

# Optimized Planning of Resources Demand Curve in Ground Handling based on Machine Learning Prediction

**Maged Mamdouh\***

Department of Computer Science Faculty of Graduate Studies for statistical Researches, Cairo University, Egypt

\*Corresponding Author E-mail: [magedmamdouh@pg.cu.edu.eg](mailto:magedmamdouh@pg.cu.edu.eg)

**Mostafa Ezzat**

Associative Professor, Department of Computer Science Faculty of Graduate Studies for statistical Researches, Cairo University, Egypt

E-mail: [mostafa.ezzat@cu.edu.eg](mailto:mostafa.ezzat@cu.edu.eg)

**Hesham Hefny**

Department of Computer Science Faculty of Graduate Studies for statistical Researches, Cairo University, Egypt. Vice Dean

E-mail: [hehefny@ieee.org](mailto:hehefny@ieee.org)

Received: 29 July 2020; Revised: 11 August 2020; Accepted: 23 August 2020; Published: 08 February 2021

**Abstract:** Determining the resource requirements at airports especially in-ground services companies is essential to successful planning in the future, which is represented in the resources demand curve according to the future flight schedule, through which staff schedules are created at the airport to cover the workload with ensuring the highest possible quality service provided. Given in the presence of variety service level agreements used on flight service vary according to many flight features, the resources assumption method makes planning difficult. For instance, flight position is not included in future flight schedule but it's efficacious in the identification of flight resources. In this regard, based on machine learning, we propose a model for building a resource demand curve for future flight schedules. It is divided into two phases, the first is the use of machine learning to predict resources of the service level agreement required on future flight schedules, and the second is the use of implement a resource allocation algorithm to build a demand curve based on predicted resources. This proposal could be applicable to airports that will provide efficient and realistic for the resources demand curve to ensure the resource planning does not deviate from the real-time resource requirements. the model has proven good accuracy when using one day of flights to measuring deviation between the proposed model predict demand curve when flights did not include the location feature and the actual demand curve when flights include location.

**Index Terms:** Machine Learning, Support Vector Machine, Resource allocation, Services Level Agreement.

## 1. Introduction

This work is concerned on-ground services in airport case study (Cairo airport) using historical flights one year of flight. The proposed work has made to support ground services companies to allow plan their resources demand curve (workload) in an efficient way, this trend to use machine learning by support vector machine to predict flight services level agreement (SLA) properly which use to build demand curve by resource allocation algorithm. Modern airports are transportation centers that provide service with a large number of passengers and aircraft every day, ground handling considers one of these tasks of the airport which provides aviation services [1]. Among many activities, Ground handling represents an important role that participates in the activity and safety of air transport [2], although ground handling is a part of services, it is can lead to huge supplemental costs for airports and airlines, because delayed flights and passengers (3).Ground handling process can be defined as activities occurring into terminal and ramp as a direct preparation of flights during waiting position between the time aircraft arrival and aircraft departure with Specifying required of functions for such services if arrival flight or departure or turnaround [3-5].

In this paper, we consider the problem of planning resources by resource demand curve during a working day at an airport in order to optimize the resources demand curve by predicting the realistic resources involved. The planning of the resource demand curve considers (workload) is a very critical first phase to build this required staff schedule. In common

practice Staff scheduling models build upon a demand curve represented in workloads[6]. Ground staff scheduling is a very complex task. Planning frequently involves several hundred Employees, several thousand work tasks per week and multitudes of constraints[7]. The better way to improve operational efficiency is resource demand planning of airports [3], many airports help airport management and planners by provides developed decision support tools based on advanced simulation-based methods or operations research techniques[8]. Sometimes, this intelligent planning software not implemented in the airport, even it exists, often has a drawback is not appropriate to the realities of that particular ground handling company using it.

In general, the ground handling agents are seeking optimal planning to provide an appropriate number of personnel and equipment resources on flight services although conditions mandatory by contract airlines represented in service level agreement (SLA), which represent sorts of resources and task duration used according to flight features as (airlines, aircraft type, orientation, location)[5, 9, 10]. But many sides challenging makes planning resources more difficult, the nature of flight location often reflects the different restriction of services level agreement resources, besides, that flight location is unknown except in operation day, and not stable on the same flight number every time it is may be changed to (hardstand\_ tube stand), therefore flight changes their number of tasks and task duration even if have same features (flight number, aircraft type, orientation)[5]. Hence the problem is very difficult planning the resource demand curve because the future flight schedule does not include location. The most Ground staff planning and scheduling studies not concerned about location feature, in [7] Considered the task generation (SLA) represents Engagement standards rules matching on specific flight events by (airline, routings, aircraft type, intervals of baggage loads. Some studies illustrated in related work, have suggested enhancing staff shift plans based on the demand curve build depends on the assumption resources, but its better enhances demand curve before.

We interest in predicting SLA for flight schedule flight based on future flight schedule features which not include location. To specified required resources rather than resource assumption for flight, when existing SLA (number of tasks - task duration) then we can build resource demand curve (workload) by using Resource Allocation Problem (RAP). The resource allocation and leveling problems are well known from project planning as which creates a workload as smooth as potential based on (number of tasks - task duration) the (RAP) works on transform these tasks demand to the workload demand curve [11]. In common, the potential planning period will be concise as the day of operation approaches, especially planning for the day of operation will affect the day itself, the basis of the resources demand is restricted to flights that generated demand curves to keep the features of the airport's flight schedule[12]. Workloads are usually visualized as a demand curve, i.e. a histogram of parallel tasks at each time of the planning horizon, as each work task requires many staff members, this representation allows for easy analysis of workforce requirements, times of high workloads are easily analyzed, and planners are enabled to assess suitable shift duties covering more than one demand peak[7]. The demand curve referred to (forecast) demand or workload can be calculated by several simple methods, such as the demand aggregated into average capacities, such as for each time slot  $t$  integer value " $dt$ " is workload required for this time, Fig.1. Shows the workload demand curve for a day of planning given many required workers during a day. For each time slot from 0 (midnight) to 24 (midnight), the number of required workers is specified. The demands curve can use demand to represent individuals as " $x$  demand resource per hour", for all resources in a similar manner to clarify the task-based workload [4].

The aim of this paper is to highlight the advantages of evolutionary machine learning techniques prediction in optimizing and solving resource allocation problems in planning ground handling resources in airports. More specifically, a support vector machine (SVM) is proposed to predict flights resource required based on service level agreement (SLA) which based on it creates demand curve by resource allocation algorithm. The important contribution of this work is the way that the proposed approach provides an accurate demand curve of workload for resources required for flights schedule. The paper is organized as follows: Section 2 Related work, Section 3 Approach model, Section 4 result discussion and Section 5 conclusion.

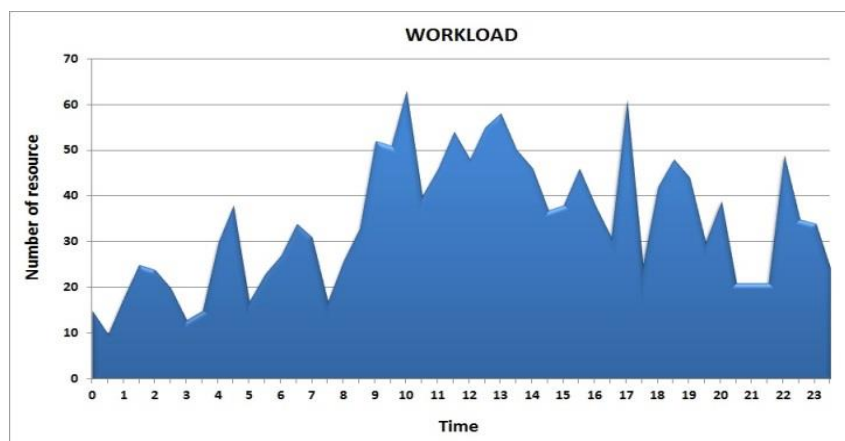


Fig.1. Workload demand curve

## 2. Related work

This section reviews the related work, which contains two parts: The first part is about demand curve solve problems, valuable and estimated, and the second part is the support vector machine to predict SLA.

