INTELIGENCIA ARTIFICIAL

Airport resource allocation using machine learning techniques

Maged Mamdouh[1,A], Mostafa Ezzat[1,2,B], Hesham A.Hefny[1,3,C]

[1]Department of Computer Science, Faculty of Graduate Studies for Statistical Researchs,
Cairo University, Egypt.
[2]Associate Professor, [3]Vice Dean.
[A]magedmamdouh@pg.cu.edu.eg, [B]mostafa.ezzat@cu.edu.eg, [C]hehefny@ieee.org

Abstract The airport ground handling has a global trend to meet the Service Level Agreement (SLA) requirements that represents resource allocation with more restrictions according to flights. That can be achieved by predicting future resource allocation. this research presents a comparison between the most used machine learning techniques implemented in many different fields for resource allocation and demand prediction. The prediction model nominated and used in this research is the Support Vector Machine (SVM) to predict the required resources for each flight, despite the restrictions imposed by airlines when contracting their services in the SLA. The approach has been trained and tested using real data from Cairo international airport. the proposed SVM technique implemented and explained with a varying accuracy of resource allocation prediction, showing that even for variations accuracy in resource prediction in different scenarios, the SVM technique can produce a good performance as resource allocation in the airport.

Keywords: Ground handling agents, Service level agreement , Resource allocation , Machine learning , Support vector machine.

1   Introduction

Airports were once seen as the main infrastructure providers of airlines. On the other hand, scientists see from the point of view of the entire air transport system that the problem is in providing high quality services to airlines depending on security and discipline, one of the main tasks of the airport is Ground handling at the airport which provides aviation services to airlines and passengers [14], Plays an important role among many activities that contribute to the activity and safety of air transport [6]. Defining ground handling process that itâs the operations of preparing flights while they are standing on the ground and identifying the requirements of equipment and personnel for such services whether the flight is arrival or departure or turnaround [19],[4]. Competition has increased in providing efficient service while allowing third party companies to provide ground handling. Based on the overall planning structure of the airport [19]. In some cases, ground handling services at airports are provided by airlines, specialized aircraft ground handling companies or Ground Handling Agents (GHA), those represents performance that affects the flight schedule time as it is likely to cause delays to the flights, the services are including passengers, baggage and aircraft maintenance before and after the flight [14],[9].

The shortage of integration of optimal planning and inefficient use of resources for various activities can lead to huge additional costs for airlines and airports, which leads to many flight delays for aircraft and passenger disturbance [6],[19]. It is not the prerogative of GHA to avoid operational density to modify flight schedules such as flight date, flight time, location, etc. So GHA want to implement two objectives together the first is to provide the best possible service to their customer's airlines at the lowest cost and the second is the needs to provide services with limited resources [4],[18]. To achieve these objectives for service agents, resources should

be prepared to provide the appropriate number of equipment and personnel resources on flights to eliminate the possibility of delays in flights or poor service delivery[9]. So airport resource allocation planning is the best way to improve the operational efficiency of an air transport system [19]. But the agents will face a problem when having a shortage of available resources in the resource allocation of service where that must provide according to a series of restrictions imposed by SLA contract airlines, depending on the different features such airlines and aircraft type [9],[18], the orientation and location of the airport terminal also reflect the sort of resources used and task duration.

The objective of this research study is to design an approached machine learning technique using SVM to train on the data to flights resource allocation with accommodating all possible constraints in SLA. The remaining paper is structured as follows: Section 2 discusses related work and introduces brief the methodology of machine learning techniques comparison. Section 3 illustrate the steps of the of proposed model. In section 4 discusses the experimental results. The conclusion and the future work are present in section 5.

2   Related Work

The resource allocation in ground handling services was not implemented by using machine learning techniques models. Despite several authors have been using machine learning techniques for making prediction models in many different fields aims to predict provisioning of resources by (predict resource allocation or predict resource demand) using machine learning techniques to satisfy services level agreement.

The author in [1] uses one way in order to meet SLA like requirements of accurate virtual machine for making prediction models, which predict future cloud resource demands that metrics for response time and throughput for resource allocation using Neural Networks (NN) Linear Regression (LR) and SVM. the SVM model gets better accuracy than both the NN and LR.

The author in [8] proposes machine learning techniques by using NN and LR to resource allocation in cloud computing, which provides online transaction systems a lot of dynamic resources for enterprises, in order to enhance manage the proactive and dynamic resources, Prevent hardware resource allocation from being delayed by several minutes. That explained The NN techniques have more accuracy than LR.

The author in [13] solved the problem of virtual data centers and cloud hosting services that it faces special in resource allocating that are considered very importantly, such as memory, CPU, and I/O bandwidth for virtual machines. Proposed this model using automated learning techniques, using NN and SVM. This model helps to improve virtual machine size, thus reducing costs, allowing cloud service providers for increased volumes configure by delivering new shipping models to increase virtual machine size performance.

The main objective of this paper [17] is to provide an innovative technique called Position Balanced Parallel Particle Swarm Optimization (PB-PPSO) to resource allocation in cloud computing and achieve a high level of user satisfaction and high profits in the environment. Both SVM and NN have used to find optimized resources for the task group while minimizing distance with minimum price. The PB-PPSO technique shows high profitability compared to current methods such as SVM and NN.

