability to emulate the cyclical fluctuations of Passengers by observing the predicted values.

The network and algorithm modeling covers the input data reshaping, model hyperparameter tuning, and network architecture. The input shape of the x component of the data for the LSTM model is adjusted to a 3D shape of (samples, steps, features), and for the SVRM and RF models to a 2D shape of (samples, steps $*$ features). For the hyperparameter tuning, each model starts with a combination of baseline parameter values, which are then tuned and tested, with the ones that yield the lowest values of MSE and RMSE, without showing signs of overfitting, being preferable. For all three of the models, a steps value of 24 is used, which is chosen as it represents one full day and a features value of 7 is used, as that is the total number of features in the finished dataset.

## 4. DATA COLLECTION & PROCESSING

The completed dataset that is used for model training contains hourly data ranging from January 1st, 2009, to 31 December 2019, with a total of 96185 data points. The dataset contains the following 7 features and their value ranges:

Time: Ranges from 0-23 and describes the hour of the day that the data point corresponds to.

Day: Ranges from 0-31, Month: Ranges from 0-12, Year: Ranges from 2009 to 2019.

Passengers: Number of passengers present at the Oslo Airport Terminal at that specific hour.

Air temperature: The air temperature at that specific hour.

Mean wind speed: The mean wind speed is the mean value of the wind speed of the 10 minutes prior to collection.

Different data Processing was done to refine and enhance the datasets and prepare for the analysis phase:

### 4.1. Passenger Data

The raw passenger data included the yearly CSVs that included the features: Date, Passengers, Arrival or Departure, Domestic or International, Hour, Airport, and IATA code, in an hourly format. The data was processed by taking only the data for Oslo Airport Gardemoen, as well as summing the amounts of passengers at each corresponding hour, regardless of whether they were domestic, international, arriving, or departing. Each yearly processed passenger csv contained 3 features, the date, hour, and passenger amount. For most of the years processed, this resulted in 8760 rows of data for 1 year, which is given by multiplying 24 hours by 365 days in a year. Some of the years had missing data, but this was very rare and, in most cases, only accounted for 2-4 days of missing hourly data from an entire year.

### 4.2. Weather Data

The raw weather data included yearly CSVs containing 3 features, Datetime, Air temperature, and Mean wind speed. Considering the passenger data was recorded at each round hour, and the weather data was in a minute format, meaning it had multiple entries for each hour, it had to be processed accordingly to accurately match it with the passenger data. This was solved by extracting the datetime components of the Datetime feature and selecting the data point with the closest datetime component to the closest round hour. For example, if there were two data points close to the round hour of 14:00, one being 14:10, and one being 13:40, the data point for 14:10 would be chosen as it is the closest one to 14:00.

The other data point, in this example for hour 13:40, would be discarded, unless it is also the closest data point for the previous hour of 13:00, in which case it would then be assigned as the 13:00 data point. Performing this processing, as with the passenger yearly data would yield 8760 rows, having one row for each hour in the year. Furthermore, when there were missing data in the passenger data, the rows corresponding to those dates were removed from that year’s weather data, and vice versa, ensuring both files matched chronologically.

The reason for using this specific weather data is that weather can have a significant impact on an airport’s operations, this can include aircraft delays, air passenger traffic, and runway availability. While numerous meteorological variables may impact the operation of airports, the air temperature, as well as the wind speed, have been shown to impact the operations of an airport, for example, causing delays or cancellations during extreme weather conditions [39]–[41].

### 4.3. Merged, Chronological Data

With the specific processing of both the passenger and weather data, the files are then merged chronologically in order to create a complete dataset starting Jan 1, 2009, and ending on 31 Dec 2019. The first 10 rows of the complete dataset are examined by the following Table 1

Table 1. The first 10 rows in the completed dataset

| Time | Day | Month | Year | Passengers | Air Temperature | Mean Wind Speed |
|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 2009 | 0 | -10 | 0.0 |
| 2 | 1 | 1 | 2009 | 0 | -10 | 0.0 |
| 3 | 1 | 1 | 2009 | 0 | -13 | 0.0 |
| 4 | 1 | 1 | 2009 | 0 | -12 | 0.0 |
| 5 | 1 | 1 | 2009 | 0 | -12 | 0.0 |
| 6 | 1 | 1 | 2009 | 187 | -13 | 0.0 |
| 7 | 1 | 1 | 2009 | 122 | -14 | 0.0 |
| 8 | 1 | 1 | 2009 | 83 | -12 | 0.0 |
| 9 | 1 | 1 | 2009 | 215 | -13 | 0.0 |
| 10 | 1 | 1 | 2009 | 524 | -11 | 0.0 |

Table 1 shows the first 10 rows in the completed dataset, where "Time" refers to the hour and ranges between 0-23, with the corresponding values of Passengers, Air temperature, and Mean wind speed corresponding to that specific hour, day, month, and year.