### 2.1. Demand curve valuable and estimated

The author in [13] formulate task-level shift planning as the problem by outline a branch-and-price solution methodology for the solution of real-world, that to aims creating cost-minimal shift duties set to fit covering a specific set of work tasks. Implemented Linear programming relaxation to measure results gets average gap 0.08% to the integer solution values and found the optimal. The author stated that used a demand curve representation of workloads is not always appropriate in ground handling. Nevertheless, the workload used has occurred from specific tasks and periods needed to the flight schedule, which joining with specific volume and type of resource requirements according to different flights. This study could have been to achieve accurate to more optimal shift planning if built on an accurate demand curve by effective predict tasks rather assumption. The author in [7] solves the presence of the shift design problem has given appropriate demand curve representation based on the task-based workload superposition of planned in flights, utilizing a leveling procedure that combines portion from resource-leveling to vehicle routing. It is a wide study but we concentrate on enhanced demand curve by approach objective function (CP-based re-optimization and repeated relaxation) to obviate acute peaks demand curve resources which very complicate for staff to covering it. The author relies on that tasks can occur movable within given boundaries time during flight service. This is premature in the presence of our study, in our study we seeking to ensure these tasks validity created at first. The author stated in [14] the problem which appears when the demand curve contains a high-level task (peak), this assures the need to look at both surplus (overstaffing) and staff shortage (understaffing), noted these different aircraft types and location types are reflected different resources required, so approached handle a hetero-generous demand and workforce using several demand curves to model this approach. The algorithm is able to generate efficient solutions for problem instances, but besides tasks assumption which main related work problems, also relies on a workforce may have multiple skills or training in several fields which extendibility to do different tasks partially overlapping areas, this does not guarantee its presence in other ground handling companies' policies. As with the study [15] The concept of demand was based on the workers who can do multiple skills that can be implemented within one shift, considers the problem of work shift scheduling which sub-editors on-demand resulting from activities, tried to find a set of work shifts that fits the demand curve and minimizes the overall cost. These multiple skills allowed implementing his approach using a large neighborhood search. This study [16] approaches the workforce planning process to achieve the required resources with time-dependent demand for check-in agents counters at airports, for guarantee covering flight schedule services contracts. Formulation binary linear programming is developed to a fortnightly tour scheduling problems with flexible employee contracts. When the model analyzed based on three real-world demand scenarios it generated convenient tours with regard to employee preferences. We just interested in how estimates workload used arrival times (flight schedule time) of passengers at check-in counters by given features of flight, departure time, passenger number, destination, etc. Despite this study related to one function represented in check-in counter employees in halls, it is another issue not related to ground service resources that related to the type of flight location in the airport area, which determines the type of resources will be executed or not. As with the study [17] the aircraft refueling activities services is provided to the plane according to flight schedule no matter the type of flights location, so the author can optimized planning resources to aircrafts' refueling in airports based on the resource leveling problem represented in the demand curve, which minimized the total cost of unused resources over time. The approach provides a mixed-integer mathematical formulation that solved with the CPLEX, increases the operators' productivity of the for airplane refueling.

In this subsection, we concern about literature that related to the demand curve optimized which affecting on staff shift plan. We concluded in all studies the demand curve is relay on hypothetical resources according to the flight schedule, this is not based on an effective accurate method.

### 2.2. Support vector machine to predict SLA

The prediction of services level agreement is addressed many in the cloud computing domain, which service providers gain benefits by models establishment based an application service-level agreement [18]. The support vector machine model can be highlighted between a set of machine learning techniques with SLA prediction. with a view to meet Service Level Agreement [19] represented in requirements of virtual machine accurate (VM), is constructed a model comparative for prediction using linear regression (LR), neural networks (NN), and support vector machine, by using the metrics throughput and response time can the model predict future cloud resource requires. SVM gets better accuracy than other techniques when measured accuracy were by MAPE, RSME the latest value is 16.84, 12.21, in PRED is .84 it's the highest technique, even when adopted response time accuracy with prediction interval 12 minute to prediction model testing the SVM is the most stable of prediction 14.17 to 14. 33.

According to service level agreements the author in [20], propose support vector machine to predict the time series of used resources, with considering cross-correlation between resources in multiple of VMs that in a multi-tier application at the same time, which helped model Automatic Proactive Resource Allocation (APRA) to reduce resource

costs in cloud computing. The author highlighted the SVM advantages when integrated for proposed model like, avoiding over-fitting leads to Better generalization, applicable to non-stationary processes and non-linear models, SVM model get the global optimum revers artificial neural networks get local minima.

The SVM model toward resource prediction [21, 22] is used to build a prediction model that can handle the changes that occurred in the cloud platform environment, addition in the second paper via a global knowledge base in advance, SVM model build for agents' behavior learning on the services, by experimental results in both papers which performed in real-time, SVM has got a prediction of demanded virtual resources high accuracy when evaluate with three different applications on real cloud systems, a high performance computing software ,a brokerage system, ,and a web server, Also in first paper the SVM model compared to other models that execute in the same problem with task size and duration execution time (ant colony algorithm- particle swarm optimization algorithm) the SVM has made high accuracy of virtual resource prediction and the real-time performance.

In the wireless network domain [23] the learning systems consist of an SVM and deep belief networks (DBNs) approach to solve the problem of capacity optimization of resources include features such as power control, link scheduling, and flow allocation, used SVM especially for classified each link to determine the completion status of power off or maximum power exhaustion, by results for simulation based on automated learning, the result obtains a rough solution to the program. In this study, The SVM in no position compared to other techniques but it is integrated with deep belief networks (DBNs) to achieve the optimal solutions that nonlinear sub problem. The performance was compared of the proposed algorithm with that of three other algorithms the genetic algorithm (GA) the greedy algorithm (Greedy) random algorithm (Random), the mean achieved capacity with 95% confidence intervals, where for each run the capacity is studied as in order to achieve after 100 Iterations.

In airport ground handling domain [5] proposed a systematic comparison of machine learning techniques, nominated the SVM model to meet the SLA requirements by resource prediction for each flight. Despite more restrictions taken for services according to flight features represented in services level agreement, however, the model accuracy achieved satisfy results. The model can predict all resources required to flight with flight schedule three features (flight number, orientation, aircraft type) with accuracy 85%. This study considers as the start point our approach, especially it's the same our data set in Cairo airport, we will use this model but for predicting each function as alone, which lead to multi and different accuracies the analysis will explain later, whatever, the uses of prediction (number of tasks – task duration) as a part of features used to build resource allocation algorithm and product the demand curve.

The earlier studies investigated a successful support vector machine for SLA prediction in different domains. In this context, we approach a new model optimizing the demand curve based on machine learning by predict SLA.