The author in [25] used the DMAA (Double Multi-Attribute Auction) mechanism for an increased utilization rate of cloud resources such as satisfying CRC (Cloud Resource Consumer) demands and increasing the CRP (Cloud Resource Provider) utilization rate in the cloud. The author used NN to transform the multiple non-prices attributes and also used SVM in the second stage within the model to support the model approach, for predicting the price depending on the historical transaction records to provide suitable pricing of resources.

The author in [20] proposed the prediction model Multilayer Perceptron (MLP), LR and support vector regression (SVR) for predicting resources which on premise data centers based on the utilization, to provide cloud resources based on applications. Where new sets of applications are not added or deleted frequently. The author also used model SVM classifier to classify resource provisioned to support the proposed model in order to predict the adequate set of resources which user requests and requirements performance.

The author in [5] provides approach a new prediction hybrid model, Data-Driven Hybrid Prediction Approach (DDHPA) that merges the Multiple Support Vector Regression (MSVR) model and the Autoregressive Integrated Moving Average (ARIMA) model. The hybrid model of predicting for consumer needs of cloud resources, which are provided for cloud consumers, to make a better decision in terms of disk storage utilization, CPU, memory, and response time. The author used the SVM to estimate parameters from the input dataset to define the relationship between each input attribute in the dataset and the target prediction class to be used by ARIMA and MSVR approach models.

The author in [15] aims to reduce resource costs in cloud computing to provide resource allocation on demand in a pay-as-you-go model to cloud consumers according to SLA. It presents an approach called Automatic Proactive Resource Allocation (APRA) that proactively allocates resources to VMs by predicting resource usage demands with two resources, CPU and memory. Used SVM to support model by time series prediction for resources, taking

into account cross correlation between resources that in multiple of VMs that at the same time in a multi-tier application.

The authors in these two papers [23],[12] they use the SVM model to build a prediction algorithm for virtual resources to handle the changes that happen in the environment cloud platform. But in the second paper was to build agents' behavior learning models on the services via a global knowledge base in advance. Both papers by empirical results were achieved to performance in real-time and got a prediction of requirements virtual resources high accuracy.

In this paper [21] provides and develops a prediction model for the selection and management of resource allocation provisioning to ensure guaranteed performance for multi-tier web applications, predict a correct set of resources by providing sufficient resources with different requirements for applications that must accept predictable performance. Used LR, SVR, and MLP, also used the SVM classifier for identifying the type of provisionings such as over-provisioning, under-provisioning, or normal provisioning and give it a label.

The author in [3] describes a proposed machine learning-based algorithm to solve the problem of wireless network resource allocation capacity optimization, which includes matters such as link scheduling, power control, and flow allocation. Used SVM for an operation classify each link to determine the status of completion of power off or maximum power consumption. DBNs (Beside Deep Belief Networks) are used for calculated for approximate optimum power distribution. The two results are collected to obtain a rough solution to the program and demonstrate the effectiveness of the algorithm based on automated learning by results for simulation.

The author in[7] includes the field of cellular networks and the tendency to improve user rates through the deployment of CRANs (cloud radio access networks), which consist of dense groups of radio headers which within them lead to the complexity of interference coordination and information overload on the channel. This study provided a design for the learning-based resource allocation system for 5G systems by used RF model for reducing the coordinated complexity of allocating resources and indicative costs, predict the party's configuration and coding system at any given location served by CRANs (cloud radio access networks). Through performance assessments, it was explained that the learning-based resource allocation technique could achieve fairly efficiency similar to CSI-based systems.

The related work gives an overview of techniques are used to compares them in order to choose the suitable technique to make an acceptable prediction for resource allocation in-ground handling of flight resources with capacity for all possible restrictions due to service level agreement (SLA). Table 1 shows comparison of machine learning techniques such as neural network (NN), linear regression (LR), random forest (RF) and support vector regression (SVR), have been widely used and shared by authors in building prediction and resource allocation models in [1],[8],[13],[17],[25],[20],[21],[3] and [7]. regard the support vector machine (SVM) recently was a powerful classification technique, has been gaining significant popularity showed in [8],[13],[17] and [12].

Table 1: Comparison table of machine learning techniques.

Feature	SVM	SVR	LR	NN	MLP	RF
A large value (cost, neurons) lead to overfitting	√	√		√		
A small value of (cost, neurons) lead to underfitting	√			√		
Accuracy is high	√	√	√	√		
Effective	√			√		√
Large training or indelicate parameters is lead to low accuracy				√		√
Powerful classification popularity and supervised	√	√				
Produce errors for complex applications				√	√	
Suitable for the type of data linear	√	√	√	√	√	
Suitable for the type of data nonlinear	√	√			√	
Suitable moderate number of tasks	√	√				
The more dimension is Positive with accuracy				√		√
Uses for prediction of future resource	√		√	√		

3   Proposed Aproach

The proposed model will be illustrated in two phases are data set and approach model.

1. Data Set

2. Proposed model

3.1   Data Set

This section aims to create the data set that contains historical flights with their specific ground handling services. The proposed dataset which will be used is real but will be private, so the airlines encoded in sample tables especially in SLA and data set which will be shown. Figure.1 shows architecture steps created of the data set approach. the steps are performed in the following steps:

1. Historical data collection from the flight schedule.

2. Converted SLA to rules.

3. Handling features flight schedule and converted to classes.

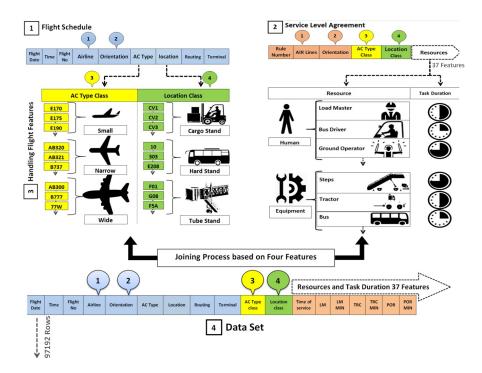4. Joining flight schedule with SLA rules to create data set.