The data is then first normalized in the range of 0-1, which is necessary because each of the data features operate on a different range. This enables the model to better understand the training data as each numeric value then operates in the same range.

Once normalized the data is split into 3 datasets, in a 60%/20%/20% ratio for training, validation, and testing data respectively. This gives 57711 data points for training, 19237 for validation, and 19237 for testing.

The last step involving the data prior to model training is reshaping. Reshaping the data for the LSTM model as an example, requires the data to be in a sequential format. This length of the sequence is the steps variable, which for the particular LSTM model is 24, meaning each sequence is 24 data points long. The reshaping for each model will be elaborated upon in each model in the next section.

## 5. MODELING & HYPERPARAMETER TUNING

In this section, we will demonstrate the modeling, hyperparameter tuning, and implementation of each of the three models, LSTM, SVRM, and RF. In each model, we will discuss the data input shape and the hyperparameter tuning. The hyperparameters will be tuned by experimenting with different combinations in order to determine which combination of hyperparameters yields the best performance, where the model's generalization ability is prioritized. Moreover, a large gap between training, validation, and testing performance where lower values for the training dataset may be an indication of the model overfitting the data, with higher values indicating underfitting. With that in mind, while lower values for MSE / RMSE are preferable, the hyperparameter tuning strives to find a balance between overfitting and underfitting, so that the final model is stable and has a good generalization ability.

### 5.1. LSTM Network Modeling

The LSTM model consists of 1 hidden LSTM layer with 128 neurons, using a tanh activation function, 1 dropout layer, dropping 20% of the neurons after the LSTM layer, and a dense, single neuron, output layer using a linear activation function. The model also uses the Adam optimizer algorithm with a learning rate of 0.001. The input data shape is defined as (samples, steps, features) where steps correspond to the sequence length of 24 and features the number of features in the dataset, which is 7.

#### 5.1.1. Input Shape

After data normalization, the data is run through a sequence function that creates the LSTM sequences, outputting the reshaped numpy arrays for the x and y components of the data, for the training, validation, and testing datasets respectively.

The x-train component of the training dataset pertains to the input sequences for the LSTM model, which have a shape of (57687, 24, 7), corresponding to (samples, steps, features).

The y-train component of the training dataset is the target parameter and corresponds to the Passengers column. The shape of this component is (57687, ) which corresponds to (samples, ), which has the same length as the x-train component because each input sequence has a corresponding target Passengers value.

The same logic is used for the validation and testing datasets. The x val component of the validation dataset has an input shape of (19213, 24, 7), corresponding to the same variables as already described in the training dataset. The y val component of the validation dataset has an input shape of (19213, ).

Similar to the validation input data, the x test and y test components of the testing dataset have the same input shape as the aforementioned validation data shapes, which is the case because both represent 20% of the total data respectively. With the data normalized, split, reshaped, and sequential, it is ready to be fit in the LSTM model.

#### 5.1.2. Hyperparameters & Network Architecture

The LSTM model has a large range of hyperparameters that can be tuned according to the data and the specific problem with the aim of improving the model's performance. For this particular data and problem, the following hyperparameters (Table 2) have been declared and tuned, with the final model using the following parameter values:

Table 2. LSTM Model Hypermeters

| Steps | Features | Batch Size | NO. of LSTM Layers | NO. of Neurons | Activation Function | Optimizer | Optimizer Learning Rate |
|---|---|---|---|---|---|---|---|
| 24 | 7 | 128 | 1 | 128 | tanh | Adam | 0.001 |

Specifying and using a set seed where possible was important as it allowed for proper hyperparameter tuning by ensuring consistency between training iterations, to precisely observe which parameters contribute to increasing model performance. For that reason, a seed number of 42 is used, this includes Numpy [42], TensorFlow [43], as well as network weight initialization and dropout layers, which are used in the LSTM model. This seed number doesn't represent anything specific and just ensures that the same values and neurons would be outputted or picked each time the code was executed.

The parameter tuning was done in the following order: activation function, optimizer algorithm, optimizer learning rate, batch size, layer amount, and neuron amount. Once a particular parameter was determined as the best performing, which for example between tanh and reLU was tanh, the tuning continued with tanh as its activation function.