### 3. Proposed Approach

The proposed approach provides for ground services companies to define the appropriate number of workers every day for covering the tasks required to serve future flight schedules daily. At first, The SVM model training on the specific function required then the model receives one day of future flight schedule as input and extracts resource demand represented in demand curve as output. Fig.2 shows the architecture of approach model, the model was built upon three phases:

#### 3.1. Future flight schedule

Usually, the agents have future flights schedule in advance; most airports have arrival flights followed by departure flight called a bank, this results in massive peaks on airport resource demand [12]. Table 1 shows a sample of the flight schedule, will be using the specific day flight schedule as inputs to the model to planning it. In most studies assume resources for flights' schedule to build demand curves, are usually not adequate in real-time [7]. In some cases, tasks may be associated with flight locations [24]. The location of flight is one of the important factors to determine the resources provided, there are types like a hard stand or tube stand [1, 4]. a flight can be with the same flight number, aircraft type, orientation, and flight time, stands on a different type of familiar location for any reason, that motivated give us the for using machine learning to predict resources due to their changed location. So, we approach the support vector machine model to predict flight resource, to be more precise SLA for each flight during the specific day.

Table 1. Sample of flight schedule

Flight Date	Flight No	Airline	Orientation	AC Type	Routing	Schedule Time
01/05/2020	SV 309	Saudi Air	Turnaround	333	JED	00:50
01/05/2020	EY 651	Etihad	Turnaround	332	AUH	01:05
01/05/2020	MS 901	Egypt Air	Departure	738	DXB	02:25
01/05/2020	NP 114	Nile Air	Arrival	321	YNB	05:30
01/05/2020	RJ 508	Royal Jordanian	Departure	318	AMM	07:00



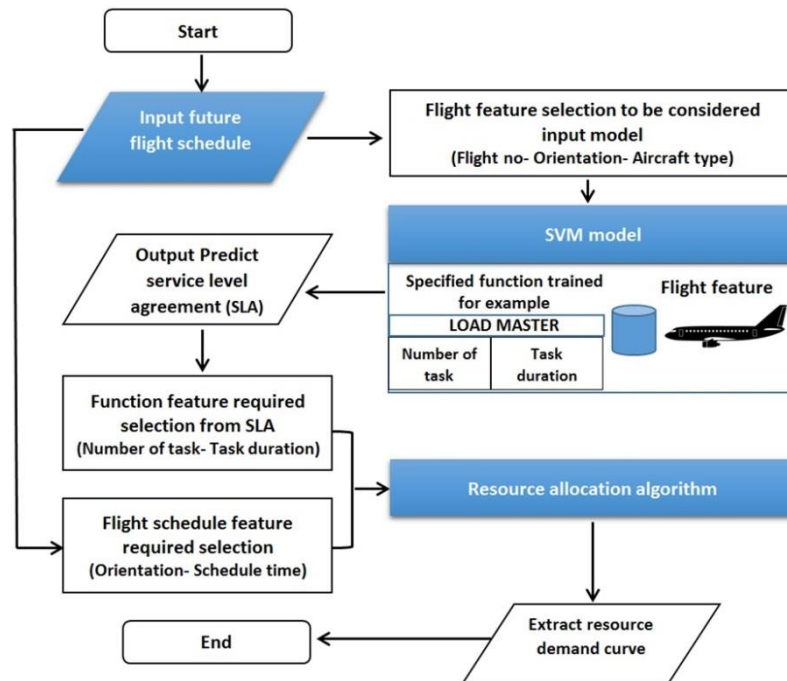


Fig.2. Architecture of approach model

### 3.2. Service level agreement prediction by support vector machine

#### A. Service level agreement

The availability of ground handling services is a concern relating to airline quality and policies [9]. The International Air Transport Association has placed those standards, called service level (LOS)[10]. Accordingly, while the service level agreement is very critical because it contains the tasks and their periods to execute, it possibly leads to build a fault demand curve. Despite many features can be drawn from flight schedule as (airline, orientation, aircraft type) to determine the SLA, but there is important factor were still unknown and not included in future flights schedule this is flight (location) or (position).

#### B. Data set

The dataset, in brief, contains a historical flight schedule for approximately 1 year includes on 97123 past flights taken already a location in the Cairo airport, where each flight joins with their service level agreement for functions required based on features of it (airline, orientation, aircraft type, location) which enough to determine the resources required SLA for each function. The dataset propose illustrated in [5] with details, especially will be using the same data set of historical flights, but in the current proposal is a new form of data outputs and prediction each type of functions to build demand curve, that will have other measures of accuracy. The resources contain all types of function whether (humans or equipment) includes (number of tasks, task duration) this is to be ready for preparation for the next step when building a resource allocation algorithm, Table 3 shows the sample of dataset the “NA” variable it means not applicable.

#### C. Support vector machine

The support vector machine is recommended for use as a prediction model, have been used successfully to predict SLA in different domains, especially it is the most successful supervised algorithms implemented [25]. we will use this model as a starting point for predict resources for each function to extract (number of tasks, task duration) as output to build demand curve by implement resource allocation algorithm for all-day flight. The support vector machine classification of multi-dimensional implemented based on the library of SVM [26]. The function has many valuable parameters that affect training and testing accuracy. The SVM model accuracy is very influenced by parameters gamma and cost values, the gamma parameters affect SVM that can lead to classifying new data correctly [27]. In the kernel, functions include ('linear', 'poly', 'rbf', 'sigmoid', 'precomputed'), In this work uses 'rbf' radial basis function is used extensively and proposed in several implementation studies [28]. The parameter Cost is value inversely proportional reliable balance between the margin maximization and the minimization of error. The Cost value is inversely proportional to margins. When the value is high the margin width will be tight and lead to overfitting, conversely, when the value is small the margin width will be big and lead to underfitting [29]. The parameter has been tuning and Table 2 shows SVM model parameters, Fig.3. shows the Architecture of the support vector machine implement model.

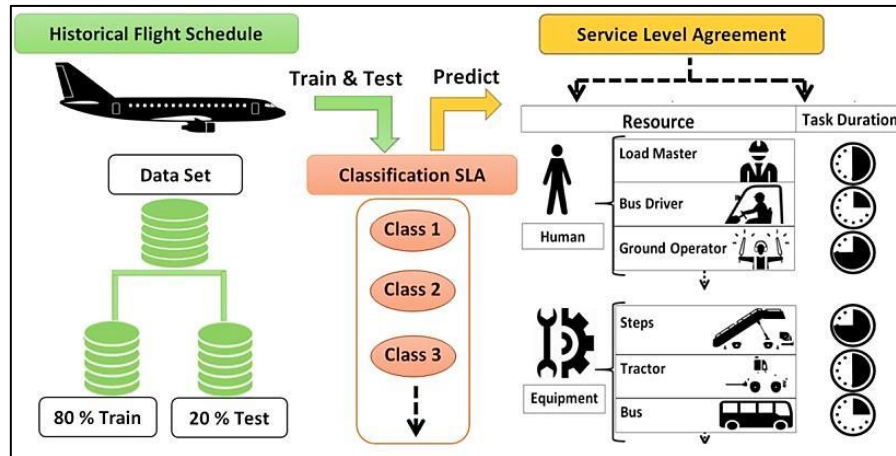


Fig.3. Architecture of the support vector machine

Table 2. SVM model parameters

Parameter	Value
cost	80
cache_size	200
class_weight	None
coef0	0.0
decision_function_shape	'ovr'
degree	3
gamma	0.05
Kernel	rbf
max_iter	-1
probability	False
random_state	None
shrinking	True
degree	3
tol	0.001
verbose	False

#### D. Training and testing SVM with the data set

The same SVM model has been trained and testing for each function as alone, by using three features of flight as inputs (flight number, aircraft type, orientation) as same as a future flights schedule features, to predict (number of tasks - task duration) for each function. The training implemented in 3 scenarios, scenario 1 includes all Airlines contains (foreign airlines - national airline) 97192 of rows, scenario 2 includes only national airline from dataset contains 71155 of rows represent 70 % of operation flights, scenario 3 includes only foreign airlines from dataset contains 25758 of rows. Clarifying the results will become clearer later in the results and discussion section. The SVM training process has been used the Scikit\_Learn library and specifically, the (train size = 0.80) and (test size = 0.20) divided into 80% for training and 20% for testing. The (Train / Test Split) tool used that training and test data set can be broken up well, there are three flight description features to predict number workers or tasks duration in minutes for each function alone. all data features have been encoded to be suitable for training and testing, excepted flight resource SLA features because they are already represented numbers whether number workers or tasks duration in minutes.