Figure 1: Architecture steps of the data set creation

3.1.1   Historical data collection from the flight schedule

Most airports have flight arrivals followed by flight departure called a bank that leads to huge peaks on demand for airport resources [11]. Normally every six months flight schedules updates after all flight schedules are determined by airlines and the timing is reset dynamically to reduce delays [14]. Anyway, the historical data contains a flight schedule over 12 months that contains all the flights served by our ground handling agent. Table 2 shows the sample from flight schedule data which contains 97192 of rows represent historical flights in 12 months.

3.1.2   Convert SLA to form of rules

providing ground handling services is an affair to the standards and policies of the airline [9]. The international air transport association (IATA) decide some principles and indicators that are called as level of service (LOS) [18]. Accordingly, the SLA would be created in the form of rules for compatibility with each possible flight, which specific demands ground handling resources such as equipment and personnel resources with task duration, according to four features, these are airlines, aircraft type, orientation and location in the airport. The SLA

Table 2: Sample of flight schedule.

Flight date	Flight No	Airline	Orientation	Aircraft type	location	Route	Time	Terminal
01/05/2018	SV 309	Saudi Air	Turnaround	333	E1	JED	00:50	2
01/05/2018	EY 651	Etihad	Turnaround	332	E6	AUH	05:05	2
01/05/2018	MS 901	Egypt Air	Departure	738	G1	DXB	05:25	3
01/05/2018	NP 114	Nile Air	Arrival	321	22	YNB	05:30	1
01/05/2018	SM 1009	Air Cairo	Arrival	320	38	SSH	05:30	1
01/05/2018	RJ 508	Royal Jord	Departure	318	E202	AMM	07:00	2
01/05/2018	MS 661	Egypt Air	Departure	332	307	JED	07:20	S

contracts have assembled and formed it in 293 bases. It contains all the possibilities for the occurrence of flight with different restrictions. Table 3 shows the SLA rules we presented it in a longitudinal shape. Note the value ("NA") in table means the resource not applicable or not required.

Table 3: Sample of SLA

Rule number	2	34	46	50	73	88
Airlines	Airline A3	Airline B	Airline B2	Airline D1	Airline F3	Airline G1
Orientation	Arrival	Arrival	Departure	Turnaround	Turnaround	Arrival
Class aircraft type	Wide	EMP	Narrow	Wide	Wide	Narrow
Class location	Hard	Hard	Tube	Tube	Tube	Hard
Time of services	90	35	45	110	60	45
Terminal	1	3	3	2	3	2
Turnround coordinator	1	NA	NA	1	1	1
Turnround coordinator minutes	90	NA	NA	110	60	45
Load master	1	1	1	1	1	1
Load master minutes	90	35	45	110	60	45
Porter	6	4	4	6	4	4
Porter minutes	90	35	45	110	60	45
Load driver	2	2	2	2	2	2
Load driver minutes	90	35	45	110	60	45
Bus driver arrival	6	3	NA	NA	NA	4
Bus driver arrival minutes	15	15	NA	NA	NA	15
Bus driver departure	NA	NA	NA	NA	NA	NA
Bus driver departure minutes	NA	NA	NA	NA	NA	NA
Crew	2	1	1	NA	NA	1
Crew minutes	25	20	20	NA	NA	20
West driver por	1	1	NA	1	1	1
West driver porter minutes	20	10	NA	20	15	15
Water driver porter	NA	NA	1	1	1	
Water driver porter minutes	NA	NA	15	20	15	NA
Ground services operation	3	2	2	3	2	2
Ground services operation minut	90	35	45	110	60	45
First step	1	1	NA	NA	NA	1
First step minutes	30	20	NA	NA	NA	25
Second step	1	1	NA	NA	NA	1
Second step minutes	90	35	NA	NA	NA	45
Ground power unite	1	1	NA	NA	NA	1
Ground power unite minutes	90	35	NA	NA	NA	45
Air condition	1	1	NA	NA	NA	1
Air condition minutes	90	35	NA	NA	NA	45
Conveyor belt	1	2	2	1	2	2
Conveyor belt minutes	90	35	45	110	60	45
High loader	2	NA	NA	2	NA	NA
High loader minutes	90	NA	NA	110	NA	NA
Cabin clean	6	4	NA	8	4	4
Cabin clean minutes	30	15	NA	30	25	25
Upper deck	NA	NA	NA	NA	NA	NA
Upper deck minutes	NA	NA	NA	NA	NA	NA

3.1.3   Handling features flight schedule and converted to classes

The joining process between historical flight schedule and SLA is based on 4 basic features, which already exist in SLA. The flight schedule needs handling the first two features for converted into classes to be similar to SLA features and make the joining process between both successfully.

- Air craft type : The air craft type is the first feature that must be converted in the flight schedules to AC type class. Each type of aircraft in the aircraft handling manual is different in determining the procedures for carrying out ground handling on it [14]. In our case, air craft type is divided into small, narrow and wide for each type of them needs specific resources and times. Table 4 shows sample air craft type.