The hyperparameter tuning was done by tuning a single parameter at a time, beginning with a single-layered LSTM architecture with 128 neurons, followed by a Dropout layer with a rate of 0.2 and a dense output single neuron layer using a linear activation function, which remained as the output layer also in the final model.

### 5.2. SVRM Algorithm Modeling

The trained and tuned SVRM model that was trained on the training dataset, and evaluated on the training, validation, and testing dataset uses the following parameters (Table 3):

Table 3. SVRM Model Hypermeters

| Steps | Features | Kernel | Degree | Gamma | Coef0 Term | Epsilon | Regularization parameter C |
|---|---|---|---|---|---|---|---|
| 24 | 7 | Polynomial | 2 | Scale | 1 | 0.01 | 1 |

#### 5.2.1. Input Shape

The SVRM model uses the same process as the other models in regards to data normalization in the range of 0-1, after which the data is run through a sequence function that outputs the numpy arrays of the x and y components of the data, for the training, validation, and testing datasets respectively.

As with the LSTM model input shape, the sequence function creates the x and y components of each dataset using the data and the number of steps, which is 24, after which it is reshaped into a 2D format in order to be used with the SVM model.

The reshaping of the data results in the format of (samples, n steps $*$ n features). The data has to be flattened to 2D in order for it to be fit into the SVRM model. In contrast to simply fitting the data in the format of (samples, n features), multiplying the number of steps by the features may allow the model to also train on sequential patterns that may exist within the data.

After the reshaping, and right before being fit to the model, examining specifically x train, its shape output is (57687, 168), the first number of which is the number of samples, which corresponds to 60% of the total data, and the second number being the product of n steps $*$ n features. The y component of the training dataset corresponds to (57687, ).

#### 5.2.2. Hyperparameter Tuning

The first evaluation examined each kernel, including rbf, polynomial 2nd degree, polynomial 3rd degree, and linear, each of these kernels were paired with different combinations of gamma and/or coef0 term values, where applicable. The best performing kernel combinations are then compared against each other, with the best performing configuration being the polynomial 2nd degree kernel using 'scale' for gamma, and a coef0 term of 1. The final two parameters that are tuned after that are the regularization parameter, C, and the value for the epsilon. 7 different values of C were tried, including 0.1,0.5,1,2,3,4, and 5, with the overall best performing one not changing from C=1, as it provided the best performance on the testing dataset. The values that were tested for epsilon include 0.01,0.05,0.1,0.2, and 0.3, with 0.01 being the one that showed the best performance for all 3 dataset evaluations. Attempting lower values for epsilon was not computationally feasible, as lower values of epsilon exponentially increased training time, without significantly improving performance.

### 5.3. Random Forest Algorithm Modeling

Table 4 shows the final trained and tuned Random Forest model with the following parameter values:

Table 4. Random Forest Model Hypermeters

| Steps | Features | Estimators | Max Depth | Min. Samples Split | Min. Samples Leaf | Max. Features |
|---|---|---|---|---|---|---|
| 24 | 7 | 100 | 10 | 2 | 2 | 0.75 |

### 5.3.1. Input Shape

The input shape processing for the Random Forest (RF) model is identical to the SVRM input shape discussed in section 5.2.1.

### 5.3.2. Hyperparameter Tuning

The RF model tuning starts with the max depth parameter, for which a value of 10 produced the best testing performance without a large risk of overfitting the training data, which the model did when a value of 'None' was used. The following parameter is the estimators parameter, which refers to the number of decision trees within the RF model, with negligible differences between the tested values, a value of 100 is used. Different combinations of min samples split and min samples leaf were tested, with again negligible differences, with the final model using 2 for each respectively as that configuration yields good overall performance without exaggerating model complexity. The final parameter that was tuned is the max features, where a value of 0.75 means 75% of the features were considered when the trees were looking to split the nodes.

## 6. COMPARATIVE ANALYSIS

The comparative analysis chapter will analyze and compare the results of the LSTM, SVRM, and RF models that were trained on the same training dataset, and evaluated on the same validation and testing datasets, with the performance evaluation based on the lowest MSE / RMSE values. In addition to the actual performance, a comparison of each of the models' predicted values for the first 48 hours of testing Passenger data will be compared against each other, based on the actual first 48 hours of Passenger data. This comparison will not only observe how large the differences between the actual and predicted values are for each model respectively but also discuss each model's ability to generalize and follow the daily fluctuations which are exhibited in the real data.