#### 3.3. Built resource allocation algorithm based on prediction

Task-based demand, in this case, demand is obtained from lists of individual tasks to be performed, tasks are usually defined in terms of a starting time and duration, or a time window within which the task must be completed, and the skills required to perform the task [24]. When the tasks with time windows have existed, then the problem of undefined demand appears which needs optimization or modeling to transform these tasks demand to the workload demand curve, the resource allocation and leveling problems are well known from project planning as which creates a workload as smooth as potential [11] In that perspective, approach algorithm 1 implemented resource allocation algorithm. Algorithm 1 shows the steps represent by program pseudo-code, Fig.4. Shows the resource allocation algorithm works. Algorithm 1 has main advancing features as follow:

- The required service level agreement of flights schedule is based on prediction from SVM model directly instead of assumptions and expertise methods, through using future flight schedules to the model as new inputs and makes a prediction SLA as output represented in (number of task- task duration), steps (2 to 6). We can connect predictions a fit machine learning model with new Inputs [30].
- The algorithm can be implemented for each function as alone, just by SVM model training previously on the specific function required (number of the task- task duration).
- The algorithm allows obtaining of resource demand represented in the demand curve to predict the number of workers or equipment each hour, steps (60 to 66).

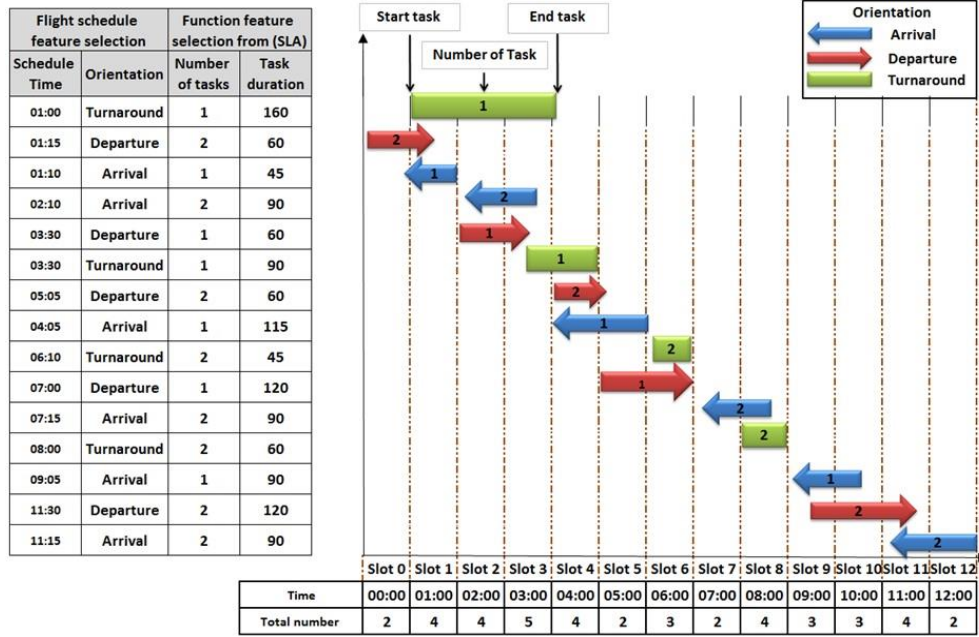


Fig.4. Architecture of resource allocation algorithm

**Algorithm 1** Begin (resource allocation algorithm based on prediction SLA)

```

1- Start
2- Input a table of future flight schedule day
3- Read columns [flight number, Aircraft type, orientation, schedule time]
4- Convert [flight number, Aircraft type, orientation] into the same model value encoder
5- new Input  $\leftarrow$  [flight number, Aircraft type, orientation]
6- new output  $\leftarrow$  prediction SVM model based on new Input for specific function has been trained before.
7- number  $\leftarrow$  first column of new output
8- duration  $\leftarrow$  second column of new output
9- orientation  $\leftarrow$  future flight schedule day [orientation]
10- schedule time  $\leftarrow$  future flight schedule day [schedule time]
11- hour  $\leftarrow$  schedule time. hour
12- minute  $\leftarrow$  schedule time. minute
13- Define array schedule minute, start minute, end minute, start, end  $\leftarrow$  []
14- for i in range length hour [] do
15-   total minute = hour [i] * 60 + minute [i]
16-   Schedule minute [i]  $\leftarrow$  total minute
17- end for
18- for i in range length orientation [] do
19-   If orientation [i] = ['Arrival'] or orientation [i] = ['Turnaround']
20-   then   start minute [i]  $\leftarrow$  schedule minute [i]
21-         end minute [i]  $\leftarrow$  schedule minute [i] + duration [i]
22-   else
23-   then   start minute [i]  $\leftarrow$  schedule minute [i] - duration [i]
24-         end minute [i]  $\leftarrow$  schedule minute [i]
25-   end if
26- end for
27- for i in range length start minute [] do

```

```

28-   Calculate hour ← 0
29-   Calculate minute ← 0
30-   Calculate hour ← integer (start minute [i] /60)*100
31-   Calculate minute ← start minute [i] – integer (start minute [i]/60)*60
32-   Start [i] ← Calculate hour + Calculate minute
33- end for
34- for i in range length end minute[] do
35-   Calculate hour ← 0
36-   Calculate minute ← 0
37-   Calculate hour ← integer (end minute [i] /60)*100
38-   Calculate minute ← end minute [i] – integer (end minute [i]/60)*60
39-   end [i] ← Calculate hour + Calculate minute
40- end for
41- Defined new integer variable slot0, slot1, slot2, slot3..... slot23
42- for i in range length start[] do
43-   If start [i] < 100 and end [i] >= 0
44-     then slot0 = slot0 + number [i]
45-   end if
46-   If start [i] < 200 and end [i] >= 100
47-     then slot1 ← slot1 + number [i]
48-   end if
49-   If start [i] < 300 and end [i] >= 200
50-     then slot2 ← slot2 + number [i]
51-   end if
52-   If start [i] < 400 and end [i] >= 300
53-     then slot3 ← slot3 + number [i]
54-   end if
55-   ..... # doing the same step number 52 to 54 until slot23
56-   If start [i] < 2400 and end [i] >= 2300
57-     then slot23 ← slot23 + number [i]
58-   end if
59- Define new array predict ← [ ]
60- predict [0] ← slot0
61- predict [1] ← slot1
62- predict [2] ← slot2
63- predict [3] ← slot3
64- ..... # doing the same step until predict [23] ← slot23
65- predict [23] ← slot23
66- Create chart histogram set predict [0] to predict [23] to be legend series (resource demand) and from 0 to
    23 to be axis label (time)
End (resource allocation algorithm)

```



Table 3. 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
Schedule 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
Turnaround coordinator	NA	NA	1	1	NA	NA
Turnaround 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 porter	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 minute	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

#### 4. Result and discussion

The results of the approached model can be divided into two phases, the first is main measure prediction SVM model by accuracy method, illustrated in subsection A, and the second method is used R-squared to measure if exists deviation between resources demand curves in case flight resources prediction based on features (flight no - orientation - aircraft type) and in case actual flight resources demand when features completed (flight no - orientation - aircraft type -

location), illustrated in subsection B.