Table 4: Sample of AC type class

Aircraft type	Aircraft type class
332	Wide
738	Narrow
E170	Small
B772	Wide
A321	Narrow

- Location : The location is the second feature that must be converted in the flight schedules to location class. Ground service operations are usually carried out at airports in tube areas and ground areas, the ground areas called hardstand in which case passengers must be taken by bus to and from the gate. The tube areas called tubes stand in these tubes passengers can walk directly from the station to the flight and some equipment is automatically supplied [14],[4]. In our case,the location class is divided into hard, tube and cargo locations, where each of them needs specific resources and times. Table5 shows sample of location class.

Table 5: Sample of location class

Location	Location class
42	Cargo
314	Hard
327	Tube
E09	Tube
E202	Hard

- Orientation : The orientation is the third feature that effects on tasks and operations that occur in the aircraft while they are on the ground, affected by the aircraft orientation if its arrival or departure and turnaround. The orientation feature didn't need to convert to classes.

- Airline : The airline is the fourth feature that divided into low cost airlines and large airlines offering full services [18]. Each of them requires different time and resources for services. The airline feature didn't need to convert to classes.

3.1.4   Joining flight schedule with SLA rules to create data set

The data set by created by combining both flight schedules with the SLA this to make each flight assigns all the required resources to be served according to the major four features. The dataset becomes 50 features as columns and containing 97192 of a row, each row contains the flight features joined with its SLA resource allocated after the joining process. Table 6 shows the sample of data set after joining.

Table 6: Sample of data set

Flight date	05/01/2018	05/01/2018	05/01/2018	05/01/2018	05/01/2018	05/01/2018
Flight No	Z 634	A 670	R 621	Y 1042	F 505	S 113
Airline	Airline Z	Airline A	Airline R	Airline Y	Airline F	Airline S
Orientation	Arrival	Arrival	Turnaround	Departure	Arrival	Departure
Aircraft type	738	738	320	320	300	321
location	314	327	16	22	CV4	22
Routing	ELQ	JED	SHJ	RMF	SHJ	YNB
Time	00:01	00:10	00:15	00:30	00:40	00:45
Terminal	3	S	1	1	1	1
Class aircraft type	Narrow	Narrow	Narrow	Narrow	Wide	Narrow
Class location	HARD	HARD	HARD	HARD	Cargo	HARD
Rules number	35	35	13	3	252	248
Time of service	45	45	60	60	90	45
Terminal	3	3	1	1	3	1
Turnround coordinator	NA	NA	1	1	NA	NA
Turnround coordinator minutes	NA	NA	60	60	NA	NA
Load master	1	1	1	1	1	NA
Load master minutes	45	45	60	60	90	NA
Porter	4	4	4	4	8	NA
Porter minutes	45	45	60	60	90	NA
Load driver	2	2	2	2	2	NA
Load driver minutes	45	45	60	60	90	NA
Bus driver arrival	4	4	4	NA	NA	NA
Bus driver arrival minutes	15	15	15	NA	NA	NA
Bus driver departure	NA	NA	4	4	NA	4
Bus driver departure minutes	NA	NA	15	15	NA	15
Crew	1	1	NA	1	1	1
Crew minutes	20	20	NA	20	20	20
West driver por	1	1	1	NA	1	NA
West driver porter minutes	15	15	15	NA	10	NA
Water driver porter	NA	NA	1	1	NA	NA
Water driver porter minutes	NA	NA	15	15	NA	NA
Ground services operation	2	2	2	2	2	NA
Ground services operation minut	45	45	60	60	90	NA
First step	1	1	2	1	1	NA
First step minutes	25	25	60	60	90	NA
Second step	1	1	NA	1	NA	NA
Second step minutes	45	45	NA	25	NA	NA
Ground power unite	1	1	1	1	1	NA
Ground power unite minutes	45	45	60	60	90	NA
Air condition	1	1	1	1	NA	NA
Air condition minutes	45	45	60	60	NA	NA
Conveyor belt	2	2	2	2	1	NA
Conveyor belt minutes	45	45	60	60	90	NA
High loader	NA	NA	NA	NA	2	NA
High loader minutes	NA	NA	NA	NA	90	NA
Cabin clean	4	4	4	NA	NA	NA
Cabin clean minutes	25	25	25	NA	NA	NA
Upper deck	NA	NA	NA	NA	1	NA
Upper deck minutes	NA	NA	NA	NA	90	NA

3.2 Proposed model

In this section recommends using a SVM that mentioned in section 2 as a prediction model in the ground handling because it's the most successful supervised algorithms implemented [16], besides its usage to support approaches models used in different studies in [25],[20],[5],[15],[21],[3]. It avoids overfitting and is guaranteed to discover the upper optimum rather than any other technique that can get a bolt to local minima [15]. It doesnt accumulate prediction errors and reduces the amount of training and testing data [5]. It has an SMO function that prevents the programming quadratic problem (QP) that arises during the training [21]. Figure.2 shows architecture steps of model approach.The implementation process explains in the following parts:
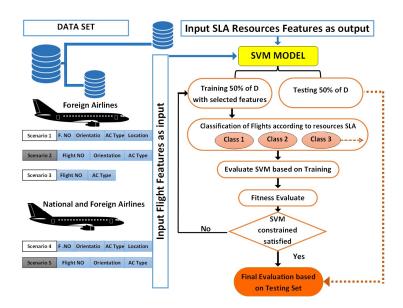
Figure 2: Architecture of the SVM model approach.