### 6.1. Performance Analysis

The performance analysis will discuss the results of each model and how each model compares to the others, both to the extent of their generalization ability and the risk of overfitting. While each model has had extensive hyperparameter tuning, it is important to note that, while the research has done its best to tune each model to make sure the comparison was fair, it is possible that the performance shown here is not the absolute best possible that each of these algorithms can attain. Further performance improvements may be possible through various methods such as different hyperparameter combinations or alternative feature engineering.

Table 5 Performance Comparison Between LSTM, SVRM, and RF models

| Model | Evaluation |
|---|---|
| Training Performance MSE/RMSE | |
| LSTM | 0.00268 / 0.05178 |
| SVRM | 0.00264 / 0.05135 |
| RF | **0.00209 / 0.04569** |
| Validation Performance MSE/RMSE | |
| LSTM | **0.00377 / 0.06141** |
| SVRM | 0.00395 / 0.06283 |
| RF | 0.00409 / 0.06396 |
| Testing Performance MSE/RMSE | |
| LSTM | **0.00445 / 0.06667** |
| SVRM | 0.00511 / 0.07147 |
| RF | 0.00543 / 0.07368 |

Table 5 shows the performance comparison between the LSTM, SVRM, and RF based on MSE / RMSE, with the lowest values in bold. These observations show that, for this dataset, the LSTM model is showing the best performance for both the validation and testing datasets, with 0.00377/0.06141 and 0.00445/0.06667 MSE / RMSE respectively. For the training dataset, the RF model has the lowest scores of 0.00209/0.04569 MSE / RMSE. The LSTM and SVRM models have a negligible difference in MSE / RMSE values for the training dataset, with only a value of 0.00004/0.00043 MSE / RMSE difference between the two, in favor of the SVRM model.

**Risk of overfitting**

The LSTM model has the lowest indication of overfitting the data between the three models, as it has the smallest gap between the training and testing performance, with a difference of 0.00177/0.01489 MSE / RMSE. The RF model has the largest indication of overfitting on the data, considering it has the lowest MSE / RMSE values for the training dataset, at the same time as it has the highest MSE / RMSE values for the validation and testing datasets.

**Generalization Capability**

As briefly mentioned, the LSTM model has the best performance, with the lowest values of MSE / RMSE for both the validation and testing datasets, while still maintaining comparable training performance. There are numerous reasons why the LSTM model ends up having the best performance, with one of the bigger reasons, compared to the other two models, is the LSTM model has the ability to learn from 3-Dimensional input data. This is an advantage that, to the extent of this paper, dataset, and task, a neural network such as LSTM has over the other two algorithms. This means that the input data for the LSTM model can be in the shape of *(samples, steps, features)*, while for the other two, the input shape had to be flattened to 2-Dimensions, in the shape of *(samples, steps* $*$ *features)*. This advantage allows the LSTM to be able to train on and learn time-dependencies that exist within the data which is, for this time-series dataset, one of the driving factors that make the LSTM model the best performing one on the basis of MSE / RMSE.

### 6.2. Actual vs. Predicted Values Analysis

This section will look at how a sample of 48 hours of actual data, corresponding to the first 48 hours of the testing dataset compares to the predicted Passenger values of each of the models for the same data points. The purpose of this comparison is threefold, for one it is to see a small example of how the models are predicting the number of Passengers based on unseen testing data, the second is to discuss and observe how far these predictions are from the actual passenger amounts, and the third is to see the model's ability to follow the time fluctuations exhibited in the actual data. For this paper, the third of the aforementioned purposes of the comparison is the most impactful one, because while 48 hours of data is a very small example in terms of the amount of data points, it is possible to observe if each of the models are capable, and to what degree to follow the daily fluctuations of Passengers based on the hour of the day.

Table 6 Actual vs. Predicted values on 48 hours of testing data, all three models