#### 4.1. Model support vector machine of prediction for service level agreement

The appropriate measure for the SVM model is accuracy method to predict SLA (number of tasks – task duration) depends on features of future flight (flight no - orientation - aircraft type) for training and testing. The accuracy method defined in the following equation (1):

$$Accuracy = (TP + NP) = (TP + TN + FP + FN) \quad (1)$$

Where: given a model and an instance, there are four feasible outputs. Two based instances originally positive, true positive (TP) is enumerated when categorized as positive, false negative (FN) is enumerated when categorized as negative. On other hand are two based instances originally negative, true negative (TN) is enumerated when categorized as negative, false positive (FP) is enumerated when categorized as positive.

Table 4 shows the accuracy results ordered according to the great accuracy of 3 scenarios with 2 groups of functions. Regarded 3 scenarios, scenario 1 includes all Airlines flights contains (foreign airlines - national airline) 97192 of rows, scenario 2 includes only national airline flights from the dataset and contains 71155 of rows represent 70 % of operation flights, scenario 3 includes only foreign airlines flights from the dataset and contains 25758 of rows represent 30 %. Regarded functions, the resources to be clear were divided into group A represents general resources and group B represents resources related to the type of location which means the resources will change (number of tasks - task duration) for serves flights according to the type of location. Even more clearly, ground service activities different in service provision for flights according to tube location and ground location, in the tube location, the passengers can walk directly from the station to the flight so no need for steps and buses, and some equipment is automatically already supplied in tube location like the air-conditioning unit and ground power unit [1, 4].

Table 4. Accuracy of all scenarios with all function

Function		Scenario 1		Scenario 2		Scenario 3	
		All airline		National airline		Foreign airlines	
		Train	Test	Train	Test	Train	Test
Group A	Turnaround coordinator	99.99	99.78	NA	NA	99.98	99.74
	Water car & driver	99.91	97.69	99.99	97.65	99.69	99.45
	West car & driver	99.90	98.06	99.99	97.69	99.71	99.52
	Crew car & driver	99.85	99.59	99.92	98.96	99.62	99.51
	Load master	99.80	99.67	99.78	98.81	99.76	99.63
	Conveyor belt	99.80	99.63	99.93	99.13	99.70	99.25
	Tractor & driver	99.78	99.52	99.65	98.99	99.72	99.78
	Ground services operator	99.65	99.26	99.34	99.13	99.73	99.43
	Porter	99.62	99.26	99.86	99.68	99.41	99.44
	High loader & driver	99.52	99.05	99.35	99.25	99.41	99.44
	Cabin clean	99.46	97.53	99.91	97.06	98.52	96.83
	Upper deck	89.30	84.30	95.35	95.48	72.49	53.89
Group B	Bus Departure & driver	75.93	70.70	65.15	58.03	98.28	96.97
	First passenger steps	60.96	55.34	37.58	30.69	97.60	94.71
	Second passenger steps	51.47	47.23	38.63	34.22	96.82	95.31
	Bus arrival & driver	51.37	45.57	32.52	34.22	98.42	97.91
	Ground power unite	48.73	42.20	26.92	21.32	98.20	95.96
	Air condition unite	46.23	41.08	23.19	18.51	97.94	95.35

From this comes the functions divides into 2 groups A, B. Group A general resources represent (Turnaround coordinator, Water car & driver, West car & driver, Crew car & driver, Load master, Conveyor belt, Tractor & driver, Ground services operator, Porter, High loader & driver, Cabin clean, Upper deck). Group B related to location represent (bus arrival, bus departure, first passenger steps, second passenger steps, ground power unit, air-condition unit).

- Scenario 1: in scenario 1 the data set includes all airlines (national airline, foreign airlines) contains 97123 rows it is noted that accuracy is very high when the model trained to predict functions group A, except functions in group B started to decrease at the orange range when starting the accuracy of train 75.93 % to 46.23 %. Notes that they are equipment strongly related to flight location type, especially whether (hardstand, tube stand), which made to find an explanation, therefore a data set has been split to the national airline in scenario 2 and foreign airlines in scenario 3.

- Scenario 2: in scenario 2 the data set includes only national airline contains 71155 of rows which has 70% of operation flight, it also notes that accuracy is very high when the model trained to predict functions group A, except functions in group B started to more decrease at the red range than scenario 1 when starting the accuracy of train 65.15% to 23.19%, that means the model is very confusing to predict group B, we still need an explanation decreased accuracy.
- Scenario 3: in scenario 3 the data set includes only foreign airlines contains 25758 of rows which has 30% of operation flight, it also notes that accuracy is very high when the model trained to predict functions group A and group B, all functions located in green range when starting the accuracy of train range 99.98 % to 96.82 %. Upper deck function in group A is exception has accuracy 72.49 %, however, it is strangely all group B functions have high accuracies as opposed to scenario 1 and scenario2. According to scenarios accuracy in table 4 and an examination into Fig. 5 shows training accuracy and Fig. 6 shows testing accuracy, we can analyze decreased accuracy that happened in Group B from a functions perspective as follows.
- Group A: the model appeared very strongly prediction these functions in scenarios 1, 2 and 3, but the only function has lowest and unsatisfied accuracy is (Upper deck) especially in scenario 3 was 72.49 % in train and 53.89% in test, and explanation for this is that equipment completely related to cargo flights, and in scenario 3 the foreign cargo flights are very few because it is seasonally and sometimes changes their flight number, therefore not allows for the SVM model to predict correctly. Whatever, it is solved in case of getting more foreign flights dataset for training. Overall, the model predicts functions group A perfectly.
- Group B: the model appeared it almost seemed very weak with these functions related to the location in group B, despite the model have trained already on 1 year of flights which got already locations joined with perfect required of resources according to services level agreement, when group B predicted in scenario 1 (national airline - foreign airlines) the accuracy decreased to the range from 75.93% to 46.23% in the train and the range from 70.70% to 41.08% to in the test. When group B predicted in scenario 2 (national airline) decreased to the range from 65.15% to 23.19% the train and the range from 58.03% to 18.51% to in the test. But group B in scenario 3 (foreign airlines) stark contrasts, notes the accuracy when predicting functions, the range from 98.42% to 96.82% in train and in the range from 97.91 % to 94.71% in the test.

That's what indicates the model works well, but the problem lies at the national airline that influence prediction in scenario 1 and 2, their flight can have the same flight number with the same aircraft type with the same orientation at the same time and found this flight stood on a different type of familiar location whether (hard or tube) for any reason. We found the national airline flight numbers always are changing the type of location in coordinate with aviation authorities these are things regarding national airline cost with aviation authorities rather than ground services agent, that means the national airline which has daily approximately 70 % of the operations is changing their flight resources daily with the ground service agent, resulting in weak the model when predicting functions related to the location group B with the--national airline in scenario 1 and 2. This is a special case in Cairo airport not possibly occur in other airports. Even with the lowest predict accuracy in scenario 1, 2 it is can be relay on better than resource assumptions. Otherwise, the model can predict properly with high accuracy and rely upon predicting functions number and task duration for building a resource allocation algorithm properly as in scenario 3.