1. Configuration of SVM parameters

2. Training of the prediction SVM model with the data set

3. Testing and result of the training model

3.2.1 Configuration of SVM parameters

The implementation of multi-dimensional support vector classification is based on libsvm. There are existing several parameters for a purpose to tune for acquiring accurate modeling with SVM are kernel, gamma, and cost. In this work uses kernel functions especially focus is on radial basis function (RBF), which has been proposed in several works and widely used [22]. The accuracy of an SVM model is very affected by these two parameters gamma and cost values. The parameter C is responsible for the balance between the minimization of error and the margin maximization. When C is a small value the margin will be maximized and leads to misclassified for a high number of samples It follows severe under-fitting. Conversely, very large C values lead to minimizing the margin's width result in the weight of the no separable samples increased [24]. The SVM algorithm influenced by the gamma parameter that can lead classify new data correctly [10]. Gamma and cost values were determined according to the calculated values tuning experiment on a part of the dataset that will be used. Table.7 shows accuracy of the experiment of tuning parameters. Figure.3 shows accuracy experiment of tuning parameters.
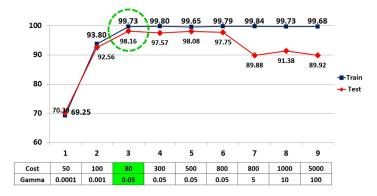

Figure 3: Accuracy of tuning.

Table 7: Configuration of SVM parameters.

Kernel	rbf
C	80
Gamma	0.05
degree	3
Cash size	200

3.2.2   Training of the prediction SVM with the data set

In the SVM training process, we do this using the Scikit_Learn library and specifically the (Train/Test Split) method so can, well, split to training and test data set, There are existing several parameters are (arrays) mapping of flight description features as array 1 to predict their resource allocation as array 2,(train_size = 0.50) and (test_size = 0.50) divided into 50% for training and 50% for testing, used the default value in remains of parameters (random_state = None), (shuffle = True) and (stratify = None). All data features have been encoded to be suitable for training and testing, excepted flight resource SLA features because they are already in numbers represented resource count and task duration in minutes. The 5 scenarios have been created in each scenario chosen different flight description features to predict flight resources by measuring training and testing accuracy. In scenarios 1, 2 and 3 includes foreign airlines to measure the results for other airlines separately and excludes national airline from the data set because this airline has approximately 70% of the total operations of the flights. In scenarios 4 and 5 includes all airlines.

- Scenario 1: Includes foreign airlines and use these flight features (Flight number, Orientation, Aircraft type, location) in order to check model accuracy for predict flight resources if with the selected major features of flight.

- Scenario 2: Includes foreign airlines and use these flight features (Flight number, Orientation, Aircraft type) in this scenario, had excluded the location feature, In order to match these features used with the features available in real practice. The future flight schedule does not include the location because it is unknown until a short time before flight time.

- Scenario 3: Includes foreign airlines and use these flight features (Flight number, Aircraft type) in this scenario, had excluded the orientation feature because it is common knowledge that flight number expresses the case of orientation if its arrival or departure. In order to if we can use these features to predict resources with future flight schedules with satisfying accuracy.

- Scenario 4: Includes all airlines with these flight features (Airlines, Orientation, Aircraft type, location) in order to check model accuracy for predict flight resources if with the selected major features of flight in case all airlines foreign and national.

- Scenario 5: Includes all airlines with these flight features (Airlines, Orientation, Aircraft type) In order to match these features used with future flight schedules as mentioned before.

3.2.3   Testing and result of the training model

The SVM model has been trained and testing by 50% of the data set for each process. That all data set was implemented in the first three scenarios 1,2 and 3 on four different sizes of data included foreign airlines in 1 month, 3 months, 6 months and 12 months, respectively 2437, 7671, 14737 and 25822 of rows. The implementation in scenarios 4 and 5 on the same duration time 1 month, 3 months, 6 months and 12 months, but with different sizes of data when including all airlines, respectively 8727, 26848, 53854 and 97192 of rows.

4   Results and discussion

The SVM model was appropriate to measure its accuracy for training and testing by used accuracy and is defined in the following equation:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \tag{1}$$

The tool of the model used is (Train/Test Split) as mentioned before, is split the dataset into training data and testing, the trained fit too nicely in order to be generalized on other data for later use testing data[2]. It will be found later the training accuracy in scenarios achieves 99.99% and 100% maybe this is related to this reason, according to our testing data may few new flights not be trained before, so it may lead to lack of precision in the new testing dataset. Besides overfitting generally occurs when the model is too complicated when including too many features and variables [2]. Also, our approach model includes many outputs as resources 37 features. However, the accuracy of the test is increased to get closely accuracy of training when the data set increased 12 months that means the model training on all potential flights and well functioning.

In 5 scenarios, resource allocation can be predicted for each flight correctly. Additionally, the scenarios are different in selected flight description features which can predict resources. The prediction results for accuracy with the different flight features are shown in the scenarios Table 8 to Table 12 and details are illustrated as follows:

- Scenario 1: When used only foreign airlines data and used flight features (Flight number, Orientation, AC type, Location) the accuracy started when the data size was 1 month 99.99% in the data training and 92.10% in data testing, when the data size was 12 months the accuracy was 99.16%in data training and 97.20% in data testing. The model predicts flights resource considerable accuracy with foreign airlines with these major features.