| Index | Actual | LSTM | SVRM | RF | Index | Actual | LSTM | SVRM | RF |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1541 | 1310 | 1701 | 1152 | 24 | 5304 | 4100 | 4544 | 4566 |
| 1 | 2097 | 1743 | 1859 | 1928 | 25 | 7067 | 7774 | 6701 | 6582 |
| 2 | 2398 | 2273 | 1801 | 2324 | 26 | 5458 | 5513 | 5203 | 5059 |
| 3 | 2540 | 2889 | 2749 | 2580 | 27 | 4727 | 4664 | 4909 | 4557 |
| 4 | 4294 | 3405 | 3357 | 2973 | 28 | 4122 | 4252 | 4059 | 3722 |
| 5 | 5761 | 4508 | 4147 | 3738 | 29 | 4826 | 3696 | 3603 | 3428 |
| 6 | 5273 | 5883 | 6533 | 4776 | 30 | 4337 | 4456 | 4678 | 4744 |
| 7 | 6506 | 6495 | 5609 | 4822 | 31 | 5399 | 4761 | 4575 | 4417 |
| 8 | 6654 | 6960 | 6395 | 5697 | 32 | 5547 | 5495 | 5663 | 5163 |
| 9 | 6876 | 7280 | 7080 | 6510 | 33 | 5517 | 6436 | 6622 | 6550 |
| 10 | 7036 | 7388 | 7590 | 6690 | 34 | 5506 | 6681 | 6595 | 6522 |
| 11 | 7919 | 7206 | 6912 | 6337 | 35 | 5210 | 5596 | 5259 | 4834 |
| 12 | 7723 | 6866 | 6582 | 6432 | 36 | 5218 | 4710 | 4740 | 4647 |
| 13 | 6387 | 6264 | 5708 | 5592 | 37 | 3667 | 3585 | 3402 | 3704 |
| 14 | 5894 | 5286 | 5746 | 5069 | 38 | 3236 | 2568 | 2759 | 2915 |
| 15 | 4796 | 4393 | 4433 | 4248 | 39 | 1881 | 1906 | 2040 | 2172 |
| 16 | 2184 | 2331 | 2071 | 2144 | 40 | 943 | 1181 | 1398 | 1438 |
| 17 | 499 | 513 | 297 | 639 | 41 | 583 | 312 | 313 | 335 |
| 18 | 181 | 115 | -147 | 58 | 42 | 0 | 86 | -43 | 66 |
| 19 | 0 | 70 | -63 | 58 | 43 | 0 | 28 | -121 | 23 |
| 20 | 0 | 79 | -110 | 17 | 44 | 0 | 3 | -121 | 4 |
| 21 | 0 | 13 | 8 | 18 | 45 | 0 | -110 | -87 | 3 |
| 22 | 99 | 11 | 275 | 71 | 46 | 0 | -29 | -100 | 80 |
| 23 | 1272 | 1075 | 1561 | 1383 | 47 | 1333 | 1057 | 1202 | 1379 |

Table 6 shows the comparison of actual vs. predicted values by each of the three models for the first 48 hours of the testing data. As already covered during the Performance Analysis, the LSTM model was the model with the lowest scores of MSE / RMSE on the testing dataset, which in turn means that it would on average have a smaller error margin when comparing its predictions to the actual values, when compared to the other two models.

Two distinct but important observations that can be made from Table 2, are how well the models are able to follow the daily fluctuations of Passengers, which will be the first one to be discussed, and the issues of negative number predictions. While, as covered in the Performance Analysis, all three models had satisfactory performance, with the LSTM model having the best testing performance and the RF model having the worst, it is evident by the table that all three can roughly follow the hourly fluctuations of Passengers, for these 48 hours of data.

The actual passenger values increase from Index 0 to Index 11, then decrease from Index 11 to Index 22, then they start to increase again at Index 23, and so on. The same behavior, with the accuracy of it being to a varying degree, can be observed by all three models, which, for these hours, roughly follow the same pattern of increasing and decreasing the passenger amounts as the actual values do. The actual values for each day drop to 0 for approximately 3-5 hours a day, which can vary between days, months, and years, but in general, there seem to be 2 or more hours in each day with no passengers present at the terminal. Observing how each of the models deals with these 0 hours, it is possible to observe a significantly different way they are being handled.

For the LSTM model, for these particular 48 hours, it managed to stay above 0 for the majority, only going below 0 in two instances at Index 45 and 46. For the SVRM model, it exhibited a tendency to undershoot in this example, predicting negative passenger values for almost every 0 hour, except for 1 at index 21 where it predicted 8 Passengers. And the RF model, in terms of not predicting negative values, performed the best of all three without dropping under 0 once, for these 48 hours. This may be an indicator that the RF model, while having the worst performance based on MSE / RMSE, might be more resilient than the other two in its ability to learn that there shouldn't be negative passenger values predicted.

### 6.3. Model Complexity & Scalability

The choice of models to be compared was done to encompass two different algorithms, and one neural network, to explore how different architectures would perform when trained on this dataset. The SVRM and RF models used the appropriate libraries from Scikit-Learn [44], and the LSTM model used the appropriate libraries from TensorFlow [43]. The LSTM model, in contrast to the SVRM and RF models, is a neural network, and while this particular one, through the hyperparameter tuning only uses a single hidden LSTM layer, has a large number of architecture combinations possible in terms of layer and neuron configurations. With that said, in terms of complexity, without counting the steps and features parameters, as they are shared for all three models, the LSTM and SVRM models each had 6 hyperparameters that were tuned respectively, and the RF model had

5 hyperparameters. In terms of model complexity and structures, while all three models can potentially be scaled to handle larger datasets, the LSTM model has in comparison a broad potential to be scaled further by the addition of additional layers and neurons, as well as the most complex one due to the nature of the LSTM's different gating mechanism contained within each LSTM layer.