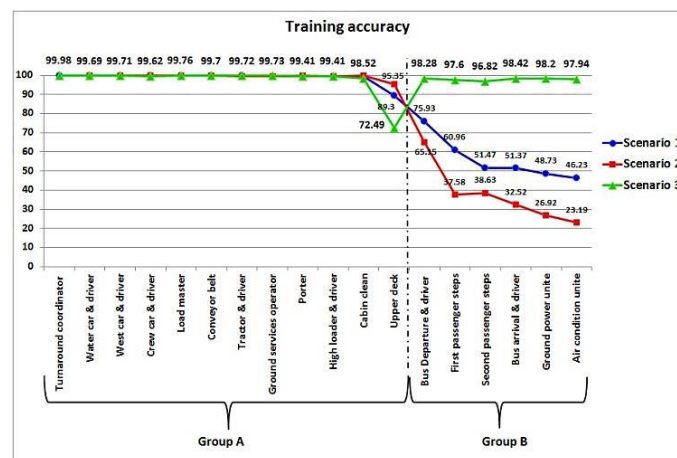


Fig.5. Training accuracy

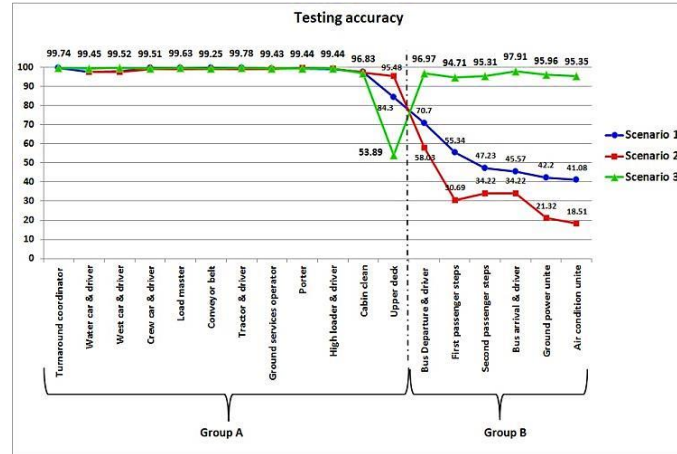


Fig.6. Testing accuracy

#### 4.2. Build a resource allocation algorithm based on the prediction model

The implementation of the resource allocation algorithm program for each function based on SVM prediction is considering the research target, but the model quality is the greater target, so we can be trust on model by the measured deviation between two resources demand curves (predict - actual) by Root squared ( $R^2$ ) method, that implemented when using one day previous of flights joined with actual resources required based on full features even location (flight no - orientation - aircraft type - location) and compared to when using the same day of flights as future flights schedule to be as input into the approaching model to predict resources based on features (flight no - orientation - aircraft type). R-squared method formulated as  $y$ ,  $\hat{y}$  refer to actual and predict demand respectively, R-squared computed in equation (2):

$$R^2_{Classical} = \frac{\text{var } \hat{y}}{\text{var } y} \quad (2)$$

Table 5. R-squared accuracies for group A, B with scenario 1

Function		Scenario 1 All airline
		R-squared
Group A	Turnaround coordinator	0.95
	Water car & driver	0.97
	West car & driver	0.97
	Crew car & driver	0.99
	Load master	1.00
	Conveyor belt	1.00
	Tractor & driver	1.00
	Ground services operator	0.99
	Porter	0.99
	High loader & driver	0.72
	Cabin clean	0.96
	Upper deck	0.83
Group B	Bus Departure & driver	0.85
	First passenger steps	0.82
	Second passenger steps	0.89
	Bus arrival & driver	0.51
	Ground power unite	0.79
	Air condition unite	0.81

$R^2$ , also recognized as the coefficient of determination it's supplemented with equivalent ways e.g., correlations, variance. The general formula explains metric of which is the ratio of the variance in the predicted values in model is  $\hat{y}$ , to the variance in the actual  $y$  [31]. We will show at first all resources group A, B (predict - actual) demand curves when using all airline datasets in scenario 1, as a target required to implement. The workload representing in Fig.7 shows group A, B with scenario 1 means all airlines with all resources. Table 5 shows R-squared for group A, B with scenario 1 to measure the deviation between (predict - actual). they predict demand curve results for group A is expected to more

be close to the actual demand curve according to the high group A accuracy in scenarios 1,2 and 3 in the SVM model in table 4. This really happened when we measured the deviation between them. It is clear the R-squared range starts 95% to 100%, except two functions are (Upper deck) is 83% and (High loader & driver) is 72%. Since the model implementation on scenario 1 is our object because it includes on all airlines (national airline – foreign airlines) and the functions have satisfactory R-squared accuracy, so we just do it implemented only group A with scenario 1. the group B with scenario 1 reflects acceptable accuracy for all functions when R-squared in range 79 % to 89% except only one function (Bus arrival & driver) was 51%. After dividing the data set into two scenarios, scenario 2 national airline and scenario 3 as foreign airlines as mentioned before.