- Scenario 2: This scenario aims to ignore the location feature make a prediction of the resources for foreign airlines depending on the same flight description features available in real practice in future flight schedules which agents had in advance. The location unknown until a short time before the flight, So when used foreign airlines data and used flight features (Flight number, Orientation, AC type) the accuracy started when the data size was 1 month 98.43% in the data training and 95.35% in data testing when the data size was 12 months the accuracy was 97.55% in data training and 95.31% in data testing. This degree of accuracy due to the foreign airline flight numbers in coordination with aviation authorities are always fixed on one location class. So, exclude the location feature doesn't impact on change the allocation of resources on foreign airline flights.

- Scenario 3: When excluding the orientation feature in order to try to be the use of these features only (Flight number, AC type) we found when the data size was 12 months the testing accuracy decreased to 92.73% rather than the previous scenario 97.55%. So would prefer to used orientation feature the same as scenario 2 when predicting resources, especially as we already have an orientation feature in flight schedule.

- Scenario 4: When used all airlines data including national airline and used flight major features (Flight number, Orientation, AC type, Location) the accuracy started when the data size was 1 month 100% in the data training and 84.50% in data testing, when the data size was 12 months accuracy was 100% in data training and 95.50% in data testing. The model predicts flight resources considerable accuracy for all airlines with these major features the same as scenario 1.

- Scenario 5: When predicting resources for all airlines including national airline which has approximately 70% of the total operations and used flight features as in scenario 2 are (Flight number, Orientation, AC type) as mentioned that ignoring location feature because it is the common typical situation similar to the real case when GHA having flight schedule and needs a prediction of resources without knowing what the location. The accuracy when the data size was 12 months of data the accuracy was 85.19% in data training and 81.75% in data testing. Despite the expected accuracy should be in the same scenario 2 when the data size was 12 months of data the accuracy was 97.55% in data training and 95.31% in data testing. But The degree of accuracy was dropped because of the location class of the flight numbers national airline always is changing in coordination with aviation authorities this matters related to national airlines cost, that means the national airline has approximately 70% of the operations and changing their flight resources daily.

Figure. 4 shows measuring accuracy training and testing and in Figure. 5 shows combined accuracy. Figure 6 illustrates the different testing accuracy with the same flight features in two scenarios 2 and 5. Tables 13, 14 show the sample of flight schedule to view the location class for both national airline and other foreign airlines. Note the national airline flight numbers do not stand on one location class permanently this is leading to changing of resources, but on the other hand, the other airlines' flight numbers always stable in one location class type.

5   Conclusion

Recently, machine learning techniques significant advances achieved and have powerful methods in a wide range for prediction resource allocation. The main purpose of this work is to show that there is a possibility with machine learning techniques for predicting the resource allocation at the airport, to the allocation ground handling required

Table 8: Scenario 1 Foreign airlines with features (Flight number, Orientation, AC Type, location).

period	size of flights	time for executing	training accuracy	testing accuracy
1 month	2437 rows	20 second	99.99	92.10
3 months	7671 rows	30 second	99.99	95.40
6 months	14737 rows	1 minute	99.85	96.92
12 months	25822 rows	3 minute	99.16	97.20

Table 9: Scenario 2 Foreign airlines with features (Flight number, Orientation, AC Type).

period	size of flights	time for executing	training accuracy	testing accuracy
1 month	2437 rows	20 second	98.43	95.35
3 months	7671 rows	30 second	100	95.40
6 months	14737 rows	1 minute	98.75	95.57
12 months	25822 rows	3 minute	97.55	95.31

Table 10: Scenario 3 Foreign airlines with features (Flight number, AC Type).

period	size of flights	time for executing	training accuracy	testing accuracy
1 month	2437 rows	20 second	96.20	90.25
3 months	7671 rows	30 second	94.66	91.08
6 months	14737 rows	1 minute	93.09	90.70
12 months	25822 rows	3 minute	92.73	90.25

Table 11: Scenario 4 All airlines with features (Flight number, Orientation, AC Type, location).

period	size of flights	time for executing	training accuracy	testing accuracy
1 month	8727 rows	2 minute	100	84.50
3 months	26848 rows	5 minute	100	92.35
6 months	53854 rows	42 minute	100	94.63
12 months	97192 rows	03:30 hour	100	95.50

Table 12: Scenario 5 All airlines with features (Flight number, Orientation, AC Type).

period	size of flights	time for executing	training accuracy	testing accuracy
1 month	8727 rows	2 minute	88.56	79.93
3 months	26848 rows	4 minute	86.75	82.27
6 months	53854 rows	13 minute	86.57	82.32
12 months	97192 rows	01:10 hour	85.19	81.75

Table 13: National airline.

Flight Date	Airline	Flight No	Orientation	AC Type	location	Route	Time	Terminal	AC Class	location class
03/06/2019	Egypt Air	MS 181	Arrival	320	330	HBE	07:00	3	Narrow	Hard
03/07/2019	Egypt Air	MS 181	Arrival	320	F3	HBE	07:00	3	Narrow	Tube
03/17/2019	Egypt Air	MS 644	Arrival	738	311	JED	23:05	4	Narrow	Hard
03/18/2019	Egypt Air	MS 644	Arrival	738	G3	JED	23:05	4	Narrow	Tube

of flights with capacity for all possible restrictions due Service Level Agreement (SLA) including features airline,

Table 14: Foreign airlines.