## 7. CONCLUSION

The aim of the research, including the order it is discussed, was done to first provide an overview of the aviation industry, underlining the current and future challenges to emphasize the reason the research within these fields is important. Building on that, the main topic and research objectives, in addition to the impactful literature review, were the creation, training, evaluation, and comparative analysis of the SVRM, RF, and LSTM models. In order to present the comparative analysis and evaluation findings in a manner that made sense within the extent of this paper, in addition to why these algorithms were chosen, an overview of relevant literature that discusses machine learning, neural networks, and the SVRM, RF, and LSTM architectures, in particular is done.

Further, numerous applications of machine learning algorithms and neural networks for the purpose of time-series forecasting were discussed, in addition to relevant literature about these applications within the aviation industry, with research that emphasizes the value of the research. The aim of the research was to contribute to existing research which has adapted machine learning algorithms for the purpose of aiding the aviation industry. The research was done to provide additional alternatives that may be beneficial to the aviation industry. Further, providing a broad overview and comparing two machine learning algorithms and one neural network against each other, may be beneficial as it provides a distinct overview of how these different architectures can learn from this format of passenger and weather data.

While the performance of the models, in terms of accurate passenger number predictions, would have varying degrees of error, the important observation of the models' prediction values is, were they able to follow the hourly fluctuations of passengers. For this dataset, all three models, within the extent of the tested values, being the 48 hours of testing data, were able to roughly follow the fluctuations of passengers, with each model having varying degrees of error.

## 8. FUTURE WORK

Numerous paths may be of interest for future work to the extent of enhancing the aviation industry's ability to predict future air traffic flow. To begin, there are numerous other machine learning algorithms and neural networks that could be trained and evaluated to determine if there may be better alternatives to the LSTM, SVRM, or RF models. Further, one of the main methods of conducting a similar analysis that builds on top of this paper is to use different data. There are two possible avenues of this, one of them may be the use of univariate data where only the passenger value and datetime feature are included, which can be combined with the training of additional algorithms. The second avenue and the recommended one is to use multivariate data and attempt to collect data pertaining to highly impactful features such as a country's GDP and population growth, or similar relevant economic or social factors. The main challenge with this approach is the granularity of data, which for the dataset in this paper was hourly, which greatly limited the amount of available relevant data that could realistically be combined with the hourly passenger data.

## ACKNOWLEDGEMENTS

A special thanks to Avinor, who provided the Passenger data that aided in making this project possible.

**Funding: This research received no external funding.**

**Conflict of interest: The authors declare no conflicts of interest.**

## REFERENCES


[1] J. E. Wang and F. J. Jin, "China's Air Passenger Transport: An Analysis of Recent Trends," http://dx.doi.org/10.2747/1538-7216.48.4.469, vol. 48, no. 4, pp. 469–480, Jul. 2013, doi: 10.2747/1538-7216.48.4.469.

[2] S. M. Phyoe, N. Y. Nguyen, S. Aneeka, and Z. W. Zhong, "The Impact of Population Growth on the Future Air Traffic Demand in Singapore," pp. 688–694, Dec. 2016, doi: 10.2991/CNCT-16.2017.96.

[3] P. Fontanet-Pérez, X. H. Vázquez, and D. Carou, "The impact of the COVID-19 crisis on the US airline market: Are current business models equipped for upcoming changes in the air transport sector?," Case Stud Transp Policy, vol. 10, no. 1, pp. 647–656, Mar. 2022, doi: 10.1016/J.CSTP.2022.01.025.

[4] "Norway Current & Historical Population," Statistics Norway. https://www.ssb.no/en/befolkning/folketall/statistikk/befolkning (accessed Jun. 01, 2023).

[5] M. Thomas and A. Tømmerås, "Norway's 2022 national population projections," 2022. https://www.ssb.no/en/befolkning/befolkningsframskrivinger/artikler/norways-2022-national-population-projections (accessed Jun. 01, 2023).

[6] P. D. United Nations Department of Economic and Social Affairs, "World Population Prospects 2022: Summary of Results," 2022.