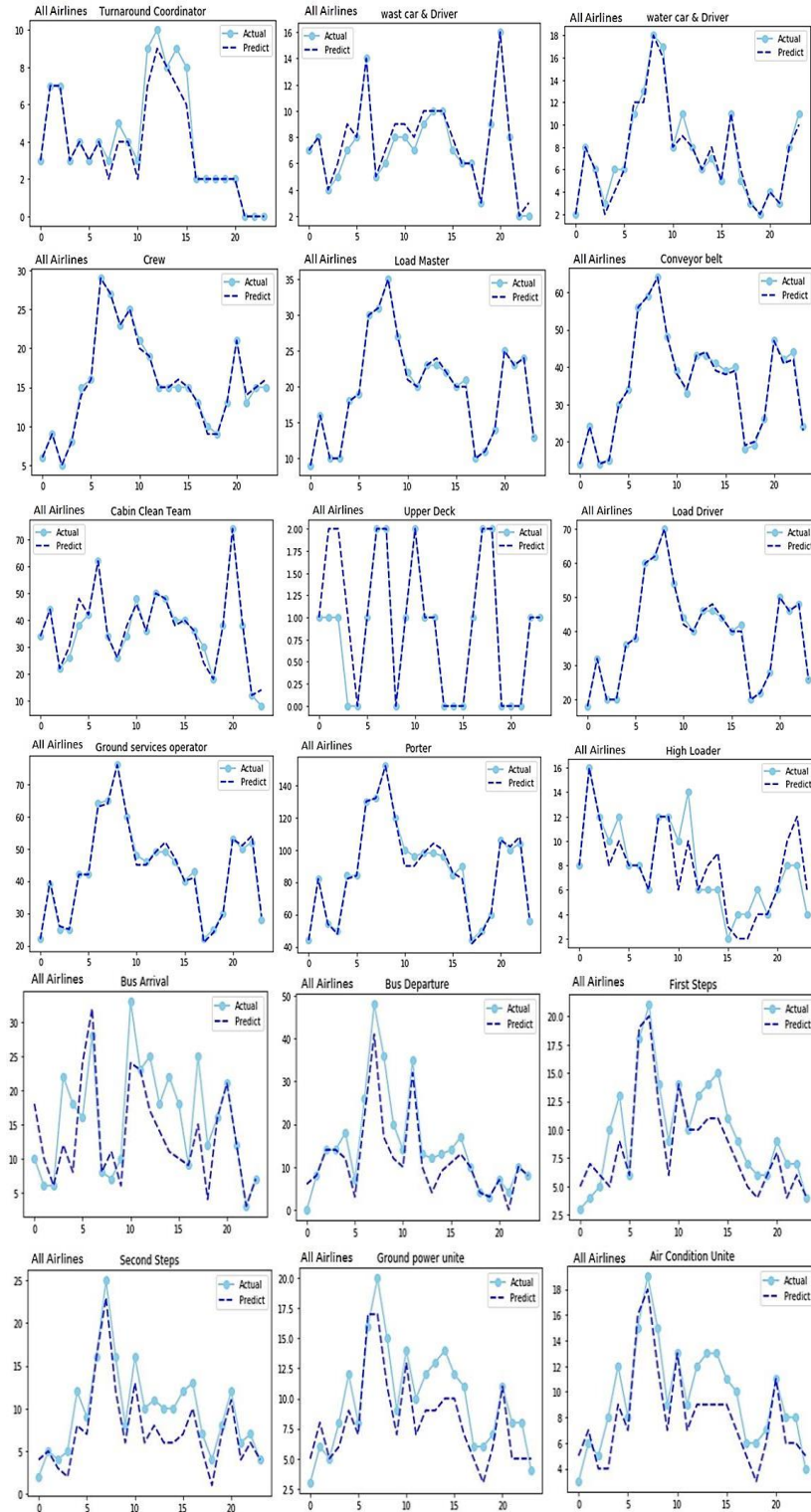


Fig.7. Group A, B with Scenario 1



We will implement the model on group B with scenarios 2 and 3 to see the amount of deviation (predict - actual) of demand curves. Fig.8 shows group B with scenario 2, Fig.9 shows group B with scenario 3. Table 5 shows R-squared for group B with scenario 2 and 3. In scenario 2 we found the model cannot predict group B functions well using a national airline dataset when R-squared in the range 36 % to 90 %. Therefore, this reflects in the resource allocation algorithm highlighted the deviation between prediction resources and actual resources, which represented in figure 8 can observe the deviation between two curves the predict resources and actual required resources. This is not a weakness of the model, this is due to national airline flight instability have taken locations type whether hardstand or tube type, although the flights are had the same features (flight number, aircraft type, orientation) but take different location type, it's regarding cost between national airline with aviation authorities who distribute flights on locations, this is an exception in the Cairo airport unlikely happens in other airports. Otherwise, in scenario 3 in the foreign airlines dataset when the model predicts group B functions that related to the location were have high accuracies when R-squared in the range 91 % to 98 % except for only one function (Second passenger steps) was 67%. In figure 9 we can observe a perfect match between two curves (predict - actual) for required resources. Scenario 3 proves the model is strongly cooperative through measure accuracy with group B.

Table 6. R-squared accuracies for group B with scenario 2 and 3

Function		Scenario 2	Scenario 3
		National airline	Foreign airlines
		R-squared	R-squared
Group B	Bus Departure & driver	0.36	0.97
	First passenger steps	0.75	0.94
	Second passenger steps	0.90	0.67
	Bus arrival & driver	0.41	0.98
	Ground power unite	0.77	0.91
	Air condition unite	0.78	0.91

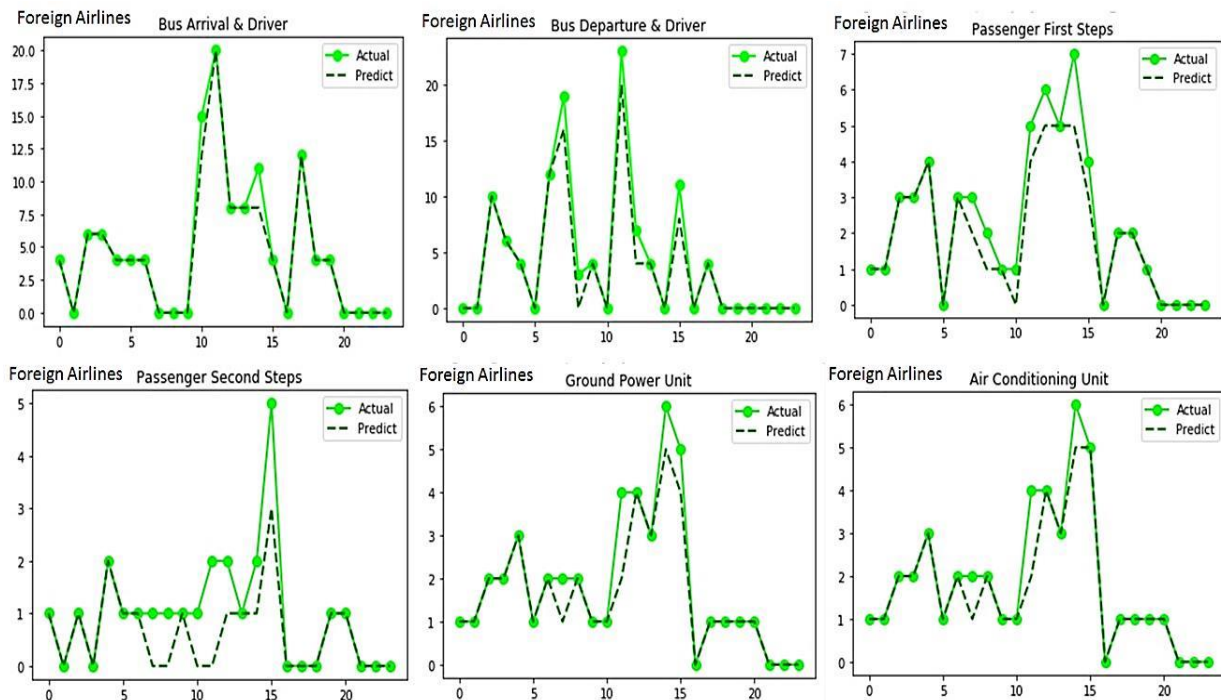


Fig.9. Group B Scenario 3

## 5. Conclusion

This work concentrating on ground handling activities in airports. Using real historical Cairo airport data contains flights engaged services level agreement (number of tasks, task duration) according to four features (flight number, aircraft type, orientation, location). We proposed a mixed model that includes a resource allocation algorithm based on support vector machine to resources prediction for build successful planning to the future resources demand curve according to the future flight schedule, rather than the resources assumption method to build demand curve, whether in manually methods or intelligent planning software methods. Three different scenarios are analyzed when model

predicate functions, scenario 1 include all airline flights, scenario 2 includes national airline flights, and scenario 3 includes foreign airline flights. the quality of the model approach measured in two phases, the first is the SVM model trained on historical flights with their functions required based on four flights features (flight number, aircraft type, orientation, location), and measured by prediction accuracy for each function required to flights based on only three future flight schedule features (flight number, aircraft type, orientation). Although, the flight location not exist it is accuracy refers to rely on the SVM model especially in Scenario 3 with foreign airlines was in range 94.71 % to 99.78% in the test for all prediction function except only on function (Upper Deck) was 53.89% and mentioned reason. The second is the resource allocation algorithm when used this predicted functions to build a future resource demand curve based on future flight schedule that three features, and compare it with the actual resource demand curve for the same day of the flights after taken location with their resources required according to four flights features, then we measured deviation by methods R-squared between two demand curves (predict - actual). The R-squared accuracy in scenario 1 which is all airlines and targets for implementation has acceptable accuracy with functions in the range 72% to 100% except one function (Bus arrival & driver) was 51%.