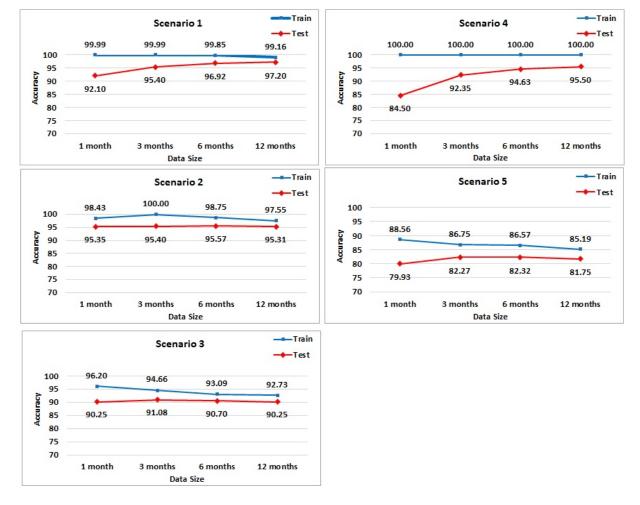| Flight Date	Airline	Flight No	Orientation	AC Type	location	Route	Time	Term-inal	AC Class	location class
03/01/2019	Emirates	EK 923	Turnaround	77W	E9	DXB	17:15	2	Wide	Tube
03/02/2019	Emirates	EK 923	Turnaround	773	E7	DXB	17:15	2	Wide	Tube
03/01/2019	Royal Air	AT 272	Turnaround	738	E6	CMN	04:25	2	Narrow	Tube
03/02/2019	Royal Air	AT 272	Turnaround	738	E2	CMN	04:25	2	Narrow	Tube

Figure 4: Measuring training and testing accuracy.

aircraft type, orientation, and location.

In this paper, we provide a study with a review of related work to make comparisons between the machine learning techniques used in the allocation of resources in the different previous studies, the support vector machine regards the recently was a powerful classification technique increasingly popular and outperformed standard machine learning algorithms, applying the SVM has been filtered the best technique for use in our approach model. Additionally, the proposed data set has been created contains historical flight schedules with their resource allocation that represent SLA to make each flight combined with all the resources required to serve it according to major flight features. The SVM model was trained and testing on 5 scenarios applied on four data set sizes and each scenario has used different flight features as inputs model to predict the resources.

The accuracies are prove outperformed in scenarios, particularly, in scenario 5, when the data include all airlines. Despite the national airline has approximately 70% total operations of the flights and don't have permanent service level agreement and requirement varying of resources, however, it has proved with three features are

Scenario 1	Train	99.99	99.99	99.85	99.16
	Test	92.10	95.40	96.92	97.20
Scenario 2	Train	98.43	100	98.75	97.55
	Test	95.35	95.40	95.57	95.31
Scenario 3	Train	96.20	94.66	93.09	92.73
	Test	90.25	91.08	90.70	90.25
Scenario 4	Train	100	100	100	100
	Test	84.50	92.35	94.63	95.50
Scenario 5	Train	88.56	86.75	86.57	85.19
	Test	79.93	82.27	82.32	81.75
		1 month	3 month	6 month	12 month

Figure 5: Combined the accuracies of scenarios.

Figure 6: Test the accuracy in scenarios 3 and 5.

(Flight number, Orientation, Aircraft type) the accuracy was 85.19% in data training and 81.75% in data testing. That means the SVM model can predict resource allocation for each flight by three features that correspond in future flight schedule form which possessed by in advance ground handling agents. Since SVM has an impact on predicting resource allocation in ground handling according to the performance of the model, it is worth implement other types of machine learning techniques or deep learning as future work. Also, it is worth investigating ways to predict resource demand in airports using machine learning techniques.

References

[1] Akindele A Bankole and Samuel A Ajila. Predicting cloud resource provisioning using machine learning techniques. In *2013 26th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–4. IEEE, 2013. DOI: http://dx.doi.org/10.1109/CCECE.2013.6567848.

[2] Adi Bronshtein. Train/test split and cross validation in python, towards data science, may 17, 2017. [online]. available: https://towardsdatascience.com/train-test-split-and-cross-validation-in-python-80b61beca4b6.

[3] Xianghui Cao, Rui Ma, Lu Liu, Hongbao Shi, Yu Cheng, and Changyin Sun. A machine learning-based algorithm for joint scheduling and power control in wireless networks. *IEEE Internet of Things Journal*, 5(6):4308–4318, 2018. DOI :http://dx.doi.org/10.1109/JIOT.2018.2853661.

[4] Tommy Clausen and David Pisinger. *Airport ground staff scheduling*. Technical University of Denmark (DTU), 2010.

[5] Abdelkarim Erradi and Hisham A Kholidy. An efficient hybrid prediction approach for predicting cloud consumer resource needs. In *2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*, pages 1–8. IEEE, 2016. DOI: 10.1109/AICCSA.2016.7945639.

[6] Salma Fitouri-Trabelsi, Felix Mora-Camino, Carlos Alberto Nunes-Cosenza, and Li Weigang. Integrated decision making for ground handling management. 2015.