[7] "Statistics - Avinor," Avinor Fly Data Archive, 2022. https://avinor.no/en/corporate/about-us/statistics/archive (accessed Jun. 01, 2023).

[8] "Norwegian Centre for Climate Services." https://seklima.met.no/observations/ (accessed Jun. 01, 2023).

[9] World Bank, "Air transport, passengers carried | Data," 2022. https://data.worldbank.org/indicator/IS.AIR.PSGR (accessed Jun. 01, 2023).

[10] "The World of Air Transport in 2019," International Civil Aviation Organisation, 2019. https://www.icao.int/annual-report-2019/Pages/the-world-of-air-transport-in-2019.aspx (accessed Jun. 01, 2023).

[11] S. Addepalli, G. Pagalday, K. Salonitis, and R. Roy, "Socio-economic and demographic factors that contribute to the growth of the civil aviation industry," Procedia Manuf, vol. 19, pp. 2–9, Jan. 2018, doi: 10.1016/J.PROMFG.2018.01.002.

[12] A. Andreoni and M. N. Postorino, "A MULTIVARIATE ARIMA MODEL TO FORECAST AIR TRANSPORT DEMAND," 2006.

[13] P. Forsyth, C. Guiomard, and H. M. Niemeier, "Covid −19, the collapse in passenger demand and airport charges," J Air Transp Manag, vol. 89, p. 101932, Oct. 2020, doi: 10.1016/J.JAIRTRAMAN.2020.101932.

[14] S. Shparberg and B. Lange, "Airbus Global Market Forecast 2022," 2022. https://www.airbus.com/en/products-services/commercial-aircraft/market/global-market-forecast (accessed Jun. 01, 2023).

[15] D. Pérez-Campuzano, P. M. Ortega, L. Rubio Andrada, and A. López-Lázaro, "Artificial Intelligence potential within airlines: a review on how AI can enhance strategic decision-making in times of COVID-19," Journal of Airline and Airport Management, vol. 11, no. 2, p. 53, 2021, doi: 10.3926/jairm.182.

[16] E. Suryani, S. Y. Chou, and C. H. Chen, "Air passenger demand forecasting and passenger terminal capacity expansion: A system dynamics framework," Expert Syst Appl, vol. 37, no. 3, pp. 2324–2339, Mar. 2010, doi: 10.1016/J.ESWA.2009.07.041.

[17] B. Mahesh, "Machine Learning Algorithms -A Review," 2020. https://www.researchgate.net/publication/344717762_Machine_Learning_Algorithms_-A_Review (accessed Jun. 01, 2023).

[18] S. Diksha and K. Neeraj, "A Review on Machine Learning Algorithms, Tasks and Applications," 2017. https://www.researchgate.net/publication/320609700_A_Review_on_Machine_Learning_Algorithms_Tasks_and_Application s (accessed Jun. 01, 2023).

[19] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. E. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," Heliyon, vol. 4, no. 11, p. e00938, Nov. 2018, doi: 10.1016/J.HELIYON.2018.E00938/ATTACHMENT/7F6968D1-2173-4A69-BC61-62AD41135D5C/MMC2.

[20] C. Cortes, V. Vapnik, and L. Saitta, "Support-vector networks," Machine Learning 1995 20:3, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.

[21] A. Raj, "Unlocking the True Power of Support Vector Regression," 2020. https://towardsdatascience.com/unlocking-the-true-power-of-support-vector-regression-847fd123a4a0# (accessed Jun. 01, 2023).

[22] A. Swarnakar, "SVM Simplified – KDAG," 2015. https://kgpdag.wordpress.com/2015/08/12/svm-simplified/ (accessed Jun. 01, 2023).

[23] H. Drucker·, C. J. C. Burges, L. Kaufman, A. Smola··, and V. Vapoik, "Support Vector Regression Machines," 1996.

[24] W. He, Z. Wang, and H. Jiang, "Model optimizing and feature selecting for support vector regression in time series forecasting," Neurocomputing, vol. 72, no. 1–3, pp. 600–611, Dec. 2008, doi: 10.1016/J.NEUCOM.2007.11.010.

[25] K. R. Müller, A. J. Smoła, G. Rätsch, B. Schölkopf, J. Kohlmorgen, and V. Vapnik, "Predicting time series with support vector machines," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 1327, pp. 999–1004, 1997, doi: 10.1007/BFB0020283/COVER.

[26] C. Olah, "Understanding LSTM Networks," 2015. http://colah.github.io/posts/2015-08-Understanding-LSTMs/ (accessed Jun. 01, 2023).