Finally, the model with acceptable accuracy proves that gives us a very clear direction to machine learning implementation for predict planning flight resource demand curve rather than assumption methods, A future work can be devoted to must be resource planning systems implemented at airports enhance with machine learning to guarantee reliable demand curve for leads to successful staff schedules.

## References

- [1] Marintseva, K., G. Yun, and S. Kachur, Resource allocation improvement in the tasks of airport ground handling operations. *Aviation*, 2015. **19**(1): p. 7-13.
- [2] Fitouri-Trabelsi, S., et al., Integrated decision making for ground handling management. 2015.
- [3] Padrón, S., et al., A bi-objective approach for scheduling ground-handling vehicles in airports. *Computers & Operations Research*, 2016. **71**: p. 34-53.
- [4] Clausen, T. and D. Pisinger, Airport ground staff scheduling. 2011: DTU Management Engineering.
- [5] Mamdouh, M., M. Ezzat, and H.A. Hefny, Airport resource allocation using machine learning techniques. *Inteligencia Artificial*, 2020. **23**(65): p. 19-32.
- [6] Tien, J.M. and A. Kamiyama, On manpower scheduling algorithms. *SIAM review*, 1982. **24**(3): p. 275-287.
- [7] Herbers, J. and J. Hromkovic, Models and algorithms for ground staff scheduling on airports. 2005, Fakultät für Mathematik, Informatik und Naturwissenschaften.
- [8] Justesen, T.F., Allocation of Ground Handling Resources at Copenhagen Airport. DTU Management Engineering, 2014.
- [9] Kierzkowski, A. and T. Kisiel. Simulation model of logistic support for functioning of ground handling agent, taking into account a random time of aircrafts arrival. in *International Conference on Military Technologies (ICMT) 2015*. 2015. IEEE.
- [10] Mota, M.M. and C.Z. Alcaraz, Allocation of airport check-in counters using a simulation-optimization approach, in *Applied Simulation and Optimization*. 2015, Springer. p. 203-229.
- [11] Brucker, P., et al., Resource-constrained project scheduling: Notation, classification, models, and methods. *European journal of operational research*, 1999. **112**(1): p. 3-41.
- [12] Kohl, N., et al., Airline disruption management—perspectives, experiences and outlook. *Journal of Air Transport Management*, 2007. **13**(3): p. 149-162.
- [13] Herbers, J., Representing labor demands in airport ground staff scheduling, in *Operations Research Proceedings 2005*. 2006, Springer. p. 15-20.
- [14] Clausen, T., A dynamic programming-based heuristic for the shift design problem in airport ground handling. (DTU Management 2010).
- [15] Quimper, C.-G. and L.-M. Rousseau, A large neighbourhood search approach to the multi-activity shift scheduling problem. *Journal of Heuristics*, 2010. **16**(3): p. 373-392.
- [16] Stolletz, R., Operational workforce planning for check-in counters at airports. *Transportation Research Part E: Logistics and Transportation Review*, 2010. **46**(3): p. 414-425.
- [17] Carotenuto, P., et al., Resource planning for aircraft refueling in airport parking area. *Transportation Research Procedia*, 2019. **37**: p. 250-257.
- [18] Nathuji, R., A. Kansal, and A. Ghaffarkhah. Q-clouds: managing performance interference effects for qos-aware clouds. in *Proceedings of the 5th European conference on Computer systems*. 2010.
- [19] Bankole, A.A. and S.A. Ajila. Predicting cloud resource provisioning using machine learning techniques. in *2013 26th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*. 2013. IEEE.
- [20] Minarolli, D. and B. 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*. 2014. IEEE.
- [21] Shen, Y. Virtual resource scheduling prediction based on a support vector machine in cloud computing. in *2015*

- 8th International Symposium on Computational Intelligence and Design (ISCID). 2015. IEEE.
- [22] Niehorster, O., et al. Autonomic resource management with support vector machines. in 2011 IEEE/ACM 12th International Conference on Grid Computing. 2011. IEEE.
- [23] Cao, X., et al., A machine learning-based algorithm for joint scheduling and power control in wireless networks. IEEE Internet of Things Journal, 2018. **5**(6): p. 4308-4318.
- [24] Ernst, A.T., et al., Staff scheduling and rostering: A review of applications, methods and models. European journal of operational research, 2004. **153**(1): p. 3-27.
- [25] Mohammed, A. and R. Kora, Deep learning approaches for Arabic sentiment analysis. Social Network Analysis and Mining, 2019. **9**(1): p. 52.
- [26] Online. sklearn.svm.SVC. 2019; Available from: <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html>.
- [27] Kilany, M., A.E. Hassanien, and A. Badr. Accelerometer-based human activity classification using water wave optimization approach. in 2015 11th International Computer Engineering Conference (ICENCO). 2015. IEEE.
- [28] Ren, Y. and G. Bai, Determination of optimal SVM parameters by using GA/PSO. JCP, 2010. **5**(8): p. 1160-1168.
- [29] Tharwat, A., A.E. Hassanien, and B.E. Elnaghi, A BA-based algorithm for parameter optimization of support vector machine. Pattern Recognition Letters, 2017. **93**: p. 13-22.
- [30] Brownlee, J. Machinelearningmastery. 2019; Available from: <https://machinelearningmastery.com/how-to-connect-model-input-data-with-predictions-for-machine-learning/>.
- [31] Redell, N., *Shapley Decomposition of R-Squared in Machine Learning Models*. arXiv preprint arXiv:1908.09718, 2019.

## Authors' Profiles



**Maged Mamdouh**, is a M.Sc. candidate at the Department of Computer science from the Faculty of Graduate Studies for Statistical Research (ISSR), Cairo University, Egypt. he was born in Egypt, Cairo in 1982. He received B.Sc. of business administration and information system from, Cairo University, Egypt, in 2004. In 2009, he had worked with Egypt Air Ground Services, in Cairo Airport, as a Business intelligence report designer and developer. His research interests are in artificial intelligence, resource allocation, machine learning.



**Mostafa Ezzat**, received a Masters and Ph.D in Computer Science degree Faculty of Graduate Studies for Statistical Research (ISSR), Cairo University, Egypt in 2010 and 2015, respectively. He is currently a faculty Associative Professor of the Department of Computer Science Faculty of Graduate Studies for Statistical Research, Cairo University, Egypt. His research interests are in artificial intelligence and Natural Language Processing.



**Hesham A Hefny**, received the B.Sc., M.Sc. and Ph.D. all in Electronics and Communication Engineering from Cairo University in 1987, 1991 and 1998 respectively. He is currently a professor of Computer Science at Faculty of Graduate Studies for Statistical Research (ISSR), Cairo University, Egypt. He is also the vice dean of graduate studies and researches of (ISSR). Prof. Hefny has authored more than 150 papers in international conferences, journals and book chapters. His major research interest includes: computational intelligence (neural networks – Fuzzy systems-genetic algorithms – swarm intelligence), data mining, uncertain decision making. He is a member in the following professional societies: IEEE Computer, IEEE Computational Intelligence, and IEEE System, Man and Cybernetics.

**How to cite this paper:** Maged Mamdouh, Mostafa Ezzat, Hesham Hefny, "Optimized Planning of Resources Demand Curve in Ground Handling based on Machine Learning Prediction", International Journal of Intelligent Systems and Applications(IJISA), Vol.13, No.1, pp.1-16, 2021. DOI: 10.5815/ijisa.2021.01.01