[7] Sahar Imtiaz, Hadi Ghauch, George P Koudouridis, and James Gross. Random forests resource allocation for 5g systems: Performance and robustness study. In *2018 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pages 326–331. IEEE, 2018. DOI: https://doi.org/10.1109/WCNCW.2018.8369028.

[8] Sadeka Islam, Jacky Keung, Kevin Lee, and Anna Liu. Empirical prediction models for adaptive resource provisioning in the cloud. *Future Generation Computer Systems*, 28(1):155–162, 2012. DOI: http://dx.doi.org/10.1016/j.future.2011.05.027.

[9] A Kierzkowski and T Kisiel. Simulation model of logistic support for functioning of ground handling agent, taking into account a random time of aircrafts arrival (2015) icmt 2015-international conference on military technologies 2015, DOI: http://dx.doi.org/10.1109/MILTECHS.2015.7153694.

[10] Moataz Kilany, Aboul Ella Hassanien, and Amr Badr. Accelerometer-based human activity classification using water wave optimization approach. In *2015 11th International Computer Engineering Conference (ICENCO)*, pages 175–180. IEEE, 2015. DOI: 10.1109/ICENCO.2015.7416344.

[11] Niklas Kohl, Allan Larsen, Jesper Larsen, Alex Ross, and Sergey Tiourine. Airline disruption managementâperspectives, experiences and outlook. *Journal of Air Transport Management*, 13(3):149–162, 2007. DOI: http://dx.doi.org/10.1016/j.jairtraman.2007.01.001.

[12] A Krieger and J Simon. Autonomic resource management with support vector machines. In *Proc. of the Lyon Conference*, 2011. DOI: https://doi.org/10.1109/Grid.2011.28.

[13] Sajib Kundu, Raju Rangaswami, Ajay Gulati, Ming Zhao, and Kaushik Dutta. Modeling virtualized applications using machine learning techniques. In *ACM Sigplan Notices*, volume 47, pages 3–14. ACM, 2012. DOI: 10.1145/2151024.2151028.

[14] Kristina Marintseva, Gennadiy Yun, and Sviatoslav Kachur. Resource allocation improvement in the tasks of airport ground handling operations. *Aviation*, 19(1):7–13, 2015. DOI: https://doi.org/10.3846/16487788.2015.1015291.

[15] Dorian Minarolli and Bernd Freisleben. Cross-correlation prediction of resource demand for virtual machine resource allocation in clouds. In *2014 Sixth International Conference on Computational Intelligence, Communication Systems and Networks*, pages 119–124. IEEE, 2014. DOI: https://doi.org/10.1109/CICSyN.2014.36.

[16] Ammar Mohammed and Rania Kora. Deep learning approaches for arabic sentiment analysis. *Social Network Analysis and Mining*, 9(1):52, 2019. DOI: https://doi.org/10.1007/s13278-019-0596-4.

[17] RS Mohana. A position balanced parallel particle swarm optimization method for resource allocation in cloud. *Indian Journal of Science and Technology*, 8(S3):182–188, 2015. DOI: 10.17485/ijst/2015/v8iS3/60501.

[18] Miguel Mujica Mota and Catya Zuniga Alcaraz. Allocation of airport check-in counters using a simulation-optimization approach. In *Applied Simulation and Optimization*, pages 203–229. Springer, 2015. DOI: http://dx.doi.org/10.1007/978-3-319-15033-8_7.

[19] Silvia Padrón, Daniel Guimarans, Juan José Ramos, and Salma Fitouri-Trabelsi. A bi-objective approach for scheduling ground-handling vehicles in airports. *Computers & Operations Research*, 71:34–53, 2016. DOI: http://dx.doi.org/10.1016/j.cor.2015.12.010.

[20] Kanchana Rajaram and MP Malarvizhi. A model for predicting resources for on-premise applications. In *2017 International Conference on Computer, Communication and Signal Processing (ICCCSP)*, pages 1–6. IEEE, 2017. DOI: 10.1109/ICCCSP.2017.7944070.

[21] Kanchana Rajaram and MP Malarvizhi. Utilization based prediction model for resource provisioning. In *2017 International Conference on Computer, Communication and Signal Processing (ICCCSP)*, pages 1–6. IEEE, 2017. DOI: 10.1109/ICCCSP.2017.7944099.

[22] Yuan Ren and Guangchen Bai. Determination of optimal svm parameters by using ga/pso. *JCP*, 5(8):1160–1168, 2010. DOI: http://dx.doi.org/10.4304/jcp.5.8.1160-1168.

[23] Yuan Shen. Virtual resource scheduling prediction based on a support vector machine in cloud computing. In *2015 8th International Symposium on Computational Intelligence and Design (ISCID)*, volume 1, pages 110–113. IEEE, 2015. DOI: https://doi.org/10.1109/ISCID.2015.303.

[24] Alaa Tharwat, Aboul Ella Hassanien, and Basem E Elnaghi. A ba-based algorithm for parameter optimization of support vector machine. *Pattern Recognition Letters*, 93:13–22, 2017. DOI: https://doi.org/10.1016/j.patrec.2016.10.007.

[25] Xingwei Wang, Xueyi Wang, Cho-Li Wang, Keqin Li, and Min Huang. Resource allocation in cloud environment: a model based on double multi-attribute auction mechanism. In *2014 IEEE 6th International Conference on Cloud Computing Technology and Science*, pages 599–604. IEEE, 2014. DOI: https://doi.org/10.1109/CloudCom.2014.103.