[27] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A Critical Review of Recurrent Neural Networks for Sequence Learning," May 2015, Accessed: Jun. 01, 2023. [Online]. Available: https://arxiv.org/abs/1506.00019v4

[28] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Comput, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/NECO.1997.9.8.1735.

[29] L. Breiman, "Random forests," Mach Learn, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324/METRICS.

[30] L. Breiman, "Bagging predictors," Mach Learn, vol. 24, no. 2, pp. 123–140, 1996, doi: 10.1007/BF00058655/METRICS.

[31] O. G. Ojo, O. N. Ogunnusi, O. Ojo, O. Ogunnusi, and O. Ogunnusi, “COMPARATIVE ALGORITHMS OF SUPPORT VECTOR REGRESSION MACHINE (SVRM) IN MODELING OF AIR PASSENGERS TRAFFIC,” 2020. [Online]. Available: https://www.researchgate.net/publication/358351514

[32] J. C. P. Putra and Safrilah, “Application of Artificial Neural Network to Predict the use of Runway at Juanda International Airport,” IOP Conf Ser Mater Sci Eng, vol. 209, no. 1, p. 012107, Jun. 2017, doi: 10.1088/1757-899X/209/1/012107.

[33] T. O. Blinova, “Analysis of possibility of using neural network to forecast passenger traffic flows in Russia,” Aviation, vol. 11, no. 1, pp. 28–34, 2007, doi: 10.1080/16487788.2007.9635952.

[34] D. Salinas, V. Flunkert, J. Gasthaus, and T. Januschowski, “DeepAR: Probabilistic forecasting with autoregressive recurrent networks,” Int J Forecast, vol. 36, no. 3, pp. 1181–1191, Jul. 2020, doi: 10.1016/J.IJFORECAST.2019.07.001.

[35] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, “Time Series Analysis, Fourth Edition.,” 2013, Accessed: Jun. 01, 2023. [Online]. Available: https://books.google.com/books/about/Time_Series_Analysis.html?id=rNt5CgAAQBAJ

[36] B. Artley, “Time Series Forecasting with ARIMA, SARIMA and SARIMAX,” 2022. https://towardsdatascience.com/time-series-forecasting-with-arima-sarima-and-sarimax-ee61099e78f6 (accessed Jun. 01, 2023).

[37] A. Bajaj, “ARIMA & SARIMA: Real-World Time Series Forecasting,” 2023. https://neptune.ai/blog/arima-sarima-real-world-time-series-forecasting-guide (accessed Jun. 01, 2023).

[38] S. Xu, H. K. Chan, and T. Zhang, “Forecasting the demand of the aviation industry using hybrid time series SARIMA-SVR approach,” Transp Res E Logist Transp Rev, vol. 122, pp. 169–180, Feb. 2019, doi: 10.1016/J.TRE.2018.12.005.

[39] Á. Rodríguez-Sanz, J. Cano, and B. Rubio Fernández, “Impact of weather conditions on airport arrival delay and throughput,” Aircraft Engineering and Aerospace Technology, vol. 94, no. 1, pp. 60–78, Jan. 2022, doi: 10.1108/AEAT-12-2020-0318/FULL/XML.

[40] M. Schultz, S. Lorenz, R. Schmitz, and L. Delgado, “Weather Impact on Airport Performance,” Aerospace 2018, Vol. 5, Page 109, vol. 5, no. 4, p. 109, Oct. 2018, doi: 10.3390/AEROSPACE5040109.

[41] S. Reitmann, S. Alam, and M. Schultz, “Advanced Quantification of Weather Impact on Air Traffic Management - Intelligent Weather Categorization with Machine Learning,” 2019. https://www.researchgate.net/publication/339017452_Advanced_Quantification_of_Weather_Impact_on_Air_Traffic_Management_-_Intelligent_Weather_Categorization_with_Machine_Learning (accessed Jun. 01, 2023).

[42] C. R. Harris et al., “Array programming with NumPy,” Nature, vol. 585, no. 7825, pp. 357–362, 2020, doi: 10.1038/s41586-020-2649-2.

[43] M. Abadi et al., “TensorFlow: A System for Large-Scale Machine Learning TensorFlow: A system for large-scale machine learning,” 2015, Accessed: Jun. 01, 2023. [Online]. Available: https://tensorflow.org.

[44] F. Pedregosa FABIANPEDREGOSA et al., “Scikit-learn: Machine Learning in Python,” Journal of Machine Learning Research, vol. 12, no. 85, pp. 2825–2830, 2011, Accessed: Jun. 01, 2023. [Online]. Available: http://jmlr.org/papers/v12/pedregosa11a.html